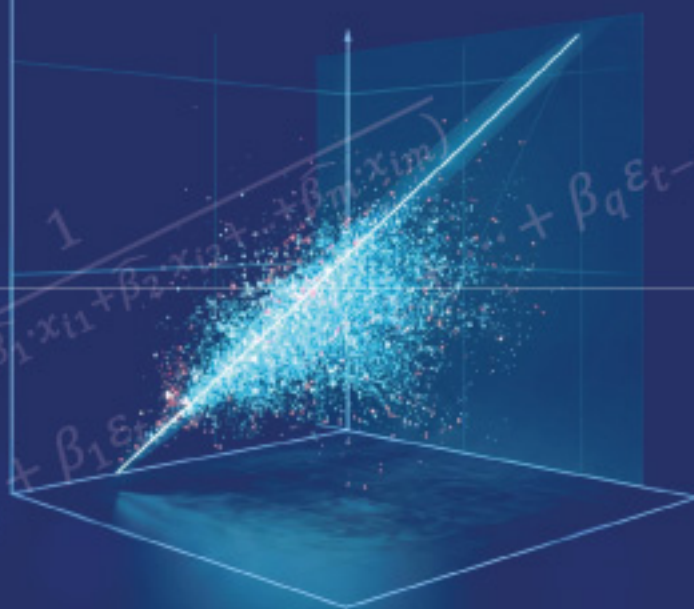


ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ
ՎԱՄԱԼՍԱՐԱՆ

Նաիրուհի Ջրբաշյան

ԷԿՈՆՈՄԵՏՐԻԿԱ



ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

ՆԱԽՐՈՒՀԻ ՋՐՔԱՇՅԱՆ

ԷԿՈՆՈՄԵՏՐԻԿԱ

(Մաս 1, Հիմունքներ)

ԵՐԵՎԱՆ

ԵՊՀ ՀՐԱՏԱՐԱԿԶՈՒԹՅՈՒՆ

2026

ՀՏԴ 330.43(075.8)

ԳՄԴ 65գ631գ73

Ջ 864

*Հրատարակության է երաշխավորվել
ԵՊՀ գիտական խորհրդի կողմից:*

Գիտական խմբագիր՝

տեխ. գիտությունների դոկտոր, պրոֆեսոր Արամ Առաքելյան

Գրախոսներ՝

տնտ. գիտությունների թեկնածու, դոցենտ Ալվարդ Խառատյան
ֆիզմաթ. գիտությունների թեկնածու, դոցենտ Աշոտ Կակոսյան

Ջրբաշյան Նաիրուհի

Ջ 864 Էկոնոմետրիկա (Մաս 1, Հիմունքներ) / Ն. Ջրբաշյան: -Եր.,
ԵՊՀ հրատ., 2026, 302 էջ:

Դասագիրքը Էկոնոմետրիկայի հիմունքների առաջին հայալեզու համապարփակ ներկայացումն է: Այն ընդգրկում է տարածական տվյալների ու ժամանակային շարքերի էկոնոմետրիկ մոդելավորման տեսական հիմունքներն ու դրանց գործնական կիրառությունները՝ համադրելով գիտական ներկայացումը մատչելի օրինակների հետ: Դասագրքի յուրաքանչյուր բաժին պարունակում է հարցեր ու առաջադրանքներ՝ ձեռք բերված տեսական գիտելիքների ստուգման և ամրապնդման համար, մաս կիրառական հմտությունների զարգացմանն ուղղված խնդիրներ:

Դասագրքի տեսական ու կիրառական բովանդակությունը համապատասխանում է ԵՊՀ տնտեսագիտության և կառավարման ֆակուլտետում դասավանդվող «Էկոնոմետրիկա» առարկայի ծրագրին՝ ապահովելով թե՛ ակադեմիական արդիականություն, թե՛ մեթոդական ամբողջականություն: Այն նախատեսված է ինչպես տնտեսագիտական, այնպես էլ այլ մասնագիտությունների բակալավրական ու մագիստրոսական ծրագրերի ուսանողների համար, որոնք ցանկանում են զարգացնել տվյալների հետ աշխատելու ու դրանց էկոնոմետրիկ մոդելավորման հմտությունները:

ՀՏԴ 330.43(075.8)

ԳՄԴ 65գ631գ73

ISBN 978-5-8084-2756-3

<https://doi.org/10.46991/YSUPH/9785808427563>

© ԵՊՀ հրատ., 2026

© Ջրբաշյան Ն., 2026

ԲՈՎԱՆԴԱԿՈՒԹՅՈՒՆ

ԴԱՍԱԳՐՔԻ ԿԱՌՈՒՑՎԱԾՔԸ.....	10
ԷԿՈՆՈՄԵՏՐԻԿ ՀԵՏԱԶՈՏՈՒԹՅԱՆ ՀԻՄՆԱԿԱՆ	
ԳՐՈՒՅԹՆԵՐԸ.....	12
1. Էկոնոմետրիկան որպես գիտություն, դրա նպատակը, առարկան ու խնդիրները	12
2. Էկոնոմետրիկ հետազոտության դասական մեթոդաբանությունը	18
3. Էմպիրիկ տվյալների տիպերը.....	19
ԳԼՈՒԽ 1. ՊԱՐԶ ԳԾԱՅԻՆ ՌԵԳՐԵՍԻԱ	21
1. Կախվածությունների տիպերը.....	21
2. Պարզ կամ միագործոն գծային ռեգրեսիայի տեսական ու Էմպիրիկ հավասարումներ ու մոդելներ	25
3. Պարզ կամ միագործոն գծային ռեգրեսիայի Էմպիրիկ հավասարման գնահատումը փոքրագույն քառակուսիների եղանակով (ՓՔԵ).....	36
4. Փոքրագույն քառակուսիների եղանակով գնահատված գծային պարզ ռեգրեսիայի Էմպիրիկ հավասարման որակի ստուգումը, դետերմինացիայի գործակից	45
5. Պարզ գծային ռեգրեսիայի Էմպիրիկ հավասարման ՓՔԵ-ով գնահատման և արդյունքների մեկնաբանության օրինակ.....	50
6. Գաուս-Մարկովի պայմաններն ու թեորեմը պարզ գծային ռեգրեսիայի համար. պարզ գծային ռեգրեսիայի նորմալ դասական մոդելի սահմանումը	53
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	59
Խնդիրներ	61
7. ՓՔԵ-ով գնահատված պարզ գծային ռեգրեսիայի Էմպիրիկ հավասարման վիճակագրական գնահատականների ճշգրտության վերլուծությունը	65
7.1. Ռեգրեսիայի Էմպիրիկ հավասարման գործակիցների վարիացիաներն ու ստանդարտ սխալները	65
7.2. Ռեգրեսիայի տեսական հավասարման պարամետրերի մասին վարկածների ստուգում. t թեստ	72
7.3. Ռեգրեսիայի տեսական հավասարման պարամետրերի միջակայքային գնահատականները	79

7.4. Ռեգրեսիայի գործակիցների վիճակագրական նշանակալիության ստուգման օրինակ	82
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	85
Խնդիրներ	86
ԳԼՈՒԽ 2. ԲԱԶՄԱԿԻ ԳԾԱՅԻՆ ՌԵԳՐԵՍԻԱ	90
1. Բազմակի գծային ռեգրեսիայի տեսական ու էմպիրիկ մոդելներն ու հավասարումները	90
2. Բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատումը ՓՔԵ-ով	96
3. Գաուս-Մարկովի պայմաններն ու թեորեմը բազմակի գծային ռեգրեսիայի համար, բազմակի գծային ռեգրեսիայի նորմալ դասական մոդելի սահմանումը	97
4. Գծային բազմակի ռեգրեսիայի էմպիրիկ հավասարման որակի ստուգումը. ճշգրտված դետերմինացիայի գործակից	100
5. Գծային բազմակի ռեգրեսիայի հավասարման կամ դետերմինացիայի գործակցի վիճակագրական նշանակալիության ստուգումը. F թեստ	101
6. Բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատման արդյունքների վերլուծության օրինակ	105
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	110
Խնդիրներ	111
ԳԼՈՒԽ 3. ԳԱՈՒՄ-ՄԱՐԿՈՎԻ ՊԱՅՄԱՆՆԵՐԻ ԽԱԽՏՈՒՄԸ ..	117
1. Պատահական շեղումների նորմալ բաշխվածության պայման. խախտման հետևանքներն ու հայտնաբերումը	117
2. Պատահական շեղումների ավտոկոռելացվածության բացակայության պայման. խախտման հետևանքներն ու հայտնաբերումը	120
2.1. Առաջին կարգի ավտոկոռելյացիայի առկայության ստուգման Դարբին-Ուոթսոնի թեստը	122
2.2. Բրեուշ-Գոդֆրիի՝ բարձր կարգերի ավտոկոռելյացիայի առկայության ստուգման Լագրանժի բազմապատկչի LM թեստը	125
3. Պատահական շեղումների հոմոսկեդաստիկության պայման. խախտման հետևանքները, հայտնաբերումն ու շտկումը	127
3.1. Հետերոսկեդաստիկության էությունը, հետևանքներն ու գրաֆիկական հայտնաբերումը	127

3.2. Ուայթի ու Բրեուշ-Պագանի՝ հետերոսկեղաստիկության առկայության ստուգման թեստերը	130
3.3. Հետերոսկեղաստիկության շտկումը. փոքրագույն քառակուսիների կշռված ու ընդհանրացված եղանակներ.....	133
4. Բազմակողմնեարության բացակայության պայման. խախտման հետևանքներն ու հայտնաբերումը.....	136
4.1. Լիակատար կողմնեարություն ու բազմակողմնեարություն	136
4.2. Բազմակողմնեարության հայտնաբերման VIF թեստը	139
5. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալներ.....	141
5.1. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ. կարևոր բացատրող փոփոխականը ներառված չէ մոդելում ..	141
5.2. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ. մոդելում ներառված է ավելորդ բացատրող փոփոխական	143
5.3. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ. կախվածության սխալ տեսքի հատկորոշում. հատկորոշման սխալի ստուգումը Ռամսեյի RESET թեստի միջոցով	145
6. Գաուս-Մարկովի պայմանների ստուգումը «EViews» փաթեթում և արդյունքների վերլուծությունը	147
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	153
Խնդիրներ	155
ԳԼՈՒԽ 4. ՈՉ ԳԾԱՅԻՆ ԿԱԽՎԱԾՈՒԹՅՈՒՆՆԵՐԻ	
ԳՆԱՀԱՏՈՒՄԸ.....	162
1. Ոչ գծային կախվածությունների տեսակները էկոնոմետրիկ մոդելավորման համատեքստում	162
2. Պարամետրերի նկատմամբ ոչ գծային, բայց գծայնացվող կախվածություններ ու դրանց մոդելավորումը	165
2.1. Երկկողմանի լոգարիթմական ռեգրեսիա	165
2.2. Կիսալոգարիթմական ռեգրեսիա. գծալոգարիթմական մոդել.....	169
2.3. Կիսալոգարիթմական ռեգրեսիա. լոգ-գծային մոդել.....	172
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	177
Խնդիրներ	178
ԳԼՈՒԽ 5. ՖԻԿՏԻՎ ՓՈՓՈԽԱԿԱՆՆԵՐԻ ՕԳՏԱԳՈՐԾՈՒՄԸ	
ԷԿՈՆՈՄԵՏՐԻԿ ՄՈԴԵԼՆԵՐՈՒՄ	182

1. Բացատրող երկարժեք ֆիկտիվ փոփոխականներով մոդելներ.....	185
2. Կախյալ երկարժեք ֆիկտիվ փոփոխականով մոդելներ. գծային հավանականական ու լոգիթ մոդելներ.....	190
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	199
Խնդիրներ	200
ԳԼՈՒԽ 6. ՊԱՏԱՀԱԿԱՆ ԳՈՐԾՐՆԹԱՑՆԵՐ ՈՒ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐ. ՀԻՄՆԱԿԱՆ ՀԱՏԿՈՒԹՅՈՒՆՆԵՐ ՈՒ ՆԱԽՆԱԿԱՆ ՈՒՍՈՒՄՆԱՍԻՐՈՒԹՅՈՒՆ	203
1. Պատահական գործընթացներ ու ժամանակային շարքեր. սահմանումներ ու տիպեր	203
2. Ժամանակային շարքերի բաղկացուցիչները.....	206
3. Ստացիոնար պատահական գործընթաց ու ժամանակային շարք, սպիտակ աղմուկ պատահական գործընթաց	209
4. Ոչ ստացիոնար պատահական գործընթաց ու ժամանակային շարք	213
5. Տարբերական օպերատոր ու տարբերությամբ վերցված ոչ ստացիոնար ժամանակային շարքեր.....	216
6. Տրենդային ստացիոնար (TS) ու տարբերությամբ ստացիոնար (DS) ժամանակային շարքերն ու դրանց ստացիոնարացումը, ինտեգրացվող ժամանակային շարքեր	217
7. Պատահական գործընթացի ու ժամանակային շարքի ավտոկոռելյացիայի գործակից, ավտոկոռելյացիայի ու մասնակի ավտոկոռելյացիայի ֆունկցիաներ, կոռելյորամ.....	221
8. Պատահական գործընթացի ավտոկոռելյացիայի գործակիցների մասին վիճակագրական վարկածների ստուգում	225
9. Դիկի-Ֆուլերի՝ ժամանակային շարքի ստացիոնարության ստուգման միավոր արմատի վիճակագրական թեստերը.....	228
10. Ժամանակային շարքի նախնական ուսումնասիրության օրինակ	231
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	238
Խնդիրներ	239
ԳԼՈՒԽ 7. ՍՏԱՑԻՈՆԱՐ ՄԻԱԶԱՓ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐԻ ԷԿՈՆՈՄԵՏՐԻԿ ՄՈԴԵԼԱՎՈՐՈՒՄ	243

1. Առաջին կարգի ավտոռեգրեսիայի AR(1) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները	243
2. Երկրորդ կարգի ավտոռեգրեսիայի AR(2) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները	247
3. Ընդհանուր դեպք՝ p -րդ կարգի ավտոռեգրեսիայի AR(p) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները	249
4. Յուլ-Ուոքերի հավասարումները	251
5. Ժամանակային շարքի մասնակի ավտոկոռելյացիայի ֆունկցիան ու AR(p) մոդելի p կարգի որոշումը.....	252
6. Սահող միջինների MA(1), MA(2) ու MA(q) մոդելներով բնութագրվող ժամանակային շարքեր.....	254
6.1. Առաջին կարգի սահող միջինների MA(1) մոդելով բնութագրվող ժամանակային շարքի հատկությունները	255
6.2. Երկրորդ կարգի սահող միջինների MA(2) մոդելով բնութագրվող ժամանակային շարքի հատկությունները	256
6.3. Ընդհանուր դեպք՝ q -րդ կարգի սահող միջինների MA(q) մոդելով բնութագրվող ժամանակային շարքի հատկություններն ու սահող միջինների MA(q) մոդելի q կարգի որոշումը	258
7. Ավտոռեգրեսիայի ու սահող միջինների ARMA(p , q) մոդելով բնութագրվող ժամանակային շարք, դրա հատկությունները	260
8. Ավտոռեգրեսիայի ու սահող միջինների ինտեգրացվող ARIMA(p , d , q) մոդելն ու դրա գնահատման Բոքս-Ջենքինսի մեթոդաբանությունը	262
9. Ժամանակային շարքի ARIMA(p , d , q) մոդելավորման օրինակ	266
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	274
Խնդիրներ	275
ԳԼՈՒԽ 8. ԲԱԶՄԱԶԱԲ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐԻ ՄՈԴԵԼԱՎՈՐՄԱՆ ՀԻՄՆԱՀԱՐՑԵՐՆ ՈՒ ՈՐՈՇ ՄՈԴԵԼՆԵՐԸ	278
1. Ժամանակային շարքերի կեղծ ռեգրեսիա	278
2. Ժամանակային շարքերի կոինտեգրացիան ու դրա ստուգումը	280

3. Կոինտեգրացվածությունն ու սխալներով ճշգրտման մեխանիզմը	284
4. Ժամանակային շարքերի դիմամիկ մոդելներ՝ բաշխված լագերի DL(p) ու ավտոռեգրեսիայի և բաշխված լագերի ARDL(p,q) մոդելների համառոտ նկարագրություն.....	287
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով	295
SUMMARY.....	297
АВНОТАЦИЯ.....	298
ՕԳՏԱԳՈՐԾՎԱԾ ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ.....	299

ՆԵՐԳԻՐՆԵՐ

Ներդիր 1. «Ռեգրեսիա» տերմինի ծագումնաբանությունն ու գործածումը էկոնոմետրիկայում.....	24
Ներդիր 2. Երկու փոփոխականների զծային կախվածություն կամ զծային ֆունկցիա.....	28
Ներդիր 3. Հետազոտության օբյեկտների գլխավոր համախմբություն ու նմուշ.....	30
Ներդիր 4. Նմուշից սրացվող վիճակագրական գնահատականներն ու դրանց հատկությունները	32
Ներդիր 5. Նմուշի թվային բնութագրիչները	39
Ներդիր 6. Նմուշից սրացված արժեքների փոխկախվածության բնութագրիչը՝ կոռելյացիայի գործակից.....	42
Ներդիր 7. Դերերմինացիայի գործակից ներկայացումը ընդունված նշանակումների միջոցով.....	49
Ներդիր 8. Գաուս-Մարկովի թեորեմի ապացույցը. անշեղություն.....	58
Ներդիր 9. Գաուս-Մարկովի թեորեմի ապացույցը. արդյունավետություն	67
Ներդիր 10. Նորմալ բաշխման մի քանի կարևոր հատկություններ.....	70
Ներդիր 11. Վիճակագրական վարկածների ստուգման որոշ հիմնական դրույթներ.....	73

Ներդիր 12. Վիճակագրական վարկածների ստուգման համար օգտագործվող որոշ քաշխումների հիմնական հատկությունները.....	78
Ներդիր 13. Առանց ազատ անդամի գծային ռեգրեսիայի մոդելի գնահատման առանձնահատկությունները	81
Ներդիր 14. Բազմակի ու մասնակի կոռելյացիայի գործակիցներ.....	95
Ներդիր 15. Կիսալոգարիթմական մոդելների գործակիցների մեկնաբանության բացատրությունը	175
Ներդիր 16. Ֆիկրիվ փոփոխականների «քակարդը».....	183
Ներդիր 17. Սեզոնայնության հեռացումը ԺՇ սրացիոնարացման նպատակով.....	220
Ներդիր 18. Կեղծ ռեգրեսիայի երևույթի հայտնաբերման պայմանությունը	278
Ներդիր 19. Գրեյնջերի պատճառականության թեստը.....	292

ԴԱՍԱԳՐՔԻ ԿԱՌՈՒՅՎԱԾՔԸ

Գլուխ 1-ում նկարագրված են պարզ գծային ռեգրեսիայի տեսական ու էմպիրիկ մոդելները, պարզ գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատման փոքրագույն քառակուսիների եղանակը, գնահատման որակի ու վավերականության ստուգման չափանիշները, ինչպես նաև տեսական մոդելի վրա դրվող պայմաններն (Գաուս-Մարկովի պայմաններ) ու գնահատականների հատկություններն այդ պայմանների տեղի ունենալու դեպքում:

Գլուխ 2-ում քննության են առնվում բազմակի գծային ռեգրեսիայի տեսական ու էմպիրիկ մոդելները, բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատման փոքրագույն քառակուսիների եղանակը, գնահատման որակի ու վավերականության ստուգման չափանիշները, տեսական մոդելի վրա դրվող պայմաններն ու գնահատականների հատկություններն այդ պայմանների տեղի ունենալու դեպքում:

Գլուխ 3-ը ներկայացնում է տեսական մոդելի վրա դրվող Գաուս-Մարկովի պայմանների խախտումները, էմպիրիկ մոդելի գնահատականների վրա դրանց խախտման ունեցած հետևանքները, ինչպես նաև Գաուս-Մարկովի պայմանների ստուգման համար օգտագործվող հիմնական վիճակագրական թեստերը:

Գլուխ 4-ը վերաբերում է ոչ գծային կախվածությունների էկոնոմետրիկ մոդելավորմանն ու դրանց առանձնահատկություններին: Մասնավորապես դիտարկվում են երկկողմանի լոգարիթմական ու կիսալոգարիթմական ռեգրեսիայի մոդելները և դրանց գնահատականների հատկությունները:

Գլուխ 5-ում ներկայացված է երկարժեք ֆիկտիվ փոփոխականների օգտագործումը էկոնոմետրիկ մոդելներում: Դիտարկված են էկոնոմետրիկ մոդելներում երկարժեք ֆիկտիվ փոփոխականների՝ որպես բացատրող փոփոխականներ օգտագործման առանձնահատկություններն ու դեպքերը: Ներկայացված են կախյալ երկարժեք ֆիկտիվ փոփոխականով գծային հավանականական ու Լոգիթ մոդելները:

Գլուխ 6-ը անդրադառնում է ժամանակային շարքերին, դրանց բաղադրիչներին ու հատկություններին, դիտարկվում է ժամանակային շարքերի ստացիոնարությունը, ստացիոնար ու ոչ ստացիոնար շարքերի մոդելների դասական օրինակները (սպիտակ աղմուկի ու պա-

տահական քայքի մոդելներ), ժամանակային շարքի ստացիոնարության ստուգման ու ստացիոնարացման մեխանիզմները:

Գլուխ 7-ում քննության են առնվում միաչափ ժամանակային շարքերի մոդելները, դրանց հատկությունները, ARIMA մոդելավորման տեխնիկան:

Գլուխ 8-ը անդրադառնում է բազմաչափ ժամանակային շարքերի մոդելավորման որոշ հիմնադրույթների՝ կեղծ ռեգրեսիային, կոինտեգրացիային ու սխալների ճշգրտման մեխանիզմին: Անդրադարձ է կատարվում նաև բաշխված լագերի ու ավտոռեգրեսիայի և բաշխված լագերի մոդելներին:

Բոլոր գլուխները ամփոփվում են խնդիրների լուծման օրինակներով, գիտելիքների ստուգման նպատակով տրվող հարցերով ու առաջադրանքներով, ինչպես նաև խնդիրներով: Չեռնարկում առկա են ներդիրներ, որոնք մանրամասնում են տվյալ թեման լրացուցիչ բացատրություններով:

Էկոնոմետրիկ մոդելավորման խնդիրներն ու դրանց լուծումները դիտարկվում են ժամանակակից ծրագրային այնպիսի փաթեթների օգտագործմամբ, ինչպիսիք են SPSS ու EViews փաթեթները:

ԷԿՈՆՈՄԵՏՐԻԿ ՀԵՏԱԶՈՏՈՒԹՅԱՆ ՀԻՄՆԱԿԱՆ ԴՐՈՒՅԹՆԵՐԸ

1. Էկոնոմետրիկան որպես գիտություն, դրա նպատակը, առարկան ու խնդիրները

Էկոնոմետրիկան տնտեսագիտական կրթության առանցքային գիտակարգերից է, մակրոտնտեսագիտության ու միկրոտնտեսագիտության հետ միասին այն երեք հիմնապայուներից մեկը, որոնց վրա կանգնած է տնտեսագիտական կրթությունը:

Էկոնոմետրիկան գիտության համեմատաբար նոր, սակայն արագ զարգացող ճյուղ է: «Էկոնոմետրիկա» տերմինը առաջին անգամ գործածվել է 1910 թ. լեհ տնտեսագետ Պավել Յիոմայայի կողմից: Բառացի թարգմանությամբ «էկոնոմետրիկա» տերմինը նշանակում է «տնտեսության (econo) չափում (metrica)», սակայն էկոնոմետրիկայի շրջանակը շատ ավելի լայն է, քան տնտեսության մեջ չափումների իրականացումը:

Էկոնոմետրիկան՝ որպես գիտություն, սկսել է զարգանալ 1930-ական թվականներից: Գիտական իմաստով էկոնոմետրիկան 1926 թ. ձևակերպել է նորվեգացի տնտեսագետ Ռագնար Ֆրիշը: 1930 թ. դեկտեմբերին Ռ. Ֆրիշի, Ի. Ֆիշերի, Թ. Թիմբերգենի, Շ. Շումպիտերի, Օ. Սնդերսոնի և այլ գիտնականների մախաձեռնությամբ ստեղծվեց Էկոնոմետրիկ ընկերությունը՝ ներառելով էկոնոմետրիկայով զբաղվող գիտնականներին: 1933 թ. Ռ. Ֆրիշը հիմնադրեց «Էկոնոմետրիկա» գիտական ամսագիրը, որը մինչ այժմ մեծ նշանակություն ունի տնտեսագիտության զարգացման համար: Արդեն 1941 թ. տպագրվեց հոլանդացի գիտնական Յան Թիմբերգենի հեղինակած նոր գիտական ճյուղի՝ էկոնոմետրիկայի առաջին դասագիրքը: Դրանից հետո էկոնոմետրիկան արագ զարգացում է ապրել: Բավական է նշել էկոնոմետրիկայի բնագավառում կատարած գիտական աշխատանքների համար Նոբելյան մրցանակ ստացած գիտնականների ցանկը.

- 1969 թ. **Ռ. Ֆրիշն** ու **Յ. Թիմբերգենը**՝ տնտեսական ցիկլերի վերլուծության նպատակով դինամիկ էկոնոմետրիկ մոդելների ստեղծման ու կիրառման համար: Ֆրիշն ու Թիմբերգենը Էկոնոմետրիկայի՝ որպես գիտության հիմնադիրներն են:
- 1975 թ. ծագումով հոլանդացի, ամերիկացի տնտեսագետ **Թյալինգ Ս. Կուլամանս** ու ռուս տնտեսագետ մաթեմատի-

կոս **Լեոնիդ Կանտորովիչը** պարզևատրվել են Նոբելյան մրցանակով ռեսուրսների օպտիմալ տեղաբաշխման տեսության մեջ ունեցած ներդրման համար: Չնայած նրան, որ մրցանակը ուղղակիորեն ուղղված չէր էկոնոմետրիկ հետազոտություններին, Կուպմանսի աշխատանքը էկոնոմետրիկ խոր արմատներ ուներ և իմտեցրում էր էկոնոմետրիկ մոդելավորումը: Թյալինգ Ս. Կուպմանսի էկոնոմետրիկ տեսության մրցանակը հիմնադրվել է նրա պատվին՝ ընդգծելով նրա երկարատև ազդեցությունը էկոնոմետրիկայի՝ որպես գիտության զարգացման վրա:

- 1980 թ. **Լոուրենս Քլայնը**՝ տնտեսության մոդելներ ստեղծելու, տնտեսության մեջ տատանումների ու տնտեսական քաղաքականության վերլուծության նպատակով դրանց կիրառման համար: Գոլդբերգի հետ միասին նա ստեղծեց ամերիկյան տնտեսության ամենահայտնի մոդելներից մեկը, որը հայտնի է որպես Քլայն-Գոլդբերգի մոդել: Այս մոդելը փոխկապակցված էկոնոմետրիկ հավասարումների համակարգ է, որի լուծումը տալիս էր ԱՄՆ տնտեսության զարգացման պատկերն ու կանխատեսումները:
- 1989 թ. **Տրիզվե Հասվելմոն**՝ ֆինն գիտնականը, որը զարգացրել է էկոնոմետրիկայի տեսական ու հավանականական հիմքերը և մրցանակը ստացել է՝ էկոնոմետրիկայի հիմունքները պարզաբանելու ու էկոնոմետրիկ մոդելավորման հավանականական մոտեցումները սահմանելու համար:
- 2000 թ. ամերիկացի տնտեսագետներ **Ջեյմս Հեքմանն** ու **Գանիել ՄաքՖադենը**. Հեքմանը՝ նմուշների վերլուծության տեսության ու մեթոդների մշակման համար, ՄաքՖադենը՝ դիսկրետ ընտրության տեսության ու մեթոդների, մասնավորապես լոգիթ մոդելի մշակման համար:
- 2003 թ. **Ռոբերտ Էնգլն** ու **Քլայվ Գրեյնջերը**՝ փոփոխական անկայունությամբ ժամանակային շարքերի մոդելավորման մեթոդների մշակման համար (ARCH մոդելներ ու կոինտեգրացվող ժամանակային շարքերի մոդելավորման մեթոդներ): Նրանց աշխատությունները հեղափոխական նվաճումներ են ֆինանսական էկոնոմետրիկայի ու մակրոտնտեսական կանխատեսման բնագավառներում:
- 2011 թ. **Քրիստոֆեր Սիմսն** ու **Թոմաս Սարգենբըր**՝ մակրոտնտեսության մեջ պատճառահետևանքային կապերի վեր-

լուծության համար: Նրանք զարգացրել են վեկտորական ավտոռեգրեսիայի (VAR) ու ռացիոնալ սպասումների էկոնոմետրիկ մոդելների շրջանակը:

- 2013 թ. ամերիկացի տնտեսագետներ **Եվգենի Ֆաման, Լարս Պիտեր Հանսեն** ու **Ռոբերտ Շիլլերը**՝ ֆինանսական շուկայի ակտիվների գների էմպիրիկ էկոնոմետրիկ վերլուծության ու կանխատեսման մեթոդների մշակման համար: Հանսենը ընդհանրացված մոմենտների մեթոդի (GMM) հեղինակն է: Վերջինս հնարավոր է դարձնում դասական ենթադրությունները չբավարարող տվյալների էկոնոմետրիկ մոդելավորումը: Հանսենը էկոնոմետրիկայի առաջատար տեսաբաններից է, նրա ներդրումը հատկապես կարևոր է դինամիկ մոդելների գնահատման բնագավառում: Ֆաման շուկայի արդյունավետության վարկածի (Efficient Market Hypothesis) հիմնադիրներից է, մշակել է իրադարձությունների վերլուծության (event study) մեթոդաբանությունը՝ նորություններին շուկայի արձագանքը չափելու համար: Շիլլերը ներդրել է վարքաբանական ֆինանսների (behavioral finance) գաղափարները՝ շեշտելով շուկայի ոչ ռացիոնալ վարքագիծը, մշակել է Case-Shiller Home Price Index-ը անշարժ գույքի գների վերլուծության համար: Նրա աշխատանքները հիմնված են ժամանակային շարքերի էկոնոմետրիկ վերլուծության վրա:
- 2021 թ. **Ջոշուա Էնգրիստը, Գուիդո Իմբենսն** ու **Դեյվիդ Քարդը**՝ պատճառահետևանքային կապերի վերլուծության մեթոդաբանական ներդրումների համար: Նրանք կիրառել են գործիքային փոփոխականներն ու բնական փորձերը էկոնոմետրիկ վերլուծություններում:

Գոյություն ունեն էկոնոմետրիկայի՝ որպես գիտության տարբեր սահմանումներ, որոնք տրվել են հիմնադիր, ինչպես նաև հանրահայտ տնտեսագետների կողմից: Դիտարկենք մի քանիսը:

Ըսր Ջոնզի (1933) էկոնոմետրիկական չափաք է դիտարկվի որպես մաթեմատիկական մեթոդների կիրառություն տնտեսագիտության մեջ: Վիճակագրությունը, տնտեսագիտության տեսությունն ու մաթեմատիկական մեթոդները, առանձին վերցված, անհրաժեշտ են, բայց ոչ բավարար ժամանակակից տնտեսության քանակական հարաբերությունների իրական պատկերացման համար: Այս երեքի համակ-

ցությունն է, որ տալիս է հզոր արդյունք: Եվ հենց այս համակցությունն է էկոնոմետրիկան¹:

Հասպեմոյի (1944) քնորոշմամբ էկոնոմետրիկ հետազոտության մեթոդը, ըստ էության, ուղղված է տնտեսագիտության տեսության ու փաստացի չափումների կամրջմանը՝ օգտագործելով վիճակագրական եզրակացության տեսությունն ու տեխնիկան որպես կամրջի հենարան²:

Ըստ Սեմյոնյանի (1954)՝ էկոնոմետրիկան իրական տնտեսական երևույթների քանակական վերլուծություն է՝ հիմնված տեսության ու դիտարկումների միաժամանակյա զարգացման վրա, որոնք կապված են եզրակացությունների համապատասխան մեթոդներով³:

Թեյլի քնորոշմամբ «էկոնոմետրիկան զբաղվում է տնտեսական օրենքների էմպիրիկ ոլորտում»:

Մայենվոյի սահմանմամբ «էկոնոմետրիկան տնտեսական օրենքների էմպիրիկ/փորձական բացահայտումն է»:

Ըստ Ալվազյանի էկոնոմետրիկան գիտակարգ է, որը, հիմնվելով տնտեսագիտության տեսության, տնտեսական վիճակագրության ու մաթեմատիկական գործիքամիջոցների վրա, քանակական արտահայտություն է տալիս տնտեսագիտության տեսության ընդհանուր (որակական) օրինաչափություններին⁴:

Ընդհանրացնելով եղած սահմանումները՝ կարող ենք ամփոփել հետևյալ կերպ. ***էկոնոմետրիկան տնտեսագիտության տեսության, տնտեսական ու մաթեմատիկական վիճակագրության ու մաթեմատիկական մոդելավորման մեթոդների վրա հիմնված գիտություն է, որը, օգտագործելով իրական տվյալները, գնահատում և քանակական արտահայտություն է տալիս տնտեսական երևույթների միջոցառություն ունեցող կապերին ու կախվածություններին:***

Էկոնոմետրիկան զբաղեցնում է գիտական ու կիրառական հետազոտությունների իր տեղը տնտեսագիտության տեսության, տնտեսական վիճակագրության, մաթեմատիկական վիճակագրու-

¹ Frisch, R. (1933). Editor's Note // *Econometrica*, 1(1), pp. 1-4.

² Haavelmo, T. (1944). The Probability Approach in Econometrics // *Econometrica*, 12, pp. Iii-115.

³ J. R. N. Stone, P. A. Samuelson, T. C. Koopmans. Report of the Evaluative Committee for Econometrica // *Econometrica*: Apr, 1954, Volume 22, Issue 2 , pp. 141-146.

⁴ Айвазян, Сергей Артемьевич. Методы эконометрики: учебник. (2010), стр. 19-20:

թյան, հավանականության տեսության ու մաթեմատիկական մոդելավորման հատման կետում: Էկոնոմետրիկայի կապը նշված գիտական ուղղությունների հետ արտահայտվում է տարբեր ձևերով.

1) տնտեսագիտության տեսությունը սահմանում է տնտեսական երևույթների միջև եղած կախվածությունները կամ դնում է դրանց մասին որոշակի վարկածներ,

2) տնտեսական վիճակագրությունը տեղեկատվական հիմք է տրամադրում տնտեսագիտական հետազոտությունների համար՝ ներկայացնելով ընտրված տնտեսական ցուցանիշների էմպիրիկ, դիտարկված վիճակագրական տվյալներն ու ապահովելով դրանց առաջնային վիճակագրական մշակումը,

3) մաթեմատիկական մոդելավորումը ձևակերպում է տնտեսական կախվածությունները մաթեմատիկական լեզվով, մաթեմատիկական մոդելի տեսքով,

4) հավանականությունների տեսությունն ու մաթեմատիկական վիճակագրությունը հիմք են այն գործիքների մշակման համար, որոնք օգտագործվում են էկոնոմետրիկայում:

Այսպիսով՝ էկոնոմետրիկան ուսումնասիրում է տնտեսական օրինաչափությունների կամ կախվածությունների քանակական կողմը: Տնտեսագիտության տեսությունը սահմանում է տնտեսական երևույթների որակական օրինաչափությունները: Օրինակ՝ միկրոտնտեսագիտությունը սահմանում է, որ այլ հավասար պայմաններում որևէ ապրանքի գնի նվազումն առաջ կբերի դրա պահանջարկի աճ: Այսինքն՝ կա հակադարձ կախվածություն ապրանքի գնի ու պահանջարկի միջև: Սակայն տեսությունը չի առաջարկում այդ կախվածության որևէ քանակական չափում, այսինքն՝ կոնկրետ որևէ ապրանքի համար դրա գնի նվազումը պահանջարկի որքան ավելացում կառաջացնի: Կամ՝ տնտեսագիտության տեսությունից հայտնի է (Ֆիլիփսի կոր), որ կարճաժամկետ հեռանկարում գնաճի մակարդակն ու գործազրկությունը հակադարձ կախվածություն մեջ են. գործազրկության մակարդակի նվազման հետ գներն աճում են: Սակայն հարց է առաջանում, թե գործազրկության մակարդակի միավոր նվազումը գների որքան աճ է պայմանավորում: Կամ՝ ըստ Քեյնսի՝ տնային տնտեսությունների եկամտի աճի հետ դրանց սպառումն աճում է ավելի փոքր համամասնությամբ, սակայն ինչպիսի՞ն է այդ համամասնությունը: Կամ՝ Օուքենի օրենքն ասում է, որ երկրի ՀՆԱ-ի իրական մակարդակի շեղումը պոտենցյալ ՀՆԱ-ից հակադարձ համեմատական է գործազրկության իրական մակարդակի շեղմանը իր բնա-

կան մակարդակից: Որքա՞ն $\dot{Z}$ ՆԱ է կորցնում երկիրը գործազրկության բնական մակարդակի նկատմամբ դրա իրական մակարդակի 1 % աճի դեպքում. այս հարցերի պատասխանները տեսությունը չի տալիս:

Մաթեմատիկական մոդելավորման գործիքները հնարավորություն են տալիս մաթեմատիկական արտահայտություն տալու նշված տեսական օրինաչափություններին: Օրինակ՝ որևէ ապրանքատեսակի պահանջարկի (Q^d) կախումը դրա գնից (P), մրցակցող ապրանքի գնից (P^c) ու եկամտից (I) կարելի է մաթեմատիկորեն արտահայտել այսպես՝ $Q^d=f(P, P^c, I)$: Կամ՝ տնային տնտեսությունների եկամտից (I) դրանց սպառման (C) կախումն այսպես՝ $C=f(I)$: Սակայն մաթեմատիկական մոդելավորման գործիքները հնարավորություն չեն տալիս չափելու այդ կախվածության քանակական կողմը:

Տնտեսական վիճակագրությունը տեղեկատվություն է տրամադրում, օրինակ, ապրանքի գնի ու սպառման կամ տնային տնտեսությունների եկամտի ու սպառման մասին, ներկայացնում է դրանք աղյուսակների ու գծապատկերների տեսքով: Մաթեմատիկական վիճակագրությունը վիճակագրական վերլուծություններ է կատարում տնտեսական ցուցանիշների կապի ուժգնության ու ուղղության, դրանց համատեղ բաշխման ու վարքի մասին:

Էկոնոմետրիկան, օգտագործելով վերը նշված բոլոր մեթոդներն ու տվյալները, գնում է դրանից առաջ և պատասխան է տալիս «որքանո՞վ» հարցին, օրինակ՝ գնահատում է, թե այլ հավասար պայմաններում տվյալ ապրանքի գնի նվազումը միջին հաշվով որքանով կավելացնի դրա պահանջարկը:

Այսպիսով՝ ***Էկոնոմետրիկայի ուսումնասիրության առարկան փոքրահասկան երևույթների միջև գոյություն ունեցող կապերի ու կախվածությունների քանակական կողմն է:***

Էկոնոմետրիկայի նպատակը ցանկացած մակարդակում կառավարման որոշումների ընդունման, ինչպես նաև տնտեսական զարգացման քաղաքականությունների ու ռազմավարությունների մշակման համար գնահատականների, կանխատեսումների ու սցենարների մշակումն ու ստացումն է տնտեսական օրինաչափությունների քանակական կողմի գնահատման միջոցով:

Էկոնոմետրիկայի խնդիրը Էկոնոմետրիկ մոդելների ձևակերպումը, գնահատումը, դրա վավերականության ու որակի ստուգումն ու եզրակացությունների ձևակերպումն է:

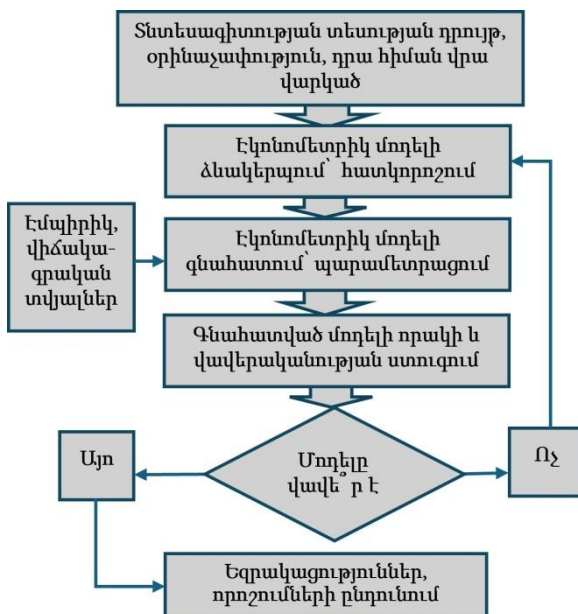
2. Էկոնոմետրիկ հետազոտության դասական մեթոդաբանությունը

Եթե էկոնոմետրիկ հետազոտությունը ներկայացնենք սխեմատիկորեն, որպես առանձին փուլերի հաջորդականություն, ապա կունենանք 6 փուլ:

- **Փուլ 1:** Հետազոտական խնդրի՝ տնտեսագիտության տեսության դրույթի կամ կախվածությունների ու դրանից բխող վարկածի սահմանում: Սա էկոնոմետրիկ հետազոտության նպատակների որոշման փուլն է:
- **Փուլ 2:** Դրանից բխող մաթեմատիկական ու էկոնոմետրիկ մոդելի ձևակերպում՝ մոդելի **հատկորոշում**: Այս քայլում որոշվում, սահմանվում են էկոնոմետրիկ մոդելի տեսքը՝ ցուցանիշներն ու գործոնները, որոնք ներառվելու են մոդելում, ինչպես նաև կախվածության բնույթն ու տեսքը:
- **Փուլ 3:** Վիճակագրական, էմպիրիկ տվյալների հավաքագրում և/կամ մշակում՝ այնուհետև դրանք մոդելի գնահատման համար օգտագործելու նպատակով:
- **Փուլ 4:** Վիճակագրական, էմպիրիկ տվյալների հիման վրա էկոնոմետրիկ մոդելի գնահատում՝ **պարամետրացում**:
- **Փուլ 5:** Գնահատված մոդելի **վավերականության ու որակի ստուգում**: Ստուգվում է, թե որքանով է գնահատված մոդելը համապատասխանում իրական տնտեսական օրինաչափություններին: Բացի դրանից՝ ստուգվում է, թե ինչպիսին են գնահատված մոդելի վիճակագրական բնութագրիչները, և որքանով ճիշտ են տեղի ունեցել մոդելի հատկորոշումն ու պարամետրացումը: Այս փուլում որոշվում է, թե արդյոք գնահատված մոդելը կարելի է օգտագործել՝ եզրակացություններ կատարելու համար:
- **Փուլ 6:** Տնտեսագիտության տեսության դրույթի կամ օրինաչափության ու դրանից բխող վարկածի մասին եզրակացությունների ձևակերպում, ինչպես նաև կանխատեսումներ ու առաջարկություններ՝ կառավարման որոշումների ընդունման ու զարգացման քաղաքականությունների մշակման համար:

Այս փուլերը ներկայացնում են էկոնոմետրիկ վերլուծության դասական մեթոդաբանությունը (տե՛ս Գծանկար 1):

Գծանկար 1. Էկոնոմետրիկ հետազոտության փուլերի սխեմայիկ ներկայացումը



3. Էմպիրիկ տվյալների տիպերը

Էկոնոմետրիկ հետազոտության փուլերից մեկը վերաբերում է վիճակագրական, էմպիրիկ տվյալների հավաքագրմանն ու մշակմանը՝ իրական տվյալների հիման վրա մոդելի գնահատման նպատակով: Այդ տվյալները կարելի է խմբավորել ներքոհիշյալ տիպերում:

1. **Տարածական տվյալները (cross-sectional data)** ժամանակի որևէ որոշակի պահին օբյեկտների համախմբության (երկրներ, ընկերություններ, տնային տնտեսություններ, սպառողներ և այլն) մեկ կամ մի քանի ցուցանիշների արժեքների համախմբություն են: Այս տվյալները ժամանակի որոշակի մեկ ամրագրված պահի համար են և վերաբերում են «տարածության մեջ» տարբեր օբյեկտների, օրինակ՝ ԵՄ 27 երկրների ՀՆԱ-ները, ներդրումներն ու արտահանումը 2026 թ.: Այս տիպի տվյալների հիման վրա էկոնոմետրիկ մոդելավորումն իրականացնելու համար սովորաբար ստեղծվում են մոդելի մեջ մտնող ցուցանիշներին վերաբերող տվյալների երկչափ շտեմարան-

ներ՝ աղյուսակներ, որոնցում օբյեկտները տողերում են (օրինակ՝ Ֆրանսիա, Գերմանիա, Իսպանիա և այլն), իսկ դրանց վերաբերող տվյալները (օրինակ՝ ՀՆԱ, ներդրումներ, արտահանում)՝ սյուներում:

2. Ժամանակային շարքերը (time series) Ժամանակի հավասարահեռ պահերին որևէ մեկ (կամ մի քանի) ցուցանիշի(ների) արժեքների համախմբություն են մեկ օբյեկտի համար: Այս տվյալները վերաբերում են ժամանակի տարբեր պահերի համար որոշակի մեկ ամրագրված օբյեկտի, օրինակ՝ Հայաստանի ՀՆԱ-ի տվյալները եռամսյակային կտրվածքով՝ 2000-2026 թթ.: Այս տիպի տվյալների հիման վրա էկոնոմետրիկ մոդելավորումն իրականացնելու համար սովորաբար ստեղծվում են մոդելի մեջ մտնող ցուցանիշի(ների) արժեքների տվյալների աղյուսակներ, որոնցում օբյեկտը մեկն է: Աղյուսակի տողերում ժամանակի հավասարահեռ պահերն են (օրինակ՝ 2000 թ. առաջին եռամսյակ, 2000 թ. երկրորդ եռամսյակ և այլն), իսկ սյունում(երում)՝ տվյալ ցուցանիշի(ներին) վերաբերող տվյալները (օրինակ՝ Հայաստանի ՀՆԱ-ն, ներդրումներն ու արտահանումը):

3. Պանելային տվյալները (panel data) Ժամանակի հավասարահեռ պահերին օբյեկտների որոշակի համախմբությանը վերաբերող ցուցանիշի(ների) արժեքների համախմբություն են: Այս տվյալները վերաբերում են օբյեկտների համախմբությանը ժամանակի տարբեր պահերին: Այսինքն՝ սրանք և՛ տարածության մեջ են, և՛ ժամանակի, օրինակ՝ եվրոպական 27 երկրների ՀՆԱ-ները, ներդրումները և արտահանումը՝ 2000- 2026 թթ.: Այս տիպի տվյալների հիման վրա էկոնոմետրիկ մոդելավորումն իրականացնելու համար սովորաբար ստեղծվում են մոդելի մեջ մտնող ցուցանիշի(ների) արժեքների տվյալների աղյուսակներ, որոնցում օբյեկտները մեկից ավելին են, իսկ դրանց վերաբերող տվյալները ժամանակի մեջ են: Աղյուսակի տողերում յուրաքանչյուր օբյեկտն է ժամանակի հավասարահեռ պահերին (օրինակ՝ Ֆրանսիան 2000 թ., 2001 թ. և այլն, Գերմանիան 2000-ին, 2001-ին և այլն, այդպես բոլոր օբյեկտների համար), իսկ տվյալ ցուցանիշին վերաբերող տվյալները սյուներում են (օրինակ՝ ՀՆԱ-ն, ներդրումներն ու արտահանումը):

1-ին տիպի տվյալների էկոնոմետրիկ մոդելավորման լավագույն մոտեցումը ռեգրեսիայի մոդելներն են: Դրանք նկարագրվում են Գ-լուխ 1-5-ում: 2-րդ տիպի տվյալների էկոնոմետրիկ մոդելավորման համար օգտագործվում են այլ մեթոդներ, որոնցից այս դասընթացում հիմնականում ուսումնասիրված է ARIMA մոդելավորումը: Դրանք նկարագրվում են Գ-լուխ 6-8-ում: 3-րդ տիպը այս դասընթացում չի ընդգրկվել:

ԳԼՈՒԽ 1. ՊԱՐԶ ԳԾԱՅԻՆ ՌԵԳՐԵՍԻՎ

1. Կախվածությունների տիպերը

Ընդհանուր առմամբ տարբերում են տարբեր մեծությունների կամ փոփոխականների միջև կախվածությունների երկու տիպ՝ ֆունկցիոնալ ու վիճակագրական:

Ֆունկցիոնալ կախվածության դեպքում մի մեծության յուրաքանչյուր արժեքի համապատասխանում է մյուս մեծության խիստ որոշակի միակ արժեք: Այսպիսի կախվածությունները հատուկ են բնական գիտություններին: Օրինակ՝ ֆիզիկական մարմնի էներգիան հանգստի պահին հավասար է նրա զանգվածի ու վակուումում լույսի արագության քառակուսու արտադրյալին՝ $E=mc^2$: Այսպիսի կախվածությունները **դետերմինացված** կախվածություններ են. դրանք խիստ որոշակի են: Տնտեսության մեջ այդպիսի կախվածության օրինակ կարող է լինել, օրինակ, արտադրական եղանակով հաշվարկված ՀՆԱ-ի կախումը իր բաղադրիչներից, որոնց հիման վրա այն հաշվարկվում է որպես գումար (օրինակ՝ ՀՆԱ=սպառում + համախառն ներքին մասնավոր ներդրումներ + կառավարության ծախսեր + գուտ արտահանում):

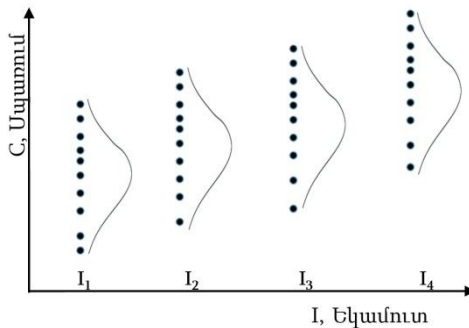
Վիճակագրական կախվածության դեպքում մի մեծության կամ փոփոխականի յուրաքանչյուր արժեքին համապատասխանում է մյուս մեծության հնարավոր արժեքների բազմություն, արժեքների որոշակի բաշխում: Վիճակագրական կախվածությունները **ստոխաստիկ** կամ **հավանականական** բնույթ ունեն, որովհետև մի մեծության յուրաքանչյուր արժեքի դեպքում մյուս մեծության արժեքը ամեն մի կոնկրետ դեպքում կախված է նաև այլ՝ պատահական բնույթ ունեցող ազդեցություններից (ներառյալ նաև չվերահսկվող, չհաշվառված ազդեցությունները): Օրինակ՝ ենթադրենք ունենք Հայաստանի 5000 տնային տնտեսությունների ամսական եկամուտների ու սպառողական ծախսերի մասին տվյալներ: Դիտարկում ենք սպառողական ծախսերի կախումը տնօրինվող եկամտից: Նույն ամսական եկամուտն ունեցող տնային տնտեսությունները ճշգրիտ նույն ամսական սպառումը չեն ունենա. դրանց սպառումները կտարբերվեն՝ կախված այլ գործոններից: Նույն եկամուտն ունեցող տնային տնտեսությունների սպառումները կարող են տարբերվել պատահական ազդեցությունների պատճառով, օրինակ՝ տվյալ ամսում տվյալ տնային տնտեսության սպառման մեծության վրա ազդել են առող-

ջապահական անապատելի ծախսերը, ավտոմեքենայի վթարին առնչվող ծախսերը կամ նրա որևէ անդամի նշանադրությանն առնչվող ծախսերը և այլն: Սպառողական ծախսերի տարբերությունները կարող են պայմանավորված լինել նաև այլ՝ չհաշվառված ազդեցություններով, օրինակ՝ տնային տնտեսությունների անդամների թվով, երեխաների ու ծերերի թվով, բնակության վայրով և այլն: Այս գործոնները մեր դիտարկած սպառում-եկամուտ կախվածության մեջ չեն հաշվառվում:

Այսպիսով՝ տնային տնտեսությունների ամսական եկամտի յուրաքանչյուր արժեքի կհամապատասխանի դրանց ամսական սպառումների տարբեր արժեքների բազմություն, որը կունենա որոշակի բաշխում ու մաթեմատիկական սպասում (այսուհետ՝ մաթ. սպասում) կամ միջին արժեք:

Հետևաբար, եթե փորձենք ձևակերպել տնային տնտեսությունների սպառման կախումը դրանց տնօրինվող եկամտից բնութագրող մաթեմատիկական մոդելը, հավասարումը որպես **դետերմինացված, ֆունկցիոնալ** կախվածություն, այն կունենա հետևյալ տեսքը՝ $C = f(I)$, սակայն սա չի համապատասխանի իրական կյանքին, քանի որ իրականում տնային տնտեսությունների սպառման վրա ազդում են այլ՝ պատահական ու չհաշվառված գործոններ ևս (տե՛ս Գծանկար 2):

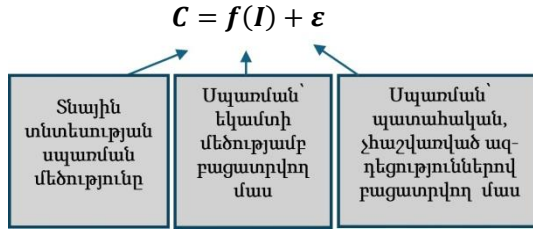
Գծանկար 2. Տնային տնտեսությունների սպառման կախումը եկամուրներից



Տնտեսության մեջ հանդիպող կախվածությունների գերակշիռ մեծամասնությունը իրատեսորեն նկարագրելու ու մոդելավորելու համար դրանք պետք է ներկայացվեն **վիճակագրական կախվածության**

տեսքով՝ ընդգրկելով նաև պատահական, ներառյալ չհաշվառված ազդեցությունները (տե՛ս Գ-ծանկար 3):

Գ-ծանկար 3. Տնային տնտեսությունների սպառողական ծախսերի վիճակագրական կախումը եկամուրներից



Վիճակագրական կախվածության տեսակ է կոռելյացիոն կախվածությունը, որի դեպքում մի մեծության, փոփոխականի յուրաքանչյուր արժեքի համապատասխանում է մյուսի պայմանական մաթ. սպասման որոշակի արժեք կամ որոշակի պայմանական միջին, և հակառակը (X և Y երկու փոփոխականների պայմանական մաթ. սպասումը Y -ի միջին արժեքն է նախապայմանով, որ X -ը ընդունում է խիստ որոշակի x արժեք՝ $X=x$): Կոռելյացիոն կախվածությունը երկկողմանի վիճակագրական կախվածություն է: Այն ընդհանուր դեպքում կարելի է արտահայտել հետևյալ տեսքով՝

$$\mathbb{E}(Y|X = x) = f(x) \text{ և } \mathbb{E}(X|Y = y) = \psi(y): \quad (1)$$

Կոռելյացիոն կախվածության երկկողմանի բնույթը նշանակում է, որ դիտարկվող կապը փոխադարձ է. եթե X -ը ուժեղ կոռելյացիա ունի Y -ի հետ, ապա Y -ը հավասարապես ուժեղ կոռելյացիա ունի X -ի հետ: Այստեղ չկան «կախյալ» ու «բացատրող» փոփոխականների, «պատճառի» ու «հետևանքի» հասկացություններ: Կոռելյացիոն կախվածություններն ուսումնասիրվում են կոռելյացիոն վերլուծության միջոցով, որի հիմնական նպատակը փոփոխականների միջև գծային կապի ուժգնությունն ու ուղղությունը չափելն է: Դրանք չափվում են կոռելյացիայի գործակցի միջոցով:

Ռեգրեսիոն կախվածությունը միակողմանի կոռելյացիոն կախվածությունն է, այսինքն՝ դա էլ է վիճակագրական կախվածություն և ունի ստոխաստիկ կամ հավանականական բնույթ: Ընդհանուր դեպքում այն ունի հետևյալ տեսքը՝

$$\mathbb{E}(Y|X = x) = f(x): \quad (2)$$

Ռեգրեսիոն կախվածությունում մի փոփոխականի (Y-ի) պայմանական մաթ. սպասումը կախված է մյուս փոփոխականի (X-ի) արժեքից: Այստեղ Y-ը կոչվում է **բացատրվող, կախյալ, էնդոգեն կամ երբային փոփոխական**: Իսկ X-ը կոչվում է **բացատրող, էկզոգեն, մուտքային կամ անկախ փոփոխական**: Ռեգրեսիոն կախվածություններն ուսումնասիրվում են ռեգրեսիոն վերլուծության՝ ռեգրեսիայի մոդելների գնահատման միջոցով: Ռեգրեսիոն վերլուծության նպատակը, ի տարբերություն կոռելյացիոն վերլուծության, մի փոփոխականի միջին արժեքի գնահատումը կամ կանխատեսումն է՝ կախված մյուս փոփոխականի(ների) արժեքներից:

Տնտեսական երևույթների կախվածությունները սովորաբար հենց վիճակագրական կախվածություններ են. դրանք ստոխաստիկ բաղադրիչ ունեն, ինչը պայմանավորված է կախյալ փոփոխականի վրա պատահական, չհաշվառված ազդեցություններով:

Ներդիր 1. «Ռեգրեսիա» տերմինի ծագումնաբանությունն ու գործածումը էկոնոմետրիկայում

«Ռեգրեսիա» տերմինը առաջին անգամ կիրառվել է Ֆրենսիս Գալթոնի կողմից 1886 թ. «Ռեգրեսիա դեպի միջին՝ ժառանգական հասակի պարագայում» ուսումնասիրության մեջ: Գալթոնը նկատել էր, որ միջինից ավելի բարձրահասակ ծնողների երեխաները հակված են լինում ավելի կարճահասակ լինելու, իսկ կարճահասակ ծնողների երեխաները՝ ավելի բարձրահասակ: Այս երևույթը նա նկարագրել է որպես **«ռեգրեսիա դեպի միջին»**, այսինքն՝ միջինից էապես տարբերվող հատկությունները հաջորդ սերնդում հակված են մոտենալ միջինին: Այս կենսաբանական դիտարկումն էլ հիմք է դարձել «ռեգրեսիա» տերմինի ի հայտ գալուն: Գալթոնի աշխատանքը հետագայում մաթեմատիկորեն զարգացվեց Կառլ Պիրսոնի ու Յուդին Յուլի կողմից: Վերջիններս ձևակերպեցին գծային ռեգրեսիայի մաթեմատիկական հիմքերը: Այս մոտեցման հիմքում ընկած է փոքրագույն քառակուսիների եղանակը՝ ձևավորված Լեգանդրի ու Գաուսի կողմից դեռևս XIX դարում: Ժամանակի ընթացքում ռեգրեսիան վերածվեց ընդհանուր վիճակագրական մեթոդի, որը կիրառվում է գլխական տարբեր ոլորտներում: Էկոնոմետրիկայում ռեգրեսիան հիմնական գործիք է՝ տնտեսական փոփոխականների միջև գոյություն ունեցող կապի քանակական գնահատման, տնտեսագիտության տեսության դրույթների ստուգման, կանխատեսումներ կատարելու նպատակով:

2. Պարզ կամ միագործոն գծային ռեգրեսիայի տեսական ու էմպիրիկ հավասարումներ ու մոդելներ

Այսպիսով՝ ռեգրեսիոն կախվածությունն ընդհանուր դեպքում ունի (2) բանաձևով ներկայացվող տեսքը՝

$$\mathbb{E}(Y|X = x) = f(x):$$

(2)-ը կոչվում է **ռեգրեսիայի տեսական հավասարում**, իսկ $f(x)$ -ը՝ **ռեգրեսիայի ֆունկցիա**: Ռեգրեսիայի ֆունկցիան նկարագրում է կախյալ փոփոխականի (Y -ի) պայմանական մաթ. սպասման կախվածությունը բացատրող մեկ փոփոխականից՝ X -ից: Ռեգրեսիայի հավասարումը ցույց է տալիս, թե ինչպես է բացատրող կամ անկախ փոփոխականը ազդում կախյալ փոփոխականի սպասվող կամ միջին արժեքի վրա: (2) տեսքի ռեգրեսիայի հավասարումը կոչվում է **պարզ կամ միագործոն ռեգրեսիայի հավասարում**, քանի որ բացատրող փոփոխականը (կամ կախյալ փոփոխականի վրա ազդող գործոնը) մեկն է: Եթե Y -ին բացատրող փոփոխականները մեկից ավելին լինեն, կունենանք բազմակի կամ բազմագործոն ռեգրեսիայի հավասարում, որն ընդհանուր դեպքում կունենա հետևյալ տեսքը՝

$$\mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) = f(x_1, x_2, \dots, x_m): \quad (3)$$

Եթե $f(x)$ ֆունկցիան ունի գծային տեսք, գծային ֆունկցիա է, ապա (2)-ը կոչվում է **գծային պարզ ռեգրեսիայի տեսական հավասարում և ունի հետևյալ տեսքը՝**

$$\mathbb{E}(Y|X = x) = f(x) = \beta_0 + \beta_1 \cdot x: \quad (4)$$

Y և X փոփոխականներն ունեն խիստ որոշակի արժեքներ, որոնք կկոչենք **դիտարկումներ**, և որոնք հավաքագրված կամ մշակված վիճակագրական, էմպիրիկ տվյալներ են:

Վերադառնանք տնային տնտեսությունների սպառողական ծախսերի ու տնօրինվող եկամտի օրինակին: Արդեն իսկ նշվել է, որ նույն ամսական եկամտով ունեցող տնային տնտեսությունները ճշգրիտ նույն ամսական սպառումը չեն ունենա. դրանց սպառումները կտարբերվեն: Այսինքն՝ տնային տնտեսությունների եկամտի յուրաքանչյուր արժեքի կհամապատասխանի դրանց ամսական սպառումների արժեքների բազմություն, որը կունենա որոշակի բաշխում, ինչպես նաև մաթ.սպասում կամ միջին: Ուրեմն սա վիճակագրական կախվածություն է և քանի որ միակողմանի է, այն ռեգրեսիոն կախ-

վածություն է, որի հավասարումն է (2)-ը: Սակայն տվյալ եկամտի դեպքում ամեն մի տնային տնտեսության սպառման իրական, փաստացի արժեքը կտարբերվի այդ մաթ. սպասումից կամ միջինից՝ կախված որոշակի գործոններից, որոնք պատահական են, ներառյալ նաև տվյալ ռեգրեսիայի հավասարման մեջ հաշվի չառնված, չներառված գործոնները:

Կամ ենթադրենք՝ որոշել եք վաճառել ձեր ավտոմեքենան: Դրա համար պետք է որոշեք ավտոմեքենայի վաճառքի գինը: Դուք նայում եք շուկայական հայտարարությունները՝ գտնելով, թե ձեր ավտոմեքենայի բնութագրիչներն ունեցող ավտոմեքենաների համար միջինում ինչ վաճառքի գին են սահմանում դրանց տերերը, կամ միջինում ինչպես է կախված ավտոմեքենայի վաճառքի համար սահմանված գինը դրա բնութագրիչներից: Իրականում դուք մտքում գնահատում եք ռեգրեսիայի (3) տեսքի հավասարում: Սակայն գործարքն իրականանում է ոչ թե ձեր որոշած վաճառքի միջին գնով, այլ դրանից տարբերվող գնով, ինչը պայմանավորված է լինում որոշակի, ձեր կողմից չհաշվառված (օրինակ՝ մեքենայի առջևի ապակին վնասված է, կամ թափքի վրա խազվածքներ կան, կամ անվաղողերը նոր չեն և այլն) ու պատահական գործոններով (օրինակ՝ գործարքը կատարելու օրը փոխարժեքի կուրսը կտրուկ ու անսպասելի փոխվեց, կամ պարզվեց, որ գնորդը ձեզ ծանոթ է, կամ գնորդը շատ շտապում էր և այլն):

Այսպիսով՝ կախյալ փոփոխականի **իրական, փաստացի** արժեքը կոնկրետ դիտարկման համար հիմնականում տարբերվում է, շեղված է դրա պայմանական միջին կամ սպասվող արժեքից ինչ-որ պատահական՝ դրական կամ բացասական ε մեծությամբ: Այս հանգամանքը կարելի է ներկայացնել այսպես՝

$$y = E(Y|X = x) + \varepsilon = f(x) + \varepsilon: \quad (5)$$

(5)-ում y -ը կախյալ փոփոխականի իրական, փաստացի արժեքն է (մեր օրինակներում՝ տվյալ տնային տնտեսության փաստացի սպառողական ծախսերը կամ ավտոմեքենայի փաստացի վաճառքի գինը), $E(Y|X = x)$ -ը՝ կախյալ Y փոփոխականի միջինը կամ մաթ. սպասումը, երբ անկախ կամ բացատրող X փոփոխականն ընդունում է տվյալ x արժեքը, իսկ ε -ը կախյալ փոփոխականի փաստացի արժեքի՝ այդ մաթ. սպասումից կամ միջինից ունեցած շեղումն է: **ε -ն ունի պատահական բնույթ, պատահական մեծություն է:**

(5)-ը կկոչենք **պարզ կամ միագործոն ռեգրեսիայի տեսական մոդել**, իսկ ε -ը, որով ռեգրեսիայի տեսական մոդելը տարբերվում է

տեսական հավասարումից, կոչվում է պատահական շեղում: Այն արտահայտում է կախյալ փոփոխականի վրա իրականում գոյություն ունեցող պատահական, ներառյալ նաև չհաշվառված ազդեցությունները:

Ռեգրեսիայի տեսական մոդելում X բացատրող փոփոխականը դետերմինացված մեծություն է, այն հավանականական, պատահական բնույթ չունի, իսկ ε -ը պատահական, ստոխաստիկ կամ հավանականական բնույթ ունի: Եվ քանի որ կախյալ Y փոփոխականը դրանց գումարն է, այն ևս ունի պատահական բնույթ:

Ենթադրենք՝ Y և X փոփոխականների համար ունենք n թվով տվյալներ կամ դիտարկումներ: Յուրաքանչյուր դիտարկման համար պարզ ռեգրեսիայի տեսական մոդելը կունենա հետևյալ տեսքը՝

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n: \quad (6)$$

Եթե $f(x)$ ֆունկցիան ունի գծային տեսք, գծային ֆունկցիա է, ապա (5)-(6)-ը կկոչվի **պարզ կամ միագործոն գծային ռեգրեսիայի տեսական մոդել:**

Պարզ գծային ռեգրեսիայի տեսական մոդելն ունի հետևյալ տեսքը՝

$$y = E(Y|X = x) + \varepsilon = f(x) + \varepsilon = \beta_0 + \beta_1 \cdot x + \varepsilon: \quad (7)$$

Այստեղ ε -ը պատահական շեղումներն են, իսկ β_0 -ն և β_1 -ը **ռեգրեսիայի տեսական մոդելի պարամետրերն են:**

Նշենք, որ ռեգրեսիայի տեսական մոդելում բացատրող փոփոխականը, ինչպես նաև β_0 և β_1 **պարամետրերը դետերմինացված մեծություններ են:** Դրան հակառակ՝ կախյալ փոփոխականն ունի ստոխաստիկ, հավանականական բնույթ, քանի որ այն կախված է նաև պատահական ε -ից:

Էկոնոմետրիկ մոդելավորման տեսանկյունից փոփոխականների գծային կախվածության ընտրությունը Էկոնոմետրիկ մոդելի (7) տեսքով Էկոնոմետրիկ մոդելի **հատկորոշման փուլն է:** Այս դեպքում ունենք Էկոնոմետրիկ մոդելի գծային հատկորոշում:

Ներդիր 2. Երկու փոփոխականների գծային կախվածություն կամ գծային ֆունկցիա

Գծային կախվածությունը փոփոխականների միջև գծային կապը նկարագրող ֆունկցիա է, որը երկու փոփոխականի դեպքում ունի հետևյալ տեսքը՝ $y = \beta_0 + \beta_1 \cdot x$: Ֆունկցիան որոշված է բոլոր իրական թվերի բազմության վրա, β_0 -ն և β_1 -ը թվեր են, հաստատուններ: Այս ֆունկցիայի գրաֆիկը ուղիղ գիծն է երկչափ հարթության վրա: β_0 -ն կոչվում է ազատ անդամ կամ բարձրության գործակից (intercept) և բնութագրում է ուղիղ գծի «բարձրությունը»: Գրաֆիկորեն դա ուղղի ու օրդինատների (Oy) առանցքի հատման կետի օրդինատն է և ցույց է տալիս, թե ինչի է հավասար կախյալ y փոփոխականը, երբ $x=0$: β_1 -ը կոչվում է անկյունային կամ թեքության գործակից (slope coefficient) և բնութագրում է ուղիղ գծի «թեքությունը»: Գծային ֆունկցիայով բնութագրվող ուղիղ գիծը արեւմտյան (Ox) առանցքի հետ կազմում է α անկյուն, որի տանգենսը հավասար է β_1 -ի՝ եղ $\alpha = \beta_1$: Միաժամանակ β_1 -ը գծային ֆունկցիայի ածանցյալն է ըստ x-ի՝ $\beta_1 = \frac{dy}{dx} = \frac{\Delta y}{\Delta x}$, և ցույց է տալիս, թե քանի միավորով է փոխվում y-ը x-ի միավոր փոփոխության դեպքում: Դրական թեքությունը ցույց է տալիս աճի միտում, բացասականը՝ նվազման, իսկ զրոյական թեքությունը՝ (Ox) առանցքին զուգահեռ հորիզոնական գիծ և x-ի և y-ի միջև կախման բացակայություն:

Էկոնոմետրիկայի խնդիրն է, հիմնվելով Y և X փոփոխականների էմպիրիկ տվյալների վրա, գտնել այդ փոփոխականների գծային կախվածության քանակական արտահայտությունը: Դա նշանակում է, որ (7) տեսքի մոդելում, ունենալով Y և X փոփոխականների համար n թվով (x_i , y_i) դիտարկումներ (ժամանակի միևնույն, ամրագրված պահի համար, այսինքն՝ տարածական տվյալներ), պետք է վիճակագրական մեթոդներով գնահատենք β_0 և β_1 անհայտ պարամետրերը: Էկոնոմետրիկ մոդելավորման տեսանկյունից սա **պարամետրացման փուլն է**:

Դիտարկենք էկոնոմետրիկ մոդելավորման այս փուլերը հետևյալ օրինակի վրա: Ենթադրենք՝ Քեյնսի սպառման ֆունկցիան հատկորոշվել է որպես պարզ գծային ռեգրեսիայի տեսական մոդել: Այդ տեսական մոդելը կունենա հետևյալ տեսքը՝

$$C = \beta_0 + \beta_1 I + \varepsilon,$$

որում C-ն տնային տնտեսությունների սպառողական ծախսերն են (կախյալ փոփոխական), I-ն՝ տնօրինվող եկամուտը (անկախ կամ բացատրող փոփոխական), β_0 -ն բնութագրում է ավտոնոմ սպառումը, իսկ β_1 -ը՝ սպառման սահմանային հակվածությունը:

Տվյալ դեպքում հետազոտության օբյեկտը ՀՀ-ում բնակվող տնային տնտեսությունն է, որի հատկորոշված սպառման ֆունկցիան պետք է պարամետրացվի, այսինքն՝ բանակապես գնահատվի ՀՀ-ում տնային տնտեսությունների ավտոնոմ սպառումն ու սպառման սահմանային հակվածությունը:

Էկոնոմետրիկ մոդելավորման հաջորդ փուլի՝ պարամետրացման իրականացման համար անհրաժեշտ է էմպիրիկ տվյալների առկայություն: Դրա համար պետք է ունենալ ՀՀ բոլոր, մոտավորապես 840,000 տնային տնտեսությունների տնօրինվող եկամուտների ու սպառողական ծախսերի տվյալները:

Հետազոտվող օբյեկտների ամբողջ բազմությունը կոչվում է օբյեկտների **գլխավոր համախմբություն (population, universe, այսուհետ՝ ԳՀ)**, օրինակ՝ ՀՀ բոլոր տնային տնտեսությունները, ՀՀ ամբողջ բնակչությունը, ՀՀ-ում բոլոր զբաղվածները, ՀՀ-ում գործող բոլոր միջին ու փոքր ձեռնարկությունները և այլն: Եթե ունենանք ԳՀ բոլոր օբյեկտների (մեր օրինակում՝ ՀՀ բոլոր տնային տնտեսությունների) տվյալները, հնարավոր կլինի վիճակագրական մեթոդներով գնահատել գծային ռեգրեսիայի **տեսական** մոդելը, այսինքն՝ գտնել տեսական մոդելի անհայտ β_0 և β_1 պարամետրերը: Գրաֆիկորեն սա կնշանակի, որ կոորդինատային հարթության վրա կորոշվի երկու փոփոխականների գծային կախումը բնութագրող ուղիղ գիծը: Սակայն բնական է, որ ԳՀ-ները սովորաբար մեծ են լինում, և դրանց յուրաքանչյուր օբյեկտին համապատասխանող ու մեզ հետաքրքրող x_i և y_i հատկանիշների մասին տվյալներ ունենալը գործնականում հնարավոր չի լինում: Այդ իսկ պատճառով էկոնոմետրիկայում հաճախ օգտագործվում են գլխավոր համախմբությունից կատարված ընտրանքի կամ **նմուշի (sample) վիճակագրական, էմպիրիկ** տվյալները: Եթե նմուշը ներկայացուցչական է և ունի փոքր սխալ, դրա ուսումնասիրության արդյունքները կարելի է տարածել ԳՀ-ի վրա:

Ներդիր 3. Հեղափոխության օբյեկտների գլխավոր համախմբություն ու նմուշ

Գլխավոր համախմբությունը (ԳՀ) ուսումնասիրվող օբյեկտների ամբողջ բազմությունն է: Դրա միավորը ուսումնասիրվող օբյեկտների ամբողջ բազմության առանձին կոնկրետ միավորն է: **Ընտրանքը կամ նմուշը** ԳՀ միավորներից ընտրված ավելի փոքր ենթաբազմություն է, որը ներկայացուցչական է ամբողջ ԳՀ-ի համար և օգտագործվում է ամբողջ ԳՀ-ի մասին եզրակացություններ անելու համար:

Նմուշը պետք է ճշգրտորեն ներկայացնի ԳՀ կառուցվածքն ու բազմազանությունը, այսինքն՝ պետք է լինի ներկայացուցչական ԳՀ-ի համար: Նմուշի ներկայացուցչականությունը մեծապես կախված է նմուշառման մեթոդից ու ընտրանքի չափից (այն անվանում են ընտրանքի ծավալ):

Սովորաբար նմուշները բաժանվում են երկու խմբի՝ **պատահական** ու **ոչ պատահական**: Եթե պատահական նմուշառումը ԳՀ յուրաքանչյուր միավորի համար ապահովում է ընտրվելու ոչ զրոյական, հայտնի հավանականություն, ապա ոչ պատահական նմուշում ԳՀ որոշ միավորներ ունեն ընտրվելու զրոյական հավանականություն, իսկ միավորների ընտրվելու հավանականություններն անհայտ են: Պատահական նմուշառման հիմնական մեթոդներն են՝ պարզ պատահական, սխտեմատիկ, ստրատիֆիկացված, կլաստերային, բազմափուլ նմուշառումները: Ոչ պատահական նմուշառման հիմնական մեթոդներն են՝ ձնագնդի, հարմարության, քվոտային, տիպական և այլն:

Նմուշառման նպատակը ԳՀ անհայտ պարամետրերի գնահատումն է նմուշից ստացվող վիճակագրական գնահատականների միջոցով: ԳՀ որևէ պարամետրի իրական, սակայն անհայտ արժեքի ու դրա՝ նմուշից ստացվող արժեքի միջև եղած տարբերությունը կոչվում է նմուշառման սխալ: Այն առաջանում է ամբողջ ԳՀ-ի փոխարեն նմուշը դիտարկելու հետևանքով և նվազում է ընտրանքի չափի աճի հետ: Պատահական նմուշառման դեպքում նմուշառման սխալը հնարավոր է գնահատել: Ոչ պատահական նմուշառման դեպքում նմուշառման սխալը հաշվարկել կամ գնահատել հնարավոր չէ:

Այսպիսով՝ խնդիր է առաջանում, ունենալով նմուշից ստացված վիճակագրական տվյալներ (այսուհետ այդ տվյալները կանվանենք էմպիրիկ տվյալներ), գնահատել X -ից Y -ի կախվածության ռեգրեսիայի մոդելը:

Էկոնոմետրիկ մոդելավորման ընթացքում էմպիրիկ տվյալների հիման վրա գնահատվում են գծային ռեգրեսիայի **էմպիրիկ** հավասարումն ու մոդելը: Նմուշից ստացված տվյալները չեն ծածկում Q - Z -ն ամբողջապես. դրա մի մասն են: Միայն նմուշից ստացված էմպիրիկ տվյալների հիման վրա պարամետրիզացումն իրականացնելիս վիճակագրական մեթոդներով ստացվում են տեսական մոդելի β_0 և β_1 անհայտ պարամետրերի **վիճակագրական գնահատականները**, որոնք տարբերվում են տեսական մոդելի β_0 և β_1 պարամետրերից: Այդ վիճակագրական գնահատականները այսուհետ կանվանենք **էմպիրիկ հավասարման գործակիցներ** և կնշանակենք այսպես՝ $\widehat{\beta}_0$ և $\widehat{\beta}_1$: Էմպիրիկ տվյալների հիման վրա տեսական մոդելի β_0 և β_1 պարամետրերի վիճակագրական գնահատականների՝ $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների ստացումը հենց գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման գնահատումն է:

Գծային պարզ ռեգրեսիայի էմպիրիկ հավասարումն ունի հետևյալ տեսքը՝

$$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x: \quad (8)$$

n թվով էմպիրիկ տվյալների, կամ n թվով դիտարկումների առկայության դեպքում ամեն մի դիտարկման համար կունենանք՝

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i, \quad i=1,2,\dots,n: \quad (9)$$

(8)-(9)-ում.

ա) $\widehat{y}$ -ը տեսական մոդելի Y կախյալ փոփոխականի պայմանական մաթ. սպասման՝ $E(Y|X=x) = f(x) = \beta_0 + \beta_1 \cdot x$ -ի վիճակագրական գնահատականն է: Այն կկոչենք կախյալ փոփոխականի գնահատված արժեք: $\widehat{y}$ -ը կախյալ փոփոխականի՝ նմուշային, **էմպիրիկ** տվյալների հիման վրա ստացված պայմանական մաթ. սպասման արժեքն է՝

$$\widehat{y} = E(y_{\text{ս}} | x = x_{\text{ս}}): \quad (10)$$

Յուրաքանչյուր դիտարկման համար կունենանք.

$$\widehat{y}_i = E(y_{\text{ս}} | x = x_i^{\text{ս}}):$$

բ) $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները տեսական մոդելի β_0 և β_1 պարամետրերի վիճակագրական գնահատականներն են, որոնք ստացվել են էմպիրիկ տվյալների հիման վրա գծային ռեգրեսիայի էմպիրիկ հավասարումը գնահատելու արդյունքում: **Գ-րանք պատահական բնույթ ունեն:**

Գծային պարզ ռեգրեսիայի էմպիրիկ մոդելն ունի հետևյալ տեսքը՝

$$y = \widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\varepsilon} = \widehat{y} + \widehat{\varepsilon}: \quad (11)$$

Կամ յուրաքանչյուր դիտարկման համար՝

$$y_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i + \widehat{\varepsilon}_i = \widehat{y}_i + \widehat{\varepsilon}_i, \quad i = 1, 2, \dots, n, \quad (12)$$

որում $\widehat{\varepsilon}$ -ը տեսական մոդելի ε պատահական շեղման վիճակագրական գնահատականն է և կոչվում է պատահական մնացորդ: (12)-ից ու (9)-ից ստացվում է, որ յուրաքանչյուր դիտարկման համար՝

$$\widehat{\varepsilon}_i = y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i) = y_i - \widehat{y}_i: \quad (13)$$

(13)-ում y_i -ն y կախյալ փոփոխականի i -րդ դիտարկմանը համապատասխանող փաստացի նմուշային, էմպիրիկ արժեքն է, իսկ $\widehat{y}_i$ -ը՝ դրա՝ ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված արժեքը: Այսինքն՝ $\widehat{\varepsilon}$ պատահական մնացորդը կախյալ փոփոխականի **փաստացի էմպիրիկ տվյալի ու դրա՝ ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված արժեքի տարբերությունն է:**

Նշենք, որ տեսական մոդելում պատահական ε_i շեղումը կախյալ փոփոխականի i -րդ դիտարկմանը համապատասխանող արժեքի ու ռեգրեսիայի տեսական հավասարմամբ որոշվող Y -ի սպասվող արժեքի միջև տարբերությունն է՝ $\varepsilon_i = Y_i - E(Y_i | X = x_i)$: ε_i պատահական շեղումը տեսական հասկացություն է, որը երբեք չի կարող դիտվել՝ ի տարբերություն պատահական մնացորդի, որը հաշվարկվում է յուրաքանչյուր դիտարկման համար ամեն անգամ, երբ ռեգրեսիայի էմպիրիկ հավասարում է գնահատվում:

Ներդիր 4. Նմուշից սրացվող վիճակագրական գնահատականներն ու դրանց հատկությունները

Նմուշառման նպատակը Գ-Հ անհայտ պարամետրերի գնահատումն է այդ պարամետրերի՝ նմուշից ստացվող վիճակագրական գնահատականների ստացման միջոցով: Տարբերում են նմուշից ստացվող վիճակագրական գնահատականների երկու տեսակ՝ կեն-

տալին ու միջակայքային:

Կետային գնահատականը $Q-Z$ -ի անհայտ պարամետրը մոտարկող մեկ արժեք է, որը հիմնված է նմուշային տվյալների վրա, օրինակ՝ տնային տնտեսությունների՝ նմուշից ստացվող միջին սպառումը՝ որպես Z բոլոր տնային տնտեսությունների միջին սպառման գնահատական: Միևնույն $Q-Z$ -ից կարելի է ընտրել տարբեր նմուշներ: Դրանցից յուրաքանչյուրը կունենա իր վիճակագրական գնահատականները, որոնք միմյանցից կտարբերվեն, քանի որ կախված են նմուշից ու այդ նմուշից ստացված տվյալներից: Քանի որ նմուշները կատարված են պատահականորեն, դրանցից ստացվող վիճակագրական գնահատականները պատահական մեծություններ են, որոնք ունեն արժեքների բազմություն, բաշխում ու թվային բնութագրիչներ (միջին, վարիացիա կամ ցրվածք և այլն):

$Q-Z$ անհայտ պարամետրը նշանակենք θ -ով, իսկ դրա՝ նմուշից ստացված վիճակագրական գնահատականը՝ $\hat{\theta}$ -ով: $Q-Z$ -ի θ պարամետրի $\hat{\theta}$ վիճակագրական գնահատականը որպեսզի որակյալ լինի, պետք է ունենա որոշակի հատկություններ.

Անշեղություն $\hat{\theta}$ վիճակագրական գնահատականը կոչվում է անշեղ, եթե դրա մաթ. սպասումը հավասար է անհայտ պարամետրի իրական արժեքին՝ $E(\hat{\theta}) = \theta$:

Արդյունավետություն $\hat{\theta}$ վիճակագրական գնահատականը կոչվում է արդյունավետ, եթե կամայական նույնատիպ այլ $\tilde{\theta}$ վիճակագրական գնահատականի համեմատ այն ունի ամենավտրք ցրումն իր միջինի շուրջը կամ վարիացիան՝ $Var(\hat{\theta}) = \min Var(\tilde{\theta})$:

Ունակություն $\hat{\theta}$ վիճակագրական գնահատականը կոչվում է ունակ, եթե նմուշի ծավալի մեծացման հետ այն ձգտում է անհայտ θ պարամետրին (ըստ հավանականության)՝ $plim_{n \rightarrow \infty} \hat{\theta} = \theta$:

Նշենք, որ կետային գնահատականը, եթե նույնիսկ ունի վերը բերված հատկությունները և որակյալ գնահատական է, այնուամենայնիվ միայն մոտարկում է անհայտ պարամետրը:

Միջակայքային գնահատականը որոշում է արժեքների (θ_1 , θ_2) միջակայք, որում որոշակի մեկին մոտ Y հավանականությամբ ընկած է անհայտ պարամետրի իրական արժեքը: Այդ միջակայքը կոչվում է վստահելիության միջակայք, իսկ Y -ն կոչվում է վստահելիության հավանականություն:

Վերը ասվածի ամփոփումը բերված է Աղյուսակ 1-ում:

Աղյուսակ 1. Պարզ գծային ռեգրեսիայի տեսական ու էմպիրիկ հավասարումների ու մոդելների բնութագրող տեսքերը:

Գծային պարզ ռեգրեսիայի տեսական հավասարում	$\mathbb{E}(Y X = x) = f(x) = \beta_0 + \beta_1 \cdot x$
Գծային պարզ ռեգրեսիայի տեսական մոդել	$y = \mathbb{E}(Y X = x) + \varepsilon = f(x) + \varepsilon = \beta_0 + \beta_1 x + \varepsilon$
Գծային պարզ ռեգրեսիայի էմպիրիկ հավասարում	$\hat{y} = \mathbb{E}(y_{\text{սւ}} x = x_{\text{սւ}}) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x$
Գծային պարզ ռեգրեսիայի էմպիրիկ մոդել	$y = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x + \varepsilon = \hat{y} + \varepsilon$

Այսպիսով՝ ռեգրեսիայի էմպիրիկ հավասարման գնահատման արդյունքում ստացվում են տեսական մոդելի վիճակագրական կետային գնահատականները, որոնք ամփոփված են Աղյուսակ 2-ում:

Աղյուսակ 2. Պարզ գծային ռեգրեսիայի տեսական մոդելի բաղադրիչներն ու ռեգրեսիայի էմպիրիկ մոդելում դրանց վիճակագրական գնահատականները

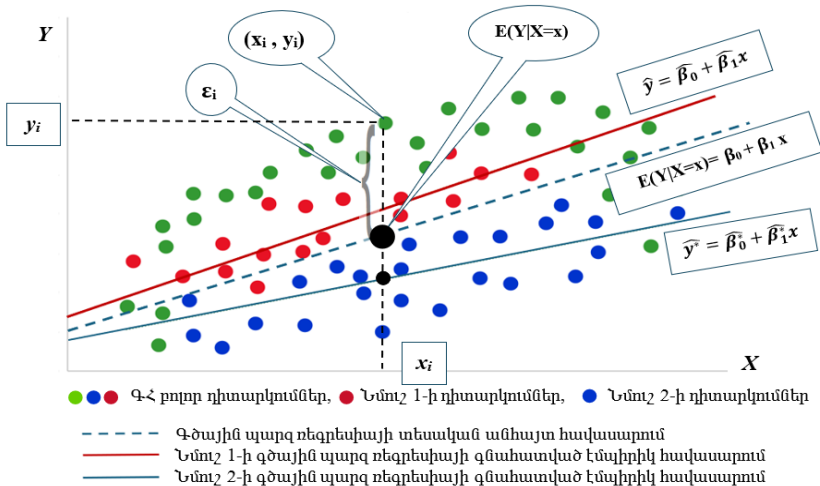
Տեսական մոդել	Էմպիրիկ մոդելում տեսական մոդելի վիճակագրական գնահատականներ
Տեսական մոդելի β_0 և β_1 պարամետրեր (դետերմինացված բնույթի են)	Էմպիրիկ մոդելի $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցներ (պատահական բնույթի են)
Y կախյալ փոփոխականի պայմանական մաթ. սպասում՝ $\mathbb{E}(Y X = x) = \beta_0 + \beta_1 x$	Կախյալ փոփոխականի գնահատված արժեք՝ $\hat{y} = \mathbb{E}(y_{\text{սւ}} x = x_{\text{սւ}}) = \widehat{\beta}_0 + \widehat{\beta}_1 x$
Տեսական մոդելի ε պատահական շեղում՝ $\varepsilon_i = Y_i - \mathbb{E}(Y_i X = x_i)$	Էմպիրիկ մոդելի պատահական մնացորդ՝ $\varepsilon_i = y_i - \mathbb{E}(y_{\text{սւ}} x = x_i^{\text{սւ}}) = y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i) = y_i - \hat{y}_i$

Եթե նմուշի n թվով դիտարկումների x և y հատկանիշներին վերաբերող էմպիրիկ տվյալները տեղադրենք կոորդինատային հարթության վրա, կստանանք (x_i, y_i) կոորդինատներով n հատ կետ: Այդ կետերի ամբողջությունը կամ ամպը կոչվում է կոռելյացիայի դաշտ կամ ցրման գրաֆիկ: Կոռելյացիայի դաշտը տվյալների գրաֆիկական ներկայացում է, որը ցույց է տալիս երկու փոփոխականների միջև եղած կապը: Կոռելյացիայի դաշտը հնարավորություն է տալիս

պատկերի վրա տեսնելու փոփոխականների միջև եղած կապի բնույթը, ուղղությունն ու ուժգնությունը:

Եթե տվյալ $Q-Z$ -ից կատարվեն տարբեր նմուշառումներ, դրանց համար ռեգրեսիայի տարբեր էմպիրիկ հավասարումներ կգնահատվեն: Տեսական հավասարումը միակն է, որն իրականում անհայտ է, քանի որ $Q-Z$ բոլոր տվյալներն անհասանելի են: Գրաֆիկորեն դա կնշանակի, որ տարբեր նմուշներից ստացված տվյալների հիման վրա կգնահատվեն երկու փոփոխականների գծային կապը բնութագրող տարբեր ուղիղ գծեր (տե՛ս Գծանկար 4):

Գծանկար 4. Գլխավոր համահարքության ու նմուշների դիսպրիսիոնների կետերի ամպը և գծային պարզ ռեգրեսիայի տեսական ու էմպիրիկ հավասարումները



Էկոնոմետրիկ մոդելավորման խնդիրն է էմպիրիկ տվյալների հիման վրա քանակապես գնահատել էմպիրիկ հավասարումն այնպես, որ այն հնարավորինս ճշգրիտ ներկայացնի, մոտարկի անհայտ տեսական հավասարումը: Գրաֆիկորեն դա կնշանակի, որ մեր խնդիրն է էմպիրիկ տվյալներից ստացված կետերի ամպը մոտարկել ուղիղ գծով այնպես, որ այն հնարավորինս մոտ լինի $Q-Z$ տեսական հավասարումը ներկայացնող անհայտ ուղիղ գծին:

Միևնույն ժամանակ խնդիր է առաջանում էմպիրիկ դիսպրիսիոններից ստացված կետերի ամպը մոտարկելու ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված ուղիղ գծով այնպես, որ այն հնարավորինս

ճշգրիտ ներկայացնի հենց այդ էմպիրիկ տվյալների միջև եղած գծային կախվածությունը:

Այսպիսով՝ էմպիրիկ տվյալների հիման վրա ռեգրեսիայի էմպիրիկ հավասարումը գնահատելիս առաջանում է երկու մակարդակի խնդիր՝

1-ին կամ վերին մակարդակի խնդիր՝ էմպիրիկ տվյալների հիման վրա ռեգրեսիայի էմպիրիկ հավասարումը գնահատել այնպես, որ այն լավագույն ձևով մոտարկի ռեգրեսիայի տեսական հավասարումը, որն անհայտ է:

2-րդ կամ ներքևի մակարդակի խնդիր՝ էմպիրիկ տվյալների հիման վրա էմպիրիկ հավասարումը գնահատել այնպես, որ այն լավագույն ձևով մոտարկի այդ տվյալների բազմությունը, կամ, որ նույնն է, գնահատված ուղիղ գիծը լավագույն ձևով մոտարկի էմպիրիկ տվյալների կետերի ամպը:

2-րդ մակարդակի խնդիրը լուծելու համար գոյություն ունի վիճակագրական մեթոդ, որը կոչվում է փոքրագույն քառակուսիների եղանակ (ՓՔԵ): Այն հնարավորություն է տալիս գտնելու այնպիսի ուղիղ գիծ կամ ստանալու այնպիսի $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականներ, որ էմպիրիկ տվյալների (x_i, y_i) կետերի ցրումը գնահատված գծի շուրջը լինի փոքրագույնը:

3. Պարզ կամ միագործոն գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատումը փոքրագույն քառակուսիների եղանակով (ՓՔԵ)

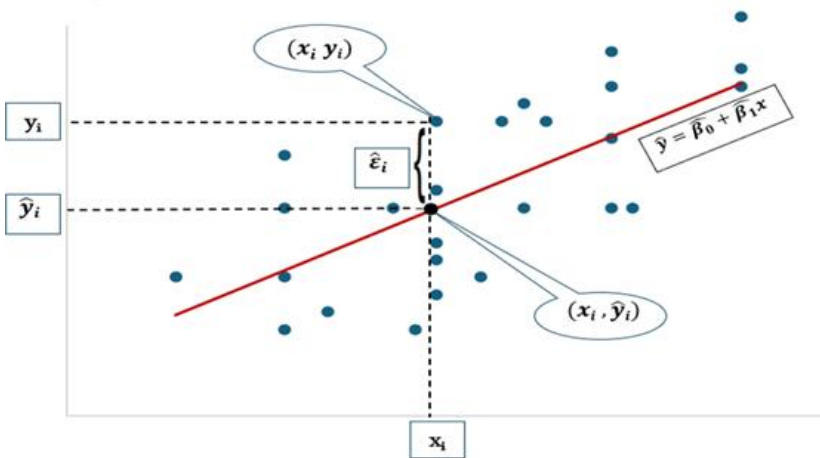
Այս գլխի նախորդ ենթավերնագրում արդեն նշվել է, որ գծային պարզ ռեգրեսիայի (11) էմպիրիկ մոդելում ε -ը տեսական մոդելի ε պատահական շեղման վիճակագրական գնահատականն է, պատահական մնացորդը, որը կարելի է արտահայտել հետևյալ բանաձևով՝ $\varepsilon_i = y_i - (\widehat{\beta}_0 + \widehat{\beta}_1 x_i) = y_i - \widehat{y}_i$: ε -ը կախյալ փոփոխականի նմուշային տվյալի ու ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված արժեքի տարբերությունն է:

Ակնհայտ է, որ որքան փոքր լինի այդ տարբերությունը, այնքան էմպիրիկ տվյալների ընդհանուր ցրվածությունը գնահատված ուղիղի շուրջը, այն է՝ ε պատահական մնացորդների հանրագումարը փոքր

կլինի, և գնահատված ուղիղ գիծն ավելի ճշգրտորեն կմոտարկի կետերի ամպը (տե՛ս Գծանկար 5):

Խնդիր է առաջանում ռեգրեսիայի էմպիրիկ հավասարումը գնահատելու այնպես, որ պատահական մնացորդների հանրագումարը լինի փոքրագույնը: Դա վերևում նշված 2-րդ՝ ներքևի մակարդակի խնդրի լուծումն է: Այդ նպատակով պետք է գտնել այնպիսի $\widehat{\beta}_0$ և $\widehat{\beta}_1$, որ՝ $\sum_{i=1}^n \widehat{\varepsilon}_i = \sum_{i=1}^n (y_i - \widehat{y}_i) \rightarrow \min$:

Գծանկար 5. Էմպիրիկ տվյալների կետերի ամպը (կոռելյացիայի դաշտը) ու գծային պարզ ռեգրեսիայի էմպիրիկ հավասարումը



Բայց քանի որ $\widehat{\varepsilon}$ պատահական մնացորդները կարող են լինել թե՛ բացասական, թե՛ դրական, ապա որպեսզի գումարելիներն իրար չմարեն, վերցվում են պատահական մնացորդների քառակուսիները: Այսինքն՝

$$\sum_{i=1}^n \widehat{\varepsilon}_i^2 = \sum_{i=1}^n (y_i - \widehat{y}_i)^2:$$

Ենթադրենք որոշված է էկոնոմետրիկ մոդելի գծային հատկորոշում՝ միագործոն գծային ռեգրեսիայի մոդել: ՓԲԵ-ն խնդիր է դնում էմպիրիկ տվյալների հիման վրա գնահատելու պարզ գծային ռեգրեսիայի $\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$ էմպիրիկ հավասարումն այնպես, որ պատահական մնացորդների քառակուսիների գումարը լինի փոքրագույնը.

$$Q = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min: \quad (14)$$

Q-ն ֆունկցիա է $\widehat{\beta}_0$ և $\widehat{\beta}_1$ -ից, որոնք ռեգրեսիայի էմպիրիկ հավասարման անհայտ գործակիցներն են և պետք է որոշվեն (14)-ից:

$$Q = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 \rightarrow \min: \quad (15)$$

Q-ն ուռուցիկ ֆունկցիա է, այն մաս անընդհատ է և ներքևից սահմանափակ է զրոյով, հետևաբար ունի մինիմում: Գտնենք այդ մինիմումը՝ դրա մասնակի ածանցյալները, ըստ $\widehat{\beta}_0$ և $\widehat{\beta}_1$ -ի, հավասարեցնելով զրոյի: Կատանանք նորմալ հավասարումների համակարգը՝

$$\begin{cases} \frac{\partial Q}{\partial \widehat{\beta}_0} = 0 \\ \frac{\partial Q}{\partial \widehat{\beta}_1} = 0 \end{cases} \rightarrow \begin{cases} \frac{\partial \left(\sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 \right)}{\partial \widehat{\beta}_0} = 0 \\ \frac{\partial \left(\sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2 \right)}{\partial \widehat{\beta}_1} = 0 \end{cases} \rightarrow \begin{cases} -2 \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) = 0 \\ -2 \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) x_i = 0 \end{cases}$$

Լուծենք առաջին հավասարումը՝

$$\begin{aligned} -2 \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) &= 0 \rightarrow \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) = \sum_{i=1}^n y_i - \\ &- n\widehat{\beta}_0 - \widehat{\beta}_1 \sum_{i=1}^n x_i = 0 \rightarrow n\widehat{\beta}_0 = \sum_{i=1}^n y_i - \widehat{\beta}_1 \sum_{i=1}^n x_i \rightarrow \\ \rightarrow \widehat{\beta}_0 &= \frac{1}{n} \sum_{i=1}^n y_i - \widehat{\beta}_1 \frac{1}{n} \sum_{i=1}^n x_i \rightarrow \boxed{\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}} \quad (16) \end{aligned}$$

Լուծենք երկրորդ հավասարումը՝

$$\begin{aligned} -2 \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) x_i &= 0 \rightarrow \sum_{i=1}^n (y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i) x_i = \\ &= \sum_{i=1}^n y_i x_i - \widehat{\beta}_0 \cdot \sum_{i=1}^n x_i - \widehat{\beta}_1 \cdot \sum_{i=1}^n x_i^2 = 0 \rightarrow \\ &\sum_{i=1}^n y_i x_i = \widehat{\beta}_0 \sum_{i=1}^n x_i + \widehat{\beta}_1 \sum_{i=1}^n x_i^2 \end{aligned}$$

Վերջին հավասարման երկու կողմերը բաժանենք n -ի: Կստանանք՝

$$\frac{1}{n} \sum_{i=1}^n x_i y_i = \widehat{\beta}_0 \frac{1}{n} \sum_{i=1}^n x_i + \widehat{\beta}_1 \frac{1}{n} \sum_{i=1}^n x_i^2 \rightarrow$$

$$\overline{x \cdot y} = \widehat{\beta}_0 \cdot \bar{x} + \widehat{\beta}_1 \cdot \overline{x^2} \rightarrow \widehat{\beta}_1 = \frac{\overline{x \cdot y} - \widehat{\beta}_0 \cdot \bar{x}}{\overline{x^2}}$$

Տեղադրենք $\widehat{\beta}_0$ -ի արժեքը (16)-ից: Կստանանք՝

$$\widehat{\beta}_1 = \frac{\overline{x \cdot y} - \widehat{\beta}_0 \cdot \bar{x}}{\overline{x^2}} = \frac{\overline{x \cdot y} - (\bar{y} - \widehat{\beta}_1 \cdot \bar{x}) \bar{x}}{\overline{x^2}} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y} + \widehat{\beta}_1 \cdot (\bar{x})^2}{\overline{x^2}} \rightarrow$$

$$\widehat{\beta}_1 \cdot \left(1 - \frac{(\bar{x})^2}{\overline{x^2}}\right) = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2}} \rightarrow$$

$$\widehat{\beta}_1 \cdot \left(\frac{\overline{x^2} - (\bar{x})^2}{\overline{x^2}}\right) = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2}} \rightarrow$$

$$\widehat{\beta}_1 = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2}} \cdot \frac{\overline{x^2}}{\overline{x^2} - (\bar{x})^2} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} \rightarrow$$

$$\widehat{\beta}_1 = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2}$$

(17)

Ներդիր 5. Նմուշի թվային բնութագրիչները

Ենթադրենք ունենք n ծավալի մնուշային տվյալներ որևէ x ցուցանիշի կամ փոփոխականի մասին: Դրանք x -ի n թվով դիտարկումներն են:

Նմուշային միջինը մնուշի տվյալների միջին թվաբանականն է: Այն Գ-Հ անհայտ միջինի՝ մաթ. սպասման ($E(X) = \mu$) կետային վիճակագրական գնահատականն է և մնուշի n ծավալի դեպքում որոշվում է հետևյալ բանաձևով՝

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i: \quad (18)$$

Պատահական մնուշի դեպքում մնուշային միջինը $\bar{Q}$ -ի միջինն համար անշեղ և ունակ գնահատական է (այն դեպքում, եթե $\bar{Q}$ -ն ունի վերջավոր վարիացիա):

Նմուշային վարիացիան բնութագրում է միջինի շուրջը տվյալների ցրվածքը: Այն $\bar{Q}$ -ն անհայտ վարիացիայի՝ $VAR(X) = E(X - E(X))^2 = E(X^2) - [E(X)]^2$ կետային վիճակագրական գնահատականն է և որոշվում է հետևյալ բանաձևով՝

$$S_x^2 = var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \overline{x^2} - (\bar{x})^2: \quad (19)$$

Փոքր ծավալի մնուշների համար օգտագործվում է մնուշային ճշգրտված վարիացիան՝ $var(x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$: Մեծ ծավալի մնուշի ($n > 30$) վարիացիան $\bar{Q}$ -ի վարիացիայի անշեղ և ունակ գնահատականն է:

Նմուշային ստանդարտ շեղումը (հաճախ անվանվում է ստանդարտ սխալ) քառակուսի արմատ է մնուշային վարիացիայից և որոշվում է հետևյալ բանաձևով՝

$$S_x = \sqrt{S_x^2} = \sqrt{var(x)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\overline{x^2} - (\bar{x})^2}: \quad (20)$$

Այն բնութագրում է x -ի միջինի շուրջը մնուշային արժեքների բազմության ցրվածքը, այդպիսով՝ նաև մնուշային միջինի՝ որպես վիճակագրական գնահատականի որակը:

Ենթադրենք ունենք n ծավալի մնուշային տվյալներ որևէ x և y փոփոխականների համար կամ, որ մույնն է, x -ի և y -ի n թվով դիտարկումներ:

Նմուշային կովարիացիան բնութագրում է երկու փոփոխականների միջև գծային կապի ուղղությունը, այսինքն՝ ինչպես են դրանք փոփոխվում միասին: Այն հետևյալ բանաձևով որոշվող $\bar{Q}$ -ն անհայտ կովարիացիայի վիճակագրական գնահատականն է $COV(X, Y) = E[(X - E(X))(Y - E(Y))] = E(X \cdot Y) - E(X) \cdot E(Y)$: Նմուշային կովարիացիան որոշվում է հետևյալ բանաձևով՝

$$cov(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \overline{x \cdot y} - \bar{x} \cdot \bar{y}: \quad (21)$$

Կովարիացիան տեղեկատվություն է տալիս՝ փոփոխականների գծային կապը ուղի՞ղ է (մեկի աճը զուգակցվում է մյուսի աճով, դրանք «շարժվում են» մույն ուղղությամբ), թե՞ հակադիր (մեկի աճը զուգակցվում է մյուսի նվազմամբ, դրանք «շարժվում են» տարբեր ուղղություններով): Սակայն կովարիացիայի մեծությունը

կախված է փոփոխականների չափման միավորներից, և դրա բացարձակ արժեքը տարբեր չափման միավորներ ընտրելիս տարբեր է լինում: Ուստի այն փոփոխականների գծային կապի, փոխադարձ կախվածության ուժգնության մասին եզրակացություն անելու հնարավորություն չի տալիս: Այդ պատճառով կովարիացիայի գործակիցը ստանդարտացնում են՝ բերելով $[-1,1]$ միջակայքին: Այդ ստանդարտացման արդյունքը երկու փոփոխականների գույգային կոռելյացիայի գործակիցն է:

Այսպիսով՝ (17)-ից, (19)-ից ու (21)-ից հետևում է, որ՝

$$\widehat{\beta}_1 = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2} = \frac{cov(x, y)}{var(x)} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}: \quad (22)$$

Աղյուսակ 3-ում ամփոփված են էմպիրիկ տվյալների հիման վրա պարզ գծային ռեգրեսիայի՝ ՓԲԵ-ով գնահատված էմպիրիկ (9) հավասարման $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների հաշվարկային բանաձևերը, որոնք ապահովում են էմպիրիկ տվյալների ամենափոքր ցրումը գնահատված ուղիղ գծի շուրջը.

Աղյուսակ 3. ՓԲԵ-ով գնահատված պարզ գծային ռեգրեսիայի էմպիրիկ հավասարման գործակիցների հաշվարկային բանաձևերը

ՓԲԵ-ով գնահատված ռեգրեսիայի էմպիրիկ հավասարման ռեգրեսիայի կամ «թեքության» գործակից $\widehat{\beta}_1$	$\widehat{\beta}_1 = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2} = \frac{cov(x, y)}{var(x)} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$
ՓԲԵ-ով գնահատված ռեգրեսիայի էմպիրիկ հավասարման ազատ անդամ կամ ռեգրեսիայի հավասարման «բարձրության» գործակից $\widehat{\beta}_0$	$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}$

Այս բանաձևերով ՓԲԵ-ով որոշված $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները հնարավորություն են տալիս գտնելու պարզ գծային ռեգրեսիայի այն հավասարումը, որի համար էմպիրիկ տվյալների կետերի ամպն ունի ամենափոքր գումարային շեղումը գնահատված ուղիղ գծից և, այդպիսով, լուծելու վերը նշված 2-րդ կամ ներքևի մակարդակի խնդիրը:

Ինչպե՞ս են մեկնաբանվում ՓԲԵ-ով գնահատված էմպիրիկ $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ հավասարման գնահատված $\hat{\beta}$ գործակիցները: $\hat{\beta}_0$ -ն հավասարման ազատ անդամն է և ցույց է տալիս, թե ինչի հավասար

կլինի կախյալ փոփոխականի սպասվող կամ միջին արժեքը, եթե բացատրող փոփոխականը հավասար լինի զրոյի: Գնահատված β_1 գործակիցը ցույց է տալիս, թե քանի միավորով կփոխվի (կաճի կամ կնվազի) կախյալ փոփոխականի սպասվող կամ միջին արժեքը, եթե x բացատրող փոփոխականի արժեքը աճի մեկ միավորով:

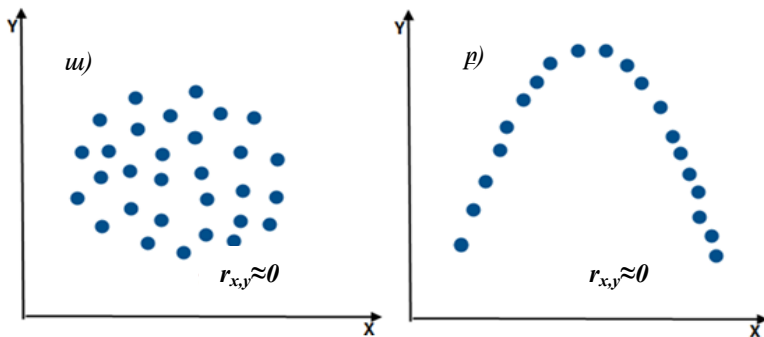
Ներդիր 6. Նմուշից սրացված արժեքների փոխկախվածության բնութագրիչը՝ կոռելյացիայի գործակից

Նմուշային զույգային կոռելյացիայի գործակիցը բնութագրում է երկու փոփոխականների զծային կապի ուժգնությունը և որոշվում է հետևյալ բանաձևով՝

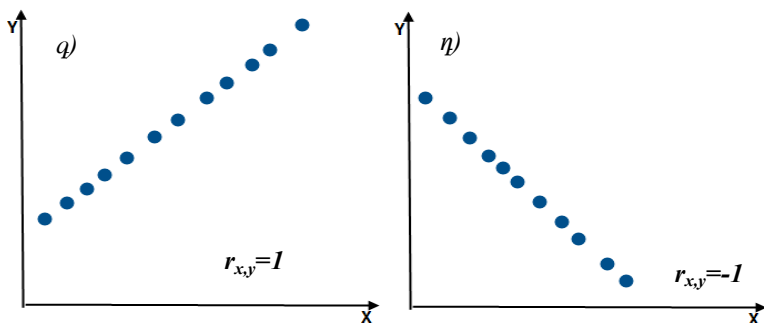
$$r_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)} \cdot \sqrt{\text{var}(y)}} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{S_x \cdot S_y} \quad (23)$$

Այն ԳՀ-ի երկու՝ X և Y պատահական մեծությունների կոռելյացիայի գործակից՝ $\rho_{xy} = \frac{\text{COV}(X,Y)}{\sqrt{\text{VAR}(X) \cdot \text{VAR}(Y)}}$ վիճակագրական կետային գնահատականն է:

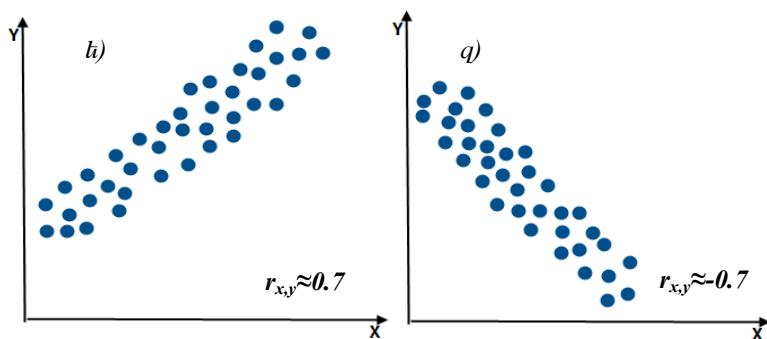
Զույգային կոռելյացիայի գործակիցն ընկած է $[-1, 1]$ միջակայքում: Եթե $r_{xy} = 0$, նշանակում է x -ի և y -ի միջև **զծային** կախվածություն չկա, և ասում ենք, որ դրանք կոռելյացված չեն: Սակայն դա չի նշանակում, որ դրանք անպայման անկախ են. դրանց միջև կարող է լինել այլ տեսքի կախվածություն: Օրինակ, եթե $r_{xy} \approx 0$, ապա դրանց կոռելյացիոն դաշտը կունենա ստորև բերված ա) տեսքը, եթե դրանք անկախ են, և բ) տեսքը, եթե դրանք զծային կախվածություն չունեն, սակայն, օրինակ, քառակուսային կախվածության մեջ են:



Եթե $r_{xy} = |1|$, ապա x -ն և y -ը ճշգրիտ ուղիղ կամ հակադիր գծային կախվածության մեջ են, և դրանց դիտարկումների արժեքներին համապատասխանող բոլոր կետերը գտնվում են ուղիղ գծի վրա:



Եթե կոռելյացիայի գործակիցը մոտ է 1-ին կամ -1-ին, ապա x -ի և y -ի միջև կա ուժեղ գծային կախվածություն: Եթե կոռելյացիայի գործակիցը դրական է, ապա երկու փոփոխականների միջև գծային կապն ուղիղ է (կամ դրական է). մեկի աճը զուգակցվում է մյուսի աճով (ե): Եթե կոռելյացիայի գործակիցը բացասական է, ապա երկու փոփոխականների միջև գծային կապը հակադիր կամ բացասական է. մեկի աճը զուգակցվում է մյուսի նվազմամբ (զ):



Գոյություն ունեն գծային կախվածության ուժգնության դասակարգման տարբեր սանդղակներ, որոնք հիմնված են կոռելյացիայի գործակցի արժեքի վրա:

Խորհուրդ տրվող սանդղակներից ամենատարածվածը հետևյալն է.

Զույգային կոռելյացիայի գործակցի բացարձակ արժեք՝ $ r_{xy} $	Մեկնաբանություն. x-ի և y-ի միջև՝
0.00 - 0.29	չկա գծային կախվածություն
0.30 to 0.49	կա թույլ գծային կախվածություն
0.50 to 0.69	կա ուժեղ գծային կախվածություն
0.70 to 1.00	կա շատ ուժեղ գծային կախվածություն

Նմուշային կոռելյացիայի գործակցի (23) ու $\widehat{\beta}_1$ գործակցի (22) քանադաներից հետևում է, որ x և y փոփոխականների զույգային կոռելյացիայի գործակցը և պարզ գծային ռեգրեսիայի հավասարման՝ ՓԲԵ-ով գնահատված $\widehat{\beta}_1$ գործակցը **ուղիղ համեմատական են՝**

$$\widehat{\beta}_1 = \frac{cov(x,y)}{var(x)} = r_{x,y} \cdot \frac{\sqrt{var(y)}}{\sqrt{var(x)}} = r_{x,y} \cdot \frac{s_y}{s_x}: \quad (24)$$

Այսպիսով՝ ՓԲԵ-ով $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակցների քանակական գնահատումը նշանակում է գծային պարզ ռեգրեսիայի $\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x$ հավասարման գնահատում, հետևաբար $\hat{y}$ -ի արժեքի, ինչպես նաև կախյալ y փոփոխականի փաստացի նմուշային տվյալի ու ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված արժեքների տարբերության՝ $\hat{\varepsilon}_i = y_i - \hat{y}_i$ պատահական մնացորդների ստացում $i=1,2,...n$ բոլոր նմուշային դիտարկումների համար, ինչը նշանակում է, որ ՓԲԵ-ով գնահատվում է ինչպես ռեգրեսիայի էմպիրիկ հավասարումը, այնպես էլ էմպիրիկ $y = \widehat{\beta}_0 + \widehat{\beta}_1 x + \hat{\varepsilon}$ մոդելը:

ՓԲԵ կիրառման արդյունքները

1. ՓԲԵ-ով ստացվող $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակցները ռեգրեսիայի տեսական մոդելի β_0 և β_1 անհայտ պարամետրերի կետային վիճակագրական գնահատականներն են:

2. ՓԲԵ-ով ստացվող $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակցների հաշվարկային քանադաներից հետևում է, որ դրանք կախված են էմպիրիկ տվյալներից,

այսինքն՝ ֆունկցիա են նմուշից: Տարբեր նմուշների համար դրանք տարբեր են, ուստի ունեն պատահական բնույթ:

3. ՓԲԵ-ով գնահատված գծային պարզ ռեգրեսիայի էմպիրիկ հավասարումը կառուցվում է այնպես, որ այն անպայման անցնում է $(\bar{x}, \bar{y})$ կետով, ինչը բխում է $\widehat{\beta}_0$ գործակցի հաշվարկային բանաձևից:

4. ՓԲԵ-ով գծային պարզ ռեգրեսիայի հավասարումը կառուցվում է այնպես, որ $\sum_{i=1}^n \hat{\varepsilon}_i = 0$, հետևաբար պատահական մնացորդների միջինը հավասար է զրոյի՝ $\frac{1}{n} \sum_{i=1}^n \hat{\varepsilon}_i = 0$: Սա հետևում է ՓԲԵ նորմալ հավասարումների համակարգի առաջին հավասարումից:

5. $\hat{\varepsilon}_i$ պատահական մնացորդներն ու անկախ x փոփոխականը կոռելացված չեն՝ $cov(\hat{\varepsilon}, x) = 0$: Սա հետևում է ՓԲԵ նորմալ հավասարումների համակարգի երկրորդ հավասարումից:

6. $\hat{\varepsilon}_i$ պատահական մնացորդներն ու կախյալ y -ի գնահատված արժեքները կոռելացված չեն՝ $cov(\hat{\varepsilon}, \hat{y}) = 0$: Սա հետևում է ՓԲԵ նորմալ հավասարումների համակարգի երկրորդ հավասարումից:

4. Փորրագույն քառակուսիների եղանակով գնահատված գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման որակի ստուգումը. դետերմինացիայի գործակից

Գծային ռեգրեսիայի էմպիրիկ հավասարման ՓԲԵ-ով գնահատումը, այսպիսով, ապահովում է էմպիրիկ տվյալների կետերի ամպի մոտարկումն ուղիղ գծով: Որքան փոքր է գումարային պատահական մնացորդը էմպիրիկ մոդելում, այնքան դիտարկումներից ձևավորված կետերն ավելի քիչ են ցրված գնահատված գծի շուրջը, և ռեգրեսիայի էմպիրիկ հավասարումն ավելի լավ է մոտարկում y -ի կախվածությունը x -ից: Կարելի է գնահատել, թե այդ մոտարկումն ընդհանուր առմամբ որքանով որակյալ է: Էմպիրիկ հավասարման որակի ստուգումը էկոնոմետրիկ հետազոտության 5-րդ փուլի (գնահատված մոդելի վավերականության ու որակի ստուգում) բաղկացուցիչներից է:

Հիշենք պարզ գծային ռեգրեսիայի էմպիրիկ մոդելի տեսքը կանոնական i -րդ դիտարկման համար.

$$y_i = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_i + \hat{\varepsilon}_i = \hat{y}_i + \hat{\varepsilon}_i: \quad (25)$$

(25)-ից երևում է, որ կախյալ փոփոխականի արժեքը կախված է x -երով բացատրվող մասից ու x -երով չբացատրվող՝ $\hat{\varepsilon}$ պատահական մնացորդներով բացատրվող մասից: Որքան մեծ է y -ի՝ x -երով բացատրվող մասը, այնքան փոքր է պատահական մնացորդներով բա-

ցատրվող մասը, և այնքան մոտ է y -ի գնահատված արժեք $\hat{y}$ -ը դրա դիտարկված մնուշային արժեք y -ին: Չնափոխենք (25)-ը՝ հավասարման երկու կողմից հանելով կախյալ փոփոխականի միջինն ու տեղադրելով դրա մեջ $\hat{\varepsilon}$ պատահական մնացորդի արտահայտությունը՝ $\hat{\varepsilon}_i = y_i - \hat{y}_i$:

$$y_i - \bar{y} = \hat{y}_i + (y_i - \hat{y}_i) - \bar{y} = (\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i): \quad (26)$$

(26) հավասարման ձախ կողմում կախյալ փոփոխականի i -րդ դիտարկման արժեքի տարբերությունն է կախյալ փոփոխականի սպասվող արժեքից, այսինքն՝ կախյալ փոփոխականի i -րդ դիտարկման շեղումն իր միջինից: Հավասարման աջ կողմի առաջին գումարելին i -րդ դիտարկման կախյալ փոփոխականի գնահատված արժեքի շեղումն է կախյալ փոփոխականի սպասվող արժեքից, իսկ երկրորդ գումարելին պատահական մնացորդ $\hat{\varepsilon}_i$ -ն է:

Չնափոխենք (26)-ը՝ հավասարման երկու կողմը քառակուսի բարձրացնելով ու գումարելով բոլոր դիտարկումների համար:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + 2 \sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \hat{y}_i) + \sum_{i=1}^n (y_i - \hat{y}_i)^2: \quad (27)$$

Հիշենք, որ ռեգրեսիայի էմպիրիկ մոդելը գնահատված է ՓՔԵ-ով, իսկ դրա արդյունքները վկայում են, որ $\sum_{i=1}^n \hat{\varepsilon}_i = \sum_{i=1}^n (y_i - \hat{y}_i) = 0$: Կարելի է ցույց տալ, որ նաև $\sum_{i=1}^n (\hat{y}_i - \bar{y}) \cdot (y_i - \hat{y}_i) = \sum_{i=1}^n (\hat{y}_i - \bar{y}) \cdot \hat{\varepsilon}_i = 0$: Հետևաբար կստանանք՝

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (\hat{\varepsilon}_i)^2: \quad (28)$$

(28)-ի ձախ կողմում գրված է y -ի ընդհանուր, գումարային վարիացիան, տատանումն իր միջինի շուրջը (եթե սա բաժանենք n -ի, կստանանք մնուշային վարիացիան): Ստացվում է, որ y -ի ընդհանուր, գումարային վարիացիան կամ ընդհանուր ցրումն իր միջինի շուրջը բաղկացած է երկու բաղկացուցից. առաջին բաղկացուցիչն այն ցրումն է, որը բացատրվում է գնահատված ռեգրեսիայի էմպիրիկ հավասարմամբ, իսկ երկրորդ բաղկացուցիչն այն ցրումն է, որը բացատրվում է պատահական մնացորդներով, այսինքն՝ ռեգրեսիայի հավասարմամբ չի բացատրվում:

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

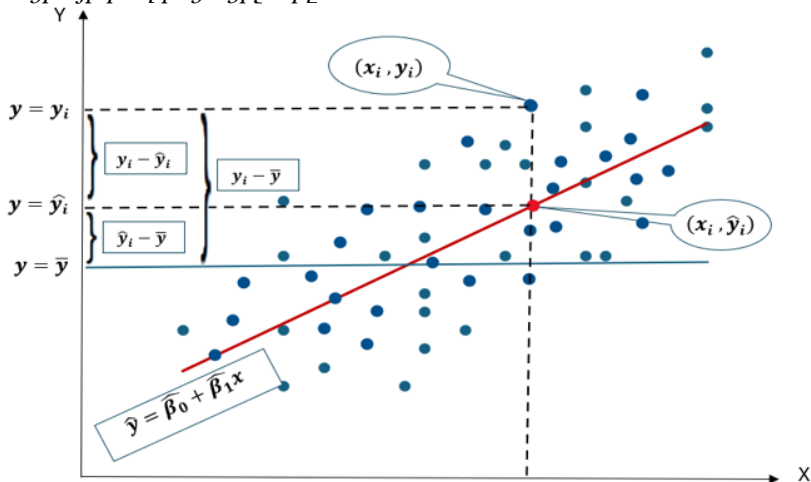
Կախյալ
փոփոխականի
ընդհանուր
վարիացիան իր
միջինի շուրջը

Կախյալ
փոփոխականի
վարիացիայի՝
ռեգրեսիայի
հավասարմամբ
բացատրվող
բաղկացուցիչ

Կախյալ փոփոխականի
վարիացիայի՝
ռեգրեսիայի
հավասարմամբ
չբացատրվող
բաղկացուցիչ

y -ի ընդհանուր վարիացիայի կամ իր միջինի շուրջը ունեցած ցրման այս երկու բաղկացուցիչները կարելի է ցույց տալ գրաֆիկի վրա որևէ կոնկրետ i -րդ դիտարկման համար (տես Գծանկար 6):

Գծանկար 6. Կախյալ փոփոխականի՝ իր միջինի շուրջն ունեցած վարիացիայի բաղկացուցիչները



(28) հավասարման երկու կողմը բաժանենք y -ի ընդհանուր վարիացիայի վրա: Կստանանք՝

$$1 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} + \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} =$$

$$= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} + \frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (29)$$

(29)-ի աջ կողմի առաջին գումարելին y -ի ընդհանուր վարիացիայի այն մասն է, որը բացատրվում է ռեգրեսիայի հավասարմամբ, իսկ երկրորդը՝ այն մասը, որը բացատրվում է պատահական մնացորդներով: Չնափոխենք (29)-ը.

$$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (30)$$

(30) հավասարման ձախ մասում y -ի՝ իր միջինի շուրջն ունեցած ընդհանուր վարիացիայի այն մասն է, որը բացատրվում է գնահատված ռեգրեսիայի էմպիրիկ հավասարմամբ:

Մահմանում. y -ի՝ իր միջինի շուրջն ունեցած ընդհանուր վարիացիայի այն մասը, որը բացատրվում է գնահատված ռեգրեսիայի էմպիրիկ հավասարմամբ, կանվանենք նմուշային դետերմինացիայի գործակից և կնշանակենք R^2 :

Այսպիսով՝ դետերմինացիայի գործակիցը որոշվում է հետևյալ բանաձևով՝

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (31)$$

Որպես y -ի ամբողջ վարիացիայի մաս՝ դետերմինացիայի գործակիցն ընկած է $[0, 1]$ միջակայքում՝

$$0 \leq R^2 \leq 1:$$

Որքան ուժեղ է y -ի գծային կախվածությունը x -ից, այնքան R^2 -ն մոտ է մեկին, որովհետև կախյալ փոփոխականի վարիացիայի ավելի մեծ մասն է բացատրվում գնահատված էմպիրիկ հավասարմամբ, և ավելի փոքր մասը՝ պատահական մնացորդներով՝ $\hat{\varepsilon}_i$ -երով: Հետևաբար ավելի բարձր է գնահատված ռեգրեսիայի էմպիրիկ հավասարման որակը, և (x_i, y_i) կետերից կազմված կետերի ամպը մոտ է գնահատված ուղիղ գծին: Եթե $R^2=1$, ապա բոլոր (x_i, y_i) կոորդինատներով կետերն ընկած են էմպիրիկ հավասարմամբ որոշված գծի վրա: Եթե y -ի գծային կախվածությունը x -ից թույլ է, R^2 -ն մոտ է զրոյին (հարկ է նշել, որ ասվածը վերաբերում է տարածական տվյալներին, ժամանակային շարքերի դեպքում դա միանշանակ չէ: Բացի դրանից՝ հնարավոր են դեպքեր, երբ դետերմինացիայի գործակիցը նշված մի-

ջակայքում չէ: Սակայն այդ դեպքերը ռեգրեսիայի դասական մոդելին չեն վերաբերում): Եթե, օրինակ, $R^2=0.75$, նշանակում է կախյալ փոփոխականի վարիացիայի 75 տոկոսը բացատրվում ռեգրեսիայի էմպիրիկ հավասարմամբ, և ընդամենը 25 տոկոսը՝ պատահական սխալներով: Եզրակացությունը հետևյալն է. գնահատված էմպիրիկ հավասարման որակը բարձր է: Սովորաբար R^2 -ն պետք է 0.5-ից մեծ լինի, որ գնահատված ռեգրեսիան ընդունելի որակ ունենա:

Պարզ գծային ռեգրեսիայի պարագայում դետերմինացիայի գործակիցը հավասար է x -ի և y -ի գույգային կոռելյացիայի գործակցի քառակուսուն.

$$\begin{aligned} R^2 &= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (\hat{\beta}_0 + \hat{\beta}_1 x_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (\hat{\beta}_1 (x_i - \bar{x}))^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \\ &= \hat{\beta}_1^2 \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)^2 \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \\ &= \frac{(\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}))^2}{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \right)^2 = r_{x,y}^2 \\ &\rightarrow R^2 = r_{xy}^2 : \end{aligned} \quad (32)$$

Ներդիր 7. Դետերմինացիայի գործակցի ներկայացումը ընդունված նշանակումների միջոցով

Էկոնոմետրիկ գրականությունում հաճախ դետերմինացիայի գործակիցը ներկայացվում է որոշակի նշանակումներ կատարելու միջոցով: Կատարենք նշանակումներ: Կախյալ փոփոխական y -ի՝ իր միջինի շուրջը ունեցած ընդհանուր վարիացիայի այն բաղկացուցիչը, որը բացատրվում է ռեգրեսիայի հավասարմամբ, նշանակենք ESS-ով (ESS – explained sum of squares, բացատրվող քառակուսիների գումար). $ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$:

Կախյալ փոփոխականի՝ իր միջինի շուրջը ունեցած ընդհանուր վարիացիայի այն բաղկացուցիչը, որը բացատրվում է պատահական $\hat{\varepsilon}_i$ մնացորդներով, նշանակենք RSS-ով (RSS – residual sum of squares, մնացորդներով բացատրվող քառակուսիների գումար). $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (\hat{\varepsilon}_i)^2$: Կախյալ փոփոխականի՝ իր միջինի շուրջը ունեցած ընդհանուր վարիացիան նշանակենք TSS-ով (Total sum of squares, քառակուսիների ընդհանուր գումար). $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$: Այդ դեպքում՝ $TSS=ESS+RSS$:

Քանի որ դետերմինացիայի գործակիցը կախյալ փոփոխականի ընդհանուր վարիացիայի այն մասն է, որը բացատրվում է գնահատված ռեգրեսիայի էմպիրիկ հավասարմամբ, ապա այս նշանակումներով դետերմինացիայի գործակիցն արտահայտվում է հետևյալ բանաձևով՝

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}.$$

5. Պարզ գծային ռեգրեսիայի էմպիրիկ հավասարման ՓՔԵ-ով գնահատման ու արդյունքների մեկնաբանության օրինակ

Անշարժ գույքի գործակալությունը, դնելով վարկած, որ Երևանում բնակարանի վաճառքի շուկայական գինը կախված է դրա մակերեսից, նպատակ ունի պարզելու, թե միջինում ինչ գնով կվաճառվի 75 քառակուսի մետր մակերեսով բնակարանը Երևանում: Գործակալությունը հատկորոշել է պարզ գծային ռեգրեսիայի տեսական մոդել՝ $y = \beta_0 + \beta_1 x + \varepsilon$, այսինքն՝ գծային կախվածություն, որտեղ y -ը բնակարանի շուկայական գինն է, իսկ x -ը՝ մակերեսը: ε պատահական շեղումների առկայությունը մոդելում բացատրվում է տարբեր պատահական ազդեցություններով, ներառյալ նաև բնակարանի գնի վրա ազդող այլ գործոնները, որոնք մոդելում հաշվի չեն առնված (օրինակ՝ թաղամաս, կանգառին մոտիկություն, հարկայնություն և այլն): Գործակալությունը հավաքագրել է էմպիրիկ տվյալներ Երևանում բազմաբնակարան շենքերի բնակարանների շուկայական գների (y , 1000 ԱՄՆ դոլար) ու դրանց մակերեսի (x , քառակուսի մետր) մասին՝ 10 բնակարանի համար (տե՛ս աղյուսակի 1-2 սյուները): Նպատակը էմպիրիկ հավասարման ու մոդելի գնահատումն ու արդյունքների մեկնաբանությունն է:

Էմպիրիկ $\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x$ հավասարումը գնահատելու համար պետք է (16) և (17) բանաձևերով հաշվարկենք $\widehat{\beta}_0$ ու $\widehat{\beta}_1$ գործակիցները: Հաշվարկները կարելի է կատարել՝ էմպիրիկ տվյալների աղյուսակը լրացնելով հաշվարկների սյուներով, օրինակ՝ MS Excel-ում:

$\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները հաշվարկելու համար նախ հաշվարկվում են դրանց բանաձևերի մեջ մտնող բաղադրիչները՝ $\bar{x}$ -ն ու $\bar{y}$ -ը, $x_i \cdot y_i$ -ը և դրա միջինը, ինչպես նաև x_i^2 -ն ու դրա միջինը: Դրանց հիման վրա կստացվի՝

$$\widehat{\beta}_1 = \frac{\bar{x} \cdot \bar{y} - \bar{x} \cdot \bar{y}}{\bar{x}^2 - (\bar{x})^2} = \frac{5904.8 - 71 \cdot 77.8}{5455.6 - 71^2} = 0.92;$$

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x} = 77.8 - 0.92 \cdot 71 = 12.5:$$

Հաշվարկների հաջորդականությունը բերված է Աղյուսակ 4-ում:

Աղյուսակ 4. Էմպիրիկ տվյալների հիման վրա ՓԶԵ-ով պարզ գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատման հաշվարկների կատարման քայլերը

Էմպիրիկ տվյալներ			Հաշվարկված մեծություններ				
Դիտարկումներ	Բնականորեն մակերես x_i	Բնականորեն շուկայական գին y_i	$x_i \cdot y_i$	x_i^2	$y_i = 12.5 + 0.92 \cdot x_i$ $\hat{y}_i$	$\varepsilon_i = y_i - \hat{y}_i$	y_i^2
	1	2	3	4	5	6	7
1	62	75.0	4650.0	3844	69.5	5.5	5625.0
2	56	60.2	3371.2	3136	64.0	-3.8	3624.0
3	99	105.0	10395.0	9801	103.6	1.4	11025.0
4	74	81.0	5994.0	5476	80.6	0.4	6561.0
5	81	95.0	7695.0	6561	87.0	8.0	9025.0
6	68	67.0	4556.0	4624	75.1	-8.1	4489.0
7	71	76.9	5459.9	5041	77.8	-0.9	5913.6
8	110	110.0	12100.0	12100	113.7	-3.7	12100.0
9	42	48.2	2025.2	1764	51.1	-2.9	2325.2
10	47	59.6	2801.2	2209	55.7	3.9	3552.2
	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 71$	$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = 77.8$	$\overline{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i = 5904.8$	$\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2 = 5455.6$			$\overline{y^2} = \frac{1}{n} \sum_{i=1}^n y_i^2 = 6424.0$

Այսպիսով՝ գնահատվեց ռեգրեսիայի հետևյալ էմպիրիկ հավասարումը՝

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot x = 12.5 + 0.92 \cdot x:$$

$\hat{\beta}_1$ -ի մեկնաբանությունը հետևյալն է. x -ի միավոր աճը հանգեցնում է y -ի սպասվող կամ միջին արժեքի 0.92 միավոր աճի: Այսինքն՝ Երևանում բնակարանի մակերեսի ավելացումը մեկ քառակուսի մետրով հանգեցնում է դրա գնի սպասվող արժեքի 0.92 հազար դոլարով կամ, որ նույնն է, 920 դոլարով աճի: $\hat{\beta}_0$ -ն ընդհանրապես ցույց է տալիս, թե միջինում ինչի հավասար կլինի կախյալ փոփոխականի սպասվող կամ միջին արժեքը, եթե բացառորդ փոփոխականը հավասար լինի զրոյի: Սակայն հաճախ այս մեկնաբանությունը տնտեսագիտական իմաստ ուղղակի չի ունենում: Տվյալ դեպքում մեկնաբանությունն այն է, որ եթե բնակարանը չունենա մակերես, ապա դրա գնի միջին արժեքը հավասար կլինի 12.5 հազար դոլարի, ինչն անիմաստ է: ՓԶԵ կիրառումը հնարավորություն է տալիս x -ի յուրաքանչյուր դիտարկման համար հաշվարկելու y -ի գնահատված արժեքը՝ $\hat{y}_i$ -ը, ինչպես նաև պատահական մնացորդը՝ ε_i -ը, և այդպիսով ստանալու գնահատված էմպիրիկ ռեգրեսիայի մոդելը:

Գնահատված ռեգրեսիայի հավասարումը հնարավորություն է տալիս նաև կանխատեսելու y -ի սպասվող արժեքը x -ի կամայական արժեքի դեպքում: Տվյալ խնդրի դեպքում, եթե $x=75$, ապա բնակարանի վաճառքի սպասվող, միջին գինը 81500 ԱՄՆ դոլարն է՝ $\hat{y} = 12.5 + 0.92 \cdot x = 12.5 + 0.92 \cdot 75 = 81.5$:

Աղյուսակում հաշվարկված տվյալներից կարելի է ստանալ նաև x -ի և y -ի գույգային կոռելյացիայի գործակիցը՝

$$r_{x,y} = \frac{\bar{x}\bar{y} - \bar{x} \cdot \bar{y}}{\sqrt{(\bar{x}^2 - (\bar{x})^2) \cdot (\bar{y}^2 - (\bar{y})^2)}} = \frac{5904.8 - 71 \cdot 77.8}{\sqrt{(5455.6 - 71^2)(6424 - 77.8^2)}} = 0.97:$$

Եզրակացությունը հետևյալն է. կա շատ ուժեղ դրական, ուղիղ գծային կախվածություն բնակարանի վաճառքի գնի ու դրա մակերեսի միջև:

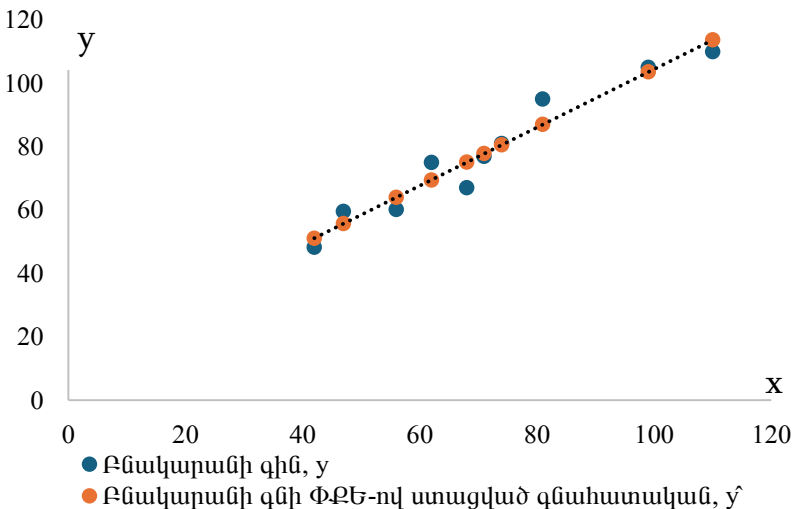
Օգտվելով (32) բանաձևից՝ պարզ ռեգրեսիայի համար կարելի է հեշտությամբ հաշվարկել դետերմինացիայի գործակիցը՝ $R^2 = r_{x,y}^2 = 0.97^2 = 0.94$:

Ընդհանուր եզրակացությունն է՝ ռեգրեսիայի էմպիրիկ հավասարումն ունի շատ բարձր որակ, վաճառքի գների՝ իր միջինի շուրջն

ունեցած վարիացիայի 94 տոկոսը բացատրվում է ռեգրեսիայի հավասարմամբ կամ, որ նույնն է, x -ով՝ բնակարանի մակերեսով:

Աղյուսակի հաշվարկներով կարելի է կառուցել (x_i, y_i) կոորդինատներով կետերի ամպի գծապատկերը՝ զուգակցելով այն x -ի ու կախյալ փոփոխականի գնահատված $\hat{y}$ արժեքներով ձևավորված $(x_i, \hat{y}_i)$ կետերի բազմության ու դրանցով գնահատված էմպիրիկ հավասարման՝ ուղիղ գծի պատկերների հետ:

Գծանկար 7. Էմպիրիկ տվյալները, կախյալ փոփոխականի՝ ՓԲԵ-ով գնահատված արժեքներն ու գնահատված գիծը



6. Գաուս-Մարկովի պայմաններն ու թեորեմը պարզ գծային ռեգրեսիայի համար. պարզ գծային ռեգրեսիայի նորմալ դասական մոդելի սահմանումը

Այսպիսով՝ ՓԲԵ կիրառումը հնարավորություն է տալիս գտնելու էմպիրիկ մոդելի $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցներն ու ε_i պատահական մնացորդները, որոնք էմպիրիկ տվյալների հիման վրա որոշում են գծային ռեգրեսիայի այն հավասարումը, այն ուղիղ գիծը, որը մոտարկում է դիտարկումների բազմությունն այնպես, որ դիտարկումների գումարային շեղումն այդ ուղղից լինի փոքրագույնը:

ՓԶԵ-ն լուծում է նախորդիվ նշած 2-րդ կամ ներքևի մակարդակի խնդիրը: Ինչպես լուծել նաև 1-ին կամ վերևի մակարդակի խնդիրը, այն է՝ գնահատել էմպիրիկ հավասարումն այնպես, որ այն լավագույնս մոտարկի ռեգրեսիայի տեսական հավասարումը, որն անհայտ է: Կամ, որ նույնն է, ՓԶԵ-ով ստանալ այնպիսի $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցներ, որոնք տեսական մոդելի β_0, β_1 անհայտ պարամետրերին հնարավորինս մոտ են: Դա, որպեսզի տեղի ունենա, տեսական մոդելի վրա պետք է դնել որոշակի պայմաններ, որոնք կոչվում են Գաուս-Մարկովի պայմաններ և հետևյալն են.

Պայման 1. Ռեգրեսիայի տեսական մոդելը գծային է β_0, β_1 պարամետրերի նկատմամբ, այսինքն՝ կախվածությունն ունի հետևյալ տեսքը՝

$$y = E(Y|X = x) + \varepsilon = \beta_0 + \beta_1 x + \varepsilon,$$

որում բացատրող փոփոխականը՝ x -ը, դետերմինացված մեծություն է և ε պատահական շեղումների հետ կոռելացված չէ՝ $COV(x, \varepsilon) = 0$:

Պայման 2. Պատահական շեղումների մաթ. սպասումը հավասար է զրոյի՝ կամայական i -ի համար՝

$$E(\varepsilon_i) = 0, \quad \forall i = 1, 2 \dots N:$$

Այս պայմանը նշանակում է, որ պատահական շեղումները, որոնք պայմանավորված են պատահական ու չներառված, մոդելում չհաշվառված ազդեցություններով, կախյալ փոփոխականի միջին արժեքի վրա սխատեմատիկ ազդեցություն չունեն: Սա նաև նշանակում է, որ կախյալ փոփոխականի պայմանական մաթ. սպասումը կախված չէ պատահական շեղումներից՝ $E(Y|X = x_i) = \beta_0 + \beta_1 x_i$:

Պայման 3. Կամայական ε_i և ε_j պատահական շեղումները կոռելացված չեն՝

$$COV(\varepsilon_i, \varepsilon_j) = 0, \quad \forall i \neq j$$

Այս պայմանը կոչվում է ε_i պատահական շեղումների **ավտոկոռելյացիայի բացակայության պայման**, իսկ դրա խախտումը՝ **ավտոկոռելյացիա**: Նշենք, որ եթե Պայման 2-ը տեղի ունի, ապա՝ $COV(\varepsilon_i, \varepsilon_j) = E[(\varepsilon_i - E(\varepsilon_i)) \cdot (\varepsilon_j - E(\varepsilon_j))] = E(\varepsilon_i \cdot \varepsilon_j)$: Ուստի, եթե Պայման 3-ը ևս տեղի ունի, ապա՝ $E(\varepsilon_i \cdot \varepsilon_j) = 0$:

Պայման 4. Պատահական շեղումների վարիացիան հաստատուն է: Այսինքն՝

$$VAR(\varepsilon_i) = \sigma_\varepsilon^2 = const, \quad \forall i = 1, 2 \dots N:$$

Այս պայմանը կոչվում է ε_i պատահական շեղումների հոմոսկեդաստիկության կամ միացրվածքայնության պայման: Պայմանի խախտումը կոչվում է հետերոսկեդաստիկություն կամ բազմացրվածքայնություն: Նշենք, որ եթե Պայման 2-ը տեղի ունի, ապա՝ $VAR(\varepsilon_i) = \mathbb{E}(\varepsilon_i - \mathbb{E}(\varepsilon_i))^2 = \mathbb{E}(\varepsilon_i^2)$: Եթե Պայման 4-ը ևս տեղի ունի, ապա՝ կամայական $i=1, 2, \dots, n$ դիտարկման համար $\mathbb{E}(\varepsilon_i^2) = \sigma_\varepsilon^2$: Այս պայմանից հետևում է, որ բացատրող փոփոխականի յուրաքանչյուր արժեքի համար կախյալ փոփոխականի վարիացիան իր միջինի շուրջը ևս հաստատուն է՝ $VAR(Y|X = x_i) = VAR(\beta_0 + \beta_1 x_i + \varepsilon_i) = VAR(\varepsilon_i) = \sigma_\varepsilon^2 = const$:

Այսպիսով՝ հոմոսկեդաստիկության պայմանը նշանակում է, որ x -ի յուրաքանչյուր արժեքին համապատասխանող y -ի արժեքների ցրումը դրանց միջինի շուրջը, հետևաբար ռեգրեսիայի տեսական հավասարման շուրջը չի փոխվում: Դա նշանակում է նաև, որ x -ի փոփոխման հետ պատահական շեղումների ցրվածքը չի փոփոխվում:

Պայման 5. Գծային ռեգրեսիայի տեսական մոդելի պատահական շեղումներն ունեն նորմալ բաշխում՝

$$\varepsilon \sim N(0, \sigma_\varepsilon^2):$$

Գաուս-Մարկովի թեորեմ

Եթե Գաուս-Մարկովի վերը նշված 1-4 պայմանները տեղի ունեն, ապա գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման ՓԲԵ-ով գնահատման արդյունքում ստացված $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականները տեսական մոդելի β_0, β_1 պարամետրերի լավագույն գծային անշեղ գնահատականներն են բոլոր գծային անշեղ գնահատականների դասում (կամ որ նույնն է՝ BLUE (best linear unbiased estimators) գնահատականներ են): Դա նշանակում է, որ դրանք՝

1. անշեղ գնահատականներ են՝ $\mathbb{E}(\widehat{\beta}_0) = \beta_0$ և $\mathbb{E}(\widehat{\beta}_1) = \beta_1$,
2. արդյունավետ գնահատականներ են. մնացած բոլոր գծային այլ անշեղ գնահատականների համեմատ ունեն փոքրագույն

վարիացիան՝ $var(\widehat{\beta}_0) < var(\widetilde{\beta}_0)$ և $var(\widehat{\beta}_1) < var(\widetilde{\beta}_1)$
**($\widehat{\beta}$ -երը կամայական այլ եղանակով ստացված գծային ան-
 շեղ գնահատականներ են):**

Մահմանում 1. Եթե Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն,
 գծային պարզ ռեգրեսիայի մոդելը կոչվում է **գծային պարզ ռեգրե-
 սիայի նորմալ դասական մոդել**: Այս մոդելի անհայտ պարամետրերն
 են β_0, β_1 -ը և պատահական շեղումների վարիացիան՝ σ_ε^2 -ն:

Մահմանում 2. Եթե տեղի ունեն Գաուս-Մարկովի 1-4 պայմանները,
 ապա գծային պարզ ռեգրեսիայի մոդելը կոչվում է **գծային պարզ
 ռեգրեսիայի դասական մոդել**:

Գաուս-Մարկովի թեորեմը չի պահանջում ε -երի նորմալ բաշխ-
 վածություն, չի պահանջում, որ 5-րդ պայմանը տեղի ունենա:
 Այնուամենայնիվ, այն շատ կարևոր է ՓԲԵ-ով գնահատված $\widehat{\beta}_0$ և $\widehat{\beta}_1$
 գնահատականների հատկությունների ու դրանց ստուգման համար:
 Ուստի ցանկալի է, որ 5-րդ պայմանը ևս տեղի ունենա, այսինքն՝
 ε -երը նորմալ բաշխված լինեն:

Հիշենք, որ ՓԲԵ-ով ստացված $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները կախված
 են x_i -երի և y_i -երի արժեքներից, հետևաբար՝ նմուշից: Նմուշից նմուշ
 էմպիրիկ հավասարման $\widehat{\beta}_0, \widehat{\beta}_1$ գործակիցները տարբերվում են, ուս-
 տի դրանք ունեն արժեքների բազմություն, բաշխում, միջին և ցրում
 այդ միջինի շուրջը: **Դրանք պատահական բնույթ ունեն**: Գաուս-
 Մարկովի 1-4 պայմանների իրականացման դեպքում դրանց միջինը
 տեսական մոդելի β պարամետրերն են, իսկ վարիացիաները ամենա-
 վոքրն են որոշակի դասում: Եթե տեսական մոդելի պատահական շե-
 ղումները նաև նորմալ բաշխված լինեն, ապա $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատա-
 կանները ևս նորմալ բաշխված կլինեն, քանի որ դրանք կախված են
 տեսական մոդելի ε -երից: Դա երևում է $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականների
 ՓԲԵ-ով ստացման (16) և (17) բանաձևերի ձևափոխումից՝

$$\begin{aligned}\widehat{\beta}_1 &= \frac{\sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i + \varepsilon_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2},\end{aligned}\tag{33}$$

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x} = \bar{y} - \frac{\bar{x} \sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i + \varepsilon_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (34)$$

Եթե $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականները նորմալ բաշխված լինեն, դրանց՝ որպես վիճակագրական գնահատականների մասին կարելի է որոշ վարկածներ ստուգել ու դրանց հատկությունների մասին որոշակի պետքական պնդումներ անել:

Ուրեմն ի՞նչ կփոխվի թեորեմի պնդման մեջ, եթե Գաուս-Մարկովի 5-րդ պայմանը ևս տեղի ունենա:

Թեորեմ*. Եթե Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն, ապա գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման ՓՔԵ-ով գնահատման արդյունքում ստացված $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականները տեսական մոդելի β_0 , β_1 պարամետրերի լավագույն անշեղ գնահատականներն են՝ բոլոր, ինչպես գծային, այնպես էլ ոչ-գծային անշեղ գնահատականների դասում, կամ, որ նույնն է, դրանք BUE (best unbiased estimators) գնահատականներ են: Դա նշանակում է, որ դրանք՝

1. անշեղ գնահատականներ են՝ $\mathbb{E}(\widehat{\beta}_0) = \beta_0$ և $\mathbb{E}(\widehat{\beta}_1) = \beta_1$,
2. արդյունավետ գնահատականներ են, դրանք մնացած բոլոր այլ անշեղ՝ ինչպես գծային, այնպես էլ ոչ գծային գնահատականների համեմատ ունեն փոքրագույն վարիացիան՝ $var(\widehat{\beta}_0) < var(\widetilde{\beta}_0)$ և $var(\widehat{\beta}_1) < var(\widetilde{\beta}_1)$,
3. ունակ են, այսինքն՝ դիտարկումների թվի մեծացման հետ այդ գնահատականները ձգտում են տեսական մոդելի β_0, β_1 պարամետրերի արժեքներին՝

$$\lim_{n \rightarrow \infty} \widehat{\beta}_0 = \beta_0 \text{ և } \lim_{n \rightarrow \infty} \widehat{\beta}_1 = \beta_1,$$

4. նորմալ են բաշխված՝

$$\widehat{\beta}_0 \sim N(\beta_0, var(\widehat{\beta}_0)) \text{ և } \widehat{\beta}_1 \sim N(\beta_1, var(\widehat{\beta}_1)):$$

Բացի դրանից, եթե Գաուս-Մարկովի 5-րդ պայմանը ևս տեղի ունի, ապա տեսական մոդելի պատահական շեղումները ոչ միայն կոռելացված չեն (Գաուս-Մարկովի պայման 3), այլև անկախ են, և y կախյալ փոփոխականը ևս ունի նորմալ բաշխում:

Այսպիսով, եթե ՓՋԵ-ն ապահովում է, որ էմպիրիկ տվյալների հիման վրա ռեգրեսիայի էմպիրիկ հավասարումը գնահատվի այնպես, որ լավագույն ձևով մոտարկի այդ տվյալների բազմությունը, ապա Գաուս-Մարկովի պայմանների տեղի ունենալն ապահովում է, որ ՓՋԵ-ով գնահատված էմպիրիկ հավասարումը նաև լավագույնս գնահատի ռեգրեսիայի անհայտ տեսական հավասարումը: Սա լուծում է առաջադրված թե՛ ներքևի, թե՛ վերին մակարդակի խնդիրները:

Ներդիր 8. Գաուս-Մարկովի թեորեմի ապացույցը. անշեղություն

Անշեղություն. Հաշվենք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների մաթ. սպասումները: Օգտվենք (33) ու (34)-ից:

$$\begin{aligned} (\widehat{\beta}_1) &= \mathbb{E} \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \mathbb{E} \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i + \varepsilon_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \\ &= \mathbb{E} \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i + \varepsilon_i - \beta_0 - \beta_1 \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \mathbb{E} \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) (\beta_1 x_i + \varepsilon_i - \beta_1 \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \\ &= \mathbb{E} \left(\frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (\beta_1 (x_i - \bar{x}) + \varepsilon_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \mathbb{E} \left(\frac{\sum_{i=1}^n \beta_1 (x_i - \bar{x}) ((x_i - \bar{x}) + \varepsilon_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \\ &= \mathbb{E} \left(\frac{\beta_1 \sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (x_i - \bar{x}) \varepsilon_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \mathbb{E} \left(\beta_1 + \frac{\sum_{i=1}^n (x_i - \bar{x}) \varepsilon_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = \\ &= \beta_1 + \frac{\sum_{i=1}^n (x_i - \bar{x}) \mathbb{E}(\varepsilon_i)}{\sum_{i=1}^n (x_i - \bar{x})^2} = \beta_1 \end{aligned}$$

$$\text{Ստացանք՝ } \mathbb{E}(\widehat{\beta}_1) = \beta_1:$$

$$\begin{aligned} \mathbb{E}(\widehat{\beta}_0) &= \mathbb{E}(\bar{y} - \widehat{\beta}_1 \bar{x}) = \mathbb{E}(\beta_0 + \beta_1 \bar{x} - \widehat{\beta}_1 \bar{x}) = \\ &= \mathbb{E}(\beta_0) + \mathbb{E}(\beta_1 \bar{x} - \widehat{\beta}_1 \bar{x}) = \beta_0 + \beta_1 \bar{x} - \beta_1 \bar{x} = \beta_0 \end{aligned}$$

$$\text{Ստացանք՝ } \mathbb{E}(\widehat{\beta}_0) = \beta_0:$$

Արդյունավետության ապացույցը բերված է Ներդիր 9-ում:

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Սահմանե՛ք էկոնոմետրիկան որպես գիտություն, տվե՛ք ձեր սահմանումը և բացատրե՛ք:
2. Բացատրե՛ք, թե ինչ է ուսումնասիրում էկոնոմետրիկան, որո՞նք են էկոնոմետրիկայի ուսումնասիրության առարկան ու խնդիրները, ո՞րն է նպատակը:
3. Բացատրե՛ք, թե ինչպես է էկոնոմետրիկան կապված տնտեսագիտության տեսության, տնտեսական վիճակագրության, մաթեմատիկական վիճակագրության, հավանականության տեսության ու տնտեսամաթեմատիկական մոդելավորման հետ:
4. Նկարագրե՛ք էկոնոմետրիկ հետազոտության փուլերն ու տրամաբանական սխեման:
5. Բացատրե՛ք, թե ինչ են էկոնոմետրիկ մոդելի հատկորոշումն ու պարամետրացումը:
6. Էմպիրիկ տվյալների ի՞նչ տիպեր գիտեք: Սահմանե՛ք և նկարագրե՛ք դրանք:
7. Փոփոխականների միջև կախվածությունների ի՞նչ տիպեր կան, ինչպե՞ս են դրանք սահմանվում, և ինչո՞վ են դրանք տարբերվում:
8. Սահմանե՛ք կոռելյացիոն ու ռեգրեսիոն կախվածությունները: Ի՞նչ տիպի կախվածություններ են դրանք և ինչո՞վ են տարբերվում:
9. Ինչո՞վ են տարբերվում ռեգրեսիայի տեսական ու էմպիրիկ մոդելները:
10. Ինչո՞վ է բացատրվում պատահական շեղման առկայությունը ռեգրեսիայի տեսական մոդելում:
11. Ի՞նչ ընդհանուր տեսք ունեն գծային պարզ ռեգրեսիայի տեսական հավասարումն ու մոդելը, էմպիրիկ հավասարումն ու մոդելը:
12. Ինչպե՞ս են սահմանվում գլխավոր համախմբությունն ու մնուշը: Նմուշների ի՞նչ տեսակներ կան: Ո՞րն է ավելի նախընտրելի:
13. Որո՞նք են մնուշի թվային բնութագրիչները, և ի՞նչ բանաձևերով են դրանք որոշվում:
14. Ի՞նչ են բնութագրում մնուշային կովարիացիան ու կոռելյացիան, ի՞նչ բանաձևեր ու հատկություններ ունեն:
15. Ի՞նչ է կետային վիճակագրական գնահատականը, և ի՞նչ հատկություններ պետք է ունենա որակյալ կետային գնահատականը:
16. Ի՞նչ է պատահական մնացորդը, ի՞նչ բանաձևով է այն որոշվում և ինչո՞վ է տարբերվում պատահական շեղումից:

17. Ռեգրեսիայի էմպիրիկ հավասարման մեջ n° րն է տեսական մոդելի Y կախյալ փոփոխականի պայմանական մաթ. սպասման վիճակագրական գնահատականը: Նշեք, թե ինչ այլ վիճակագրական գնահատականներ է պարունակում ռեգրեսիայի էմպիրիկ մոդելը:
18. H° ն բնույթ ունեն ռեգրեսիայի տեսական ու էմպիրիկ մոդելների կախյալ ու անկախ փոփոխականները, տեսական մոդելի պարամետրերը, էմպիրիկ մոդելի գործակիցները, պատահական շեղումներն ու պատահական մնացորդները (ստոխաստիկ/հավանականական կամ դետերմինացված):
19. Ռեգրեսիայի էմպիրիկ հավասարման գնահատման տեսանկյունից H° ն էրկու մակարդակի խնդիրներ են առաջանում: H° ն խնդիր է լուծում ՓԲԵ-ն էկոնոմետրիկ մոդելավորման գործընթացում:
20. Որո՞նք են ՓԲԵ էությունն ու դրա կիրառման արդյունքները պարզ գծային ռեգրեսիայի մոդելի համար:
21. H° ն բանաձևերով և ինչպե՞ս են ՓԲԵ-ով որոշվում գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները, և ինչի՞նչ վիճակագրական գնահատականներն են դրանք:
22. Մեկնաբանե՛ք գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների իմաստը:
23. Որո՞նք են ՓԲԵ կիրառման արդյունքները:
24. H° ն առնչություն գոյություն ունի գույգային կոռելյացիայի գործակցի ու գծային պարզ ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_1$ գործակցի միջև:
25. Երկու փոփոխականների գծային կախվածության ուժգնության դասակարգման H° ն սանդղակ է խորհուրդ տրվում օգտագործել:
26. Ինչպե՞ս է սահմանվում դետերմինացիայի գործակիցը, H° ն միջակայքում է այն գտնվում, H° ն բանաձևով է որոշվում և ինչպե՞ս է մեկնաբանվում:
27. Ինչպե՞ս են կապված դետերմինացիայի ու կոռելյացիայի գործակիցները պարզ գծային ռեգրեսիայի դեպքում:
28. Որո՞նք են Գաուս-Մարկովի պայմանները պարզ գծային ռեգրեսիայի մոդելի համար: Բացատրե՛ք դրանցից յուրաքանչյուրի էությունը:
29. Գաուս-Մարկովի n° ր պայմանն է կոչվում հոմոսկեդաստիկություն, n° րն է դրա էությունը, և H° ն է նշանակում դրա խախտումը:

30. Գաուս-Մարկովի n° ր պայմանի խախտումն է կոչվում ավտոկորելյացիա, n° րն է դրա էությունը:
31. Ինչո՞ւ է կարևոր պատահական շեղումների նորմալ բաշխման պայմանը Գաուս-Մարկովի պայմանների թվում:
32. Ձևակերպե՛ք Գաուս-Մարկովի թեորեմը այն դեպքերում, երբ տեղի ունեն Գաուս-Մարկովի 1-4 ու 1-5 պայմանները: Ի՞նչ է նշանակում լավագույն գծային անշեղելի (BLUE) գնահատական: Ի՞նչ է նշանակում լավագույն անշեղելի (BUE) գնահատական:
33. Ինչպե՞ս են փոխվում պարզ գծային ռեգրեսիայի էմպիրիկ մոդելի գնահատված գործակիցների հատկությունները, եթե Գաուս-Մարկովի՝ պատահական շեղումների նորմալ բաշխման պայմանը ևս տեղի ունի:
34. Սահմանե՛ք պարզ գծային ռեգրեսիայի դասական ու նորմալ դասական մոդելները:
35. Էկոնոմետրիկ հետազոտության համատեքստում ի՞նչ խնդիր է լուծում գծային ռեգրեսիայի գնահատման փոքրագույն քառակուսիների եղանակը, և ի՞նչ խնդիր է լուծում Գաուս-Մարկովի պայմանների տեղի ունենալը:

Խնդիրներ

1. Տրված է երկու ցուցանիշների մմուշային տվյալների հիման վրա հաշվարկված նկարագրական վիճակագրությունների աղյուսակը: Մեկնաբանե՛ք ստացված արդյունքները յուրաքանչյուր ցուցանիշի համար:

	N	Range, քափ	Minimum, մինիմում	Maximum, մաքսիմում	Mean, միջին	Variance, վարիացիա	Std. Deviation, ստանդարտ շեղում
Ընկերության աշխատողների ամսական աշխատավարձ, 1000 ՀՀ դրամ	164	212.0	75.8	287.8	125.5	289.0	17.0
Ընկերության աշխատողների աշխատանքային ստաժ, տարի	164	23.8	1.2	25.0	7.8	3.8	1.9

2. Տրված է երկու ցուցանիշների գույգային կոռելյացիայի գործակիցների կոռելյացիոն մատրիցը: Մեկնաբանե՞ք ստացված արդյունքները:

		Գրագիտության մակարդակ, %	Մանկամահացություն՝ 1,000 ծնվածի հաշվով
Գրագիտության մակարդակ, % բնակչությունից	Չույգային կոռելյացիայի գործակից N	1 220	-0.519 220
Մանկամահացություն՝ 1,000 ծնվածի հաշվով	Չույգային կոռելյացիայի գործակից N	-0.519 220	1 220

3. Տրված են 10 կազմակերպությունների գովազդային ծախսերի (x) ու վաճառքների ծավալների (y) մասին էմպիրիկ տվյալներ՝ 100 հազար դրամներով:

Էմպիրիկ տվյալներ

i	x_i	y_i
1	2.9	3.0
2	1.7	2.5
3	1.2	2.3
4	0.3	1.3
5	0.2	1.7
6	0.8	1.6
7	1.4	2.9
8	0.5	1.3
9	1.4	2.6
10	2.1	3.8

1. Հաշվե՞ք x-ի և y-ի մնուշային կովարիացիան և եզրակացություններ արե՞ք գծային կախվածության ուղղության մասին:

2. Հաշվե՞ք x-ի և y-ի մնուշային վարիացիան և ստանդարտ շեղումը:

3. Հաշվե՞ք x-ի և y-ի մնուշային կոռելյացիայի գործակիցը և եզրակացություն արե՞ք նրանց միջև կախվածության ուղղության ու ուժգնության մասին:

4. Կառուցե՞ք x-ի և y-ի մնուշային կոռելյացիոն դաշտի կետապատկերը:

5. ՓԶԵ-ով գնահատե՞ք գովազդային ծախսերից վաճառքների ծավալի կախվածության միագործոն գծային ռեգրես-

սիայի էմպիրիկ հավասարման $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները:

6. Գրե՞ք միագործոն գծային ռեգրեսիայի գնահատված էմպիրիկ հավասարումը և մեկնաբանե՞ք հավասարման $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների իմաստը:

7. Հաշվե՞ք կախյալ փոփոխականի՝ գծային ռեգրեսիայի էմպիրիկ հավասարմամբ գնահատված արժեքները՝ $\widehat{y}_i$ -երն ու պատասխական մնացորդները՝ $\widehat{\varepsilon}_i$ -երը:

8. Հաշվենք ռեգրեսիայի էմպիրիկ հավասարման դետերմինացիայի գործակիցը և եզրակացությունն արենք հավասարման որակի մասին:
8. Կառուցենք ռեգրեսիայի էմպիրիկ հավասարման ուղիղ գիծը՝ կոռելյացիոն դաշտի կետապատկերի հետ միևնույն գրաֆիկի վրա համադրելով:
9. Ինչի՞ հավասար կլինի վաճառքների ծավալը միջինում, եթե գովազդային ծախսերը հավասար լինեն 1.5-ի:
4. 160 երկրների 2026 թ. տվյալների հիման վրա ՓԲԵ-ով գնահատվել է մեկ շնչի հաշվով ՀՆԱ-ի (y , 1000 ԱՄՆ դոլար) կախումը գործազրկության մակարդակից (x , տոկոս) պարզ գծային ռեգրեսիայի մոդելով: Ստացվել է հետևյալ էմպիրիկ հավասարումը՝ $\hat{y} = 50 - 2.8x$: Ինչի՞ հավասար պատահական մնացորդը՝ ε -ը i -րդ դիտարկման համար, եթե այդ դիտարկմանը համապատասխանող $y_i = 8$; $x_i = 13$:
5. 120 երկրների 2026 թ. տվյալների հիման վրա ՓԲԵ-ով գնահատվել է գնաճի (y , տոկոս) կախումը գործազրկության մակարդակից (x , տոկոս) պարզ գծային ռեգրեսիայի մոդելով: Ստացվել է հետևյալ էմպիրիկ հավասարումը՝ $\hat{y} = \widehat{\beta}_0 - 0.15x$: Ինչի՞ հավասար հավասարման ազատ անդամը՝ $\widehat{\beta}_0$ -ն, եթե հայտնի է, որ $\bar{y} = 7.1$; $\bar{x} = 5.8$:
6. 190 երկրների 2026 թ. տվյալների հիման վրա ՓԲԵ-ով գնահատվել է աղքատության մակարդակի (y , տոկոս) կախումը մեկ շնչի հաշվով ՀՆԱ-ից (x , 1000 ԱՄՆ դոլար) պարզ գծային ռեգրեսիայի մոդելով: Ստացվել է հետևյալ էմպիրիկ հավասարումը՝ $\hat{y} = 89 + \widehat{\beta}_1 x$: Ինչի՞ հավասար հավասարման $\widehat{\beta}_1$ ռեգրեսիայի գործակիցը, եթե հայտնի է, որ $\bar{y} = 12.6$; $\bar{x} = 19.9$; $\overline{xy} = 248.6$; $\overline{x^2} = 410$:
7. 130 երկրների 2026 թ. տվյալների հիման վրա ՓԲԵ-ով գնահատվել է բուհական կրթության աստիճանում ընդգրկվածության մակարդակի (y , տոկոս համապատասխան տարիքի բնակչության մեջ) կախումը միջնակարգ կրթության աստիճանում մեկ սովորողի հաշվով կատարվող տարեկան պետական ծախսերից (x , 1000 ԱՄՆ դոլար) պարզ գծային ռեգրեսիայի մոդելով: Ստացվել է

հետևյալ էմպիրիկ հավասարումը՝ $\hat{y} = 1.5 + 3.75x$: Ինչի՞նչ է հավասար կախյալ փոփոխականի փաստացի էմպիրիկ արժեքը՝ y -ը i -րդ դիտարկման համար, եթե այդ դիտարկմանը համապատասխանող $\hat{\epsilon}_i = -6.8$; $x_i = 13$:

8. 145 երկրների 2026 թ. տվյալների հիման վրա ՓՔԵ-ով գնահատվել է ՕՈՒՆ-ի՝ երկրի ինստիտուցիոնալ ներուժը բնութագրող «Օրենքի ուժի ինդեքսից» կախվածության պարզ գծային ռեգրեսիայի էմպիրիկ հավասարումը, և ստացվել է, որ $\widehat{\beta}_1 = 0.56$: Ինչի՞նչ է հավասար x -ի և y -ի մոլուշային կոռելյացիայի գործակիցը՝ $r_{x,y}$ -ը, եթե հայտնի է, որ y -ի մոլուշային վարիացիան՝ $var(y) = 144$, իսկ x -ինը՝ $var(x) = 25$:
9. 120 երկրների 2026 թ. տվյալների հիման վրա ՓՔԵ-ով գնահատվել է Կոբ Դուգլասի արտադրական ֆունկցիան որոշ ձևափոխությամբ հետո պարզ գծային ռեգրեսիայի էմպիրիկ հավասարում գնահատելու միջոցով: Ստացվել է, որ $RSS=375$ և $TSS=1250$: Հաշվե՞ք հավասարման դետերմինացիայի գործակիցը և մեկնաբանե՞ք ստացված արդյունքը:
10. Պարզ գծային ռեգրեսիայի մոդելը ՓՔԵ-ով գնահատելու արդյունքում ստացվել է, որ կախյալ փոփոխականի՝ y -ի՝ իր միջինի շուրջը ունեցած ընդհանուր վարիացիան հավասար է 120-ի: Գտե՞ք այդ ընդհանուր վարիացիայի՝ պատահական մնացորդներով՝ $\hat{\epsilon}_i$ -երով բացատրվող մասը՝ RSS -ը, եթե դետերմինացիայի գործակիցը հավասար է 0.5-ի:

7. ՓՔԵ-ով գնահատված պարզ գծային ռեգրեսիայի էմպիրիկ հավասարման վիճակագրական գնահատականների ճշգրտության վերլուծությունը

7.1. Ռեգրեսիայի էմպիրիկ հավասարման գործակիցների վարիացիաներն ու ստանդարտ սխալները

Ենթադրենք՝ Գաուս-Մարկովի 1-5 պայմանները, ներառյալ պատահական շեղումների նորմալ բաշխվածությունը, տեղի ունեն։ Այդ դեպքում գծային ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցները տեսական մոդելի β_0 և β_1 պարամետրերի լավագույն գնահատականներն են։ Մասնավորապես դրանց միջինները հավասար են տեսական մոդելի անհայտ պարամետրերին, և դրանք նորմալ բաշխում ունեն։ Միջինի՝ որպես վիճակագրական գնահատականի ճշգրտությունը կախված է այդ միջինի շուրջը բազմության ցրվածքից։ Ուրեմն $\widehat{\beta}_0$, $\widehat{\beta}_1$ գնահատականների ճշգրտությունը կամ որակը (որքանով որակյալ գնահատականներ են $\widehat{\beta}_0$, $\widehat{\beta}_1$ -ի միջինները տեսական մոդելի անհայտ β -երի համար) կարելի է որոշել դրանց՝ իրենց միջինների շուրջը ունեցած ցրվածքի միջոցով։ Որքան փոքր է նշված ցրվածքը, այնքան ճշգրիտ կամ որակյալ գնահատականներ են ՓՔԵ-ով գնահատված հավասարման $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցները։ Միջինի շուրջը ցրվածքը չափվում է մնուշային վարիացիայի ու ստանդարտ շեղման միջոցով։

Հաշվարկենք $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցների վարիացիաները։ Դիցուք ՓՔԵ-ով գնահատվել է պարզ գծային ռեգրեսիայի էմպիրիկ մոդել՝ $y_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i + \widehat{\varepsilon}_i$, որի համար բավարարված են Գաուս-Մարկովի 1-5 պայմանները։ Այդ դեպքում նկատի ունենալով, որ $\sum_{i=1}^n (x_i - \bar{x}) = 0$, կստանանք՝

$$\begin{aligned}\widehat{\beta}_1 &= \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} - \frac{\bar{y} \cdot \sum_{i=1}^n (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} : \end{aligned}$$

Նշանակենք՝

$$\frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} = k_i :$$

Հետևաբար՝ $\widehat{\beta}_1 = \sum_{i=1}^n k_i \cdot y_i$: (35)

k_i -երը դետերմինացված մեծություններ են, պատահական բնույթ չունեն, քանի որ կախված են միայն x -երից, որոնք նույնպես դետերմինացված մեծություններ են: Այդ դեպքում $\widehat{\beta}_1$ -ի վարիացիան կլինի՝

$$\begin{aligned} var(\widehat{\beta}_1) &= var\left(\sum_{i=1}^n k_i \cdot y_i\right) = var\left(\sum_{i=1}^n k_i \cdot (\beta_0 + \beta_1 x_i + \varepsilon_i)\right) = \\ &= var\left(\beta_0 \cdot \sum_{i=1}^n k_i + \beta_1 \cdot \sum_{i=1}^n k_i x_i + \sum_{i=1}^n k_i \varepsilon_i\right) = var\left(\sum_{i=1}^n k_i \varepsilon_i\right) = \\ &= var\left(\frac{\sum_{i=1}^n (x_i - \bar{x}) \varepsilon_i}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) = \frac{(\sum_{i=1}^n (x_i - \bar{x}))^2}{\sum_{i=1}^n (x_i - \bar{x})^4} \cdot var(\varepsilon_i) = \\ &= \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} : \end{aligned}$$

- $var(\widehat{\beta}_1)$ -ը սրանալու համար օգտագործվել են վարիացիայի հեղուկայի հատկությունները. եթե a -ն հաստատուն է, ապա՝ $var(a) = 0$, $var(ax) = a^2 \cdot var(x)$:

Ստացվեց՝

$$var(\widehat{\beta}_1) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} : \quad (36)$$

Նույն կերպ կարելի է ստանալ $\widehat{\beta}_0$ -ի վարիացիան: Ունենք՝

$$\begin{aligned} \widehat{\beta}_0 &= \bar{y} - \widehat{\beta}_1 \bar{x} \rightarrow \\ var(\widehat{\beta}_0) &= var(\bar{y} - \widehat{\beta}_1 \bar{x}) = var(\bar{y}) + var(\widehat{\beta}_1 \cdot \bar{x}) - 2cov(\bar{y}, \widehat{\beta}_1 \bar{x}) = \\ &= var\left(\frac{1}{n} \sum_{i=1}^n y_i\right) + (\bar{x})^2 \cdot var(\widehat{\beta}_1) = \frac{1}{n^2} \cdot var\left(\sum_{i=1}^n (\beta_0 + \beta_1 x_i + \varepsilon_i)\right) + \\ &+ \frac{\sigma_\varepsilon^2 \cdot (\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{1}{n^2} \cdot n \cdot \sigma_\varepsilon^2 + \frac{\sigma_\varepsilon^2 \cdot (\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \frac{\sigma_\varepsilon^2}{n} + \frac{\sigma_\varepsilon^2 \cdot (\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = \sigma_\varepsilon^2 \cdot \left(\frac{1}{n} + \frac{(\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) : \end{aligned}$$

- $var(\widehat{\beta}_0)$ -ն սրանալու համար օգտագործվել են վարիացիայի հետևյալ հատկությունները՝ $var(x - y) = var(x) + var(y) - 2 \cdot cov(x, y)$, $var(a) = 0$, $var(ax) = a^2 \cdot var(x)$:
Ինչպես նաև կարելի է ցույց տալ, որ $cov(\bar{y}, \widehat{\beta}_1 \bar{x}) = 0$:

Ստացվեց՝

$$var(\widehat{\beta}_0) = \sigma_\varepsilon^2 \cdot \left(\frac{1}{n} + \frac{(\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right): \quad (37)$$

Այսպիսով՝ Գաուս-Մարկովի 5-րդ պայմանի տեղի ունենալու պարագայում էմպիրիկ մոդելի $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցներն ունեն նորմալ բաշխում հետևյալ պարամետրերով՝

$$\begin{aligned} \widehat{\beta}_0 &\sim N \left(\beta_0, \sigma_\varepsilon^2 \cdot \left(\frac{1}{n} + \frac{(\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) \right) \\ \widehat{\beta}_1 &\sim N \left(\beta_1, \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right): \end{aligned}$$

(36)-(37)-ից հետևում է, որ $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցների վարիացիաները ուղիղ համեմատական են տեսական մոդելի պատահական շեղումների՝ ε_i -երի վարիացիային: Որքան մեծ է ε_i -երի վարիացիան, այնքան մեծ է $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցների՝ որպես վիճակագրական գնահատականների ցրումը իրենց միջինների՝ տեսական մոդելի β_0 և β_1 պարամետրերի շուրջը, հետևաբար այնքան փոքր են այդ գործակիցների ճշգրտությունը, որակը: Մյուս կողմից, որքան մեծ է դիտարկումների քանակը՝ n -ը, այնքան փոքր են $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականների վարիացիաները, հետևաբար մեծ է այդ գնահատականների ճշգրտությունը:

Ներդիր 9. Գաուս-Մարկովի թեորեմի ապացույցը. արդյունավետություն

Օգտվենք ՓԲԵ-ով ստացված $\widehat{\beta}_1$ գնահատականի (35) տեսքից: Ենթադրենք, որ գոյություն ունի ռեգրեսիայի գործակցի մեկ այլ մեթոդով ստացված գծային անշեղ $\widetilde{\beta}_1$ գնահատական, որն արտահայտվում է այսպես՝ $\widetilde{\beta}_1 = \sum_{i=1}^n \omega_i \cdot y_i$, ըստ որում՝ $\omega_i = k_i + \theta_i$ (որում θ -ն ևս, ինչպես k -ն, դետերմինացված մեծություն է): Զանի որ θ թե՛ β_1 -ը, թե՛ $\widetilde{\beta}_1$ -ը անշեղ գնահատականներ են, ապա՝ $E(\widehat{\beta}_1) = E(\widetilde{\beta}_1) = \beta_1$ և $E(\widetilde{\beta}_1) - E(\widehat{\beta}_1) = 0$:

Հաշվենք.

$$\begin{aligned}\mathbb{E}(\widetilde{\beta}_1) - \mathbb{E}(\hat{\beta}_1) &= \mathbb{E}(\widetilde{\beta}_1 - \hat{\beta}_1) = \mathbb{E}(\sum_{i=1}^n \omega_i \cdot y_i - \sum_{i=1}^n k_i \cdot y_i) = \\ &= \mathbb{E}(\sum_{i=1}^n k_i \cdot y_i + \sum_{i=1}^n \theta_i \cdot y_i - \sum_{i=1}^n k_i \cdot y_i) = \mathbb{E}(\sum_{i=1}^n \theta_i \cdot y_i) = \\ &= \mathbb{E}(\sum_{i=1}^n \theta_i \cdot (\beta_0 + \beta_1 x_i + \varepsilon_i)) = \beta_0 \sum_{i=1}^n \theta_i + \beta_1 \sum_{i=1}^n \theta_i x_i : \end{aligned}$$

Եթե $\beta_0 \neq 0$ և $\beta_1 \neq 0$, ապա $\mathbb{E}(\widetilde{\beta}_1) - \mathbb{E}(\hat{\beta}_1) = 0 \rightarrow \sum_{i=1}^n \theta_i = 0$ և $\sum_{i=1}^n \theta_i x_i = 0$: Հաշվենք.

$$\begin{aligned}Var(\widetilde{\beta}_1) &= Var(\sum_{i=1}^n \omega_i \cdot y_i) = \sum_{i=1}^n \omega_i^2 \cdot Var(y_i) = \\ &= \sum_{i=1}^n (\omega_i^2 \cdot Var(\beta_0 + \beta_1 x_i + \varepsilon_i)) = \sum_{i=1}^n (\omega_i^2 \cdot Var(\varepsilon_i)) = \\ &= \sum_{i=1}^n (\omega_i^2 \cdot \sigma_\varepsilon^2) = \sum_{i=1}^n (k_i + \theta_i)^2 \cdot \sigma_\varepsilon^2 = \\ &= \sigma_\varepsilon^2 \cdot \sum_{i=1}^n (k_i^2 + 2 \cdot k_i \theta_i + \theta_i^2) = \\ &= \sigma_\varepsilon^2 \cdot \sum_{i=1}^n k_i^2 + 2\sigma_\varepsilon^2 \cdot \sum_{i=1}^n k_i \theta_i + \sigma_\varepsilon^2 \cdot \sum_{i=1}^n \theta_i^2 = \\ &= \sigma_\varepsilon^2 \cdot \sum_{i=1}^n k_i^2 + \sigma_\varepsilon^2 \cdot \sum_{i=1}^n \theta_i^2 = \frac{\sigma_\varepsilon^2 \sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^4} + \sigma_\varepsilon^2 \cdot \sum_{i=1}^n \theta_i^2 = \\ &= \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (x_i - \bar{x})^2} + \sigma_\varepsilon^2 \cdot \sum_{i=1}^n \theta_i^2 = var(\widetilde{\beta}_1) + \sigma_\varepsilon^2 \cdot \sum_{i=1}^n \theta_i^2 > var(\hat{\beta}_1): \end{aligned}$$

Ստացվեց՝ $Var(\widetilde{\beta}_1) > var(\hat{\beta}_1)$:

- Այս հաշվարկները ստացվեցին՝ հաշվի առնելով, որ

$$k_i = \frac{x_i - \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2} \text{ և }$$

$$\sum_{i=1}^n k_i \theta_i = \frac{\sum_{i=1}^n (x_i - \bar{x}) \theta_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n \theta_i x_i - \bar{x} \sum_{i=1}^n \theta_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = 0$$

և օգտագործելով (36)-ը:

Նույն ձևով կարելի է ցույց տալ նաև, որ $Var(\widetilde{\beta}_0) > var(\hat{\beta}_0)$:

$\widehat{\beta}_0$ և $\widehat{\beta}_1$ գնահատականների վարիացիաները որոշվում են տեսական մոդելի ε_i պատահական շեղումների վարիացիայի, σ_ε^2 -ու միջոցով, որն անհայտ է: Գաուս-Մարկովի պայմանները պահանջում են միայն, որ այն լինի հաստատուն բոլոր դիտարկումների համար, սակայն այն գծային պարզ ռեգրեսիայի նորմալ դասական մոդելի անհայտ պարամետրերից մեկն է:

Տեսական մոդելի պատահական ε_i շեղումների անհայտ σ_ε^2 վարիացիան կարելի է գնահատել էմպիրիկ մոդելի $\hat{\varepsilon}_i$ պատահական մնացորդների վարիացիայի միջոցով: Ենթադրենք՝ Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն: Կարելի է ցույց տալ, որ այդ դեպքում $\hat{\varepsilon}_i$ պատահական մնացորդները տեսական մոդելի ε_i պատահական շեղումների անշեղ գնահատականներն են, իսկ դրանց $var(\hat{\varepsilon}_i)$ վարիացիան տեսական մոդելի ε_i -երի վարիացիայի վիճակագրական անշեղ գնահատականն է: Բանի որ $\hat{\varepsilon}_i$ -երը հայտնի են, դրանց վարիացիան կարելի է հաշվել:

$\hat{\varepsilon}_i$ -երի վարիացիան որոշվում է որպես նմուշային ճշգրտված վարիացիա հետևյալ բանաձևով՝

$$Var(\hat{\varepsilon}_i) = \frac{1}{n-2} \sum_{i=1}^n (\hat{\varepsilon}_i - \mathbb{E}(\hat{\varepsilon}_i))^2 = \frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{n-2},$$

որում $n-2 = n - (m+1)$ – մոդելի ազատության աստիճանն է, n -ը նմուշի ծավալը կամ դիտարկումների քանակն է, իսկ m -ը՝ մոդելի բացատրող փոփոխականների քանակը, որը տվյալ դեպքում հավասար է 1-ի (համապատասխանաբար, $m+1$ -ը էկոնոմետրիկ մոդելի գնահատվող պարամետրերի քանակն է):

Նշանակենք՝ $var(\hat{\varepsilon}_i) = S_{\hat{\varepsilon}}^2$: $S_{\hat{\varepsilon}}^2$ -ը տեսական մոդելի ε_i -երի σ_ε^2 անհայտ վարիացիայի վիճակագրական անշեղ գնահատականն է: Հետևաբար $\widehat{\beta_0}$ ու $\widehat{\beta_1}$ գործակիցների վարիացիաները և ստանդարտ սխալները կարելի մոտարկել հետևյալ բանաձևերով՝

$$var(\widehat{\beta_1}) \approx \frac{S_{\hat{\varepsilon}}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} = S_{\hat{\beta_1}}^2 \quad (38), \quad S_{\hat{\beta_1}} \approx \sqrt{\frac{S_{\hat{\varepsilon}}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}, \quad (39)$$

$$var(\widehat{\beta_0}) \approx S_{\hat{\varepsilon}}^2 \cdot \left(\frac{1}{n} + \frac{(\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right) = S_{\hat{\beta_0}}^2, \quad (40)$$

$$S_{\hat{\beta_0}} \approx \sqrt{S_{\hat{\varepsilon}}^2 \cdot \left(\frac{1}{n} + \frac{(\bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)}: \quad (41)$$

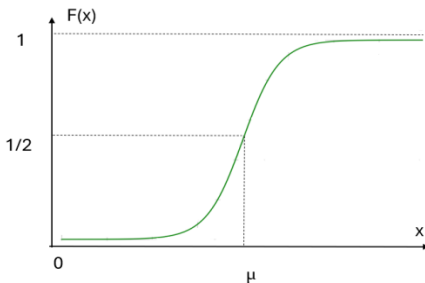
$S_{\hat{\varepsilon}}^2$ -ն ռեգրեսիայի մոդելի չբացատրվող մասն է և ցույց է տալիս, թե նմուշային դիտարկումները որքան են ցրված ռեգրեսիայի գնահատված էմպիրիկ հավասարման շուրջը: $S_{\hat{\varepsilon}} = \sqrt{S_{\hat{\varepsilon}}^2} = \sqrt{\frac{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}{n-2}}$ -ն կոչվում

է գնահատման ստանդարտ սխալ (Standard Error of the Estimate, SEE): Եթե Գաուս-Մարկովի բոլոր (այդ թվում՝ 5-րդ) պայմանները տեղի ունեն, ապա ռեգրեսիայի էմպիրիկ հավասարմամբ որոշված գծից $\pm 2 \cdot S_{\varepsilon}$ -ով հեռացած միջակայքում ընկած է նմուշային դիտարկումներով որոշված կետերի 95, իսկ $\pm S_{\varepsilon}$ -ով հեռացած միջակայքում՝ 68 տոկոսը: Որքան փոքր է գնահատման ստանդարտ սխալը, այնքան կետերի ցրումը գնահատված գծի շուրջը փոքր է, և գիծը որակով է գնահատված:

Ներդիր 10. Նորմալ բաշխման մի քանի կարևոր հատկություններ

X պատահական մեծությունը (այսուհետ՝ ՊՄ) նորմալ է բաշխված μ ու σ պարամետրերով, եթե դրա բաշխման խտության ֆունկցիան որոշվում է հետևյալ բանաձևով՝ $f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, որում $E(X) = \mu$, $\text{var}(X) = \sigma^2$: Նորմալ բաշխման կուտակային $F(x)$ ֆունկցիայի գրաֆիկն ունի հետևյալ տեսքը՝

Նորմալ բաշխման կուտակային ֆունկցիան



Նորմալ բաշխման խտության ֆունկցիան հասնում է իր առավելագույն արժեքին $x = \mu$ կետում և սիմետրիկ է $x = \mu$ ուղղի, այսինքն՝ ՊՄ միջինի նկատմամբ: Ըստ որում՝ նորմալ բաշխման խտության կորի տակի մակերեսը հավասար է մեկի՝ $\int_{-\infty}^{+\infty} f(x) dx = 1$: Նաև $\int_a^b f(x) dx = P(a \leq x \leq b)$, այ-

սինքն՝ խտության ֆունկցիայի միջոցով կարելի է որոշել X -ի՝ տրված $[a, b]$ միջակայքից որևէ արժեք ընդունելու հավանականությունը: $f(x)$ ֆունկցիայի գրաֆիկի թեքության կտրուկ փոփոխությունը տեղի է ունենում $x = \mu - \sigma$ և $x = \mu + \sigma$ արժեքների համար, և որքան փոքր է ՊՄ σ ստանդարտ շեղումը (նաև վարիացիան իր միջինի շուրջը), այնքան մեծ են ՊՄ արժեքների խտությունը կամ կուտակվածությունն իր միջինի շուրջը:

Երևք սիզմայի կանոնը

Նորմալ բաշխում ունեցող ՊՄ արժեքների շեղումը դրա միջին արժե-

քից 99.7 % հավանականությամբ չի գերազանցում երեք սիգման: Այսինքն՝ դրա արժեքների 99.7 %-ն ընկած են $[\mu-3\sigma; \mu+3\sigma]$ միջակայքում: Ըստ որում՝

1. $P(|X-\mu|<\sigma)=P(\mu-\sigma<X<\mu+\sigma)=0,6826$ – ՊՄ արժեքների մոտավորապես 68 %-ը դրա միջին արժեքից ± 1 անգամ ստանդարտ շեղման սահմաններում են,

2. $P(|X-\mu|<2\sigma)=P(\mu-2\sigma<X<\mu+2\sigma)=0,9545$ – ՊՄ արժեքների մոտավորապես 95 %-ը դրա միջին արժեքից ± 2 անգամ ստանդարտ շեղման սահմաններում են,

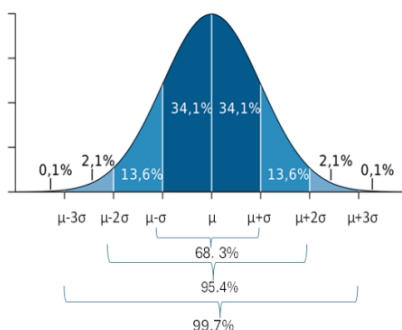
3. $P(|X-\mu|<3\sigma)=P(\mu-3\sigma<X<\mu+3\sigma)=0,9973$ – ՊՄ արժեքների մոտավորապես 99.7 %-ը դրա միջին արժեքից ± 3 անգամ ստանդարտ շեղման սահմաններում են:

Նորմալ բաշխում ունեցող ՊՄ երրորդ կարգի կենտրոնական մոմենտը, որը բնութագրում է դրա միջինի շուրջը արժեքների բաշխվածության ոչ սիմետրիկությունը, հավասար է գրոյի՝ $\mu_3(X) = E(X - E(X))^3 = E(X - \mu)^3 = 0$: Համապատասխանաբար ասիմետրիայի գործակիցը (Skewness)՝

$$S(X) = \frac{\mu_3(X)}{\sigma^3(X)} = 0:$$

Նորմալ բաշխում ունեցող ՊՄ չորրորդ կարգի կենտրոնական մոմենտը, որը բնութագրում է դրա միջինի շուրջը արժեքների կուտակվածությունը (Kurtosis)՝ $\mu_4(X) = E(X - E(X))^4 = E(X - \mu)^4 = 3$: Համապատասխանաբար կուտակվածության գործակիցը (Excess Kurtosis)՝ $K(X) = \frac{\mu_4(X)}{\sigma^4(X)} - 3 = 0$:

Նորմալ բաշխման խտության ֆունկցիայի գրաֆիկը



7.2. Ռեգրեսիայի տեսական հավասարման պարամետրերի մասին վարկածների ստուգում. t թեստ

Ենթադրենք՝ Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն, այսինքն՝ ունենք գծային միագործոն ռեգրեսիայի նորմալ դասական մոդել: Տեսական մոդելը գնահատելուց ու էմպիրիկ մոդելի գործակիցները ստանալուց հետո անհրաժեշտ է ստուգել գնահատված մոդելի վավերականությունն ու որակը: Դրա բաղկացուցիչներից մեկը ստուգումն է, թե արդյոք տեսական մոդելի անհայտ պարամետրերը զրոյից նշանակալիորեն տարբերվող մեծություններ են:

Տեսական մոդելի β_1 պարամետրի պարագայում դա նշանակում է, որ պետք է ստուգել, թե տեսական մոդելում իրականում x -ը y -ին բացատրում է, թե՞ ոչ, կամ, որ նույնն է, y -ը գծային կախվածության մեջ է x -ից, թե՞ ոչ: β_0 պարամետրի պարագայում դա նշանակում է, որ պետք է ստուգել, թե արդյոք ռեգրեսիայի տեսական հավասարումն անցնում է կոորդինատային հարթության սկզբնակետով:

Դրվում են հետևյալ վիճակագրական զրոյական վարկածները՝

$$H_0: \beta_0 = 0 \text{ և } H_0: \beta_1 = 0:$$

Մրցակցող վարկածներն են՝

$$H_1: \beta_0 \neq 0 \text{ և } H_1: \beta_1 \neq 0:$$

Այս զրոյական վարկածները կոչվում են **ռեգրեսիայի հավասարման գործակիցների վիճակագրական նշանակալիության մասին վարկածներ**: Եթե H_0 -ն ընդունվի, ապա կասենք, որ ա) ռեգրեսիայի տեսական հավասարման անհայտ β_1 գործակիցը զրոյից նշանակալիորեն չի տարբերվում, և x -ը y -ին չի բացատրում, կամ, որ նույնն է, ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_1$ գործակիցը վիճակագրորեն նշանակալի չէ, բ) β_0 -ի պարագայում եթե H_0 -ն ընդունվի, ապա կասենք, որ ռեգրեսիայի տեսական հավասարումը ազատ անդամ չի պարունակում, կամ, որ նույնն է, ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_0$ գործակիցը վիճակագրորեն նշանակալի չէ: Եթե H_0 -ն ընդունվի, ապա ռեգրեսիայի էմպիրիկ հավասարումը վավերական չի լինի, ուստի այն կիրառելի չի լինի:

Եթե H_0 -ն մերժվի, կասենք, որ ա) ռեգրեսիայի էմպիրիկ հավասարման $\widehat{\beta}_1$ գործակիցը վիճակագրորեն նշանակալի է, և տեսական հավասարման մեջ x -ը բացատրում է y -ին, բ) էմպիրիկ հավասարման $\widehat{\beta}_0$ գործակիցը վիճակագրորեն նշանակալի է, և ռեգրեսիայի տեսական հավասարումն ունի ազատ անդամ:

Ստուգելու համար, քե արդյոք տեսական մոդելի անհայտ β_1 պարամետրը իրականում հավասար է զրոյի, օգտվում են դրա անշեղ գնահատականից՝ էմպիրիկ մոդելի $\widehat{\beta}_1$ գործակցից: Իսկ ստուգելու համար, քե արդյոք տեսական հավասարումն իրականում անցնում է սկզբնականով, օգտվում են տեսական մոդելի անհայտ β_0 պարամետրի անշեղ գնահատականից՝ էմպիրիկ մոդելի $\widehat{\beta}_0$ գործակցից:

Ներդիր 11. Վիճակագրական վարկածների ստուգման որոշ հիմնական դրույթներ

Վիճակագրական վարկածը որոշակի ենթադրություն է ուսումնասիրվող գլխավոր համախմբության հատկությունների և բնութագրերի վերաբերյալ: Վիճակագրական վարկածների ստուգումը հնարավորություն է տալիս որոշելու, քե արդյոք ուսումնասիրվող տվյալները համապատասխանում են առաջ քաշված վարկածին, և արդյոք վարկածի ու նմուշային դիտարկումների միջև անհամապատասխանությունը կարող է վերագրվել պատահական սխալներին, որոնք պայմանավորված են պատահական ընտրության մեխանիզմով: Այս խնդիրը լուծվում է մաթեմատիկական վիճակագրության հատուկ մեթոդների օգնությամբ:

Ստուգվող հիմնական վիճակագրական վարկածը կոչվում է զրոյական վարկած և նշանակվում է H_0 : Չրոյական վարկածի հետ մեկտեղ դիտարկվում է նաև այլընտրանքային կամ մրցակցող H_1 վարկածը, որը կրնոունվի, եթե H_0 -ն մերժվի: Մրցակցող վարկածի ընտրությունը կախված է ստուգվող խնդրի դրվածքից: Օրինակ, եթե առաջադրված զրոյական վարկածն է՝ $H_0 : \beta=0$, մրցակցող վարկածը կարող է լինել՝ $H_1 : \beta \neq 0$ կամ $H_1 : \beta > 0$ կամ $H_1 : \beta < 0$:

Վարկածը ստուգելիս նմուշային տվյալները կարող են հակասել զրոյական վարկածին: Այդ դեպքում այն կմերժվի: Հակառակ դեպքում զրոյական վարկածը չի մերժվի: Նմուշային տվյալների վրա հիմնված վիճակագրական վարկածի ստուգումը պարունակում է սխալ որոշում կայացնելու ռիսկ: Ըստ որում, հնարավոր են երկու տիպի սխալներ՝ ա) առաջին տիպ. մերժվել է իրականում ճիշտ զրոյական վարկածը, բ) երկրորդ տիպ. ընդունվել է զրոյական վարկածը, որն իրականում պետք է չընդունվեր, քանի որ ճիշտ է մրցակցող վարկածը: Առաջին տիպի սխալ թույլ տալու հավանականությունը կոչվում է **արժեքականության մակարդակ** և սովորաբար նշանակվում է α -ով: Երկրորդ տիպի սխալ թույլ տալու

հավանականությունը նշանակվում է β -ով, իսկ երկրորդ տիպի սխալ թույլ չտալու հավանականությունը՝ **1- β , կոչվում է վիճակագրական չափանիշի կամ վիճականու հզորություն**:

H_0 վարկածը ստուգելու համար ընտրվում է որոշակի հայտնի բաշխում ունեցող պատահական մեծություն, որը կոչվում է չափանիշ կամ վիճականի: Վիճականու բոլոր հնարավոր արժեքների բազմությունը բաժանվում է երկու չհատվող ենթաբազմությունների, որոնցից մեկը պարունակում է վիճականու այն արժեքները, որոնց դեպքում H_0 վարկածը մերժվում է (կոչվում է **կրիտիկական տիրույթ**), իսկ մյուսը՝ վիճականու այն արժեքները, որոնց դեպքում H_0 վարկածը չի մերժվում (կոչվում է **թույլատրելի տիրույթ**): Վիճականու կրիտիկական ու թույլատրելի տիրույթները միմյանցից տարանջատող արժեքը կոչվում է **կրիտիկական կետ**: Կախված նրանից, թե ինչպիսին է մրցակցող վարկածը, կրիտիկական տիրույթը կարող է լինել **միակողմանի** կամ **երկկողմանի**: Օրինակ, եթե $H_0: \beta=0$ և $H_1: \beta \neq 0$, ապա կրիտիկական տիրույթը երկկողմանի է, իսկ եթե $H_0: \beta=0$ և $H_1: \beta < 0$ ՝ միակողմանի:

Վարկածի ստուգման համար մմուշային տվյալների հիման վրա հաշվարկվում է ընտրված վիճականու արժեքը, որը կոչվում է **վիճականու դիտարկված արժեք**: Ընտրվում է հնարավորինս փոքր α հավանականություն (0.01, 0.05, կամ 0.1), որը գրոյական վարկածը մերժելիս սխալվելու հավանականությունը կամ արժեքականության մակարդակն է: Հետևաբար գրոյական վարկածը մերժելիս սխալված չլինելու կամ, որ նույնն է, ճիշտ գրոյական վարկածն ընդունելու հավանականությունը հավասար է $\gamma=1-\alpha$:

Եթե դիտարկված արժեքն ընկած է կրիտիկական տիրույթում, ապա H_0 վարկածը մերժելիս սխալվելու հավանականությունը բավականաչափ փոքր է, և այն մերժվում է տվյալ արժեքականության մակարդակում: Եթե դիտարկված արժեքն ընկած է թույլատրելի տիրույթում, ապա ճիշտ H_0 վարկածն ընդունելու հավանականությունը բավականաչափ մեծ է, և այն չի մերժվում (ընդունվում է) տվյալ արժեքականության մակարդակում:

Դրված գրոյական վարկածները (H_0) ստուգելու համար օգտագործվում է t վիճականին, որը որոշվում է որպես ռեգրեսիայի էմպիրիկ հավասարման գործակցի հարաբերություն իր ստանդարտ սխալին՝

$$t_{\widehat{\beta_0}} = \frac{\widehat{\beta_0}}{S_{\widehat{\beta_0}}} \quad (42) \quad \text{և} \quad t_{\widehat{\beta_1}} = \frac{\widehat{\beta_1}}{S_{\widehat{\beta_1}}}: \quad (43)$$

Եթե Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն, և մոդելի պատահական շեղումներն ունեն նորմալ բաշխում, ապա $\widehat{\beta_0}$ և $\widehat{\beta_1}$ գործակիցները նորմալ բաշխում ունեն: Այդ դեպքում (42) և (43) բանաձևերով որոշվող t վիճականին ունի Ստյուդենտի բաշխում $\nu = n - (m + 1)$ ազատության աստիճանով (որում n -ը դիտարկումների, իսկ m -ը՝ բացատրող փոփոխականների քանակն է): Այդ պատճառով այն հաճախ անվանվում է Ստյուդենտի t վիճականի: Նկատի ունենալով, որ մրցակցող վարկածները հետևյալն են՝ $H_1: \beta_0 \neq 0$ և $H_1: \beta_1 \neq 0$ վիճականու կրիտիկական տիրույթը երկ-կողմանի է:

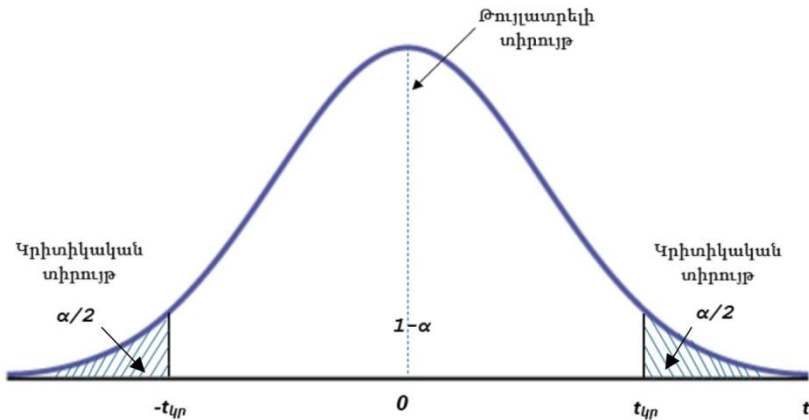
Ընդունենք, որ α -ն արժեքականության մակարդակն է, իսկ $t_{\text{կր}}$ -ը՝ Ստյուդենտի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքն α արժեքականության մակարդակով և $\nu = n - p$ ազատության աստիճանով (պարզ ռեգրեսիայի դեպքում $\nu = n - 2$): Այդ դեպքում, եթե մնուշային տվյալների հիման վրա (42) և (43) բանաձևերով հաշվարկված t վիճականու դիտարկված $t_{\widehat{\beta_0}}^{\text{նիս}}$ ու $t_{\widehat{\beta_1}}^{\text{նիս}}$ արժեքներն ընկած են Գծանկար 8-ի շեղագծերով նշված տիրույթներում, ապա H_0 -ն մերժվում է. այն մերժելիս սխալված լինելու հավանականությունը փոքր է. $P(t^{\text{նիս}} < -t_{\text{կր}}) = P(t^{\text{նիս}} > t_{\text{կր}}) = \frac{\alpha}{2}$: Այսինքն՝ H_0 -ն մերժելու ընդհանուր, երկկողմանի հավանականությունը հավասար է α -ի: Իսկ եթե դրանք գտնվում են գծանկարի սպիտակ տիրույթում, H_0 -ն չի մերժվում, ընդունվում է. այն մերժելիս սխալված լինելու հավանականությունը մեծ է՝ $P(-t_{\text{կր}} \leq t^{\text{նիս}} \leq t_{\text{կր}}) = 1 - \alpha = \gamma$:

Վարկածը ստուգվում է t վիճականու (42) և (43) բանաձևերով հաշվարկված արժեքները արժեքականության ընտրված α մակարդակում Ստյուդենտի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքի ($t_{\frac{\alpha}{2}, n-2}^{\text{կր}}$) հետ համեմատելու միջոցով: Կրիտիկական արժեքը ստացվում է Ստյուդենտի բաշխման կրիտիկական արժեքների աղյուսակից՝ ընտրելով α արժեքականության մակարդակն ու ν ազատության աստիճանը: Եթե մնուշային տվյալների հիման վրա (42) և (43) բանաձևերով հաշվարկված t վիճականու դիտարկված $t_{\widehat{\beta_0}}^{\text{նիս}}$ և $t_{\widehat{\beta_1}}^{\text{նիս}}$ արժեքների համար՝

$$\left| t_{\hat{\beta}_0}^{\eta_{\text{իտ}}} \right| > t_{\frac{\alpha}{2}, n-2}^{\text{կր}} \quad (44) \text{ և } \left| t_{\hat{\beta}_1}^{\eta_{\text{իտ}}} \right| > t_{\frac{\alpha}{2}, n-2}^{\text{կր}} \quad (45) \rightarrow H_0\text{-ն մերժվում է:}$$

$$\left| t_{\hat{\beta}_0}^{\eta_{\text{իտ}}} \right| \leq t_{\frac{\alpha}{2}, n-2}^{\text{կր}} \quad (46) \text{ և } \left| t_{\hat{\beta}_1}^{\eta_{\text{իտ}}} \right| \leq t_{\frac{\alpha}{2}, n-2}^{\text{կր}} \quad (47) \rightarrow H_0\text{-ն չի կարող մերժվել. այն ընդունվում է:}$$

Գծանկար 8. t վիճականու բաշխման խրուխյան ֆունկցիան, թույլատրելի ու կրիտիկական տիրույթները և կրիտիկական արժեքները



Եթե տեղի ունի (44)-ը կամ (45)-ը, ապա H_0 -ի մերժվելը նշանակում է, որ արժեքականության α մակարդակում տեսական հավասարման տվյալ β անհայտ պարամետրը նշանակալիորեն զրոյից տարբերվող մեծություն է, իսկ ռեգրեսիայի էմպիրիկ հավասարման համապատասխան $\hat{\beta}$ գնահատված գործակիցը վիճակագրորեն նշանակալի գնահատական է: β_1 -ի պարագայում դա նշանակում է, որ տեսական մոդելում x -ը բացատրում է կախյալ փոփոխականին՝ y -ին: β_0 -ի պարագայում նշանակում է, որ ռեգրեսիայի տեսական հավասարումն ունի ազատ անդամ, սկզբնակետով չի անցնում:

Եթե տեղի ունի (46)-ը կամ (47)-ը, նշանակում է արժեքականության α մակարդակում տեսական հավասարման տվյալ β անհայտ պարամետրը զրոյից նշանակալիորեն չի տարբերվում, իսկ ռեգրեսիայի էմպիրիկ հավասարման համապատասխան $\hat{\beta}$ գնահատված գործակիցը վիճակագրորեն նշանակալի չէ: β_1 -ի պարագայում դա նշանակում է, որ տեսական մոդելում x -ը չի բացատրում կախյալ փոփոխականին՝ y -ին: Այս դեպքում ռեգրեսիայի գնահատված հավա-

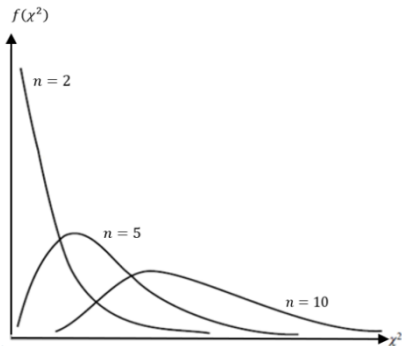
սարումը կիրառելի չէ: β_0 -ի պարագայում նշանակում է, որ ռեգրեսիայի տեսական հավասարումը չունի ազատ անդամ:

Չրոյական վարկածի ստուգման համար կարելի է կիրառել կոպիտ կանոն: Նկատի ունենալով, որ Ստյուդենտի բաշխման $t_{\alpha/2, n-2}^{\text{կր}}$ կրիտիկական արժեքները 80-ից մեծ ազատության աստիճանների ու $\alpha=0.05$ ($\alpha/2=0.025$) արժեքականության մակարդակի համար փոքր են երկուսից, ապա եթե $|t^{\text{դիտ}}| \geq 2$, կարելի է զրոյական վարկածը մերժել, իսկ հակառակ դեպքում՝ ընդունել:

Չրոյական վարկածի ստուգման ընթացքում առաջանում է α արժեքականության մակարդակի ընտրության խնդիր: Այդ պատճառով կիրառվում է մեկ այլ մոտեցում ևս, որով որոշվում է վիճականու դիտարկված արժեքն ընդունելու իրական, ճշգրիտ հավանականությունը: Այն կոչվում է p արժեք, որը հայտնի է մաս որպես դիտարկված կամ ճշգրիտ նշանակալիության մակարդակ (p -value, significance) և առաջին տիպի սխալ թույլ տալու (զրոյական վարկածը մերժելու դեպքում սխալվելու) ճշգրիտ հավանականությունն է: Այլ խոսքով՝ p արժեքը սահմանվում է որպես ամենացածր նշանակալիության մակարդակ, որի դեպքում զրոյական վարկածը պետք է մերժվի: Գործնական վերլուծություններում հաճախ p արժեքը համեմատության մեջ է դրվում 0.05 -ի հետ: Սա համապատասխանում է $100\% \cdot \alpha = 100\% \cdot 0.05 = 5\%$ արժեքականության մակարդակին: Ըստ ընդունված մոտեցման, եթե $p\text{-value} \geq 0.05$, ապա զրոյական վարկածը մերժելու դեպքում սխալվելու հավանականությունը մեծ է, և զրոյական վարկածն ընդունվում է: Իսկ եթե $p\text{-value} < 0.05$, ապա զրոյական վարկածը մերժելու դեպքում սխալվելու հավանականությունը փոքր է, և այն մերժվում է: p արժեքը հաճախ նաև անվանում են «զրոյական վարկածի հավանականություն»: Այսպիսով՝ որքան փոքր է p արժեքը, այնքան քիչ հավանական է զրոյական վարկածը:

Ներդիր 12. Վիճակագրական վարկածների ստուգման համար օգտագործվող որոշ բաշխումների հիմնական հատկությունները

χ^2 բաշխում: Ենթադրենք ունենք $i=1,2,\dots,n$ քվով նորմալ բաշխված անկախ պատահական X_i մեծություններ՝ m_i մաթ. սպասումներով ու σ_i ստանդարտ շեղումներով: Այսինքն՝ $X_i \sim N(m_i, \sigma_i)$: Այդ դեպքում $Z_i = \frac{X_i - m_i}{\sigma_i}$, $i=1,2,\dots,n$ պատահական մեծություններն ունեն ստանդարտ նորմալ բաշխում հետևյալ պարամետրերով՝ $Z_i \sim N(0, 1)$:

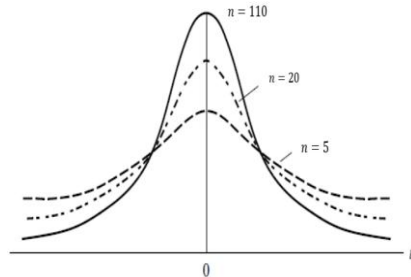


Ստանդարտ նորմալ բաշխում ունեցող անկախ n պատահական մեծությունների քառակուսիների գումար հանդիսացող $\chi^2_n = \sum_{i=1}^n Z_i^2$ պատահական մեծությունն ունի χ^2 բաշխում՝ n ազատության աստիճանով:

χ^2 բաշխում ունեցող պատահական մեծության միջինը՝ $\mu = n$, վարիացիան՝ $\sigma^2 = 2n$, իսկ ստանդարտ շեղումը՝ $\sigma = \sqrt{2n}$: Դրա հավանականության խտության ֆունկցիայի գրաֆիկը գտնվում է կոորդինատային հարթության առաջին քառորդում և ունի ասիմետրիկ տեսք՝ երկար աջ «պոչով»:

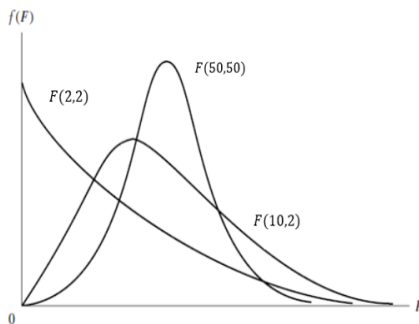
Ստյուդենտի t բաշխում:

Ենթադրենք՝ Z_1 -ը ստանդարտ նորմալ բաշխում ունեցող պատահական մեծություն է, իսկ Z_2 -ը՝ χ^2 բաշխում ունեցող պատահական մեծություն՝ n ազատության աստիճանով, ըստ որում՝ Z_1 -ը և Z_2 -ը անկախ են: Այդ դեպքում $t = \frac{Z_1}{\sqrt{Z_2/n}}$ պատահական մեծությունն ունի Ստյուդենտի t բաշխում՝ n ազատության աստիճանով: t բաշխման հավանականությունների խտության ֆունկցիան սիմետրիկ է իր միջինի նկատմամբ:



Այն նման է նորմալ բաշխման խտության ֆունկցիային, բայց ավելի հարթ է կամ տափակ է: Ազատության աստիճանի աճին գուցենքաց է բաշխումը մոտենում է նորմալ բաշխմանը: Ստյուդենտի բաշխման միջինը՝ $\mu = 0$, վարիացիան՝ $\sigma^2 = n/n - 2$, իսկ ստանդարտ շեղումը՝ $\sigma = \sqrt{\frac{n}{n-2}}$:

Ֆիշերի F բաշխում: Ենթադրենք՝ Z_1 -ն ու Z_2 -ը անկախ, χ^2 բաշխում ունեցող պատահական մեծություններ են՝ համապատասխանաբար n ու m ազատության աստիճաններով: Այդ դեպքում $F = \frac{Z_1/n}{Z_2/m}$ պատահական մեծությունն ունի Ֆիշերի բաշխում՝ n ու m ազատության աստիճաններով: F բաշխման հավանականության խտության ֆունկցիայի գրաֆիկը ևս ասիմետրիկ է, երկար աջ «պոչով»: Ազատության աստիճանների աճի հետ այն ևս մոտենում է նորմալ բաշխմանը: Ֆիշերի բաշխման միջինը որոշված է, երբ $m > 2$ և հավասար է՝ $\mu = \frac{m}{m-2}$: Ֆիշերի բաշխման վարիացիան որոշված է, երբ $m > 4$, այն հավասար է՝ $\sigma^2 = \frac{2m^2(n+m-2)}{m(n-2)^2(m-4)}$:



7.3. Ռեգրեսիայի տեսական հավասարման պարամետրերի միջակայքային գնահատականները

Ենթադրենք՝ ունենք գծային միագործոն ռեգրեսիայի նորմալ դասական մոդել, և էմպիրիկ հավասարման $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցները նորմալ են բաշխված և վիճակագրորեն նշանակալի գնահատականներ են: $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցները տեսական մոդելի անհայտ պարամետրերի **կետային** գնահատականներն են:

Կարելի է որոշել նաև տեսական մոդելի անհայտ β պարամետրերի **միջակայքային** գնահատականները: Միջակայքային գնահատականը այն (θ_1, θ_2) միջակայքն է, որում որևէ Y հավանականու-

թյամբ ընկած է տեսական մոդելի անհայտ β պարամետրը: Այդ (θ_1, θ_2) միջակայքը կոչվում է **վստահելիության միջակայք**, իսկ Y -ն կոչվում է **վստահելիության հավանականություն**:

Տեսական մոդելի անհայտ β պարամետրերի **միջակայքային** գնահատականները որոշելու համար դրվում է գրոյական վարկած այն մասին, որ էմպիրիկ մոդելի գնահատված $\hat{\beta}$ -երը հավասար են տեսական մոդելի β անհայտ պարամետրերին: Այդ վարկածներն են՝ $H_0: \widehat{\beta}_0 = \beta_0$ և $H_0: \widehat{\beta}_1 = \beta_1$: Միջակցող վարկածներն են՝ $H_1: \widehat{\beta}_0 \neq \beta_0$ և $H_1: \widehat{\beta}_1 \neq \beta_1$: Ջրոյական վարկածները ստուգելու համար օգտագործվում է Ստյուդենտի բաշխում ունեցող t վիճականին, որն այս դեպքում որոշվում է հետևյալ բանաձևով՝

$$t_{\widehat{\beta}_0} = \frac{\widehat{\beta}_0 - \beta_0}{S_{\widehat{\beta}_0}} \quad (48)$$

$$\text{և } t_{\widehat{\beta}_1} = \frac{\widehat{\beta}_1 - \beta_1}{S_{\widehat{\beta}_1}}: \quad (49)$$

Դիտարկենք $\widehat{\beta}_1$ -ին վերաբերող գրոյական վարկածը: Վերը արդեն ցույց ենք տվել, որ եթե H_0 -ն չի մերժվում, ընդունվում է, ապա՝ $P(-t_{\text{կր}} \leq t^{\text{նիւ}} \leq t_{\text{կր}}) = 1 - \alpha = \gamma$: Տեղադրենք (49)-ը այս արտահայտության մեջ.

$$P\left(-t_{\text{կր}} \leq \frac{\widehat{\beta}_1 - \beta_1}{S_{\widehat{\beta}_1}} \leq t_{\text{կր}}\right) = 1 - \alpha \rightarrow$$

$$P(-t_{\text{կր}} \cdot S_{\widehat{\beta}_1} \leq \widehat{\beta}_1 - \beta_1 \leq t_{\text{կր}} \cdot S_{\widehat{\beta}_1}) = 1 - \alpha \rightarrow$$

$$P(\widehat{\beta}_1 - t_{\text{կր}} \cdot S_{\widehat{\beta}_1} \leq \beta_1 \leq \widehat{\beta}_1 + t_{\text{կր}} \cdot S_{\widehat{\beta}_1}) = 1 - \alpha: \quad (50)$$

Նշանակենք. $\theta_1 = \widehat{\beta}_1 - t_{\text{կր}} \cdot S_{\widehat{\beta}_1}$ (51) և $\theta_2 = \widehat{\beta}_1 + t_{\text{կր}} \cdot S_{\widehat{\beta}_1}$ (52) և ձևափոխենք (50)-ը: Կունենանք.

$$P(\theta_1 \leq \beta_1 \leq \theta_2) = 1 - \alpha = \gamma: \quad (53)$$

Ստացվեց, որ տեսական մոդելի անհայտ β_1 պարամետրը $\gamma = 1 - \alpha$ հավանականությամբ ընկած է (θ_1, θ_2) վստահելիության միջակայքում, որը կարելի է հաշվարկել (51) ու (52) բանաձևերով, եթե հայտնի են էմպիրիկ մոդելի $\widehat{\beta}_1$ գործակիցը և դրա ստանդարտ սխալը, իսկ t վիճականու կրիտիկական արժեքը որոշված է Ստյուդենտի բաշխման աղյուսակից:

Նույն տրամաբանությամբ կարելի է դուրս բերել նաև տեսական մոդելի անհայտ β_0 պարամետրի վստահելիության (θ_1, θ_2) միջակայքը.

$$P(\widehat{\beta}_0 - t_{\text{կր}} \cdot S_{\widehat{\beta}_0} \leq \beta_0 \leq \widehat{\beta}_0 + t_{\text{կր}} \cdot S_{\widehat{\beta}_0}) = 1 - \alpha: \quad (54)$$

$$\theta_1 = \widehat{\beta}_0 - t_{\text{կր}} \cdot S_{\widehat{\beta}_0} \quad (55) \text{ և } \theta_2 = \widehat{\beta}_0 + t_{\text{կր}} \cdot S_{\widehat{\beta}_0}: \quad (56)$$

(θ_1, θ_2) վստահելիության միջակայքը որոշելու համար նախ պետք է ընտրել պնդման արժեքականության մակարդակը՝ α -ն: Դա հնարավորություն կտա որոշելու γ վստահելիության հավանականությունը, ինչպես նաև t վիճականու կրիտիկական աղյուսակային արժեքը: Եթե վերցնենք $\alpha=0.05$, ապա՝ $\gamma=1-\alpha=0.95$, հետևաբար տեսական մոդելի β պարամետրի վստահելիության միջակայքը կորոշվի 95 % հավանականությամբ: Այդ դեպքում 95 % հավանականությամբ տեսական մոդելի β պարամետրն ընկած կլինի $\widehat{\beta} \pm t_{\text{կր}} \cdot S_{\widehat{\beta}}$ միջակայքում:

Ներդիր 13. Առանց ազատ անդամի գծային ռեգրեսիայի մոդելի գնահատման առանձնահատկությունները

Ի՞նչ անել, եթե գծային ռեգրեսիայի մոդելում ազատ անդամը վիճակագրորեն նշանակալի չէ, կա՞ն տեսությունը, կա՞ն էլ համատեքստը ենթադրում են, որ x և y փոփոխականների միջև կապն այնպիսին է, որ երբ $x=0$, ապա $y=0$: Արդյո՞ք ռեգրեսիայի մոդելում ազատ անդամ չմտցնելը հնարավորություն է տալիս գնահատելու այն ՓԲԵ-ով ու ստանդարտ մոտեցումներով ստուգելու գնահատված հավասարման որակը: Հեշտության համար դիտարկենք այս խնդիրը պարզ գծային ռեգրեսիայի դեպքի համար:

Փաստորեն՝ ռեգրեսիայի տեսական մոդելն առանց ազատ անդամի ստանում է այս տեսքը. $y = \beta_1 x + \varepsilon$, էմպիրիկ մոդելը՝ $y = \hat{y} + \hat{\varepsilon} = \widehat{\beta}_1 x + \hat{\varepsilon}$, իսկ էմպիրիկ հավասարումը՝ $\hat{y} = \widehat{\beta}_1 x$: Այս հավասարումը կարելի է գնահատել ՓԲԵ-ով, սակայն $\widehat{\beta}_1$ -ի որոշման բանաձևը կփոխվի: ՓԲԵ-ն այս դեպքում մինիմիզացնում է $Q = \sum_{i=1}^n (\hat{\varepsilon}_i)^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \widehat{\beta}_1 x_i)^2$ ֆունկցիան: Էքստրեմումը գտնվում է հետևյալ կերպ.

$$\frac{\partial Q}{\partial \widehat{\beta}_1} = \frac{\partial (\sum_{i=1}^n (y_i - \widehat{\beta}_1 x_i)^2)}{\partial \widehat{\beta}_1} = 0 \rightarrow -2 \sum_{i=1}^n (y_i - \widehat{\beta}_1 x_i) x_i = 0 \rightarrow$$

$$\sum_{i=1}^n y_i x_i - \widehat{\beta}_1 \sum_{i=1}^n x_i^2 = 0 \rightarrow \widehat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i}{\sum_{i=1}^n x_i^2}:$$

Այս արդյունքը տարբերվում է ՓԲԵ-ով որոշված $\widehat{\beta}_1$ -ի արժեքից, երբ մոդելն ունի ազատ անդամ: Եթե իրականում մոդելը պետք է

ունենա ազատ անդամ, բայց գնահատվել է առանց ազատ անդամի, ապա $\widehat{\beta}_1$ -ն անշեղ գնահատական չի լինի տեսական մոդելի անհայտ պարամետրի համար, ուստի այսպիսի մոդելի գնահատումը պետք է ըստ ամենայնի հիմնավորված լինի:

Մյուս կողմից, առանց ազատ անդամի մոդելի պարագայում դետերմինացիայի գործակիցը կորցնում է իր ստանդարտ բովանդակությունը, և դրա մեկնաբանությունն անիմաստ է դառնում: Եթե ազատ անդամով ռեգրեսիայի համար ունեինք $R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} =$

$$= 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

ապա առանց ազատ անդամի ռեգրեսիայի համար ունենք $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n y_i^2} = 1 - \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sum_{i=1}^n y_i^2}.$

Այսինքն՝ կոտորակի համարիչը նույնն է, բայց հայտարարը կարող է լինել y -ի՝ իր միջինի նկատմամբ ընդհանուր վարիացիա (TSS) միայն այն դեպքում, եթե $\bar{y} = 0$, այնինչ ռեգրեսիայի մոդելի գնահատումը հենց y -ի սպասվող արժեքի գնահատումն է՝ կախված բացատրող փոփոխականներից: Նշենք նաև, որ առանց ազատ անդամի ռեգրեսիայի դեպքում R^2 -ն կարող է լինել բացասական:

Այս նկատառումներով, ինչպես նաև հաշվի առնելով, որ ազատ անդամը մոդելների մեծ մասում տնտեսագիտական մեկնաբանություն չի ունենում, հաճախ գնահատված էմպիրիկ հավասարման ազատ անդամի վիճակագրական ոչ նշանակալիությունը հաշվի չի առնվում:

7.4. Ռեգրեսիայի գործակիցների վիճակագրական նշանակալիության ստուգման օրինակ

Ենթադրենք՝ ընկերությունն ուզում է վերանայել աշխատանքի վարձատրության իր մոտեցումները և ուսումնասիրում է, թե իր աշխատակիցների՝ ընկերությունում ունեցած աշխատանքային փորձի տևողությունն ինչպես է ազդում աշխատավարձի միջին մակարդակի վրա: Հատկորոշվել է գծային պարզ ռեգրեսիայի տեսական մոդել. $y = \beta_0 + \beta_1 \cdot x + \varepsilon$, որում y -ը ընկերության աշխատակցի ամսական աշխատավարձն է (1000 ՀՀ դրամ), իսկ x -ը՝ ընկերությունում ունեցած աշխատանքային փորձի տևողությունը (տարի): Ընկերության

250 աշխատակիցների մասին հավաքագրվել է նմուշային տեղեկատվություն, և մոդելը պարամետրացվել է ՓԶԵ-ով: Գնահատված էմպիրիկ հավասարումն է՝ $\hat{y} = 320 + 8.5 \cdot x$: Մեկնաբանությունը հետևյալն է. եթե աշխատակիցը նոր է ընդունվում տվյալ ընկերությունում աշխատանքի (աշխատանքային փորձի տևողությունը հավասար է զրոյի), ապա նրան առաջարկվող ամսական աշխատավարձի սպասվող կամ միջին արժեքը ընդունվելու պահին 320 հազար դրամ է: Ընկերությունում աշխատելու ամեն լրացուցիչ մեկ տարին ավելացնում է աշխատակցի վարձատրության սպասվող կամ միջին արժեքը 8500 դրամով: Պարամետրացման փուլից հետո անհրաժեշտ է դետերմինացիայի գործակցի միջոցով ստուգել գնահատված էմպիրիկ հավասարման որակը, ինչպես նաև գնահատված գործակիցների վիճակագրական նշանակալիությունն ու տեսական հավասարման անհայտ պարամետրերի վստահելիության միջակայքերը: Ենթադրենք՝ Գաուս-Մարկովի բոլոր պայմանները, ներառյալ պատահական շեղումների նորմալ բաշխումը, տեղի ունեն: Հաշվարկվել է էմպիրիկ հավասարման պատահական մնացորդների քառակուսիների գումարը. $RSS = \sum_{i=1}^n (\hat{\varepsilon}_i)^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 6.3$, կախյալ փոփոխականի՝ իր միջինի շուրջը ունեցած ընդհանուր վարիացիան. $TSS = \sum_{i=1}^n (y_i - \bar{y})^2 = 14.2$, էմպիրիկ հավասարման գործակիցների ստանդարտ սխալները. $S_{\hat{\beta}_0} = 125$, $S_{\hat{\beta}_1} = 1.6$: Դետերմինացիայի գործակիցը կհաշվենք հետևյալ բանաձևով՝

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{6.3}{14.2} = 1 - 0.44 = 0.56:$$

Հետևաբար ռեգրեսիայի էմպիրիկ հավասարումն ունի ընդունելի, բայց ոչ շատ բարձր որակ: Կախյալ փոփոխականի (աշխատավարձի)՝ իր միջինի շուրջն ունեցած վարիացիայի կեսից ավելին՝ 56 %-ը, բացատրվում է ռեգրեսիայի հավասարմամբ կամ, որ նույնն է, x-ով (ընկերությունում ունեցած աշխատանքային փորձով), իսկ պատահական մնացորդներով բացատրվում է դրա 44 %-ը: Գործակիցների վիճակագրական նշանակալիությունը ստուգելու համար պետք է հաշվել դրանց t վիճականիները.

$$t_{\hat{\beta}_0}^{\eta \text{ hun}} = \frac{\widehat{\beta}_0}{S_{\hat{\beta}_0}} = \frac{320}{125} = 2.56; \quad t_{\hat{\beta}_1}^{\eta \text{ hun}} = \frac{\widehat{\beta}_1}{S_{\hat{\beta}_1}} = \frac{8.5}{1.6} = 5.31:$$

Ստյուդենտի t բաշխման կրիտիկական արժեքները (երկկողմանի տիպով)			
$\alpha/2$ v	0.05	0.025	0.005
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
15	1.753	2.131	2.947
20	1.725	2.086	2.845
25	1.708	2.060	2.787
30	1.697	2.042	2.750
35	1.690	2.030	2.724
40	1.684	2.021	2.704
50	1.676	2.009	2.678
60	1.671	2.000	2.660
70	1.667	1.994	2.648
80	1.664	1.990	2.639
100	1.660	1.984	2.626
200	1.653	1.972	2.601
201+	1.646	1.962	2.581

Դրանից հետո պետք է ընտրել α արժեքականության մակարդակը և որոշել t վիճականու կրիտիկական արժեքը:

Վերցնենք $\alpha=0.05$: Ազատության աստիճանը՝ $v = n - (m + 1) = 250 - 2 = 248$: Ստյուդենտի բաշխման կրիտիկական արժեքների աղյուսակը վկայում է, որ $t_{\frac{\alpha}{2}, n-2}^{\text{կր}} = 1.96$:

Ստյուդենտի բաշխման կրիտիկական արժեքների աղյուսակը հասանելի է հետևյալ QR կոդով:



Կիրառենք թեստը.

$$|t_{\beta_0}^{\text{դիտ}}| = |2.56| = 2.56 > 1.96 = t_{\frac{\alpha}{2}, n-2}^{\text{կր}},$$

$$|t_{\beta_1}^{\text{դիտ}}| = |5.31| = 5.31 > 1.96 = t_{\frac{\alpha}{2}, n-2}^{\text{կր}}:$$

Հետևաբար զրոյական վարկածը մերժվում է թե՛ $\widehat{\beta}_0$, թե՛ $\widehat{\beta}_1$ գործա-

կիցների համար: Եզրակացությունը հետևյալն է. $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները վիճակագրորեն նշանակալի են: Տեսական մոդելում x -ը բացատրում է կախյալ փոփոխականին՝ y -ին, և ռեգրեսիայի տեսական հավասարումն ունի ազատ անդամ:

Որոշենք տեսական մոդելի անհայտ պարամետրերի վստահելիության միջակայքերը:

$$\theta_1 = \widehat{\beta}_0 - t_{\text{կր}} \cdot S_{\widehat{\beta}_0} = 320 - 1.96 \cdot 125 = 75,$$

$$\theta_2 = \widehat{\beta}_0 + t_{\text{կր}} \cdot S_{\widehat{\beta}_0} = 320 + 1.96 \cdot 125 = 565:$$

Նշանակում է՝ 95 % հավանականությամբ տեսական մոդելի β_0 պարամետրն ընկած է $[75; 565]$ միջակայքում:

$$\theta_1 = \widehat{\beta}_1 - t_{\text{կր}} \cdot S_{\widehat{\beta}_1} = 8.5 - 1.96 \cdot 3.25 = 6.37,$$

$$\theta_2 = \widehat{\beta}_0 + t_{\text{կր}} \cdot S_{\widehat{\beta}_0} = 8.5 + 1.96 \cdot 3.25 = 14.87:$$

Նշանակում է՝ 95 % հավանականությամբ տեսական մոդելի β_1 պարամետրն ընկած է [6.37; 14.87] միջակայքում:

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Ի՞նչ բնույթ ունեն ռեգրեսիայի տեսական մոդելի β_0, β_1 պարամետրերը և էմպիրիկ մոդելի $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցները՝ հավանականակա՞ն, թե՞ դետերմինացված, և ինչո՞ւ:
2. Դո՞ւրս բերեք գծային պարզ ռեգրեսիայի էմպիրիկ մոդելի $\widehat{\beta}_1$ գործակցի վարիացիայի ու ստանդարտ սխալի բանաձևերը:
3. Դո՞ւրս բերեք գծային պարզ ռեգրեսիայի էմպիրիկ մոդելի $\widehat{\beta}_0$ գործակցի վարիացիայի ու ստանդարտ սխալի բանաձևերը:
4. Ինչպե՞ս են առնչվում գծային ռեգրեսիայի էմպիրիկ մոդելի $\widehat{\beta}_0, \widehat{\beta}_1$ գործակիցների վարիացիաները տեսական մոդելի ε_i պատահական շեղումների վարիացիային:
5. Ի՞նչ է պարզ գծային ռեգրեսիայի էմպիրիկ մոդելի գնահատման ստանդարտ սխալը (SEE), ինչպե՞ս է այն որոշվում և ի՞նչ է ցույց տալիս:
6. Ձևակերպե՛ք 3 սիգմայի կանոնը և բացատրե՛ք դրա էությունը: Ինչի՞նչ են հավասար նորմալ բաշխման ասիմետրիայի ու կուտակվածության գործակիցները:
7. Թվարկե՛ք վիճակագրական վարկածների ստուգման հիմնական դրույթները:
8. Նկարագրե՛ք ռեգրեսիայի գործակիցների վիճակագրական նշանակալիության վերաբերյալ վարկածների ստուգման էթեստը:
9. Ո՞րն է ռեգրեսիայի գործակիցների վիճակագրական նշանակալիության էությունը:
10. Ո՞րն է ռեգրեսիայի գործակիցների վիճակագրական նշանակալիության որոշման «կոպիտ» կանոնը:
11. Ի՞նչ է նշանակում միջակայքային գնահատական, և ինչպե՞ս են որոշվում ռեգրեսիայի տեսական մոդելի β_0, β_1 պարամետրերի միջակայքային գնահատականները:

Խնդիրներ

1. $n=500$ մարդկանց տվյալների հիման վրա գնահատվել է սովորած տարիների քանակից աշխատավարձի կախումը բնութագրող պարզ գծային ռեգրեսիայի էմպիրիկ հավասարումը, և ստացվել է $\widehat{\beta}_1 = 25000$ ռեգրեսիայի գործակիցը: Հաշվենք $\widehat{\beta}_1$ -ի վարիացիան ու ստանդարտ սխալը, եթե հայտնի է, որ $\sum_{i=1}^n (x_i - \bar{x})^2 = 100$, $var(\hat{\varepsilon}_i) = 16000$: Հաշվենք $\widehat{\beta}_1$ -ի 95 % վստահելիության միջակայքը, եթե $t_{կր} = 1.96$:
2. $n=100$ սպառողների տվյալների հիման վրա գնահատվել է նրանց եկամուտներից (1000 դրամով) «Յուրոմ» օպերատորի տրամադրած բջջային ինտերնետ կապի սպառման կախումը բնութագրող պարզ գծային ռեգրեսիայի էմպիրիկ հավասարումը, և ստացվել է $\widehat{\beta}_0 = 300$ գործակիցը: Հաշվենք դրա վարիացիան ու ստանդարտ սխալը, եթե հայտնի է, որ $\sum_{i=1}^n (x_i - \bar{x})^2 = 80$, $var(\hat{\varepsilon}_i) = 144$, $\bar{x} = 170$: Հաշվենք $\widehat{\beta}_0$ -ի 95 % վստահելիության միջակայքը, եթե $t_{կր} = 1.96$:
3. $n=202$ ուսանողների տվյալների հիման վրա գնահատվել է «Էկո-նոմետրիկա» առարկայի ընթացիկ քննություններից հավաքած միավորների կախումը առարկայի ուսումնասիրության վրա ուսանողների ծախսած ժամանակից: Պարզ գծային ռեգրեսիայի էմպիրիկ մոդելը ՓԲԵ-ով գնահատելու արդյունքում ստացվել է էմպիրիկ հավասարման պատահական մնացորդների քառակուսիների գումարը՝ $\sum_{i=1}^n (\hat{\varepsilon}_i)^2 = 400$: Հաշվենք $\hat{\varepsilon}_i$ -երի վարիացիան ու գնահատման ստանդարտ սխալը (SEE): Պայմանով, որ Գաուս-Մարկովի 5-րդ պայմանը տեղի ունի, որոշենք, թե գնահատված ուղղից ի՞նչ միջակայքով է հեռացած նմուշային դիտարկումներով որոշված կետերի 95 ու 68 %-ը:
4. Միևնույն ճյուղի 350 ընկերությունների տվյալների հիման վրա գնահատված է դրանց թողարկման ծավալի (OUTPUT, ԱՄՆ դոլար) գծային կախումը ծախսված ռեսուրսներից՝ հիմնական միջոցներից (FUNDS, ԱՄՆ դոլար): Ստացվել է հետևյալ էմպիրիկ հավասարումը՝ $\widehat{OUTPUT} = -3700 + 16264 \cdot FUNDS$: Հաշվարկվել են $\widehat{\beta}_0$, $\widehat{\beta}_1$ գործակիցների ստանդարտ սխալները. $S_{\widehat{\beta}_0} = 1270$ և $S_{\widehat{\beta}_1} = 7925$: Հաշվենք t վիճականու դիտարկված արժեքները $\widehat{\beta}_0$ ու

$\widehat{\beta}_1$ գործակիցների համար: Եզրակացություններ կատարեք էմպիրիկ մոդելի $\widehat{\beta}_0, \widehat{\beta}_1$ գործակիցների վիճակագրական նշանակալիության ու տեսական մոդելի β_0, β_1 պարամետրերի՝ զրոյից նշանակալիորեն տարբերվելու մասին՝ նկատի ունենալով, որ $\alpha=0.05$: Արդյո՞ք ռեգրեսիայի էմպիրիկ հավասարումը կիրառելի է, եթե ենթադրենք, որ Գաուս-Մարկովի պայմանները տեղի ունեն:

5. ՓԶԵ-ով գնահատվել է պարզ գծային ռեգրեսիայի էմպիրիկ հավասարումը, և ստացվել են $\widehat{\beta}_0 = 10$; $\widehat{\beta}_1 = -20$: Ենթադրելով, որ Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն՝ հաշվե՞ք, թե որոնք են β_0, β_1 գործակիցների 95 % վստահելիության միջակայքերը, եթե հայտնի է, որ $S_{\widehat{\beta}_0} = 3$; $S_{\widehat{\beta}_1} = 6$ և $t_{lp} = 1.96$:
6. Գնահատվել են գծային ռեգրեսիայի հետևյալ 4 էմպիրիկ հավասարումները, և ստացվել են հետևյալ արդյունքները: Մեկնաբանե՞ք յուրաքանչյուր հավասարումը: Մոդելներից ո՞րը (որո՞նք) են կիրառելի՝ գործակիցների վիճակագրական նշանակալիության տեսանկյունից: Դրանցից ո՞րն ունի ամենաբարձր որակը:

1. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = -458 + 198.5x$ $t_{\hat{\beta}_0} = -3.9$; $t_{\hat{\beta}_1} = 5.01$; $R^2 = 0.67$	2. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 65 + 72.5x$ $t_{\hat{\beta}_0} = 8.1$; $t_{\hat{\beta}_1} = 0.8$; $R^2 = 0.75$
3. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 1.65 - 160.5x$ $t_{\hat{\beta}_0} = 2.6$; $t_{\hat{\beta}_1} = -30.6$; $R^2 = 0.33$	4. $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x = 268 + 22.7x$ $t_{\hat{\beta}_0} = 4.3$; $t_{\hat{\beta}_1} = 5.6$; $R^2 = 0.78$

7. Միևնույն ճյուղի 230 ընկերությունների տվյալների հիման վրա գնահատված է դրանց թողարկման ծավալի գծային կախումը ծախսված ռեսուրսից՝ ներգրավված աշխատուժի քանակից: Ստացվել են հետևյալ արդյունքները՝

$$\hat{y} = -3200 + 1657 \cdot x; \quad S_{\varepsilon}^2 = 1550; \quad \sum_{i=1}^n (x_i - \bar{x})^2 = 225; \\ \bar{x} = 275; \quad R^2 = 0.67$$

1. Մեկնաբանե՞ք գնահատված էմպիրիկ հավասարման բոլոր գործակիցների իմաստը:

2. Հաշվեք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների վարիացիաները և ստանդարտ սխալները:
 3. Հաշվեք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների t վիճականիները:
 4. Ստուգեք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարեք եզրակացություն գնահատված $\widehat{\beta}$ -ների վիճակագրական նշանակալիության մասին:
 5. Կատարեք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
 6. Կատարեք ընդհանուր եզրակացություն. արդյո՞ք գնահատված ռեգրեսիայի մոդելը կիրառելի է գործակիցների վիճակագրական նշանակալիության տեսանկյունից և բավարարողակ ունի՞:
8. Համաշխարհային բանկի «Աշխարհի զարգացման ցուցանիշներ» (World development indicators) օնլայն տեղեկատվական հարթակից (էլեկտրոնային հասցեն՝ <https://databank.worldbank.org/source/world-development-indicators>) MS Excel ձևաչափով ներբեռնեք գյուղատնտեսությունից առաջացող ածխաթթու գազի արտանետումների ծավալի (մլն տոննայով՝ Carbon dioxide (CO2) emissions from Agriculture (Mt CO2e)) ու գյուղատնտեսության ոլորտում ավելացված արժեքի (ԱՄՆ դոլարով՝ Agriculture, forestry, and fishing, value added (current US\$)) ցուցանիշների ամենավերջին հասանելի տարվան վերաբերող տվյալները աշխարհի 217 երկրների համար:
1. Ձևավորեք Excel-ի տվյալներն այնպես, որ դրանք համապատասխանեն էկոնոմետրիկ մոդելավորման համար օգտագործվող տարածական տվյալների երկչափ շտեմարանների պահանջներին:
 2. Ձևավորեք գյուղատնտեսության ոլորտում ավելացված արժեքի ցուցանիշը բոլոր երկրների համար՝ դրանք բերելով մլն ԱՄՆ դոլարի չափողականության:
 3. Տեղավորեք տվյալները SPSS կամ EViews փաթեթ և գնահատեք միագործոն գծային ռեգրեսիայի մոդել, որը գնահատում է գյուղատնտեսությունից առաջացող ածխաթթու գազի արտանետումների ծավալի կախումը գյուղատնտեսության ոլորտում ավելացված արժեքից:

4. Գրե՛ք միագործոն գծային ռեգրեսիայի գնահատված էմպիրիկ հավասարումը և մեկնաբանե՛ք հավասարման գործակիցների իմաստը:
5. Ստուգե՛ք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարե՛ք եզրակացություն գնահատված $\hat{\beta}$ -ների վիճակագրական նշանակալիության մասին:
6. Կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
7. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն՝ եզրակացությո՛ւն կատարեք մոդելի գնահատման արդյունքների կիրառելիության մասին:

ԳԼՈՒԽ 2. ԲԱԶՄԱԿԻ ԳԾԱՅԻՆ ՌԵԳՐԵՍԻԱ

1. Բազմակի գծային ռեգրեսիայի տեսական ու էմպիրիկ մոդելներն ու հավասարումները

Ցանկացած տնտեսական ցուցանիշի վրա գրեթե միշտ ազդեցություն են ունենում մեկից ավելի գործոններ: Օրինակ՝ որոշակի ապրանքի պահանջարկը կախված է ոչ միայն դրա գնից, այլև փոխարինող ապրանքների գներից, սպառողների եկամտից ու այլ գործոններից: Բնակչության կենսամակարդակը կախված է ոչ միայն մեկ շնչի հաշվով ՀՆԱ-ի մեծությունից, այլև աշխատուժի մասնակցության մակարդակից, չափահաս բնակչության կրթամակարդակից, եկամուտների բաշխման անհավասարությունից և այլն: ՀՆԱ-ն կախված է ոչ միայն ներդրումներից, այլև զուտ արտահանումից, վերջնական սպառումից և այլն: Այդ պատճառով տնտեսության մեջ կախվածությունները ռեգրեսիոն կախվածությամբ արտահայտելու համար օգտագործում են բազմակի ռեգրեսիայի մոդելները:

Ենթադրենք՝ ունենք ոչ թե մեկ, այլ մի քանի բացատրող փոփոխականներ՝ գործոններ, և ուզում ենք գնահատել դրանց բոլորի ազդեցությունը կախյալ փոփոխականի վրա: Այդ դեպքում կունենանք ոչ թե պարզ ռեգրեսիայի մոդել, այլ բազմագործոն կամ բազմակի ռեգրեսիայի մոդել:

Օրինակ՝

$$Q^D = f(P, X, PM),$$

որում Q^D -ը որևէ ապրանքի պահանջարկն է, X -ը՝ եկամուտը, P -ն՝ ապրանքի միավորի գինը, PM -ը՝ մրցակցող ապրանքի միավորի գինը:

$X_1, X_2, \dots, X_m$ ու թվով բացատրող փոփոխականներ կամ գործոններ պարունակող բազմակի ռեգրեսիայի **տեսական հավասարումը** պարզ ռեգրեսիայի նմանությամբ կունենա հետևյալ տեսքը՝

$$\mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) = f(X_1, X_2, \dots, X_m), \quad (57)$$

իսկ **տեսական մոդելը** կունենա հետևյալ տեսքը՝

$$\begin{aligned} y &= \mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) + \varepsilon = \\ &= f(X_1, X_2, \dots, X_m) + \varepsilon: \end{aligned} \quad (58)$$

Ենթադրենք՝ գործոնների ազդեցությունը կախյալ փոփոխականի վրա գծային օրենքի է ենթարկվում: Այդ դեպքում կունենանք **բազմակի գծային ռեգրեսիայի** (այսուհետ՝ **ԲԳՌ**) մոդել:

$$y = \mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) + \varepsilon = f(X_1, X_2, \dots, X_m) + \varepsilon = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon: \quad (59)$$

ԲԳՌ տեսական հավասարումն այդ դեպքում կստանա հետևյալ տեսքը՝

$$\mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) = f(X_1, X_2, \dots, X_m) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m: \quad (60)$$

(60)-ում $\beta_0, \beta_1, \dots, \beta_m$ -ը ԲԳՌ տեսական մոդելի անհայտ պարամետրերն են՝ թվով $m+1$ հատ: $\beta_1, \dots, \beta_m$ պարամետրերից յուրաքանչյուրը բնութագրում է համապատասխան բացատրող փոփոխականի ազդեցությունը կախյալ փոփոխականի միջինի կամ սպասվող արժեքի վրա:

Ռեգրեսիոն մոդելավորումը հնարավորություն է տալիս Y -ի և X -երի բոլոր հնարավոր դիտարկումների գլխավոր համախմբությունից կատարված նմուշային տվյալների հիման վրա գնահատելու **ԲԳՌ էմպիրիկ հավասարումը**, որն ունի հետևյալ տեսքը՝

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_m x_m, \quad (61)$$

իսկ **ԲԳՌ էմպիրիկ մոդելը**՝

$$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_m x_m + \varepsilon = \hat{y} + \varepsilon: \quad (62)$$

$\hat{y}$ -ը կախյալ փոփոխականի գնահատված արժեքն է: Այն տեսական մոդելի Y կախյալ փոփոխականի պայմանական մաթ. սպասման՝ $\mathbb{E}(Y|X_1 = x_1, X_2 = x_2, \dots, X_m = x_m)$ -ի վիճակագրական գնահատականն է: $\hat{\beta}$ գործակիցները տեսական մոդելի β անհայտ պարամետրերի վիճակագրական գնահատականներն են, որոնք ստացվել են էմպիրիկ տվյալների հիման վրա գծային ռեգրեսիայի էմպիրիկ հավասարումը գնահատելու արդյունքում: $\hat{\varepsilon}$ -ը պատահական մնացորդն է, տեսական մոդելի ε պատահական շեղման վիճակագրական գնահատականը:

n թվով էմպիրիկ տվյալների կամ n թվով դիտարկումների առկայության դեպքում ամեն մի դիտարկման համար ԲԳՌ էմպիրիկ մոդելը կձևակերպվի հետևյալ կերպ.

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_m x_{im} + \varepsilon_i = \hat{y}_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (63)$$

իսկ էմպիրիկ հավասարումը՝

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times 1} \quad B = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}_{(m+1) \times 1} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & x_{23} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nm} \end{bmatrix}_{n \times (m+1)} \quad :$$

Այդ դեպքում ԲԳՌ (59) տեսական մոդելը կարելի է ներկայացնել մատրիցային տեսքով.

$$Y = XB + \varepsilon: \quad (65)$$

ԲԳՌ էմպիրիկ մոդելը մատրիցային տեսքով կլինի հետևյալը.

$$Y = X\hat{B} + \hat{\varepsilon}, \quad (66)$$

իսկ էմպիրիկ հավասարումը՝

$$\hat{Y} = X\hat{B}: \quad (67)$$

Բազմակի գծային ռեգրեսիայի տեսական ու էմպիրիկ հավասարումների ու մոդելների տեսքերը ամփոփված են Աղյուսակ 5-ում:

Աղյուսակ 5. Բազմակի գծային ռեգրեսիայի տեսական ու էմպիրիկ հավասարումների ու մոդելների տեսքերը:

ԲԳՌ	Ոչ մատրիցային տեսքով	Մատրիցային տեսքով
տեսական հավասարում	$\begin{aligned} \mathbb{E}(Y X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) &= \\ &= f(X_1, X_2, \dots, X_m) = \\ &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m \end{aligned}$	$\begin{aligned} \mathbb{E}(Y X_1 = x_1, \\ X_2 = x_2, \dots, X_m = x_m) &= \\ &= XB \end{aligned}$
տեսական մոդել	$\begin{aligned} y &= \mathbb{E}(Y X_1 = x_1, X_2 = x_2, \dots, X_m = x_m) + \\ &+ \varepsilon = f(X_1, X_2, \dots, X_m) + \varepsilon = \\ &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_m X_m + \varepsilon \end{aligned}$	$Y = XB + \varepsilon$
էմպիրիկ հավասարում	$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_m x_m$	$\hat{Y} = X\hat{B}$
էմպիրիկ մոդել	$y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_m x_m + \hat{\varepsilon} = \hat{y} + \hat{\varepsilon}$	$Y = X\hat{B} + \hat{\varepsilon}$

Ինչպե՞ս են մեկնաբանվում ԲԳՌ- $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_m x_m$ էմպիրիկ հավասարման գնահատված $\hat{\beta}$ գործակիցները: $\hat{\beta}_0$ -ը հավասարման ազատ անդամն է և ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականի միջինը, եթե բոլոր բացատրող փոփոխականները հավասար լինեն զրոյի: Գնահատված $\hat{\beta}_1$ գործակիցը ցույց է տալիս, թե քանի միավորով կփոխվի կախյալ փոփոխականի միջին արժեքը, եթե x_1 բացատրող փոփոխականի արժեքը փոխվի մեկ միավորով այն պայմանով, որ մնացած բոլոր բացատրող փոփոխականների արժեքները ամրագրված են, չեն փոխվում: $\hat{\beta}_2$ գործակիցը ցույց է տալիս, թե քանի միավորով կփոխվի կախյալ փոփոխականի միջին արժեքը, եթե x_2 բացատրող փոփոխականի արժեքը փոխվի մեկ միավորով այն պայմանով, որ մնացած բոլոր բացատրող փոփոխականների արժեքները չեն փոխվում: Նմանապես են մեկնաբանվում նաև մնացած բոլոր $\hat{\beta}$ գործակիցները: Էմպիրիկ հավասարումը մեկնաբանելիս հաճախ «պայմանով, որ մնացած բոլոր բացատրող փոփոխականների արժեքները ամրագրված են» արտահայտությունը փոխարինում են «այլ հավասար պայմաններում» (*ceteris paribus*) արտահայտությամբ: Ինչո՞ւ է պայմանավորված այսպիսի մեկնաբանությունը: Դիտարկենք պարզագույն ԲԳՌ- էմպիրիկ հավասարումը երկու բացատրող փոփոխականով՝ $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$: $\hat{y}$ -ը որոշում է y -ի նմուշային սպասվող արժեքն այն դեպքում, երբ X_1 փոփոխականն ընդունում է x_1 արժեք, իսկ X_2 -ը՝ x_2 արժեք: Ենթադրենք X_1 փոփոխականն աճում է մեկ միավորով և ընդունում $x_1 + \Delta x_1$ արժեք, իսկ X_2 -ը չի փոխվում, ամրագրված է և ընդունում է x_2 արժեք: Այդ պարագայում $\hat{y}$ -ը ևս կփոխվի ինչ-որ $\Delta \hat{y}$ միավորով, և հավասարումը կընդունի հետևյալ տեսքը՝ $\hat{y} + \Delta \hat{y} = \hat{\beta}_0 + \hat{\beta}_1(x_1 + \Delta x_1) + \hat{\beta}_2 x_2$: Այս երկրորդ հավասարումից հանենք առաջինը: Կստանանք՝ $\Delta \hat{y} = \hat{\beta}_0 + \hat{\beta}_1(x_1 + \Delta x_1) + \hat{\beta}_2 x_2 - \hat{\beta}_0 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_2 = \hat{\beta}_1 \Delta x_1 \rightarrow \hat{\beta}_1 = \frac{\Delta \hat{y}}{\Delta x_1}$: Այսինքն՝ $\hat{\beta}_1$ -ը ցույց է տալիս, թե քանի միավորով է փոփոխվում y -ի սպասվող արժեքը, երբ x_1 -ը փոխվում է մեկ միավորով, իսկ x_2 -ը մնում է անփոփոխ կամ ամրագրված: Նմանօրինակ ձևով կստացվի նաև $\hat{\beta}_2$ -ի մեկնաբանությունը:

Ներդիր 14. Բազմակի ու մասնակի կոռելյացիայի գործակիցներ

Բազմակի կոռելյացիան վիճակագրական չափանիշ է, որը ցույց է տալիս մի փոփոխականի և միաժամանակ մեկից ավելի այլ փոփոխականների միջև գծային կապի ուժգնությունը: Բազմակի կոռելյացիայի գործակիցն օգտագործվում է որպես կախյալ փոփոխականի ու բացատրող փոփոխականների բազմության միջև գծային վիճակագրական ընդհանուր կապի ուժգնության աստիճանի չափանիշ, կամ, այլ կերպ ասած, այն գնահատում է կախյալ փոփոխականի վրա բացատրող փոփոխականների համատեղ ազդեցության ուժգնությունը: Բազմակի կոռելյացիայի գործակիցն ընկած է $[0;1]$ միջակայքում և քառակուսի արմատ է գծային բազմակի ռեգրեսիայի դետերմինացիայի գործակցից: Այն որոշվում է բոլոր փոփոխականների զույգային կոռելյացիայի գործակիցների միջոցով: Օրինակ՝ մնուշային բազմակի կոռելյացիայի գործակիցը y, x_1, x_2 երեք փոփոխականի դեպքում որոշվում է հետևյալ բանաձևով՝

$$r_{y|x_1, x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2 \cdot r_{yx_1} \cdot r_{yx_2} \cdot r_{x_1x_2}}{1 - r_{x_1x_2}^2}}:$$

Մասնակի կոռելյացիան վիճակագրական չափանիշ է, որը ցույց է տալիս երկու փոփոխականների միջև գծային կապի ուժգնությունն ու ուղղությունը մեկ կամ մի քանի այլ փոփոխականների ազդեցությունը վերացնելուց կամ մաքրելուց հետո: Այն թույլ է տալիս հասկանալ՝ արդյոք երկու փոփոխականների միջև կա ուժեղ իրական կապ, թե դրանց կապը պայմանավորված է այլ գործոնների ընդհանուր ազդեցությամբ: Այսպիսով՝ մասնակի կոռելյացիայի գործակիցը ցույց է տալիս, թե որքան ուժեղ է երկու փոփոխականների «մաքուր» գծային կախվածության ուժգնությունը՝ առանց հաշվի առնելու այլ գործոնների միջնորդավորված ազդեցությունները: Այն որոշվում է բոլոր փոփոխականների զույգային կոռելյացիայի գործակիցների միջոցով: Օրինակ՝ x, y, z երեք փոփոխականի դեպքում x -ի և y -ի մասնակի, z -ի ազդեցությունից մաքրված մնուշային կոռելյացիայի գործակիցը որոշվում է հետևյալ բանաձևով՝

$$r_{x,y|z} = \frac{r_{x,y} - r_{x,z} \cdot r_{y,z}}{\sqrt{(1 - r_{x,z}^2)} \cdot \sqrt{(1 - r_{y,z}^2)}}:$$

2. Բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատումը ՓԲԵ-ով

Մատրիցային տեսքով ԲԳՌ էմպիրիկ մոդելի պատահական մնացորդների վեկտորը կարտահայտվի այսպես.

$$\widehat{\varepsilon} = Y - X\widehat{B} = Y - \widehat{Y}:$$

ՓԲԵ-ով ԲԳՌ էմպիրիկ հավասարման գնահատումը խնդիր է դնում գտնելու այնպիսի $\widehat{B}$ վեկտոր-սյուն, որ պատահական մնացորդների քառակուսիների գումարը լինի փոքրագույնը: Այսինքն՝ պետք է գտնենք այն $\widehat{B}$ վեկտոր-սյունը, որի դեպքում՝

$$Q = \widehat{\varepsilon}^T \cdot \widehat{\varepsilon} \rightarrow \min:$$

Բացենք $Q = \widehat{\varepsilon}^T \cdot \widehat{\varepsilon}$ -ն:

$$\begin{aligned} Q &= \widehat{\varepsilon}^T \cdot \widehat{\varepsilon} = (Y - \widehat{Y})^T (Y - \widehat{Y}) = (Y - X\widehat{B})^T (Y - X\widehat{B}) = \\ &= Y^T Y - Y^T X\widehat{B} - \widehat{B}^T X^T Y + \widehat{B}^T X^T X\widehat{B} = Y^T Y - 2\widehat{B}^T X^T Y + \widehat{B}^T X^T X\widehat{B}: \quad (68) \end{aligned}$$

- Սա սրացվեց՝ հաշվի առնելով, որ $Y^T X\widehat{B}$ և $\widehat{B}^T X^T Y$ մաթրիցաները սկալյարներ են, $\widehat{B}^T X^T Y = Y^T X\widehat{B}$, ինչպես նաև, որ $(Y - X\widehat{B})^T = Y^T - (X\widehat{B})^T$; $(X\widehat{B})^T = \widehat{B}^T X^T$:

Ածանցենք (68)-ը ըստ $\widehat{B}$ -ի՝

$$Q' = \frac{\partial Q}{\partial \widehat{B}} = -2X^T Y + 2X^T X\widehat{B}:$$

Նորմալ հավասարումների համակարգը մատրիցային տեսքով կլինի հետևյալը.

$$-2X^T Y + 2X^T X\widehat{B} = 0 \quad \rightarrow \quad X^T X\widehat{B} = X^T Y \quad \rightarrow \quad \widehat{B} = (X^T X)^{-1} X^T Y:$$

Ստացվեց այն բանաձևը, որով փոքրագույն քառակուսիների մեթոդով որոշվում են ԲԳՌ էմպիրիկ մոդելի $\widehat{\beta}_0, \widehat{\beta}_1, \dots, \widehat{\beta}_m$ գործակիցները՝

$$\widehat{B} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \vdots \\ \widehat{\beta}_m \end{bmatrix} = (X^T X)^{-1} X^T Y: \quad (69)$$

3. Գաուս-Մարկովի պայմաններն ու թեորեմը բազմակի գծային ռեգրեսիայի համար, բազմակի գծային ռեգրեսիայի նորմալ դասական մոդելի սահմանումը

Ինչպես արդեն ասվել է, ՓԶԵ-ով գնահատված էմպիրիկ հավասարումը, որպեսզի լավագույն ձևով մոտարկի գծային ռեգրեսիայի տեսական մոդելը, անհրաժեշտ է, որ տեղի ունենան Գաուս-Մարկովի պայմանները: Բազմագործոն գծային ռեգրեսիայի մոդելի համար դրանք հետևյալն են.

Պայման 1. Ռեգրեսիայի տեսական մոդելը գծային է β պարամետրերի նկատմամբ, այսինքն՝

$$Y = XB + \varepsilon,$$

որում X -ը դետերմինացված մատրից է, և բացատրող փոփոխականները պատահական շեղումների հետ կոռելացված չեն. $\text{COV}(X, \varepsilon)=0$:

Պայման 2. Բացատրող $X_1, \dots, X_m$ փոփոխականները գծորեն անկախ են, և X մատրիցն ունի առավելագույն ռանգ. $\text{rang}(X)=m+1$: Այս պայմանը կոչվում է բազմակողմնաբարության բացակայության պայման, իսկ դրա խախտումը՝ **բազմակողմնաբարություն**:

Պայման 3. Պատահական շեղումների մաթ. սպասումը բոլոր դիտարկումների համար հավասար է զրոյի. $E(\varepsilon) = \theta_{n \times 1}, \forall i$, որում θ -ն զրոյական n -շափանի վեկտոր-սյուն է:

Այսպիսով՝

$$E(\varepsilon) = E \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} = \begin{bmatrix} E(\varepsilon_1) \\ E(\varepsilon_2) \\ \vdots \\ E(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \theta:$$

Պայման 4. Պատահական շեղումների կովարիացիայի մատրիցը հավասար է՝ $\Omega_\varepsilon = \sigma_\varepsilon^2 I_{n \times n}$, որում I -ն միավոր մատրից է (անկյունագծի վրա 1-եր են, իսկ դրանից դուրս՝ 0-ներ):

Պատահական շեղումների կովարիացիայի մատրիցի անկյունագծի վրա ε -երի վարիացիաներն են, իսկ դրանից դուրս՝ տարբեր դիտարկումներին համապատասխանող ε -երի կովարիացիաները.

$$\begin{aligned}\Omega_{\varepsilon} &= \mathbb{E}(\varepsilon \cdot \varepsilon^T) = \mathbb{E} \left[\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}^T \right] = \begin{bmatrix} \mathbb{E}(\varepsilon_1^2) & \mathbb{E}(\varepsilon_1 \cdot \varepsilon_2) & \dots & \mathbb{E}(\varepsilon_1 \cdot \varepsilon_n) \\ \mathbb{E}(\varepsilon_2 \cdot \varepsilon_1) & \mathbb{E}(\varepsilon_2^2) & \dots & \mathbb{E}(\varepsilon_2 \cdot \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(\varepsilon_n \cdot \varepsilon_1) & \mathbb{E}(\varepsilon_n \cdot \varepsilon_2) & \dots & \mathbb{E}(\varepsilon_n^2) \end{bmatrix} = \\ &= \begin{bmatrix} \text{Var}(\varepsilon_1) & \text{Cov}(\varepsilon_1 \cdot \varepsilon_2) & \dots & \text{Cov}(\varepsilon_1 \cdot \varepsilon_n) \\ \text{Cov}(\varepsilon_2 \cdot \varepsilon_1) & \text{Var}(\varepsilon_2) & \dots & \text{Cov}(\varepsilon_2 \cdot \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(\varepsilon_n \cdot \varepsilon_1) & \text{Cov}(\varepsilon_n \cdot \varepsilon_2) & \dots & \text{Var}(\varepsilon_n) \end{bmatrix}_{n \times n}\end{aligned}$$

$$\text{Պայմանը պահանջում է, որ } \Omega_{\varepsilon} = \sigma_{\varepsilon}^2 I = \sigma_{\varepsilon}^2 \cdot \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \sigma_{\varepsilon}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\varepsilon}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\varepsilon}^2 \end{bmatrix}_{n \times n}, \text{ ուրեմն պայման 4-ը նշանակում է՝}$$

$$\Omega_{\varepsilon} = \sigma_{\varepsilon}^2 I = \begin{bmatrix} \text{Var}(\varepsilon_1) & \text{Cov}(\varepsilon_1 \cdot \varepsilon_2) & \dots & \text{Cov}(\varepsilon_1 \cdot \varepsilon_n) \\ \text{Cov}(\varepsilon_2 \cdot \varepsilon_1) & \text{Var}(\varepsilon_2) & \dots & \text{Cov}(\varepsilon_2 \cdot \varepsilon_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(\varepsilon_n \cdot \varepsilon_1) & \text{Cov}(\varepsilon_n \cdot \varepsilon_2) & \dots & \text{Var}(\varepsilon_n) \end{bmatrix} = \begin{bmatrix} \sigma_{\varepsilon}^2 & 0 & \dots & 0 \\ 0 & \sigma_{\varepsilon}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\varepsilon}^2 \end{bmatrix}.$$

Հետևաբար այս պայմանը նշանակում է, որ.

ա) տարբեր դիտարկումների համապատասխանող պատահական շեղումները կոռելացված չեն. $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$ (ավտոկոռելյացիայի բացակայության պայման),

բ) պատահական շեղումների վարիացիան հաստատուն է. $\text{Var}(\varepsilon_i) = \sigma_{\varepsilon}^2, \forall i$ (հոմոսկեդաստիկության պայման):

Պայման 5. Պատահական շեղումները մորմալ են բաշխված՝ $\varepsilon \sim N(\theta, \Omega_{\varepsilon})$:

Գաուս-Մարկովի թեորեմ

Եթե Գաուս-Մարկովի 1-4 պայմանները տեղի ունեն, ապա ՓՔԵ-ով ստացված $\hat{B}$ գնահատականները լավագույն գծային անշե-

ղելի գնահատականներն են բոլոր գծային անշեղելի գնահատականների դասում կամ BLUE գնահատականներ են: Այսինքն՝

1. անշեղ գնահատականներ են. $\mathbb{E}(\hat{B}) = B$,
2. արդյունավետ գնահատականներ են. ունեն ամենափոքր կովարիացիոն մատրիցը՝ որևէ այլ մեթոդով գնահատված կամայական այլ $\tilde{B}$ գնահատականների համեմատ. $\Omega_{\hat{B}} < \Omega_{\tilde{B}}$:

* Եթե տեղի ունի նաև Պայման 5-ը, ապա $\hat{B}$ -երը լավագույն անշեղ գնահատականներն են բոլոր ինչպես գծային, այնպես էլ ոչ գծային անշեղ գնահատականների դասում (կամ, որ նույնն է, BLUE գնահատականներ են), դրանք նաև ունակ գնահատականներ են՝ $\lim_{n \rightarrow \infty} \hat{B} = B$, և նորմալ են բաշխված. $\hat{B} \sim N(B, \Omega_{\hat{B}})$:

Սահմանում 1. Եթե Գաուս-Մարկովի 1-5 պայմանները տեղի ունեն, ԲԳՌ-ն մոդելը կոչվում է **բազմակի գծային ռեգրեսիայի նորմալ դասական մոդել**:

Սահմանում 2. Եթե տեղի ունեն Գաուս-Մարկովի 1-4 պայմանները, ապա ԲԳՌ-ն մոդելը կոչվում է **բազմակի գծային ռեգրեսիայի դասական մոդել**:

ԲԳՌ- տեսական մոդելի անհայտ պարամետրերն են ոչ միայն B վեկտոր-սյան էլեմենտները, այլև պատահական շեղումների կովարիացիոն մատրիցը՝ Ω_{ε} -ն: Այս մոդելը գնահատելու համար պետք է գնահատել նաև Ω_{ε} մատրիցը:

Եթե Գաուս-Մարկովի պայմանները տեղի ունեն, ապա ՓՔԵ-ով ստացված $\hat{\beta}$ -երը անշեղ գնահատականներ են, այսինքն՝ $\mathbb{E}(\hat{B}) = B$: Այդ դեպքում $\Omega_{\hat{B}}$ -ն կունենա հետևյալ տեսքը՝

$$\Omega_{\hat{B}} = \mathbb{E}[(\hat{B} - \mathbb{E}(\hat{B})) \cdot (\hat{B} - \mathbb{E}(\hat{B}))^T] = \mathbb{E}[(\hat{B} - B) \cdot (\hat{B} - B)^T]$$

Հաշվենք $(\hat{B} - B)$ -ը: Նկատի ունենալով, որ ԲԳՌ-ն հավասարումը ՓՔԵ-ով գնահատելու դեպքում ստացվող $\hat{B}$ վեկտոր-սյունը որոշվում էր հետևյալ բանաձևով՝ $\hat{B} = (X^T X)^{-1} X^T Y$ և $Y = XB + \varepsilon$, կունենանք.

$$\begin{aligned} (\hat{B} - B) &= ((X^T X)^{-1} X^T Y - B) = ((X^T X)^{-1} X^T (XB + \varepsilon) - B) = \\ &= ((X^T X)^{-1} X^T XB + (X^T X)^{-1} X^T \varepsilon - B) = B + (X^T X)^{-1} X^T \varepsilon - B = \\ &= (X^T X)^{-1} X^T \varepsilon: \end{aligned}$$

Հետևաբար՝

$$(\hat{B} - B)^T = ((X^T X)^{-1} X^T \mathcal{E})^T = \mathcal{E}^T X (X^T X)^{-1}.$$

Եվ

$$(\hat{B} - B) \cdot (\hat{B} - B)^T = (X^T X)^{-1} X^T \mathcal{E} \mathcal{E}^T X (X^T X)^{-1}.$$

Հաշվի առնելով, որ X մատրիցը դետերմինացված է, կունենանք՝

$$\begin{aligned} \Omega_{\hat{B}} &= \mathbb{E}[(\hat{B} - B) \cdot (\hat{B} - B)^T] = (X^T X)^{-1} X^T \mathbb{E}(\mathcal{E} \mathcal{E}^T) X (X^T X)^{-1} = \\ &= \mathbb{E}(\mathcal{E} \mathcal{E}^T) (X^T X)^{-1} X^T X (X^T X)^{-1} = \Omega_{\mathcal{E}} \cdot (X^T X)^{-1} \end{aligned} \quad (70)$$

Ստացվեց, որ $\hat{B}$ վեկտոր-սյան $\Omega_{\hat{B}}$ կովարիացիոն մատրիցն ուղիղ համեմատական է $\mathcal{E}$ վեկտոր-սյան $\Omega_{\mathcal{E}}$ կովարիացիոն մատրիցին: Որքան փոքր լինի վերջինը, այնքան փոքր կլինեն $\hat{B}$ -երի վարիացիաներն ու կովարիացիաները, հետևաբար այնքան ճշգրիտ գնահատականներ կլինեն դրանք:

4. Գծային բազմակի ռեգրեսիայի էմպիրիկ հավասարման որակի ստուգումը. ճշգրտված դետերմինացիայի գործակից

Բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման որակը նույնպես ստուգվում է դետերմինացիայի գործակցի միջոցով: Սակայն ԲԳՌ- պարագայում դետերմինացիայի գործակցի արժեքը կախված է մոդելի մեջ մտցված բացատրող փոփոխականների քանակից: Բացատրող փոփոխականների քանակի աճի հետ (31) բանաձևով որոշվող դետերմինացիայի գործակցի արժեքը մեծանում է, և դրա ճշգրտությունը բազմակի ռեգրեսիայի պարագայում փոքրանում է: Դրանից խուսափելու համար ԲԳՌ- դեպքում հաշվարկվում է ճշգրտված դետերմինացիայի գործակիցը, որը չեզոքացնում է նշված ազդեցությունը ազատության աստիճանների օգտագործմամբ: Ճշգրտված դետերմինացիայի գործակիցը որոշվում է հետևյալ բանաձևով՝

$$R_{adj}^2 = 1 - (1 - R^2) \cdot \frac{n - 1}{n - (m + 1)}, \quad (71)$$

որում n -ը դիտարկումների, իսկ m -ը բացատրող փոփոխականների քանակն է:

Ակնհայտորեն ճշգրտված դետերմինացիայի գործակցի մեկնաբանությունը նույնն է, ինչ դետերմինացիայի գործակցինը: Բազմակի ռեգրեսիայի մոդելի t վիճակահանները հաշվարկվում և մեկնաբանվում են պարզ ռեգրեսիայի նմանությամբ:

5. Գծային բազմակի ռեգրեսիայի հավասարման կամ դետերմինացիայի գործակցի վիճակագրական նշանակալիության ստուգումը. F թեստ

Ենթադրենք՝ ունենք բազմակի գծային ռեգրեսիայի տեսական մոդել.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m + \varepsilon:$$

ՓԶԵ-ով այն գնահատելուց հետո ԲԳՌ- էմպիրիկ հավասարման որակը ստուգվում է ճշգրտված դետերմինացիայի գործակցի միջոցով, իսկ դրա գնահատված բոլոր β գործակիցների առանձին-առանձին վիճակագրորեն նշանակալի լինելը ստուգվում է t վիճակահան միջոցով: Եթե դրանք բոլորը վիճակագրորեն նշանակալի են, մոդելի վավերականության ստուգումը կարելի է շարունակել: Եթե դրանցից որևէ մեկը նշանակալի չէ, մոդելը կիրառելի չէ, դրա հատկորոշումը պետք է փոխել:

ՓԶԵ-ով ԲԳՌ- էմպիրիկ հավասարումը գնահատելու արդյունքում կարելի է ստուգել ոչ միայն առանձին-առանձին բոլոր β -ների վիճակագրական նշանակալիությունը, այլ նաև՝ միաժամանակ բոլոր β -ների վիճակագրական նշանակալիությունը: Այլ խոսքով՝ կարելի է ստուգել, թե տեսական հավասարման մեջ մտցված բոլոր բացատրող փոփոխականները կախյալ փոփոխականի վրա ազդում են, թե ոչ: Սա հավասարագոր է նրան, որ ստուգվի ռեգրեսիայի տեսական հավասարման վիճակագրական նշանակալիությունը: Դա ստուգելու համար դրվում է զրոյական վարկած այն մասին, որ տեսական մոդելի բոլոր β պարամետրերը, բացի β_0 -ից, զրոյից նշանակալիորեն չեն տարբերվում, այսինքն՝ ոչ մի x կախյալ փոփոխականին չի բացատրում: Ունենք՝

$$H_0: \beta_1 = \beta_2 = \dots = \beta_m = 0: \quad (72)$$

Միջակցող վարկածն այն է, որ առնվազն մեկ β պարամետր հավասար չէ զրոյի, տեսական մոդելում առնվազն մեկ x բացատրում է y -ին: Եթե այս զրոյական վարկածը մերժվի, կարող ենք եզրակացնել,

որ ռեգրեսիայի գնահատված էմպիրիկ հավասարումը վիճակագրորեն նշանակալի է, և դրա արդյունքները վստահելի են: Վարկածը ստուգվում է ԲԳՌ-էմպիրիկ մոդելի ՓԲԵ-ով գնահատման արդյունքներն օգտագործելու միջոցով:

Չրոյական վարկածը ստուգվում է հետևյալ վիճականու միջոցով.

$$F = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / m}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - (m + 1))} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / m}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2 / (n - (m + 1))} =$$

$$= \frac{ESS/m}{RSS/(n - (m + 1))} = \frac{ESS}{RSS} \cdot \frac{(n - (m + 1))}{m}; \quad (73)$$

Եթե Գաուս-Մարկովի բոլոր պայմանները տեղի ունեն, F վիճականին ունի Ֆիշերի բաշխում՝ $v_1 = m$, $v_2 = n - (m + 1)$ ազատության աստիճաններով: F վիճականու համարիչում կախյալ փոփոխականի ընդհանուր վարիացիայի՝ ռեգրեսիայի գնահատված հավասարմամբ բացատրվող բաղկացուցիչն է, իսկ հայտարարում՝ ռեգրեսիայի գնահատված հավասարմամբ չբացատրվող ու պատահական մնացորդներով բացատրվող բաղկացուցիչը: Որքան համարիչը մեծ լինի, այնքան կախյալ փոփոխականի ընդհանուր վարիացիայի՝ ռեգրեսիայի էմպիրիկ հավասարմամբ բացատրվող բաղկացուցիչը մեծ կլինի, այնքան մեծ կլինի F վիճականին, և այնքան մեծ կլինի գնահատված հավասարման որակը: Եթե H_0 -ն տեղի ունի, ապա կախյալ փոփոխականի ընդհանուր վարիացիայի՝ ռեգրեսիայի գնահատված հավասարմամբ բացատրվող բաղկացուցիչը նշանակալիորեն ավելի փոքր է, քան դրանով չբացատրվող բաղկացուցիչը: Իսկ եթե մրցակցող H_1 -ը տեղի ունի, ապա կախյալ փոփոխականի ընդհանուր վարիացիայի՝ ռեգրեսիայի գնահատված հավասարմամբ բացատրվող մասը նշանակալիորեն ավելի մեծ է, քան դրանով չբացատրվող մասը:

Վարկածի ստուգման համար α արժեքականության մակարդակում F վիճականու (73) բանաձևով հաշվարկված արժեքը (անվանում ենք դիտարկված արժեք) համեմատվում է Ֆիշերի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքի հետ՝ $v_1 = m$, $v_2 = n - (m + 1)$ ազատության աստիճաններով ու α արժեքականության մակարդակում, որը որոշվում է Ֆիշերի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքների աղյուսակից: Եթե՝

$$F^{\eta\text{հս}} > F^{\text{հր}}_{m,n-(m+1)} \rightarrow H_0\text{-ն մերժվում է:}$$

Հետևաբար ԲԳՌ- էմպիրիկ հավասարումը վիճակագրորեն նշանակալի է, իսկ տեսական հավասարման անկախ փոփոխականներից առնվազն մեկը բացատրում է կախյալ փոփոխականին: Իսկ եթե՝

$$F^{\eta\text{հս}} \leq F^{\text{հր}}_{m,n-(m+1)} \rightarrow H_0\text{-ն չի կարող մերժվել, այն ընդունվում է:}$$

Հետևաբար ԲԳՌ- էմպիրիկ հավասարումը վիճակագրորեն նշանակալի չէ, իսկ տեսական հավասարման անկախ փոփոխականներից ոչ մեկը չի բացատրում կախյալ փոփոխականին:

Այս գրոյական վարկածը համարժեք է դետերմինացիայի գործակցի վիճակագրական նշանակալիության մասին վարկածին: Այն է՝

$$H_0: R^2 = 0; H_1: R^2 > 0:$$

Վարկածը ստուգվում է F վիճականու միջոցով, որն էլ որոշվում է հետևյալ բանաձևով.

$$F = \frac{R^2/m}{(1-R^2)/(n-(m+1))}, \quad (74)$$

որում m -ը բացատրող փոփոխականների թիվն է, n -ը՝ դիտարկումների թիվը, իսկ R^2 -ն՝ գնահատված ռեգրեսիայի էմպիրիկ հավասարման դետերմինացիայի գործակցը:

Հաշվի առնելով, որ՝

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{ESS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{RSS}{TSS}$$

և

$$1 - R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

ձևափոխենք (74)-ը, կստանանք.

$$F = \frac{\frac{R^2}{m}}{\frac{1-R^2}{n-(m+1)}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{m \cdot \sum_{i=1}^n (y_i - \bar{y})^2} \cdot \frac{(n-(m+1)) \cdot \sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} =$$

$$= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \cdot \frac{(n - (m + 1))}{m} = \frac{ESS}{RSS} \cdot \frac{(n - (m + 1))}{m}.$$

Նաև պետք է նշել, որ (74) բանաձևից հետևում է, որ R^2 -ն ու F վիճակագրի միաժամանակ են զրոյի հավասարվում: Այսինքն՝ եթե $R^2 = F = 0$, ապա կնշանակի, որ կախյալ փոփոխականի վարիացիան բոլորովին չի բացատրվում գնահատված ռեգրեսիայի հավասարմամբ: Հետևաբար y -ը x -երով չի բացատրվում, ռեգրեսիայի հավասարումը նշանակալի չէ:

Այս դեպքում ևս F վիճակագրի ունի Ֆիշերի բաշխում՝ $v_1 = m$, $v_2 = n - (m + 1)$ ազատության աստիճաններով: Վարկածի ստուգման համար α արժեքականության մակարդակում F վիճակագրի համեմատվում է Ֆիշերի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքի հետ՝ $v_1 = m$, $v_2 = n - (m + 1)$ ազատության աստիճաններով ու α արժեքականության մակարդակով:

Եթե՝

$$F^{\text{դիտ}} > F^{\text{կր}}_{m, n-(m+1)} \rightarrow H_0\text{-ն մերժվում է:}$$

Նշանակում է՝ R^2 դետերմինացիայի գործակիցը վիճակագրորեն նշանակալի է, ինչը հավասարագոր է նրան, որ ռեգրեսիայի հավասարումն ընդհանուր առմամբ նշանակալի է:

Իսկ եթե՝

$$F^{\text{դիտ}} \leq F^{\text{կր}}_{m, n-(m+1)} \rightarrow H_0\text{-ն չի կարող մերժվել, այն ընդունվում է:}$$

Նշանակում է՝ R^2 դետերմինացիայի գործակիցը վիճակագրորեն նշանակալի չէ, ինչը հավասարագոր է նրան, որ ռեգրեսիայի հավասարումն ընդհանուր առմամբ նշանակալի չէ:

Գնահատված ռեգրեսիայի հավասարման վիճակագրական նշանակալիությունը կարելի է ստուգել մաս H_0 վարկածի տեղի ունենալու հավանականության՝ p արժեքի միջոցով: Դա, միասին վերցված, բոլոր ռեգրեսիայի գործակիցների (բացի ազատ անդամից) զրո լինելու մասին վարկածը մերժելու դեպքում սխալված լինելու հավանականությունն է: Հեշտության համար կասենք՝ « H_0 վարկածի տեղի ունենալու հավանականությունն է»: Այն համեմատում են ընտրված α արժեքականության մակարդակի հետ (0.1, կամ՝ 0.05, կամ՝ 0.01): Եթե $p\text{-value} < \alpha$, ապա զրոյական վարկածը պետք է մերժել α արժեքականության մակարդակով: Հակառակ դեպքում, եթե $p\text{-value} \geq \alpha$, ապա զրոյական վարկածը չի կարող մերժվել, ընդունվում է:

6. Բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատման արդյունքների վերլուծության օրինակ

Ուսումնասիրվում է, թե երկրի մեկ շնչի հաշվով ՀՆԱ-ի մեծությամբ վրա ինչ ազդեցություն ունեն երկրի բնակչության կրթամակարդակը, աշխատանքի շուկայում նրանց մասնակցությունն ու ուրբանիզացիայի մակարդակը: Հատկորոշվել է գծային բազմակի ռեգրեսիայի էկոնոմետրիկ մոդել: Համաշխարհային բանկի ու ՄԱԿ-ի զարգացման ծրագրի պաշտոնական կայքէջերից ձեռք են բերվել հետևյալ ցուցանիշներին վերաբերող ամենաթարմ տարեկան տվյալներ աշխարհի 160 երկրների համար՝ աշխատուժի մասնակցության մակարդակը (x_1 , զբաղվածների մասնաբաժինը 16-64 տարեկան աշխատանքային ռեսուրսների թվում, տոկոս), ուրբանիզացիայի մակարդակը (x_2 , քաղաքաբնակ բնակչության մասնաբաժինը երկրի բնակչության թվում, տոկոս), 25 ու ավելի տարեկան բնակչության՝ սովորելու տարիների միջին քանակը (x_3), մեկ շնչի հաշվով ՀՆԱ-ն ընթացիկ ԱՄՆ դոլարով (y): Մոդելի պարամետրիզացումն իրականացվել է ՓԷԵ-ով, SPSS վիճակագրական վերլուծությունների ծրագրային փաթեթով: Գնահատվել է հետևյալ էմպիրիկ մոդելը՝ $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$: Ստացվել են մոդելի գնահատման արդյունքները ներկայացնող ելքային աղյուսակներ, որոնք պետք է մեկնաբանվեն եզրակացություններ անելու համար (նշենք, որ վիճակագրական վերլուծությունների ու էկոնոմետրիկ մոդելավորման համար նախատեսված բոլոր փաթեթները տարբեր ձևավորմամբ գեներացնում են ելքային աղյուսակներ, որոնք պարունակում են այս ելքային աղյուսակներում բերված մեծությունների գերակշիռ մեծամասնությունը):

Model Summary			
R	R Square	Adjusted R Square	Std. Error of the Estimate
0.760	0.578	0.567	14335.9

Model Summary աղյուսակում հաշվարկված է բազմակի կոռելյացիայի գործակիցը, որը ցույց է տալիս մոդելի բացատրող փոփոխականների ամբողջ համախմբի ու կախյալ փոփոխականի միջև գծային կախվածության ուժգնությունը: $R=0.760$, ուստի մոդելի մեջ մտցված բացատրող փոփոխականների համախմբությունը շատ ուժեղ կոռելյացիոն կապի մեջ է կախյալ փոփոխականի հետ: Բազ-

մակի կոռելյացիայի գործակցի քառակուսին՝ $R^2 = 0.578$: Սա մոդելի դետերմինացիայի գործակիցն է (R Square): Աղյուսակում հաշվարկված է նաև ճշգրտված դետերմինացիայի (Adjusted R Square) գործակիցը: Ակնհայտորեն երևում է, որ $R_{adj}^2 = 0.567 < 0.578 = R^2$: ԲԳՌ մոդելի որակի մասին եզրակացությունն արվում է ճշգրտված դետերմինացիայի գործակցի հիման վրա. $R_{adj}^2 = 0.567$: Եզրակացությունը հետևյալն է. կախյալ փոփոխականի ընդհանուր վարիացիայի 56.7 տոկոսը բացատրվում է ռեգրեսիայի հավասարմամբ, իսկ 43.3 տոկոսը՝ պատահական մնացորդներով: Գնահատված հավասարման որակն ընդունելի է, սակայն ոչ շատ բարձր: Std. Error of the Estimate-ը $S_{\hat{\varepsilon}} = \sqrt{S_{\hat{\varepsilon}}^2}$ գնահատման ստանդարտ սխալն է:

ANOVA				
	Sum of Squares	df	F	Sig.
Regression	35466115729.1	3	71.223	0.000
Residual	25893946085.9	156		
Total	61360061815.0	159		

ANOVA աղյուսակի Sum of Squares սյան մեջ բերված են կախյալ փոփոխականի ընդհանուր վարիացիան ու դրա բաղադրիչները: Regression տողում y-ի ընդհանուր վարիացիայի այն բաղկացուցիչն է, որը բացատրվում է ռեգրեսիայի հավասարմամբ. $ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = 35466115729.1$: Նույն սյան Residual տողում RSS-ն է՝ y-ի ընդհանուր վարիացիայի այն բաղկացուցիչը, որը չի բացատրվում ռեգրեսիայի հավասարմամբ, այլ բացատրվում է պատահական ու չհաշվառված ազդեցություններով՝ $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n \hat{\varepsilon}^2 = 25893946085.9$: Այս երկու մեծությունների գումարը y-ի ընդհանուր վարիացիան է, TSS=61360061815.0, որը բերված է նույն սյան Total տողում: Այս սյան միջոցով կարելի է հաշվարկել դետերմինացիայի գործակիցը.

$$R^2 = ESS/TSS = 35466115729.1/61360061815.0 = 0.578:$$

Այս աղյուսակի տվյալների հիման վրա կարելի է հաշվարկել նաև ճշգրտված դետերմինացիայի գործակիցը՝ օգտագործելով df սյան մեջ բերված ազատության աստիճանները: Այդ սյան առաջին տողում ռեգրեսիայի հավասարման բացատրող փոփոխականների թիվն է՝ $m=3$, երկրորդ տողում մոդելի ազատության աստիճանն է՝ $n-(m+1)=160-4=156$, իսկ երրորդ տողում դիտարկումների թիվն է՝ համաձայն մեկ. $n-1=160-1=159$:

$$\text{Ուրեմն՝ } R_{adj}^2 = \frac{ESS}{TSS} \cdot \frac{n-(m+1)}{n-1} = \frac{35466115729.1}{61360061815.0} \cdot \frac{156}{159} = 0.578 \cdot 0.981 = 0.567:$$

Այս տվյալների հիման վրա կարելի է հաշվարկել հավասարման վիճակագրական նշանակալիությունը ստուգող F վիճականին հետևյալ բանաձևով՝

$$F^{\eta_{\text{հս}}} = \frac{ESS}{RSS} \cdot \frac{n-(m+1)}{m} = \frac{35466115729.1}{25893946085.9} \cdot \frac{156}{3} = 71.223:$$

Ֆիշերի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքների աղյուսակից $\alpha=0.05$ արժեքականության մակարդակի և $m=3$, $n-(m+1)=156$ ազատության աստիճանների համար գտնում ենք $F_{m,n-(m+1)}^{\text{կր}} = 2.60$: Ստացվեց. $F^{\eta_{\text{հս}}} = 71.223 > 2.60 = F_{m,n-(m+1)}^{\text{կր}}$: Հետևաբար զրոյական վարկածը պետք է մերժել: Եզրակացություն այն է, որ գնահատված հավասարումը վիճակագրորեն նշանակալի է:

Ֆիշերի բաշխման կրիտիկական արժեքները, $\alpha=0.05$				
$\nu_1 \backslash \nu_2$	1	2	3	4+
1	161.5	199.5	215.7	254.3
2	18.51	19.00	19.16	19.50
3	10.13	9.55	9.28	8.53
4	7.71	6.94	6.59	5.63
5	6.61	5.79	5.41	4.36
10	4.96	4.10	3.71	2.54
15	4.54	3.68	3.29	2.07
20	4.35	3.49	3.10	1.84
25	4.24	3.38	2.99	1.71
30	4.17	3.32	2.92	1.62
40	4.08	3.23	2.84	1.51
45	4.06	3.21	2.81	1.48
50	4.03	3.18	2.79	1.44
60	4.00	3.15	2.76	1.39
70	3.98	3.13	2.74	1.35
80	3.96	3.11	2.72	1.31
90	3.95	3.10	2.71	1.28
100	3.94	3.09	2.70	1.26
125	3.92	3.07	2.68	1.21
150	3.90	3.06	2.66	1.18
151+	3.84	2.99	2.60	1.00

Ֆիշերի բաշխում ունեցող պատահական մեծության կրիտիկական արժեքների աղյուսակը հասանելի է հետևյալ QR կոդով:



Ելքային ANOVA աղյուսակում արդեն իսկ բերված է հաշվարկված F վիճականին, իսկ դրա կողքին՝ p արժեքը (Significance, Sig. սյունում): Դա H_0 -ն մերժելիս սխալվելու կամ, որ նույնն է, H_0 -ի հավանականությունն է: Քանի որ $\text{Sig.}=p\text{-value}=0.0<0.05$, ապա H_0 -ի հավանականությունը փոքր է, և այն մերժվում է: Եզ-

բակացությունն այն է, որ էմպիրիկ հավասարումը վիճակագրորեն նշանակալի է, իսկ տեսական հավասարման բացատրող փոփոխականներից առնվազն մեկը բացատրում է կախյալ փոփոխականին:

Coefficients աղյուսակի B սյան մեջ բերված են գնահատված $\hat{\beta}$ գործակիցները բացատրող փոփոխականներին համապատասխանող տողերում, իսկ (Constant)-ի տողում գրված է հավասարման ազատ անդամը: Std. Error սյունում բերված են $\hat{\beta}$ գործակիցների ստանդարտ սխալները:

Coefficients						
Model	Unstandardized Coefficients		t	Sig.	95.0 % Confidence Interval for B	
	B	Std. Error			Lower Bound	Upper Bound
(Constant)	-47592.68	6811.228	-6.987	0.000	-61046.809	-34138.542
x_1 – աշխատուժի մասնակցության մակարդակ	405.96	108.426	3.744	0.000	191.783	620.127
x_2 – ուրբանիզացիայի մակարդակ	273.37	62.874	4.348	0.000	149.173	397.562
x_3 – սովորած տարիների միջին բանակ	1933.75	464.605	4.162	0.000	1016.024	2851.482

Այսպիսով՝ գնահատված էմպիրիկ հավասարումն է՝

$$\hat{y} = -47592.7 + 405.9x_1 + 273.4x_2 + 1933.8x_3$$

$$(S_{\hat{\beta}_0} = 6811.2) \quad (S_{\hat{\beta}_1} = 108.4) \quad (S_{\hat{\beta}_2} = 62.9) \quad (S_{\hat{\beta}_3} = 464.6)$$

Աղյուսակի t սյան մեջ բերված են տվյալների հիման վրա հաշվարկված t վիճականիները:

$$t_{\hat{\beta}_0}^{\eta\text{իս}} = \frac{\hat{\beta}_0}{S_{\hat{\beta}_0}} = \frac{-47592.7}{6811.2} = -6.987; \quad t_{\hat{\beta}_1}^{\eta\text{իս}} = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}} = \frac{405.9}{108.4} = 3.744:$$

$$t_{\hat{\beta}_2}^{\eta\text{իս}} = \frac{\hat{\beta}_2}{S_{\hat{\beta}_2}} = \frac{273.4}{62.9} = 4.348; \quad t_{\hat{\beta}_3}^{\eta\text{իս}} = \frac{\hat{\beta}_3}{S_{\hat{\beta}_3}} = \frac{1933.8}{464.6} = 4.162:$$

Սյուդենտի բաշխման կրիտիկական արժեքների աղյուսակից կգտնենք, որ $t_{\frac{\alpha}{2}, n-2}^{\text{կր}} \approx 1.96$: Հետևաբար գնահատված բոլոր գործակիցների համար $|t_{\hat{\beta}}^{\eta\text{իս}}| > t_{\frac{\alpha}{2}, n-2}^{\text{կր}}$ պայմանը բավարարվում է, և դրանց բոլորի համար H_0 -ն մերժվում է: Եզրակացությունն այն է, որ տեսական հավասարման բոլոր β անհայտ պարամետրը նշանակալիորեն

գրոյից տարբերվող մեծություններ են, իսկ ռեգրեսիայի էմպիրիկ հավասարման համապատասխան բոլոր β գնահատված գործակիցը վիճակագրորեն նշանակալի գնահատականներ են:

Նույնն են վկայում նաև աղյուսակի Sig. սյունում բերված p արժեքները: Դրանք H_0 -ն մերժելիս սխալվելու (կարելի է ասել՝ H_0 -ի տեղի ունենալու) հավանականություններն են: Քանի որ բոլոր β գործակիցների համար Sig.=p-value=0.0<0.05, ապա H_0 -ի հավանականությունը փոքր է, և այն մերժվում է առանձին-առանձին յուրաքանչյուր β -ի համար:

Ընդհանուր եզրակացությունը հետևյալն է. ՓԶԵ-ով գնահատված էմպիրիկ հավասարման որակը բավական բարձր է, էմպիրիկ հավասարումը վիճակագրորեն նշանակալի է, ռեգրեսիայի էմպիրիկ հավասարման բոլոր β գնահատված գործակիցները վիճակագրորեն նշանակալի գնահատականներ են: Արդեն կարող ենք մեկնաբանել գնահատված հավասարումը:

$\widehat{\beta}_0$ -ն, ինչպես արդեն նշվել է, տնտեսագիտական իմաստովից բովանդակություն կարող է չունենալ: Տվյալ դեպքում դրա մեկնաբանությունն է. եթե երկրում աշխատուժի մասնակցություն աշխատաշուկայում և քաղաքաբնակ բնակչություն ընդհանրապես չլինի, իսկ մարդիկ ոչ մի կրթություն ընդհանրապես չստանան, ապա երկրի մեկ շնչի հաշվով ՀՆԱ-ի սպասվող կամ միջին արժեքը հավասար կլինի -47592.7 -ի: $\widehat{\beta}_1$ -ի մեկնաբանությունը հետևյալն է. այլ հավասար պայմաններում աշխատուժի մասնակցության մակարդակի մեկ միավոր աճը կհանգեցնի մեկ շնչի հաշվով ՀՆԱ-ի սպասվող կամ միջին արժեքի՝ 405.9 ԱՄՆ դոլարով աճի: $\widehat{\beta}_2$ -ի մեկնաբանությունը հետևյալն է. այլ հավասար պայմաններում ուրբանիզացիայի մակարդակի մեկ միավոր աճը կհանգեցնի մեկ շնչի հաշվով ՀՆԱ-ի սպասվող արժեքի 273.4 ԱՄՆ դոլարով աճի: $\widehat{\beta}_3$ -ի մեկնաբանությունն է՝ այլ հավասար պայմաններում սովորած տարիների միջին քանակի մեկ միավոր աճը կհանգեցնի մեկ շնչի հաշվով ՀՆԱ-ի սպասվող արժեքի 1933.8 ԱՄՆ դոլարով աճի:

Coefficient աղյուսակի Lower Bound և Upper Bound սյուններում բերված են արդեն իսկ հաշվարկված ռեգրեսիայի բոլոր գործակիցների վստահելիության միջակայքերի ստորին ու վերին սահմանները: Դրանք փաստում են, որ եթե Գաուս-Մարկովի 5-րդ պայմանը տեղի ունի, ապա 95 % հավանականությամբ տեսական մոդելի պարամետրերն ընկած են հետևյալ միջակայքում՝

$\beta_0 \in [-61046.809; -34138.542]$, $\beta_1 \in [191.783; 620.127]$,

$\beta_2 \in [149.173; 397.562]$, $\beta_3 \in [1016.024; 2851.482]$:

ՓՔԵ-ով գնահատումը հանգեցրել է կիրառելի գնահատականների: Սակայն արդյո՞ք դրանք նաև տեսական անհայտ հավասարման լավագույն գնահատականներն են: Սա պարզելու ու գնահատված էմպիրիկ մոդելի կիրառելիությունը վերջնականորեն պարզելու համար պետք է ստուգվի, թե արդյոք Գաուս-Մարկովի պայմանները ևս տեղի ունեն:

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Ի՞նչ ընդհանուր տեսք ունեն բազմակի գծային ռեգրեսիայի տեսական հավասարումն ու մոդելը, էմպիրիկ հավասարումն ու մոդելը (ոչ մատրիցային տեսքով)
2. Գրե՛ք բազմակի գծային ռեգրեսիայի էմպիրիկ հավասարումների ընդհանուր տեսքը յուրաքանչյուր դիտարկման համար առանձին-առանձին: Քանի՞ հավասարում և քանի՞ անհայտ կունենաք, և ինչպիսի՞ն պետք է լինի դրանց հարաբերակցությունը, որ ՓՔԵ-ով մոդելի գնահատումը հնարավոր լինի:
3. Ի՞նչ ընդհանուր տեսք ունեն բազմակի գծային ռեգրեսիայի տեսական հավասարումն ու մոդելը, էմպիրիկ հավասարումն ու մոդելը (մատրիցային տեսքով):
4. Որո՞նք են ՓՔԵ էությունն ու դրա կիրառման արդյունքները բազմակի գծային ռեգրեսիայի մոդելի համար: Դո՞ւրս բերե՛ք ՓՔԵ-ով գնահատված ռեգրեսիայի գործակիցների բանաձևը մատրիցային տեսքով:
5. Ինչպե՞ս են մեկնաբանվում ԲԳՌ- էմպիրիկ հավասարման գնահատված $\hat{\beta}$ գործակիցներից յուրաքանչյուրը:
6. Ձևակերպե՛ք Գաուս-Մարկովի պայմանները բազմակի գծային ռեգրեսիայի համար: Ի՞նչ պայման է ավելանում այդ պայմաններին պարզ գծային ռեգրեսիայի դեպքի համեմատ:
7. Ձևակերպե՛ք Գաուս-Մարկովի թեորեմը բազմակի գծային ռեգրեսիայի համար Գաուս-Մարկովի 1-4 ու 1-5 պայմանների իրականացման դեպքում:
8. Ի՞նչ տեսք ունի ԲԳՌ- մոդելի պատահական շեղումների կովարիացիայի մատրիցը, եթե Գաուս-Մարկովի պայմանները տեղի ունեն, և ի՞նչ տեսք այն կունենա, եթե հոմոսկեդաստիկության պայմանը տեղի չունենա:

9. Ինչպե՞ս են փոխվում բազմակի գծային ռեգրեսիայի էմպիրիկ մոդելի գնահատված գործակիցների հատկությունները, եթե Գաուս-Մարկովի՝ պատահական շեղումների նորմալ բաշխման պայմանը ևս տեղի ունի:
10. Ինչպիսի՞ հարաբերակցության մեջ են բազմակի ռեգրեսիայի $\hat{B}$ գնահատականների վեկտոր-սյան կովարիացիոն մատրիցը՝ $\Omega_{\hat{B}}$ -ն, և պատահական մնացորդների ε վեկտոր-սյան Ω_{ε} կովարիացիոն մատրիցը: Դո՞ւրս բերե՞ք այդ հարաբերակցությունը:
11. Ի՞նչ է ճշգրտված դետերմինացիայի գործակիցը, ո՞ր մոդելների համար է այն կիրառվում և ինչո՞ւ ու ի՞նչ բանաձևով է հաշվարկվում և ի՞նչ հատկություններ ունի:
12. Արդյո՞ք առանց ազատ անդամի ռեգրեսիայի մոդելի դետերմինացիայի գործակիցը կիրառելի է որպես գնահատված մոդելի որակի չափանիշ: Մեկնաբանե՞ք պատասխանը:
13. Նկարագրե՞ք ռեգրեսիայի հավասարման վիճակագրական նշանակալիության ստուգման թեստի իրականացման սխեման: Ի՞նչ վարկած է դրվում, ինչպե՞ս է ստուգվում, և որո՞նք են թեստի հնարավոր ելքերը:

Խնդիրներ

1. Աշխարհի 168 երկրների որոշակի տարվա տվյալների հիման վրա գնահատվել է դրանց մեկ շնչի հաշվով ՀՆԱ-ի (p_c_GDP , ԱՄՆ դոլար) գծային կախումը արդյունաբերության ոլորտում ավելացված արժեքից ($Manuf_VA$, տոկոս ՀՆԱ-ի նկատմամբ) և երկրում ոչ ֆորմալ ինստիտուտների (խաղի կանոնների) մակարդակը բնութագրող կոռուպցիայի ընկալման համաթվից ($Corrupt$, տոկոս): SPSS փաթեթով մոդելը գնահատելու արդյունքում ստացվել են հետևյալ ելքային աղյուսակները:

Model Summary			
R	R Square	Adjusted R Square	Std. Error of the Estimate
0.715	0.512	0.506	19077.91

ANOVA				
	Sum of Squares	df	F	Sig.
Regression	62950984580.4	2	86.48	0.000
Residual	60054548112.9	165		
Total	123005532693.3	167		

Coefficients

Model	Unstandardized Coefficients		t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error			Lower Bound	Upper Bound
(Constant)	-16297.8	4491.2	-3.629	0.000	-25165.4	-7430.2
Corrupt	724.3	55.1	13.151	0.000	615.5	833.1
Manuf VA	631.5	218.8	2.887	0.004	199.5	1063.5

1. Ինչպիսի՞ն է բացատրող փոփոխականների ամբողջ համախմբի ու կախյալ փոփոխականի միջև գծային կախվածության ուժգնությունը, աղյուսակներում բերված ցուցանիշներից որո՞վ է այն բնութագրվում:
2. Model Summary ու ANOVA աղյուսակներում բերված տվյալների հիման վրա երկու եղանակով հաշվարկե՛ք մոդելի ղետերմինացիայի գործակիցը և ստուգե՛ք՝ արդյոք ձեր հաշվարկները համապատասխանում են Model Summary աղյուսակում բերված ղետերմինացիայի գործակցի արժեքին:
3. Օգտվելով Model Summary աղյուսակում բերված ղետերմինացիայի գործակցի արժեքից ու ANOVA աղյուսակում բերված ազատության աստիճաններից՝ հաշվարկե՛ք մոդելի ճշգրտված ղետերմինացիայի գործակիցը և ստուգե՛ք՝ արդյոք ձեր հաշվարկները համապատասխանում են Model Summary աղյուսակում բերված ճշգրտված ղետերմինացիայի գործակցի արժեքին:
4. Կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
5. Հիմնվելով ղետերմինացիայի գործակցի ու ազատության աստիճանների վրա՝ հաշվարկե՛ք F վիճականին և ստուգե՛ք՝ արդյոք ձեր հաշվարկները համապատասխանում են ANOVA աղյուսակում բերված F վիճականու արժեքին:
6. Ստուգե՛ք ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին վարկածը և կատարե՛ք եզրակացություն ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին:
7. Գրե՛ք ռեգրեսիայի էմպիրիկ հավասարումը և մեկնաբանե՛ք դրա բոլոր գնահատված գործակիցների իմաստը:
8. Հաշվարկե՛ք ռեգրեսիայի յուրաքանչյուր գնահատված գործակցի t վիճականին և ստուգե՛ք՝ արդյոք ձեր հաշվարկները

համապատասխանում են Coefficient աղյուսակում բերված t վիճականիների արժեքներին:

9. Ստուգեք β գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարեք եզրակացություններ այն մասին, թե գնահատված β -ները վիճակագրորեն նշանակալի են, թե ոչ:
 10. Որոշեք տեսական մոդելի անհայտ պարամետրերի միջակայքային գնահատականները կամ վստահելիության միջակայքերը՝ ենթադրելով, որ Գաուս-Մարկովի բոլոր պայմանները տեղի ունեն:
 11. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն, կատարեք ընդհանուր եզրակացություն գնահատված մոդելի կիրառելիության մասին:
2. 150 երկրների տվյալների հիման վրա գնահատվել է մեկ շնչի հաշվով ՀՆԱ-ի գծային կախումը տնտեսության բարդության համաթվից⁵: Ստացվել են հետևյալ արդյունքները.
- $$\hat{y} = 983.7 + 0.8x_1; \quad R^2 = 0.69; \quad F_{\eta}^{h_{\text{un}}} = 4.6$$
- $$(S_{\hat{\beta}_0} = 122.3) \quad (S_{\hat{\beta}_1} = 0.32):$$
1. Գրեք ռեգրեսիայի էմպիրիկ հավասարումն ու մեկնաբանեք դրա բոլոր գնահատված գործակիցների իմաստը:
 2. Հաշվարկեք ռեգրեսիայի յուրաքանչյուր գնահատված գործակցի t վիճականին, ստուգեք β գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարեք եզրակացություն գնահատված β -ների վիճակագրական նշանակալիության մասին:
 3. Կատարեք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
 4. Ստուգեք ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին վարկածը և կատարեք եզրակացու-

⁵ Տնտեսական բարդության համաթիվը (Economic Complexity Index, ECI) ցուցանիշ է, որը գնահատում է տվյալ տնտեսության (երկրի) արտադրական կարողությունների ու տեխնոլոգիաների զարգացման մակարդակը՝ հիմնվելով արտահանվող ապրանքների բազմազանության ու գիտելիքահեղձության/տեխնոլոգիական բարդության վրա: Մանրամասները տես հետևյալ հղմամբ՝

<https://oec.world/en/rankings/eci/hs6/hs96?tab=ranking>:

թյուն ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին:

5. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն` կատարենք ընդհանուր եզրակացություն գնահատված մոդելի կիրառելիության մասին:

3. 5000 տնային տնտեսությունների տվյալների հիման վրա գնահատված է դրանց սպառողական ծախսերի կախումը եկամտից (x_1) ու տնային տնտեսությունում 0-18 տարեկան երեխաների քվաքանակից (x_2): Ստացվել են հետևյալ արդյունքները.

$$\hat{y} = -55752.1 + 0.76x_1 + 15720x_2; \quad R^2 = 0.569$$

$$(S_{\hat{\beta}_0} = 43879.2) \quad (S_{\hat{\beta}_1} = 0.23) \quad (S_{\hat{\beta}_2} = 4563.9):$$

1. Գրենք ռեգրեսիայի էմպիրիկ հավասարումը և մեկնաբանենք դրա բոլոր գնահատված գործակիցների իմաստը:
 2. Հաշվարկենք ռեգրեսիայի յուրաքանչյուր գնահատված գործակցի t վիճականին, ստուգենք $\hat{\beta}$ գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարենք եզրակացություն գնահատված $\hat{\beta}$ -ների վիճակագրական նշանակալիության մասին:
 3. Հաշվարկենք մոդելի ճշգրտված դետերմինացիայի գործակիցը և կատարենք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
 4. Նկատի ունենալով, որ $ESS=14669$ $RSS=11111$ ` հաշվարկենք $F^{դիտ}$ -ը և կատարենք եզրակացություն ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին:
 5. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն, կատարենք ընդհանուր եզրակացություն գնահատված մոդելի կիրառելիության մասին:
4. 146 երկրների տվյալների հիման վրա գնահատված է դրանց մեկ շնչի հաշվով ՀՆԱ-ի (GDP_P_C) գծային կախումը գնաճի մակարդակից ($INFLATION$) ու աշխատուժի մասնակցության մակարդակից ($PARTICIPATION$): EViews փաթեթով մոդելը գնահատելու արդյունքում ստացվել է հետևյալ ելքային արդյունակը`

Dependent Variable: GDP_P_C

Method: Least Squares

Included observations: 146 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INFLATION	-281.16	161.07	-1.75	0.083
PARTICIPATION	811.03	137.49	5.89	0.000
C	-39066.5	9670.73	-4.04	0.000
R-squared	0.238	Mean dependent var		15356.03
Adjusted R-squared	0.227	S.D. dependent var		20294.70
S.E. of regression	17843.21	Akaike info criterion		22.437
Sum squared resid	4.5E ¹⁰	Schwarz criterion		22.498
Log likelihood	-1634.89	Hannan-Quinn crit.		22.462
F-statistic	22.290	Durbin-Watson stat		1.712
Prob(F-statistic)	0.000			

1. Գրե՛ք գնահատված ռեգրեսիայի էմպիրիկ հավասարումը և մեկնաբանե՛ք $\hat{\beta}$ գործակիցների իմաստը:
 2. Կատարե՛ք եզրակացություն գնահատված $\hat{\beta}$ -ների վիճակագրական նշանակալիության մասին: Տեսականորեն հիմնավորե՛ք պատասխանը:
 3. Կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի մոդելի որակի մասին: Տեսականորեն հիմնավորե՛ք պատասխանը:
 4. Կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի հավասարման կամ, որ նույնն է, դետերմինացիայի գործակցի վիճակագրական նշանակալիության մասին:
 5. Կատարե՛ք ընդհանուր եզրակացություն՝ արդյոք գնահատված հավասարումը կիրառելի է, եթե ենթադրենք, որ Գաուս-Մարկովի պայմանները տեղի ունեն:
5. Համաշխարհային բանկի «Աշխարհի զարգացման ցուցանիշներ» (World development indicators) օնլայն տեղեկատվական հարթակից (էլեկտրոնային հասցեն՝ <https://databank.worldbank.org/source/world-development-indicators>) MS Excel ձևաչափով ներբեռնե՛ք գյուղատնտեսությունից առաջացող ածխաթթու գազի արտանետումների ծավալի (մլն տոն-

նայով՝ Carbon dioxide (CO₂) emissions from Agriculture (Mt CO₂e)), գյուղատնտեսության ոլորտում ավելացված արժեքի (ԱՄՆ դոլարով՝ Agriculture, forestry and fishing, value added (current US\$)) ու գյուղատնտեսական հողատարածքների (քառ. կմ՝ Agricultural land (sq. Km)) ցուցանիշների ամենավերջին հասանելի տարվան վերաբերող տվյալները աշխարհի 217 երկրների համար:

1. Ձևափոխելք Excel-ի տվյալներն այնպես, որ դրանք համապատասխանեն էկոմոնետորիկ մոդելավորման համար օգտագործվող տարածական տվյալների երկչափ շտեմարանների պահանջներին:
2. Ձևափոխելք գյուղատնտեսության ոլորտում ավելացված արժեքի ցուցանիշը բոլոր երկրների համար՝ դրանք բերելով մլն ԱՄՆ դոլարի չափողականության:
3. Տեղափոխելք տվյալները SPSS կամ EViews փաթեթ, հետագոտելք ցուցանիշների զույգային կոռելյացիաները և ստացելք կոռելյացիայի դաշտերը:
4. Գնահատելք բազմակի գծային ռեգրեսիայի մոդել՝ գյուղատնտեսությունից առաջացող ածխաթթու գազի արտանետումների ծավալի կախումը գյուղատնտեսության ոլորտում ավելացված արժեքից և գյուղատնտեսական հողատարածքների չափից:
5. Գրելք բազմակի գծային ռեգրեսիայի գնահատված էմպիրիկ հավասարումը և մեկնաբանելք հավասարման գնահատված գործակիցների իմաստը:
6. Ստուգելք β գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարելք եզրակացություն գնահատված β -ների վիճակագրական նշանակալիության մասին:
7. Կատարելք եզրակացություն գնահատված գծային ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
8. Կատարելք եզրակացություն գնահատված գծային ռեգրեսիայի էմպիրիկ հավասարման վիճակագրական նշանակալիության մասին:
9. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն՝ եզրակացություն արելք մոդելի գնահատման արդյունքների կիրառելիության մասին:

ԳԼՈՒԽ 3. ԳԱՌՈՒՍ-ՄԱՐԿՈՎԻ ՊԼՅՄԱՆՆԵՐԻ ԽԱՆՏՈՒՄԸ

ՓԶԵ-ով գծային ռեգրեսիայի էմպիրիկ հավասարման գնահատումը հնարավորություն է տալիս եզրակացություններ անելու հավասարման որակի ու վիճակագրական նշանակալիության, ինչպես նաև գնահատված գործակիցների վիճակագրական նշանակալիության մասին: Այդ եզրակացությունների հիման վրա կարելի է եզրահանգման գալ, թե արդյոք մոդելի ՓԶԵ-ով իրականացված գնահատումը որակյալ ու կիրառելի է: Սակայն էկոնոմետրիկ հետազոտության նպատակներից մեկն էլ այն է, որ գնահատված մոդելը հնարավորինս ճշգրիտ մոտարկի տեսական անհայտ մոդելին: Խնդիր է առաջանում ստուգելու՝ արդյոք ՓԶԵ կիրառման արդյունքում ստացված գնահատականները տեսական անհայտ հավասարման լավագույն գնահատականներն են, և գնահատված մոդելի արդյունքներն ու եզրակացությունները վավերական, վստահելի ու կիրառելի են: Վերջինի տեղի ունենալու համար պետք է բավարարվեն Գաուս-Մարկովի պայմանները: Էմպիրիկ մոդելի կիրառելիությունը վերջնականորեն պարզելու համար պետք է ստուգել՝ արդյո՞ք Գաուս-Մարկովի պայմանները տեղի ունեն: Ստորև կանոդադառնանք այդ պայմաններից յուրաքանչյուրի խախտմանը, դրանց հետևանքներին և ստուգման հնարավորություններին: Եթե դրանցից որևէ մեկը տեղի չունի, ապա գնահատման արդյունքները վստահելի չեն, և մոդելի հատկորոշումը պետք է վերանայել:

1. Պատահական շեղումների նորմալ բաշխվածության պայման. խախտման հետևանքներն ու հայտնաբերումը

Գաուս-Մարկովի թեորեմը չի պահանջում, որ ռեգրեսիայի տեսական մոդելի պատահական շեղումների նորմալ բաշխվածության պայմանը տեղի ունենա: Նույնիսկ եթե այն տեղի չունի, միևնույն է, բոլոր մնացած պայմանների տեղի ունենալն ապահովում է, որ ՓԶԵ-ով ստացված գնահատականները լինեն լավագույն՝ BLUE գնահատականներ, այսինքն՝ լինեն լավագույնը բոլոր գծային անշեղ գնահատականների դասում: Սակայն եթե Գաուս-Մարկովի 5-րդ պայմանը չբավարարվի, դրանց, ինչպես նաև գնահատված հավասարման վիճակագրական նշանակալիության թեստերն այլևս լավ չեն աշխատի, դրանց արդյունքները վստահելի չեն լինի: Շեղումների նորմալ բաշխման պայմանն օգտագործվում է վարկածների ստուգ-

ման ժամանակ, քանի որ t ու F թեստերը հիմնվում են այն ենթադրության վրա, որ շեղումները, հետևաբար նաև β գործակիցները նորմալ են բաշխված:

Եթե շեղումները նորմալ չեն բաշխված, ապա՝

- t -թեստի ու F -թեստի արդյունքները կարող են ունենալ սխալ p արժեքներ և այդպիսով վատահեղի չլինել,
- գնահատված գործակիցների վատահեղիության միջակայքերը կարող են սխալ որոշվել:

Պետք է նշել, որ եթե դիտարկումների թիվը մեծ է, այսինքն՝ մեծ նմուշների համար այս վտանգները չեզոքացվում են՝ ելնելով Կենտրոնական սահմանային թեորեմից: Մեծ նմուշների պարագայում պատահական շեղումները, հետևաբար նաև ռեգրեսիայի գործակիցներն ունենում են ասիմպտոտիկ նորմալ բաշխում, իսկ t ու F վիճակահաները վատահեղի են լինում: Սակայն սահմանելը, թե որն է «մեծ նմուշը», բավական խնդրահարույց է:

Սա նշանակում է, որ Գաուս-Մարկովի այս պայմանի ստուգումը կարևոր է: Քանի որ տեսական մոդելն անհայտ է, պատահական շեղումների նորմալ բաշխածության ստուգման համար կարելի է օգտագործել ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդները, որոնք պատահական շեղումների՝ ՓԲԵ կիրառման արդյունքում ստացված վիճակագրական գնահատականներն են:

Նորմալ բաշխման ստուգման համար վիճակագրական տարբեր փաթեթներ առաջարկում են Ժարկ-Բերրայի թեստը: Ժարկ-Բերրայի թեստը ստուգում է, թե արդյոք ուսումնասիրվող փոփոխականի դիտարկումների տվյալ բազմությունը նմուշ է նորմալ բաշխում ունեցող պատահական մեծությունից: Դրվում են գրոյական ու մրցակցող հետևյալ վարկածները.

H_0 : Ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդները նմուշ են տեսական մոդելի՝ նորմալ բաշխում ունեցող պատահական շեղումներից,

H_1 : Ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդները նմուշ են տեսական մոդելի՝ որևէ այլ, ոչ նորմալ բաշխում ունեցող պատահական շեղումներից:

Թեստը ստուգվում է տեսական մոդելի պատահական շեղումների վիճակագրական գնահատականների՝ էմպիրիկ մոդելի պատահական մնացորդների միջոցով: Ձրոյական վարկածը ստուգելու համար օգտագործվում է Ժարկ-Բերրայի վիճակահան հետևյալ բանաձևով.

$$JB = \frac{n}{6} \cdot \left(S^2 + \frac{(K-3)^2}{4} \right): \quad (75)$$

(75)-ում n -ը նմուշի դիտարկումների թիվն է, S -ը՝ ՓՔԵ-ով էմպիրիկ մոդելի գնահատման արդյունքում հաշվարկված $\hat{\varepsilon}$ պատահական մնացորդների բաշխման ասիմետրիայի գործակիցը, իսկ K -ն՝ կուտակվածությունը: Վիճականին էմպիրիկ տվյալների ասիմետրիան ու կուտակվածությունը համեմատում է նորմալ բաշխման նույն պարամետրերի հետ: Ժարկ-Բերրայի վիճականին ունի ասիմպտոտիկ χ^2 բաշխում՝ 2 ազատության աստիճանով (քանի որ ունենք երկու ստուգվող պարամետր՝ ասիմետրիայի գործակից ու կուտակվածություն): Եթե տվյալ α արժեքականության մակարդակում Ժարկ-Բերրայի վիճականու հաշվարկային արժեքը մեծ է 2 ազատության աստիճանով χ^2 բաշխման կրիտիկական արժեքից, ապա պատահական շեղումների նորմալ բաշխման մասին գրոյական վարկածը մերժվում է, իսկ հակառակ դեպքում չի մերժվում, ընդունվում է.

$JB > \chi_{\alpha}^2 \rightarrow H_0$ -ն մերժվում է, պատահական շեղումները նորմալ չեն բաշխված,

$JB \leq \chi_{\alpha}^2 \rightarrow H_0$ -ն չի կարող մերժվել, պատահական շեղումները նորմալ են բաշխված:

2 ազատության աստիճանով χ^2 բաշխման կրիտիկական արժեքները տարբեր α արժեքականության մակարդակների համար բերված են Աղյուսակ 6-ում:

Աղյուսակ 6. χ^2 բաշխման կրիտիկական արժեքները տարբեր արժեքականության մակարդակների համար

	χ^2 բաշխման կրիտիկական արժեք
0.10 (90%)	4.61
0.05 (95%)	5.99
0.01 (99%)	9.21

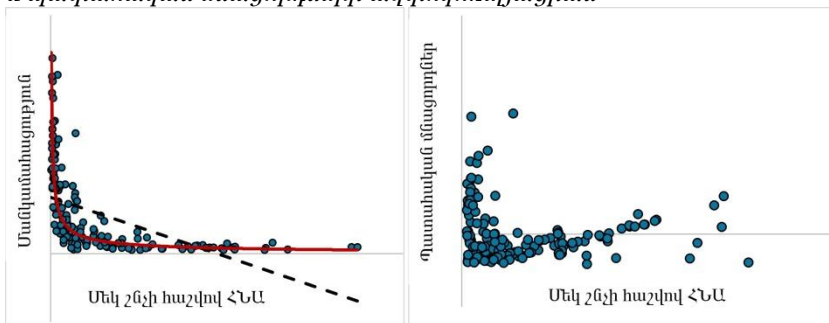
Թեստի ելքը կարող է մատնանշել մաս p արժեքը, որը գրոյական վարկածը մերժելիս սխալվելու կամ, կարելի է ասել, գրոյական վարկածի հավանականությունն է: Եթե $p\text{-value} < \alpha$, պատահական շեղումների նորմալ բաշխման մասին գրոյական վարկածը պետք է մերժել α արժեքականության մակարդակով: Ժարկ-Բերրայի թեստը ենթադրում է մեծ ծավալի նմուշի, այսինքն՝ դիտարկումների մեծ քանակի առկայություն: Սա թեստի թերությունն է, 100-ից պակաս քանակի դիտարկումների համար թեստի արդյունքները վստահելի չեն:

2. Պատահական շեղումների ավտոկորոշվածության բացակայության պայման. խախտման հետևանքներն ու հայտնաբերումը

Գաուս-Մարկովի հաջորդ պայմանը, որի խախտումը կդիտարկենք, պատահական շեղումների ավտոկորոշվածության բացակայության պայմանն է: Այն պահանջում է, որ ռեգրեսիայի տեսական մոդելի պատահական շեղումները կորոշված չլինեն. $COV(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$: Պարզ ասած՝ այս պայմանը ենթադրում է, որ որևէ դիտարկմանը վերաբերող պատահական շեղումը չի ազդում մեկ այլ դիտարկմանը վերաբերող պատահական շեղման վրա: Այս պայմանի խախտումը կոչվում է ավտոկորոշվածություն: Ավտոկորոշվածության կարող է լինել իրական և խաբուսիկ: Իրական ավտոկորոշվածություն սովորաբար հանդիպում է ժամանակային շարքերի մոդելներում. այն բխում է տվյալների հատկություններից: Որոշ դեպքերում ավտոկորոշվածություն կարող է հանդիպել նաև տարածական տվյալները մոդելավորելիս: Այս դեպքում դա «խաբուսիկ» ավտոկորոշվածություն է: Այն պայմանավորված է լինում մոդելի սխալ հատկորոշմամբ (հատկորոշման սխալով)՝ ա) y -ի և x -ի միջև կախվածության սխալ տեսքի ընտրությամբ կամ բ) կարևոր բացատրող փոփոխականը մոդելում չներառելու պատճառով:

Դիտարկենք առաջին դեպքը: Ենթադրենք 150 երկրների տվյալների հիման վրա ուսումնասիրվում է մանկամահացության կախվածությունը երկրի մեկ շնչի հավով ՀՆԱ-ից: Կոռելյացիայի դաշտը հստակորեն վկայում է ուժեղ, սակայն ոչ գծային կախվածության մասին, բայց սխալմամբ հատկորոշվել և գնահատվել է գծային ռեգրեսիայի մոդել (տես Գծանկար 9): Այդ դեպքում x -ի՝ y -ի վրա ունեցած ազդեցության, կախման օրինաչափության մի մասը, որը սխալ հատկորոշման պատճառով չի վերագրվել բացատրող փոփոխականին, կմնա էմպիրիկ մոդելի ε պատահական մնացորդների մեջ, և տարբեր դիտարկումներին համապատասխանող պատահական մնացորդները որոշակի օրինաչափությամբ միմյանցից կախվածություն կունենան: Քանի որ ε -երը տեսական մոդելի ε -երի վիճակագրական գնահատականներն են, սա կվերաբերի նաև տեսական մոդելի պատահական շեղումներին:

Փծանկար 9. y -ի և x -ի միջև կախվածության սխալ տեսքի գնահատումը և պատահական մնացորդների ավտոկորելյացիան



Նույնը տեղի կունենա, եթե y -ի վրա մեծ ազդեցություն ունեցող բացատրող փոփոխականը հատկորոշման սխալի պատճառով չի մտցվել գնահատվող մոդելի մեջ:

Ի՞նչ հետևանքներ է առաջացնում պատահական շեղումների ավտոկորելյացիայի բացակայության պայմանի խախտումը ՓԲԵ-ով գնահատված գործակիցների հատկությունների վրա:

Ավտոկորելյացիայի հետևանքները.

1. Եթե Գաուս-Մարկովի մնացած պայմանները տեղի ունեն, ապա ՓԲԵ-ով գնահատված β գործակիցները մնում են գծային անշեղ գնահատականներ, բայց դրանք արդյունավետ չեն լինում: Նշանակում է՝ դրանք լավագույն գծային գնահատականներ (BLUE գնահատականներ) չեն լինում:

2. β գործակիցների ստանդարտ շեղումներն ավելի փոքր են որոշվում, քան դրանք իրականում պետք է լինեին, և որպես արդյունք՝ դրանց t վիճականիներն ավելի մեծ են լինում: Այդ պատճառով t ու F թեստերի արդյունքները, ինչպես նաև դետերմինացիայի գործակիցը ճշգրիտ չեն լինում և անվստահելի են:

Նշենք, որ մնացորդների ավտոկորելյացիան կարելի է չեզոքացնել կամ շտկել, դրա համար կան ընդունված մեթոդներ (օրինակ՝ Cochrane–Orcutt-ի, Prais–Winsten-ի մեթոդները, փոքրագույն քառակուսիների ընդհանրացված եղանակը): Քանի որ տարածական տվյալների պարագայում մնացորդների ավտոկորելյացիան սովորաբար տվյալների բնույթով պայմանավորված չի լինում, այստեղ այդ մեթոդները քննության չենք առնի:

Ավտոկոռելյացիայի առկայությունը ստուգելու համար կան ընդունված վիճակագրական թեստեր: Դրանցից ամենատարածվածն են բերված են ստորև:

2.1. Առաջին կարգի ավտոկոռելյացիայի առկայության ստուգման Դարբին-Ուոթսոնի թեստը

Ենթադրենք՝ ունենք գծային ռեգրեսիայի մոդել, որտեղ կոռելյացված են անմիջապես իրար հաջորդող դիտարկումների պատահական շեղումները: Այդ դեպքում ասում են, որ **կա առաջին կարգի ավտոկոռելյացիա**:

Դարբին-Ուոթսոնի թեստը ստուգում է առաջին կարգի ավտոկոռելյացիայի առկայությունը: Եթե կա առաջին կարգի ավտոկոռելյացիա, նշանակում է՝ իրար անմիջապես հարևան դիտարկումների ε պատահական շեղումները կոռելյացված են, միմյանցից կախված են: Այդ կոռելյացիան կարելի է ներկայացնել գծային կախվածության տեսքով: Դիտարկենք պարզագույն դեպքը՝ միագործոն գծային ռեգրեսիայի $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ տեսական մոդելը: ՓԲԵ-ով գնահատվել է էմպիրիկ մոդելը: Եթե տեսական մոդելում իրար հաջորդող ε պատահական շեղումները կոռելյացված են, կոռելյացված են նաև էմպիրիկ մոդելի $\hat{\varepsilon}$ -երը, դրանց միջև կա գծային կախվածություն, որը կարելի է ներկայացնել հետևյալ տեսքով՝

$$\varepsilon_i = \rho \cdot \varepsilon_{i-1} + u_i: \quad (76)$$

(76)-ը կոչվում է առաջին կարգի ավտոկոռելյացիայի հավասարում՝ AR(1): ρ -ն կոչվում է **առաջին կարգի ավտոկոռելյացիայի գործակից**: Սովորաբար այն նշանակվում է այսպես՝ $\rho(1)$: (76) մոդելը կարելի է գնահատել ՓԲԵ-ով և գտնել առաջին կարգի ավտոկոռելյացիայի գործակիցի վիճակագրական գնահատականը, որը կնշանակենք $\hat{\rho}(1)$ -ով: Այդ դեպքում, եթե Գաուս-Մարկովի մնացած պայմանները տեղի ունեն, ապա ՓԲԵ-ով ստացված գնահատականը՝ $\hat{\rho}(1)$ -ը, կոռոշվի (22) բանաձևի հիման վրա հետևյալ կերպ՝

$$\hat{\rho}(1) = \frac{\text{cov}(\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1})}{\text{var}(\hat{\varepsilon}_i)} = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i - \bar{\hat{\varepsilon}}) \cdot (\hat{\varepsilon}_{i-1} - \bar{\hat{\varepsilon}})}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2} = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1})}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2}, \quad (77)$$

քանի $\bar{\hat{\varepsilon}} = 0$ բոլոր $i=1 \dots n$ դիտարկումների համար:

Յույց տանք, որ (77) բանաձևով որոշված առաջին կարգի ավտոկոռելյացիայի գործակցի վիճակագրական գնահատականը իսկապես իրար հաջորդող դիտարկումներին համապատասխանող պատահական մնացորդների կոռելյացիայի գործակիցն է:

$$r_{\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1}} = \frac{\text{cov}(\hat{\varepsilon}_i, \hat{\varepsilon}_{i-1})}{\sqrt{\text{var}(\hat{\varepsilon}_i) \cdot \text{var}(\hat{\varepsilon}_{i-1})}} = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i - \bar{\hat{\varepsilon}}_i) \cdot (\hat{\varepsilon}_{i-1} - \bar{\hat{\varepsilon}}_{i-1})}{\sqrt{\sum_{i=1}^n \hat{\varepsilon}_i^2 \cdot \sum_{i=2}^n \hat{\varepsilon}_{i-1}^2}} \approx \frac{\sum_{i=2}^n (\hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1})}{\sum_{i=1}^n \hat{\varepsilon}_i^2} = \hat{\rho}(1):$$

Այսպիսով՝ (77) բանաձևով որոշվող $\hat{\rho}(1)$ -ը (76) մոդելի առաջին կարգի ավտոկոռելյացիայի գործակցի նմուշային վիճակագրական գնահատականն է:

Առաջադրենք հետևյալ գրոյական վարկածը՝

H_0 : $\rho(1)=0$, տեսական մոդելի պատահական շեղումների միջև չկա առաջին կարգի ավտոկոռելյացիա,

H_1 : $\rho(1) \neq 0$, տեսական մոդելի պատահական շեղումների միջև կա առաջին կարգի ավտոկոռելյացիա:

H_0 վարկածը կարելի է ստուգել $\hat{\rho}(1)$ -ի օգտագործմամբ: Վարկածը ստուգվում է Դարբին-Ուոթսոնի վիճականու միջոցով, որը հաշվարկվում է հետևյալ բանաձևով՝

$$DW = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i - \hat{\varepsilon}_{i-1})^2}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2} = \frac{\sum_{i=2}^n (\hat{\varepsilon}_i^2 - 2 \cdot \hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1} + \hat{\varepsilon}_{i-1}^2)}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2} \approx$$

$$\approx 1 - 2 \cdot \frac{\sum_{i=2}^n (\hat{\varepsilon}_i \cdot \hat{\varepsilon}_{i-1})}{\sum_{i=1}^n (\hat{\varepsilon}_i)^2} + 1 \approx 2 \cdot (1 - \hat{\rho}(1)): \quad (78)$$

Այսպիսով՝

$$DW \approx 2 \cdot (1 - \hat{\rho}(1)): \quad (79)$$

Քանի որ $\hat{\rho}(1)$ -ը էմպիրիկ մոդելից ստացված նմուշային կոռելյացիայի գործակից է, այն ունի կոռելյացիայի գործակցի բոլոր հատկությունները, մասնավորապես՝ $-1 \leq \hat{\rho}(1) \leq 1$, հետևաբար Դարբին-Ուոթսոնի վիճականին ընկած է $[0, 4]$ միջակայքում՝

$$0 \leq DW \leq 4: \quad (80)$$

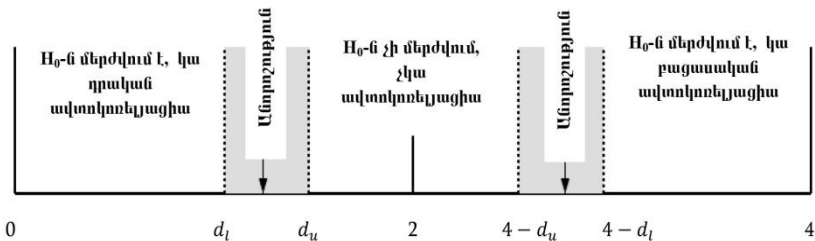
(79)-ից հետևում է, որ եթե՝

- կա բավական ուժեղ $\rightarrow \hat{\rho}(1) \approx 1 \rightarrow DW \approx 2 \cdot (1 - 1) \rightarrow DW \approx 0$
դրական ավտոկոռելյացիա
- չկա ավտոկոռելյացիա $\rightarrow \hat{\rho}(1) \approx 0 \rightarrow DW \approx 2 \cdot (1 - 0) \rightarrow DW \approx 2$
- կա բավական ուժեղ $\rightarrow \hat{\rho}(1) \approx -1 \rightarrow DW \approx 2 \cdot (1 + 1) \rightarrow DW \approx 4$
բացասական ավտոկոռելյացիա

Դարբին-Ուոթսոնի վիճականու միջոցով զրոյական վարկածը ստուգելու համար մշակված են Դարբին-Ուոթսոնի վիճականու կրիտիկական արժեքների աղյուսակներ: Սակայն դրանցով որոշվում է ոչ թե մեկ որոշակի կրիտիկական արժեք, այլ կրիտիկական արժեքների միջակայք՝ դրանց ստորին ու վերին սահմանների արժեքներով: Նշված աղյուսակներից α արժեքականության մակարդակում α ու դիտարկումների ու m բացատրող փոփոխականների համար որոշվում է երկու թիվ՝ d_l - կրիտիկական միջակայքի ստորին սահման (lower bound) ու d_u - կրիտիկական միջակայքի վերին սահման (upper bound): Դարբին-Ուոթսոնի վիճականու հաշվարկային դիտարկված արժեքը $DW_{\text{դիտ}}$ -ը, α արժեքականությամբ ընկած է $d_l \leq DW_{\text{դիտ}} \leq d_u$ կրիտիկական միջակայքում:

Զրոյական վարկածը ստուգելու համար DW վիճականու հաշվարկված արժեքը պետք է համեմատել Դարբին-Ուոթսոնի բաշխման կրիտիկական արժեքների վերին ու ստորին սահմանների հետ: Որոշման կայացման մոտեցումը կամ Դարբին-Ուոթսոնի թեստի ելքերը ներկայացված են Գծանկար 10-ում:

Գծանկար 10. Դարբին-Ուոթսոնի վիճականու հնարավոր արժեքների տիրույթն ու թեստի ելքերը



Դարբին-Ուոթսոնի թեստի պրակտիկ օգտագործման համար կարելի է օգտվել նաև Աղյուսակ 7-ից:

Աղյուսակ 7. Դարբին-Ուոթսոնի վիճականու հաշվարկային արժեքներն ու թեստի ելքերը

Եթե DW վիճականու հաշվարկ- ված արժեքը.	Ապա թեստի ելքն է.
$0 \leq DW_{\eta_{\text{իտ}}} < d_l$	Որոշակի հավանականությամբ մոդելում առկա է առաջին կարգի դրական ավտոկորելյացիա:
$d_l \leq DW_{\eta_{\text{իտ}}} < d_u$	Թեստի արդյունքն անորոշ է:
$d_u \leq DW_{\eta_{\text{իտ}}} < 4 - d_u$	Որոշակի հավանականությամբ մոդելում առկա չէ առաջին կարգի ավտոկորելյացիա:
$4 - d_u \leq DW_{\eta_{\text{իտ}}} < 4 - d_l$	Թեստի արդյունքն անորոշ է:
$4 - d_l \leq DW_{\eta_{\text{իտ}}} < 4$	Որոշակի հավանականությամբ մոդելում առկա է առաջին կարգի բացասական ավտոկորելյացիա:

Բավականաչափ մեծ նմուշների համար կոպիտ կանոնն այն է, որ եթե DW վիճականու հաշվարկային արժեքն ընկած է մոտավորապես 1.8-2.2 միջակայքում, ապա առաջին կարգի ավտոկորելյացիա չկա:

Իսկ ինչպե՞ս ստուգել ավելի բարձր կարգի ավտոկորելյացիայի առկայությունը, այսինքն՝ ոչ թե իրար անմիջապես հաջորդող դիտարկումներին համապատասխանող պատահական շեղումների, այլ 2, 3 ու ավելի քայլերով իրարից հեռացած դիտարկումներին համապատասխանող պատահական շեղումների կոռելյացվածությունը:

2.2. Բրեուշ-Գոդֆրիի՝ բարձր կարգերի ավտոկորելյացիայի առկայության ստուգման Լագրանժի բազմապատկչի LM թեստը

Ավելի բարձր կարգի ավտոկորելյացիայի գոյության կասկածի դեպքում կարելի է օգտվել Բրեուշ-Գոդֆրիի Լագրանժի բազմապատկչի թեստից (Serial correlation LM test): Թեստը դնում է զրոյական վարկած այն մասին, որ տեսական մոդելի տարրեր դիտարկումներին համապատասխանող պատահական շեղումների միջև առաջինից մինչև որևէ p -րդ կարգի ավտոկորելյացիա չկա: Մրցակցող վարկածը 1-ից p -րդ կարգի առնվազն մեկ կարգի ավտոկորելյացիայի առկայությունն է:

$H_0 : \rho(1) = \rho(2) = \dots = \rho(p) = 0$, տեսական մոդելի պատահական շեղումների միջև չկա մինչև p -րդ կարգի ավտոկորելյացիա,

H_1 : տեսական մոդելի պատահական շեղումների միջև կա միմյնային կարգի առնվազն մեկ ավտոկոռելյացիա՝ որևէ $\rho(k) \neq 0, k \leq p$:

Թեստը ստուգվում է տեսական մոդելի պատահական շեղումների վիճակագրական գնահատականների՝ ՓԲԵ-ով ստացված պատահական մնացորդների միջոցով: Ընդհանուր դեպքում ենթադրենք ունենք բազմակի գծային ռեգրեսիայի մոդել, որը գնահատվել է ՓԲԵ-ով, և ստացվել են էմպիրիկ մոդելի պատահական մնացորդները:

Թեստի առաջին քայլում ՓԲԵ-ով գնահատվում է ռեգրեսիայի էմպիրիկ հավասարումը՝

$$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \widehat{\beta}_2 x_{i2} + \dots + \widehat{\beta}_m x_{im},$$

և հաշվարկվում են էմպիրիկ մոդելի պատահական $\widehat{\varepsilon}_i$ մնացորդները՝ $\widehat{\varepsilon}_i = y_i - \widehat{y}_i$:

Երկրորդ քայլում գնահատվում է օժանդակ ռեգրեսիա, որտեղ $\widehat{\varepsilon}_i$ -երը կախված են բոլոր բացատրող փոփոխականներից ու $\widehat{\varepsilon}_i$ -երի p թվով այլ դիտարկումներից: Օժանդակ ռեգրեսիայի հավասարումը հետևյալն է՝

$$\widehat{\varepsilon}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \widehat{\beta}_2 x_{i2} + \dots + \widehat{\beta}_m x_{im} + \sum_{s=1}^p \gamma_s \widehat{\varepsilon}_{i-s} + u_i: \quad (81)$$

Չրոյական վարկածը ստուգելու նպատակով գնահատված օժանդակ ռեգրեսիայի համար հաշվարկվում են երկու վիճականիներ՝ F վիճականին ու Բրեուշ-Գոդֆրիի $Obs \cdot R^2$ վիճականին: Եթե F վիճականու բաշխումն իրականում անհայտ է, ապա $Obs \cdot R^2$ վիճականին ունի ասիմպտոտիկ χ^2 բաշխում՝ p ազատության աստիճանով, և հաշվարկվում է որպես (81) օժանդակ ռեգրեսիայի դիտարկումների թվի (Obs) և դետերմինացիայի գործակցի (R^2) արտադրյալ:

Եթե $Obs \cdot R^2 > \chi_{\alpha p}^2 \longrightarrow$ ավտոկոռելյացիայի բացակայության գրոյական վարկածը մերժվում է տվյալ արժեքականության α մակարդակում,

Եթե $Obs \cdot R^2 \leq \chi_{\alpha p}^2 \longrightarrow$ ավտոկոռելյացիայի բացակայության գրոյական վարկածը չի կարող մերժվել, ընդունվում է արժեքականության α մակարդակում:

Թեստի ելքը որոշելու համար կարելի է մաս օգտագործել p արժեքը: Եթե $p\text{-value} < \alpha$, մինչև p -րդ կարգի ավտոկորելյացիայի բացակայության մասին զրոյական վարկածը պետք է մերժել՝ α արժեքականության մակարդակով, հակառակ դեպքում՝ ընդունել այն:

3. Պատահական շեղումների հոմոսկեդաստիկության պայման. խախտման հետևանքները, հայտնաբերումն ու շտկումը

3.1. Հետերոսկեդաստիկության էությունը, հետևանքներն ու գրաֆիկական հայտնաբերումը

Գաուս-Մարկովի պայմանները պահանջում են, որ ռեգրեսիայի տեսական մոդելի պատահական շեղումների վարիացիան լինի հաստատուն՝ $Var(\varepsilon_i) = \sigma_\varepsilon^2$: Սա հոմոսկեդաստիկության պայմանն է: Ընդհանուր դեպքում բազմակի ռեգրեսիայի համար դա նշանակում է, որ պատահական շեղումների կովարիացիոն մատրիցի անկյունագծի վրա մույն σ_ε^2 -ներն են մատրիցի բոլոր տողերում: Այսինքն՝

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & \dots & 0 \\ 0 & \sigma_\varepsilon^2 & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & \sigma_\varepsilon^2 \end{bmatrix}$$

Եթե հոմոսկեդաստիկության պայմանը չի բավարարվում, և տեսական մոդելի ε -երը հետերոսկեդաստիկ կան, որ մույնն է, բազմացվածքային են, իսկ ԳՄ մյուս պայմանները բավարարված են, ապա տեսական մոդելի ε -երի կովարիացիոն մատրիցի անկյունագծի վրա կլինեն միմյանցից տարբերվող վարիացիաներ՝

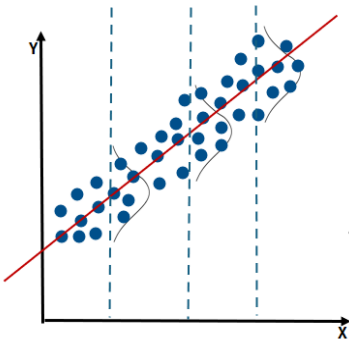
$$\Omega_\varepsilon = \sigma_\varepsilon^2 I = \begin{bmatrix} \sigma_{\varepsilon_1}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{\varepsilon_2}^2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \sigma_{\varepsilon_n}^2 \end{bmatrix}$$

Հոմոսկեդաստիկության պայմանի էությունն այն է, որ չպետք է լինի որևէ պատճառ, որ որոշակի դիտարկումների համար ε պատահական շեղումները մեծ լինեն, իսկ մնացածի համար՝ փոքր:

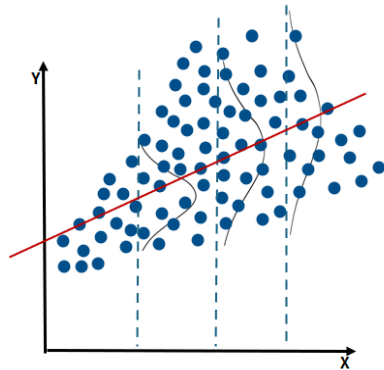
Քանի որ ՓԲԵ-ով գնահատված հավասարման $\hat{\varepsilon}$ մնացորդները տեսական մոդելի ε -երի վիճակագրական գնահատականներն են, իսկ վերջիններս անհայտ են, տեսական մոդելի ε -երի հետերոսկեդաստիկության առկայությունը արտահայտվում է նաև $\hat{\varepsilon}$ մնացորդների մեջ և ստուգվում է դրանց միջոցով: Օրինակ՝ Գծանկար 11-ում առաջին դեպքում ունենք հոմոսկեդաստիկություն, դիտարկումներից կազմված կետերի ցրումը գնահատված գծի շուրջը ($\hat{\varepsilon}$ մնացորդները) գնահատված գծի ամբողջ երկայնքով հաստատուն է: Դա նշանակում է, որ չնայած x -ի աճի հետ y -ն աճում է, բայց x -ի տարբեր կոնկրետ արժեքների համար y -ի ցրումն իր միջինի շուրջը մնում է նույնը, x -ի աճի հետ $\hat{\varepsilon}$ մնացորդների ցրվածքը (վարիացիան) մնում է հաստատուն: Երկրորդ դեպքում x -ի աճի հետ x -ի ամեն կոնկրետ արժեքի դեպքում y -ի ցրումն իր միջինի շուրջը ևս աճում է: Արդյունքում x -ի աճի հետ աճում են նաև $\hat{\varepsilon}$ մնացորդները և դրանց ցրումն իրենց միջինի շուրջը: Սա հետերոսկեդաստիկության օրինակ է:

Գծանկար 11. x և y փոփոխականների կոռելյացիայի դաշտն ու հետերոսկեդաստիկությունը

Հոմոսկեդաստիկություն



Հետերոսկեդաստիկություն



Տարածական տվյալների դեպքում $\hat{\varepsilon}$ մնացորդների հետերոսկեդաստիկությունը կարող է լինել իրական, այսինքն՝ տվյալների հատկություններից բխող, ինչպես նաև՝ խաբուսիկ: Վերջինը պայմանավորված է լինում հատկորոշման սխալներով:

Ի՞նչ տեղի կունենա, եթե սովորական ՓԲԵ-ով գնահատվի մի հավասարում, որի համար խախտված է հոմոսկեդաստիկության

պայմանը, սակայն Գաուս-Մարկովի մյուս բոլոր պայմանները բավարարված են.

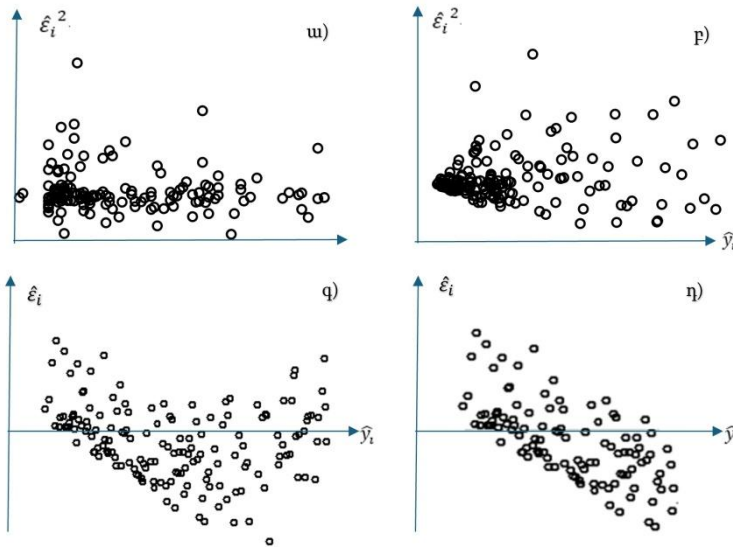
1. β գնահատականները գծային անշեղ գնահատականներ են մնում, բայց դրանք այլևս արդյունավետ չեն լինում, այսինքն՝ չեն ունենում ամենավորը վարիացիան գծային անշեղ գնահատականների դասում, ուրեմն դրանք BLUE գնահատականներ չեն լինում:

2. β գնահատականների ստանդարտ սխալները հաշվարկվում են ավելի փոքր, քան դրանք կան իրականում: Սա կարող է հանգեցնել t ու F թեստերի հիման վրա կատարված եզրակացությունների անարժանահավատության:

Հետերոսկեդաստիկությունը կարելի է հայտնաբերել մի քանի եղանակներով: Եթե հետերոսկեդաստիկության բնույթի մասին չկա նախնական տեղեկատվություն, գործնականում հնարավոր է իրականացնել հետահայաց վերլուծություն՝ գնահատել մոդելը՝ ենթադրելով, որ հետերոսկեդաստիկություն չկա, այնուհետև՝ իրականացնել $\hat{\varepsilon}$ մնացորդների ուսումնասիրություն՝ տեսնելու համար՝ արդյոք դրանք որևէ օրինաչափություն են դրսևորում:

Հետերոսկեդաստիկության հայտնաբերման ամենապարզ եղանակը ՓԶԵ-ով մոդելի գնահատման արդյունքում հաշվարկված $\hat{\varepsilon}$ մնացորդների գրաֆիկական հետազոտությունն է: Սովորաբար հետերոսկեդաստիկության դեպքում $\hat{\varepsilon}$ մնացորդների ցրվածքը կախված է լինում բացատրող փոփոխականներից որևէ մեկից կամ մի քանիսից: Այդ պատճառով հետերոսկեդաստիկության հայտնաբերման համար գրաֆիկորեն դիտարկում են $\hat{\varepsilon}_i$ -երի կամ դրանց քառակուսիների կախումը $\hat{y}_i$ -երից, քանի որ վերջիններս գծային կոմբինացիա են բոլոր x -երից՝ $\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2 + \dots + \widehat{\beta}_m x_m$: Եթե գրաֆիկում $\hat{\varepsilon}$ մնացորդների ցրվածքը ոչ մի ակնհայտ օրինաչափություն, ցրվածքի փոփոխություն, կախում չի դրսևորում, ապա կարելի է ենթադրել, որ հետերոսկեդաստիկություն չկա: Հակառակ դեպքում եթե նկատվում է մնացորդների (կամ դրանց քառակուսիների) կամ դրանց ցրվածքի որոշակի կախում $\hat{y}_i$ -երից, կարելի է ենթադրել հետերոսկեդաստիկության առկայություն: Կարելի է դիտարկել նաև $\hat{\varepsilon}_i$ -երի քառակուսիների կախումը բոլոր x -երից առանձին-առանձին պատկերող գրաֆիկները (տես Գծանկար 12):

Քծանկար 12. Պարահական մնացորդների հոմոսկետաստիկության (ա) ու հետերոսկետաստիկության (բ, գ, դ) օրինակների գրաֆիկական պարկերը



Հոմոսկետաստիկությունը կամ դրա խախտումը ավելի վստահաբար կարելի է ստուգել վիճակագրական թեստերի միջոցով:

3.2. Ուայթի ու Բրեուշ-Պագանի՝ հետերոսկետաստիկության առկայության ստուգման թեստերը

Ուայթի թեստ: Ուայթի թեստում դրվում է գրոյական վարկած այն մասին, որ տեսական մոդելի պատահական շեղումների վարիացիան հաստատուն է, այսինքն՝ դրանք հոմոսկետաստիկ են: Մյուսակցող վարկածն այն է, որ վերջիններս հետերոսկետաստիկ կամ բազմացրվածքային են:

$H_0: Var(\varepsilon_i) = \sigma_\varepsilon^2 = const, \forall i = 1, 2 \dots n$ – հոմոսկետաստիկություն,

$H_1: Var(\varepsilon_i) = \sigma_{\varepsilon_i}^2 \neq const$ – հետերոսկետաստիկություն:

Թեստն իրականացնում է հետևյալ քայլերով.

Քայլ 1. Գնահատվում է ռեգրեսիայի էմպիրիկ հավասարումը՝ $\hat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2 + \dots + \widehat{\beta}_m x_m$, և հաշվարկվում են էմպիրիկ մոդելի մնացորդները՝ $\hat{\varepsilon}_i$ -ները:

Քայլ 2. Գնահատվում է օժանդակ ռեգրեսիայի մոդել, որում $\hat{\varepsilon}_i$ -երի քառակուսիները կախված են բոլոր x -երից, դրանց քառակուսիներից և (ընտրողաբար) գույգ առ գույգ արտադրյալներից: Օրինակ, եթե առաջին քայլում գնահատվել է երկու բացատրող փոփոխականով գծային բազմագործոն ռեգրեսիայի մոդել, ապա օժանդակ ռեգրեսիայի գնահատվող մոդելն ունի այս տեսքը՝

$$\hat{\varepsilon}_i^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 x_{i1} + \widehat{\alpha}_2 x_{i2} + \widehat{\alpha}_3 x_{i1}^2 + \widehat{\alpha}_4 x_{i2}^2 + \widehat{\alpha}_5 x_{i1} x_{i2} + \hat{u}_i: \quad (82)$$

Քայլ 3. Հաշվարկվում է օժանդակ ռեգրեսիայի դիտարկումների թվի ու դետերմինացիայի գործակցի արտադրյալը՝ $Obs \cdot R^2$, և ստուգվում է H_0 -ն: Ուայթի թեստի վիճականին $Obs \cdot R^2$ -ն է, այն ունի ասիմպտոտիկ χ^2 բաշխում՝ օժանդակ ռեգրեսիայի գնահատվող պարամետրերից մեկով պակաս (կամ, որ նույնն է, օժանդակ ռեգրեսիայի բացատրող փոփոխականների քանակին հավասար) ազատության աստիճանով՝ $Obs \cdot R^2 \sim \chi_m^2$: Այս թեստում հաշվարկվում է մաս F վիճականի: Հարկ է նշել, որ գրոյական վարկածի դեպքում F վիճականու բաշխումն անհայտ է: Այդ պատճառով նախընտրելի է օգտվել $Obs \cdot R^2$ վիճականուց:

Եթե $Obs \cdot R^2 > \chi_{\alpha}^2 \longrightarrow$ հոմոսկեդաստիկության գրոյական վարկածը մերժվում է տվյալ արժեքականության α մակարդակում,

Եթե $Obs \cdot R^2 \leq \chi_{\alpha}^2 \longrightarrow$ հոմոսկեդաստիկության գրոյական վարկածը չի կարող մերժվել, ընդունվում է արժեքականության α մակարդակում:

Թեստի ելքը որոշելու համար կարելի է մաս օգտագործել p արժեքը: Եթե $p\text{-value} < \alpha$, ապա հոմոսկեդաստիկության գրոյական վարկածը պետք է մերժել α արժեքականության մակարդակով: Հակառակ դեպքում պետք է ընդունել այն:

Բրեուշ-Պազանի թեստ: Այս թեստում ևս դրվում է H_0 վարկած. տեսական մոդելի պատահական շեղումների վարիացիան հաստատուն է, այսինքն՝ դրանք հոմոսկեդաստիկ են: Մրցակցող վարկածն այն է, որ վերջիններս հետերոսկեդաստիկ կամ բազմացրվածքային են:

$$H_0 : Var(\varepsilon_i) = \sigma_\varepsilon^2 = const, \forall i = 1, 2 \dots n$$

$$H_1 : Var(\varepsilon_i) = \sigma_{\varepsilon_i}^2 \neq const:$$

Թեստն իրականացնում է հետևյալ քայլերով.

Քայլ 1. Գնահատվում է ռեգրեսիայի էմպիրիկ հավասարումը՝ $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_m x_m$, և հաշվարկվում են էմպիրիկ մոդելի մնացորդները՝ $\hat{\varepsilon}_i$ -երը:

Քայլ 2. Գնահատվում է օժանդակ ռեգրեսիայի մոդել, որում $\hat{\varepsilon}_i$ -երի քառակուսիները կախված են բոլոր x -երից՝ $\hat{\varepsilon}_i^2 = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \hat{\alpha}_2 x_2 + \dots + \hat{\alpha}_m x_m + \hat{u}_i$: Հոմոսկեդաստիկության գրոյական վարկածը համարժեք է նրան, որ օժանդակ մոդելի բոլոր գործակիցները, բացի $\hat{\alpha}_0$ -ից, և դրա դետերմինացիայի գործակիցը հավասար լինեն զրոյի:

Քայլ 3. Հաշվարկվում է օժանդակ ռեգրեսիայի դիտարկումների թվի ու դետերմինացիայի գործակցի արտադրյալը՝ $Obs \cdot R^2$, և ստուգվում է H_0 -ն: Այս թեստի վիճականին ևս $Obs \cdot R^2$ -ն է, այն ունի ասիմպտոտիկ χ^2 բաշխում՝ օժանդակ ռեգրեսիայի գնահատվող պարամետրերից մեկով պակաս (դրա բացատրող փոփոխականների քանակին հավասար) ազատության աստիճանով՝ $Obs \cdot R^2 \sim \chi^2_m$:

Եթե $Obs \cdot R^2 > \chi_{\alpha}^2 \longrightarrow$ հոմոսկեդաստիկության գրոյական վարկածը մերժվում է տվյալ արժեքականության α մակարդակում,

Եթե $Obs \cdot R^2 \leq \chi_{\alpha}^2 \longrightarrow$ հոմոսկեդաստիկության գրոյական վարկածը չի կարող մերժվել, ընդունվում է արժեքականության α մակարդակում:

Թեստի ելքը որոշելու համար կարելի է մաս օգտագործել p արժեքը: Եթե $p\text{-value} < \alpha$, հոմոսկեդաստիկության մասին գրոյական վարկածը պետք է մերժել α արժեքականության մակարդակով: Հակառակ դեպքում պետք է ընդունել այն:

3.3. Հետերոսկեդաստիկության շտկումը. փոքրագույն քառակուսիների կշռված և բնօրինարացված եղանակներ

Ի՞նչ կարելի է անել, եթե տվյալների բնույթն այնպիսին է, որ հո-
մոսկեղաատիկության պայմանը խախտվում է, և ՓԶԵ-ը լավագույն
գնահատականներ չի գեներացնում: Կամ մի քանի ընդունված մոտե-
ցումներ: Դիտարկենք դրանք:

1. ՓՔ կշռված եղանակ

Այս եղանակը կիրառելի է այն դեպքում, եթե $\hat{\varepsilon}_i$ պատահական մնացորդների ստանդարտ սխալները՝ $\hat{\sigma}_{\varepsilon_i}$ -երը, բոլոր դիտարկումների համար ($i=1, \dots, n$) հայտնի են: Այդ դեպքում կարելի է էմպիրիկ տվյալների բոլոր դիտարկումները կշռել նույն այդ դիտարկմանը համապատասխանող $\hat{\varepsilon}_i$ -ի ստանդարտ սխալով՝ $\hat{\sigma}_{\varepsilon_i}$ -ով: Այդ դեպքում էմպիրիկ տվյալների դիտարկումները՝

$$(x_{1,i}, x_{2,i}, x_{3,i} \dots x_{m,i}, y_i), i = 1, 2, \dots, n,$$

ԿՎԻՄԱՐԻՆՎԵՆ ԻԵՏԼՅԱԼ ՌԻՏԱՐԿՈՒՄՆԵՐՈՎ՝

$$\left(\frac{x_{1,i}}{\hat{\sigma}_{\varepsilon_i}}, \frac{x_{2,i}}{\hat{\sigma}_{\varepsilon_i}}, \frac{x_{3,i}}{\hat{\sigma}_{\varepsilon_i}} \dots \frac{x_{m,i}}{\hat{\sigma}_{\varepsilon_i}}, \frac{y_i}{\hat{\sigma}_{\varepsilon_i}}\right), i = 1, 2, \dots, n:$$

Դա կհանգեցնի նրան, որ չի-երի մեծ վարիացիաներին համապատասխանող դիտարկումները կբաժանվեն մեծ ստանդարտ սխալների վրա և տվյալների նոր համախմբության վրա գնահատվող մոդելի մեջ կմտնեն փոքր կշիռներով, իսկ փոքր վարիացիաներին համապատասխանող դիտարկումները կբաժանվեն փոքր ստանդարտ սխալների վրա և տվյալների վրա գնահատվող մոդելի մեջ կմտնեն մեծ կշիռներով:

Յուրաքանչյուր դիտարկման համար ռեգրեսիայի հավասարումը կրաժանվի $\hat{\sigma}_{\varepsilon_i}$ -ու վրա՝

$$\begin{aligned} \text{i} = 1 & \quad \frac{y_1}{\widehat{\sigma}_{\varepsilon_1}} = \frac{\beta_0}{\widehat{\sigma}_{\varepsilon_1}} + \beta_1 \frac{x_{1,1}}{\widehat{\sigma}_{\varepsilon_1}} + \beta_2 \frac{x_{2,1}}{\widehat{\sigma}_{\varepsilon_1}} + \cdots + \beta_m \frac{x_{m,1}}{\widehat{\sigma}_{\varepsilon_1}} + \frac{\varepsilon_1}{\widehat{\sigma}_{\varepsilon_1}} \\ \text{i} = 2 & \quad \frac{y_2}{\widehat{\sigma}_{\varepsilon_2}} = \frac{\beta_0}{\widehat{\sigma}_{\varepsilon_2}} + \beta_1 \frac{x_{1,2}}{\widehat{\sigma}_{\varepsilon_2}} + \beta_2 \frac{x_{2,2}}{\widehat{\sigma}_{\varepsilon_2}} + \cdots + \beta_m \frac{x_{m,2}}{\widehat{\sigma}_{\varepsilon_2}} + \frac{\varepsilon_2}{\widehat{\sigma}_{\varepsilon_2}} \\ & \dots\dots\dots \\ \text{i} = n & \quad \frac{y_n}{\widehat{\sigma}_{\varepsilon_n}} = \frac{\beta_0}{\widehat{\sigma}_{\varepsilon_n}} + \beta_1 \frac{x_{1,n}}{\widehat{\sigma}_{\varepsilon_n}} + \beta_2 \frac{x_{2,n}}{\widehat{\sigma}_{\varepsilon_n}} + \cdots + \beta_m \frac{x_{m,n}}{\widehat{\sigma}_{\varepsilon_n}} + \frac{\varepsilon_n}{\widehat{\sigma}_{\varepsilon_n}} \end{aligned}$$

Ընդհանուր դեպքում կունենանք՝

$$\frac{y_i}{\hat{\sigma}_{\varepsilon_i}} = \frac{\beta_0}{\hat{\sigma}_{\varepsilon_i}} + \beta_1 \frac{x_{1,i}}{\hat{\sigma}_{\varepsilon_i}} + \beta_2 \frac{x_{2,i}}{\hat{\sigma}_{\varepsilon_i}} + \dots + \beta_m \frac{x_{m,i}}{\hat{\sigma}_{\varepsilon_i}} + \frac{\varepsilon_i}{\hat{\sigma}_{\varepsilon_i}}, \quad i = 1, 2, \dots, n$$

Նշանակենք՝

$$\frac{y_i}{\hat{\sigma}_{\varepsilon_i}} = \acute{y}_i, \quad \frac{1}{\hat{\sigma}_{\varepsilon_i}} = z_i, \quad \frac{x_{1,i}}{\hat{\sigma}_{\varepsilon_i}} = \acute{x}_{1,i} \dots \frac{x_{m,i}}{\hat{\sigma}_{\varepsilon_i}} = \acute{x}_{m,i}, \quad \frac{\varepsilon_i}{\hat{\sigma}_{\varepsilon_i}} = \acute{\varepsilon}_i:$$

Կունենանք՝

$$\acute{y}_i = \beta_0 z_i + \beta_1 \acute{x}_{1,i} + \beta_2 \acute{x}_{2,i} + \dots + \beta_m \acute{x}_{m,i} + \acute{\varepsilon}_i, \quad i = 1, 2, \dots, n \quad (83)$$

Մոդելը ձևափոխվեց առանց ազատ անդամի (83) տեսքի մոդելի, որն արդեն հոմոսկեդաստիկ է: Այս մոդելի ε -երի կովարիացիայի մատրիցի անկյունագծի վրա կլինեն 1-եր, իսկ դրանից դուրս՝ զրոներ, եթե ԳՄ այլ պայմանները տեղի ունեն: Այսինքն՝ այդ մատրիցը կստանա հետևյալ տեսքը՝

$$\Omega_{\varepsilon} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Յույց տանք, որ (83) մոդելում մնացորդների հոմոսկեդաստիկության պայմանն ապահովվում է՝

$$\begin{aligned} var(\acute{\varepsilon}_i) &= var\left(\frac{\hat{\varepsilon}_i}{\hat{\sigma}_{\varepsilon_i}}\right) = E\left[\frac{\hat{\varepsilon}_i}{\hat{\sigma}_{\varepsilon_i}} - E\left(\frac{\hat{\varepsilon}_i}{\hat{\sigma}_{\varepsilon_i}}\right)\right]^2 = E\left[\frac{\hat{\varepsilon}_i}{\hat{\sigma}_{\varepsilon_i}}\right]^2 = \frac{E(\hat{\varepsilon}_i^2)}{E(\hat{\sigma}_{\varepsilon_i}^2)} = \\ &= \frac{\hat{\sigma}_{\varepsilon_i}^2}{\hat{\sigma}_{\varepsilon_i}^2} = 1, \quad i = 1, 2, \dots, n: \end{aligned}$$

Պրակտիկայում $\hat{\sigma}_{\varepsilon_i}$ -երը անհայտ են լինում, ուստի այս մեթոդը կիրառելու համար մախ պետք է $\hat{\sigma}_{\varepsilon_i}$ -երը գնահատել: Մյուս կողմից, ստացված մոդելը չունի ազատ անդամ, որը որոշակիորեն խանգարում է արդյունքների վերլուծությանը: Այդ իսկ պատճառով ավելի հարմար է լինում հետերոսկեդաստիկության շտկման համար կիրառել **ՓՔ ընդհանրացված եղանակը**:

2. ՓՔ քնդհանրացված եղանակ

Քանի որ $\hat{\sigma}_{\varepsilon_i}$ -երը անհայտ են լինում, ենթադրություն է արվում, որ դրանք ուղիղ համեմատական են բացատրող փոփոխականներից որևէ մեկի արժեքներին: Այսինքն՝ ենթադրվում է, որ յուրաքանչյուր դիտարկման համար՝

$$\text{var}(\hat{\varepsilon}_i) = \hat{\sigma}_{\varepsilon_i}^2 = \delta \cdot x_{i,k}, \quad \delta = \text{const}:$$

Այդ դեպքում ռեգրեսիայի յուրաքանչյուր i -րդ դիտարկմանը համապատասխանող հավասարումը բաժանվում է k -րդ բացատրող փոփոխականի՝ այդ i -րդ դիտարկման արժեքի քառակուսի արմատի վրա: Եվ ստացվում է նոր հավասարում, որը պետք է գնահատվի: Այն է՝

$$\frac{y_i}{\sqrt{x_{i,k}}} = \frac{\beta_0}{\sqrt{x_{i,k}}} + \beta_1 \frac{x_{1,i}}{\sqrt{x_{i,k}}} + \beta_2 \frac{x_{2,i}}{\sqrt{x_{i,k}}} + \dots + \beta_m \frac{x_{m,i}}{\sqrt{x_{i,k}}} + \frac{\varepsilon_i}{\sqrt{x_{i,k}}},$$

$$i = 1, 2, \dots, n \quad (84):$$

Նշանակենք՝

$$\frac{y_i}{\sqrt{x_{i,k}}} = \tilde{y}_i, \quad \frac{1}{\sqrt{x_{i,k}}} = \tilde{z}_i, \quad \frac{x_{1,i}}{\sqrt{x_{i,k}}} = \tilde{x}_{1,i} \dots \frac{x_{m,i}}{\sqrt{x_{i,k}}} = \tilde{x}_{m,i}, \quad \frac{\varepsilon_i}{\sqrt{x_{i,k}}} = \tilde{\varepsilon}_i$$

Կունենանք՝

$$\tilde{y}_i = \beta_0 \tilde{z}_i + \beta_1 \tilde{x}_{1,i} + \beta_2 \tilde{x}_{2,i} + \dots + \beta_m \tilde{x}_{m,i} + \tilde{\varepsilon}_i, \quad i = 1, 2, \dots, n:$$

Կարելի է ցույց տալ, որ $\tilde{\varepsilon}_i$ -երը հոմոսկեդաստիկ են: Իրոք.

$$\text{var}(\tilde{\varepsilon}_i) = \text{var}\left(\frac{\varepsilon_i}{\sqrt{x_{i,k}}}\right) = \frac{1}{x_{i,k}} \text{var}(\varepsilon_i) = \frac{\delta \cdot x_{i,k}}{x_{i,k}} = \delta = \text{const}:$$

Եթե Գաուս-Մարկովի այլ պայմանները տեղի ունեն, (84) հավասարումը կարելի է գնահատել ՓՔԵ-ով:

4. Բազմակողմնաբարության բացակայության պայման. խախտման հետևանքներն ու հայտնաբերումը

4.1. Լիակատար կողմնաբարություն ու բազմակողմնաբարություն

Գաուս-Մարկովի պայմանները ԲԳՌ դեպքում պահանջում են, որ $Y = XB + \varepsilon$ տեսական մոդելում $X_1, \dots, X_m$ բացատրող փոփոխականները գծորեն անկախ լինեն, հետևաբար մոդելի X մատրիցի ռանգը հավասար լինի դրա սյուների քանակին: Այս պայմանը կոչվում է բազմակողմնաբարության բացակայության պայման, իսկ դրա խախտումը՝ բազմակողմնաբարություն:

Ենթադրենք՝ X մատրիցի սյուներից մեկը մյուսների (կամ մյուսների որոշ մասի) գծային կոմբինացիան է, դրա սյուների միջև (կամ որոշ մասի միջև) կա գծային ֆունկցիոնալ կախվածություն: Այդ դեպքում մոդելում առկա է **լիակատար կողմնաբարություն**:

Դիտարկենք երկու բացատրող փոփոխականներով ԲԳՌ մոդել՝

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i = 1, 2, \dots, n: \quad (85)$$

Ենթադրենք x_1 -ը և x_2 -ը գծային կախվածության մեջ են, և այդ կախվածության տեսքը հայտնի է.

$$x_{i2} = \gamma_0 + \gamma_1 x_{i1} : \quad (86)$$

Տեղադրենք (86)-ը (85)-ի մեջ: Կունենանք՝

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i = \beta_0 + \beta_1 x_{i1} + \beta_2 \gamma_0 + \beta_2 \gamma_1 x_{i1} + \varepsilon_i = \\ &= (\beta_0 + \beta_2 \gamma_0) + (\beta_1 + \beta_2 \gamma_1) x_{i1} + \varepsilon_i = \alpha_0 + \alpha_1 x_{i1} + \varepsilon_i \rightarrow \\ y_i &= \alpha_0 + \alpha_1 x_{i1} + \varepsilon_i, \end{aligned} \quad (87)$$

որում՝ $\alpha_0 = \beta_0 + \beta_2 \gamma_0$; $\alpha_1 = \beta_1 + \beta_2 \gamma_1$:

Եթե (87)-ը գնահատենք ՓՔԵ-ով, կստանանք $\widehat{\alpha}_0$ և $\widehat{\alpha}_1$ գործակիցները: Սկզբնական մոդելի $\widehat{\beta}_0$; $\widehat{\beta}_1$; $\widehat{\beta}_2$ գործակիցները գտնելու համար պետք է գնահատենք հետևյալ հավասարումների համակարգը՝

$$\begin{aligned} \widehat{\alpha}_0 &= \widehat{\beta}_0 + \widehat{\beta}_2 \gamma_0, \\ \widehat{\alpha}_1 &= \widehat{\beta}_1 + \widehat{\beta}_2 \gamma_1: \end{aligned}$$

Քանի որ x_1 -ի և x_2 -ի գծային կախվածության տեսքը, հետևաբար նաև γ_0 ; γ_1 գործակիցները հայտնի են, ապա նշված համակարգն

ունի երկու հավասարում ու երեք անհայտ՝ $\widehat{\beta}_0; \widehat{\beta}_1; \widehat{\beta}_2$: Այսպիսի համակարգն ունի անվերջ թվով լուծումներ, ուստի լիակատար կոլինեարության դեպքում դրանք միարժեքորեն որոշել հնարավոր չէ:

Ընդհանուր դեպքում, ինչպես հայտնի է, ՓԶԵ-ով ԲԳՌ գնահատված գործակիցներից կազմված վեկտոր-սյան տեսքը հատկապես է՝ $\widehat{B} = (X^T X)^{-1} X^T Y$: $\widehat{B}$ վեկտորը հնարավոր կլինի որոշել, եթե $X^T X$ մատրիցի որոշիչը հավասար չէ գրոյի, ինչը ենթադրում է, որ X մատրիցի սյուները գծորեն անկախ են, և դրա ռանգը հավասար է սյունների քանակին: Լիակատար կոլինեարության դեպքում $\text{rang}(X) < m+1$:

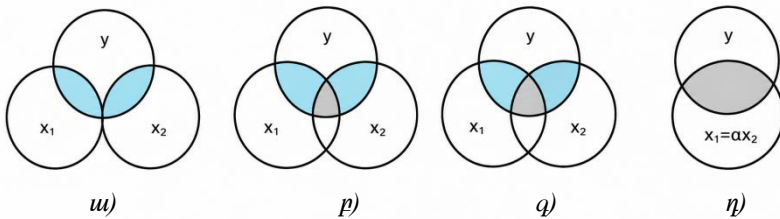
Այսպիսով, եթե մոդելում առկա է լիակատար կոլինեարություն, հնարավոր չէ այն գնահատել ՓԶԵ-ով: Պետք է մոդելից հեռացնել գծային կախվածության մեջ գտնվող բացատրող փոփոխականներից որևէ մեկը:

Գործնականում լիակատար կոլինեարությունը հազվադեպ է հանդիպում: Ավելի հաճախ հանդիպում է **բազմակոլինեարությունը: Բազմակոլինեարությունը մի իրավիճակ է, երբ բացատրող փոփոխականների միջև ուժեղ կոռելյացիոն կախվածությունը հանգեցնում է ռեգրեսիայի անվստահելի գնահատականների ստացման:** Բազմակոլինեարության դեպքում բացատրող փոփոխականների միջև գծային ֆունկցիոնալ կապ չկա, սակայն կա ուժեղ կոռելյացիոն կապ: Այդ պատճառով բազմակոլինեարությունը հաճախ անվանում են **ոչ լիակատար կոլինեարություն:** Իրականում բացատրող փոփոխականները միշտ էլ ինչ-որ չափով կոռելյացված են լինում: Սակայն Գաուս-Մարկովի պայմանների խախտում առաջանում է միայն դրանց շատ ուժեղ կոռելյացիոն կախվածության պարագայում:

Բազմակոլինեարությունը հնարավորություն չի տալիս առանձնացնելու տարբեր բացատրող փոփոխականների ազդեցությունը կախյալ փոփոխականի վրա և գնահատված գործակիցները դարձնում է անվստահելի: Յույց տանք բազմակոլինեարության հետևանքը Վենի գծապատկերների (դիագրամի) միջոցով: Գծանկար 13-ում (*ա*) x_1 -ի և x_2 -ի միջև չկա ոչ մի տեսակի կոլինեարություն, կոռելյացիոն կախվածություն: Դրանցից ամեն մեկի առանձին-առանձին ազդեցությունը y -ի վրա կարելի է հստակորեն որոշել, և β գործակիցները յույց կտան տվյալ x -ի մաքուր ազդեցությունը y -ի վրա: Դրանք գծանկարի վրա y -ի ու x_1 -ի, նաև y -ի և x_2 -ի հատման մասերն են: Եթե x_1 -ի և x_2 -ի միջև կա թույլ կոռելյացվածություն (*բ*), ապա դրանցից ամեն մեկի առանձին-առանձին ազդեցությունը y -ի վրա կարելի է համարյա հստակորեն որոշել: Դրանց ազդեցությունը y -ի վրա

առանձնացնել հնարավոր չէ միայն մոխրագույնով նշված փոքր հատման մասի համար: x_1 -ի և x_2 -ի միջև կոռելյացիայի աճին զուգընթաց դրանց ազդեցությունը y -ի վրա առանձնացնելն ավելի ու ավելի դժվար է դառնում, որովհետև մեծանում է դրանց համատեղ ազդեցության մասը (զ): Լիակատար կոլինեարության դեպքում x_1 -ի և x_2 -ի ազդեցությունը y -ի վրա առանձնացնելը դառնում է անհնար (դ):

Գծանկար 13. Վեճի գծապարկերները երկու բացատրող փոփոխականի համար



ՓԲԵ գնահատման արդյունքների վրա ոչ լիակատար կոլինեարության կամ բազմակոլինեարության հետևանքներն են.

Եթե Գաուս-Մարկովի մնացած պայմանները տեղի ունեն, ՓԲԵ-ով ստացված գնահատականները մնում են անշեղելի և արդյունավետ, BLUE գնահատականներ, սակայն՝

1. $\hat{\beta}$ գնահատականներն ունենում են մեծ ստանդարտ սխալներ, հետևաբար փոքր t վիճականիներ: Սա կարող է հանգեցնել նրան, որ բացատրող փոփոխականի ազդեցությունը կախյալի վրա գնահատվում է ոչ նշանակալի, չնայած իրականում դա այդպես չէ,
2. $\hat{\beta}$ գնահատականները ճշգրիտ ցույց չեն տալիս բացատրող փոփոխականի ազդեցությունը կախյալի վրա,
3. $\hat{\beta}$ գնահատականները ու դրանց ստանդարտ սխալները շատ զգալուն են լինում տվյալների նկատմամբ: Մի քանի դիտարկում ավելացնելու կամ պակասեցնելու դեպքում դրանք շատ մեծ փոփոխությունների են ենթարկվում,
4. Բազմակոլինեարության դեպքում հնարավոր է, որ F ու t վիճականիների թեստերը իրարամերժ արդյունքներ տան,
5. հնարավոր է $\hat{\beta}$ գնահատականների սխալ, տեսությանը հակասող նշանների ստացում:

4.2. Բազմակողմնաբարձր հայտնաբերման VIF թեստը

Ինչպե՞ս որոշել բազմակողմնաբարձրի առկայությունը:

Նախնական ախտորոշում: Եթե առկա են հետևյալ երևույթները.

1. մոդելի դետերմինացիայի գործակիցը բավականաչափ մեծ է, և F վիճականին փաստում է դրա վիճակագրական նշանակալիությունը, իսկ որոշ բացատրող փոփոխականների t վիճականները փաստում են համապատասխան β -երի ոչ արժեքականության մասին,
2. բացատրող փոփոխականները շատ ուժեղ կոռելյացիոն կախվածության մեջ են, $|r| > 0.80$, ապա կարելի է ենթադրել բազմակողմնաբարձրի առկայություն:

Վերջնական եզրակացությունը կարելի է անել վիճակագրական թեստի միջոցով, որը կոչվում է VIF (Variation inflation factor) թեստ:

VIF թեստով կարելի է որոշել, թե որքան է տվյալ բացատրող փոփոխականը կոռելացված մոդելում առկա մյուս բացատրող փոփոխականների հետ: Այդ նպատակով յուրաքանչյուր բացատրող փոփոխականի համար թեստը գնահատում է գծային ռեգրեսիայի մոդել, որտեղ այդ փոփոխականը կախված է մյուս բոլոր բացատրող փոփոխականներից:

Ենթադրենք Φ -ով գնահատվում է բազմագործոն գծային ռեգրեսիայի տեսական մոդելը.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j x_j + \dots + \beta_m x_m + \varepsilon:$$

VIF թեստն իրականացվում է յուրաքանչյուր բացատրող փոփոխականի համար: Յուրաքանչյուր j -րդ բացատրող փոփոխականի համար գնահատվում է օժանդակ ռեգրեսիայի հետևյալ տեսքի մոդել, որտեղ այն կախված է մյուս բոլոր բացատրող փոփոխականներից՝

$$x_j = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_{j-1} x_{j-1} + \alpha_{j+1} x_{j+1} \dots + \alpha_m x_m + \delta, \\ j = 1, \dots, m: \quad (88)$$

Յուրաքանչյուր օժանդակ ռեգրեսիայի գնահատման արդյունքում ստացվում է (88) մոդելի դետերմինացիայի գործակիցը, R_j^2 -ն, որի հիման վրա էլ հաշվարկվում է VIF վիճականին j -րդ բացատրող փոփոխականի համար՝

$$VIF_j = \frac{1}{1 - R_j^2}: \quad (89)$$

VIF վիճականին չափում է, թե որքանով է ռեգրեսիայի գործակցի վարիացիան ուռճացված բազմակոլինեարության պատճառով: Որքան մեծ է R_j^2 -ն, այսինքն՝ տվյալ բացատրող փոփոխականի կախվածությունը մնացած բացատրող փոփոխականներից ուժեղ է, այնքան մեծ է VIF վիճականին: Եվ հակառակը. եթե $R_j^2=0$, այսինքն՝ տվյալ բացատրող փոփոխականը մյուսների հետ կոռելացված չէ, $VIF=1$: Համարվում է, որ եթե $VIF>5$, ապա առկա է բազմակոլինեարություն: Սովորաբար օգտագործում են VIF վիճականու արժեքների՝ Աղյուսակ 8-ում ներկայացված դասակարգումը:

Աղյուսակ 8. VIF վիճականու արժեքների դասակարգումն ու բազմակոլինեարության առկայության մասին եզրակացությունները

VIF վիճականու արժեք	Մեկնաբանություն	Եզրակացություն
$VIF = 1$	Տվյալ բացատրող փոփոխականը մնացածների հետ կոռելացված չէ:	Չկա բազմակոլինեարություն:
$1 < VIF < 5$	Տվյալ բացատրող փոփոխականը մնացածների հետ բազմակոլինեարություն առաջացնող կոռելյացիա չունի:	Չկա բազմակոլինեարություն:
$5 \leq VIF < 10$	Տվյալ բացատրող փոփոխականը մնացածների հետ ուժեղ կոռելացված է:	Առկա է բազմակոլինեարություն, պետք է վերանայել մոդելի հատկորոշումը:
$VIF \geq 10$	Տվյալ բացատրող փոփոխականը մնացածների հետ շատ ուժեղ կոռելացված է:	Առկա է բազմակոլինեարություն, պետք է վերանայել մոդելի հատկորոշումը:

Բազմակոլինեարությունը վերացնելու նպատակով կարելի է միավորել կոռելացված բացատրող փոփոխականները (օրինակ՝ ինդեքսներ, համաթվեր ստեղծելու միջոցով), ինչպես մաս հեռացնել մնացածի հետ մեծ կոռելյացիա ունեցող բացատրող փոփոխականը մոդելի հատկորոշումից:

5. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալներ

Նախքան ցանկացած հավասարման գնահատումը այն պետք է ամբողջությամբ հատկորոշվի: Էկոնոմետրիկ հավասարման հատկորոշումը ենթադրում է բացատրող փոփոխականների ու կախվածության տեսքի ճիշտ ընտրություն:

Եթե այս ընտրություններից որևէ մեկը սխալ է արվել, ապա առաջանում է այսպես կոչված հատկորոշման սխալ, որը ՓՔԵ-ով գնահատված մոդելի գործակիցների վրա ունենում է որոշակի բացասական ազդեցություն:

Անդրադառնանք բացատրող փոփոխականների ընտրությանը: Դրանց ոչ ճիշտ ընտրությունը կարող է երկու տարբերակ ունենալ՝ կա՛մ կարևոր բացատրող փոփոխականը(ները) ներառված չէ(չեն) մոդելում, կա՛մ մոդելում ներառված են ավելորդ, կախյալ փոփոխականի վրա նշանակալի ազդեցություն չունեցող բացատրող փոփոխականներ: Դիտարկենք այս երկու դեպքերն առանձին:

5.1. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ. կարևոր բացատրող փոփոխականը ներառված չէ մոդելում

Դիտարկենք այն դեպքը, երբ գնահատվում է գծային բազմակի ռեգրեսիայի մոդել, և ռեգրեսիայի հավասարման մեջ հետազոտողը չի ներառել կախյալ փոփոխականի վրա կարևոր ազդեցություն ունեցող որևէ անկախ փոփոխական: Այդ դեպքում գնահատված հավասարման մեկնաբանությունն ու կիրառելիությունը խնդրահարույց են դառնում:

Դիցուք իրականում գծային ռեգրեսիայի տեսական մոդելը, որը հաշվի է առնում բոլոր կարևոր ազդեցությունները, ունի հետևյալ տեսքը՝

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon: \quad (90)$$

Այդ դեպքում ՓՔԵ-ով գնահատված էմպիրիկ հավասարումը կունենա հետևյալ տեսքը՝

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2: \quad (91)$$

Սակայն հետազոտողը ինչ-ինչ պատճառներով (տեսության կամ իրականության ոչ լիարժեք հաշվառում, տեղեկատվության բա-

ցակայություն և այլն) գնահատել է հետևյալ տեսքի էմպիրիկ մոդել՝ անտեսելով x_2 կարևոր փոփոխականի ազդեցությունը.

$$\hat{y} = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + v: \quad (92)$$

Կարևոր x_2 բացատրող փոփոխականի չներառելը (92) մոդելում հանգեցնում է նրան, որ դրա ազդեցությունը կախյալ փոփոխականի վրա ընդգրկվում է այդ մոդելի v պատահական մնացորդի մեջ՝ ըստ դրա սահմանման: Ելնելով նրանից, որ միևնույն կախյալ փոփոխականի վրա ազդող բացատրող փոփոխականները ինչ-որ չափով կոռելացված են լինում՝ խախտվում է Գաուս-Մարկովի այն պայմանը, որը պահանջում է, որ բացատրող փոփոխականները կոռելացված չլինեն պատահական շեղման հետ՝ $\text{COV}(x_1, v) \neq 0$: Դրա հետևանքով ՓԲԵ-ով գնահատված $\hat{\alpha}_1$ գործակիցը դառնում է շեղված՝ $E(\hat{\alpha}_1) \neq \beta_1$, և BLUE գնահատական չի լինում: Գնահատված սխալ (92) տեսքի $\hat{\alpha}_1$ գործակցի շեղումը ճիշտ տեսքի տեսական մոդելի β_1 պարամետրից համեմատական է x_1 և x_2 բացատրող փոփոխականների կոռելյացիային: $\hat{\alpha}_1$ գնահատված գործակցի ստանդարտ շեղումը ևս լինում է շեղված գնահատական, ինչի հետևանքով t վիճականին վստահելի արդյունք չի տալիս: Ուրեմն այսպիսի մոդելի գնահատման արդյունքները վստահելի չեն:

Տեսնելու համար, թե ինչպես կարող է բաց թողնված բացատրող փոփոխականն առաջացնել $\hat{\alpha}_1$ -ի շեղելիություն, դիտարկենք արտադրական ֆունկցիա, որում թողարկումը (Y) կախված է օգտագործված ռեսուրսներից՝ աշխատուժից (L) ու կապիտալից (K): Սա տեսությամբ թելադրվող մեր իրական տեսական մոդելն է՝

$$y = \beta_0 + \beta_1 L + \beta_2 K + \varepsilon:$$

Ի՞նչ կլինի, եթե կապիտալի վերաբերյալ տվյալները ինչ-ինչ պատճառներով անհասանելի լինեն, և հետագոտողը կապիտալը չներառի գնահատվող հավասարման մեջ: Այդ դեպքում կգնահատվի $y = \beta_0 + \beta_1 L + u$ մոդելը: Այստեղ կապիտալն ու աշխատուժը որոշակիորեն դրականորեն փոխկապակցված են (կապիտալի աճը սովորաբար պահանջում է այն օգտագործող աշխատուժի որոշակի աճ և հակառակը): Այդ պատճառով իրականում կապիտալի ազդեցությունը թողարկման վրա կմնա մոդելի մեջ և կարտահայտվի մոդելի պատահական շեղումներում, չնայած K -ն բացակայում է մոդելից: K -ի ազդեցությունը անուղղակիորեն կարտահայտվի աշխատուժի β_1 գործակցի միջոցով այնչափ, որչափ որ կապիտալն ու աշխատու-

ժը կոռելացված են, ինչը կհանգեցնի β_1 -ի շեղված գնահատականի ստացմանը:

Թվում է, թե լուծումը շատ հեշտ է՝ ավելացնել բացակայող փոփոխականը գնահատվող հավասարմանը: Բայց դա միշտ չէ, որ հնարավոր է անել: Նախ գտնելը, որ ունենք չներառված փոփոխական, բավական բարդ է: Սակայն այսպիսի հատկորոշման սխալը կարելի է հայտնաբերել պատահական մնացորդների՝ կախյալ փոփոխականից ունեցած կախման գրաֆիկի ուսումնասիրության միջոցով: Եթե այն ցույց տա որևէ օրինաչափություն, կարելի է կասկածել, որ դա չներառված փոփոխականի ազդեցությունն է:

5.2. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ. մոդելում ներառված է ավելորդ բացատրող փոփոխական

Ի՞նչ տեղի կունենա, եթե ռեգրեսիայի հավասարման հատկորոշումն ընդգրկի այնպիսի փոփոխականներ, որոնք իրականում ազդեցություն չունեն կախյալ փոփոխականի վրա:

Ենթադրենք իրականում տեսական մոդելն ունի հետևյալ հատկորոշումը՝

$$y = \beta_0 + \beta_1 x_1 + \varepsilon: \quad (93)$$

Սակայն հետազոտողը դիտարկել է հետևյալ հատկորոշմամբ տեսական մոդել՝

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + u: \quad (94)$$

Այստեղ x_2 -ը մի փոփոխական է, որն y -ի վրա էական ազդեցություն չունի: Այս դեպքում հետևանքները շատ ավելի թեթև են: Եթե Գաուս-Մարկովի մնացած պայմանները տեղի ունեն, ապա **$\widehat{\alpha}_1$ -ը անշեղ գնահատական է մնում**, սակայն ունենում է ավելի մեծ վարիացիա, քան այն $\widehat{\beta}_1$ -ը, որը կստացվեր, եթե գնահատեինք կախվածության ճիշտ տեսք: Արդյունքում (94)-ի գնահատման միջոցով ստացված **$\widehat{\alpha}_1$ -ը լինում է անշեղ, բայց ոչ արդյունավետ գնահատական**:

Գործնականում ավելորդ փոփոխականը ռեգրեսիայի հավասարման մեջ կարելի է փորձել գտնել, օրինակ, $\widehat{\alpha}_2$ -ի է վիճականու միջոցով, որը կհաստատի $\widehat{\alpha}_2$ -ի՝ գրոյից նշանակալիորեն չտարբերվելու վարկածը: Կարելի է ստուգել նաև F վարկածը՝ առանց ավելորդ փոփոխականի ու ավելորդ փոփոխականով մոդելների համար: Հիշենք, որ F վիճականին, փաստորեն, ցույց է տալիս, թե կախյալ փոփոխականի ընդհանուր վարիացիայի՝ ռեգրեսիայի գնահատված հավա-

սարմամբ բացատրվող մասը որքանով մեծ է ռեգրեսիայի գնահատված հավասարմամբ չբացատրվող մասից: Եթե F վիճականին (93) և (94) մոդելների համար էական տարբերություն չի դրսևորում, ուրեմն (94) մոդելի գնահատումն անիմաստ է, պետք է վերցնել (93) տեսքի հատկորոշումը: Նույնը կարելի է դիտարկել նաև երկու մոդելների դե-տերմինացիայի գործակիցների համեմատության միջոցով: Կա նաև այլ մեթոդ, որով կարելի է որոշել, թե արդյոք մոդելի մեջ մտցված է ավելորդ տեղեկատվություն՝ բացատրող փոփոխականների տեսքով: Այդ մեթոդը, այսպես կոչված, տեղեկատվական չափանիշների հաշվարկն է: Ամենատարածված տեղեկատվական չափանիշներից են Ակայկեի ու Շվարցի բայեսյան չափանիշները: Դրանք օգտագործվում են էկոնոմետրիկ մոդելավորման ամենատարբեր իրավիճակներում:

Ակայկեի տեղեկատվական չափանիշ (Akaike's Information Criterion (AIC))

Այս չափանիշը հաշվարկվում է հետևյալ բանաձևով՝

$$AIC = e^{2 \cdot (m+1)/n} \cdot \frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n} = e^{2 \cdot \frac{m+1}{n}} \cdot \frac{RSS}{n}, \quad (95)$$

որում m -ը բացատրող փոփոխականների քանակն է, $m+1$ -ը՝ գնահատվող պարամետրերի, իսկ n -ը՝ դիտարկումների: Հաճախ օգտագործում են այս չափանիշի բնական հիմքով լոգարիթմված տեսքը՝

$$\ln(AIC) = 2 \cdot \frac{m+1}{n} + \ln\left(\frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n}\right) = 2 \cdot \frac{m+1}{n} + \ln\left(\frac{RSS}{n}\right):$$

Երկու կամ մի քանի էկոնոմետրիկ մոդելներն իրար հետ համեմատելիս ընտրվում է այն մոդելը, որի համար այս չափանիշի արժեքը փոքրագույնն է:

Շվարցի բայեսյան տեղեկատվական չափանիշ (Schwarz's Bayesian Information Criterion (SIC կամ BIC))

$$SIC = n^{\frac{m+1}{n}} \cdot \frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n} = n^{\frac{m+1}{n}} \cdot \frac{RSS}{n}, \quad (96)$$

որում m -ը բացատրող փոփոխականների քանակն է, $m+1$ -ը՝ գնահատվող պարամետրերի, իսկ n -ը՝ դիտարկումների: Հաճախ օգտագործում են այս չափանիշի բնական հիմքով լոգարիթմված տեսքը՝

$$\ln(SIC) = \frac{m+1}{n} \cdot \ln(n) + \ln\left(\frac{\sum_{i=1}^n \hat{\varepsilon}_i^2}{n}\right) = \frac{m+1}{n} \cdot \ln(n) + \ln\left(\frac{RSS}{n}\right):$$

Երկու կամ մի քանի էկոնոմետրիկ մոդելներն իրար հետ համեմատելիս ընտրվում է այն մոդելը, որի համար այս չափանիշի արժեքը փոքրագույնն է:

**5.3. Գծային ռեգրեսիայի մոդելի հատկորոշման սխալ.
կախվածության սխալ տեսքի հատկորոշում. հատկորոշման սխալի
ստուգումը Ռ-ամսեյի RESET թեստի միջոցով**

Էկոնոմետրիկ մոդելավորման ընթացքում կախվածության տեսքի ճիշտ ընտրությունը կարևորագույն նշանակություն ունի: Երբ իրական տվյալների ստեղծման գործընթացը ոչ գծային է, բայց հետազոտողը սխալմամբ հատկորոշել և գնահատել է գծային մոդել, դա հանգեցնում է հատկորոշման սխալի կարևորագույն տեսակի առաջացման:

Ենթադրենք իրականում կախվածության տեսքն այսպիսին է՝ $y = f(x) + \varepsilon$, որում $f(x)$ -ը ոչ գծային ֆունկցիա է: Եթե հատկորոշվել է $y = XB + u$ տեսքի գծային կախվածություն, ապա այս մոդելի պատահական շեղումները կլանում են, իրենց մեջ են վերցնում իրական կախվածության ոչ գծային կառուցվածքը՝ $u = f(x) + \varepsilon - XB$ ՝ այսպիսով պարունակելով ոչ միայն պատահական, այլև բացատրող փոփոխականներից կախված բաղադրիչ: Ստացվում է, որ խախտվում է Գաուս-Մարկովի այն պայմանը, ըստ որի՝ պատահական շեղումներն ու բացատրող փոփոխականները պետք է անկախ լինեն՝ $COV(x, u) = 0$: Սրա հետևանքով կարող են առաջանալ նաև խաբուսիկ (ոչ թե տվյալների բնույթի հետ կապված, այլ կախվածության սխալ տեսքի գնահատումով պայմանավորված) ավտոկոռելյացիա ու հետերոսկեդաստիկություն:

Հատկորոշման այսպիսի սխալը հանգեցնում է նրան, որ ՓԲԵ-ով գնահատված գործակիցները լինում են **շեղված** ($E(\hat{\beta}) \neq \beta$), դրանք այլևս լավագույն, BLUE գնահատականներ չեն լինում:

Սխալ տեսքի կախվածության ընտրության նախնական ախտորոշման համար կարելի է գրաֆիկորեն ուսումնասիրել ՓԲԵ-ով գնահատված մոդելի պատահական մնացորդների կախվածությունը բոլոր բացատրող փոփոխականներից առանձին-առանձին կամ ջի-ից: Եթե երևան ակնհայտ կախվածություններ, ուրեմն կարելի է ենթադրել հատկորոշման սխալի առկայություն (տես Գծանկար 9):

Էկոնոմետրիկ մոդելի հատկորոշման սխալի առկայության մասին վերջնական եզրակացություն անելու համար կան վիճակագրա-

կան թեստեր, որոնցից ամենատարածվածը Ռամսեյի **RESET** (regression equation specification error test) թեստն է: Ենթադրենք հասկորոշվել է ԲԳՌ տեսական մոդել՝

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon:$$

RESET թեստում դրվում է գրոյական վարկած այն մասին, որ ռեգրեսիայի մոդելի հասկորոշումը կախվածության ճիշտ տեսքն է: Մրցակցող վարկածն այն է, որ իրականում ընտրված է սխալ հատկորոշում:

H_0 : ընտրված հասկորոշումը կախվածության ճիշտ տեսքն է,

H_1 : ընտրված հասկորոշումը կախվածության ճիշտ տեսքը չէ:

Թեստը գնահատում է ԲԳՌ մոդելը ՓԲԵ-ով և ստանում էմպիրիկ մոդելը.

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_m x_{im} + \hat{\varepsilon}_i:$$

Այս թեստի համատեքստում այն կանվանենք առանց սահմանափակման մոդել և կնշանակենք UR-ով (unrestricted): Որոշվում են գնահատված էմպիրիկ UR մոդելի պատահական $\hat{\varepsilon}_i$ մնացորդները և կախյալ փոփոխականի գնահատված $\hat{y}_i$ արժեքները: Այնուհետև գնահատվում է կախյալ փոփոխականի՝ բացատրող փոփոխականներից ու կախյալ փոփոխականի գնահատված արժեքներից կախվածության հետևյալ տեսքի օժանդակ ռեգրեսիա.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \alpha_1 \hat{y}_i^2 + \alpha_1 \hat{y}_i^3 + \dots + \alpha_1 \hat{y}_i^k + u_i, \quad (97)$$

որում k -ն ընտրվում է հետագոտողի կողմից: Օժանդակ ռեգրեսիայի մոդելը կանվանենք սահմանափակումով մոդել և կնշանակենք R-ով (restricted): Թեստի վիճականին F վիճականին է, որը որոշվում է հետևյալ բանաձևով.

$$F_{\eta\text{խ}} = \frac{(\sum_{i=1}^n \hat{u}_i - \sum_{i=1}^n \hat{\varepsilon}_i)/k}{\sum_{i=1}^n \hat{\varepsilon}_i / (n - (m + k + 1))} = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR} / (n - (m + k + 1))}, \quad (98)$$

որում RSS_R -ը օժանդակ ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդների քառակուսիների գումարն է, RSS_{UR} -ը՝ սկզբնական, առանց սահմանափակման ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդների քառակուսիների գումարը, k -ն՝ օժանդակ ռեգրեսիայում սահմանափակումների կամ կախյալ փոփոխականի գնահատված արժեքների աստիճանների քանակը, $m + k$ -ն՝ օժանդակ

ռեգրեսիայի բացատրող փոփոխականների թիվը, իսկ $m + k + 1$ -ը՝ օժանդակ ռեգրեսիայի գնահատվող պարամետրերի թիվը:

Չրոյական վարկածը ստուգվում է (98)-ով հաշվարկված $F_{\eta\text{իտ}}\text{-ը}$ Ֆիշերի բաշխում ունեցող պատահական մեծության $F_{\text{կրիտ}}$ կրիտիկական արժեքի հետ համեմատելու միջոցով: $F_{\text{կրիտ}}$ -ը բնութագրվում է α արժեքականության մակարդակով ու $(m + k - 1; n - (m + k + 1))$ ազատության աստիճաններով:

Եթե՝

$F_{\eta\text{իտ}} > F_{\text{կրիտ}} \longrightarrow H_0\text{-ն մերժվում է տվյալ } \alpha \text{ արժեքականության մակարդակում, ռեգրեսիայի հավասարման հատկորոշումը սխալ է ընտրված:}$

$F_{\eta\text{իտ}} \leq F_{\text{կրիտ}} \longrightarrow \alpha \text{ արժեքականության մակարդակում } H_0\text{-ն չի կարող մերժվել, ընդունվում է: Ռեգրեսիայի հավասարման հատկորոշումը ճիշտ է ընտրված:}$

Թեստի ելքը որոշելու համար կարելի է մաս օգտագործել p արժեքը: Եթե $p\text{-value} < \alpha$, ապա H_0 -ն մերժվում է տվյալ α արժեքականության մակարդակում: Հակառակ դեպքում H_0 -ն չի կարող մերժվել, ընդունվում է:

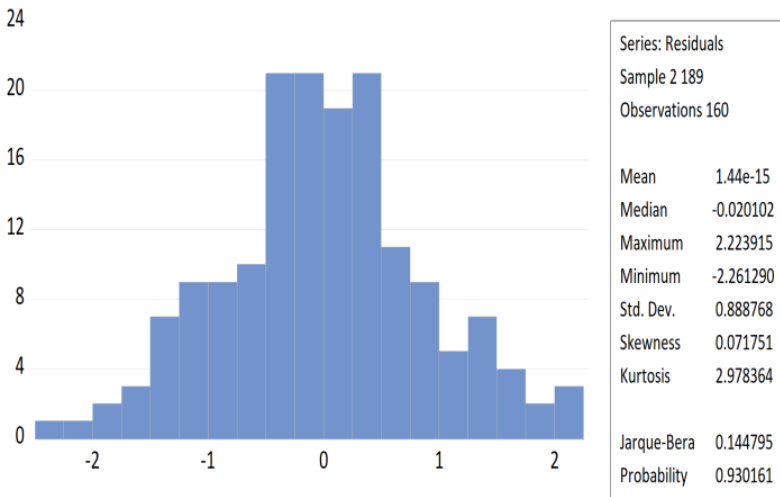
6. Գաուս-Մարկովի պայմանների ստուգումը «EViews» փաթեթում ու արդյունքների վերլուծությունը

Վերադառնանք Գլուխ 2-ի կետ 6-ում բերված օրինակին: Ուսումնասիրվում էր, թե երկրի մեկ շնչի հաշվով ՀՆԱ-ի մեծության վրա ինչ ազդեցություն ունեն երկրի բնակչության կրթամակարդակը, աշխատանքի շուկայում նրանց մասնակցությունը ու ուրբանիզացիայի մակարդակը: Հատկորոշվել էր գծային բազմակի ռեգրեսիայի էկոնոմետրիկ մոդել, և ստացվել էին գնահատված մոդելի բավարար որակ և գնահատված գործակիցների ու հավասարման վիճակագրական նշանակալիություն: Մոդելի օտագործելիության մասին վերջնական եզրակացություն անելու համար պետք է ստուգել Գաուս-Մարկովի պայմանները: Դրանց ստուգման արդյունքները ներկայացնենք «EViews» փաթեթի ելքային աղյուսակների միջոցով:

1. Ժարկ-Բերրայի՝ նորմալ բաշխման թեստը

Թեստի զրոյական վարկածը հետևյալն է. ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդները նմուշ են տեսական մոդելի՝ նորմալ բաշխում ունեցող պատահական շեղումներից: Թեստը ստուգվում է տեսական մոդելի պատահական շեղումների վիճակագրական գնահատականների՝ էմպիրիկ մոդելի պատահական մնացորդների միջոցով:

Քանի որ էմպիրիկ մոդելը գնահատված է, կարելի է կառուցել պատահական մնացորդների հաճախությունների բաշխման հիստոգրամը, ինչպես նաև հաշվել թեստի վիճականին:



Ըստ ստացված արդյունքների՝ $JB = 0.144795$: $\alpha=0.05$ արժեքականության մակարդակի համար $\chi_{\text{կր}}^2 = 5.99$ (տես Գլուխ 3, կետ 1, Աղյուսակ 6): Հետևաբար $JB = 0.144795 < 5.99 = \chi_{\text{կր}}^2$: Եզրակացությունը հետևյալն է. H_0 -ն չի կարող մերժվել, պատահական շեղումները նորմալ են բաշխված: Թեստի այդ ելքը հաստատում է նաև p արժեքը, որը զրոյական վարկածը մերժելիս սխալվելու կամ, որ նույնն է, զրոյական վարկածի հավանականությունն է: $p\text{-value} = \text{Probability} = 0.930161$: Քանի որ $p\text{-value} > \alpha = 0.05$, պատահական շեղումների նորմալ բաշխման մասին զրոյական վարկածի տեղի ունենալու հավանականությունը մեծ է, և այն պետք է ընդունել:

2. Դարբին-Ուոթսոնի՝ առաջին կարգի ավտոկոռելյացիայի առկայության ստուգման թեստը

Թեստի գրոյական վարկածը հետևյալն է. տեսական մոդելի պատահական շեղումների միջև չկա առաջին կարգի ավտոկոռելյացիա: Փաթեթի հաշվարկներով՝ $DW=2.06$: Դարբին-Ուոթսոնի վիճականու կրիտիկական միջակայքի ստորին ու վերին սահմանները $n=160$ $m=3$ համար կարող ենք որոշել աղյուսակից՝ նշված QR կոդի հղումով: Ըստ այդ աղյուսակի՝ $d_l = 1.74$ և $d_u = 1.8$:

Հետևաբար $4 - d_l = 2.26$, $4 - d_u = 2.2$ և $d_u \leq 2.06 < 4 - d_u$:

Ըստ Աղյուսակ 7-ի (Գլուխ 3, կետ 2.1)՝ եզրակացությունն այն է, որ մոդելում առկա չէ առաջին կարգի ավտոկոռելյացիա:



3. Բրեուշ-Գոդֆրիի՝ բարձր կարգերի ավտոկոռելյացիայի առկայության ստուգման LM թեստը

Թեստը ստուգվել է մինչև 3-րդ կարգի ավտոկոռելյացիայի համար: Ջրոյական վարկածն է՝ տեսական մոդելի պատահական շեղումների միջև չկա մինչև 3-րդ կարգի ավտոկոռելյացիա: Թեստի վիճականին՝ $Obs \cdot R-squared=1.196396$: χ^2 բաշխման կրիտիկական արժեքների աղյուսակը հասանելի է նշված QR կոդի հղումով: $p=3$ ազատության աստիճանի և $\alpha=0.05$ արժեքականության մակարդակի համար $\chi_{\text{կր}}^2 = 7.815$: Ստացանք՝ $Obs \cdot R-squared=1.196396 < 7.815$, հետևաբար գրոյական վարկածը չենք կարող մերժել, այն ընդունվում է և տեսական մոդելի պատահական շեղումների միջև չկա մինչև 3-րդ կարգի ավտոկոռելյացիա:



Breusch-Godfrey Serial Correlation LM Test:

Null hypothesis: No serial correlation at up to 3 lags

F-statistic	0.387104	Prob. F(3,159)	0.7624
Obs*R-squared	1.196396	Prob. Chi-Square(3)	0.7539

Թեստի այդ ելքը հաստատում է մաս p արժեքը, որը գրոյական վարկածը մերժելիս սխալվելու կամ, որ նույնն է, գրոյական վարկա-

ծի հավանականությունն է: $p\text{-value} = \text{Prob. Chi-Square} = 0.7539$: Քանի որ $p\text{-value} > \alpha = 0.05$, մինչև 3-րդ կարգի ավտոկոռելյացիա չլինելու մասին գրոյական վարկածի տեղի ունենալու հավանականությունը մեծ է, և այն պետք է ընդունել:

4. Հետերոսկեդաստիկության Ուայթի թեստը

Ուայթի թեստի գրոյական վարկածը հետևյալն է. տեսական մոդելի պատահական շեղումների վարիացիան հաստատուն է, այսինքն՝ դրանք հոմոսկեդաստիկ են: Թեստի վիճականին՝ $\text{Obs} \cdot \text{R-squared} = 18.70956$: Թեստի օժանդակ ռեգրեսիայի բացատրող փոփոխականների քանակն հավասար է 6-ի (տվյալ դեպքում դրանք 3 բացատրող փոփոխականներն են և դրանց քառակուսիները: Դրանց արտադրյալները չեն մտցվել օժանդակ ռեգրեսիայի մեջ՝ մոդելի ազատության աստիճանը չփոքրացնելու համար): χ^2 բաշխման կրիտիկական արժեքների աղյուսակը հասանելի է վերը նշված QR կոդի հղումով: $p = 6$ ազատության աստիճանի և $\alpha = 0.05$ արժեքականության մակարդակի համար $\chi_{\text{կր}}^2 = 12.592$: Ստացանք $\text{Obs} \cdot \text{R-squared} = 18.70956 > 12.592$, հետևաբար գրոյական վարկածը մերժվում է, և տեսական մոդելի պատահական շեղումները հոմոսկեդաստիկ չեն: Մոդելում առկա է հետերոսկեդաստիկություն, ուստի այս տեսքով մոդելն անօգտագործելի է:

Heteroskedasticity Test: White

Null hypothesis: Homoskedasticity

F-statistic	4.067006	Prob. F(6,159)	0.0017
Obs*R-squared	18.70956	Prob. Chi-Square(6)	0.0022
Scaled explained SS	17.84029	Prob. Chi-Square(6)	0.0032

Թեստի այդ ելքը հաստատում է նաև p արժեքը, որը գրոյական վարկածը մերժելիս սխալվելու կամ, որ նույնն է, գրոյական վարկածի հավանականությունն է: $p\text{-value} = \text{Prob. Chi-Square} = 0.0022$: Քանի որ $p\text{-value} < \alpha = 0.05$, պատահական շեղումների հոմոսկեդաստիկության գրոյական վարկածի տեղի ունենալու հավանականությունը փոքր է, և այն պետք է մերժել:

5. Բրեսուշ-Պագանի՝ հետերոսկեդաստիկության թեստը

Այս թեստի գրոյական վարկածը ևս տեսական մոդելի պատահական շեղումների վարիացիայի հաստատունության, այսինքն՝ հոմոսկեդաստիկության վարկածն է: Թեստի վիճականի՝ $\text{Obs}^*\text{R-squared}=15.87419$: Թեստի օժանդակ ռեգրեսիայի բացատրող փոփոխականների քանակն հավասար է 3-ի: χ^2 բաշխման կրիտիկական արժեքների աղյուսակը հասանելի է վերը նշված QR կոդի հղումով: $p=3$ ազատության աստիճանի և $\alpha=0.05$ արժեքականության մակարդակի համար $\chi_{\text{կր}}^2 = 7.815$: Ստացանք՝ $\text{Obs}^*\text{R-squared}=15.87419 > 7.815$, հետևաբար գրոյական վարկածը մերժվում է, և տեսական մոդելի պատահական շեղումները հոմոսկեդաստիկ չեն:

Heteroskedasticity Test: Breusch-Pagan-Godfrey

Null hypothesis: Homoskedasticity

F-statistic	8.622315	Prob. F(3,162)	0.0003
Obs*R-squared	15.87419	Prob. Chi-Square(3)	0.0004
Scaled explained SS	15.13666	Prob. Chi-Square(3)	0.0005

Թեստի այդ ելքը հաստատում է նաև p արժեքը, որը գրոյական վարկածի հավանականությունն է: $p\text{-value}=\text{Prob. Chi-Square}=0.0004$: Քանի որ $p\text{-value} < \alpha=0.05$, պատահական շեղումների հոմոսկեդաստիկության գրոյական վարկածի տեղի ունենալու հավանականությունը փոքր է, և այն պետք է մերժել:

6. Բազմակոլինեարության հայտնաբերման VIF թեստը

VIF թեստն իրականացվում է յուրաքանչյուր բացատրող փոփոխականի համար: Յուրաքանչյուր բացատրող փոփոխականի համար գնահատվում է օժանդակ ռեգրեսիայի մոդել, որում այն կախված է մյուս բոլոր բացատրող փոփոխականներից և հաշվարկվում է այդ օժանդակ ռեգրեսիայի դետերմինացիայի գործակիցը, որի հիման վրա (89) բանաձևով հաշվարկվում է VIF վիճականին: Ստացված արդյունքների ու Աղյուսակ 8-ի (Գլուխ 3, կետ 4.2) հիման վրա կարող ենք եզրակացնել, որ երեք բացատրող փոփոխականներն էլ մնացածների հետ բազմակոլինեարություն առաջացնող կոռելյացիա չունեն, քանի որ դրանց բոլորի համար էլ $1 < \text{VIF} < 5$: Ուստի մոդելում բազմակոլինեարություն չկա:

Variance Inflation Factors
Included observations: 165

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
PARTICIPATION	3.98E-05	38.38773	1.083437
URBAN	1.04E-05	9.060343	1.233437
EDUC	2.03E-05	8.160343	1.310893
C	0.177897	36.70694	NA

7. Գծային հատկորոշման սխալի ստուգման Ռ-ամսեյի RESET քեստ

Այս քեստը դնում է գրոյական վարկած, որ գնահատված մոդելի հատկորոշումը կախվածության ճիշտ տեսքն է: Օժանդակ ռեգրեսիայի մեջ մտցվել են կախյալ փոփոխականի գնահատված արժեքների քառակուսիներն ու խորանարդները: Հետևաբար $k=2$, իսկ օժանդակ ռեգրեսիայի բացատրող փոփոխականների թիվն է $m+k=3+2=5$: Թեստի վիճականին F վիճականին է, որը որոշվում է (98) բանաձևով.

$$F_{\eta\text{իտ}} = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n - (m + k + 1))} = \frac{(129.5449 - 113.3904)/2}{113.3904/(160 - (3 + 2 + 1))} = \frac{8.077}{0.736} = 10.97:$$

$(m + k - 1; n - (m + k + 1)) = (4; 154)$
ազատության աստիճաններով Ֆիշերի բաշխման կրիտիկական արժեքը կարելի է որոշել աղյուսակից, որը հասանելի է նշված QR կոդի հղումով: Աղյուսակից գտնում ենք, որ $\alpha=0.05$ արժեքականության մակարդակի համար $F_{կրիտ}=5.63$: Ստացվեց, որ $F_{\eta\text{իտ}} = 10.97 > 5.63 = F_{կրիտ}$: Հետևաբար գրոյական վարկածը պետք է մերժել, և գնահատված մոդելի հատկորոշումը ճիշտ չէ:

Դրա մասին է վկայում նաև p արժեքը, որը գրոյական վարկածի հավանականությունն է: $p\text{-value}=\text{Probabilty F-statistic}=0.000$: Զանի որ $p\text{-value}<\alpha=0.05$, կախվածության գծային տեսքի գրոյական վարկածի տեղի ունենալու հավանականությունը փոքր է, և այն պետք է մերժել:



	Value	df	Probability
F-statistic	10.97418	(2, 154)	0.0000
Likelihood ratio	21.97644	2	0.0000

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	16.15447	2	8.077235
Restricted SSR	129.5449	157	0.799660
Unrestricted SSR	113.3904	154	0.708690

Այսպիսով՝ գնահատված մոդելում առկա է պատահական շեղումների հետերոսկեդաստիկություն ու գծային հատկորոշման սխալ: Մոդելը կիրառելի չէ: Հետերոսկեդաստիկությունը շտկելու համար կարելի է օգտագործել փոքրագույն քառակուսիների ընդհանրացված եղանակը (տե՛ս Գլուխ 3, կետ 3.3):

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Ի՞նչ հետևանքներ է ունենում ՓԶԵ-ով գնահատված ռեգրեսիայի մոդելի արդյունքների վրա Գաուս-Մարկովի այն պայմանի խախտումը, որը պահանջում է պատահական շեղումների նորմալ բաշխվածություն:
2. Նկարագրե՛ք Ժարկ-Բերրայի թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
3. Տարածական էմպիրիկ տվյալների պարագայում հիմնականում ինչո՞ւ է պայմանավորված լինում պատահական շեղումների ավտոկոռելացվածության բացակայության պայմանի խախտումը:
4. Ի՞նչ հետևանքներ է առաջացնում պատահական շեղումների ավտոկոռելյացիայի բացակայության պայմանի խախտումը ՓԶԵ-ով գնահատված գործակիցների հատկությունների վրա:
5. Ի՞նչ է պատահական շեղումների առաջին կարգի ավտոկոռելյացիան, ի՞նչ բանաձևով է որոշվում առաջին կարգի ավտոկոռելյացիայի գործակիցը:

6. Ինչպե՞ս է որոշվում և ի՞նչ միջակայքում է ընկած Դարբին-Ուոթսոնի վիճականին, ինչպիսի՞ն են դրա արժեքները՝ կախված առաջին կարգի ավտոկոռելյացիայի գործակցի արժեքներից:
7. Նկարագրե՛ք Դարբին-Ուոթսոնի թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
8. Նկարագրե՛ք Բրեուշ-Գոդֆրիի Լագրանժի բազմապատկչի թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
9. Ինչպե՞ս կարելի է գրաֆիկորեն ուսումնասիրել հետերոսկեդաստիկության առկայությունը:
10. Նկարագրե՛ք Ուայթի թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
11. Նկարագրե՛ք Բրեուշ-Պագանի թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
12. Հետերոսկեդաստիկության շտկման ի՞նչ մեթոդներ կան: Նկարագրե՛ք դրանք և ներկայացրե՛ք, թե ձեր կարծիքով դրանցից որի կիրառումն է ավելի իրատեսական:
13. Ի՞նչ է լիակատար կոլինեարությունը: Հիմնավորե՛ք, թե հնարավոր է արդյոք կիրառել ՓԶԵ-ն ու ստանալ ռեգրեսիայի գնահատված գործակիցները լիակատար կոլինեարության պայմաններում:
14. Ինչո՞վ է տարբերվում բազմակոլինեարությունը լիակատար կոլինեարությունից, և ո՞րն է դրա էությունը:
15. Որո՞նք են բազմակոլինեարության հետևանքները:
16. Ինչպե՞ս որոշել բազմակոլինեարության առկայությունը: Նկարագրե՛ք VIF թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:
17. Էկոնոմետրիկ մոդելի հատկորոշման ինչպիսի՞ սխալներ կարող են լինել:
18. Ի՞նչ հետևանքներ է առաջացնում կախյալ փոփոխականի վրա կարևոր ազդեցություն ունեցող բացատրող փոփոխականը հավասարման մեջ չներառելու հատկորոշման սխալը:
19. Ի՞նչ հետևանքներ է առաջացնում ավելորդ բացատրող փոփոխականը հավասարման մեջ ներառելու հատկորոշման սխալը:
20. Ինչպե՞ս կարելի է ստուգել էկոնոմետրիկ մոդելում ավելորդ բացատրող փոփոխականի առկայությունը:

21. Ինչպիսի՞ տեղեկատվական չափանիշներ գիտեք, ինչպե՞ս են դրանք որոշվում, և դրանց միջոցով ինչպե՞ս կարելի է որոշել լավագույն մոդելը:
22. Ի՞նչ հետևանքներ է ունենում կախվածության սխալ տեսքի գնահատումը ՓՋԵ-ով գնահատված մոդելի պատահական մնացորդների ու գնահատված գործակիցների հատկությունների վրա:
23. Նկարագրե՛ք Ռամսեյի RESET թեստը՝ դրվող վարկածը, վիճականին, ստուգման մեթոդն ու եզրակացությունները:

Խնդիրներ

1. 170 երկրների տվյալների հիման վրա ՓՋԵ-ով գնահատվել է գծային ռեգրեսիայի մոդել, որում դրանց մարդկային զարգացման համաթիվը կախված է կրթության վրա կատարվող պետական ծախսերից (տոկոս՝ ՀՆԱ-ի նկատմամբ), առողջապահության վրա կատարվող պետական ծախսերից (տոկոս՝ ՀՆԱ-ի նկատմամբ) ու եկամուտների բաշխման անհավասարության Ջինիի գործակցից: Ստացվել են ռեգրեսիայի էմպիրիկ հավասարման պատահական մնացորդները, և վիճակագրական վերլուծությամբ որոշվել են դրանց ասիմետրիայի գործակիցը՝ $S = \frac{1}{n} \sum_{i=1}^n \frac{(\varepsilon_i - \bar{\varepsilon})^3}{S_{\varepsilon}} = -0.018$ և կուտակվածությունը՝ $K = \frac{1}{n} \sum_{i=1}^n \frac{(\varepsilon_i - \bar{\varepsilon})^4}{S_{\varepsilon}} = 3.359$: Օգտագործելով χ^2 բաշխման կրիտիկական արժեքը $\alpha=0.05$ արժեքականության մակարդակի համար՝ կիրառե՛ք պատահական շեղումների նորմալ բաշխման Ժարկ-Բերայի թեստը և կատարե՛ք եզրակացություն:
2. Գնահատված է գծային ռեգրեսիայի մոդել, որը բնութագրում է աշխատավարձի (Y) կախումը սովորած տարիների քանակից (X)՝ $Y = 75000 + 5000X + \varepsilon$: Հաշվարկվել են՝ $DW = 1.83$; $dl=1.07$; $du=1.60$: Կիրառե՛ք Դարբին-Ուոթսոնի թեստը և պարզե՛ք, թե արդյոք տեսական մոդելի պատահական շեղումների միջև կա՞ 1-ին կարգի ավտոկոռելյացիա. եթե այո, ապա դրակա՞ն, թե՞ բացասական: Հիմնավորե՛ք պատասխանը:
3. 200 էմպիրիկ տվյալների հիման վրա գնահատված է գծային ռեգրեսիայի մոդել, որը բնութագրում է վաճառքի ծավալների կախումը ծախսված ռեսուրսներից՝ հիմնական միջոցներից ու աշխա-

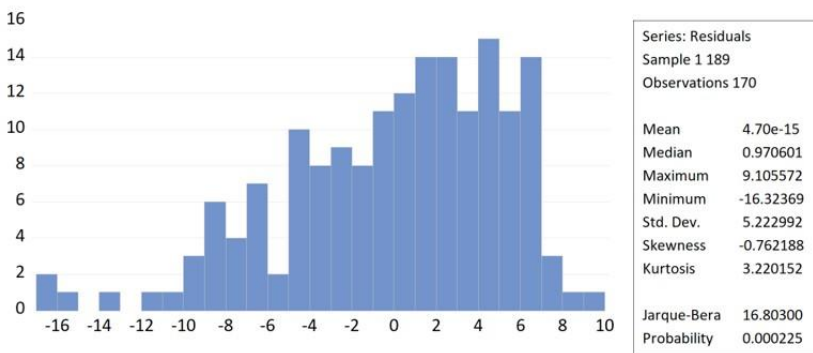
տուժի օգտագործումից, ինչպես նաև գովազդային ծախսերից: Հաշվարկվել է էմպիրիկ մոդելի պատահական մնացորդների առաջին կարգի ավտոկոռելյացիայի գործակիցը՝ $\hat{\rho}(1) = -0.457$: Հաշվեք Դարբին-Ուոթսոնի վիճականին և ստուգեք տեսական մոդելի պատահական շեղումների 1-ին կարգի ավտոկոռելյացիայի գոյությունը՝ օգտվելով Դարբին-Ուոթսոնի վիճականու կրիտիկական արժեքների աղյուսակից: Կատարեք եզրակացություն:

4. 170 երկրների տվյալների հիման վրա գնահատվել է գնաճի (ինֆլյացիայի) կախումը գործազրկության մակարդակից բնութագրող գծային ռեգրեսիայի մոդել: Կիրառվել է տեսական մոդելի պատահական մնացորդների մինչև 3-րդ կարգի ավտոկոռելյացիաների գոյության ստուգման Բրեուշ-Գոդֆրիի Լագրանժի բազմապատկչի թեստը: Թեստի օժանդակ ռեգրեսիայի դետերմինացիայի գործակիցը՝ $R^2=0.035$: Օգտվելով χ^2 բաշխման կրիտիկական արժեքների աղյուսակից $\alpha=0.05$ արժեքականության մակարդակի համար՝ ստուգեք մինչև 3-րդ կարգի ավտոկոռելյացիաների գոյության մասին վարկածը և կատարեք եզրակացություն:
5. 169 տնային տնտեսությունների տվյալների հիման վրա գնահատվել է դրանց սպառողական ծախսերի կախումը եկամուտներից ու տնային տնտեսության անդամների թվից բնութագրող գծային ռեգրեսիայի էմպիրիկ մոդել: Կիրառվել է Ուայթի՝ հետերոսկեդաստիկության թեստը և ստացվել է թեստի օժանդակ ռեգրեսիայի դետերմինացիայի գործակիցը՝ $R^2=0.1057$, օժանդակ ռեգրեսիայի մոդելում 5 բացատրող փոփոխականների պարագայում: Օգտվելով χ^2 բաշխման կրիտիկական արժեքների աղյուսակից $\alpha=0.05$ արժեքականության մակարդակի համար՝ ստուգեք տեսական մոդելի պատահական շեղումների հոմոսկեդաստիկության վարկածը և կատարեք եզրակացություն:
6. 500 տնային տնտեսությունների տվյալների հիման վրա գնահատվել է դրանց էլեկտրաէներգիայի սպառման եկամուտներից ու բնակարանի մակերեսից կախվածությունը բնութագրող գծային ռեգրեսիայի էմպիրիկ մոդել: Կիրառվել է հետերոսկեդաստիկության Բրեուշ-Պագանի թեստը, և ստացվել է թեստի օժանդակ ռեգրեսիայի դետերմինացիայի գործակիցը՝ $R^2=0.097$, օժանդակ ռեգրեսիայի մոդելում 2 բացատրող փոփոխականների պարագայում: Օգտվելով χ^2 բաշխման կրիտիկական արժեքների աղյուսակից $\alpha=0.05$ արժեքականության մակարդակի համար՝

ստուգե՞ք տեսական մոդելի պատահական շեղումների հոմոսկեդաստիկության վարկածը և կատարե՞ք եզրակացություն:

7. 170 երկրների տվյալների հիման վրա գնահատվել է մանկամահացության կախումը մեկ շնչի հաշվով ՀՆԱ-ից բնութագրող գծային ռեգրեսիայի էմպիրիկ մոդել: Կիրառվել է հատկորոշման սխալի ստուգման Ռ-ամսեյի RESET թեստը, որի օժանդակ ռեգրեսիայի մեջ մտցվել են կախյալ փոփոխականի գնահատված արժեքի քառակուսիներն ու խորանարդները: Ինչի՞նչ են հավասար k -ն, $m + k$ -ն և $n - (m + k + 1)$ -ը: Հաշվարկվել է օժանդակ ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդների քառակուսիների գումարը՝ $ESS_R = 78540.86$, առանց սահմանափակման ռեգրեսիայի էմպիրիկ մոդելի պատահական մնացորդների քառակուսիների գումարը՝ $ESS_{UR} = 55745.41$: Հաշվե՞ք թեստի $F_{\eta_{\text{իտ}}}$ վիճականին և համեմատելով այն Ֆիշերի բաշխում ունեցող պատահական մեծության $F_{\eta_{\text{իտ}}}$ կրիտիկական արժեքի հետ $\alpha=0.05$ արժեքականության մակարդակի համար՝ եզրակացություն կատարե՞ք գծային հատկորոշման տեսքի ճիշտ լինելու մասին:

8. «EViews» փաթեթով գնահատվել է գծային բազմակի ռեգրեսիայի մոդել, և գեներացվել է դրա պատահական մնացորդների Ժարկ-Բերրայի՝ նորմալ բաշխման թեստը: Մեկնաբանե՞ք ստացված արդյունքները և կատարե՞ք եզրակացություն մնացորդների նորմալ բաշխման մասին:



9. «EViews» փաթեթով գնահատվել է գծային պարզ ռեգրեսիայի մոդել, և ստացվել է հետևյալ ելքային աղյուսակը: Գրեք գնահատված հավասարումը, մեկնաբանեք գնահատված գործակիցների իմաստը, եզրակացություն կատարեք մոդելի որակի, գնահատված գործակիցների ու հավասարման վիճակագրական նշանակալիության մասին: Դիտարկեք պատահական մնացորդների առաջին կարգի ավտոկոռելյացիայի թեստը: Օգտվելով Դարբին-Ուոթսոնի վիճականու արժեքից ու դրա կրիտիկական միջակայքի սահմաններից (որոնք $\alpha=0.05$ արժեքականության մակարդակի համար կստանաք համապատասխան աղյուսակից)՝ եզրակացություն կատարեք պատահական մնացորդների առաջին կարգի ավտոկոռելյացիայի գոյության վերաբերյալ:

Dependent Variable: Y

Included observations: 170

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.000243	2.09E-05	11.61870	0.0000
C	70.20951	0.495202	141.7796	0.0000
R-squared	0.575534	Mean dependent var		73.57294
Adjusted R-squared	0.572234	S.D. dependent var		7.014261
S.E. of regression	5.238514	Akaike info criterion		6.161648
Sum squared resid	4610.261	Schwarz criterion		6.198539
Log likelihood	-521.7401	Hannan-Quinn criter.		6.176618
F-statistic	134.9942	Durbin-Watson stat		2.126580
Prob(F-statistic)	0.000000			

10. «EViews» փաթեթով գնահատվել է գծային բազմակի ռեգրեսիայի մոդել, և գեներացվել է դրա պատահական մնացորդների՝ մինչև 3-րդ կարգի ավտոկոռելյացիայի ստուգման Բրեուշ-Գոֆֆրիի Լագրանժի բազմապատկչի թեստը: Մեկնաբանեք թեստի ելքային աղյուսակը և եզրակացություն կատարեք պատահական մնացորդների՝ մինչև 3-րդ կարգի ավտոկոռելյացիայի գոյության մասին:

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.640254	Prob. F(3,165)	0.5901
Obs*R-squared	1.956195	Prob. Chi-Square(3)	0.5815

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Included observations: 170

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	6.24E-08	2.10E-05	0.002969	0.9976
C	0.004970	0.497163	0.009996	0.9920
RESID(-1)	-0.103384	0.086670	-1.192846	0.2346
RESID(-2)	-0.147248	0.082430	-1.786343	0.0759
RESID(-3)	-0.039124	0.083629	-0.467831	0.6405

11. «EViews» փաթեթով գնահատվել է երկու բացատրող փոփոխականներով բազմակի գծային ռեգրեսիայի մոդել, և գեներացվել է դրանց միջև բազմակողմնադրության առկայության VIF թեստը: Մեկնաբանե՞ք թեստի ելքային աղյուսակը և եզրակացություն կատարե՞ք բացատրող փոփոխականների բազմակողմնադրության մասին:

Variance Inflation Factors

Date: 11/09/25 Time: 15:53

Sample: 1 189

Included observations: 162

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
X1	5.87E-10	1.931430	1.277137
X2	0.001618	43.31660	1.277137
C	6.774695	38.80753	NA

12. «EViews» փաթեթով գնահատվել է պարզ գծային ռեգրեսիայի մոդել, և գեներացվել է Ուայթի՝ դրա պատահական մնացորդների հոմոսկեդաստիկության ստուգման թեստը: Մեկնաբանե՞ք թեստի ելքային աղյուսակը և եզրակացություն կատարե՞ք պատահական մնացորդների հոմոսկեդաստիկության մասին:

Heteroskedasticity Test: White

F-statistic	26.10012	Prob. F(2,167)	0.0000
Obs*R-squared	40.48372	Prob. Chi-Square(2)	0.0000
Scaled explained SS	43.88881	Prob. Chi-Square(2)	0.0000

Test Equation:

Dependent Variable: RESID^2

Included observations: 170

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	38.50475	3.865535	9.961040	0.0000
X^2	3.14E-08	4.40E-09	7.137342	0.0000
X	-0.002094	0.000346	-6.046281	0.0000

R-squared	0.238140	Mean dependent var	27.11918
Adjusted R-squared	0.229015	S.D. dependent var	40.52742
S.E. of regression	35.58540	Akaike info criterion	9.999237
Sum squared resid	211475.5	Schwarz criterion	10.05457
Log likelihood	-846.9352	Hannan-Quinn criter.	10.02169
F-statistic	26.10012	Durbin-Watson stat	1.649793
Prob(F-statistic)	0.000000		

13. «EViews» փաթեթով գնահատվել է պարզ գծային ռեգրեսիայի մոդել, և գեներացվել է Բրեուշ-Պագանի՝ դրա պատահական մնացորդների հոմոսկեդաստիկության ստուգման թեստը: Մեկ-նախանձ թեստի ելքային աղյուսակը և եզրակացությունը կատարելք պատահական մնացորդների հոմոսկեդաստիկության մասին:

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	0.970191	Prob. F(1,168)	0.3260
Obs*R-squared	0.976104	Prob. Chi-Square(1)	0.3232
Scaled explained SS	1.058205	Prob. Chi-Square(1)	0.3036

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Included observations: 170

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	24.91304	3.831434	6.502277	0.0000
X	0.000160	0.000162	0.984983	0.3260
R-squared	0.005742	Mean dependent var	27.11918	
Adjusted R-squared	-0.000176	S.D. dependent var	40.52742	
S.E. of regression	40.53100	Akaike info criterion	10.25371	
Sum squared resid	275984.0	Schwarz criterion	10.29060	
Log likelihood	-869.5650	Hannan-Quinn criter.	10.26868	
F-statistic	0.970191	Durbin-Watson stat	1.687339	
Prob(F-statistic)	0.326049			

14. «EViews» փաթեթով գնահատվել է երկու բացատրող փոփոխականներով բազմակի գծային ռեգրեսիայի մոդել, և իրականացվել է հատկորոշման սխալի Ռամսեյի RESET թեստը: Մեկնաբանե՞ք թեստի ելքային աղյուսակը և եզրակացությունը կատարե՞ք մոդելի հատկորոշման սխալի առկայության մասին:

Ramsey RESET Test
Specification: Y X1 X2 C

	Value	df	Probability
F-statistic	57.40210	(2, 157)	0.0000
Likelihood ratio	88.91145	2	0.0000
F-test summary:			
	Sum of Sq.	df	Mean Squares
Test SSR	1899.273	2	949.6367
Restricted SSR	4496.617	159	28.28061
Unrestricted SSR	2597.343	157	16.54359
LR test summary:			
	Value		
Restricted LogL	-499.0703		
Unrestricted LogL	-454.6145		

ԳԼՈՒԽ 4. ՈՉ ԳԾԱՅԻՆ ԿԱԽՎԱԾՈՒԹՅՈՒՆՆԵՐԻ ԳՆԱՀԱՏՈՒՄԸ

Մինչ այժմ դիտարկված էկոնոմետրիկ մոդելավորման տեխնիկան վերաբերում էր գծային կախվածություններին, որովհետև դրանց համար մշակված է հստակ մեթոդաբանություն: Իրականում տնտեսության մեջ ուսումնասիրվող կախվածությունների հսկայական մասը ոչ գծային բնույթի են: Այդ դեպքում ինչպե՞ս կարելի է կիրառել գծային ռեգրեսիայի մոդելավորման մոտեցումն այդ կախվածությունների համար: Պարզվում է, որ ոչ գծային կախվածությունների գերակշիռ մասը կարելի է գծայնացնել և այնուհետև կիրառել գծային ռեգրեսիայի էկոնոմետրիկ մոդելավորման տեխնիկան: Այս մոտեցումը կիրառելու համար նախ պետք է որոշակի մոտեցմամբ խմբավորել ոչ գծային կախվածությունները:

1. Ոչ գծային կախվածությունների տեսակները էկոնոմետրիկ մոդելավորման համատեքստում

Ինչպես հայտնի է, գծային ռեգրեսիայի դասական մոդելը ենթադրում է գծային կախվածություն պարամետրերի նկատմամբ: Ի՞նչ տեսակի ոչ գծային կախվածություններ կարող են լինել էկոնոմետրիկ մոդելների հատկորոշումներում: Ոչ գծային կախվածությունները կարելի է խմբավորել երկու հիմնական խմբերում՝

1. ոչ գծային կախվածություն բացատրող փոփոխականների նկատմամբ,
2. ոչ գծային կախվածություն էկոնոմետրիկ մոդելի պարամետրերի նկատմամբ:

Եթե կախվածությունը ոչ գծային է բացատրող փոփոխականների նկատմամբ, սակայն գծային է պարամետրերի նկատմամբ, գծային ռեգրեսիայի դասական մոդելի տեսանկյունից այն դիտարկվում է որպես գծային մոդել՝ ելնելով Գաուս-Մարկովի առաջին պայմանից, որը պահանջում է գծային կախվածություն միայն պարամետրերի նկատմամբ: Պատճառն այն է, որ 1-ին խմբի մոդելները միշտ կարելի է բերել գծային տեսքի և գնահատել ՓԶԵ-ով:

1-ին խմբի կախվածությունների օրինակներ են բազմանդամի տեսքով կախվածությունները՝

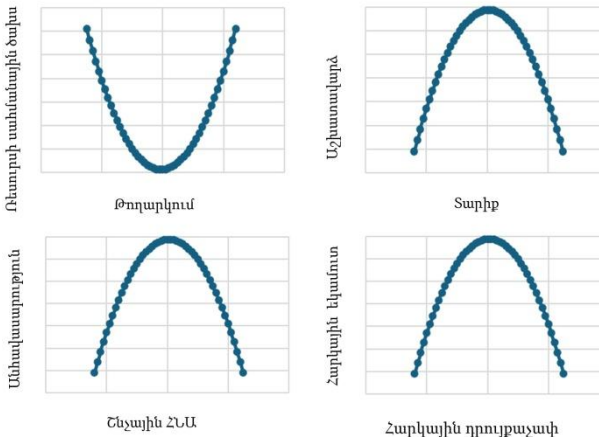
$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon: \quad (99)$$

Կամ ընդհանուր դեպքում՝ $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_k x^k + \varepsilon:$

Այս տեսքի ոչ գծային կախվածությունները հեշտությամբ գծայնացվում են՝ պարզապես նշանակումներ կատարելով: Օրինակ՝ (99)-ում կատարելով $x^2 = z$ նշանակումը՝ կստանանք գծային կախվածություն՝ $y = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon$, որն արդեն հնարավոր է գնահատել ՓԶԵ-ով:

(99) տեսքի կախվածությունները (կոչվում են U տեսքի կորեր կամ շրջված U տեսքի կորեր, U-sharp curves, inverted U-sharp curves) հաճախ են հանդիպում տնտեսագիտության տեսությունում: Դրանք կարող են օգտագործվել թողարկման ծավալ (x) - ռեսուրսի սահմանային ծախս (y), գովազդի վրա ծախսեր (x) - վաճառքների ծավալ (y), ֆիզիկական աշխատանք կատարող անձի տարիք (x) - աշխատավարձ (y), երկրի մեկ շնչի հաշվով ՀՆԱ (x) - անհավասարության մակարդակ (y) (Կուզնեցի կոր), հարկային դրույքաչափ (x) - պետական բյուջեի հարկային եկամուտներ (y) (Լաֆերի կոր) կախվածությունների գնահատման համար, որոնցում x-ի ինչ-որ մակարդակից հետո y-ի՝ x-ից կախվածության ուղղությունը փոխվում է (տե՛ս Գծանկար 14): Այդ մակարդակը կորի ուղղության շրջադարձային կետն է, որում $x = -\frac{\beta_1}{2\beta_2}$: Ըստ որում, եթե $\beta_2 < 0$, ապա կորը շրջադարձից հետո անկում է ապրում (շրջված U տեսքի կոր), իսկ եթե $\beta_2 > 0$, սկսում է աճել (U տեսքի կոր): Նշենք, որ այս կախվածությունների մի մասը կարելի է ներկայացնել նաև 3-րդ կարգի բազմանդամի տեսքով:

Գծանկար 14. U տեսքի ու շրջված U տեսքի կորերի օրինակները տնտեսագիտության տեսությունում



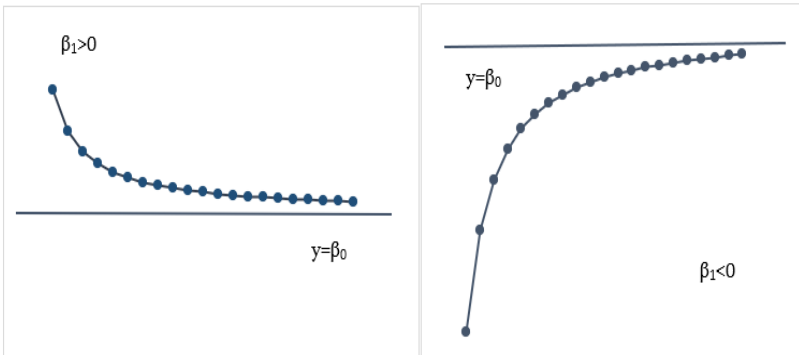
1-ին խմբի կախվածություն է մաս հակադարձ կախվածությունը՝

$$y = \beta_0 + \beta_1 \frac{1}{x} + \varepsilon: \quad (100)$$

Այս տեսքի կախվածությունները ևս կարելի է գծայնացնել՝ կատարելով նշանակում՝ $\frac{1}{x} = z$: Այդ դեպքում կունենանք գծային կախվածություն, որը հնարավոր է գնահատել ՓԶԵ-ով՝ $y = \beta_0 + \beta_1 z + \varepsilon$:

(100) տեսքի կախվածությունները բնութագրվում են նրանով, որ x -ի անընդհատ աճի հետ y -ը ասիմպտոտիկ մոտենում է ինչ-որ սահմանային արժեքի (տե՛ս Գծանկար 15): Եթե $\beta_1 > 0$, ապա երբ $x \rightarrow \infty$, y -ը ասիմպտոտիկ մոտենում է իր ներքին սահմանին, որը որոշվում է β_0 -ի միջոցով: Եթե $\beta_1 < 0$, ապա երբ $x \rightarrow \infty$, y -ը ասիմպտոտիկ մոտենում է իր վերին սահմանին, որը որոշվում է β_0 -ի միջոցով: Այս տիպի կախվածությունները ենթադրում են հագեցվածության որոշակի մակարդակի գոյություն, որը համապատասխանում է β_0 -ին:

Գծանկար 15. Հակադարձ կախվածության կորերը



Այսպիսի կախվածության տեսքը կիրառելի է այնպիսի տնտեսական կախվածությունների համար, ինչպիսիք են, օրինակ, եկամտից ռաջին անհրաժեշտության ապրանքների սպառման կախվածության էնգելի կորը, սպառման՝ ապրանքի գնից կախվածության կորը, Ֆիլիպսի կորը, որը սահմանում է կախվածություն գործազրկության մակարդակի ու աշխատավարձի աճի տեմպի միջև և այլն:

Այսպիսով՝ այս 1-ին տիպի կախվածությունները հեշտությամբ գծայնացվում են, որից հետո կարող են գնահատվել ՓՔԵ-ով: Այդ իսկ պատճառով գծային ռեգրեսիայի դասական մոդելը պահանջում է գծային կախում միայն պարամետրերի նկատմամբ:

Ի տարբերություն դրա՝ 2-րդ տիպի կախվածությունները, որոնք ոչ գծային են պարամետրերի նկատմամբ, որոշ դեպքերում հնարավոր չէ գծայնացնել: Դրանք ՓՔԵ-ով գնահատել հնարավոր չի լինի: Ստորև կանդրադառնանք պարամետրերի նկատմամբ ոչ գծային այն կախվածություններին, որոնք հնարավոր է գծայնացնել:

2. Պարամետրերի նկատմամբ ոչ գծային, բայց գծայնացվող կախվածություններ ու դրանց մոդելավորումը

2.1. Երկկողմանի լոգարիթմական ռեգրեսիա

Ենթադրենք ունենք մի կախվածություն, որում պարամետրերը բացատրող փոփոխականի ցուցիչում են: Օրինակ՝

$$y = \beta_0 \cdot x^{\beta_1} \cdot \varepsilon: \quad (101)$$

Այսպիսի կախվածությունը ոչ գծային է β_1 պարամետրի նկատմամբ, բայց գծային է բացատրող փոփոխականի՝ x -ի նկատմամբ: Այս տեսքի կախվածություններով կարող են արտահայտվել, օրինակ, պահանջարկի ու առաջարկի կորերը, արտադրական ֆունկցիաները: Օրինակ՝ Կոր-Դուգլասի արտադրական ֆունկցիան՝ $y = A \cdot K^\alpha \cdot L^\beta \cdot \varepsilon$:

Այս տեսքի կախվածությունները կարելի է գծայնացնել դրանք բնական հիմքով լոգարիթմելու միջոցով: Եթե (101)-ը բնական հիմքով լոգարիթմենք, կստանանք.

$$\ln(y) = \ln(\beta_0) + \beta_1 \ln(x) + \ln(\varepsilon): \quad (102)$$

(102)-ը երկկողմանի լոգարիթմական մոդել է, քանի որ $\ln$ կախյալ, $\ln$ բացատրող փոփոխականները մոդելում ներառված են բնական հիմքով լոգարիթմով: Այն դարձել է գծային β_1 պարամետրի նկատմամբ:

Ընդհանուր դեպքում երկկողմանի լոգարիթմական մոդելն ունի հետևյալ տեսքը՝

$$\ln(y) = \beta_0 + \beta_1 \ln(x) + \varepsilon: \quad (103)$$

(103)-ը կարելի է գնահատել ՓՋԵ-ով: Իսկապես, եթե կատարենք նշանակումներ՝ $y^* = \ln(y)$; $x^* = \ln(x)$, ապա կունենանք սովորական գծային ռեգրեսիայի մոդել.

$$y^* = \beta_0 + \beta_1 x^* + \varepsilon: \quad (104)$$

Սակայն (103) տեսքի մոդելում գնահատվող պարամետրերի իմաստն ու մեկնաբանությունը փոխվում են:

Դիտարկենք պարզ գծային ռեգրեսիայի մոդելը.

$$y = \beta_0 + \beta_1 x + \varepsilon:$$

Ինչպես նշվել է, այստեղ β_1 -ը ցույց է տալիս կախյալ փոփոխականի սպասվող արժեքի փոփոխությունը՝ անկախ x փոփոխականի միավոր փոփոխության դեպքում: Այդ փոփոխությունը ռեգրեսիայի գծի թեքությունն է և հաստատուն է.

$$\beta_1 = y_x = \frac{dy}{dx} = \frac{\Delta y}{\Delta x}: \quad (105)$$

Պարզ գծային ռեգրեսիայի պարագայում y -ի առաձգականությունն ըստ x -ի հաստատուն չէ. այն կախված է x -ի և y -ի արժեքներից.

$$E_x(y) = \frac{dy/y}{dx/x} = \beta_1 \cdot \frac{x}{y}: \quad (106)$$

Ածանցենք երկկողմանի լոգարիթմական ռեգրեսիայի (103) մոդելն ըստ x -ի.

$$\begin{aligned} \ln(y) = \beta_0 + \beta_1 \ln(x) + \varepsilon &\rightarrow \frac{1}{y} \cdot \frac{dy}{dx} = \beta_1 \cdot \frac{1}{x} \rightarrow \\ \rightarrow \beta_1 = \frac{x}{y} \cdot \frac{dy}{dx} = \frac{dy/y}{dx/x} = E_x(y) &\rightarrow \beta_1 = E_x(y): \end{aligned} \quad (107)$$

Ստացվեց, որ **երկկողմանի լոգարիթմական (103) մոդելում β_1 -ը առաձգականության իմաստ է ստանում: Այս դեպքում β_1 -ը y -ի առաձգականությունն է՝ ըստ x -ի: Այն ցույց է տալիս y -ի սպասվող արժեքի հարաբերական փոփոխությունը x -ի միավոր հարաբերական փոփոխության դեպքում:** Կամ, որ նույնն է, այն ցույց է տալիս,

թե x -ի 1 % փոփոխության (աճի) դեպքում y -ի սպասվող արժեքը քանի տոկոսով է փոխվում (աճում կամ նվազում):

Երկկողմանի լոգարիթմական մոդելում, ի տարբերություն գծային ռեգրեսիայի մոդելի, y -ի առաձգականությունն ըստ x -ի հաստատուն է, այն հավասար է β_1 -ին: Այդ պատճառով երկկողմանի լոգարիթմական մոդելները կիրառվում են այն դեպքերի համար, երբ ենթադրվում է, որ դիտարկվող կախվածությունն ունի հաստատուն առաձգականություն:

Ինչպես հետևում է (107)-ից, երկկողմանի լոգարիթմական մոդելով գնահատված կորի թեքությունը հաստատուն չէ, այն կախված է x -ի և y -ի արժեքներից.

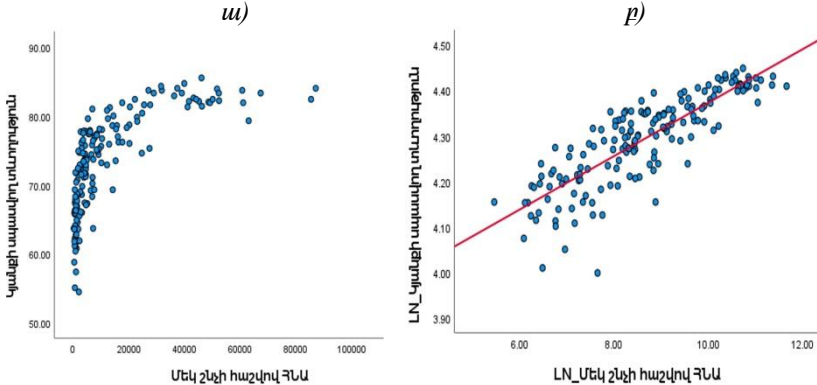
$$\beta_1 = \frac{x}{y} \cdot \frac{dy}{dx} \rightarrow \frac{dy}{dx} = \beta_1 \cdot \frac{y}{x}: \quad (108)$$

Այսպիսով՝

1. երկկողմանի լոգարիթմական մոդելը հաստատուն առաձգականությամբ մոդել է,
2. դրա β_1 պարամետրը y -ի առաձգականությունն է՝ ըստ x -ի, և այն ցույց է տալիս, թե x -ի 1 % փոփոխության դեպքում y -ի սպասվող արժեքը քանի տոկոսով է փոխվում: β_0 -ի մեկնաբանությունը հետևյալն է. եթե $x=1$ (հետևաբար $\ln(x)=0$), y -ի սպասվող արժեքը հավասար կլինի $\hat{y} = e^{\hat{\beta}_0}$,
3. Այս մոդելով գնահատված կորի թեքությունը հաստատուն չէ:

Դիտարկենք օրինակ: Ուսումնասիրվում է կյանքի միջին սպասվող տևողության կախումը երկրի մեկ շնչի հաշվով ՀՆԱ-ից: 170 երկրի վերջին հասանելի տվյալների հիման վրա կառուցվել է այդ ցուցանիշների կոռելյացիայի դաշտը, որն ակնհայտորեն վկայում է ոչ գծային կախվածության մասին (տե՛ս Գծանկար 16 ա): Գծային ռեգրեսիայի $y = \beta_0 + \beta_1 x + \varepsilon$ մոդելի գնահատումն այս պարագայում կառաջացնի հատկորոշման սխալ՝ բոլոր անցանկալի հետևանքներով: Կարող են նաև առաջանալ «խաբուսիկ» ավտոկոռելյացիա ու հետերոսկեդաստիկություն, որոնք կխախտեն Գաուս-Մարկովի պայմանները և կհանգեցնեն մոդելավորման անարժանահավատ արդյունքների: Երկու փոփոխականները բնական հիմքով լոգարիթմելու արդյունքում $\ln(y)$ -ի կախվածությունը $\ln(x)$ -ից դառնում է գծային, ինչը երևում է դրանց կոռելյացիայի դաշտից (տե՛ս Գծանկար 16 բ):

Գծանկար 16. Ոչ գծային ու երկկողմանի լոգարիթմական կախվածությունների կոռելյացիայի դաշտերը



Երկկողմանի լոգարիթմական մոդելի գնահատումը ՓՔԵ-ով տվել է հետևյալ արդյունքները՝

$$\ln(\bar{y}) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot \ln(x) = 3.8 + 0.06 \cdot \ln(x);$$

$$R^2 = 0.71; F = 411.7 (p_{value} = 0 < 0.05)$$

$$(S_{\widehat{\beta}_0} = 0.025) (S_{\widehat{\beta}_1} = 0.003):$$

Մեկնաբանություն

1. Ելնելով դետերմինացիայի գործակցի արժեքից՝ գնահատված հավասարման որակը բարձր է, կախյալ փոփոխականի վարիացիայի 71 %-ը բացատրվում է գնահատված հավասարմամբ:

2. Հավասարումը վիճակագրորեն նշանակալի է, քանի որ F վիճականու p արժեքը հավասար է զրոյի: Դա նշանակում է, որ այն վարկածի հավանականությունը, որ իրականում դետերմինացիայի գործակիցը հավասար է զրոյի, կամ, որ նույնն է, իրականում x-ը y-ին չի բացատրում, և գնահատված հավասարումը վիճակագրորեն նշանակալի չէ, փոքր է 5 տոկոսից և վարկածը մերժվում է:

3. Գնահատված գործակիցները վիճակագրորեն նշանակալի են, քանի որ $t_{\widehat{\beta}_0}^{\eta_{\text{իտ}}} = \frac{\widehat{\beta}_0}{S_{\widehat{\beta}_0}} = \frac{3.8}{0.025} = 152 > 1.96$ և $t_{\widehat{\beta}_1}^{\eta_{\text{իտ}}} = \frac{\widehat{\beta}_1}{S_{\widehat{\beta}_1}} = \frac{0.06}{0.003} = 20 > 1.96$:

4. $\widehat{\beta}_1$ -ի մեկնաբանությունը. մեկ շնչի հաշվով ՀՆԱ-ի աճը 1 %-ով հանգեցնում է կյանքի տևողության սպասվող արժեքի աճի 0.06%-ով: $\widehat{\beta}_0$ -ի մեկնաբանությունը. եթե ՀՆԱ-ն հավասար լինի մեկ դոլարի

(այսինքն՝ $\ln x = 0$), ապա՝ $\widehat{\ln(y)} = \widehat{\beta_0} \rightarrow \hat{y} = e^{\widehat{\beta_0}} = e^{3.8} = 44.7$, բնակչության կյանքի սպասվող միջին տևողությունը կլինի 44.7 տարին (արդեն նշել ենք, որ հաճախ մոդելի ազատ անդամը տնտեսագիտական իմաստ չունի):

2.2. Կիսալոգարիթմական ռեգրեսիա. գծալոգարիթմական մոդել

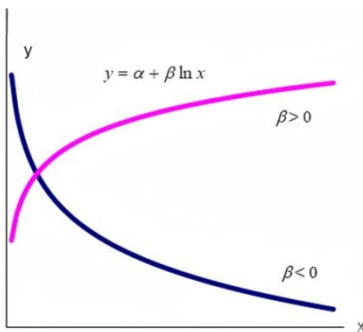
Գծալոգարիթմական մոդելում կախյալ փոփոխականն ուղիղ տեսքով է մասնակցում, իսկ բացատրող փոփոխականը՝ բնական հիմքով լոգարիթմով: Այդ պատճառով այս տեսքի մոդելները կիսալոգարիթմական մոդելներ են:

Ընդհանուր դեպքում գծալոգարիթմական մոդելն ունի հետևյալ տեսքը՝

$$y = \beta_0 + \beta_1 \cdot \ln(x) + \varepsilon: \quad (109)$$

(109) տեսքի մոդելները կարելի է գնահատել ՓԲԵ-ով և ստանալ $\widehat{\beta_0}$ -ն ու $\widehat{\beta_1}$ -ը:

Գծանկար 17. Գծալոգարիթմական կախվածության կորերը



Գծալոգարիթմական մոդելներն օգտագործվում են այն դեպքերում, երբ բացատրող փոփոխականի աճի համեմատ կախյալ փոփոխականն աճում (կամ նվազում) է նվազող արագությամբ, ինչը հնարավորություն է տալիս մոդելավորելու հազեցվածության էֆեկտը:

Գծալոգարիթմական մոդելում գնահատվող պարամետրերի, մասնավորապես ռեգրեսիայի β_1 գործակցի իմաստը տարբերվում է գծային ու երկկողմանի լոգարիթմական ռեգրեսիայի մոդելներից:

Աժանցենք (109)-ն ըստ x -ի, կունենանք՝

$$\begin{aligned} y = \beta_0 + \beta_1 \cdot \ln(x) + \varepsilon &\rightarrow \frac{dy}{dx} = \beta_1 \cdot \frac{1}{x} \rightarrow \beta_1 = x \cdot \frac{dy}{dx} = \frac{dy}{dx/x} \rightarrow \\ &\rightarrow \beta_1 = \frac{dy}{dx/x}: \end{aligned} \quad (110)$$

β_1 -ը այս դեպքում ցույց է տալիս y կախյալ փոփոխականի բացարձակ փոփոխությունը բացատրող x փոփոխականի միավոր հարաբերական փոփոխության դեպքում: Այսինքն՝ այն ցույց է տալիս, թե y -ը քանի միավորով է փոխվում x -ի 1 % փոփոխության դեպքում: Այս մոդելում **β_1 -ի մեկնաբանությունը հետևյալն է. եթե x -ը աճի 1%-ով, y -ը կաճի (կամ կնվազի) $\frac{\beta_1}{100} = \beta_1 \cdot 0.01$ միավորով:** β_0 -ն ցույց է տալիս, որ եթե $x=1$, y -ի սպասվող արժեքը հավասար կլինի β_0 -ի:

Այս դեպքում արդեն (110)-ից հետևում է, որ y -ի առաձգականությունն ըստ x -ի կախված է y կախյալ փոփոխականի արժեքներից. այն նվազում է վերջինիս աճի համեմատ: Այսինքն՝ **այդ առաձգականությունը հաստատուն չէ.**

$$E_x(y) = \frac{dy/y}{dx/x} = \beta_1 \cdot \frac{1}{x} \cdot \frac{x}{y} = \beta_1 \cdot \frac{1}{y}; \quad (111)$$

Այս մոդելով նկարագրվող կորի թեքությունը ևս հաստատուն չէ, կախված է բացատրող փոփոխականի արժեքներից և նվազում է վերջինիս աճի համեմատ՝

$$\frac{dy}{dx} = \beta_1 \cdot \frac{1}{x}:$$

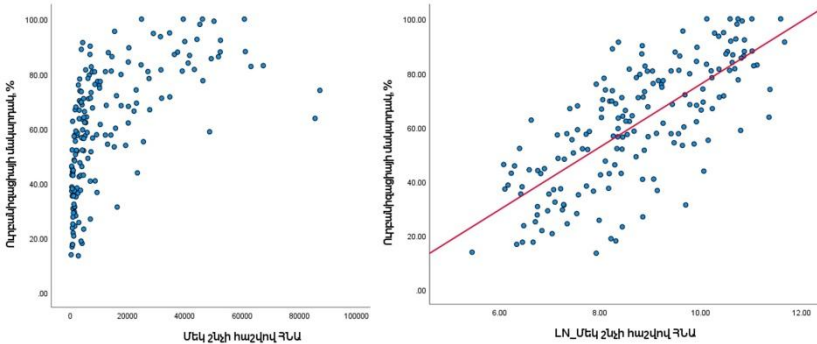
Այսպիսով՝

1. գծադրարիթմական մոդելը փոփոխական առաձգականությամբ մոդել է, y -ի առաձգականությունն ըստ x -ի նվազում է y կախյալ փոփոխականի արժեքների աճի հետ,
2. β_1 գործակցի մեկնաբանությունն է՝ x -ի 1 % աճի դեպքում y -ի սպասվող արժեքը կփոխվի $\frac{\beta_1}{100} = \beta_1 \cdot 0.01$ միավորով: β_0 գործակցի մեկնաբանությունն է՝ եթե $x=1$ ($\ln(x)=0$), ապա y -ի սպասվող արժեքը հավասար կլինի β_0 -ի,
3. այս մոդելով գնահատված կորի թեքությունը հաստատուն չէ. այն նվազում է բացատրող փոփոխականի՝ x -ի արժեքների աճի համեմատ:

Դիտարկենք օրինակ: Ուսումնասիրվում է ուրբանիզացիայի մակարդակի կախումը մեկ շնչի հաշվով ՀՆԱ-ից: 170 երկրի վերջին հասանելի տվյալների կոռելյացիայի դաշտը փաստում է ոչ գծային կախվածության մասին (տե՛ս Գծանկար 18 ա), ըստ որում՝ ՀՆԱ-ի աճի հետ ուրբանիզացիայի մակարդակն աճում է նվազող արագու-

թյամբ: Բացատրող փոփոխականը բնական հիմքով լոգարիթմելու արդյունքում կախվածությունն ուղղվում է (տե՛ս Գծանկար 18 ք):

Գծանկար 18. Ոչ գծային ու գծալոգարիթմական կախվածությունների կոռելյացիայի դաշտերը



Ենթադրենք՝ ՓԲԵ-ով գնահատվել է ուրբանիզացիայի մակարդակի կախումը մեկ շնչի հաշվով ՀՆԱ-ի բնական հիմքով լոգարիթմից ներկայացնող գծալոգարիթմական մոդել: Ստացվել են հետևյալ արդյունքները.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \ln(x) = -40.2 + 11.6 \cdot \ln(x);$$

$$R^2 = 0.54; F = 194.3 (p_{value} = 0 < 0.05)$$

$$(t_{\hat{\beta}_0} = -5.5) \quad (t_{\hat{\beta}_1} = 13.9):$$

Մեկնաբանությունը

1. Ելնելով դետերմինացիայի գործակցի արժեքից՝ գնահատված հավասարման որակն ընդունելի է, կախյալ փոփոխականի վարիացիայի կեսից քիչ ավելին՝ 54 %-ը, բացատրվում է գնահատված հավասարմամբ:
2. Հավասարումն ու դետերմինացիայի գործակիցը վիճակագրորեն նշանակալի են, քանի որ F վիճականու p արժեքը հավասար է զրոյի: Դա նշանակում է, որ այն վարկածի հավանականությունը, որ իրականում x-ը y-ին չի բացատրում, և գնահատված հավասարումը վիճակագրորեն նշանակալի չէ, ունի 5 տոկոսից փոքր հավանականություն, և վարկածը մերժվում է:
3. Գնահատված գործակիցները վիճակագրորեն նշանակալի են, քանի որ $t_{\hat{\beta}_0}^{nhun} > 1.96$ և $t_{\hat{\beta}_1}^{nhun} > 1.96$:

4. $\widehat{\beta}_1$ -ի մեկնաբանությունը. մեկ շնչի հաշվով ՀՆԱ-ի աճը 1 %-ով հանգեցնում է ուրբանիզացիայի մակարդակի սպասվող արժեքի աճի $11.6/100=0.11$ միավորով: $\widehat{\beta}_0$ -ն տվյալ դեպքում ցույց է տալիս, թե եթե մեկ շնչի հաշվով ՀՆԱ-ն հավասար լինի մեկ դոլարի (այսինքն՝ $\ln x = 0$), ապա ուրբանիզացիայի սպասվող մակարդակը հավասար կլինի -40.2 -ի:

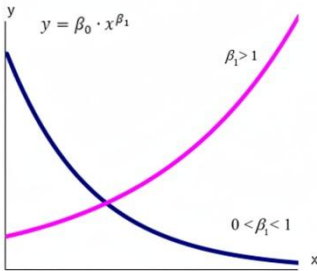
2.3. Կիսալոգարիթմական ռեգրեսիա. լոգ-գծային մոդել

Լոգ-գծային մոդելում կախյալ փոփոխականը մասնակցում է բնական հիմքով լոգարիթմով, իսկ բացատրող փոփոխականը՝ ուղիղ տեսքով:

Լոգ-գծային մոդելների ընդհանուր տեսքը հետևյալն է՝

$$\ln(y) = \beta_0 + \beta_1 x + \varepsilon: \quad (112)$$

Գծանկար 19. Լոգ-գծային կախվածության կորերը



Գծալոգարիթմական մոդելներն կիրառվում են այն դեպքերում, երբ բացատրող փոփոխականի աճի հետ կախյալ փոփոխականն աճում (կամ նվազում) է մեծացող արագությամբ, ավելի ու ավելի արագ:

Լոգ-գծային մոդելում գնահատվող պարամետրերի բովանդակությունը տարբերվում է վերը քննարկված դեպքերից:

Ածանցենք (112)-ն ըստ x -ի կոմենանք՝

$$\ln(y) = \beta_0 + \beta_1 \cdot x + \varepsilon \rightarrow \frac{1}{y} \cdot \frac{dy}{dx} = \beta_1 \rightarrow \beta_1 = \frac{dy/y}{dx}: \quad (113)$$

β_1 -ը այս մոդելում ցույց է տալիս կախյալ y փոփոխականի հարաբերական փոփոխությունը բացատրող x փոփոխականի միավոր բացարձակ փոփոխության դեպքում: Այսինքն՝ այն ցույց է տալիս, թե քանի տոկոսով է փոխվում y -ը x -ի մեկ միավոր բացարձակ փոփոխության դեպքում: Այս մոդելում β_1 -ի մեկնաբանությունը հետևյալն է. եթե x -ը աճի 1 միավորով, y -ի սպասվող արժեքը կաճի $\beta_1 \cdot 100$ տոկոսով:

սով: β_0 -ի մեկնաբանությունը հետևյալն է. եթե $x=0$, ապա կախյալ փոփոխականի սպասվող արժեքը հավասար կլինի $\hat{y} = e^{\beta_0}$:

Այս դեպքում արդեն, ելնելով (113)-ից, y -ի առաձգականությունն ըստ x -ի հավասար է՝

$$E_x(y) = \frac{dy/y}{dx/x} = \beta_1 \cdot x: \quad (114)$$

Սա նշանակում է, որ այդ առաձգականությունը կախված է բացատրող փոփոխականի՝ x -ի արժեքներից, այն աճում է վերջինիս աճի համեմատ: Այսինքն՝ այդ առաձգականությունը հաստատում չէ:

Այս մոդելով նկարագրվող կորի թեքությունը հաստատում չէ. այն կախված է կախյալ փոփոխականի՝ y -ի արժեքներից և աճում է վերջինիս աճի համեմատ՝

$$\frac{dy}{dx} = \beta_1 \cdot y: \quad (115)$$

Այսպիսով՝

1. լոգ-գծային մոդելը փոփոխական առաձգականությամբ մոդել է, y -ի առաձգականությունը, ըստ x -ի, աճում է բացատրող փոփոխականի՝ x -ի արժեքների աճի համեմատ,
2. այս մոդելի β_1 գործակիցը ցույց է տալիս, թե x -ի 1 միավոր փոփոխության դեպքում y -ի սպասվող արժեքը քանի տոկոսով է փոխվում՝ **$\beta_1 \cdot 100$ տոկոսով**: Մոդելի β_0 գործակիցը մեկնաբանվում է այսպես. եթե $x=0$, ապա կախյալ փոփոխականի սպասվող արժեքը հավասար կլինի $\hat{y} = e^{\beta_0}$,
3. այս տեսքի մոդելով գնահատված կորի թեքությունը հաստատում չէ, այն աճում է կախյալ փոփոխականի՝ y -ի արժեքների աճի համեմատ:

Դիտարկենք օրինակ: 104 բարձր մարդկային զարգացման համաթիվ ունեցող երկրների վերջին հասանելի տվյալների հիման վրա գնահատվել է մեկ շնչի հաշվով ՀՆԱ-ի բնական հիմքով լոգարիթմի կախումը աշխատուժի մասնակցության մակարդակից՝ $\ln y = \beta_0 + \beta_1 \cdot x + \varepsilon$: Սա լոգ-գծային մոդել է, որի ՓԶԵ-ով գնահատումը տվել է հետևյալ արդյունքները՝

$$\begin{aligned} \widehat{\ln(y)} &= \widehat{\beta_0} + \widehat{\beta_1} \cdot x = 4.8 + 0.07 \cdot x; \\ R^2 &= 0.51; F = 66.5 (p_{value} = 0 < 0.05) \\ (t_{\widehat{\beta_0}} &= 8.5) \quad (t_{\widehat{\beta_1}} = 8.2): \end{aligned}$$

Մեկնաբանությունը

1. Ելնելով դետերմինացիայի գործակցի արժեքից՝ գնահատված հավասարման որակն ընդունելի է, բայց բարձր չէ: Կախյալ փոփոխականի վարիացիայի կեսից քիչ ավելին՝ 51 %-ը, բացատրվում է գնահատված հավասարմամբ:
2. Հավասարումն ու դետերմինացիայի գործակիցը վիճակագրորեն նշանակալի են, քանի որ F վիճականու p արժեքը հավասար է զրոյի: Դա նշանակում է, որ այն վարկածի հավանականությունը, որ իրականում x -ը y -ին չի բացատրում, և գնահատված հավասարումը վիճակագրորեն նշանակալի չէ, ունի 5 %-ից փոքր հավանականություն, և վարկածը մերժվում է:
3. Գնահատված գործակիցները վիճակագրորեն նշանակալի են, քանի որ $t_{\widehat{\beta}_0}^{\eta_{\text{hun}}} > 1.96$ և $t_{\widehat{\beta}_1}^{\eta_{\text{hun}}} > 1.96$:
4. $\widehat{\beta}_1$ -ի մեկնաբանությունը հետևյալն է. աշխատուժի մասնակցության մակարդակի աճը մեկ միավորով հանգեցնում է մեկ շնչի հաշվով ՀՆԱ-ի սպասվող արժեքի աճի՝ $0.07 \cdot 100 = 7\%$ -ով: $\widehat{\beta}_0$ -ի մեկնաբանությունը հետևյալն է. եթե աշխատուժի մասնակցության մակարդակը հավասար լինի 0-ի, մեկ շնչի հաշվով ՀՆԱ-ի սպասվող արժեքը հավասար կլինի $e^{\widehat{\beta}_0} = e^{4.8} = 121.5$:

Նշենք, որ երկկողմանի լոգարիթմական ու կիսալոգարիթմական մոդելներն ունեն սահմանափակումներ այն իմաստով, որ եթե լոգարիթմվող փոփոխականները որոշ դիտարկումների համար ընդունում են զրոյին հավասար կամ բացասական արժեքներ, այդ դիտարկումների համար լոգարիթմած փոփոխականները չեն որոշվում, և մոդելի ազատության աստիճանը փոքրանում է:

Աղյուսակ 9-ում ամփոփված են գծային, երկկողմանի ու կիսալոգարիթմական մոդելների հատկությունները:

Աղյուսակ 9. Գծային, երկկողմանի ու կիսալոգարիթմական մոդելների հարկությունները

Ռեգրեսիայի մոդել	Գծային	Երկկողմանի լոգարիթմական	Գծա-լոգարիթմական	Լոգ-գծային
Էնսիթիկ հավասարում	$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$	$\ln(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1 \ln(x)$	$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \ln(x)$	$\ln(\hat{y}) = \hat{\beta}_0 + \hat{\beta}_1 x$
Կորի բերու-թյուն	$\frac{dy}{dx} = \beta_1 = const$	$\frac{dy}{dx} = \beta_1 \cdot \frac{y}{x} \neq const$	$\frac{dy}{dx} = \beta_1 \frac{1}{x} \neq const$	$\frac{dy}{dx} = \beta_1 y \neq const$
Առաձգականություն $E_x(y)$	$E_x(y) = \beta_1 \cdot \frac{x}{y} \neq const$	$E_x(y) = \beta_1 = const$	$E_x(y) = \beta_1 \cdot \frac{1}{y} \neq const$	$E_x(y) = \beta_1 \cdot x \neq const$
β_1	$\beta_1 = \frac{dy}{dx}$	$\beta_1 = \frac{dy/y}{dx/x} = E_x(y)$	$\beta_1 = \frac{dy}{dx/x}$	$\beta_1 = \frac{dy/y}{dx}$
β_1 -ի մեկնաբանությունը	x-ի 1 միավոր աճ $\rightarrow$ y-ի սպասվող արժեքի β_1 միավոր աճ	x-ի 1% աճ $\rightarrow$ y-ի սպասվող արժեքի β_1 % աճ	x-ի 1% աճ $\rightarrow$ y-ի սպասվող արժեքի $\beta_1/100$ միավոր աճ	x-ի 1 միավոր աճ y-ի $\rightarrow$ սպասվող արժեքի $\beta_1 \cdot 100$ տոկոս աճ
β_0 -ի մեկնաբանությունը	$x=0 \rightarrow \hat{y} = \hat{\beta}_0$	$x=1 \rightarrow \hat{y} = e^{\hat{\beta}_0}$	$x=1 \rightarrow \hat{y} = \hat{\beta}_0$	$x=0 \rightarrow \hat{y} = e^{\hat{\beta}_0}$

Ներդիր 15. Կիսալոգարիթմական մոդելների գործակիցների մեկնաբանության բացատրությունը

Գծալոգարիթմական մոդել. x-ի 1% աճը պայմանավորում է y-ի սպասվող արժեքի $\beta_1/100$ միավոր աճ

Ունենք՝ $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \ln(x)$: Ենթադրենք ունենք որևէ դիտարկման համար էմպիրիկ տվյալների (x_1, y_1) զույգը: Էմպիրիկ հավասարումն այդ դիտարկման համար կլինի հետևյալը՝ $\hat{y}_1 = \hat{\beta}_0 +$

$+\widehat{\beta}_1 \ln(x_1)$: Ենթադրենք x_1 -ն աճել է 1 %-ով և ընդունել է x_2 արժեքը: Սա նշանակում է, որ $x_2 = x_1 \cdot \left(1 + \frac{1}{100}\right) = 1.01 \cdot x_1$: x_2 -ին համապատասխանող կախյալ փոփոխականի սպասվող արժեքը կլինի՝ $\widehat{y}_2 = \widehat{\beta}_0 + \widehat{\beta}_1 \ln(x_2) = \widehat{\beta}_0 + \widehat{\beta}_1 \ln(1.01 \cdot x_1)$: x_1 -ի 1 %-ի աճի դեպքում կախյալ փոփոխականի գնահատված արժեքի փոփոխությունը կհաշվարկվի այսպես. $\Delta \widehat{y} = \widehat{y}_2 - \widehat{y}_1 = \widehat{\beta}_0 + \widehat{\beta}_1 \ln(1.01 \cdot x_1) - \widehat{\beta}_0 - \widehat{\beta}_1 \ln(x_1) = \widehat{\beta}_1 \cdot (\ln(1.01 \cdot x_1) - \ln(x_1)) = \widehat{\beta}_1 \cdot (\ln(1.01) + \ln(x_1) - \ln(x_1)) = \widehat{\beta}_1 \cdot \ln(1.01) = \widehat{\beta}_1 \cdot 0.00995 \approx 0.01 \widehat{\beta}_1 \approx \frac{\widehat{\beta}_1}{100}$:

Լոգ-գծային մոդել. x-ի 1 %-ի աճը պայմանավորում է y-ի սպասվող արժեքի $\widehat{\beta}_1 \cdot 100$ միավոր աճ

Ունենք՝ $\ln(\widehat{y}) = \widehat{\beta}_0 + \widehat{\beta}_1 x$: Ենթադրենք ունենք որևէ դիտարկման համար էմպիրիկ տվյալների (x_1, y_1) զույգը: Էմպիրիկ հավասարումն այդ դիտարկման համար կլինի հետևյալը՝ $\ln(\widehat{y}_1) = \widehat{\beta}_0 + \widehat{\beta}_1 x_1$:

Ավելացնենք x-ը մեկ միավորով՝ $x_2 = x_1 + 1$: x_2 -ին համապատասխանող կախյալ փոփոխականի սպասվող արժեքը կլինի՝ $\ln(\widehat{y}_2) = \widehat{\beta}_0 + \widehat{\beta}_1(x_1 + 1) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_1 + \widehat{\beta}_1$: x_1 -ի միավոր աճի դեպքում կախյալ փոփոխականի գնահատված արժեքի լոգարիթմի փոփոխությունը կհաշվարկվի այսպես. $\Delta \ln(\widehat{y}) = \ln(\widehat{y}_2) - \ln(\widehat{y}_1) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_1 + \widehat{\beta}_1 - \widehat{\beta}_0 - \widehat{\beta}_1 x_1 = \widehat{\beta}_1$: Մյուս կողմից, լոգարիթմի հատկություններից բխում է, որ $\Delta \ln(\widehat{y}) = \ln(\widehat{y}_2) - \ln(\widehat{y}_1) = \ln\left(\frac{\widehat{y}_2}{\widehat{y}_1}\right)$: Այս վերջին հավասարման երկու կողմը էքսպոնենցիա-

լացնենք, կստացվի՝ $\Delta \widehat{y} = \frac{\widehat{y}_2}{\widehat{y}_1} = e^{\widehat{\beta}_1}$: Ստացվեց, որ x_1 -ի միավոր աճի դեպքում կախյալ փոփոխականի գնահատված արժեքը կաճի $e^{\widehat{\beta}_1}$ միավորով: x_1 -ի միավոր աճի դեպքում կախյալ փոփոխականի տոկոսային փոփոխությունը կհաշվարկվի հետևյալ բանաձևով՝ $\frac{\widehat{y}_2 - \widehat{y}_1}{\widehat{y}_1} \cdot 100 = (e^{\widehat{\beta}_1} - 1) \cdot 100$: Երբ $\widehat{\beta}_1$ -ը բավականաչափ փոքր է, $e^{\widehat{\beta}_1}$ -ի Թեյլորի շարքի վերածումը վկայում է, որ $e^{\widehat{\beta}_1} \approx 1 + \widehat{\beta}_1$: Հետևաբար $\frac{\widehat{y}_2 - \widehat{y}_1}{\widehat{y}_1} \cdot 100 \approx \widehat{\beta}_1 \cdot 100$, ինչը նշանակում է, որ x_1 -ի միավոր աճի դեպքում կախյալ փոփոխականը կաճի մոտավորապես $\widehat{\beta}_1 \cdot 100$ %-ով:

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Բացատրող փոփոխականների նկատմամբ ոչ գծային կախվածության ինչպիսի՞ տեսքերի օրինակներ կարող եք բերել: Տնտեսագիտական ինչպիսի՞ իմաստ կարող են ունենալ այդ կախվածությունները:
2. Ընդհանուր դեպքում ի՞նչ տեսք ունի երկկողմանի լոգարիթմական ռեգրեսիայի մոդելը:
3. Ինչի՞ է հավասար երկկողմանի լոգարիթմական մոդելով գնահատված կորի թեքությունը, ինչի՞ց է այն կախված:
4. Ի՞նչ մեկնաբանություն ունեն երկկողմանի լոգարիթմական ռեգրեսիայի մոդելի գնահատված գործակիցները:
5. Ամփոփե՛ք երկկողմանի լոգարիթմական ռեգրեսիայի մոդելի հիմնական հատկությունները:
6. Ընդհանուր դեպքում ի՞նչ տեսք ունի գծալոգարիթմական ռեգրեսիայի մոդելը: Ինչի՞ է հավասար այդ մոդելով գնահատված կորի թեքությունը, ինչի՞ց է այն կախված: Ինչի՞ է հավասար այդ մոդելում y-ի առաձգականությունն ըստ x-ի:
7. Ի՞նչ մեկնաբանություն ունեն գծալոգարիթմական ռեգրեսիայի մոդելի գնահատված գործակիցները:
8. Ամփոփե՛ք գծալոգարիթմական ռեգրեսիայի մոդելի հիմնական հատկությունները:
9. Ընդհանուր դեպքում ի՞նչ տեսք ունի լոգ-գծային ռեգրեսիայի մոդելը: Ինչի՞ է հավասար այդ մոդելով գնահատված կորի թեքությունը, ինչի՞ց է այն կախված: Ինչի՞ է հավասար այդ մոդելում y-ի առաձգականությունն ըստ x-ի:
10. Ի՞նչ մեկնաբանություն ունեն լոգ-գծային ռեգրեսիայի մոդելի գնահատված գործակիցները:
11. Ամփոփե՛ք լոգ-գծային ռեգրեսիայի մոդելի հիմնական հատկությունները:
12. Բացատրե՛ք գծալոգարիթմական ու լոգ-գծային ռեգրեսիայի մոդելների գնահատված գործակիցների մեկնաբանության ստացումը:

Խնդիրներ

1. Հիմնվելով որոշակի երկրի վիճակագրական տվյալների վրա՝ կառուցվել է Կոր-Դուգլասի արտադրական ֆունկցիան՝ հետևյալ տեսքի մոդելով.

$$\begin{aligned}\widehat{\ln(y)} &= -3.52 + 0.53\ln(K) + 0.47\ln(L); \quad R_{adj}^2 = 0.815; \\ F^{\eta_{\text{hun}}} &= 286.7 \\ t_{\widehat{\beta}_0}^{\eta_{\text{hun}}} &= -1.45; \quad t_{\widehat{\beta}_1}^{\eta_{\text{hun}}} = 3.16; \quad t_{\widehat{\beta}_2}^{\eta_{\text{hun}}} = 5.31,\end{aligned}$$

որում y -ը իրական ՀՆԱ-ն է (միլիոն դոլար), K -ն՝ կապիտալ ծախսերի ծավալը (միլիոն դոլար), L -ը՝ աշխատանքային ծախսերի ծավալը (մարդ-օր):

1. Ի՞նչ տիպի կախվածություն է գնահատված այս մոդելով:
 2. Մեկնաբանե՞ք ռեգրեսիայի գործակիցների իմաստը:
 3. Կատարե՞ք եզրակացություն գնահատված ռեգրեսիայի էմպիրիկ մոդելի որակի մասին:
 4. Ստուգե՞ք ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին վարկածը և կատարե՞ք եզրակացություն ռեգրեսիայի հավասարման վիճակագրական նշանակալիության մասին:
 5. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն՝ կատարե՞ք ընդհանուր եզրակացություն գնահատված մոդելի կիրառելիության մասին:
2. Գնահատվել է աղքատության մակարդակի (Poverty) կախումը շնչային ՀՆԱ-ից (GDP_PC) և անհավասարության Ջինիի գործակցից ($Gini$) հետևյալ տեսքի էկոնոմետրիկ մոդելով՝

$$\ln(Poverty) = \beta_0 - \beta_1 \ln(GDP_PC) + \beta_2 Gini + \varepsilon:$$

1. Ի՞նչ տիպի կախվածություն է սա՝ ա) շնչային ՀՆԱ-ի նկատմամբ, բ) Ջինիի գործակցի նկատմամբ:
 2. Տվե՞ք β_1 և β_2 գործակիցների մեկնաբանությունները:
3. Միևնույն ճյուղի 230 ֆիրմաների տվյալների հիման վրա գնահատված է դրանց թողարկման ծավալի բնական հիմքով լոգարիթմի ($\ln_OUTPUT$, ԱՄՆ \$) գծային կախումը ծախսված ռե-

սուրսներից՝ հիմնական միջոցների լոգարիթմից (LN_FUNDS, ԱՄՆ \$) ու աշխատավարձից (LABOR, ԱՄՆ \$):

SPSS փաթեթով մոդելը գնահատելու արդյունքում ստացվել են հետևյալ ելքային աղյուսակները՝

Model Summary

Model	R	R Square	Adjusted R Square	F	Sig.
1	0.631	0.501	0.497	113.9	0.000

Model		Coefficients		t	Sig.
		B	Std. Error		
1	(Constant)	-3700.4	641.5	-5.77	0.000
	LN_FUNDS	1.92	0.88	2.18	0.000
	LABOR	0.006	0.0012	5.0	0.000

1. Գրե՛ք գնահատված հավասարումը և մեկնաբանե՛ք բոլոր գործակիցների իմաստը:
 2. Կատարե՛ք եզրակացություններ այն մասին, թե գնահատված β -ները վիճակագրորեն նշանակալի են, թե ոչ: Հիմնավորե՛ք եզրակացությունները:
 3. Կատարե՛ք եզրակացություններ այն մասին, թե հավասարումը վիճակագրորեն նշանակալի է, թե ոչ, հիմնավորե՛ք եզրակացությունները:
 4. Կատարե՛ք եզրակացություն գնահատված հավասարման որակի մասին:
 5. Ենթադրելով, որ Գաուս-Մարկովի պայմանները տեղի ունեն՝ կատարե՛ք ընդհանուր եզրակացություն գնահատված մոդելի կիրառելիության մասին:
4. Գնահատվել է ագրեգացված սպառման (Y) կախումը տնօրինվող եկամտից (X) հետևյալ կիսալոգարիթմական տեսքի էկոնոմետրիկ մոդելով՝

$$y = \beta_0 + \beta_1 \ln(x) + \varepsilon:$$

Ընտրե՛ք առաջարկված տարբերակներից միայն մեկը՝ օղակի մեջ վերցնելով տարբերակի համարը, և հիմնավորե՛ք ձեր ընտրությունը տեսական մեկնաբանությամբ:

- ա) տնօրինվող եկամտի ավելացումը 1 %-ով միջինում կավելացնի ագրեգացված սպառումը β_1 միավորով,
- բ) տնօրինվող եկամտի ավելացումը մեկ միավորով միջինում կավելացնի ագրեգացված սպառումը β_1 միավորով,
- գ) տնօրինվող եկամտի ավելացումը մեկ միավորով միջինում կավելացնի ագրեգացված սպառումը $\beta_1 \cdot 100$ %-ով,
- դ) տնօրինվող եկամտի ավելացումը 1 %-ով միջինում կավելացնի ագրեգացված սպառումը $\beta_1 / 100$ միավորով,
- ե) տնօրինվող եկամտի ավելացումը 1 %-ով միջինում միջինում կավելացնի ագրեգացված սպառումը β_1 %-ով:

5. Համաշխարհային բանկի «Աշխարհի զարգացման ցուցանիշներ» (World development indicators) օնլայն տեղեկատվական հարթակից (էլեկտրոնային հասցեն՝

<https://databank.worldbank.org/source/world-development-indicators>) MS Excel ձևաչափով ներբեռնել համացանցային ամրագրված լայնաշերտ կապի բաժանորդների քանակի (100 մարդու հաշվով՝ Fixed broadband subscriptions (per 100 people)) ու մեկ շնչի հաշվով ՀՆԱ-ի (ԱՄՆ դոլարով՝ GDP per capita (constant 2015 US\$)) ցուցանիշների ամենավերջին հասանելի տարվան վերաբերող տվյալները աշխարհի 217 երկրների համար:

1. Չնափոխել Excel-ի տվյալներն այնպես, որ դրանք համապատասխանեն էկոնոմետրիկ մոդելավորման համար օգտագործվող տարածական տվյալների երկչափ շտեմարանների պահանջներին:
2. Տեղափոխել տվյալները «EViews» փաթեթ և հետազոտել ցուցանիշների գույգային կոռելյացիան, ստացել կոռելյացիայի դաշտը: Կատարել եզրակացություն կախվածության տեսքի մասին:
3. Գնահատել միագործոն գծային ռեգրեսիայի մոդել՝ համացանցային ամրագրված լայնաշերտ կապի բաժանորդների քանակի կախումը մեկ շնչի հաշվով ՀՆԱ-ից:
4. Գրել գծային ռեգրեսիայի գնահատված էմպիրիկ հավասարումը և մեկնաբանել հավասարման գնահատված գործակիցների իմաստը: Ստուգել β_0 և β_1 գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարել եզրակացություն գնահատված β -ների վիճակագրական նշանակալիության մասին: Կատարել եզրակացություն գնահատված

գծային ռեգրեսիայի էմպիրիկ մոդելի որակի, գնահատված գծային ռեգրեսիայի էմպիրիկ հավասարման վիճակագրական նշանակալիության մասին:

5. Ստուգե՛ք՝ Գաուս-Մարկովի պայմանները տեղի ունեն, թե ոչ վերը նշված բոլոր թեստերով, նաև եզրակացություն կատարեք մոդելի գնահատման արդյունքների կիրառելիության մասին:
6. Գնահատե՛ք միագործոն գծալոգարիթմական ռեգրեսիայի մոդել՝ համացանցային ամրագրված լայնաշերտ կապի բաժանորդների քանակի կախումը մեկ շնչի հաշվով ՀՆԱ-ի բնական հիմքով լոգարիթմից:
7. Գրե՛ք գծալոգարիթմական ռեգրեսիայի գնահատված էմպիրիկ հավասարումը և մեկնաբանե՛ք հավասարման գնահատված $\widehat{\beta}_0$ ու $\widehat{\beta}_1$ գործակիցների իմաստը: Ստուգե՛ք $\widehat{\beta}_0$ և $\widehat{\beta}_1$ գործակիցների վիճակագրական նշանակալիության մասին վարկածները և կատարե՛ք եզրակացություններ այն մասին, թե գնահատված β -ները վիճակագրորեն նշանակալի են, թե ոչ: Կատարե՛ք եզրակացություն գնահատված գծային ռեգրեսիայի էմպիրիկ մոդելի որակի, գնահատված գծային ռեգրեսիայի էմպիրիկ հավասարման վիճակագրական նշանակալիության մասին:
8. Ստուգե՛ք՝ Գաուս-Մարկովի պայմանները տեղի ունեն, թե ոչ վերը նշված բոլոր թեստերով, և եզրակացություն կատարեք մոդելի գնահատման արդյունքների կիրառելիության մասին:
9. Համեմատե՛ք գնահատված երկու մոդելները Գաուս-Մարկովի պայմանների ու մոդելի կիրառելիության իմաստով:

ԳԼՈՒԽ 5. ՖԻԿՏԻՎ ՓՈՓՈԽԱԿԱՆՆԵՐԻ ՕԳՏԱԳՈՐԾՈՒՄԸ ԷԿՈՆՈՄԵՏՐԻԿ ՄՈԴԵԼՆԵՐՈՒՄ

Մինչ այժմ դիտարկված էկոնոմետրիկ մոդելներում նկատի ունենք, որ էմպիրիկ տվյալները քանակական բնույթ ունեն: Սակայն շատ հաճախ էկոնոմետրիկ հետազոտությունը պահանջում է որակական տվյալների ներառում մոդելում: Այս փոփոխականները չեն կարող ուղղակիորեն ներառվել ռեգրեսիայի մոդելում՝ առանց դրանց քանակական, թվային ներկայացման: Այդ պատճառով որակական փոփոխականներին վերագրվում են թվային արժեքներ, որոնք արտահայտում են այս կամ այն որակական հատկանիշի առկայությունը: Օրինակ՝ ենթադրենք ուսումնասիրվում է աշխատավարձի կախվածությունը կրթամակարդակից: Կրթամակարդակը հնարավոր է մոտարկել սովորած տարիների քանակով, որը քանակական փոփոխական է և չափվում է տարիներով: Սակայն ավելի խոր վերլուծության համար կարելի է օգտագործել որակական փոփոխական, որն ունի կրթամակարդակը բնութագրող որակական կատեգորիաներ՝ տարրական, հիմնական, լրիվ միջնակարգ, միջին մասնագիտական, բարձրագույն, հետբարձրագույն: Այս որակական փոփոխականը մոդելի մեջ մտցնելու համար անհրաժեշտ է որակի 6 կատեգորիայից յուրաքանչյուրին վերագրել դրական ամբողջ թիվ՝ այդպիսով «կոդավորելով» որակները: Կունենանք արհեստականորեն ստեղծված քանակական փոփոխական, որի հնարավոր արժեքները կլինեն $[1;6]$ միջակայքի ամբողջ թվերը: Այսպիսի փոփոխականները կոչվում են «ֆիկտիվ» փոփոխականներ, քանի որ ստեղծվել են արհեստականորեն: Այսպիսով, ֆիկտիվ փոփոխականները ընդունում են որոշակի քանակությամբ արժեքներ, որոնք դրական ամբողջ թվեր են՝ $0, 1, 2, 3, \dots$, և ներկայացնում են դիտարկվող օբյեկտի որոշակի որակական հատկանիշին վերագրված թվային արժեքներ՝ կոդեր: Ֆիկտիվ փոփոխականները էկոնոմետրիկ մոդելներում կարող են օգտագործվել թե՛ որպես կախյալ, թե՛ որպես անկախ փոփոխականներ: Այս դեպքերի մոդելավորման տեխնիկաները տարբեր են:

Մասնավորապես, ֆիկտիվ փոփոխականը կարող է արտահայտել որոշակի որակի առկայություն կամ բացակայություն: Այսպիսի որակական փոփոխականները կոչվում են երկարժեք ֆիկտիվ փոփոխականներ (Dummy) և ընդունում են երկու արժեք՝ 0 ու 1՝ կախված նրանից, թե տվյալ որակը առկա է, թե բացակա:

Նշանակենք երկարժեք ֆիկտիվ փոփոխականը D -ով և բերենք օրինակներ էմպիրիկ տվյալների դիտարկումների կտրվածքով.

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումը համապատասխանում է կնոջ} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը համապատասխանում է տղամարդու} \end{cases}$$

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումն ունի բարձրագույն կրթություն} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը չունի բարձրագույն կրթություն} \end{cases}$$

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումը } \text{ԵՄ երկիր է} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը } \text{ԵՄ երկիր չէ} \end{cases}$$

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումը տվյալ ճյուղին պատկանող} \\ & \text{ձեռնարկություն է} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը տվյալ ճյուղին պատկանող} \\ & \text{ձեռնարկություն չէ} \end{cases}$$

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումը տվյալ խմբին պատկանող} \\ & \text{ապրանքատեսակ է} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը տվյալ խմբին պատկանող} \\ & \text{ապրանքատեսակ չէ} \end{cases}$$

Ընդհանուր դեպքում

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումն ունի տվյալ որակական հատկանիշը} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը չունի տվյալ որակական հատկանիշը} \end{cases}$$

Նշենք, որ կամայական m որակական կատեգորիա ունեցող փոփոխականն էկոնոմետրիկ մոդելում կարելի է ներկայացնել $m-1$ երկարժեք ֆիկտիվ փոփոխականների միջոցով՝ մոդելում լիակատար կոլլինեարություն չառաջացնելու նկատառմամբ:

Ներդիր 16. Ֆիկտիվ փոփոխականների «թակարդը»

Ռեգրեսիայի մոդելում m թվով որակական կատեգորիա ունեցող փոփոխականի ներկայացումը m երկարժեք ֆիկտիվ փոփոխականների միջոցով կոչվում է երկարժեք ֆիկտիվ փոփոխականի թակարդ (dummy trap): Օրինակ՝ վերը նշված կրթամակարդակի փոփոխականը, որն ունի 6 կատեգորիա, կարելի է էկոնոմետրիկ մոդելում ներկայացնել $m-1=5$ երկարժեք ֆիկտիվ փոփոխականի միջոցով:

Դիտարկենք սա օրինակի վրա: Ենթադրենք բացատրող որակական փոփոխականն է սպառողի բնակության վայրը՝ հետևյալ տարանջատմամբ՝ Երևան, այլ քաղաքներ, գյուղական բնակա-

վայրեր: Այս փոփոխականը մոդելում ներառելու համար դրա նշված երեք կատեգորիային կարող ենք վերագրել արժեքներ՝

$$x_i = \begin{cases} 1, \text{ եթե } i - \text{րդ սպառողը բնակվում է Երևանում} \\ 2, \text{ եթե } i - \text{րդ սպառողը բնակվում է ՀՀ այլ քաղաքում} \\ 3, \text{ եթե } i - \text{րդ սպառողը բնակվում է գյուղում} \end{cases}$$

Այս փոփոխականը կարելի է ներկայացնել մաս երկարժեք ֆիկտիվ փոփոխականների տեսքով՝ ստեղծելով դրա որակական կատեգորիաներին համապատասխանող երեք երկարժեք ֆիկտիվ փոփոխական՝

$$x_{i1} = \begin{cases} 1, \text{ եթե } i - \text{րդ սպառողը բնակվում է Երևանում} \\ 0, \text{ եթե } i - \text{րդ սպառողը չի բնակվում Երևանում} \end{cases}$$

$$x_{i2} = \begin{cases} 1, \text{ եթե } i - \text{րդ սպառողը բնակվում է ՀՀ այլ քաղաքում} \\ 0, \text{ եթե } i - \text{րդ սպառողը չի բնակվում ՀՀ այլ քաղաքում} \end{cases}$$

$$x_{i3} = \begin{cases} 1, \text{ եթե } i - \text{րդ սպառողը բնակվում է գյուղում} \\ 0, \text{ եթե } i - \text{րդ սպառողը չի բնակվում գյուղում} \end{cases}$$

Դիցուք ունենք չորս սպառողի մասին տվյալներ, որոնցից առաջին երկուսը բնակվում են ՀՀ այլ քաղաքներում, երրորդը՝ Երևանում, իսկ չորրորդը՝ գյուղում: Եթե մոդելի մեջ մտցվեն այս երեք երկարժեք փոփոխականներն էլ՝ որպես բացատրող փոփոխականներ, ապա $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$ մոդելի X մատրիքը կունենա հետևյալ տեսքը.

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ 1 & x_{31} & x_{32} & x_{33} \\ 1 & x_{41} & x_{42} & x_{43} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}:$$

Երեք որակական կատեգորիա ունեցող փոփոխականը մոդելում երեք երկարժեք փոփոխականի միջոցով ներկայացնելու դեպքում բացատրող փոփոխականների X մատրիցի երեք սյունների գումարը հավասար է լինում 1-ի (մատրիցի բոլոր տողերում), ինչը հավասար է X մատրիցի առաջին սյանը: Առաջանում է լիակատար կոլինեարություն: Այդ պատճառով բավարար է մոդելում մտցնել միայն X_1 -ն և X_2 -ը: Եթե դրանք երկուսն էլ միաժամանակ հավասար լինեն զրոյի, կստացվի, որ սպառողը բնակվում է գյուղում, և այդպիսով երրորդ կատեգորիայի՝ մոդելից բացակայող դեպքը ևս անուղղակիորեն կարտահայտվի մոդելում:

Այս գլխում կդիտարկենք միայն երկարժեք ֆիկտիվ փոփոխականների օգտագործումը էկոնոմետրիկ մոդելներում որպես բացատրող ու կախյալ փոփոխականներ, ինչպես նաև հեշտության համար այսուհետ երկարժեք ֆիկտիվ փոփոխականը կանվանենք dummy:

1. Բացատրող երկարժեք ֆիկտիվ փոփոխականներով մոդելներ

Դիտարկենք պարզագույն մոդել՝ պարզ գծային ռեգրեսիայի մոդել, որում միակ բացատրող փոփոխականը dummy է: Մոդելն ունի հետևյալ տեսքը՝

$$y_i = \beta_0 + \beta_1 D_i + \varepsilon_i: \quad (117)$$

Ենթադրենք հետազոտվում է որևէ ֆիրմայի աշխատողների ամսական աշխատավարձի կախումը նրանց սեռից: Այդ դեպքում y -ը ֆիրմայի աշխատողների ամսական աշխատավարձն է, իսկ D -ն dummy է.

$$D_i = \begin{cases} 1, & \text{եթե } i - \text{րդ աշխատողը տղամարդ է} \\ 0, & \text{եթե } i - \text{րդ աշխատողը կին է} \end{cases}$$

(117)-ը կարող ենք գնահատել ՓՋԵ-ով, և dummy-ի առկայությունը չի վնասի Գաուս-Մարկովի թեորեմի պայմանների տեղի ունենալուն:

Եթե $D=0$, ֆիրմայի աշխատողը կին է, ապա նրա միջին աշխատավարձը հավասար կլինի $\widehat{\beta}_0$ -ի, քանի որ՝

$$\hat{y} = E(y_{\text{ս}} | x = x_{\text{ս}}) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot 0 = \widehat{\beta}_0:$$

Գնահատված $\widehat{\beta}_0$ գործակիցը ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե $D=0$, այսինքն՝ երբ դիտարկվող որակական հատկանիշը բացակայում է:

Եթե $D=1$, ֆիրմայի աշխատողը տղամարդ է, ապա նրա միջին աշխատավարձը հավասար կլինի $\widehat{\beta}_0 + \widehat{\beta}_1$, քանի որ՝

$$\hat{y} = E(y_{\text{ս}} | x = x_{\text{ս}}) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot 1 = \widehat{\beta}_0 + \widehat{\beta}_1$$

$\widehat{\beta}_1$ գործակիցը ցույց է տալիս, թե որքան է տղամարդկանց միջին աշխատավարձի տարբերությունը կանանց համեմատ: Կամ, որ նույնն է, $\widehat{\beta}_1$ գործակիցը ցույց է տալիս դիտարկվող հատկանիշի առկայության դեպքում ($D=1$) կախյալ փոփոխականի միջին արժեքի

տարբերությունը այդ հատկանիշի բացակայության ($D=0$) համեմատ:

Բնական է դիտարկել այն տարբերակը, երբ մոդելում կա քանակական բացատրող փոփոխական: Պարզագույն դեպքում այն մեկն է: Դիտարկենք նույն մոդելը, որում ներառված է x քանակական բացատրող փոփոխական՝ աշխատանքային ստաժը ամիսներով: Մոդելն ունի հետևյալ տեսքը՝

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 D_i + \varepsilon_i: \quad (118)$$

Այս մոդելում, եթե $D=0$, կամ, որ նույնն է, ֆիրմայում աշխատող կանանց համար՝

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x + \widehat{\beta}_2 \cdot 0 = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x:$$

$\widehat{\beta}_0$ գործակիցը ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե դիտարկվող որակական հատկանիշը բացակայում է, և $x=0$: $\widehat{\beta}_1$ -ը ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով: Քանի որ գնահատվել է ուղիղ գիծ, եթե $D=0$, գծի բարձրությունը հավասար է $\widehat{\beta}_0$ -ի, իսկ թեքությունը՝ $\widehat{\beta}_1$ -ի:

Եթե $D=1$ (ընկերությունում աշխատող տղամարդկանց համար)՝

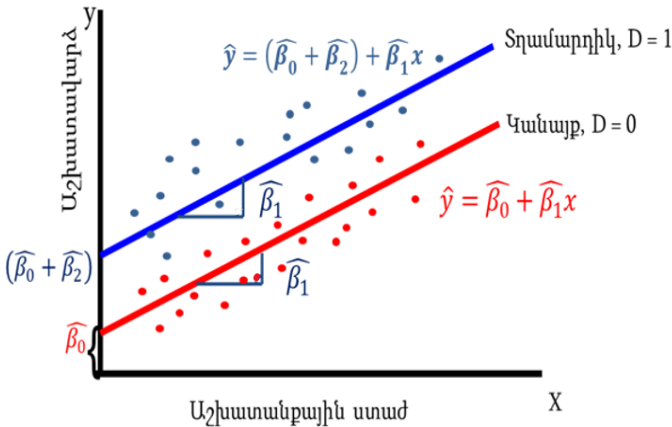
$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x + \widehat{\beta}_2 \cdot 1 = (\widehat{\beta}_0 + \widehat{\beta}_2) + \widehat{\beta}_1 \cdot x:$$

Այս դեպքում $\widehat{\beta}_0 + \widehat{\beta}_2$ -ն ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե դիտարկվող որակական հատկանիշը առկա է, և $x=0$: $\widehat{\beta}_1$ -ը նորից ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով: $D=1$ դեպքում գնահատվել է երկրորդ ուղիղ գիծը, որի բարձրությունը հավասար է $\widehat{\beta}_0 + \widehat{\beta}_2$ -ի, իսկ թեքությունը նույնն է, ինչ $D=0$ դեպքում, և հավասար է $\widehat{\beta}_1$ -ի: Գնահատվել են երկու գույքահեռ ուղիղներ, որոնք ունեն տարբեր բարձրություններ, և այդ բարձրությունների տարբերությունը $\widehat{\beta}_2$ -ն է: Այդ պատճառով (118) տեսքի մոդելներում dummy-ն կոչվում է **բարձրության dummy**: $\widehat{\beta}_2$ -ը նույն աշխատանքային ստաժն ունեցող կանանց ու տղամարդկանց միջին ամսական աշխատավարձերի տարբերությունն է:

Ընդհանուր դեպքում $\widehat{\beta}_2$ -ը ցույց է տալիս, թե x -ի միևնույն արժեքի համար որքան է կախյալ փոփոխականի միջինների տարբերությունը որակական հատկանիշի առկայության ու բացակայության

դեպքերի միջև: Եթե $\widehat{\beta}_2$ -ը վիճակագրորեն նշանակալի է, ապա այդ տարբերությունը ևս վիճակագրորեն նշանակալի է:

Գծանկար 20. Բարձրության երկարժեք ֆիկտիվ բացարձող փոփոխականով զծային ռեգրեսիայի էմպիրիկ հավասարումները



Հաճախ որակական հատկանիշի առկայությունը ուժեղացնում կամ թուլացնում է կախյալ փոփոխականի վրա քանակական բացարձող փոփոխականի ունեցած ազդեցությունը: Օրինակ՝ կարելի է ենթադրել, որ աշխատանքային ստաժի ազդեցությունն աշխատավարձի վրա կանանց ու տղամարդկանց պարագայում տարբեր է: Այդպիսի դեպքերում dummy-ի օգտագործմամբ գնահատվում են հետևյալ տեսքի մոդելներ՝

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 D_i x_i + \varepsilon_i: \quad (119)$$

Մոդելում, եթե $D=0$, այսինքն՝ ընկերությունում աշխատող կանանց համար՝

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x + \widehat{\beta}_2 \cdot 0 \cdot x = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x:$$

Այս դեպքում $\widehat{\beta}_0$ գործակիցը ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե $x=0$: $\widehat{\beta}_1$ -ը ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով որակական հատկանիշի բացա-

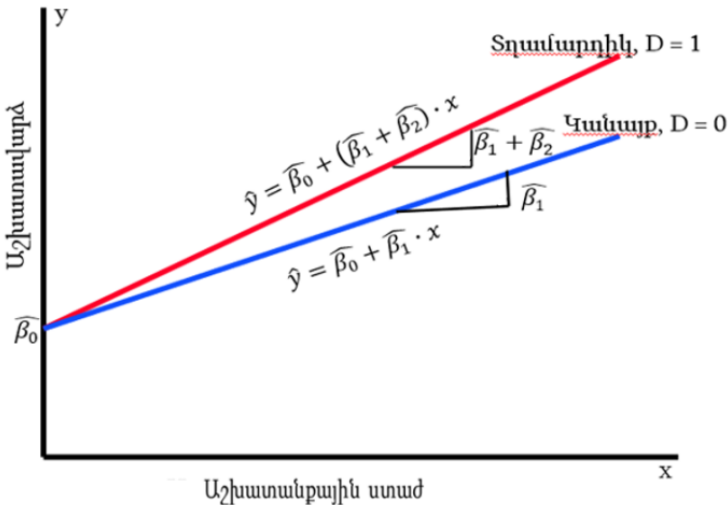
կայության դեպքում, այսինքն՝ եթե $D=0$ (կանանց համար) գծի բարձրությունը հավասար է $\widehat{\beta}_0$ -ի, իսկ թեքությունը՝ $\widehat{\beta}_1$ -ի:

Եթե $D=1$ (տղամարդկանց համար).

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x + \widehat{\beta}_2 \cdot 1 \cdot x = \widehat{\beta}_0 + (\widehat{\beta}_1 + \widehat{\beta}_2) \cdot x:$$

Այս դեպքում $\widehat{\beta}_0$ -ն նորից ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե $x=0$: $\widehat{\beta}_1 + \widehat{\beta}_2$ -ը ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով որակական հատկանիշի առկայության դեպքում: $D=1$ դեպքում գնահատվել է երկրորդ ուղիղ գիծը, որի բարձրությունը նույնն է, ինչ $D=0$ դեպքում, և հավասար է $\widehat{\beta}_0$ -ի, իսկ թեքությունը տարբերվում է $D=0$ դեպքից և հավասար է $\widehat{\beta}_1 + \widehat{\beta}_2$ -ի: Գնահատվել են երկու ուղիղներ, որոնք ունեն նույն բարձրությունն ու տարբեր թեքություններ: Դրանց թեքությունների տարբերությունը $\widehat{\beta}_2$ -ն է: Այդ պատճառով (119) տեսքի մոդելներում dummy-ն կոչվում է **բերության dummy**: $\widehat{\beta}_2$ -ը աշխատանքային ստաժի միավոր փոփոխությունն ունեցող կանանց ու տղամարդկանց միջին ամսական աշխատավարձերի փոփոխության տարբերությունն է:

Գծանկար 21. Թեքության երկարժեք ֆիկտիվ բացարձող փոփոխականով գծային ռեգրեսիայի էմպիրիկ հավասարումները



Ընդհանուր դեպքում $\widehat{\beta}_2$ -ը ցույց է տալիս, թե x -ի միավոր փոփոխության դեպքում որքան է y -ի փոփոխության տարբերությունը որակական հատկանիշի առկայության ու բացակայության դեպքերի միջև: Եթե $\widehat{\beta}_2$ -ը վիճակագրորեն նշանակալի է, ապա այդ տարբերությունը ևս վիճակագրորեն նշանակալի է:

Որոշ դեպքերում որակական հատկանիշի ազդեցությունը կախյալ փոփոխականի վրա անհրաժեշտ է լինում դիտարկել թե՛ առանձին, թե՛ քանակական բացատրող փոփոխականի հետ ունեցած փոխազդեցությամբ: Այդպիսի դեպքերում dummy-ի օգտագործմամբ գնահատվում են հետևյալ տեսքի մոդելներ՝

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 D_i + \beta_3 D_i x_i + \varepsilon_i: \quad (120)$$

Մոդելում, եթե $D=0$, այսինքն՝ ֆիրմայում աշխատող կանանց համար՝

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 \cdot 0 + \widehat{\beta}_3 \cdot 0 \cdot x = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x:$$

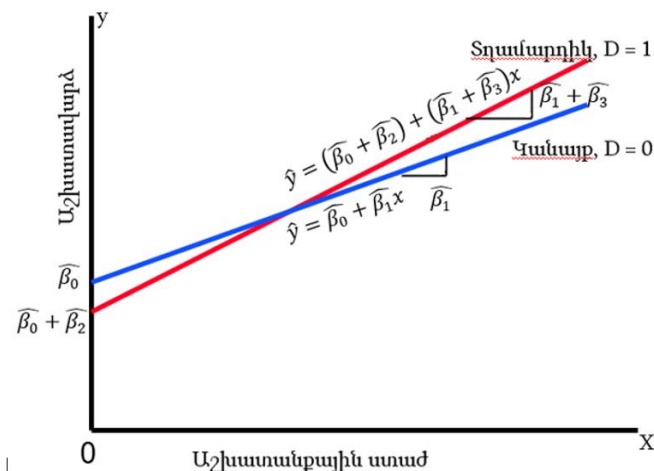
Այս դեպքում $\widehat{\beta}_0$ գործակիցը ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե $x=0$ և որակական հատկանիշը բացակայում է: $\widehat{\beta}_1$ -ը ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով՝ որակական հատկանիշի բացակայության դեպքում, այսինքն՝ եթե $D=0$ (կանանց համար) գծի բարձրությունը հավասար է $\widehat{\beta}_0$ -ի, իսկ թեքությունը՝ $\widehat{\beta}_1$ -ի:

Եթե $D=1$, այսինքն՝ ընկերությունում աշխատող տղամարդկանց համար՝

$$\hat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 \cdot 1 + \widehat{\beta}_3 \cdot 1 \cdot x = (\widehat{\beta}_0 + \widehat{\beta}_2) + (\widehat{\beta}_1 + \widehat{\beta}_3) \cdot x:$$

Այստեղ $\widehat{\beta}_0 + \widehat{\beta}_2$ -ը ցույց է տալիս, թե ինչի հավասար կլինի կախյալ փոփոխականը միջինում, եթե $x=0$, որակական հատկանիշի առկայության դեպքում: $\widehat{\beta}_1 + \widehat{\beta}_3$ -ը ցույց է տալիս, թե միջինում քանի միավորով է փոխվում կախյալ փոփոխականը, եթե x -ը փոխվում է մեկ միավորով՝ որակական հատկանիշի առկայության դեպքում: $D=1$ դեպքում գնահատվել է երկրորդ ուղիղ գիծը, որի թե՛ բարձրությունը, թե՛ թեքությունը տարբերվում են $D=0$ դեպքից: Դրանց բարձրությունների տարբերությունը $\widehat{\beta}_2$ -ն է, իսկ թեքությունների տարբերությունը $\widehat{\beta}_3$ -ն է: Այդ պատճառով (120) տեսքի մոդելներում dummy-ն կոչվում է **թեքության և բարձրության dummy**:

Գծանկար 22. Թեքության ու բարձրության երկարժեք ֆիկտիվ բացա-
 րող փոփոխականով գծային ռեգրեսիայի էմպիրիկ հավասարումները



Ընդհանուր դեպքում $\widehat{\beta}_2$ -ը ցույց է տալիս, թե որքան է կախյալ փոփոխականի միջինների տարբերությունը որակական հատկանիշի առկայության ու բացակայության դեպքերի միջև, եթե $x=0$: $\widehat{\beta}_3$ -ը ցույց է տալիս, թե x -ի միավոր փոփոխության դեպքում որքան է կախյալ փոփոխականի փոփոխության տարբերությունը որակական հատկանիշի առկայության ու բացակայության դեպքերի միջև: Եթե $\widehat{\beta}_2$ -ը և $\widehat{\beta}_3$ -ը վիճակագրորեն նշանակալի են, ապա այդ տարբերությունները ևս վիճակագրորեն նշանակալի են:

2. Կախյալ երկարժեք ֆիկտիվ փոփոխականով մոդելներ. գծային հավանականական ու լոգիթ մոդելներ

Հաճախ հետազոտողին անհրաժեշտ է լինում գնահատել այն-
 պիսի կախվածություններ, որոնցում կախյալ փոփոխականն է որա-
 կական: Մասնավորապես, այն երկարժեք ֆիկտիվ փոփոխական է,
 որը բնութագրում է հետազոտողին հետաքրքրող որևէ հատկանիշի
 առկայություն կամ բացակայություն: Այդպիսի կախվածությունները
 նկարագրող էկոնոմետրիկ մոդելները կոչվում են **երկարժեք կամ բի-**

նար ընտրության մոդելներ: Օրինակ՝ էկոնոմետրիկ հետազոտության նպատակն է պարզել, թե՝

- տվյալ եկամուտը, ընտանիքի կազմը, տարիքն ու սեռն ունեցող անհատը կգնի՞ մեզ հետաքրքրող ապրանքը,
- տվյալ առաջադիմությունը, եկամուտը, ընտանիքի կազմը, տարիքը ունեցող բակալավրիատի ուսանողը կշարունակի՞ ուսումը մագիստրոսական ծրագրում,
- տվյալ եկամուտը, ընտանիքի կազմը, տարիքն ու սեռն ունեցող անհատը իր ձայնը կտա՞ տվյալ թեկնածուին,
- տվյալ թողարկման ծավալը, հիմնական ու շրջանառու ֆոնդերն ունեցող ֆիրման կկատարի՞ ներդրումներ տեխնիկական վերազինման վրա,
- տվյալ ՀՆԱ մակարդակը, ինստիտուտների որակը, ՀՆԱ-ի նկատմամբ առողջապահության ու կրթության ոլորտներում պետական ծախսերի մակարդակն ունեցող երկիրը կմտնի՞ բարձր մարդկային զարգացում ունեցող երկրների խումբ:

Այսպիսի կախվածությունների գնահատումը ենթադրում է ընդհանուր ու ամենապարզ ձևակերպմամբ հետևյալ տեսքի էկոնոմետրիկ մոդելների հատկորոշում՝

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad (121)$$

$$y_i = \begin{cases} 1, & \text{եթե } i - \text{րդ դիտարկումն ունի տվյալ հատկանիշը} \\ 0, & \text{եթե } i - \text{րդ դիտարկումը չունի տվյալ հատկանիշը} \end{cases}$$

(121)-ը կոչվում է **գծային հավանականական մոդել (Linear Probability Model, LPM):**

Եթե (121)-ը գնահատվի ՓԶԵ-ով, ապա կգնահատվի կախյալ փոփոխականի 1 արժեք ընդունելու հավանականությունը կամ, որ նույնն է, տվյալ հատկանիշն ունենալու հավանականությունը: Իսկապես, ենթադրենք յուրաքանչյուր դիտարկման համար $y=1$ արժեքը կախյալ փոփոխականն ընդունում է P_i հավանականությամբ: Այդ դեպքում $y=0$ արժեքը կընդունի $1-P_i$ հավանականությամբ: Կախյալ փոփոխականի սպասվող արժեքն այդ դեպքում կլինի հետևյալը.

$$\hat{y}_i = \mathbb{E}(y_{i\text{us}} | x = x_{i\text{us}}) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_i = 1 \cdot P_i + 0 \cdot (1 - P_i) = P_i:$$

Հետևաբար գծային հավանականական մոդելով գնահատվում է կախյալ փոփոխականի 1 արժեքն ընդունելու հավանականությունը, որը որոշվում է հետևյալ կերպ՝

$$\hat{y}_i = P_i(y_i = 1) = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_i: \quad (122)$$

(121)-ի ՓԲԵ-ով գնահատված հավասարման մեջ $\widehat{\beta}_0$ -ն $x=0$ դեպքում կախյալ փոփոխականի 1 արժեք ընդունելու հավանականությունն է, իսկ $\widehat{\beta}_1$ -ը ցույց է տալիս, թե միջինում քանի միավորով կփոխվի $y=1$ արժեքը ընդունելու հավանականությունը, եթե x -ը փոխվի մեկ միավորով:

Սակայն գծային հավանականական մոդելի գնահատումը ՓԲԵ-ով խնդրահարույց է և վստահելի չէ, քանի որ խախտվում են Գաուս-Մարկովի պայմանները, և ստացված գնահատականները լավագույն գծային գնահատականներ չեն լինում:

Ի՞նչ խնդիրներ են առաջանում գծային հավանականական մոդելը ՓԲԵ-ով գնահատելիս.

1. Մոդելի պատահական շեղումները նորմալ բաշխում չունեն: Դրանք յուրաքանչյուր դիտարկման համար կարող են ընդունել երկու արժեք: Դրա մասին են վկայում տեսական մոդելի վիճակագրական գնահատականները՝ պատահական մնացորդները.

$$\hat{y}_i = P_i = 1 \quad \rightarrow \quad \varepsilon_i = y_i - \hat{y}_i = 1 - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

$$\hat{y}_i = P_i = 0 \quad \rightarrow \quad \varepsilon_i = y_i - \hat{y}_i = -\hat{\beta}_0 - \hat{\beta}_1 x_i:$$

2. Մոդելի պատահական շեղումները հետերոսկեդաստիկ են: Կարելի է ցույց տալ, որ $\text{var}(\varepsilon_i) = P_i \cdot (1 - P_i)$: Քանի որ P -ն կախված է y -ից, որն էլ կախված է x -ից, ուրեմն պատահական շեղումների վարիացիան կախված է բացատրող փոփոխականից, այսինքն՝ հաստատուն չէ:

3. Ելնելով x , y կոռելյացիայի դաշտից՝ այսպիսի մոդելն ունի վոքըր դետերմինացիայի գործակից:

4. Քանի որ գնահատվում է ուղիղ գիծ, ապա կախյալ փոփոխականի գնահատված $\hat{y}_i$ արժեքները (ինչպես ցույց է տրվել, դրանք $y=1$ դեպքի հավանականությունների գնահատված արժեքներն են), կարող են գրոշից վոքըր ու մեկից մեծ թվեր լինել, ինչը հակասում է հավանականության գաղափարին:

Արդյունքում ՓԲԵ-ով գծային հավանականական մոդելի գհահատմամբ ստացված $\hat{\beta}$ գործակիցները անշեղելի են, բայց ոչ արդյունավետ: Այս հանգամանքները խիստ նեղացնում են գծային հավանականական մոդելի կիրառման շրջանակը: Այն նպատակահարմար է կիրառել մեծ նմուշների դեպքում և հիմնականում որպես տվյալներ-

րի նախնական մշակման գործիք՝ հետազայում այլ մեթոդներով ստացված արդյունքների հետ համեմատելու համար:

Ինչպես արդեն նշվել է, պատահական շեղումների նորմալ բաշխումը չի խանգարում ՓԶԵ-ով BLUE գնահատականներ ստանալուն, եթե Գաուս-Մարկովի մնացած պայմանները տեղի ունեն, իսկ հետերոսկեդաստիկությունը կարելի է շտկել՝ կշռված կամ ընդհանրացված ՓԶԵ կիրառելով: Սակայն $\hat{y}$ -ի՝ որպես հավանականության, արժեքների $[0, 1]$ միջակայքին պատկանելու խնդիրը մնում է չլուծված:

Նշված խնդրի վերացման համար երկարժեք ֆիկտիվ կախյալ փոփոխականով կախվածությունները նպատակահարմար է մոտարկել ոչ թե ուղիղ գծով, այլ այնպիսի կորով, որն ընկած է $[0, 1]$ միջակայքում: Նման կորով է նկարագրվում, մասնավորապես, լոգիստիկ բաշխումը: Լոգիստիկ բաշխումը նման է նորմալ բաշխմանը, բայց ունի ավելի սուր գագաթ և ավելի «ծանր» պոչեր: Լոգիստիկ բաշխման ֆունկցիան (կամ լոգիստիկ ֆունկցիան) ունի հետևյալ տեսքը՝ $F(X) = \frac{1}{1+e^{-\frac{x-\mu}{s}}}$, μ միջինով ու $\frac{\pi^2 \cdot s^2}{3}$ վարիացիով: Լոգիստիկ բաշխման կուտակային ֆունկցիայի կորը, այսպես կոչված, «սիգմոիդ» է կամ S տառի ձևը ունեցող կոր (S sharp curve): Այս կորը հարմար է օգտագործել նշված խնդրի լուծման համար: Երկարժեք ընտրության ամենատարածված մոդելներից մեկը, որը կոչվում է լոգիթ մոդել, օգտագործում է լոգիստիկ բաշխումը:

Լոգիթ մոդելում ենթադրվում է, որ մոդելով գնահատված կախյալ փոփոխականի $y=1$ արժեքն ընդունելու P հավանականությունն արտահայտվում է ոչ թե (122) բանաձևով, այլ ենթարկվում է լոգիստիկ բաշխմանը և ունի հետևյալ տեսքը՝

$$\hat{y}_i = \mathbb{P}_i(y_i = 1) = \frac{1}{1+e^{-(\beta_0+\beta_1 \cdot x_i)}} : \quad (123)$$

Ընդհանուր դեպքում՝

$$\hat{y}_i = \mathbb{P}_i(y_i = 1) = \frac{1}{1+e^{-(\hat{\beta}_0+\hat{\beta}_1 \cdot x_{i1}+\hat{\beta}_2 \cdot x_{i2}+...+\hat{\beta}_m \cdot x_{im})}}:$$

Կամ մատրիցային տեսքով՝

$$\hat{Y} = \mathbb{P}(Y = 1) = 1/(1+e^{-XB}), \quad (124)$$

որում $\mathbb{P}$ -ն յուրաքանչյուր դիտարկման $y=1$ արժեք ընդունելու հավանականություններից կազմված n -չափանի վեկտոր է:

(123)-(124)-ից երևում է, որ երբ $x_i \rightarrow +\infty$, $P_i \rightarrow 1$, իսկ երբ $x_i \rightarrow -\infty$, $P_i \rightarrow 0$: Այսինքն՝ այս տեսքի մոդելի գնահատումը լուծում է $\mathbb{P}_i \in [0, 1]$ խնդիրը: Սակայն (123)-(124)-ում կախյալ փոփոխականի մեկ արժեք ընդունելու հավանականությունը ոչ գծային կախվածության մեջ է ոչ միայն բացատրող փոփոխականներից, այլև β պարամետրերից, և ՓԶԵ-ով գնահատումը կիրառելի չէ:

Այդ պատճառով կատարվում է ձևափոխություն.

$$\begin{aligned}\mathbb{P}(Y=1) &= \frac{1}{1+e^{-X\hat{B}}} \rightarrow 1+e^{-X\hat{B}} = \frac{1}{\mathbb{P}} \rightarrow e^{-X\hat{B}} = \frac{1}{\mathbb{P}} - 1 \\ &\rightarrow e^{-X\hat{B}} = \frac{1-\mathbb{P}}{\mathbb{P}} \rightarrow e^{X\hat{B}} = \frac{\mathbb{P}}{1-\mathbb{P}} \rightarrow \ln(e^{X\hat{B}}) = \ln\left(\frac{\mathbb{P}}{1-\mathbb{P}}\right) \\ &\rightarrow X\hat{B} = \ln\left(\frac{\mathbb{P}}{1-\mathbb{P}}\right): \end{aligned}$$

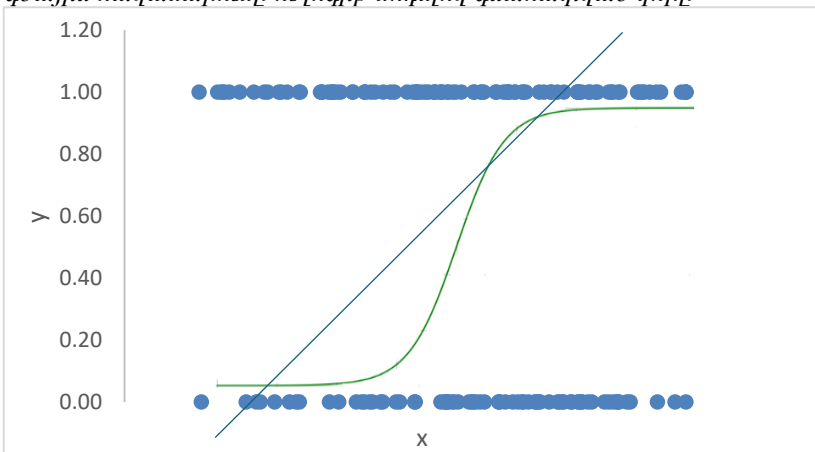
Ստացվեց, որ լոգիթ մոդելը գնահատում է հետևյալ կախվածությունը.

$$\hat{y} = \ln\left(\frac{\mathbb{P}}{1-\mathbb{P}}\right) = X\hat{B} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot x_1 + \widehat{\beta}_2 \cdot x_2 + \dots + \widehat{\beta}_m \cdot x_m: \quad (125)$$

(125)-ում $\frac{\mathbb{P}}{1-\mathbb{P}}$ -ը կոչվում է շանսերի հարաբերակցություն և ցույց է տալիս, քե $y=1$ արժեքն ընդունելու հավանականությունը քանի անգամ է մեծ $y=0$ արժեքն ընդունելու հավանականությունից: (125)-ից երևում է, որ $\hat{\beta}$ գործակիցների մեկնաբանությունն այսպիսին է՝ տվյալ x_j բացատրող փոփոխականի աճը մեկ միավորով, այլ հավասար պայմաններում հանգեցնում է շանսերի հարաբերակցության (կամ $y=0$ արժեքն ընդունելու հավանականության նկատմամբ $y=1$ արժեքն ընդունելու հավանականության) աճի $e^{\hat{\beta}_j}$ անգամ:

Լոգիթ մոդելում $\ln\left(\frac{\mathbb{P}}{1-\mathbb{P}}\right)$ -ն կոչվում է լոգիթ և հավասար է շանսերի հարաբերակցության բնական հիմքով լոգարիթմին: (125) լոգիթ մոդելի գնահատումը շանսերի հարաբերակցության լոգարիթմի կամ լոգիթի գնահատումն է: Լոգիթի կախումը բացատրող փոփոխականներից ու պարամետրերից գծային է: Ըստ որում, քանի որ $\mathbb{P}_i \in [0, 1]$, ապա $\ln\left(\frac{\mathbb{P}}{1-\mathbb{P}}\right) \in (-\infty, +\infty)$:

Գծանկար 23. Երկարժեք կախյալ փոփոխականի ու քանակական բացատրող փոփոխականի կոռելյացիայի դաշտը, ՓԶԵ-ով գնահատված զծային հավասարումը ու լոգիթ մոդելով գնահատված կորը



Այնուամենայնիվ, (125) լոգիթ մոդելը ևս ՓԶԵ-ով գնահատել հնարավոր չէ, քանի որ դրա համար անհրաժեշտ է ունենալ $\frac{\mathbb{P}}{1-\mathbb{P}}$ շանսերի հարաբերակցության համար էմպիրիկ տվյալներ: Այս մոդելը գնահատվում է առավելագույն ճշմարտամամուրթյան եղանակով:

Դիտարկենք օրինակ: Ենթադրենք բանկի վարկային մասնագետը խնդիր ունի բացահայտելու, թե որ գործոններն են ազդում անհատ վարկառուների կողմից սպառողական վարկի ձախողման վրա կամ, որ նույնն է, վարկային ռիսկի վրա: Այդ նպատակով նա որոշակի տեղեկատվություն է հավաքել նախկինում վարկ վերցրած 700 անհատների մասին: Դրանք են՝ վարկը ձախողելու փաստը (երկարժեք ֆիկտիվ փոփոխական, $D=0$, եթե տվյալ անձը նախկինում վերցրած վարկը չի ձախողել, և $D=1$, եթե տվյալ անձը նախկինում վերցրած վարկը ձախողել է), վարկառուի տարիքը (Age), անձի վերցրած վարկի գումարի հարաբերակցությունը նրա տարեկան եկամտի նկատմամբ (debtinc, %), անձի աշխատանքային եկամուտների կայունությունը՝ տվյալ աշխատավայրում անձի աշխատելու տևողությունը (employment, ամիսներ): Վարկային մասնագետն ուզում է գնահատել վարկի ձախողման հավանականությունը՝ կախված Age, debtinc ու employment փոփոխականներից: Դիտարկվում է երկարժեք կախյալ փոփոխականի (D) կախումը նշված երեք բացատրող փոփոխականներից: Գնահատվում է լոգիթ մոդել՝ SPSS

փաթեթի համապատասխան գործիքով: Գնահատվող մոդելը հետևյալն է՝

$$\hat{y} = \ln\left(\frac{\mathbb{P}}{1 - \mathbb{P}}\right) = X\hat{B} = \hat{\beta}_0 + \hat{\beta}_1 \cdot Age + \hat{\beta}_2 \cdot deptinc + \hat{\beta}_m \cdot employment:$$

Գնահատումն իրականացվում է առավելագույն ճշմարտանմանության եղանակով, որի նպատակը ճշմարտանմանության ֆունկցիայի մաքսիմալացումն է: Ալգորիթմը առաջին փուլում գնահատում է բացատրող մասում միայն ազատ անդամի առկայությամբ մոդել: Հաջորդ փուլում մոդելի մեջ են մտցվում մաս բացատրող փոփոխականները՝ միանգամից կամ հերթականությամբ: Այս փուլերի համար հաշվարկվում է ճշմարտանմանության ֆունկցիայի լոգարիթմի կրկնակին՝ բացասական նշանով (-2Log liklihood), որպես նպատակային ֆունկցիա: Լավագույնը այն մոդելն է, որի համար -2Log liklihood-ի արժեքն ավելի փոքր է: Այդ նկատառումով հաշվարկվում է երկու հաջորդական մոդելների նպատակային ֆունկցիաների տարբերությունը, որն ունի ասիմպոտիկ χ^2 բաշխում և կոչվում է Chi-square վիճականի: Դրվում է գրոյական վարկած այն մասին, որ բացատրող փոփոխականների ամբողջությունը վիճակագրորեն նշանակալի ազդեցություն չունի կախյալ փոփոխականի վրա: Sig-ը (p արժեք) այդ վարկածը մերժելիս սխալված լինելու հավանականությունն է: Եթե այն փոքր է $\alpha=0.05$ -ից, ապա գրոյական վարկածը պետք է մերժել, և անկախ փոփոխականները՝ որպես խումբ, վիճակագրորեն նշանակալի ազդեցություն ունեն կախյալ փոփոխականի վրա: Գնահատած մոդելի համար այս վիճականին է (Omnibus Tests of Model Coefficients աղյուսակից)՝

Chi-square	Sig.
188.911	0.000

Սա նշանակում է, որ բացատրող փոփոխականների ավելացումը մոդելում փոքրացրել է նպատակային ֆունկցիան 188.9 միավորով, և այդ փոքրացումը վիճակագրորեն նշանակալի է, այսինքն՝ անկախ փոփոխականները՝ որպես խումբ, վիճակագրորեն նշանակալի ազդեցություն ունեն կախյալ փոփոխականի բացատրության վրա:

Գնահատման որակի՝ գծային ռեգրեսիայի մոդելի դեպքում կիրառվող ստանդարտ չափանիշը՝ դետերմինացիայի գործակիցը, այս դեպքում հնարավոր չէ հաշվարկել: Այդ պատճառով հաշվարկվում են դրան մոտարկող չափանիշներ: Դրանք են՝ ա) Կոքսի և Սնելի կեղծ R^2 -ն, որն ընկած է $[0,1)$ միջակայքում, սակայն մույնիսկ «կա-

տարյալ» մոդելի համար այն ունի 1-ից փոքր առավելագույն արժեք, ք) Նազելկերկեի կեղծ R^2 -ն, որը Կոքսի ու Սնելի R^2 -ի ճշգրտված տարբերակն է և ներառում է 0-ից մինչև 1 ամբողջ միջակայքը: Ուստի նախընտրելի է որպես գնահատման որակի չափանիշ դիտարկել Նազելկերկեի R^2 -ն: Թեև կեղծ R^2 -ն կարող է որոշակիորեն նկարագրական լինել, սակայն այն առավել օգտակար է մրցակցող մոդելները համեմատելիս: Այս չափանիշի համաձայն՝ «լավագույնը» ամենամեծ R^2 ունեցող մոդելն է:

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	615.454 ^a	0.237	0.346

Գնահատված մոդելի համար ունենք Նազելկերկեի $R^2=34.6$: Սա նշանակում է, որ առաջին փուլում գնահատված մոդելի համեմատելիորդ փուլի մոդելը, որի մեջ մտցված են բացատրող փոփոխականները, բարելավել է մոդելի գնահատման արդյունքների ճշմարտամանությունը:

Լոգիթ մոդելի գնահատման որակի ավելի բնութագրական ցուցանիշ է Հոսմեր-Լեմեշի թեստը: Այս թեստը գնահատում է, թե որքանով են Լոգիթ մոդելի կիրառման արդյունքում գնահատված հավանականությունները համապատասխանում կախյալ փոփոխականի իրական արժեքներին: Դրվում է H_0 զրոյական վարկած այն մասին, որ Լոգիթ մոդելը բավարար կերպով համապատասխանում է տվյալներին, և գնահատված հավանականությունների ու կախյալ փոփոխականի իրական արժեքների միջև տարբերությունը վիճակագրորեն նշանակալի չէ: Մրցակցող վարկածը հետևյալն է. մոդելը բավարար չափով չի համապատասխանում տվյալներին, այսինքն՝ կանխատեսված հավանականությունների տարբերությունը կախյալ փոփոխականի դիտարկված արժեքներից վիճակագրորեն նշանակալի է: Թեստի վիճականին Հոսմեր-Լեմեշի վիճականին է, որն ունի χ^2 բաշխում: Թեստի ելքը կարելի է որոշել p արժեքի միջոցով: Եթե p արժեքը $> 0.05 \rightarrow H_0$ -ն չենք կարող մերժել, այն ընդունվում է, այսինքն՝ գնահատման արդյունքները բավարար կերպով համապատասխանում են տվյալներին, մոդելը որակյալ է: Եթե p արժեքը $\leq 0.05 \rightarrow H_0$ -ն մերժվում է, գնահատման որակը բավարար չէ:

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.183	8	0.840

Գնահատված մոդելի համար p արժեք= $\text{Sig.}=0.840$: Եզրակացությունը հետևյալն է. H_0 -ն չենք կարող մերժել, այն ընդունվում է, այսինքն՝ գնահատման արդյունքները բավարար կերպով համապատասխանում են տվյալներին, մոդելը որակյալ է:

Գնահատված հավասարումն է՝

$$\hat{y} = \ln\left(\frac{\mathbb{P}}{1 - \mathbb{P}}\right) = -1.725 + 0.001 \cdot \text{Age} + 0.145 \cdot \text{deptinc} - 0.142 \cdot \text{employment}:$$

Գնահատված գործակիցները երևում են ստորև բերված Աղյուսակ 10-ի B սյան մեջ: S.E. սյան ցուցանիշները գնահատված գործակիցների ստանդարտ սխալներն են: Wald սյան մեջ ներկայացված է Վալդի վիճականին, որը մաթեմատիկորեն Ստյուդենտի t վիճականու քառակուսին է, սակայն այն ունի ասիմպտոտիկ χ^2 բաշխում մեկ ազատության աստիճանով: Վալդի վիճականին օգտագործվում է տվյալ գնահատված գործակցի վիճակագրական նշանակալիությունը ստուգելու համար: Դրվում է H_0 զրոյական վարկած այն մասին, որ տվյալ անկախ փոփոխականն իրականում չի բացատրում կախյալին, այսինքն՝ տեսական մոդելում տվյալ β պարամետրը հավասար է զրոյի: Հաշվարկվում է Վալդի վիճականու դիտարկված արժեքը հետևյալ բանաձևով՝ $Wald_{\beta_j} = \left(\frac{\hat{\beta}_j}{S_{\hat{\beta}_j}}\right)^2$, և համեմատվում χ^2

բաշխման կրիտիկական արժեքի հետ: Եզրակացության գալու համար կարելի է օգտագործել p արժեքը (Sig.): Եթե $\text{Sig.} > 0.05 \rightarrow H_0$ -ն մերժվում է, հակառակ դեպքում՝ ընդունվում:

Աղյուսակ 10. SPSS փաթեթում Lnqhp մոդելի գնահատված գործակիցների ու դրանց բնութագրիչների էլքային աղյուսակը

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	<i>deptinc</i>	0.145	0.016	82.129	1	0.000	1.156
	<i>Age</i>	0.001	0.014	0.005	1	0.939	1.001
	<i>employment</i>	-0.142	0.021	45.723	1	0.000	0.868
	Constant	-1.725	0.463	13.880	1	0.000	0.178

a. Variable(s) entered on step 1: Debt to income ratio (x100), Age in years, Years with current employer.

Տվյալ դեպքում Վալդի վիճականին վկայում է, որ *deptinc* և *employment* փոփոխականների գնահատված գործակիցները վիճակագրորեն նշանակալի են, իսկ Age փոփոխականինը՝ ոչ, այդ փոփոխականը չի բացատրում կախյալին ու պետք է հանել մոդելից:

Ինչպե՞ս են մեկնաբանվում մոդելի գնահատված գործակիցները: *Deptinc* փոփոխականի միավոր աճն այլ հավասար պայմաններում հանգեցնում է կախյալ փոփոխականի $y=0$ արժեքն ընդունելու հավանականության նկատմամբ $y=1$ արժեքն ընդունելու (վարկը ձախողելու) հավանականության աճ՝ $e^{\beta_1} = 1.156$ անգամ կամ $1.156 \cdot 100 - 100 = 15.6$ տոկոսով: Age փոփոխականի միավոր աճը այլ հավասար պայմաններում հանգեցնում է կախյալ փոփոխականի $y=1$ արժեքն ընդունելու (վարկը ձախողելու) հավանականության աճ՝ $e^{\beta_2} = 1.001$ անգամ կամ $1.001 \cdot 100 - 100 = 0.1$ տոկոսով: *Employment* փոփոխականի միավոր աճը այլ հավասար պայմաններում հանգեցնում է կախյալ փոփոխականի $y=1$ արժեքն ընդունելու (վարկը ձախողելու) հավանականության նվազում՝ $e^{\beta_3} = 0.868$ անգամ կամ $100 - 0.868 \cdot 100 = 13.2$ տոկոսով:

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Ի՞նչ նպատակով են ֆիկտիվ փոփոխականները ներառվում էկոնոմետրիկ մոդելներում: Բերե՛ք օրինակներ:
2. Ի՞նչ է երկարժեք ֆիկտիվ փոփոխականը, ի՞նչ արժեքներ է այն ընդունում և ի՞նչ բովանդակությամբ: Բերե՛ք օրինակներ:
3. Ի՞նչ է ֆիկտիվ փոփոխականների թակարդը, և ինչպե՞ս պետք է դրանից խուսափել:
4. Նկարագրե՛ք բարձրության երկարժեք բացատրող փոփոխականով ռեգրեսիայի էկոնոմետրիկ մոդելը և բացատրե՛ք գնահատվող գործակիցների իմաստը:
5. Նկարագրե՛ք թեքության երկարժեք բացատրող փոփոխականով ռեգրեսիայի էկոնոմետրիկ մոդելը և բացատրե՛ք գնահատվող գործակիցների իմաստը:
6. Նկարագրե՛ք բարձրության ու թեքության երկարժեք բացատրող փոփոխականով ռեգրեսիայի էկոնոմետրիկ մոդելը և բացատրե՛ք գնահատվող գործակիցների իմաստը:

7. Նկարագրե՛ք բիմար ընտրության գծային հավանականական մոդելը, դրա հատկություններն ու թերությունները:
8. Նկարագրե՛ք բիմար ընտրության լոգիթ մոդելը, բացատրե՛ք դրա հատկություններն ու գործակիցների մեկնաբանությունը:

Խնդիրներ

1. ՀՀ-ում գործող «Glovo» կազմակերպությունը հավաքագրել է տեղեկատվություն վեջին մեկ տարվա ընթացքում իր 670 հաճախորդների կողմից կատարվող պատվերների վերաբերյալ: Դրանց հիման վրա գնահատվել է ամսական պատվերի գնի (Price, ՀՀ դրամ) կախումը պատվիրատուի ընտանիքի միջին ամսական եկամտից (Inc, ՀՀ դրամ) ու պատվիրատուի սեռից (D, սա երկարժեք DUMMY է՝ D=1, եթե պատվիրատուն տղամարդ է, և D=0, եթե կին է): «EViews» փաթեթով $\widehat{Price} = \widehat{\beta}_0 + \widehat{\beta}_1 \cdot Inc + \widehat{\beta}_2 D \cdot Inc$ մոդելը գնահատելու արդյունքում ստացվել է հետևյալ ելքային աղյուսակը՝

Included observations: 670

Dependent Variable: Price

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Inc	3 200	87.1	36.7	0.0000
$D \cdot Inc$	580	17.9	32.4	0.0000
C	10 000	233.5	42.8	0.0000
R-squared	0.611			
Adjusted R-squared	0.610			
S.E. of regression	4247.7			
F-statistic	1009.7			
Prob(F-statistic)	0.000			

1. Գրե՛ք գնահատված հավասարումը և կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի որակի մասին:
2. Հիմնավորե՛ք և պատասխանե՛ք այն հարցին, թե գնահատված ռեգրեսիայի հավասարումը վիճակագրորեն նշանակալի է, թե ոչ:
3. Կատարե՛ք հիմնավորված եզրակացություններ այն մասին, թե գնահատված $\hat{\beta}_0$, $\hat{\beta}_1$ ու $\hat{\beta}_2$ գործակիցները վիճակագրորեն նշանակալի են, թե ոչ:

4. Մեկնաբանե՞ք β_0 , β_1 և β_2 գործակիցների բովանդակությունը /իմաստը: Արդյո՞ք երկարժեք DUMMY-ն փոխում է գծի թեքությունը, բարձրությունը, թե՞ երկուսն էլ: Դո՞ւրս բերեք և հիմնավորե՞ք պատասխանը:
2. Աշխարհի 178 երկրի սոցիալ-տնտեսական տվյալների հիման վրա գնահատվել է բնակչության կյանքի սպասվող միջին տևողության (Life_exp, տարի) կախումը երկրի տնտեսական զարգացման մակարդակից (GDP, շնչային ՀՆԱ, ԱՄՆ դոլար) ու առողջապահական ծառայությունների մատուցման միջազգային կոնվենցիային երկրի պատկանելությունից (D, սա երկարժեք DUMMY է՝ D=1, եթե երկիրը անդամակցում է կոնվենցիային, և D=0՝ հակառակ դեպքում): SPSS փաթեթով $Life_exp = \beta_0 + \beta_1 GDP + \beta_2 D + \beta_3 D \cdot GDP$ մոդելը գնահատելու արդյունքում ստացվել են հետևյալ ելքային աղյուսակները.

Model Summary

Model	R	R Square	Adjusted R Square	F	Sig.
1	0.725 ^a	0.769	0.761	3270.1	0.01

Coefficients					
Model		Unstandardized Coefficients		t	Sig.
		B	Std. Error		
1	(Constant)	40.5	19.5	2.10	0.000
	GDP	3.5	0.26	13.1	0.000
	D	1.2	0.01	87.5	0.000
	D*GDP	5.3	0.13	40.01	0.000

1. Գրե՛ք գնահատված հավասարումը և կատարե՛ք եզրակացություն գնահատված ռեգրեսիայի որակի մասին:
2. Հիմնավորե՛ք և պատասխանե՛ք այն հարցին, թե գնահատված ռեգրեսիայի հավասարումը վիճակագրորեն նշանակալի է, թե ոչ:
3. Կատարե՛ք հիմնավորված եզրակացություններ այն մասին, թե գնահատված β_0 , β_1 , β_2 , β_3 գործակիցները վիճակագրորեն նշանակալի են, թե ոչ:

4. Մեկնաբանե՛ք β_0 , β_1 , β_2 , β_3 գործակիցների բովանդակությունը/իմաստը: Արդյո՞ք երկարժեք DUMMY-ն փոխում է գծի թեքությունը, բարձրությունը, թե՞ երկուսն էլ: Դո՞ւրս բերեք և հիմնավորե՛ք պատասխանը:
3. Ենթադրենք ուսումնասիրվում է տվյալ օնլայն վաճառքների հարթակում գրանցված օգտատերերի սպառողական վարքագիծը՝ կախված նրանց որոշակի բնութագրիչներից: Գնահատվել է Լոգիթ մոդել, որում կախյալ փոփոխականը երկարժեք ֆիկտիվ փոփոխական է՝ $D=0$, եթե տվյալ օգտատերը 6 ամսվա ընթացքում օնլայն հարթակում կատարել է 50 հազար դրամից պակաս արժուրոյքամբ գնումներ, և $D=1$, եթե կատարել է 50 հազար դրամից ավել արժուրոյքամբ գնումներ: Մոդելի բացատրող փոփոխականներն են՝ օգտատերերի ամսական եկամուտը, ընտանիքի անդամների թիվը, տարիքը: Մեկնաբանե՛ք SPSS փաթեթում մոդելի գնահատման արդյունքները՝ հիմնվելով հետևյալ ելքային աղյուսակների վրա:

Omnibus Tests of Model Coefficients				
		Chi-square	Df	Sig.
Step 1	Step	173.287	3	0.000

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	626.587 ^a	0.224	0.328

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	4.474	8	0.812

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	Եկամուտ, դրամ	0.164	0.022	55.570	1	0.000	1.178
	Տարիք	0.091	0.044	4.277	1	0.020	1.095
	Ընտ. անդամներ	-0.164	0.023	50.843	1	0.000	0.849
	Constant	-1.567	0.226	48.075	1	0.000	0.209

ԳԼՈՒԽ 6. ՊԱՏԱՀԱԿԱՆ ԳՈՐԾԸՆԹԱՑՆԵՐ ՈՒ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐ. ՀԻՄՆԱԿԱՆ ՀԱՏԿՈՒԹՅՈՒՆՆԵՐ ՈՒ ՆԱԽՆԱԿԱՆ ՈՒՍՈՒՄՆԱՍԻՐՈՒԹՅՈՒՆ

1. Պատահական գործընթացներ ու ժամանակային շարքեր. սահմանումներ ու տիպեր

Ինչպես արդեն նշվել է, էկոնոմետրիկ հետազոտության հիմքում ընկած էմպիրիկ տվյալները կարող են լինել ինչպես ժամանակի որևէ ամրագրված պահի համար (տարածական տվյալներ), այնպես էլ ժամանակի մեջ (ժամանակային շարքեր): Եթե նախորդ գլուխներում ուսումնասիրվում էին տարածական տվյալների էկոնոմետրիկ մոդելավորման հիմունքները, ապա այս և հաջորդ գլուխներում անդրադարձ է կատարվում ժամանակային շարքերի մոդելավորմանը: Այդ նպատակով ձևակերպենք մի քանի սահմանումներ, որոնք կօգտագործվեն ժամանակային շարքերի էկոնոմետրիկ մոդելավորման հաստեքստում:

Սահմանում 1. Ստոխաստիկ կամ պատահական $\{Y_t\}$ գործընթացը $Y_1, Y_2, \dots, Y_n$ պատահական մեծությունների ընտանիք է, որոնք ինդեքսավորված կամ դասավորված են ժամանակի մեջ ($t, t \in T$) և սահմանվում են որոշակի հավանականական տարածության մեջ:

Սահմանում 2. Ժամանակային շարքը ժամանակի որոշակի միջակայքում դրա հավասարահեռ պահերին պատահական գործընթացի ընդունած արժեքների (դիտարկումների) հաջորդականությունն է՝ դասավորված ըստ ժամանակի:

Ժամանակային շարքը, այսինքն՝ փոփոխականի դիտարկումների հաջորդականությունը ժամանակի $t=1, \dots, n$ պահերին կնշանակենք $y_t = y_1, y_2, \dots, y_n$:

Ժամանակային շարքի $y_1, y_2, \dots, y_n$ դիտարկումների հաջորդականությունը ժամանակի մեջ՝ $Y_1, Y_2, \dots, Y_n$ պատահական մեծությունների հաջորդականության կամ, որ նույնն է, $\{Y_t\}$ պատահական գործընթացի իրականացում (ռեալիզացիա) է: Ժամանակային շարքը ժամանակի յուրաքանչյուր t պահին $\{Y_t\}$ պատահական գործընթացի իրականացված, դիտարկված արժեքն է: Պատահական գործընթացը, որը գեներացնում է իր իրականացումը հանդիսացող ժամա-

նակային շարքը, անհայտ է: Ժամանակային շարքերի էկոնոմետրիկ մոդելավորումը հնարավորություն է տալիս եզրակացություններ կատարելու այն գեներացնող պատահական գործընթացի մասին: Օրինակ՝ ՀՀ ՀՆԱ-ն ժամանակի մեջ կարելի է դիտարկել որպես պատահական գործընթաց: Ժամանակի յուրաքանչյուր t պահին այն պատահական մեծություն է, որը որոշակի հավանականությամբ կարող է ընդունել այս կամ արժեքն իր հնարավոր արժեքների բազմությունից: Հենց որ այդ արժեքն արդեն ընդունում է ժամանակի t պահին, այն դառնում է պատահական գործընթացի իրականացում, ժամանակային շարքի դիտարկում: Տնտեսագիտական ուսումնասիրություններում մենք կարող ենք տեսնել պատահական գործընթացի միայն մեկ իրականացում, քանի որ չենք կարող ժամանակի մեջ հետ գնալ և գործընթացը նորից սկսել. այն արդեն պատմություն է: Այնուամենայնիվ, եթե պատմության որոշակի պայմաններ տարբեր լինեին, հնարավոր է, որ ունենայինք պատահական գործընթացի այլ իրականացում: Այդ պատճառով ժամանակային շարքերի դիտարկումները պատահական բնույթ ունեն: Պատահական գործընթացի բոլոր հնարավոր իրականացումների բազմությունը մնան է տարածական տվյալների գլխավոր համախմբությանը, որն անհայտ է, իսկ ժամանակային շարքը, փաստորեն, մնուշ է պատահական ընթացքի հնարավոր իրականացումներից:

Ժամանակային շարքերի էկոնոմետրիկ մոդելավորման հիմնական նպատակը այն գեներացնող պատահական գործընթացի ապագա արժեքների կանխատեսումն է կարճաժամկետ ու միջնաժամկետ հեռանկարում, ինչպես նաև մի քանի պատահական գործընթացի փոխազդեցությունների ուսումնասիրությունը ժամանակի մեջ:

Ինչպես պատահական գործընթացները, այնպես էլ դրանց իրականացումները՝ ժամանակային շարքերը, կարող են լինել **դիսկրետ** (ժամանակի ընդհատվող պահերի համար) ու **անընդհատ**, երբ ժամանակի պահերը, որոնցով ինդեքսավորված է գործընթացը կամ շարքը, անընդհատ են: Անընդհատ ժամանակային շարքերը էկոնոմետրիկայում հիմնականում չեն ուսումնասիրվում, սակայն կան որոշ բացառություններ՝ կապված ֆինանսական շուկայի ուսումնասիրությունների հետ:

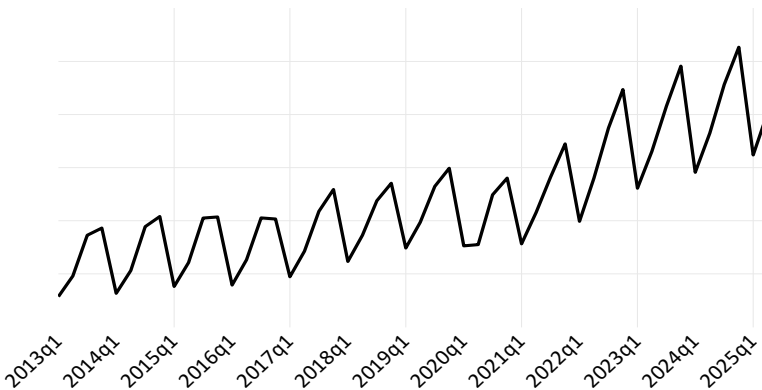
Այս ու հաջորդ գլուխներում կուսումնասիրվեն դիսկրետ ժամանակային շարքերը, ինչպես նաև կենթադրենք, որ ժամանակային շարքի դիտարկումները վերաբերում են ժամանակի հավասարահեռ պահերին: Սովորաբար ժամանակի հավասարահեռ պահերն են՝

տարին, եռամսյակը, շաբաթը, օրը, ժամը և այլն: Ժամանակային շարքերը կարելի է ներկայացնել գրաֆիկական տեսքով:

Որպես էկոնոմետրիկ մոդելավորման օբյեկտ՝ ժամանակային շարքերը (այսուհետ՝ ԺՇ) կարելի է դասել երկու խմբի՝ **միաչափ** և **բազմաչափ** ԺՇ-եր: Միաչափ ԺՇ-ն միայն մեկ փոփոխականի դիտարկումներն են ժամանակի հավասարահեռ պահերին ու ժամանակի որոշակի միջակայքում: Դրանց էկոնոմետրիկ վերլուծությունը նպատակ ունի մոդելավորելու **մեկ** փոփոխականի վարքը ժամանակի մեջ և դրա հիման վրա կանխատեսելու պատահական գործընթացի ապագա արժեքները: Բազմաչափ ԺՇ-ն մեկից ավելի փոփոխականների դիտարկումներն են ժամանակի հավասարահեռ պահերին և ժամանակի որոշակի միջակայքում: Դրանց ուսումնասիրությունը ենթադրում է մոդելավորել ու կանխատեսել **մեկ ցուցանիշի վարքը ժամանակի մեջ՝ կախված նաև ժամանակի մեջ մեկ կամ մի քանի այլ ցուցանիշների վարքից**: Գծանկար 24-ում ներկայացված է միաչափ ժամանակային շարք՝ գրաֆիկական տեսքով:

Գծանկար 24. ՀՀ ՀՆԱ-ն 2013-2025 թթ.՝ ըստ եռամսյակների, ընթացիկ գներով, մլն ՀՀ դրամ

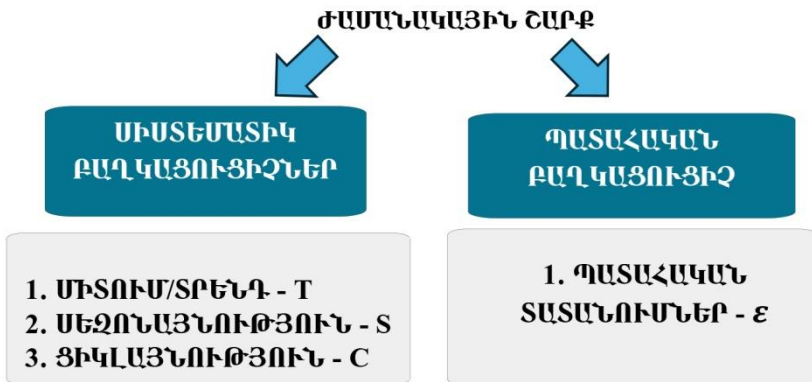
GDP



2. Ժամանակային շարքերի բաղկացուցիչները

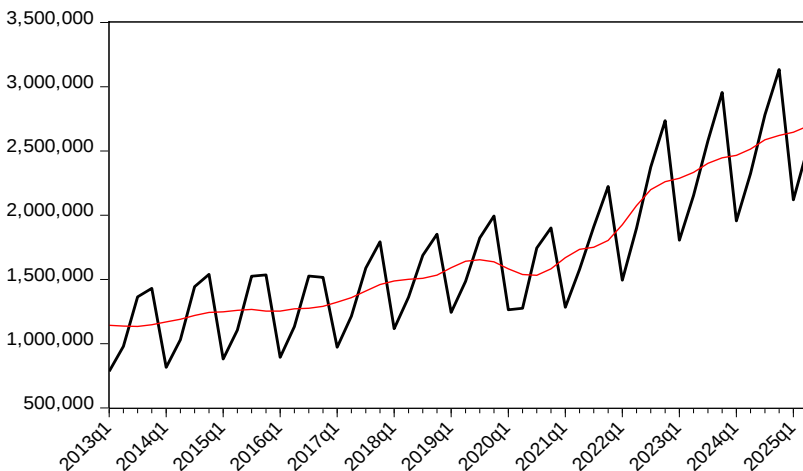
Ժամանակային շարքի արժեքները յուրաքանչյուր պահին ձևավորվում են մի շարք գործոնների ազդեցությամբ, որոնցով պայմանավորվում են շարքի բաղկացուցիչները: Ժամանակային շարքի բաղկացուցիչները (կամ շարքը ձևավորող գործոնները) կարելի է դասել երկու խմբի՝ **սխտեմատիկ** ու **պատահական**: ԺՇ սխտեմատիկ բաղադրիչներն են՝ **միտումը (տրենդը)**, **սեզոնայնությունը** և **ցիկլայնությունը** (տե՛ս Գծանկար 25):

Գծանկար 25. Ժամանակային շարքի բաղկացուցիչները



Միտումը (տրենդը) ԺՇ վարքի՝ ժամանակի մեջ ունեցած կայուն օրինաչափությունն է, ժամանակի մեջ շարքի փոփոխման հիմնական օրենքը, որը պայմանավորված է կայուն, երկարատև ազդեցություններով: Օրինակ՝ Գծանկար 26-ում բերված ՀՀ ՀՆԱ-ի ԺՇ-ն ունի աճի ընդհանուր միտում, ժամանակի մեջ աճելու օրինաչափություն, տրենդ:

Գծանկար 26. ՀՀ ՀՆԱ-ն 2013-2025 թթ.՝ ըստ հռանայականների, ընթացիկ գներով, մլն դրամ և դրա միջումը՝ տրենդային բաղադրիչը



Կարճաժամկետ հեռանկարում ժամանակային շարքի խիստ որոշակի պարբերությամբ կրկնվող տատանումները կոչվում են **սեզոնայնություն**: Դրանք պայմանավորված են ժամանակի մեջ որոշակի պարբերականությամբ կրկնվող ազդեցություններով՝ բնակլիմայական, տնտեսական, մշակութային և այլն:

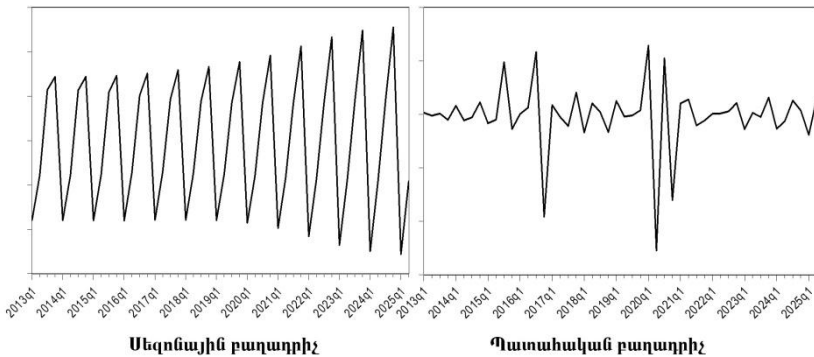
Ցիկլայնությունը ժամանակային շարքի պարբերական տատանումներն են երկարաժամկետ հեռանկարում: Դրանք պայմանավորված են ժամանակի որոշ երկարաժամկետ պարբերականությամբ կրկնվող ազդեցություններով (օրինակ, բիզնես ցիկլերով) և ի հայտ են գալիս պատմական տվյալների, երկար, պատմական շարքերի առկայության պայմաններում:

Պատահական տատանումները շարքի այն մասն է, որը կմնա սիստեմատիկ բաղկացուցիչները շարքից հեռացնելուց հետո: Այս պատահական տատանումները պայմանավորված են միայն պատահական ազդեցություններով: Հաճախ այս բաղադրիչն անվանում են պատահական աղմուկ (random noise):

ԺՇ-ն կարող է պարունակել որևէ մեկ կամ մի քանի կամ բոլոր 3 տեսակի սիստեմատիկ բաղկացուցիչները, սակայն կարող է նաև ընդհանրապես սիստեմատիկ բաղկացուցիչ չպարունակել: ԺՇ-ն միշտ պարունակում է պատահական բաղկացուցիչ:

Գծանկար 26-ում բերված ՀՀ ՀՆԱ-ի ԺՇ-ն ունի ինչպես աճման միտում, այնպես էլ սեզոնայնություն. մեկ տարվա եռամսյակների մեջ որոշակի տատանում կա, որը պարբերաբար տարեցտարի կրկնվում է: Գծանկար 27-ում ներկայացված են ՀՀ ՀՆԱ-ի ԺՇ առանձնացված սեզոնային ու պատահական բաղադրիչները:

Գծանկար 27. ՀՀ ՀՆԱ-ի 2013-2025 թթ. եռամսյակային տվյալների ԺՇ սեզոնային ու պատահական բաղադրիչները



ԺՇ միտումը կարող է լինել **դետերմինացված կամ ստոխաստիկ**. դա կքննվի հետագայում:

ԺՇ-ի վրա ազդող գործոնների ազդեցությամբ պայմանավորված՝ ԺՇ բաղկացուցիչները կարող են ձևավորել ԺՇ-ն տարբեր ձևերով: Եթե ԺՇ-ն ձևավորվում է իր բաղկացուցիչների գումարով, այդպիսի շարքը կոչվում է **ադիտիվ**: Ադիտիվ ԺՇ-ն ընդհանուր դեպքում ունի հետևյալ տեսքը՝ $y_t = T_t + S_t + C_t + \varepsilon_t$: Եթե ԺՇ-ն ձևավորվում է որպես իր բաղկացուցիչների արտադրյալ, այդպիսի շարքը կոչվում է **մուլտիպլիկատիվ**: Այն ընդհանուր դեպքում կարելի է ներկայացնել հետևյալ տեսքով՝ $y_t = T_t \cdot S_t \cdot C_t \cdot \varepsilon_t$: Եթե ԺՇ-ի բաղկացուցիչները տարբեր կերպ են մասնակցում ԺՇ ձևավորմանը, կոմենսանք **խառը** ԺՇ, օրինակ՝ $y_t = T_t \cdot S_t \cdot C_t + \varepsilon_t$:

3. Ստացիոնար պատահական գործընթաց ու ժամանակային շարք, սպիտակ աղմուկ պատահական գործընթաց

Քանի որ $\mathcal{F}$ -երն ու դրանք ձևավորող բաղկացուցիչները բազմատեսակ են, $\mathcal{F}$ -երի վիճակագրական վերլուծության, մոդելավորման ու կանխատեսման խնդիրը գործնականորեն լուծելու համար անհրաժեշտ է ինչ-որ կերպ սահմանափակել մոդելավորվող $\mathcal{F}$ -երի դասը՝ ներկայացնելով որոշակի ենթադրություններ շարքի կառուցվածքի ու դրա հավանականական բնութագրերի վերաբերյալ: Այս սահմանափակումներից մեկը ենթադրում է ժամանակային շարքի **կայունություն ժամանակի մեջ, որը կանվանենք ստացիոնարություն**:

Ստացիոնար $\mathcal{F}$ -երի համար մշակված են էկոնոմետրիկ մոդելավորման որոշակի տեխնիկաներ, իսկ ոչ ստացիոնար $\mathcal{F}$ -երի գերակշիռ մեծամասնությունը կարելի է ստացիոնարացնել ու մոդելավորել եղած տեխնիկաներով:

Այսպիսով՝ $\mathcal{F}$ -երի մոդելավորման հիմքում դրա մի կարևորագույն հատկությունն է, որը կոչվում է ստացիոնարություն, և որի առկայությունը կամ բացակայությունը պայմանավորում է $\mathcal{F}$ մոդելավորման տեխնիկան:

Կան պատահական գործընթացի ու դրա իրականացումը հանդիսացող $\mathcal{F}$ -ի ստացիոնարության երկու սահմանումներ՝ խիստ (նեղ) ու թույլ (լայն):

ՍԱՀՄԱՆՈՒՄ 1. $\{Y_t: t=1,2,\dots,T\}$ պատահական գործընթացը կոչվում է **նեղ կամ խիստ իմաստով ստացիոնար**, եթե դրա կամայական m հաջորդական պատահական մեծությունների համատեղ բաշխման ֆունկցիան նույնն է, ինչ կամայական τ պահով դրանից հեռացած m հաջորդական պատահական մեծությունների համատեղ բաշխման ֆունկցիան: Այսինքն՝

$$F(Y_{t_1}, Y_{t_2}, \dots, Y_{t_m}) = F(Y_{t_1+\tau}, Y_{t_2+\tau}, \dots, Y_{t_m+\tau}), \forall m, \tau: \quad (126)$$

Սա նշանակում է, որ ստացիոնար պատահական գործընթացի բաշխումը, հատկությունները կախված չեն դրա սկզբնակետից. դրանք ժամանակի մեջ կայուն են: Քանի որ պատահական գործընթացն ուսումնասիրելիս իրականում գործ ունենք դրա իրականացումը հանդիսացող $\mathcal{F}$ -ի հետ, ապա գործնականում դա նշանակում է, որ $\mathcal{F}$ -ի՝ ժամանակի մեջ որ մասից էլ դիտարկենք որևէ երկարության հատված, այն կունենա նույն հատկությունները, ինչ ժամանա-

կի կամայական այլ մասից վերցված նույն երկարությամբ հատվածը: Եթե պատահական գործընթացի m հաջորդական պատահական մեծությունների համատեղ բաշխումը նորմալ բաշխում է, ապա այդ պատահական գործընթացը կոչվում է **Գաուսյան**:

Նկատի ունենալով, որ գործնականում պատահական գործընթացի բաշխման ֆունկցիան հաշվարկելն ու դրա ստացիոնարությունը ստուգելը բավական բարդ է՝ օգտագործվում է ստացիոնարության թույլ սահմանումը՝

ՍԱՀՄԱՆՈՒՄ 2. $\{Y_t: t=1,2,...T\}$ պատահական գործընթացը կամ դրա իրականացումը հանդիսացող y_t ժամանակային շարքը կոչվում է լայն կամ թույլ իմաստով ստացիոնար, եթե դրա թվային բնութագրիչները՝ միջինը, վարիացիան ու կովարիացիան, կախված չեն ժամանակից, այսինքն՝ ԺՇ-ի համար.

$$1. E(y_t) = \mu, \forall t = 1, 2, \dots, n, \quad (127)$$

$$2. Var(y_t) = E(y_t - \mu)^2 = \sigma_y^2, \quad \forall t = 1, 2, \dots, n, \quad (128)$$

$$3. Cov(y_t, y_{t+\tau}) = E((y_t - \mu) \cdot (y_{t+\tau} - \mu)) = \gamma(\tau), \forall t, \tau = 1, \dots, n, \tau \neq 0: \quad (129)$$

(127)-ը նշանակում է, որ ԺՇ-ն ժամանակի մեջ տատանվում է հաստատուն միջինի շուրջը, ԺՇ միջինը ժամանակի մեջ չի փոխվում, կախված չէ ժամանակից: (128)-ը նշանակում է, որ իր հաստատուն միջինի շուրջը ԺՇ-ն տատանվում է հաստատուն վարիացիայով, միջինից ԺՇ շեղումների ամպլիտուդը կախված չէ ժամանակից: (129)-ը նշանակում է, որ ԺՇ կամայական τ ժամանակային պահով իրարից հեռացած երկու անդամի կովարիացիան կախված չէ ժամանակից, այլ կախված է միայն այդ անդամների՝ ժամանակի մեջ իրարից ունեցած հեռավորությունից, որն այսուհետ կանվանենք **լագ**: Քանի որ կովարիացիան բնութագրում է երկու մեծությունների գծային կախումը, դա նշանակում է, որ ԺՇ կամայական τ լագով իրարից հեռացած երկու անդամների գծային կապը կախված չէ նրանից, թե ԺՇ որ հատվածից ենք վերցնում այդ երկու անդամները, այլ կախված է այդ երկու անդամների՝ ժամանակի մեջ միմյանցից ունեցած հեռավորությունից:

Ստացիոնարության (127)-(129) պայմանները բավարարող ԺՇ-ն տատանվելով վերադառնում է իր միջին արժեքին, և այդ միջինի շուրջ տատանումներն ունեն ընդհանուր առմամբ հաստատուն

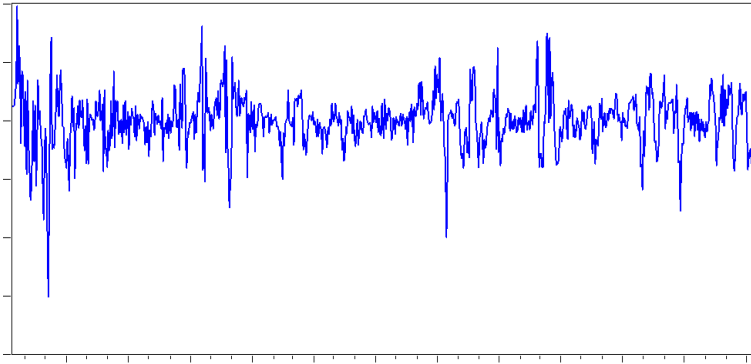
ամալիտուդ: Այլ կերպ ասած՝ ստացիոնար ԺՇ-ն շատ չի շեղվում իր միջին արժեքից հաստատուն ու վերջավոր վարիացիայի շնորհիվ:

- ✓ Այսուհետ ստացիոնարություն ասելով՝ կհասկանանք թույլ իմաստով ստացիոնարությունը:
- ✓ Ենթադրվում է, որ ԺՇ-ն ունի թվային բնութագրիչներ:
- ✓ Եթե պատահական գործընթացը կամ ԺՇ-ն ստացիոնար է նեղ իմաստով, ապա այն նաև լայն իմաստով ստացիոնար է (ԺՇ-ի դեպքում, եթե ԺՇ-ն ունի թվային բնութագրիչներ): Հակառակը միշտ չէ, որ ճիշտ է:
- ✓ Եթե պատահական գործընթացը Գ-աուսյան է, ապա թույլ/լայն իմաստով ստացիոնարությունը նշանակում է նաև նեղ/խիստ իմաստով ստացիոնարություն:

Եթե պատահական գործընթացի կամ դրա իրականացումը հանդիսացող ԺՇ-ի համար խախտված է դրա ստացիոնարության վերը նշված (127)-(129) պայմաններից մեկը (կամ մի քանիսը), ապա ԺՇ-ն կամ պատահական գործընթացը **ստացիոնար չէ: Դա նշանակում է, որ ոչ ստացիոնար պատահական գործընթացի ու ԺՇ միջինը, վարիացիան կամ կովարիացիան կախված են ժամանակից, այսինքն՝ ժամանակի մեջ փոփոխվում են:** Այստեղից հետևում է, որ **սիստեմատիկ բաղադրիչներ ունեցող ԺՇ-ն ստացիոնար լինել չի կարող:** Իսկապես, եթե ԺՇ-ն ունի միտում, ժամանակի մեջ փոփոխման օրինաչափություն (աճ, նվազում, աճից հետո նվազում, սեզոնայնություն և այլն), նշանակում է՝ դրա միջինը հաստատուն չէ: Խախտվում է առնվազն ստացիոնարության 1-ին պայմանը: Գծանկար 24-ում պատկերված ՀՀ ՀՆԱ ԺՇ-ն ստացիոնար չէ: Այն ունի աճի միտում, դրա միջինը ժամանակի մեջ աճում է: Միևնույն ժամանակ այն ունի նաև այլ սիստեմատիկ բաղադրիչ՝ սեզոնայնություն:

Գծապատկեր 28-ում ներկայացված է ստացիոնար միաչափ ԺՇ գրաֆիկի օրինակ:

Գծանկար 28. Ստացիոնար ԺՇ գրաֆիկի օրինակ



Ստացիոնար պատահական գործընթացի ու դրանով գեներացված ԺՇ դասական օրինակ է «սպիտակ աղմուկ» ԺՇ-ն:

Սպիտակ աղմուկ պատահական գործընթացը կամ ԺՇ-ն բավարարում է հետևյալ պայմանները.

- 1) այն ունի հաստատուն միջին, որը հավասար է 0-ի՝

$$\mathbb{E}(y_t) = \mu = 0, \forall t = 1, 2, \dots, n, \quad (130)$$

- 2) այն իր միջինի շուրջը տատանվում է հաստատուն վարիացիայով՝

$$\text{Var}(y_t) = \mathbb{E}(y_t - \mu)^2 = \sigma_y^2, \quad \forall t = 1, 2, \dots, n, \quad (131)$$

- 3) դրա կամայական երկու անդամների կովարիացիան ոչ միայն կախված չէ ժամանակից. այն հավասար է զրոյի, այսինքն՝ շարքի ոչ մի անդամ որևէ մյուս անդամի հետ կոռելացված չէ (ԺՇ-ն չունի «հիշողություն», դրա կամայական պահի դիտարկումը, բնորոշվում արժեքը «չի հիշում» նախորդ կամ հաջորդ որևէ պահի անդամին).

$$\begin{aligned} \text{Cov}(y_t, y_{t \pm \tau}) &= \mathbb{E}((y_t - \mu) \cdot (y_{t \pm \tau} - \mu)) = 0, \\ \forall t, \tau &= 1, \dots, n, \tau \neq 0: \end{aligned} \quad (132)$$

Նշված երեք պայմանները նշանակում են, որ սպիտակ աղմուկ ԺՇ-ն ստացիոնար է, ինչպես նաև երրորդ պայմանից հետևում է, որ այն մաքուր պատահական ԺՇ է: Եթե սպիտակ աղմուկ գործընթացը կամ ԺՇ-ն նաև նորմալ բաշխում ունի, ապա այն կոչվում է Գաուսյան սպիտակ աղմուկ:

4. Ոչ ստացիոնար պատահական գործընթաց և ժամանակային շարք

Տնտեսական ցուցանիշներ բնութագրող ԺՇ-երը հազվադեպ են լինում ստացիոնար: Ոչ ստացիոնար պատահական գործընթացի դասական օրինակ է «պատահական քայլը» կոչվող պատահական գործընթացը: Բնական է ենթադրել, որ պատահական գործընթացի կամ ԺՇ տարբեր պահերի անդամները որոշակի կոռելյացիայի մեջ են միմյանց հետ: Դրանք որոշ չափով «հիշում են» միմյանց: Դա նշանակում է, որ միաչափ ԺՇ-երի մոդելներում շարքի յուրաքանչյուր t -րդ պահի արժեքը կարելի է մոդելավորել՝ վերջինս կախման մեջ դնելով շարքի նախորդ անդամների արժեքներից: Ենթադրենք այդ կախումը գծային տեսք ունի:

Ենթադրենք ԺՇ յուրաքանչյուր t -րդ պահի արժեքը գծային կախման մեջ է միայն իր նախորդ պահի արժեքից ու մաքուր պատահական տատանումից, և ԺՇ-ն կարելի է ներկայացնել հետևյալ գծային տեսքի մոդելով՝

$$y_t = y_{t-1} + \varepsilon_t, t=1,2,\dots,n, \quad (133)$$

որում ε_t -ն սպիտակ աղմուկ ԺՇ է, այսինքն՝ դրա համար տեղի ունեն (130)-(132) պայմանները:

(133) տեսքով արտահայտվող ԺՇ-ն կոչվում է «պատահական քայլը» ԺՇ (random walk- RW): (133)-ը կարելի է ներկայացնել հետևյալ կերպ՝ $y_t = \beta_1 \cdot y_{t-1} + \varepsilon_t$, $\beta_1 = 1$, ուստի այս շարքը հաճախ անվանում են նաև **միավոր արժանո ԺՇ**: Քանի որ (133) տեսքի գծային կախման մեջ ազատ անդամը՝ $\beta_0 = 0$, այսպիսի պատահական քայլով արտահայտվող ԺՇ-ն կոչվում է առանց տեղաշարժի պատահական քայլով ԺՇ (RW without drift): Առանց տեղաշարժի RW ԺՇ յուրաքանչյուր t -րդ դիտարկում ԺՇ նախորդ պահի դիտարկումից տարբերվում է միայն պատահական աղմուկով: Այս RW-ն ոչ ստացիոնար ԺՇ դասական օրինակ է: Յուրաքանչյուր, որ RW ԺՇ-ն ստացիոնար չէ, այսինքն՝ նրա համար խախտված է լայն իմաստով ստացիոնարության 3 պայմաններից առնվազն մեկը:

Գրենք (133) հավասարումը ԺՇ հաջորդական անդամների համար առանձին-առանձին՝

$$\begin{aligned} t = 1 \quad y_1 &= y_0 + \varepsilon_1, \\ t = 2 \quad y_2 &= y_1 + \varepsilon_2 = y_0 + \varepsilon_1 + \varepsilon_2, \\ t = 3 \quad y_3 &= y_2 + \varepsilon_3 = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3: \\ &\dots\dots\dots \end{aligned}$$

Ընդհանուր առմամբ կամայական t պահի համար կունենանք.

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t = y_0 + \sum_{i=1}^t \varepsilon_i: \quad (134)$$

Ստացվեց, որ RW ԺՇ կամայական t պահի արժեքը կախված է միայն ԺՇ սկզբնական անդամից ու առաջինից մինչև t պահերի պատահական տատանումների գումարից: Այսինքն՝ այս ԺՇ յուրաքանչյուր t պահի անդամ «հիշում է» իր իսկ՝ ԺՇ միայն սկզբնական անդամին, բայց դրա փոխարեն, «հիշում է» մինչև t պահը ներառյալ բոլոր պատահական տատանումները: Կամայական պահի պատահական տատանման ազդեցությունը ԺՇ-ի վրա չի «մոռացվում», ինչքան էլ ժամանակի մեջ տեղաշարժվենք: **Պատահական տատանումների նկատմամբ այս ԺՇ-ն ունի շատ երկար «հիշողություն», որը չի մարում, ինչքան էլ ժամանակի մեջ հեռանանք շարքի սկզբնական արժեքից: Սա նշանակում է, որ RW-ն ունի ժամանակի մեջ փոփոխման օրինաչափություն՝ ԺՇ անդամների վրա պատահական տատանումների ունեցած ազդեցության առումով:** Այդ իսկ պատճառով RW-ում $\sum_{i=1}^t \varepsilon_i$ -ն անվանում են ստոխաստիկ միտում կամ ստոխաստիկ տրենդ:

Ցույց տանք, որ RW-ն ստացիոնար չէ, և դա պայմանավորված է ստոխաստիկ միտումի առկայությամբ: Հաշվենք դրա միջինը:

$$\begin{aligned} \mathbb{E}(y_t) &= \mathbb{E}(y_{t-1} + \varepsilon_t) = \mathbb{E}(y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = \\ &= \mathbb{E}(y_0 + \sum_{i=1}^t \varepsilon_i) = y_0: \end{aligned}$$

Առանց տեղաշարժի RW ԺՇ միջինը հավասար է այդ ԺՇ սկզբնական անդամին, որը թիվ է, այսինքն՝ կախված չէ ժամանակից: Ստացիոնարության առաջին պայմանը բավարարված է: Նշենք, որ պրակտիկայում սովորաբար շարքի սկզբնակետը կամ սկզբնական անդամը՝ $y_0 = 0$, այսինքն՝ RW -ի միջինը 0-ն է: Հաշվենք RW ԺՇ-ի վարիացիան:

$$\begin{aligned} \text{Var}(y_t) &= \text{Var}(y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t) = \\ &= \text{Var}(y_0 + \sum_{i=1}^t \varepsilon_i) = t \cdot \sigma_\varepsilon^2, \end{aligned}$$

որում σ_ε^2 -ն սպիտակ աղմուկ ε_t ԺՇ-ի հաստատուն վարիացիան է:

Ստացանք, որ RW ԺՇ վարիացիան կախված է ժամանակից, այն աճում է ժամանակի մեջ, այսինքն՝ RW ԺՇ ստացիոնարության երկրորդ պայմանը խախտված է: Հետևաբար այս շարքը ստացիո-

նար z_t , և դա պայմանավորված է ստոխաստիկ միտումի՝ $\sum_{i=1}^t \varepsilon_t$ -ի, առկայությամբ:

Ենթադրենք ԺՇ յուրաքանչյուր t -րդ պահի արժեքը կախված է միայն իր նախորդ պահի արժեքից ու մաքուր պատահական տատանումներից հետևյալ գծային տեսքով՝

$$y_t = \beta_0 + y_{t-1} + \varepsilon_t, t=1,2,\dots,n, \quad (135)$$

որում ε_t -ն սպիտակ աղմուկ ԺՇ է:

(135)-ով արտահայտվող ԺՇ -ն կոչվում է «տեղաշարժով պատահական քայլք ԺՇ » (RW with drift - RWWD), քանի որ այս տեսքում ազատ անդամը՝ $\beta_0 \neq 0$: Յույց տանք, որ RWWD ԺՇ -ն ստացիոնար չէ, այսինքն՝ նրա համար խախտված է լայն իմաստով ստացիոնարության 3 պայմաններից առնվազն մեկը:

Գրենք RWWD-ի հավասարումը ԺՇ հաջորդական անդամների համար՝

$$\begin{aligned} y_1 &= \beta_0 + y_0 + \varepsilon_1 \\ y_2 &= \beta_0 + y_1 + \varepsilon_2 = \beta_0 + \beta_0 + y_0 + \varepsilon_1 + \varepsilon_2 = 2\beta_0 + y_0 + \varepsilon_1 + \varepsilon_2 \\ y_3 &= \beta_0 + y_2 + \varepsilon_3 = \beta_0 + \beta_0 + \beta_0 + y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \\ &= 3\beta_0 + y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3: \end{aligned}$$

.....

Ընդհանուր առմամբ կամայական t պահի համար կունենանք՝

$$\begin{aligned} y_t &= \beta_0 + y_{t-1} + \varepsilon_t = t \cdot \beta_0 + y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \dots + \varepsilon_t = \\ &= t \cdot \beta_0 + y_0 + \sum_{i=1}^t \varepsilon_t: \end{aligned}$$

Հաշվենք RWWD ԺՇ -ի միջինը՝

$$\mathbb{E}(y_t) = \mathbb{E}(\beta_0 + y_{t-1} + \varepsilon_t) = \mathbb{E}\left(t \cdot \beta_0 + y_0 + \sum_{i=1}^t \varepsilon_t\right) = t \cdot \beta_0 + y_0:$$

RWWD ԺՇ միջինը կախված է ժամանակից, ստացիոնարության առաջին պայմանը չի բավարարվում:

$$\text{Var}(y_t) = \text{Var}\left(t \cdot \beta_0 + y_0 + \sum_{i=1}^t \varepsilon_t\right) = t \cdot \sigma_\varepsilon^2:$$

RWWD ԺՇ վարիացիան ևս կախված է ժամանակից, այն աճում է ժամանակի մեջ, այսինքն՝ RWWD ԺՇ ստացիոնարության երկրորդ պայմանը ևս խախտված է: Այս շարքը ևս ստացիոնար չէ, և դա պայմանավորված է ստոխաստիկ միտումի առկայությամբ:

5. Տարբերական օպերատոր ու տարբերությամբ վերցված ոչ ստացիոնար ժամանակային շարքեր

Ինչպես նշվել է, պատահական քայլքով արտահայտվող ԺՇ-երը ստացիոնար չեն՝ պայմանավորված նրանով, որ դրանք ունեն ստոխաստիկ միտում: Դիտարկենք՝ արդյոք ստոխաստիկ միտում ունեցող ԺՇ-երը կարելի է ձևավորել ու ստացիոնարացնել:

Ենթադրենք y_t ԺՇ-ն ստացիոնար չէ և ունի ստոխաստիկ միտում:

Սահմանում. Δ -ն կանվանենք տարբերական օպերատոր, եթե դրա կիրառումը ԺՇ-ի նկատմամբ հաշվարկում է ԺՇ յուրաքանչյուր t -րդ ու $t-1$ -րդ պահերի արժեքների տարբերությունը՝

$$\Delta(y_t) = y_t - y_{t-1}:$$

Այլ կերպ ասած՝ ԺՇ յուրաքանչյուր t պահի անդամի նկատմամբ Δ օպերատորի կիրառումը գեներացնում է մի նոր ԺՇ, որը կոչվում է y_t ԺՇ առաջին տարբերությամբ վերցված ԺՇ կամ y_t ԺՇ առաջին տարբերություն (first difference), և որի կամայական t -րդ անդամը հավասար է y_t ԺՇ-ի t -րդ ու $t-1$ -րդ պահերի արժեքների տարբերությանը (կամ, որ նույնն է, պարզ աճին):

Նշենք, որ տարբերական օպերատորը ժամանակային շարքի նկատմամբ կարելի է կիրառել մեկից ավելի անգամ: Եթե այն կիրառվում է մեկ անգամ, ստացված նոր շարքը նշանակվում է $\Delta(y_t)$ -ով: Երկու անգամ տարբերական օպերատորի կիրառումը հանգեցնում է հետևյալ շարքին՝

$$\begin{aligned}\Delta^2(y_t) &= \Delta(\Delta(y_t)) = \Delta(y_t - y_{t-1}) = y_t - y_{t-1} - (y_{t-1} - y_{t-2}) = \\ &= y_t - 2y_{t-1} + y_{t-2}:\end{aligned}$$

d անգամ տարբերական օպերատորի կիրառումը հանգեցնում է $\Delta^d(y_t)$ շարքի ստացմանը:

Կիրառենք տարբերական օպերատորը առանց տեղաշարժի RW ԺՇ-ի նկատմամբ.

$$\begin{aligned}y_t = y_{t-1} + \varepsilon_t &\rightarrow \Delta(y_t) = y_t - y_{t-1} = y_{t-1} + \varepsilon_t - y_{t-1} = \varepsilon_t \\ &\rightarrow \Delta(y_t) = \varepsilon_t:\end{aligned}$$

Ստացվեց, որ չնայած առանց տեղաշարժի RW-ն ստացիոնար չէ, որովհետև ունի ստոխաստիկ միտում, դրա առաջին տարբերու-

թյունը ստացիոնար է, քանի որ այն հավասար է ε_t -ի, որը սպիտակ աղմուկ $\mathcal{F}_t$ է, ըստ սահմանման՝ ստացիոնար է: Տարբերական օպերատորի կիրառումը հնարավորություն տվեց վերացնելու ստոխաստիկ միտումը առանց տեղաշարժի RW $\mathcal{F}_t$ -ից և այդպիսով ստացիոնարացնելու այն:

Դիտարկենք տեղաշարժով պատահական քայլքի RWWD մոդելով բնութագրվող $\mathcal{F}_t$, որը նույնպես ստացիոնար չէ՝ պայմանավորված ստոխաստիկ տրենդի առկայությամբ, ինչպես ցույց ենք տվել: Կիրառենք այս $\mathcal{F}_t$ -ի վրա տարբերական օպերատորը կամ, որ նույնն է, հաշվենք դրա առաջին տարբերությունը՝

$$\begin{aligned} y_t &= \beta_0 + y_{t-1} + \varepsilon_t \rightarrow \Delta(y_t) = y_t - y_{t-1} = \\ &= \beta_0 + y_{t-1} + \varepsilon_t - y_{t-1} = \beta_0 + \varepsilon_t: \end{aligned}$$

$\Delta(y_t)$ -ն ստացիոնար է, քանի որ ստացիոնար սպիտակ աղմուկից տարբերվում է հաստատուն β_0 -ով: Այս շարքը ստացիոնար է: Այս դեպքում ևս տարբերական օպերատորի կիրառումը հնարավորություն տվեց վերացնելու ստոխաստիկ միտումը նաև տեղաշարժով RWWD $\mathcal{F}_t$ -ից ու այդպիսով ստացիոնարացնելու այն: Ընդհանուր եզրակացությունը հետևյալն է. ստոխաստիկ միտում ունեցող ոչ ստացիոնար $\mathcal{F}_t$ -ն բերվում է ստացիոնար $\mathcal{F}_t$ -ի՝ տարբերական օպերատորի կիրառմամբ:

6. Տրենդային ստացիոնար (TS) ու տարբերությամբ ստացիոնար (DS) ժամանակային շարքեր ու դրանց ստացիոնարացումը, ինտեգրացվող ժամանակային շարքեր

Ինչպես արդեն նշվել է, սիստեմատիկ բաղադրիչ ունեցող, այդ թվում՝ միտում ունեցող $\mathcal{F}_t$ -ն ստացիոնար չէ:

Սահմանում. y_t $\mathcal{F}_t$ -ն կոչվում է $f(t)$ դետերմինացված միտումի նկատմամբ ստացիոնար կամ տրենդային ստացիոնար (TS, trend stationary), եթե $y_t - f(t)$ $\mathcal{F}_t$ -ն ստացիոնար է, որում $f(t)$ -ն y_t $\mathcal{F}_t$ -ի՝ ժամանակից ունեցած կախվածության դետերմինացված ֆունկցիան է (մասնավոր դեպքում եթե $\mathcal{F}_t$ միտումը մուլտիպլիկատիվ է, ստացիոնար է նաև $y_t / f(t)$ $\mathcal{F}_t$ -ն):

Այսպիսով՝ դետերմինացված $f(t)$ միտումը $\mathcal{F}_t$ դիտարկումների ֆունկցիոնալ կախումն է ժամանակից: Դետերմինացված միտումը

կարող է լինել ինչպես գծային, այնպես էլ ոչ գծային: Օրինակ՝ դետերմինացված միտում ունեցող ԺՇ-երը կարող են արտահայտվել հետևյալ տեսքի մոդելներով՝

$$\begin{aligned} y_t &= \beta_0 + \beta_1 \cdot t + \varepsilon_t & y_t &= \beta_0 + \beta_1 \cdot t + \beta_1 \cdot t^2 + \varepsilon_t \\ y_t &= \beta_0 + \frac{\beta_1}{t} + \varepsilon_t & y_t &= \beta_0 + t^{\beta_1} + \varepsilon_t \\ y_t &= \beta_0 \cdot \beta_1^t \cdot \varepsilon_t & y_t &= e^{(\beta_0 + \beta_1 \cdot t + \varepsilon_t)} \end{aligned}$$

Սահմանումից երևում է, որ TS ԺՇ-ն կարելի է բերել ստացիոնար ԺՇ-ի՝ դետերմինացված միտումը դրանից հեռացնելով: $f(t)$ դետերմինացված միտումի հեռացումը ԺՇ-ից կոչվում է **սպատրենդա-ցում**:

Դիտարկենք ամենապարզ դեպքը, երբ $f(t)$ -ն պարզ գծային կախում է ժամանակից, և ԺՇ-ն կարելի է արտահայտել այս բանաձևով՝

$$y_t = f(t) + \varepsilon_t = \beta_0 + \beta_1 t + \varepsilon_t,$$

որում ε_t -ն սպիտակ աղմուկ ստացիոնար ԺՇ է: y_t շարքը ստացիոնար չէ, բանի որ $E(y_t) = \beta_0 + \beta_1 t$, ԺՇ միջինը կախված է ժամանակից: Եթե y_t ԺՇ-ից հանենք, հեռացնենք դետերմինացված միտումը, կստանանք ստացիոնար շարք.

$$y_t - f(t) = \beta_0 + \beta_1 t + \varepsilon_t - (\beta_0 + \beta_1 t) = \varepsilon_t:$$

Սպատրենդացման համար նախ պետք է գտնել ԺՇ-ի՝ ժամանակից ունեցած կախվածության ճիշտ տեսքը, սպա գնահատել $y_t = f(t) + \varepsilon_t$ կախումը ՓՋԵ-ով և արդյունքում ստացված գնահատված $\hat{y}$ -ը հեռացնել ԺՇ-ից:

Սահմանում. Ստոխաստիկ միտում ունեցող ոչ ստացիոնար ԺՇ-երը կանվանենք DS (difference stationary), տարբերություններով ստացիոնար կամ ինտեգրացվող ԺՇ-եր:

Այս գլխի նախորդ կետում ցույց ենք տվել, որ այս ԺՇ-երը կարելի է ստացիոնարացնել **տարբերական օպերատոր կիրառելու միջոցով**:

Ընդհանրապես ասած՝ տարբերական օպերատորի կիրառումը ստացիոնարացնում է թե՛ DS և թե՛ TS ժամանակային շարքերը: Սակայն եթե TS ԺՇ-ն ստացիոնարացվում է ոչ թե սպատրենդացման միջոցով, այլ տարբերական օպերատորի կիրառմամբ, այդ շարքը ձեռք է բերում այնպիսի ոչ ցանկալի հատկություններ, որոնք հետա-

գայում խոչընդոտում են դրա մոդելավորմանն ու վստահելի արդյունքների ստացմանը:

Նշենք, որ կամ ժամանակային շարքեր, որոնք ունեն n ստոխաստիկ, n դետերմինացված միտում, օրինակ՝

$$y_t = f(t) + y_{t-1} + \varepsilon_t = \beta_0 + \beta_1 t + y_{t-1} + \varepsilon_t:$$

Այսպիսի շարքը ստացիոնարացնելու համար դրա նկատմամբ n պետք է տարբերական օպերատոր կիրառել, n ապատրենդացնել՝ հեռացնել դետերմինացված տրենդը:

$$\Delta(y_t) - f(t) = \Delta(\beta_0 + \beta_1 t + y_{t-1} + \varepsilon_t) - (\beta_0 + \beta_1 t) = \beta_0 + \beta_1 t + y_{t-1} + \varepsilon_t - y_{t-1} - (\beta_0 + \beta_1 t) = \varepsilon_t:$$

ԺՇ-ն կարող է ունենալ միայն դետերմինացված միտում կամ միայն ստոխաստիկ միտում, այն կարող է ունենալ n դետերմինացված, n ստոխաստիկ միտում, կարող է նաև չունենալ դրանցից ոչ մեկը և ի սկզբանե լինել ստացիոնար:

Սահմանում. y_t ԺՇ-ն կոչվում է d -րդ կարգի ինտեգրացվող, եթե d անգամ y_t -ի նկատմամբ տարբերական օպերատոր կիրառելուց հետո ստացված $\Delta^d(y_t)$ ԺՇ-ն ստացիոնար է, իսկ $d-1$ անգամ տարբերական օպերատոր կիրառելուց հետո ստացված $\Delta^{d-1}(y_t)$ ԺՇ-ն ստացիոնար չէ:

Ընդհանուր առմամբ եթե ոչ ստացիոնար, **DS** ժամանակային շարքի նկատմամբ պետք է d անգամ կիրառել տարբերական օպերատոր, որպեսզի այն ստացիոնար դառնա, ապա այդ ժամանակային շարքը կոչվում է d -րդ կարգի ինտեգրացվող: d -րդ կարգի ինտեգրացվող Y_t ժամանակային շարքը նշանակվում է՝ $y_t \sim I(d)$: Եթե y_t ժամանակային շարքը ստացիոնար է, այն անվանում են զրոյական կարգի ինտեգրացվող և նշանակում $y_t \sim I(0)$:

Ինտեգրացվող ԺՇ-ի հատկությունները

1. Եթե ունենք երկու ԺՇ՝ $x_t \sim I(0)$ և $y_t \sim I(1)$, ապա $z_t = (x_t + y_t) \sim I(1)$: Այսինքն՝ երկու (ստացիոնար ու ոչ ստացիոնար) ԺՇ-երի գծային կոմբինացիան ոչ ստացիոնար է:
2. Եթե ԺՇ-ն d -րդ կարգի ինտեգրացվող է՝ $x_t \sim I(d)$, իսկ a -ն ու b -ն հաստատուններ են, ապա $z_t = (a + bx_t) \sim I(d)$:

3. Եթե ունենք երկու $\mathcal{F}\mathcal{C}$ ՝ $x_t \sim I(d_1)$ ու $y_t \sim I(d_2)$ և $d_1 < d_2$, իսկ a -ն ու b -ն հաստատուններ են, ապա $z_t = (ax_t + by_t) \sim I(d_2)$:
4. Եթե ունենք երկու $\mathcal{F}\mathcal{C}$ ՝ $x_t \sim I(d)$ և $y_t \sim I(d)$, ապա $z_t = (ax_t + by_t) \sim I(d^*)$, որում, բացառությամբ որոշ դեպքերի (որանք կոինտեգրացվածության դեպքերն են, որոնց համար $d^* < d$, ինչը կդիտարկենք հետագայում), $d^* = d$:

Ներդիր 17. Սեզոնայնության հեռացումը $\mathcal{F}\mathcal{C}$ սպացիոնարացման նպատակով

Ինչպես արդեն նշվել է, սիստեմատիկ բաղադրիչներ ունեցող $\mathcal{F}\mathcal{C}$ -երը ստացիոնար չեն: Այդ սիստեմատիկ բաղադրիչներից է սեզոնայնությունը, ուստի $\mathcal{F}\mathcal{C}$ ստացիոնարացման համար անհրաժեշտ է նաև սեզոնային բաղադրիչի հեռացումը $\mathcal{F}\mathcal{C}$ -ից:

Գոյություն ունեն սեզոնայնության հեռացման ամենատարբեր գործիքներ, որոնք ներդրված են էկոնոմետրիկ մոդելավորման համար օգտագործվող փաթեթներում:

Ապասեզոնայնացման ամենապարզ եղանակը $\mathcal{F}\mathcal{C}$ սեզոնային տարբերության կիրառումն է: Դրա էությունը հետևյալն է. եթե y_t $\mathcal{F}\mathcal{C}$ սեզոնայնության պարբերությունը հավասար է s -ի (օրինակ՝ եռամսյակային սեզոնայնության դեպքում $s=4$, ամսականի դեպքում՝ $s=12$ և այլն), ապա գեներացվում է նոր $\mathcal{F}\mathcal{C}$, որի յուրաքանչյուր t պահի անդամը որոշվում է հետևյալ կերպ՝ $y_t^* = y_t - y_{t-s}$:

Սեկ այլ եղանակ է $\mathcal{F}\mathcal{C}$ ապասեզոնայնացնումը՝ այն տարրալուծելով բաղադրիչների՝ միտում, սեզոնայնություն, պատահական բաղադրիչ: Մասնավորապես $\mathcal{F}\mathcal{C}$ -ից առանձնացվում է սեզոնայնությունը՝ ապասեզոնայնացնելով այն, սակայն պահպանելով մնացած բաղադրիչներն ու պատահական տատանումները: Սեզոնային բաղադրիչի որոշման ու $\mathcal{F}\mathcal{C}$ -ից հեռացման նպատակով յուրաքանչյուր ժամանակահատվածի համար հաշվարկվում է «սեզոնային ինդեքս»:

Ընդունված մոտեցում է նաև սահող միջինների եղանակով շարքի սեզոնայնության հարթեցման մոտեցումը: Ամենաընդհանուր դեպքում սահող միջիններով հարթեցումը ենթադրում է հարթեցման միջակայքի ℓ երկարության ընտրություն, որը համապատասխանում է սեզոնայնության պարբերությանը: $\mathcal{F}\mathcal{C}$ -ն բաժանվում է ℓ երկարությամբ միջակայքերի, և յուրաքանչյուր միջակայքի համար որոշվում է $\mathcal{F}\mathcal{C}$ միջինը: Այսպիսով՝ հարթեցման միջակայքը

«սահեցվում» է ԺՇ երկայնքով, և արդյունքում գեներացվում է նոր ԺՇ, որի յուրաքանչյուր անդամը սկզբնական շարքի յուրաքանչյուր ℓ թվով անդամների միջինն է: Օրինակ, եթե $\ell = 2p + 1$, այսինքն՝ կենտ թիվ է, ապա նոր ԺՇ յուրաքանչյուր անդամ կորոշվի հետևյալ բանաձևով՝ $\bar{y}_t = \frac{\sum_{i=t-p}^{t+p} y_i}{2p+1}$: Հաճախ օգտագործվում է նաև ԺՇ էքսպոնենցյալ հարթեցման մոտեցումը: Նշենք, որ հարթեցված շարքը կարճանում է:

7. Պատահական գործընթացի ու ժամանակային շարքի ավտոկոռելյացիայի գործակից, ավտոկոռելյացիայի ու մասնակի ավտոկոռելյացիայի ֆունկցիաներ, կոռելյոգրամ

Ինչպես արդեն ասվել է, $\{Y_t\}$ պատահական գործընթացը $Y_1, Y_2, \dots, Y_{t+\tau} \dots Y_n$ պատահական մեծությունների համախմբություն է՝ դասավորված ժամանակի մեջ: Այդ պատահական մեծությունները միմյանց հետ որոշակի վիճակագրական կապ, որոշակի կախվածություն ունեն: Նույնը վերաբերում է նաև պատահական գործընթացի իրականացմանը՝ ԺՇ-ին: Օրինակ՝ ՀՆԱ-ի 2025-ի մեծությունը որոշակի կախում ունի դրա նախորդ արժեքներից: Պատահական գործընթացը ձևավորող պատահական մեծությունների կամայական գույգի գծային կախվածության ուժգնությունը կարելի է չափել դրանց գույգային կովարիացիայի ու կոռելյացիայի գործակիցների միջոցով:

$\{Y_t\}$ պատահական գործընթացը ձևավորող իրարից կամայական τ լագով հեռացած երկու պատահական մեծությունների կովարիացիայի գործակիցը կանվանենք ավտոկովարիացիայի գործակից: Այն որոշվում է կովարիացիայի

$$Cov(Y_t, Y_{t+\tau}) = E[(Y_t - E(Y_t))(Y_{t+\tau} - E(Y_{t+\tau}))] \quad (136)$$

բանաձևով, իսկ դրանց գույգային կոռելյացիան՝ հետևյալ բանաձևով՝

$$r_{Y_t, Y_{t+\tau}} = \frac{E[(Y_t - E(Y_t))(Y_{t+\tau} - E(Y_{t+\tau}))]}{\sqrt{E(Y_t - E(Y_t))^2} \cdot \sqrt{E(Y_{t+\tau} - E(Y_{t+\tau}))^2}} = \rho(\tau): \quad (137)$$

Այսուհետ պատահական գործընթացի (137) բանաձևով որոշվող կոռելյացիայի գործակիցը կանվանենք τ -րդ կարգի ավտոկոռելյացիայի գործակից և կնշանակենք $\rho(\tau)$ -ով: Որպես կոռելյացիայի գործակից՝ $\rho(\tau) \in [-1; 1]$:

Ստացիոնար պատահական գործընթացի մաթ. սպասումը, վարիացիան ու կովարիացիան կախված չեն ժամանակից, կովարիացիան կախված է միայն ժամանակի τ լագից: Ուստի հաշվի առնելով (127)-(129)-ը՝ ստացիոնար պատահական գործընթացի համար կունենանք.

$$\begin{aligned} Cov(Y_t, Y_{t+\tau}) &= \mathbb{E}[(Y_t - \mathbb{E}(Y_t))(Y_{t+\tau} - \mathbb{E}(Y_{t+\tau}))] = \\ &= \mathbb{E}[(Y_t - \mu)(Y_{t+\tau} - \mu)] = \gamma(\tau): \end{aligned} \quad (138)$$

Ըստ որում՝ կովարիացիայի (138) բանաձևով որոշվող $\gamma(\tau)$ գործակիցը կախված չէ նրանից, քե պատահական մեծությունները ժամանակի մեջ որ ուղղությամբ են իրարից հեռացած, այսինքն՝ $\gamma(\tau) = \gamma(-\tau)$:

Մասնավորապես, եթե $\tau=0$, կունենանք պատահական գործընթացի վարիացիան.

$$Cov(Y_t, Y_t) = \mathbb{E}(Y_t - \mathbb{E}(Y_t))^2 = Var(Y_t) = \gamma(0): \quad (139)$$

Ստացիոնար պատահական գործընթացի՝ ժամանակի մեջ կամայական լագով իրարից հեռացած երկու պատահական մեծությունների ավտոկոռելյացիայի գործակիցը ևս կախված չէ ժամանակից. այն կախված է ժամանակի մեջ դրանց հեռավորությունից՝

$$\rho(\tau) = \frac{\mathbb{E}[(Y_t - \mu)(Y_{t+\tau} - \mu)]}{Var(Y_t)} = \frac{\gamma(\tau)}{\gamma(0)}: \quad (140)$$

Այսպիսով՝ ստացիոնար պատահական գործընթացի համար $\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$: Նշենք, որ եթե $\tau=0$, ապա $\rho(\tau) = \rho(0) = 1$, և $\rho(\tau) = \rho(-\tau)$:

Սակայն պատահական գործընթացն անհայտ է, հետագոտողին հասանելի է միայն դրա մեկ իրականացման մասին տեղեկատվությունը, ուստի այս թվային բնութագրիչները գործնականում հնարավոր է հաշվել $\mathcal{F}_t$ -ի համար: Դրանք կհանդիսանան պատահական

գործընթացի թվային բնութագրիչների վիճակագրական գնահատականները:

Ստացիոնար ԺՇ կամայական τ լագով իրարից հեռացած երկու անդամների ավտոկովարիացիայի գործակիցը որոշվում է հետևյալ բանաձևով՝

$$\text{cov}(y_t, y_{t+\tau}) = \frac{1}{n-\tau} \cdot \sum_{t=1}^{n-\tau} (y_t - \mu) \cdot (y_{t+\tau} - \mu) = \hat{\gamma}(\tau): \quad (141)$$

$\hat{\gamma}(\tau)$ -ն պատահական գործընթացի $\gamma(\tau)$ ավտոկովարիացիայի գործակցի վիճակագրական գնահատականն է: Ըստ որում՝ $\hat{\gamma}(\tau) = \hat{\gamma}(-\tau)$:

Ստացիոնար ԺՇ կամայական τ լագով իրարից հեռացած երկու անդամների զույգային ավտոկոռելյացիայի գործակիցը որոշվում է նմուշային զույգային կոռելյացիայի բանաձևով՝

$$\begin{aligned} r_{y_t, y_{t+\tau}} &= \frac{\text{cov}(y_t, y_{t+\tau})}{\sqrt{\text{var}(y_t)} \cdot \sqrt{\text{var}(y_{t+\tau})}} = \frac{\text{cov}(y_t, y_{t+\tau})}{\text{var}(y_t)} = \\ &= \frac{1}{n-\tau} \cdot \frac{\sum_{t=1}^{n-\tau} (y_t - \mu) \cdot (y_{t+\tau} - \mu)}{\frac{1}{n} \cdot \sum_{t=1}^n (y_t - \mu)^2} = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)} = \hat{\rho}(\tau): \quad (142) \end{aligned}$$

$\hat{\rho}(\tau)$ -ն պատահական գործընթացի $\rho(\tau)$ ավտոկոռելյացիայի գործակցի վիճակագրական գնահատականն է:

Ենթադրենք (141) բանաձևով հաշվարկվել են ԺՇ-ի $\tau=1, 2, 3, \dots, k$ լագով իրարից հեռացած երկու անդամների ավտոկովարիացիայի գործակիցները՝ $\hat{\gamma}(1)$, $\hat{\gamma}(2)$, $\hat{\gamma}(3)$, ..., $\hat{\gamma}(k)$: Դրանք համապատասխանաբար կկոչենք ԺՇ 1-ին, 2-րդ, 3-րդ ... k -րդ կարգի ավտոկովարիացիաների գործակիցներ:

Սահմանում. ԺՇ ավտոկովարիացիայի ֆունկցիան դրա 1-ին, 2-րդ, 3-րդ ... k -րդ կարգի ավտոկովարիացիաների գործակիցների համախմբությունն է՝ որպես ֆունկցիա $\tau=1, 2, \dots, k$ լագերից:

ԺՇ ավտոկովարիացիայի ֆունկցիան պատահական գործընթացի անհայտ ավտոկովարիացիայի ֆունկցիայի վիճակագրական գնահատականն է:

Ենթադրենք՝ (142) բանաձևով հաշվարկված են ԺՇ-ի 1-ին, 2-րդ, 3-րդ ... k -րդ կարգի ավտոկոռելյացիաների գործակիցները՝ $\hat{\rho}(1)$, $\hat{\rho}(2)$, $\hat{\rho}(3)$, ..., $\hat{\rho}(k)$:

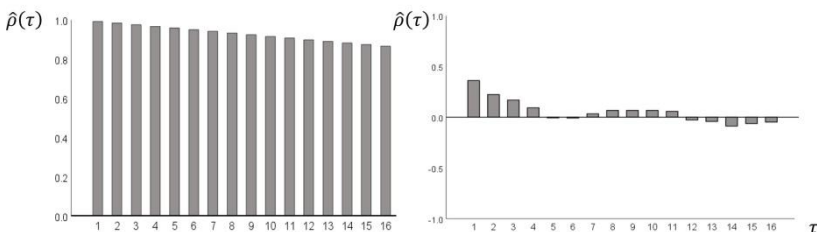
Սահմանում. ԺՇ ավտոկոռելյացիայի ֆունկցիան (Autocorrelation function, ACF) դրա 1-ին, 2-րդ, ... k -րդ կարգի զույ-

գային ավտոկոռելյացիայի գործակիցների համախմբությունն է՝ որպես ֆունկցիա $\tau=1,2,\dots,k$ լագերից:

ԺՇ ավտոկոռելյացիայի ֆունկցիան պատահական գործընթացի անհայտ ավտոկոռելյացիայի ֆունկցիայի վիճակագրական գնահատականն է: Ավտոկոռելյացիայի ֆունկցիայի գրաֆիկական ներկայացումը, կախված լագերից, կոչվում է կոռելոգրամ:

Ավտոկոռելյացիայի ֆունկցիան ԺՇ ու պատահական գործընթացի կարևորագույն բնութագրիչներից է, և դրա միջոցով կարելի է նախնական եզրակացություն կատարել ԺՇ և այն գեներացրած պատահական գործընթացի ստացիոնարության մասին: Ավտոկոռելյացիայի ֆունկցիան բնութագրում է, թե շարքի յուրաքանչյուր հաջորդ անդամ որքանով է կոռելացված, որքանով է «հիշում» իր նախորդ անդամներին, կրում դրանց ազդեցությունը: Եթե ԺՇ-ն ստացիոնար չէ, ապա այն ունի երկար «հիշողություն», ինչը նշանակում է, որ ժամանակի մեջ իրարից ավելի ու ավելի հեռացող անդամների կոռելյացիաներն ու ԺՇ ավտոկոռելյացիայի ֆունկցիան լագերի աճի հետ դանդաղ են նվազում: Բանի որ ստացիոնար ԺՇ-ն ունի ժամանակի մեջ հաստատուն բնութագրիչներ, այն պետք է ունենա կարճ «հիշողություն», ինչը նշանակում է, որ իրարից ժամանակի մեջ ավելի ու ավելի հեռացող անդամների կոռելյացիաները, միմյանց վրա ունեցած ազդեցությունը պետք է արագ նվազեն: Ստացիոնար ԺՇ ավտոկոռելյացիայի ֆունկցիան լագերի աճի հետ պետք է արագ, էքսպոնենցիալ օրենքով կամ նշանը փոխելով նվազի և սկսի տատանվել 0-ի շուրջը: Արագ չմարող, չնվազող ավտոկոռելյացիայի ֆունկցիան կվկայի ԺՇ ոչ ստացիոնարության մասին: Գծանկար 29-ում ներկայացված են ոչ ստացիոնար (ա) և ստացիոնար (բ) ԺՇ-երի կոռելոգրամների օրինակներ:

Գծանկար 29. Ոչ ստացիոնար և ստացիոնար ԺՇ-երի կոռելոգրամներ
ա) բ)



Սպիտակ աղմուկ պատահական գործընթացի բոլոր կարգի ավտոկովարիացիայի գործակիցները հավասար են 0-ի՝ ըստ սպիտակ աղմուկի սահմանման: Դա նշանակում է, որ դրա ավտոռելյացիայի գործակիցները, որոնց համարիչում ավտոկովարիացիաներ են, հավասար կլինեն 0-ի բոլոր $\tau=1,2,\dots, k$ լագերի համար: Իսկ եթե $\tau=0$, ավտոռելյացիայի գործակիցը հավասար կլինի 1-ի: Նշանակում է՝ սպիտակ աղմուկ $\mathcal{C}$ կամայական τ լագով իրարից հեռացած երկու անդամները իրար հետ կոռելացված չեն (ինչը բխում է նաև սպիտակ աղմուկի սահմանումից):

Կարելի է հաշվել $\mathcal{C}$ անդամների ոչ միայն գույգային կոռելյացիաների գործակիցները և ստանալ դրա ավտոկոռելյացիայի ֆունկցիան, այլ նաև հաշվարկել $\mathcal{C}$ անդամների մասնակի ավտոկոռելյացիայի գործակիցները (տես՝ Ներդիր 14): Դրանք ցույց են տալիս $\mathcal{C}$ կամայական τ լագով իրարից հեռացած անդամների «մաքուր», միջանկյալ անդամների միջնորդավորված ազդեցություններից զերծ գծային կապի ուժգնությունը:

Սահմանում. $\mathcal{C}$ մասնակի ավտոկոռելյացիայի ֆունկցիան (Partial Autocorrelation function, PACF) դրա 1-ին, 2-րդ, ... k -րդ կարգի մասնակի ավտոկոռելյացիայի գործակիցների համախմբությունն է՝ որպես ֆունկցիա $\tau=1,2,\dots, k$ լագերից:

$\mathcal{C}$ մասնակի ավտոկոռելյացիայի ֆունկցիան տարբերվում է դրա ավտոկոռելյացիայի ֆունկցիայից, որը բնութագրող ավտոկոռելյացիայի գործակիցներն արտահայտում են $\mathcal{C}$ որևէ լագով իրարից հեռացված անդամների ընդհանուր կոռելյացիան՝ ներառյալ ինչպես ուղղակի, այնպես էլ անուղղակի, միջանկյալ լագերի միջոցով փոխանցվող ազդեցությունները:

8. Պատահական գործընթացի ավտոկոռելյացիայի գործակիցների մասին վիճակագրական վարկածների ստուգում

Ինչպես արդեն նշվել է, պատահական գործընթացի ավտոկոռելյացիայի ֆունկցիան անհայտ է, իսկ դրա իրականացումը հանդիսացող $\mathcal{C}$ ավտոկոռելյացիայի ֆունկցիան կարելի է հաշվարկել: Այն կներկայացնի անհայտ պատահական գործընթացի ավտոկոռելյացիայի ֆունկցիան, քանի որ դրա վիճակագրական գնահատականն է: Ստացիոնարության ստուգման տեսանկյունից մեզ հետաքրքրում է, որ պատահական գործընթացի ավտոկոռելյացիայի գործակիցներ-

րը լազերի աճի հետ արագ նվազեն և հավասարվեն գրոյի: Սակայն հետազոտողը կարող է տեսնել միայն σ ավտոկոռելյացիաները: Հետևաբար անհրաժեշտ են մեխանիզմներ σ ավտոկոռելյացիայի գործակիցների միջոցով պատահական գործընթացի ավտոկոռելյացիայի գործակիցների 0-ից նշանակալիորեն չտարբերվելու փաստը ստուգելու համար: Եթե պատահական գործընթացի ավտոկոռելյացիայի գործակիցը 0-ից նշանակալիորեն չի տարբերվում, ապա σ համապատասխան ավտոկոռելյացիայի գործակիցը վիճակագրորեն նշանակալի չէ: Պատահական գործընթացի ավտոկոռելյացիայի գործակիցների 0-ից նշանակալիորեն չտարբերվելու փաստը կամ σ ավտոկոռելյացիայի գործակցի վիճակագրական նշանակալիությունը կարելի է ստուգել վիճակագրական քեստով:

Ապացուցված է, որ սպիտակ աղմուկ մաքուր պատահական, ստացվող գործընթացը ներկայացնող σ ընտրանքային ավտոկոռելյացիաները մեծ n -երի համար ունեն նորմալ բաշխում՝ 0 միջինով և $1/n$ վարիացիայով (n -ը ժամանակային շարքի երկարությունն է)

$$\hat{\rho}(\tau) \sim N(0, 1/n):$$

Ստանդարտ նորմալ բաշխման հատկություններից ելնելով՝ դա կնշանակի, որ այդ σ -ն գեներացրած, անհայտ պատահական գործընթացի անհայտ $\rho(\tau)$ ավտոկոռելյացիայի գործակիցը 95 % հավանականությամբ ընկած կլինի հետևյալ միջակայքում՝

$$\hat{\rho}(\tau) - t_{\alpha} \cdot \sqrt{\frac{1}{n}} \leq \rho(\tau) \leq \hat{\rho}(\tau) + t_{\alpha} \cdot \sqrt{\frac{1}{n}}: \quad (143)$$

Այս արդյունքը հնարավորություն է տալիս ստուգելու պատահական գործընթացի անհայտ ավտոկոռելյացիայի գործակցի՝ գրոյին հավասար լինելու կամ σ համապատասխան գործակցի վիճակագրական նշանակալիության մասին վարկածը:

Դրվում է գրոյական վարկած այն մասին, որ σ -ն գեներացրած անհայտ պատահական գործընթացի տվյալ կարգի ավտոկոռելյացիայի գործակիցը՝ $\rho(\tau)$ -ն, իրականում հավասար է գրոյի: Վարկածը դրվում է առանձին-առանձին $\tau=1, 2, \dots, k$ կարգի ավտոկոռելյացիայի գործակիցների համար՝

$$H_0: \rho(\tau) = 0; H_1: \rho(\tau) \neq 0:$$

Եթե տվյալ τ -ի համար՝

$$-t_{\text{կր}} \cdot \sqrt{\frac{1}{n}} \leq \hat{\rho}(\tau) \leq t_{\text{կր}} \cdot \sqrt{\frac{1}{n}},$$

ապա H_0 -ն ընդունվում է, հետևաբար 95 % հավանականությամբ կարելի է պնդել, որ պատահական գործընթացի տվյալ կարգի ավտոկոռելյացիայի անհայտ $\rho(\tau)$ գործակիցը 0-ից նշանակալիորեն չի տարբերվում, իսկ $\mathcal{F}\mathcal{C}$ -ով որոշված $\hat{\rho}(\tau)$ գործակիցը վիճակագրորեն նշանակալի չէ: Հակառակ դեպքում, եթե $\hat{\rho}(\tau)$ -ն ընկած է նշված միջակայքից դուրս, ապա H_0 -ն մերժվում է, և 95 % հավանականությամբ կարելի է պնդել, որ պատահական գործընթացի տվյալ կարգի ավտոկոռելյացիայի անհայտ գործակիցը նշանակալիորեն տարբերվում է 0-ից, և $\mathcal{F}\mathcal{C}$ -ով որոշված $\hat{\rho}(\tau)$ գործակիցը վիճակագրորեն նշանակալի է:

Հնարավոր է մաս ստուգել ոչ թե առանձին-առանձին $\tau=1,2,\dots,k$ կարգի ավտոկոռելյացիայի գործակիցների վիճակագրական նշանակալիությունը, այլ պատահական գործընթացի առաջինից մինչև φ -րդ կարգի բոլոր ավտոկոռելյացիայի գործակիցների վիճակագրական նշանակալիությունը: Այս դեպքում դրվում է գրոյական վարկած այն մասին, որ $\mathcal{F}\mathcal{C}$ -ն գեներացրած անհայտ պատահական գործընթացի առաջինից մինչև φ -րդ կարգի բոլոր ավտոկոռելյացիայի գործակիցները իրականում հավասար են զրոյի՝

$$H_0: \rho(1) = \rho(2) = \dots = \rho(\varphi) = 0:$$

Մրցակցող վարկածն այն է, որ դրանցից առնվազն մեկը հավասար չէ զրոյի: Սա ստուգվում է **Քոքս-Լյունգի Q վիճականու** միջոցով, որը որոշվում է հետևյալ բանաձևով և ունի ասիմպտոտիկ χ^2 բաշխում φ ազատության աստիճանով՝

$$Q = n(n+2) \sum_{\tau=1}^{\varphi} \frac{(\hat{\rho}(\tau))^2}{n-\tau}: \quad (144)$$

Եթե $Q > \chi_{\text{կր}}^2$, ապա H_0 -ն մերժվում է, և պատահական գործընթացի առաջինից մինչև φ -րդ կարգի ավտոկոռելյացիայի գործակիցներից առնվազն մեկը իրականում զրոյից նշանակալիորեն տարբերվում է: Եթե $Q \leq \chi_{\text{կր}}^2$, ապա H_0 -ն չի մերժվում, ընդունվում է, ինչը նշանակում է, որ ընդունվում է $\rho(1) = \rho(2) = \dots = \rho(\varphi) = 0$ վարկածը:

Թեստի ելքը որոշելու համար կարելի է նաև օգտագործել p արժեքը: Եթե $p\text{-value} < \alpha$, մինչև p -րդ կարգի ավտոկորելյացիայի բացակայության մասին զրոյական վարկածը պետք է մերժել α արժեքականության մակարդակով, հակառակ դեպքում՝ ընդունել այն: Նշենք, որ այս թեստը հիմնականում կիրառվում է ԺՇ սպիտակ աղմուկ լինելու փաստը ստուգելու համար:

9. Դիկի-Ֆուլերի՝ ժամանակային շարքի ստացիոնարության ստուգման միավոր արմատի վիճակագրական թեստերը

1. Միավոր արմատի Դիկի-Ֆուլերի թեստ (DF թեստ)

Այս թեստի հիմքում ընկած են միավոր արմատ ԺՇ հատկությունները: Ենթադրենք ԺՇ-ն ունի հետևյալ տեսքը՝

$$y_t = \rho \cdot y_{t-1} + \varepsilon_t, \quad \rho = 1, \quad (145)$$

որում ε_t -ն սպիտակ աղմուկ ստացիոնար ԺՇ է: Այս շարքը առանց տեղաշարժի պատահական քայլքի մոդելով արտահայտվող ԺՇ է՝ միավոր արմատ ԺՇ: Արդեն իսկ հայտնի է, որ այն ստացիոնար չէ և ունի ստոխաստիկ տրենդ: Հետևաբար կարելի է ԺՇ-ն արտահայտել այս մոդելով և ստուգել, թե արդյոք $\rho=1$: Եթե այո, ապա ԺՇ-ն ստացիոնար չէ, այն RW է: Ձևափոխենք (145) հավասարումը՝ հավասարման երկու կողմից հանելով y_{t-1} :

$$y_t - y_{t-1} = \rho \cdot y_{t-1} - y_{t-1} + \varepsilon_t = (\rho - 1) \cdot y_{t-1} + \varepsilon_t: \quad (146)$$

(146) -ը կարելի է գրել այսպես՝

$$\Delta(y_t) = \delta \cdot y_{t-1} + \varepsilon_t, \quad (147)$$

որում $\Delta(y_t)$ -ն y_t ԺՇ-ի առաջին տարբերությունն է, իսկ $\delta = \rho - 1$:

Այդ դեպքում (147)-ը կարելի է գնահատել ՓԶԵ-ով և ստանալ δ -ը, այնուհետև՝ ստուգել վիճակագրական վարկած, թե արդյոք $\delta = 0$: Եթե $\delta = 0$, ապա $\rho = 1$, և ԺՇ-ն ստացիոնար չէ: Այսպիսով՝ դրվում է հետևյալ զրոյական վարկածը՝

$$H_0: \delta = 0:$$

$$\text{Մրցակցող վարկածն է՝ } H_1: \delta < 0:$$

Այս թեստը միակողմանի է: Չրոյական վարկածը համարժեք է y_t ԺՇ ոչ ստացիոնարության կամ միավոր արմատ ԺՇ լինելու վարկածին: Եթե զրոյական վարկածը ընդունվում է, ապա $\rho = 1$, և ԺՇ-ն ստացիոնար չէ. այն առաջին կարգի ինտեգրացվող է՝ $I(1)$ է: Եթե զրոյական վարկածը մերժվում է, և ընդունվում է մրցակցողը, այսինքն՝ $\delta < 0, \rightarrow |\rho| < 1$, ապա ԺՇ-ն ստացիոնար է՝ $I(0)$ է:

Սովորական t վիճականու միջոցով այս վարկածը ստուգել հնարավոր չէ, քանի որ այս զրոյական վարկածի պարագայում այն չունի Ստյուդենտի բաշխում: Դիկին ու Ֆուլերը ցույց են տվել, որ զրոյական $\delta = 0$ վարկածը կարելի է ստուգել τ (տաու) վիճականու միջոցով, որի կրիտիկական արժեքները նրանք հաշվարկել են (կոչվում են Դիկի-Ֆուլերի կրիտիկական արժեքներ, իսկ թեստը՝ Դիկի-Ֆուլերի թեստ): t վիճականու սովորական բանաձևով ($t_{\eta_{\text{խ}}} = \frac{\hat{\delta}}{s_{\hat{\delta}}}$) հաշվարկվում է դրա դիտարկված արժեքը, ընտրվում է α արժեքականության մակարդակը, և $t_{\eta_{\text{խ}}}$ -ը համեմատվում է Դիկի-Ֆուլերի կրիտիկական արժեքի հետ: Եթե $t_{\eta_{\text{խ}}} < \tau_{\text{կր}}$, ապա H_0 -ն մերժվում է. ԺՇ-ն ստացիոնար, $I(0)$ շարք է: $t_{\eta_{\text{խ}}} > \tau_{\text{կր}}$, H_0 -ն ընդունվում է, ԺՇ-ն ստացիոնար չէ, $I(1)$ է:

Դիկի-Ֆուլերի թեստը գնահատվում է երեք տարբերակով երեք տարբեր տեսքի մոդելների համար՝

1. y_t ԺՇ-ն կարելի է ներկայացնել առանց տեղաշարժի RW-ով՝

$$\Delta(y_t) = \delta \cdot y_{t-1} + \varepsilon_t,$$
2. y_t ԺՇ-ն կարելի է ներկայացնել տեղաշարժով RW-ով՝

$$\Delta(y_t) = \beta_0 + \delta \cdot y_{t-1} + \varepsilon_t,$$
3. y_t ԺՇ-ն կարելի է ներկայացնել տեղաշարժով RW-ով, որը նաև ունի գծային դետերմինացված տրենդ՝ $\Delta(y_t) = \beta_0 + \beta_1 \cdot t + \delta \cdot y_{t-1} + \varepsilon_t$:

Յուրաքանչյուր դեպքում առաջադրվում է վերը բերված զրոյական վարկածը և ստուգվում արդեն իսկ նշված ձևով:

Այս թեստը հնարավորություն է տալիս ստուգելու ինչպես ստոխաստիկ, այնպես էլ գծային դետերմինացված տրենդի առկայությունը ԺՇ-ում, ինչպես նաև ստուգել, թե ԺՇ-ն որ կարգի ինտեգրացվող է տարբերական օպերատորը մի քանի անգամ ԺՇ նկատմամբ կիրառելու և թեստը նորից կիրառելու միջոցով:

Այս թեստն ունի կիրառման որոշ առանձնահատկություններ ու սահմանափակումներ, ինչի պատճառով կիրառվում է ոչ թե DF թեստը, այլ Դիկի-Ֆուլերի ընդհանրացված ADF թեստը:

Դիֆի-Ֆոլերի ընդհանրացված թեստը (ADF թեստ)

DF թեստը ենթադրում է, որ գնահատվող $\Delta(y_t) = \delta \cdot y_{t-1} + \varepsilon_t$ մոդելում ε_t -ն սպիտակ աղմուկ $\mathcal{O}(1)$ է, այսինքն՝ դրա կամայական կարգի ավտոկոռելյացիաների գործակիցները հավասար են զրոյի: Իրականում այս պայմանը կարող է չբավարարվել: Այդ պատճառով կիրառվում է Դիֆի-Ֆոլերի ընդհանրացված թեստը (ADF թեստ), որում այս խնդիրը լուծվել է:

ADF թեստում ենթադրվում է, որ y_t $\mathcal{O}(1)$ -ն արտահայտվում է հետևյալ մոդելով՝

$$\Delta(y_t) = \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t: \quad (148)$$

Այս թեստը ևս գնահատվում է երեք տարբերակով երեք տարբեր տեսքի մոդելների համար՝

1. y_t $\mathcal{O}(1)$ -ն կարելի է ներկայացնել առանց տեղաշարժի RW-ով՝

$$\Delta(y_t) = \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t:$$

2. y_t $\mathcal{O}(1)$ -ն կարելի է ներկայացնել տեղաշարժով RW-ով՝

$$\Delta(y_t) = \beta_0 + \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t:$$

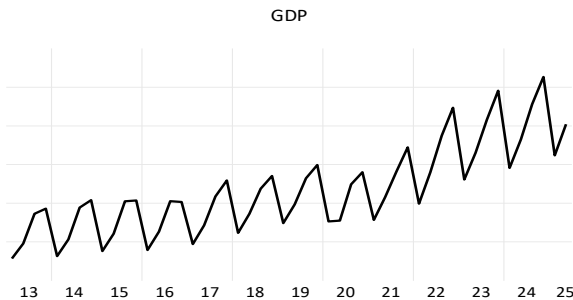
3. y_t $\mathcal{O}(1)$ -ն կարելի է ներկայացնել տեղաշարժով RW-ով, որը մաս ունի գծային դետերմինացված տրենդ՝

$$\Delta(y_t) = \beta_0 + \beta_1 \cdot t + \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t:$$

Թեստը օգտագործում է նույն վարկածը, կրիտիկական արժեքներն ու ստուգման նույն մեխանիզմը, ինչ DF թեստը: Մասնագիտական գրականությունում խորհուրդ է տրվում $\mathcal{O}(1)$ ստացիոնարության ստուգման ADF թեստը սկսելու ամենաբարդ՝ 3-րդ տարբերակից՝ ստուգելու համար թե՛ դետերմինիստիկ, թե՛ ստոխաստիկ միտման առկայությունը, այնուհետև, եթե ոչ ստացիոնարությունը չի ժխտվում, իրականացնել 2-րդ ու 1-ին տարբերակները:

10. Ժամանակային շարքի նախնական ուսումնասիրության օրինակ

Ենթադրենք ուսումնասիրվում է ՀՀ արտադրական եղանակով հաշվարկված ՀՆԱ ԺՇ-ն: Առկա է ԺՇ 50 տվյալ 2013-2025 թթ. համար՝ եռամսյակային կտրվածքով: Մոդելավորման համար ԺՇ նախապատրաստական ուսումնասիրությունը սկսվում է դրա գրաֆիկից: Այն փաստում է սեզոնայնության ու աճման միտումի առկայությունը մասին:



Ակնհայտորեն սեզոնայնության պարբերականությունը չորսն է. այն կրկնվում է չորս եռամսյակը մեկ: Էկոնոմետրիկ մոդելավորման տարբեր փաթեթներ հնարավորություն ունեն առանց լրացուցիչ կոդերի օգտագործման կառուցելու ԺՇ ավտոկոռելյացիայի ֆունկցիան:

Օգտվելով EViews փաթեթի հնարավորությունից՝ կառուցվել են այս ԺՇ ավտոկոռելյացիայի (AC) ու մասնակի ավտոկոռելյացիայի (PAC) ֆունկցիաները $\tau=1,2,\dots,24$ լագերի համար, ինչպես նաև ստացվել են դրանց կոռելյացիաները:

Ավտոկոռելյացիաները արագ չեն նվազում, ինչպես նաև դրսևորում են որոշակի սեզոնայնություն: Կոռելյացիանում կետագծերով նշված են ԺՇ-ով որոշված $\hat{\rho}(\tau)$ գործակցի վիճակագրական նշանակալիության $\pm 1.96 \cdot \sqrt{\frac{1}{n}}$ միջակայքի եզրերը:

Մինչև 8 լագով (8 եռամսյակով կամ 2 տարի), ինչպես նաև 12 լագով իրարից հեռացած անդամների՝ AC սյունում բերված ավտոկոռելյացիայի $\hat{\rho}(\tau)$ գործակիցներն ընկած են այդ միջակայքից դուրս: Դա փաստում է, որ դրանք վիճակագրորեն նշանակալի են, և պատահական գործընթացի համապատասխան ավտոկոռելյացիայի գործակիցները 0-ից նշանակալիորեն տարբերվում են: Բոքս-Լյունգի

Q վիճականու (Q-Stat) արժեքները ևս հաշվարկված են բոլոր τ լագերի համար: Դրանց p արժեքները փաստում են, որ մինչև 24-րդ կարգի ավտոկոռելյացիայի բացակայության մասին գրոյական վարկածը պետք է մերժել՝ $\alpha=0.05$ արժեքականության մակարդակով կամ, որ նույնն է. 95 % հավանականությամբ: Այս բոլոր վերլուծությունները փաստում են շարքի ոչ ստացիոնարության մասին:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.635	0.635	21.378	0.000
		2	0.437	0.057	31.708	0.000
		3	0.536	0.401	47.604	0.000
		4	0.774	0.582	81.477	0.000
		5	0.443	-0.545	92.799	0.000
		6	0.256	-0.029	96.680	0.000
		7	0.336	-0.013	103.51	0.000
		8	0.545	0.105	121.90	0.000
		9	0.252	-0.137	125.91	0.000
		10	0.080	-0.040	126.34	0.000
		11	0.151	-0.028	127.86	0.000
		12	0.336	-0.007	135.59	0.000
		13	0.088	-0.010	136.14	0.000
		14	-0.055	-0.004	136.36	0.000
		15	0.019	0.004	136.39	0.000
		16	0.200	0.031	139.44	0.000
		17	-0.008	-0.030	139.44	0.000
		18	-0.139	-0.048	141.02	0.000
		19	-0.063	0.013	141.35	0.000
		20	0.111	0.004	142.41	0.000
		21	-0.059	0.010	142.73	0.000
		22	-0.178	-0.030	145.66	0.000
		23	-0.132	-0.137	147.34	0.000
		24	0.017	-0.036	147.37	0.000

Այնուամենայնիվ, վերջնական եզրակացությունը պետք է կայացվի վիճակագրական թեստերի իրականացման միջոցով: ADF թեստի իրականացումը սկսենք 3-րդ տարբերակից՝ ստուգելու համար դետերմինիստիկ ու ստոխաստիկ միտումների առկայությունը:

Թեստի գրոյական վարկածը D ոչ ստացիոնարությունն է: Թեստի վիճականին՝ $t_{\eta_{\text{խտ}}} = -1.442$ (տե՛ս հաջորդ էջի աղյուսակը):

Դրա ներքևում բերված են ԴՖ բաշխման կրիտիկական արժեքները՝ $\alpha=0.01$, $\alpha=0.05$, $\alpha=0.1$ արժեքականության մակարդակների համար: Դրանց բոլորի համար էլ $t_{\eta\text{իտ}} > \tau_{կր}$: Հետևաբար զրոյական վարկածը չենք կարող մերժել. այն ընդունվում է, ԺՇ-ն ստացիոնար չէ: ԺՇ ոչ ստացիոնարության մասին է փաստում մասն քեստի վիճականուր արժեքը. Prob.=p-value=0.835>0.05, հետևաբար զրոյական վարկածը մերժելիս սխալվելու հավանականությունը մեծ է, և այն չենք կարող մերժել:

Null Hypothesis: GDP has a unit root

Exogenous: Constant, Linear Trend

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.442	0.835
Test critical values:	1% level	-4.176	
	5% level	-3.513	
	10% level	-3.187	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(GDP)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP(-1)	-0.138	0.096	-1.442	0.158
D(GDP(-1))	-0.205	0.149	-1.371	0.178
D(GDP(-2))	-0.222	0.140	-1.578	0.123
D(GDP(-3))	-0.279	0.129	-2.159	0.037
D(GDP(-4))	0.730	0.125	5.824	0.000
C	113856.414	84045.588	1.355	0.184
@TREND("2013Q1")	5807.326	3143.763	1.847	0.073
R-squared	0.967	Akaike info criterion	25.839	
Adjusted R-squared	0.962	Schwarz criterion	26.120	
F-statistic	184.403	Hannan-Quinn crit.	25.944	
Prob(F-statistic)	0.000	Durbin-Watson stat	2.106	

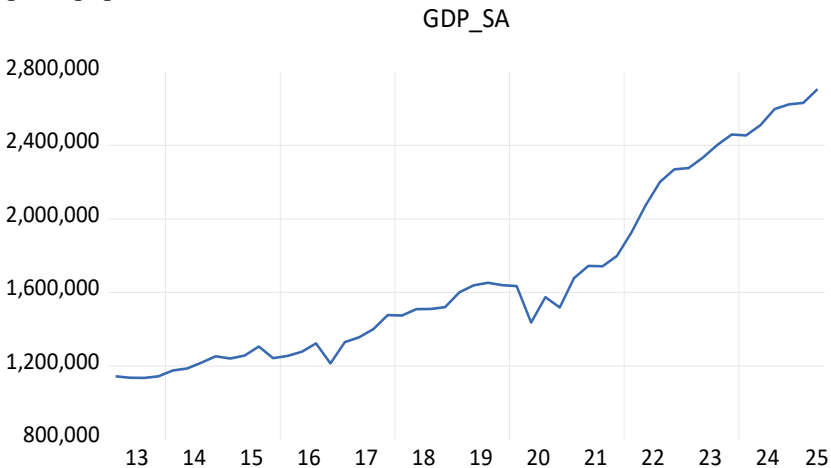
Հիշենք, որ ADF թեստի 3-րդ տարբերակը գնահատում է հետևյալ տեսքի մոդել.

$$\Delta(y_t) = \beta_0 + \beta_1 \cdot t + \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t:$$

Եթե $\delta=0$, ապա $\rho = 1$, և շարքն ունի ստոխաստիկ միտում: Եթե $\beta_1 \neq 0$, ապա շարքն ունի դետերմինիստիկ գծային միտում:

Հետևաբար պետք է ստուգել նաև ADF թեստի՝ ՓԲԵ-ով գնահատված մոդելի գործակիցների վիճակագրական նշանակալիությունը: Մոդելի գնահատման արդյունքները փաստում են, որ δ -ն տեսական մոդելում նշանակալիորեն զրոյից չի տարբերվում, և $\hat{\delta}$ -ը (GDP(-1)-ի գործակիցը) վիճակագրորեն նշանակալի չէ, դրա t վիճակահանին՝ $|t_{\text{դիտ}}| = |-1.441504| < 1.96$: Այսինքն՝ շարքն ունի ստոխաստիկ միտում: Նույն կերպ գծային միտումի (@TREND) գործակիցը ևս վիճակագրորեն նշանակալի չէ, այսինքն՝ տեսական մոդելում $\beta_1 = 0$: Հետևաբար շարքը չունի գծային դետերմինիստիկ միտում: Ստոխաստիկ միտումի առկայությունը փաստվում է նաև ADF թեստի մյուս երկու տարբերակների իրականացմամբ:

Ուսումնասիրությունը շարունակելու համար որևէ մեթոդով շարքից հեռացնենք միայն սեզոնայնությունը՝ ապասեզոնայնացնենք ԺՇ-ն: Ապասեզոնայնացված շարքի աճի միտումն ավելի հստակորեն է երևում, իսկ պարբերական տատանումները հարթվել են: Ապասեզոնայնացված ԺՇ գրաֆիկը ներկայացված է ստորև՝ GDP_SA գծանկարում:



Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.933	0.933	46.208	0.000
		2	0.870	-0.007	87.198	0.000
		3	0.799	-0.091	122.53	0.000
		4	0.727	-0.053	152.42	0.000
		5	0.660	-0.004	177.55	0.000
		6	0.593	-0.032	198.30	0.000
		7	0.521	-0.084	214.70	0.000
		8	0.452	-0.028	227.33	0.000
		9	0.386	-0.017	236.78	0.000
		10	0.323	-0.025	243.56	0.000
		11	0.262	-0.039	248.13	0.000
		12	0.199	-0.059	250.85	0.000
		13	0.145	0.015	252.33	0.000
		14	0.103	0.048	253.10	0.000
		15	0.067	0.010	253.44	0.000
		16	0.032	-0.051	253.51	0.000
		17	-0.003	-0.037	253.51	0.000
		18	-0.032	0.014	253.60	0.000
		19	-0.048	0.067	253.79	0.000
		20	-0.068	-0.069	254.20	0.000

Ստանանք ապաստեղծայնացված ԺՇ ավտոկոռելյացիայի ֆունկցիան: Այն ևս դանդաղ է նվազում: Մինչև 10 լագով իրարից հեռացած անդամների՝ AC սյունում բերված ավտոկոռելյացիայի $\hat{\rho}(\tau)$ գործակիցներն ընկած են վիճակագրական նշանակալիության $\pm 1.96 \cdot \sqrt{\frac{1}{n}}$ միջակայքից դուրս: Դա փաստում է, որ դրանք վիճակագրորեն նշանակալի են, և պատահական գործընթացի համապատասխան ավտոկոռելյացիայի գործակիցները 0-ից նշանակալիորեն տարբերվում են: Բոքս-Լյունգի Q վիճականու (Q-Stat) արժեքների p արժեքները փաստում են, որ մինչև 20-րդ կարգի ավտոկոռելյացիայի բացակայության մասին զրոյական վարկածը պետք է մերժել $\alpha=0.05$ արժեքականության մակարդակով: Այս վերլուծությունները փաստում են շարքի ոչ ստացիոնարության մասին:

Նորից կիրառենք ADF թեստն արդեն ապաստեղծայնացված ԺՇ-ի համար՝ սկսելով 3-րդ տարբերակից: Դրա արդյունքները փաստում են, որ այս ԺՇ-ն ևս ստացիոնար չէ՝ ելնելով t վիճականու ու p արժեքի մեծություններից: ADF թեստի՝ մոդելի գնահատման մասը ևս փաստում է այն մասին, որ այս ԺՇ-ն ունի ստոխաստիկ միտում $\delta=0$, հետևաբար $\rho = 1$ վարկածը չի կարող մերժվել. ընդունվում է:

Null Hypothesis: GDP_SA has a unit root
Exogenous: Constant, Linear Trend

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.241	0.891
Test critical values:	1% level	-4.157	
	5% level	-3.504	
	10% level	-3.182	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GDP_SA)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP_SA(-1)	-0.067	0.054	-1.241	0.221
C	60304.631	50776.185	1.188	0.241
@TREND("2013Q1")	3298.976	1778.019	1.855	0.070
R-squared	0.108	Akaike info criterion		24.918
Adjusted R-squared	0.069	Schwarz criterion		25.033
F-statistic	2.788	Hannan-Quinn criter.		24.961
Prob(F-statistic)	0.072	Durbin-Watson stat		2.279

Այսինքն՝ շարքն ունի ստոխաստիկ միտում: Նույն կերպ գծային միտումի (@TREND) գործակիցը ևս վիճակագրորեն նշանակալի չէ, այսինքն՝ տեսական մոդելում $\beta_1 = 0$: Հետևաբար շարքը չունի գծային դետերմինիստիկ միտում: Ստոխաստիկ միտումի առկայությունը փաստում են նաև ADF թեստի մյուս երկու տարբերակների արդյունքները:

Այսպիսով՝ ապաստեգոնայնացված շարքն ունի ստոխաստիկ միտում և չունի դետերմինացված միտում: Նշանակում է՝ այս ԺՇ ստացիոնարացնելու համար անհրաժեշտ է դրա նկատմամբ կիրառել տարբերական օպերատորը:

Null Hypothesis: D(GDP_SA) has a unit root
Exogenous: Constant, Linear Trend

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.046	0.000
Test critical values:	1% level	-4.161	
	5% level	-3.506	
	10% level	-3.183	

*MacKinnon (1996) one-sided p-values.

Տարբերական օպերատորի կիրառումից հետո ՀՆԱ-ի ապասե-
զոնայնացված ԺՇ ADF թեստը 2-րդ տարբերակով փաստում է, որ
շարքն արդեն ստացիոնար է: Թեստի վիճականին՝ $t_{\eta_{\text{իտ}}} = -8.046$: Այն
փոքր է Դիկի-Ֆուլերի բաշխման երեք կրիտիկական արժեքներից՝
 $t_{\eta_{\text{իտ}}} < \tau_{\eta_{\text{իտ}}}$, $\alpha=0.01$, $\alpha=0.05$ $\alpha=0.1$ արժեքականության մակարդակնե-
րում: Հետևաբար զրոյական վարկածը մերժվում է, այս ԺՇ-ն ստա-
ցիոնար է:

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.096	-0.096	0.4808	0.488
		2	0.239	0.232	3.5306	0.171
		3	-0.010	0.032	3.5364	0.316
		4	0.043	-0.012	3.6380	0.457
		5	0.043	0.043	3.7412	0.587
		6	0.132	0.141	4.7539	0.576
		7	-0.009	-0.009	4.7585	0.689
		8	-0.082	-0.160	5.1644	0.740
		9	-0.042	-0.066	5.2745	0.810
		10	-0.149	-0.116	6.7054	0.753
		11	0.036	0.027	6.7917	0.816
		12	-0.005	0.054	6.7932	0.871
		13	-0.032	-0.027	6.8640	0.909
		14	0.239	0.298	10.935	0.691
		15	-0.168	-0.094	13.009	0.602
		16	0.105	-0.018	13.847	0.610
		17	-0.075	-0.049	14.280	0.647
		18	0.070	-0.009	14.676	0.684
		19	0.040	0.052	14.806	0.735
		20	0.043	-0.063	14.966	0.778

ԺՇ ստացիոնարության մասին է փաստում մաս թեստի վիճա-
կանու p արժեքը. $\text{Prob.} = p\text{-value} = 0 < 0.05$, հետևաբար զրոյական վար-
կածը մերժելիս սխալվելու հավանականությունը հավասար է զրոյի,
այն պետք է մերժել: Դրա մասին է վկայում մաս այս շարքի ACF-ը:
Այսպիսով, ՀՆԱ-ի ԺՇ-ը ունի սեզոնայնություն, որի պարբերությունը
4 է: Ապասեզոնայնացված ՀՆԱ-ի ԺՇ-ը չունի դետերմինիստիկ
միտում և ունի ստոխաստիկ միտում, այն DS շարք է:
Ապասեզոնայնացված ՀՆԱ ԺՇ-ը մեկ անգամ տարբերական
օպերատոր կիրառելուց հետո ստացիոնարանում է, այսինքն՝ այն

առաջին կարգի ինտեգրացվող I(1) շարք է: Այս ստացիոնարացված շարքն արդեն կարելի է մոդելավորել:

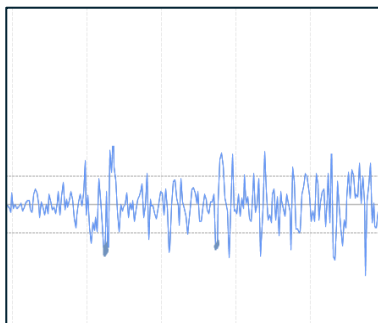
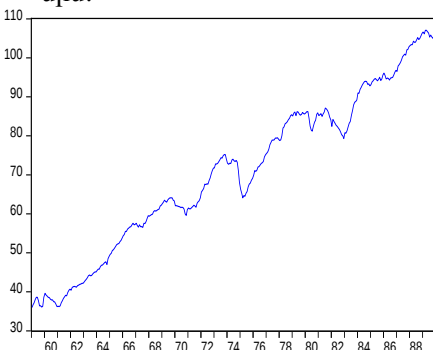
Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Սահմանե՛ք պատահական գործընթացն ու ժամանակային շարքը: Բնութագրե՛ք, թե ինչով են դրանք տարբերվում:
2. Ի՞նչ բաղկացուցիչներ կարող են ունենալ պատահական գործընթացն ու ժամանակային շարքը: Ի՞նչ տեսքով կարող են դրանք մտնել ԺՇ-ի մեջ:
3. Սահմանե՛ք ու բացատրե՛ք ԺՇ բաղկացուցիչները:
4. Ներկայացրե՛ք ստացիոնարության երկու սահմանումները և բացատրե՛ք դրանց իմաստը:
5. Սահմանե՛ք «սպիտակ աղմուկ» ԺՇ-ն և բնութագրե՛ք դրա հիմնական հատկությունները:
6. Ներկայացրե՛ք առանց տեղաշարժի պատահական քայլի մոդելով արտահայտվող ԺՇ տեսքը: Ստուգե՛ք դրա թույլ իմաստով ստացիոնարությունը:
7. Ներկայացրե՛ք տեղաշարժով պատահական քայլի մոդելով արտահայտվող ԺՇ տեսքը: Ստուգե՛ք դրա թույլ իմաստով ստացիոնարությունը:
8. Ինչո՞ւ է պատահական քայլի մոդելով արտահայտվող ԺՇ-ն կոչվում «միավոր արմատ»:
9. Ինչո՞վ է պայմանավորված պատահական քայլի մոդելով արտահայտվող ԺՇ ոչ ստացիոնարությունը: Ի՞նչ է ստոխաստիկ միտումը/տրենդը:
10. Սահմանե՛ք տարբերական օպերատորը և կիրառե՛ք այն առանց տեղաշարժի ու տեղաշարժով պատահական քայլի մոդելով արտահայտվող ԺՇ-ի նկատմամբ:
11. Սահմանե՛ք տրենդային ստացիոնար և տարբերությամբ ստացիոնար ԺՇ-ն: Բերե՛ք օրինակներ: Ինչպե՞ս է պետք ստացիոնարացնել դրանցից յուրաքանչյուրը:
12. Բերե՛ք թե՛ դետերմինիստիկ և թե՛ ստոխաստիկ միտում ունեցող ԺՇ օրինակ և ստացիոնարացրե՛ք դրանք:
13. Սահմանե՛ք ինտեգրացվող ԺՇ-ն: Նշե՛ք դրա հատկությունները:
14. Սահմանե՛ք պատահական գործընթացի ու ԺՇ ավտոկովարիացիայի ու ավտոկոռելյացիայի ֆունկցիաները, մասնակի ավտոկոռելյացիայի ֆունկցիան:

15. Ի՞նչ հատկություններ ունի ստացիոնար պատահական գործընթացի ու $\mathcal{F}_t$ ավտոկոռելյացիայի ֆունկցիան: Լագերի աճի հետ ինչպե՞ս են փոխվում ստացիոնար և ոչ ստացիոնար $\mathcal{F}_t$ -երի ավտոկոռելյացիայի ֆունկցիաները, ի՞նչ տեսք կունենան դրանց կոռելոգրամները:
16. Ի՞նչ թեստեր կան պատահական գործընթացի ավտոկոռելյացիայի գործակիցների՝ զրոյից տարբեր արժեքներ ունենալը ստուգելու համար: Նկարագրե՛ք այդ թեստերն ու դրանց հնարավոր ելքերը:
17. Նկարագրե՛ք $\mathcal{F}_t$ ստացիոնարության ստուգման Դիկի-Ֆուլերի թեստերը, դրանց տարբերակներն ու ելքերը:

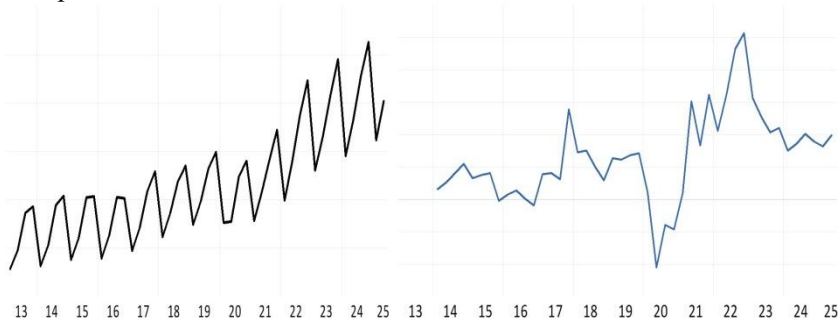
Խնդիրներ

1. Տրված են երկու տարբեր $\mathcal{F}_t$ գրաֆիկները ժամանակի մեջ: Ելնելով ժամանակի մեջ դրանց ունեցած վարքից՝ ի՞նչ ենթադրություն կարելի է անել դրանց բաղադրիչների ու ստացիոնարության մասին:



2. Տրված է $\mathcal{F}_t$, որն արտահայտվում է հետևյալ բանաձևով՝ $y_t = \beta_0 + \varepsilon_t$, որում ε_t -ն սպիտակ աղմուկ $\mathcal{F}_t$ է: Ելնելով թույլ իմաստով ստացիոնարության պայմաններից՝ ստուգե՛ք y_t $\mathcal{F}_t$ ստացիոնարությունը:
3. Տրված է $\mathcal{F}_t$, որն արտահայտվում է հետևյալ բանաձևով՝ $y_t = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot t^2 + y_{t-2} + \varepsilon_t$, որում ε_t -ն սպիտակ աղմուկ $\mathcal{F}_t$ է: Ի՞նչ տիպի միտում ունի այս շարքը: Ստացիոնարացրե՛ք այն:

4. Տրված են երկու տարբեր ԺՇ գրաֆիկները ժամանակի մեջ: Ելնելով ժամանակի մեջ դրանց ունեցած վարքից՝ ի՞նչ ենթադրություն կարելի է անել դրանց բաղադրիչների ու ստացիոնարության մասին:



5. Տրված են երկու տարբեր ԺՇ ավտոկոռելյացիայի ու մասնակի ավտոկոռելյացիայի ֆունկցիայի գրաֆիկները՝ կոռելյոգրամները, 24 լագի համար: Ելնելով ավտոկոռելյացիայի ֆունկցիայի վարքից՝ կախված լագերից, դրանց վիճակագրական նշանակալիության ու Բոքս-Լյունգի թեստերից՝ ի՞նչ ենթադրություն կարելի է անել դրանց ստացիոնարության մասին:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.835	0.835	57.900	0.000
		2 0.777	0.263	108.67	0.000
		3 0.744	0.164	155.84	0.000
		4 0.681	-0.027	195.90	0.000
		5 0.624	-0.037	229.92	0.000
		6 0.573	-0.024	259.04	0.000
		7 0.531	0.013	284.40	0.000
		8 0.467	-0.078	304.25	0.000
		9 0.384	-0.146	317.90	0.000
		10 0.321	-0.069	327.58	0.000
		11 0.258	-0.049	333.90	0.000
		12 0.171	-0.121	336.71	0.000
		13 0.156	0.144	339.09	0.000
		14 0.099	-0.048	340.06	0.000
		15 0.094	0.171	340.95	0.000
		16 0.071	0.020	341.48	0.000
		17 0.017	-0.095	341.51	0.000
		18 -0.003	-0.008	341.51	0.000
		19 -0.028	-0.016	341.59	0.000
		20 -0.039	0.030	341.76	0.000
		21 -0.040	0.022	341.94	0.000
		22 -0.061	-0.080	342.36	0.000
		23 -0.080	-0.088	343.10	0.000
		24 -0.066	0.069	343.61	0.000

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.335	-0.335	9.3352	0.002
		2 -0.088	-0.226	9.9915	0.007
		3 0.127	0.017	11.355	0.010
		4 -0.042	-0.005	11.506	0.021
		5 -0.018	-0.008	11.533	0.042
		6 -0.028	-0.059	11.602	0.071
		7 0.078	0.051	12.146	0.096
		8 0.069	0.132	12.584	0.127
		9 -0.043	0.068	12.754	0.174
		10 0.016	0.040	12.779	0.236
		11 0.060	0.075	13.119	0.286
		12 -0.242	-0.224	18.763	0.094
		13 0.107	-0.061	19.890	0.098
		14 -0.165	-0.278	22.594	0.067
		15 0.075	-0.075	23.163	0.081
		16 0.102	0.051	24.228	0.085
		17 -0.121	-0.024	25.740	0.079
		18 0.042	0.001	25.929	0.101
		19 -0.012	0.008	25.944	0.132
		20 -0.083	-0.035	26.694	0.144
		21 0.073	0.101	27.283	0.162
		22 -0.027	0.066	27.366	0.198
		23 -0.066	-0.048	27.868	0.221
		24 0.053	-0.096	28.200	0.252

6. ՀՀ կենտրոնական բանկի պաշտոնական կայքէջից ներբեռնե՛ք արտադրական եղանակով հաշվարկված ՀՆԱ-ի ու դրա բաղադրիչների ժամանակային շարքերը 2013-2025 թթ. համար՝ եռամսյակային կտրվածքով (հասցեն՝

<https://www.cba.am/hy/statistics/gdp-components-calculated-by-production-approach-nace-2/28>): Տեղափոխե՛ք ՀՆԱ-ի՝ «Էլեկտրականության, գազի, գոլորշու և լավորակ օդի մատակարարում» բաղադրիչը «EViews» փաթեթ: Ուսումնասիրե՛ք դրա գրաֆիկը, անհրաժեշտության պարագայում ապասեզոնայնացրե՛ք այն: Ուսումնասիրե՛ք (ապասեզոնայնացված) շարքի անավտոկոռելյացիայի ֆունկցիայի գրաֆիկը, իրականացրե՛ք Դիկի-Ֆուլերի ստացիոնարության թեստը, անհրաժեշտության դեպքում կիրառե՛ք տարբերական օպերատորը և ուսումնասիրե՛ք տարբերության վերցված շարքի կոռելյոգրամը և ստացիոնարությունը: Եկե՛ք եզրակացության շարքի ստացիոնարության, դետերմինացված ու ստոխաստիկ տրենդների առկայության ու շարքի իմսեզոգացման կարգի մասին:

7. Ժամանակային շարքի ստացիոնարությունը ստուգելու նպատակով «EViews» փաթեթում իրականացվել է Դիկի-Ֆուլերի ընդհանրացված թեստի 3-րդ տարբերակը $\Delta(y_t) = \beta_0 + \beta_1 \cdot t + \delta \cdot y_{t-1} + \sum_{i=1}^p \alpha_i \cdot \Delta(y_{t-i}) + \varepsilon_t$: Ելնելով բերված ելքային

աղյուսակից՝ եզրակացությունն արեք ԺՇ ստացիոնարության, դե-
տերմինիստիկ ու ստոխաստիկ միտումների առկայության մասին:

Null Hypothesis: HS has a unit root

	t-Statistic	Prob.
Augmented Dickey-Fuller test statistic	-2.293	0.433
Test critical values:		
1% level	-4.077	
5% level	-3.467	
10% level	-3.160	

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(HS)

Method: Least Squares

Variable	Coefficient	Std. Error	t-Statistic	Prob.
HS(-1)	-0.155	0.067	-2.293	0.025
D(HS(-1))	-0.269	0.109	-2.461	0.016
C	1.080	0.475	2.273	0.026
@TREND(“1989M06”)	0.001	0.000	1.664	0.100
R-squared	0.176	Mean dependent var		0.001
Adjusted R-squared	0.144	S.D. dependent var		0.076
S.E. of regression	0.070	Akaike info criterion		-2.424
Sum squared resid	0.375	Schwarz criterion		-2.305
Log likelihood	100.955	Hannan-Quinn criter.		-2.376
F-statistic	5.418	Durbin-Watson stat		2.100
Prob(F-statistic)	0.002			

8. Համաշխարհային բանկի «Աշխարհի զարգացման ցուցանիշներ» առցանց տվյալների շտեմարանից (հասցեն՝ <https://databank.worldbank.org/source/world-development-indicators>) ներբեռնեք Կոլումբիայի գյուղատնտեսության, անտառտնտեսության ու ձկնորսության ավելացված արժեքի ցուցանիշը (ընթացիկ ԱՄՆ դոլարով) 1965-2024 թթ. համար՝ տարեկան կտրվածքով: Ուսումնասիրեք դրա գրաֆիկը, անհրաժեշտության պարագայում ապաստեզոնայնացրեք այն: Ուսումնասիրեք (ապաստեզոնայնացված) շարքի անավտոկոռելյացիայի ֆունկցիայի գրաֆիկը, իրականացրեք Դիկի-Ֆուլերի ստացիոնարության թեստը, անհրաժեշտության դեպքում կիրառեք տարբերական օպերատորը և ուսումնասիրեք տարբերությամբ վերցված շարքի կոռելյորամը ու ստացիոնարությունը: Եկեք եզրակացության շարքի ստացիոնարության, դետերմինացված և ստոխաստիկ տրենդների առկայության ու շարքի ինտեգրացման կարգի մասին:

ԳԼՈՒԽ 7. ՄՏԱՑԻՈՆԱՐ ՄԻԱԶԱՓ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐԻ ԷԿՈՆՈՄԵՏՐԻԿ ՄՈԳԵԼԱՎՈՐՈՒՄ

Միաչափ ԺՇ-երի էկոնոմետրիկ մոդելավորումը, այսպիսով, սկսվում է դրանց նախնական ուսումնասիրությունից, բաղադրիչների որոշումից, ավտոկոռելյացիայի ֆունկցիայի հետազոտումից, ստացիոնարության ստուգումից ու ինտեգրացման կարգի որոշումից: Արդյունքում պարզվում է, թե շարքն ինչպես պետք է ստացիոնարացնել: Ստացիոնար միաչափ շարքերի մոդելավորման համար մշակված է մեթոդաբանություն, որը կոչվում է ARIMA մոդելավորում: Այն կարելի է կիրառել ստացիոնար կամ ստացիոնարացված ԺՇ-երի համար: Ստորև ներկայացված են ստացիոնար (կամ ստացիոնարացված) միաչափ ԺՇ-երի էկոնոմետրիկ մոդելները՝ ըստ նշված մեթոդաբանության:

1. Առաջին կարգի ավտոռեգրեսիայի AR(1) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները

Ստացիոնար ԺՇ-երի ամենատարածված մոդելները ավտոռեգրեսիայի մոդելներն են: Դրանք գծային կախվածության մոդելներ են, ուստի կարող են գնահատվել ՓՋԵ-ով:

Ենթադրենք՝ ունենք y_t ստացիոնար ժամանակային շարք, որի յուրաքանչյուր t պահի արժեքը գծային կախման մեջ է դրա նախորդ $t-1$ պահի արժեքից ու սպիտակ աղմուկ ստացիոնար ժամանակային շարքի t պահի արժեքից: Ժամանակային շարքի այս ներկայացումը, մոդելը կոչվում է առաջին կարգի ավտոռեգրեսիայի մոդել կամ Մարկովի մոդել՝ AR(1): AR(1) մոդելի ընդհանուր տեսքն է՝

$$y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \varepsilon_t, \quad t = 1, \dots, n, \quad (148)$$

որում ε_t -ն սպիտակ աղմուկ ԺՇ է և բավարարում է հետևյալ պայմանները՝

$$\mathbb{E}(\varepsilon_t) = 0, \quad \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 = \text{const}, \quad \text{cov}(\varepsilon_t, \varepsilon_{t+\tau}) = 0 \quad \forall t \neq \tau:$$

Ինչպիսի՞ն են AR(1)-ով բնութագրվող ԺՇ ու այն գեներացնող պատահական գործընթացի հատկությունները: Հաշվենք AR(1)-ով բնութագրվող ԺՇ միջինը: Հաշվի առնելով, որ ԺՇ-ն ստացիոնար է, և

որա համար բավարարված են թույլ իմաստով ստացիոնարության երեք պայմանները, մասնավորապես $\mathbb{E}(y_t) = \mu = const$, կունենանք՝

$$\begin{aligned}\mathbb{E}(y_t) &= \mathbb{E}(\alpha_0 + \alpha_1 \cdot y_{t-1} + \varepsilon_t) = \mathbb{E}(\alpha_0 + \alpha_1 \cdot y_{t-1}) + \mathbb{E}(\varepsilon_t) = \\ &= \alpha_0 + \alpha_1 \cdot \mathbb{E}(y_{t-1}): \end{aligned}$$

Քանի որ ստացիոնար $\mathcal{C}$ -ի համար $\mathbb{E}(y_t) = \mu$, $\forall t = 1, 2, \dots, n$, ապա՝

$$\begin{aligned}\mu &= \alpha_0 + \alpha_1 \cdot \mu \rightarrow \mu - \alpha_1 \cdot \mu = \alpha_0 \rightarrow \mu(1 - \alpha_1) = \alpha_0 \\ &\rightarrow \mu = \frac{\alpha_0}{(1 - \alpha_1)}: \end{aligned} \quad (149)$$

Նշենք, որ եթե $\alpha_0 = 0$, ապա $\mathbb{E}(y_t) = \mu = 0$:

Հաշվենք $AR(1)$ -ով բնութագրվող ժամանակային շարքի վարիացիան:

$$\begin{aligned}\hat{\gamma}(0) &= Var(y_t) = Var(\alpha_0 + \alpha_1 \cdot y_{t-1} + \varepsilon_t) = \\ &= \alpha_1^2 \cdot Var(y_{t-1}) + \sigma_\varepsilon^2: \end{aligned} \quad (150)$$

Քանի որ ստացիոնար $\mathcal{C}$ -ի համար $Var(y_t) = \sigma_y^2$, $\forall t = 1, 2, \dots, n$, ուրեմն՝

$$\sigma_y^2 = \alpha_1^2 \cdot \sigma_y^2 + \sigma_\varepsilon^2 \rightarrow Var(y_t) = \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - \alpha_1^2}: \quad (151)$$

Նկատի ունենալով, որ վարիացիան դրական մեծություն է՝ (150)-ից հետևում է, որ՝

$$1 - \alpha_1^2 > 0 \rightarrow \alpha_1^2 < 1 \rightarrow |\alpha_1| < 1:$$

Մյուս կողմից՝ $1 - \alpha_1^2 \neq 0 \rightarrow |\alpha_1| \neq 1$:

Հետևաբար $y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \varepsilon_t$ տեսքի $AR(1)$ մոդելով բնութագրվող ստացիոնար $\mathcal{C}$ -ի համար $|\alpha_1| < 1$: Սա $AR(1)$ -ով մոդելավորվող $\mathcal{C}$ ստացիոնարության պայմանն է:

Հաշվենք շարքի անդամների ավտոկովարիացիաները: Առաջին կարգի ավտոկովարիացիան.

$$\begin{aligned}\hat{\gamma}(1) &= cov(y_t, y_{t-1}) = \mathbb{E}((y_t - \mathbb{E}(y_t)) \cdot (y_{t-1} - \mathbb{E}(y_{t-1}))) = \\ &= \mathbb{E}((y_t - \mu) \cdot (y_{t-1} - \mu)): \end{aligned}$$

Եթե $\alpha_0 = 0$, ապա $\mu = \mathbb{E}(y_t) = 0$, ուրեմն կունենանք.

$$\begin{aligned}
\hat{\gamma}(1) &= \text{cov}(y_t, y_{t-1}) = \mathbb{E}(y_t \cdot y_{t-1}) = \mathbb{E}((\alpha_1 \cdot y_{t-1} + \varepsilon_t) \cdot y_{t-1}) = \\
&= \mathbb{E}(\alpha_1 \cdot y_{t-1}^2 + \varepsilon_t \cdot y_{t-1}) = \mathbb{E}(\alpha_1 \cdot y_{t-1}^2) + \mathbb{E}(\varepsilon_t) \cdot \mathbb{E}(y_{t-1}) = \\
&= \alpha_1 \cdot \mathbb{E}(y_{t-1}^2) = \alpha_1 \cdot \sigma_y^2 = \alpha_1 \cdot \text{Var}(y_t):
\end{aligned} \tag{152}$$

Հետևաբար՝

$$\hat{\gamma}(1) = \alpha_1 \cdot \hat{\gamma}(0) \rightarrow \alpha_1 = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\text{cov}(y_t, y_{t-1})}{\text{var}(y_t)} = \hat{\rho}(1): \tag{153}$$

Նշանակում է՝ AR(1)-ով բնութագրվող ստացիոնար ժամանակային շարքի առաջին կարգի ավտոկոռելյացիայի գործակիցը հավասար է AR(1) մոդելի α_1 գործակցին՝ $\hat{\rho}(1) = \alpha_1$:

Եթե հաշվենք AR(1)-ով բնութագրվող ստացիոնար ժամանակային շարքի երկրորդ կարգի ավտոկովարիացիան, կստանանք՝

$$\begin{aligned}
\hat{\gamma}(2) &= \text{cov}(y_t, y_{t-2}) = \mathbb{E}(y_t \cdot y_{t-2}) = \mathbb{E}((\alpha_1 \cdot y_{t-1} + \varepsilon_t) \cdot y_{t-2}) = \\
&= \mathbb{E}(\alpha_1 \cdot y_{t-1} \cdot y_{t-2} + \varepsilon_t \cdot y_{t-2}) = \alpha_1 \cdot \mathbb{E}(y_{t-1} \cdot y_{t-2}) = \\
&= \alpha_1 \cdot \hat{\gamma}(1) = \alpha_1 \cdot \alpha_1 \cdot \hat{\gamma}(0) = \alpha_1^2 \cdot \hat{\gamma}(0) \rightarrow
\end{aligned}$$

$$\alpha_1^2 = \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} = \frac{\text{cov}(y_t, y_{t-2})}{\text{var}(y_t)} = \hat{\rho}(2): \tag{154}$$

Եթե հաշվենք AR(1)-ով բնութագրվող ստացիոնար ժամանակային շարքի կամայական k -րդ կարգի (k քայլով կամ լագով իրարից հեռացված անդամների) ավտոկովարիացիան, կստանանք՝

$$\hat{\gamma}(k) = \text{cov}(y_t, y_{t-k}) = \alpha_1^k \cdot \sigma_y^2 = \alpha_1^k \cdot \hat{\gamma}(0) \rightarrow \alpha_1^k = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \hat{\rho}(k): \tag{155}$$

Այսպիսով՝ AR(1)-ով բնութագրվող ստացիոնար ժամանակային շարքի կամայական k -րդ կարգի ավտոկոռելյացիաները ներկայացվում են մոդելի α_1 գործակցի միջոցով՝

$$\begin{aligned}
\hat{\rho}(1) &= \frac{\text{cov}(y_t, y_{t-1})}{\text{var}(y_t)} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \alpha_1 \\
\hat{\rho}(2) &= \frac{\text{cov}(y_t, y_{t-2})}{\text{var}(y_t)} = \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} = \alpha_1^2 \\
&\dots\dots\dots \\
\hat{\rho}(k) &= \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \alpha_1^k:
\end{aligned}$$

Եթե y_t -ն ստացիոնար է, ապա $|\alpha_1| < 1$, հետևաբար $AR(1)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի $\hat{\rho}(k)$ ավտոկոռելյացիայի գործակիցները լազի ավելացման հետ միասին արագ նվազում են:

Մյուս կողմից, $AR(1)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի բոլոր ավտոկոռելյացիայի գործակիցները կարելի է ներկայացնել առաջին կարգի ավտոկոռելյացիայի գործակցի միջոցով՝

$$\begin{aligned}\hat{\rho}(k) = \alpha_1^k &= \alpha_1 \cdot \alpha_1^{k-1} = \hat{\rho}(1) \cdot \alpha_1^{k-1} = \hat{\rho}(1) \cdot \alpha_1 \cdot \alpha_1^{k-2} = \\ &= \hat{\rho}(1)^2 \cdot \alpha_1^{k-2} = \dots = \hat{\rho}(1)^k:\end{aligned}\quad (156)$$

Մասնավորապես՝

$$\hat{\rho}(k) = \alpha_1^k = \alpha_1 \cdot \alpha_1^{k-1} = \alpha_1 \cdot \hat{\rho}(k-1): \quad (157)$$

Ամփոփենք $AR(1)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի հատկությունները.

1. $AR(1)$ -ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է՝ $|\alpha_1| < 1$:
2. $AR(1)$ -ով բնութագրվող ժամանակային շարքի թվային բնութագրիչներն են՝
 - $E(y_t) = 0$, եթե $\alpha_0 = 0$ և $E(y_t) = \frac{\alpha_0}{(1-\alpha_1)}$, եթե $\alpha_0 \neq 0$,
 - $Var(y_t) = \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\alpha_1^2}$,
 - $\hat{\rho}(k) = cov(y_t, y_{t-k}) = \alpha_1^k \cdot \sigma_y^2$:
3. $AR(1)$ -ով բնութագրվող ժամանակային շարքի ավտոկոռելյացիայի գործակիցները որոշվում են հետևյալ բանաձևով՝

$$\hat{\rho}(k) = \alpha_1^k = \alpha_1 \cdot \hat{\rho}(k-1):$$

4. Քանի որ $|\alpha_1| < 1$ և $\hat{\rho}(k) = \alpha_1^k$, ապա $AR(1)$ -ով բնութագրվող ժամանակային շարքի ավտոկոռելյացիայի ֆունկցիան արագ ու սահուն, առանց կտրուկ անկման նվազում է:

2. Երկրորդ կարգի ավտոռեգրեսիայի AR(2) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները

Ենթադրենք՝ ունենք y_t ստացիոնար ժամանակային շարք, որի յուրաքանչյուր t պահի արժեքը ներկայացվում է որպես դրա նախորդ երկու՝ $t-1$ ու $t-2$ պահերի ու սպիտակ աղմուկ ստացիոնար ժամանակային շարքի t պահի արժեքների գծային կոմբինացիա: Ժամանակային շարքի այս մոդելը կոչվում է երկրորդ կարգի ավտոռեգրեսիայի մոդել կամ Յուլի մոդել՝ AR(2): Ավտոռեգրեսիայի այս մոդելի ընդհանուր տեսքն է՝

$$y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \varepsilon_t, \quad t = 1, \dots, n, \quad (158)$$

որում՝ $E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma_\varepsilon^2 = \text{const}$, $\text{cov}(\varepsilon_t, \varepsilon_{t+\tau}) = 0$, $\forall t, \tau$:

Հաշվենք $E(y_t)$ -ն հաշվի առնելով, որ y_t -ն ստացիոնար է՝ μ հաստատուն միջինով ու σ_y^2 հաստատուն վարիացիայով:

$$\mu = E(y_t) = E(\alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \varepsilon_t) =$$

$$= \alpha_0 + E(\alpha_1 \cdot y_{t-1}) + E(\alpha_2 \cdot y_{t-2}) + E(\varepsilon_t) =$$

$$= \alpha_0 + \alpha_1 \cdot E(y_{t-1}) + \alpha_2 \cdot E(y_{t-2}) \rightarrow$$

$$\mu = \alpha_0 + \alpha_1 \cdot \mu + \alpha_2 \cdot \mu \rightarrow \mu(1 - (\alpha_1 + \alpha_2)) = \alpha_0 \rightarrow$$

$$\mu = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2)}: \quad (159)$$

(159)-ից հետևում է, որ եթե $\alpha_0 = 0$, ապա $E(y_t) = \mu = 0$, հետևում է նաև, որ $\alpha_1 + \alpha_2 \neq 1$: Հաշվենք AR(2)-ով բնութագրվող ԺՇ վարիացիան:

$$\begin{aligned} \hat{\gamma}(0) &= \text{Var}(y_t) = \sigma_y^2 = \text{Var}(\alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \varepsilon_t) = \\ &= \alpha_1^2 \cdot \text{Var}(y_{t-1}) + \alpha_2^2 \cdot \text{Var}(y_{t-2}) + \sigma_\varepsilon^2: \end{aligned}$$

Քանի որ y_t -ն ստացիոնար է, դրա վարիացիան կախված չէ ժամանակից՝ $\text{Var}(y_t) = \sigma_y^2 = \text{const}$: Հետևաբար՝

$$\sigma_y^2 = \alpha_1^2 \cdot \sigma_y^2 + \alpha_2^2 \cdot \sigma_y^2 + \sigma_\varepsilon^2 \rightarrow \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - (\alpha_1^2 + \alpha_2^2)}: \quad (160)$$

(160)-ից հետևում է, որ $\alpha_1^2 + \alpha_2^2 < 1 \rightarrow |\alpha_1| < 1$ և $|\alpha_2| < 1$: Սա AR(2)-ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է: $y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \varepsilon_t$ տեսքի մոդելով բնութագրվող ԺՇ-ն ստացիոնար է, եթե $|\alpha_1| < 1$ և $|\alpha_2| < 1$:

Հաշվենք $AR(2)$ -ով բնութագրվող ժամանակային շարքի անդամների ավտոկովարիացիաները: Ենթադրենք $\alpha_0 = 0$: Այդ դեպքում k -րդ կարգի ավտոկովարիացիան՝

$$\hat{\gamma}(k) = \text{cov}(y_t, y_{t-k}) = \mathbb{E}[y_t - \mathbb{E}(y_t)](y_{t-k} - \mathbb{E}(y_{t-k})) = \mathbb{E}(y_t \cdot y_{t-k}):$$

Նախ հաշվենք հետևյալը՝

$$y_t \cdot y_{t-k} = (\alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \varepsilon_t) \cdot y_{t-k} = \alpha_1 \cdot y_{t-1} \cdot y_{t-k} + \alpha_2 \cdot y_{t-2} \cdot y_{t-k} + \varepsilon_t \cdot y_{t-k}:$$

Այդ դեպքում.

$$\begin{aligned} \hat{\gamma}(k) &= \text{cov}(y_t, y_{t-k}) = \mathbb{E}(y_t \cdot y_{t-k}) = \\ &= \mathbb{E}(\alpha_1 \cdot y_{t-1} \cdot y_{t-k} + \alpha_2 \cdot y_{t-2} \cdot y_{t-k} + \varepsilon_t \cdot y_{t-k}) = \alpha_1 \cdot \mathbb{E}(y_{t-1} \cdot y_{t-k}) + \\ &+ \alpha_2 \cdot \mathbb{E}(y_{t-2} \cdot y_{t-k}) = \alpha_1 \cdot \text{cov}(y_{t-1}, y_{t-k}) + \alpha_2 \cdot \text{cov}(y_{t-2}, y_{t-k}) = \\ &= \alpha_1 \cdot \hat{\gamma}(k-1) + \alpha_2 \cdot \hat{\gamma}(k-2): \end{aligned} \quad (161)$$

(161)-ից հետևում է, որ.

$$\begin{aligned} \hat{\rho}(k) &= \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\alpha_1 \cdot \hat{\gamma}(k-1) + \alpha_2 \cdot \hat{\gamma}(k-2)}{\hat{\gamma}(0)} = \\ &= \alpha_1 \cdot \frac{\hat{\gamma}(k-1)}{\hat{\gamma}(0)} + \alpha_2 \cdot \frac{\hat{\gamma}(k-2)}{\hat{\gamma}(0)} = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2): \end{aligned} \quad (162)$$

Ամփոփենք $AR(2)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի հատկությունները՝

1. $AR(2)$ -ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է՝ $|\alpha_1| < 1$ և $|\alpha_2| < 1$,
2. $AR(2)$ -ով բնութագրվող ժամանակային շարքի թվային բնութագրիչներն են՝

- $\mathbb{E}(y_t) = 0$, եթե $\alpha_0 = 0$ և

$$\mathbb{E}(y_t) = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2)}, \text{ եթե } \alpha_0 \neq 0 \rightarrow \alpha_1 + \alpha_2 \neq 1$$

- $\text{Var}(y_t) = \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - (\alpha_1^2 + \alpha_2^2)} \rightarrow \alpha_1^2 + \alpha_2^2 < 1 \rightarrow$

$$|\alpha_1| < 1 \text{ և } |\alpha_2| < 1$$

- $\hat{\gamma}(k) = \text{cov}(y_t, y_{t-k}) = \alpha_1 \cdot \hat{\gamma}(k-1) + \alpha_2 \cdot \hat{\gamma}(k-2):$

3. AR(2)-ով բնութագրվող ժամանակային շարքի ավտոկոռելյացիայի գործակիցները որոշվում են հետևյալ բանաձևով՝

$$\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2):$$

4. Վերը բերված 1-ին ու 3-րդ հատկություններից բխում է, որ AR(2)-ով բնութագրվող ԺՇ ավտոկոռելյացիայի ֆունկցիան արագ ու սահուն, առանց կտրուկ անկման նվազում է:

3. Ընդհանուր դեպք՝ p-րդ կարգի ավտոռեգրեսիայի AR(p) մոդելով բնութագրվող ժամանակային շարքերն ու դրանց հատկությունները

Ենթադրենք՝ ունենք y_t ստացիոնար ժամանակային շարք, որի յուրաքանչյուր t պահի արժեքը ներկայացվում է որպես դրա նախորդ $t-1$ ու $t-2$... $t-p$ պահերի արժեքների ու սպիտակ աղմուկ ստացիոնար ժամանակային շարքի t պահի արժեքի գծային կոմբինացիա: Ժամանակային շարքի այս ներկայացումը, մոդելը կոչվում է p -րդ կարգի ավտոռեգրեսիայի մոդել՝ **AR(p)**: Ավտոռեգրեսիայի այս մոդելի ընդհանուր տեսքն է՝

$$y_t = \alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \dots + \alpha_p \cdot y_{t-p} + \varepsilon_t, \quad (163)$$

որում՝ $E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma_\varepsilon^2 = \text{const}$, $\text{cov}(\varepsilon_t, \varepsilon_{t+\tau}) = 0$, $\forall t \neq \tau$:

Նույն մոտեցմամբ կարող ենք ստանալ AR(p)-ով բնութագրվող ստացիոնար ժամանակային շարքի հատկությունները: Ամփոփենք դրանք:

Հաշվենք $E(y_t)$ -ն՝ հաշվի առնելով, որ y_t -ն ստացիոնար է՝ μ հաստատուն միջինով ու σ_y^2 հաստատուն դիսպերսիայով, իսկ ε_t -ն սպիտակ աղմուկ ստացիոնար ԺՇ է:

$$\begin{aligned} E(y_t) &= E(\alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \dots + \alpha_p \cdot y_{t-p} + \varepsilon_t) = \alpha_0 + \\ &+ E(\alpha_1 \cdot y_{t-1}) + E(\alpha_2 \cdot y_{t-2}) + \dots + E(\alpha_p \cdot y_{t-p}) + E(\varepsilon_t) = \\ &= \alpha_0 + \alpha_1 \cdot E(y_{t-1}) + \alpha_2 \cdot E(y_{t-2}) + \dots + \alpha_p \cdot E(y_{t-p}) \rightarrow \\ \mu &= \alpha_0 + \alpha_1 \cdot \mu + \alpha_2 \cdot \mu + \dots + \alpha_p \cdot \mu \rightarrow \mu = \alpha_0 + \mu(\alpha_1 + \alpha_2 + \dots + \alpha_p) \\ \rightarrow \mu - \mu(\alpha_1 + \alpha_2 + \dots + \alpha_p) &= \alpha_0 \rightarrow \mu(1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)) = \alpha_0 \\ \rightarrow \mu &= \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)}: \end{aligned}$$

Հետևաբար՝ $\alpha_1 + \alpha_2 + \dots + \alpha_p \neq 1$:

Այստեղից ստացվում է, որ եթե $\alpha_0 = 0$, ապա $\mu = 0$:

Հաշվենք $AR(p)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի վարիացիան:

$$\begin{aligned}\sigma_y^2 &= \text{Var}(y_t) = \text{Var}(\alpha_0 + \alpha_1 \cdot y_{t-1} + \alpha_2 \cdot y_{t-2} + \dots + \alpha_p \cdot y_{t-p} + \varepsilon_t) = \\ &= \alpha_1^2 \cdot \text{Var}(y_{t-1}) + \alpha_2^2 \cdot \text{Var}(y_{t-2}) + \dots + \alpha_p^2 \cdot \text{Var}(y_{t-p}) + \sigma_\varepsilon^2\end{aligned}$$

Քանի որ y_t -ն ստացիոնար է, դրա վարիացիան կախված չէ ժամանակից՝ $\text{Var}(y_t) = \sigma_y^2 = \text{const}$, ապա՝

$$\sigma_y^2 = \alpha_1^2 \cdot \sigma_y^2 + \alpha_2^2 \cdot \sigma_y^2 + \dots + \alpha_p^2 \cdot \sigma_y^2 + \sigma_\varepsilon^2 \rightarrow$$

$$\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - (\alpha_1^2 + \alpha_2^2 + \dots + \alpha_p^2)} :$$

Այստեղից հետևում է, որ $(\alpha_1^2 + \alpha_2^2 \dots + \alpha_p^2) < 1 \rightarrow |\alpha_1| < 1, |\alpha_2| < 1 \dots |\alpha_p| < 1$: Սա $AR(p)$ -ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է:

Եթե հաշվենք $AR(p)$ -ով բնութագրվող ժամանակային շարքի անդամների կամայական k -րդ կարգի ավտոկովարիացիան, ապա կստանանք՝

$$\hat{\gamma}(k) = \text{cov}(y_t, y_{t-k}) = \alpha_1 \cdot \hat{\gamma}(k-1) + \alpha_2 \cdot \hat{\gamma}(k-2) + \dots + \alpha_p \cdot \hat{\gamma}(k-p):$$

Հետևաբար՝

$$\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2) + \dots + \alpha_p \cdot \hat{\rho}(k-p):$$

Ամփոփենք $AR(p)$ -ով բնութագրվող ստացիոնար ժամանակային շարքի հատկություններ.

1. $AR(p)$ -ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է՝ $|\alpha_1| < 1, |\alpha_2| < 1 \dots |\alpha_p| < 1$,
2. $AR(p)$ -ով բնութագրվող ժամանակային շարքի թվային բնութագրիչներն են՝

- $\mathbb{E}(y_t) = 0$, եթե $\alpha_0 = 0$ և $\mathbb{E}(y_t) = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)}$, եթե

$$\alpha_0 \neq 0 \rightarrow \alpha_1 + \alpha_2 + \dots + \alpha_p \neq 1,$$

- $\text{Var}(y_t) = \sigma_y^2 = \frac{\sigma_\varepsilon^2}{1 - (\alpha_1^2 + \alpha_2^2 + \dots + \alpha_p^2)} \rightarrow$

$$\alpha_1^2 + \alpha_2^2 + \dots + \alpha_p^2 < 1,$$

- $\hat{\gamma}(k) = \text{cov}(y_t, y_{t-k}) = \alpha_1 \cdot \hat{\gamma}(k-1) + \alpha_2 \cdot \hat{\gamma}(k-2) + \dots + \alpha_p \cdot \hat{\gamma}(k-p):$

3. AR(p)-ով բնութագրվող ժամանակային շարքի ավտոկոռելյացիայի գործակիցները որոշվում են այսպես՝

$$\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2) + \dots + \alpha_p \cdot \hat{\rho}(k-p):$$

4. Վերը բերված 1-ին ու 3-րդ հատկություններից բխում է, որ AR(p)-ով բնութագրվող ժամանակային շարքի ավտոկոռելյացիայի ֆունկցիան արագ ու սահուն, առանց կտրուկ անկման նվազում է:

4. Յուլ-Ուոքերի հավասարումները

Հիշենք, որ AR(1) մոդելով ներկայացվող ԺՇ-ի համար ստացել էինք՝ $\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1)$, AR(2) մոդելով ներկայացվող ԺՇ-ի համար՝ $\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2)$, իսկ ընդհանուր դեպքում՝ AR(p) մոդելով ներկայացվող ԺՇ-ի համար՝

$$\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2) + \dots + \alpha_p \cdot \hat{\rho}(k-p): \quad (164)$$

(164)-ը կոչվում է Յուլ-Ուոքերի հավասարումներ: Դրանք գծային հավասարումների համակարգ են, որոնք կապում են AR(p) մոդելով բնութագրվող ստացիոնար ԺՇ ավտոկոռելյացիայի գործակիցները մոդելի α գործակիցների հետ: Օրինակ, եթե ԺՇ-ն արտահայտվում է AR(1) մոդելով ($p=1$), ապա $\hat{\rho}(1) = \alpha_1 \cdot \hat{\rho}(0) = \alpha_1$, $\hat{\rho}(2) = \alpha_1 \cdot \hat{\rho}(1) = \alpha_1^2$ $\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) = \alpha_1^k$: Եթե ԺՇ-ն արտահայտվում է AR(2) մոդելով ($p=2$), ապա Յուլ-Ուոքերի հավասարումներն են՝

$$\hat{\rho}(1) = \alpha_1 \cdot \hat{\rho}(0) + \alpha_2 \cdot \hat{\rho}(-1) = \alpha_1 + \alpha_2 \cdot \hat{\rho}(1),$$

$$\hat{\rho}(2) = \alpha_1 \cdot \hat{\rho}(1) + \alpha_2 \cdot \hat{\rho}(0) = \alpha_1 \cdot \hat{\rho}(1) + \alpha_2$$

....

$$\hat{\rho}(k) = \alpha_1 \cdot \hat{\rho}(k-1) + \alpha_2 \cdot \hat{\rho}(k-2):$$

Եթե հաշվարկված են ԺՇ ավտոկոռելյացիայի գործակիցները, ապա այս հավասարումների միջոցով կարելի է գտնել AR մոդելի α գործակիցները և հակառակը. եթե գնահատված են $\hat{\rho}$ գործակիցները, դրանց միջոցով կարելի է գտնել կամայական կարգի ավտոկոռելյացիայի գործակիցները:

5. Ժամանակային շարքի մասնակի ավտոկոռելյացիայի ֆունկցիան և AR(p) մոդելի p կարգի որոշումը

Ստացիոնար միաչափ ԺՇ-ն ավտոռեգրեսիայի մոդելով մոդելավորելիս AR մոդելի p կարգը անհայտ է և պետք է որոշվի մոդելավորման ընթացքում: p-ի արժեքը որոշելու համար սովորաբար կիրառվում է երկու հիմնական մոտեցում կամ դրանց զուգակցումը: Առաջին մոտեցումը մասնակի ավտոկոռելյացիայի ֆունկցիայի օգտագործումն է, իսկ երկրորդ մոտեցումն օգտագործում է տեղեկատվական չափանիշները:

Նախ անդրադառնանք մասնակի ավտոկոռելյացիայի ֆունկցիային: Ներկայացնենք y_t ԺՇ-ն հերթականությամբ AR(1), AR(2), AR(3), AR(p) մոդելների միջոցով՝

$$y_t = \alpha_{0,1} + \alpha_{1,1} \cdot y_{t-1} + \varepsilon_t, \quad (165)$$

$$y_t = \alpha_{0,2} + \alpha_{1,2} \cdot y_{t-1} + \alpha_{2,2} \cdot y_{t-2} + \varepsilon_t, \quad (166)$$

$$y_t = \alpha_{0,2} + \alpha_{1,3} \cdot y_{t-1} + \alpha_{2,3} \cdot y_{t-2} + \alpha_{3,3} \cdot y_{t-3} + \varepsilon_t, \quad (167)$$

.....

$$y_t = \alpha_{0,p} + \alpha_{1,p} \cdot y_{t-1} + \alpha_{2,p} \cdot y_{t-2} + \dots + \alpha_{p,p} \cdot y_{t-p} + \varepsilon_t: \quad (168)$$

Այս մոդելներից յուրաքանչյուրը հերթականությամբ կարելի է գնահատել ՓԶԵ-ով և գտնել համապատասխան $\hat{\alpha}$ գործակիցները: (166) մոդելից սկսած՝ դրանք, փաստորեն, բազմագործոն գծային ռեգրեսիայի մոդելներ են: Այդ դեպքում (165) մոդելի ՓԶԵ-ով գնահատումից ստացված $\hat{\alpha}_{1,1}$ գործակիցը կարտահայտի y_{t-1} -ի ուղղակի ազդեցությունը y_t կախյալ փոփոխականի վրա: Ըստ սահմանման՝ $\hat{\alpha}_{1,1}$ գործակիցը կներկայացնի y_t ԺՇ մեկ լագով իրարից հեռացված անդամների առաջին կարգի մասնակի ավտոկոռելյացիան: (166) մոդելի գնահատումից ստացված $\hat{\alpha}_{2,2}$ գործակիցը կարտահայտի y_{t-2} -ի ուղղակի, y_{t-1} -ի միջնորդավորված (անուղղակի) ազդեցությունից «մաքրված» ազդեցությունը y_t կախյալ փոփոխականի վրա (y_{t-1} -ի ամրագրված արժեքի դեպքում կամ այլ հավասար պայմաններում): Ուրեմն $\hat{\alpha}_{2,2}$ -ը y_t ԺՇ-ի երկրորդ կարգի մասնակի ավտոկոռելյացիան է: Նույն տրամաբանությամբ (167) մոդելի գնահատումից ստացված $\hat{\alpha}_{3,3}$ գործակիցը ԺՇ երրորդ կարգի մասնակի ավտոկոռելյացիան է, իսկ (168)-ից ստացված $\hat{\alpha}_{p,p}$ -ը՝ p-րդ կարգի մասնակի ավտոկոռելյացիան:

Այսպես կարելի է ստանալ y_t ԺՇ-ի մասնակի ավտոկոռելյացիայի ֆունկցիան՝ PACF-ը 1-ից մինչև p լագերի համար: PACF-ը կարելի է օգտագործել ԺՇ ավտոռեգրեսիայի մոդելի կարգի որոշման համար: Եթե (165) մոդելի գնահատումից ստացված $\hat{\alpha}_{1,1}$ գործակիցը՝ առաջին կարգի մասնակի ավտոկոռելյացիան, վիճակագրորեն նշանակալի է, ապա ԺՇ-ն պետք է արտահայտվի առնվազն առաջին կարգի AR(1) մոդելով: Հիշենք, որ ՓՔԵ-ով գնահատված գործակցի վիճակագրական նշանակալիությունը համարժեք է տեսական մոդելի (տվյալ դեպքում պատահական գործընթացի) համապատասխան պարամետրի՝ գրոյին հավասար չլինելուն: Եթե (166) մոդելի գնահատումից ստացված $\hat{\alpha}_{2,2}$ գործակիցը՝ երկրորդ կարգի մասնակի ավտոկոռելյացիան, ևս վիճակագրորեն նշանակալի է, ապա ԺՇ-ն պետք է արտահայտվի առնվազն երկրորդ կարգի AR(2) մոդելով: Հակառակ դեպքում, եթե երկրորդ կարգի մասնակի ավտոկոռելյացիան վիճակագրորեն նշանակալի չէ, ապա ԺՇ-ն պետք է արտահայտվի առաջին կարգի AR(1) մոդելով: Եթե նախորդ երկու մասնակի ավտոկոռելյացիաները նշանակալի են, և (167) մոդելի գնահատումից ստացված $\hat{\alpha}_{3,3}$ գործակիցը՝ երրորդ կարգի մասնակի ավտոկոռելյացիան, ևս վիճակագրորեն նշանակալի է, ապա ԺՇ-ն պետք է արտահայտվի առնվազն երրորդ կարգի AR(3) մոդելով: Հակառակ դեպքում, եթե երրորդ կարգի մասնակի ավտոկոռելյացիան վիճակագրորեն նշանակալի չէ, ապա ԺՇ-ն պետք է արտահայտվի երկրորդ կարգի AR(2) մոդելով: Եվ այդպես շարունակ: Ընդհանուր դեպքում, եթե ԺՇ-ն պետք է արտահայտվի p -րդ կարգի AR(p) մոդելով, ապա դրա մասնակի ավտոկոռելյացիայի ֆունկցիայի 1-ից մինչև p -րդ կարգի ավտոկոռելյացիայի գործակիցները վիճակագրորեն նշանակալի են, իսկ կամայական $j > p$ կարգի ավտոկոռելյացիայի գործակիցները վիճակագրորեն նշանակալի չեն:

Այսպիսով՝ ստացիոնար ԺՇ մասնակի ավտոկոռելյացիայի ֆունկցիայի միջոցով կարելի է որոշել այդ ԺՇ ավտոռեգրեսիայի մոդելի կարգը՝ p -ն: Եթե AR կարգը հավասար է p -ի, ապա $j > p$ լագերի համար մասնակի ավտոկոռելյացիայի գործակիցները 0-ից նշանակալիորեն չեն տարբերվում: Վերջին հանգամանքը կարելի է ստուգել՝ յուրաքանչյուր մասնակի ավտոկոռելյացիայի գործակցի համար կիրառելով դրա վիճակագրական նշանակալիության մասին վարկածի վերը բերված միջակայքային թեստը, ըստ որի, եթե $\hat{\alpha}_j > \left| 1.96 \cdot \sqrt{\frac{1}{n}} \right|$,

ապա $\hat{\alpha}_j$ -ն նշանակալիորեն 0-ից տարբերվում է, իսկ եթե $\hat{\alpha}_j \leq \left| 1.96 \cdot \sqrt{\frac{1}{n}} \right|$, ապա $\hat{\alpha}_j$ -ն 0-ից նշանակալիորեն չի տարբերվում 95 % հավանականությամբ:

Այսինքն՝ **AR մոդելի p կարգը որոշելու համար պետք է դիտարկել դրա մասնակի ավտոկոռելյացիայի ֆունկցիան՝ PACF-ը: p-ն համապատասխանում է այն լագին, որից հետո PACF-ը կտրուկ անկում է ապրում և սկսում տատանվել 0-ի շուրջ՝ դրանից նշանակալիորեն չտարբերվելով:**

6. Սահող միջինների MA(1), MA(2) ու MA(q) մոդելներով բնութագրվող ժամանակային շարքեր

Ենթադրենք $\mathcal{F}$ Ը կամայական t պահի արժեքը կարելի է արտահայտել ոչ թե իր նախորդ արժեքների, այլ սպիտակ աղմուկ $\mathcal{F}$ Ը տվյալ t պահի ու նախորդ q պահերի արժեքների կշռված գումարի տեսքով՝

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} + \varepsilon_t, \quad (169)$$

որում ε_t -ն սպիտակ աղմուկ $\mathcal{F}$ Ը է: Ինչպես գիտենք, սպիտակ աղմուկ $\mathcal{F}$ Ը-ն ստացիոնար է. այն բավարարում է թույլ իմաստով ստացիոնարության բոլոր 3 պայմանները:

(169)-ը կոչվում է q -րդ կարգի սահող միջինների (moving average) մոդել և նշանակվում է՝ MA(q):

MA(q) մոդելով արտահայտվող ժամանակային շարքը ստացիոնար է, քանի որ այն միայն ստացիոնար սպիտակ աղմուկ $\mathcal{F}$ Ը-ով է արտահայտվում: Մասնավոր դեպքում, եթե $q=0$, ապա $y_t = \beta_0 + \varepsilon_t$, երբ $\beta_0 \neq 0$ և $y_t = \varepsilon_t$, երբ $\beta_0 = 0$:

Եթե $q=1$, ունենք առաջին կարգի սահող միջիններով մոդել կամ MA(1)՝

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \varepsilon_t: \quad (170)$$

Եթե $q=2$, ունենք երկրորդ կարգի սահող միջիններով մոդել կամ MA(2)՝

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t: \quad (171)$$

6.1. Առաջին կարգի սահող միջինների MA(1) մոդելով բնութագրվող ժամանակային շարքի հատկությունները

Հաշվենք (170) բանաձևով կամ MA(1) մոդելով արտահայտվող ԺՇ թվային բնութագրիչները: Ելնելով սպիտակ աղմուկ ԺՇ հատկություններից՝ MA(1) մոդելով արտահայտվող ԺՇ միջինը կորոշվի հետևյալ կերպ՝

$$\mathbb{E}(y_t) = \mathbb{E}(\beta_0 + \beta_1 \varepsilon_{t-1} + \varepsilon_t) = \beta_0: \quad (172)$$

Նշենք, որ եթե $\beta_0 = 0$, ապա $\mathbb{E}(y_t) = 0$: Այսպիսով՝ MA(1) մոդելով արտահայտվող ԺՇ-ն տատանվում է կա՛մ β_0 -ի, կա՛մ 0-ի շուրջ:

Հաշվենք y_t ԺՇ-ի վարիացիան.

$$\begin{aligned} \hat{\gamma}(0) &= \text{Var}(y_t) = \text{Var}(\beta_0 + \beta_1 \varepsilon_{t-1} + \varepsilon_t) = \beta_1^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = \\ &= (1 + \beta_1^2) \cdot \sigma_\varepsilon^2 = \text{const}: \end{aligned}$$

Հաշվենք MA(1) մոդելով արտահայտվող ԺՇ տարբեր կարգի ավտոկովարիացիաները՝ պարզության համար ենթադրելով, որ $\beta_0 = 0$, կամ դիտարկելով $y_t - \beta_0$ ԺՇ-ն: Նկատի ունենալով, որ սպիտակ աղմուկ ԺՇ կամայական լագով իրարից հեռացված անդամների կովարիացիան հավասար է զրոյի՝ հաշվենք ԺՇ $k=1$ լագով իրարից հեռացած անդամների ավտոկովարիացիան՝

$$\begin{aligned} \hat{\gamma}(1) &= \text{cov}(y_t, y_{t-1}) = \mathbb{E}((y_t - \mathbb{E}(y_t))(y_{t-1} - \mathbb{E}(y_{t-1}))) = \\ &= \mathbb{E}(y_t \cdot y_{t-1}) = \mathbb{E}[(\beta_1 \varepsilon_{t-1} + \varepsilon_t)(\beta_1 \varepsilon_{t-2} + \varepsilon_{t-1})] = \\ &= \mathbb{E}[\beta_1^2 \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-2} + \beta_1 \cdot \varepsilon_{t-1}^2 + \beta_1 \cdot \varepsilon_t \varepsilon_{t-2} + \varepsilon_t \cdot \varepsilon_{t-1}] = \\ &= \beta_1^2 \cdot \mathbb{E}(\varepsilon_{t-1} \cdot \varepsilon_{t-2}) + \beta_1 \cdot \mathbb{E}(\varepsilon_{t-1}^2) + \beta_1 \cdot \mathbb{E}(\varepsilon_t \cdot \varepsilon_{t-2}) + \\ &+ \mathbb{E}(\varepsilon_t \cdot \varepsilon_{t-1}) = \beta_1 \cdot \sigma_\varepsilon^2: \end{aligned}$$

Կամայական $k>1$ լագով իրարից հեռացած անդամների ավտոկովարիացիան՝

$$\begin{aligned} \hat{\gamma}(k) &= \text{cov}(y_t, y_{t-k}) = \mathbb{E}(y_t \cdot y_{t-k}) = \mathbb{E}[(\beta_1 \varepsilon_{t-1} + \varepsilon_t)(\beta_1 \varepsilon_{t-k-1} + \varepsilon_{t-k})] = \\ &= \mathbb{E}[\beta_1^2 \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-k-1} + \beta_1 \cdot \varepsilon_{t-1} \cdot \varepsilon_{t-k} + \beta_1 \cdot \varepsilon_t \varepsilon_{t-k-1} + \varepsilon_t \cdot \varepsilon_{t-k}] = 0: \end{aligned}$$

Այսպիսով՝ MA(1) մոդելով արտահայտվող ԺՇ ավտոկովարիացիայի ֆունկցիան կարտահայտվի հետևյալ կերպ՝

$$\hat{\gamma}(k) = \begin{cases} (1 + \beta_1^2) \cdot \sigma_\varepsilon^2, \text{ երբ } k = 0 \\ \beta_1 \cdot \sigma_\varepsilon^2, \text{ երբ } k = 1 \\ 0, \text{ երբ } k > 1 \end{cases} : \quad (173)$$

Այսինքն՝ MA(1) մոդելով արտահայտվող y_t ԺՇ միայն առաջին կարգի ավտոկովարիացիայի գործակիցն է զրոյից տարբերվում. դրա 1-ից ավելի լագերով իրարից հեռացած անդամների ավտոկովարիացիայի գործակիցները հավասար են զրոյի:

Օգտվելով (173)-ից՝ MA(1) մոդելով արտահայտվող ԺՇ ավտոկոռելյացիայի ֆունկցիան կարտահայտվի հետևյալ կերպ՝

$$\frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \hat{\rho}(k) = \begin{cases} \frac{(1+\beta_1^2) \cdot \sigma_\varepsilon^2}{(1+\beta_1^2) \cdot \sigma_\varepsilon^2} = 1, \text{ երբ } k = 0 \\ \frac{\beta_1 \cdot \sigma_\varepsilon^2}{(1+\beta_1^2) \sigma_\varepsilon^2} = \frac{\beta_1}{1+\beta_1^2}, \text{ երբ } k = 1 \\ 0, \text{ երբ } k > 1 \end{cases} : \quad (174)$$

(174)-ից հետևում է, որ կոռելյացված են MA(1) մոդելով արտահայտվող ժամանակային շարքի միայն իրար հարևան, մի լագով իրարից հեռացված անդամները: Դրանց ավտոկոռելյացիան դրական է, եթե $\beta_1 > 0$, իսկ եթե $\beta_1 < 0$, դրանց ավտոկոռելյացիան բացասական է: Բացի դրանից՝ կամայական MA(1) մոդելով արտահայտվող ԺՇ-ի համար $|\hat{\rho}(1)| \leq 0.5$: Այսինքն՝ MA(1)-ով արտահայտվող ԺՇ իրար հարևան երկու անդամները բավական թույլ են կոռելյացված՝ ի տարբերություն AR(1)-ի:

Այսպիսով՝ MA(1)-ով բնութագրվող ԺՇ ավտոկոռելյացիայի ֆունկցիան առաջին լագից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջը՝ դրանից նշանակալիորեն չտարբերվելով: Կարելի է ցույց տալ, որ MA(1) մոդելով արտահայտվող ԺՇ մասնակի ավտոկոռելյացիայի ֆունկցիան արագ ու սահուն, առանց կտրուկ անկման նվազող ֆունկցիա է:

6.2. Երկրորդ կարգի սահող միջինների MA(2) մոդելով բնութագրվող ժամանակային շարքի հատկությունները

Ունենք y_t ԺՇ-ն, որը մոդելավորվում է երկրորդ կարգի սահող միջինների մոդելով՝

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t,$$

որում ε_t -ն սպիտակ աղմուկ ԺՇ է: Հաշվենք MA(2) մոդելով արտահայտվող ժամանակային շարքի թվային բնութագրիչները: Հաշվի

առնելով սպիտակ աղմուկ σ^2 ստացիոնարությունը՝ y_t σ^2 -ի միջինը կհաշվարկվի այսպես՝

$$\mathbb{E}(y_t) = \mathbb{E}(\beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t) = \beta_0 = \text{const:}$$

Ուրեմն $\text{MA}(2)$ մոդելով արտահայտվող σ^2 -ն տատանվում է կա՛մ β_0 -ի, կա՛մ 0 -ի շուրջ, եթե $\beta_0 = 0$:

Հաշվենք y_t σ^2 -ի վարիացիան՝

$$\begin{aligned} \hat{\gamma}(0) &= \text{Var}(y_t) = \text{Var}(\beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t) = \\ &= \beta_1^2 \sigma_\varepsilon^2 + \beta_2^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = (1 + \beta_1^2 + \beta_2^2) \sigma_\varepsilon^2 = \text{const:} \end{aligned}$$

Հաշվենք $\text{MA}(2)$ մոդելով արտահայտվող σ^2 տարբեր կարգի ավտոկովարիացիաները՝ պարզության համար ենթադրելով, որ $\beta_0 = 0$, կամ դիտարկելով $y_t - \beta_0$ σ^2 -ն: Նախ հաշվենք σ^2 $k=1$ լագով իրարից հեռացած անդամների ավտոկովարիացիան՝

$$\begin{aligned} \hat{\gamma}(1) &= \text{cov}(y_t, y_{t-1}) = \mathbb{E}(y_t \cdot y_{t-1}) = \\ &= \mathbb{E}[(\beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t)(\beta_1 \varepsilon_{t-2} + \beta_2 \varepsilon_{t-3} + \varepsilon_{t-1})] = \beta_1 \mathbb{E}(\varepsilon_{t-1}^2) + \\ &+ \beta_1 \beta_2 \mathbb{E}(\varepsilon_{t-2}^2) = \beta_1 \sigma_\varepsilon^2 + \beta_1 \beta_2 \sigma_\varepsilon^2 = (\beta_1 + \beta_1 \beta_2) \cdot \sigma_\varepsilon^2 = \text{const:} \end{aligned}$$

Հաշվենք σ^2 $k=2$ լագով իրարից հեռացած անդամների ավտոկովարիացիան՝

$$\begin{aligned} \hat{\gamma}(2) &= \text{cov}(y_t, y_{t-2}) = \mathbb{E}(y_t \cdot y_{t-2}) = \\ &= \mathbb{E}[(\beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t)(\beta_1 \varepsilon_{t-3} + \beta_2 \varepsilon_{t-4} + \varepsilon_{t-2})] = \beta_2 \mathbb{E}(\varepsilon_{t-2}^2) = \\ &= \beta_2 \cdot \sigma_\varepsilon^2 = \text{const:} \end{aligned}$$

Հաշվենք կամայական $k>2$ լագով իրարից հեռացած անդամների ավտոկովարիացիան՝

$$\begin{aligned} \hat{\gamma}(k) &= \text{cov}(y_t, y_{t-k}) = \mathbb{E}(y_t \cdot y_{t-k}) = \\ &= \mathbb{E}[(\beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \varepsilon_t)(\beta_1 \varepsilon_{t-k-1} + \beta_2 \varepsilon_{t-k-2} + \varepsilon_{t-k})] = 0: \end{aligned}$$

Այսպիսով՝ $\text{MA}(2)$ մոդելով արտահայտվող σ^2 ավտոկովարիացիայի ֆունկցիան կարտահայտվի հետևյալ կերպ՝

$$\hat{\gamma}(k) = \begin{cases} (1 + \beta_1^2 + \beta_2^2) \cdot \sigma_\varepsilon^2, & \text{երբ } k = 0 \\ (\beta_1 + \beta_1 \beta_2) \cdot \sigma_\varepsilon^2, & \text{երբ } k = 1 \\ \beta_2 \sigma_\varepsilon^2, & \text{երբ } k = 2 \\ 0, & \text{երբ } k > 2 \end{cases} : \quad (175)$$

Այսինքն՝ MA(2)-ով արտահայտվող ԺՇ ավտոկովարիացիայի գործակիցները գրոյից տարբերվում են միայն առաջին ու երկրորդ կարգի ավտոկովարիացիայի գործակիցների համար: y_t ԺՇ 2-ից ավելի լագերով իրարից հեռացած անդամների համար ավտոկովարիացիայի գործակիցները հավասար են գրոյի:

Օգտվելով (175)-ից՝ MA(2) մոդելով արտահայտվող ԺՇ ավտոկոռելյացիայի ֆունկցիան կարտահայտվի հետևյալ կերպ՝

$$\frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \hat{\rho}(k) = \begin{cases} \frac{(1+\beta_1^2+\beta_2^2) \cdot \sigma_{\varepsilon}^2}{(1+\beta_1^2+\beta_2^2) \cdot \sigma_{\varepsilon}^2} = 1, \text{ երբ } k = 0 \\ \frac{(\beta_1+\beta_1\beta_2) \cdot \sigma_{\varepsilon}^2}{(1+\beta_1^2+\beta_2^2) \cdot \sigma_{\varepsilon}^2} = \frac{\beta_1+\beta_1\beta_2}{1+\beta_1^2+\beta_2^2}, \text{ երբ } k = 1 \\ \frac{\beta_2 \sigma_{\varepsilon}^2}{(1+\beta_1^2+\beta_2^2) \cdot \sigma_{\varepsilon}^2} = \frac{\beta_2}{1+\beta_1^2+\beta_2^2}, \text{ երբ } k = 2 \\ 0, \text{ երբ } k > 2 \end{cases} \quad (176)$$

(176)-ից երևում է, որ MA(2) մոդելով բնութագրվող ԺՇ ավտոկոռելյացիայի ֆունկցիան երկրորդ լագից հետո կտրուկ անկում է ապրում և սկսում տատանվել գրոյի շուրջ՝ դրանից նշանակալիորեն չտարբերվելով: Կարելի է ցույց տալ, որ սահող միջինների մոդելով արտահայտվող ԺՇ մասնակի ավտոկոռելյացիայի ֆունկցիան արագ ու սահուն, առանց կտրում անկման նվազում է:

6.3. Ընդհանուր դեպք՝ q-րդ կարգի սահող միջինների MA(q) մոդելով բնութագրվող ժամանակային շարքի հատկությունները և սահող միջինների MA(q) մոդելի q կարգի որոշումը

Ունենք y_t ԺՇ-ն, որը մոդելավորվում է q-րդ կարգի սահող միջինների մոդելով՝

$$y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} + \varepsilon_t,$$

որտեղ ε_t -ն սպիտակ աղմուկ ԺՇ է:

Հաշվենք MA(q) մոդելով արտահայտվող ժամանակային շարքի թվային բնութագրիչները: y_t ԺՇ-ի միջինը կունենա հետևյալ տեսքը՝

$$\mathbb{E}(y_t) = \mathbb{E}(\beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} + \varepsilon_t) = \beta_0 = \text{const}$$

MA(q) մոդելով արտահայտվող ԺՇ-ն β_0 -ի, կամ 0-ի շուրջը՝ եթե $\beta_0=0$: Այս ԺՇ-ի վարիացիան՝

$$\begin{aligned}\hat{\gamma}(0) &= \text{Var}(y_t) = \text{Var}(\beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} + \varepsilon_t) = \\ &= \beta_1^2 \sigma_\varepsilon^2 + \beta_2^2 \sigma_\varepsilon^2 + \dots + \beta_q^2 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = (1 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2) \cdot \sigma_\varepsilon^2 = \\ &= \sigma_\varepsilon^2 \cdot \sum_{j=0}^q \beta_j^2 = \text{const:}\end{aligned}$$

Եթե հաշվենք MA(q)-ով արտահայտվող ԺՇ-ի տարրեր կարգի ավտոկովարիացիաները, ապա կստանանք (ենթադրենք $\beta_0 = 0$)`

$$\begin{aligned}\text{ա) } 0 < k \leq q \text{ դեպքում } \gamma(k) &= \text{cov}(y_t, y_{t-k}) = \mathbb{E}(y_t \cdot y_{t-k}) = \\ &= \left\{ \sum_{j=0}^{q-k} \beta_j \beta_{j+k} \right\} \cdot \sigma_\varepsilon^2:\end{aligned}$$

$$\text{բ) } k > q \text{ դեպքում } \gamma(k) = 0:$$

Հետևաբար MA(q) մոդելով արտահայտվող ԺՇ-ի ավտոկովարիացիայի և ավտոկոռելյացիայի ֆունկցիաները կարտահայտվեն հետևյալ կերպ`

Ավտոկովարիացիայի ֆունկցիան`

$$\hat{\gamma}(k) = \begin{cases} \sigma_\varepsilon^2 \cdot \sum_{j=0}^q \beta_j^2, & \text{երբ } k = 0 \\ \left\{ \sum_{j=0}^{q-k} \beta_j \beta_{j+k} \right\} \cdot \sigma_\varepsilon^2, & \text{երբ } 0 < k \leq q: \\ 0, & \text{երբ } k > q \end{cases} \quad (177)$$

Ավտոկոռելյացիայի ֆունկցիան`

$$\frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \hat{\rho}(k) = \begin{cases} 1, & \text{երբ } k = 0 \\ \frac{\sum_{j=0}^{q-k} \beta_j \beta_{j+k}}{\sum_{j=0}^q \beta_j^2}, & \text{երբ } 0 < k \leq q: \\ 0, & \text{երբ } k > q \end{cases} \quad (178)$$

Այսինքն` MA(q) մոդելով արտահայտվող ԺՇ-ի ավտոկոռելյացիայի ֆունկցիայում գ-րդ լազից ավելի մեծ լազերի համար ավտոկոռելյացիայի գործակիցները զրոյից նշանակալիորեն չեն տարբերվում: Դրա ավտոկոռելյացիայի ֆունկցիան գ-րդ լազից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջը:

Հետևաբար, եթե ստացիոնար ԺՇ-ն բնութագրվում է սահող միջինների մոդելով, ապա այդ մոդելի կարգը` q-ն, կարելի է որոշել շար-

քի ավտոկոռելյացիոն ֆունկցիայի միջոցով: Այդ կարգը հավասար է այն լագին, որից հետո ավտոկոռելյացիայի ֆունկցիան կտրուկ անկում է ապրում և սկսում տատանվել գրոյի շուրջը՝ դրանից նշանակալիորեն չտարբերվելով: Սահող միջինների MA(q) մոդելով արտահայտվող ԺՇ-ի մասնակի ավտոկոռելյացիայի ֆունկցիան արագ և սահուն, առանց կտրուկ անկման նվազում է:

7. Ավտոռեգրեսիայի և սահող միջինների ARMA(p, q) մոդելով բնութագրվող ժամանակային շարք, դրա հատկությունները

Հաճախ ԺՇ-ի վարքը ժամանակի մեջ կախված է լինում քե՛ իր նախորդ արժեքներից և քե՛ սպիտակ աղմուկ ԺՇ-ի տվյալ պահի ու նախորդ արժեքներից: Այդ դեպքում ԺՇ-ն արտահայտվում է ավելի բարդ տեսքի ARMA(p, q) մոդելով: Այս մոդելը բաղկացած է p կարգի ավտոռեգրեսիայի մասից և q կարգի սահող միջինների մասից: Ժամանակային շարքը բնութագրող ARMA(p, q) մոդելն ունի հետևյալ տեսքը՝

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q} + \varepsilon_t, \quad (179)$$

որտեղ ε_t -ն սպիտակ աղմուկ ԺՇ է:

Դիտարկենք պարզագույն դեպքը՝ ARMA(1, 1) մոդելը, որում y_t ԺՇ-ի յուրաքանչյուր t պահի արժեքը գծորեն կախված է իր նախորդ պահի, սպիտակ աղմուկ ε_t ԺՇ-ի նախորդ և t պահերի արժեքներից՝

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t: \quad (180)$$

Հաշվենք ARMA(1, 1) մոդելով արտահայտվող ստացիոնար ԺՇ-ի միջինը և վարիացիան՝

$$\begin{aligned} \mathbb{E}(y_t) &= \mathbb{E}(\alpha_0 + \alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t) = \\ &= \alpha_0 + \alpha_1 \mathbb{E}(y_{t-1}) + \beta_1 \mathbb{E}(\varepsilon_{t-1}) + \mathbb{E}(\varepsilon_t) = \alpha_0 + \alpha_1 \mathbb{E}(y_{t-1}) \end{aligned}$$

Քանի որ y_t -ն ստացիոնար ԺՇ է՝ հաստատուն μ միջինով, ապա՝

$$\mathbb{E}(y_t) = \mu = \alpha_0 + \alpha_1 \mu \rightarrow (1 - \alpha_1) \mu = \alpha_0 \rightarrow \mu = \frac{\alpha_0}{1 - \alpha_1}: \quad (181)$$

Ստացանք, որ ARMA(1, 1) մոդելով արտահայտվող ստացիոնար ԺՇ-ի միջինը AR(1) մոդելի միջինն է (տե՛ս (148)):

ARMA(1, 1) մոդելով արտահայտվող ստացիոնար ԺՇ-ի վարիացիան՝

$$\hat{\gamma}(0) = \text{Var}(y_t) = \text{Var}(\alpha_0 + \alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t) = \\ = \alpha_1^2 \text{Var}(y_{t-1}) + \beta_1^2 \text{Var}(\varepsilon_{t-1}) + \text{Var}(\varepsilon_t)$$

Քանի որ y_t -ն ստացիոնարն ԺՇ է՝ հաստատուն σ_y^2 վարիացիայով, ապա՝

$$\hat{\gamma}(0) = \sigma_y^2 = \alpha_1^2 \cdot \sigma_y^2 + \beta_1^2 \cdot \sigma_\varepsilon^2 + \sigma_\varepsilon^2 = \alpha_1^2 \cdot \sigma_y^2 + (1 + \beta_1^2) \cdot \sigma_\varepsilon^2 \rightarrow \\ (1 - \alpha_1^2) \cdot \sigma_y^2 = (1 + \beta_1^2) \cdot \sigma_\varepsilon^2 \rightarrow \hat{\gamma}(0) = \sigma_y^2 = \frac{(1 + \beta_1^2) \cdot \sigma_\varepsilon^2}{(1 - \alpha_1^2)} : \quad (182)$$

(182)-ից հետևում է, որ ARMA(1,1) մոդելում $\alpha_1^2 < 1 \rightarrow |\alpha_1| < 1$: Այս արդյունքը նույնական է AR(1) մոդելով արտահայտվող ԺՇ-ի ստացիոնարության պայմանին: Եթե $|\alpha_1| < 1$, ապա ARMA(1, 1)-ով բնութագրվող ժամանակային շարքը ստացիոնար է:

Հաշվենք ARMA(1,1)-ով բնութագրվող ժամանակային շարքի անդամների ավտոկովարիացիաները՝ պարզության համար ենթադրենք, որ $\alpha_0 = 0$:

$$\begin{aligned} k=1 \quad \hat{\gamma}(1) &= \mathbb{E}(y_t \cdot y_{t-1}) = \mathbb{E}((\alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t) \cdot y_{t-1}) = \\ &= \mathbb{E}(\alpha_1 y_{t-1} \cdot y_{t-1} + \beta_1 \varepsilon_{t-1} \cdot y_{t-1} + \varepsilon_t \cdot y_{t-1}) = \alpha_1 \hat{\gamma}(0) + \beta_1 \sigma_\varepsilon^2 \\ k=2 \quad \hat{\gamma}(2) &= \mathbb{E}(y_t \cdot y_{t-2}) = \mathbb{E}((\alpha_1 y_{t-1} + \beta_1 \varepsilon_{t-1} + \varepsilon_t) \cdot y_{t-2}) = \\ &= \mathbb{E}(\alpha_1 y_{t-1} \cdot y_{t-2} + \beta_1 \varepsilon_{t-1} \cdot y_{t-2} + \varepsilon_t \cdot y_{t-2}) = \alpha_1 \hat{\gamma}(1) \\ k \geq 2 \quad \hat{\gamma}(k) &= \alpha_1 \hat{\gamma}(k-1) \end{aligned} \quad (183)$$

(183)-ից կստանանք.

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\alpha_1 \hat{\gamma}(k-1)}{\hat{\gamma}(0)} = \alpha_1 \cdot \hat{\rho}(k-1) \quad (184)$$

(184)-ից և $|\alpha_1| < 1$ պայմանից հետևում է, որ ARMA(1,1)-ով բնութագրվող ԺՇ-ի ավտոկոռելյացիայի ֆունկցիան $k \geq 2$ լագերի դեպքում արագ, էքսպոնենցիալ օրենքով կամ նշանը փոխելով՝ նվազում է: Սա նշանակում է նաև, որ ARMA(1,1) մոդելի սահող միջինների մասի կարգը՝ $q=1$:

Կարելի է ցույց տալ, որ ARMA(1,1)-ով բնութագրվող ԺՇ-ի մասնակի ավտոկոռելյացիայի ֆունկցիան նույնպես $k \geq 2$ լագերի դեպքում արագ, էքսպոնենցիալ օրենքով կամ նշանը փոխելով՝ նվազում

է: Սա նշանակում է նաև, որ $ARMA(1,1)$ մոդելի ավտոռեգրեսիայի մասի կարգը՝ $p=1$:

Ընդհանուր դեպքում $ARMA(p, q)$ մոդելով բնութագրվող y_t ԺՇ-ի հատկություններն են՝

1. $ARMA(p, q)$ -ով բնութագրվող ժամանակային շարքի ստացիոնարության պայմանն է՝ $|\alpha_1| < 1, |\alpha_2| < 1 \dots |\alpha_p| < 1$,
2. $E(y_t) = 0$, եթե $\alpha_0 = 0$ և $E(y_t) = \frac{\alpha_0}{1 - (\alpha_1 + \alpha_2 + \dots + \alpha_p)}$, եթե $\alpha_0 \neq 0$
 $\rightarrow \alpha_1 + \alpha_2 + \dots + \alpha_p \neq 1$,
3. $ARMA(p, q)$ մոդելով արտահայտվող ԺՇ-ի ավտոկոռելյացիայի ֆունկցիան արագ՝ առանց կտրուկ անկման, նվազում է և q լագից հետո սկսում տատանվել զրոյի շուրջը:
4. $ARMA(p, q)$ մոդելով արտահայտվող ԺՇ-ի մասնակի ավտոկոռելյացիայի ֆունկցիան արագ, առանց կտրուկ անկման նվազում է և p լագից հետո սկսում տատանվել զրոյի շուրջը:

Այսպիսով՝ $ARMA(p, q)$ մոդելով բնութագրվող ԺՇ ավտոկոռելյացիայի ու մասնակի ավտոկոռելյացիայի ֆունկցիաները միևնույն վարքն ունեն, դրանք արագ նվազում են՝ առանց կտրուկ անկման, և որոշակի լագից հետո սկսում տատանվել զրոյի շուրջը:

8. Ավտոռեգրեսիայի ու սահող միջինների ինտեգրացվող $ARIMA(p, d, q)$ մոդելն ու դրա գնահատման Բոքս-Ջենքինսի մեթոդաբանությունը

Ենթադրենք y_t -ն d -րդ կարգի ինտեգրացվող շարք է՝ $I(d)$, այսինքն՝ այն ունի ստոխաստիկ տրենդ և d անգամ տարբերական օպերատոր կիրառելուց հետո ստացիոնարանում է: Եթե ստացված ստացիոնար շարքը բնութագրվում է $ARMA(p, q)$ մոդելով, ապա ասում են, որ սկզբնական y_t շարքը բնութագրվում է $ARIMA(p, d, q)$ կամ, որ նույնն է, d -րդ կարգի ինտեգրացվող, p -րդ կարգի ավտոռեգրեսիայի ու q -րդ կարգի սահող միջինների մոդելով:

Եթե y_t -ն ստացիոնար է և նկարագրվում է $ARMA(p, q)$ մոդելով, ապա այն կներկայացվի $ARIMA(p, 0, q)$ մոդելով: Եթե y_t -ն նկարագրվում է $MA(q)$ մոդելով, ապա այն կներկայացվի $ARIMA(0, 0, q)$ մոդելով: Եթե y_t -ն d -րդ կարգի ինտեգրացվող է և նկարագրվում է $AR(p)$ մոդելով, ապա կներկայացվի $ARIMA(p, d, 0)$ մոդելով:

Միաշափ ԺՇ-երի ARIMA մոդելավորման համար մշակված է Բոքս-Ջենքինսի մեթոդաբանությունը, որի քայլերի հակիրճ նկարագրությունն է՝

1. ԺՇ-ի նախնական ուսումնասիրություն,
2. մոդելի հատկորոշում՝ ARIMA մոդելի p , d , q կարգերի որոշում,
3. ARIMA(p , d , q) մոդելի գնահատում,
4. մոդելի գնահատման արդյունքների վավերականության ու որակի ստուգում,
5. կանխատեսում:

ԺՇ նախնական ուսումնասիրությունն ընդգրկում է դրա գրաֆիկի, ավտոկոռելյացիայի ֆունկցիայի, սեզոնայնության, դետերմինացված ու ստոխաստիկ միտումների առկայության ուսումնասիրությունը, ինչպես նաև ստացիոնարության թեստերի իրականացումը: Նախնական ուսումնասիրության արդյունքում որոշվում է շարքի ինտեգրացման կարգը՝ d -ն:

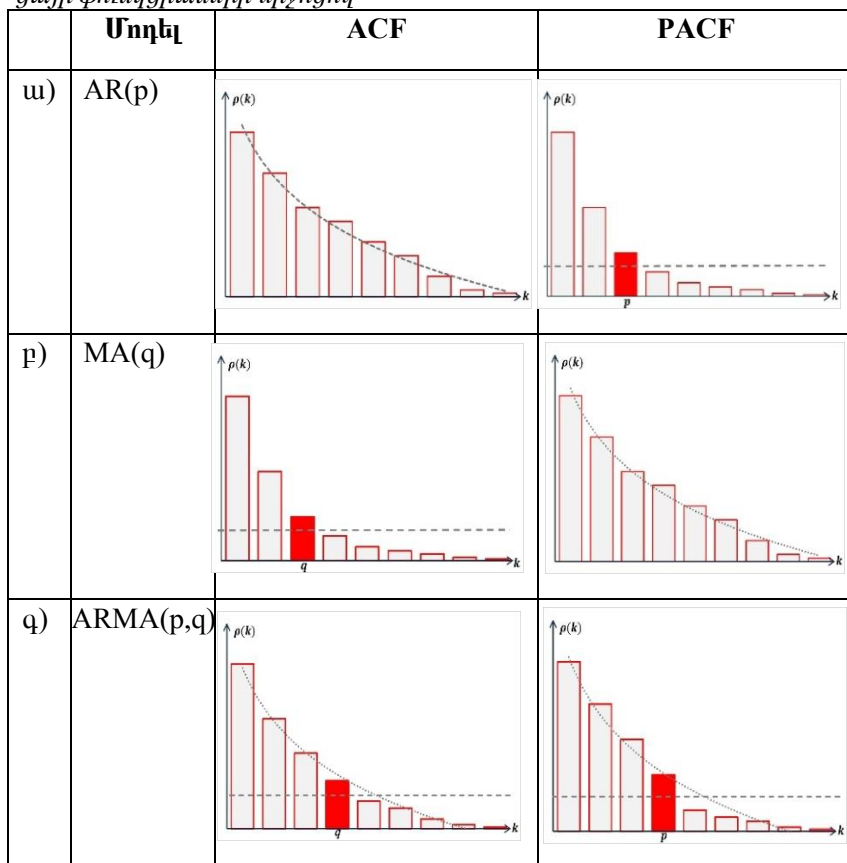
Մոդելի հատկորոշումը ենթադրում է նաև դրա p ու q կարգերի որոշում: Այդ նպատակով շարքը ստացիոնարացվում է, և կառուցվում են ստացիոնարացված շարքի ավտոկոռելյացիայի ու մասնակի ավտոկոռելյացիայի ֆունկցիաները: Դրանց վարքից ելնելով՝ p ու q կարգերը որոշվում են Այուսակ 11-ում ներկայացված մեխանիզմով:

Այուսակ 11. AR, MA ու ARMA մոդելների կարգերի որոշման մեխանիզմը

	Ավտոկոռելյացիայի ֆունկցիա՝ ACF	Մասնակի ավտոկոռելյացիայի ֆունկցիա՝ PACF
AR(p)	Արագ և սահուն նվազում է, առանց կտրուկ անկման:	p լազից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջը՝ բոլոր $k > p$ լազերի համար:
MA(q)	q լազից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջ՝ բոլոր $k > q$ լազերի համար:	Արագ և սահուն նվազում է, առանց կտրուկ անկման:
ARMA(p , q)	Արագ նվազում է՝ առանց կտրուկ անկման և $k > q$ լազից հետո սկսում է տատանվել զրոյի շուրջ:	Արագ նվազում է՝ առանց կտրուկ անկման և $k > p$ լազից հետո սկսում է տատանվել զրոյի շուրջ:

Այս պարզ սխեման ներկայացված է Գծանկար 30-ում: Այսպիսի կոռելագրամներ ունեցող ԺՇ-ն կրնոթագրվի. ա) $ARIMA(3,0,0)$, բ) $ARIMA(0,0,3)$, գ) $ARIMA(5,0,5)$ մոդելով:

Գծանկար 30. Սրահցիոնար ժամանակային շարքի ARMA մոդելի p , q կարգերի որոշումը ավրոկոռելյացիայի ու մասնակի ավրոկոռելյացայի ֆունկցիաների միջոցով



Այնուհետև գնահատվում է ընտրված $ARIMA(p, d, q)$ մոդելը ՓՋԵ-ով կամ առավելագույն ճշմարտամամուրթյան եղանակով: Հետագոտողը կարող է դիտարկել նաև p ու q կարգերի՝ մեկից ավելի նախընտրելի տարբերակներ: Այդ դեպքում գնահատվում են մեկից

ավելի մոդելներ, և դրանցից լավագույնն ընտրվում է՝ ելնելով մոդելի բնութագրիչներից ու տեղեկատվական չափանիշներից: Նշենք նաև, որ ARIMA մոդելը գծային է, հետևաբար, եթե ժամանակի մեջ շարքը չունի գծային վարք, անհրաժեշտ է նախքան մոդելավորելը այն բնական հիմքով լուծարիքմել:

Դրանից հետո ստուգվում են գնահատված մոդելի որակն ու վավերականությունը: Գնահատված հավասարման, ինչպես նաև յուրաքանչյուր գործակցի վիճակագրական նշանակալիությունը ստուգվում է ՓԶԵ ընդհանուր մոտեցմամբ՝ օգտագործելով F ու t վիճականիները: Իսկ մոդելի որակը ստուգվում է դետերմինացիայի գործակցի միջոցով, ինչպես նաև մոդելի կայունությունը բնութագրող որոշ չափանիշներով: Վավերականության ստուգումը շարունակվում է մոդելի պատահական մնացորդների հատկությունների ուսումնասիրությամբ: Ստուգվում է, թե մոդելի պատահական մնացորդները բավարարո՞ւմ են սպիտակ աղմուկ σ^2 հատկությունները՝ զրոյին հավասար միջին, հոմոսկեդաստիկություն ու ավտոկոռելացվածության բացակայություն, ստացիոնարություն: Եթե գնահատվել են մեկից ավելի մոդելներ, ապա դրանցից ընտրվում է այն, որն ունի տեղեկատվական չափանիշների փոքրագույն արժեքներ և վավերական է:

Մոդելի վերջնական ընտրությունից ու գնահատումից հետո կարող են կանխատեսվել կախյալ փոփոխականի ապագա արժեքները՝ ելնելով գնահատման արդյունքներից: Ժամանակի մեջ մեկ քայլ առաջ կանխատեսելու համար կախյալ փոփոխականի կանխատեսված արժեքը հաշվարկվում է այն սկզբունքով, որ պատահական տատանումների կամ սպիտակ աղմուկի հաջորդ պահի արժեքը հավասար է իր սպասվող արժեքին, այսինքն՝ զրոյի: Ընդհանուր դեպքում՝

$$\hat{y}_{t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 y_t + \hat{\alpha}_2 y_{t-1} + \dots + \hat{\alpha}_p y_{t-p+1} + \hat{\beta}_1 \varepsilon_t + \hat{\beta}_2 \varepsilon_{t-1} + \dots + \hat{\beta}_q \varepsilon_{t-q+1} \quad (185)$$

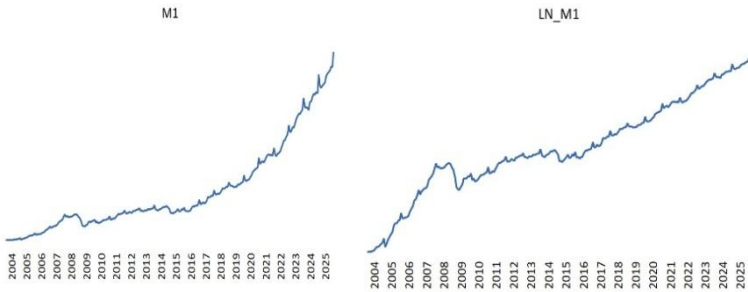
Կախյալ փոփոխականի հաջորդ պահերի արժեքները հերթականությամբ կանխատեսվում են՝ օգտագործելով դրա նախորդ պահի կանխատեսված արժեքները, նորից պատահական տատանումների կամ սպիտակ աղմուկի հաջորդ պահի արժեքները վերցնելով հավասար զրոյի: Կանխատեսման անորոշությունը մեծանում է դրա հորիզոնի աճի հետ: Կանխատեսման սխալների վարիացիան կախված է պատահական տատանումների կառուցվածքից: Մեծ կանխատեսման հորիզոնի պարագայում կանխատեսումները համընկնում են գնահատվող շարքի միջինի հետ (եթե այն ստացիոնար է):

9. Ժամանակային շարքի ARIMA(p, d, q) մոդելավորման օրինակ

Ուսումնասիրվում է Հայաստանում փողի ագրեգատի՝ M1-ի ամսական կտրվածքով ԺՇ-ն 2004-2025 թթ. համար (264 տվյալ)՝ կիրառելով «EViews» փաթեթը: ԺՇ-ի գրաֆիկի նախնական ուսումնասիրությունը փաստում է, որ այն ունի ոչ գծային միտում ու 12 պարբերությամբ սեզոնայնություն: Քանի որ ARIMA մոդելները գծային տեսք ունեն, նախ անհրաժեշտ է գծայնացնել շարքի կախումը ժամանակից: Ոչ գծային կախվածությունները գծայնացնելու համար արդեն հայտնի է, որ կիրառվում է տվյալները լոգարիթմելու մեթոդը: ԺՇ-ն բնական հիմքով լոգարիթմելուց հետո դրա կախումը ժամանակից գծայնանում է:

M1 ԺՇ-ի գրաֆիկը

*Լոգարիթմած M1
ԺՇ-ի գրաֆիկը*



Այնուհետև շարքը ապասեզոնայնացվում է որևէ հասանելի եղանակով, և մոդելավորումը շարունակվում է լոգարիթմված ու ապասեզոնայնացված Ln_M1_SA շարքի հետ: Այս շարքն ունի ակնհայտ աճի միտում, ուստի ստացիոնար չէ:

Լոգարիթմած և ապասեզոնայնացված Ln_M1_SA ԺՇ-ի գրաֆիկը



Ստացիոնարության Դիկի-Ֆուլերի թեստը փաստում է, որ այս ԺՇ-ն ունի ստոխաստիկ միտում և առաջին կարգի ինտեգրացվող է. այն ստացիոնար է դառնում մեկ անգամ տարբերական օպերատորը կիրառելուց հետո:

Ln_M1_SA ԺՇ-ի Դիկի-Ֆուլերի սրացիոնարության թեստը

Null Hypothesis: LN M1 SA has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 3 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.528811	0.3140
Test critical values: 1% level	-3.993746	
5% level	-3.427203	
10% level	-3.136898	

*MacKinnon (1996) one-sided p-values.

Առաջին տարբերությանը d(Ln_M1_SA) ԺՇ-ի Դիկի-Ֆուլերի սրացիոնարության թեստը

Null Hypothesis: D(LN M1 SA) has a unit root

Exogenous: Constant, Linear Trend

Lag Length: 2 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-5.759085	0.0000
Test critical values: 1% level	-3.993746	
5% level	-3.427203	
10% level	-3.136898	

*MacKinnon (1996) one-sided p-values.

Նույնն է փաստում նաև ԺՇ ավտոկոռելյացիայի ֆունկցիան: Այն շատ դանդաղ է նվազում ԺՇ-ի համար, սակայն մեկ անգամ տարբերությանը ԺՇ-ի համար արդեն արագ, առանց կտրուկ անկման է նվազում: Տարբերությանը ԺՇ հիշողությունը չի հասնում նույնիսկ մեկ տարվա, տասը ամսից հետո ավտոկոռելյացիայի գործակիցները տատանվում են զրոյի շուրջ: Հետևաբար ARIMA(p,d,q) մոդելի ինտեգրացման կարգը՝ d=1:

Մեկ անգամ տարբերությամբ ԺՇ մասնակի ավտոկոռելյացիայի ֆունկցիան 3-րդ լագից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջը: Մեկ անգամ տարբերությամբ ԺՇ մասնակի ավտոկոռելյացիաների կոռելյոգրամում գործակիցները 3-րդ լագից հետո ընկած են դրանց վիճակագրական նշանակալիության $\pm 1.96 \cdot \sqrt{\frac{1}{n}}$ միջակայքի ներսում: Դա փաստում է, որ վիճակագրորեն նշանակալի են միայն առաջին երեք կարգի մասնակի ավտոկոռելյացիայի գործակիցները, իսկ դրանից մեծ լագերին համապատասխանող գործակիցները 0-ից նշանակալիորեն չեն տարբերվում: Եզրակացությունը. ստացիոնարացված ԺՇ -ն կարելի է մոդելավորել որպես $\text{AR}(p)=\text{AR}(3)$ մոդել, քանի որ դրա ավտոկոռելյացիայի ֆունկցիան սահում և արագ նվազում է, իսկ մասնակի ավտոկոռելյացիայի ֆունկցիան $p=3$ լագից հետո կտրուկ անկում է ապրում և սկսում տատանվել զրոյի շուրջ՝ բոլոր $k>3$ լագերի համար:

Ln_M1_SA ԺՇ-ի կոռելյոգրամը

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.984	0.984	258.45	0.000
		2 0.967	-0.017	509.28	0.000
		3 0.950	-0.022	752.34	0.000
		4 0.933	-0.016	987.62	0.000
		5 0.916	-0.011	1215.2	0.000
		6 0.899	-0.015	1435.0	0.000
		7 0.881	-0.022	1646.9	0.000
		8 0.863	-0.016	1851.1	0.000
		9 0.844	-0.014	2047.5	0.000
		10 0.826	-0.009	2236.2	0.000
		11 0.808	-0.012	2417.4	0.000
		12 0.790	-0.011	2591.1	0.000
		13 0.771	-0.022	2757.4	0.000
		14 0.753	0.016	2916.6	0.000
		15 0.735	0.004	3069.1	0.000
		16 0.719	0.008	3215.4	0.000

Առաջին փաթեթերության $d(Ln_M1_SA)$ ԺՇ-ի կոռելյոգրամը						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.302	0.302	24.236	0.000
		2	0.290	0.218	46.636	0.000
		3	0.285	0.174	68.417	0.000
		4	0.182	0.026	77.365	0.000
		5	0.138	-0.004	82.501	0.000
		6	0.101	-0.015	85.247	0.000
		7	0.152	0.088	91.534	0.000
		8	0.118	0.039	95.339	0.000
		9	0.140	0.063	100.71	0.000
		10	0.099	-0.012	103.39	0.000
		11	0.060	-0.039	104.39	0.000
		12	-0.057	-0.153	105.28	0.000
		13	0.055	0.071	106.14	0.000
		14	0.056	0.069	107.02	0.000
		15	-0.058	-0.080	107.98	0.000
		16	0.010	-0.010	108.01	0.000

Դիտարկենք ARIMA(3,1,q) մոդելի սահող միջինների MA բաղադրիչի հնարավոր կարգը: Եթե ստացիոնարացված ԺՇ մոդելն ունի միայն սահող միջինների բաղադրիչ, ապա դրա ավտոկոռելյացիայի ֆունկցիան պետք է որևէ q լագից հետո կտրուկ անկում ապրի և սկսի տատանվել զրոյի շուրջը, իսկ մասնակի ավտոկոռելյացիայի ֆունկցիան պետք է արագ ու սահուն նվազի, մինչդեռ ստացիոնարացված ԺՇ ավտոկոռելյացիայի ֆունկցիան սահուն է նվազում՝ առանց կտրուկ անկման, իսկ մասնակի ավտոկոռելյացիայի ֆունկցիան երրորդ լագից հետո կտրուկ անկում է ապրում: Ուրեմն այս ԺՇ-ն միայն սահող միջինների մոդելով չի ներկայացվի: Հետևաբար այս ԺՇ ARIMA մոդելի կարգերն են՝ ARIMA(3,1,0), իսկ գնահատվելիք մոդելն ունի հետևյալ տեսքը՝

$$\Delta(y_t) = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \varepsilon_t:$$

«EViews» փաթեթում այն ՓՔԵ-ով գնահատելու արդյունքում ստացվում է հետևյալ ելքային աղյուսակը՝

Sample (adjusted): 2003M05 2025M06

Included observations: 266 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.0056	0.0017	3.3772	0.0008
D(LN_M1_SA(-1))	0.1576	0.0607	2.5975	0.0099
D(LN_M1_SA(-2))	0.1814	0.0602	3.0126	0.0028
D(LN_M1_SA(-3))	0.1896	0.0605	3.1347	0.0019
R-squared	0.1425	Mean dependent var		0.0119
Adjusted R-squared	0.1327	S.D. dependent var		0.0241
S.E. of regression	0.0225	Akaike info criterion		-4.7382
Sum squared resid	0.1323	Schwarz criterion		-4.6843
Log likelihood	634.1800	Hannan-Quinn criter.		-4.7165
F-statistic	14.5183	Durbin-Watson stat		2.0017
Prob(F-statistic)	0.0000			

Գնահատված հավասարումն է՝

$$\Delta(\widehat{\ln y}_t) = 0.0056 + 0.1576 \cdot \Delta(\ln y_{t-1}) + 0.1814 \cdot \Delta(\ln y_{t-2}) + 0.1896 \cdot \Delta(\ln y_{t-3}):$$

Ln_M1_SA-ի ԺՇ մոդելի բանաձևը ստանալու համար հիշենք, որ՝

$$\Delta(y_t) = y_t - y_{t-1} :$$

Հետևաբար՝

$$\Delta(y_{t-1}) = y_{t-1} - y_{t-2},$$

$$\Delta(y_{t-2}) = y_{t-2} - y_{t-3},$$

$$\Delta(y_{t-3}) = y_{t-3} - y_{t-4}:$$

Չնափոխենք գնահատված հավասարումը՝ կիրառելով վերը բերված բանաձևերը:

$$\begin{aligned} \Delta(\widehat{\ln y}_t) &= 0.0056 + 0.1576 \cdot \Delta(\widehat{\ln y}_{t-1}) + 0.1814 \cdot \Delta(\widehat{\ln y}_{t-2}) + \\ &+ 0.1896 \cdot \Delta(\widehat{\ln y}_{t-3}) = 0.0056 + 0.1576 \cdot (\widehat{\ln y}_{t-1} - \widehat{\ln y}_{t-2}) + \\ &+ 0.1814 \cdot (\widehat{\ln y}_{t-2} - \widehat{\ln y}_{t-3}) + 0.1896 \cdot (\widehat{\ln y}_{t-3} - \widehat{\ln y}_{t-4}) = \\ &= 0.0056 + 0.1576 \cdot \widehat{\ln y}_{t-1} - 0.1576 \cdot \widehat{\ln y}_{t-2} + 0.1814 \cdot \widehat{\ln y}_{t-2} - \end{aligned}$$

$$-0.1814 \cdot \widehat{\ln y}_{t-3} + 0.1896 \cdot \widehat{\ln y}_{t-3} - 0.1896 \cdot \widehat{\ln y}_{t-4} = 0.0056 + \\ + 0.1576 \cdot \widehat{\ln y}_{t-1} + 0.0238 \widehat{\ln y}_{t-2} + 0.0082 \cdot \widehat{\ln y}_{t-3} - 0.1896 \cdot \widehat{\ln y}_{t-4}$$

Արտահայտենք $\hat{y}_t$ -ը: Ունենք՝

$$\Delta(\widehat{\ln y}_t) = \widehat{\ln y}_t - \widehat{\ln y}_{t-1} = 0.0056 + 0.1576 \cdot \widehat{\ln y}_{t-1} + \\ + 0.0238 \widehat{\ln y}_{t-2} + 0.0082 \cdot \widehat{\ln y}_{t-3} - 0.1896 \cdot \widehat{\ln y}_{t-4}$$

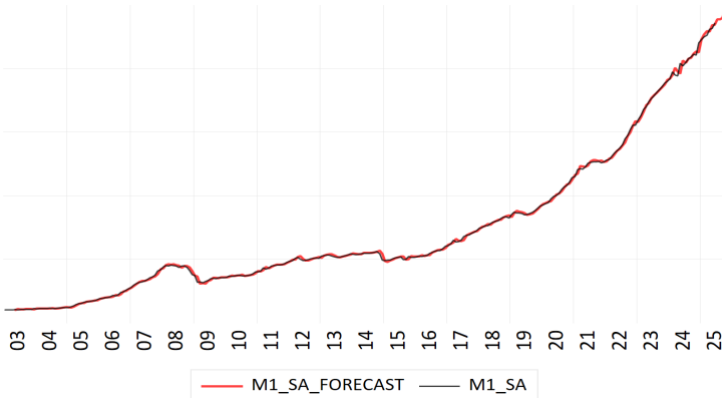
Հետևաբար՝ $\widehat{\ln y}_t = 0.0056 + 1.1576 \cdot \widehat{\ln y}_{t-1} + 0.0238 \cdot \widehat{\ln y}_{t-2} + \\ + 0.0238 \cdot \widehat{\ln y}_{t-3} - 0.1896 \cdot \widehat{\ln y}_{t-4}$ և

$$\hat{y}_t = e^{0.0056} \cdot \hat{y}_{t-1}^{1.1576} \cdot \hat{y}_{t-2}^{0.0238} \cdot \hat{y}_{t-3}^{0.0238} \cdot \hat{y}_{t-4}^{-0.1896}$$

Սա ապաստեզոնայնացված շարքի գնահատված հավասարումն է: Այս հավասարման միջոցով ընդունելով, որ մոդելի մնացորդների սպասվող արժեքները $t+1$, $t+2$ պահերին հավասար են զրոյի՝ կարելի է ստանալ կախյալ փոփոխականի կանխատեսվող կամ սպասվող արժեքները, օրինակ, հաջորդ 3 ամիսների համար:

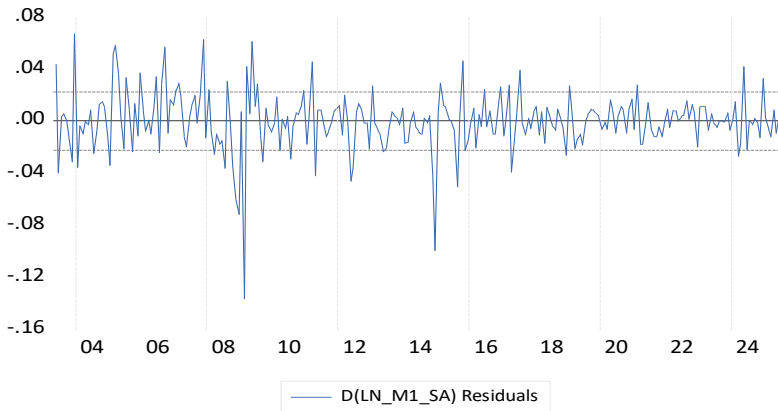
$$\hat{y}_{t+1} = e^{0.005} \cdot \hat{y}_t^{1.183} \cdot \hat{y}_{t-1}^{0.003} \cdot \hat{y}_{t-2}^{-0.009} \cdot \hat{y}_{t-3}^{-0.176} \\ \hat{y}_{t+2} = e^{0.005} \cdot \hat{y}_{t+1}^{1.183} \cdot \hat{y}_t^{0.003} \cdot \hat{y}_{t-1}^{-0.009} \cdot \hat{y}_{t-2}^{-0.176} \\ \hat{y}_{t+3} = e^{0.005} \cdot \hat{y}_{t+2}^{1.183} \cdot \hat{y}_{t+1}^{0.003} \cdot \hat{y}_t^{-0.009} \cdot \hat{y}_{t-1}^{-0.176}$$

Կարելի է ստանալ այս մոդելով գնահատված ապաստեզոնայնացված $\hat{y}_t$ -ի (M1_SA_FORECAST), դրա ապագա արժեքների կանխատեսումների ու ապաստեզոնայնացված y_t -ի իրական, դիտարկված արժեքների (M1_SA) համապատասխանությունը գրաֆիկի վրա:



Նկատելի է, որ չնայած մոդելի գնահատման որակը ցածր է, դետերմինացիայի ճշգրտված գործակիցը հավասար է 0.142, այնուամենայնիվ մոդելավորված $\hat{y}_t$ -ի արժեքները բավական լավ են մոտարկել y_t -ի իրական, դիտարկված արժեքներին, ապաստեգոնայնացված շարքին: Այնուհետև մոդելի վավերականության ստուգումը սկսենք F և t վիճականիներից: Դրանք փաստում են, որ ինչպես ամբողջ հավասարումը, այնպես էլ բոլոր գնահատված գործակիցները վիճակագրորեն նշանակալի են: Պետք է ստուգել, թե արդյոք գնահատված մոդելի պատահական մնացորդը սպիտակ աղմուկ $\mathcal{CC}$ է, այսինքն՝ դրա միջինը զրոն է, դրա ոչ մի անդամը մյուսի հետ կոռելյացված չէ, և դրանք հոմոսկեդասիկ են: Այդ ստուգումները կարելի է իրականացնել արդեն հայտնի վիճակագրական թեստերի միջոցով:

Նախ ստանանք այդ մնացորդների գրաֆիկը: Այն փաստում է, որ մնացորդների շարքը տատանվում է զրո միջինի շուրջը: Մնացորդների $\mathcal{CC}$ միջինը կարելի է ստանալ նաև դրանց հաճախությունների բաշխման հիստոգրամը ստանալու միջոցով:



Մնացորդների $\mathcal{CC}$ ավտոկոռելյացիայի ֆունկցիան ու Բոքս-Լյունգի վիճականին փաստում են, որ ավտոկոռելյացիայի բոլոր գործակիցները գտնվում են $\pm 1.96 \cdot \sqrt{\frac{1}{n}}$ միջակայքի ներսում, այսինքն՝ վիճակագրորեն նշանակալի չեն, իսկ Բոքս-Լյունգի վիճականու p արժեքը 0,05-ից բավական մեծ է բոլոր լազերի համար: Ուստի մնացորդները ավտոկոռելյացված չեն: Դրա մասին է փաստում նաև Բրեուշ-Գորդֆրիի Լագրանժի բազմապատկչի թեստը մինչև 36 լազի համար:

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob*
		1 -0.004	-0.004	0.0053	0.942
		2 -0.005	-0.005	0.0114	0.994
		3 -0.003	-0.003	0.0141	1.000
		4 0.010	0.010	0.0415	1.000
		5 -0.023	-0.023	0.1818	0.999
		6 -0.052	-0.052	0.9049	0.989
		7 0.060	0.059	1.8693	0.967
		8 0.034	0.034	2.1846	0.975
		9 0.105	0.106	5.1527	0.821
		10 0.041	0.044	5.6107	0.847
		11 0.008	0.007	5.6284	0.897
		12 -0.134	-0.136	10.542	0.569
		13 0.050	0.055	11.244	0.590
		14 0.054	0.059	12.051	0.602
		15 -0.110	-0.103	15.408	0.422
		16 -0.029	-0.038	15.635	0.479
		17 0.063	0.047	16.737	0.472
		18 0.049	0.024	17.426	0.494
		19 0.062	0.083	18.496	0.490
		20 0.027	0.032	18.698	0.542
		21 0.021	0.024	18.827	0.596
		22 -0.015	-0.010	18.887	0.652
		23 -0.020	-0.018	19.005	0.701
		24 -0.084	-0.090	21.042	0.636

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.367522	Prob. F(10,235)	0.1962
Obs*R-squared	13.69308	Prob. Chi-Square(10)	0.1875

Սակայն հոմոսկեդաստիկության Ուայթի թեստը փաստում է, որ մնացորդները հետերոսկեդաստիկ են: Սա չի բավարարում մնացորդների սպիտակ աղմուկ լինելու պայմանին, ինչից հետևում է, որ մոդելը վավերական չէ:

Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	6.7807	Prob. F(9,256)	0.0000
Obs*R-squared	51.2042	Prob. Chi-Square(9)	0.0000
Scaled explained SS	214.8812	Prob. Chi-Square(9)	0.0000

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Բնութագրե՛ք առաջին կարգի ավտոռեգրեսիայի մոդելը, դո՛րս բերե՛ք դրա հիմնական հատկությունները՝ թվային բնութագրիչներն ու ավտոկոռելյացիաները:
2. Ո՞րն է $AR(1)$ մոդելով բնութագրվող ԺՇ ստացիոնարության պայմանը, և ի՞նչ տեսք ունի դրա ավտոկոռելյացիայի ֆունկցիան:
3. Բնութագրե՛ք երկրորդ կարգի ավտոռեգրեսիայի մոդելը, դո՛րս բերե՛ք դրա հիմնական հատկությունները՝ թվային բնութագրիչներն ու ավտոկոռելյացիաները:
4. Ո՞րն է $AR(2)$ մոդելով բնութագրվող ԺՇ ստացիոնարության պայմանը, և ի՞նչ տեսք ունի դրա ավտոկոռելյացիայի ֆունկցիան:
5. Բնութագրե՛ք p -րդ կարգի ավտոռեգրեսիայի մոդելը, դո՛րս բերե՛ք դրա հիմնական հատկությունները՝ թվային բնութագրիչներն ու ավտոկոռելյացիաները:
6. Ո՞րն է $AR(p)$ մոդելով բնութագրվող ԺՇ ստացիոնարության պայմանը, և ի՞նչ տեսք ունի դրա ավտոկոռելյացիայի ֆունկցիան:
7. Ի՞նչ տեսք ունեն Ξ ուլ-Ուոքերի հավասարումները: Դո՛րս բերե՛ք դրանք $AR(1)$, $AR(2)$ ու $AR(p)$ մոդելների համար:
8. Ինչպե՞ս կարելի է ստանալ ԺՇ-ի մասնակի ավտոկոռելյացիայի ֆունկցիան, և ինչի՞նչ միջոցով կարելի է որոշել այդ ԺՇ ավտոռեգրեսիայի մոդելի կարգը՝ p -ն:
9. Բնութագրե՛ք սահող միջինների $MA(1)$, $MA(2)$ ու $MA(q)$ մոդելներով բնութագրվող ժամանակային շարքերի հիմնական հատկությունները և դո՛րս բերե՛ք դրանց ավտոկովարիացիայի ու ավտոկոռելյացիայի ֆունկցիաները:
10. Բնութագրե՛ք $ARMA(p, q)$ մոդելով բնութագրվող ԺՇ-ի հատկությունները:
11. Նկարագրե՛ք ավտոռեգրեսիայի և սահող միջինների ինտեգրացվող $ARIMA(p, d, q)$ մոդելն ու դրա գնահատման Բոքս-Ջենքինսի մեթոդաբանությունը:
12. Ի՞նչ սկզբունքով է որոշվում, թե որոնք են տվյալ ստացիոնար ԺՇ-ի $ARMA$ մոդելի p ու q կարգերը:
13. Որո՞նք են $ARIMA(p, d, q)$ գնահատված մոդելի վավերականության ստուգման քայլերը:

Խնդիրներ

1. Տրված են. ա) ստացիոնար ϕ Շ ամսական տվյալների ավտոկորելյացիայի ու մասնակի ավտոկորելյացիայի ֆունկցիաների գրաֆիկները: Որոշե՛ք գնահատվելիք $ARMA(p, q)$ մոդելի p ու q կարգերը, բ) մեկ այլ ստացիոնար ϕ Շ ամսական տվյալների ավտոկորելյացիայի ու մասնակի ավտոկորելյացիայի ֆունկցիաների գրաֆիկները: Արդյո՞ք այս ϕ Շ-ն սպիտակ աղմուկ ϕ Շ է: Եթե այո, ապա ինչո՞ւ:

ա)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.339	0.339	30.526	0.000
		2 0.103	-0.013	33.377	0.000
		3 -0.029	-0.068	33.603	0.000
		4 -0.064	-0.037	34.713	0.000
		5 -0.066	-0.030	35.904	0.000
		6 -0.153	-0.135	42.274	0.000
		7 -0.082	0.010	44.118	0.000
		8 -0.037	-0.003	44.483	0.000
		9 -0.054	-0.064	45.271	0.000
		10 -0.035	-0.017	45.612	0.000
		11 -0.005	0.010	45.619	0.000
		12 -0.006	-0.035	45.629	0.000
		13 -0.048	-0.062	46.267	0.000
		14 -0.085	-0.064	48.306	0.000
		15 0.028	0.076	48.523	0.000
		16 0.052	0.012	49.291	0.000

բ)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.004	-0.004	0.0053	0.942
		2 -0.005	-0.005	0.0114	0.994
		3 -0.003	-0.003	0.0141	1.000
		4 0.010	0.010	0.0415	1.000
		5 -0.023	-0.023	0.1818	0.999
		6 -0.052	-0.052	0.9049	0.989
		7 0.060	0.059	1.8693	0.967
		8 0.034	0.034	2.1846	0.975
		9 0.105	0.106	5.1527	0.821
		10 0.041	0.044	5.6107	0.847
		11 0.008	0.007	5.6284	0.897
		12 -0.134	-0.136	10.542	0.569
		13 0.050	0.055	11.244	0.590
		14 0.054	0.059	12.051	0.602
		15 -0.110	-0.103	15.408	0.422
		16 -0.029	-0.038	15.635	0.479

2. ՀՀ վիճակագրական կոմիտեի պաշտոնական կայքէջից ներբեռնելով ՀՀ կենտրոնական կառավարության արտաքին պարտքի ժամանակային շարքի տվյալները 2013-2023 թթ. համար՝ եռամսյակային կտրվածքով (հասցեն՝ <https://armstat.am/file/doc/99542473.xlsx>): Տեղափոխելով տվյալները «EViews» փաթեթ: Ուսումնասիրելով դրա գրաֆիկը, անհրաժեշտության պարագայում ապաստեզոնայնացրելով այն: Ուսումնասիրելով (ապաստեզոնայնացված) շարքի ավտոկոռելյացիայի ֆունկցիայի գրաֆիկը, իրականացրելով Դիկի-Ֆուլերի ստացիոնարության թեստը, անհրաժեշտության դեպքում կիրառելով տարբերական օպերատորը և ուսումնասիրելով տարբերությամբ վերցված շարքի կոռելյոգրամն ու ստացիոնարությունը: Եկելով եզրակացության շարքի ստացիոնարության, դետերմինացված և ստոխաստիկ տրենդների առկայության ու շարքի ինտեգրացման կարգի մասին: Ուսումնասիրելով ստացիոնարացված շարքի կոռելյոգրամը և Բոքս-Ջենքինսի մեթոդաբանության հիման վրա որոշելով $ARIMA(p,d,q)$ մոդելի կարգերը: Գնահատելով մոդելը, ստուգելով դրա որակն ու վավերականությունը և կանխատեսելով ապագա 3 եռամսյակների արժեքները: Կատարելով վերջնական եզրակացություն մոդելի կիրառելիության մասին:

3. ՀՀ կենտրոնական բանկի պաշտոնական կայքէջից ներբեռնելով փողի ագրեգատի՝ M2-ի ժամանակային շարքը 2013-2025 թթ. համար՝ ամսական կտրվածքով (հասցեն՝ <https://www.cba.am/hy/statistics/monetary-aggregates/50>): Տեղափոխելով տվյալները EViews փաթեթ: Ուսումնասիրելով դրա գրաֆիկը, անհրաժեշտության պարագայում ապաստեզոնայնացրելով այն: Ուսումնասիրելով (ապաստեզոնայնացված) շարքի ավտոկոռելյացիայի ֆունկցիայի գրաֆիկը, իրականացրելով Դիկի-Ֆուլերի ստացիոնարության թեստը, անհրաժեշտության դեպքում կիրառելով տարբերական օպերատորը և ուսումնասիրելով տարբերությամբ վերցված շարքի կոռելյոգրամն ու ստացիոնարությունը: Եկելով եզրակացության շարքի ստացիոնարության, դետերմինացված և ստոխաստիկ տրենդների առկայության ու շարքի ինտեգրացման կարգի մասին: Ուսումնասիրելով ստացիոնարացված շարքի կոռելյոգրամը և Բոքս-Ջենքինսի մեթոդաբանության հիման վրա որոշելով $ARIMA(p,d,q)$ մոդելի կարգերը: Գնահատելով մոդելը, ստուգելով դրա որակն ու վավերականությունը և կանխատեսելով

ապագա 6 ամիսների արժեքները: Կատարենք վերջնական եզրակացություն մոդելի կիրառելիության մասին:

4. ՀՀ կենտրոնական բանկի պաշտոնական կայքէջից ներբեռնենք ժամկետային ավանդների ժամանակային շարքը 2013-2025 թթ. համար՝ ամսական կտրվածքով (հասցեն՝ <https://www.cba.am/hy/statistics/monetary-aggregates/50>): Տեղափոխենք տվյալները «EViews» փաթեթ: Ուսումնասիրենք դրա գրաֆիկը, անհրաժեշտության պարագայում ապաստեզոնայնացրենք այն: Ուսումնասիրենք (ապաստեզոնայնացված) շարքի ավտոկոռելյացիայի ֆունկցիայի գրաֆիկը, իրականացրենք Դիկի-Ֆուլերի ստացիոնարության թեստը, անհրաժեշտության դեպքում կիրառենք տարբերական օպերատորը և ուսումնասիրենք տարբերությամբ վերցված շարքի կոռելոգրամն ու ստացիոնարությունը: Կատարենք եզրակացություն շարքի ստացիոնարության, դետերմինացված և ստոխաստիկ տրենդների առկայության ու շարքի ինտեգրացման կարգի մասին: Ուսումնասիրենք ստացիոնարացված շարքի կոռելոգրամը և Բոքս-Ջենքինսի մեթոդաբանության հիման վրա որոշենք $ARIMA(p,d,q)$ մոդելի կարգերը: Գնահատենք մոդելը, ստուգենք դրա որակը, վավերականությունը և կանխատեսենք ապագա 6 ամիսների արժեքները: Կատարենք վերջնական եզրակացություն մոդելի կիրառելիության մասին:

ԳԼՈՒԽ 8. ԲԱԶՄԱԶԱՓ ԺԱՄԱՆԱԿԱՅԻՆ ՇԱՐՔԵՐԻ ՄՈԴԵԼԱՎՈՐՄԱՆ ՀԻՄՆԱՀԱՐՑԵՐՆ ՈՒ ՈՐՈՇ ՄՈԴԵԼՆԵՐ

1. Ժամանակային շարքերի կեղծ ռեգրեսիա

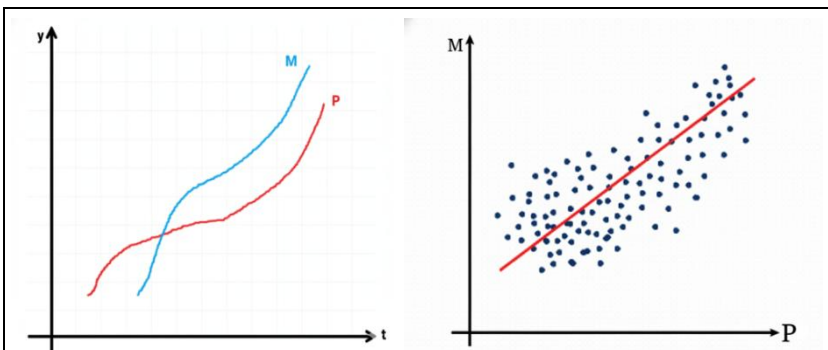
Ենթադրենք ունենք երկու կամ ավելի ԺՇ-եր և ուզում ենք գնահատել դրանցից մեկի՝ ժամանակի յուրաքանչյուր t պահի արժեքի գծային կախումը մյուսների՝ նույն t պահի արժեքից: Այդ կախումը գծային ռեգրեսիայի մոդելի տեսքով ներկայացնելու դեպքում կունենանք՝

$$y_t = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \dots + \alpha_m x_{m,t} + \varepsilon_t: \quad (186)$$

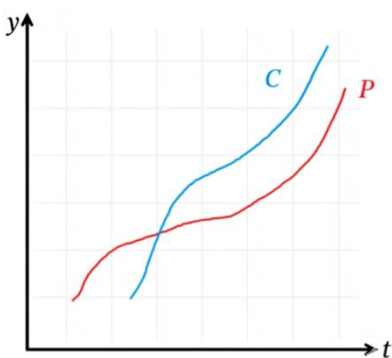
(186)-ը ստատիկ մոդել է, քանի որ կապում է ԺՇ-երի՝ ժամանակի միևնույն պահի արժեքները: Բազմաչափ ԺՇ-երի ռեգրեսիոն մոդելավորման ժամանակ կարևորվում է դրանց ստացիոնարությունը: Եթե (186) մոդելի մեջ մտցված ԺՇ-երը ստացիոնար չեն, ապա առաջանում է մի երևույթ, որը կոչվում է **կեղծ ռեգրեսիա**: Կեղծ ռեգրեսիան վիճակագրական երևույթ է, որն արտահայտվում է նրանով, որ իրարից իրականում կախում չունեցող ԺՇ-երի միջև ի հայտ է գալիս առերևույթ բարձր աստիճանի ռեգրեսիոն կախվածություն: Կեղծ ռեգրեսիան առաջանում է այն պատճառով, որ ռեգրեսիայի հավասարման մեջ մտցված ԺՇ-երը ստացիոնար չեն և ունեն կամ դետերմինացված, կամ ստոխաստիկ միտում: Բուն երկու ԺՇ-երը մեկը մյուսից կախվածություն չունեն, սակայն դրանց միտումները՝ ժամանակի մեջ փոփոխման օրինաչափությունները, ունեն կախվածություն, որն էլ առաջացնում է կեղծ ռեգրեսիա:

Ներդիր 18. Կեղծ ռեգրեսիայի երևույթի հայտնաբերման պայմանությունը

1980 թ. Դեյվիդ Հենդրին գնահատեց այսպիսի ստատիկ մոդել երկու ժամանակային շարքերի միջև՝ զների մակարդակ (P_t) ու փողի առաջարկ (M_t): Այդ շարքերը ժամանակի մեջ ունեին աճման միտում, տրենդ: Այս երկու ԺՇ-երի կոռելյացիայի դաշտը փաստում էր, որ ԺՇ-երն ուժեղ կոռելյացիոն կապի մեջ են, իսկ $\hat{P}_t = \alpha_0 + \alpha_1 M_t$ գծային ռեգրեսիայի մոդելի գնահատման արդյունքում ստացված դետերմինացիայի գործակիցը մոտ էր մեկին՝ 0,995: Գնահատված գործակիցներն ու հավասարումը վիճակագրորեն նշանակալի էին:



Այնուհետև Հենդրին գների մակարդակի հետ միասին դիտարկեց դրա հետ բոլորովին ոչ մի կապ չունեցող, սակայն ժամանակի մեջ փողի առաջարկի վարքին շատ մնան վարքով մի այլ մեծություն՝ տեղումների մակարդակը (C_t): Գների մակարդակի՝ տեղումների մակարդակից կախվածությունը գնահատող $\hat{P}_t = \alpha_0 + \alpha_1 C_t$ մոդելի գնահատման արդյունքում ստացված դետերմինացիայի



գործակիցը նույնիսկ ավելի մեծ ստացվեց՝ 0,998, իսկ գնահատված գործակիցներն ու հավասարումը վիճակագրորեն նշանակալի էին: Ստացվեց, որ տեղումների քանակը բացատրում է գների մակարդակը: Բնականաբար, գնահատված ռեգրեսիայի հավասարումն ու այս կախումը կեղծ էին: Կեղծ ռեգրեսիայի պատճառն այն էր, որ գծային ռեգրեսիայի ստատիկ

մոդելի մեջ մտցված գների մակարդակի ու տեղումների մակարդակի ԺՇ-երը ստացիոնար չէին:

Գրեյնջերը նույնիսկ ստացել է կեղծ ռեգրեսիայի ստուգման ձև: Եթե ԺՇ-երի միջև գնահատվում է գծային ռեգրեսիայի մոդել, և ՓԶԵ-ով գնահատված մոդելի **դետերմինացիայի գործակիցը մեծ է Դարբին Ուոթսոնի վիճականուց՝ $R^2 > DW$, ապա գնահատված ռեգրեսիան կեղծ է:** Այս պայմանը նշանակում է, որ Դարբին Ուոթսոնի վիճականին մոտ է 0-ին, այսինքն՝ առկա է դրական առաջին

կարգի ավտոկոռելյացիա մոդելի պատահական մնացորդների միջև: Ավտոկոռելյացիայի հետևանքով գնահատված պարամետրերի t վիճականիներն ունենում են շատ բարձր արժեքներ, որոնք փաստում են, որ բացատրող փոփոխականը նշանակալի ազդեցություն ունի կախվածի վրա, սակայն դա իրականությանը չի համապատասխանում: Այսինքն՝ t վիճականիները գերազնահատված են լինում: Սա կախվածության կեղծ պատկեր է ստեղծում:

2. Ժամանակային շարքերի կոինտեգրացիան ու դրա ստուգումը

Կեղծ ռեգրեսիայից խուսափելու համար մի քանի միաչափ ԺՇ-երի կախվածությունը բնութագրող գծային ռեգրեսիայի մոդել կարելի է գնահատել եթե՝

1. մոդելում ներառվող ԺՇ-երը ստացիոնար կամ ստացիոնարացված են,
2. մոդելում ներառվող ԺՇ-երը ստացիոնար կամ ստացիոնարացված չեն, բայց **կոինտեգրացվող են:**

Սահմանում: Երկու՝ X_t և Y_t պատահական գործընթացները կամ դրանց իրականացումը հանդիսացող ԺՇ-երը կոինտեգրացվող են, եթե՝

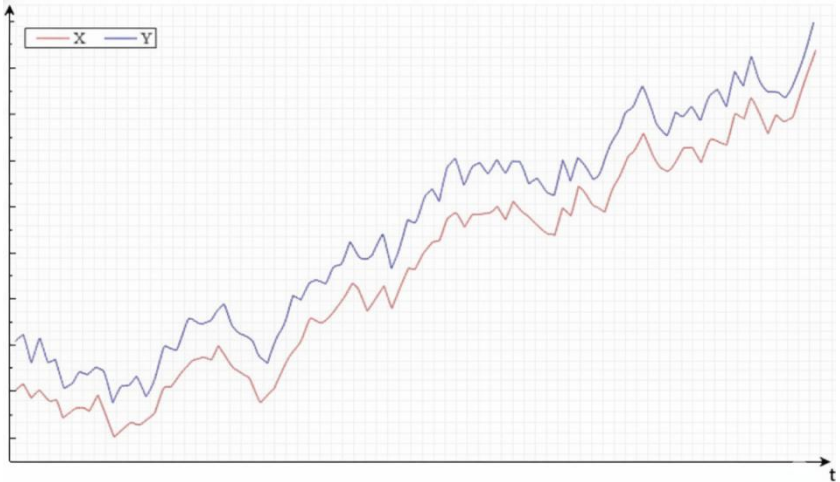
1. դրանք երկուսն էլ միևնույն՝ առաջին կարգի ինտեգրացվող են, այսինքն՝ $Y_t \sim I(1)$ և $X_t \sim I(1)$,
2. գոյություն ունի (առնվազն) մեկ հաստատուն՝ β , այնպիսին, որ $Y_t - \beta X_t$ գործընթացը ստացիոնար է, այսինքն՝ $Y_t - \beta X_t \sim I(0)$, որում β գործակիցը կոչվում է կոինտեգրացման գործակից:

Այս սահմանումը նշանակում է, որ կոինտեգրացվող ԺՇ-երը ստացիոնար չեն, առաջին կարգի ինտեգրացվող են, սակայն գոյություն ունի դրանց գծային կոմբինացիա, որը ստացիոնար է:

Երկու ԺՇ-երի կոինտեգրացվածությունը նշանակում է, որ նրանց միջև կա երկարաժամկետ հավասարակշռություն, դրանք երկարաժամկետ կայուն կապի, կախվածության մեջ են և ժամանակի մեջ «միասին են շարժվում», իսկ այդ երկարաժամկետ կայուն կախվածությունից շեղումը ժամանակի մեջ ստաբիլ է, այսինքն ստացիոնար է: Սա ԺՇ-երի համատեղ գրաֆիկական պատկերի մեջ նշանա-

կում է, որ երկու ԺՇ-երի միջև եղած «հեռավորությունը» ժամանակի մեջ հաստատուն է, չի փոխվում:

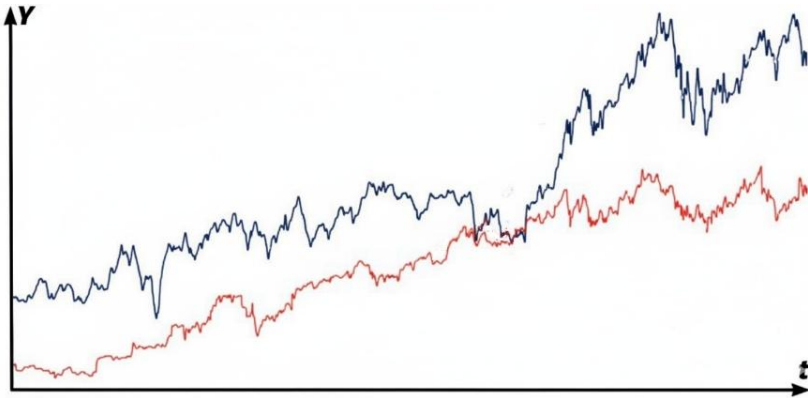
Գծանկար 31. Երկու կոինտեգրացվող ոչ սրացիոնար ժամանակային շարքերի վարքը ժամանակի մեջ



Գծանկար 31-ում բերված են երկու ԺՇ-եր, որոնք կոինտեգրվող են. դրանք երկուսն էլ ստացիոնար չեն, որովհետև աճման միտում ունեն, և դրանց միջև եղած հեռավորությունը ժամանակի մեջ չի փոխվում:

Գծանկար 32-ում բերված ԺՇ-երի միջև հեռավորությունը փոխվում է ժամանակի մեջ, և, ի վերջո, դրանք գնալով հեռանում են իրարից: Այս շարքերը կոինտեգրացվող չեն: **Կոինտեգրացվածության բացակայությունը նշանակում է, որ երկարաժամկետ հեռանկարում շարքերն իրարից հեռանում են:**

Գծանկար 32. Երկու չկոինտեգրացվող ոչ սրացիոնար ժամանակային շարքերի վարքը ժամանակի մեջ



Ինչպե՞ս կարելի է ստուգել, թե արդյոք երկու ԺՇ-եր կոինտեգրացվող են: Ամենատարածված մոտեցումը Էնգել-Գրեյնջերի մոտեցումն է: Դրա էությունը հետևյալն է. ենթադրենք ունենք y_t և x_t ԺՇ-երը, որոնք ստացիոնար չեն, և երկուսն էլ առաջին կարգի ինտեգրացվող են՝ $y_t \sim I(1)$ և $x_t \sim I(1)$:

Գնահատվում է այս ԺՇ-երի միջև կախվածության հետևյալ գծային ռեգրեսիայի հավասարումը՝

$$y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t: \quad (187)$$

Այստեղից՝

$$\varepsilon_t = y_t - \alpha_0 - \alpha_1 x_t: \quad (188)$$

Ենթադրենք՝ ՓԲԵ-ով գնահատվել է (187)-ը, և ստացվել են $\hat{\alpha}$ -երը և $\hat{\varepsilon}_t$ -երը: Ինչպես երևում է (188)-ից, $\hat{\varepsilon}_t$ -ը y_t և x_t ԺՇ-երի գծային կոմբինացիա է: Ուրեմն կարող ենք ստուգել, թե արդյոք այդ գծային կոմբինացիան ստացիոնար է, այսինքն՝ $\hat{\varepsilon}_t \sim I(0)$: Եթե այո, ապա y_t ու x_t ԺՇ-երը կոինտեգրացվող են:

ԺՇ-երի կոինտեգրացվածության Էնգել-Գրեյնջերի մոտեցման քայլերը հետևյալն են.

1. մախ ստուգվում է՝ շարքերը ստացիոնար են, թե ոչ,
2. եթե այո, դրանք կարելի է մոդելավորել (187)-ով և ՓԲԵ-ով գնահատել մոդելը: Եթե ոչ, ստուգվում է՝ արդյո՞ք $y_t \sim I(1)$ և $x_t \sim I(1)$,
3. եթե այո, ՓԲԵ-ով գնահատվում է (187) մոդելը, և գտնվում են $\hat{\varepsilon}_t$ -երը,

4. ստուգվում է $\hat{\varepsilon}_t$ -երի ստացիոնարությունը,

5. եթե $\hat{\varepsilon}_t$ -ն ստացիոնար է, շարքերը կոինտեգրացվող են:

$\hat{\varepsilon}_t$ -երի ստացիոնարությունը կարելի է ստուգել Դիկի-Ֆոլերի թեստի միջոցով: Սակայն այս թեստի վիճակահին տվյալ դեպքում օգտագործելի կլինի, իսկ 1, 5 ու 10 % արժեքականության դեպքում ստացվող կրիտիկական արժեքները օգտագործելի չեն լինի: Հետևաբար $\hat{\varepsilon}_t$ -երի ստացիոնարությունը ստուգել հնարավոր չի լինի: Այդ պատճառով $\hat{\varepsilon}_t$ -երի ստացիոնարության ստուգման համար կիրառվում են Էնգել-Գրեյնջերի պարամետրական ու Ֆիլիպս-Օուլիարիսի ոչ պարամետրական թեստերը: Դրանք իրականացվում են կոինտեգրացվող ռեգրեսիայի գնահատման պայմանով, իսկ դրվող գրոյական վարկածն է՝ ԺԸ-երը կոինտեգրացվող չեն: Եթե թեստերի վիճակահանք արժեքը մեծ է 0.05-ից, ապա այն ընդունվում է, և ԺԸ-երը կոինտեգրացվող չեն: Հակառակ դեպքում կոինտեգրացվող են: Նշենք, որ այս թեստերը պահանջում են, որ՝ $y_t \sim I(1)$ և $x_t \sim I(1)$: ԺԸ-երի կոինտեգրացվածության նշված թեստերն առկա են, օրինակ, E«views» ծրագրային փաթեթում:

Այն դեպքում, երբ դիտարկվում է երկու ԺԸ՝ y_t և x_t , կարող է լինել y_t -ի և x_t -ի առավելագույնը միայն մեկ գծային կոմբինացիա, որը ստացիոնար է, այսինքն՝ առավելագույնը մեկ կոինտեգրացիոն հարաբերություն:

Ենթադրենք, որ գնահատվող հավասարումը ներառում է k թվով փոփոխականներ՝ կախյալ փոփոխականը՝ y_t -ն և $k-1$ բացատրող փոփոխականներ՝ $x_{1,t}, x_{2,t} \dots x_{k-1,t}$: Այս դեպքում կարող են գոյություն ունենալ մինչև r գծային անկախ կոինտեգրացիոն հարաբերություններ ($r \leq k - 1$): Սա պոտենցիալ խնդիր է ներկայացնում վերը նկարագրված ՓԶԵ-ով գնահատվող ռեգրեսիայի ու դրա մնացորդների ստացիոնարության ստուգման մոտեցման համար, որն ի վիճակի է գտնել առավելագույնը մեկ կոինտեգրացնող հարաբերություն՝ անկախ նրանից, թե քանի փոփոխական կա հավասարման մեջ: Սակայն տվյալ դեպքում, եթե կան բազմաթիվ կոինտեգրացնող հարաբերություններ, ինչպես՝ կարելի է իմանալ, թե բացի ռեգրեսիայի մոդելի մնացորդներից՝ արդյոք կան նաև այլ կոինտեգրացնող հարաբերություններ, կամ արդյո՞ք գտնվել է «լավագույն» կամ ամենաուժեղ կոինտեգրացնող հարաբերությունը: Այս խնդրի պատասխանը համակարգային մոտեցման օգտագործումն է կոինտեգրման թեստավորման համար, որը թույլ կտա որոշել բոլոր կոինտեգրացնող հարա-

բերությունները: Նման մոտեցումներից մեկը **Յոհանսենի մեթոդն է**: Այն ևս առկա է «EViews» ծրագրային փաթեթում:

3. Կոինտեգրացվածությունն ու սխալներով ճշգրտման մեխանիզմը

Երբ 1970-ական թթ. գիտական շրջանակները սկսեցին զբաղվել ոչ ստացիոնար ժամանակային շարքերի մոդելավորման հիմնահարցերով, առաջին արձագանքն այն էր, որ ոչ ստացիոնար ԺՇ-երի գծային կախվածությունը մոդելավորելու համար ուղղակի պետք է ստացիոնարացնել ԺՇ-երը՝ վերցնելով դրանց առաջին տարբերությունները, այնուհետև օգտագործել այդ առաջին տարբերությունները հետագա մոդելավորման գործընթացում, այսինքն՝ գնահատել, օրինակ, այս տեսքի մոդել՝

$$\Delta(y_t) = \alpha_0 + \alpha_1 \Delta(x_t) + \varepsilon_t: \quad (189)$$

Ինչպես արդեն գիտենք, միաչափ ԺՇ-երի մոդելավորման պարագայում դա լիովին արդարացված և բացարձակապես ճիշտ մոտեցում է: Սակայն երբ ունենք մի քանի ԺՇ-եր և մոդելավորում ենք դրանց փոխադարձ կապը ժամանակի մեջ, ապա նման ընթացակարգը նպատակահարմար չէ, քանի որ, մի կողմից, ստանում ենք ոչ թե բուն երևույթների, այլ դրանց աճերի միջև կապը, իսկ մյուս կողմից՝ տարբերությամբ վերցված բազմաչափ ԺՇ-երի մոդելները փոփոխականների երկարաժամկետ կապը չեն կարողանում արտացոլել: Սակայն եթե շարքերը կոինտեգրացվող են, դրանք հենց երկարաժամկետ կապի կամ հավասարակշռության մեջ են: Ուրեմն կոինտեգրացվող ԺՇ-երի միջև կախվածությունը գնահատելիս նպատակահարմար է գնահատել այնպիսի մոդել, որը հնարավորություն կտա գնահատելու դրանց միջև ինչպես կարճաժամկետ, այնպես էլ երկարաժամկետ կախվածությունները:

Այս նկատառումով մշակվեց մոդելավորման այնպիսի մոտեցում, որը հաշվի է առնում ԺՇ-երի կոինտեգրացվածությունը: Այդպիսի մոդել է սխալների (կամ հավասարակշռության) ճշգրտման մոդելը (այսուհետ՝ ECM, error correction model): Գրեյնջերի թեորեմը պնդում է, որ եթե երկու փոփոխականներ կոինտեգրացվող են, ապա դրանց միջև կախվածությունը կարելի է արտահայտել **սխալների ճշգրտման մոդելով**: Եթե ունենք երկու ԺՇ՝ y_t և x_t , ապա այդ մոդելի պարզագույն տեսքն է՝

$$\Delta(y_t) = \beta_0 + \beta_1 \cdot \Delta(x_t) + \beta_2 \cdot (y_{t-1} - \alpha_0 - \gamma \cdot x_{t-1}) + u_t: \quad (190)$$

Այսինքն՝ ԺՇ-երի ոչ ստացիոնարության ու կոինտեգրացման պարագայում (189) մոդելի փոխարեն նպատակահարմար է օգտագործել (190)-ը:

Եթե y_t և x_t փոփոխականները երկուսն էլ առաջին կարգի ինտեգրացվող են և կոինտեգրացվող են՝ γ կոինտեգրացման գործակցով, ապա $y_{t-1} - \gamma \cdot x_{t-1}$ -ը ստացիոնար է, $I(0)$ է:

(190) մոդելում $y_{t-1} - \gamma \cdot x_{t-1}$ -ը y_t և x_t փոփոխականների միջև գնահատված $y_t = \alpha_0 + \gamma x_t + \varepsilon_t$ ռեգրեսիայի մոդելի մնացորդն է ժամանակի նախորդ պահի համար:

$$\varepsilon_{t-1} = y_{t-1} - \alpha_0 - \gamma x_{t-1}: \quad (191)$$

Հիշենք, որ եթե y_t և x_t փոփոխականները կոինտեգրացվող են, ապա դրանք երկարաժամկետ կայուն կապի մեջ են, որը բնութագրվում է հետևյալ հավասարմամբ՝

$$y_t = \alpha_0 + \gamma x_t + \varepsilon_t, \quad (192)$$

իսկ ε_{t-1} -ը բնութագրում է y_t և x_t փոփոխականների երկարաժամկետ կայուն կախվածությունից, երկարաժամկետ հավասարակշռությունից շեղումը ժամանակի $t-1$ պահին:

Այսպիսով՝ (190)-ը կարելի է ներկայացնել հետևյալ տեսքով՝

$$\Delta(y_t) = \beta_0 + \beta_1 \cdot \Delta(x_t) + \beta_2 \cdot \varepsilon_{t-1} + u_t: \quad (193)$$

(193) մոդելում կախյալ փոփոխականի աճը կախման մեջ է դրվում ոչ միայն բացատրող փոփոխականի աճից (երկուսն էլ տարբերությամբ վերցված, ստացիոնարացված են), այլև $y_t = \alpha_0 + \gamma x_t + \varepsilon_t$ (191) տեսքի մոդելի՝ ժամանակի նախորդ պահի մնացորդից՝ ε_{t-1} -ից, որը բնութագրում է y_t և x_t **փոփոխականների երկարաժամկետ կայուն փոխկապվածությունից, երկարաժամկետ հավասարակշռությունից շեղումը ժամանակի $t-1$ պահին:**

ε_{t-1} -ը կոչվում է սխալների ճշգրտման անդամ, քանի որ այն ճշգրտում, ուղղում է y_t -ի շեղումը իր երկարաժամկետ հավասարակշռային արժեքից, որը որոշվում է կոինտեգրացիայի (191) հավասարմամբ: Իսկ β_2 գործակիցը կոչվում է սխալների ճշգրտման գործակից: Եթե ԺՇ-երը կոինտեգրացվող են, ապա β_2 -ը կամ սխալների ճշգրտման գործակիցը պետք է բացասական լինի: Այն ցույց է տա-

լիս այն արագությունը, որով մոդելի կախյալ փոփոխականը վերադառնում է իր հավասարակշռության վիճակին: Եթե այն բացասական է, նշանակում է, որ հավասարակշռությունից ցանկացած շեղում ժամանակի ընթացքում կուղղվի, կճշգրտվի, քանի որ համակարգը պետք է վերադառնա հավասարակշռությանը, ինչն էլ կոինտեգրացվածության էությունն է: Դրական գործակիցը ցույց կտա անկայունություն, քանի որ այդ դեպքում հավասարակշռությունից շեղումները պարզապես անընդհատ կավելանան, ինչը կնշանակի, որ շարքերը կոինտեգրացվող չեն:

Եթե սխալների ճշգրտման անդամը՝ ε_{t-1} -ը, հավասար է զրոյի, ապա երկու փոփոխականների միջև անհավասարակշռություն չի առաջանում, և այդ դեպքում երկարաժամկետ հարաբերությունը կտրվի կոինտեգրացիայի հավասարմամբ՝ $y_t = \alpha_0 + \gamma x_t$: Եթե վերջինս զրոյից տարբեր է, ուրեմն երկու փոփոխականների միջև կախվածությունը իր հավասարակշռային վիճակից շեղվել է, դուրս է եկել:

Մխալների ուղղման (190) մոդելները մեկնաբանվում են հետևյալ կերպ: Ենթադրվում է, որ y -ը ժամանակի նախորդ ու հաջորդ պահերի՝ $t-1$ -ի ու t -ի միջև փոխվում է $t-1$ -ի ու t -ի միջև x բացատրող փոփոխականի արժեքների փոփոխության արդյունքում, և նախորդ ժամանակաշրջանում գոյություն ունեցող ցանկացած անհավասարակշռություն մասամբ ճշգրտվում է: β_1 -ը բնութագրում է y_t և x_t փոփոխականների աճերի կարճաժամկետ կախվածությունը՝ $\Delta(x_t)$ -ի ազդեցությունը նույն պահի $\Delta(y_t)$ -ի վրա: Այսինքն՝ β_1 -ը նկարագրում է x -ի և y -ի կարճաժամկետ կապը: β_2 -ը ցույց է տալիս այն արագությունը, որով մոդելի կախյալ փոփոխականը վերադառնում է իր հավասարակշռության վիճակին վերը նշված փոփոխությունից հետո: γ -ը բնութագրում է x -ի և y -ի միջև երկարաժամկետ կապը:

Կեղծ ռեգրեսիայի, կոինտեգրացիայի ու սխալների ճշգրտման մեխանիզմի գաղափարներն ու տեխնիկաներն օգտագործվում են բազմաչափ ԺՇ-երի էկոնոմետրիկ մոդելավորման ամենատարբեր մոտեցումներում, մասնավորապես՝ բազմաչափ ԺՇ-երի **դինամիկ** մոդելներում:

**4. Ժամանակային շարքերի դինամիկ մոդելներ՝ բաշխված լագերի
DL(p) և ավտոռեգրեսիայի ու բաշխված լագերի ARDL(p,q)
մոդելների համառոտ նկարագրություն**

Պարզ է, որ տնտեսությունում որևէ փոփոխականի տվյալ պահի արժեքի վրա ազդեցություն ունեն այլ փոփոխականների ինչպես ընթացիկ, տվյալ պահի արժեքները, այնպես էլ նախորդ պահերի արժեքները: Այսինքն՝ կախյալ փոփոխականի տվյալ պահի արժեքն իր վրա կրում է նաև ուշացած ազդեցություններ: Ինչպես արդեն ասվել է, այդ «ուշացումը» կոչվում է լագ: Բացատրող փոփոխականների արժեքները ժամանակի նախորդ պահերին կոչվում են դրանց լագային արժեքներ կամ լագային փոփոխականներ: Միևնույն ժամանակ արդեն դիտարկվել է, որ կախյալ փոփոխականի տվյալ պահի արժեքը ձևավորվում է նաև իր իսկ նախորդ վարքի կամ իր իսկ նախորդ պահերի արժեքների ազդեցության տակ: Այս բոլոր ազդեցությունները պայմանավորում են էկոնոմետրիկ մոդելների դինամիկ հատկորոշման նպատակահարմարությունը: ARIMA մոդելները ևս դինամիկ մոդելների տարատեսակ են, սակայն միաչափ ԺՇ-երի համար: ԺՇ-երի էկոնոմետրիկ մոդելը կոչվում է դինամիկ, եթե այն ընդգրկում է ԺՇ-ի ինչպես ժամանակի տվյալ t պահի, այնպես էլ նախորդ պահերի արժեքները:

Դինամիկ էկոնոմետրիկ մոդելները կարելի է բաժանել երկու հիմնական տիպի՝

- I. մոդելներ, որոնցում բացատրող փոփոխականների լագային արժեքները անմիջականորեն մտցված են մոդելի մեջ, և դրանք բացատրում են կախյալ փոփոխականի ընթացիկ t պահի արժեքը: Այդպիսիք են, օրինակ, ARIMA մոդելները՝ միաչափ ժամանակային շարքերի համար, բաշխված լագերի մոդելները և ավտոռեգրեսիայի ու բաշխված լագերի ARDL մոդելները, սխալների ճշգրտման, վեկտորական ավտոռեգրեսիայի մոդելները՝ բազմաչափ ԺՇ-երի համար,
- II. մոդելներ, որոնցում դինամիկ ազդեցությունները լատենտ, չերևացող են: Այս մոդելները ներառում են փոփոխականներ, որոնք բնութագրում են կախյալ փոփոխականի անկնկալվող կամ ցանկալի արժեքը ժամանակի t պահին: Այդպիսիք են, օրինակ, մասնակի ճշգրտման, ադապտիվ սպասումների, ռացիոնալ սպասումների մոդելները:

ԺՇ-երի միջև դինամիկ գծային ռեգրեսիայի մոդելների մեծ մասը կարելի է գնահատել ՓԶԵ-ով որոշակի պայմանների իրականացման դեպքում: Ամեն դեպքում **ԺՇ-երի դինամիկ մոդելների գնահատման համար անհրաժեշտ է, որ մոդելի մեջ մտցվող ԺՇ-երը լինեն ստացիոնար կամ ստացիոնարացված, եթե դրանք կոինտեգրացվող չեն:**

Այստեղ կոինտարկենք բաշխված լագերի մոդելները: Դրանք դիտարկում են կախյալ փոփոխականի (ԺՇ-ի) t պահի արժեքի գծային ռեգրեսիոն կախվածությունը բացատրող փոփոխականների ընթացիկ t պահի ու լագային արժեքներից: Այսպիսի մոդելները հաշվի են առնում այն իրավիճակները, երբ կախյալ փոփոխականի՝ որևէ տնտեսական ցուցանիշի ընթացիկ t պահի արժեքի վրա ազդում են այլ ցուցանիշի կամ ցուցանիշների ինչպես ընթացիկ t պահի, այնպես էլ նախորդ պահերի արժեքները: Այսինքն՝ եթե բացատրող փոփոխականի արժեքը ժամանակի ինչ-որ t պահին փոխվի, ապա այդ փոփոխությունը կազդի կախված փոփոխականի հաջորդ արժեքների վրա:

Օրինակ՝ որևէ ընկերության հասույթը ժամանակի t պահին կախված է ընկերության գովազդային կամ մարքեթինգային ծախսերի մեծությունից, որոնք ընկերությունն արել է նախկինում՝ ժամանակի $t-1$, $t-2$... պահերին: Կամ երկրի ՀՆԱ-ն ժամանակի t պահին կախված է ներդրումներից, որոնք կատարվել են ինչպես ժամանակի նույն պահին, այնպես էլ դրանից առաջ՝ ժամանակի նախորդ պահերին: Ստացվում է, որ ժամանակի ինչ-որ t պահին կատարված ներդրումների ազդեցությունը ՀՆԱ-ի վրա «բաշխվում» է տվյալ t պահի, հաջորդ $t+1$, $t+2$... պահերի ՀՆԱ-ի արժեքների վրա որոշակի ուշացմամբ կամ լագով: Դրա համար այսպիսի ազդեցություններ բնութագրող էկոնոմետրիկ մոդելները կոչվում են **բաշխված լագերի մոդելներ (DL, distributed lags):**

DL մոդելները կարող են պարունակել բացատրող փոփոխականների վերջավոր կամ անվերջ թվով լագեր: Վերջավոր թվով լագեր պարունակող մոդելները կոչվում են **վերջավոր բաշխված լագերի մոդելներ, իսկ դրանց կարգը որոշվում է անկախ փոփոխականների լագերի քանակով:** Այստեղ կանոնադաշտանք վերջավոր մոդելներին: p -րդ կարգի DL(p) մոդելը մեկ բացատրող փոփոխականի դեպքում ունի հետևյալ տեսքը՝

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_p x_{t-p} + \varepsilon_t, \quad (194)$$

որում՝ $t=1, 2, \dots, n$, ε_t -ն ստացիոնար, սպիտակ աղմուկ σ^2 է:

Այս մոդելում առկա են անկախ փոփոխականի ինչպես t պահի, այնպես էլ նախորդ p պահերի արժեքները: Այս մոդելը արտացոլում է այնպիսի կախվածություն, որն ապահովում է, որ երբ բացատրող փոփոխականը՝ x_t -ը, ժամանակի t պահին փոփոխվի, դա կունենա ազդեցություն y կախյալ փոփոխականի տվյալ t պահի ու հաջորդ p հատ՝ $t+1$, $t+2$, $t+p$ պահերի արժեքների վրա: Կամ, այլ ձևով ասած, բացատրող փոփոխականի t - p պահի փոփոխությունը ազդում է կախյալ փոփոխականի բոլոր հաջորդ $t-p+1$, $t-p+2$... $t-1$, և t ՝ վերջավոր թվով պահերի արժեքների վրա: Այսինքն՝ x փոփոխականի ազդեցությունը y -ի վրա բաշխվում է, պահպանվում է ժամանակի վերջավոր թվով p պահերի ընթացքում:

Այս մոդելում բացատրող փոփոխականի t պահի արժեքի՝ x_t -ի գործակիցը՝ β_0 -ն, կոչվում է **կարճաժամկետ մուլտիպլիկատոր** (short-run) և ցույց է տալիս, թե միջինում քանի միավորով կփոխվի y կախյալ փոփոխականի t պահի արժեքը, եթե բացատրող փոփոխականը ժամանակի նույն t պահին փոխվի մեկ միավորով, եթե հաշվի չառնենք վերջինիս «ուշացած», լագային արժեքների ազդեցությունը y_t -ի վրա:

$\beta_0 + \beta_1$ -ը ցույց կտա բացատրող x փոփոխականի մեկ նախորդող $t-1$ ու ընթացիկ t պահերի ընդհանուր, գումարային ազդեցությունը y կախված փոփոխականի t պահի արժեքի վրա և կբնութագրի, թե միջինում քանի միավորով կփոխվի y -ը t պահին, եթե x -ը ժամանակի $t-1$ պահին ու t պահին փոխվի մեկ միավորով, եթե հաշվի չառնենք վերջինիս ավելի ուշացած լագային արժեքների ազդեցությունը: $\beta_0 + \beta_1$ -ը կոչվում է **միջնաժամկետ մուլտիպլիկատոր**: Նույն տրամաբանությամբ միջանկյալ մուլտիպլիկատորներ են մաս $\beta_0 + \beta_1 + \beta_2$ -ը, $\beta_0 + \beta_1 + \beta_2 + \beta_3$ -ը և այլն:

Երկարաժամկետ հեռանկարում ժամանակի p պահ հետո y -ի t պահի արժեքի վրա ազդում են x -ի t , $t-1$, $t-p$ պահերի արժեքները: Ուստի՝

$$\beta = \beta_0 + \beta_1 + \dots + \beta_p = \sum_{j=0}^p \beta_j$$

գումարը ցույց կտա, թե միջինում քանի միավորով կփոխվի y -ը t պահին, եթե x -ը ժամանակի բոլոր նշված պահերին փոխվի մեկ միավորով: β -ն կոչվում է **երկարաժամկետ մոփոխականություն**:

Նշանակենք՝ $b_j = \frac{\beta_j}{\beta}$; $j=1, \dots, p$: b_j -երը կոչվում են (194) մոդելի հարաբերական գործակիցներ: Եթե (194) մոդելի բոլոր β_j ռեգրեսիայի գործակիցներն ունեն նույն նշանը, ապա բոլոր j -երի համար՝

$$0 \leq b_j \leq 1 \text{ և } \sum_{j=0}^p b_j = 1:$$

Այդ դեպքում b_j հարաբերական գործակիցները կշիռներ են՝ β_j ռեգրեսիայի գործակիցների համար: b_j հարաբերական գործակիցները ցույց են տալիս, թե x -ի y -ի վրա ունեցած ընդհանուր երկարաժամկետ ազդեցության որ մասն է ապահովվել x -ի j -րդ պահի ազդեցությամբ:

Եթե b_j հարաբերական գործակիցները հայտնի են, կարելի է հաշվարկել DL մոդելի երկու կարևոր բնութագրիչներ՝ մեդիանային ու միջին լագերի մեծությունները:

Մեդիանային լագը այն ժամանակն է, որն անհրաժեշտ է y կախյալ փոփոխականի՝ x բացատրող փոփոխականի երկարաժամկետ ազդեցությամբ տեղի ունեցող ընդհանուր փոփոխության կեսին կամ 50 տոկոսին հասնելու համար: Այսինքն՝ մեդիանային լագը ժամանակային այն k_{med} լագի մեծությունն է, որի համար՝

$$\sum_{j=0}^{k_{med}} b_j = 0.5:$$

Միջին լագը այն միջին ժամանակն է, որի ընթացքում բացատրող փոփոխականի փոփոխությունը ժամանակի t պահին առաջ է բերում կախյալ փոփոխականի փոփոխություն: Այսինքն, եթե ժամանակի որևէ t պահին բացատրող փոփոխականը փոփոխվում է, որքա՞ն է միջին հաշվով տևում այդ փոփոխության ազդեցությունը կախյալ փոփոխականի վրա: Այն որոշվում է հետևյալ բանաձևով՝

$$\bar{p} = \sum_{j=0}^p p \cdot b_j = \sum_{j=0}^p \frac{p \cdot \beta_j}{\beta}:$$

Բերենք օրինակ: Ենթադրենք ուսումնասիրվում է ձեռնարկության վաճառքների ծավալի կախումը գովազդի վրա կատարված

ծախսերից: Ամսական տվյալների հիման վրա սպեցիֆիկացվել ու գնահատվել է DL(3) մոդելը, որի էմպիրիկ հավասարումն է՝

$$\hat{y}_t = -0.67 + 4.5x_t + 3.0x_{t-1} + 1.5x_{t-2} + 0.5x_{t-3}:$$

Այստեղ **կարճաժամկետ մուլտիպլիկատորը** հավասար է 4.5-ի: Դա նշանակում է, որ ժամանակի որևէ t պահին գովազդի ծախսերի աճը մեկ միավորով հանգեցնում է նույն t պահին նույն ամսում վաճառքների ծավալի աճի՝ միջին հաշվով 4.5 միավորով, եթե հաշվի չառնենք գովազդային ծախսերի «ուշացած», լազային արժեքների ազդեցությունը վաճառքների ծավալի վրա: **Առաջին միջնաժամկետ մուլտիպլիկատորը** հավասար է՝ $4.5+3.0=7.5$: Դա նշանակում է, որ գովազդային ծախսերի աճը մեկ միավորով ժամանակի նախորդ $t-1$ պահին ու ընթացիկ t պահին, ժամանակի ընթացիկ ամսում առաջ կբերի վաճառքների **գումարային** աճ՝ ընդհանուր առմամբ միջին հաշվով 7.5 միավորով: **Երկրորդ միջնաժամկետ մուլտիպլիկատորը** հավասար է՝ $4.5+3.0+1.5=9.0$: Դա նշանակում է, որ գովազդային ծախսերի աճը մեկ միավորով ժամանակի $t-2$, $t-1$ ու t պահերին, ժամանակի t պահին կամ ընթացիկ ամսում առաջ կբերի վաճառքների **գումարային** աճ՝ ընդհանուր առմամբ միջին հաշվով 9 միավորով: **Երկարաժամկետ մուլտիպլիկատորը** հավասար է 9.5: Դա նշանակում է, որ գովազդային ծախսերի աճը մեկ միավորով ժամանակի որևէ t պահին երկարաժամկետ հեռանկարում առաջ կբերի վաճառքների աճ՝ ընդհանուր առմամբ միջին հաշվով 9.5 միավորով: Գնահատված մոդելի հարաբերական գործակիցներն են՝ $b_1=4.5/9.5=0.474$, $b_2=3.0/9.5=0.316$, $b_3=1.5/9.5=0.158$, $b_4=0.5/9.5=0.053$: Սա նշանակում է, որ վաճառքների ընթացիկ ամսվա աճի 47.7 տոկոսը պայմանավորված է գովազդի ծախսերի ընթացիկ պահի/ամսվա աճով, 31.6 տոկոսը՝ նախորդ ամսվա, 15.8 տոկոսը՝ երկու ամիս առաջվա, 5 տոկոսը՝ երեք ամիս առաջվա աճով:

Միջին լազը՝ $\bar{p} = \sum_{j=0}^3 p \cdot b_j = 0 \cdot 0.474 + 1 \cdot 0.316 + 2 \cdot 0.518 + 3 \cdot 0.053 = 0.791$: Միջին լազը փոքր է. այն մեկ ամիս էլ չի կազմում, ինչը հաստատում է այն, որ գովազդի ծախսերի աճը ամենամեծ ազդեցությունը վաճառքների աճի վրա ունենում է ընթացիկ ամսում: **Մեդիանային լազը** k_{med} այն լազն է, որի համար $\sum_{j=0}^{k_{\text{med}}} b_j = 0.5$: Դա մի փոքր մեծ է մեկ ամսից, քանի որ $b_1=0.474$:

Ներդիր 19. Գրեյնջերի պատճառականության թեստը

Բազմաչափ ժամանակային շարքերի դիմամիկ մոդելները գնահատելու համար անհրաժեշտ է պարզել, թե բացատրող փոփոխականների լազային արժեքները, որոնք մտցվելու են մոդելի մեջ, օգտակա՞ր են կախված փոփոխականի ապագա արժեքների կանխատեսման համար:

Գրեյնջերի պատճառականության թեստը առաջին անգամ առաջարկվել է Քլայվ Գրեյնջերի կողմից 1969 թ.: Թեստի հիմքը այն է, որ եթե y փոփոխականի ապագա արժեքները բացատրելու գործում x փոփոխականի նախորդ պահերի արժեքները պարունակում են վիճակագրական նշանակություն ունեցող տեղեկատվություն, ապա կարելի է ասել, որ x -ը ըստ Գրեյնջերի պատճառ է y -ի համար (x is "Granger-cause" y):

Գրեյնջերի մոտեցումն այն հարցին, թե արդյոք x -ը պատճառ է y -ի համար, այն է, որ պարզվում է, թե որքանով y -ի ներկա պահի արժեքը կարող է բացատրվել x -ի լազային կամ անցյալ արժեքներով, ինչպես նաև տեսնել, թե արդյոք x -ի լազային արժեքների ավելացումը կարող է բարելավել y -ի ներկա պահի արժեքի բացատրությունը: Ենթադրվում է, որ x -ը ըստ Գրեյնջերի պատճառ է y -ի համար, եթե x -ի լազային արժեքները օգնում են կանխատեսել y -ի ներկա պահի արժեքը: Պետք է նկատի ունենալ, որ հաճախ առկա է լինում երկկողմանի պատճառահետևանքային կապ. x -ը ըստ Գրեյնջերի պատճառ է y -ի համար, իսկ y -ը ըստ Գրեյնջերի պատճառ է x -ի համար:

Թեստի հիմնական ենթադրությունները

1. Ժամանակային առաջնություն. պատճառը պետք է տեղի ունենա ազդեցությունից առաջ:
2. Կանխատեսող տեղեկատվություն. պատճառը պետք է հատուկ տեղեկատվություն պարունակի «ազդվող» փոփոխականի ապագա արժեքների մասին:
3. Թեստը ենթադրում է ԺՇ-երի ստացիոնարության ստուգում. նախ պետք է ստուգվի մոդելի մեջ մտցվող ԺՇ-երի ստացիոնարությունը, և եթե դրանք ստացիոնար չեն, պետք է ստացիոնարացվեն: Թեստն աշխատում է ստացիոնար ԺՇ-երի հետ:

Ենթադրենք՝ ունենք երկու փոփոխական՝ x , y : Գրեյնջերի թեստում դրվում են հետևյալ վարկածները՝

1. x -ը ըստ Գրեյնջերի պատճառ չէ y -ի համար (X does not Granger Cause Y)

2. y-ը ըստ Գրեյնջերի պատճառ չէ x-ի համար (Y does not Granger Cause X)

Գրանց ստուգման համար գնահատվում է երկու օժանդակ ռեգրեսիայի հավասարում՝

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_l y_{t-l} + \beta_1 x_{t-1} + \dots + \beta_l x_{t-l} + \varepsilon_t,$$

$$x_t = \alpha_0 + \alpha_1 x_{t-1} + \dots + \alpha_l x_{t-l} + \beta_1 y_{t-1} + \dots + \beta_l y_{t-l} + u_t,$$

և երկու հավասարումների համար առանձին-առանձին դրվում է հետևյալ զրոյական վարկածը՝ $H_0: \beta_1 = \beta_2 = \dots = \beta_l = 0$: Այս վարկածը նշանակում է, որ.

- առաջին օժանդակ հավասարման համար՝ x-ը, ըստ Գրեյնջերի, պատճառ չէ y-ի համար,
- երկրորդ օժանդակ հավասարման համար՝ y-ը, ըստ Գրեյնջերի, պատճառ չէ x-ի համար:

Վարկածները ստուգվում են Վալդի F վիճականու միջոցով, ինչպես նաև H_0 վարկածի տեղի ունենալու հավանականության՝ p արժեքի միջոցով: Եթե $p < 0.05$, ապա H_0 -ն մերժվում է: Հակառակ դեպքում H_0 -ն մերժել չենք կարող:

Եթե քեստի արդյունքում ստացվի, որ երկու զրոյական վարկածներն էլ մերժվում են, ապա ստացվում է երկկողմանի պատճառականություն. x-ը ըստ Գրեյնջերի պատճառ է y-ի համար, և y-ը ըստ Գրեյնջերի պատճառ է x-ի համար:

Թեստի սահմանափակումները/թերությունները

- Կռեելյացիա, այլ ոչ իրական պատճառականություն, երբ Գրեյնջերի պատճառականությունը չի ենթադրում իրական պատճառականություն՝ պատճառ ու հետևանք իմաստներով: Այն միայն մատնանշում է կանխատեսումային պատճառականությունը. մի ԺՇ-ի լագային, անցյալ արժեքները օգնում են բացատրել մյուս ԺՇ-ի ներկա պահի արժեքը: Այսինքն՝ մի ԺՇ-ի լագային, անցյալ արժեքները կոռելացված են մյուս ԺՇ-ի ներկա պահի արժեքի հետ:
- Ոչ գծային կախվածություն չի ենթադրվում: Թեստը ենթադրում է գծային հարաբերություններ և կարող է չբացահայտել ոչ գծային պատճառական հարաբերությունները:
- Կոինտեգրացվածությունը հաշվի չի առնվում, այսինքն՝ փոփոխականների միջև երկարաժամկետ կապերը, կոինտեգրացվածությունը թեստում հաշվի չեն առնվում:

Ավտոռեգրեսիայի ու բաշխված լագերի (ARDL, Autoregressive Distributed Lag) մոդելը բազմաչափ ԺՇ -երի էկոնոմետրիկ մոդելավորման մեկ այլ գործիք է: Այն դիտարկում է կախյալ փոփոխականի t պահի արժեքի գծային ռեգրեսիոն կախվածությունը իր սեփական լագային արժեքներից և բացատրող փոփոխականների մույն t պահի ու լագային արժեքներից: ARDL մոդելները լայնորեն կիրառվում են ինչպես կարճաժամկետ, այնպես էլ երկարաժամկետ կապերի վերլուծության համար: ARDL մոդել կարելի է գնահատել ինչպես ստացիոնար և մույն՝ առաջին կարգի $I(1)$ ինտեգրացվող, այնպես էլ ինտեգրացման տարբեր $I(0)$ ու $I(1)$ կարգեր ունեցող ԺՇ -երի միջև (ինտեգրացման կարգը պետք է փոքր լինի 2-ից): Ամեն դեպքում այս մոդելի մեջ մտցվող փոփոխականները ևս պետք է ստացիոնար կամ ստացիոնարացված լինեն, եթե կոինտեգրացվող չեն:

ARDL մոդելները ընդլայնում են ավտոռեգրեսիայի $\text{AR}(p)$ մոդելները՝ ներառելով բացատրող փոփոխականներ ու դրանց լագային արժեքներ: Պարզագույն ARDL(p, q) մոդելը, որը պարունակում է մեկ բացատրող փոփոխական, ունի հետևյալ տեսքը՝

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_l y_{t-p} + \beta_0 x_t + \beta_1 x_{t-1} + \dots + \beta_l x_{t-q} + \varepsilon_t, \quad (195)$$

որում՝ $t=1, 2, \dots, n$, ε_t -ն ստացիոնար, սպիտակ աղմուկ ԺՇ է: Այս մոդելը $\text{AR}(p)$ և $\text{DL}(q)$ մոդելների համակցությունն է, որտեղ p -ն բնութագրում է մոդելի ավտոռեգրեսիայի մասի կարգը, իսկ q -ն՝ բաշխված լագերի մասի կարգը:

Եթե այս մոդելի մեջ մտցվող ԺՇ -երը կոինտեգրացվող են, ապա գնահատվում է ընդլայնված $\text{ARDL}+\text{ECM}$ մոդել, որը ներառում է սխալների ճշգրտման բաղադրիչ և գնահատում է ինչպես կարճաժամկետ, այնպես էլ երկարաժամկետ ազդեցությունները: Ինչպե՞ս ստուգել ԺՇ -երի կոինտեգրացվածությունն այս մոդելի համար: Եթե մոդելի ԺՇ -երը բոլորն առաջին կարգի ինտեգրացվող են, կարելի է օգտագործել Էնգել-Գրեյնջերի կամ Ֆիլիպս-Օուլիարուսի թեստերը: Իսկ եթե դրանք տարբեր կարգի ինտեգրացվող են, այս թեստերը չեն աշխատում, որովհետև դրանք պահանջում են, որ բոլոր ԺՇ -երը լինեն $I(1)$: Այդ դեպքում կոինտեգրացնող հավասարման գոյությունը կամ ԺՇ -երի կոինտեգրացվածությունը ստուգվում է «Սահմանների» թեստով (Bounds test, Pesaran, Shin and Smith, PSS): Թեսթն իրականացվում է առանց ԺՇ -երը ստացիոնարացնելու: Այն դնում է հետևյալ զրոյական վարկածը՝

H_0 : ԺՇ-երի միջև ոչ մի կոինտեգրացնող հավասարում գոյություն չունի,

H_1 : ԺՇ-երի միջև գոյություն ունի կոինտեգրացնող հավասարում:

Թեստի վիճականին F վիճականին է, որը ենթարկվում է որոշակի ոչ ստանդարտ բաշխման և համեմատվում այդ բաշխման կրիտիկական արժեքների հետ՝ 10, 5 և 1 % արժեքականությամբ:

Թեստի ելքերը

- Եթե F վիճականին փոքր է թեստի վիճականու ներքին սահմանի կրիտիկական արժեքից ($I(0)$ -ից), ապա H_0 -ն չենք կարող մերժել, այն ընդունվում է: Հետևաբար ԺՇ-երի միջև գոյություն չունի կոինտեգրացնող հավասարում, դրանք երկարաժամկետ հավասարակշռության մեջ չեն: Այս դեպքում գնահատվում է հետևյալ $ARDL(p,q)$ մոդելը ստացիոնարացված փոփոխականների միջև.

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \dots + \alpha_l \Delta y_{t-p} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + \dots + \beta_l \Delta x_{t-q} + \varepsilon_t: \quad (196)$$

- Եթե F վիճականին գտնվում է թեստի վիճականու վերին սահմանի ու ներքին սահմանի կրիտիկական արժեքների միջև, ապա թեստի արդյունքն անորոշ է:
- Եթե F վիճականին մեծ է թեստի վիճականու վերին սահմանի կրիտիկական արժեքից ($I(1)$ -ից), ապա H_0 -ն մերժվում է, ԺՇ-երի միջև գոյություն ունի կոինտեգրացնող հավասարում, դրանք երկարաժամկետ հավասարակշռության մեջ են: Այս դեպքում գնահատվում է $ARDL(p,q)+ECM$ ընդլայնված մոդելը, որն ընդգրկում է երկարաժամկետ ազդեցությունների մաս և ունի հետևյալ ընդհանուր տեսքը՝

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \dots + \alpha_l \Delta y_{t-p} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + \dots + \beta_l \Delta x_{t-q} + \theta_1 \cdot y_{t-1} + \theta_2 \cdot x_{t-1} + \varepsilon_t: \quad (197)$$

Հարցեր ու առաջադրանքներ գիտելիքների ստուգման նպատակով

1. Ի՞նչ է կեղծ ռեգրեսիան, և ի՞նչ դեպքերում ու պատճառով է այն առաջանում:
2. Ինչպե՞ս կարելի է ստուգել, որ ԺՇ-երի միջև գնահատված ռեգրեսիան կեղծ է:

3. Սահմանենք ԺՇ-երի կոինտեգրացիան և բացատրենք, թե ինչպես է այն մեկնաբանվում:
4. Ինչպե՞ս կարելի է ստուգել ԺՇ-երի կոինտեգրացվածությունը:
5. Նշենք էնգել-Գրեյնջերի մոտեցման հիմնական քայլերը:
6. Ինչպե՞ս է կոինտեգրացիան արտահայտում երկարաժամկետ հավասարակշռություն:
7. Ինչո՞ւ է սխալների ճշգրտման մոդելը (ECM) կարևոր կոինտեգրացված ԺՇ-երի մոդելավորման համար:
8. Գրենք սխալների ճշգրտման մոդելի ընդհանուր տեսքը և մեկնաբանենք դրա գործակիցների իմաստը:
9. Գրենք բաշխված լագերի մոդելի (DL) ընդհանուր տեսքը և բացատրենք դրանց գործակիցների իմաստը:
10. Բացատրենք բաշխված լագերի մոդելի մուլտիպլիկատորները, հարաբերական գործակիցները, մեդիանային ու միջին լագերը:
11. Նկարագրենք Գրեյնջերի պատճառականության թեստը: Ինչո՞ւ է այն կարևոր դինամիկ մոդելները գնահատելիս:
12. Նկարագրենք ավտոռեգրեսիայի ու բաշխված լագերի ARDL մոդելը և դրա ARDL+ECM տարբերակը:
13. Նկարագրենք «Սահմանների» թեստը: Ինչի՞նչ համար և ե՞րբ է այն օգտագործվում:

SUMMARY

Econometrics (Part 1: Foundations)

Nairuhi M. Jrbashyan

This textbook offers the first comprehensive Armenian-language introduction to the fundamentals of econometrics. It presents the theoretical foundations and practical applications of econometric modeling with spatial data and time series, combining rigorous exposition with accessible examples. Each chapter integrates review questions and exercises to reinforce theoretical knowledge and develop applied skills.

The content is closely aligned with the *Econometrics* curriculum taught at the Faculty of Economics and Management of Yerevan State University, ensuring both academic relevance and pedagogical coherence. Designed for undergraduate and graduate students in economics and related disciplines, the textbook equips readers with essential tools for working with data and applying econometric methods in empirical research.

АННОТАЦИЯ

Эконометрика (Часть 1: Основы)

Наируи М. Джрбашян

Учебник представляет собой первое всеобъемлющее изложение основ эконометрики на армянском языке. Он охватывает как теоретические основы, так и практические приложения эконометрического моделирования пространственных данных и временных рядов, сочетая научное изложение с доступными примерами. Каждый раздел включает вопросы и задания для проверки и закрепления полученных теоретических знаний, а также задачи, направленные на развитие прикладных навыков.

Теоретическое и прикладное содержание учебника соответствует программе курса «Эконометрика», преподаваемого на факультете экономики и управления Ереванского государственного университета, обеспечивая как академическую актуальность, так и методическую целостность. Учебник предназначен для студентов бакалавриата и магистратуры по экономическим и смежным специальностям, которые стремятся развить навыки работы с данными и их эконометрического моделирования.

ՕԳՏԱԳՈՐԾՎԱԾ ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ

1. Հարությունյան Ե., Ղազանչյան Տ., Մեսրոպյան Ն. և այլք, Հավանականություն և կիրառական վիճակագրություն տնտեսագետների ու գործարարների համար, Երևան, «Գիտություն», 2000:
2. Ղուշչյան Լ., Թերզյան Տ., Դավթյան Լ., Տարրական էկոնոմետրիկա, Երևան, «ԶԵՄԹ», 2002:
3. Վարդանյան Գ., էկոնոմետրիկայի հիմունքները, Երևան, «Սարվարդ» հրատ., 2003:
4. Айвазян С. А., Методы эконометрики: учебник, М.: «Магистр» ИНФРА-М, 2010:
5. Артамонов Н. В., Ивин Е. А., Курбацкий А. Н., Фантащин Д., Введение в анализ временных рядов. Вологда: ВолНЦ РАН, 2021.
6. Ивин Е. А., Артамонов Н. В., Курбацкий А. Н., Методическое пособие по эконометрике для социально-экономических специальностей. Вологда: ИСЭРТ РАН, 2016.
7. Кремер Н. Ш., Эконометрика : учебник и практикум для вузов. 4-е изд., М.: «Юрайт», 2026.
8. Магнус Я.Р., Катышев П.К., Пересецкий А.А., Эконометрика. Начальный курс. 8-е издание, М., «Дело», 2007.
9. Мхитарян В. С., Архипова М. Ю., Сиротин В. П., Эконометрика, М.: «ЕАОИ», 2008.
10. Носко В. П., Эконометрика. Книга 1, 2, М.: «Дело», 2011.
11. Agung I. Gusti Ngurah, Advanced Time Series Data Analysis: Forecasting Using EViews. John Wiley & Sons, 2019.
12. Aljandali Abdulkader and Motasam Tatahi, Economic and Financial Modelling with EViews: A Guide for Students and Professionals. Springer International Publishing, 2018.
13. Baltagi Badi H., Econometrics. 6th ed., Springer, 2021.
14. Békés Gábor and Gábor Kézdi, Data Analysis for Business, Economics, and Policy. Cambridge University Press, 2021.
15. Brockwell Peter J., and Richard A. Davis, Introduction to Time Series and Forecasting. 3rd ed., Springer, 2016.
16. Brooks C., Introductory Econometrics for Finance. 4th ed., Cambridge University Press, 2019.

17. Christensen Ronald, Log-Linear Models and Logistic Regression. 3rd ed., Springer, 2025.
18. Das Panchanan, Econometrics in Theory and Practice. Springer, 2019.
19. Greene William H., Econometric Analysis. 7th ed., Pearson Education, 2012.
20. Gujarati Damodar N., Porter Dawn C., Basic econometrics. 5th ed. McGraw Hill. 2009.
21. Hansen Bruce E., Econometrics. Princeton University Press, 2022.
22. Hill R. Carter, and William E. Griffiths. Using EViews for Principles of Econometrics. 5th ed. John Wiley & Sons, 2018.
23. Hill R. Carter, William E. Griffiths Guay C. Lim. Principles of Econometrics. 5th ed, John Wiley & Sons, 2018.
24. Levendis John D., Time Series Econometrics: Learning Through Replication. 2nd ed., Springer, 2023.
25. Mignon Valérie, Principles of Econometrics: Theory and Applications. Springer, 2024.
26. Neusser Klaus, Time Series Econometrics. 2nd ed., Springer Nature, 2025.
27. Stockemer Daniel, Quantitative Methods for the Social Sciences: A Practical Introduction with Examples in SPSS and Stata. Springer, 2019.
28. Stock James H., and Mark W. Watson, Introduction to Econometrics. 4th ed., Pearson, 2020.
29. Studenmund A. H., Using Econometrics: A Practical Guide. 7th ed., Pearson, 2017.
30. Wooldridge Jeffrey M., Introductory Econometrics: A Modern Approach. 7th ed., Cengage Learning, 2020.

ԵՐԵՎԱՆԻ ՊԵՏԱԿԱՆ ՀԱՄԱԼՍԱՐԱՆ

ՆԱԻՐՈՒՀԻ ՄԵԻԹԱՐԻ ՋՐԲԱՇՅԱՆ

ԷԿՈՆՈՄԵՏՐԻԿԱ

(Մաս 1, Հիմունքներ)

Համակարգչային ձևավորումը՝ Կ. Չալաբյանի
Կազմի ձևավորումը՝ Ա. Պատվականյանի
Հրատ. խմբագիր՝ Մ. Հովհաննիսյան

Հեղինակը հաստատում է, որ ծանոթ է «ԵՊՀ գրահրատարակչական քաղաքականությանը», և գրքում առկա փաստերը, դիրքորոշումները, կարծիքները շարադրված են հեղինակային իրավունքի և էթիկայի միջազգայնորեն ընդունված սկզբունքների պահպանմամբ:

Տպագրված է «Արփինե Բեգլարյան» ԱԶ-ում:
Հասցե՝ ք. Եղվարդ, Զովունի 11փ. 812.

Հրատարակվել է՝ 10.07.2026:
Չափսը՝ 60x84 ¹/₁₆: Տպ. մամուլը՝ 18,875:
Տպաքանակը՝ 100:

ԵՊՀ հրատարակչություն
ք. Երևան, 0025, Ալեք Մանուկյան 1
www.publishing.yu.am



publishing.ysu.am