QO‘QON UNIVERSITETI

O‘zbekiston Respublikasi
davlat mustaqilligining
30 yilligiga bag‘ishlanadi!

SH.I. MUSTAFAKULOV,
H.N. SABIROV

EKONOMETRIKA I

(O‘quv qo‘llanma)

TOSHKENT – 2022

UO‘K: 330.42 (07.58)
KBK 65.012.2ya7
SH 59

Sh.I.Mustafakulov, H.N.Sabirov. Ekonometrika I. O‘quv qo‘llanma. – T.: «Ilm-fan va innovatsiya» 2022, 232 bet.

ISBN 978-9943-7965-9-1

Mazkur o‘quv qo‘llanma "Ekonometrika I" fani o‘quv dasturiga muvofiq yozilgan. Unda ekonometrik modellashtirish asoslari, ekonometrik modellarning axborot ta’minoti, juft va ko‘p omilli ekonometrik tahlil, ekonometrik modellarni baholash mazmun-mohiyati yoritib berilgan. O‘quv qo‘llanma iqtisodiyot sohasi ta’lim muassasalari bakalavriat va magistratura dasturi bo‘yicha ta’lim olayotgan talabalar, shuningdek, ekonometrikani mustaqil o‘rganuvchilarga mo‘ljallangan bo‘lib, u ekonometrika sohasidagi fanlarni o‘rganishda tavsiya etiladi.

Bundan tashqari mazkur o‘quv qo‘llanmadan professor - o‘qituvchilar hamda tadqiqotchilar ham foydalanishlari mumkin.

***

Данное учебное пособие написан в соответствии с учебной программой «Эконометрика I». Он охватывает основы эконометрического моделирования, информационного обеспечения эконометрических моделей, двойного и многофак торного эконометрического анализа, оценки эконометрических моделей. Учебное пособие предназначено для студентов и аспирантов экономических специаль ностей, а также самостоятельных студентов, изучающих эконометрику, и рекомендуется для изучения эконометрики.

Пособие также может быть использовано профессорами, преподавателями и исследователями.

UO‘K: 330.43(07.58)
KBK 65.012.2ya7

Mas’ul muharrir:

Baxtiyor Salimov – Toshkent davlat iqtisodiyot universiteti professori, iqtisodiyot fanlari doktori.

Taqrizchilar:

Muyassar Mirzakarimova – Qo‘qon universiteti professori, iqtisodiyot fanlari doktori;
Abdulla Almuradov – Toshkent davlat iqtisodiyot universiteti, "Iqtisodiyotda matematik metodlar" kafedrasi dotsenti.

Ushbu o‘quv qo‘llanma Qo‘qon universiteti Kengashining 2022 yil 25 apreldagi 8- son qarori asosida nashr qilindi.

ISBN 978-9943-7965-9-1

Mualliflardan kitobxonlarga!

Qadirdon kitobxonlar, avvalambor, Ekonometrika I nomli o‘quv qo‘llanmadan foydalanayotganingiz uchun sizga minnatdorchilik bildiramiz. Ushbu o‘quv qo‘llanma talabalar va professor-o‘qituvchilarning talabi, ta’lim ehtiyojidan kelib chiqqan holda yaratildi. O‘quv qo‘llanma ekonometrika bilan bog‘liq amaliy muammolarni hal qilish va ma'lumotlar to‘plamlari bilan ishonchli ishlashga imkon beruvchi amaliy masala-misollar bilan boyitilgan.

Bugungi kunda birorbir soha, yo‘nalish yo‘q-ki, ekonometrik tahlillar kirib bormagan bo‘lsa. Ijtimoiy-iqtisodiy hayotimizda sodir bo‘layotgan voqea-hodisalar sabab va oqibat tamoyiliga bog‘langandir. Biror bir nazariya yoki amaliyotda yuz beradigan o‘zgarishlar tafsiloti, albatta, gipotezalar orqali tushuntiriladi. Shuningdek, iqtisodiy jarayonlarni modellar orqali izohlashda ekonometrik instrumentlardan keng foydalaniladi.

Umid qilamizki, ushbu o‘quv qo‘llanmadan foydalangan holda ekonometrika haqida tushunchaga ega bo‘lasiz. Bu qo‘llanmaning bir turi bo‘lib, Excel dasturidan foydalangan holda qo‘llanmadagi formulalarni qanday bajarish kerakligini misollar va dastur menyulari asosida ko‘rsatib beradi. Ushbu kitob ekonometrika fanini o‘rganayotgan talabalar, shuningdek, o‘qituvchilar va Excel dasturidan ekonometrik va statistik tahlil uchun foydalanmoqchi bo‘lganlar uchun foydalidir. Qo‘llanmada Excel kompyuter dasturidan tashqari, shunga o‘xshash EViews, Stata, Gretl va Shazam dasturiy ta'minot paketlaridan ham foydalanilgan.

Biz ushbu o‘quv qo‘llanma bo‘yicha izohlarni va takomillashtirish bo‘yicha fikr-mulohazalarni, taklif-tavsiyalarni mamnunyat bilan qabul qilamiz. Savollarimizga javob berish va qo‘llanmani takomillashtirishga beradigan takliflaringizni mualliflarning elektron pochta manziliga yuborishingizni so‘raymiz. (mustafaqulov_sh@mail.ru; khasansabirov19@gmail.com). Navbatdagi izlanishlarimizda Stata dasturiy paketining ekonometrik modellashtirishdagi imkoniyatlarini yoritish, shunga o‘xshash dasturiy ta’minot paketlarining qiyosiy tahlili rejalashtirilgan va sizlar bergan takliflaringiz orqali takomillashtiriladi.

KIRISH

Bugungi kunda ekonometrik usullar iqtisodiy tahlil va prognozlashning eng muhim instrumentiga aylanib bormoqda. Ekonometrik tahlilda statistik axborotlarni qayta ishlash va tahlil qilishda bir qator universal dasturiy mahsulotlar mavjud. Koʻrib chiqilayotgan masalalar koʻlamiga qarab, ular ekonometrik usullarni oʻrganishda nafaqat talabalarga, balki statistik maʼlumotlardan foydalanib, iqtisodiy tahlil va prognozlash masalalarini hal qilishda ilmiy izlanuvchilar hamda iqtisodchilarga ham foydali boʻlishi mumkin.

Mazkur oʻquv qoʻllanmada barcha asosiy ekonometrik hisoblarni Excel amaliy paketida amalga oshirish koʻzda tutilga. Excel dasturini tanlash undan foydalanishning qulayligi va ekonometrik modellashtirish amaliyotida keng qoʻllanishi bilan asoslanadi.

Excel dasturi ekonometrik modellashtirishning barcha bosqichlarini oʻzida mujassamlashtirgan. Unda maʼlumotlarni kiritish, ular asosida tavsifiy statistikalarni oʻtkazish, korrelyatsion tahlil, juft va koʻplikdagi regression tahlillarni oʻtkazish, fazoviy va vaqtli qatorlar boʻyicha turli xil grafiklarni olish, olingan natijalarni turli testlar yordamida tekshirish imkoniyati mavjud. Yanada aniqroq toʻxtaladigan boʻlsak, ushbu qoʻllanma 13ta bobdan iborat. Birinchi bob ekonometrikaga kirish deb nomlanadi. Bu bobda ekonometrikaning maqsad va vazifalari, iqtisodiyotni ekonometrik modellashtirishning zarurligi, ekonometrik model tushunchasi, turlari va undagi oʻzgaruvchilar haqida tushunchalar ifodalanilgan.

Ikkinchi bob ekonometrik modellarning axborot taʼminoti deb nomlangan va bu qismda iqtisodiy maʼlumotlarning statistik tabiati, bogʻliq va bogʻliq boʻlmagan oʻzgaruvchilarni tanlash va ekonometrik modellarni tuzishda qatnashadigan iqtisodiy maʼlumotlarga qoʻyiladigan talablar tushuntirib oʻtilgan. Uchinchi bob esa ekonometrikada ehtimollar nazariyasi va matematik statistikaning asosiy tushunchalari deb nomlandi hamda ehtimollar nazariyasi va matematik statistikaning asosiy tushunchalari, toʻplamlar va ularning xossalari va diskret va uzluksiz tasodifiy miqdorlar, tasodifiy miqdorlarning xarakteristikalarini hisoblash jarayonlariga to‘xtalib o‘tilgan. To‘rtinchi bob Juft korrelyatsion tahlil deb nomlandi. Bu bobda asosan siz funksional va statistik bog‘liqlik tushunchalari, korrelyatsion tahlil tushunchasi, bog‘lanish turlari va korrelyatsiya koeffitsiyentini hisoblash usullari, korrelyatsiya koeffitsiyentini o‘zgarish intervallari va baholanish haqida bilim ko‘nikmalarga ega bo‘lasiz. Beshinchi bobni juft regression tahlil deb nomladik va bu bobda regression tahlil tushunchasi, bir omilli va ko‘p omilli regressiya, chiziqli va chiziqsiz regressiya, korrelyatsion-regression tahlilda eng kichik kvadratlar usulining qo‘llanilishi va o‘rtacha elastiklik koeffitsiyenti tushunchalari ifodalangan.

Shuningdek, keyingi boblarda standart normal taqsimot, Z - statistika va uning mohiyati, normal taqsimlash zichligi, styudent t-taqsimot mezoni, ijtimoiy-iqtisodiy jarayonlar o‘rtasida bog‘lanishlarni o‘rganishda chiziqsiz funksiyalar bilan foydalanish, chiziqsiz regressiya modellari, chiziqsiz bog‘lanishlar uchun korrelyatsiya indeksini hisoblash, "Eng kichik kvadratlar" (EKK) yordamida chiziqsiz regressiya koeffitsiyentlarini hisoblash, ekonometrik modellarning iqtisodiy tahlilida verifikatsiya bosqichining ahamiyati, ekonometrik modellar sifati va ahamiyatini mezonlar bo‘yicha baholash, regressiya tenglamaning parametrlarni baholarining xususiyatlari, ekonometrik modellarning iqtisodiy tahlilida verifikatsiya bosqichining ahamiyati, Gauss-Markov teoremasi, gipoteza testlari, binar o‘zgaruvchilar, multikollinearlik, Laglar, birinchi ayirma, logarifm va o‘sish darajalari, avtoregressiya modellari, avtoregressiv taqsimlangan lag (ADL) modeli va nostatsionarlik to‘g‘risida ketma-ket mavzular yoritilgan.

I BOB. EKONOMETRIKAGA KIRISH

1.1. Ekonometrika fanining maqsad va vazifalari

1.2. Iqtisodiyotni ekonometrik modellashtirishning zarurligi

1.3. Ekonometrik model tushunchasi, turlari va undagi o‘zgaruvchilar

1.1. Ekonometrika fanining maqsadi va vazifalari

Ekonometrik bilimlar iqtisodiy nazariya, iqtisodiy matematika, iqtisodiy statistika, ehtimollar nazariyasi va matematik statistika kabi fanlarning o‘zaro bog‘liqligi va rivojlanishining natijasi sifatida ajralib chiqqan va shakllangan.

Ekonometrika o‘zining predmeti, maqsadi va tadqiqot masalalarini shakllantiradi. Shu bilan birga ekonometrikaning mazmuni, uning tarkibi va qo‘llanilish sohasi yuqorida keltirilgan fanlar bilan doimo aloqada bo‘ladi.[1]

Ekonometrikaning boshqa fanlar bilan o‘zaro aloqasi quyidagilarda namoyon bo‘ladi (1-jadval).

1.1-jadval

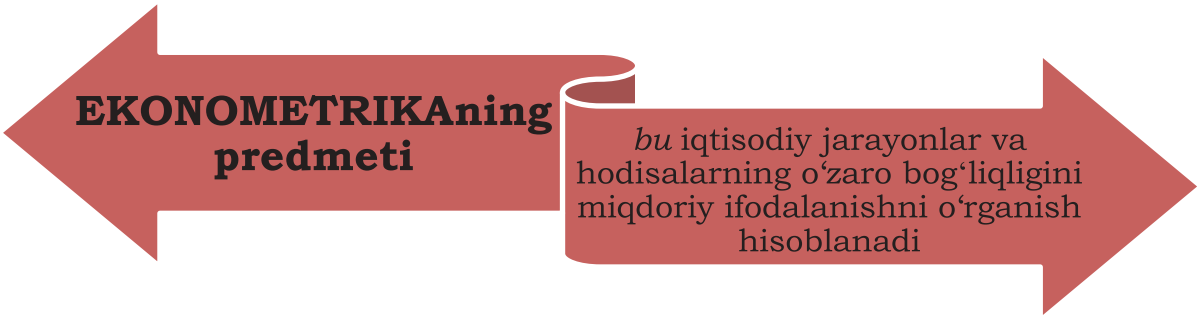
Ekonometrikaning boshqa fanlar bilan o‘zaro aloqasi[2]

Ekonometrika	Boshqa fanlar
Iqtisodiy hodisalar miqdoriy xarakteristikalar nuqtai nazaridan o‘rganiladi. Iqtisodiy qonunlarning amaldagi jarayonlarga mos kelishi tekshiriladi.	Iqtisodiy nazariya. Iqtisodiy hodisalarning sifat jihatlari o‘rganiladi. Matematik iqtisodiyot. Iqtisodiy qonunlarning ifodasi matematik modellar shaklida olinadi.
Iqtisodiy statistikaning instrumentariylari iqtisodiy o‘zaro aloqalarni tahlil qilish va bashorat qilish uchun qo‘llaniladi. Iqtisodiy ko‘rsatkichlarning katta qismi tasodifiy xarakterga ega bo‘lganligi uchun matematik statistikaning apparatidan foydalaniladi.	Iqtisodiy statistika. Iqtisodiy ma’lumotlar ko‘rgazmali shaklda namoyish etish uchun to‘planadi va qayta ishlanadi Matematik statistika. Tadqiqot maqsadidan kelib chiqib, ma’lumotlarni tahlil qilish usullari ishlab chiqiladi.

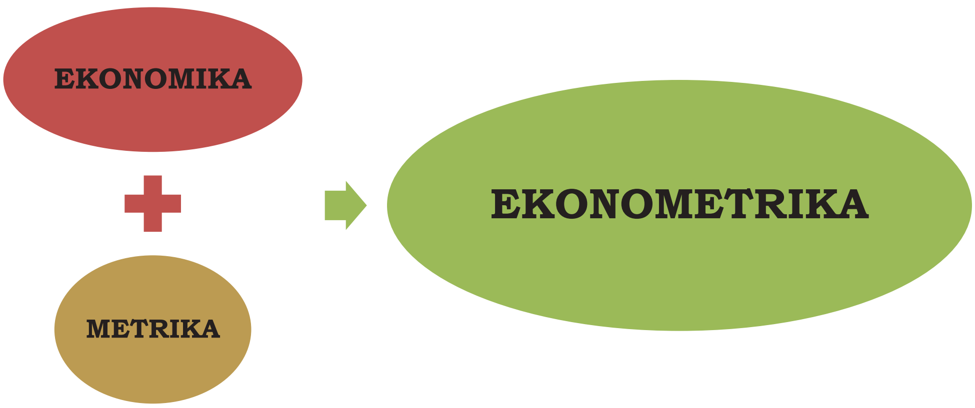
[1] Gujarati D.N. Basic Econometrics. McGraw-Hill, 4th edition, 2003 (Gu), Inc.p. 7

[2] http://www.studfiles.ru/preview/6327597/

Iqtisodchilar “EKONOMETRIKA” terminidan P. Sempa (1910), Y.Shumpeter (1923), R.Frish (1930) larning tadqiqotlari natijasida qo‘llay boshladilar.



Ekonometrika sohasidagi tadqiqotchi olimlar “ekonometrika” tushunchasining mazmunini belgilashda turli yondashuvlarga asoslanishgan. Grek tilidan tarjima qilganda OIKONOMOS (ekonomist) – bu uy boshqaruvchisi, METRIKA *(metrihe, metron)* – o‘lchov ma’nolarini bildiradi.



Ekonometrik modellash iqtisodiy ko‘rsatkichlarni o‘zgarish qonuniyatlarini, tendensiyalarni aniqlash natijasida ekonometrik modellar yordamida iqtisodiy jarayonlarni rivojlanish va prognozlash yo‘llarini belgilaydi.

Ekonometrikani aniqlash bo‘yicha yondashuvlar tahlili hamda ekonometrika fanining holati ayrim masalalarni yechishga erishishda ushbu fanning maqsadini shakllantiradi.

Ekonometrikaning asosiy maqsadi – omillararo bog‘lanishlarni, o‘zgarish qonuniyatlarini va tendensiyalarni o‘rganish hisoblanadi.

1.2-jadval

Ekonometrika tushunchasining mazmuni[3]

Muallif	"Ekonometrika" tushunchasining mazmuni
R. Frish 	«...uchta tashkil etuvchi - statistika, iqtisodiy nazariya va matematika fanlarining birlashuvidir»
S. Grilixes 	«...bizni o'rab turgan iqtisodiy dunyoni o'rganish uchun bir vaqtning o'zida bizning teleskopimiz hamda mikroskopimizdir»
E. Malenvo 	«...bizning xayoliy iqtisodiy tasavvurlarimizni empirik mazmun bilan to'ldiradi»
S. Fisher 	«...iqtisodiy o'zgaruvchilar o'rtasida o'zaro aloqalarni o'lchash uchun statistik usullarni ishlab chiqish va qo'llash bilan shug'ullanadi»
S. Ayvazyan 	«...sifat jihatdan o'zaro bog'lanishlarga miqdoriy ifodani berishga imkon beruvchi usullar va modellar to'plamini birlashtiradi»

3 http://www.studfiles.ru/preview/6327597/

Ekonometrikaning maqsadi - bu real iqtisodiy obyektlarni modellashtirish va miqdoriy tahlil qilishning usullarini ishlab chiqishdan iborat.

Ekonometrik modellarni tuzish, ya'ni amaliy jihatdan tekshirish mumkin bo'lgan modelni ifodalash. Bu ekonometrika tilida spetsifikatsiya deb ataladi. Bu modelni tanlangan va kuzatilgan ma'lumotlar asosida tekshirish. Bu qism ekonometrik modelning adekvatligi va to'g'riligi bilan bog'liq. Modelni prognoz va qarorlar qabul qilish uchun ishlatiladi.

Ekonometrikaning vazifalari:

Modelni spetsifikatsiya qilish - empirik tahlil uchun ekonometrik modellarni tuzish.

Modelni parametrlashtirish - tuzilgan model parametrlarini baholash.

Modelni verifikatsiya qilish - model parametrlari sifatini va butun modelning o'zini tekshirish

Model asosida prognoz qilish - ekonometrik modellashtirish natijalari bo'yicha aniq iqtisodiy hodisalar uchun prognozlar tuzish va takliflar ishlab chiqish

Ekonometrikaning metodologiyasi. Ekonometristlar iqtisodiy muammoni tahlil qilishda qanday yo'l tutadilar, ya'ni ularning metodologiyasi nimalardan iborat? Ekonometrika metodologiyasi bo'yicha bir necha maktablar mavjud, lekin biz bu yerda hozirgacha iqtisodiyot va boshqa ijtimoiy fanlarning empirik tadqiqotlarida ustunlik qilib kelayotgan an'anaviy yoki klassik metodologiyani keltirib o'tamiz[4].

An'anaviy ekonometrik metodologiyalar quyidagi yo'nalishlarda olib boriladi:

1. Nazariya yoki gipotezaning qo'yilishi.

---

[4] Ekonometrika metodologiyasi bilan batafsilroq tanishish uchun quyidagiga qarang. David F. Hendry, *Dynamic Econometrics*, Oxford University Press, New York, 1995. See also Aris Spanos, *op. cit*.

2. Nazariyaning matematik modelini aniqlashtirish.
3. Statistik yoki ekonometrik modelni aniqlashtirish.
4. Ma’lumotlarni to‘plash.
5. Ekonometrik model parametrlarini baholash.
6. Gipotezalarni testdan o‘tkazish.
7. Prognozlash yoki oldindan aytib berish.
8. Boshqarish maqsadlari uchun modeldan foydalanish.

Yuqorida keltirilgan qadamlarni izohlash uchun J.M.Keynsning iste’mol nazariyasini ko‘rib chiqamiz.

1. Nazariya yoki gipotezaning qo‘yilishi.

Keyns ta’kidlaydiki "...fundamental psixologik qonun shundan iboratki, qoidaga ko‘ra erkaklar (ayollar) o‘zlarining o‘rtacha daromadlari ortishi bilan daromadga nisbatan unchalik katta bo‘lmagan darajada o‘zlarining iste’mollarini oshirishga harakat qiladilar"[5].

Qisqacha aytganda, Keyns iste’molga bo‘lgan chekli moyillik (MRS) bu daromadning bir birlikka (1 dollarga) o‘zgarishi bilan iste’moldagi o‘zgarishning tezligi bo‘lib, noldan katta, ammo 1 dan kichik deb faraz qilgan.

2. Iste’molning matematik modelini spetsifikatsiya qilish.

Keyns daromad va iste’mol o‘rtasida musbat o‘zaro bog‘liqlik mavjudligini aytgan bo‘lsada, ular o‘rtasida bog‘liqlik qaysi shaklda ekanligini aniqlashtirmagan. Soddalik uchun, iqtisodchi-matematik Keynsning iste’mol funksiyasini quyidagi ko‘rinishda taklif qilishi mumkin:

$$Y = \beta_1 + \beta_2 X, \qquad 0 < \beta_2 < 1 \qquad (1.1)$$

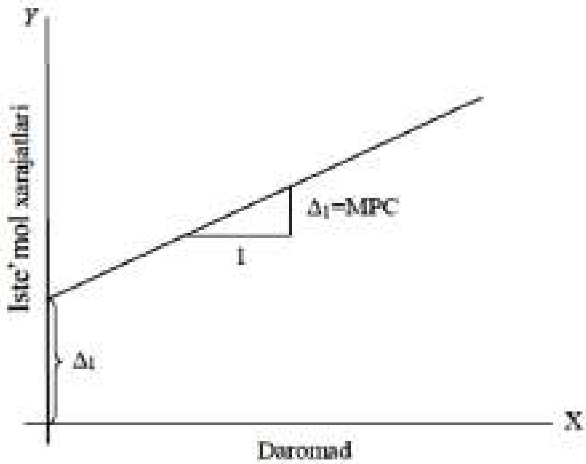
bu yerda, Y - iste’mol uchun xarajatlar va X - daromad, $\beta_1$va $\beta_2$lar modelning ma’lum parametrlari va mos ravishda burchak koeffitsiyentlari hisoblanadi.

β_2 burchak koeffitsiyenti iste’molga bo‘lgan chekli moyillikni o‘lchaydi. (1.1) tenglama geometrik tarzda 1.1-rasmda keltirilgan.

[5] John Maynard Keynes, *The General Theory of Employment, Interest and Money*, Harcourt Brace Jovanovich, New York, 1936, p. 96.

Ushbu tenglama iste’mol daromad bilan chiziqli bog‘langan bo‘lib, bu iste’mol va daromad o‘rtasida o‘zaro bog‘liqlikning matematik modeliga misol bo‘ladi hamda u iqtisodiyotda iste’mol funksiyasi deb ataladi. Model matematik tenglamalar to‘plami sifatida namoyon bo‘ladi. Agar model bitta tenglamaga ega bo‘lsa, u bir tenglamali model deyiladi, agar bittadan ortiq tenglamaga ega bo‘lsa, u holda ko‘plikdagi model deyiladi.

(1.1) formuladagi tenglik belgisidan chap tomonda turgan o‘zgaruvchi bog‘liq o‘zgaruvchi, o‘ng tomonidagi esa bog‘liq bo‘lmagan yoki tushuntirib beruvchi o‘zgaruvchi deyiladi. Shunday qilib, Keynsning iste’mol funksiyasida (1.1) tenglamadagi iste’mol funksiyasi bog‘liq o‘zgaruvchi va daromad manbai bog‘liq bo‘lmagan yoki tushuntirib beruvchi o‘zgaruvchi hisoblanadi.



1-rasm. Keynsning iste’mol funksiyasi

1.2. Iqtisodiyotni ekonometrik modellashtirishning zarurligi

Ekonometrik usullar oddiy an’anaviy usullarni inkor etmasdan, balki ularni yanada rivojlantirishga va obyektiv o‘zgaruvchan natija ko‘rsatkichlarini boshqa ko‘rsatkichlar orqali muayyan tahlil qilishga yordam beradi. Ekonometrik usullarning va kompyuterlarning milliy iqtisodiyotni boshqarishda afzalliklaridan biri shundaki, ular yordamida modellashtiruvchi obyektga omillarning ta’sirini, natija ko‘rsatkichiga resurslarning o‘zaro munosabatlarini ko‘rsatish mumkin. Bu esa o‘nlab tarmoqlar va minglab korxonalarda ishlab chiqarish natijalari va milliy iqtisodiyotni ilmiy asosda prognozlashtirish va boshqarishga imkon beradi.

Ekonometrik modellash iqtisodiy ko‘rsatkichlarni o‘zgarish qonuniyatlarini, tendensiyalarni aniqlash natijasida ekonometrik modellar yordamida iqtisodiy jarayonlarni rivojlanish va prognozlash yo‘llarini belgilaydi.

Ekonometrik modellashtirish va modellarning ahamiyati quyidagilarda namoyon bo‘ladi:

1) Ekonometrik usullar yordamida moddiy, mehnat va pul resurslaridan oqilona foydalaniladi.

2) Ekonometrik usullar va modellar iqtisodiy va tabiiy fanlarni rivojlantirishda yetakchi vosita bo‘lib xizmat qiladi.

3) Ekonometrik usullar va modellar yordamida tuzilgan prognozlarni umumiy amalga oshirish vaqtida ayrim tuzatishlarni kiritish mumkin bo‘ladi.

4) Ekonometrik modellar yordamida iqtisodiy jarayonlar faqat chuqur tahlil qilibgina qolmasdan, balki ularning yangi o‘rganilmagan qonuniyatlarini ham ochishga imkon yaratiladi. Shuningdek, ular yordamida iqtisodiyotning kelgusidagi rivojlanishini oldindan aytib berish mumkin.

5) Ekonometrik usullar va modellar hisoblash ishlarini avtomatlashtirish bilan birga, aqliy mehnatni yengillashtiradi, iqtisodiy soha xodimlarining mehnatini ilmiy asosda tashkil etadi va boshqaradi.

Ekonometrikani o‘rganish jarayoni – bu iqtisodiyot, iqtisodiy jarayonlarning ekonometrik modellarini tuzish jarayonidir.

Asosiy qo‘llanadigan usuli – korrelyatsion-regression tahlil usuli.

Ekonometrik modellashtirish quyidagi ilmiy yo‘nalishlar kompleksidir:

- iqtisodiy nazariya;
- ehtimollar nazariyasi;
- matematik statistika;
- kompyuter texnologiyalari.

1.3. Ekonometrik model tushunchasi, turlari va undagi o‘zgaruvchilar

Kuzatilayotgan obyektlarni chuqur va har tomonlama o‘rganish maqsadida tabiatda va jamiyatda ro‘y beradigan jarayonlarning modellari yaratiladi. Buning uchun obyektlar hamda ularni xossalari kuzatiladi va ular to‘g‘risida dastlabki tushunchalar hosil bo‘ladi. Bu tushunchalar oddiy so‘zlashuv tilida, turli rasmlar, sxemalar, belgilar, grafiklar orqali ifodalanishi mumkin. Ushbu tushunchalar **model** deb aytiladi.

Model so‘zi lotincha *modulus* so‘zidan olingan bo‘lib, o‘lchov, me’yor degan ma’noni anglatadi. Keng ma’noda model biror obyektni yoki obyektlar sistemasini namunasidir. Model tushunchasi biologiya, meditsina, fizika va boshqa fanlarda ham qo‘llaniladi. Jamiyatdagi va iqtisodiyotdagi obyektlarni matematik modellar yordamida kuzatish mumkin. Bu tushuncha modellashtirish deyiladi.

Tuzilgan barcha modellarni 2 turga bo‘lish mumkin: moddiy modellar va ideal modellar (1.2-rasm).

MODELLAR

MODDIY MODELLAR

Moddiy modellar real obyektlarni tabiiy va sun’iy materiallar yordamida aks ettiradi: mel bilan doskaga yozish, karton bilan maket tuzish, qalam bilan formula yozish, metalldan aviamodel yasash.

IDEAL MODELLAR

Ideal modellar odamni fikrlash jarayoni bilan chambarchas bog‘langandir. Bunday modellar bilan operatsiyalar miyada amalga oshiriladi.

1.2-rasm. Modellar turlari[6]

Iqtisodiy model - iqtisodiy obyektlarning soddalashtirilgan nusxasidir. Bunda modelning hayotiyligi, uning modellashtiriladigan obyektga aynan mos kelishi muhim ahamiyatga egadir. Lekin yagona modelda o‘rganilayotgan obyektning hamma tomonini aks ettirish mumkin emas. Shunda jarayonning eng xarakterli va eng muhim belgilari aks ettiriladi.

Ekonometrik model – bu ehtimollik-stoxastik model. Bu model yordamida iqtisodiy ko‘rsatkichlarni o‘zgarish qonuniyatlarini matematik ko‘rinishida tenglamalar, tengsizliklar va tenglamalar tizimi

[6] Shodiyev T.Sh. va boshqalar. Iqtisodiy-matematik usullar va modellar. O‘quv qo‘llanma. –T.: TDIU, 2010, 8 b.

ko'rinishda ifodalash mumkin. Umumiy ko'rinishida ekonometrik model quyidagicha yoziladi:

$$Y = f(x_1, x_2, \dots x_n)$$

Ekonometrik modelda Y – asosiy **endogen ko'rsatkich**, modelda Y o'zgarish qonuniyatlarini $(x_1, x_2, \dots x_n)$ yordamida o'rganish mumkin.

$(x_1, x_2, \dots x_n)$– ta'sir etuvchi, **ekzogen ko'rsatkichlar**.[7]

Ekonometrik modelda fiktiv ko'rsatkichlar qatnashishi mumkin. Fiktiv ko'rsatkichlar – bu sifatli ko'rsatkichlar miqdoriy ko'rsatkichlarga o'tkazilgan ko'rsatkichlar.

Ekonometrik model chiziqli va chiziqsiz ko'rinishda tuzilishi mumkin.

Chiziqsiz modellar parabola, giperbola, darajali funksiya, ko'rsatkichli funksiya, trigonometrik funksiya va boshqalar ko'rinishida bo'lishi mumkin.

Nazorat uchun savollar

1. Ekonometrika fanining maqsadi nimalardan iborat?
2. Ekonometrik modellashtirishning zarurligi?
3. Ekonometrikaning qo'llanish sohalarini tushuntirib bering?
4. Ekonometrik modellashtirish usullari tasnifi qanday?
5. Ekonometrik modellarni tuzish bosqichlarini aytib bering?
6. Iqtisodiy model so'zini tushintirib bering?
7. Iqtisodiy-matematik modellarga ta'rif bering?
8. «Model» tushunchasiga ta'rif bering?

[7]Gujarati D.N. Basic Econometrics. McGraw-Hill, 4th edition, 2003 (Gu), Inc.p. 29

II BOB. EKONOMETRIK MODELLARNING AXBOROT TA’MINOTI

2.1. Iqtisodiy ma’lumotlarning statistik tabiati
2.2. Bog‘liq va bog‘liq bo‘lmagan o‘zgaruvchilarni tanlash
2.3. Ekonometrik modellarni tuzishda qatnashadigan iqtisodiy ma’lumotlarga qo‘yiladigan talablar

2.1. Iqtisodiy ma’lumotlarning statistik tabiati

Iqtisodiy jarayonlarni vaqt davomida o‘zgarishini o‘rganish muhim ahamiyatga ega. Chunki barcha iqtisodiy jarayonlar va hodisalar vaqt davomida o‘zgaruvchan bo‘ladi. Iqtisodiyotda barcha iqtisodiy jarayonlarni iqtisodiy-statistik .modellar orqali o‘rganish natijasida u yoki bu iqtisodiy ko‘rsatkichning hozirgi holati va kelajakdagi o‘zgarishini ilmiy asosda tahlil qilish va prognozlash mumkin bo‘ladi.

Iqtisodiy-statistik modellashtirish iqtisodiy ko‘rsatkichlar va ishlab chiqarish omillari o‘rtasidagi aloqalar o‘z mohiyatiga ko‘ra stoxastik bo‘lgan asosga tayanadi. Iqtisodiy subyektlar faoliyatini statistik modellashtirish zamon va makonda ularning rivojlanish jarayonini o‘rganishda asosiy o‘rin egallaydi. Bu modellar ishlab chiqarish tendensiyalari va qonuniyatlarini aniqlash uchun moslashgandir.

Iqtisodiy-statistik modelashtirishni noaniq bo‘lishligining sabablari quyidagi hollarda sodir bo‘lishi mumkin:

1. Axborotli - axborotning xatoligi, uning ko‘rsatkichlari, omillar va obyektlar majmuining noaniqligi.

2. Tarkibiy - aniqlanmagan xilma-xilliklarning mavjudligi

3. Modelli - ko‘rsatkichlar va dalillar o‘rtasida bog‘lanish shakllaridan noto‘g‘ri foydalanish

Iqtisodiy-statistik kuzatuvlar olib borilganda, texnik-iqtisodiy ko‘rsatkichlar ko‘rinishidagi, materiallar oqimidagi axborotlarga duch kelamiz. Shu nuqtai nazardan, ishlab chiqarishga - kirish axborotini, chiqish axborotiga o‘zgartirgich sifatida qaraladi.

Ekonometrik modellarni tuzishda muhim bosqichlaridan biri modelda qatnashadigan omillar va ko‘rsatkichlarni tanlashdir.

Ko‘p hollarda o‘rganilayotgan ko‘rsatkichlarga juda ko‘p omillar ta’sir etmoqda. Shu jumladan, ularning hammasi modelda qatnashishi mumkin emas yoki iqtisodiy jihatdan maqsadga muvofiq emas.

Ko‘rsatkichlar va omillarni to‘liq qator sifatida quyidagicha tasvirlash mumkin:

$$y = f(/\, x_1,...,x_k \,/\, x_{(k+1)},...,x_m \,/\, x_{(m+1)},...,x_n)$$

1) Birinchi omillar guruhi $(x_1, \dots, x_k)$ – bu modelga kiritiladigan o‘zgaruvchilar;

2) Ikkinchi omillar guruhi $(x_{(k+1),}\ \dots, x_m)$ – modelda qatnashmaydi, lekin ulardan har biri tadqiqotchi tomonidan kuzatilayotgan statistik jamlanmada u yoki bu qiymatlarda nazorat qilinadi;

3) Uchinchi omillar guruhi $(x_{(m+1)}, \dots, x_n)$ – tasodifiy o‘zgaruvchilar, ular tadqiqotchi tomonidan nazorat qilinmaydi, lekin “y”ning o‘zgarishiga ta’sir etmoqda.

Agar birinchi guruhga soni bo‘yicha ko‘p bo‘lmagan, lekin “y” ning o‘zgarishiga kuchli ta’sir qilgan omillar kirsa, ushbu ekonometrik model ahamiyatli deb hisoblanadi.

Bundan tashqari, qolgan omillardan ko‘proq soni 2 chi guruhga va kamroq soni 3 chi guruhga kirgani maqsadga muvofiqdir.

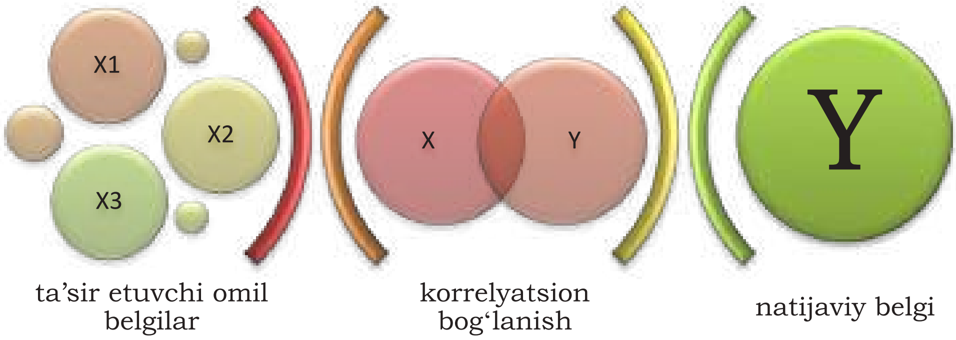
2.2. Bog‘liq va bog‘liq bo‘lmagan o‘zgaruvchilarni tanlash

Hodisalar orasidagi o‘zaro bog‘lanishlarni o‘rganish ekonometrika fanining muhim vazifasidir. Bu jarayonda ikki xil belgilar yoki ko‘rsatkichlar ishtirok etadi, biri erkli o‘zgaruvchilar, ikkinchisi erksiz o‘zgaruvchilar hisoblanadi. Birinchi toifadagi belgilar boshqalariga ta’sir etadi, ularning o‘zgarishiga sababchi bo‘ladi, shuning uchun ular omil belgilar deb yuritiladi, ikkinchi toifadagilar esa natijaviy belgilar deyiladi. Masalan, paxta yoki bug‘doyga suv, mineral o‘g‘itlar va ishlov berish natijasida ularning hosildorligi oshadi. Bu bog‘lanishda hosildorlik natijaviy belgi, unga ta’sir etuvchi kuchlar (suv, o‘g‘it, ishlov berish va h.k.) omil belgilardir.

Yoki, iste’molchining daromadi ortib borishi natijasida uning tovar va xizmatlarga bo‘lgan talabi oshadi. Bu bog‘lanishda talabning ortishi natijaviy belgi, unga ta’sir etuvchi omil, ya’ni daromad esa omil belgidir.

Omillarning har bir qiymatiga turli sharoitlarida natijaviy belgining har xil qiymatlari mos keladigan bog‘lanish korrelyatsion bog‘lanish yoki munosabat deyiladi. Korrelyatsion bog‘lanishning xarakterli xususiyati shundan iboratki, bunda omillarning to‘liq soni noma’lumdir. Shuning uchun bunday bog‘lanishlar to‘liqsiz hisoblanadi va ularni formulalar orqali taqriban ifodalash mumkin, xolos.

Umumiy holda qaralsa, korrelyatsion munosabatda erkin o‘zgaruvchi X belgining har bir qiymatiga ($x_i \; i=\overline{1...k}$) erksiz o‘zgaruvchi Y belgining ($y_j \; j=\overline{1..s}$) taqsimoti mos keladi. O‘z-o‘zidan ravshanki, bu holda ikkinchi Y belgining har bir qiymati (y_j) ham birinchi X belgining (x_i) taqsimoti bilan xarakterlanadi. Agar to‘plam hajmi katta bo‘lsa, belgi X va Y larning juft qiymatlari x_i va y_j ham ko‘p bo‘ladi va ulardan ayrimlari tez-tez takrorlanishi mumkin. bu holda korrelyatsion bog‘lanish kombinatsion jadval (korrelyatsiya to‘ri) shaklida tasvirlanadi.



Bog‘lanishlar to‘g‘ri chiziqli va egri chiziqli bo‘ladi. Agar bog‘lanishning tenglamasida omil belgilar (X_1, X_2, ..., X_K) faqat birinchi daraja bilan ishtirok etib, ularning yuqori darajalari va aralash ko‘paytmalari qatnashmasa, ya’ni $y_x = a_0 + \sum_{i=1}^{K} a_i X_i$ ko‘rinishda bo‘lsa, chiziqli bog‘lanish yoki xususiy holda, omil bitta bo‘lganda $y = a_0 + a_1 x$ to‘g‘ri chiziqli bog‘lanish deyiladi.[8]

Ifodasi to‘g‘ri chiziqli tenglama bo‘lmagan bog‘lanish egri chiziqli bog‘lanish deb ataladi. Xususan,

parabola $y = a_0 + a_1 x + a_2 x^2$

giperbola $y_x = a_0 + \dfrac{a_1}{x}$

darajali $y_x = a_0 x^a$ va boshqa ko‘rinishlarda ifodalanadigan bog‘lanishlar egri chiziqsiz bog‘lanishga misol bo‘la oladi.

2.3. Ekonometrik modellarni tuzishda qatnashadigan iqtisodiy ma’lumotlarga qo‘yiladigan talablar

Korrelyatsion va regression tahlilni qo‘llash vaqtida, omillarni tanlab olish va ulardan modellarda foydalanish hamda baholashdagi asosiy qoidalar quyidagilardan iborat:

---

[8]Gujarati D.N. Basic Econometrics. McGraw-Hill, 4th edition, 2003 (Gu),Inc.p. 10

1. Omillarni o‘rganish bilan qamrab olinadigan ro‘yxat chegaralangan, omillar esa nazariy asoslangan bo‘lishi lozim.

2. Modelga kiritilgan barcha omillar miqdor o‘zgarishlarga ega bo‘lishi kerak.

3. Tadqiq qilinayotgan to‘plam sifatli bir jinsli bo‘lishi lozim.

4. Omillar o‘zaro funksional bog‘lanmasliklari shart.

5. Kelajakda omillar o‘zaro ta’sirini ekstrapolyatsiya qilish uchun modellardan foydalanilayotgan vaqtda xarakter jiddiy o‘zgarmasligi, statistik mustahkam va barqaror bo‘lishi lozim.

6. Regression tahlilda har bir omilning (x) qiymatiga bir xil regressiyali natijaviy o‘zgaruvchi (y) taqsimoti normal yoki yaqin darajada mos kelish lozim.

7. O‘rganilayotgan omillar tadqiq etilgan, natijaviy ko‘rsatkichli, mantiqan davriy bo‘lishi lozim.

8. Natijaviy ko‘rsatkichga jiddiy ta’sir ko‘rsatadigan faqat muhim omillar ta’sirini ko‘rib chiqish lozim.

9. Regressiya tenglamalariga kiritilgan omillar soni katta bo‘lmasligi lozim. Chunki omillar sonining katta bo‘lishi, asosiy omillardan chetga olib kelishi mumkin. Omillar soni kuzatishlar sonidan 3-5 marta kam bo‘lishi kerak.

10. Regressiya tenglamasining omillari turli xil xatolar ta’sirida buzilishga olib keladigan xatoliklar bo‘lmasligi kerak. Omillar o‘rtasida funksional yoki shunga yaqin bog‘lanishlarning mavjudligi – multikollenearlik borligini ko‘rsatadi.

11. Kuzatuvlar sonini oshirish uchun ularning makonda takrorlanishidan foydalanish mumkin emas. Makonda hodisalarning o‘zgarishi avtoregressiyani vujudga keltirishi mumkin. Avtoregressiya esa statistikadagi mavjud o‘zgaruvchilar o‘rtasidagi bog‘lanishni ma’lum darajada buzadi. Shuning uchun ko‘rsatkichlar dinamik qatorlarida regression bog‘lanishni o‘rganish statistikadagi bog‘lanishni o‘rganishdan tubdan farq qiladi.

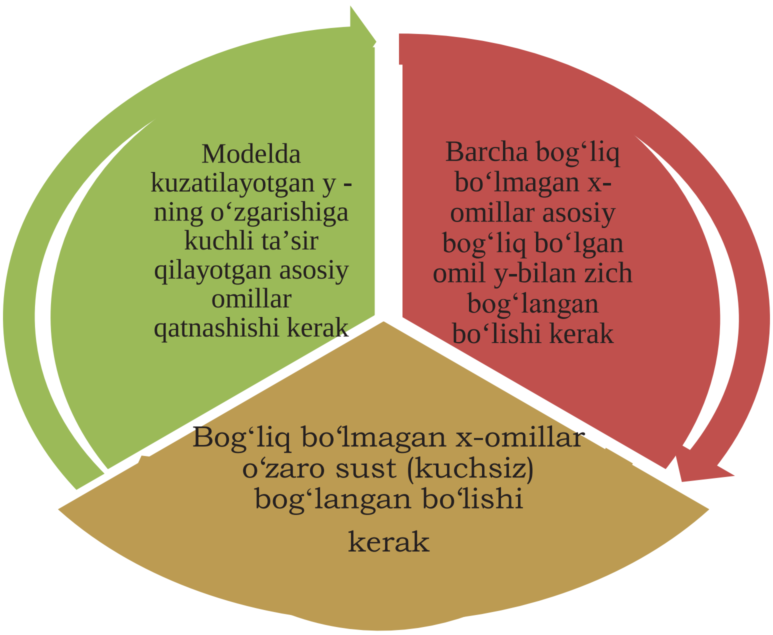
12. Har bir omil bo‘yicha taqsimot normal taqsimotga ega bo‘lishi shart emas. Bu regression tahlilni natijaviy, alomatli qiymat va tasodifsiz qiymatli omillar o‘rtasidagi bog‘lanishni ifodalovchi sifatida ta’riflashdan kelib chiqadi.

13. Omillarni natural birlikda o‘lchashda nisbiy qiymatlarga nisbatan ortiqroq ko‘rish lozim. Nisbiy qiymatlar o‘rtasidagi korrelyatsiya, regressiya tenglamasi parametrlari qiymati bog‘lanish mazmunini buzishi mumkin, omillar o‘rtasidagi bog‘lanishni ifodalovchi sifatida ta’riflashdan kelib chiqadi.

Iqtisodiy jarayonlar dinamikasini aks ettirish mohiyatiga ko‘ra, statik va dinamik modellar mavjud.

Statik modellar o‘zida vaqtning ayrim, qayd qilingan oralig‘ini qamrab oladi. Dinamik model vaqtning izchil oraliq tizimi holatini aks ettiradi. O‘zgaruvchan xarakterga ko‘ra, boshlang‘ich iqtisodiy ishlab chiqarish omillari yoki aralash omillarni o‘z ichiga olgan modellarni ko‘rsatish mumkin.

Demak, ekonometrik modellarga qo‘yiladigan asosiy talablar:



Ishlab chiqarishning boshlang‘ich omillari deganda, keyinchalik taqsimlab bo‘lmaydigan oddiy omillar, masalan, resurslar xarajati - jonli mehnat, vosita, mehnat qurollari tushuniladi. Modelning tuzilishiga qarab, ularni modelga turli o‘lchov birligi (natural, qiymat) va turli aniqlik darajasi bilan kiritish mumkin. Bunday holda ularning boshlang‘ich xarakteri saqlanadi.

Quyidagi modellar turi boshlang‘ich va ishlab chiqarish omillarining turli kombinatsiyalarini beradi:

❖ ishlab chiqarish natijalarining boshlang‘ich resurslar xarajati darajasi va tarkibiga hamda ishlab chiqarish ehtiyojlari sharoitiga bog‘liqligini xarakterlaydigan to‘liq modellar;

❖ ishlab chiqarish ehtiyojlari sharoiti obyektlari guruhi yoki vaqt bo‘yicha barqaror hisoblangan paytlarda qo‘llaniladigan “vazifalar - mahsulot ishlab chiqarish” modeli;

❖ ishlab chiqarish texnik-iqtisodiy ko‘rsatkichlar o‘rtasidagi o‘zaro va boshlang‘ich ishlab chiqarish omillari bilan aloqalarini xarakterlovchi turli xil modellar.

Modellar o‘zgaruvchanligiga ko‘ra, umumiy va xususiy modellarga bo‘linadi. Umumiy model o‘lchanadigan alomatlarning barchasini hamda o‘rganilayotgan ishlab chiqarish jarayonining bir tomonini, masalan, tabiiy sharoit belgilarini qisman o‘z ichiga oladi. Alomatlarning barchasini o‘z ichiga olgan model bilan xususiy (masalan, faqat tabiiy sharoit omillari) modelni taqqoslab, ishlab chiqarish tabiiy iqlim omillarining ta’siri qaysi vaqtda ko‘proq, qaysi vaqtda kamroq bo‘lishini aniqlash mumkin.

Umumiylik darajasi bo‘yicha iqtisodiy ko‘rsatkichlar avtonom tizimidagi farqlarni ajrata bilish lozim. Birinchi xil modellar mustaqil foydalanish, ikkinchi xil modellar esa qandaydir tizimdagi modellarning organik tarkibiy qismi hisoblanadi. va ularni qo‘llash xarakterini aniqlaydi.

Tasniflashning mana shu turiga modellarning bir sathli, pog‘onali va ko‘p sathli bo‘linishi ham kiradi. Ayrim hollarda ishlab chiqarish boshlang‘ich omillarining katta sonlarni hisobga olish va xususiy texnik-iqtisodiy ko‘rsatkichlar orqali ularni samaradorlikning umumiy sintetik ko‘rsatkichlariga ta’sirini tekshirish xususiyati bilan ikkinchi sxema ustun turadi.

Pog‘onali, ko‘p sathli modellar faqat turli darajadagi iqtisodiy aloqalarni aks ettirish uchun tuzilmay, balki turli davrlarga mansub bo‘lgan iqtisodiy ko‘rsatkichlarni modellashtirish bilan aniqlash uchun ham tuziladi.

Modellarni tuzilishi bo‘yicha tasniflash jarayonini modellar yordamida ifodalash va boshlang‘ich axborotdan foydalanish xarakteri alomati bo‘yicha tasniflashdan iborat. Birinchi xil alomat (belgi) bo‘yicha ikki xil statistik modellarni ko‘rsatish mumkin. Ular bashoratlarni tavsiflash va tushuntirish modellaridir.

Tavsiflash modellari - o‘zgaruvchan o‘zaro aloqalarni eng yaxshi tarzda tavsiflaydigan regressiyalarni tenglashtirish modeli hisoblanadi. Bunday hollarda modellar parametri mazmundor ma’noga ega bo‘lmaydi. Mazkur parametrlar qiymatini belgilashda approksimatsiya, ya’ni tavsiflanayotgan o‘zgaruvchan kirish bilan tavsiflanayotgan chiqish o‘rtasidagi statistik muvofiqlik barqarorlik vazifalari hal eiladi.

Tavsiflash modellarini tuzish paytida ko‘pincha belgilangan muddatdagi iqtisodiy ko‘rsatkichlarning aralashma faktlaridan foydalaniladi. Bunday hollarda ko‘rsatkichlar harakatidagi ketma-ketlik va aloqalar mavjudligi to‘g‘risidagi statistik ma’lumotlar tadqiqotchilarni qiziqtiradi.

Ko‘pincha tavsiflash modellarini tuzish vaqtida iqtisodiy ko‘rsatkichlarning aralash faktlaridan foydalaniladi. Bunday hollarda tadqiqotchilarni dalil sifatida tanlab olingan ko‘rsatkichlar funksiyalarning o‘zgarishiga sabab bo‘lgan yoki bo‘lmaganligi haqidagi statistik dalil qiziqtiradi. Tushuntirish - bashoratlash modelining nomi, uning milliy iqtisodiyotda qanday rol tutishini aniq tushuntiradi. Ular belgilangan faktlar majmui, gipotezalar o‘rtasidagi muvofiqlikni aniqlaydi. Bunday omillar - dalillarni taqqoslash asosida prognozlashtirilayotgan ko‘rsatkich shakllanish mexanizmini o‘rganish, ya’ni sanoat obyekti rivojlanishining harakatlantiruvchi kuchlarini aniqlash masalasi turadi.

Tushuntirish - bashoratlash modeli parametrlarini baholashda aynan tenglashtirish masalasi hal qilinadi. Masalaning mohiyati qandaydir to‘g‘ri keladigan statistik usullar yordamida chuqur ma’noli farazlar asosida tuzilgan tenglamalarning noma’lum parametrlarini qidirib topishdan iborat. Binobarin, identifikatsiya masalalarining approksimatsiya masalalaridan farqi shundaki, unda oldindan o‘zgaruvchan bog‘lanish tarkibi berilgan bo‘ladi.

Nazorat uchun savollar

1. Iqtisodiy ko‘rsatkichlarni qanday shakllarda namoyon etish mumkin?
2. Iqtisodiy ma’lumotlarni qayta ishlashning qanday usullarini bilasiz?
3. Talab va taklif modelida qaysi o‘zgaruvchi bog‘liq va qaysi o‘zgaruvchi bog‘liq emas?
4. Ekonometrik modellarni tuzishda qanday talablar qo‘yiladi?
5. Omillar o‘lchov birligini tanlashda qanday muammolarga duch kelinadi?
6. Ekonometrik modellarning qanday shakllari mavjud?

III BOB. EKONOMETRIKADA EHTIMOLLAR NAZARIYASI VA MATEMATIK STATISTIKANING ASOSIY TUSHUNCHALARI

3.1. Ehtimollar nazariyasi va matematik statistikaning asosiy tushunchalari

3.2. To‘plamlar va ularning xossalari

3.3. Diskret va uzluksiz tasodifiy miqdorlar

3.4. Tasodifiy miqdorlarning xarakteristikalarini hisoblash

3.1. Ehtimollar nazariyasi va matematik statistikaning asosiy tushunchalari

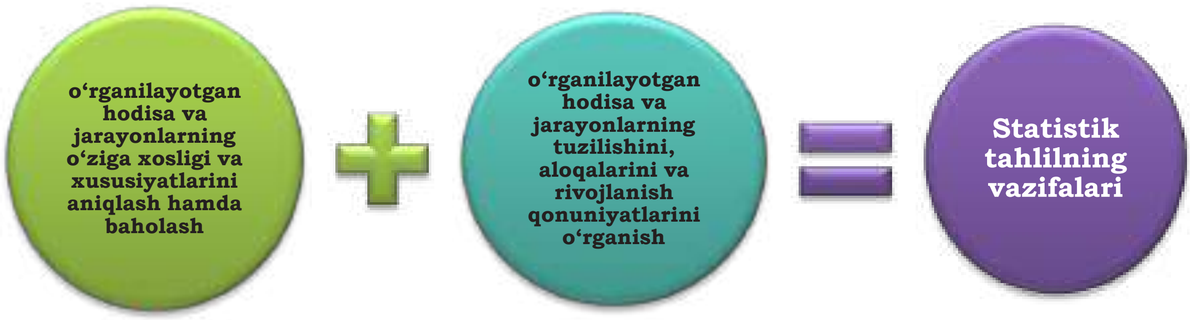
Statistik ma’lumotlarni tahlil qilish va umumlashtirish - bu statistik tadqiqotlarning yakuniy bosqichi bo‘lib, uning pirovard maqsadi o‘rganilayotgan ijtimoiy-iqtisodiy hodisalar va jarayonlarning tendensiyalari va qonuniyatlari to‘g‘risida nazariy xulosalar va amaliy xulosalar olishdir.

Ehtimollar nazariyasida nisbiy chastota, ehtimollik tushunchasi, chegaraviy va shartli ehtimollik tushunchalarini bilish, Bayes teoremasi haqida axborotga ega bo‘lish zarur.

Tahlil - bu obyektni individual jihatlari va tarkibiy qismlarini hisobga olgan holda ilmiy tadqiq etish usuli.

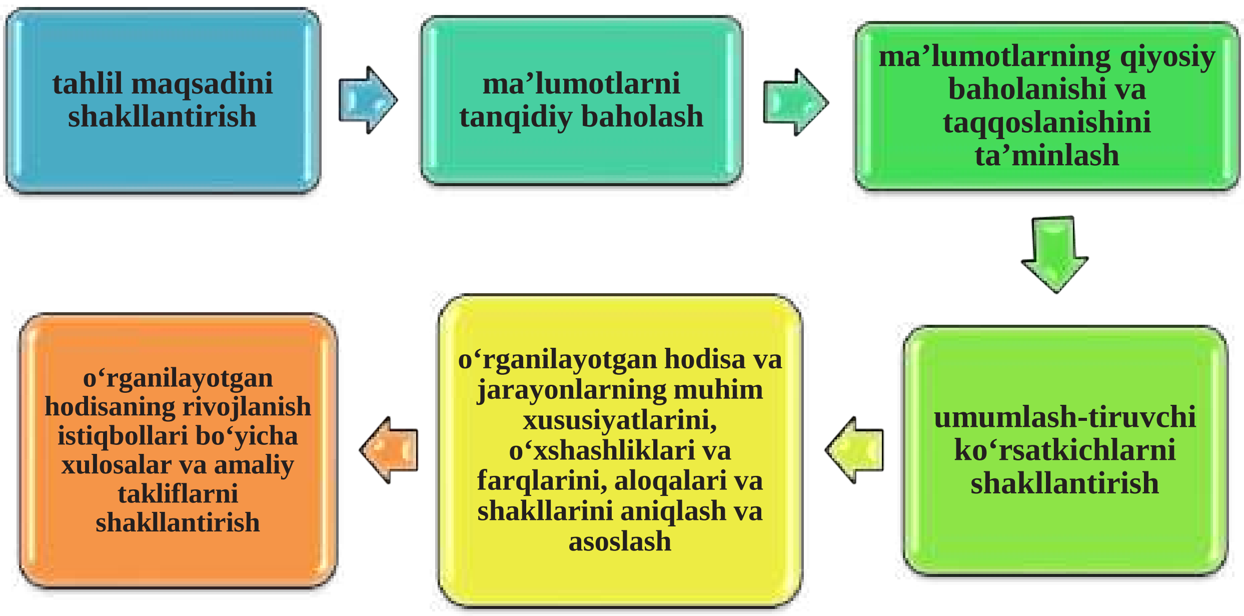
Iqtisodiy va statistik tahlil - bu o‘rganilayotgan hodisa va jarayonlarning yetarli darajada aks etishini nazorat qilish maqsadida an’anaviy statistik va matematik usullardan keng foydalanishga asoslangan metodologiyani ishlab chiqish.

Tahlil usullari o‘rganilayotgan jarayonlarning tabiati, ularning o‘ziga xosligi, xususiyatlari va namoyon bo‘lish shakllariga qarab o‘zgarishi kerak.



Ma’lumotlarning statistik tahlili o‘rganilayotgan hodisalarning mohiyatini nazariy, sifatli tahlil qilish va ularning tuzilishi, aloqalari va dinamikasini o‘rganish uchun mos keladigan miqdoriy vositalar bilan uzviy bog‘liqlikda amalga oshiriladi.

Statistik tahlil bosqichlari:



Matematik statistika - bu matematikaning ilmiy va amaliy xulosalar uchun tizimlashtirish, qayta ishlash va statistik ma’lumotlarni ishlatishning matematik usullariga bag‘ishlangan bo‘limi.

Matematik statistika - bu eksperimental ma‘lumotlarni qayta ishlash usullari bilan shug‘ullanadigan fan. Har qanday fan murakkabligi va ahamiyatini oshirish maqsadida quyidagi vazifalarni hal qiladi:

1) hodisaning tavsifi;
2) tahlil qilish va prognoz qilish;
3) optimal yechimni izlash.

Matematik statistika ham bunday masalani hal qiladi:

1) olingan statistik materiallarni tizimlashtirish;
2) olingan eksperimental ma’lumotlar asosida kuzatilgan tasodifiy o‘zgaruvchisining raqamli xususiyatlarini baholash;
3) minimal o‘lchov xatolar bilan ishonchli natijalarni olish uchun yetarli bo‘lgan tajribalar sonini aniqlash.

Tasodifiy o‘zgaruvchilar ehtimollik taqsimoti va ehtimollik kutulish normal taqsimotining xossalarini va binominal taqsimotni bilish zarur.

Statistik tanlovda oddiy tanlama usulini bilish yetarli.

Baholash xususida uning usullarini, dispersiya, dispersiyani hisoblash va xatosiz ma’lumotlarni bilish kerak.

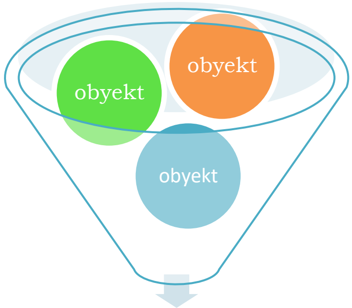
Statistik xulosa qilish uchun t- va F testlarni o‘tkazishni, ishonchlilik intervalini, taxminlar ma’nosi va ahamiyatini bilish kerak.

Asosiy statistik ko‘rsatkichlar 2 guruhga bo‘linadi: o‘rtacha darajasini o‘lchaydigan va dispersiyani o‘lchaydigan.

Matematik statistika barcha obyektlarni tahlil qilmaydi, ammo faqat, bunday obyektlarning umumiy xarakteristikalari asosida yaratilgan katta guruhdan tanlangan bir nechtasi. Matematik statistikadagi ushbu hodisa tahlilning tanlanma usuli deb ataladi.

3.2. To‘plamlar va ularning xossalari

Statistikada *to‘plam* iborasi juda keng qo‘llaniladi. To‘plam, agar uning obyektlarining bir yoki bir nechta o‘rganilgan muhim xususiyatlari barcha birliklar uchun umumiy bo‘lsa, bir xil deb nomlanadi.



To‘plam hajmi

Turli xil hodisalarni o‘z ichiga olgan jami, bir jinsli emas deb hisoblanadi. Jami bir jihatdan bir xil, ikkinchisida turli xil bo‘lishi mumkin. Har bir alohida holatda, to‘plamning bir xilligi o‘rganilayotgan ijtimoiy hodisaning mazmunini oydinlashtirgan holda sifatli tahlilni o‘tkazish yo‘li bilan o‘rnatiladi.

Tanlanma to‘plam, yoki oddiy qilib, tanlanma deb tasodifiy ravishda tanlab olingan obyektlar to‘plamiga aytiladi.

Bosh to‘plam deb tanlanma ajratilgan obyektlar to‘plamiga aytiladi.

Masalan, 1000 ta detaldan tekshirish uchun 100 ta detal olingan bo‘lsa, u holda bosh to‘plam hajmiga N=1000, tanlanma hajmi esa n=100.

To‘plamning quyidagi turlari mavjud:

- asosiy
- tanlama
- cheklangan
- cheksiz

Bosh to‘plam ko‘pincha *chekli* sondagi elementlarni o‘z ichiga oladi. Ammo bu son ancha katta bo‘lsa, u holda hisoblashlarni soddalashtirish yoki nazariy xulosalarni ixchamlash maqsadini ko‘zda tutib, ba’zan bosh to‘plam *cheksiz* ko‘p sondagi obyektlardan iborat deb faraz qilinadi. Bunday yo‘l qo‘yish shu bilan oqlanadaki bosh to‘plam hajmini orttirish tanlanma ma’lumotlarini ishlab chiqish natijalariga amalda ta’sir etmaydi.

Matematik statistikaning asosiy vazifasi tanlanmalar uchun bosh to‘plamning xususiyatlarini baholashdir.

Butun bosh to‘plamning haqida biz, qoida tariqasida, hech narsani aniq bilmaymiz va faqat taxminlar-farazlar qilishimiz mumkin. Gipotezalarimizni sinab ko‘rish uchun biz umumiy populyatsiyadan mustaqil namunani tekshiramiz va noma’lum nazariy parametrlarning namunaviy baholari asosida tuzamiz.

To‘plam birligi – kuzatish talab etiladigan element.

Belgi – to‘plam birligining belgilar turlari:

Variatsiya – belgining o‘zgarishidir.

Variant – o‘zgaruvchi belgining konkret ifodasi. Variantlar lotin harflarida belgilanadi. Masalan:

X1, X2,…,Xk

Y1, Y2,….,Yk

O‘zgaruvchi belgining miqdorlari majmuasi *variatsion qator* deb ataladi.

Agar variantlarni ko‘payish yoki kamayish bo‘yicha joylashtirsak, *tartibli variatsion qatorni* tuzamiz.

3.3. Diskret va uzluksiz tasodifiy miqdorlar

20 ta talabalar ichida o‘g‘il bolalar soni 0,1,2,...,20 qiymatlarni qabul qilishi mumkin bo‘lgan tasodifiy miqdordir.

Uzluksiz tasodifiy miqdor deb chekli yoki cheksiz oraliqdagi barcha qiymatlarni qabul qilishi mumkin bo‘lgan miqdorga aytiladi.

Tasodifiy miqdor **- avvaldan noma'lum bo'lgan va oldindan inobatga olib bo'lmaydigan tasodifiy sabablarga bog'liq bo'lgan hamda sinash natijasida bitta mumkin bo'lgan qiymat qabul qiluvchi miqdorga aytiladi.**

Diskret (uzlukli) tasodifiy miqdor **- ayrim, ajralgan qiymatlarni ma'lum ehtimollar bilan qabul qiluvchi miqdorga aytiladi. Diskret tasodifiy miqdorning mumkin bo'lgan qiymatlari soni chekli yoki cheksiz bo'lishi mumkin.**

To'pdan otilgan snaryadning uchib o'tgan masofasi tasodifiy miqdordir. Bu miqdorning mumkin bo'lgan qiymatlari [a, b] oraliqqa tegishlidir.

Tasodifiy o'zgaruvchi taqsimot qonuni bilan berilgan.

Taqsimot qonuni → tasodifiy o'zgaruvchining barcha mumkin bo'lgan qiymatlari va ularning ehtimolliklari o'rtasidagi bog'liqlik.

Diskret tasodifiy o'zgaruvchining taqsimot qonuni bu mumkin bo'lgan qiymatlar va ularning ehtimolliklari o'rtasidagi moslikdir.

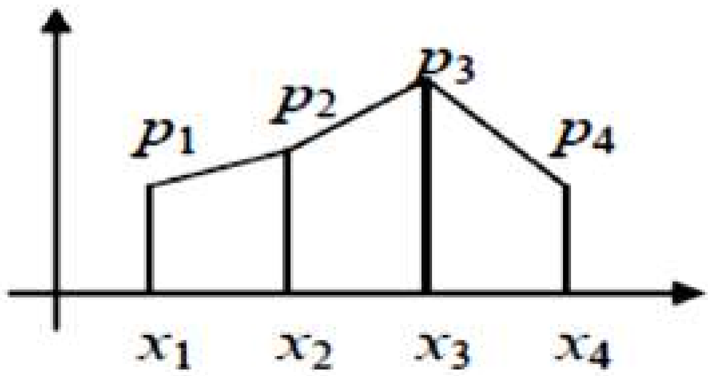
Jadvalda:

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

Analitik:

$$p_1 + p_2 + \ldots + p_n = 1$$

Grafik jihatdan: taqsimot ko'pburchagi



Diskret tasodifiy miqdorning matematik kutilishi deb, uning barcha mumkin bo‘lgan qiymatlarini mos ehtimollarga ko‘paytmalari yig‘indisiga aytiladi:

$$M(X)=x_1p_1+ x_2p_2+\ldots+x_np_n=\sum_{i=1}^{n} x_i\, p_i$$

Matematik kutilishning xossalari.

1. O‘zgarmas miqdorning matematik kutilishi shu o‘zgarmasning o‘ziga teng:

$$M(C)=C$$

2. O‘zgarmas ko‘paytuvchini matematik kutilishi belgisidan tashqariga chiqarish mumkin:

$$M(CX)=CM(X)$$

3. Ikkita erkli *X* va *Y* tasodifiy miqdorlar ko‘paytmasining matematik kutilishi ularning matematik kutilishlari ko‘paytmasiga teng:

$$M(XY)=M(X)M(Y)$$

4. Ikkita tasodifiy miqdor yig‘indisining matematik kutilishi qo‘shiluvchilarning matematik kutilishlari yig‘indisiga teng:

$$M(X+Y)=M(X)$$

Mumkin bo‘lgan qiymatlari [*a*, *b*] segmentiga tegishli bo‘lgan uzluksiz *X* tasodifiy o‘zgaruvchining matematik kutilishi aniq integral hisoblanadi:

$$M(X)=\int_a^b xf(x)dx$$

Agar tasodifiy o‘zgaruvchining mumkin bo‘lgan qiymatlari butun *Ox* o‘qi bo‘ylab taqsimlansa, u holda

$$M(X)=\int_{-\infty}^{+\infty} xf(x)dx$$

Dispersiya tasodifiy o‘zgaruvchining kuzatilgan qiymatlarining uning matematik kutilish atrofida tarqalish darajasini tavsiflaydi. Tasodifiy o‘zgaruvchining dispersiyasi tasodifiy o‘zgaruvchining o‘rtacha qiymatidan chetlanishini tavsiflaydi va diskret tasodifiy o‘zgaruvchi uchun quyidagi formula bo‘yicha topiladi:

$$D(X)=M[X-M(X)]^2$$

Dispersiyaning ikkinchi formulasi mavjud bo‘lib, uni qo‘lda hisoblash uchun ishlatish qulayroq

$$D(X)=M(X^2)-(M(X))^2$$

Uzluksiz X tasodifiy o‘zgaruvchining dispersiyasi quyidagi formula bo‘yicha topiladi:

$$D(X)=\int_a^b x^2 f(x)dx-[M(X)]^2$$

$$D(X)=\int_a^b x^2 f(x)dx-[M(X)]^2$$

Dispersiyaning xossalari.

1. O‘zgarmas miqdorning dispersiyasi nolga teng:

$$D(C)=0$$

2. O‘zgarmas ko‘paytuvchini kvadrati yordamida dispersiyasi belgisidan tashqariga chiqarish mumkin:

$$D(CX)=C^2D(X)$$

3. Ikkita tasodifiy miqdor yig‘indisining dispersiyasi qo‘shiluvchilarning dispersiyasi yig‘indisiga teng:

$$D(X+Y)=D(X)+D(Y)$$

Matematik kutilishning o‘lchami tasodifiy o‘zgaruvchining o‘zi bilan bir xil.

Dispersiyaning o‘lchami tasodifiy o‘zgaruvchining o‘lchovi kvadratiga teng.

Tasodifiy o‘zgaruvchining o‘zi bilan bir xil birliklarda tarqalish darajasini taxmin qilish uchun uchinchi raqamli xususiyat kiritiladi.

Tasodifiy o‘zgaruvchini ko‘rsatishning turli usullari mavjud:

- taqsimot funksiyasi - har qanday tasodifiy o‘zgaruvchilar uchun
- taqsimot qatori - diskret miqdorlar uchun
- ehtimollik zichligi - uzluksiz miqdorlar *r* uchun

Tasodifiy o‘zgaruvchining o‘rtacha kvadratik farqi (standart chetlanishi) – bu dispersiyaning kvadrat ildizi bo‘ladi:

$$\sigma = \sqrt{D(x)}$$

Dispersiya singari, standart chetlanishi ham tasodifiy o‘zgaruvchining kuzatilgan qiymatlarining uning matematik kutilish atrofida tarqalish darajasini tavsiflaydi.

Ammo bu holda σ ning o‘lchami tasodifiy o‘zgaruvchining o‘lchamiga teng bo‘ladi.

3.4. Tasodifiy miqdorlarning xarakteristikalarini hisoblash

O‘rtacha qiymatlar ommaviy ijtimoiy hodisalarning xulosaviy (yakuniy) xarakteristikasini ta‘minlovchi umumlashtiruvchi statistik ko‘rsatkichlarga ishora qiladi, chunki ular turli xil atributlarning ko‘p sonli individual qiymatlari asosida tuzilgan.

Statistikada ikkita katta sinfga bo‘linadigan har xil o‘rtacha ko‘rsatkichlar qo‘llaniladi:

Statistikada ikkita katta sinfga bo‘linadigan har xil o‘rtacha ko‘rsatkichlar qo‘llaniladi:	
darajali o‘rtachalar (o‘rtacha garmonik, o‘rtacha geometrik, o‘rtacha arifmetik, o‘rtacha kvadratik, o‘rtacha kubik)	tarkibiy o‘rtachalar (moda, mediana)

Darajali o‘rtacha qiymatlarini hisoblash uchun siz mavjud bo‘lgan barcha belgi qiymatlaridan foydalanishingiz kerak.

Moda va medianalar faqat taqsimot tuzilishi bilan belgilanadi, shuning uchun ular tarkibiy, pozitsion o‘rtacha ko‘rsatkichlar deyiladi. Moda va medianalar ko‘pincha o‘rtacha ko‘rsatkich sifatida darajali o‘rtachalarni hisoblash imkonsiz yoki amaliy bo‘lmagan to‘plamlarda qo‘llaniladi.

O‘rtachaning eng keng tarqalgan turi - bu o‘rtacha arifmetikdir.

Oddiy arifmetik o‘rtacha bilan birga o‘rtacha vaznli (torttirilgan) arifmetik o‘rtacha ham o‘rganiladi.

***Arifmetik o‘rtacha*:**

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Chastota (m) - absolyut miqdor bo‘lib, har variantning to‘plamda necha bor uchrashuvini ko‘rsatadi.

Chastotaning nisbiy ko‘rinishi *chastota ulushi* deb ataladi.

$$w_i = \frac{m_i}{\sum_{i=1}^{n} m_i}, \qquad \sum_{i=1}^{n} w_i = 1$$

$$\sum w_i \cdot 100 = 100\%$$

Tanlanmaning statistik taqsimoti deb variantalar va ularga mos chastotalar yoki nisbiy chastotalar ro‘yxatiga aytiladi.

***Variatsiya chegarasi* (*R*)** - variatsion qatorning ekstremal qiymatlari farqiga aytiladi.

$$R = X_{\max} - X_{\min}$$

O‘rtacha chiziqli farq (ρ)**:**

$$\rho = \frac{\sum\left|X - \overline{X}\right|}{n} \qquad \text{(torttirilmagan),}$$

$$\rho = \frac{\sum\left|X - \overline{X}\right| \cdot m}{\sum m} \qquad \text{(torttirilgan)}$$

Dispersiya (σ^2) - variantlarning arifmetik o‘rtachadan farqlarining o‘rtacha kvadrati.

$$\sigma^2 = \frac{\sum\left(X - \overline{X}\right)^2}{n} \qquad \text{(torttirilmagan),}$$

$$\sigma^2 = \frac{\sum\left(X - \overline{X}\right)^2 \cdot m}{\sum m} \qquad \text{(torttirilgan)}$$

O‘rtacha kvadratik farq (σ) - belgining o‘zgarishini ifodalaydi va quyidagicha hisoblanadi:

$$\sigma = \sqrt{\frac{\sum\left(X - \overline{X}\right)^2}{n}} \qquad \text{- (torttirilmagan),}$$

$$\sigma = \sqrt{\frac{\sum\left(X - \overline{X}\right)^2 \cdot m}{\sum m}} \qquad \text{- (torttirilgan)}$$

Variatsiya koeffitsiyenti *(V)* - nisbiy ko‘rsatkich bo‘lib, belgining o‘zgarishini ifodalaydi va protsentlarda ifodalanadi.

$V_R = \frac{R}{\overline{X}} \cdot 100\%$ - variatsiya chegarasi bo‘yicha variatsiya koeffitsiyenti, *ossillyatsiya koeffitsiyenti*.

$V_\rho = \frac{\rho}{\overline{X}} \cdot 100\%$ - o‘rtacha chiziq farq bo‘yicha variatsiya koeffitsiyenti.

$V_\sigma = \frac{\sigma}{\overline{X}} \cdot 100\%$ - kvadrat farq bo‘yicha variatsiya koeffitsiyenti.

Moda M_0 deb eng katta chastotaga ega bo‘lgan variantaga aytiladi. Masalan, ushbu

variant	1	4	7	9
chastota	5	1	20	6

qator uchun moda 7 ga teng.

Mediana M_e deb variatsion qatorni variantalar soni teng bo‘lgan ikki qismga ajratadigan variantaga aytiladi. Agar variantalar soni toq, ya'ni $n=2k+1$, bo'lsa, u holda $M_e = X_{k+1}$; n juft, ya'ni $n=2k$ da mediana:

$$M_e = \frac{X_k + X_{k+1}}{2}$$

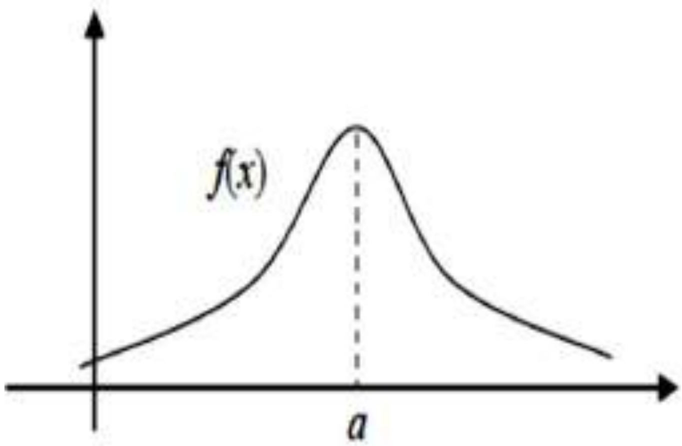
Ko'pgina iqtisodiy ko'rsatkichlar taqsimotning normal yoki normal holatiga yaqin. Masalan, aholi daromadlari, sohadagi firmalarning foydasi, iste'mol hajmi va boshqalar normal taqsimotga ega.

Normal taqsimot qonuni (Gauss qonuni) bu amaliyotda eng keng tarqalgan taqsimot qonuni bo'lib, u jarayonlar va tizimlarning asosiy xarakteristikalarining tasodifiy buzilishlarini va og'ishlarini, o'lchov xatolarini va boshqalarni tavsiflaydi.

Normal taqsimot deb (3.1-rasm) differensial funksiya bilan tavsiflanadigan uzluksiz tasodifiy miqdor taqsimotiga aytiladi (a- normal taqsimotning matematik kutilishi, σ- o'rtacha kvadratik chetlanishi).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

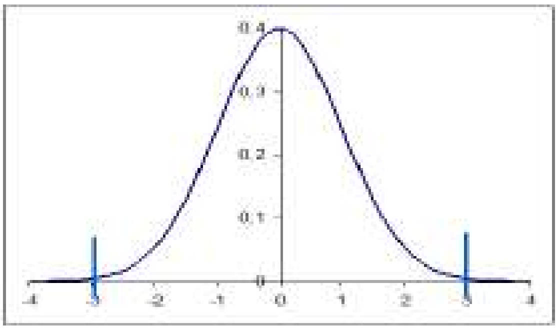
$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(x-a)^2}{2\sigma^2}} dx$$



3.1-rasm. Normal taqsimot grafigi

«Uchta sigma» (3.2-rasm) qoidasida shunday deyilgan: agar tasodifiy o'zgaruvchining normal taqsimoti bo'lsa, u holda uning matematik kutilishdan chetlanishining absolyut qiymati 0,9973 ehtimollik bilan standart chetlanishining uch baravaridan oshmaydi.

Agar tasodifiy o‘zgaruvchining taqsimoti qonuni noma’lum bo‘lsa va faqat matematik kutilishi va standart chetlanishi ma’lum bo‘lsa, amalda segment odatda tasodifiy o‘zgaruvchining amaldagi mumkin bo‘lgan qiymatlari bo‘limi sifatida qaraladi.



3.2-rasm. «Uchta sigma» qoidasi

Nazariy taqsimot assimetriyasi deb uchinchi tartibli markaziy momentning o‘rta kvadratik chetlanish kubi nisbatiga aytiladi:

$$A_s = \frac{\mu_3}{\sigma^3}$$

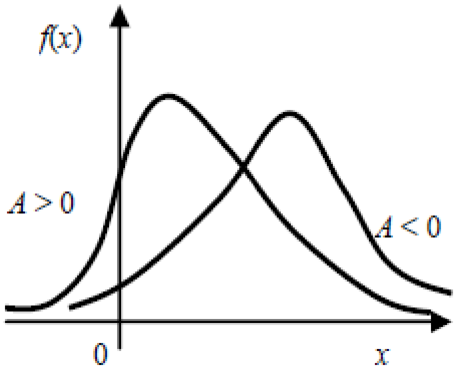
Nazariy taqsimot ekssesi deb

$$E_k = \frac{\mu_4}{\sigma^4} - 3$$

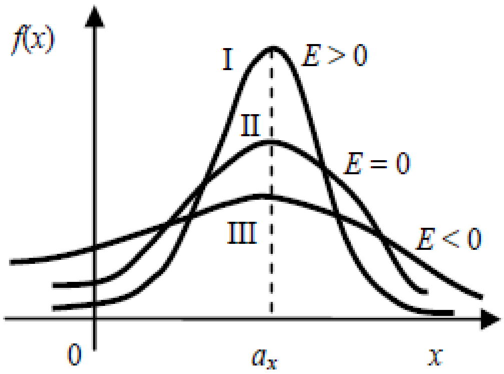
tenglik bilan aniqladigan xarakteristikaga aytiladi.

Agar eksses musbat bo‘lsa, u holda egri chiziq normal egri chiziqqa qaraganda balandroq va «o‘tkirroq» uchga ega bo‘ladi, agar eksses manfiy bo‘lsa, u holda taqqoslanayotgan egri chiziq normal egri chiziqqa qaraganda pastroq va «yassiroq» uchga ega bo‘ladi.

3.3-rasmda ikkita taqsimot egri chizig‘i ko‘rsatilgan. Ulardan biri musbat (o‘ng tomonlama) asimmetriyaga ega (A>0), ikkinchisi egri chizig‘i - manfiy (chap tomonli) (A<0).



3.3-rasm. Nazariy taqsimot asimmetriyasi



3.4-rasm. Nazariy taqsimot ekssessi

3.4-rasmda odatdagidan yuqori cho‘qqilarga ega bo‘lgan egri chiziqlar musbat ekssessiga, ko‘proq tekis tepaliklar manfiy ekssessiga ega.

Nazorat uchun savollar

1. Tasodifiy miqdorlarning qanday turlarini bilasiz?
2. Tanlama deganda nimani tushunasiz?
3. Diskret va uzluksiz tasodifiy miqdorlarga misol keltiring?
4. Tasodifiy miqdorning asosiy statistik xarakteristikalarini aytib bering?
5. Dispersiya nimani ko‘rsatadi?
6. Kovariatsiya koeffitsiyenti qanday hisoblanadi?
7. Ekssessning musbatligi yoki manfiyligi nimalarni bildiradi?
8. Moda va mediana nima uchun hisoblanadi?
9. Nazariy taqsimot asimmetriyasi deganda nimani tushunasiz?
10. «Uchta sigma» qoidasi nimalarni bildiradi?

Masala. Variatsion qator uchun aniqlansin:

Variant, X	1	4	7	9
Chastota, Y	5	20	20	6

1. Arifmetik o‘rtacha
2. O‘rtacha chiziqli farq
3. Dispersiya
4. O‘rtacha kvadratik farq
5. Moda
6. Mediana

Yechilishi:

T	Chastota, Y	Variant, X	Y^2	X^2	Y*X	$(X-X_{o‘rt})^2$	$(Y-Y_{o‘rt})^2$
1	5	1	25	1	5	18,06	60,06
2	20	4	400	16	80	1,56	52,56
3	20	7	400	49	140	3,06	52,56
4	6	9	36	81	54	14,06	45,56
Jami	**51**	**21**	**861**	**147**	**279**	**36,75**	**210,75**
O‘rtacha	**12,75**	**5,25**	**215,25**	**36,75**	**69,75**	**9,1875**	**52,6875**

1. **Arifmetik o‘rtacha:** $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$;

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{4}\sum_{1}^{4}(1+4+7+9) = \frac{1}{4}21 = 5{,}25$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{4}\sum_{1}^{4}(5+20+20+6) = \frac{1}{4}51 = 12{,}75$$

2. **O‘rtacha chiziqli farq:** $\rho = \frac{\sum_{i=1}^{n}\left|X_i - \bar{X}\right|}{n}$;

$$\rho = \frac{\sum_{i=1}^{n}\left|X_i - \bar{X}\right|}{n} = \frac{|1-5{,}25| + |4-5{,}25| + |7-5{,}25| + |9-5{,}25|}{4} = \frac{11}{4} = 2{,}25$$

$$\rho = \frac{\sum_{i=1}^{n}\left|Y_i - \bar{Y}\right|}{n} = \frac{|5-12{,}75| + |20-12{,}75| + |20-12{,}75| + |6-12{,}75|}{4} = \frac{29}{4} = 7{,}25$$

3. **Dispersiya** σ^2- variantlarning arifmetik o‘rtachadan farqlarining o‘rtacha kvadrati: $\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$;

$$\sigma_x^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 = 36{,}75 \quad \sigma_y^2 = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 210{,}75$$

4. Variantlar oralig‘i uchun dipersiya namunasi:

$$\sigma_x^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{36{,}75}{4-1} = 12{,}25 \quad \sigma_y^2 = \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n-1} = \frac{210{,}75}{3} = 70{,}25$$

5. O‘rtacha kvadratik farq (standart og‘ish) (σ) - belgining o‘zgarishini ifodalaydi va quyidagicha hisoblanadi: $\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}}$;

$$\sigma_x = \sqrt{12{,}25} = 3{,}5 \quad \sigma_y = \sqrt{70{,}25} = 8{,}38$$

6. *Moda* - M_0 deb eng katta chastotaga ega bo‘lgan variantaga aytiladi. Masalan, ushbu moda – to‘plamda songa yoki salmoqqa ega bo‘lgan ko‘rsatkich. U oraliq va oraliq bo‘lmagan qatorlar uchun aniqlanishi mumkin. Oraliq qatorlarda modani hisoblash formulasi quyidagicha

$$M_0 = X_0 + d \cdot \frac{(f_2 - f_1)}{(f_2 - f_1) + (f_2 - f_3)}$$

Bu yerda M_0 **- moda;** X_0 **- moda oralig‘ining quyi chegarasi;**

d - moda oralig‘i kattaligi; f_1 - moda oralig‘ining quyi chegarasidagi vazn;

f_2 - modani o‘z ichiga olgan oraliqning vazn (varianti, uchrashish tezligi);

f_3 - moda oralig‘ining yuqori chegaradagi vazn.

Variant, X	1	4	7	9
Chastota, Y	5	20	20	6

Chastota uchun Y_{moda}=20 ga teng.

7. *Mediana* - M_e deb to‘plamni teng ikkiga bo‘luvchi ko‘rsatkich. U quyidagi formula yordamida topiladi. $M_e = X_0 + d\frac{\sum\frac{f}{2} - S_{m-1}}{f_m}$

Bu yerda: M_e - mediana; X_0 - mediana oralig‘ining quyi chegarasi;

d - mediana oralig‘i; $\sum f$ - variantlar soni yig‘indisi;

S_{m-1} - mediana oralig‘idan oldingi oraliqlar;

f_m - Medinani o‘z ichiga olgan oraliq vazni.

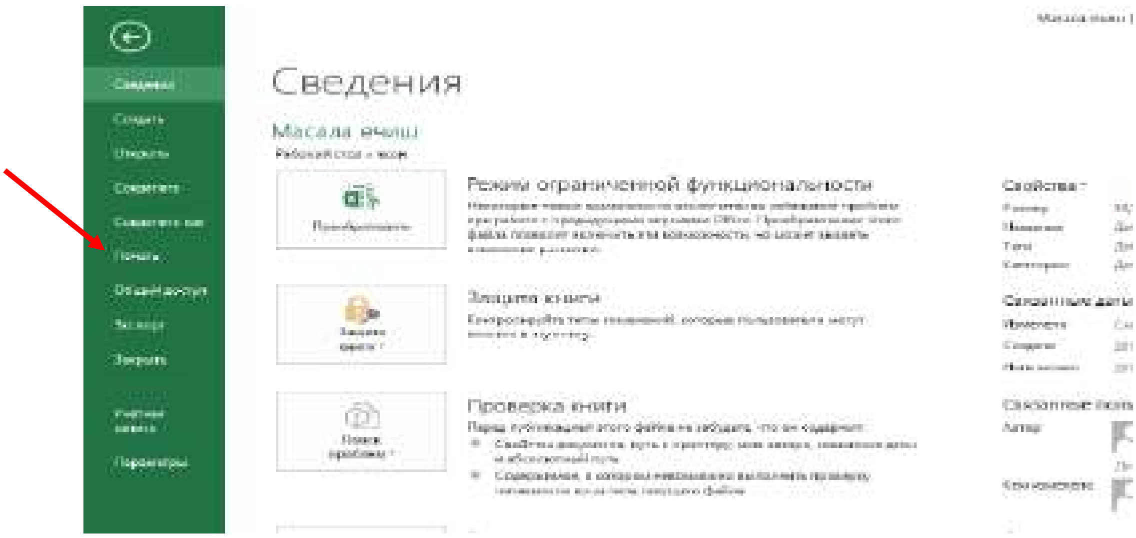
Agar variantalar soni toq, ya'ni $n = 2k + 1$, bo‘lsa, u holda $M_e = X_{k+1}$; n juft, ya'ni $n = 2k$ da mediana:

$M_e = \dfrac{X_k + X_{k+1}}{2}$ n=4 bo‘lgani uchun k=2bo‘ladi, bundan:

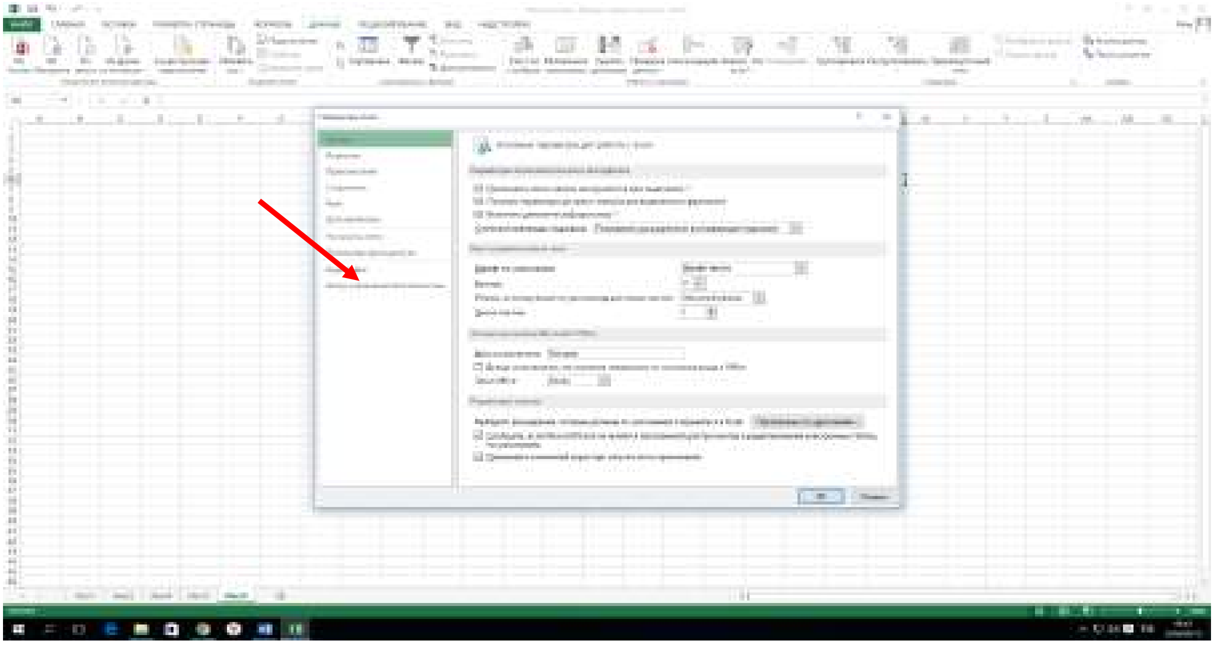
$$M_{x_e} = \frac{X_2 + X_3}{2} = \frac{4+7}{2} = \frac{11}{2} = 5{,}5 \quad M_{y_e} = \frac{Y_1 + Y_3}{2} = \frac{5+20}{2} \approx 13$$

Ushbu topshiriqni bir vaqtning o‘zida qisqa muddat ichida kompyuter texnologiyasidan foydalanib aniqlash mumkin. Bu o‘z navbatida vaqtning tejalashi va natijalar aniqligiga erishish imkonini beradi. Buning uchun kompyuterning EXCEL dasturi ochiladi.

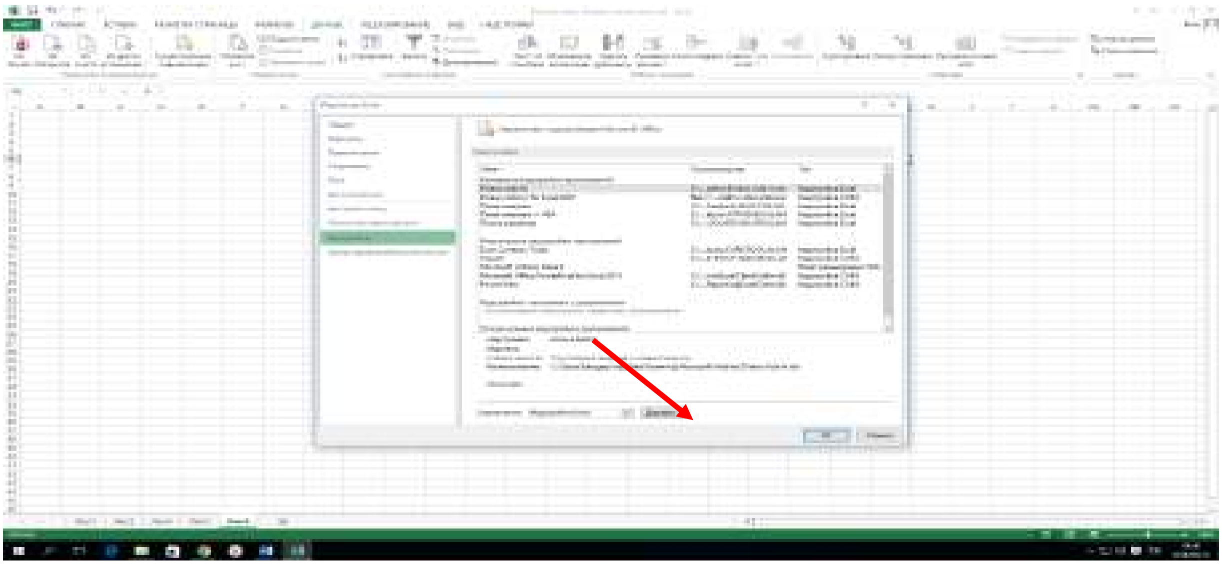
Ko‘pgina kompyuterlarda foydalanilmaganligi uchun tahlil paketlari o‘rnatilmagan bo‘ladi. Shuning uchun EXCEL dasturi ochilgandan keyin asosiy menyudan <fayl> ni bosamiz. Natijada quyida keltirilganidek, ma'lumotlar oynasi ochiladi.



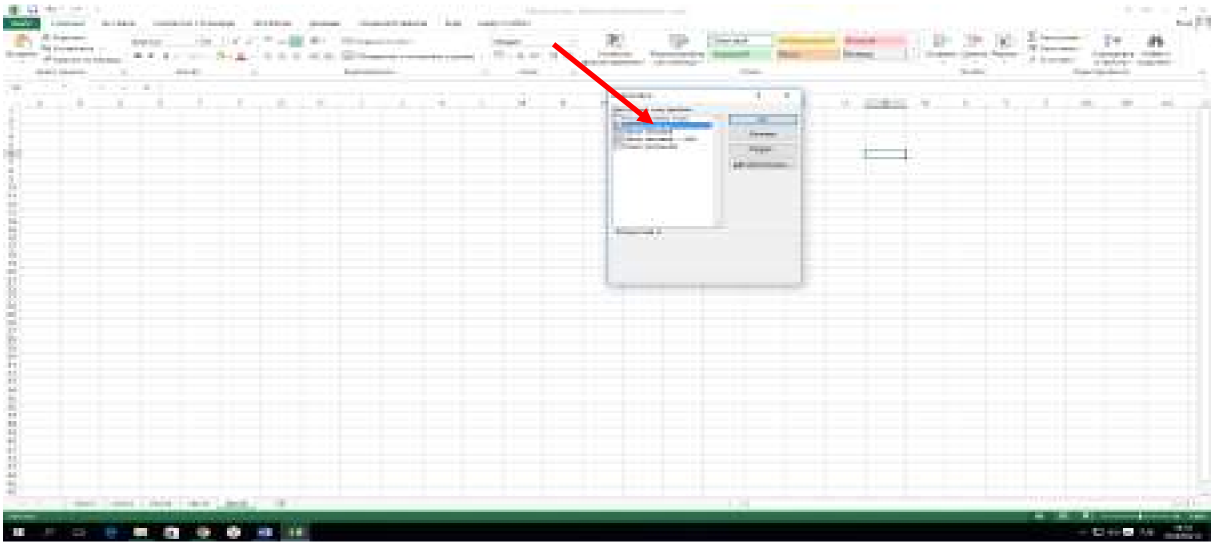
Ma'lumotlar oynasida keltirilgan instrumentlardan <parametri> tugmachasini bosamiz va quyidagi keltirilgan EXCEL parametrlari oynasi hosil bo‘ladi.



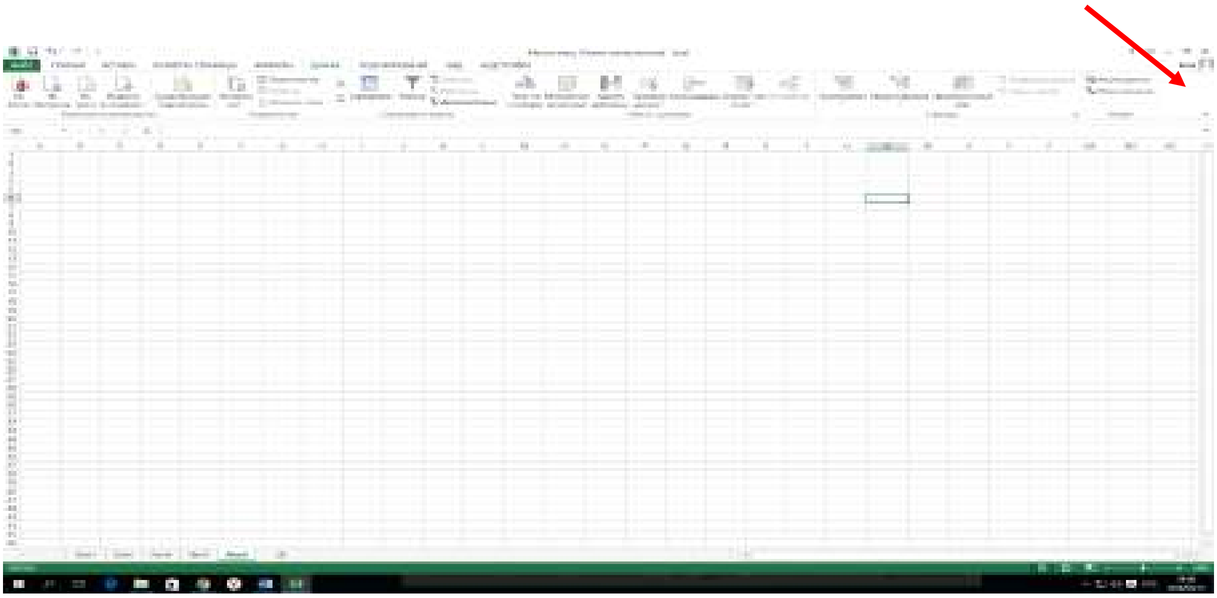
EXCEL parametrlari oynasida keltirilgan parametrlardan <nadstroyka> ni bosamiz. Kompyuter ekranida Microsoft offise qo‘shimcha parametrlarini boshqarish oynasi hosil bo‘ladi.



Ushbu oynaning pastki qismida joylashgan <pereyti> tugmasini bosamiz va ekranda quyidagi Dostupnie nadstroyka ni ko‘rinishda oyna hosil bo‘ladi.

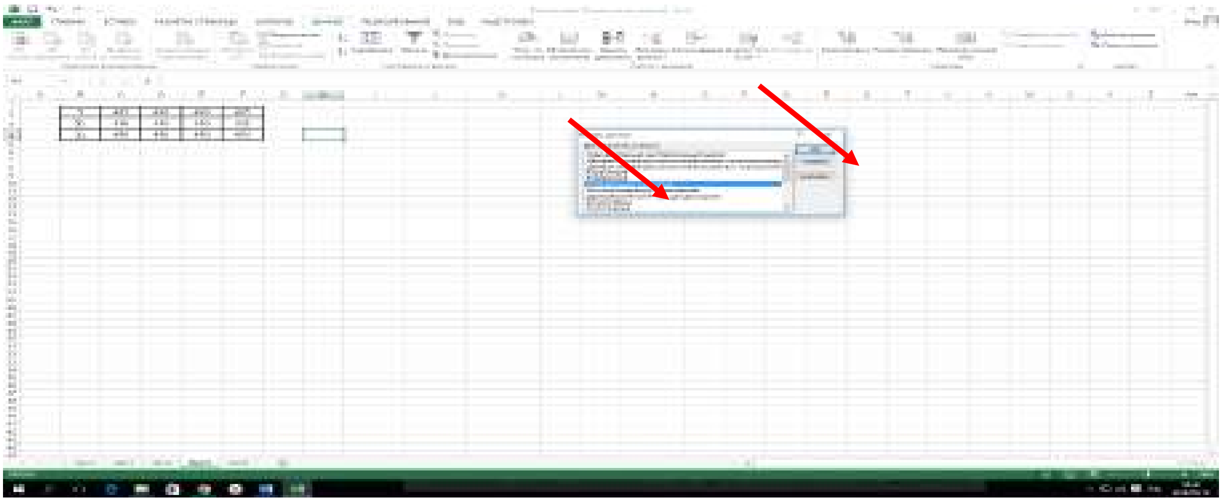


Oynada keltirilgan qo‘shimchalarning barchasiga bayroqchalarni belgilab, OK tugmasi bosiladi va kompyuter yuklab olishi uchun kutiladi. Natija asosiy Menyuda keltirilgan <Dannie> tugmasini bosish bilan tekshiriladi.

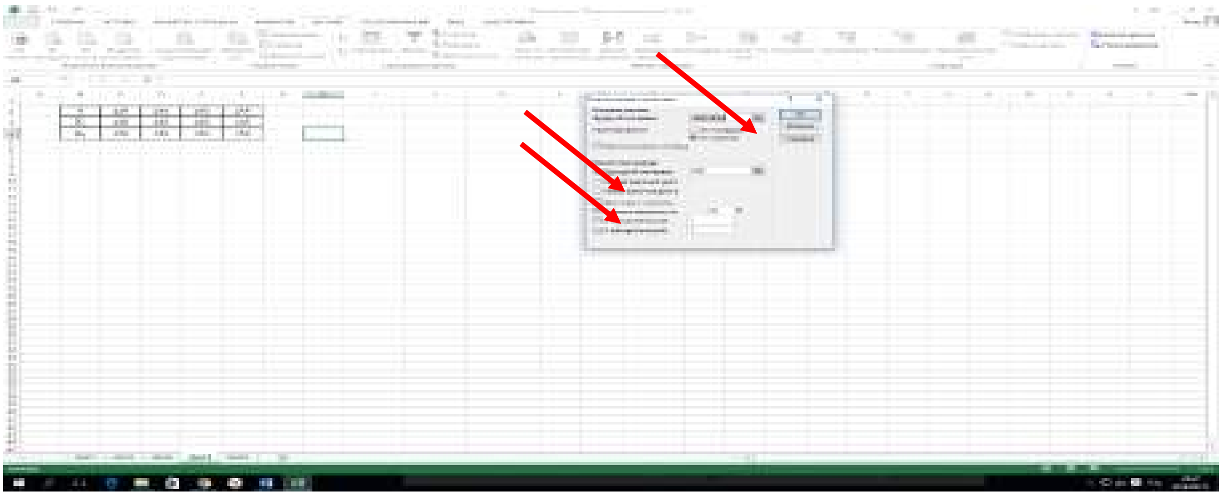


Agar sizning kompyuteringizda ham rasmlarda keltirilgan ko‘rinishda oyna ochilgan bo‘lsa, demak kompteringizda ekonometrik tahlillarni olib borishingiz mumkin bo‘ladi.

Yuqorida keltirilgan misollarni kompyuter texnologiyasida bajarilishini amalga oshirish uchun avvalambor, ma'lumotlar EXCEL dasturi oynasiga joylashtiriladi va <Dannie+analiz dannix> tugmasini bosiladi, natijada quyidagi oyna ochiladi.



Ushbu 7-rasmdan ko‘rish mumkinki, tahlil vositalari (instrumenti analiza) da ko‘plab tahlil usullari berilgan bo‘lib, ulardan qo‘yilgan maqsadga qarab tanlab olinadi, bizga berilgan masala shartiga ko‘ra ko‘rsatkichlarni tavsiflash so‘ralgan. Shu munosabat bilan biz <Opisatelnaya statistika>ni belgilab OK tugmasi bosiladi.



Rasmda keltirilgan <vxodnoy interval> katakchasiga berilgan ma'lumotlarni barchasini belgilash bilan kiritamiz. <Metki> katakchasini belgilaymiz. Ta'kidlash joizki, guruhlashda keltirilgan <po stolbtsam> yoki <po strokam> ma'lumotlarni qanday holatda joylashganligiga qarab belgilanadi. Shundan so‘ng o‘z ixtiyoringizga ko‘ra, natijalarni chiqish parametrlarini belgilaysiz:

<**Выходной интервал**> - ish bajarilayotgan oynaning o’ziga;

<**Новый рабочий лист**> - yangi ish varag‘iga;

<**Новая рабочая книга**> - yangi EXCEL dasturi oynasiga

<**Итоговая статистика**> - xulosa statistikada keltirilgan ma'lumotlarni belgilash olinadigan natijalarni to‘liqligi va ishonchliligini yanada oshirish imkonini beradi. Keltirilgan barcha amallarni bajarish orqali quyidagi natijani beradi.

Sonli qatorlarning statistik tasniflanishi

Частота, Y		*Вариант, X*	
Среднее	12,75	Среднее	5,25
Стандартная ошибка	4,19076365	Стандартная ошибка	1,75
Медиана	13	Медиана	5,5
Мода	20	Мода	#Н/Д
Стандартное отклонение	8,38152731	Стандартное отклонение	3,5
Дисперсия выборки	70,25	Дисперсия выборки	12,25
Эксцесс	-5,92895227	Эксцесс	-1,59766764
Асимметричность	-0,01231313	Асимметричность	-0,32069971
Интервал	15	Интервал	8
Минимум	5	Минимум	1
Максимум	20	Максимум	9
Сумма	51	Сумма	21
Счет	4	Счет	4
Наибольший (1)	20	Наибольший (1)	9
Наименьший (1)	5	Наименьший (1)	1
Уровень надежности(95,0%)	13,3368803	Уровень надежности (95,0%)	5,569281034

IV BOB. JUFT KORRELYATSION TAHLIL

4.1. Funksional va statistik bog‘liqlik tushunchalari

4.2. Korrelyatsion tahlil tushunchasi

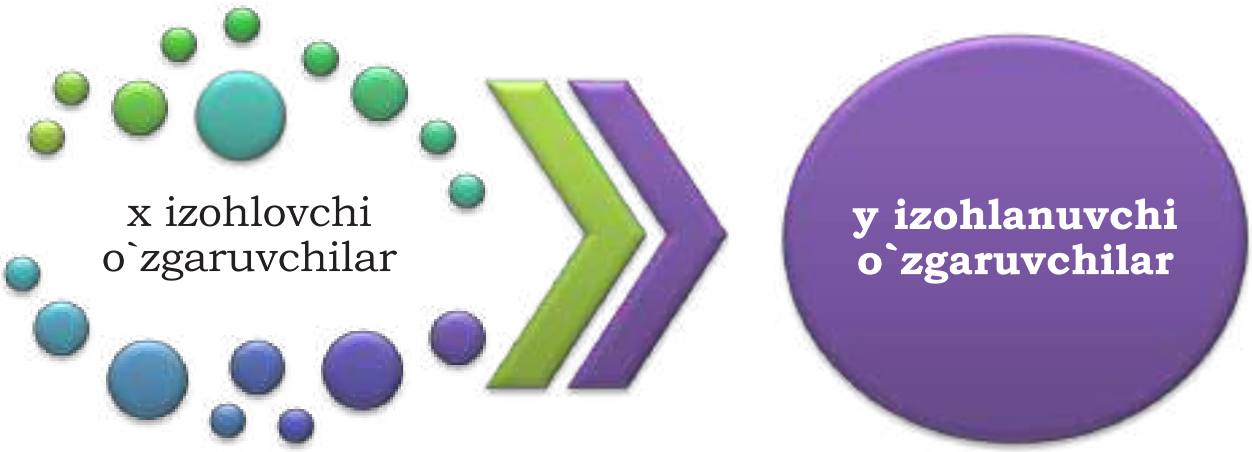
4.3. Bog‘lanish turlari va korrelyatsiya koeffitsiyentini hisoblash usullari

4.4. Korrelyatsiya koeffitsiyentini o‘zgarish intervallari va baholanishi

4.1. Funksional va statistik bog‘liqlik tushunchalari

Iqtisodiy hodisalar, g‘oyat xilma-xil bo‘lgani holda, ular o‘zining u yoki bu xususiyatlarini aks ettiruvchi ko‘plab belgilar bilan tavsiflanadi. Ushbu belgilar vaqtga ko‘ra va makonda o‘zgarib turadi. Ko‘pincha belgi (omil)larning o‘zgarishi o‘zaro bog‘langan va o‘zaro shartlangan. Bir sharoitda omillar o‘rtasidagi bog‘liqlik uzviy (masalan, soatbay ishlab chiqarish va ish haqi), boshqa holatlarda esa omillar o‘rtasidagi bog‘liqlik umuman ko‘zga tashlanmaydi yoki juda sust ifodalanadi (masalan, talabalarning jinsi va ularning o‘zlashtirishi). Belgi (omil)lar o‘rtasidagi bog‘liqlik qanchalik uzviy bo‘lsa, qabul qilinayotgan qarorlar shunchalik aniq va tizimlarni boshqarish shunchalik oson bo‘ladi.

Hodisalar bog‘liqligining ko‘plab shakllari ichida barcha boshqa shakllarni belgilab beruvchi sababli bog‘lanish muhim rol o‘ynaydi. Sabablilikning mohiyati bir hodisaning boshqa hodisaga sabab bo‘lishidan (uni keltirib chiqarishidan) iborat. Har qanday muayyan bog‘lanishda bir belgilar boshqalariga ta'sir etuvchi va ularning o‘zgarishini belgilab beruvchi omillar sifatida, boshqa belgilar esa ushbu omillar ta'sirining natijasi sifatida ishtirok etadi. Boshqacha qilib aytganda, bir belgilar sababni, boshqalari esa oqibatni o‘zida namoyon etadi. Oqibatni tavsiflovchi belgilar, natijali (erksiz) belgilar (y izohlanuvchi o‘zgaruvchilar) deb, sababni tavsiflovchi belgilar esa omilli (mustaqil) belgilar (x izohlovchi o‘zgaruvchilar) deb nomlanadi.



Hodisalar va ularning belgilari o‘rtasidagi bog‘liqlikning ikkita turi mavjud:

Funksional, yoki qat'iy determinatsiyalangan bog‘liqlik (masalan, bir ishchiga to‘g‘ri keladigan mahsulot ishlab chiqarish hajmining ishlab chiqarilgan mahsulot hajmiga va ishchilar soniga bog‘liqligi) va statistik, yoxud stoxastik determinatsiyalangan bog‘liqli (masalan, mehnat unumdorligi bilan mahsulot birligining tannarxi o‘rtasidagi bog‘liqlik).

Funksional bog‘liqlik - bu unda x mustaqil o‘zgaruvchining har bir qiymatiga y erksiz o‘zgaruvchining aniq belgilangan qiymati mos keladigan bog‘lanish. Funksional bog‘liqlik ko‘pincha tabiiy fanlarda uchraydi. Bunday bog‘lanishlar ijtimoiy turmushda, xususan iqtisodiy jarayonlarda kamroq kuzatiladi.

Ijtimoiy-iqtisodiy hodisalar shu bilan tavsiflanadiki, ularga muhim omillar bilan bir qatorda ko‘plab boshqa omillar, shu jumladan tasodifiy omillar ta'sir ko‘rsatadi. Shu munosabat bilan mavjud bog‘liqlik bu yerda funksional bog‘lanishlardagi kabi har bir alohida holatda, balki faqat ko‘p sonli kuzatishlarda "umuman olganda va o‘rtacha darajada" namoyon bo‘ladi. Mazkur holatda statistik bog‘liqlik haqida so‘z boradi.

Statistik bog‘liqlik - bu unda x mustaqil o‘zgaruvchining har bir qiymatiga y erksiz o‘zgaruvchining ko‘plab qiymatlari mos keladigan bog‘lanish, bunda y aynan qanaqa qiymatni qabul qilishi oldindan ma'lum emas.

Statistik bog‘liqlikning alohida holati sifatida korrelyatsion bog‘liqlik ishtirok etadi.

Korrelyatsion bog‘lanish "to‘liqsiz" bog‘liqlik bo‘lib, u har bir alohida holatda emas, balki ancha ko‘p holatlarda faqat o‘rtacha kattaliklarda namoyon bo‘ladi. Ma'lumki, masalan, xodim malakasining oshishi mehnat unumdorligining oshishiga olib keladi. Bu hol ko‘p holatlarda o‘z tasdig‘ini topadi va aynan bir xil jarayon bilan band bo‘lgan bir toifadagi ikki yoki undan ko‘p ishchida bir xil mehnat unumdorligi bo‘lishini anglatmaydi. Mehnat unumdorligi darajalari va ish mahsullari, kam bo‘lsa-da, farq qiladi, chunki bunday ishchilarda ish staji, dastgohning texnik holati, salomatligining holati va hokazolar turlicha bo‘lishi mumkin.

Bundan kelib chiqadiki, statistik bog‘liqlik bu alohida bitta yig‘indining emas, balki u butun yig‘indining xossasi hisoblanadi. Funksional bog‘liqlik — hamma vaqt formulalar bilan ifodalanadi, bu ko‘proq aniq fanlar (matematika, fizika)ga xos.

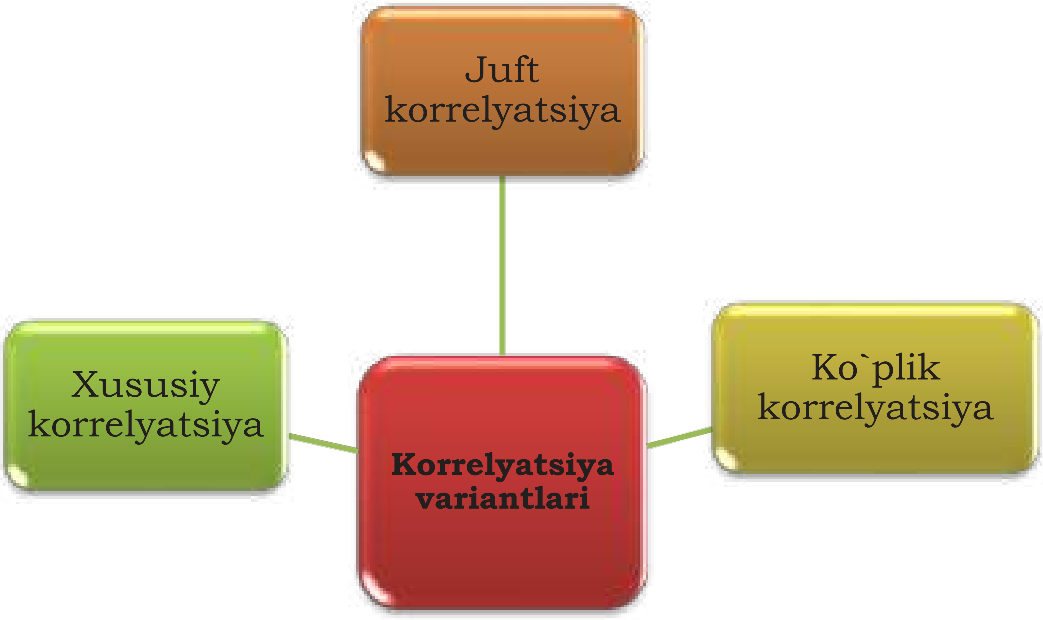
Yig‘indining barcha birliklarida bir xil kuch bilan namoyon bo‘ladi. To‘liq va aniq hisoblanadi, chunki odatda barcha omillar ro‘yxati va ularning tenglama ko‘rinishidagi o‘zgaruvchiga ta'sir etish mexanizmi ma'lum.

Korrelyatsion bog‘liqlik - omillarning xilma-xilligi, ularning o‘zaro bog‘liqligi va qarama-qarshi harakatlar y o‘zgaruvchining keng variantlarda o‘zgarishini keltirib chiqaradi. Alohida holatlarda emas, balki ko‘p holatlarda namoyon bo‘ladi va uni o‘rganish uchun ommaviy kuzatuvlar talab qilinadi. x va y o‘zgaruvchilar o‘rtasidagi bog‘liqlik to‘liqsiz bo‘lib, faqat o‘rtacha kattaliklarda namoyon bo‘ladi. Korrelyatsion bog‘liqlik korrelyatsion va regression tahlil usullari yordamida tadqiq etiladi.

4.2. Korrelyatsion tahlil tushunchasi

Korrelyatsion tahlil ikkita omil o‘rtasidagi (juft bog‘liqlikda) hamda natijaviy omillar bilan boshqa ko‘p omillar o‘rtasidagi (ko‘p omilli bog‘liqlikda) bog‘liqlikning zichligini miqdoriy jihatdan aniqlashdan iborat.

Korrelyatsiya - bu tasodifiy kattaliklar o‘rtasidagi unda tasodifiy kattaliklardan birining o‘zgarishi boshqasining matematik kutishi o‘zgarishiga olib keluvchi statistik bog‘liqlik.



Juft korrelyatsiya - ikkita omil (natijaviy va omillar yoki ikkita omil) o‘rtasidagi bog‘liqlik.

Xususiy korrelyatsiya - boshqa omillarning qat'iy belgilangan qiymatida natijali omil bilan bitta omil yoki ikkita omil o‘rtasidagi bog‘liqlikdir.

Ko‘plik korrelyatsiyasi - natijaviy omil va tadqiqotga kiritilgan ikkita yoki undan ko‘p omillar o‘rtasidagi bog‘liqlik hisoblanadi.

Bog‘liqlikning zichligi miqdoriy jihatdan korrelyatsiya koeffitsiyentlari qiymati bilan ifodalanadi. Korrelyatsiya koeffitsiyentlari qiymatini topish x_i va y_i omillari yakka tartibdagi qiymatlarining ularning $\bar{x}$ va $\bar{y}$ o‘rtacha qiymatlaridan og‘ishlari ko‘paytmasining yig‘indisiga asoslangan:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \tag{4.1}$$

Ushbu kattalik n kuzatuvlar soniga bo‘lindi va chiqqan natija kovariatsiya deb nomlanadi. U ikki belgi variatsiyasining bog‘langanligini tavsiflaydi va ikkita tasodifiy o‘zgaruvchi o‘zaro ta'sirining statistik o‘lchamini o‘zida namoyon etadi. Kovariatsiyani aniqlash formulasi quyidagi ko‘rinishga ega:

$$Cov(y,x)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n} \tag{4.2}$$

bu yerda, n-tadqiq etilayotgan kuzatuvlarning umumiy soni;
x_i-mustaqil o‘zgaruvchining / qiymati (i= 1, 2,…., n);
y_i-erksiz o‘zgaruvchining i qiymati (i= 1, 2,..., n);
$\bar{x}$-mustaqil o‘zgaruvchining o‘rtacha qiymati.

4.3. Bog‘lanish turlari va korrelyatsiya koeffitsiyentini hisoblash usullari

To‘g‘ri chiziqli funksional va korrelyatsion bog‘liqlik - omil miqdorining ortishi bilan natijaviy omil miqdorining bir me'yorda ortishi (yoki kamayishi) yuz beradi (to‘g‘ri chiziq tenglamasi bilan ifodalanadi). Egri chiziqli funksional va korrelyatsion bog‘liqlik-omil miqdorining ortishi bilan natijaviy omil miqdorining ortishi (yoki kamayishi) bir me'yorda yuz bermaydi (egri chiziqlar tenglamalari bilan ifodalanadi).

Harakat yo‘nalishiga qarab funksional va korrelyatsion bog‘liqlik to‘g‘ridan-to‘g‘ri va teskari turlarga bo‘linadi.

To‘g‘ridan-to‘g‘ri funksional va korrelyatsion bog‘liqlik, bu omilli belgi qiymatlarining ortishi (kamayishi) bilan natijali belgining ortishi (kamayishi) yuz beradi.

Teskari bog‘liqlik funksional va korrelyatsion omilli belgi qiymatlarining ortishi (kamayishi) bilan natijali omilning kamayishi (ortishi) yuz beradi.

Modelga kiritilgan omillarning soniga qarab korrelyatsion bog‘liqliklar bir omilli va ko‘p omilli bog‘liqliklarga bo‘linadi. Bir omilli (juft) korrelyatsion bog‘liqliklar bir belgi – omil bilan natijaviy omil o‘rtasidagi bog‘liqlik (boshqa omillarning ta'siri mavxumlashganda) hisoblanadi. Ko‘p omilli (ko‘plik) korrelyatsion bog‘liqliklar esa bir necha omillar (belgilar) bilan natijaviy omil (belgi) o‘rtasidagi bog‘liqlik (omillar birgalikda, ya'ni bir vaqtning o‘zida va o‘zaro bog‘liqlikda ta'sir ko‘rsatadi).

Tahliliy ifodaga ko‘ra bog‘liqlik

To‘g‘ri chiziqli funksional va korrelyatsion bog‘liqlik - omil miqdorining ortishi bilan natijaviy omil miqdorining bir me'yorda ortishi (yoki kamayishi) yuz beradi (to‘g‘ri chiziq tenglamasi bilan ifodalanadi).

Egri chiziqli funksional va korrelyatsion bog‘liqlik-omil miqdorining ortishi bilan natijaviy omil miqdorining ortishi (yoki kamayishi) bir me'yorda yuz bermaydi (egri chiziqlar tenglamalari bilan ifodalanadi).

Korrelyatsion tahlilda korrelyatsiya koeffitsiyentlarini aniqlash va ularning muhimligini, ishonchliligini baholashga asoslanadi. [9]

Chiziqli korrelyatsiya koeffitsiyentining hisoblash formulasi:

$$r_{yx} = \frac{\overline{yx} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y} \tag{4.3}$$

bu yerda, σ_x va σ_y mos ravishda x va y o‘zgaruvchilarning o‘rtacha kvadratik chetlanishidir va ular quyidagi formulalar yordamida hisoblanadi:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} \quad \sigma_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}} \tag{4.4}$$

yoki

$$r_{yx} = \frac{\text{cov}(x,y)}{\sigma_x \cdot \sigma_y} = \frac{\frac{1}{n}\sum(x_i - \overline{x})(y_i - \overline{y})}{\sigma_x \cdot \sigma_y} \tag{4.5}$$

$$r_{\frac{y}{x}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{4.6}$$

$$r_{\frac{y}{x}} = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}} \tag{4.7}$$

Korrelyatsion tahlil o‘tkazilganda quyidagi korrelyatsiya koef-fitsiyentlari hisoblanadi:

1. Xususiy korrelyatsiya koeffitsiyentlari. Xususiy korrelyatsiya koeffitsiyenti asosiy va unga ta’sir etuvchi omillar o‘rtasidagi bog‘lanish zichligini bildiradi.

[9] Gujarati D.N. Basic Econometrics. McGraw-Hill, 4th edition, 2003 (Gu), Inc. p. 90

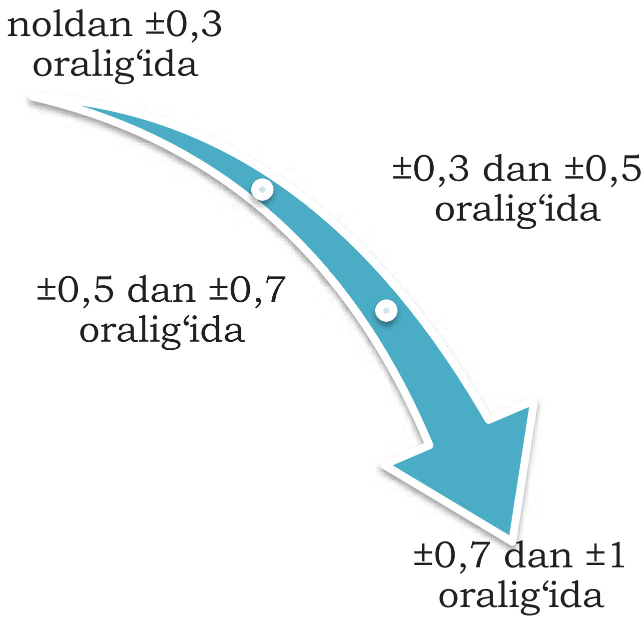
2. Juft korrelyatsiya koeffitsiyentlari asosiy omil inobatga olinmagan nuqtada hisoblanadi. Agar juft korrelyatsiya koeffitsiyenti 0,6 dan katta bo‘lsa, unda omillararo bog‘lanish kuchli deb hisoblanadi va erkin omillar ma’lum darajada bir birini takrorlaydi. Agar modelda o‘zaro bog‘langan omillar qatnashsa, model yordamida qilingan hisoblar noto‘g‘ri chiqishi mumkin va omillar ta’siri ikki barovar hisoblanishi mumkin. O‘zaro bog‘langan ta’sir etuvchi omillardan bittasi modeldan chiqarib tashlanadi. Albatta modelda kuchliroq va mustahkamroq omil qoladi.

3. Ko‘p omilli modellarda agar natijaviy omilga bir necha omillar ta’sir ko‘rsatsa, unda omillar orasida ko‘plikdagi korrelyatsiya koeffitsiyenti hisoblanadi.

4.4. Korrelyatsiya koeffitsiyentini o‘zgarish intervallari va baholanishi

Korrelyatsiya koeffitsiyenti (r) -1 dan +1 oralig‘ida bo‘ladi. Koeffitsiyentning musbat qiymati to‘g‘ridan-to‘g‘ri bog‘liqlikning, manfiy qiymati esa teskari bog‘liqlikning mavjudligidan dalolat beradi. Agar r= ±1 bo‘lsa, korrelyatsion bog‘liqlik chiziqli funksional bog‘liqlik bilan ifodalanadi. Agar $r=0$ bo‘lsa omillar o‘rtasida bog‘lanish mavjud emas deb hisoblanadi. Yuqorida ta’kidlanganidek, agarda $0<r<1$ bo‘lsa, bunda to‘g‘ri bog‘lanish mavjud bo‘ladi va $-1<r<0$ oralig‘da teskari bog‘lanish mavjud hisoblanadi. $r=1$ga bo‘lsa funksional bog‘lanish mavjud bo‘ladi.

Korrelyatsiya koeffitsiyenti $r_{y/x}$, omillar o‘rtasidagi bog‘liqlikni sifat jihatidan tavsiflaydi:



Korrelyatsiya koeffitsiyentlari statistik kattaliklar sifatida ishonchlilik nuqtai nazaridan tahlil qilinadi va baholanadi. Bu shu bilan izohlanadiki, kuzatuvlarning har qanday to‘plami ayrim tanlashni o‘zida namoyon etadi, demak, tanlash asosida hisoblab chiqilgan har qanday ko‘rsatkichning qiymati haqiqiy qiymat sifatida ko‘rib chiqilishi mumkin emas, balki uning ozmi yoki ko‘pmi aniq bahosi hisoblanadi. Shu munosabat bilan ko‘rsatkichlarning ahamiyatliligi (muhimligi)ni tekshirish zarurati paydo bo‘ladi.

Ba’zi tadqiqotlarda korrelyatsiyaning juft koeffitsiyentlarini tahlil qilishni omilli belgilar o‘rtasidagi multikollinear bog‘liqlikni aniqlash usuli deb biladilar. Ikkita argument, agar ular o‘rtasidagi korrelyatsiyaning juft koeffitsiyentlari absolyut kattaligi bo‘yicha 0,8 dan katta bo‘lsa. kollinear deb hisoblaydilar. Amaliyotda esa ikki argument o‘rtasidagi korrelyatsiyaning juft koeffitsiyentlari absolyut kattaligi bo‘yicha 0,8 dan katta bo‘lsada, multikollinearlik mavjud bo‘lmasligi mumkin.

Shu o‘rinda ta'kidlash lozimki, tanlangan omilliy ko‘rsatkichlar orasida multikolleniarlik mavjud yoki yo‘qligini VIF testi (ing. Variance inflation factor) va qabul qilish (ing. Tolarence) testlariga tekshirish orqali aniqlanadi.

VIF testi quyidagi formula orqali hisoblanadi[10]:

$$VIF_k = \frac{1}{1 - R_k^2}$$

Bu yerda, VIF_k - k o‘zgaruvchi uchun VIF qiymati, - k o‘zgaruvchi uchun determinatsiya koeffitsiyenti

Odatda VIF testining 5 dan kichik qiymatlari tahlil uchun eng munosib deb topiladi, lekin adabiyotlarda VIF ning 10 dan kichik qiymatlari regressiya tenglamasini tuzishda tanlangan omillar ishtirokini ta'minlash mumkinligini ko‘rsatadi, 10 dan katta qiymatlari esa tanlangan omillar regressiya tenglamasini tuzishda munosib emasligini ifodalaydi.

Shuningdek, ko‘pincha tadqiqotchilar tomondan kolleniarlikning darajasini aniqlash uchun 1/ VIF orqali aniqlanadigan qabul qilish testidan foydalaniladi. VIF ning 10 qiymatiga mos ravishda, qabul qilish testining 0,1 dan kichik qiymatlari tahlil uchun tanlangan bir o‘zgaruvchining boshqa o‘zgaruvchilar bilan chiziqli bog‘liq ekanligini ifodalaydi.

Chiziqsiz regressiya uchun o‘rganilayotgan hodisalar o‘rtasidagi bog‘lanishlarning zichligi korrelyatsiya koeffitsiyenti bilan baholanadi ρ_{xy}

$$0 \leq \rho \leq 1$$

$$\rho_{yx} = \sqrt{1 - \frac{\sigma_{ocm}^2}{\sigma_y^2}} = \sqrt{1 - \frac{\sum(y - y_x)^2}{\sum(y - \bar{y})^2}}$$

Natijaviy ko‘rsatkichga ta’sir etuvchi omillarning umumiy ta’siri ko‘plikdagi korrelyatsiya indeksi bilan baholanadi:

$$R_{yx_1x_2} = \sqrt{1 - \frac{\sigma_y^2}{\sigma_m^2}}$$

[10] Jamal I.Daoud (2017) Multicollinearity and regression analysis, Journal of Physics:Conference Series

Ko‘plikdagi korrelyatsiya indeksning qiymati 0 dan 1 gacha oraliqda o‘zgaradi va maksimal juft korrelyatsiyasi indeksidan katta yoki teng bo‘lishi kerak:

$$R_{yx_1x_2,..,x_p} \qquad \text{(i=1,p).}$$

Chiziqli bog‘lanish uchun *ko‘plikdagi korrelyatsiya koeffitsiyenti* juft korrelyatsiya koeffitsiyentlar matritsasi orqali aniqlanishi mumkin:

$$R_{yx_1x_2,\dots x_p} = \sqrt{1 - \frac{\Delta r}{\Delta r_{11}}}$$

bu yerda

$$\Delta = \begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_p} \\ r_{yx_1} & 1 & r_{x_1x_2} & \dots & r_{x_1x_2} \\ r_{yx_2} & r_{yx_1} & 1 & \dots & r_{x_2x_p} \\ \dots & \dots & \dots & & \dots \\ r_{yx_p} & r_{x_px_1} & r_{x_px_2} & \dots & 1 \end{vmatrix}$$ - juft korrelyatsiya koeffitsiyentini aniqlash matritsasi.

Xususiy korrelyatsiya koeffitsiyentlari ga omilning boshqa faktorlar o‘zgarmas darajada bo‘lganda ta’sirini o‘lchaydi va quyidagi formula orqali aniqlanadi:

$$r_{yx_1x_1x_2\dots.x_{i-1}x_{i+1}\dots.x_p} = \sqrt{1 - \frac{1 - R^2_{yx_1x_1x_2\dots.x_{i-1}x_{i+1}\dots.x_p}}{1 - R^2_{yx_1x_1x_2\dots.x_{i-1}x_{i+1}\dots.x_p}}}$$

yoki quyidagi rekurrent formuladan foydalanib:

$$r_{yx_1.x_1x_2\dots x_p} = \frac{r_{yx_1.x_1x_2\dots x_p} - r_{yx_p.x_1x_2\dots x_{p-1}}\, r_{yx_p.x_1x_2\dots x_{p-1}}}{\sqrt{(1 - r^2_{yx_1x_1x_2\dots.x_{p-1}})(1 - r^2_{yx_1x_1x_2\dots.x_{p-1}}}}$$

Xususiy korrelyatsiya koeffitsiyentlari -1 dan 1 gacha oraliqda o‘zgaradi.

Tuzilgan modelning umumiy sifatini determinatsiya koeffitsiyenti baholaydi. Ko‘plikdagi determinatsiya koeffitsiyenti ko‘plikdagi korrelyatsiya indeksining kvadratiga teng:

$$R^2_{yx_1x_2,\dots.x_p}$$

Nazorat uchun savollar

1. Korrelyatsion tahlilning maqsadlari nimalardan iborat?

2. Juft, xususiy va koʻplikdagi korrelyatsiya koeffitsiyentlarining farqi nimadan iborat?

3. Qaysi hollarda korrelyatsiya indeksi qoʻllaniladi?

4. Toʻgʻri va teskari korrelyatsiya koeffitsiyentlarining farqi nimadan iborat?

5. Korrelyatsiya koeffitsiyentlari qaysi oraliqda boʻladi?

6. Modelga kiritilgan omillarning soniga qarab korrelyatsion bogʻliqliklar qanday turlari mavjud?

Masala. Avtomobil ishlab chiqaruvchi kompaniyaning sotilgan avtomobillar sonini aholining oʻrtacha daromadiga qay darajada bogʻliqligini tahlil qiling va oʻz xulosa va takliflaringizni bayon eting?

№	Sotilgan avtomobillar soni	Aholining oʻrtacha daromadi
1	220	17,7
2	213	15,1
3	205	14,6
4	197	12,4
5	194	12,1
6	185	11,7
7	181	11,2
8	174	10,6
9	168	9,8
10	163	9,7
11	152	9,5
12	146	9,1
13	139	8,9
14	134	8,5
15	127	8,1

Topshiriq. Berilgan ma’lumotlar asosida korrelyatsion tahlil amalga oshirilsin.

Yechim. Excel dasturini ishga tushirib, ma’lumotlarni yuqoridagi tartibda kiritamiz. Bunda asosiy omil (y) sifatida sotilgan avtomobillar sonini hamda unga ta’sir e’tuvchi omil (x) sifatida aholining o‘rtacha daromadini belgilaymiz.

1. Avvalo, korrelyatsion tahlil formulasini esga olamiz: $r_{yx} = \dfrac{\overline{yx} - \bar{x} \cdot \bar{y}}{\sigma_x \cdot \sigma_y}$. So‘ngra formulada ko‘rsatilgan natijalarni excel funksiya yordamida topib olamiz.

2. Bunda avvalo barcha o‘zgaruvchilarimiz uchun o‘rtacha qiymatini topamiz. Buning uchun barcha o‘zgaruchilarimizni belgilab, exceldagi Среднее buyrug‘i yordamida o‘rtacha qiymatni topib olamiz:

3.Endilikda keyingi bo‘sh ustunga ta’sir etuvchi omil (x) ning kvadratga oshirilgan qiymatini hisoblaymiz, ushbu amalni bir katakcha uchun qo‘llaymiz va buyruq orqali qolgan barcha x lar uchun ham hisobni amalga oshiramiz:

	A	B	C	D
1	№	Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2
2	1	220	17,7	=C2^2
3	2	213	15,1	
4	3	205	14,6	
5	4	197	12,4	
6	5	194	12,1	
7	6	185	11,7	
8	7	181	11,2	
9	8	174	10,6	

4. Ana endi yangi ustunga y va x ko‘paytmalarini hisoblaymiz:

	A	B	C	D	E
1	№	Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x
2	1	220	17,7	313,29	=B2*C2
3	2	213	15,1	228,01	
4	3	205	14,6	213,16	
5	4	197	12,4	153,76	
6	5	194	12,1	146,41	
7	6	185	11,7	136,89	
8	7	181	11,2	125,44	
9	8	174	10,6	112,36	
10	9	168	9,8	96,04	
11	10	163	9,7	94,09	
12	11	152	9,5	90,25	
13	12	146	9,1	82,81	
14	13	139	8,9	79,21	
15	14	134	8,5	72,25	
16	15	127	8,1	65,61	
17	O'rtacha	173,2	11,26666667	133,972	

5. Keyingi ustunimizda har bir y qiymatlaridan y o'rtacha ayirmasining kvadratini hisoblab olamiz:

6.

I7 =(B2-B17)^2

A	B	C	D	E	F
№	Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x	(y-y_)^2
1	220	17,7	313,29	3894	=(B2-B17)^2
2	213	15,1	228,01	3216,3	
3	205	14,6	213,16	2993	
4	197	12,4	153,76	2442,8	
5	194	12,1	146,41	2347,4	
6	185	11,7	136,89	2164,5	
7	181	11,2	125,44	2027,2	
8	174	10,6	112,36	1844,4	
9	168	9,8	96,04	1646,4	
10	163	9,7	94,09	1581,1	
11	152	9,5	90,25	1444	
12	146	9,1	82,81	1328,6	
13	139	8,9	79,21	1237,1	
14	134	8,5	72,25	1139	
15	127	8,1	65,61	1028,7	
O'rtacha	173,2	11,26666667	133,972	2022,3	

Bunda barcha o‘zgaruvchilarimiz uchun alohida hisobni amalga oshirmasligimiz uchun o‘rtacha qiymatni kiritayotganda F4 tugmasini bosib qo‘yamiz va shunda uni shu ustun bo‘yicha o‘zgartirmasdan saqlab qoladi, va albatta bu hisobimiz uchun ham o‘rtacha qiymatni hisoblaymiz.

7. Shuningdek, yuqoridagi amallarni x uchun ham takrorlaymiz:

№	Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x	(y-y_)^2	(x-x_)^2
1	220	17,7	313,29	3894	2190,24	41,38778
2	213	15,1	228,01	3216,3	1584,04	14,69444
3	205	14,6	213,16	2993	1011,24	11,11111
4	197	12,4	153,76	2442,8	566,44	1,284444
5	194	12,1	146,41	2347,4	432,64	0,694444
6	185	11,7	136,89	2164,5	139,24	0,187778
7	181	11,2	125,44	2027,2	60,84	0,004444
8	174	10,6	112,36	1844,4	0,64	0,444444
9	168	9,8	96,04	1646,4	27,04	2,151111
10	163	9,7	94,09	1581,1	104,04	2,454444
11	152	9,5	90,25	1444	449,44	3,121111
12	146	9,1	82,81	1328,6	739,84	4,694444
13	139	8,9	79,21	1237,1	1169,64	5,601111
14	134	8,5	72,25	1139	1536,64	7,654444
15	127	8,1	65,61	1028,7	2134,44	10,02778
O'rtacha	173,2	11,26666667	133,972	2022,3	809,76	7,034222

8. Ana endi formulada ko‘rsatilgan standart og‘ishlarni topishimiz uchun (y-y_)^2 va (x-x_)^2 qiymatlarni excel funksiyasidan foydalanib, ya’ni bo‘sh katakga КОРЕНЬ buyrug‘ini yozib yoniga (y-y_)^2 va (x-x_)^2 qiymatlarni belgilash orqali ildizdan chiqarib, natijalarni topib olamiz:

B	C	D	E	F	G	H
Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x	(y-y_)^2	(x-x_)^2	
220	17,7	313,29	3894	2190,24	41,38778	
213	15,1	228,01	3216,3	1584,04	14,69444	
205	14,6	213,16	2993	1011,24	11,11111	
197	12,4	153,76	2442,8	566,44	1,284444	
194	12,1	146,41	2347,4	432,64	0,694444	
185	11,7	136,89	2164,5	139,24	0,187778	
181	11,2	125,44	2027,2	60,84	0,004444	
174	10,6	112,36	1844,4	0,64	0,444444	
168	9,8	96,04	1646,4	27,04	2,151111	
163	9,7	94,09	1581,1	104,04	2,454444	
152	9,5	90,25	1444	449,44	3,121111	
146	9,1	82,81	1328,6	739,84	4,694444	
139	8,9	79,21	1237,1	1169,64	5,601111	
134	8,5	72,25	1139	1536,64	7,654444	
127	8,1	65,61	1028,7	2134,44	10,02778	
173,2	11,26666667	133,972	2022,3	809,76	7,034222	
				=КОРЕНЬ(F17)		

Demak bu buyruq orqali quyidagi natijalarga erishdik:

Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x	(y-y_)^2	(x-x_)^2
220	17,7	313,29	3894	2190,24	41,38778
213	15,1	228,01	3216,3	1584,04	14,69444
205	14,6	213,16	2993	1011,24	11,11111
197	12,4	153,76	2442,8	566,44	1,284444
194	12,1	146,41	2347,4	432,64	0,694444
185	11,7	136,89	2164,5	139,24	0,187778
181	11,2	125,44	2027,2	60,84	0,004444
174	10,6	112,36	1844,4	0,64	0,444444
168	9,8	96,04	1646,4	27,04	2,151111
163	9,7	94,09	1581,1	104,04	2,454444
152	9,5	90,25	1444	449,44	3,121111
146	9,1	82,81	1328,6	739,84	4,694444
139	8,9	79,21	1237,1	1169,64	5,601111
134	8,5	72,25	1139	1536,64	7,654444
127	8,1	65,61	1028,7	2134,44	10,02778
173,2	11,26666667	133,972	2022,3	809,76	7,034222
				28,45628	2,652211

9. Shunday qilib formula uchun kerak bo‘lgan barcha natijalar topildi, endi topilgan natijalarni formula o‘rniga kiritib, korrelyatsiyani hisoblaymiz:

Sotilgan avtomobillar soni (y)	Aholining o'rtacha daromadi (x)	x^2	y*x	(y-y_)^2	(x-x_)^2
220	17,7	313,29	3894	2190,24	41,38778
213	15,1	228,01	3216,3	1584,04	14,69444
205	14,6	213,16	2993	1011,24	11,11111
197	12,4	153,76	2442,8	566,44	1,284444
194	12,1	146,41	2347,4	432,64	0,694444
185	11,7	136,89	2164,5	139,24	0,187778
181	11,2	125,44	2027,2	60,84	0,004444
174	10,6	112,36	1844,4	0,64	0,444444
168	9,8	96,04	1646,4	27,04	2,151111
163	9,7	94,09	1581,1	104,04	2,454444
152	9,5	90,25	1444	449,44	3,121111
146	9,1	82,81	1328,6	739,84	4,694444
139	8,9	79,21	1237,1	1169,64	5,601111
134	8,5	72,25	1139	1536,64	7,654444
127	8,1	65,61	1028,7	2134,44	10,02778
173,2	11,26666667	133,972	2022,3	809,76	7,034222
				28,45628	2,652211
		r(y/x)=	=(E17-(C17*B17))/(G18*F18)		

Demak, natijamiz r(y/x)= 0,939597.

Xulosa: korrelyatsion tahlil natijasidan ko‘rinadiki, sotilgan avtomobillar soni aholining o‘rtacha daromadiga to‘g‘ri va kuchli bog‘langan. Empirik tahlil natijasiga asosan aholining o‘rtacha daromadining ortishi mashinalar sotuv hajmini ortishiga olib keladi.

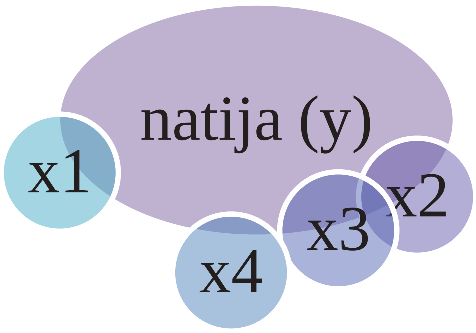
V BOB. JUFT REGRESSION TAHLIL

5.1. Regression tahlil tushunchasi

Regression tahlil bog‘liqlikning unda natijaviy omilning o‘zgarishi bir yoki bir necha omillarning ta'siri bilan shartlangan, natijaviy omilga ta'sir ko‘rsatuvchi boshqa barcha omillar ko‘pligi esa doimiy va o‘rtacha qiymat sifatida qabul qilinadigan tahliliy shaklini aniqlashdan iborat.

Regression tahlilning maqsadi - natijaviy omil shartli o‘rtacha qiymatining omilli belgilarga funksional bog‘liqligini baholashdan iborat, regression tahlilning asosiy omilli shundan iboratki, faqat natijaviy omil taqsimlashning normal qonuniga, ta'sir etuvchi omillar esa taqsimlashning ixtiyoriy qonuniga bo‘ysunadi.

Bunda regression tahlilda natija (y) va omillar (x) o‘rtasidagi sabab-oqibat bog‘liqlikning mavjudligi oldindan nazarda tutiladi.



Regressiya tenglamasi yoki ijtimoiy-iqtisodiy hodisalar bog‘liqlik modeli quyidagi funksiya bilan ifodalanadi:

$$\hat{y}_x = f(x) \qquad (5.1)$$

Bunda juft regressiya: natijaviy va bitta omil o‘rtasidagi bog‘liqlikni tavsiflaydi.

$$\hat{y}_x = f(x_1, x_2, \dots x_k) \qquad (5.2)$$

bu yerda, k -omillar soni.

Bunda ko‘plik regressiyasi mavjud bo‘lib, y natijaviy omil ($\hat{y}$) bilan ikki va undan ko‘p omil o‘rtasidagi bog‘liqlikni tavsiflaydi. Tenglama uni tuzishda talablarga amal qilingan taqdirda real modellashtiriladigan hodisa yoki jarayonga mos keladi.

Regressiya tenglamasini tuzishga nisbatan quyidagi talablar qo‘yiladi.

1) Boshlang`ich ma'lumotlar yig`indisi bir xil bo`lishi va matematik jihatdan uzluksiz funksiyalar bilan ta'riflanishi kerak.

2) Ancha katta hajmdagi tadqiq etiladigan tanlangan yig`indining mavjudligi.

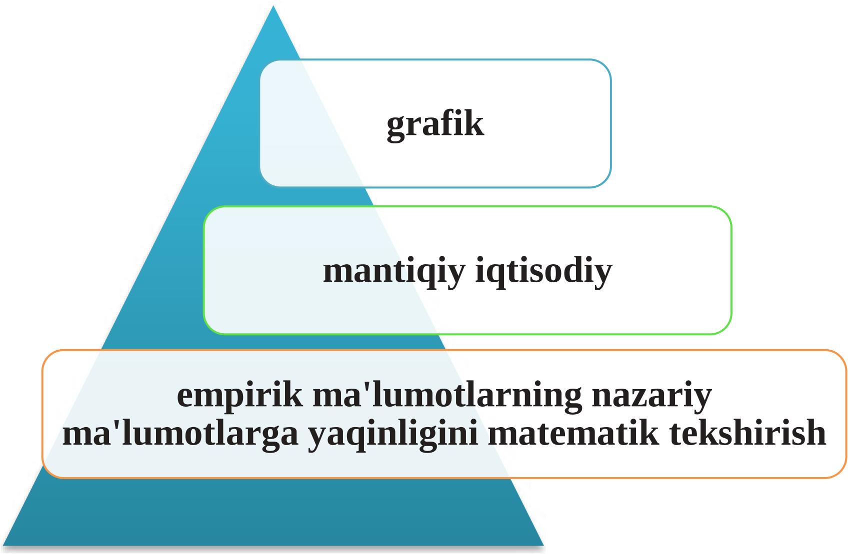
3) Hodisalar va jarayonlar o`rtasidagi sababli-oqibatli bog`liqliklarni, imkon qadar, bog`liqlikning chiziqli (yoki chiziqli holatga keltiriladigan) shakl bilan ta'riflash lozim.

4) Model parametrlariga nisbatan miqdoriy cheklovlarning mavjud emasligi.

5) Omillarning miqdoriy ifodasi.

Bog‘liqlik shakli chiziqli funksiya bilan ham (to‘g‘ri tenglama), chiziqsiz funksiyalar bilan ham (turli tartiblar polinomlari, giperbola, darajali funksiya va b.) ifodalanishi mumkin.

Belgilar o‘rtasidagi bog‘liqlik shaklini ifodalash uchun funksiyalarni tanlash bir necha bosqichda kechadi:



Ko‘pincha korrelyatsion bog‘liqlik shaklini ifodalash uchun bir vaqtning o‘zida bir necha funksiya mos keladi, shuning uchun bog‘liqlik shaklini ifodalash uchun funksiyalarni muqobil asosda tanlashni yakuniy asoslagan ma'qul. Regressiyaning chiziqli shakli tushunish, talqin etish va hisob-kitoblar texnikasi nuqtai nazaridan eng oddiy shakl hisoblanadi.

Hodisalarning o‘zaro bog‘liqligi modellarini nazariy jihatdan asoslash muayyan shartlarga amal qilish orqali ta'minlanadi. Ular jumlasiga quyidagilar kiradi:

Chiziqli juft regressiya tenglamasi umumiy holda quyidagi ko‘rinishga ega:

$$y(x) = a_0 - a_1 x_1 - \varepsilon_i$$

bu yerda, a_0, a_1 - model parametrlari;

εi - tasodifiy kattalik (qoldiq miqdori).

Chiziqli juft regressiya modeli parametrlarining mazmuni:

a_0 - regression tenglamaning erkin koeffitsiyenti (a'zosi).

Iqtisodiy ma'noga ega emas va, agar omil $x = 0$ bo‘lsa, natijaviy omilning belgining qiymatini ko‘rsatadi.

a_1 - regressiya koeffitsiyenti, agar x o‘zgaruvchi bir o‘lchov birligiga oshirilsa, y natijaviy omil o‘rtacha qancha miqdorga o‘zgarishini ko‘rsatadi. Regressiya koeffitsiyentidagi belgi bog‘liqlikning yo‘nalishini ko‘rsatadi:

$a_1 > 0$ bo‘lganida – bog‘liqlik to‘g‘ri; $a_1 < 0$ bo‘lganida – bog‘liqlik teskari.

ε_i - mustaqil, normal taqsimlangan tasodifiy kattalik, nolli matematik kutishli ($M_\varepsilon = 0$) va doimiy dispersiyali ($D_\varepsilon = \sigma^2$) qoldiq. y ning o‘zgarishi x ning o‘zgarishi bilan noaniq ta'riflanishini aks ettiradi, chunki ushbu modelda hisobga olinmagan boshqa omillar ham ishtirok etadi.

a_0 va a_1 modelining parametrlarini baholash eng kichik kvadratlar usuli bilan amalga oshiriladi.

5.2. Bir omilli va ko‘p omilli regressiya

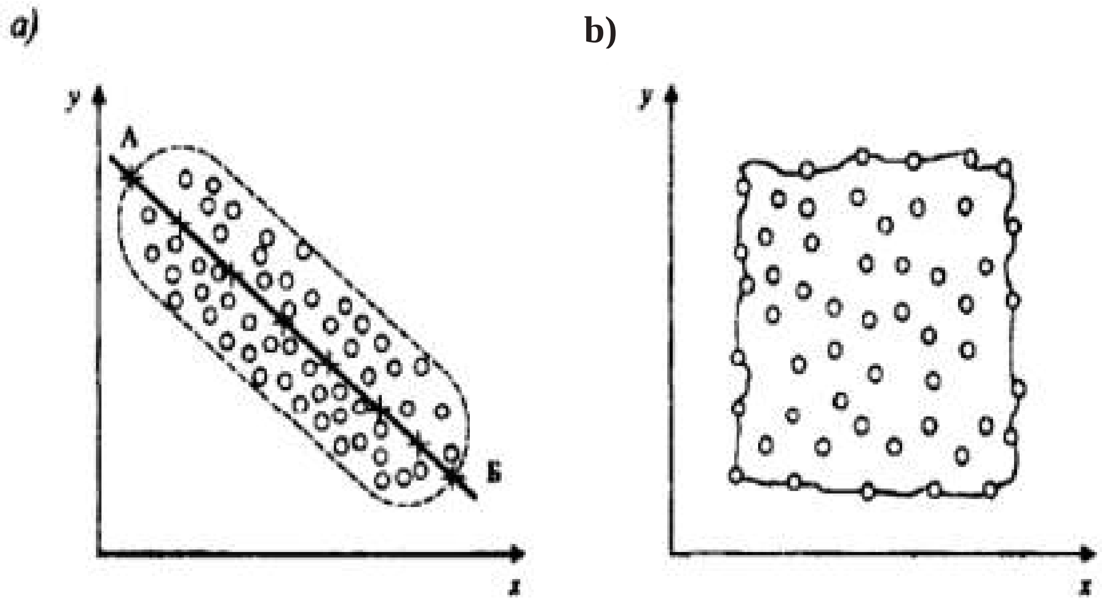
Regression tahlili deganda hodisalar (jarayonlar) o‘rtasidagi munosabatlarni o‘rganish tushuniladi, ular ko‘p, ba’zan noma'lum omillarga bog‘liq.

Ko‘pincha x va y o‘zgaruvchilari o‘rtasida munosabatlar mavjud, lekin aniq emas, bunda y ning bir nechta qiymatlari (to‘plami) x ning bitta qiymatiga to‘g‘ri keladi. Bunday hollarda munosabatlar regressiya deb ataladi. Shunday qilib, y=f(x) funksiya regressiya (korrelyatsiya), agar argumentning har bir qiymati y statistik taqsimot qatoriga to‘g‘ri kelsa.

Regressiya tahlilining mohiyati regressiya tenglamasini o‘rnatishga kamayadi, ya'ni tasodifiy o‘zgaruvchilar orasidagi egri shakli (argumentlar x va y funksiyasi), ular orasidagi munosabatlarning mustahkamligini baholash, o‘lchov natijalarining ishonchliligi va adekvatligi.

X va Y o‘rtasidagi bunday aloqaning mavjudligini oldindan aniqlash uchun grafikda nuqta chiziladi va korrelyatsiya deb ataladigan maydon quriladi (5.1-rasm). Korrelyatsiya maydonining turiga qarab, korrelyatsiya mavjudligini baholash mumkin. Shunday qilib, a-grafikda shuni ko‘rsatadiki, eksperimental ma'lumotlar x va y rasmdagi o‘lchovlar o‘rtasida ma'lum bog‘liqlikka ega. b-grafikdagi bunday aloqani ko‘rsatmaydi.

Bir o‘zgaruvchan (juftlashgan) va ko‘p o‘zgaruvchan regressiya bog‘liqliklarini ajrating. Juftlik qaramligi bilan juftlik regressiyasini to‘g‘ri chiziq, parabola, giperbola, logarifmik, darajali yoki eksponentli funksiya, polinom va boshqalar bilan taxmin qilish mumkin. Ikki faktorli maydonni tekislik, ikkinchi darajali parabola, giperbola bilan taxmin qilish mumkin.



5.1-rasm. Korrelyatsion maydon

Juft regressiya modelini (yoki bitta faktorli modelni) tuzish y va x ikkita ko‘rsatkich o‘rtasidagi bog‘liqlik tenglamasini topishdan iborat, ya'ni bitta omilning o‘zgarishi boshqasiga qanday ta'sir qilishi aniqlanadi.

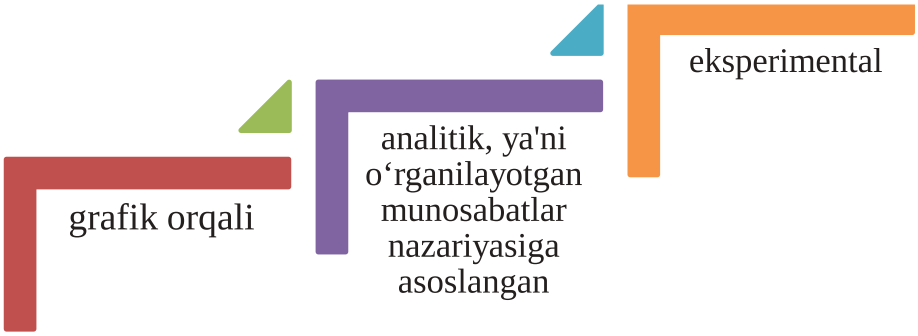
Ekonometrika muammolarida asosiy bosqich - bu model parametrlarini topish va ularning sifatini baholash. Juft regressiya modelining tenglamasini umumiy shaklda yozish mumkin:

$$y = \hat{f}(x)$$

bu yerda, y - bog‘liq ko‘rsatkich (natijaviy ko‘rsatkich);

x - mustaqil omil.

Juft regressiyada $y = \hat{f}(x)$ matematik funksiya turini tanlash uchta usul bilan amalga oshirilishi mumkin:



Ikkala xususiyat o‘rtasidagi bog‘liqlikni o‘rganayotganda, regressiya tenglamasining turini tanlashning grafik usuli juda aniq. U korrelyatsiya maydoniga asoslangan.

Regressiya tenglamasining turini tanlashning analitik usuli katta qiziqish uyg‘otadi. U o‘rganilayotgan xususiyatlarning bog‘lanishining moddiy tabiatini o‘rganishga asoslangan.

Kompyuterda ma'lumotlarni qayta ishlashda regressiya tenglamasi turini tanlash odatda eksperimental usul bilan amalga oshiriladi.

Chiziqli regressiyada kuzatilgan miqdoriy y o‘zgaruvchining kuzatilgan x o‘zgaruvchilarga bog‘liqligini quyidagicha ifodalash mumkin:

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

Yuqorida ifodalanilgan funksiya chiziqli (ko‘p omilli) ko‘plikdagi regressiya deb ataladi. Bu x o‘zgaruvchilari va $\beta_1, \beta_2 .. \beta_k$ parametrlari bo‘yicha ham chiziqli va qo‘shimcha ravishda tasodifiy o‘zgaruvchiga bog‘liq (xatolar, yashirin-yashirin tasodifiy o‘zgaruvchilar)u. Y o‘zgaruvchisi tasodifiy o‘zgaruvchiga bog‘liq bo‘lgani uchun u ham tasodifiydir. Klassik chiziqli regressiya modelidagi x o‘zgaruvchilari tasodifiy deb hisoblanmaydi. Agar x o‘zgaruvchilar tasodifiy bo‘lishi mumkin deb taxmin qilinsa, bunday model umumlashtirilgan chiziqli regressiya modeliga aylanadi. Regressorlarning tasodifiy yoki tasodifiy bo‘lishidan qat'iy nazar, regressiya modeli stoxastik (ehtimolli) modellar sinfiga kiradi.

Yuqorida muhokama qilingan bir faktorli modelning ijobiy xususiyati uning soddaligi. Uning asosiy kamchiligi - bu narxlanish jarayonining yetarli darajada adekvat tavsifi. Aniqrog‘i, qimmatli qog‘ozlar rentabelligini shakllantirish jarayonlarini ko‘p chiziqli regressiya modellari sinfiga mansub ko‘p o‘zgaruvchan modellar yordamida tasvirlash mumkin.

Iqtisodiyotning holati ko‘plab omillarga bog‘liq bo‘lib, ular orasida iqtisodiyotning barcha sohalariga ta'sir ko‘rsatadigan bir nechta asosiy omillar bor:

5.3. Chiziqli va chiziqsiz regressiya

Analitik bog‘liqlik turiga qarab, chiziqli va chiziqsiz regressiyalar bo‘linadi. Chiziqli juft regressiya quyidagi tenglama bilan tavsiflanadi:

$$y = a_0 + a_1 x$$

Ijtimoiy-iqtisodiy jarayonlar o‘rtasida bog‘lanishlarni o‘rganishda quyidagi chiziqsiz funksiyalar bilan foydalaniladi:

Ikkinchi darajali parabola – $y = a_0 + a_1 x + a_2 x^2$

Uchinchi darajali parabola – $y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$

n-darajali parabola – $y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$

Giperbola – $y = a_0 + \frac{a_1}{x}$

b- darajali giperbola – $y = a_0 + \frac{a_1}{x^b}$

Logarifmik – $logy = a_0 + a_1 x$

Yarim logarifmik – $y = a_0 + a_1 lnx$

Ko‘rsatkichli funksiya – $y = a_0 \cdot a_1^x$

Darajali funksiya – $y = a_0 \cdot x_1^{a_1}$

Logistik funksiya – $y = \frac{a_0}{1+a_1e^{-bx}}$

Regressiya funksiyasining turini tanlab, ya'ni ko‘rib chiqilayotgan modelning Y ga bog‘liqligi X (yoki X ning Y ga bog‘liqligi), masalan, chiziqli model $y = a + bx$, model koeffitsiyentlarining o‘ziga xos qiymatlarini aniqlash kerak.

a va b ning har xil qiymatlari uchun $y = a + bx$ shaklidagi cheksiz ko‘p bog‘liqliklar qurilishi mumkin, ya'ni koordinata tekisligida cheksiz ko‘p to‘g‘ri chiziqlar bor, lekin bizga shunday bog‘liqlik kerak eng yaxshi tarzda kuzatilgan qiymatlarga mos keladi.

Shunday qilib, vazifa eng yaxshi koeffitsiyentlarni tanlashgacha kamayadi. a+bx chiziqli funksiyani faqat ma'lum miqdordagi kuzatuvlar asosida qidiramiz. Kuzatilgan qiymatlarga eng mos keladigan funksiyani topish uchun biz eng kichik kvadratlar usulidan foydalanamiz.

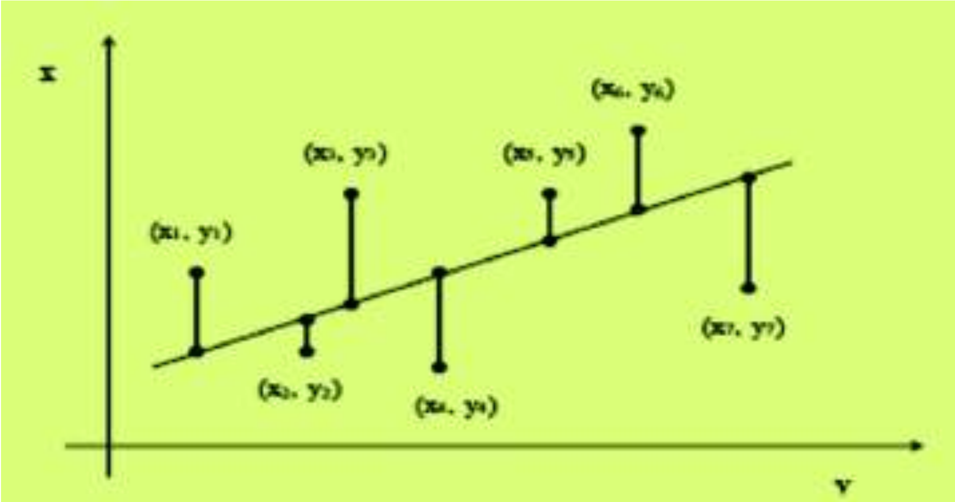
LSMning mohiyati quyidagicha: korrelyatsiya maydonidagi eksperimental nuqtalar orqali chizish mumkin bo‘lgan chiziqlar to‘plamidan y=b1+b0x regressiya chizig‘i tanlanadi, shunda vertikal masofalar kvadratlari yig‘indisi tajriba nuqtalari va bu chiziq eng kichigidir.

Tajriba nuqtalari va regressiya chizig‘i orasidagi masofalar ei og‘ishidir.

5.4. Korrelyatsion-regression tahlilda eng kichik kvadratlar usulining qo‘llanilishi

Funksiyalar parametrlari odatda **“eng kichik kvadratlar”** usuli bilan aniqlanadi. Eng kichik kvadratlar usulini mazmuni quyidagicha: haqiqiy miqdorlarning tekislangan miqdorlardan farqining kvadratlari yig‘indisi eng kam bo‘lishi zarur (5.2-rasm):

$$\sum (y - y_x)^2 \rightarrow min$$



2-rasm. Eng kichik kvadratlar usulining grafikli ko‘rinishi

Bir omilli chiziqli bog‘lanishni olaylik:

$$y_x = a_0 + a_1 x$$

Qiymat $\sum(y - y_x)^2$ eng kam bo‘lishi uchun birinchi darajali xosilalar nolga teng bo‘lishi kerak:

S=$\sum(y - y_x)^2 - \sum(y - a_0 - a_1\, x)^2 \rightarrow min$

$$\frac{\partial S}{\partial a_0} = 0 \qquad \frac{\partial S}{\partial a_1} = 0$$

Bir necha o‘zgarishlardan so‘ng eng kichik kvadratlar usulining normal tenglamalar tizimi hosil bo‘ladi.

$$\begin{cases} n \cdot a_0 + a_1 \cdot \sum x = \sum y \\ a_0 \cdot \sum x + a_1 \cdot \sum x^2 = \sum y \cdot x \end{cases}$$

Quyidagi ifoda bilan foydalanib

$$n\bar{x} - \sum x_1 n\bar{y} - \sum ynx^2 - \sum x^2 nx\,\bar{y} - \sum xy$$

$$\begin{cases} a_0 + a_1\bar{x} - \bar{y} \\ ax + a_1 x^2 - xy \end{cases}$$

Yuqoridan $a0$ va $a1$ parametrlar quyidagi formulalar bilan aniqlanadi:

$$a_0 = \bar{y} - a_1\bar{x}a_1 - \frac{x \cdot \bar{y} - \bar{x}\bar{y}}{x^2 - x^{-7}}$$

Chiziqli regressiyaning parametrlar sharxini ko‘rib chiqamiz.

Omilli o‘zgaruvchi b koeffitsiyenti, u x omili bir birlikka o‘zgarsa y ning o‘rtacha hisobda qanchalarga o‘zgarishini ko‘rsatadi.

Misol uchun tasavur qilaylik, xarajat bilan ishlab chiqarilgan mahsulot hajmi o‘rtasida bog‘liqligi quyidagini tashkil etsa:

$$y = 35000 + 0{,}58x$$

U holda, ishlab chiqarish hajmni 1 birlikka oshirish uchun bizdan 580 so‘m qo‘shimcha xarajatni talab etadi.

5.5. O‘rtacha elastiklik koeffitsiyenti

Tadqiq etilayotgan ko‘rsatkichlar birliklarining bir-biridan farq qilishi tufayli a_1 parametrdan omilning natijaviy omil belgiga ta'sirini bevosita baholash uchun foydalanib bo‘lmaydi. Ushbu maqsadlarda elastiklik koeffitsiyenti va beta-koeffitsiyent hisoblab chiqiladi. Elastiklik koeffitsiyentini aniqlash formulasi quyidagicha:

$$Э_{yx} = a_1 \frac{\bar{x}}{\bar{y}}$$

Elastiklik koeffitsiyenti x omil bir foizga o‘zgarganda y natijaviy omil qancha foizga o‘zgarishini ko‘rsatadi. Beta-koeffitsiyentni aniqlash formulasi:

$$\beta_{yx} = a_1 \frac{\sigma_x}{\sigma_y}$$

bu yerda, σ_x va σ_y - x va y tasodifiy kattaliklarning o‘rtacha kvadratik og‘ishlari.

Beta-koeffitsiyent omili o‘zining o‘rtacha kvadratik og‘ishi miqdoriga o‘zgarganda, natijaviy omilning qiymati o‘zining o‘rtacha kvadratik og‘ishining o‘rtacha qancha qismiga o‘zgarishini ko‘rsatadi.

Nazorat uchun savollar

1. Regressiya koeffitsiyentlarining iqtisodiy mohiyati nimadan iborat?
2. “Eng kichik kvadratlar usuli” ning mohiyatini tushuntirib bering.
3. Normal tenglamasini yechish usullarini tushuntirib bering.
4. Yalpi korrelyatsion-regression tahlilni o‘tkazish bosqichlari.
5. Bir darajali regression model va uni tuzishga nisbatan qo‘yiladigan talablar.

Masala. Davlat yangi turdagi lotoreya biletlarini sotuvga chiqarishdan avval olinadigan o‘rtacha foyda taxminan qancha bo‘lishini bilmoqchi. Nazorat bo‘limida avval sotilgan ana shu lotoreyalardan olingan foyda miqdori haqida ma’lumotlar bor. Ularni tahlil qilib, hukumat nazorat bo‘limiga o‘z maslahatingizni berishingizni so‘raymiz.

№	Y	X
1	142,3	51,6
2	1400	504
3	113,3	35,7
4	489	214
5	45,9	44578
6	118	40
7	1130	553
8	94,5	44830
9	58	44610
10	760	332
11	1270	410
12	1005	410
13	177	80
14	58,7	44762
15	1100	472,2
16	1460	666,8
17	1070	376
18	100,4	33,2
19	1338	570
20	57,9	44702
21	44645	8
22	194	78
23	66	28

Topshiriq. Berilgan ma’lumotlar asosida regression tahlil amalga oshirilsin.

Yechim. Ushbu masalani regression tahlil qilib, model tuzishimiz uchun, biz eng kichik kvadratlar usulining normal tenglamalar sistemasidan foydalanamiz:

$$\begin{cases} n \cdot a_0 + a_1 \cdot \sum x = \sum y \\ a_0 \cdot \sum x + a_1 \cdot \sum x^2 = \sum y \cdot x \end{cases}$$

Buning uchun avvalo, excelni ishga tushirib ma’lumotlarni yuqoridagi ketma-ketlikda kiritib chiqamiz. Bu yerda asosiy omil sifatida avval olingan foydani va unga ta’sir etuvchi omil sifatida sotilgan biletlar sonini ishlatamiz.

1. Ma’lumotlarni kiritganimizdan so‘ng excel funksiyasidan foydalanib, ya’ni Сумма buyrug‘i yordamida x va y qiymatlarimizning summasi (yig‘indisi) ni topib olamiz.

2. Endilikda keyingi bo‘sh ustunga ta’sir etuvchi omil (x) ning kvadratga oshirilgan qiymatini hisoblaymiz, ushbu amalni bir katakcha uchun qo‘llaymiz va buyruq orqali qolgan barcha x lar uchun ham hisobni amalga oshiramiz:

C2 =C2^2

	A	B	C	D	E
1	№	Y	X	X^2	
2	1	142,3	51,6	=C2^2	
3	2	1400	504		
4	3	113,3	35,7		
5	4	489	214		
6	5	45,9	44578		
7	6	118	40		
8	7	1130	553		
9	8	94,5	44830		
10	9	58	44610		
11	10	760	332		
12	11	1270	410		
13	12	1005	410		
14	13	177	80		
15	14	58,7	44762		
16	15	1100	472,2		

3. Ana endi yangi ustunga x va y ko‘paytmalarini hisoblaymiz:

	A	B	C	D	E	F
1	№	Y	X	X^2	X*Y	
2	1	142,3	51,6	2662,56	=C2*B2	
3	2	1400	504	254016		
4	3	113,3	35,7	1274,49		
5	4	489	214	45796		
6	5	45,9	44578	1987198084		
7	6	118	40	1600		
8	7	1130	553	305809		
9	8	94,5	44830	2009728900		
10	9	58	44610	1990052100		
11	10	760	332	110224		
12	11	1270	410	168100		
13	12	1005	410	168100		
14	13	177	80	6400		
15	14	58,7	44762	2003636644		
16	15	1100	472,2	222972,84		
17	16	1460	666,8	444622,24		

4. Shuningdek, x va y ko‘paytmalarini ham summasini hisoblaymiz. So‘ngra, biz uchun kerak bo‘lgan barcha natijalar topildi, ana endi bu natijalarni eng kichik kvadratlar usulining normal tenglamalar sistemasi o‘rniga qo‘yib chiqamiz:

$$\left\{\begin{array}{l} 23*a0+a1*228344,5=56893 \\ 228344,5*a0*9991*a1=19771 \end{array}\right.$$

5. So‘ngra, ushbu sistemani Kramer usulidan foydalanib hisoblaymiz. Bu usul quyidagicha hisoblaniladi:

$$\begin{cases} 23*a0+a1*228344,5=56893 \\ 228344,5*a0*9991*a1=19771 \end{cases}$$

delta=	23	228344,5
	228344,5	9991090419
delta1=	56893	228344,5
	19771595	9991090419
delta2=	23	56893
	228344,5	19771595,2

6. Deltalarni topishda excel funksiyasidan foydalanamiz, ya’ni МОПРЕД buyrug‘i orqali har bir deltalar qiymatlarini belgilab hisoblaymiz:

$$\begin{cases} 23*a0+a1*228344,5=56893 \\ 228344,5*a0*9991*a1=19771 \end{cases}$$

delta=	23	228344,5	=МОПРЕД(I6:J7)
	228344,5	9991090419	
delta1=	56893	228344,5	
	19771595	9991090419	
delta2=	23	56893	
	228344,5	19771595,2	

7. Ana endi ushbu a0=delta1/delta va a1=delta2/delta formulalar yordamida parametrlarimizni topib olamiz. Buning uchun quyidagi amallarni bajaramiz:

$$\begin{cases} 23*a0+a1*228344{,}5=56893 \\ 228344{,}5*a0*9991*a1=19771 \end{cases}$$

delta=	23	228344,5	1,77654E+11
	228344,5	9991090419	
delta1=	56893	228344,5	5,63908E+14
	19771595	9991090419	
delta2=	23	56893	-12536456950
	228344,5	19771595,2	
a0=	=K9/K6		
a1=			

$$\begin{cases} 23*a0+a1*228344{,}5=56893 \\ 228344{,}5*a0*9991*a1=19771 \end{cases}$$

delta=	23	228344,5	1,77654E+11
	228344,5	9991090419	
delta1=	56893	228344,5	5,63908E+14
	19771595	9991090419	
delta2=	23	56893	-12536456950
	228344,5	19771595,2	
a0=	3174,197		
a1=	=K12/K6		

Shunday qilib parametrlarimizning natijalarni oldik va bular quyidagicha:

a0=	3174,197
a1=	-0,07057

Shundan kelib chiqib, modelimiz ko‘rinishi quyidagicha bo‘ladi:

$$y=3174{,}19-0{,}07*x+\varepsilon$$

8. Ana endi elastiklik qiymatini topamiz va buning uchun quyidagi formuladan foydalanamiz:

$$Э_{yx}=a_1\frac{\bar{x}}{\bar{y}}$$

Ushbu formuladan kelib chiqib x va y uchun o‘rtacha qiymatlarini topishda x va y ning summalarini jami kuzatuvlar soniga bo‘lamiz:

19	18	100,4	33,2	1102,24	3333,28
20	19	1338	570	324900	762660
21	20	57,9	44702	1998268804	2588246
22	21	44645	8	64	357160
23	22	194	78	6084	15132
24	23	66	28	784	1848
25	Summa	56893	228344,5	9991090419	19771595
26	O'rtacha	=B25/A24			

Natijada quyidagi qiymatlarga erishamiz:

№	Y	X
Summa	56893	228344,5
O‘rtacha	2473,609	9928,022

9. Endilikda barcha hisoblangan qiymatlarimizni elastiklik formulasiga qo‘yib hisoblaymiz:

A	B	C	D	E	F	G	H	I	J	K
5	45,9	44578	1987198084	2046130			delta=	23	228344,5	1,77654E+11
6	118	40	1600	4720				228344,5	9991090419	
7	1130	553	305809	624890						
8	94,5	44830	2009728900	4236435			delta1=	56893	228344,5	5,63908E+14
9	58	44610	1990052100	2587380				19771595	9991090419	
10	760	332	110224	252320						
11	1270	410	168100	520700			delta2=	23	56893	-12536456950
12	1005	410	168100	412050				228344,5	19771595,2	
13	177	80	6400	14160						
14	58,7	44762	2003636644	2627529			a0=	3174,197		
15	1100	472,2	222972,84	519420			a1=	-0,07057		
16	1460	666,8	444622,24	973528						
17	1070	376	141376	402320			y=3174,19-0,07*x+ε			
18	100,4	33,2	1102,24	3333,28						
19	1338	570	324900	762660			Elastiklik=	=I16*(C26/B26)		
20	57,9	44702	1998268804	2588246						
21	44645	8	64	357160						
22	194	78	6084	15132						
23	66	28	784	1848						
Summa	56893	228344,5	9991090419	19771595						
O'rtacha	2473,609	9928,022								

10. So‘ngra quyidagi natijaga erishamiz:

Elastiklik koeffitsiyenti =-0,28323

Xulosa: Agar biletlarning soni 1 birlikka oshsa, avval sotilgan biletlardan olingan foyda 0,07ga kamayadi.

Bu yerda a0 modelimizning konstantasi hisoblanadi. Agar bilet soni 0 ga tenglashsa, avval olingan foyda 3174 ga tenglashadi. Shuningdek, a0 modelimizda ishtirok etmagan ta’sir etuvchi omillarning miqdori ham hisoblanadi.

Shuningdek, bilet sonining 1 foizga oshishi, avval olingan foydaning 0,28 % ga kamayishini izohlaydi.

VI BOB. NORMAL TAQSIMOT

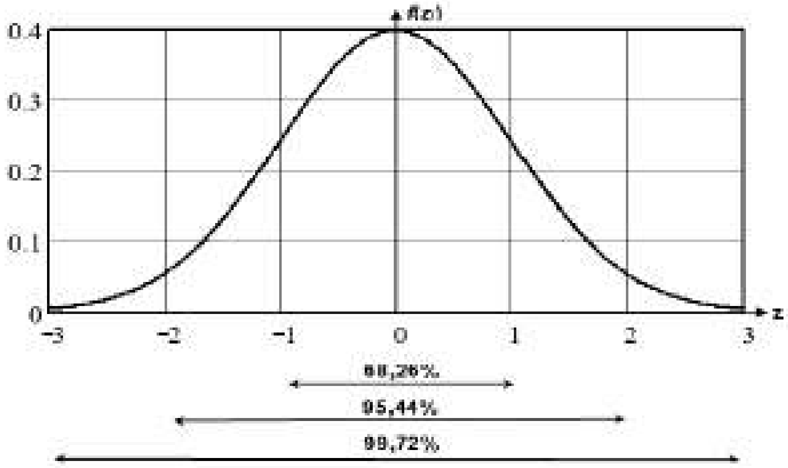
6.1. Standart normal taqsimot

Normal taqsimot - ehtimollar nazariyasidagi muhim taqsimotlardan biri bo‘lgan tasodifiy miqdorlar taqsimoti (a — ixtiyoriy haqiqiy son, a>0). Normal taqsimot (1) ga bo‘ysungan % tasodifiy miqdorning o‘rta qiymati a ga, dispersiyasi a2 ga teng bo‘ladi: M2; = a, Dt, = a2. Normal taqsimot x = a nuqtaga nisbatan simmetriyaga ega. O‘zaro bog‘liq bo‘lmagan i;,, i;2, ..., !;,, ning taqsimoti (juda keng shartlarda) Normal taqsimotga yaqin bo‘lishi isbotlangan (qarang Limit teoremalar).

Biror tasodifiy miqdorni katta sondagi o‘zaro bog‘liqmas sabablarning natijasi deb qarash tatbiqlarda ko‘p uchraganligi uchun Normal taqsimot ehtimollar nazariyasi va tabiatshunoslikda katta ahamiyatga ega. Normal taqsimotning vujudga kelishiga klassik namunalar K.Gauss (kuzatish xatolari taqsimoti qonuni) va J.Maksvell (molekulalar tezliklari taqsimoti qonuni) ga tegishli.[1]

Ommaviy ma'lumotlardan foydalanganda axborotga qo‘yiladigan eng muhim talab uning sifat va miqdoriy bir xilligi. Sifatli bir xillik kuzatuvlar yoki bir-biriga o‘xshash narsalar tekshirilishini nazarda tutadi. O‘xshash bo‘lmagan narsalardan foydalanish individual xususiyatlar o‘rtasidagi munosabatlarning xususiyatlarini buzadi.

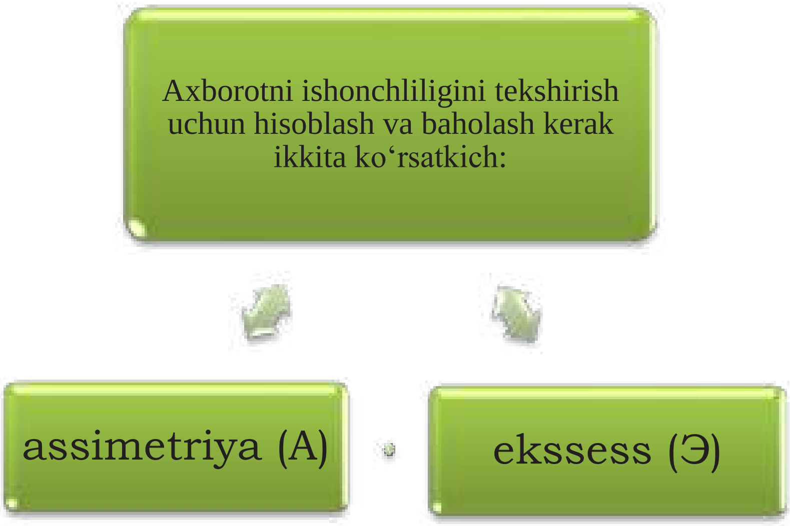
Ko‘pgina iqtisodiy ko‘rsatkichlar bo‘yicha ma'lumot taqsimoti odatdagiga yaqin. Oddiy taqsimot bir qator kuzatuvlardan olinadi, ularning o‘zgarishi ko‘p sonli kichik, tasodifiy yoki tasodifiy ta'sirlarning ta'siri bilan bog‘liq (6.1-rasm).



6.1-rasm. Oddiy taqsimot maydoni

Normal taqsimot belgining ko‘pincha uning o‘rtacha qiymatiga eng ko‘p yaqin ekanligini ko‘rsatadi

O‘rtacha qiymatdan uzoqlashganda, kuzatishlar soni yoki voqea yuzaga kelishi ehtimoli kamayadi. Bundan tashqari 68,26 % holatlar $(\bar{x}-\sigma_x)$dan $(\bar{x}-3\sigma_x)$ gacha, 95,46 % - $(\bar{x}-2\sigma_x)$ dan $(\bar{x}+2\sigma_x)$ gacha, 99,73 %- $(\bar{x}-3\sigma_x)$ dan $(\bar{x}+3\sigma_x)$ intervalgacha tushadi. Ishlarning aksariyati normal taqsimot bilan so‘nggi intervalgacha tushadi. Ekonometrik modellarni yaratish uchun foydalaniladigan dastlabki ma'lumotlar ishonchli bo‘lishi kerak.



Assimetriya (A) va ekssess (Э). Ushbu ko‘rsatkichlar quyidagi formulalar bilan aniqlanadi:

$$A = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^3}{n\sigma_x^3};$$

$$Э = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^4}{n\sigma_x^4} - 3,$$

Bu yerda:

$X_i -$ indikatorning haqiqiy qiymati;

$X -$ koʻrsatkichning oʻrtacha qiymati;

$n -$ tajribalar soni;

$\sigma_x -$ hisoblangan standart ogʻish formulaga muvofiq

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n}}.$$

Agar A va Э parametrlari nolga teng boʻlsa, unda dastlabki ma'lumotlar toʻliq ishonchli hisoblanadi. Bunday mukammal variant kamdan-kam hollarda iqtisodiy hodisa va jarayonlarni oʻrganishda duch keladi.

Agar assimetriya ijobiy qiymatga ega boʻlsa, unda mos keladigan ehtimollik maydoni normal taqsimot grafigida nisbatan oʻng tomonga siljiydi. Salbiy assimmetriya koeffitsiyenti bilan grafik chapga siljiydi.

Ekssessning oʻziga xos qiymati vertikal ehtimollik taqsimoti grafigini siljitish orqali aniqlanadi. Xususan, agar ehtimollik maydoni eng yuqori darajaga koʻtariladi, keyin Э > 0. Oʻz navbatida,Э koeffitsiyentining pasayishi oʻrganilayotgan grafik tobora tekislanib borishiga olib keladi.

Shu munosabat bilan, ogʻishlarning mumkin boʻlgan chegaralari toʻgʻrisida savol tugʻiladi nol qiymatlardan A va Э koeffitsiyentlari. Axborot mumkin ishonchli deb hisoblaydi va agar keyingi ishlov berish uchun mos boʻlsa quyidagi ikkita tengsizlik mavjud:

$$|A| \leq 3\sigma_A$$
$$|Э| \leq 5\sigma_Э$$

Yuqoridagi formulalarda σ_A va $\sigma_Э$ xatolar formulalar bilan belgilanadigan assimetriya va ekssess:

$$\sigma_A = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}};$$

$$\sigma_Э = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}.$$

Asimmetriya va ekssess xatolari faqat n tajribalar soniga bog‘liq. Axborotni ishonchliligini tekshirish bosqichida tavsiya etiladi. O‘rganilayotgan populyatsiyadan keskin farq qiladigan ma'lumotlarni olib tashlash - ko‘rsatkichlarning juda yuqori yoki past ko‘rsatkichlari. Bunday qadriyatlarning mavjudligi haqiqat tomonidan ko‘rsatiladi. Egri chiziq va kurtozning ruxsat etilgan chegaralar koeffitsiyentlari. Uchta sigma qoidasi ushbu "nostandart" qiymatlarni ta'kidlashga imkon beradi:

$$|x_i - \bar{x}| \leq 3\sigma_x$$

Agar biron bir qiymat uch sigma qoidasini qondirmasa, u o‘chirilishi kerak. Barcha 8 qator bir vaqtning o‘zida o‘chirilishi kerak ushbu qiymatga tegishli bo‘lgan (tajriba, obyekt, korxona va boshqalar). O‘rganilayotgan populyatsiyadan olib tashlanishi kerak bo‘lgan obyektlar batafsil monografik tahlildan o‘tkaziladi. Keraksiz tajribalarni olib tashlaganingizdan so‘ng, yangisini hisoblash maqsadga muvofiqdir skewness va kurtosis qiymatlari va qolgan ma'lumotlarning ishonchliligini tekshirish.

Qabul qilingan taqsimot qonuni haqidagi gipotezani sinovdan o‘tkazish rozilik mezonlari deb nomlangan holda amalga oshirilishi mumkin. Umumjahon x^2 Pirson testi, asosan, kuzatilgan namunaning ba'zi nazariy qonunlarga tegishli ekanligi haqidagi gipotezani tekshirish uchun ishlatiladigan mezon sifatida tarqatish. Pearsonning mezonlari savolga javob beradi, farq qiladi empirik va nazariy chastotalar bo‘ladimi. Pearson mezonining mohiyati x^2 mezonini hisoblashdan iborat formulaga muvofiq.

$$x^2 = \frac{\sum_{i=1}^{k}(n_i - n_i')^2}{n_i'}$$

Bu yerda: k - kuzatilgan qiymatlarning bitlar soni;

n_i - tanlanmadan olingan empirik chastotalar;

n_i' - nazariy jihatdan topilgan nazariy chastotalar.

Farq qancha kichik bo‘lsa $(n_i - n_i')$, empirik taqsimot nazariy jihatdan shunchalik yaqin bo‘ladi.

Pearson mezonidan foydalangan holda hisoblash algoritmini bajarish quyidagilardan iborat:

1. Namunaviy ma'lumotlarga asoslanib, statistik taqsimotni oling, kuzatilgan xususiyat.

2. Xususiyatning nazariy chastotalarini hisoblang, ular nima agar atribut haqiqatan ham ushbu qonunga muvofiq taqsimlangan bo‘lsa.

3. Yuqoridagi formuladan foydalanib, x_{his}^2 mezonning empirik qiymatini hisoblang.

4. Pirson mezonining kritik qiymatlari jadvaliga ko‘ra aniqlang qiymat -x_{jad}^2 zarur bo‘lgan ahamiyatlilik darajasida α berilgan raqam uchun erkinlik darajasi s. Erkinlik darajasi soni formula bo‘yicha hisoblanadi:

$$s = k - 1 - r,$$

Bu yerda: k - kuzatilgan qiymatlarning bitlar soni;

r - faraz qilingan taqsimot parametrlari soni (holda) normal yoki bir xil taqsimot (r=2).

5. Agar $x_{his}^2 < x_{jad}^2$ asosiy gipotezani qabul qiling, bu holda ma'lum bir ahamiyatlilik darajasida, o‘rganilayotgan parametrning statistik taqsimoti normal taqsimot qonuniga bo‘ysunadi deb ta'kidlash mumkin.

Agar bu $x_{his}^2 \geq x_{jad}^2$ tengsizlikning aksi bo‘lsa, alternativ gipotezani qabul qiladi: statistik taqsimot odatdagidan farq qiladi.

6.2. Z - statistika va uning mohiyati

Z - statistikasi - bu noaniq gipotezaga qarshi muqobil gipotezani sinash uchun ishlatiladigan statistik jarayondir. Bu har xil statistik gipoteza, agar dispersiya ma'lum bo‘lsa va namuna katta bo‘lsa, ikkita namunaning vositasi boshqacha ekanligini aniqlash uchun ishlatiladi. Tanlov va populyatsiya o‘rtasida sezilarli farq borligini Z testi aniqlaydi.

Z-test odatda katta namunalar bilan bog‘liq muammolarni hal qilish uchun ishlatiladi. Bu aralashuvning "z testi" drayveri standart normal taqsimotdan iborat va "Z" standart oddiy tasodifiy o‘zgaruvchini belgilash uchun ishlatiladigan an'anaviy belgidir. Z-test formulasi tanlangan o‘rtacha minus populyatsiyaning standart og‘ish va namunaviy o‘lchamiga bo‘linadi. Agar namuna o‘lchami 30 birlikdan oshsa, u holda z-testini o‘tkazish kerak. Matematik jihatdan, test formulasi z shaklda berilgan

$$Z - Test\ = (\bar{x} - \mu)/\left(\sigma/\sqrt{n}\right)$$

Bu yerda,

$\bar{x}$ = namunaning o‘rtacha qiymati

μ = o‘rtacha aholi soni

σ = aholi standart og‘ishi

n = kuzatuvlar soni

Z -test statistikasi formulasiga misol

Aytaylik, investor bitta kompaniyaning kunlik o‘rtacha daromadliligini 1%dan ortiq tahlil qilmoqchi, yoki emasmi? Shunday qilib, investorlar 50 tasodifiy tanlovni tanlab, daromadni o‘rtacha 0,02 bilan hisobladilar, sarmoyadorlar esa o‘rtacha og‘ishning 0,025 deb hisobladilar.

Shunday qilib, bu holda, nol gipoteza - o‘rtacha 3%, muqobil gipoteza - o‘rtacha daromad 3%dan katta. Investorlar, alfa 0,05% ikki tomonlama test sifatida tanlanadi va har bir quyruqdagi namunaning

0,025% ni tanlaydi va kritik alfa qiymati 1,96 yoki -1,96 ni tashkil qiladi.

	A	B
1	Sample Mean ($\bar{x}$)	0.02
2	Population Mean (μ)	1%
3	Standard Deviation of population (σ)	0.025
4	Number of Observation (n)	50
5		

Z testi = (0,02 - 1%) / (0,025 / $\sqrt{50}$)

Z testi = 2,83

Shunday qilib, yuqoridagi hisob-kitoblardan kelib chiqib, investorlar xulosaga keladi va u bekor qilingan gipotezani rad etadi, chunki z natijasi 1,96 dan katta va bir kunlik oʻrtacha daromadlilik 1%dan ortiq boʻlgan tahlilga keladi.

Tushuntirish

- Birinchidan, namuna oʻrtacha qiymatini aniqlang (bu barcha tasodifiy namunalarning oʻrtacha ogʻirligi).
- Populyatsiyaning oʻrtacha qiymatini aniqlang va undan namunaviy oʻrtacha qiymatni olib tashlang.
- Keyin bu qiymatni kuzatuvlar sonining kvadrat ildiziga boʻlinadigan standart ogʻish bilan ajrating.
- Yuqoridagi amallarni bajargandan soʻng, testning z statistikasi natijalari hisoblab chiqiladi.

Z-testi oddiy tasodifiy oʻzgaruvchining oʻrtacha qiymatini berilgan qiymat bilan solishtirish uchun ishlatiladi. Z-testi foydalidir yoki namuna 30 dan oshganda va populyatsiyaning dispersiyasi ma'lum boʻlganda ishlatilishi kerak. Z-testi oʻrtacha namunaning taqsimlanishi normal boʻlsa yaxshi boʻladi. Agar z-testda ma'lum shartlar bajarilgan boʻlsa qoʻllaniladi, aks holda biz boshqa testlardan foydalanishimiz kerak va z-testda hech qanday tebranishlar yoʻq. Bir agentli z-testi ma'lum bir populyatsiyaning oʻrtacha gipotezasini tekshirish uchun ishlatiladi. Z-testi statistik gipotezani tekshirishning asoslaridan biri bo‘lib, ko‘pincha kirish darajasida o‘qitiladi. Ma'lumki, z-testlar binomial va Puasson kabi boshqa taqsimotdan ma'lumotlar ishlab chiqarilganda qo‘llanilishi mumkin.

6.3. Normal taqsimlash zichligi

Oddiy taqsimlash eng keng tarqalgan taqsimot turi hisoblanadi. Oddiy taqsimotning asosiy xususiyati shundaki, bu boshqa tarqatishlar yaqinlashadigan chegara. Faqat doimiy tasodifiy o‘zgaruvchilar normal qonunga bo‘ysunadi. Oddiy tarqatish qonunining zichligi ko‘rib chiqiladi.

Oddiy taqsimotni ratsion bilan standartga aylantirish kerak:

z - buning o‘rniga ishlatiladigan yangi o‘zgaruvchi *x;*

m - kutilayotgan qiymat;

σ - standart og‘ish.

Ma'lumot namunasi uchun baholash quyidagilar orqali hisoblanadi:

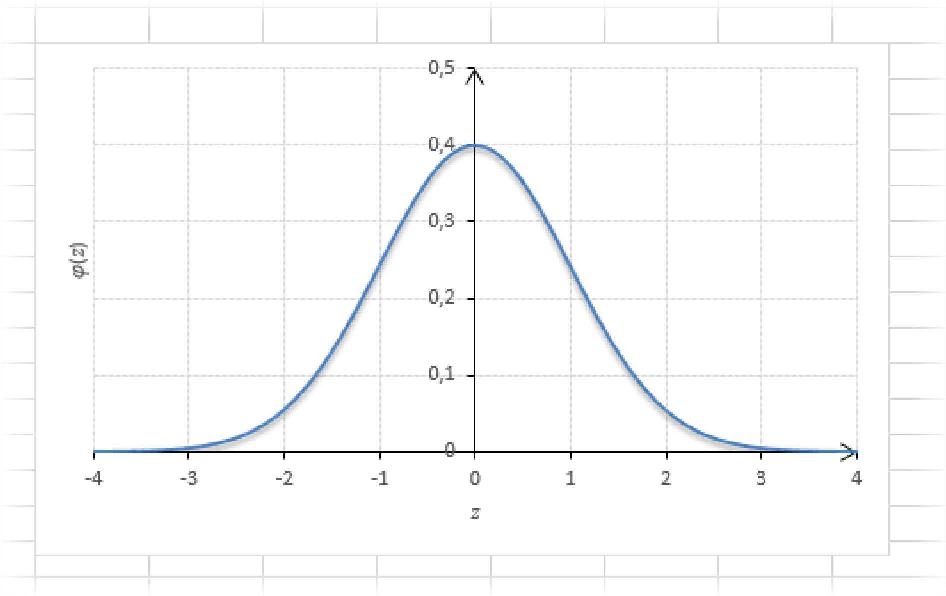
O‘rtacha o‘zgaruvchini o‘rtacha arifmetik va tarqalish (*z*). Hozir u mos ravishda 0 va 1 ga teng. Bu oddiy algebraik o‘zgarishlar yordamida oson ta'minlanadi. Namunaviy tarqatish ko‘rinishi zichligi (uchun *z-hisob-kitoblar*) ya’ni Gauss funksiyasi:

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

taqsimlash zichligi funksiyasi:

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Zichlik grafigi:



Markaz kutilganidek, 0 nuqtada joylashgan. Xuddi shu nuqtada, Gauss funksiyasi maksimal darajada yetib boradi, bu esa uning o‘rtacha qiymatining tasodifiy qiymatini qabul qilishga mos keladi. Ushbu nuqtada zichligi 0,3989.

Shunday qilib, grafikka muvofiq, o‘rtadan kichik og‘ishlar boshqalarga qaraganda tez-tez tushish va markazdan uzoqroq bo‘lgan qadriyatlar kamroq keng tarqalgan ekanligi aniq ko‘rinadi. Absissa o‘qining shkalasi standart og‘ishlarda o‘lchanadi, bu o‘lchov birliklaridan tortib olinadi va normal taqsimotning umumbashariy tuzilishini oladi. Gauss egri normallashtirilgan ma'lumotlar normal taqsimotning boshqa xususiyatlarini mukammal darajada namoyish etadi.

Oddiy tarqatish funksiyasi sizga ehtimollikni hisoblash imkonini beradi.

$$P(Z < z) = \Phi(z) = \int_{-\infty}^{z} \varphi(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$$

Bevosita hammasi hisoblab chiqilgan va aniqlangan maxsus jadvallarda statistikadagi har qanday darslik oxirida joylashgan.

Значение функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
	Сотые доли x									
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3721	3605	3588	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3411	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1827	1804	1781	1759	1736
1,3	[illegible]	1692	[illegible]	[illegible]	1626	[illegible]	1582	1561	1540	[illegible]

Gauss funksiyasi belgilangan o‘qqa nisbatan nosimmetrikdir.

6.4. Styudent t-taqsimot mezoni

Regressiya va korrelyatsiya koeffitsiyentlarining statistik ahamiyatini baholash uchun Styudentning t-mezoni va har bir ko‘rsatkich uchun ishonchlilik intervallari aniqlanadi.

Ko‘rsatkichlarning tasodifiy tabiatiga nisbatan H_0 gipoteza, ya’ni ularning noldan ahamiyatsiz farqliligi to‘g‘risida. Regressiya va korrelyatsiya koeffitsiyentlarining ahamiyati Styudent t-mezonidan foydalangan holda ularning qiymatlarini ularning tasodifiy xatosi bilan taqqoslash yo‘li bilan amalga oshiriladi:

$$t_b = \frac{b}{m_b};\ t_a = \frac{a}{m_a};\ t_r = \frac{r}{m_r}$$

Tasodifiy xatolar chiziqli regressiya parametrlarining va korrelyatsiya koeffitsiyenti formulalar bilan aniqlanadi:

$$m_b = \sqrt{\frac{\sum \frac{(y - y_x)^2}{n-2}}{\sum (x - \bar{x})^2}} = \sqrt{\frac{S_{qol}^2}{\sum (x - \bar{x})^2}} = \frac{S_{qol}}{\sigma_x \sqrt{n}};$$

$$m_a = \sqrt{\frac{\sum(y-y_x)^2}{(n-2)}\frac{\sum x^2}{n\sum(x-\bar{x})^2}} = \sqrt{S_{qol}^2\frac{\sum x^2}{n^2\sigma_x^2}} = S_{qol}\frac{\sqrt{\sum x^2}}{n\sigma_x};$$

$$m_{r_{xy}} = \sqrt{\frac{1-r_{xy}^2}{n-2}}$$

t-statistikani haqiqiy t_{haq} va kritik (jadval) t_{jad}qiymatini solishtirib H_0 gipotezani qabul qilamiz yoki rad etamiz.

Fisherning F-mezoni va Styudent t-statistika o‘rtasidagi bog‘liqlik tenglik bilan ifodalanadi.

$$t_r^2 = t_b^2 = \sqrt{F}$$

Agar $t_{jad} < t_{haq}$ bo‘lsa H_0 rad etiladi, ya’ni a, b va r_{xy} tasodifan noldan farq qilmaydi va sistematik ta’sir qiluvchi omil x ta’siri ostida hosil bo‘ladi. Agar $t_{jad} > t_{haq}$, H_0 gipoteza rad etilmaydi va a, b va r_{xy} ning shakllanishining tasodifiy tabiati tan olinmaydi.

Ishonch oralig‘ini hisoblash uchun har bir ko‘rsatkich uchun chekli xato aniqlanadi:

$$\Delta_a = t_{jad}m_a, \quad \Delta_b = t_{jad}m_b$$

Ishonch oralig‘ini hisoblash uchun formulalar quyidagicha:

$$\gamma_a = a \pm \Delta_a;\ \ \gamma_{a_{min}} = a - \Delta_a;\ \ \gamma_{a_{min}} = a + \Delta_a$$

$$\gamma_b = b \pm \Delta_b;\ \ \gamma_{b_{min}} = b - \Delta_b;\ \ \gamma_{b_{min}} = b + \Delta_b$$

Agar ishonch oralig‘i chegarasiga nol tushib qolsa, ya’ni pastki chegara salbiy bo‘lsa va yuqori chegara ijobiy bo‘lsa, unda baholanadigan parametr nolga teng, chunki u bir vaqtning o‘zida ham ijobiy, ham manfiy qiymatlarni qabul qilaolmaydi.

y_p prognoz qiymatlari $y_x = a + bx$ regressiya tenglamasida mos keladigan (prognozlangan) $x_r$qiymatini almashtirish bilan aniqlanadi. m_{y_p} prognozning o‘rtacha standart xatosi quyidagicha hisoblanadi:

$$m_{y_p} = \sigma_{qold}\sqrt{1 + \frac{1}{n} + \frac{\left(x_p - \bar{x}\right)^2}{\sum(x - \bar{x})^2}}$$

bu yerda:

$$\sigma_{qold} = \sqrt{\frac{\sum(y - \bar{y})^2}{n - m - 1}}$$

Va prognozlash uchun ishonch intrevali tuziladi:

$$\gamma_{y_p} = y_p \pm \Delta_{y_p};\ \gamma_{y_{p_{min}}} = y_p - \Delta_{y_p};\ \gamma_{y_{p_{max}}} = y_p + \Delta_{y_p}$$

u yerda

$$\Delta_{y_p} = t_{jadv} * m_{y_p}$$

Nazorat uchun savollar

1. Axborotni ishonchliligini tekshirish uchun qanday ko‘rsatkichlar hisoblanadi va baholanadi?
2. Z - statistikasi bu qanday jarayondir?
3. Normal taqsimlash zichligi funksiyasini ifodalang.
4. Regressiya va korrelyatsiya koeffitsiyentlarining statistik ahamiyatini baholash uchun qanday mezon qo‘llaniladi?
5. Ishonch oralig‘i to‘g‘risida tushuncha bering.

Masala. O‘zbekiston Respublikasi hududlarining 7 ta viloyatlarida ikkita iqtisodiy ko‘rsatkich qiymati keltirilgan: y-umumiy xarajatlarda oziq-ovqatlarga qilingan xarajatlar ulushi, %; x- bir ishchining o‘rtacha ish haqi, so‘m.

O‘zbekiston Respublikasi hududlarining iqtisodiy ko‘rsatkichlari:

N	Viloyatlar	y	x
1	Qoraqalpog‘iston Respublikasi	68,8	45,1
2	Toshkent shahri	61,2	59
3	Toshkent viloyati	59,9	57,2
4	Andijon viloyati	56,7	61,8
5	Namangan viloyati	55	58,8
6	Samarqand viloyati	54,3	47,2
7	Xorazm viloyati	49,3	55,2

Topshiriq. Yuqoridagi jadval asosida:

1. Korrelyatsiya koeffitsiyenti topilsin.
2. Chiziqli regressiya modeli tuzib uning parametrlari t-styudent mezoni orqali baholansin.

Yechim. Excel dasturini ishga tushirib, ma’lumotlarni yuqoridagi tartibda kiritamiz. Bunda asosiy omil (y) sifatida umumiy xarajatlarda oziq-ovqatlarga qilingan xarajatlar ulushini, ta’sir etuvchi omil (x) sifatida bir ishchining o‘rtacha ish haqini belgilaymiz.

1. Avvalo, y va x ning summasi hamda o‘rtacha qiymatlarini excel funksiyasi yordamida hisoblaymiz. So‘ngra, yangi ustunga x ning kvadratga oshirilgan qiymatini hamda yana bir yangi ustunga y va x ning ko‘paytmasini topib olamiz. Shuningdek, ushbu natijalarning summa va o‘rtacha qiymatlarini excel funksiyasi yordamida hisoblaymiz:

B	C	D	E	F
N	y	x	x^2	y*x
1	68,8	45,1	2034,01	3102,88
2	61,2	59	3481	3610,8
3	59,9	57,2	3271,84	3426,28
4	56,7	61,8	3819,24	3504,06
5	55	58,8	3457,44	3234
6	54,3	47,2	2227,84	2562,96
7	49,3	55,2	3047,04	2721,36
Summa	405,2	384,3	21338,41	22162,34
O'rtacha	57,88571	54,9	3048,344	3166,049

2. Keyingi ustunga har bir y qiymatlaridan y o‘rtacha ayirmasining kvadratini hisoblab olamiz:

B	C	D	E	F	G
N	y	x	x^2	y*x	(y-y_)^2
1	68,8	45,1	2034,01	3102,88	=(C2-C10)^2
2	61,2	59	3481	3610,8	
3	59,9	57,2	3271,84	3426,28	
4	56,7	61,8	3819,24	3504,06	
5	55	58,8	3457,44	3234	
6	54,3	47,2	2227,84	2562,96	
7	49,3	55,2	3047,04	2721,36	
Summa	405,2	384,3	21338,41	22162,34	
O'rtacha	57,88571	54,9	3048,344	3166,049	

Bunda barcha qiymatlarimiz uchun alohida hisobni amalga oshirmasligimiz uchun o‘rtacha qiymatni kiritayotganda F4 tugmasini bosib qo‘yamiz va shunda uni shu ustun bo‘yicha o‘zgartirmasdan saqlab qoladi va albatta bu hisobimiz uchun ham summa hamda o‘rtacha qiymatlarini hisoblaymiz:

N	y	x	x^2	y*x	(y-y_)^2
1	68,8	45,1	2034,01	3102,88	119,12163
2	61,2	59	3481	3610,8	10,98449
3	59,9	57,2	3271,84	3426,28	4,0573469
4	56,7	61,8	3819,24	3504,06	1,4059184
5	55	58,8	3457,44	3234	8,3273469
6	54,3	47,2	2227,84	2562,96	12,857347
7	49,3	55,2	3047,04	2721,36	73,71449
Summa	405,2	384,3	21338,41	22162,34	230,46857
O‘rtacha	57,88571	54,9	3048,344	3166,049	32,924082

3. Shuningdek, yuqoridagi amallarni x uchun ham takrorlaymiz:

N	y	x	x^2	y*x	(y-y_)^2	(x-x_)^2
1	68,8	45,1	2034,01	3102,88	119,12163	96,04
2	61,2	59	3481	3610,8	10,98449	16,81
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61
5	55	58,8	3457,44	3234	8,3273469	15,21
6	54,3	47,2	2227,84	2562,96	12,857347	59,29
7	49,3	55,2	3047,04	2721,36	73,71449	0,09
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34
O‘rtacha	57,88571	54,9	3048,344	3166,049	32,924082	34,334286

4. Ana endi korrelyatsiya formulasida ko‘rsatilgan standart og‘ishlarni topish uchun (y-y_)^2 va (x-x_)^2 qiymatlarni excel funksiyasidan foydalanib, ya’ni bo‘sh katakka КОРЕНЬ buyrug‘ini yozib yoniga uchun (y-y_)^2 va (x-x_)^2 qiymatlarni belgilash orqali ildizdan chiqarib, natijalarni topib olamiz:

B	C	D	E	F	G	H
N	y	x	x^2	y*x	(y-y_)^2	(x-x_)^2
1	68,8	45,1	2034,01	3102,88	119,12163	96,04
2	61,2	59	3481	3610,8	10,98449	16,81
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61
5	55	58,8	3457,44	3234	8,3273469	15,21
6	54,3	47,2	2227,84	2562,96	12,857347	59,29
7	49,3	55,2	3047,04	2721,36	73,71449	0,09
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34
O'rtacha	57,88571	54,9	3048,344	3166,0486	32,924082	34,334286
					5,737951	5,8595465

5. Ana endi hisoblangan natijalarimizni korrelyatsiya formulasi o‘rniga qo‘yib hisoblaymiz:

B	C	D	E	F	G	H	I
N	y	x	x^2	y*x	(y-y_)^2	(x-x_)^2	
1	68,8	45,1	2034,01	3102,88	119,12163	96,04	
2	61,2	59	3481	3610,8	10,98449	16,81	
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29	
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61	
5	55	58,8	3457,44	3234	8,3273469	15,21	
6	54,3	47,2	2227,84	2562,96	12,857347	59,29	
7	49,3	55,2	3047,04	2721,36	73,71449	0,09	
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34	
O'rtacha	57,88571	54,9	3048,344	3166,0486	32,924082	34,334286	
					5,737951	5,8595465	
				r(y/x)=	=(F10-(D10*C10))/(H11*G11)		

Demak, natijamiz:

r(y/x)= -0,353257

6. Korrelyatsiya koeffitsiyentini topib oldik, ana endi ushbu masalani regression tahlil qilamiz. Biz bunda eng kichik kvadratlar usulining normal tenglamalar sistemasidan foydalanamiz. Buning uchun avvalo, hisoblangan natijalarni eng kichik kvadratlar usulining normal tenglamalar sistemasi o‘rniga qo‘yib chiqamiz:

$$\begin{cases} 7*a0+a1*384,3=405,2 \\ a0*384,3+a1*21338,4=22162,3 \end{cases}$$

7. So‘ngra ushbu sistemani Kramer usulidan foydalanib hisoblaymiz. Bu usul quyidagicha hisoblanadi:

{ 7*a0+a1*384,3=405,2		
a0*384,3+a1*21338,4=22162,3		
delta=	7	384,3
	384,3	21338,41
delta1=	405,2	384,3
	22162,34	21338,41
delta2=	7	405,2
	384,3	22162,34

8. Deltalarni topishda excel funksiyasidan foydalanamiz, ya’ni МОПРЕД buyrug‘i orqali har bir deltalar qiymatlarini belgilab hisoblaymiz. Ana endi ushbu a0=delta1/delta va a1=delta2/delta formulalar yordamida parametrlarimizni topib olamiz:

{ 7*a0+a1*384,3=405,2			
a0*384,3+a1*21338,4=22162,3			
delta=	7	384,3	1682,38
	384,3	21338,41	
delta1=	405,2	384,3	129336,5
	22162,34	21338,41	
delta2=	7	405,2	-581,98
	384,3	22162,34	
a0=	76,87708		
a1=	-0,34593		

Shundan kelib chiqib, modelimiz ko‘rinishi quyidagicha bo‘ladi:

y^=76,87-0,34*x+ε

9. Endilikda topgan parametrlarimizning ishonchliligini t-styudent mezoni orqali tekshiramiz. Buning uchun quyidagi formulalardan foydalanamiz:

$$t_{a1} = \frac{a1}{m_{a1}};\ t_{a0} = \frac{a0}{m_{a0}};\ t_r = \frac{r}{m_r}$$

Tasodifiy xatolar chiziqli regressiya parametrlarining va korrelyatsiya koeffitsiyenti formulalar bilan aniqlanadi:

$$m_{a1} = \sqrt{\frac{\sum\frac{(y-y_x)^2}{n-2}}{\sum(x-\bar{x})^2}} = \sqrt{\frac{S_{qol}^2}{\sum(x-\bar{x})^2}} = \frac{S_{qol}}{\sigma_x\sqrt{n}};$$

$$m_{a0} = \sqrt{\frac{\sum(y-y_x)^2}{(n-2)}\ \frac{\sum x^2}{n\sum(x-\bar{x})^2}} = \sqrt{S_{qol}^2\frac{\sum x^2}{n^2\sigma_x^2}} = S_{qol}\frac{\sqrt{\sum x^2}}{n\sigma_x};$$

$$m_{r_{xy}} = \sqrt{\frac{1-r_{xy}^2}{n-2}}$$

10. Avvalo, a1 parametrimizni tekshiramiz. Buning uchun regression modelimizdan foydalanib, hisoblangan y qiymatimizni topib olamiz:

N	**y**	**X**	**x^2**	**y*x**	**(y-y_)^2**	**(x-x_)^2**	**y^**
1	68,8	45,1	2034,01	3102,88	119,12163	96,04	61,275795
2	61,2	59	3481	3610,8	10,98449	16,81	56,467415
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29	57,090083
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61	55,498821
5	55	58,8	3457,44	3234	8,3273469	15,21	56,536601
6	54,3	47,2	2227,84	2562,96	12,857347	59,29	60,549349
7	49,3	55,2	3047,04	2721,36	73,71449	0,09	57,781936

11. So‘ngra (y-y^)^2 qiymatni, ya’ni qoldiqlarimizni hisoblab, summasini topib olamiz:

N	y	X	x^2	y*x	(y-y_)^2	(x-x_)^2	y^	(y-y^)^2
1	68,8	45,1	2034,01	3102,88	119,12163	96,04	61,275795	56,6136608
2	61,2	59	3481	3610,8	10,98449	16,81	56,467415	22,3973588
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34		201,708234

12. Ana endi m_{a1} ni excel funksiyasidan foydalanib, КОРЕНЬ buyrug‘i yordamida quyidagicha topib olamiz:

N	y	x	x^2	y*x	(y-y_)^2	(x-x_)^2	y^	(y-y^)^2
1	68,8	45,1	2034,01	3102,88	119,12163	96,04	61,275795	56,6136608
2	61,2	59	3481	3610,8	10,98449	16,81	56,467415	22,3973588
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29	57,090083	7,895633
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61	55,498821	1,44283167
5	55	58,8	3457,44	3234	8,3273469	15,21	56,536601	2,36114119
6	54,3	47,2	2227,84	2562,96	12,857347	59,29	60,549349	39,0543646
7	49,3	55,2	3047,04	2721,36	73,71449	0,09	57,781936	71,9432435
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34		201,708234
O'rtacha	57,88571	54,9	3048,344	3166,0486	32,924082	34,334286		
					5,737951	5,8595465		
				r(y/x)=	-0,353257			
ma1=	=КОРЕНЬ((J9/5)/H9)							

Demak, natija $m_{a1} = 0{,}409698$.

13. Endilikda ma1 qiymatni t-styudent mezoniga qo‘yamiz va a1 parametrning ishonchliligini tekshiramiz. Demak, natija quyidagicha bo‘ldi:

$$t_{a1} = -0{,}84435$$

14. So‘ngra, a0 parametrning ishonchliligini tekshiramiz, buning uchun avvalo, ma0 qiymatni yuqorida keltirilgan formula yordamida, excelning КОРЕНЬ funksiyasi yordamida topib olamiz:

ma1=	0,409698		ma0=	=КОРЕНЬ((J9*E9)/(5*7*H9)	
ta1=	-0,84435			КОРЕНЬ(число)	

Demak, natija $m_{a0} = 22{,}620166$

15. Shuningdek, m_{a0} qiymatni t-styudent mezoni formulasi o‘rniga qo‘yamiz va a0 parametrning ishonchliligini tekshiramiz. Demak, natija quyidagicha bo‘ldi:

$$t_{a0} = 3{,}3986083$$

16. Ana endi $m_{r(y,x)}$ qiymatni hisoblaymiz. Buning uchun avvalo, korrelyatsiya koeffitsiyentini kvadratga oshiramiz:

$$r_{(y,x)} = -0{,}353257$$

$$r_{(y,x)}{}^2 = 0{,}1247907$$

17. So‘ngra, $m_{r(y/x)}$ qiymatni ham m_{a1} va m_{a0} qiymatlarni topganimizdek, excel va formula yordamida hisoblaymiz:

N	y	x	x^2	y*x	(y-y_)^2	(x-x_)^2	y^	(y-y^)^2
1	68,8	45,1	2034,01	3102,88	119,12163	96,04	61,275795	56,6136608
2	61,2	59	3481	3610,8	10,98449	16,81	56,467415	22,3973588
3	59,9	57,2	3271,84	3426,28	4,0573469	5,29	57,090083	7,895633
4	56,7	61,8	3819,24	3504,06	1,4059184	47,61	55,498821	1,44283167
5	55	58,8	3457,44	3234	8,3273469	15,21	56,536601	2,36114119
6	54,3	47,2	2227,84	2562,96	12,857347	59,29	60,549349	39,0543646
7	49,3	55,2	3047,04	2721,36	73,71449	0,09	57,781936	71,9432435
Summa	405,2	384,3	21338,41	22162,34	230,46857	240,34		201,708234
O'rtacha	57,88571	54,9	3048,344	3166,0486	32,924082	34,334286		
					5,737951	5,8595465		
				r(y/x)=	-0,353257			
				r(y/x)^2=	0,1247907			
ma1=	0,409698		ma0=	22,620166		mr(y/x)=	=КОРЕНЬ((1-G14)/5)	
ta1=	-0,84435		ta0=	3,3986083				

Demak, natija $m_{r(y/x)}$=0,41838

1. Shuningdek, $m_{r(y/x)}$ qiymatni t-styudent mezoni formulasi o‘rniga qo‘yamiz va ishonchliligini tekshiramiz. Demak, natija quyidagicha bo‘ldi:

$$t_{r(y/x)}=-0,8443455$$

Xulosa:

tjadval=	1,895
ta1=	-0,844345
ta0=	3,3986083
tr(y/x)=	-0,844345

Demak, natijalarimizni t-styudentning jadval qiymatiga solishtiradigan bo‘lsak, bunda a0 parametrimiz t-jadval qiymatidan katta bo‘lganligi uchun ishonchli. Shuningdek, a1 va r esa t-jadval qiymatidan kichik bo‘lgani uchun ishonchsizligini izohladi.

VII BOB. CHIZIQSIZ REGRESSIYA

7.1. Ijtimoiy-iqtisodiy jarayonlar o‘rtasida bog‘lanishlarni o‘rganishda chiziqsiz funksiyalar bilan foydalanish

7.2. Chiziqsiz regressiya modellari

7.3. Chiziqsiz bog‘lanishlar uchun korrelyasiya indeksini hisoblash

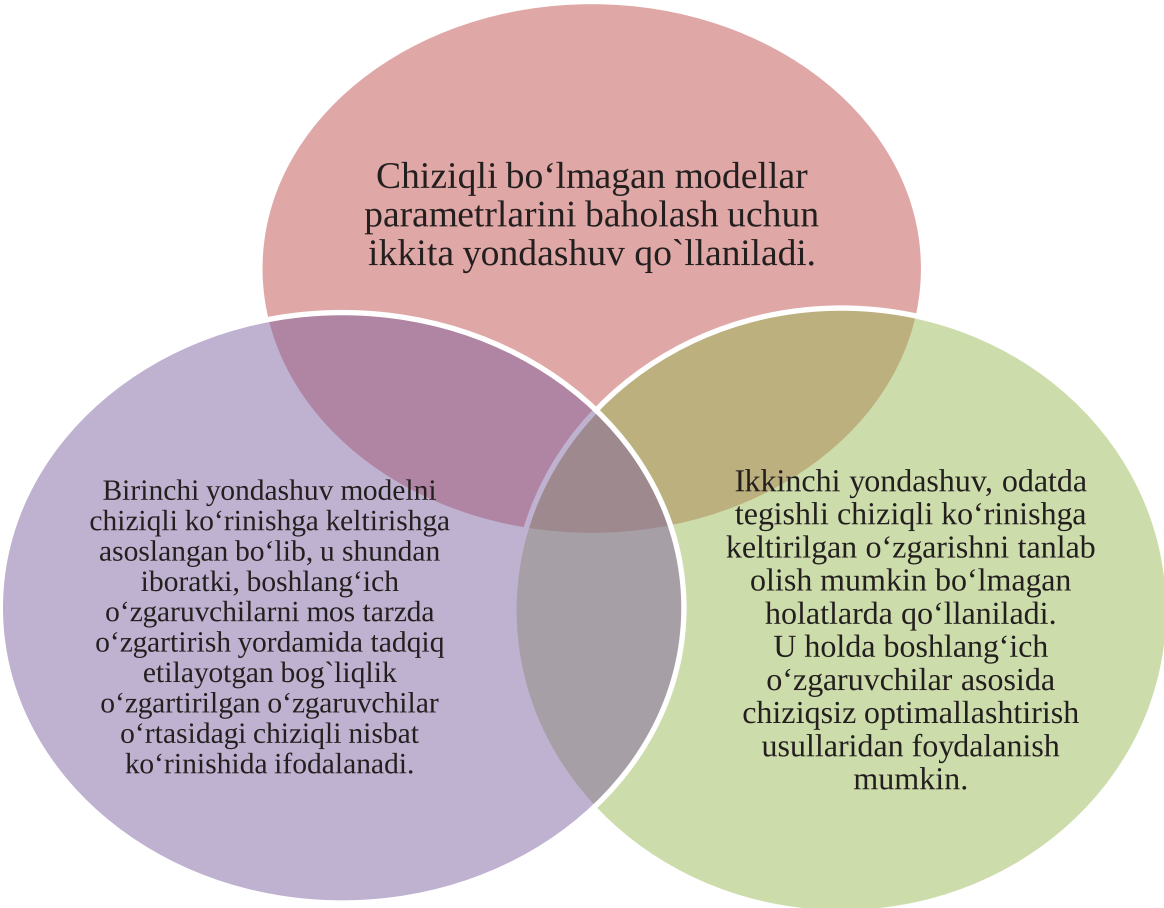
7.4. "Eng kichik kvadratlar" (EKK) yordamida chiziqsiz regressiya koeffitsiyentlarini hisoblash

7.1. Ijtimoiy-iqtisodiy jarayonlar o‘rtasida bog‘lanishlarni o‘rganishda chiziqsiz funksiyalar bilan foydalanish

Ijtimoiy-iqtisodiy hodisalar va jarayonlar o‘rtasidagi nisbatni hamma vaqt ham chiziqli funksiyalar bilan ifodalab bo‘lmaydi. Masalan, ishlab chiqarish funksiyalari (ishlab chiqarilgan mahsulotning hajmi bilan asosiy ishlab chiqarish omillari - mehnat, kapital va h.k. o‘rtasidagi bog‘liqliklar), talab funksiyalari (tovarlar, xizmatlarga bo‘lgan talab bilan ularning narxlari yoki daromad o‘rtasidagi bog‘liqlik) va hokazolar chiziqsiz bo‘lib chiqadi.

Agar iqtisodiy hodisalar o‘rtasida chiziqsiz nisbatlar mavjud bo‘lsa, u holda ular tegishli chiziqsiz funksiyalar bilan ifodalanadi. Chiziqsizlik o‘zgaruvchilarga nisbatan ham, funksiyaga kiruvchi koeffitsiyentlar (parametrlar)ga nisbatan ham ifodalanishi mumkin. Chiziqsiz regressiyalarning ikkita sinfi mavjud. Chiziqsiz regressiyalar sinflarining birinchisiga tahlilga kiritilgan o‘zgaruvchilar bo‘yicha chiziqsiz, lekin baholanayotgan parametrlar bo‘yicha chiziqli regressiyalar (turli polinomlar, giperbola) kiradi. Ikkinchi sinfi baholanayotgan parametrlar bo‘yicha chiziqsiz regressiyalar (darajali, ko‘rsatkichli, eksponensial funksiyalar) dan tashkil topadi.

Ko‘pincha iqtisodiy tahlilda qo‘llaniladigan chiziqli bo‘lmagan regressiyalarning turlari quyidagilar: ikkinchi tartibli polinom, giperbola, darajali funksiya va ko‘rsatkichli funksiya.



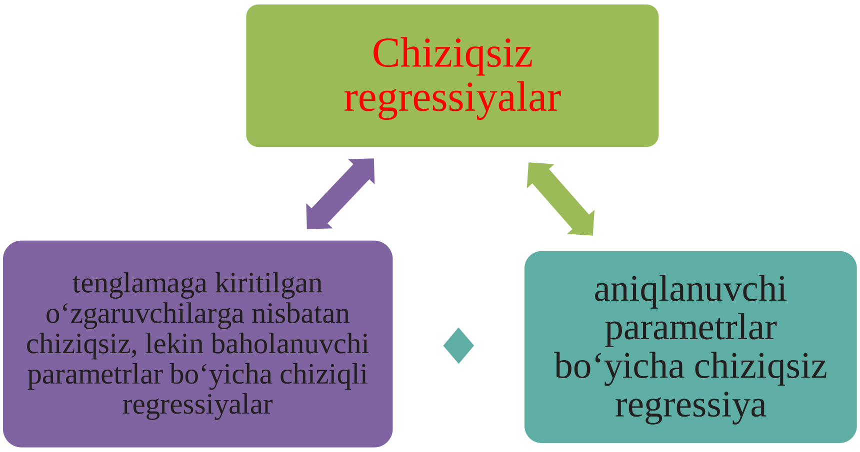
Tahlilga kiritilgan o‘zgaruvchilar bo‘yicha chiziqli bo‘lmagan, lekin baholanayotgan parametrlar bo‘yicha chiziqli regressiya parameterlarini baholash normal tenglamalarni hal etish yo‘li bilan eng kichik kvadratlar usuli yordamida amalga oshiriladi.

Egri chiziqli korrelyatsion bog‘liqlikning har qanday shaklidan foydalanishda o‘zgaruvchilar o‘rtasidagi bog‘liqlikning jipsligi xuddi bog‘liqlikning chiziqli shakli uchun korrelyatsiya koeffitsiyenti singari aniqlanadigan korrelyatsiya indeksi yordamida o‘lchanishi mumkin.

Korrelyatsion bog‘liqlik tenglamasi o‘rganilayotgan o‘zgaruvchilar o‘rtasidagi bog‘liqlikning mohiyati aniq namoyon bo‘lishi, tenglamaning parametrlari esa muayyan tarzda iqtisodiy talqin etilishi uchun imkon qadar soddaroq bo‘lishi kerak. Tegishli bog‘liqlik tenglamasini tanlash masalasi har bir holatda alohida tarzda hal etiladi.

7.2. Chiziqsiz regressiya modellari

Agar iqtisodiy jarayonlar orasida chiziqsiz munosabatlar mavjud bo‘lsa, u holda ular mos ravishda chiziqsiz funksiyalar orqali ifodalanadi: masalan; teng tomonli giperbola, $y = a + \frac{b}{x} + \varepsilon$; ikkinchi tartibli parabola, $y = a + b \cdot x + c \cdot x^2 + \varepsilon$ va boshqalar.



Kiritilgan o‘zgaruvchilarga nisbattan chiziqsiz regressiyaga quyidagi funksiyalar misol bo‘la oladi:

- turli darajali polinomlar

$$y = a + b \cdot x + c \cdot x^2 + \varepsilon; \; y = a + b \cdot x + c \cdot x^2 + d \cdot x^3 + \varepsilon;$$

- teng tomonli giperbola - $y = a + \frac{b}{x} + \varepsilon.$ funksiyalar misol bo‘la oladi.

Tenglamaga kiritilgan o‘zgaruvchilar bo‘yicha chiziqsiz regressiyaning parametrlarini baholash ko‘p qiyinchiliklarni yuzaga keltirmaydi. Ular chiziqli regressiyadagi kabi eng kichik kvadratlar usuli (EKKU) bilan aniqlanadi.

Ikkinchi darajali parabola tenglamasida

$$y = a_0 + a_1 x + a_2 \cdot x^2 + \varepsilon,$$

o'zgaruvchilarni $x = x_1,\ x^2 = x_2$, deb almashtirib quydagi ikki omilli chiziqli regressiya tenglamasini olamiz;

$$y = a_0 + a_1 x_1 + a_2 x_2 + \varepsilon.$$

Mos ravishda uchinchi, to'rtinchi va hokazo k–tartibli polinomlarda ushbu usulni qo'llab, uch, to'rt va hokazo k omilli chiziqli regressiya modellarini olish mumkin.

Misol uchun $y = a_0 + a_1 \cdot x + a_2 \cdot x^2 + ... + a_k \cdot x^k + \varepsilon$, k–tartibli polinomda $y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + ... + a_k \cdot x_k + \varepsilon$, ko'p omilli chiziqli regressiya modelini hosil qilamiz. Ushbu tenglamalarning parametrlarni EKKU bilan hech qanday qiyinchiliksiz aniqlash mumkin.

7.3. Chiziqsiz bog'lanishlar uchun korrelyatsiya indeksini hisoblash

Chiziqsiz regressiya tenglamasi chiziqli bog'lanish kabi korrelyatsiya ko'rsatkichlari, aynan quyidagi korrelyatsiya indeksi(R) bilan to'ldiriladi.

$$R = \sqrt{\left(1 - \frac{\sigma_{qol}^2}{\sigma_y^2}\right)},$$

bu yerda: σ_y^2 - y natijaviy belgining umumiy dispersiyasi;

σ^2_{qol} - $\hat{y}_x = f(x)$ regressiya tenglamasidan kelib chiqib aniqlaniladigan qoldiq dispersiya.

Korrelyatsiya indeksini quyidagi ko‘rinishda ham yozish mumkin:

$$R = \sqrt{1 - \frac{\sum (y - \hat{y}_x)^2}{\sum (y - \bar{y})^2}}.$$

Ushbu ko‘rsatkichning qiymati $[0{,}1]$ oralig‘ida yotadi, ya’ni $0 \le R \le 1$, ko‘rsatkich qanchalik 1ga yaqin bo‘lsa o‘rganilayotgan belgilar orasidagi bog‘lanish shunchalik zich bo‘ladi va tuzilgan regressiya tenglamasi shunchalik haqiqatga yaqin bo‘ladi.

Agar korrelyatsiya indeksini kvadratga aylantirsak, natijada aniqlanadigan determinatsiya indeksi deyiladi:

$$R^2 = \frac{\sum(\hat{y}_x - \bar{y})^2}{\sum(y - \bar{y})^2}$$

Determinatsiya indeksi, regressiya modeli bilan izohlangan y belgining dispersiyasi bo‘lgan natijaviy belgining umumiy dispersiyasining qaysi ulushini ko‘rsatadi. Korrelyatsiya va determinatsiya indekslardan tashqari, elastiklik koeffitsiyentlari va o‘zgaruvchilar o‘rtasidagi bog‘liqlikning zichligini baholashga imkon beradi. Elastiklikning umumiy koeffitsiyenti omil belgisi 1% ga o‘zgarganda natijaviy ko‘rsatkichni taxminan necha foiz o‘zgarishini ko‘rsatadi.

$$Э = f'(x)\frac{\bar{x}}{\bar{y}}$$

Umumlashtiruvchi (o‘rtacha) va nuqtali elastiklik koeffitsiyentlari mavjud. Elastiklikning umumlashtiruvchi koeffitsiyenti o‘rtacha uchun hisoblab chiqiladi va o‘rtacha darajasiga nisbatan x 1% ga o‘sishi bilan o‘rtacha foizga nisbatan necha foiz o‘zgarishini ko‘rsatadi. Nuqtali elastiklik koeffitsiyenti ma'lum bir qiymat uchun hisoblanadi $x = x_0$ va x nuqta x_0 darajadan 1% ga oshganda y o‘zgaruvchi $y_0(x)$ darajadan necha foiz o‘zgarishini ko‘rsatadi. x va y o‘rtasida o‘zaro bog‘liqlik turiga qarab elastiklik koeffitsiyentlarini hisoblash formulalari o‘zgaradi. Asosiy formulalar 7.1-rasmda keltirilgan.

Funksiya tiri, $y = f(x)$	Elastiklik koeffitsiyenti
Chiziqli $y = b_0 + b_1 \cdot x$	$Э(x_0) = \frac{b_1 \cdot x_0}{b_0 + b_1 \cdot x_0}$
Parabola $y = a + b \cdot x + c \cdot x_2$	$Э(x_0) = \frac{(a + 2c + bx_0) \cdot x_0}{a + b \cdot x_0 + c \cdot x_0^2}$
Teng tomonli giperbola $y = a + b / x$	$Э(x_0) = \frac{-b}{a \cdot x_0 + b}$
Darajali $y = a \cdot x^b$	$Э(x_0) = b$
Ko'rsatkichli $y = a \cdot b^x$	$Э(x_0) = x_0 \cdot \ln b$

7.1-rasm. Elastiklik koeffitsiyentlarini hisoblash formulalari

Faqat darajali funksiyalari $(y = a * x^b)$ uchun elastiklik koeffitsiyenti x qiymatdan doimiy mustaqil miqdorni o‘zida aks ettiradi (bu holda b parametrga teng). Shuning uchun ekonometrik tadqiqotlarda darajali funksiyalari keng qo‘llaniladi. Bunday funksiyalardagi b parametr aniq iqtisodiy talqinga ega - bu omil 1% ga oshganda natijaning foiz o‘zgarishini ko‘rsatadi.

7.4. “Eng kichik kvadratlar” (EKK) yordamida chiziqsiz regressiya koeffitsiyentlarini hisoblash

Tajribalar shuni ko‘rsatadiki chiziqsiz regressiyalar ichida ko‘proq ikkinchi tartibli parabola, ayrim hollarda uchinchi tartibli parabola ishlatiladi. Yuqori tartibli polinomlarni qo‘llashdagi chegaralanishlar o‘rganilayotgan to‘plamning bir jinsliligi bilan bog‘liq, polinom darajasi qancha yuqori bo‘lsa egri chiziqdagi sinishlar shuncha ko‘p bo‘ladi va mos ravishda natijaviy belgi to‘plami ham bir jinsli bo‘lmaydi. Undan tashqari ma’lumotlarni to‘plashda va hisoblashlarda noaniqliklar keltirib chiqaradi.

Ikkinchi tartibli parabolani omil belgi qiymatlarining ma'lum bir oraliqda qaralayotgan o'zgaruvchining bog'lanish xususiyatini o'zgarishiga: ya'ni to'g'ri bog'lanishni teskari bog'lanishga, teskari bog'lanishni to'g'ri bog'lanishga olib keladigan holatlarda qo'llash maqsadga muvofiq. Bunday holatlarda omil belgining natijaviy belgini ekstrimal (maksimal yoki minimal) qiymatga erishtiruvchi qiymati aniqlanadi.

Buning uchun ikkinchi darajali parabolaning hosilasi nolga tenglashtiriladi; ya'ni $\hat{y}_x = a + b \cdot x + c \cdot x^2$ dan hosila olamiz va $b + 2 \cdot c \cdot x = 0$, bundan $x = -\frac{b}{2c}$ hosil bo'ladi.

Agar berilgan ma'lumatlar bog'lanish yo'nalishini o'zgarishini ta'minlay olmasa, u holda ikkinchi tartibli parabola parametrlarining ma'nosini tushinish qiyin bo'ladi. Bunday holatda bog'lanish shakli boshqa chiziqsiz model bilan almashtiriladi.

Ikkinchi darajali parabolaning $a,b,c,$ paramerlarining qiymatlarini topish EKKUni qo'llab quydagi normal tenglamalar sistemasini matematikaning biror bir usulini qo'llab yechishga olib keladi:

$$\begin{cases} \sum y = n \cdot a + b \cdot \sum x + c \cdot \sum x^2, \\ \sum y \cdot x = a \cdot \sum x + b \cdot \sum x^2 + c \cdot \sum x^3, \\ \sum \mathrm{y} \cdot \mathrm{x}^2 = a \cdot \sum x^2 + b \cdot \sum x^3 + c \cdot \sum x^4. \end{cases}$$

$b > 0$ va $c < 0$ bo'lganda egri chiziq eng yuqori nuqtaga, ya'ni egri chiziqning sinish, bog'lanish yo'nalishini o'zgartirish nuqtasiga nisbatan simmetrik bo'ladi, aynan o'sish pasayishga o'zgaradi. Bunday funksiyalarnini iqtisodiyotda jismoniy mehnat bilan shug'ullanuvchi ishchilarning ish haqini ularning yoshiga bog'liqligini o'rganishda kuzatish mumkin. Ishchilarning yoshi kattalashib borgan sari ularning tajribasi ortishi bilan birga ularning malakasi ham yuqorilashib ish haqi ko'payib boradi. Lekin ma'lum bir yoshdan boshlab organizimni qarishi natijasida mehnat samaradorligini pasayishi ishchining ish haqini pasayishiga olib kelishi mumkin.

Agar o‘zaro bog‘lanishning parabolik shakli natijaviy ko‘rsatkichni avval o‘sishini, so‘ngra pasayishini namoish etsa, u holda omil belgining natijani maksimumga erishtiradigan qiymati topiladi. Masalan, oilada A mahsulot (birligini) daromad darajasiga bog‘liq holda iste’mol qilinishi $\hat{y}_x = 5 + 60 \cdot x - x^2$ tenglama bilan tavsiflansin. Tenglamaning birinchi tartibli hosilasini nolga tenglab $\hat{y}_x^1 = 60 - 2x = 0$, maksimal iste’mol miqdorini beruvchi daromad qiymatini topamiz, ya’ni $x = 30$ ming so‘mda iste’mol maksimal darajaga yetadi.

$b < 0$ va $c > 0$ bo‘lganda ikkinchi darajali parabola o‘zining eng quyi nuqtasiga simmetrik bo‘ladi. Bunday holat funksiyaning bog‘lanish yo‘nalishini (kamayishni o‘sishga) o‘zgartiruvchi eng kichik qiymatni topish imkonini beradi. Faraz qilaylik ishlab chiqarish xarajatlarini ishlab chiqarilgan mahsulot hajmiga bog‘liqligi quyidagi tenglama bilan tavsiflansin:

$$\hat{y}_x = 1200 - 60 \cdot x + 2 \cdot x^2,$$

bu holatda eng kam xarajatga $x = 15$ mahsulot birligi ishlab chiqarilganda erishiladi $(-60 + 2 \cdot 2 \cdot x = 0)$.

Ikkinchi tartibli parabola egri chizig‘i simmetrik bo‘lganligi sababli u aniq tadqiqotlarda har doim ham qo‘llanilavermaydi. Tadqiqotchi ko‘pincha parabolaning to‘liq shakli bilan emas balki, uning ayrim segmentidan foydalanib ish yuritadi. Parabolik bog‘lanishning parametrlari har doim ham mantiqqa ega bo‘lavermaydi. Shuning uchun bog‘lanish grafigi ikkinchi tartibli parabolani aniq ifodalamasa, u boshqa chiziqsiz funksiyaga almashtiriladi, masalan darajali funksiyaga. Ikkinchi tartibli parabola ko‘proq qishloq xo‘jaligida hosildorlikni berilgan o‘g‘itlar miqdoriga bog‘liqligini tavsiflash uchun qo‘llaniladi. Bog‘lanishning bu shakli quyidagicha asoslanadi -o‘simlikka berilayotgan o‘g‘itning miqdori ortishi bilan hosildorlik, faqat berilayotgan o‘g‘itning miqdori optimal dozasiga yetgunga qadar oshib boradi, deyiladi. Dozaning keyingi ortishi o‘simlik uchun zarar va hosildorlikni kamayishiga olib keladi. Shuning uchun amalda bunday bog‘lanish ko‘proq parabolaning segmenti ko‘rinishida beriladi.

Chiziqsiz funksiyalar sinfida parametrlarning qiymati hech qanday qiyinchiliksiz EKKU bilan aniqlanadigan funksiya sifatida, ekonometrikada ma’lum bo‘lgan, teng tomonli giperbola $\hat{y}_x = a + \frac{b}{x} + \varepsilon$ ni ko‘rish mumkin. Bunga klassik misol sifatida ishsizlik me’yori (x) va ish haqi (y)ning o‘sish foizi orasidagi munosabatini tavsiflovchi Fillips egri chizig‘i keltiriladi:

$$\hat{y} = a + \frac{b}{x} + \varepsilon \ .$$

Ingliz iqtisodchisi A.V.Fillips 100 yildan ko‘proq davrdagi ma’lumotlarni tahlil qilib XX asrning 50-yillari oxirida ish haqini o‘sib borishi darajasi, ishsizlik darajasiga nisbatan teskari bog‘langanligini aniqlagan.

$y = a + \frac{b}{x} + \varepsilon$ ko‘rinishidagi teng tomonli giperbolada $\frac{1}{x}$ ni z bilan almashtirib $y = a + b \cdot z + \varepsilon$ chiziqli regressiya tenglamasini olamiz. Uning parametrlarini EKKU bilan aniqlash mumkin .

Ushbu funksiya uchun normal tenglamalar sistemasi quydagidan iborat:

$$\begin{cases} \sum y = n \cdot a + b \cdot \sum \frac{1}{x}, \\ \sum \frac{y}{x} = a \cdot \sum \frac{1}{x} + b \cdot \sum \frac{1}{x^2} \end{cases}$$

$b > 0$ bo‘lganda teskari bog‘lanish bo‘lib, $x \to \infty$ bo‘lganda y a parametr bilan baholanadigan o‘zinig eng kichik qiymatiga erishadi.

$$\hat{y}_x = 0{,}00679 + 0{,}1842 \frac{1}{x}$$

funksiyasi bilan ifodalanuvchi Fillips egri chizig‘ida a parametrning qiymati 0,00679ga teng, bu ishsizlik darajasining o‘sishi bilan ish haqining qo‘shimcha o‘sishi sur’ati nolga intilishini ko‘rsatadi.

$b < 0$ bo‘lib x cheksizga intilganda ($x \to \infty$) yuqori asimptotaga ega bo‘lgan sekin o‘suvchi, ya’ni $\hat{y}_x = a + \frac{b}{x}$ tenglamada a parametr baho beradigan y ning maksimumga erishuvchi, funksiyaga ega bo‘lamiz.

Misol sifatida umumiy xarajatlar (yoki daromadlar) bilan uzoq muddatli tovarlarga xarajatlar ulushi orasidagi bog‘lanishni ko‘rish mumkin. Bunday bog‘lanishning matematik yozuvi ***Engel egri chizig‘i*** deb nom olgan. 1857-yilda nemis statistik olimi E. Engel oila xarajatlarini o‘rganish asosida: daromadni ortishi bilan daromadning oziq-ovqatlarga sarf qilinadigan ulushi kamayib borish qonuniyatini aniqlagan. Mos ravishda daromadning ortib borishi bilan daromadning nooziq-ovqat mahsulotlariga sarf qilinadigan ulushi ortib boradi. Lekin bu o‘sish chegarasiz bo‘lmaydi, ya’ni birdan katta yoki 100%dan ko‘p bo‘lmaydi. Ayrim tovarlar uchun bu chegara $\hat{y}_x = a - \frac{b}{x}$ tenglamaning a parametri bilan tavsiflanadi. Ushbu tenglamada:

y - nooziq-ovqat tovarlariga xarajatlar ulushi;

x - daromad.

Teng tomonli giperbolada a va b parametrlar quyidagicha hisoblanadi:

$$a = \frac{\sum \frac{1}{x^2} \sum y - \sum \frac{y}{x} \sum \frac{1}{x}}{\sum \frac{1}{x^2} - \left(\sum \frac{1}{x}\right)^2}, \quad b = \frac{n \sum \frac{y}{x} - \sum y \sum \frac{1}{x}}{n \sum \frac{1}{x^2} - \left(\sum \frac{1}{x}\right)^2}.$$

Engel egri chizig‘ining modelini yozish uchun $y = a + b \cdot \ln x + \varepsilon$ ko‘rinishdagi yarim logarifmik funksiyalar ham qo‘llaniladi (1943-y. Uorking va 1964-y. Lizer).

$\ln x$ ni z bilan almashtirsak yana $y = a + b \cdot z + \varepsilon$ ko‘rinishidagi chiziqli tenglamani olamiz. Ushbu funksiya avvalgi funksiya kabi parametrlar bo‘yicha chiziqli, asosiy x o‘zgaruvchi bo‘yicha esa chiziqli emas. a va b parametrlarni EKKU yordamida aniqlash mumkin. Bunda normal tenglamalar sistemasi quyidagicha bo‘ladi:

$$\begin{cases} \sum y = n \cdot a + b \cdot \sum \ln x, \\ \sum y \cdot \ln x = a \cdot \sum \ln x + b \cdot \sum (\ln x)^2. \end{cases}$$

Nazorat uchun savollar

1. Chiziqli va chiziqsiz regressiyaning farqi nimada?
2. Chiziqli bo‘lmagan munosabatlar uchun korrelyatsiya qanday aniqlanadi?
3. Determinatsiya koeffitsiyenti qanday aniqlanadi?
4. Umumiy elastiklik koeffitsiyentining mohiyatini tushuntiring?
5. Barcha chiziqli bo‘lmagan regressiyalar qaysi ikki turga bo‘linadi?
6. Darajali chiziqsiz regressiya uchun xos bo‘lgan narsa.
7. Parametrlari bo‘yicha chiziqli bo‘lmagan regressiyalar.
8. Agar korrelyatsiya indeksini kvadratga aylantirsak, natijada aniqlanadigan indeksi nima deyiladi?

Masala

N	x	y
1	50,2	5,3
2	50,3	6,2
3	50,4	5
4	51	5,7
5	51,2	6,3
6	51,4	5,4
7	51,6	6,8
8	51,9	7,3
9	52,3	6,6
10	52,4	6,2
11	62,4	6,6
12	52,9	6,3
13	53	6,7
14	53,2	7,8
15	53,5	6,8
16	53,6	7,2
17	54,3	7,7
18	54,7	8,5
19	55,2	7,8
20	55,5	8,4
21	56,5	8,8

Topshiriq. Berilgan ma’lumotlar asosida teng tomonli giperbola (Filips egri chizig‘i) uchun chiziqsiz regressiya modelini tuzing.

Yechim. Excel dasturini ishga tushirib, ma’lumotlarni yuqoridagi tartibda kiritamiz. Bunda asosiy omil sifatida – y ni hamda ta’sir e’tuvchi omil sifatida – x ni belgilaymiz.

1. Tenglamaga kiritilgan o‘zgaruvchilar bo‘yicha chiziqsiz regressiyaning parametrlarini baholash ko‘p qiyinchiliklarni yuzaga keltirmaydi. Ular chiziqli regressiyadagi kabi eng kichik kvadratlar usuli (EKK) bilan aniqlanadi. Ushbu funksiya uchun normal tenglamalar sistemasi quyidagidan iborat:

$$\begin{cases} \sum y = n \cdot a + b \cdot \sum \frac{1}{x}, \\ \sum \frac{y}{x} = a \cdot \sum \frac{1}{x} + b \cdot \sum \frac{1}{x^2} \end{cases}$$

Biz avvalo, ushbu tenglamalar sistemasidagi natijalarni, ya’ni y, 1/x, 1/x^2 va y/x qiymatlar hamda ularning summalarini excel funksiyasidan foydalanib topib olamiz:

N	**x**	**y**	**1/x**	**1/x^2**	**y/x**
1	50,2	5,3	0,019920319	0,00039682	0,105577689
2	50,3	6,2	0,019880716	0,00039524	0,123260437
3	50,4	5	0,01984127	0,00039368	0,099206349
.	.	.	.	.	.
.	.	.	.	.	.
20	55,5	8,4	0,018018018	0,00032465	0,151351351
21	56,5	8,8	0,017699115	0,00031326	0,155752212
Summa	1117,5	143,4	0,395562018	0,00746721	2,689798522

2. Demak, biz uchun kerak bo‘lgan natijalar topildi, ana endi bu natijalarni eng kichik kvadratlar usulining normal tenglamalar sistemasi o‘rniga qo‘yib chiqamiz:

21*a0+0,395*a1=143,4
{ 0,395*a0+0,0074*a1=2,689

3. So‘ngra, ushbu sistemani Kramer usulidan foydalanib hisoblaymiz. Deltalarni topishda esa excelda МОПРЕД buyrug‘i orqali har bir deltalar qiymatlarini belgilab hisoblaymiz:

delta=	21	0,395562	0,000342
	0,395562	0,007467	
delta1=	143,4	0,395562	0,006815
	2,689799	0,007467	
delta2=	21	143,4	-0,23782
	0,395562	2,689799	

4. Ana endi a0=delta1/delta va a1=delta2/delta formulalar yordamida parametrlarimizni topib olamiz:

a0= 19,92621

a1= -695,341

Shundan kelib chiqib, chiziqsiz modelimiz ko‘rinishi quyidagicha bo‘ladi:

y^=19,9-695,3/x+ε

Xulosa: Agar ta’sir etuvchi omil 1 birlikka oshsa, natijaviy omilning 695,3 ga kamayishini izohladi. a0 parametrimiz modelimizning konstantasi hisoblanadi. Agar ta’sir etuvchi omil 0ga tenglashsa, natijaviy omil 19,9 ga tenglashadi. Shuningdek, a0 modelimizda ishtirok etmagan ta’sir etuvchi omillarning miqdori ham hisoblanadi.

VIII BOB. EKONOMETRIK MODELLARNI BAHOLASH

8.1. Ekonometrik modellarning iqtisodiy tahlilida verifikatsiya bosqichining ahamiyati

Ekonometrik modellarning iqtisodiy tahlilida verifikatsiya bosqichining ahamiyati identifikatsiya qilish bosqichidan keyin quyidagi savollar tug‘iladi: Tuzilgan modeli maqsadga muvofiqmi, ya’ni uning prognozlash va imitatsion hisoblar uchun ishlatilishi kutilayotgan natijalar haqiqatga adekvatli natijalarni beradimi? Tuzilgan modelga asoslangan prognozlash va imitatsion hisoblarning aniqligi nimadan iborat? Ushbu savollarga javob olish ekonometrik modelning verifikatsiya muammosi mazmunidir.

Verifikatsiya usullari gipotezalarning statistik tekshiruvi va statistik baholashning turli usullarining aniqlik xususiyatlarini statistik tahlil qilishga asoslangan. Bu, shuningdek, ekonometrik modellarda qo‘llaniladigan verifikatsiya bosqichida retrospektiv hisoblash tamoyilini ta’kidlash lozim.

Tamoyilini mohiyati bo‘yicha dastlabki statistik ma’lumotlar ikki qismga bo‘linadi: haqiqiy, ma’lumotlar qoldig‘ini tashkil topgan kuzatuvlar va ko‘riklar namunadagi ba’zi tashkil topgan ta’lim majmui: Bundan tashqari qadamlar ta’lim namuna uchun spetsifikatsiya va identfikatsiya amalga oshiriladi. Olingan model misol ekzogen o‘zgaruvchilarni barpo etilgan va (ilgari orqaga) endogen o‘zgaruvchilarni olingan model qadriyatlarni o‘rganib etiladi. Ko‘rib chiqish namunasi haqiqiy qadriyatlar bilan mos keladigan model bilan ushbu qadriyatlar solishtirish model topilmalar haqiqatga va aniqlik malakasi tahlil qilish bizga beradi.

Tahlil qilinayotgan qatorlar dinamikasi har doim anchagina uzunroq qatorlarning tanlamasi hisoblanadi. Shuning uchun korrelyatsion-regerssion tahlil asosida olingan ekonometrik modellarning ishonchliligini har tomonlama tekshirish va baholash lozim. Tuzilgan ekonometrik modelning ahamiyatliligi, ishonchliligi va keyinchalik bashoratlashda qo‘llash mumkinligi quyidagi mezonlar asosida baholanadi:

Ekonometrik modellarning iqtisodiy tahlilida verifikatsiya bosqichining ahamiyati

Ekonometrik modellarni ahamiyatini Fisher mezoni va approksimatsiya xatoligi yordamida baholash	Ekonometrik modellar sifatini ko‘p omilli korrelyatsiya koeffitsiyenti va determinatsiya koeffitsiyenti yordamida baholash	Ekonometrik model parametrlarini Styudent mezoni yordamida baholash	Qatorlarda qoldiq avtokorrelyatsiyani Darbin-Uotson mezoni bo‘yicha baholash.

8.2. Ekonometrik modellar sifati va ahamiyatini mezonlar bo‘yicha baholash

Regressiya tenglamasi sifatini baholashda F-Fisher mezonidan foydalaniladi. Olingan regressiya tenglamasining sifatini baholash dispersion tahlil qilish usullariga asoslangan. Natijaviy ko‘rsatkich y_i

ning qiymatlari ikkita y_i va e_i komponentlarning yig‘indisi sifatida ifodalanishi mumkin

$$y_i = y_i + e_i \qquad (8.1)$$

Kattalik $y_i = a + b \cdot x_i$ kuzatuv i uchun y ning hisoblangan qiymati. Qoldiq e_i natijaviy ko‘rsatkich u ning kuzatiladigan va hisoblangan qiymatlari orasidagi farq yoki regressiya tenglamasi yordamida tushuntirilmagan u o‘zgaruvchining qismi.

8.1-formuladan o‘zgaruvchining kuzatilgan qiymatlari $D(y)$ dispersiyaning, uning hisoblangan qiymatlari $D(\hat{y})$ ning va $D(e)$ qoldiqlari (qoldiq dispersiyalar $D_{qoldiq} = D(e)$) o‘rtasidagi quyidagi munosabati kelib chiqadi:

$$D(y) = D(\hat{y}) + D(e)$$

$$D(y) = \frac{1}{n}\sum(y_i - \bar{y})^2$$

$$D(e) = D_{qoldiq} = \frac{1}{n}\sum(y_i - \hat{y}_i)^2 \quad va\ M(e) = 0$$

bo‘lsa quyidagi ko‘rinishda bo‘ladi:

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (8.2)$$

O‘zgaruvchi y ning ta’riflangan qismi $D(y)$ ning umumiy dispersiyasi $D(\hat{y})$ ga munosabati:

$$R^2 = \frac{D(\hat{y})}{D(y)} \quad yoki \quad R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (8.3)$$

determinatsiya koeffitsiyenti deb ataladi va regressiya tengamasi sifati yoki bog‘lanish modelini xarakterlash uchun ishlatiladi. Ushbu (8.3) munosabati quyidagi shaklda ifodalanishi mumkin:

$$R^2 = 1 - \frac{D_{qoldiq}}{D(y)} \quad yoki \quad R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad (8.4)$$

Determinatsiya koeffitsiyenti 0 va 1 oralig'ida o'zgaradi.

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

Quyidagi savol tug'iladi: R^2 ning qaysi qiymatlari uchun regressiya tenglamasi statistika jihatidan ahamiyatsiz deb hisoblanadigan bo'lib, uni tahlil qilishda asossiz deb biladi? Bu savolga Fisherning F −mezonida javob berilgan.

Quyidagi belgilashlarni kiritamiz:

$TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2$ - chetlanish kvadratlarning to'liq yig'indisi;

$ESS = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ - chetlanish kvadratlarning tushuntirilgan yig'indisi;

$RSS = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2$ - chetlanish kvadratlarning tushuntirilmagan yig'indisi.

Ma'lumki,

$$F = \frac{\frac{ESS}{k}}{\frac{RSS}{n-m-1}}$$

bu yerda k -regressiya tenglamasining bog‘liq bo‘lmagan o‘zgaruvchilar soni (juft regresiya uchun $k = 1$), ya’ni normal taqsimlangan xatolik uchun ε_i Fisherning F-statistikasi (Fisher qonuniga muvofiq taqsimlangan tasodifiy miqdor) $k_1 = k,\ k_2 = n - k - 1$ erkinlik darajalari bilan.

Fisherning F-mezoniga ko‘ra, regressiya tenglamasining statistik ahamiyatsizligi (ya’ni F qiymatining noldan statistik jihatdan ahamiyatsiz farqligi) haqida H_0 «nollik» gipoteza. Bu gipoteza $F > F_{jad}$ shartni qondirganda rad etilmoqda, bu yerda F_{jad} Fisher F-mezonining jadvalidan α darajaga ega bo‘lgan va $k_1 = k,\ k_2 = n - k - 1$ erkinlik darajalari bilan aniqlanadi.

Statistik gipotezalarda ahamiyatlilik darajasi (α belgisi) to‘g‘ri gipotezani rad etish ehtimoli deb ataladi (bu birinchi turdagi xatolik). Ahamiyatlilik darajasi α odatda 0,05 va 0,01 qiymatlarni qabul qiladi, bu esa birinchi turdagi xatolarning 5% va 1% gacha bo‘lish ehtimoliga to‘g‘ri keladi.

F qiymatini R^2 determinatsiya koeffitsiyenti bo‘yicha ifodalanishi mumkin:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - k - 1}{k}$$

Masalan, 30 ta kuzatuvlar bo‘yicha qo‘yidagi regressiya teglamasi olingan.

$$y = 50{,}5 + 3{,}2x \quad \text{va} \quad R^2 = 0{,}6$$

Uning muhimligini $\alpha = 0{,}05$ ahamiyatliligi darajasida tekshirish kerak. $k = 1$ ni hisobga olgan holda F-statistikasining qiymatini aniqlaylik

$$F = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k} = \frac{0{,}6}{1-0{,}6} \cdot \frac{30-1-1}{1} = \frac{0{,}6}{0{,}4} \cdot 28 = 42$$

Fisherning F- kriteriyasi jadvali bo‘yicha:

$k_1 = k,\ k_2 = n - k - 1 = 30 - 1 - 1 = 28$ va $\alpha = 0{,}05$

$F_{jad} = 4{,}2$. $F = 42 > F_{jad} = 4{,}2$ bo‘lgani uchun regressiya tenglamaning statistik ahamiyatliligi to‘g‘risida xulosa qilishimiz mumkin.

Approksimatsiya xatoligi – bu nazariy y ning haqiqiy Y qiymatlaridan o‘rtacha nisbiy chetlanishi:

$$\varepsilon = \frac{1}{n}\sum\left|\frac{y_i - \hat{y}}{y_i}\right| * 100\%$$

bu yerda, n- kuzatuvlar soni

y_i- asosiy omilni haqiqiy qiymatlari

$\hat{y}$ - asosiy omilni tekislangan nazariy qiymatlari.

Agar ε qiymat 10-12 foizdan oshmasa, tuzilgan regressiya tenglamasi qoniqarli deb hisoblanadi.

Darbin–Uotson mezoni Darbina-Uotson mezoni yoki d – mezoni (qoldiqlarning bog‘liq bo‘lmaganligi xossasi ya’ni avtokorrelyatsiya mavjud emasligi).

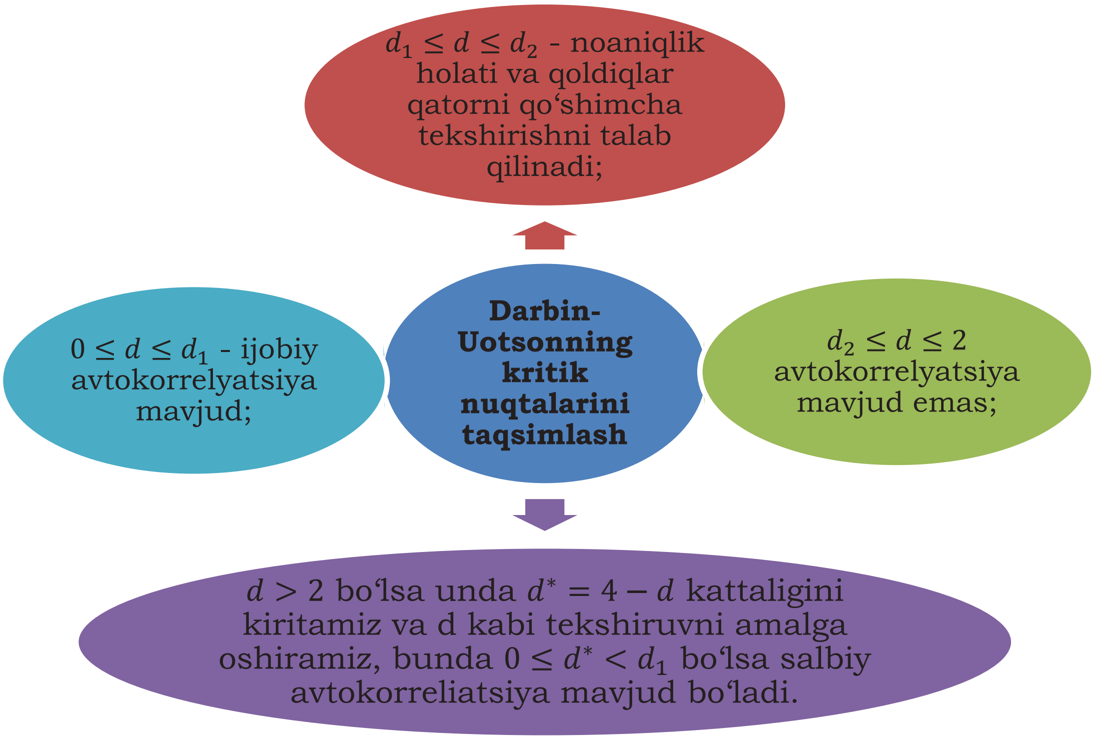
$$dw = \frac{\sum_{i=1}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

bu yerda, ei=yhaq-yhis

dw – mezoni $0 \leq d \leq 4$ oraliqda yotadi.

Agar $dw < 2$ bo‘lsa, qoldiqlar qatori ijobiy avtokorrelyatsiyaga ega va agar $dw > 2$ bo‘lsa salbiy avtokorrelyatsiyaga ega. dw – mezoni 4ga yaqin bo‘lsa salbiy avtokorrelyatsiyani mavjudligini bildiradi. dw – mezoni 0 ga yaqin bo‘lsa ijobiy avtokorrelyatsiya mavjudligini bildiradi. Darbin-Uotsonning kritik nuqtalarini taqsimlash jadvallari mavjud.

Unda berilgan α statistik ahamiyatliligi qiymat darajasi, n kuzatilishlar soni va omilli o‘zgaruvchilar soni uchun 2ta kattaliklar aniqlaydi: d_1 quyi chegara va d_2 yuqori chegara. Hisoblangan kattaligi d ushbu ikki qiymatlar bilan taqqoslanadi.



8.3. Regressiya tenglamaning parametrlarni baholarining xususiyatlari

Chiziqli bir omilli model qurishda uning ayrim kamchiliklariga e’tiborni qaratmoq lozim. Modelni jarayonning bitta omil yordamida, u hatto hal qiluvchi omil bo‘lgan taqdirda ham haqqoniy yoritib berishi mumkin emas. Masalan, paxta xomashyosini yalpi yig‘ib olishni o‘rganishda asosiy omil sifatida hosildorlikni olish mumkin, lekin sinchiklab o‘rganish natijasida yer miqdori va sifati, o‘g‘itlar (ularni miqdori, sifati, quritish muddati), sug‘orish harakat tartibi va boshqa omillarni ham e’tiborga olish zarur.

Shunday qilib, «asosiy» omillar miqdori cheksiz o‘zgarishi mumkin. Bunday masalani hal etish bir omilli modeldan ko‘p omilligacha o‘tishni taqozo etadi. Ammo bu ham funksiyaga asosiy omillardan tashqari yana ko‘p sonli ikkinchi darajali omillar ta’sir qilishi hisobiga hisoblashda xatolik bo‘lishini rad etmaydi. Ko‘pincha ularning ta’siri sezilarsiz va qarama-qarshi xarakterga ega. Ushbu omillarning barcha samarasi, ham musbat ham manfiy qiymatlarni qabul qiluvchi Y tasodifiy o‘zgaruvchi bilan baholanadi. Chiziqli bog‘liqlik:

$Y = f(X_1, U)$ $yoki$ $Y = f(X_1, X_2, \dots X_n, U)$ ko‘rinishda bo‘ladi.

Y o‘zgaruvchi quyidagi stoxastik xususiyatlarga ega bo‘lgan xato sifatida namoyon bo‘ladi:

- ehtimoliy me’yoriy taqsimotga ega bo‘ladi;
- nolli o‘rtachaga ega;
- chekli dispersiyaga ega;
- o‘lchash xatosi hisoblanadi.

Statistik ma’lumot yig‘ishda ko‘p hollarda parametrning haqiqiy qiymatlari o‘rniga yashirin xatoga ega o‘lchamlar kiritiladi (ular obyektiv, subyektiv xarakterga ega bo‘lishlari, o‘lcham hisoblarining noaniqligi, noaniq hujjat aylanishi, alohida o‘lchamlarini subyektiv bahosi va boshqalar). Barcha yuqorida sanab o‘tilgan kamchiliklar o‘lchash xatolarini tenglama xatolariga o‘tishiga olib keladi, ya’ni:

$$Y = a_0 + a_1 X + W$$
$$W = U + V$$

bunda, W-jami xato; U-stoxastik e’tiroz bildirish; V-o‘lchash xatosi. Nisbatan oddiy bog‘liqlik deb chiziqli bir omilli bog‘liqlik yoki chiziqli ko‘p omilli model, u tasodifiy xatoga nisbatan bir necha taxminlarni qabul qilganda hisoblanadi: o‘rtacha nolga teng; dispersiya sust va asosiy omillarga bog‘liq emas va tasodifiy xato bir-biriga bog‘liq emas.

Ko‘p omilli holatda: $Y = a_{0i} + a_{1i} X_i + U_i, a_0$ va a_1 koeffitsiyentlarni quyidagi shartlardan kelib chiqqan holda aniqlash mumkin:

$$E(U) = 0, i \in N$$

$$E(U_i U_j) = \begin{cases} 0 & \text{agar } \; i \neq j, \; i,j \in N \\ \sigma_u^2 & \text{agar } \; i = j, \; i,j \in N \end{cases}$$

Sodda iqtisodiy modellarni ko‘rib chiqishda bu masalani standart usuli yordamida yechish mumkin. Eng kichik kvadrat usuli klassik hisoblanadi. Lekin nisbatan murakkabroq vaziyatlarda murakkab ekonometrik modelni koʻrib chiqishda murakkab texnik yoʻllardan foydalangan holda yangi usullarni ishlab chiqish zarur.

Oddiy chiziqli regression modelning toʻliq spetsifikatsiyasi regression tenglamadan va 5 ta birlamchi yoʻl qoʻyishlardan tashkil topgan. Shu yoʻl qoʻyishlarni koʻrib chiqamiz. Birinchi ikki taxmin shundan iboratki, X ning har bir qiymati uchun ε xato nol qiymat atrofida me’yoriy taqsimlangan. Taxmin qilinadiki, εi uzluksiz kattalik hisoblanib, oʻrtacha atrofida simmetrik taqsimlangan $-\infty$ dan $+\infty$ gacha oʻzgaradi va uning taqsimlanishi 2 oʻlcham oʻrtacha va variatsiya yordamida aniqlanadi.

Demak:

Birinchi taxmin: ε_i- me’yoriy taqsimlangan.

Ikkinchi taxmin: - $E(e_i) = 0$ oʻrtacha xato nolga teng.

Haqiqatda biz stoxastik xatoni har bir qiymatini, koʻpgina sabablar natijasi sifatida koʻrishimiz mumkinki, bunda har bir sabab bogʻliq oʻzgaruvchini, u deterministik hisoblanishi mumkin boʻlgan qiymatdan sezilarsiz tarzda ogʻdiradi. Bunday koʻzdan kechirishda oʻlchash xatosi oʻxshashi bilan taqsimot xatosi toʻgʻri va shuning uchun oʻrtacha xatoni me’yoriyligini va nolga tengligi haqida taxminlar oʻxshash.

Uchinchi taxmin gomoskedatlikka tegishli boʻlib, u har bir xatoσ^2 ning qiymati noma’lum boʻlgan bir xil variatsiya ekanligini anglatadi. Bu taxmin, masalan X ning katta qiymatlari uchun xato dispersiyasini imkoni, xuddi kichik qiymatlardagi kabi degan tasdiq bilan kelishiladi. Yuqorida koʻrib oʻtilgan ishlab chiqarish funksiyasida, bu taxminga asosan ishlab chiqarishdagi variatsiya ham, ish kuchi qiymatiga bogʻliq emas.

“Eng kichik kvadratlar” usulining ekonometrik modellardagi parametrlarni baholashda qoldiqlar kvadratlari yigʻindisining minimumga intilishiga asoslanadi. Shuning uchun regressiyaning qoldiq qiymatlarini koʻrib chiqish muhim ahmiyat kasb etadi.

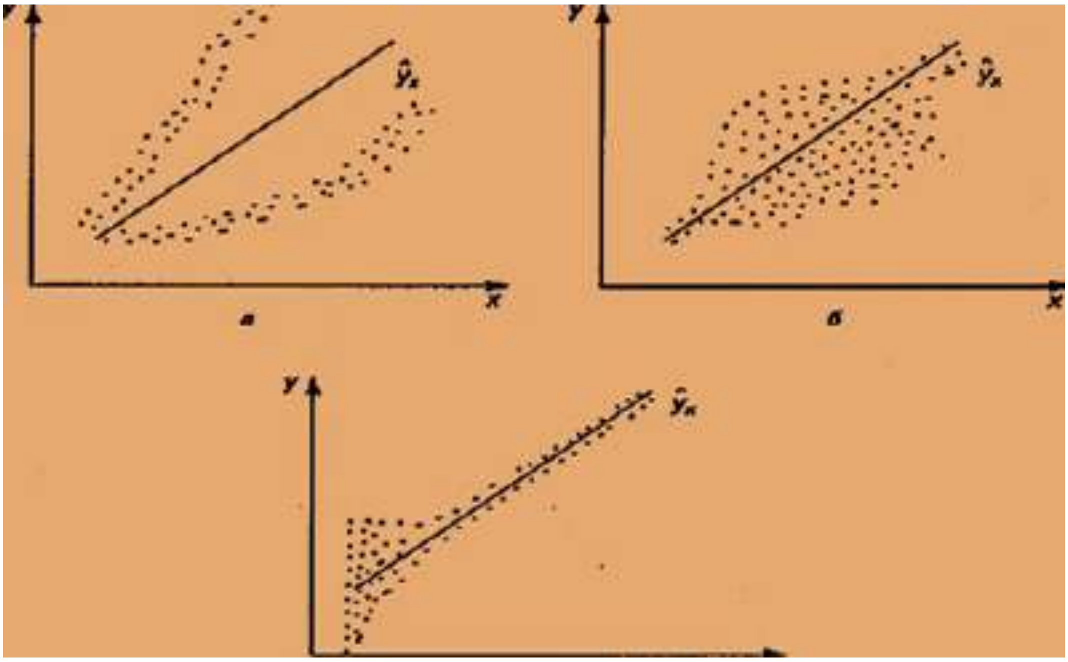
“Eng kichik kvadratlarining” uchinchi taxmini gomoskedatlikka tegishli boʻlib, u har bir X uchun qoldiqning dispersiyasi bir xil boʻlishi ekanligini anglatadi. Bu taxmin, masalan X ning katta qiymatlari uchun qoldiq dispersiyasini imkoni, xuddi kichik qiymatlardagi kabi degan tasdiq bilan kelishiladi.

Uchinchi taxmin: Gomoskediklik

$$Var(\varepsilon_i) = \sigma^2$$

Gomoskedatlik sharti:

Agar yuqoridagi “Eng kichik kvadratlar” usulining qo‘llanish sharti bajarilmasa, bunda geteroskedatlik holati hosil bo‘ladi (8.1-rasm.). Geteroskedatlik regressiya tenglamasining parametrlari samaradorligini pasayishiga ta’sir qilmoqda.



8.1-rasm. Geteroskedatlik holatlari

To‘rtinchi taxmin: qoldiqdagi avtokorrelyatsiya bilan bog‘liq. Taxmin qilinadiki, xatolar orasida avtokorrelyatsiya yo‘q, ya’ni avtokorrelyatsiya mavjud emas.

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

Bu taxmin shuni anglatadiki, agar bugun natijadagi ishlab chiqarish kutilgandan ko‘p bo‘lsa, bundan ertaga ishlab chiqarish ko‘p (yoki kam) bo‘ladi degan xulosaga kelish kerak emas. Birinchi va to‘rtinchi taxmin birgalikda ehtimollik nuqtai-nazaridan, taqsimot xatolariga bog‘liq emas deyish imkonini beradi. Shuning uchun $\varepsilon_1, \varepsilon_2, \dots \varepsilon_n$ o‘zgaruvchini o‘xshash va erkin taqsimlanishi sifatida qaralishi mumkin. $E(\varepsilon_i) = 0$ bo‘lgani uchun:

$$Var(\varepsilon_i) = E(\varepsilon)^2$$

Bundan

$$Cov(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j)$$

Beshinchi taxmin: X erkin o‘zgaruvchi stoxastik emasligini tasdiqlaydi. Boshqacha qilib aytganda, X ning qiymatlari nazorat qilinadi yoki butunlay bashorat qilinadi. Bu taxminni muhim qo‘llanilishi shundan iboratki, i va j ning barcha qiymatlari uchun:

$$E(\varepsilon_i, X_j) = X_j E(\varepsilon_i) = 0$$

Beshinchi taxmin: X qiymatlari stoxastik emas, ular tanlashda tanlov miqyosidan qat’iy nazar o‘xshash

$$\left(\frac{1}{n}\right)\sum_{n=1}(X_i - X)^2$$

noldan farq qiladi va uning n→∞ limiti chekli son.

To‘g‘ri, amaliyotda ko‘rsatilgan taxminlarni mutloq mavjudligiga aniq erishish qiyin, lekin biz agar bu taxminlarga taxminan amal qilinsa qoniqish hosil qilamiz. Yuqorida keltirib o‘tilgan taxminlar klassik chiziqli regression model tuzish, regressiya parametrlarini hisoblash uchun zarur. Regression tenglama va besh taxmin bilan keltirilgan regression modelning to‘liq spetsifikatsiyasidan so‘ng, endi uni ayrim o‘ziga xos tomonlarini ko‘rib chiqamiz. Avvalombor, Y bog‘liq o‘zgaruvchining taqsimot ehtimoliga qaytamiz.

Y_i funksiyaning birinchi o‘rtachasi, tenglamaning ikki qismini matematik kutilishi sifatida olinishi mumkin:

$$E(Y_i) = E(\alpha + \beta X_i + \varepsilon_i) = \alpha + \beta X_i$$

Bu, α va β parametrlar spetsifikatsiyasidan, X_i ning stoxastik emasligidan (bu berilgan son) va $e_i = 0$ o‘rtachadan (ikkinchi taxmin) kelib chiqadi. Keyin Y_i variatsiya bo‘lmish:

$$Var(Y_i) = E[Y_i - E(Y_i)]^2 = E[(\alpha + \beta X_i + \varepsilon_i) - (\alpha + \beta X_i)]^2 = E(\varepsilon_i^2) = \sigma^2$$

Har bir X bog‘liq o‘zgaruvchiga Y o‘zgaruvchini o‘rtacha qiymatini beruvchi tenglama regressiyaning empirik chizig‘i deyiladi. Bu chiziqni ordinata bilan kesishishi, X ning nolga teng qiymatida Y bahosini o‘lchaydigan α kattalikka mos keladi. β ning og‘ishi, Y qiymatni X qiymatning har bir qo‘shimcha birligiga og‘ishdagi o‘zgarishini o‘lchaydi. Masalan, agar Y yalpi iste’mol, X yalpi daromad ko‘rinishida bo‘lsa, u holda β nolga teng daromadda iste’mol darajasining chegaraviy og‘ishini namoyon qiladi. Bu o‘lchamlar qiymatlari noma’lum bo‘lgani uchun regressiyaning empirik chizig‘i ma’lum emas. α va β ning o‘lchamlari qiymatlarini hisoblab, regressiyaning nazariy chizig‘ini olamiz. α va β ning qiymatlari $\widehat{\alpha}$ va $\widehat{\beta}$ hisoblangandek mos hisoblangan bo‘lsa, mos xolda, bunda regressiyaning nazariy chizig‘i quyidagi tenglama orqali berilgan:

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} x_i$$

bunda, $\hat{Y}_i -$ Y ning tekislangan qiymati. Barchasi bo‘lmasa ham, ko‘pchiligi Y empirik qiymatlar nazariy chiziqda yotmaydi, shuning uchun Y_i va $\hat{Y}_i$ qiymatlar mos kelmaydi. Bu farq qoldiq deb ataladi va Ei bilan belgilanadi. Shuning uchun quyidagi tenglamalar farqlanadi:

$$Y_i = \alpha + \beta X_i + \varepsilon_i \qquad (empirik)$$
$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i + \varepsilon_i \qquad (nazariy)$$

Nazorat uchun savollar

1. Avtokorrelyatsiya qachon vujudga keladi?
2. Avtokorrelyatsiyani necha xil usul yordamida bartaraf etish mumkin?
3. Ekonometrik modelni real o‘rganilayotgan jarayonga mos kelishini qaysi mezon yordamida aniqlash mumkin?
4. Ekonometrik modeldagi parametrlardan birortasi ishonchsiz bo‘lsa, uni nima qiilish mumkin?
5. Darbin-Uotson mezoni qiymati qaysi oraliqda o‘zgaradi?

6. Bashorat modelini adekvatligini baholovchi mezonlari.

7. Omillarni tanlash va bosqichini asosiy shartlarini aytib bering.

8. Korrelyatsiya koeffitsiyentini mustahkamlashni aniqlashda Styudent mezonini qo‘llanilishi.

9. Bashorat modelini tanlashda qanday mezonlar qo‘llanadi?

Masala. O‘zbekiston Respublikasi kelt jamiyatining iqtisodiy madaniyati bo‘yicha 7 ta viloyat markazlari bilan uchrashdi: oziq-ovqat milliy organi, manfaatlarni qondirish, %; X-Milliy Kengash.

Viloyatlar	y	x
Andijon	800	25
Buxoro	6200	260
Navoiy	400	20
Namangan	600	30
Jizzax	300	10
Qashqadaryo	700	40
Qoraqalpog‘iston	400	15
Samarqand	6000	270
Sirdaryo	350	13
Surxondaryo	700	25
Farg‘ona	1200	40
Xorazm	3000	200
Toshkent Shahri	4000	300
Toshkent Viloyati	700	25

Topshiriq. Berilgan ma’lumotlar asosida ekonometrik model tuzilsin va mezonlar bo‘yicha tekshirilsin.

Yechim. Excel dasturini ishga tushirib, ma’lumotlarni yuqoridagi tartibda kiritamiz. Bunda asosiy omil sifatida – y ni, ta’sir etuvchi omil sifatida – x ni belgilaymiz.

1. Avvalo, x va y ning summalari va o‘rtacha qiymatlarini excel funksiyasi yordamida topib olamiz:

Viloyatlar	y	x
Andijon	800	25
.	.	.
.	.	.
Toshkent Viloyati	700	25
Summa	25350	1273
O‘rtacha	1810,714	90,92857

2. Keyin x^2, x*y va (y-y_)^2 qiymatlarini topib, ularning summasini hisoblaymiz. So‘ngra, excel funksiyasidan foydalanib, korrelyatsion-regression tahlillarni amalga oshiramiz va modelimiz parametrlarini eng kichik kvadratlar usulining normal tenglamalar sistemasi orqali topib olamiz. Demak, natijalar quyidagicha:

r(y/x)=	***Столбец 1***	***Столбец 2***
Столбец 1	1	
Столбец 2	0,946540447	1

α0=	180,52299
α1=	17,928262

Bundan ko‘rinadiki, modelimiz ko‘rinishi quyidagicha:

y^=180,52+17,92*x+ε

3. Endilikda, y hisoblangan qiymatlarini topamiz. Buning uchun yuqoridagi modeldan foydalanamiz:

$\hat{y}$	628,73	4841,87	539,088	718,371	359,806	897,653	449,447	5021,15	413,59	628,73	897,653	3766,18	5559	628,73

4. Ana endi topgan parametrlarimizning ishonchliligini t-styudent mezoni orqali tekshiramiz:

m(α0)=	**248,266272**	**m(α1)=**	**1,763814**	**m(r)=**	**0,093122313**
t(α0)=	0,72713458	t(α1)=	10,164486	t(r)=	10,16448598

5. So‘ngra, modelimiz sifatini tekshirish uchun determinatsiya koeffitsiyenti orqali baholaymiz. Buning uchun quyidagi formuladan foydalanamiz:

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

Ushbu formuladan foydalanishimiz uchun bizga, (y-y^)^2 qiymat kerak bo‘ladi. Shuning uchun biz (y-y^)^2 qiymat (y-y^=e) ni va uning summasini excel funksiyasi yordamida topib olamiz:

e^2
29333,566
1844513,75
19345,5387
14011,6618
3576,71173
39066,9022
2444,99874
958139,771
4043,73943
5079,47659
91413,412
587024,873
2430486,4
5079,47659
6033560,28

6. Ana endi topilgan qiymatlarni determinatsiya formulasi o‘rniga qo‘yib, natijani olamiz:

7.

R^2= 0,895938817

8. Endilikda modelimizning ahamiyatini Fisher mezoni orqali tekshiramiz va buni quyidagi formula yordamida excel orqali amalga oshiramiz:

$$F = \frac{R^2}{1 - R^2} \cdot \frac{n - k - 1}{k}$$

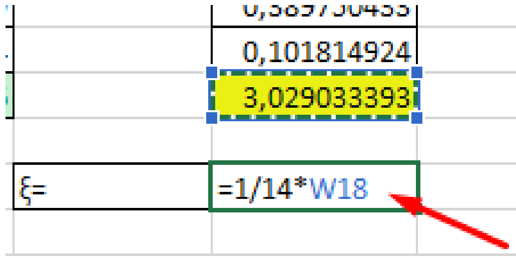
Demak, natija quyidagicha:

F= 103,3167753

9. So‘ngra, modelimizning ahamiyatini yana bir testga, ya’ni approksimatsiya xatoligiga tekshirib ko‘ramiz. Buning uchun yangi ustunga (y-y^)/y qiymatni excel funksiyasi yordamida ya’ni ABS buyrug‘i orqali moduldan chiqarib, qiymatlarni topib olamiz. Hamda ushbu o‘zgaruvchilarning ham summasini hisoblaymiz:

ABS((y-y^)/y)
=ABS((C4-N4)/C4)
0,219053027
0,347720602
0,197284776
0,199352054
0,282362129
0,123617321
0,163141024
0,181686867
0,101814924
0,251955425
0,255391828
0,389750433
0,101814924
3,029033393

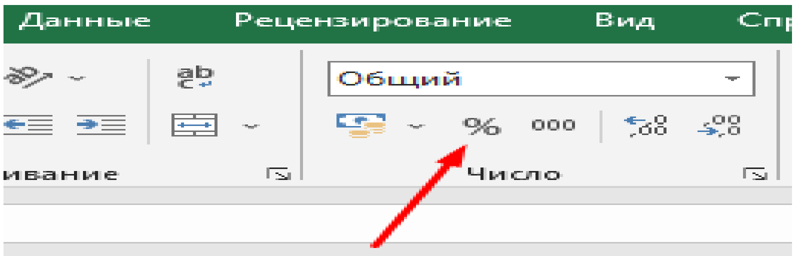
10. Keyin topilgan natijalarni quyidagicha hisoblaymiz:



Va quyidagi natijaga erishamiz:

$$\xi = 0{,}216359528$$

11. Shuningdek, ushbu topilgan natijani 100 % ga ko‘paytiramiz. Bu amalni excel yordamida osonlikcha amalga oshiramiz:



Demak rasmda ko‘rinib turganidek, natijamizni belgilab yuqorida ko‘rsatilgan % belgisini bossak, approksimatsiya xatoligi natijasiga erishamiz va u quyidagicha:

$$\xi = 22\%$$

12. Ana endi so‘nggi testimiz, ya’ni qatorlarda qoldiq avtokorrelyatsiyasini Darbin-Uotson mezoni bo‘yicha tekshiramiz. Buning uchun quyidagi formuladan foydalanamiz:

$$dw = \frac{\sum_{i=1}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$$

Formuladan maxraj qismi topilgan. Demak endi surat qismini topib olamiz. Buning uchun yangi ustunga y-y^=e qiymatni topamiz.

Ana endi (e-e(i-1))^2 qiymatni excelda topib olamiz, ya'ni bitta katak qoldirib quyidagicha hisoblaymiz. Shuningdek, summasini topamiz:

e	(e-e(i-1))^2
171,27045	
1358,1288	=(AC5-AC4)^2
-139,0882	2241658,77
-118,3709	429,2096412
-59,80562	3429,888427
-197,6535	19002,03636
-49,44693	21965,18491
978,84614	1057386,641
-63,5904	1086673,953
71,270447	18187,44897
302,34651	53396,1469
-766,1755	1141739,253
-1559,002	628573,4572
71,270447	2657787,374
	10338862,04

13. Endilikda oxirgi qiladigan amalimiz, bu topgan natijalarimizni formula o'rniga qo'yib hisoblaymiz va quyidagi natijaga erishamiz:

$$dw = 1{,}713559087$$

Xulosa: Demak, natijalarimizni t-styudentning jadval qiymatiga solishtiradigan bo'lsak, bunda r va a_1 parametrlarimiz t-jadval qiymatidan katta bo'lganligi uchun ishonchli.

$$t_{jadval} = 2{,}1448$$
$$t_{\alpha 0} = 0{,}7271346$$
$$t_{\alpha 1} = 10{,}164486$$
$$t_r = 10{,}164486$$

Shuningdek, a_0 parametrimiz esa t-jadval qiymatidan kichik bo'lgani uchun ishonchsizligini izohladi.

$$R\wedge 2 = 0{,}895938817$$

Determinatsiya koeffitsiyenti natijasidan ko'rinadiki, regressiya tenglamasi natijaviy ko'rsatkich dispersiyasini 89% tashkil qiladi.

$$F_{jadval} = 3,74$$
$$F = 103,31678$$

Fisher natijasi shuni ko‘rsatdiki, F-qiymat ahamiyatli. Ya’ni bunda cheklanmagan modeldan cheklangan modelga kvadrat xatolar yig‘indisini solishtirishimiz mumkin.

$$\xi = 22\%$$

Modelimizda approksimatsiya xatoligimiz 22 % ni tashkil qildi. Ya’ni bu nazariy y ning haqiqiy y qiymatlaridan o‘rtacha nisbiy chetlanishni izohlab berdi. Demak, o‘rtacha hisoblangan qiymatlar haqiqiy qiymatga nisbatan 22 % ga nisbatan og‘ganligini ko‘rsatmoqda.

$$dw = 1,713559087$$
$$d2=1,35 \le d=1,71 \le 2$$

Biz topgan natijamiz 1,71 ga teng bo‘lganligi sababli, bu 2-qoidamizga, ya’ni biz topgan natijamiz d2 dan katta va 2 dan kichik degan qoidamizga to‘g‘ri keladi. Bundan ko‘rinadiki, qatorlarda qoldiq avtolorrelyatsiya mavjud emas.

IX BOB. ENG KICHIK KVADRATLAR USULI VA GAUSS-MARKOV TEOREMASI

9.1. Eng kichik kvadratlar (EKK) usuli

Regressiya modelining $\hat{a}_0$ va $\hat{a}_1$ parametrlarini hisoblash uchun ma'lum bir qoida yoki metod kerak. Ilmiy adabiyotlarda bir nechta metod va usullar mavjud bo'lsada, ularda eng ko'p tarqalgani eng kichik kvadratlar (EKK) usulidir. Bu usul yordamida o'tkazilgan regressiya chizig'i, ya'ni ($\hat{a}_0$,$\hat{a}_1$) parametrlari shunday tanlanadiki, unda har bir nuqtada chiziqqacha bo'lgan vertical masofalar ($\hat{e}_i$) kvadratlari yig'indisi minimallashadi, ya'ni:

$$S(\beta_0, \beta_1) = \hat{\beta}_0^{min} \hat{\beta}_1 \sum \hat{u}_i^2 = \sum \left[y_i - \left(\hat{\beta}_0 + \hat{\beta}_0 x_i\right)\right]^2$$

Muammo esa $\hat{a}_0$ va $\hat{a}_1$ larni qulay yo'l bilan topishdadir. Agar y va x bo'yicha n nafar kuzatishlar yig'ilgan bo'lsa, unda biz "kvadratlar yig'indisi" funksiyasini kamaytiradigan noma'lum parametirlar $\hat{a}_0$ va $\hat{a}_1$ uchun qiymatlarni topishni istaymiz va quyidagi formula orqali topishimiz ham mumkin:

$$\hat{\beta}_1 = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Bu yerda $\bar{y} = \sum \frac{y_i}{n}$ va $\bar{x} = \sum \frac{x_i}{n}$.

Masofalarning kvadratlari olinishining sababi shundaki, kvadrat olinmasdan yig'indi olinsa, musbat (nuqta chiziqdan yuqorida bo'lsa) va manfiy (nuqta chiziqdan pastda bo'lsa) masofalar bir-birini neytral-lashtiradi.

EKK usuli yordamida hisoblangan regressiya chizig‘ini quyidagicha yozish mumkin

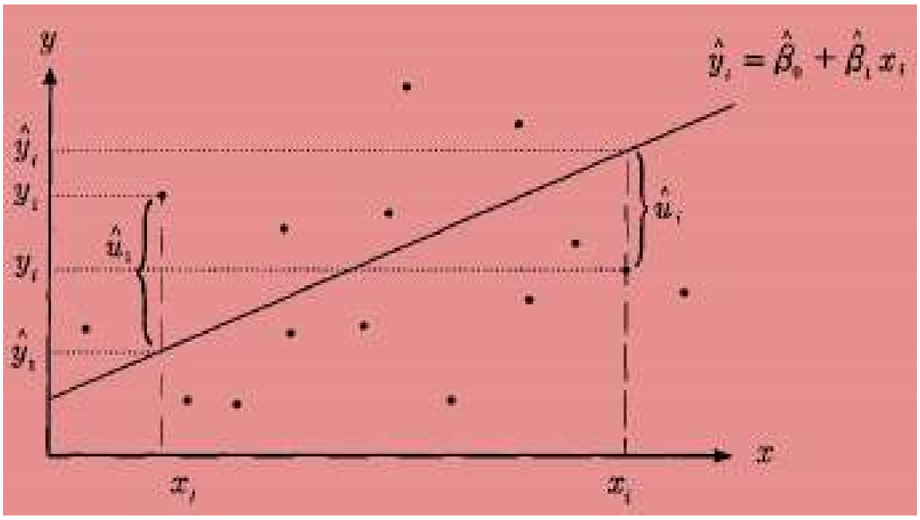
$$\hat{C}_i = \hat{\beta}_0 + \hat{\beta}_1 INC_i$$

Bunda, $\hat{a}_0$ va $\hat{a}_1$ tanlamadan hisoblangan parametrlar, $\hat{C}_i$ esa i tartibdagi oilaning modellashtirilgan yoki hisoblangan oziq-ovqat iste’moli.

Har bir nuqtadan hisoblangan regressiya chizig‘igacha vertikal masofa eng kichik kvadratlar qoldig‘i yoki qoldiq deyiladi va ular quyidagicha ifodalanadi:

$$\hat{u}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

Regressiya parametrlarini hisoblashdan chiqqan qoldiqlar $\hat{u}_i$ har bir qoldiq bilan bog‘liq $\hat{y}_i$ va y o‘rtasidagi munosabat quyidagi rasmda tasvirlangan.



Bugungi kunda bu kabi hisob-kitoblar Microsoft Excel, Stata, EViews kabi kompyuter dasturlari yordamida amalga oshiriladi. Har bir dasturiy to‘plamning regressiya natijasi bir biridan biroz farq qiladi, chunki ishlatiladigan atamalar turlicha.

Hisoblangan parametrlarni iqtisodiy talqini zarur masala sifatida ko‘riladi. Ko‘p tadqiqotchilar amaliyotda murakkab hisob-kitoblarni amalga oshirsa-da, ularni talqin qilishda yetarli ahamiyat berishmaydi.

Regressiya chizig‘ining ordinata o‘qi bilan kesishgan nuqtasi $\hat{a}_0$ va qiyaligi $\hat{a}_1$ hisoblanadi.

Ekonometrik amaliyotda $\hat{a}_0 < 0$ bo‘lib chiqish holatlari ko‘p uchraydi. Statistik nuqtai nazardan $E(u) = 0$ shartini qoniqtirish uchun $\hat{a}_0 < 0$ bo‘lishi kerak, ammo uning iqtisodiy talqini mavjud emas. Shu nuqtai nazardan regression hisob natijalarini tahlil qilishda $\hat{a}_0$ qandaydir iqtisodiy ma’noga ega bo‘lishi yoki bo‘lmasligi muhim emas, $\hat{a}_1$ qiyalik parametrining iqtisodiy talqini esa muhim.

9.2. Gauss-Markov teoremasi

Stoxastik tahlil bilvosita aloqalarni o‘rganishga yo‘naltirilgan, ya'ni ta'siri to‘g‘ri bo‘lmagan, boshqalar orqali ta'sir etadigan (uzluksiz zanjir bo‘yicha to‘g‘ri aloqalarni aniqlash imkoni bo‘lmaganida) omillarni aniqlashga qaratilgan. Bundan determinallashgan va stoxastik tahlil o‘rtasidagi munosabat (nisbat) haqida muhim xulosa kelib chiqadi: to‘g‘ri bog‘lanishlarni birinchi navbatda o‘rganish kerakligi uchun, stoxastik tahlil yordamchi xarakter kasb etadi. Stoxastik tahlil omillar bo‘yicha determinallashgan model tuzib bo‘lmaganida, ularni chuqur determinallashgan tahlilni o‘rganish quroli sifatida foydalaniladi.

Xo‘jalik faoliyatining ba'zi bir tomonlarini o‘zaro bog‘lanishlarini omilli tahlilining stoxastik modellashtirish xo‘jalik faoliyati omillari va natijalarining miqdoriy xarakteristikalari - iqtisodiy ko‘rsatkichlarning qiymatlarini tebranish qonuniyatlarini umumlashtirishga tayanadi. Bog‘lanishlarning miqdoriy parametrlari xo‘jalik obyektlari to‘plami yoki davrlarida o‘rganilayotgan ko‘rsatkichlarning qiymatlarini qiyoslash (solishtirish) asosida aniqlanadi. Shunday qilib, stoxastik modellashtirishning birinchi asosi bo‘lib, kuzatishlar to‘plamini tashkil eta olish, ya'ni bir hodisa parametrini turli sharoitlarda qaytadan o‘lchash imkoniyatlari hisoblanadi.

Determinallashgan omilli tahlilda o‘rganilayotgan obyektning modeli xo‘jalik obyektlari va davrlari bo‘yicha o‘zgarmaydi (negaki, mos keluvchi asosiy kategoriyalarning nisbati barqarordir). Alohida xo‘jaliklar yoki bir xo‘jalikni turli, alohida davrlardagi faoliyatlari natijalarini qiyoslash zaruriyati tug‘ilganida model asosida aniqlangan miqdoriy analitik natijalarni qiyoslash haqida savol tug‘ilishi mumkin. Stoxastik tahlilda modelning o‘zi empirik ma'lumotlar to‘plami asosida tuzilgani uchun, haqiqiy modelni hosil qilishning asosi bo‘lib barcha dastlabki kuzatishlar bo‘yicha bog‘lanishlarning miqdoriy xarakteristikalarini mos kelishi hisoblanadi. Bundan kelib chiqadiki, ko‘rsatkichlarning qiymatini o‘zgarishi hodisalarni bir xildagi aniqlik chegarasida amalga oshishi kerak, ularning xarakteristikalari bo‘lib modellashtirilayotgan iqtisodiy ko‘rsatkichlar hisoblanadi (o‘zgarish chegarasida ifodalanayotgan hodisaning xarakterida sifatning keskin o‘zgarishi (sakrashi) ro‘y bermasligi kerak). Shunday ekan, bog‘lanishlarni modellashtirishda stoxastik yondashishning qo‘llanishini ikkinchi asosi bo‘lib, to‘plamni sifatli, bir jinsliligi hisoblanadi.

Iqtisodiy ko‘rsatkichlarning o‘rganilayotgan qonuniyatlari (modellashtirilayotgan bog‘lanish) yashirin tarzda namoyon bo‘ladi. O‘rganish, izlanish nuqtai nazardan bu ko‘rsatkichning tasodifiy o‘zgarishi va kovariatsiya komponentalari bilan aralashib ketadi. Katta sonlar qonuni bo‘yicha faqat katta to‘plamda bog‘lanish qonuniyatlari o‘zgarish yo‘nalishlariga tasodifiy mos kelishidan kuchliroq namoyon bo‘ladi (tasodifiy kovariatsiya). Bundan statistik tahlilning uchinchi asosi kelib chiqadi – o‘rganilayotgan qonuniyatlarni (modellashtirilayotgan bog‘lanishlarni) yetarli ishonchlik va aniqlikda aniqlash uchun kuzatishlar to‘plami yetarli darajada (miqdorda) bo‘lishi kerak. Modelni ishonchli aniqlik darajasi modelni ishlab chiqarish xo‘jalik faoliyatini boshqarishdagi amaliy maqsadlarda foydalanish mumkinligi bilan aniqlanadi.

Stoxastik tahlil yondashishining to‘rtinchi asosi – iqtisodiy ko‘rsatkichlarning bog‘lanishlarini miqdoriy parametrlarini ko‘rsatkich darajasini tebranishini ommaviy ma'lumotlaridan aniqlash imkonini beruvchi usullarning mavjudligi. Qo‘llanilayotgan usullarning matematik apparati ba'zida modellashtirilayotgan empirik ma'lumotlarga o‘ziga xos bo‘lgan talablarni qo‘yadi. Ushbu talablarni bajarish usullarini qo‘llash va olingan natijalarni ishonchli bo‘lishi uchun ahamiyatli asos hisoblanadi.

Stoxastik omilli tahlilning asosiy xususiyati shundan iboratki, stoxastik tahlilda modelni sifatli (nazariy) tahlil yo‘li bilan tuzib bo‘lmaydi, buning uchun empirik ma'lumotlarning miqdoriy tahlili zarur bo‘ladi.

Chiziqli regressiya modeliga qo‘yiladigan shartlarga e’tibor qaratamiz.

1. Erksiz o‘zgaruvchi y erkli o‘zgaruvchi x bilan quyidagicha bog‘langan:

$$Y = a_0 + a_1 x + u$$

2. Tasodifiy xatolik e matematik kutilishi 0 ga teng $E(u) = 0$. Ushbu shart quyidagi ifodaga ekvivalent hisoblanadi:

$$E(y|x) = a_0 + a_1 x$$

3. Tasodifiy xatolik u dispersiyasi o‘zgarmas (xomoskedastiklik)

$$var(u) = \sigma^2 = var(y)$$

4. Ikki xatolik o‘rtasidagi kovariatsiya 0ga teng

$$cov(u_i, u_j) = cov(y_i, y_j)$$

5. Erkli o‘zgaruvchi x tasodifiy emas va kamida ikkita qiymat qabul qiladi.

Gauss-Markov teoremasiga muvofiq, yuqorida keltirilgan chiziqli regressiya modelining shartlari qoniqtirilsa, EKK usuli yordamida hisoblangan $\hat{a}_0$ va $\hat{a}_1$ parametrlar barcha chiziqli va siljimagan baholar ichida eng kichik dispersiyaga ega bo‘ladi.

Agar chiziqli regressiya modeliga qo‘yilgan shartlardan biri buzilsa, Gauss-Markov teoremasi natijasi noto‘g‘ri bo‘lib chiqadi va boshqa hisoblash usullari qo‘llaniladi. Masalan, 3-shart (xomoskedastiklik) qoniqtirilmasa, hisoblangan baholar dispersiyasi noto‘g‘ri bo‘ladi va teorema natijasi Gauss-Markov teoremasiga muvofiq bo‘lmaydi. Gauss-Markov teoremasiga tegishli bo‘lmagan chiziqli regressiyaga qo‘yiladigan shartlardan yana biri bu tasodifiy xatolikning normal taqsimot qonuniga bo‘y sunishidir (normallik sharti): $u \sim N(0, \sigma^2)$

9.3. Eng kichik kvadratlar usuli hisoblash metodikasi

Mezon: haqiqiy miqdorlarning tekislangan miqdorlardan farqining kvadratlari yig‘indisi eng kam bo‘lishi zarur.

$$S = \sum \left(Y - \overline{Y}_t\right)^2 \to \min$$

Misol: $Y_t = a_0 + a_1 t$

Qiymat $\sum \left(Y - \overline{Y}_t\right)^2$ bo‘lishi uchun birinchi darajali hosilalar nolga teng bo‘lishi kerak.

$$S = \sum \left(Y - \overline{Y}_t\right)^2 = \sum \left(Y - a_0 - a_1 t\right)^2 \to \min$$

$$\frac{\partial S}{\partial a_0} = 0 \qquad \frac{\partial S}{\partial a_0} = 0 \qquad \to \begin{cases} n \cdot a_0 + a_1 \sum t = \sum y \\ a_0 \sum t + a_1 \sum t^2 = \sum y \cdot t \end{cases}$$

Normal tenglamalar tizimi.

$$S = \sum \left(Y - \overline{Y}_t\right)^2$$

Demak

$$\overline{Y} = a_0 + a_1 x + a_1 x^2 + \ldots + a_n x^n$$

$$\frac{\partial S}{\partial a_0} = \sum \left[2\left(Y - a_0 - a_1 X - a_2 X^2 - \ldots - a_n X^n\right)\right] \cdot (-1) = 0$$

$$\frac{\partial S}{\partial a_1} = \sum \left[2\left(Y - a_0 - a_1 X - a_2 X^2 - \ldots - a_n X^n\right)\right] \cdot (-X) = 0$$

....

$$\frac{\partial S}{\partial a_n} = \sum \left[2\left(Y - a_0 - a_1 X - a_2 X^2 - \ldots - a_n X^n\right)\right] \cdot \left(-X^n\right) = 0$$

Chiziqli funksiya bo‘yicha tekislanganda

$$\overline{Y} = a_0 + a_1 X$$

$$S = \sum \left(Y - a_0 - a_1 X\right)^2 \to \min$$

$$\begin{cases} \dfrac{\partial S}{\partial a_0} = \sum 2\left(Y - a_0 - a_1 X\right) \cdot (-1) = 0 \\ \dfrac{\partial S}{\partial a_1} = \sum 2\left(Y - a_0 - a_1 X\right) \cdot (-X) = 0 \end{cases}$$

Bunda

$$\begin{cases} \sum y - n \cdot a_0 - a_1 \sum X = 0 \\ \sum y \cdot X - a_0 \sum X - a_1 \sum X^2 = 0 \end{cases}$$

$$\begin{cases} n \cdot a_0 + a_1 \sum X = \sum y \\ a_0 \sum X + a_1 \sum X^2 = \sum y \cdot X \end{cases}$$

Nazorat uchun savollar

1. "Eng kichik kvadratlar usuli" ning mohiyatini tushuntirib bering.

2. Normal tenglamalar sistemasini yechish usullarini tushuntirib bering.

3. To‘g‘ri chiziq bo‘yicha eng kichik kvadratlar usuli yordamida tenglash qanday amalga oshiriladi?

4. Gauss-Markov teoremasining xususiyatlarini tushuntirib bering.

X BOB. INTERVALNI BAHOLASH VA ISHONCH INTERVALI

10.1. Intervalni baholash
10.2. t-taqsimot tushunchasi
10.3. Interval baholamalarini aniqlash

10.1. Intervalni baholash

Oddiy chiziqli regressiya modelidagi parametrlar uchun nuqta baholarini ishlab chiqish uchun eng kichik kvadratlar usullaridan foydalanamiz. Bu hisob-kitoblar iqtisodiy o‘zgaruvchilar o‘rtasidagi munosabatni tavsiflovchi E(y) = β_1+$\beta_2 x$ regressiya funksiyasi haqidagi xulosani ifodalaydi. Xulosa "ma’lum yoki taxmin qilingan narsadan xulosa chiqarish" degan ma'noni anglatadi. Ushbu lug‘at ta'rifi statistik xulosani ham tavsiflaydi. Biz iqtisodiy o‘zgaruvchilar o‘rtasidagi munosabatni taxmin qilamiz va regressiya modeli haqida turli xil taxminlarni olamiz. Ushbu taxminlarga asoslanib va regressiya parametrlarining empirik baholarini hisobga olgan holda, biz ma'lumotlar olingan populyatsiya haqida xulosa qilamiz.

Ushbu bobda biz statistik xulosa chiqarishning qo‘shimcha vositasi bilan tanishamiz: intervallarni baholash. Intervalli baholash – bu noma'lum parametrlar joylashishi mumkin bo‘lgan, ba'zan ishonch oraliqlari deb ataladigan qiymatlar diapazonlarini yaratish tartibi.

Intervallarni baholash jarayonlari oddiy chiziqli regressiya modelining faraziga va natijada eng kichik kvadratlar baholovchilarining normalligiga bog‘liq. Agar taxmini to‘g‘ri kelmasa, unda eng kichik kvadratlar hisoblagichlarining taqsimoti taxminan normal bo‘lishi uchun tanlama hajmi yetarlicha katta bo‘lishi kerak. Bunday holda, biz ushbu bobda ishlab chiqqan jarayonlardan foydalanish mumkin, ammo ular taxminiydir. Ushbu bobdagi jarayonlarni ishlab chiqishda biz "Students" t-taqsimotidan foydalanamiz. Bundan tashqari, ba'zida biz muhokama qilmoqchi bo‘lgan tushunchalarni oddiyroq sharoitda ko‘rish foydali bo‘ladi.

Intervalli baholash haqiqiy parametr β_2 tushishi mumkin bo‘lgan qiymatlar oralig‘ini taklif qiladi. Qiymatlar diapazonini taqdim etish parametr qiymati qanday bo‘lishi mumkinligini va biz uni qanday aniqlik bilan baholaganimiz haqida ma'lumot beradi. Bunday intervallarni ko‘pincha ishonch intervallari deyiladi. Biz ularni intervalli taxminlar deb atashni ma'qul ko‘ramiz, chunki "ishonch" atamasi noto‘g‘ri tushuniladi va noto‘g‘ri ishlatiladi. Ko‘rib turganimizdek, bizning ishonchimiz oraliqlarning o‘zida emas, balki intervallarni olish uchun foydalanadigan jarayonga bog‘liq.

10.2. t-taqsimot tushunchasi

Taxminlar oddiy chiziqli regressiya modeli uchun amal qiladi deb faraz qilaylik. Bu holda biz bilamizki, eng kichik kvadratlar usullari b_1 va b_2, normal taqsimotlarga ega. Masalan, b_2 ning normal taqsimoti, β_2 ning eng kichik kvadratlar usullari:

$$b^2 \sim N\left(\beta^2, \frac{\sigma^2}{\sum(x_i - \overline{x})^2}\right)$$

Standartlashtirilgan normal tasodifiy miqdor b_2 dan uning o‘rtacha va qiymatini ayirish yo‘li bilan olinadi, uning standart og‘ishiga bo‘lish mumkin.

$$Z = \frac{b^2 - \beta^2}{\sqrt{\frac{\sigma^2}{\sum(x_i - \overline{x})^2}}} \sim N\,(0{,}1) \qquad (10.1)$$

Standartlashtirilgan tasodifiy o‘zgaruvchi Z odatda o‘rtacha 0 va dispersiya 1 bilan taqsimlanadi. Oddiy ehtimollar jadvalidan foydalanib, biz buni quyidagicha ifodalaymiz:

$$P\,(-1{,}96 \le Z \le 1{,}96) = 0{,}95$$

Ushbu ifodaga (10.1) ni almashtirib, biz quyidagini hosil qilamiz:

$$P\left(-1,96 \leq \frac{b_2 - \beta_2}{\sqrt{\frac{\sigma^2}{\sum(x_i - \overline{x})^2}}} \leq 1,96\right) = 0,95$$

Qayta tartibga solish bizga quyidagi natijani beradi.

$$P\left(b_2 - 1,96\sqrt{\frac{\sigma^2}{\sum(x_i - \overline{x})^2}} \leq \beta_2 \leq b_2 + 1,96\sqrt{\frac{\sigma^2}{\sum(x_i - \overline{x})^2}}\right) = 0,95$$

Bu β_2 parametrini o‘z ichiga olishi ehtimoli 0,95 ga teng bo‘lgan intervalni belgilaydi. Ikki so‘nggi nuqtalar $\left(b_2 \pm 1,96\sqrt{\frac{\sigma^2}{\sum(x_i - \overline{x})^2}}\right)$ intervalli baholovchini taqdim eting. Takroriy namuna olishda, shu tarzda tuzilgan intervallarning 95% parametrning haqiqiy β_2 qiymatini o‘z ichiga oladi. Intervalli baholovchining bu oson chiqarilishi va bizning taxminimizga asoslanadi hamda bunda σ^2 sigma belgisi dispersiyasini bildiradi.

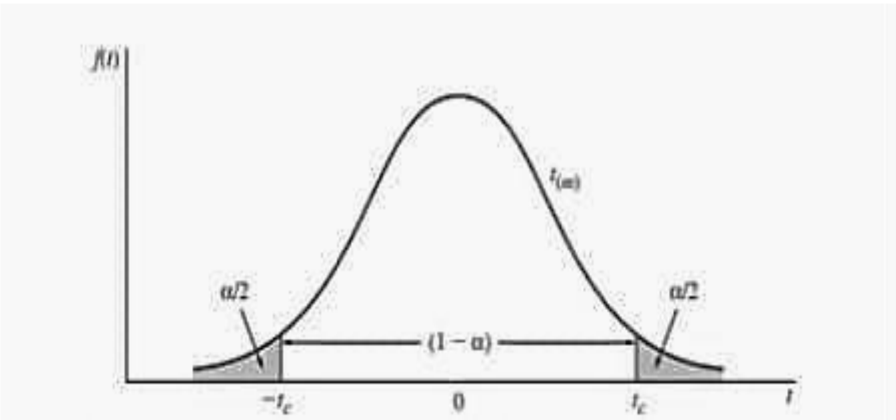
Biz σ^2 qiymatini bilmasak ham, uni taxmin qilishimiz mumkin. Eng kichik kvadratlar qoldiqlari $\hat{e}_i = y_i - b_1 - b_2 x_i$ va σ^2 ni baholovchimiz $\hat{\sigma}^2 = \frac{\sum \hat{e}_i^2}{N-2}$. (10.1) formulada σ^2 ni $\hat{\sigma}^2$ ga almashtirish biz ishlashimiz mumkin bo‘lgan tasodifiy o‘zgaruvchini yaratadi, ammo bu almashtirish ehtimollik taqsimotini standart normaldan $N - 2$ erkinlik darajasiga ega bo‘lgan t-taqsimotiga o‘zgartiradi,

$$t = \frac{b_2 - \beta_2}{\sqrt{\frac{\hat{\sigma}^2}{\sum(x_i - \overline{x})^2}}} = \frac{b_2 - \beta_2}{\sqrt{\widehat{var}(b_2)}} = \frac{b_2 - \beta_2}{se(b_2)} \sim t_{(N-2)} \qquad (10.2)$$

$t = \frac{b_2 - \beta_2}{se(b_2)}$ nisbati $N - 2$ erkinlik darajasi bilan t-taqsimotiga ega, biz buni $t \sim t_{(N-2)}$ deb belgilaymiz. Shunga o‘xshash natija b_1 uchun amal qiladi. Shuning uchun aytishimiz mumkinki, agar oddiy chiziqli regressiya modelidagi taxminlari mavjud bo‘lsa, u holda

$$t = \frac{b_k - \beta_k}{se(b_k)} \sim t_{(N-2)} \qquad uchun \quad k=1,2 \tag{10.3}$$

Ushbu tenglama oddiy chiziqli regressiya modelida intervallarni baholash va gipotezalarni tekshirish uchun asos bo‘ladi. T-taqsimot bilan ishlaganda, uning markazida nol joylashgan qo‘ng‘iroq shaklidagi egri chiziq ekanligini unutmang. Bu standart normal taqsimotga o‘xshaydi, bundan tashqari u ko‘proq tarqalgan, a bilan katta dispersiya va qalin dumlar mavjud bo‘ladi. T-taqsimot shakli bitta tomonidan boshqariladi va parametr **erkinlik darajalari** deb ataladi, odatda uni ***df*** deb qisqartiriladi. Erkinlik darajasi m bo‘lgan t-taqsimotini belgilash uchun t_m belgisidan foydalanamiz.



10.1-rasm t-taqsimotdan kritik qiymatlar

m erkinlik darajasi uchun t-taqsimotining 95-foizli $t_{(0{,}95,m)}$ bilan belgilanadi. Bu qiymat 0,95 ehtimollikning chap tomoniga to‘g‘ri keladigan xususiyatga ega, shuning uchun $P\left[t_{(m)} \le t_{(0{,}95,m)}\right] = 0{,}95$. Masalan, agar erkinlik darajalari m = 20 bo‘lsa, u holda t-styudentning taqsimot jadvalida $t_{(0.95,20)} = 1{,}725$ ga teng. Agar siz berilmagan foizlarni talab qiladigan muammoga duch kelsangiz, siz uchun interpolyatsiya qilishingiz mumkin. Aniq qiymatni olish uchun kompyuter dasturidan foydalanishingiz mumkin.

t-taqsimot hosilasi. Ushbu bobdagi intervallarni baholash va t-taqsimot jarayonlari quyidagilarni o‘z ichiga oladi. Bu yerda biz asosiy natijani ishlab chiqamiz. Eng kichik kvadratlar hisoblagichining normal taqsimlanishi kerak bo‘lgan birinchi natijadir. Masalan, $\boldsymbol{b_2}$ning eng kichik kvadratlari usullari $\boldsymbol{\beta_2}$ ning normal taqsimotini ko‘rib chiqing.

$$b_2 \sim N\left(\beta_2, \frac{\sigma^2}{\sum(x_i - \bar{x})^2}\right)$$

Standartlashtirilgan normal tasodifiy miqdor b_2 dan uning o‘rtacha qiymatini ayirish va standart og‘ishiga bo‘lish yo‘li bilan olinadi:

$$Z = \frac{b_2 - \beta_2}{\sqrt{var(b_2)}} \sim N(0,1)$$

Ya'ni, standartlashtirilgan tasodifiy o‘zgaruvchi Z odatda o‘rtacha 0 va dispersiya 1 bilan taqsimlanadi. Jumboqning ikkinchi qismi kvadrat tasodifiy o‘zgaruvchini o‘z ichiga oladi. Agar taxmini bajarilsa, tasodifiy xato atamasi e_i normal taqsimotga ega, $e_i \sim N(0, \sigma^2)$. Shunga qaramay, biz tasodifiy o‘zgaruvchini uning standart og‘ishiga bo‘lish$\frac{e_i}{\sigma} \sim N(0,1)$ orqali standartlashtirishimiz mumkin. Standart oddiy tasodifiy o‘zgaruvchining kvadrati chi-kvadrat tasodifiy o‘zgaruvchidir, shuning uchun $(e_i/\sigma)^2 \sim X^2{}_{(1)}$ bir darajadagi erkinlik bilandir. Agar barcha tasodifiy xatolar mustaqil bo‘lsa:

$$\sum\left(\frac{e_i}{\sigma}\right)^2 = \left(\frac{e_1}{\sigma}\right)^2 + \left(\frac{e_2}{\sigma}\right)^2 + \cdots + \left(\frac{e_N}{\sigma}\right)^2 \sim X^2{}_{(N)}$$

Haqiqiy tasodifiy xatolar kuzatilmaydigan bo‘lgani uchun biz ularni namunaviy tengdoshlari bilan almashtiramiz, eng kichik kvadrat qoldiqlari $\hat{e}_i = y_i - b_1 - b_2 x_i$ bilan olish mumkin.

$$V = \frac{\sum \hat{e}_i{}^2}{\sigma^2} = \frac{(N-2)\hat{\sigma}^2}{\sigma^2}$$

V tasodifiy o‘zgaruvchisi $X^2{}_{(N)}$taqsimotiga ega emas, chunki eng kichik kvadratlar qoldiqlari mustaqil tasodifiy o‘zgaruvchilar emas. Barcha N qoldiq $\hat{e}_i = y_i - b_1 - b_2 x_i$ eng kichik kvadratlar b_1 va b_2

baholovchilariga bog‘liq. Oddiy chiziqli regressiya modelida eng kichik kvadrat qoldiqlaridan faqat $N-2$ tasi mustaqil ekanligini ko‘rsatish mumkin. Binobarin, tasodifiy miqdor $N-2$ erkinlik darajasiga ega chi-kvadrat taqsimotiga ega. Ya'ni, $\frac{N-2}{\sigma^2}$ doimiysiga ko‘paytirilganda, $\hat{\sigma}^2$ tasodifiy o‘zgaruvchisi $N-2$ erkinlik darajasiga ega x-kvadrat taqsimotiga ega bo‘ladi,

$$V = \frac{(N-2)\hat{\sigma}^2}{\sigma^2} \sim X^2{}_{(N-2)}$$

Biz chi-kvadrat tasodifiy o‘zgaruvchisi V statistik jihatdan eng kichik kvadratlarning b_1 va b_2 baholovchilaridan mustaqil ekanligini aniqlamadik. V va Z ikkita tasodifiy o‘zgaruvchilardan t-tasodifiy o‘zgaruvchini hosil qilishimiz mumkin. T-tasodifiy o‘zgaruvchi standart normal tasodifiy o‘zgaruvchini $Z \sim N(0{,}1)$ ga bo‘lish yo‘li bilan hosil bo‘ladi. Mustaqil chi-kvadrat tasodifiy o‘zgaruvchining kvadrat ildizi bo‘yicha, $V \sim X^2{}_{(m)}$ erkinlik darajalari bo‘yicha, m dir.

$$t = \frac{Z}{\sqrt{\frac{V}{m}}} \sim t_{(m)}$$

T-taqsimot shakli to‘liq erkinlik darajalari m parametri bilan belgilanadi, taqsimoti esa $t_{(m)}$ ramz bilan ifodalanadi. Foydalanish Z va V dan mos ravishda bizda mavjud.

$$t = \frac{Z}{\sqrt{\frac{V}{N-2}}}$$

$$= \frac{\dfrac{b_2 - \beta_2}{\sqrt{\dfrac{\sigma^2}{\sum (x_i - \bar{x})^2}}}}{\sqrt{\dfrac{\frac{(N-2)\hat{\sigma}^2}{\sigma^2}}{N-2}}}$$

$$= \frac{b_2 - \beta_2}{\sqrt{\frac{\hat{\sigma}^2}{\sum(x_i - \bar{x})^2}}} = \frac{b_2 - \beta_2}{\sqrt{\widehat{var(b_2)}}} = \frac{b_2 - \beta_2}{se(b_2)} \sim t_{(N-2)}$$

Oxirgi satr biz (10.2) formulada umumlashtirgan holda (10.3) formula bayon qilgan asosiy natijadir.

10.3. Interval baholamalarini aniqlash

t-styudentning taqsimot jadvalidan biz t-taqsimotdan “kritik qiymat” t_c ni topishimiz mumkin, shunday qilib $P(t \geq t_c) = P(t \leq -t_c) = \frac{\alpha}{2}$, bu yerda α ehtimollik odatda α=0,01 yoki α=0,05 sifatida qabul qilinadi. m erkinlik darajalari uchun kritik qiymat t_c - foizli qiymat $t_{\left(t-\frac{\alpha}{2},m\right)}$. t_c va uning qiymatlari 10.1-rasmda tasvirlangan.

Har bir soyali “dum” maydoni α/2 ehtimollikni o‘z ichiga oladi, shuning uchun 1-α ehtimollik markaziy qismda joylashgan. Shunday qilib, biz ehtimollik bayonotini qilishimiz mumkin.

$$P(-t_c \leq t \leq t_c) = 1 - \alpha \qquad (10.4)$$

95% ishonch oralig‘i uchun kritik qiymatlar 1 - α = 0,95 ehtimolini o‘z ichiga olgan t-taqsimotining markaziy hududini belgilaydi. Bu ikki dum o‘rtasida teng bo‘lingan α = 0,05 ehtimolini qoldiradi, shuning uchun α/2=0,025 ga teng bo‘ladi. Keyin kritik qiymat $t_c = t_{(t-0,025,m)} = t_{(0,975,m)}$. Oddiy regressiya modelida erkinlik darajalari $m = N - 2$ ga teng, shuning uchun (10.4) ifodaga aylanadi

$$P\left[-t_{(0,975,N-2)} \leq t \leq t_{(0,975,N-2)}\right] = 0,95$$

Biz t-studentning taqsimot jadvalida ehtimollik qiymatlarini topamiz $t_{(0,975,N-2)}$. Keling, intervalni baholash tartibini yaratish uchun barcha formulalarni qanday qilib birlashtirishimiz mumkinligini ko‘rib chiqaylik. Olish uchun (10.3) dan (10.4) ni almashtiring

$$P\left[-t_c \leq \frac{b_k - \beta_k}{se(b_k)} \leq t_c\right] = 1 - \alpha$$

Olish uchun ushbu ifodani qayta tartiblang

$$P[b_k - t_c se(b_k) \leq \beta_k \leq b_k + t_c se(b_k)] = 1 - \alpha \qquad (10.5)$$

$b_k - t_c se(b_k)$ va $b_k + t_c se(b_k)$ oraliq so‘nggi nuqtalari tasodifiydir, chunki ular turlicha bo‘ladi. Bu so‘nggi nuqtalar b_k ning **intervalli baholovchisini** belgilaydi. (10.5) formuladagi ehtimollik bayonotida aytilishicha, $b_k \pm t_c se(b_k)$ oraliqda haqiqiy, lekin noma'lum b_k parametrni o‘z ichiga olishi ehtimoli 1-α bo‘ladi.

Agar (10.5) formuladagi b_k va se(b_k) ma'lumotlarning ma'lum namunasi asosida hisoblangan qiymatlar (raqamlar) bo‘lsa, $b_k \pm t_c se(b_k)$ b_k ning 100(1-α)% **intervalli taxmini** deyiladi. Ekvivalent tarzda u 100(1-α)% **ishonch oralig‘i** deb ataladi. Odatda α=0,01 yoki α =0,05, shuning uchun biz 99% ishonch oralig‘ini yoki 95% ishonch oralig‘ini olamiz.

Ishonch oraliqlarini talqin qilish katta e'tibor talab qiladi. Intervallarni baholash jarayonining xususiyatlari takroriy namuna olish tushunchasiga asoslanadi. Agar biz N o‘lchamdagi ko‘plab tasodifiy namunalarni tanlasak, har bir namuna uchun eng kichik kvadratlar bahosi b_k va uning standart xatosi se(b_k) ni hisoblab chiqsak va keyin har bir namuna uchun $b_k \pm t_c se(b_k)$ oraliq bahosini tuzsak, u holda 100(1-α)% bo‘ladi. Barcha tuzilgan intervallarning haqiqiy b_k parametrini o‘z ichiga oladi.

Bitta ma’lumotlar namunasi asosidagi har qanday bitta intervalli baho haqiqiy b_k parametrini o‘z ichiga olishi yoki bo‘lmasligi mumkin va b_k noma’lum bo‘lgani uchun biz uning bor yoki yo‘qligini hech qachon bilmaymiz. "Ishonch oraliqlari" muhokama qilinganda, bizning ishonchimiz intervalli baholashni tuzishda qo‘llaniladigan protseduraga bog‘liqligini unutmang; u ma'lumotlar namunasi bo‘yicha hisoblangan biron bir intervalli taxminda emas.

Masala. Oziq-ovqat xarajatlari ma'lumotlari uchun $N = 40$ va erkinlik darajalari $N - 2 = 38$. 95% ishonch oralig‘i uchun $\alpha = 0{,}05$. Kritik qiymat $t_c = t_{\left(1-\frac{\alpha}{2}, N-2\right)} = t_{(0.975,38)} = 2.024$. 38 erkinlik darajasi bilan t-taqsimlanishidan 97,5 foizdir. β_2 uchun (10.5) dagi ehtimollik quyidagi natija bo‘ladi:

$$P[b_2 - 2.2024se(b_2) \leq \beta_2 \leq b_2 + 2.2024se(b_2)] = 0.95 \qquad (10.6)$$

β_2 uchun intervalli taxminni yaratish uchun biz eng kichik kvadratlar bahosi $b_2 = 10.21$ va uning standart xatosidan foydalanamiz

$$se(b_2) = \sqrt{\widehat{var(b_2)}} = \sqrt{4.38} = 2.09$$

Ushbu qiymatlarni (10.6) ga almashtirib, biz β_2 uchun “95% ishonch oralig‘i taxminini” olamiz:

$$b_2 \pm t_c se(b_2) = 10.21 \pm 2.024(2.09) = [5.94, 14.45]$$

Ya'ni, biz "95% ishonch bilan" haftalik daromadning qo‘shimcha 100 dollaridan uy xo‘jaliklari oziq-ovqat uchun 5,97 dan 14,45 dollargacha sarflashini taxmin qilamiz. β_2 aslida [5.97, 14.45] oraliqdami? Biz bilmaymiz va hech qachon bilmaymiz. Bizga ma'lumki, biz qo‘llagan jarayon bir xil populyatsiyadan olingan ma'lumotlarning ko‘plab tasodifiy namunalariga qo‘llanilsa, ushbu protsedura yordamida tuzilgan barcha intervalli taxminlarning 95% haqiqiy parametrni o‘z ichiga oladi. Intervalni baholash protsedurasi vaqtning 95 foizida "ishlaydi". Bizning bitta namunamiz asosida oraliq baholash haqida nima deyishimiz mumkin, bu jarayonning ishonchliligini hisobga olgan holda, agar β_2 [5.97, 14.45] oraliqda bo‘lmasa, biz "ikkilanamiz".

β_2 oraliq bahosining foydaliligi nimada? Regressiya natijalari haqida hisobot berishda biz har doim nuqtali baho beramiz, masalan, $b_2 = 10{,}21$. Biroq, faqat ball bahosi uning ishonchliligini anglatmaydi. Shunday qilib, biz intervalli taxmin haqida ham xabar berishimiz mumkin. Intervalli hisob-kitoblar eng kichik kvadratlar usulining o‘zgaruvchanligi o‘lchovi bo‘lgan taxminiy bahoni ham, standart xatosini ham o‘z ichiga oladi. Intervalli taxmin, shuningdek, tanlama hajmi uchun ruxsatni o‘z ichiga oladi, chunki erkinlikning past darajalari uchun t-taqsimot kritik qiymati t_c kattaroqdir.

Agar intervalli taxmin keng bo‘lsa (katta standart xatoni nazarda tutsa), bu namunada β_2 haqida ko‘p ma'lumot yo‘qligini ko‘rsatadi. Agar intervalli taxmin tor bo‘lsa, bu biz β_2 haqida ko‘proq ma'lumot olganimizni ko‘rsatadi.

"Keng" va "tor" nima muammoga bog‘liq. Misol uchun, bizning modelimizda $b_2 = 10.21$ uy xo‘jaliklarining haftalik daromadi 100 dollarga oshgan taqdirda, haftalik oziq-ovqat xarajatlari qancha ko‘payishini taxmin qiladi. Supermarketlar tarmog‘ining direktori ushbu hisobdan daromadning o‘sishi prognozlarini hisobga olgan holda kelajakdagi do‘konlar sig‘imi talablarini rejalashtirish uchun foydalanishi mumkin. Biroq, hech qanday qaror faqat bitta raqamga asoslanmaydi. Ehtiyotkor bosh direktor 10.21 atrofida β_2 qiymatlarini hisobga olgan holda sezgirlik tahlilini amalga oshiradi. Savol shuki, ‘‘Qaysi qiymatlar?’’ Bitta javob intervalli baho bilan ta’minlanadi [5.97, 14.45]. β_2 bu oraliqda bo‘lishi yoki bo‘lmasligi mumkin bo‘lsa-da, bosh direktor intervalli taxminni olish uchun ishlatiladigan jarayon vaqtning 95% "ishlaydi" deb biladi.

Agar intervalda β_2 o‘zgarishi kompaniyaning sotuvi va foydasiga jiddiy oqibatlarga olib kelsa, bosh direktor qaror qabul qilish va yangi hamda kattaroq ma'lumotlar namunasini buyurtma qilish uchun dalillar yetarli emas degan xulosaga kelishi mumkin.

Intervalli baholovchilar regressiya natijalari haqida hisobot berishning qulay usuli hisoblanadi, chunki ular noma'lum parametrlar tushishi mumkin bo‘lgan qiymatlar oralig‘ini ta'minlash uchun nuqtalarni baholashni tanlamaning o‘zgaruvchanligi o‘lchovi bilan birlashtiradi. Agar eng kichik kvadratlar baholovchisining tanlama o‘zgaruvchanligi nisbatan kichik bo‘lsa, intervalli taxminlar nisbatan tor bo‘ladi, bu eng kichik kvadratlar taxminlari "ishonchli" ekanligini anglatadi. Agar eng kichik kvadratlar baholovchilari katta tanlama o‘zgaruvchanligidan aziyat cheksa, u holda intervalli taxminlar keng bo‘ladi, bu esa eng kichik kvadratlar taxminlari "ishonchsiz" ekanligini bildiradi.

Nazorat uchun savollar

1. «Takroriy tanlab olish nazariyasi»ning intervallarni ishlab chiqarish va gipotezalarni yetarli darajada qanday aloqasi borligini muhokama qiling.

2. Statistik xulosalar uchun b_1 va b_2 eng kichik kvadratlar usullari normal taqsimlangan miqdorlar uchun nima uchun muhimligini tushuntiring.

3. Intervalli baholovchining “ishonch aniq darajasini”ni va takroriy tanlab olish kontekstida nimani anglatishini tushuntiring va misol keltiring.

4. Intervalli baholovchi va intervalli baho o‘rtasidagi farqni tushuntiring. Intervalli taxminni qanday izohlashni tushuntiring.

XI BOB. GIPOTEZA TUSHUNCHASI. GIPOTEZALARNI TEKSHIRISH JARAYONI

11.1. Gipoteza testlari
11.2. Test statistikasi
11.3. Muayyan alternativlar uchun rad etish regioni

11.1. Gipoteza testlari

Koʻpgina biznes va iqtisodiy qarorlar bilan bogʻliq muammolar parametrning oʻziga xos qiymat ekanligi yoki yoʻqligi haqida mulohaza yuritishni talab qiladi. Oziq-ovqat xarajatlari misolida, β_2 10 dan katta yoki yoʻqligini hal qilish uchun juda katta farq qilishi mumkin, bu daromadning 100 dollarga oshishi oziq-ovqat xarajatlarini 10 dollardan koʻproqqa oshirishini koʻrsatadi. Shuningdek, iqtisodiy nazariyaga asoslanib, biz β_2 ijobiy boʻlishi kerak deb hisoblaymiz. Bizning ma'lumotlarimiz va modelimizni tekshirishdan biri bu nazariy taklifning ma'lumotlar tomonidan qoʻllab-quvvatlanishi yoki yoʻqligi.

Gipotezani tekshirish protseduralari bizda populyatsiya haqidagi taxminni ma'lumotlar namunasidagi ma'lumotlar bilan taqqoslaydi. Iqtisodiy va statistik modelni hisobga olgan holda, iqtisodiy xatti-harakatlar haqida farazlar shakllanadi. Keyinchalik bu farazlar model parametrlari haqidagi bayonotlar sifatida ifodalanadi. Gipoteza testlari gipoteza toʻgʻrisida xulosa chiqarish uchun ma'lumotlar namunasidagi parametr, uning eng kichik kvadratlari nuqtasi va standart xatosi haqidagi ma'lumotlardan foydalanadi.

Nol gipoteza. H_0 (H-nol) bilan belgilangan nol gipoteza regressiya parametrining qiymatini belgilaydi, bu umumiylik uchun biz b_k, k=1 yoki 2 uchun belgilaymiz. Nol gipoteza $H_0 : \beta_k = c$. Bu yerda c doimiy va muayyan regressiya modeli kontekstida muhim qiymatdir. Nol gipoteza - bu toʻgʻri emasligiga namunaviy dalillar orqali ishonch hosil qilmagunimizcha biz davom etadigan ishonchdir, bu holda biz nol gipotezani rad etamiz.

Alternativ gipoteza. Har bir nol gipoteza bilan birlashtirilgan mantiqiy alternativ gipoteza H1 bo‘lib, agar nol gipoteza rad etilsa, biz uni qabul qilamiz. Muqobil gipoteza moslashuvchan va ma'lum darajada iqtisodiy nazariyaga bog‘liq. $\mathrm{H}_0 : \beta_k = c$ nol gipotezasi uchun uchta mumkin bo‘lgan muqobil gipoteza mavjud:

- $H_1 : \beta_k > c$. $\beta_k = c$ degan nol gipotezani rad etish bizni $\beta_k > c$ degan xulosani qabul qilishga olib keladi. Tengsizlikning muqobil farazlari iqtisodiyotda keng qo‘llaniladi, chunki iqtisodiy nazariya ko‘pincha o‘zgaruvchilar o‘rtasidagi munosabatlar belgilari haqida ma'lumot beradi. Misol uchun, oziq-ovqat xarajatlari misolida biz $H_0 : \beta_2 = 0$ nol gipotezasini $\mathrm{H}_1 : \beta_2 > 0$ ga nisbatan sinab ko‘rishimiz mumkin, chunki iqtisodiy nazariya oziq-ovqat kabi ehtiyojlar oddiy tovarlar ekanligini va agar daromad ko‘tarilsa, oziq-ovqat xarajatlari oshishini qat'iy taklif qiladi.
- $H_1 : \beta_k < c$. Bu holda $\beta_k = c$ degan nol gipotezani rad etish bizni $\beta_k < c$ degan xulosani qabul qilishga olib keladi.
- $H_1 : \beta_k \neq c$. Bu holda $\beta_k = c$ degan nol gipotezani rad etish bizni β_k c dan kattaroq yoki kichikroq qiymatni oladi degan xulosani qabul qilishga olib keladi.

11.2. Test statistikasi

Nol gipoteza haqidagi namunaviy ma'lumot test statistikasining namunaviy qiymatida mujassamlangan. Sinov statistikasining qiymatiga asoslanib, biz nol gipotezani rad etishga yoki uni rad etmaslikka qaror qilamiz. Test statistikasi o‘ziga xos xususiyatga ega: uning ehtimollik taqsimoti nol gipoteza to‘g‘ri bo‘lganda to‘liq ma'lum bo‘ladi va agar nol gipoteza to‘g‘ri bo‘lmasa, u boshqa taqsimotga ega.

Hammasi (10.3) dagi asosiy natijadan boshlanadi, $t = \frac{b_k - \beta_k}{se(b_k)} \sim t_{(N-2)}$. Agar $H_0 : \beta_k = c$ nol gipotezasi to‘g‘ri bo‘lsa, u holda bk o‘rniga c ni qo‘yishimiz mumkin va bundan kelib chiqadiki

$$t = \frac{b_k - c}{se(b_k)} \sim t_{(N-2)}$$

Agar nol gipoteza to‘g‘ri bo‘lmasa, u holda (10.7) dagi t-statistik $N-2$ erkinlik darajasi bilan t-taqsimotga ega emas. Bu masala 3B-ilovada batafsil yoritilgan.

Rad etish regioni. Rad etish hududi muqobil shakliga bog‘liq. Bu nol gipotezani rad etishga olib keladigan test statistikasi qiymatlari diapazoni. Agar bizda mavjud bo‘lsa, rad etish hududini qurish mumkin.

- Nol gipoteza to‘g‘ri bo‘lganda taqsimlanishi ma'lum bo‘lgan test statistikasi
- Muqobil gipoteza
- Muhimlik darajasi

Rad etish hududi nol gipoteza to‘g‘ri bo‘lganida yuzaga kelishi ehtimoli kam bo‘lgan va past ehtimolga ega bo‘lgan qiymatlardan iborat. Mantiq zanjiri "Agar test statistikasining qiymati past ehtimollik mintaqasiga to‘g‘ri keladigan bo‘lsa, unda test statistikasi taxmin qilingan taqsimotga ega bo‘lishi ehtimoldan yiroq emas va shuning uchun nol gipoteza to‘g‘ri bo‘lishi ehtimoldan yiroq emas. " Agar muqobil gipoteza to‘g‘ri bo‘lsa, test statistikasining qiymatlari odatdan tashqari katta yoki g‘ayrioddiy kichik bo‘ladi. "Katta" va "kichik" atamalari testning ahamiyatlilik darajasi deb ataladigan α ehtimolini tanlash yo‘li bilan aniqlanadi, bu "ehtimol bo‘lmagan hodisa" uchun ma'no beradi. Testning ahamiyatlilik darajasi. α odatda 0.01, 0.05 yoki 0.10 sifatida tanlanadi.

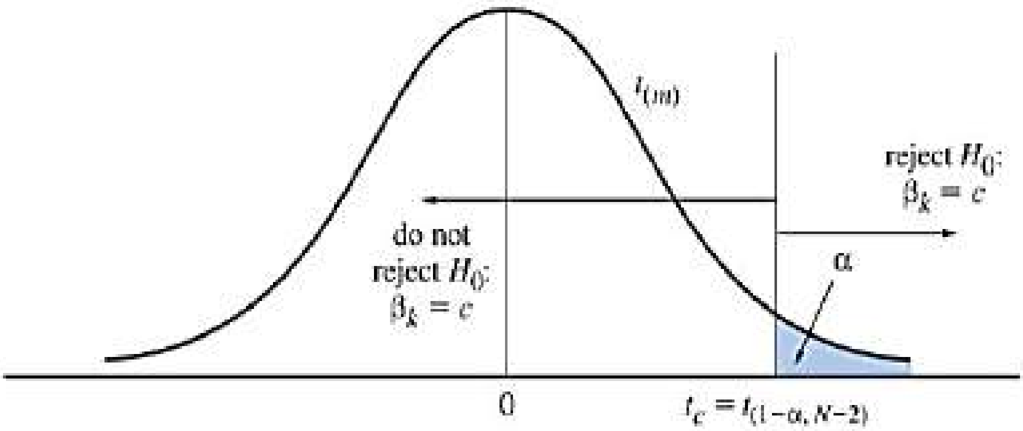
Agar biz nol gipoteza to‘g‘ri bo‘lsa, uni rad qilsak, biz I turdagi xato deb ataladigan xatoga yo‘l qo‘yamiz. Sinovning ahamiyatlilik darajasi I turdagi xatolikka yo‘l qo‘yish ehtimolidir, shuning uchun $P(Tip \text{ I } xato) = \alpha$. Har doim biz nol gipotezani rad qilsak, biz bunday xatoga yo‘l qo‘ygan bo‘lishimiz mumkin - undan qochishning iloji yo‘q. Yaxshi xabar shundaki, biz α ahamiyatlilik darajasini belgilash orqali biz bardosh beradigan I turdagi xato miqdorini belgilashimiz mumkin. Agar bunday xato qimmatga tushsa, biz kichik qilamiz. Agar biz noto‘g‘ri bo‘lgan nol gipotezani rad qilmasak, biz II turdagi xatoga yo‘l qo‘ygan bo‘lamiz. Haqiqiy vaziyatda biz ushbu turdagi xatolik ehtimolini nazorat qila olmaymiz yoki hisoblay olmaymiz, chunki u noma'lum haqiqiy parametr bk ga bog‘liq.

Xulosa. Gipotezani sinab ko‘rishni tugatgandan so‘ng, siz o‘z xulosangizni bayon qilishingiz kerak. Siz nol gipotezani rad qilasizmi yoki nol gipotezani rad qilmaysizmi? Quyida bahslashayotganimizdek, siz nol gipotezani "qabul qilaman" deb aytishdan qochishingiz kerak, bu juda noto‘g‘ri bo‘lishi mumkin. Bundan tashqari, siz ishlayotgan muammoning iqtisodiy kontekstida xulosa nimani anglatishini va topilmaning iqtisodiy ahamiyatini aytishni odatiy amaliyotga aylantirishingizni so‘raymiz. Statistik protseduralar o‘z-o‘zidan maqsad emas. Ular biron bir sababga ko‘ra amalga oshiriladi va siz tushuntirishingiz kerak bo‘lgan ma'noga ega.

11.3. Muayyan alternativlar uchun rad etish regioni

Ushbu bo‘limda biz $H_0\colon b_k = c$ nol gipotezasiga uchta mumkin bo‘lgan muqobillarning har biri uchun rad etish qoidalarining tabiati haqida juda aniq bo‘lishga umid qilamiz. Oldingi bo‘limda ta'kidlanganidek, nol gipoteza uchun rad etish hududiga ega bo‘lish uchun bizda mavjud bo‘lgan test statistikasi kerak, u (10.7) da keltirilgan. Ikkinchidan, bizga ma'lum bir muqobil kerak, b_k > c, b_k < c yoki $b_k \neq c$. Uchinchidan, testning ahamiyatlilik darajasini belgilashimiz kerak. Sinovning ahamiyatlilik darajasi, α - bu nol gipoteza haqiqatda to‘g‘ri bo‘lsa, uni rad etish ehtimoli, bu I turdagi xato deb ataladi.

“Katta” (>) alternativ bilan bir dumli sinovlar. $H_0\colon \beta_k = c$ nol gipotezasini tekshirishda, agar $H_1\colon \beta_k > c$ muqobil gipoteza to‘g‘ri bo‘lsa, u holda t-statistikaning qiymati (10.7) t-taqsimot uchun odatdagidan kattaroq bo‘lishga intiladi. Sinov statistikasi ahamiyatlilik darajasi uchun kritik qiymatdan kattaroq bo‘lsa, biz nol gipotezani rad qilamiz. α ehtimolini o‘ngda qoldiradigan kritik qiymat dum $(1-\alpha)$-foiz $t_{(1-\alpha,N-2)}$, 11.2-rasmda ko‘rsatilganidek. Masalan, agar $\alpha = 0.05$ va $N-2 = 30$ bo‘lsa, 2-jadvaldagi kritik qiymat 95-persentil qiymati $t_{(0.95,30)} = 1.697$ hisoblanadi.



11.1-rasm. $H_1: \beta_k > c$ ga nisbatan $H_0: \beta_k = c$ bir dumli sinov uchun rad etish regioni

Rad etish qoidasi. $H_0: \beta_k = c$ nol gipotezasini $H_1: \beta_k > c$ muqobil gipotezasiga qarshi tekshirishda nol gipotezani rad eting va muqobil gipotezani qabul qiling, agar $t \geq t_{(1-\alpha,N-2)}$ bo‘lganda.

Sinov "bir dumli" test deb ataladi, chunki t-statistikaning mumkin bo‘lmagan qiymatlari ehtimollik taqsimotining faqat bitta dumiga tushadi. Agar nol gipoteza to‘g‘ri bo‘lsa, test statistikasi (10.7) t-taqsimotga ega va uning qiymati taqsimot markaziga, ehtimollikning katta qismi joylashgan kritik qiymatning chap tomoniga tushadi. Muhimlik darajasi α shunday tanlanadiki, agar nol gipoteza to‘g‘ri bo‘lsa, u holda t-statistik qiymat taqsimotning o‘ta o‘ng dumiga tushish ehtimoli kichik bo‘ladi; tasodifan sodir bo‘lishi ehtimoldan yiroq bo‘lmagan hodisa.

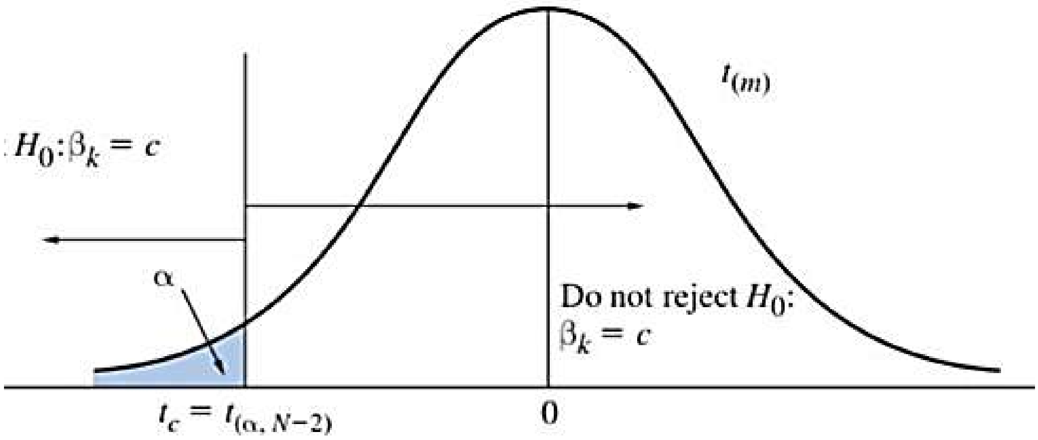
Agar biz rad etish hududida test statistik qiymatini olsak, biz uni nol gipotezaga qarshi dalil sifatida qabul qilamiz, bu bizni nol gipoteza haqiqat bo‘lishi ehtimoldan yiroq emas degan xulosaga olib keladi. Nol gipotezaga qarshi dalillar muqobil gipotezani tasdiqlovchi dalildir. Shunday qilib, agar biz nol gipotezani rad qilsak, alternativ to‘g‘ri degan xulosaga kelamiz. Agar nol gipoteza $H_0 : \beta_k = c$ to‘g‘ri bo‘lsa, test statistikasi (10.7) t-taqsimotga ega va uning qiymatlari $1 - \alpha$ ehtimollik bilan rad etilmagan mintaqaga to‘g‘ri keladi. Agar $t_{(1-\alpha, N-2)}$ bo‘lsa, unda nol gipotezaga qarshi statistik jihatdan ahamiyatli dalillar yo‘q va biz uni rad etmaymiz.

“Kamroq” (<) alternativ bilan bir dumli sinovlar. Agar $\boldsymbol{H_1 : \beta_k < c}$ muqobil gipoteza to‘g‘ri bo‘lsa, u holda t-statistikaning qiymati (10.7) t-taqsimot uchun odatdagidan kichikroq bo‘lishga intiladi. Agar test statistikasi ahamiyatlilik darajasi uchun kritik qiymatdan kichikroq bo‘lsa, biz nol gipotezani rad qilamiz. Chap dumda α ehtimolini qoldiradigan kritik qiymat α-foiz $\boldsymbol{t_{(\alpha, N-2)}}$, 11.2-rasmda ko‘rsatilganidek.

Kritik qiymatlarni aniqlash uchun t-styudent jadvalidan foydalanganda esda tutingki, t-taqsimot nolga yaqin simmetrikdir, shuning uchun α-protsentil $t_{(\alpha, N-2)}$ $(1 - \alpha)$-protsentil $t_{(1-\alpha, N-2)}$. Misol uchun, agar $\alpha = 0.05$ va $N - 2 = 20$ bo‘lsa, 2-jadvaldan t-taqsimotining 95-persentili $t_{(0.95,20)} = 1.725$ va 5-persentil qiymati $t_{(0.05,20)} = -1.725$ ga teng.

Rad etish qoidasi: $H_0 : \beta_k = c$ nol gipotezasini $H_1 : \beta_k < c$ muqobil gipotezasiga qarshi tekshirishda, agar $t \leq t_{(\alpha, N-2)}$ bo‘lsa nol gipotezani rad eting va muqobil gipotezani qabul qiling.

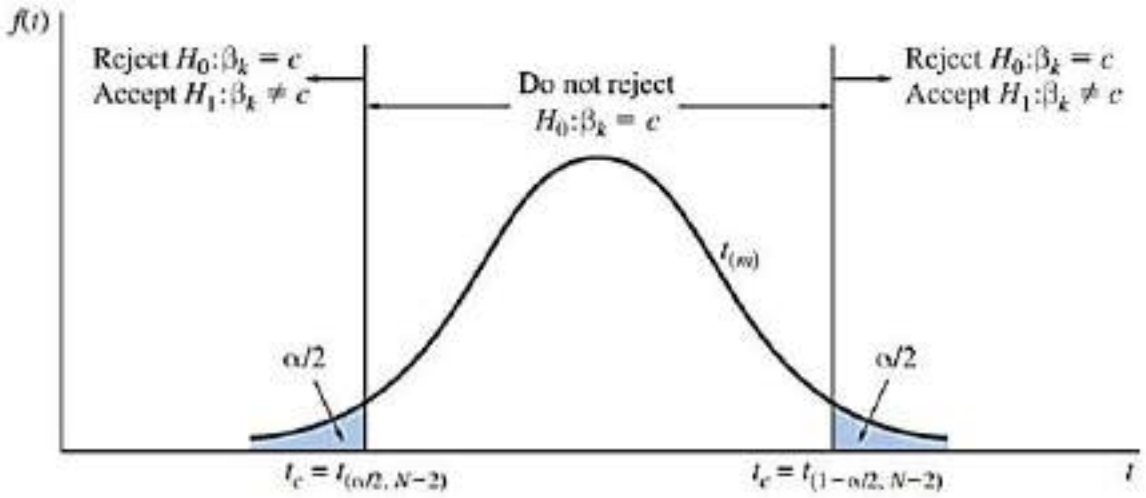
Rad etmaydigan t-statistik hudud $t_{(\alpha, N-2)}$ qiymatidan kattaroq. Agarda nol gipoteza haqiqat bo‘lsa, bunday t-qiymatini olish ehtimoli $1 - \alpha$ga teng bo‘lsa, tanlangan qiymat katta bo‘ladi. Shunday qilib, agar $t > t_{(\alpha, N-2)}$ bo‘lsa, keyin $H_0 : \beta_k = c$ ni rad etib bo‘lmaydi. Rad etish hududi qayerda joylashganligini eslab qolish quyidagi usul bilan osonlashtirilishi mumkin: Bir dumli sinov uchun rad etish hududi yo‘nalishi bo‘yicha muqobil usul. Agar muqobillik (>) katta bo‘lsa, keyin o‘ng dumda rad eting, agar kichik bo‘lsa, chap dumdagi hodisani rad eting.



11.2-rasm. $\boldsymbol{H_1: \beta_k < c}$ ga nisbatan $\boldsymbol{H_0: \beta_k = c}$ bir dumli sinov uchun rad etish hududi ko‘rsatilgan

Tengsizlik (≠) alternativi bilan ikki dumli testlar. $\boldsymbol{H_0: \beta_k = c}$ nol gipotezasini sinab ko‘rishda, agar $H_1: \beta_k \neq c$ muqobil gipoteza to‘g‘ri bo‘lsa, u holda t-statistikaning qiymati (10.7) t-taqsimoti uchun odatdagidan kattaroq yoki kichikroq bo‘lishga intiladi. Muhimlik darajasi α bo‘lgan testga ega bo‘lish uchun biz kritik qiymatlarni aniqlaymiz, shunda t-statistikaning ikkala dumga tushish ehtimoli $\frac{\alpha}{2}$ bo‘ladi.

Chap-dum kritik qiymat $\boldsymbol{t \leq t_{\left(\frac{\alpha}{2}, N-2\right)}}$ va o‘ng-dum kritik qiymati foiz $\boldsymbol{t \geq t_{\left(1-\frac{\alpha}{2}, N-2\right)}}$. $\boldsymbol{H_0: \beta_k = c}$ degan nol gipotezani rad qilamiz, agar test statistikasi $\boldsymbol{t \leq t_{\left(\frac{\alpha}{2}, N-2\right)}}$ yoki $\boldsymbol{t \geq t_{\left(1-\frac{\alpha}{2}, N-2\right)}}$ bo‘lsa, $\boldsymbol{H_1: \beta_k \neq c}$ bo‘lgan muqobil foydasiga, 10.4-rasmda ko‘rsatilganidek. Masalan, $\boldsymbol{\alpha = 0.05}$ va $\boldsymbol{N - 2 = 30}$ bo‘lsa, $\boldsymbol{\frac{\alpha}{2} = 0.025}$ va chap tomondagi kritik qiymat **2,5** foizli $\boldsymbol{t_{(0.025,30)} = -2.042;}$ o‘ng-dum kritik qiymati **97,5** foizli $\boldsymbol{t_{(0975,30)} = 2.042.}$ O‘ng-dum kritik qiymati 2-jadvalda, chap-dum kritik qiymati esa t-taqsimot simmetriyasi yordamida topiladi.



11.3-rasm. $\boldsymbol{H_1: \beta_k \neq c}$ **ga nisbatan** $\boldsymbol{H_0: \beta_k = c}$ **sinovi uchun rad etish regioni**

Rad etish hududi chap va o‘ng dumdagi t-taqsimot qismlaridan iborat bo‘lganligi sababli, bu test ikki dumli test deb ataladi. Nol gipoteza to‘g‘ri bo‘lsa, test statistikasining ikkala dum sohasiga to‘g‘ri keladigan qiymatini olish ehtimoli "kichik" bo‘ladi. Dum ehtimollarining yig‘indisi α. Dum sohalarida joylashgan test statistikasining namunaviy qiymatlari nol gipotezaga mos kelmaydi va nol gipotezaning to‘g‘riligiga dalildir. Boshqa tomondan, agar $H_0: \beta_k = c$ nol gipotezasi to‘g‘ri bo‘lsa, markaziy rad etilmagan mintaqada test statistikasi t-qiymatini olish ehtimoli yuqori bo‘ladi. Markaziy rad etmaslik sohasidagi test statistikasining namunaviy qiymatlari nol gipotezaga mos keladi va nol gipotezaning to‘g‘riligiga dalil sifatida qabul qilinmaydi. Shunday qilib, rad etish qoidasi

$H_0: \beta_k = c$ nol gipotezani $H_1: \beta_k \neq c$ muqobil gipotezaga qarshi tekshirishda, nol gipotezani rad etish, va alternativ $t \leq t_{\left(\frac{\alpha}{2}, N-2\right)}$ yoki $t \geq t_{\left(1-\frac{\alpha}{2}, N-2\right)}$ gipotezani qabul qilish mumkin.

Agar $t_{\left(\frac{\alpha}{2}, N-2\right)} < t < t_{\left(1-\frac{\alpha}{2}, N-2\right)}$ bo‘lsa, biz nol gipotezani rad etmaymiz.

Nazorat uchun savollar

1. Nol gipoteza, muqobil gipoteza va rad etish misoli atamalarini va rad etishining chizmasini keltirib chiqaradi.

2. Statistik test mantiqini tushuntiring, shu narsaning, agar nol gipoteza to‘g‘ri bo‘lsa, test statistikasi ma'lum ehtimollik taqsimotiga ega bo‘lishi nima uchun muhimligini tushuntiring.

3. P-qiymatidan qanday foydalanishni tushuntiring p-qiymatini ko‘rsatgan eskizni taqdim eting.

4. Bir dumli va ikki dumli testlarning farqini tushuntiring. Bir dumli test uchun rad etish hududini qanday aniqlashni intuitiv tarzda tushuntiring.

5. I turdagi xatoni tushuntiring va uni eskizda tasvirlang. Sinovning ahamiyatliligini aniqlang.

XII BOB. KO‘P OMILLI REGRESSIYA MODELI

12.1. Qoldirilgan o‘zgaruvchilar natijasida siljish

12.2. Ko‘p omilli regressiya modellarida parametrlarni hisoblash

12.3. Binar o‘zgaruvchilar

12.4. Multikollinearlik

12.1. Qoldirilgan o‘zgaruvchilar natijasida siljish

Avvalgi boblarda oddiy regressiya modeli (bir o‘zgaruvchi - erksiz o‘zgaruvchi, regressiyaning boshqa bir o‘zgaruvchi - erkli o‘zgaruvchi, regressorga bog‘liqligi)ni ko‘rib chiqdik. Oddiy regressiya modeli bir qator holatlarda foydali bo‘lsa-da, iqtisodiy jarayonlardagi o‘zgarishlar bir nechta omillar o‘zgarishi bilan bog‘liq bo‘ladi. Masalan, muayyan tovar yoki xizmatga bo‘ladigan talab miqdori nafaqat uning narxiga, balki iste'molchi daromadi, bozordagi o‘rnini bosuvchi va to‘ldiruvchi tovarlar narxlariga ham bog‘liq bo‘ladi. Erksiz o‘zgaruvchini bu kabi bir nechta omillar bilan bog‘liqligi ekonometrik modellashtirilsa, ko‘p omilli regressiya modeli hosil bo‘ladi:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

Bu regressiyada y x_1, x_2, … x_k regressorlar orqali tushuntiriladi va

β_0: ozod had

β_1: regressor x_1 bilan bog‘liq parametr

β_2: regressor x_2 bilan bog‘liq parametr va hokazo.

1-misol. Umumiy iqtisodiyot, xususan, marketing nazariyasiga ko‘ra, firma sotuvlari hajmi (sales) sotilayotgan Tovar narxi (price), reklama xarajatlari (ads) va boshqa omillarga bog‘liq. Bunday bog‘liqlikni quyidagicha ekonometrik modellashtirish mumkin:

$$sales = \beta_0 + \beta_1 price + \beta_2 ads + \cdots + u$$

Bunda β_1 tovar narxini sotuvlar hajmiga ta'sirini o'lchaydi, β_2 esa reklama xarajatlarning sotuvlar hajmiga xarajat qilingan davrdagi ta'sirini o'lchaydi[11].

2-misol. Mehnat bozori tadqiqotchilari kadrlar erishgan ta'lim darajasi (educ) qay darajada ularning ish haqlari (wage) ga ta'sir qilishini keng o'rganadilar. Ish haqlari ta'limdan tashqari ishchining malakasi (exper), ichki qobilyati (abil) va boshqa omillarga bog'liq. Bunday bog'liqlikni quyidagi model orqali ifodalaymiz:

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 abil + \cdots + u$$

Parametrlar tegishli regressorning ish haqlariga ceteris paribus samarasini o'lchasa-da, tadqiqotchi sifatida bizga faqatgina ta'limning ta'siri qiziqarli bo'lsin. Shu bilan birga, boshqa o'zgaruvchilar kontrol o'zgaruvchilar sifatida modelga kiritiladi; aks holda, tadqiq etilayotgan β_1 parametrining hisoblangan qiymati haqiqiysiga nisbatan siljib qolish xavfi mavjud (keyingi bobda batafsilroq muhokama qilingan!) va iqtisodiy jihatdan qilingan talqinning noto'g'ri izohi shakllanishi mumkin.

Gauss-Markov teoremasining shartlariga muvofiq tasodifiy xatolikning shartli o'rtacha qiymati nolga teng:

$$E(u|x_1, x_2, \dots x_k) = 0$$

Bu tasodifiy xatolik[12] hech bir regressorga bog'liq bo'lmasligini taqozo etadi. Oddiy regressiya modellarida bu shart aksariyat hollarda qoniqtirilmaydi. Misol uchun $wage = \beta_0 + \beta_1 educ + u$ modelida ishchilar tomonidan olinadigan ish haqlari nafaqat ta'limga, balki ularning ichki qobiliyati (abil) ga ham bog'liq ekanligi umumiy mantiq yoki iqtisodiy nazariyadan ma'lum.

Shu nuqtai nazardan $wage = \beta_0 + \beta_1 educ + \beta_2 abil + u$ modeli reallikni yaxshiroq ifodalagani tufayli shu modelni haqiqiy model deb tasavvur qiling. Ichki qobiliyat o'zgaruvchisini miqdoriy ifodalash

---

[11] Regressor tomonidan erksiz o`zgaruvchiga ko`rsatiladigan ta`sirni talqin etishda boshqa omillar o`zgarmas, ya`ni ceteris paribus farazi har doim hisobga olinishi kerak.

[12] Oddiy regressiya uchun bu E(u|x) =0 shaklini oladi.

mushkul masala **20** yoki bu o‘zgaruvchi bo‘yicha ma'lumot yig‘ilmagani sababli $wage = \tilde{\beta}_0 + \tilde{\beta}_1 educ + u$ modelini, ya'ni ichki qobiliyat (abil) o‘zgaruvchisini tushirib hisoblasak, agar ishchining ichki qobiliyati uning ta'lim darajasini belgilasa (aslida belgilaydi ham!), ya'ni abil va educ umumiy mantiq yoki iqtisodiy nazariyaga ko‘ra, bog‘liqligi o‘rnatilsa, $E(u|x) \neq 0$ bo‘ladi va natijada hisoblangan ekanligi kelib chiqadi. Bunda hisoblangan parametrlar qaror qabul qilish uchun foydali bo‘lmaydi.

Buni matematik jihatdan izohlash maqsadida quyidagi holatni ko‘rib chiqaylik. Haqiqiy model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$ bo‘lsin. Shu bilan birga, tanlamada y va x_t o‘zgaruvchilari bo‘yicha ma'lumot yig‘ilgan, x_2 bo‘yicha esa ma'lumot yo‘q yoki u raqamlarda o‘lchash qiyin o‘zgaruvchi (ichki qobiliyat kabi) bo‘lsin. Bunda model $y = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + u$ bo‘yicha hisoblangan parametrlar haqiqiy parametrdan farqli bo‘lib qolib, o‘rta hisobda siljib qoladi.

$$\tilde{\beta}_1 = \frac{\sum(x_1 - \tilde{x}_1)y}{\sum(x_1 - \tilde{x}_1)^2} = \frac{\sum(x_1 - \tilde{x}_1)(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + u)}{\sum(x_1 - \tilde{x}_1)^2}$$
$$= \frac{\sum(x_1 - \tilde{x}_1)}{\sum(x_1 - \tilde{x}_1)^2}\beta_0 + \frac{\sum(x_1 - \tilde{x}_1)x_1}{\sum(x_1 - \tilde{x}_1)^2}\beta_1 + \frac{\sum(x_1 - \tilde{x}_1)x_2}{\sum(x_1 - \tilde{x}_1)^2}\beta_2 + \frac{\sum(x_1 - \tilde{x}_1)u}{\sum(x_1 - \tilde{x}_1)^2}$$

Bunda:

$$\sum(x_1 - \tilde{x}_1) = \sum(x_1) - \sum(\tilde{x}_1) = nx_1 - n\tilde{x}_1 = 0$$
$$\sum(x_1 - \tilde{x}_1)^2 = \sum(x_1 - \tilde{x}_1)(x_1 - \tilde{x}_1)$$
$$= \sum(x_1 - \tilde{x}_1)x_1 - \sum(x_1 - \tilde{x}_1)\tilde{x}_1$$
$$= \sum(x_1 - \tilde{x}_1)x_1 - \tilde{x}_1\sum(x_1 - \tilde{x}_1)$$
$$= \sum(x_1 - \tilde{x}_1)x_1$$

hamda $\sum(x_1 - \tilde{x}_1)u = E(xu) = 0$ (Gauss-Markov shartlaridan biri) ekanligini hisobga olsak,

$$E(\tilde{\beta}_1) = \tilde{\beta}_1 + \frac{\sum(x_1 - \tilde{x}_1)x_2}{\sum(x_1 - \tilde{x}_1)^2}\beta_2$$

Ushbu tenglikda $\frac{\sum(x_1-\tilde{x}_1)x_2}{\sum(x_1-\tilde{x}_1)^2}\beta_2$ qismi $\tilde{\beta}_1$ ning o‘z haqiqiy qiymati β_1 ga nisbatan siljishini ko‘rsatadi. Ushbu siljish ikki holatda 0 ga teng: (1) agar $\beta_2 = 0$ bo‘lsa, ya’ni x_2 ning ta’siri yo‘q yoki u regressiyaning omili emas va (2) $\sum(x_1 - \tilde{x}_1)x_2 = cov(x_1, x_2) = 0$ bo‘lsa, ya’ni x_1 hamda x_2 bir-biriga umuman bog‘liq bo‘lmasa. Aksariyat hollarda bir omil ikkinchisiga bog‘liq bo‘lgani bois, qoldirilgan o‘zgaruvchi(lar) natijasida siljish hosil bo‘ladi.

12.2. Ko‘p omilli regressiya modellarida parametrlarni hisoblash

Parametrlarni hisoblash mexanikasi bir nechta amalni amalga oshirishni taqozo etadi. Muayyan tanlanma k omildan iborat regressiya tenglamasi quyidagicha berilgan bo‘lsin:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + u_i$$

Oddiy regressiya bobida muhokama qilinganidek, EKK usuli yordamida parametrlarning nuqtaviy bahosini topib olamiz. Bunda yuqoridagi tenglamadan har bir tasodifiy xatolik $u_i = y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_k x_{ik}$ ekanligini nazarda tutib,

$$\sum_{i=1}^{n} \hat{u}_i^{\,2} = \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_k x_{ik}\right)$$

Har bir parametrga nisbatan birinchi tartibdagi shartlar (hosilalar) quyidagi natijani beradi:

$$\frac{d\hat{u}_i^{\,2}}{d\hat{\beta}_0} = -2\sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_k x_{ik}\right) = 0$$

$$\frac{d\hat{u}_i^{\,2}}{d\hat{\beta}_1} = -2\sum_{i=1}^{n} x_{i1}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_k x_{ik}\right) = 0$$

$$\frac{d\hat{u}_i^{\,2}}{d\hat{\beta}_2} = -2\sum_{i=1}^{n} x_{i2}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_k x_{ik}\right) = 0$$

….

$$\frac{d\hat{u}_i^{\,2}}{d\hat{\beta}_k} = -2\sum_{i=1}^{n} x_{ik}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \cdots - \hat{\beta}_k x_{ik}\right) = 0$$

Bugungi kunda ushbu k+1 tenglamalar tizimidan $\hat{\beta}_{0,}$ $\hat{\beta}_{1,}$ $\widehat{\beta}_{2,}$ ... $\hat{\beta}_{k,}$ parametrlarini kompyuter yordamida hisoblanadi.

Shu bilan birga, uning mohiyatini anglash maqsadida quyidagi misolni keltirib o‘tamiz. $sales = \beta_0 + \beta_1 price + \beta_2 ads + u$ modelida qo‘yiladigan bir qator qiziqarli masalalardan narxni tushirilganda, sotuvlar hajmi qanchaga o‘zgaradi yoki reklamaga sarflangan mablag‘lar samaradorligi qanday bo‘ldi kabi savollarga javob olishimiz uchun bu model parametrlari hisoblanishi kerak. Berilgan ma'lumot asosida EViews dasturi yordamida hisoblangan parametrlar hisoblangan model $\widehat{sales} = \hat{\beta}_0 + \hat{\beta}_1 price + \hat{\beta}_2 ads$ quyidagi jadvalda keltirilgan. Hisoblangan parametrlarni o‘rniga qo‘ysak, $\widehat{sales} = 79.28 - 2.64 \cdot price + 1.24 \cdot ads$ ko‘rinishiga ega bo‘ldi.

Narx omili bilan bog‘liq koeffitsiyent -2.64 ga teng. Yuqori narxlar kam sotuv hajmi (yoki past narxlar yuqori sotuv hajmi) bilan bog‘liq bo‘lgani uchun kutilganidek manfiy. Firmalar reklamaga sotuv hajmini ko‘paytirish maqsadida sarf qilishlarini hisobga olsak, musbat koeffitsiyent 1.24 ham kutilganidek. Koeffisiyentning kutilgan ishoralarga ega bo‘lishi hisoblanadigan parametrga qo‘yiladigan eng zaruriy shartlardan hisoblanadi, chunki aks holda boshqa zaruriy shartlar haqida fikr yuritaolmaymiz. Bundan tashqari, standart xatoliklar yetarli past bo‘lgani uchun yuqori t-statiska va 1 % dan kichik bo‘lgan p-qiymatlar hosil bo‘ldi.

Dependent Variable: SALES
Method: Least Squares
Sample: 1 75
Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRICE	-2.635952	0.365331	-7.215243	0.0000
ADS	1.241723	0.455464	2.726283	0.0080
C	79.27575	4.234425	18.72173	0.0000

R-squared	0.448258	Mean dependent var	51.58311
Adjusted R-squared	0.432932	S.D. dependent var	4.325691
S.E. of regression	3.257416	Akaike info criterion	5.238923
Sum squared resid	763.9745	Schwarz criterion	5.331623
Log likelihood	-193.4596	Hannan-Quinn criter.	5.275937
F-statistic	29.24787	Durbin-Watson stat	2.183037
Prob (F-statistic)	0.000000		

12.1-rasm. $sales = \beta_0 + \beta_1 price + \beta_2 ads + u$ **modeli parametrlarini hisoblash**

$R^2 = 0.448$, ya'ni Sotuv hajmi variatsiyasining 44.8% narx va reklama xarajatlardagi o‘zgarishlarni tushuntirmoqda. Garchi yosh iqtisodchilar R^2 qiymatiga katta urg‘u bersalar-da, bu juda yuqori yoki juda past deyishga asosimiz yo‘q.

12.3. Binar o‘zgaruvchilar

Real hayotda binar o‘zgaruvchilar ko‘p uchraydi va ular faqatgina ikkita qiymat qabul qiladi: 0 yoki 1. Masalan, odam jinsi (erkak yoki ayol), obyekt joylashuvi (shahar yoki qishloq), bugungi ob-havo (yomg‘ir yog‘adi yoki yog‘maydi), talabaning imtihondan muvaffa-qiyatli o‘tishi (yetarli baland yoki yetarli baland bo‘lmagan ball to‘plashi) va hokazo.

Binar o‘zgaruvchilar ko‘p omilli regressiya doirasida regressor sifatida qo‘llaniladi. Uy-joy bozoridan quyidagi misolni quraylik. Toshkent shahrida uy narxlarini belgilovchi yashash joyi yuzasi, xonalar soni, etaji kabi omillardan eng asosiysi uning joylashgan joyi (markaz yoki chetda) hisoblanadi.

$$price = \beta_0 + \beta_1 D + \beta_2 area + \beta_3 rooms + u$$

modelni quraylik. Bunda, price - uy narxi, area - yashash joyi yuzasi, rooms - xonalar soni hamda D - uyni markaz yoki chetda joylashgan ekanligini ko‘rsatuvchi binar o‘zgaruvchini belgilasin, ya'ni D=1 (agar uy markazda joylashgan bo‘lsa) yoki 0 (boshqa holatda)

Bunda agar D=1 bo‘lsa,

$$price = (\beta_0 + \beta_1) + \beta_2 area + \beta_3 rooms + u$$

Agar D=0 bo‘lsa,

$$price = \beta_0 + \beta_2 area + \beta_3 rooms + u$$

Shu ikki regressiya tenglamasi faqatgina β_1 miqdori bilan farqlanadi. Bir xil yashash joyi yuzasi va xonalar soniga ega bo‘lgan uylar qurilganda, markazda joylashgan uy narxi chetda joylashgan uy narxidan o‘rta hisobda β_1 shartli birlik (so‘m, masalan) balandroq bo‘ladi. Shu nuqtai nazardan berilgan tanlama uchun $\beta_1 > 0$ bo‘lishi kutiladi.

12.4. Multikollinearlik

Multikollinearlik muammosi bir-biriga chiziqli bog‘langan regressorlar regressiya modelida ishlatilganda namoyon bo‘ladi.

Misol uchun inson jinsi bo‘yicha binar o‘zgaruvchini ko‘raylik. U ikkita qiymat qabul qiladi: erkak (male) yoki ayol (female). Oldin ko‘rilgan ish haqi regressiyasini jinslar bo‘yicha ko‘rish maqsadida $wage = \beta_0 + \beta_1 educ + \beta_2 abil + u$ modelini $wage = \beta_0 + \beta_1 male + \beta_2 female + \beta_3 educ + \beta_4 abil + u$ deb o‘zgartirsak, modelda multikollinearlik muammosi paydo bo‘ladi, chunki female va male o‘zgaruvchilari bir-biriga chiziqli funksional bog‘langan, ya'ni male = 1-female (yoki female = 1-male).

Boshqa misol, Keyns iste'mol funksiyasini olaylik. Unda aholi (oziq-ovqat iste'moli (CONS) daromad (INC) ga bog‘liq, ko‘riladi: $CONS = \beta_0 + \beta_1 INC + u$.

Agar ushbu modelda biz daromadlarni, ya'ni milliy valyuta so‘mni (INC_{uzs}) boshqa valyuta (INC_{USD}) dagi qiymatini qo‘shib, ya'ni $CONS = \beta_0 + \beta_1 INC_{uzs} + \beta_1 INC_{uzs} + u$ modelini hisoblasak, multikollinearlik muammosi vujudga keladi, chunki $INC_{uzs} = \varepsilon * INC_{uzs}$, $\varepsilon - USD/UZS$ valyuta kursi.

Multkollinearlik muammosini yanada chuqurroq tushunish maqsadida quyidagi hisoblangan regressiya modelini ko‘raylik.

Dependent Variable: Y
Method: Least Squares
Sample: 1 100
Included observations: 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PRICE	-2.635952	0.365331	-7.215243	0.0000
ADS	1.241723	0.455464	2.726283	0.0080
C	79.27575	4.234425	18.72173	0.0000

R-squared	0.062564	Mean dependent var	31.74902
Adjusted R-squared	0.043236	S.D. dependent var	7.937254
S.E. of regression	7.763773	Akaike info criterion	6.966355
Sum squared resid	5846.788	Schwarz criterion	7.044510
Log likelihood	-345.3177	Hannan-Quinn criter	6.997985
F-statistic	3.236872	Durbin-Watson stat	0.156677
Prob(F-statistic)	0.043568		

12.2-rasm. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$ modelida multikollinearlik alomatlari

E'tibor bering, ushbu regressiyada x_1 va x_2 regressorlar bilan bog‘liq t-statistika statistik muhim emas, lekin umumiy F-statistika statistik muhim. x_1 va x_2 regressorlar o‘rtasidagi korrelyatsiyani tekshirsak, corr (x_1, x_2) =0.95, ya'ni juda ham yuqori ekanligi ma'lum bo‘ladi. Yuqori korrelyatsiyaga ega bo‘lgan regressorlar modelda ishlatilganida shu kabi holat yuzaga keladi. Ko‘p omilli modellarda ixtiyoriy parametr dispersiyasi quyidagi formula orqali hisoblanadi:

$$var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$$

Bunda: $j = \{1, 2, \dots. k\}$, $SST_j = \sum(x_{ij} - \tilde{x}_j)^2$ ya’ni x_j ning tanlamadagi variatsiyasi, R_j^2 esa x_j regressor sifatida ishlatib, boshqa regressorlarga qilingan regressiyaning R^2.

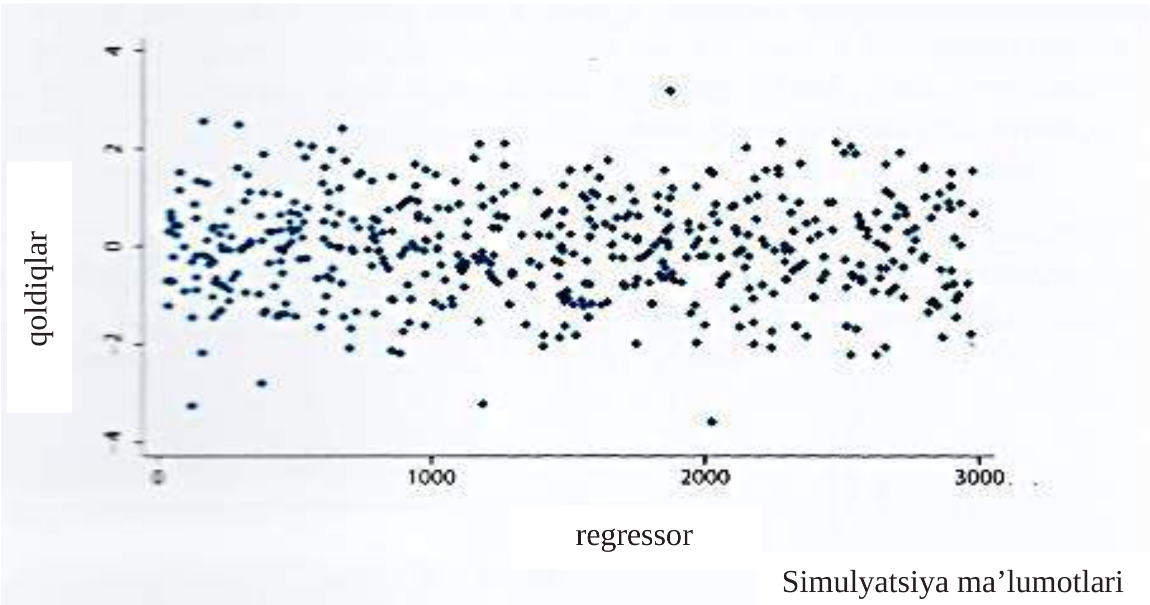
Agar corr (x_1, x_2) =1 bo‘lsa, ya'ni x_1 va x_2 chiziqli bog‘langan bo‘lsa, $R_j^2 = 1$ bo‘ladi va dispersiyani hisoblash imkoni bo‘lmaydi. Agar R_j^2 juda yuqori bo‘lsa, dispersiyani hisoblash mumkin bo‘lsa - da, dispersiyani hisoblashda ishlatiladigan maxraj kamayadi va bu dispersiyani oshirishga xizmat qiladi.

12.5. Xeteroskedarslik

Gauss-Markov teoremasi shartlaridan biri – bu tasodifiy xatolik o‘zgarmas dispersiyaga ega (xomoskedastik) bo‘lishi kerak, ya’ni $var(u_i) = \sigma^2$. Bunda EKK usuli yordamida hisoblangan parametrlarning standart xatoliklari eng kichik bo‘lishi ta’minlanadi. Muayyan modelda bu shart qoniqtirilmasa, xetroskedastiklik muammosi yuzaga keladi va $var(u_i) = \sigma_i^2$ yoki $var(u_i) \neq \sigma^2$ holat yuzaga keladi.

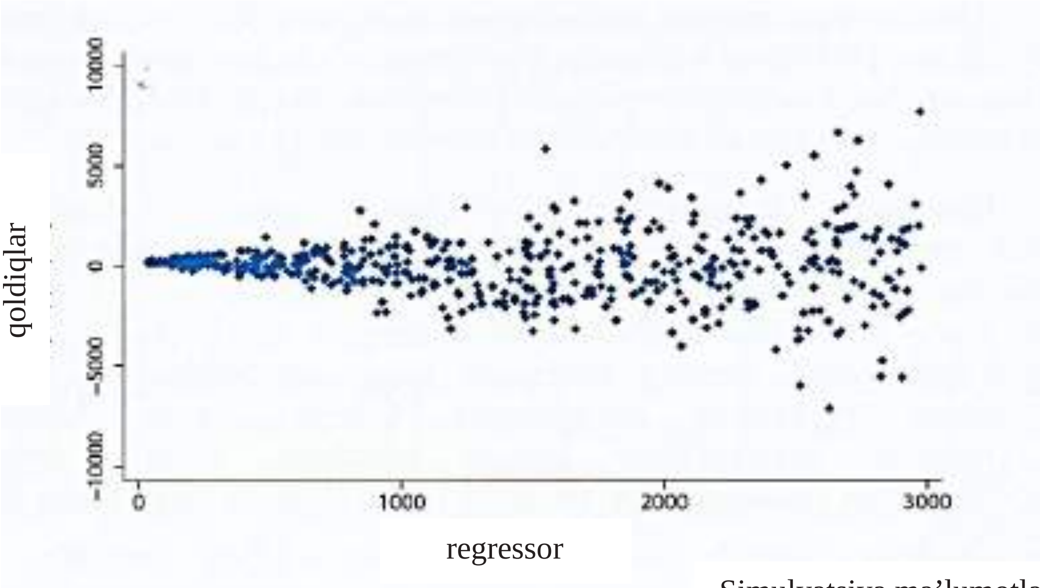
Xetroskedastiklik mammosini aniqlash. Xetroskedastiklikni tanlanmadan EKK usuli yordamida hisoblangan qoldiqlar grafigi asosida ko‘rish mumkin. Bunday usul rasmiy hisoblanmasada, xeteroskedarslik muammosi kuchli bo‘lgan holatlarni aniqlashda qo‘l keladi.

Quyidagi grafiklarga e’tibor bering. Abssissa (x^{-}) o‘qida regressor yoki ko‘p omili regressiya modeliga kiritilgan regressorlarning yoki chiziqli kombinatsiyasi, ordinata (y^{-}) o‘qida esa, EKK usuli yordamida hisoblangan modeldan qoldiqlar ko‘rsatilgan. Barcha qoldiqlar normal taqsimot qonuniga bo‘ysunadi. 12.3-rasmda xatoliklar regressor qiymatlari bo‘yicha tekis taqsimlangan va ularning dispersеyasi deyarli o‘zgarmas. Bunday xatoliklar xomoskedastik deyiladi va tegishli Gauss-Markov teoremasini shartini bajaradi.



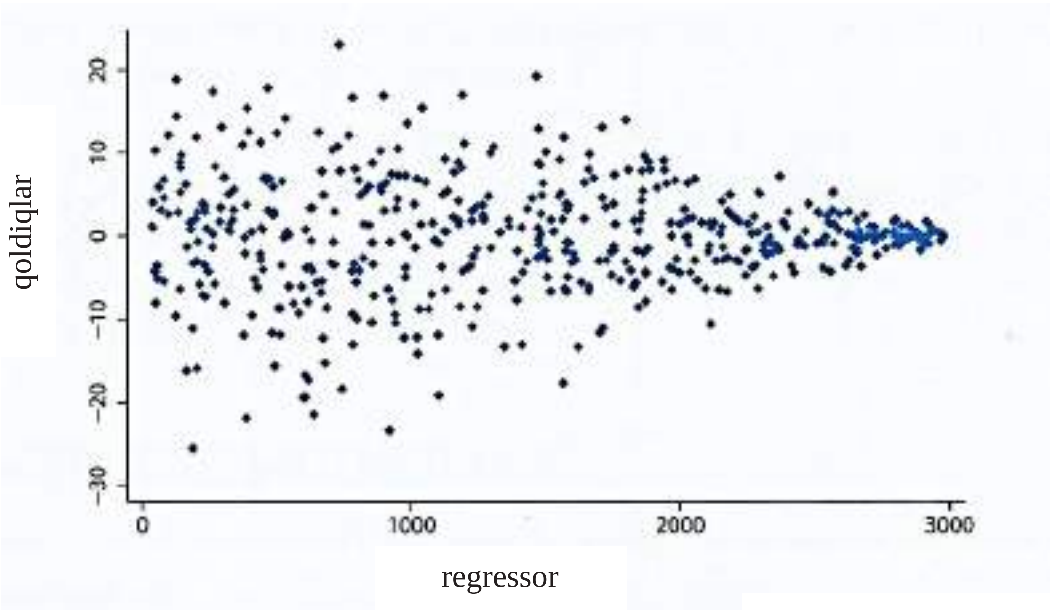
12.3-rasm. Xeteroskedastik xatoliklar (Berilgan regressor qiymatlari uchun qoldiqlar dispersiyasi o‘zgarmas)

Grafik usul foydali bo‘lsa-da, xetroskedastiklik muammosini aniqlashda rasmiy testlardan keng foydalaniladi. Shulardan birinchisi, Broysh-Pagan (BP) testidir. Unda xeteroskedastiklikni chiziqli formalari, agar mavjud bo‘lsa, ushbu test yordamida aniqlanadi. Testni amalga oshirish uchun quyidagi hisoblashlar amalga oshiriladi.



Simulyatsiya ma’lumotlari

12.4-rasm. Xeteroskedastik xatoliklar (Regressor oshishi bilan qoldiqlar dispersiyasi ham oshmoqda)



Simulyatsiya ma’lumotlari

12.5- rasm. Xeteroskedastik xatoliklar (Regressor oshishi bilan qoldiqlar dispersiyasi kamaymoqda)

1. Ko‘p omilli regressiyani EKK usulidan foydalangan holda qoldiqlar hisoblanadi.

2. Qoldiqlarning kvadiratlari va ulardan regressand sifatioda foydalangan holda, regressalarga nisbatan regressiya modeli hisoblanadi, ya'ni

$\hat{u}^2 = \delta_u + \delta_1 x_1 + \delta_2 x_2 + \cdots \delta_k x_k +$ xatolik modeli hisoblanib, uning $R^2_{\hat{u}^2}$ saqlanadi.

3. Oxirgi modeldan F-statistika (yoki LM-statistikasi) hisoblanib, uning ρ-qiymati taqsimoti (yoki X^2_k taqsimoti) dan foydalanib hisoblanadi. Bunda

$$F - \text{statika} = \frac{\frac{R^2_{\hat{u}^2}}{k}}{(1 - R^2_{\hat{u}^2})/(n - k - 1)}$$

(LM-statistika=$nR^2_{\hat{u}^2}$)

Nolinchi gipoteza xomoskedastiklik farazini tashkil etadi, shuning uchun tanlama asosida hisoblangan. F-yoki LM-statistikasi tegishli o'zining kritik qiymatlaridan baland bo'lsa (yoki $\rho -$ qiymat an'anaviy 1% 5% yoki 10% dan kam bo'lsa), nolinchi gipoteza rad etiladi va modelda xeteroskedastiklik,muammosi mavjudligi aniqlanadi.

Quyida hisoblangan $FOOD = \beta_0 + \beta_1 \cdot INCOME + u$ modelga (FOOD – muayyan uy xo'jaligining oylik oziq-ovqat iste'moli, $INCOME -$ shu uy xo'jalik oylik daromadi, ming so'mlarda) e'tibor bering.

Model parametrlarini hisoblangan parametrlar bilam aniqlashtirsak, regressiya tenglanmasini $\widehat{\text{FOOD}} = 244 + 0.6 \cdot INCOME$ ko'rinishida yozish mumkin bo'ladi.Parametrlar ishoralari kutilganidek: uy xo'jaliklari daromadlari 100 ming so'mga oshsa, o'rta hisobda ozq-ovqat iste'moli 60 ming so'mga oshmoqda (boshqa omillar o'zgarmagan sharoitda). Daromadi yo'q uy xo'jaliklari o'rta hisobda 244 ming so'mga ekvivalent oziq-ovqat iste'mol qilmoqdalar. Iqtisodiy jihatdan daromadi yo'q uy xo'jaliklari qanday qilib, oziq-ovqat iste'mol qilishi biroz mantiqsiz tuyulsada, uni tushuntirish mumkin.

Dependent Variable: FOOD
Method: Least Squares
Sample: 1 40
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INCOME	0.595784	0.124768	4.775119	0.0000
C	243.9470	194.9949	1.251043	0.2186

R-squared	0.375018	Mean dependent var	1134.294
Adjusted R-squared	0.358571	S.D. dependent var	450.7007
S.E. of regression	360.9628	Akaike info criterion	14.66413
Sum squared resid	4951178.	Schwarz criterion	14.74858
Log likelihood	-291.2827	Hannan-Quinn criter.	14.69467
F-statistic	22.80176	Durbin-Watson stat	1.863319
Prob(F-statistic)	0.000027		

12.6-rasm. $FOOD = \beta_0 + \beta_1 \cdot INCOME + u$ **modelida multikollinearlik alomatlari**

Birinchidan, bizda uy xo‘jaliklarining yashashi to‘g‘risida ma’lumot bo‘lmasada, daromadi yo‘q uy tamorqada yetishtirilgan oziq-ovqatni iste’mol qilgan bo‘lishi mumkin.

Ikkinchidan, agar bu hisoblangan parametrning statistikasiga e’tibor bersak, u statistik muhim emas. Ko‘pincha daromad oldidagi parametr iqtisodiy jihatdan qaror qabul qilishda ishlatilgani bois, aynan shu parametrning statistik xossalari muhimroq. Ushbu misolda $\hat{\beta}_1$ statistik muhimligi 1% darajada, ya’ni agar model “to‘g‘ri” hisoblangan bo‘lsa, uni qaror qabul qilish uchun foydalansa bo‘ladi.

Model “to‘g‘ri” hisoblanganligini bir qancha testlar orqali aniqlash mumkin. Broysh-Pagan (BP) testi natijalari shuni ko‘rsatadiki, F-yoki X^1 -statistikasi o‘zining kritik qiymatlaridan baland *(F-statistik=8.54, Obs·R-squared=7.34)* yoki ularga mos ravishda hisoblangan ρ-qiymatlar *(Prob. F(1,38)=0.058. Prob. Chi-Square (1)=0.0067)* an’anaviy 1% dan kichik. Bunda H_0 1% muhimlik darajasida rad etiladi va ushbu modelda BP-testiga muvofiq xeteroskedastiklik muammosi mavjudligini ko‘rsatadi.

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	8.540046	Prob. F(1,38)	0.0058
Obs*R-squared	7.339955	Prob. Chi-Square(1)	0.0067
Scaled explained SS	6.188816	Prob. Chi-Square(1)	0.0129

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample: 1 40
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-113006.7	84737.25	-1.333614	0.1903
INCOME	158.4477	54.21954	2.922336	0.0058

R-squared	0.183499	Mean dependent var	123779.5
Adjusted R-squared	0.162012	S.D. dependent var	171354.2
S.E. of regression	156860.5	Akaike info criterion	26.81281
Sum squared resid	9.35E+11	Schwarz criterion	26.89725
Log likelihood	-534.2562	Hannan-Quinn criter.	26.84334
F-statistic	8.540046	Durbin-Watson stat	2.389489
Prob(F-statistic)	0.005822		

12.7-rasm. $FOOD = \beta_0 + \beta_1 \cdot INCOME + u$ modelida Broysh-Pagan(BP) testi

Modelda xeteroskedastiklik muammosini aniqlashda Uayt testidan ham foydalaniladi. Bu testni amalga oshirish quyidagicha:

1. Ko‘p omilli regressiyani EKK usuli orqali hisoblab, qoldiqlar va model bo‘yicha regressor ($\hat{y}$) va ularni kvadratlari hisoblanadi;

2. Qoldiqlar kvadratlarini regressand sifatida ishlatib, ularni model bo‘yicha regressor va uning kvadratiga nisbatan regressiya modeli hisoblanadi, ya’ni $\hat{y}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2\, \hat{y}^2 +$ xatolik modeli hisoblanib, uning $R^2_{u^2}$ saqlanadi.

3. Oxirgi modeldan F-statistika (yoki LM-statistikasi) hisoblanib, uning $\rho -$ qiymati $F_{2,\ n-3}$ taqsimoti (yoki X^2_2 taqsimoti)dan foydalanib hisoblanadi. Qaror qabul qilish jarayoni BP – testidan farq qilmaydi.

Quyidagi jadvalda xeteroskedastiklik muammosi mavjudligi Uayt testi natijalari orqali ko‘rsatilgan. Regressor soni bitta bo‘lganligi bois, Eviews dasturi berilgan regressor (INCOME) ning kvadrati qo‘shib hisoblangan.

Ko‘rinib turganidek, Uayt testi ham BP testi natijasini yana bir marta tasdiqladi. Bunda hisoblangan F-statistikasi va LM (*Obs·R-squared*) statistikasi 5% darajada statistik muhim bo‘lganligi uchun H_0 rad etiladi va xeteroskedastiklik muammosi mavjudligi aniqlanadi.

Heteroskedasticity Test: White

F-statistic	4.164423	Prob. F(2,37)	0.0234
Obs*R-squared	7.349710	Prob. Chi-Square(2)	0.0254
Scaled explained SS	6.197041	Prob. Chi-Square(2)	0.0451

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample: 1 40
Included observations: 40

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-95485.20	187467.2	-0.509343	0.6135
INCOME^2	0.009536	0.090694	0.105141	0.9168
INCOME	131.1377	265.4938	0.493939	0.6243

R-squared	0.183743	Mean dependent var	123779.5
Adjusted R-squared	0.139621	S.D. dependent var	171354.2
S.E. of regression	158942.4	Akaike info criterion	26.86251
Sum squared resid	9.35E+11	Schwarz criterion	26.98918
Log likelihood	-534.2502	Hannan-Quinn criter.	26.90831
F-statistic	4.164423	Durbin-Watson stat	2.392116
Prob(F-statistic)	0.023378		

12.8-rasm. $\boldsymbol{FOOD = \beta_0 + \beta_1 \cdot INCOME + u}$ **modelida Uayt testi**

Boshqa misolda ko‘radigan bo‘lsak, oldinroq hisoblangan $sales = \beta_0 + \beta_1 price + \beta_2\, ads + u$ modelida xeteroskedastiklik muammosi mavjudligini tekshirib ko‘raylik.

Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	1.278517	Prob. F(2,72)	0.2847
Obs*R-squared	2.572226	Prob. Chi-Square(2)	0.2763
Scaled explained SS	2.577210	Prob. Chi-Square(2)	0.2757

Test Equation:
Dependent Variable: RESID²
Method: Least Squares
Sample: 1 75
Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	15.88790	19.58346	0.811292	0.4199
PRICE	0.044781	1.689590	0.026504	0.9789
ADS	-3.368185	2.106437	-1.598996	0.1142
R-squared	0.034296	Mean dependent var		10.18633
Adjusted R-squared	0.007471	S.D. dependent var		15.12156
S.E. of regression	15.06496	Akaike info criterion		8.301799
Sum squared resid	16340.62	Schwarz criterion		8.394498
Log likelihood	-308.3174	Hannan-Quinn criter.		8.338812
F-statistic	1.278517	Durbin-Watson stat		1.964887
Prob(F-statistic)	0.284695			

12.9-rasm. $\boldsymbol{sales = \beta_0 + \beta_1 price + \beta_2\, ads + u}$ ***modelida BP test***

BP testi natijasi shuni ko‘rsatadiki, unda H_0 rad etilmaydi, ya’ni xeteroskedastiklik muammosi bu modelda yo‘q, deb xulosa qilsak bo‘ladi. Xeteroskedastiklikka Uayt test bo‘yicha tekshirilgan ham bu natija tasdiqlanadi. Empirik ishlarda hamma vaqt ham bu xeteroskedastiklik testlari bir xil javob bermaydi. Xeteroskedastiklikning funksional ko‘rinishiga qarab, bu ikkala test har xil natija berishi mumkin.

Heteroskedasticity Test: White

F-statistic	0.888642	Prob. F(5,69)	0.4936
Obs*R-squared	4.537393	Prob. Chi-Square(5)	0.4749
Scaled explained SS	4.546185	Prob. Chi-Square(5)	0.4737

Test Equation:	
Dependent Variable: RESID^2	Sample: 1 75
Method: Least Squares	Included observations: 75

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-259.4154	276.5095	-0.938179	0.3514
PRICE^2	-2.382867	2.101433	-1.133925	0.2608
PRICE*ADS	1.416045	1.950960	0.725819	0.4704
PRICE	51.61169	48.30201	1.068520	0.2890
ADS^2	0.365920	3.137046	0.116645	0.9075
ADS	-20.53198	24.99790	-0.821348	0.4143

R-squared	0.060499	Mean dependent var	10.18633
Adjusted R-squared	-0.007581	S.D. dependent var	15.12156
S.E. of regression	15.17877	Akaike info criterion	8.354291
Sum squared resid	15897.26	Schwarz criterion	8.539690
Log likelihood	-307.2859	Hannan-Quinn criter.	8.428319
F-statistic	0.888642	Durbin-Watson stat	1.926181
Prob(F-statistic)	0.493597		

12.10-rasm. $\boldsymbol{sales = \beta_0 + \beta_1 price + \beta_2\, ads + u}$ ***modelida Uayt testi***

Xeteroskedastiklik asoslari

Xeteroskedastiklik muammosi mavjud modelda odatda, modelga tegishli boshqa muammolar (model spetsifikatsiyasi) ham bo‘ladi – ularga tegishli asoratlar boshqa bo‘lsa-da, hozir uchun boshqa muammolar bartaraf etilgan holatda EKK usuli yordamida hisoblangan parametrlar va qaror qabul qilishda qanday asoratlarga olib kelishini keltirib o‘tamiz.

➢ EKK usuli yordamida hisoblangan parametrlar o‘rta hisobda haqiqiysidan siljimaydi, ya’ni $E(\hat{\beta}_j) = \beta_j$ saqlanadi. Xeteroskedastiklik va xomoskedastiklik sharoitlarida hisoblangan parametrlar bir xil bo‘ladi.

➢ EKK usuli yordamida hisoblangan parametr standart xatoliklari esa (1) eng kichik bo‘lmaydi (xomoskedastiklik xatolar sharoitida eng kichik bo‘ladi!) va (2) haqiqiy standart xatoliklardan siljib qoladi.Bunday xatoliklar asosida hisoblangan t-statistika va p-qiymatlar ham noto‘g‘ri ma’lumotni ko‘rsatadi va natijada noto‘g‘ri ma’lumotni ko‘rsatadi va natijada noto‘g‘ri qaror qabul qilishga olib keladi.

Xeteroskedastiklikka bardosh standart xatoliklar

Agar BP testi yoki Uayt testi natijasida modelda xeteroskedastiklik muammosi aniqlangan bo‘lsa, EKK usuli yordamida hisoblangan parametrlarning standart xatoliklari noto‘g‘ri bo‘lgani uchun, shularni to‘g‘ri hisoblash maqsadga muvofiq hisoblanadi. To‘g‘ri standart xatoliklar ekonometrik adabiyotlarda Uayt, Xuber yoki Eyker standart xatoliklari nomi bilan ma’lum. Bir nechta olimlar ishtirok etgani tufayli ularni Uayt standart xatoliklari yoki xeteroskedastiklikka bardosh beruvchi standart xatoliklardan foydalanib, hisoblangan t-statistikasi xeteroskedastiklikka bardosh $t-$ statistikasi va hokazo deb yuritiladi.

Dependent Variable: FOOD
Method: Least Squares
Sample: 1 40
Included observations: 40
White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INCOME	0.595784	0.116275	5.123939	0.0000
C	243.9470	144.8731	1.683866	0.1004

R-squared	0.375018	Mean dependent var	1134.294
Adjusted R-squared	0.358571	S.D. dependent var	450.7007
S.E. of regression	350.9628	Akaike info criterion	14.66413
Sum squared resid	4951178	Schwarz criterion	14.74858
Log likelihood	-291.2827	Hannan-Quinn criter.	14.69467
F-statistic	22.80176	Durbin-Watson stat	1.863319
Prob(F-statistic)	0.000027	Wald F-statistic	26.25475
Prob(Wald F-statistic)	0.000009		

12.11-rasm. $\boldsymbol{FOOD = \beta_0 + \beta_1 \cdot INCOME + u}$ **modelida xeteroskedastiklikka bardosh standart xatoliklar**

Yuqoridagi jadvalda keltirilgan ma’lumotlar 12.6-rasmdagidan farqi shundaki, unda EKK usuli yordamida hisoblangan $FOOD = \beta_0 + \beta_1 \cdot INCOME + u$ modelda parametr standart xatoliklari xeteroskedastiklikka bardoshli hisoblanadi. Hisoblangan parametrlar qiymati bir xil, lekin ularning standart xatoliklari har xil. Shu bois, $t -$ statistika va p - qiymatlar ham o‘zgargan. Xeteroskedastiklikka bardosh standart xatoliklar kamaygani bois, t - statistika oshgan va p - qiymatlar kamaygan, ya’ni standart xatoliklar yanada aniqlashgan hisoblanadi.

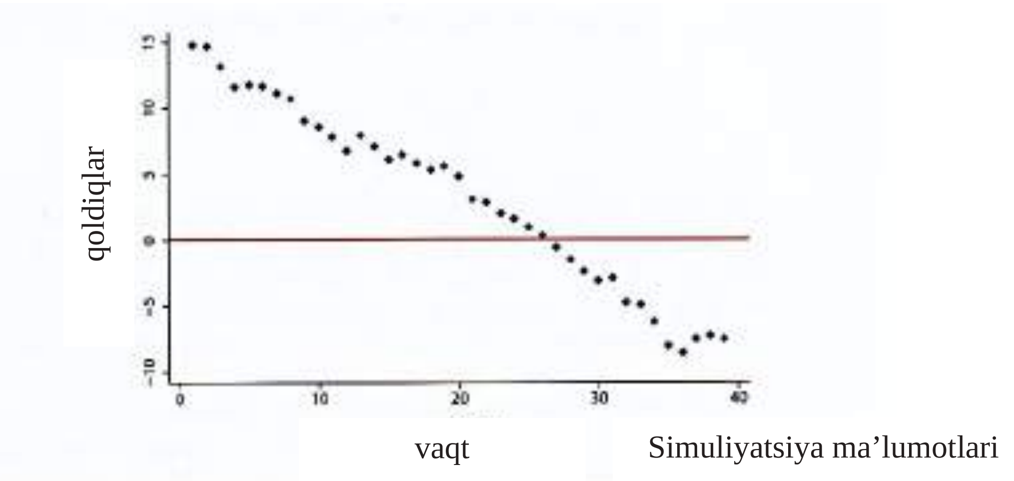
12.6. Avtokorrelyatsiya

Avtokorrelyatsiya muammosi ko‘pincha dinamik qatorlarda uchraydi. Gauss-Markov shartlaridan biri-bu $y = \beta_0 + \beta_1 x_1 + \cdots \beta_k x_k + u$ regrassiya modelining tasodifiy xatolari bir-biriga bog‘liq emas, farazidir, ya’ni $cov(u_i, u_j) = 0,\ i \neq j$. Bunday faraz kross-seksion ma’lumot turiga ko‘proq to‘g‘ri kelsa-da, dinamik qatorlarda ko‘pincha buziladi. Masalan, O‘zbekistonning iqtisodiy o‘sish ko‘rsatkichlarini olaylik.Bir yili baland o‘sish ko‘rsatkichi qayd etilgan bo‘lsa, keyingi yillarda ham shunday bo‘lishi kutiladi, ya’ni tanlanmaning ikkita qiymati orasida qandaydir bog‘liqlik bor. Bunday bog‘liqlik modelining Gauss-Markov shartlarini qoniqtirmaydi, shu sababli EKK usuli yordamida hisoblangan parametrlar kerakli xossalarga ega bo‘lmaydi. Hisoblangan parametrlar qaror qilishda keraksiz bo‘lib qoladi.

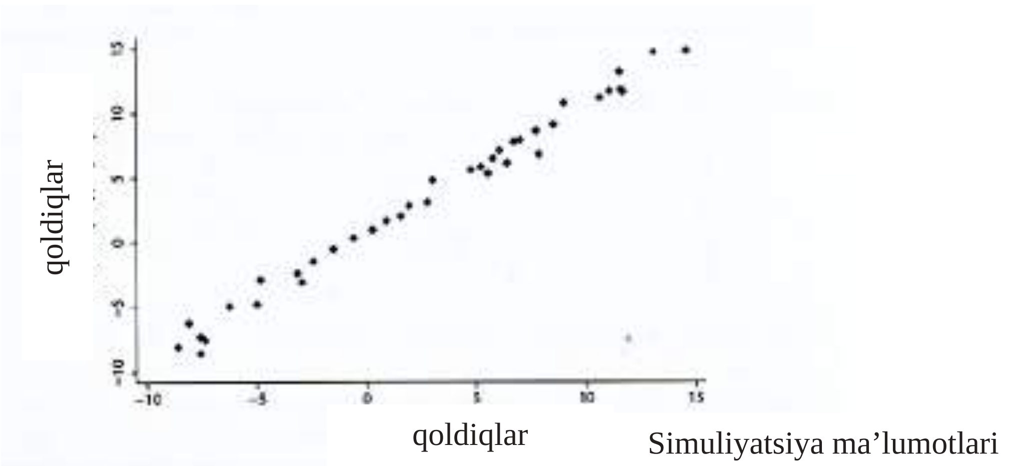
Agar bugungi kundagi tasodifiy xatolik o‘tgan davrdagi qiymati bilan bog‘langan bo‘lsa, **birinchi tartibdagi avtokorrelyatsiya**, undan oldingi davrdagi xatolik bilan bog‘langan bo‘lsa, **ikkinchi tartibdagi avtokorrelyatsiya** va hokazo deyiladi. Birinchi va undan yuqori tartibdagi avtokorrelyatsiya musbat yoki manfiy qiymatlarni qabul qiladi.

Avtokorrelyatsiya mammosini aniqlash

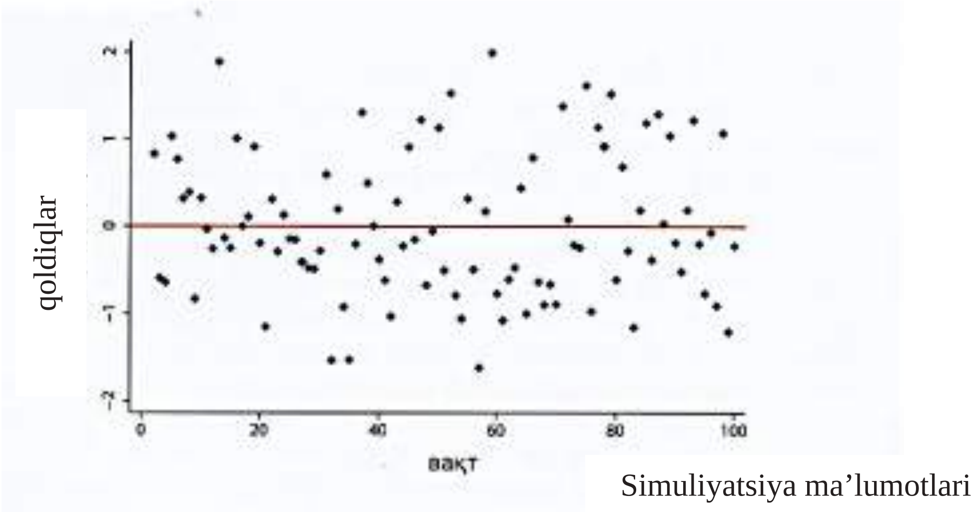
Xeteroskedastiklikka o‘xshab, avtokorrelyatsiya muammosini ham grafiklar yordamida tahlil qilish va o‘rganish mumkin. 12.12-rasmda qoldiqlar vaqt bo‘yicha tasvirlangan. Ko‘rinib turganidek, bir davrda qoldiqlar musbat bo‘lsa keyingi davrda ham musbat, manfiy bo‘lsa keyingi davrda ham manfiy qiymatlar qabul qilmoqda. Bu holat 12.13-rasmda yanada yaqqol tasvirlangan.



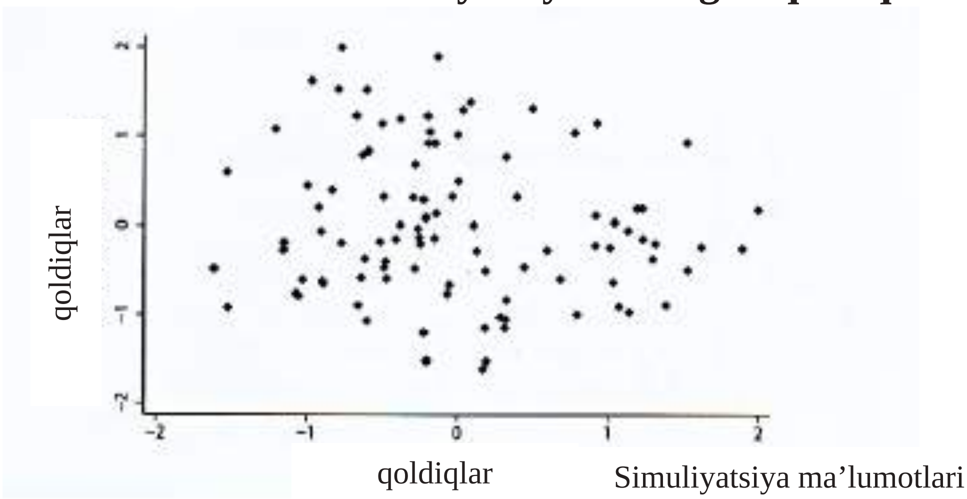
12.12-rasm. Musbat avtokorrelyatsiyaga ega qoldiqlar



12.13-rasm. Birinchi tartibdagi avtokorrelyatsiyasi



12.14-rasm. Avtokorrelyatsiyalanmagan qoldiqlar



12.15-rasm. Birinchi tartibdagi avtokorrelyatsiyasi=0

Grafik usul avtokorrelyatsiya muammosini insepsiya qilish imkoniyatini yaratsa-da, u formal test emas. Eng mashhur formal testlardan biri - bu Darbin- Uotson testidir. Darbin-Uotson statistikasi *(d)* quyidagicha hisoblanadi:

$$d = \frac{\sum_{t=2}^{n}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n}\hat{u}_t^2}$$

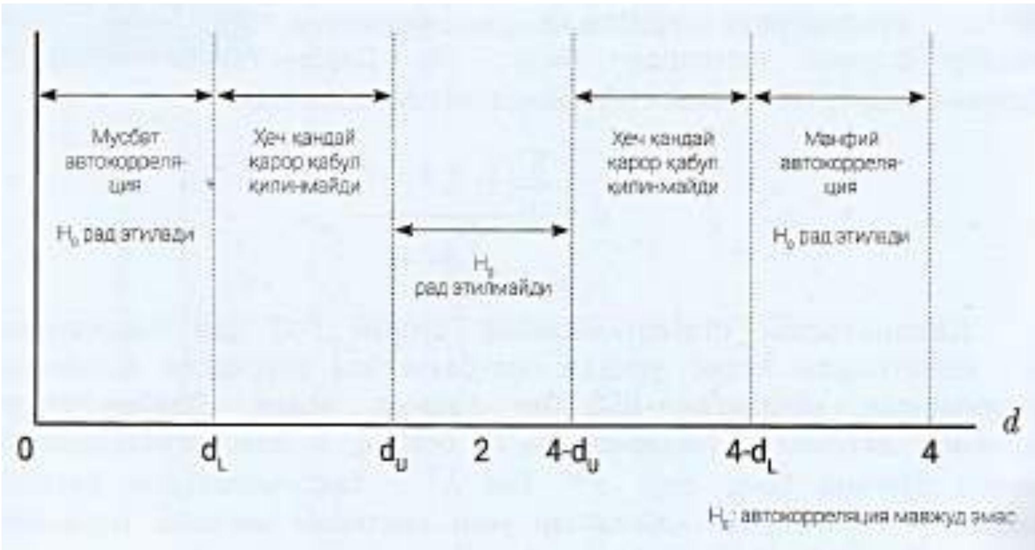
Darbin-Uotson statistikasining surati t=2 boshlanishi ishlatishdan kelib chiqadi, mahraji esa regrresiya modelini hisoblashdan keladigan RSS ni tashkil etadi. Darbin-Uotson testining ehtimoliy taqsimoti aads ga bog‘liq csdd esa regrressorlarga bog‘liq bo‘lgani bois, $t-$, $F-$ yoki $X2-$ taqsimotlardan farqli o‘laroq, uni universal qiymatlar uchun keltirib chiqarish murakkab masala. Lekin Darbin va Uotson d-statistikasini pastki (d_L) va yuqori (d_U) qiymatlarni, muvaffaqyatli hisoblab chiqishga (alohida jadvallar yaratilgan) va bu qiymatlar avtokorrelyatsiya gipotezasini tekshirishda ishlatiladi.

12.7. Darbin-Uotson statistikasi yordamida qaror qabul qilish

Darbin-Uotson statistikasi d ni bir oz o‘zgartirib, quyidagi ko‘rinishida ifodalash mumkin; bunda,

$$p = \frac{\sum_{t-2}^{n} \hat{u}_t \cdot \hat{u}_{t-1}^2}{\sum_{t=1}^{n} \hat{u}_t^2}$$

ya’ni $\hat{u}_t$va $\hat{u}_{t-1}$ o‘rtasidagi birinchi tartibdagi korrelyatsiya koeffitsiyenti $|p| \leq 1$ bo‘lgani uchun $0 \leq d \leq 4$ bo‘ladi. Agar d ning hisoblangan qiymati 2 atrofida bo‘lsa, birinchi tartibdagi avtokorrelyatsiya muammosi mavjud bo‘lmaydi, $d \leq d_L$ yoki $d \leq d_U$ bo‘lsa, birinchi tartibdagi avtokorrelyatsiya muammosi mavjud bo‘ladi.

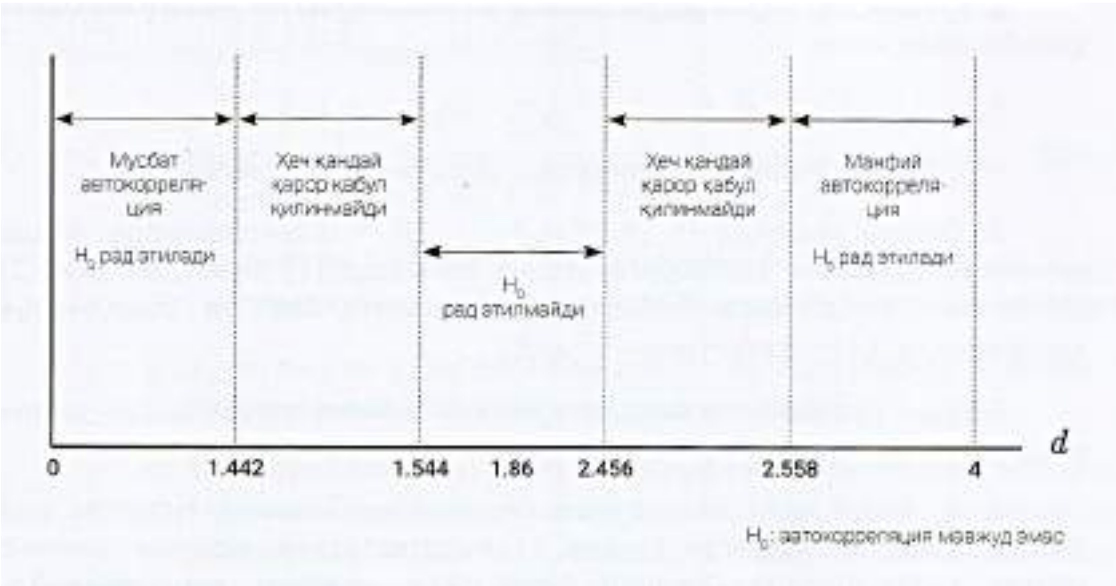


Masalan, Keyin iste’mol funksiyasi hisoblanganda (7-rasm), $d = 1.86$ ekanligini ma’lum bo‘lgan edi. Bu statistik nuqtai nazardan 2 dan farqlimi degan savol tug‘iladi, chunki farqli bo‘lsa, avtokorrelyatsiya muammosi mavjud bo‘ladi, aks holda bo‘lmaydi. Buni tekshirish maqsadida d_L va d_U ni jadvaldan bitta regrresorli model uchun topib olib, 4-d_L va 4-d_U hisoblab olamiz.

12.8. $FOOD = \beta_0 + \beta_1 \cdot INCOME + u$ modelida Darbin-uotson statistikasi yordamida avtokorrelyatsiya muammosiga tekshirish

Ko‘rinib turganidak, hisoblangan $d = 1.86$ H_0 rad etilmaydigan intervalda joylashgani tufayli birinchi tartibdagi avtokorrelyatsiya muammosi bu modelda yo‘q, deb xulosa qilinadi.

Darbin-Uotson testini amalga oshirishda bir nechta shartlar amalga oshishi kerak. Birinchidan, regressorlar deterministik bo‘lishi yoki bo‘lmasligi sharti bilan. Masalan, keyingi bobda ko‘rsatilgan AR(1) va yuqoriroq avtoregressiya modellarida regressorlar regressand-ning lagli qiymatidan iborat bo‘lgani uchun deterministi emas.



Ikkinchidan, tasodifiy xatoliklar normal taqsimot qonuniga bo‘ysunishi kerak. Bu shart bajariladigan holatlar ko‘p va uchinchidan, bu test faqat birinchi tartibdagi avtokorrelyatsiyani hisoblaydi. Agar modelda yuqoriroq tartibdagi avtokorrelyatsiya muammosi mavjud bo‘lsa, bu test uni aniqlamaydi.

Broysh-Godfri (BG) testi

Darbin-Uotson testi bilan bog‘liq kamchiliklar uni amalda ishlatishni bir oz chegaralaydi. Cheklovlari kamroq ya’ni stoxastik regressor va yuqori tartibdagi avtokorrelyatsiyani hisobga oluvchi test Broysh-Godfri tomonidan ishlab chiqilgan. Uni odatda, BG testi yoki LM testi nomi bilan yuritiladi.

U quyidagicha amalga oshiriladi:

1. Ko‘p omilli regressiyani EKK usuli orqali qoldiqlari hisoblanadi.

2. Qoldiqlar barcha regressorlar va qoldiqlardagilariga nisbatan hisoblanadi, ya’ni

$$\hat{u}_t = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \cdots + \delta_k x_k + p_1 \hat{u}_{t-1} + p_2 \hat{u}_{t-2} + \cdots p_p \hat{u}_{t-p} +$$

Xatolik modeli hisoblanib, uning $R^2_{\hat{u}^2}$ saqlanadi.

3. Oxirgi modelda $H_0: \rho_1 = \rho_2 = \cdots = \rho_p = 0$ tekshiriladi. Bunda xeteroskedastiklik testlariga o‘xshash ravishda (1) F- testdan yoki (2) LM- testdan foydalansa bo‘ladi. LM-statistika bir oz boshqacharoq hisoblanadi. LM- statistika = $(n-q)R^2_{\hat{u}^2}$.

Bundan tashqari, avtokorrelyatsiyaga BG bo‘yicha tekshirishdan oldin modelda xeteroskedastiklik muammosini yo‘qotish kerak.

Agar F- test yoki LM-test natijasiga ko‘ra nolinchi gipotez rad etilsa, ya’ni hisoblangan F- yoki LM- statistikasi tegishli o‘zining kritik qiymatlaridan baland bo‘lsa (yoki qiymat an’anaviy 1%, 5% yoki 10% dan kam bo‘lsa), modelda avtokorrelyatsiya muammosi mavjudligi aniqlanadi.

Avtokorrelyatsiya asoratlari. Avtokorrelyatsiya muammosi mavjud EKK usuli bilan hisoblangan parametrlar o‘rta hisobda siljimasa-da, ularning samaradorligi, ya’ni standart xatoligi noto‘g‘ri bo‘lib chiqadi. Agar mavjud avtokorrelyatsiya ishorasi musbat bo‘lsa, standart xatoliklar haqiqiysidan kichikroq bo‘ladi. Bu model parameterlarini statistik muhimlik darajasiga tekshirilganda, aslida “muhimroq” darajani ko‘rsatadi.

Avtokorrelyatsiya muammosini bartaraf etish. Bunday savolga javob berishdan oldin modelga avtokorrelyatsiyadan oldin boshqa muammolarini hal qilish zarur. Masalan, ko‘p hollarda model noto‘g‘ri tanlangan bo‘lsa, avtokorrelyatsiyaga tekshiruvchi testlar avtokorrelyatsiya muammosiga ishora qilishi mumkin. Modelda to‘g‘ri regressorlar tanlanishi bilan avtokorrelyatsiya yo‘qolishi mumkin.

Agar BG testi natijasida modelga avtokorrelyatsiya muammosi aniqlansa, EKK usuli yordamida hisoblangan parametrlarning standart holiqlari noto‘g‘ri bo‘lgani uchun ularni to‘g‘ri hisoblash maqsadga muvofiq hisoblanadi. To‘g‘ri standart xatoliklar ekonometrik adabiyotlarda Ney-Vest standart xatoliklari nomi bilan ma’lum. Bundan tashqari, EKK usulidan voz kechib, umumiy kvadratlar usuli (UKK) dan foydalanish ham maqsadga muvofiq.

XIII BOB. DINAMIK QATORLAR REGRESSIYASI

13.1. Laglar, birinchi ayirma, logarifm va o‘sish darajalari
13.2. Avtoregressiya modellari
13.3. Avtoregressiv taqsimlangan lag (ADL) modeli
13.4. Ma’lumot kriteriylari yordamida lag tartibini tanlash
13.5. Nostatsionarlik

13.1. Laglar, birinchi ayirma, logarifm va o‘sish darajalari

Dinamik qatorlar - muayyan o‘zgaruvchining vaqt bo‘yicha to‘plangan qiymatlaridan iborat. Dinamik qatorlardan iborat tanlanmaning cross-section ma’lumotlaridan iborat tanlanmadan asosiy farqi shundaki, cross-section kuzatishlar muayyan o‘zgaruvchi bo‘yicha bir vaqtda yig‘ilgani uchun bir-biriga ta’sir etmaydi, dinamik qatorlarda esa, aksincha, ta’sir etadi.

Ish haqi o‘zgaruvchisini ko‘raylik: bir hududda bir nechta shaxsdan ish haqi bo‘yicha ma’lumot so‘rovnoma orqali yig‘ilsa (cross-section tanlanma), tasodifiy tanlangan bir shaxsning ish haqi boshqasinikiga ta’sir etmaydi. Bir shaxsning ish haqi vaqt bo‘yicha (oyma-oy, masalan) yig‘ilganda (dinamik qatorlar), uning fevraldagi oyligi baland bo‘lsa, mart oyidagisi ham baland bo‘lishi kutiladi, ya’ni o‘zgaruvchining vaqt bo‘yicha qiymatlari bir-biriga ta’sir etadi.

Boshqa misol sifatida O‘zbekiston Respublikasining iqtisodiy o‘sish ko‘rsatkichlari yillar bo‘yicha ko‘rilganda, bir yili yuqori darajadagi o‘sish ko‘rsatkichi qayd etilgan bo‘lsa, keyingi yillarda ham yuqori darajadagi o‘sish, bir yili past bo‘lganda, keying yillarda ham past bo‘lishi kuzatiladi.

Dinamik qatorlarning ushbu xossasi ma’lumotlarni tasodifiy tanlama sifatida ko‘rilganda, iqtisodiy tahlilda bir qator qiyinchiliklarni tug‘diradi. Boshqa tarafdan, dinamik qatorlar, cross-section tanlamadan farqli o‘laroq, dinamik ta’sirni hisoblash imkoniyatini yaratadi. Keyingi oyda inflyatsiya bo‘yicha sizning eng yaxshi bashoratingiz qanday? Bu kabi savolga dinamik qatorlar javob bersa-da, ularni tahlil qilish kross-seksion tanlamalarga xos bo‘lmagan uslublarni ko‘rib chiqishga undaydi.

Dinamik qatorlarda Y o‘zgaruvchisining t vaqtda qabul qilingan qiymati Y_t, bunda kuzatishlar soni T, ya’ni $t = 1,2,..T$ deb belgilanadi. Tanlama doirasida t va undan keyingi vaqt $t+1$ kuzatishlarning vaqt bo‘yicha intervalini ko‘rsatadi. Masalan, agar l yillarda berilgan bo‘lib, Y_to‘zgaruvchining bugungi kundagi qiymatini ko‘rsatsa, uning o‘tkan yilgi qiymati Y_{t-1}, avvalgi yildagi qiymati Y_{t-2} va hokazo deb belgilanadi.

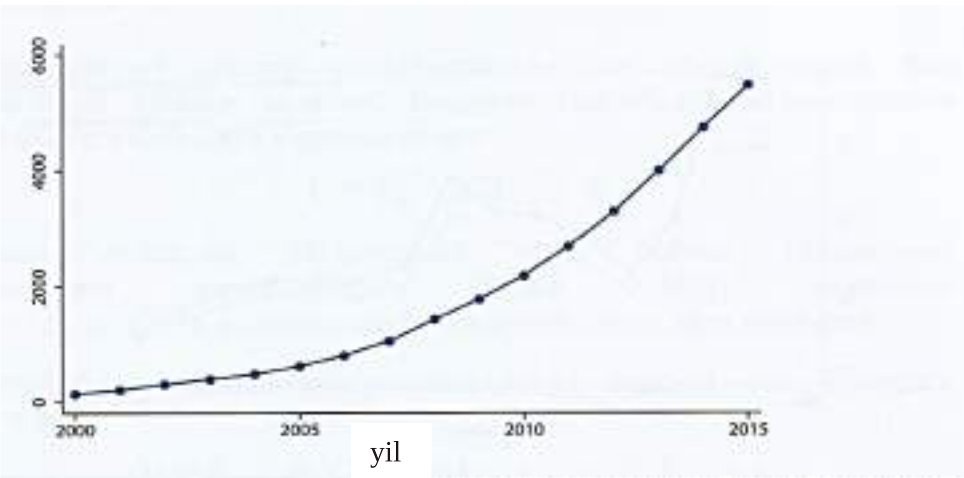
Dinamik qatorlarda o‘zgaruvchining oldingi qiymatlarini ishlatishda maxsus atamalar qo‘llaniladi. Y_{t-1}- ma’no jihatidan Y o‘zgaruvchisining o‘tgan vaqtdagi qiymatini ko‘rsatsa-da, uni Y o‘zgaruvchisining birinchi lagi yoki birinchi lagi $Y, Y_{t\;j}$ esa Y o‘zgaruvchisining j −lagi yoki j −lagli Y deb yuritiladi. Y o‘zgaruvchisining t va $t-1$ oralig‘idagi o‘zgarishi $Y_t - Y_{t-1}$ uning birinchi ayirmasi deb ataladi va $\triangle Y_t = Y_t - Y_{t-1}$ deb belgilanadi. Quyidagi jadvalda Y o‘zgaruvchining O‘zR YaIM jon boshiga misolida hisoblangan lagli va birinchi ayirmasi keltirilgan.

t, йиллар	Y_t	Y_{t-1}	ΔY_t	Y_{t-2}
2010	2 184.3	1 778.2[25]	406.1	1 427.3
2011	2 684.6	2 184.3	500.3	1 778.2
2012	3 289.0	2 684.6	604.4	2 184.3
2013	3 996.3	3 289.0	707.3	2 684.6
2014	4 741.8	3 996.3	745.5	3 289.0
2015	5 475.1	4 741.8	733.3	3 996.3

13.1-rasm. O‘zbekiston Respublikasi jon boshiga YaIM (Y_t), ming so‘mda

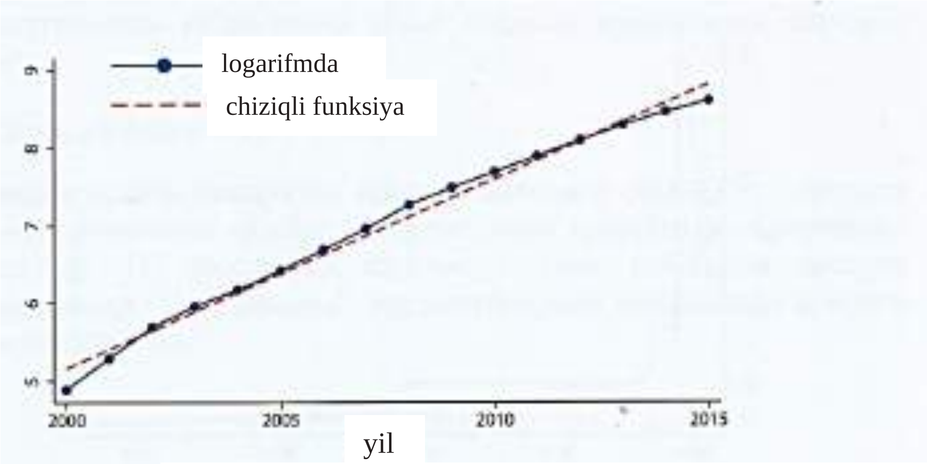
Dinamik qatorlarda iqtisodiy tahlil ko‘pincha ularning logarifmlarini yoki logarifmlar ayirmasini hisoblash orqali amalga oshiriladi. Buning bir nechta sabablari bor. Birinchidan, ko‘p iqtisodiy ko‘rsatkichlar, jumladan, YaIM taqriban eksponensial o‘sish ko‘rsatkichlariga ega, ya’ni vaqtlar oralig‘idagi o‘sish sur’atlari bir-biriga juda yaqin bo‘ladi. Bunday qatorlar logarifmlari taqriban chiziqli o‘sish ko‘rsatkichlarini ko‘rsatadi, chiziqli funksiyalar esa qulay matematik xossalarga ega. Ikkinchidan, ko‘p iqtisodiy ko‘rsatkichlar logarifmlarining standart xatoligi qabul qilgan qiymatiga proporsional bo‘ladi va shuning uchun vaqt bo‘yicha o‘zgarmaydi.

O‘zbekiston Respublikasining jon boshiga to‘g‘ri keladigan YaIM eksponensial o‘sishini namoyon qilsa-da (13.2-rasm), uning logarifmi taqriban chiziqli o‘zgarishni ko‘rsatmoqda (13.3-rasm). Bundan tashqari, iqtisodchilar o‘sish qiymatlarini taqribiy hisoblash maqsadida logarifm ayirmasidan foydalanishadi. 13.4-rasmda logarifmlar ayirmasi o‘sish qiymatlariga taqriban yaqin; ayniqsa, o‘sish ko‘rsatkichlari oldingi yillarga nisbatan pastroq qiymat qabul qilganida yanada yaqin bo‘lishi ko‘rinib turipti. 13.5- rasmda esa, boshqa dinamik qator - qayta moliyalash stavkasi tasvirlangan. O‘zbekiston mustaqilligining dastlabki yillarida ro‘y bergan giperinflyatsiya qayta moliyalash stavkasida o‘z aksini topganligi ko‘rinib turipdi.



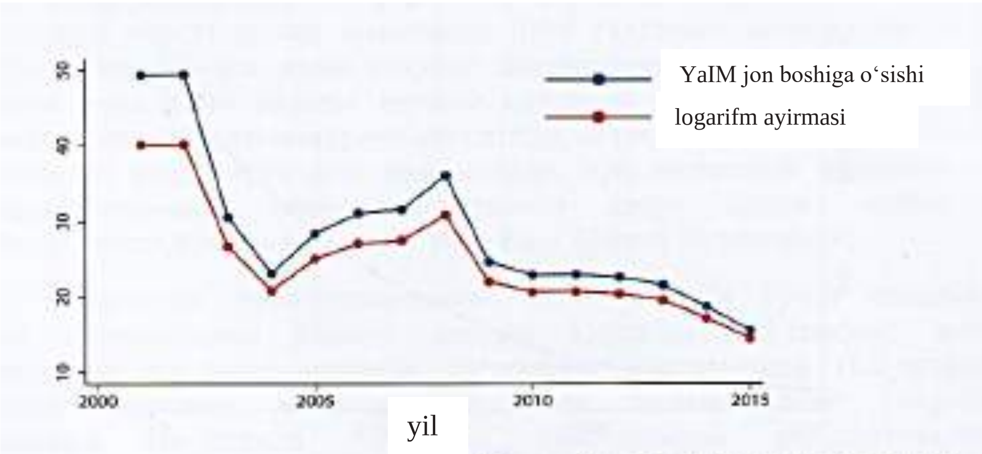
Manba: O‘zbekiston Resheblika qo‘mitasi, 2016

13. 2-rasm O‘zR YaIM jon boshiga, ming so‘mlarda



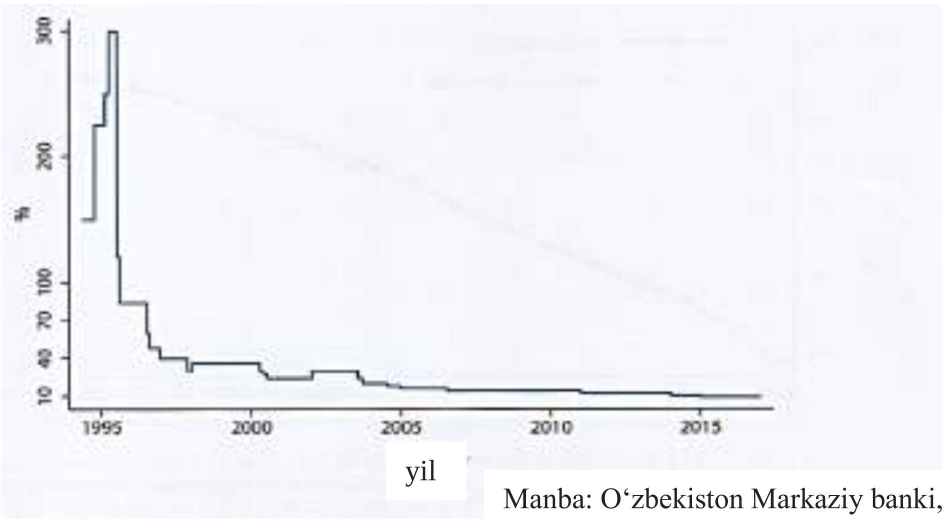
Manba: O‘zbekiston Resheblikasi Staqtisikasi qo‘mitasi, 2016

13.3-rasm O‘zR YaIM jon boshiga natural logarifmda



Manba: O‘zbekiston Resheblikasi Staqtisikasi qo‘mitasi, 2016

13.4-rasm YaIM jon boshiga: o‘sishi va logarifm ayirmasi



Manba: O‘zbekiston Markaziy banki, 2016

13.5-rasm O‘zR markaziy banki qayta moliyalash stavkasi

13.2. Avtoregressiya modellari

Avtoregressiya modeli ishlatiladigan lag soniga qarab, bir nechta tartibda bo‘lishi mumkin. Birinchi tartibda avtoregressiya yoki AR(1) modeli quyidagi ko‘rinishga ega:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_1$$

Ayrim hollarda, o‘zgaruvchini vaqt bo‘yicha o‘zgarishini modellashtrish qiziqarliroq bo‘lsa AR(1) modelini $\Delta Y_t = \beta_0 + \beta_1 \underline{\Delta} Y_{t-1} + u_1$ ko‘rinishida ifidalash maqsadga muvofiq.

p –tartibdagi avtoregressiya yoki AR(p) modeli esa quyidagi ko‘rinishga ega:

$$\Delta Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + u_t$$

Prognoz va prognoz xatoligi. Tasavvur qiling, Y_t o‘zgaruvchi bo‘yicha dinamik qator yig‘ilgan va uni AR(1) modelidan foydalanib prognozlashtrish, yani Y_{t+1}, ni topish masalasi qo‘yilgan bo‘lsin. β_0 hamda β_1 berilmaganligi uchun ularni EKK usuli yordamida hisoblab, yuqoridagi avtoregressiv tenglamadan foydalanib prognozlashtriladi: $Y\wedge_{T+1|T} = \beta\wedge_0 + \beta\wedge_1 Y_t$. Prognoz xatoligi – bu prognozlashtrilgan o‘zgaruvchini uning haqiqiy qiymatidan farqini ko‘rsatadi. Prognoz xatoligi $= Y_{T+1} - Y\wedge_{T+1\,T}$

Ildizli o‘rtacha kvadratik prognoz xatoligi (RMSE) - prognoz xatoligining hajmining ko‘rsatkichi bo‘lib, uning qiymati ikki manbadan tashkil topadi: (1) tasodifiy xatolik u_t uning noma’lum prognoz qiymatlari hamda (2) β_0 hamda β_1 parametrlarini hisoblashda vujudga keladigan xatoliklar

$$RMSE = \sqrt{}(Y_{T+1} - Y\wedge_{T+1\,|T})^2$$

Birinchi manba ikkinchisidan kattaroq bo‘lsa, ya’ni parametrlarni hisoblashda vujudga keladigan xatoliklar yetarli kichik bo‘lsa masalan, katta tanlama bilan ish ko‘rganda, yuz beradigan holat, $RMSE \approx \sqrt{Var(u_t)}$, ya’ni RSMFE regressiya standart xatoligi (SER)ga taqriban yaqinlashadi.

Inflyatsiya darajasi keyingi yilda qanday bo‘ladi? Bu savolga javob keng ommaga qiziqarli. Markaziy bank monetar siyosatni yuritishda inflyatsiya darajasidan zaruriy ma’lumot sifatida foydalanadi. Moliya vazirligi tizimida keyingi yili uchun budjet tasdiqlanayotganida inflyatsiya darajasi inobatga olinadi. Firmalar sotuv hajmlarini prognozlashtirayotganda inflyatsiya qiymatiga bog‘lashadi. Ushbu bo‘limda ushbu savolga avtoregressiya modellaridan foydalanilgan holda javob berishga harakat qilamiz.

$INFL_t = \beta_0 + \beta_1 INFL_{t-1} + u_t$ modelini ko‘raylik. Bunda $INFL_t -$ O‘zbekiston Respublikasida inflyatsiya (YAIM deflyatori) darajasining joriy yildagi qiymati $INFL_{t-1}$ esa o‘tgan davrdagi (yildagi) qiymatini ko‘rsatsin. EKK usuli yordamida modelni baholasak, 12.6-rasmda ko‘rsatilgan natija hosil bo‘ldi. Regressiya tenglamasini $INFL_t = 4.55 + 0.68 * INFL_{t-1}$ ko‘rinishda yozib olsak, uning prognoz qiymatlarini $INFL_t = 4.55 + 0.68 * INFL_T$ bo‘yicha topib olsak bo‘ladi.

Masalan, T=2015 bo‘lgani uchun

$$\begin{aligned} Y_{2016*2015} &= 4.552533 + 0.6839516 * INFL_{2015} \\ &= 4.55 + 0.68 * 8.68262 \approx 10.49 \end{aligned}$$

Inflyatsiyaning 2016-yildagi prognoz qiymatini 2017-yil prognoz qiymatiga qo‘llab, uning 2017-yildagi qiymatini prognoz qilsak bo‘ladi. Inflyatsiyaning prognozi kelajakda qanchalik yaxshi yoki yomon ekanligi tekshirishning bir yo‘li bu hisoblangan inflyatsiya prognoz qiymatlari bo‘lib o‘tgan davrdagi qiymatlarga qanchalik yaqin yoki yaqin emasligini ko‘rishimiz mumkin. Agar kelajak davr o‘tgan davrga o‘xshash bo‘lsa, ya’ni parametrlarni hisoblashda vujudga keladigan xatoliklar kam bo‘lsa, va shu bilan birga, tanlama davri uchun prognoz haqiqiy qiymatlarga qanchalik yaqin bo‘lsa, kelajak davr uchun prognoz shunchalik sifatli, ya’ni prognoz xatoligi kam bo‘ladi. 13.7-rasmda bu holat tasvirlangan.

Ko‘rinib turganidek, prognoz 2002-2015-yil oralig‘i uchun inflyatsiyaning prognoz qiymatlari haqiqiy qiymatlarga yaqin. Prognoz xatoligi shu tasvirlangan ikki chiziqlar farqidan iborat. 2016-yil uchun prognoz xatoligini hisoblab bo‘lmasada ($INFL_{2016}$ $hali\ ma'lum\ bo'lmagan\ sababli$), prognoz sifatini belgilovchi bir nechta ko‘rsatkichlarni ko‘rishimiz mumkin.

Dependent Variable: INFL

Method: Least Squares
Sample: 2002 2015
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INFL(-1)	0.683952	0.141290	4.840748	0.0004
C	4.552533	3.497576	1.301625	0.2175
R-squared	0.661331	Mean dependent var		20.04767
Adjusted R-squared	0.633109	S.D. dependent var		8.706984
S.E. of regression	5.273955	Akaike info criterion		6.295002
Sum squared resid	333.7752	Schwarz criterion		6.386296
Log likelihood	-42.06501	Hannan-Quinn criter.		6.286551
F-statistic	23.43284	Durbin-Watson stat		1.905866
Prob(F-statistic)	0.000405			

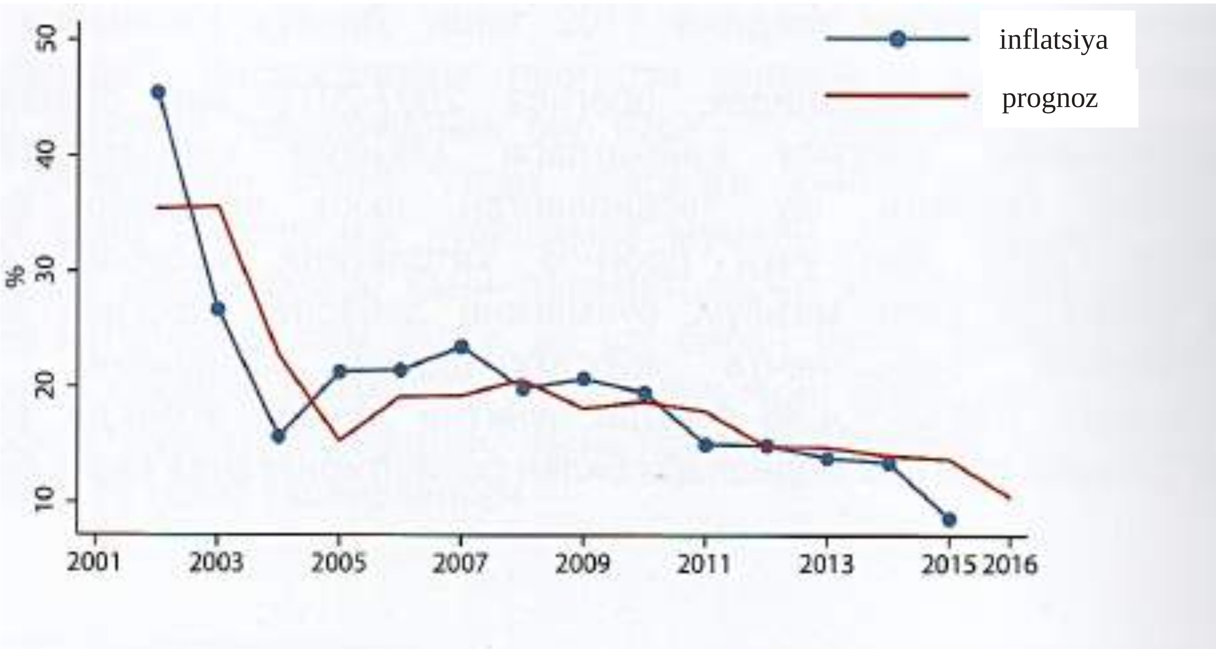
13.6-rasm. O‘zbekiston Respublikasida inflayotsiyani AR(1) asosida prognozlashtirish

Jumladan, $RMSE = 4.88$ foizli punktni tashkil etmoqda. Bu raqam AR(1) boshqa prognoz modellari bilan solishtirilganda kerak bo‘ladi.

Forecast: INFLF
Actual: INFL
Forecast sample: 2002 2016
Included observations: 14

Root Mean Squared Error	4.882734
Mean Absolute Error	3.741180
Mean Absolute Percentage Error	19.10332
Theil Inequality Coefficient	0.113791
Bias Proportion	0.000000
Variance Proportion	0.103008
Covariance Proportion	0.896992

13.7-rasm. Inflyatsiya prognozi ko‘rsatkichlari



Manba: *World Development Indicators, 2016*

13.8-rasm. O‘zbekiston Respublikasida YaIM Deflyatori va uning prognoz qiymatlari

Inflyatsiyani AR(2) asosida modellashtirsak, $INFL_t = \beta_0 + \beta_1 INFL_{t-1} + \beta_2 INFL_{t-2} + u_t$ ko‘rinishiga ega bo‘ladi. Uni EKK usuli yordamida hisoblasak, 13.11-rasmda ko‘rsatilgan natija hosil bo‘ldi. Ushbu regressiya natijasi shuni ko‘rsatadiki, inflyatsiyaning o‘tgan yilgi qiymati ta’sir etsa-da, ikki yil oldingi qiymati manfiy ta’sir qilmoqda, lekin bu ta’sir statistik muhim emas. Uning prognozini hisoblasak,

$$Y_{2016\,|2015.2014} = 4.840338 + 0.793268 * INFL_{2015} - 0.110259 = 10.24 \text{ ya'ni}$$

AR(1) modeli bilan deyarli bir xil natijani bermoqda (13.8-rasm). Ajablanarlisi shundaki, inflyatsiyaning prognoz va haqiqiy qiymatlarini ko‘rsatuvchi grafik ham deyarli bir xil natija ko‘rsatmoqda.

Forecast: INFLF
Actual: INFL
Forecast sample: 2002 2016
Included observations: 14

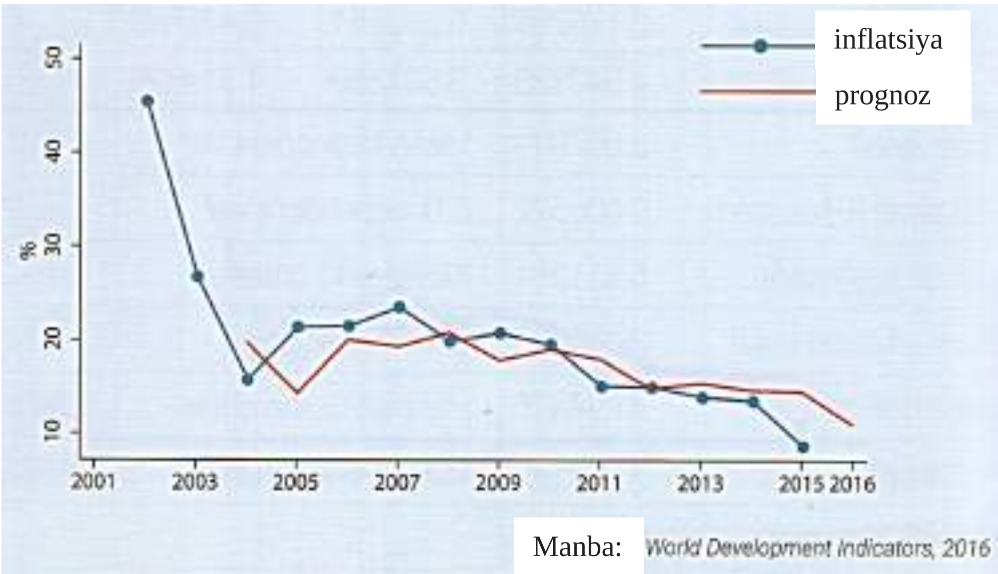
Variable	Coefficient	Std. Error	t-Statistic	Prob.
INFL(-1)	0.793268	0.290006	2.735348	0.0194
INFL(-2)	-0.110259	0.252551	-0.436580	0.6709
C	4.840338	3.681360	1.314823	0.2153
R-squared	0.667099	Mean dependent var		20.04767
Adjusted R-squared	0.606572	S.D. dependent var		8.706984
S.E. of regression	5.461354	Akaike info criterion		6.420680
Sum squared resid	328.0903	Schwarz criterion		6.557621
Log likelihood	-41.94476	Hannan-Quinn criter.		6.408004
F-statistic	11.02145	Durbin-Watson stat		2.033751
Prob(F-statistic)	0.002359			

13.9-rasm. O‘zbekiston Respublikasida inflyatsiyani AR(2) asosida prognozlashtirish

Forecast: INFLF
Actual: INFL
Forecast sample: 2002 2016
Included observations: 14

Root Mean Squared Error	4.706961
Mean Absolute Error	3.760173
Mean Absolute Percentage Error	20.66195
Theil Inequality Coefficient	0.112983
Bias Proportion	0.042831
Variance Proportion	0.308973
Covariance Proportion	0.648195

13.10-rasm. Inflyatsiya prognoz ko‘rsatkichlari



13.11-rasm. O‘zbekiston Respublikasida YaIM deflyatori va uning prognoz qiymatlari

Agar AR(2) modeli AR(1) modeli bilan solishtirganda, deyarli bir xil natijani bersa, AR(2) ga qo‘shilgan 2-lagli inflyatsiya qiymatlarini qo‘shishimiz kerakmi? Bu savolga javob qisman ijobiy bo‘lsa-da, modelning boshqa jihatlarini ham hisobga olish kerak. 2-lagli inflyatsiya parametri statistik muhim bo‘lmasa-da, model F-statistikasi statistik muhim. Ushbu o‘zgaruvchi haqiqiy modelga tegishli bo‘lsa, uni olib tashlansa, 1-lagli inflyatsiya bilan bog‘liq parametr siljib qolishi xavfi bor. Shunga lag tanlashda boshqa test natijalarini ham hisobga olish maqsadga muvofiq hisoblanadi.

13.3. Avtoregressiv taqsimlangan lag(ADL) modeli

Iqtisodiy nazariyaga ko‘ra, o‘rganilayotgan o‘zgaruvchini prognozlashtirish o‘zining laglaridan tashqari boshqa omillar va ular laglari yordamida amalga oshirish ham mumkin. Boshqa o‘zgaruvchilar $AR(p)$ modelga qo‘shilganida, **avtoregressiv taqsimlangan lag modeli** $ADL(p,q)$ hosil bo‘ladi.

$ADL(p,q)$ modeli Y o‘zgaruvchisining p −tartibgacha laglari hamda boshqa omil (X o‘zgaruvchisi) ning q −tartibgacha laglari ishlatilishi nazarda tutiladi.

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \cdots + \delta_q X_{t-q} + u_t$$

Bunda $\beta_0, \beta_1, , , , \beta_{p,}$ $\delta_1, \delta_2, , , \delta_q$ noma’lum parametrlar, u_t tasodifiy xato va $E(u_t | Y_{t-1}, Y_{t-2}, \dots, X_{t-1}, X_{t-2}, \dots) = 0$.

13.4. Ma’lumot kriteriylari yordamida lag tartibini tanlash

Avtoregressiya va avtoregressiv taqsimlangan lag modellarida lag tartibini tanlash muhim masala sifatida ko‘riladi. Modelga kam lag kiritilsa, uzoq laglardagi ma’lumot modelda aks etmaydi va prognozda ishlatilmaydi. Ikkinchi tomondan, ko‘proq lag kiritilsa, yetarlidan ortiq regrossorlar kiritilish xavfi mavjud – bunda prognoz xatoligi oshadi. Shu nuqtai nazardan, lag tartibini belgilash ko‘proq lag kiritishdan keladigan naf va prognoz xatoliklarini balanslashtirish masalasiga aylanadi. Buni amalga oshirishning 3 xil usulini taqdim etamiz.

F-test

Lag tartibini F-test orqali amalga oshirish avtoregressiya modelini yuqori tartibda oldin baholab, oxirgi tartibli lagni gipoteza testidan o‘tkazishni taqozo etadi. Masalan, berilgan regressand uchun AR(5)baholanib,5-tartibli lag oldidagi parametrni 5% (yoki an’anaviy 1%, 10%-qo‘yilgan maqsadga bog‘liq holda) darajada statistik muhimligini tekshirishimiz kerak. Agar bu darajada statistik muhimlikni o‘rnatib bo‘lmasa, AR(4) ni hisoblab, xuddi shunday test o‘tkazish orqali amalga oshiriladi. Agar 4- tartibdagi lag parametri statistik muhim bo‘lmasa, AR(3) hisoblanadi va hokazo.

Avtoregressiv taqsimlangan lag modellarida esa kiritilgan laglar oldidagi bir nechta parametrlarni F-test orqali qo‘shma gipoteza sifatida tekshiriladi. Solishtiriladigan modellar soni ko‘p bo‘lmasa, F-test oson amalga oshiriladi.

Bu usulning muammoli tomoni shundan iboratki, agar haqiqiy lag tartibi 3 bo‘lib, ya’ni 4 va undan yuqori tartibli lag parametrlari 0 bo‘lsa, test statistikasi yuqoriroq tartiblilagni tanlash kabi natijaga olib keladigan holatlar uchraydi.

Shvars ma’lumot kriteriysi (SIC)

Lag tartibini Shvars ma’lumot mezonlari orqali aniqlash p −tartibli avtoregressiya modellarida quyidagi funksiyani minimallashtirishni taqozo etadi:

$$SIC(p) = \ln\left|\frac{SSR(p)}{T}\right| + \frac{(p+1)\ln(T)}{T}$$

Bunda $SSR(p) - AR(p)$ modelini EKK usuli yordamida hisoblashda hosil bo‘ladigan qoldiqlar kvadratlari yig‘indisini ko‘rsatadi. p tartibini har xil qiymatlari uchun $SIC(p)$ hisoblanib, minimal qiymat uchun topilgan p uning lagini ifodalaydi. p^ qabul qilishi mumkin qiymatlar 0,1,2 va hokazo.

$SIC(p)$ formulasi g‘aroyib ko‘rinsa-da, uni tushunish oson. Modelga lag qo‘shilganda $\frac{SSR(p)}{T}$ qismi kamayadi, ya’ni oshmaydi, lekin $\frac{(p+1)\ln(T)}{T}$ esa oshadi.

Akaike ma’lumot mezonlari (AIC)

Shvars ma’lumot mezonidan tashqari Akaike ma’lumot mezonidan ham foydalaniladi. Ularning formulalari deyarli bir xil:

$$AIC(p) = \ln\left|\frac{SSR(p)}{T}\right| + \frac{(p+1)\,2}{T}$$

$AIC(p)\ ning\ SIC(p)$ dan farqi faqatgina $SIC(p)$dagi $\ln(T)$ qismi, 2 raqami bilan almashtirilgan. T≥8 uchun (aksariyat holat) ln $(T) > 2$ bo‘lgani uchun , qo‘shimcha lagni modelga qo‘shish uchun AIC orqali tushuntirish osonroq, chunki $SSR(p)$ da kamroq o‘zgarish talab etadi. Shu nuqtai nazardan AIC haqiqiysidan ko‘proq, SIC esa kamroq lagni ishlatishni ko‘rsatishi muqarrar.

Yuqorida keltirilgan ma’lumot me’zonlarini avval hisoblangan modellar misolida ko‘rib chiqsa bo‘ladi. Inflyatsiyani prognozlash-tirilganda, AR(1) va AR(2) modellaridan foydalandik va ular bo‘yicha prognoz mos ravishda 10.49% va 10,29% hisoblandi.

Hisob-langan kriteriy	$INFL_{2016}$	Aksike ma’lumot kriteriysi	Shvars ma’lumot kriteriysi	F-statistika	Manba
AR(1)	10.49	6.295	6.386	23.433	Jadval 5.2
AR(2)	10.29	6.421	6.558	11.021	Jadval 5.4

13.12-rasm. $\boldsymbol{INFL_t = \beta_0 + \beta_1 INFL_{t-1} + \cdots + \beta_p INFL_{t-p} + u_t}$**modelida lag uzunligini aniqlash.**

Bu ikkita modelning prognoz qiymatlaridan AR(1) modelini tanlash mqsadga muvofiq, chunki AR(2) modelidan AR(1) modeliga o‘tishda (1) hisoblangan F-statistika qiymatlari oshmoqda, ya’ni bunga bog‘liq p-qiymat kamaymoqda va (2) Akaike va Shvars ma’lumot mezonlari ham kamaymoqda (13.12-rasm)

13.5. Nostatsionarlik

Dinamik qatorlarda regressiya tahlili muayyn o‘zgaruvchining o‘tkan davrdagi kuzatishlarini ishlatish orqali amalga oshiriladi. Agar kelajak o‘tmish bilan bir xil bo‘lsa,o‘tgan davr kuzatishlari kelajakni prognozlashtirishda foydali bo‘ladi. Agar kelajak o‘tmishdan fundamental ravishda farq qiladigan bo‘lsa, o‘tgan davr kuzatishlari kelajakni prognozlashtirishda foydali bo‘lmaydi. Shu nuqtai nazardan, avtoregressiya va avtoregressiv taqsimlangan lag modellari statsionar o‘zgaruvchilarni modelda ishlatishni taqozo etadi. “O‘tmish kelajkni belgilaydi” g‘oyasi umumiy hisobda statsionarlik tushunchasi bilan umumlashtiriladi.

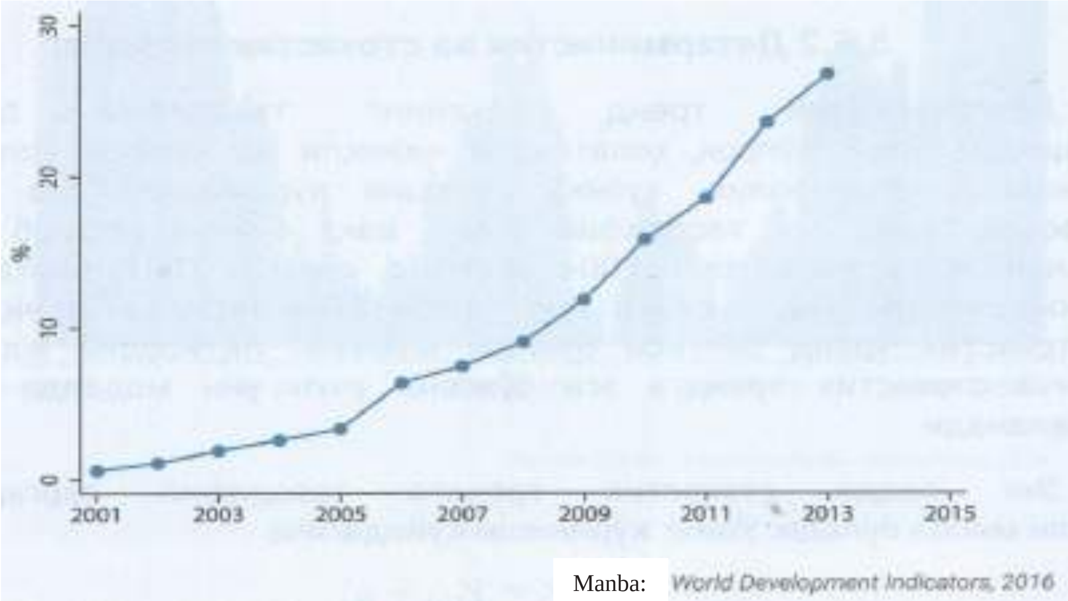
Agar Y_t o‘zgaruvchisining ehtimollik taqsimoti vaqt bo‘yicha o‘zgarmasa, ya’ni $(Y_1, Y_2, \dots Y_T)$hamda $(Y_{s+1}, Y_{s+2}, \dots, Y_{s+T})$ ehtimollik taqsimotlari s=1,2.. uchun bir xil bo‘lsa, y statsionar hisoblanadi.

Statsionarlikni kuchsizroq tariflari ham bor. Agar Y_t o‘zgaruvchisining (1) matematik kutilishi va (2) dispersiyasi vaqt bo‘yicha deyarli o‘zgarmas bo‘lib, (3) $cov(Y_t \ , Y_{t-s})$ s ga bog‘liq bo‘lib, t ga bog‘liq bo‘lmasa, dinamik qator statsionar bo‘ladi.

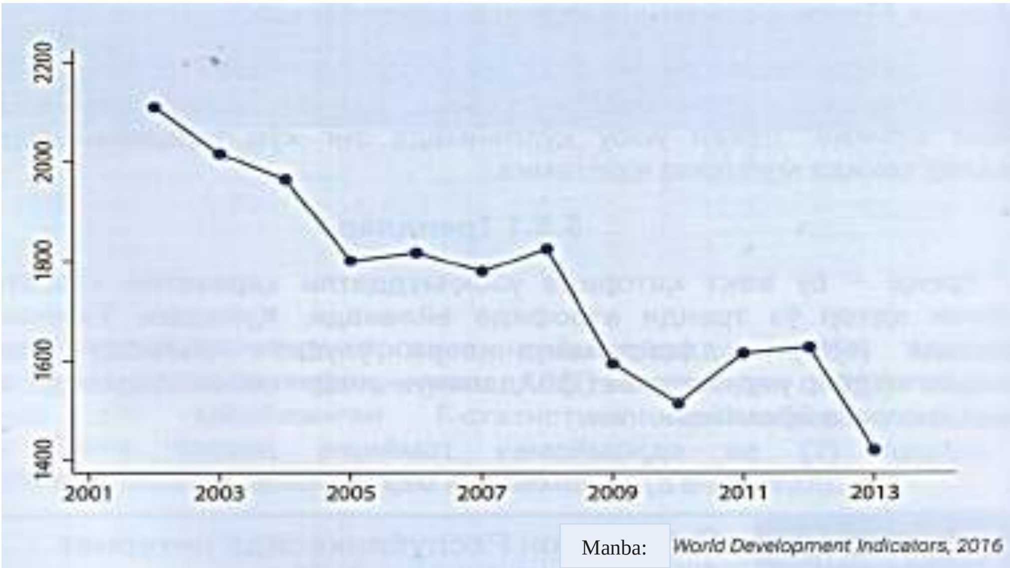
Statsionar bo‘lmagan dinamik qatorlar nostatsionar hisoblanadi. O‘zgaruvchilarda nostatsionarlik bir qancha ko‘rinishlarga ega bo‘lishi mumkin. Lekin ushbu qo‘llanmada eng ko‘p tarqalgan turi–trendlar haqida mulohaza yuritamiz.

13.5.1 Trendlar

Trendlar - bu vaqt qatorida uzoq-muddatli harakatni ko‘rsatadi. Dinamik qator o‘z trendi atrofida aylanadi. Quyidagi O‘zbekiston aholisida internet foydalanuvchilari ulushiga e’tibor bering. Ko‘rilayotgan davr uchun internet foydalanuvchilari keskin oshmoqda, ya’ni trendi yuqoriga yo‘naltirilgan.



13.13-rasm. O‘zbekiston Respublikasida internet foydalanuvchilari, % da



13.14-rasm. O‘zbekiston Respublikasida energiyadan foydalanish (neft ekvivalentidagi kg da, aholi jon boshiga)

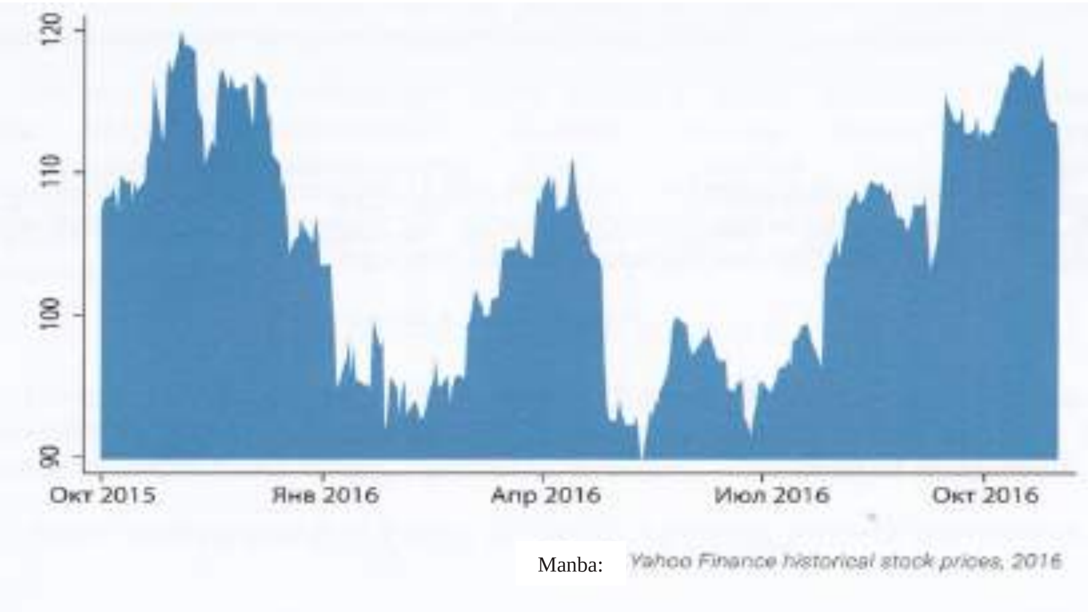
13.5.2 Deterministik va stoxastik trendlar

Deterministik trend vaqtning tasodifiy bo‘lmagan funksiyasi bo‘lib, ko‘proq holatlarda chiziqli va kamroq holatlarda polinomial (parabolik, kubik) funksiya ko‘rinishiga ega bo‘ladi. Stoxastik trend esa tasodifiy bo‘lib, vaqt bo‘yicha o‘zgarib turadi. Masalan, YaIM deflyatori 2004-yilgacha pastga, 2007-yilga qadar esa yuqoriga va undan keyin yana pastga yo‘naltirilgan trendga ega. Bu stoxastik trend belgisi bo‘lishi mumkin. Iqtisodiy jarayonlar ko‘pincha stoxastik trendga ega bo‘lgani uchun uni modelda shunday ifodalanadi.

Eng sodda stoxastik trendga tasodifiy sargardonlik modeli misol bo‘ladi. Uning ko‘rinishi quyidagicha

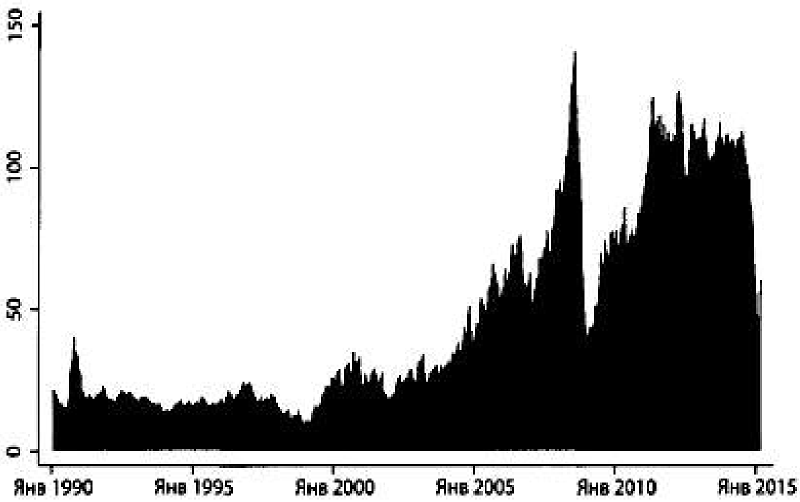
$$Y_t = Y_{t-1} + u_t$$

bunda $E\ (u_t \mid Y_{t,} Y_{t-1}, ..)=0$ va u_t bog‘liqsiz va tekis taqsimlangan tasodifiy miqdor.



13.15-rasm. Apple Inc. (AAPL) aksiyasi narxlari, AQSH dollarida
(dividendlar aksiya narxida aks ettirlgan)

Ushbu model talqini shundan iboratki, $\Delta Y_t = u_t$ bo‘lgani uchun bir muddatdan ikkinchisida o‘zgarishi bog‘liqsiz va tekis taqsimlangan tasodifiy miqdorni tashkil etadi. Shu bois, uning eng yaxshi keyingi muddat prognozi bugun qabul qilingan qiymat hisoblanadi. Tasodifiy sargardonlik jarayoniga bir misol sifatida aksiya narxlarining kunlik harakatini olaylik. Aksiyalar narxi o‘rta hisobda oshib borsa-da, keyingi muddatda qabul qiladigan qiymati bo‘yicha eng yaxshi prognoz bu uning bugungi narxi. Aslida esa, uning qiymati bugungi narx plyus bugungi narx va prognoz qilib bo‘lmaydigan qismdan iborat. Tasodifiy sargardonlik modeliga ikkinchi misol sifatida kunlik neft narxlari hamda aksiya narxlari quyidagi grafiklarda keltirilgan.



Manba: US Energy Information Administration, 2016

13.16 -rasm. Yevropa Brent barrel spot narxlari FOB (AQSH dollarida)

Tasodifiy sargardonlik modeli AR(1) modelining xususiy qismi bo‘lib, u nostatsionar model hisoblanadi. Bu ko‘rilayotgan o‘zgaruvchining dispersiyasi vaqtga bog‘liqligidan kelib chiqadi:

$$Y_t = Y_{t-1} + u_t = Y_{t-2} + u_{t-1} + u_t = \ldots$$
$$= Y_0 + u_1 + u_2 + \cdots + u_t = Y_0 \sum_{t=1}^{T} u_t.$$

Berilgan Y_0 uchun,

$$var(Y_t) = var\ (Y_0) + var\left(\sum_{t=1}^{T} u_t\right) = T\sigma_u^2$$

Ko‘rinib turganidek, dispersiya vaqtga bog‘liq, vaqt oshsa, dispersiya ham oshadi. Shu bois, agar ko‘rilayotgan o‘zgaruvchi tasodifiy sargardonlik bo‘yicha tushuntirilsa, u nostatsionar bo‘ladi.

Umuman olganda, $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$ modelida $|\beta_1|<1$ bo‘lsa, Y_t statsionar, aks holda nostatsionar deb yuritiladi. Tasodifiy sargardonlik modelida β_1=1 bo‘lgani uchun u nostatsionar hisoblanadi.

AR(1) modelida nostatsionar o‘zgaruvchilarda β_1=1 bo‘lgani uchun ularda **birlik ildiz mavjud** yoki **stoxastik trendda** ega deyiladi**.**

13.5.3 Birlik ildiz muammolari

Agar regressorning birlik ildizi mavjud bo‘lsa, EKK usuli bilan hisoblangan parametrlarda bir qancha muammolar tug‘iladi. Birinchidan, hisoblangan parametrlar nol yo‘nalishida siljigan bo‘ladi, ya’ni hisoblangan parametrlar bo‘yicha qabul qilingan qaror noto‘g‘ri bo‘ladi. Ikkinchidan, ularning hisoblangan t– statistikalari normal taqsimotga tanlanma hajmi oshirilganda ham ega bo‘lmaydi, ya’ni ko‘rsatayotgan statistik muhimlik darajasi rost ma’lumotni bo‘lmaydi. Uchinchidan, hisoblangan regressiya “ yolg‘on regressiya” hisoblanadi, ya’ni regressor va regressand orasida kuchli statistik bog‘lanish ko‘rinsa ham, aslida bu bog‘lanish “ yolg‘on” hisoblanadi.

Misol uchun O‘zbekiston jon boshiga YaIM janubiy Koreyaning o‘rtacha umr davomiyligiga qanday ta’sir etadi? Mantiqan, hech qanday! O‘zbekistonning o‘sib borayotgan YaIM O‘zbekiston hududida yashovchi aholining umr davomiyligini oshiradi, Janubiy Koreyada yashovchi aholining emas! Quyidagi regressiya modelini ko‘raylik:

$$\ln(life\_KOR) = \beta_0 + \beta_1 \ln(gdppc\_UZB) + u_t$$

Bunda ln(*life_KOR*)- Janubiy Koreya aholisining o‘rtacha umr davomiyligi (yillarda), *gdppc_UZB* esa O‘zbekiston jon boshiga YaIM, 2000 -yilga AQSH dollarida, t vaqt ko‘rsatkichi – 1987-2014-yillar.

Avval ushbu modelni avval 2000-2014-yillar uchun hisoblaylik.

Dependent Variable: LOG(LIFE_KOR)
Method: Least Squares
Sample (adjusted): 2000 2014
Included observations: 15 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(GDPPC_UZB)	0.097283	0.004307	22.58551	0.0000
C	3.685505	0.030385	121.2951	0.0000

R-squared	0.975148	Mean dependent var	4.371325
Adjusted R-squared	0.973237	S.D. dependent var	0.025549
S.E. of regression	0.004180	Akaike info criterion	-7.993603
Sum squared resid	0.000227	Schwarz criterion	-7.899196
Log likelihood	61.95202	Hannan-Quinn criter.	-7.994609
F-statistic	510.1052	Durbin-Watson stat	0.301281
Prob(F-statistic)	0.000000		

13.17-rasm. ln $(life\_KOR_t = \beta_0 + \beta_1 \ln(gdppc_{UZB}) + u_t$ modeli (yolg‘on model) parametrlarini 2000-2014 muddati uchun hisoblash

Hisoblangan regressiya parametrlari O‘zbekiston jon boshiga YaIM 10% o‘sishi Janubiy Koreya aholisining o‘rtacha umr davomiyligini deyarli 15 oshirmoqda, *ceteris paribus*. Bu hisoblangan parametrning statistik muhimligi 1% darajada, modelning R^2=0.98 va F-statistikasining p -qiymati 1% darajada. Boshqacha qilib aytganda, ko‘rilayotgan ikki dinamik qatorlar orasidagi bog‘lanish kuchli.

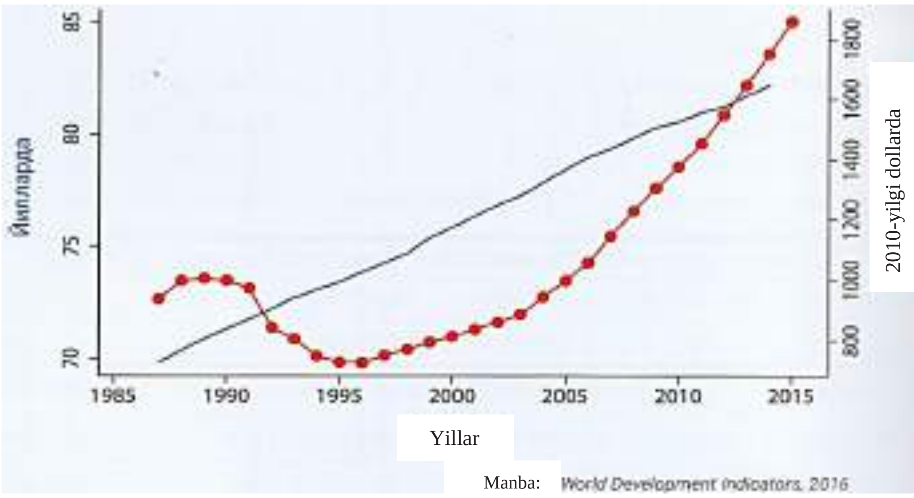
Albatta, modellarda bu kabi ko‘rsatkichlar tadqiqotchi uchun quvonchli bo‘lsa-da, regressiya ko‘rsatayotgan natijalar “yolg‘on”. Agar biz ushbu modelni 1987-2000-yillar uchun hisoblasak, umuman boshqacha natijalar chiqishi ma’lum bo‘ladi.

Ko‘rinib turganidek, regressiya natijalari bu safar shuni ko‘rsatadiki, O‘zbekiston jon boshiga YaIM 10% oshishi Janubiy Koreya aholisining o‘rtacha umr davomiyligini 1.6% foizga qisqartiradi, ceteris paribus. Shu bilan birga, hisoblangan parametr va modelning statistik muhimligi 1% darajada. EKK usulida hisoblaganimizda kuchli statistik bog‘lanish ko‘rinmoqda. Albatta, modellarda bu kabi ko‘rsatkichlar tadqiqotchi uchun quvonchli bo‘lsa-da, regressiya ko‘rsatayotgan natijalar “yolg‘on”. Bunday natijaning sababi ko‘rilayotgan o‘zgaruvchilar birlik ildizlari mavjudligidandir.

Dependent Variable: LLIFE_KOR
Method: Least Squares
Sample (adjusted): 1987 2000
Included observations: 14 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LGDPPC_UZB	-0.157458	0.037727	-4.173652	0.0013
C	5.348267	0.254191	21.04035	0.0000
R-squared	0.592105	ean dependent var		4.287533
Adjusted R-squared	0.558114	S.D. dependent var		0.025696
S.E. of regression	0.017081	Akaike info criterion		-5.170091
Sum squared resid	0.003501	Schwarz criterion		-5.078797
Log likelihood	38.19063	Hannan-Quinn criter.		-5.178542
F-statistic	17.41937	Durbin-Watson stat		0.354557
Prob(F-statistic)	0.001291			

13.18-rasm. $\ln\left(life\_KOR_t = \beta_0 + \beta_1 \ln(gdppc_{UZB}) + u_t\right.$ **modeli (yolg‘on model) parametrlarini 1987-2000 muddati uchun hisoblash**



13.19-rasm. O‘zbekiston Respublikasi jon boshiga YaIM va Janubiy Koreya aholisining o‘rtacha umr davomiyligi

Yuqoridagi grafikka e’tibor bering. Ushbu bo‘limning 2000-2015-yillar uchun regressiya natijalariga muvofiq O‘zbekiston jon boshiga

YaIM Janubiy Koreya aholisining o‘rtacha umr davomiyligini oshirgan bo‘lsa-da, nima uchun undan avvalgi, ayniqsa 1991-yilda sobiq Sovet Ittifoqi tugagandan keyingi davrda, jon boshiga YaIM kamayganida, Janubiy Koreyaning o‘rtacha umr davomiyligi kamaymadi? Aksincha, u oshdi, chunki bularning tushuntiradigan omillari har xil va shu bois, “yolg‘on” regressiyadir (13.19-rasm).

Bu kabi regressiya modellarini O‘zbekiston Respublikasiga tegishli ma’lumotlar asosida hisoblanganda ham shu holat yuz berishi tabiiy. Bunday natijalar yosh iqtisodchilarni quvontirsa-da, regressiya modelida ishlatiladigan o‘zgaruvchilar birlik ildizga ega yoki emasligiga amin bo‘lishimiz kerak. Masalan, investitsiya hajmini YaIMga dinamik qatorda regressiya qilganimizda, odatda, shu kabi holat yuz beradi, chunki ikkala o‘zgaruvchi ham bir vaqtda oshayotganligi bois, statistik muhim bog‘lanish hisoblanadi. Shu sababli regressiya modelida hisoblashdan oldin bu o‘zgaruvchilar birlik ildizga ega yoki ega emasligini tekshirish amaliyoti mavjud.

Birlik ildiz mavjud yoki emasligiga tekshirish

Birlik ildiz mavjudligini grafik usulda inspeksiya qilish mumkin bo‘lsa-da, bunday usulda uni aniqlash qiyin, ayniqsa, agar o‘zgaruvchi o‘rtachasi va dispersiyasi grafikda aniq ko‘rinmasa. Bundan tashqari, dinamik qatorning avtokorrelyatsiyasi koeffitsiyenti o‘rganish orqali ham amalga oshirish mumkin.

Ushbu bo‘limda birlik ildizga tekshirishni Diki-Fuller testi orqali amalga oshiramiz. Birlik ildizga tekshirishni bir qancha formal testlari bo‘lsa-da, Diki-Fuller amaliyotda keng tarqalgan.

AR(1) modeli uchun birlik ildiz mavjudligini aniqlash, umumiy hisobda avtoregressiya modelini EKK usuli yordamida hisoblashdan iborat. Avval muhokama etilganidek, AR(1) modelida $\beta_1 = 1$ bo‘lsa, tasodifiy sargardonlik modeliga, va $|\beta_1| < 1$ bo‘lsa, statsionar o‘zgaruvchilarga kelamiz. Shu nuqtai nazardan birlik ildiz mavjudligiga tekshirish uchun $Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t$ modelida $H_0: \beta_1 = 1$ va $H_1: \beta_1 < 1$ gipotezasini tekshirsak, yetarli bo'ladi. Bunda $H_0: \beta_1 = 1$ rad etilsa, Y_t statsionar o'zgaruvchi bo'ladi, aks holda uning birlik ildizi mavjud bo'ladi.

AR(1) modelini transformatsiya qilib, birlik ildizga tekshirish yo'li ham mavjud. Bunda AR(1) modelini $\Delta Y_t = \beta_0 + \delta\, Y_{t-1} + u_t$, $\delta = \beta_1 - 1$ ko'rinishida ifodalab, $H_0: \delta = 0$ va $H_1: \delta < 0$ gipotezasi tekshiriladi. AR(1) modelini bunday transformatsiyasi modelni statistik va ekonometrik dasturlarda hisoblaganda qulaylik tug'diradi.

AR(p) modeli uchun kengaytirilgan Diki-Fuller (KDF) testi $\Delta Y_t = \beta_0 + \delta\, Y_{t-1} + y_1\, \Delta Y_{t-1} + y_2\, \Delta Y_{t-2} + \cdots + y_p \Delta\, Y_{t-p} + u_t$ modelida $H_0: \delta = 0$ va $H_1: \delta < 0$ ni tekshirishga olib keladi. Bunda nolinchi gipotezaga muvofiq Y_t o'zgaruvchining birlik ildizi bor, muqobil gipotezaga muvofiq esa bu o'zgaruvchi statsionar bo'ladi. Shu bois, agar H_0 rad etilsa, Y_t statsionar o'zgaruvchi bo'ladi, aks holda uning birlik ildizi mavjud bo'ladi.

Ba'zi hollarda o'zgaruvchi chiziqli trend atrofida statsionarlikni namoyon qilsa, t regressiyaga qo'shiladi va bunda KDF testi $\Delta Y_t = \beta_0 + at + \delta\, Y_{t-1} + y_1\, \Delta Y_{t-1} + y_2\, \Delta Y_{t-2} + \cdots + y_p \Delta\, Y_{t-p} + u_t$ modelida $H_0: \delta = 0$ va $H_1: \delta < 0$ ni tekshirishga olib keladi.

KDF statistikasining tanqidiy qiymatlari

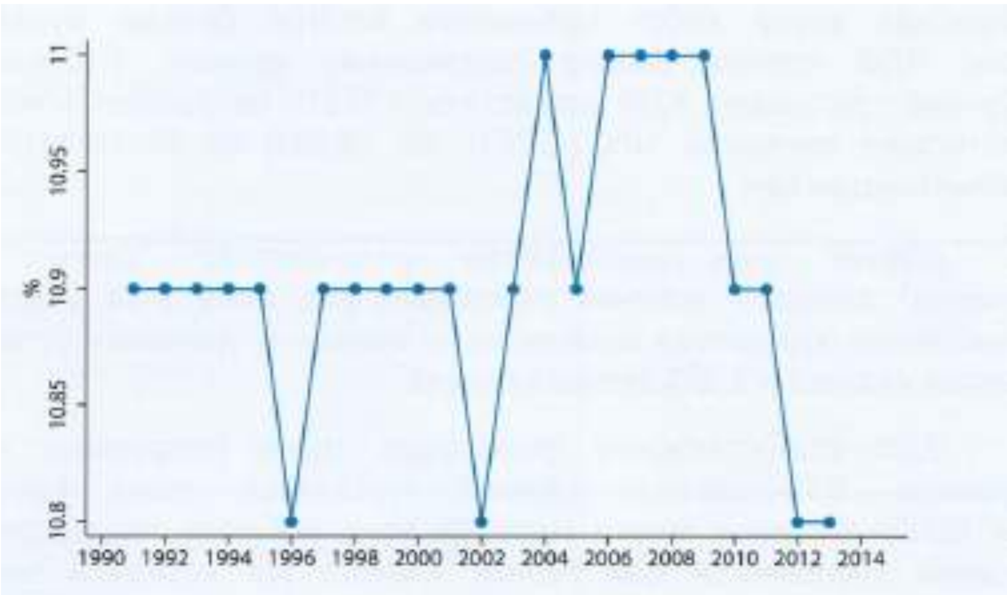
KDF statistikasi qiymati ko'rilgan modellarda hisoblangan t-statistikasi bilan bir xil qiymat qabul qilsa-da, u normal yoki Styudent taqsimoti qonuniga bo'ysunmaydi, tanlama hajmi oshirilganda ham. shuning uchun birlik ildiz mavjudligiga tekshirishda standart hisob-langan p-qiymatlar foydali bo'lmaydi. KDF kritik qiymatlari alohida jadvallar yoki ekonometrik dastur yordamida hisoblanadi.

Deterministik regressorlar	10%	5%	1%
Ozod had	-2.57	-2.86	-3.43
Ozod had va trend	-3.12	-3.41	-3.96

13.20-rasm. KDF tanqidiy qiymatlari

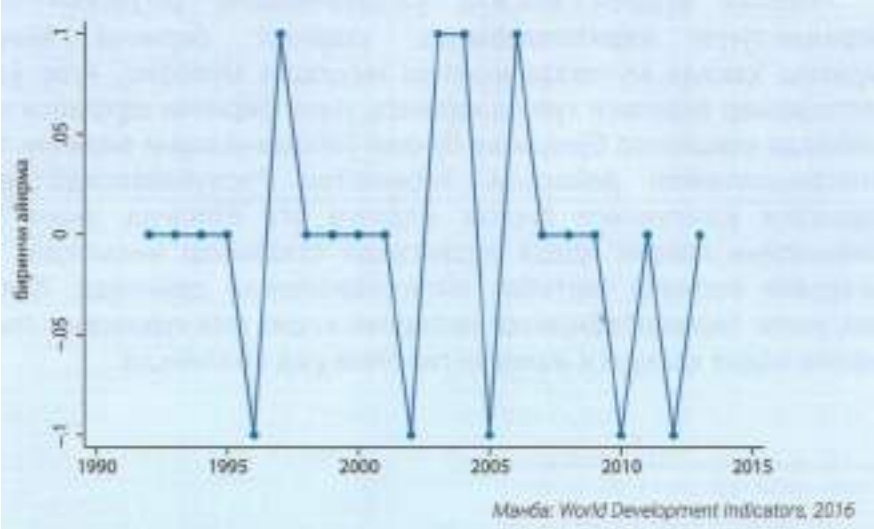
$H_1: \delta < 0$ bo'lgani uchun birlik ildiz mavjudligiga tekshirish gipotezani bir tomonlama tekshirishga olib keladi. Agar modelda, masalan, trend kiritilgan bo'lsa, KDF statistikasi hisoblanganda -2.86 dan kichik bo'lsa, $H_0: \delta = 0$ rad etiladi va dinamik qator statsionar ekanligi kelib chiqadi.

Misol uchun O'zbekistonda ishsizlik darajasini ko'raylik. Ishsizlik darajasi absolyut qiymatda ko'p o'zgarmagan bo'lmasada, ko'rilayotgan 1991-2013-yillar uchn iqtisodiy sikllarga bog'liq ravishda o'zgarib kelgan. 2002-yildan keskin oshib, 2009-yildan keskin kamayishi qayd etilgan (13.21-rasm).



Manba: *World Development Indicators, 2016*

13.21-rasm. O'zbekiston Respublikasida ishsizlik darajasi



13.22-rasm. Ishsizlik darajasining birinchi darajasi

Bu kabi holat oʻzgarayotgan oʻrtachani koʻrsatgani tufayli bu nostatsionar qator ekanligini anglashimiz mumkin. Lekin bu formal test boʻlmagani uchun statsionarlik yoki nostatsionarlik toʻgʻrisida qaror qabul qilishimiz notoʻgʻri boʻlishi mumkin. Shu bois KDF testini amalga oshirishimiz mumkin. 13.23-rasmda koʻrinib turganidek, KDF statistika= 0.315 ni tashkil etmoqda. Bu statistika tanlangan 10% (-3.255), 5% (-3.633) va 1% (-4.441) kritik qiymatlardan kam.

Shuning uchun hisoblangan p-qiymatlar “Birlik ildiz mavjud” haqidagi nolinchi gipotezani rad etish uchun yetarli past emasligini koʻrishimiz mumkin, yaʼni ishsizlik darajasi oʻzgaruvchisi birlik ildizga ega, deb xulosa qilamiz.

KDF statistikasini hisoblashda trend oʻzgaruvchisi modelga qoʻshildi. 13.23-rasmda koʻrinib turganidek, trend oʻzgaruvchisi (@TREND) statistika muhim emas. Shu bois, uni olib tashlab regressiya modeli hisoblangan ham birlik ildizga ega natijasi chiqmoqda, yaʼni ishsizlik darajasini nostatsionar oʻzgaruvchisi. Bu degani ishsizlik darajasini EKK usuli bilan hisoblanadigan modellarda toʻgʻridan-toʻgʻri ishlatish maqsadga muvofiq boʻlmaydi. Shuni ham hisobga olish kerakki, KDF statistikasini hisoblash uchun regressiya modelida regressorlarni lagini tanlash avtomatik tarzda Akaike ma’lumot mezoni asosida amalga oshirilgan.

Birlik ildizi mavjud o‘zgaruvchilarni regressiya modeliga to‘g‘ridan-to‘g‘ri kiritolmasak-da, ularning birinchi ayirmasini kiritish haqida mulohaza yuritish maqsadga muvofiq. Agar o‘zgaruvchi nostatsionar ekanligini xulosa qilinsa, uning birinchi ayirmasi aksariyat hollarda statsionar bo‘ladi va bunday o‘zgaruvchilarni birinchi tartibda integratsiyalangan deyiladi. O‘zbekiston Respublikasida ishsizlik darajasi o‘zgaruvchisi birlik ildizga ega bo‘lsa-da, uning birinchi ayirmasini grafik holda ko‘rilganda statsionar chiqmoqda, ya’ni bu o‘zgaruvchi birinchi ayirmasini birlik ildiz mavjudligini tekshirsak, birlik ildiz haqidagi nolinchi gipoteza rad etilmoqda.

Null Hypothesis: UNEMPL has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on AIC, maxlag=5)

		t-Statistic	Prob.
Augmented Dickey-Fuller test statistic		0.315272	0.9974
Test critical values:	1% level	-4.440739	
	5% level	-3.632896	
	10% level	-3.254671	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(UNEMPL)
Method: Least Squares
Sample (adjusted): 1993 2014
Included observations: 22 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
UNEMPL(-1)	0.096656	0.306611	0.315272	0.7562
D(UNEMPL(-1))	-0.466442	0.308263	-1.513127	0.1476
C	-1.014586	3.337595	-0.303987	0.7646
@TREND("1991")	-0.004421	0.002573	-1.718588	0.1028

R-squared	0.216010	Mean dependent var	-0.013636
Adjusted R-squared	0.085345	S.D. dependent var	0.077432
S.E. of regression	0.074054	Akaike info criterion	-2.205086
Sum squared resid	0.098711	Schwarz criterion	-2.006714
Log likelihood	28.25594	Hannan-Quinn criter.	-2.158355
F-statistic	1.653158	Durbin-Watson stat	1.879953
Prob(F-statistic)	0.212563		

13.23-rasm. O‘zbekiston Respublikasi ishsizlik darajasi o‘zgaruvchisini birlik ildizga tekshirish

Nazorat uchun savollar

1. O‘zbekiston Respublikasi yalpi ichki mahsuloti dinamik qatori gdp_uz.xlsx faylida joylashtirilgan.

a. Ushbu dinamik qatorni vaqtga nisbatan grafik tasvirlang. Bu qatorning o‘rtacha va standart chetlanishi vaqt bo‘yicha o‘zgarmoqdami? Bundan qanday xulosa qilish mumkin?

b. Ushbu dinamik qatorning birinchi ayirmasini hisoblang va uni grafik tasvirlang. Bu qatorning o‘rtachasi va standart chetlanishi vaqt bo‘yicha o‘zgarmoqdami? (a) da berilgan javobga bog‘liq holda bundan qanday xulosa qilish mumkin?

c. Ushbu dinamik qatorning o‘sish sur’atlarini hisoblang.

Bunda uni foizli o‘sish $\%\,\Delta\, GDP = \frac{GDP_t - GDP_{t-1}}{GDP_{t-1}}$ va logarifm ayirmasi $\%\,\Delta\ln(GDP) = \ln(GDP_t) - \ln\,(GDP_{t-1})$ formulalaridan foydalangan holda hisoblang va javoblarni solishtiring. Bundan qanday xulosa qilish mumkin?

d. O‘zbekiston Respublikasi YaIM o‘zgaruvchi sifatida unda birlik ildiz mavjudligini KDF testi yordamida aniqlang. Bunda lag uzunligini tanlash uchun AIC yordamida aniqlang.

ATAMALAR IZOHI

Akaike Information	Akaike ma’lumot kriteriysi Criterion (AIC)
autocorrelation	avtokorrelyatsiya
average	o‘rtacha
Central Limit Theory (CLT)	Markaziy limit teoremasi (MLT)
cross-sectional dataset	kross-seksion ma’lumot
descriptive statistics	tasviriy statistika
deterministic trend	deterministik trend
dummy variable	binar (Bernulli) o‘zgaruvchisi
first difference	birinchi ayirma
heteroskedasticity	xeteroskedastiklik
hypothesistesting	gipotezalarni tekshirish
integrated of order1	birinchi tartibda integratsiya-langan
interval estimation	intervalli baholash
joint distribution	qo‘shma taqsimot
lag	lag
linear functions	chiziqli funksiya
multicollinearity	multikollinearlik
nonstationarity	nostatsionarlik
omitted variable blas	tushirib qoldirilgan o‘zgaruv-chilar natijasida siljish
probability density function (PDF)	ehtimoliy zichlik funksiyasi (EZF)
random error	tasodifiy xatolik
regressand	regressand bog‘liq o‘zgaruvchi
regression	regressiya
regressor	regressor, mustaqil o‘zgaruvchi
Schwarz Information Criterion (SIC)	Shvars ma’lumot kriteriysi

standard deviation	standart chetlanish
standard error	standart xatolik
standard error of regression	regressiya standart xatoligi
stationarity	statsionarlik
statistical inference	statistik xulosa
stochastic trend	tasodifiy trend
structural break	tizimli uzilish
trend	trend
unit root	birlik ildiz
variance	dispersiya
white noise	oq shovqin

FOYDALANILGAN ADABIYOTLAR

1. Anderson, T.W. and H. Rubin (1949) “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” Annals of Mathematical Statistics, 21, pp. 46-63

2. “Some Properties of a Modification of the Limited Information Estimator,” Econometrica, 45, pp.939-953.

3. Principles of Econometrics, 4th Edition by R. Carter Hill, William E. Griffiths and Guay C. Lim (Wiley, 2011)

4. LEE C. ADKINS, R. CARTER HILL. Using Stata For Principles of Econometrics, Fourth Edition.2011. p-625

4. O‘zbekiston Davlat statistika qo‘mitasining 2010-2018-yillardagi “O‘zbekiston Respublikasi yillik statistik to‘plami” . http://www.stat.uz

2. Lee C. Adkins, R. Carter Hill. Using Stata for Principles of Econometrics, Fourth Edition. Wiley, 2011

3. Sh.I. Mustafakulov va boshqalar. Ekonometrika: o‘quv qo‘llanma. – Toshkent 2017, 155 b.

4. A.Ishnazarov, Sh.Nurullayeva, M.Mo‘minova. “Ekonometrika asoslari” fani bo‘yicha o‘quv-uslubiy majmua. Toshkent: Iqtisodiyot, 2017-yil.

5. Introduction to econometrics / James H. Stock, Mark W. Watson. 2nd ed., Boston: Pearson Addison Wesley, 2007.

6. Data analysis using regression and multilevel/hierarchical models / Andrew Gelman, Jennifer Hill. Cambridge; New York : Cambridge University Press, 2007.

7. Gujarati D.N. “Basic Econometrics”. McGraw-Hill, 4th edition, 2003 (Gu), Inc.p. 90.

8. William H. Greene. Econometric analysis, seventh edition. USA, 2012.

9. Stata basics (Stata 13 or 14), Ursina Kuhn and Oliver Lipps, FORS, Lausanne. Version August 2016

10. Sh.I. Mustafakulov. Investitsion jozibadorlikning nazariy, metodologik va amaliy talqini [Matn] /–T.: «Ma'naviyat», 2021, 352b.

11. Q.A. Isayev. Xalqaro iqtisodiyot [Matn]: darslik. T.: «Ma'naviyat», 2021. -456 bet.

12. Sh.I. Mustafakulov va boshqalar. Iqtisodiy atamalarning izohli lug'ati [Matn]: lug'at. T.: Innovatsion rivojlantirish nashriyot-matbaa uyi, 2019,-489b.

13. A.Ishnazarov, Sh.Nurullayeva, S.Xomidov. Ekonometrikaga kirish. Darslik. –Toshkent: Iqtisodiyot, 2021. - 279 bet.

14. R.Alimov, B. Bayxonov, A. Ishnazarov. Ekonometrikadan laboratoriya ishini bajarish bo'yicha xizmat ko'rsatilmoqda. - Toshkent, 2018. – B

15. Gujarati D.N. Basic Econometrics. McGraw-Hill, 4th edition, 2003 (Gu), Inc.p. 7

16. Эконометрика методологияси билан батафсилроқ танишиш учун қуйидагига қаранг. David F. Hendry, Dynamic Econometrics, Oxford University Press, New York, 1995. See also Aris Spanos, op. cit.

17. John Maynard Keynes, The General Theory of Employment, Interest and Money, Harcourt Brace Jovanovich, New York, 1936, p. 96.

18. Shodiyev T.Sh. va boshqalar. Iqtisodiy-matematik usullar va modellar. O'quv qo'llanma. –T.: TDIU, 2010, 8 b.

19. Jamal I.Daoud (2017) Multicollinearity and regression analysis, Journal of Physics: Conference Series

Statistik taqsimot jadvallari

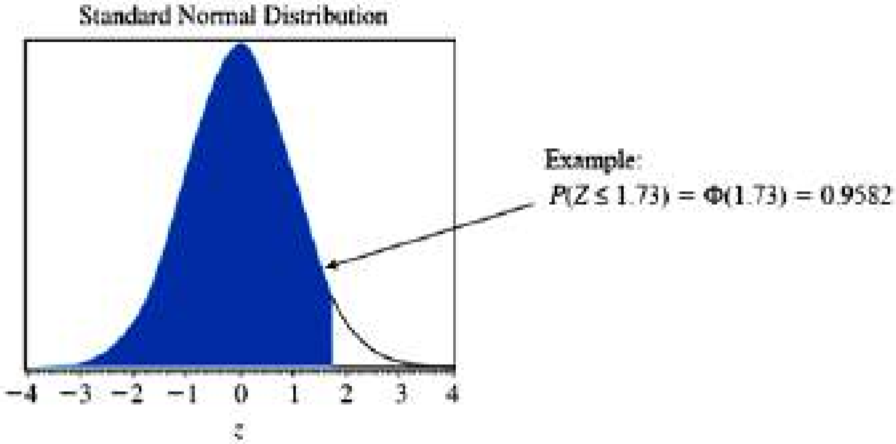


Table 1 **Cumulative Probabilities for the Standard Normal Distribution**
$\Phi(z) = P(Z \le z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Source: This table was generated using the SAS® function PROBNORM.

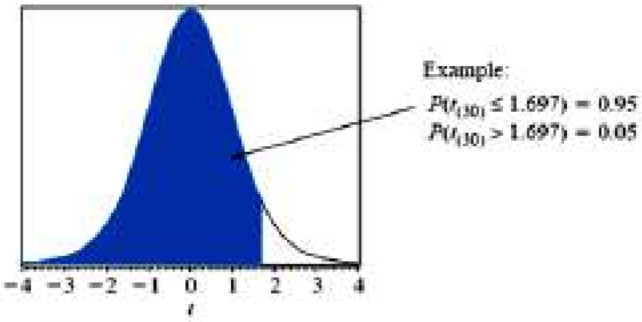
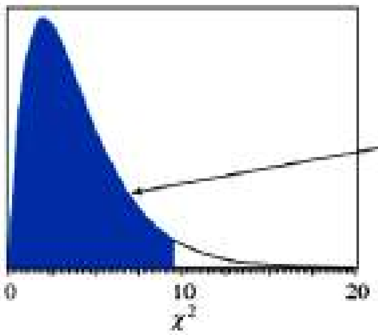


Table 2 **Percentiles of the t-distribution**

df	$t_{(0.90,df)}$	$t_{(0.95,df)}$	$t_{(0.975,df)}$	$t_{(0.99,df)}$	$t_{(0.995,df)}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
31	1.309	1.696	2.040	2.453	2.744
32	1.309	1.694	2.037	2.449	2.738
33	1.308	1.692	2.035	2.445	2.733
34	1.307	1.691	2.032	2.441	2.728
35	1.306	1.690	2.030	2.438	2.724
36	1.306	1.688	2.028	2.434	2.719
37	1.305	1.687	2.026	2.431	2.715
38	1.304	1.686	2.024	2.429	2.712
39	1.304	1.685	2.023	2.426	2.708
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
∞	1.282	1.645	1.960	2.326	2.576

Source: This table was generated using the SAS® function TINV.



Example:
$P(\chi^2_{(4)} \leq 9.488) = 0.95$
$P(\chi^2_{(4)} > 9.488) = 0.05$

Table 3 Percentiles of the Chi-square Distribution

df	$\chi^2_{(0.90,df)}$	$\chi^2_{(0.95,df)}$	$\chi^2_{(0.975,df)}$	$\chi^2_{(0.99,df)}$	$\chi^2_{(0.995,df)}$
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28.300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.144	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.289	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.652	40.646	44.314	46.928
26	35.563	38.885	41.923	45.642	48.290
27	36.741	40.113	43.195	46.963	49.645
28	37.916	41.337	44.461	48.278	50.993
29	39.087	42.557	45.722	49.588	52.336
30	40.256	43.773	46.979	50.892	53.672
35	46.059	49.802	53.203	57.342	60.275
40	51.805	55.758	59.342	63.691	66.766
50	63.167	67.505	71.420	76.154	79.490
60	74.397	79.082	83.298	88.379	91.952
70	85.527	90.531	95.023	100.425	104.215
80	96.578	101.879	106.629	112.329	116.321
90	107.565	113.145	118.136	124.116	128.299
100	118.498	124.342	129.561	135.807	140.169
110	129.385	135.480	140.917	147.414	151.948
120	140.233	146.567	152.211	158.950	163.648

Source: This table was generated using the SAS® function CINV.

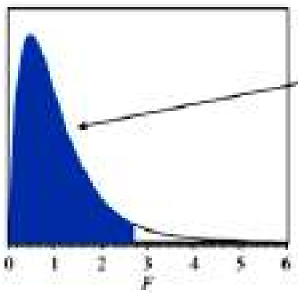


Table 4 **95th Percentile for the F-distribution**

ν_2/ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	250.10	252.20	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.46	19.48	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.62	8.57	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.69	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.43	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.74	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.38	3.30	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	3.01	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.79	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.70	2.62	2.54
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.16	2.07
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.95	1.84
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.92	1.82	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.74	1.62
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.04	1.96	1.88	1.79	1.68	1.56
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.74	1.64	1.51
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	1.97	1.89	1.81	1.71	1.60	1.47
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.87	1.78	1.69	1.58	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.65	1.53	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.55	1.43	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.32	1.00

Source: This table was generated using the SAS® function FINV.

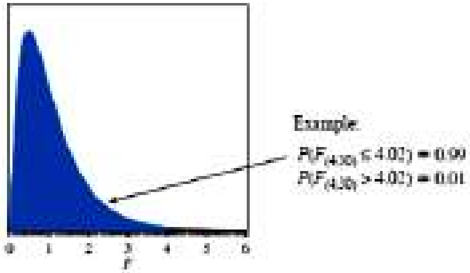


Table 5 99th Percentile for the F-distribution

ν_2/ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	30	60	∞
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6157.28	6208.73	6260.65	6313.03	6365.87
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.47	99.48	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.50	26.32	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.84	13.65	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.38	9.20	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.23	7.06	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	5.99	5.82	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.20	5.03	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.65	4.48	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.25	4.08	3.91
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.21	3.05	2.87
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.78	2.61	2.42
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.54	2.36	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.39	2.21	2.01
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88	2.74	2.60	2.44	2.28	2.10	1.89
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.20	2.02	1.80
45	7.23	5.11	4.25	3.77	3.45	3.23	3.07	2.94	2.83	2.74	2.61	2.46	2.31	2.14	1.96	1.74
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.56	2.42	2.27	2.10	1.91	1.68
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.03	1.84	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.86	1.66	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.70	1.47	1.00

Source: This table was generated using the SAS® function FINV.

TABLE D.5A Durbin–Watson d Statistic: Significance Points of d_L and d_U at 0.05 Level of Significance

	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$		$k'=6$		$k'=7$		$k'=8$		$k'=9$		$k'=10$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
6	0.610	1.400	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7	0.700	1.356	0.467	1.896	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8	0.763	1.332	0.559	1.777	0.368	2.287	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588	—	—	—	—	—	—	—	—	—	—	—	—
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822	—	—	—	—	—	—	—	—	—	—
11	0.927	1.324	0.658	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005	—	—	—	—	—	—	—	—
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149	—	—	—	—	—	—
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266	—	—	—	—
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360	—	—
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470
28	1.328	1.476	1.255	1.560	1.181	1.650	1.104	1.747	1.028	1.850	0.951	1.958	0.874	2.071	0.798	2.188	0.723	2.309	0.650	2.431
29	1.341	1.483	1.270	1.563	1.198	1.650	1.124	1.743	1.050	1.841	0.975	1.944	0.900	2.052	0.826	2.164	0.753	2.278	0.682	2.396
30	1.352	1.489	1.284	1.567	1.214	1.650	1.143	1.739	1.071	1.833	0.998	1.931	0.926	2.034	0.854	2.141	0.782	2.251	0.712	2.363
31	1.363	1.496	1.297	1.570	1.229	1.650	1.160	1.735	1.090	1.825	1.020	1.920	0.950	2.018	0.879	2.120	0.810	2.226	0.741	2.333
32	1.373	1.502	1.309	1.574	1.244	1.650	1.177	1.732	1.109	1.819	1.041	1.909	0.972	2.004	0.904	2.102	0.836	2.203	0.769	2.306
33	1.383	1.508	1.321	1.577	1.258	1.651	1.193	1.730	1.127	1.813	1.061	1.900	0.994	1.991	0.927	2.085	0.861	2.181	0.795	2.281
34	1.393	1.514	1.333	1.580	1.271	1.652	1.208	1.728	1.144	1.808	1.080	1.891	1.015	1.979	0.950	2.069	0.885	2.162	0.821	2.257
35	1.402	1.519	1.343	1.584	1.283	1.653	1.222	1.726	1.160	1.803	1.097	1.884	1.034	1.967	0.971	2.054	0.908	2.144	0.845	2.236
36	1.411	1.525	1.354	1.587	1.295	1.654	1.236	1.724	1.175	1.799	1.114	1.877	1.053	1.957	0.991	2.041	0.930	2.127	0.868	2.216
37	1.419	1.530	1.364	1.590	1.307	1.655	1.249	1.723	1.190	1.795	1.131	1.870	1.071	1.948	1.011	2.029	0.951	2.112	0.891	2.198
38	1.427	1.535	1.373	1.594	1.318	1.656	1.261	1.722	1.204	1.792	1.146	1.864	1.088	1.939	1.029	2.017	0.970	2.098	0.912	2.180
39	1.435	1.540	1.382	1.597	1.328	1.658	1.273	1.722	1.218	1.789	1.161	1.859	1.104	1.932	1.047	2.007	0.990	2.085	0.932	2.164
40	1.442	1.544	1.391	1.600	1.338	1.659	1.285	1.721	1.230	1.786	1.175	1.854	1.120	1.924	1.064	1.997	1.008	2.072	0.952	2.149
45	1.475	1.566	1.430	1.615	1.383	1.666	1.336	1.720	1.287	1.776	1.238	1.835	1.189	1.895	1.139	1.958	1.089	2.022	1.038	2.088
50	1.503	1.585	1.462	1.628	1.421	1.674	1.378	1.721	1.335	1.771	1.291	1.822	1.246	1.875	1.201	1.930	1.156	1.986	1.110	2.044
55	1.528	1.601	1.490	1.641	1.452	1.681	1.414	1.724	1.374	1.768	1.334	1.814	1.294	1.861	1.253	1.909	1.212	1.959	1.170	2.010
60	1.549	1.616	1.514	1.652	1.480	1.689	1.444	1.727	1.408	1.767	1.372	1.808	1.335	1.850	1.298	1.894	1.260	1.939	1.222	1.984
65	1.567	1.629	1.536	1.662	1.503	1.696	1.471	1.731	1.438	1.767	1.404	1.805	1.370	1.843	1.336	1.882	1.301	1.923	1.266	1.964
70	1.583	1.641	1.554	1.672	1.525	1.703	1.494	1.735	1.464	1.768	1.433	1.802	1.401	1.837	1.369	1.873	1.337	1.910	1.305	1.948
75	1.598	1.652	1.571	1.680	1.543	1.709	1.515	1.739	1.487	1.770	1.458	1.801	1.428	1.834	1.399	1.867	1.369	1.901	1.339	1.935
80	1.611	1.662	1.586	1.688	1.560	1.715	1.534	1.743	1.507	1.772	1.480	1.801	1.453	1.831	1.425	1.861	1.397	1.893	1.369	1.925
85	1.624	1.671	1.600	1.696	1.575	1.721	1.550	1.747	1.525	1.774	1.500	1.801	1.474	1.829	1.448	1.857	1.422	1.886	1.396	1.916
90	1.635	1.679	1.612	1.703	1.589	1.726	1.566	1.751	1.542	1.776	1.518	1.801	1.494	1.827	1.469	1.854	1.445	1.881	1.420	1.909
95	1.645	1.687	1.623	1.709	1.602	1.732	1.579	1.755	1.557	1.778	1.535	1.802	1.512	1.827	1.489	1.852	1.465	1.877	1.442	1.903
100	1.654	1.694	1.634	1.715	1.613	1.736	1.592	1.758	1.571	1.780	1.550	1.803	1.528	1.826	1.506	1.850	1.484	1.874	1.462	1.898
150	1.720	1.746	1.706	1.760	1.693	1.774	1.679	1.788	1.665	1.802	1.651	1.817	1.637	1.832	1.622	1.847	1.608	1.862	1.594	1.877
200	1.758	1.778	1.748	1.789	1.738	1.799	1.728	1.810	1.718	1.820	1.707	1.831	1.697	1.841	1.686	1.852	1.675	1.863	1.665	1.874

Ilovalar

The Rules of Summation

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \cdots + x_n$$

$$\sum_{i=1}^{n} a = na$$

$$\sum_{i=1}^{n} ax_i = a\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} (ax_i + by_i) = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} (a + bx_i) = na + b\sum_{i=1}^{n} x_i$$

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

$$\sum_{i=1}^{n} (x_i - \overline{x}) = 0$$

$$\sum_{i=1}^{2}\sum_{j=1}^{3} f(x_i, y_j) = \sum_{i=1}^{2} [f(x_i, y_1) + f(x_i, y_2) + f(x_i, y_3)]$$
$$= f(x_1, y_1) + f(x_1, y_2) + f(x_1, y_3) + f(x_2, y_1) + f(x_2, y_2) + f(x_2, y_3)$$

Expected Values & Variances

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \cdots + x_n f(x_n)$$
$$= \sum_{i=1}^{n} x_i f(x_i) = \sum_x x f(x)$$

$$E[g(X)] = \sum_x g(x) f(x)$$

$$E[g_1(X) + g_2(X)] = \sum_x [g_1(x) + g_2(x)] f(x)$$
$$= \sum_x g_1(x) f(x) + \sum_x g_2(x) f(x)$$
$$= E[g_1(X)] + E[g_2(X)]$$

$E(c) = c$
$E(cX) = cE(X)$
$E(a + cX) = a + cE(X)$
$\text{var}(X) = \sigma^2 = E[X - E(X)]^2 = E(X^2) - [E(X)]^2$
$\text{var}(a + cX) = E[(a + cX) - E(a + cX)]^2 = c^2\text{var}(X)$

Marginal and Conditional Distributions

$$f(x) = \sum_y f(x, y) \quad \text{for each value } X \text{ can take}$$

$$f(y) = \sum_x f(x, y) \quad \text{for each value } Y \text{ can take}$$

$$f(x|y) = P[X = x | Y = y] = \frac{f(x, y)}{f(y)}$$

If X and Y are independent random variables, then $f(x,y) = f(x)f(y)$ for each and every pair of values x and y. The converse is also true.

If X and Y are independent random variables, then the conditional probability density function of X given that

$$Y = y \text{ is } f(x|y) = \frac{f(x, y)}{f(y)} = \frac{f(x)f(y)}{f(y)} = f(x)$$

for each and every pair of values x and y. The converse is also true.

Expectations, Variances & Covariances

$$\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$$
$$= \sum_x \sum_y [x - E(X)][y - E(Y)] f(x, y)$$

$$\rho = \frac{\text{cov}(X,Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

$E(c_1X + c_2Y) = c_1E(X) + c_2E(Y)$
$E(X + Y) = E(X) + E(Y)$
$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z) + 2ab\text{cov}(X,Y) + 2ac\text{cov}(X,Z) + 2bc\text{cov}(Y,Z)$

If X, Y, and Z are independent, or uncorrelated, random variables, then the covariance terms are zero and:

$$\text{var}(aX + bY + cZ) = a^2\text{var}(X) + b^2\text{var}(Y) + c^2\text{var}(Z)$$

Normal Probabilities

If $X \sim N(\mu, \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$

If $X \sim N(\mu, \sigma^2)$ and a is a constant, then

$$P(X \geq a) = P\left(Z \geq \frac{a - \mu}{\sigma}\right)$$

If $X \sim N(\mu, \sigma^2)$ and a and b are constants, then

$$P(a \leq X \leq b) = P\left(\frac{a - \mu}{\sigma} \leq Z \leq \frac{b - \mu}{\sigma}\right)$$

Assumptions of the Simple Linear Regression Model

SR1 The value of y, for each value of x, is $y = \beta_1 + \beta_2 x + e$

SR2 The average value of the random error e is $E(e) = 0$ since we assume that $E(y) = \beta_1 + \beta_2 x$

SR3 The variance of the random error e is $\text{var}(e) = \sigma^2 = \text{var}(y)$

SR4 The covariance between any pair of random errors, e_i and e_j is $\text{cov}(e_i, e_j) = \text{cov}(y_i, y_j) = 0$

SR5 The variable x is not random and must take at least two different values.

SR6 (*optional*) The values of e are *normally distributed* about their mean $e \sim N(0, \sigma^2)$

Least Squares Estimation

If b_1 and b_2 are the least squares estimates, then

$$\hat{y}_i = b_1 + b_2 x_i$$
$$\hat{e}_i = y_i - \hat{y}_i = y_i - b_1 - b_2 x_i$$

The Normal Equations

$$Nb_1 + \Sigma x_i b_2 = \Sigma y_i$$
$$\Sigma x_i b_1 + \Sigma x_i^2 b_2 = \Sigma x_i y_i$$

Least Squares Estimators

$$b_2 = \frac{\Sigma(x_i - \overline{x})(y_i - \overline{y})}{\Sigma(x_i - \overline{x})^2}$$
$$b_1 = \overline{y} - b_2\overline{x}$$

Elasticity

$$\eta = \frac{\text{percentage change in } y}{\text{percentage change in } x} = \frac{\Delta y/y}{\Delta x/x} = \frac{\Delta y}{\Delta x}\cdot\frac{x}{y}$$

$$\eta = \frac{\Delta E(y)/E(y)}{\Delta x/x} = \frac{\Delta E(y)}{\Delta x}\cdot\frac{x}{E(y)} = \beta_2\cdot\frac{x}{E(y)}$$

Least Squares Expressions Useful for Theory

$b_2 = \beta_2 + \Sigma w_i e_i$

$$w_i = \frac{x_i - \bar{x}}{\Sigma(x_i - \bar{x})^2}$$

$\Sigma w_i = 0, \quad \Sigma w_i x_i = 1, \quad \Sigma w_i^2 = 1/\Sigma(x_i - \bar{x})^2$

Properties of the Least Squares Estimators

$$\text{var}(b_1) = \sigma^2\left[\frac{\Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2}\right] \quad \text{var}(b_2) = \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}$$

$$\text{cov}(b_1, b_2) = \sigma^2\left[\frac{-\bar{x}}{\Sigma(x_i - \bar{x})^2}\right]$$

Gauss-Markov Theorem: Under the assumptions SR1–SR5 of the linear regression model the estimators b_1 and b_2 have the *smallest variance of all linear and unbiased estimators* of β_1 and β_2. They are the Best Linear Unbiased Estimators (BLUE) of β_1 and β_2.

If we make the normality assumption, assumption SR6, about the error term, then the least squares estimators are normally distributed.

$$b_1 \sim N\left(\beta_1, \frac{\sigma^2\,\Sigma x_i^2}{N\Sigma(x_i - \bar{x})^2}\right), b_2 \sim N\left(\beta_2, \frac{\sigma^2}{\Sigma(x_i - \bar{x})^2}\right)$$

Estimated Error Variance

$$\hat{\sigma}^2 = \frac{\Sigma\hat{e}_i^2}{N-2}$$

Estimator Standard Errors

$\text{se}(b_1) = \sqrt{\widehat{\text{var}}(b_1)}, \quad \text{se}(b_2) = \sqrt{\widehat{\text{var}}(b_2)}$

***t*-distribution**

If assumptions SR1–SR6 of the simple linear regression model hold, then

$$t = \frac{b_k - \beta_k}{\text{se}(b_k)} \sim t_{(N-2)}, \quad k = 1, 2$$

Interval Estimates

$P[b_2 - t_c\text{se}(b_2) \le \beta_2 \le b_2 + t_c\text{se}(b_2)] = 1 - \alpha$

Hypothesis Testing

Components of Hypothesis Tests

1. A *null* hypothesis, H_0
2. An *alternative* hypothesis, H_1
3. A test *statistic*
4. A *rejection* region
5. A conclusion

If the null hypothesis $H_0: \beta_2 = c$ is ***true***, then

$$t = \frac{b_2 - c}{\text{se}(b_2)} \sim t_{(N-2)}$$

Rejection rule for a two-tail test: If the value of the test statistic falls in the rejection region, either tail of the *t*-distribution, then we reject the null hypothesis and accept the alternative.

Type I error: The null hypothesis is *true* and we decide to *reject* it.

Type II error: The null hypothesis is *false* and we decide *not* to reject it.

***p*-value rejection rule:** When the *p*-value of a hypothesis test is *smaller* than the chosen value of α, then the test procedure leads to *rejection* of the null hypothesis.

Prediction

$y_0 = \beta_1 + \beta_2 x_0 + e_0, \; \hat{y}_0 = b_1 + b_2 x_0, \; f = \hat{y}_0 - y_0$

$$\widehat{\text{var}(f)} = \hat{\sigma}^2\left[1 + \frac{1}{N} + \frac{(x_0 - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}\right], \text{se}(f) = \sqrt{\widehat{\text{var}(f)}}$$

A $(1-\alpha)\times 100\%$ confidence interval, or prediction interval, for y_0

$$\hat{y}_0 \pm t_c\text{se}(f)$$

Goodness of Fit

$\Sigma(y_i - \bar{y})^2 = \Sigma(\hat{y}_i - \bar{y})^2 + \Sigma\hat{e}_i^2$

$SST = SSR + SSE$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} = (\text{corr}(y, \hat{y}))^2$$

Log-Linear Model

$\ln(y) = \beta_1 + \beta_2 x + e, \quad \widehat{\ln(y)} = b_1 + b_2 x$

$100 \times \beta_2 \approx$ % change in y given a one-unit change in x.

$\hat{y}_n = \exp(b_1 + b_2 x)$

$\hat{y}_c = \exp(b_1 + b_2 x)\exp(\hat{\sigma}^2/2)$

Prediction interval:

$$\exp\left[\widehat{\ln(y)} - t_c\text{se}(f)\right], \; \exp\left[\widehat{\ln(y)} + t_c\text{se}(f)\right]$$

Generalized goodness-of-fit measure $R_g^2 = (\text{corr}(y, \hat{y}_n))^2$

Assumptions of the Multiple Regression Model

MR1 $\quad y_i = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + e_i$

MR2 $\quad E(y_i) = \beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \Leftrightarrow E(e_i) = 0.$

MR3 $\quad \text{var}(y_i) = \text{var}(e_i) = \sigma^2$

MR4 $\quad \text{cov}(y_i, y_j) = \text{cov}(e_i, e_j) = 0$

MR5 The values of x_{ik} are not random and are not exact linear functions of the other explanatory variables.

MR6 $\quad y_i \sim N[(\beta_1 + \beta_2 x_{i2} + \cdots + \beta_K x_{iK}), \sigma^2] \Leftrightarrow e_i \sim N(0, \sigma^2)$

Least Squares Estimates in MR Model

Least squares estimates $b_1, b_2, \ldots, b_K$ minimize
$S(\beta_1, \beta_2, \ldots, \beta_K) = \Sigma(y_i - \beta_1 - \beta_2 x_{i2} - \cdots - \beta_K x_{iK})^2$

Estimated Error Variance and Estimator Standard Errors

$$\hat{\sigma}^2 = \frac{\Sigma\hat{e}_i^2}{N-K} \quad \text{se}(b_k) = \sqrt{\widehat{\text{var}}(b_k)}$$

Hypothesis Tests and Interval Estimates for Single Parameters

Use t-distribution $\quad t = \dfrac{b_k - \beta_k}{\mathrm{se}(b_k)} \sim t_{(N-K)}$

t-test for More than One Parameter

$$H_0: \beta_2 + c\beta_3 = a$$

When H_0 is true $\quad t = \dfrac{b_2 + cb_3 - a}{\mathrm{se}(b_2 + cb_3)} \sim t_{(N-K)}$

$$\mathrm{se}(b_2 + cb_3) = \sqrt{\widehat{\mathrm{var}(b_2)} + c^2\widehat{\mathrm{var}(b_3)} + 2c \times \widehat{\mathrm{cov}(b_2, b_3)}}$$

Joint F-tests

To test J joint hypotheses,

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

To test the overall significance of the model the null and alternative hypotheses and F statistic are

$$H_0: \beta_2 = 0,\ \beta_3 = 0, \ldots, \beta_K = 0$$
$$H_1: \textit{at least one} \text{ of the } \beta_k \text{ is nonzero}$$
$$F = \frac{(SST - SSE)/(K-1)}{SSE/(N-K)}$$

RESET: A Specification Test

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i \qquad \hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3}$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + e_i, \qquad H_0: \gamma_1 = 0$

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + e_i, \qquad H_0: \gamma_1 = \gamma_2 = 0$

Model Selection

$AIC = \ln(SSE/N) + 2K/N$

$SC = \ln(SSE/N) + K\ln(N)/N$

Collinearity and Omitted Variables

$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + e_i$

$$\mathrm{var}(b_2) = \frac{\sigma^2}{(1 - r_{23}^2)\sum(x_{i2} - \bar{x}_2)^2}$$

When x_3 is omitted, $\mathrm{bias}(b_2^*) = E(b_2^*) - \beta_2 = \beta_3 \dfrac{\widehat{\mathrm{cov}(x_2, x_3)}}{\widehat{\mathrm{var}(x_2)}}$

Heteroskedasticity

$\mathrm{var}(y_i) = \mathrm{var}(e_i) = \sigma_i^2$

General variance function

$\sigma_i^2 = \exp(\alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS})$

Breusch-Pagan and White Tests for H_0: $\alpha_2 = \alpha_3 = \cdots = \alpha_S = 0$

When H_0 is true $\quad \chi^2 = N \times R^2 \sim \chi^2_{(S-1)}$

Goldfeld-Quandt test for $H_0: \sigma_M^2 = \sigma_R^2$ versus $H_1: \sigma_M^2 \neq \sigma_R^2$

When H_0 is true $\quad F = \hat{\sigma}_M^2 / \hat{\sigma}_R^2 \sim F_{(N_M - K_M, N_R - K_R)}$

Transformed model for $\mathrm{var}(e_i) = \sigma_i^2 = \sigma^2 x_i$

$y_i/\sqrt{x_i} = \beta_1(1/\sqrt{x_i}) + \beta_2(x_i/\sqrt{x_i}) + e_i/\sqrt{x_i}$

Estimating the variance function

$\ln(\hat{e}_i^2) = \ln(\sigma_i^2) + v_i = \alpha_1 + \alpha_2 z_{i2} + \cdots + \alpha_S z_{iS} + v_i$

Grouped data

$$\mathrm{var}(e_i) = \sigma_i^2 = \begin{cases} \sigma_M^2 & i = 1, 2, \ldots, N_M \\ \sigma_R^2 & i = 1, 2, \ldots, N_R \end{cases}$$

Transformed model for feasible generalized least squares

$$y_i \Big/ \sqrt{\hat{\sigma}_i} = \beta_1\left(1 \Big/ \sqrt{\hat{\sigma}_i}\right) + \beta_2\left(x_i \Big/ \sqrt{\hat{\sigma}_i}\right) + e_i \Big/ \sqrt{\hat{\sigma}_i}$$

Regression with Stationary Time Series Variables

Finite distributed lag model

$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + v_t$

Correlogram

$r_k = \sum(y_t - \bar{y})(y_{t-k} - \bar{y}) / \sum(y_t - \bar{y})^2$

For $H_0: \rho_k = 0, \qquad z = \sqrt{T} r_k \sim N(0, 1)$

LM test

$y_t = \beta_1 + \beta_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t \qquad$ Test $H_0: \rho = 0$ with t-test

$\hat{e}_t = \gamma_1 + \gamma_2 x_t + \rho \hat{e}_{t-1} + \hat{v}_t \qquad$ Test using $LM = T \times R^2$

AR(1) error $\quad y_t = \beta_1 + \beta_2 x_t + e_t \qquad e_t = \rho e_{t-1} + v_t$

Nonlinear least squares estimation

$$y_t = \beta_1(1 - \rho) + \beta_2 x_t + \rho y_{t-1} - \beta_2 \rho x_{t-1} + v_t$$

ARDL(p, q) model

$$y_t = \delta + \delta_0 x_t + \delta_1 x_{t-1} + \cdots + \delta_q x_{t-q} + \theta_1 y_{t-1} + \cdots + \theta_p y_{t-p} + v_t$$

AR(p) forecasting model

$$y_t = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + v_t$$

Exponential smoothing $\quad \hat{y}_t = \alpha y_{t-1} + (1 - \alpha)\hat{y}_{t-1}$

Multiplier analysis

$$\delta_0 + \delta_1 L + \delta_2 L^2 + \cdots + \delta_q L^q = (1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_p L^p) \times (\beta_0 + \beta_1 L + \beta_2 L^2 + \cdots)$$

Unit Roots and Cointegration

Unit Root Test for Stationarity: Null hypothesis:

$H_0: \gamma = 0$

Dickey-Fuller Test 1 (no constant and no trend):

$\Delta y_t = \gamma y_{t-1} + v_t$

Dickey-Fuller Test 2 (with constant but no trend):

$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$

Dickey-Fuller Test 3 (with constant and with trend):

$\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + v_t$

Augmented Dickey-Fuller Tests:

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{s=1}^{m} a_s \Delta y_{t-s} + v_t$$

Test for cointegration

$\Delta \hat{e}_t = \gamma \hat{e}_{t-1} + v_t$

Random walk: $\quad y_t = y_{t-1} + v_t$

Random walk with drift: $\quad y_t = \alpha + y_{t-1} + v_t$

Random walk model with drift and time trend:

$y_t = \alpha + \delta t + y_{t-1} + v_t$

Panel Data

Pooled least squares regression

$y_{it} = \beta_1 + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it}$

Cluster robust standard errors $\quad \mathrm{cov}(e_{it}, e_{is}) = \psi_{ts}$

Fixed effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{1i}$ not random

$y_{it} - \bar{y}_i = \beta_2(x_{2it} - \bar{x}_{2i}) + \beta_3(x_{3it} - \bar{x}_{3i}) + (e_{it} - \bar{e}_i)$

Random effects model

$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it} \qquad \beta_{it} = \bar{\beta}_1 + u_i$ random

$y_{it} - \alpha\bar{y}_i = \bar{\beta}_1(1 - \alpha) + \beta_2(x_{2it} - \alpha\bar{x}_{2i}) + \beta_3(x_{3it} - \alpha\bar{x}_{3i}) + v_{it}^*$

$$\alpha = 1 - \sigma_e \Big/ \sqrt{T\sigma_u^2 + \sigma_e^2}$$

Hausman test

$$t = (b_{FE,k} - b_{RE,k}) \Big/ \left[\widehat{\mathrm{var}(b_{FE,k})} - \widehat{\mathrm{var}(b_{RE,k})}\right]^{1/2}$$

MUNDARIJA

IX BOB. ENG KICHIK KVADRATLAR USULI VA GAUSS-MARKOV TEOREMASI

X BOB. INTERVALNI BAHOLASH VA ISHONCH INTERVALI

XI BOB. GIPOTEZA TUSHUNCHASI. GIPOTEZALARNI TEKSHIRISH JARAYONI

XII BOB. KO‘P OMILLI REGRESSIYA MODELI

XIII BOB. DINAMIK QATORLAR REGRESSIYASI

MUSTAFAKULOV SHERZOD IGAMBERDIYEVICH,
SABIROV HASAN NUSRATOVICH

EKONOMETRIKA I

Toshkent – «Ilm-fan va innovatsiya» – 2022

Muharrir: S.Abdunabiyeva
Tex. muharrir: M.Tursunov
Musavvir: A.Shushunov
Musahhih: L.Ibragimov
Kompyuterda sahifalovchi: M.Zoyirova

Nashriyot litsenziyasi 4913-3160-38f4-4927-7ad7-2416-2546, 02.06.2021.

Bosishga ruxsat etildi 25.05. 2022.
Bichimi 60x84 1/16. «Timez Uz» garniturasi.
Ofset bosma usulida bosildi.
Shartli bosma tabog‘i: 15,0. Nashriyot bosma tabog‘i 14,5.
Tiraji: 100. Buyurtma № 26

«Ilm-fan va innovatsiya» bosmaxonasida chop etildi.
100174, Toshkent sh, Olmazor tumani, Usta Shirin 116/3 - uy.