XALQARO NORDIK UNIVERSITETI

Iqtisodiyot va pedagogika fakulteti,
Iqtisodiyot va biznesni boshqarish kafedrasi

Fan o‘qituvchisi: Sabirov Xasan Nusratovich

Mavzu: Normal taqsimot

Reja:

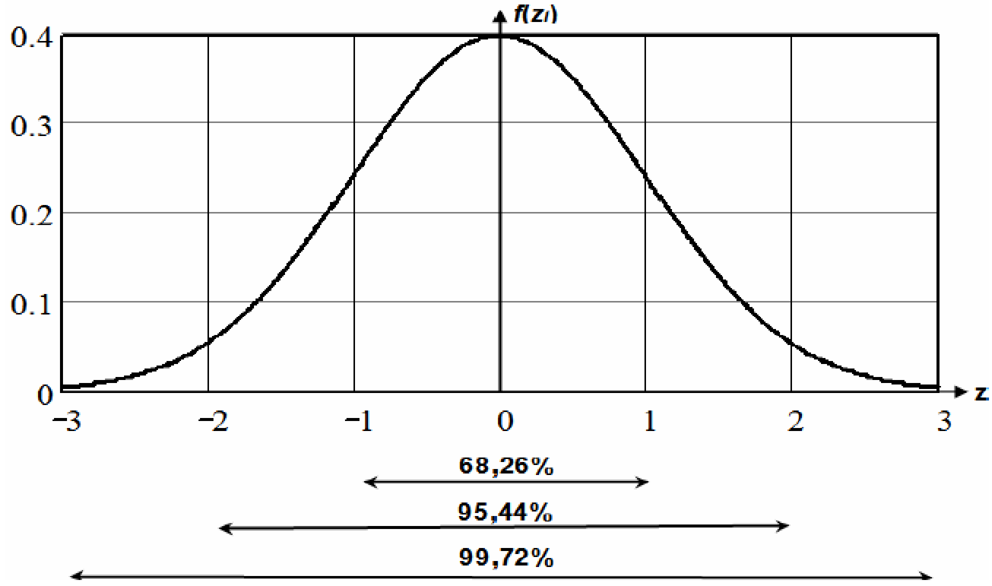
1. **Standart normal taqsimot**
2. **Z - statistika va uning mohiyati**
3. **Normal taqsimlash zichligi**
4. **Styudent t-taqsimot mezoni**

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 (1-rasm).



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 intervalgacha tushadi $(\bar{x} - \sigma_x)$ 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 $(\bar{x} + 3\sigma_x)$. gacha. Ishlarning aksariyati normal taqsimot bilan so'nggi intervalgacha tushadi. Ekonometrik modellarni yaratish uchun foydalaniladigan dastlabki ma'lumotlar ishonchli bo'lishi kerak. Axborotni ishonchliligini tekshirish uchun hisoblash va baholash kerak ikkita ko'rsatkich: assimetriya (A) va ekstsess (Э). Ushbu ko'rsatkichlar quyidagi formulalar bilan aniqlanadi:

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

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

Bu yerda:

$X_i -$ indikatorning haqiqiy qiymati;

$X -$ ko'rsatkichning o'rtacha qiymati;

$n -$ tajribalar soni;

$\sigma_x -$ hisoblangan standart og'ish formulaga muvofiq

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

Agar A va Э parametrlari nolga teng bo'lsa, unda dastlabki ma'lumotlar to'liq ishonchli hisoblanadi. Bunday mukammal variant kamdan-kam hollarda iqtisodiy hodisa va jarayonlarni o'rganishda duch keladi.

Agar assimetriya ijobiy qiymatga ega bo'lsa, unda mos keladigan ehtimollik maydoni normal taqsimot grafigida nisbatan o'ng tomonga siljiydi. Salbiy assimmetriya koeffitsienti bilan grafik chapga siljiydi.

Ekstsessning o'ziga xos qiymati vertikal ehtimollik taqsimoti grafigini siljitish orqali aniqlanadi. Xususan, agar ehtimollik maydoni eng yuqori darajaga ko'tariladi, keyin Э > 0. O'z navbatida,Э koeffitsientining pasayishi o'rganilayotgan grafik tobora tekislanib borishiga olib keladi.

Shu munosabat bilan, og'ishlarning mumkin bo'lgan chegaralari to'g'risida savol tug'iladi nol qiymatlardan A va Э koeffitsientlari. Axborot mumkin ishonchli deb hisoblaydi va agar keyingi ishlov berish uchun mos bo'lsa 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 ekstsess

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

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

Asimmetriya va ekstsess xatolari faqat n tajribalar soniga bog'liq. Axborotni ishonchliligini tekshirish bosqichida tavsiya etiladi

Asimmetriya va ekstsess 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 koeffitsientlari. 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, ob'ekt, korxona va boshqalar).O'rganilayotgan populyatsiyadan olib tashlanishi kerak bo'lgan ob'ektlar 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 dan 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, mezonning empirik qiymatini hisoblang χ^2_{emp}.

4. Pirson mezonining kritik qiymatlari jadvaliga ko'ra aniqlang qiymat χ^2_{cr} 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 $\chi^2_{emp} < \chi^2_{cr}$, 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 $\chi^2_{emp} \geq \chi^2_{cr}$ tengsizlik ning aksi bo'lsa, alternativ gipotezani qabul qiladi: statistik taqsimot odatdagidan farq qiladi.

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

$$\boldsymbol{Z - Test \;=\; (\bar{x} - \mu) \,/\, (\sigma \,/\, \sqrt{n})}$$

Bu yerda,

x̄ = namunaning o'rtacha qiymati

m = o'rtacha aholi soni

σ = aholi standart og'ishi

n = kuzatuvlar soni

Z -test statistikasi formulasi - №1 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 o'rtacha daromadlilik 1%dan ortiq bo'lgan tahlilga keladi.

Tushuntirish

- Birinchidan, namuna o'rtacha qiymatini aniqlang (bu barcha tasodifiy namunalarning o'rtacha og'irligi).
- Populyatsiyaning o'rtacha qiymatini aniqlang va undan namunaviy o'rtacha qiymatni olib tashlang.
- Keyin bu qiymatni kuzatuvlar sonining kvadrat ildiziga bo'linadigan standart og'ish bilan ajrating.
- Yuqoridagi amallarni bajargandan so'ng, testning z statistikasi natijalari hisoblab chiqiladi.

Z-testi oddiy tasodifiy o'zgaruvchining o'rtacha qiymatini berilgan qiymat bilan solishtirish uchun ishlatiladi. Z-testi foydalidir yoki namuna 30 dan oshganda va populyatsiyaning dispersiyasi ma'lum bo'lganda ishlatilishi kerak. Z-testi o'rtacha namunaning taqsimlanishi normal bo'lsa yaxshi bo'ladi. Agar z-test ma'lum shartlar bajarilgan bo'lsa qo'llaniladi, aks holda biz boshqa testlardan foydalanishimiz kerak va z-testda hech qanday tebranishlar yo'q. Bir agentli z-testi ma'lum bir populyatsiyaning o'rtacha 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.

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:

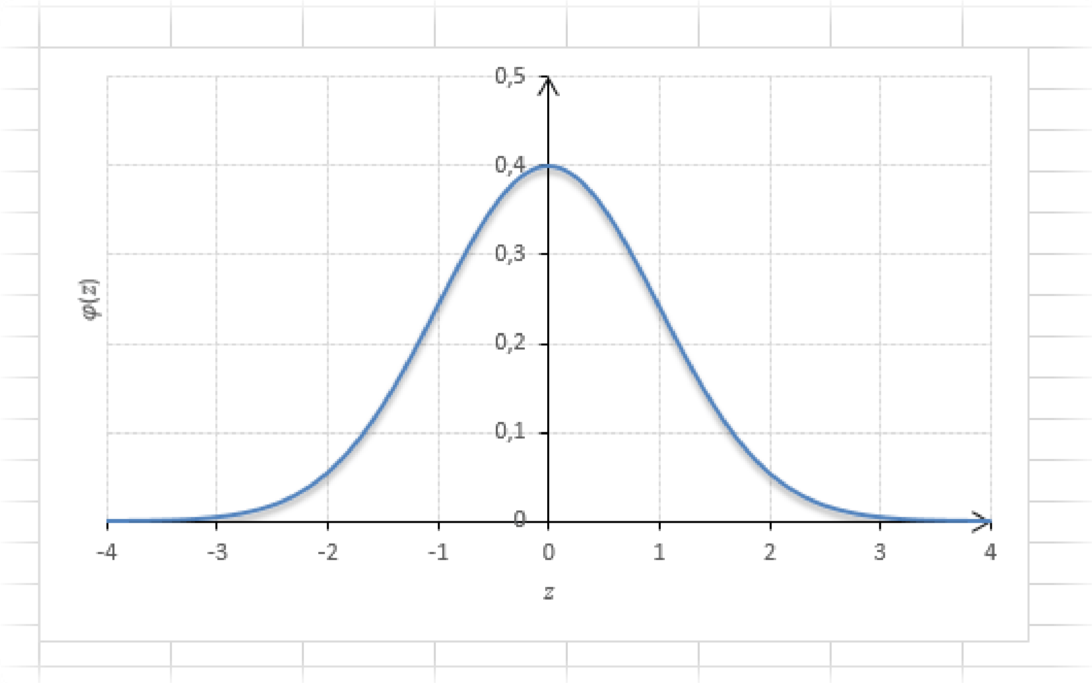
O'rtacha o'zgaruvchini o'rtacha arifmetik va tarqalash (*z)*. Hozir u mos ravishda 0 va 1 ga teng. Bu oddiy algebraik o'zgarishlar yordamida oson ta'minlanadi. Namunaviy tarqatish ko'rinishi zichligi (uchun *z-hisob-kitoblar*) ya`ni Gauss funktsiyasi:

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

Normal taqsimlash zichligi funktsiyasi:

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

Zichlik grafigi:



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

Shunday qilib, grafikga 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 funktsiyasi 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 funktsiyasi belgilangan o'qga nisbatan nosimmetrikdir.

Styudent t-taqsimot mezoni

Regressiya va korrelyatsiya koeffitsientlarining 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 koeffitsientlarining 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 koeffitsienti formulalar bilan aniqlanadi:

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

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

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

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:

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

bu yerda:

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

Va prognozlash uchun ishonch intrevali tuziladi:

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

bu 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 funktsiyasini ifodalang.
4. Regressiya va korrelyatsiya koeffitsientlarining statistik ahamiyatini baholash uchun qanday mezon qo`llaniladi.
5. Ishonch oralig`i to`g`risida tushuncha bering.