7,8-ma'ruzalar. Matematik statistika vazifalari. Statistik taqsimot. Empirik taqsimot funksiyasi.

Tayanch iboralar. *Bosh to'plam, tanlanma, statistik taqsimot, empirik taqsimot funksiyasi, poligon, gistogramma, reprezentativ, variatsion qator, varianta.*

Reja:

1-mashg'ulot

1. Matematik statistikaning vazifalari va masalalari.
2. Korrelyatsiya momenti.
3. Tanlanma metod.

2-mashg'ulot

4. Statistik taqsimot.
5. Empirik taqsimot funksiya.
6. Poligon va gistogramma.

1. Statistika lotincha so'z bo'lib, holat, vaziyat ma'nosini bildiradi. Statistika tabiatda va jamiyatda uchraydigan hodisalarni o'rganadi va ular bo'ysinadigan qonuniyatni aniqlaydi.

Buning uchun quyidagi vazifalarni bajarishi kerak:

Matematik statistikaning birinchi vazifasi-statistik ma'lumotlarni to'plash va (agar ma'lumotlar juda ko'p bo'lsa) guruhlash usullarini ko'rsatish.

Matematik statistikaning ikkinchi vazifasi-statistik ma'lumotlarni tahlil qilish metodlarini tadqiqot masalalariga muvofiq holda ishlab chiqish.

Matematik statistika yuqoridagi vazifalarni bajarish mobaynida shug'ullanadigan ba'zi masalalarni keltirib o'tamiz:

1) tasodifiy hodisa ro'y berishi ehtimolining noma'lum qiymatini baholash;

2) noma'lum taqsimot funksiyani baholash;

3) ko'rinishi ma'lum bo'lgan taqsimot funksiyasining noma'lum parametrlarini baholash;

4) tasodifiy miqdorning bir yoki bir necha tasodifiy miqdorlarga bog'liqligini va bog'liqlik darajasini aniqlash;

5) statistik gipotezalarni tekshirish.

Zamonaviy statistika fani noaniqlik sharoitida muammoning eng qulay yechimini aniqlab beradi.

Shunday qilib, matematik statistikaning vazifasi ilmiy va nazariy xulosalar chiqarish maqsadida statistik ma'lumotlarni to'plash va ularni tahlil qilish metodlarini yaratishdan iboratdir.

Bir jinsli ob'yektlar to'plamini bu ob'yektlarni xarakterlovchi biror bir sifat yoki son belgisiga nisbatan o'rganish talab qilinsin. Masalan, agar ob'yekt biror xil detallar partiyasi bo'lsa, u holda detalning sifat belgisi bo'lib, uning standartligi, son belgisi bo'lib esa detalning o'lchami xizmat qiladi.

2. *1-ta'rif.* X va Y tasodifiy miqdorlarning korrelyatsiya momenti (yoki kovariatsiyasi) deb, quyidagi songa aytiladi:

$$K_{xy} = M[(X - M(X))(Y - M(Y))].$$

X va Y tasodifiy miqdorlar diskret bo'lsa, u holda bu formula quyidagi ko'rinishini oladi:

$$K_{xy} = \sum_{i \cdot j} \left(x_i - M(X)\right)\left(y_i - M(Y)\right) p_{ij} ,$$

bunda $p_{ij} = P\left(X = x_i; Y = y_j\right)$.

Korrelyatsiya momenti ifodasini matematik kutilma xossalari asosida quyidagicha almashtirilish mumkin:

$$M(X - M(X))(Y - M(Y)) = M[XY - XM(Y) - YM(X) + M(X)M(Y)] =$$
$$= M(XY) - M(X)M(Y) - M(Y)M(X) + M(X)M(Y) = M(XY) - M(X)M(Y).$$

1-teorema. Agar tasodifiy miqdorlar o'zaro bog'liq bo'lmasa, u holda korrelyatsiya momenti nolga teng bo'ladi.

2-ta'rif. X va Y tasodifiy miqdorlarning korrelya tsiya koeffitsienti deb

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} \qquad (1)$$

tenglik bilan aniqlanadigan kattalikka aytiladi.

Korrelyatsiya momenti uchun quyidagi

$$\left|K_{xy}\right| \le \sqrt{D_x D_y}$$

tengsizlik o'rinli bo'lishini ko'rsatish mumkin, chunki $\left|r_{xy}\right| \le 1$.

Agar X va Y tasodifiy miqdorlar bog'liqmas bo'lsa, u holda ularning korrelyatsiya koeffitsienti nolga teng bo'ladi.

Quyidagi teorema tasodifiy miqdorlar orasida bog'lanishni tavsiflashda korrelyatsiya koeffitsientining ahamiyatini yana ham batafsil oydinlashtirib beradi.

2-teorema. Y tasodifiy miqdor X tasodifiy miqdorning chiziqli funksiyasi, ya'ni $Y = aX + b$ bo'lsin, u holda agar $a > 0$ bo'lsa, $r_{xy} = 1$, agar $a < 0$ bo'lsa, $r_{xy} = -1$ bo'ladi.

3. Ba'zan tekshirish yalpi o'tkaziladi, ya'ni to'plamdagi ob'yektlarning har birini o'rganilayotgan belgiga nisbatan tekshiriladi. Lekin yalpi tekshirish amaliyotda nisbatan kam qo'llaniladi. Masalan, to'plam juda ko'p ob'yektlarni o'z ichiga olgan bo'lsa, u holda yalpi tekshirish o'tkazish maqsadga muvofiq emas. Bunday hollarda to'plamdan chekli sondagi ob'yektlar tasodifiy ravishda olinadi va ular o'rganiladi.

3-ta'rif. Tanlanma to'plam (bundan keyin *tanlanma*) deb umumiy to'plamdan tasodifiy ravishda ajratib olingan ob'yektlar to'plamiga aytiladi.

4-ta'rif. Bosh to'plam deb tanlanma ajratiladigan ob'yektlar to'plamiga aytiladi.

To'plam (bosh to'plam yoki tanlanma) hajmi deb, bu to'plamdagi ob'yektlar soniga aytiladi. Masalan, 500 ta detaldan tekshirish uchun 50 ta detal olingan bo'lsa, u holda bosh to'plam hajmi $N = 500$, tanlanma hajmi esa $n = 50$.

5-ta'rif. Bosh to'plamdan olingan tanlanma bo'yicha bosh to'plam haqida hulosa qilishga asoslangan usulga, *tanlanma usul* deb ataladi.

Tanlanmani ajratib olish ikki xil yo'l bilan amalga oshirilishi mumkin: ob'ekt ajratib olinib uning ustida kuzatish o'tkazilgandan so'ng, u bosh to'plamga qaytarilishi yoki qaytarilmasligi mumkin.

6-ta'rif. Takroriy tanlanma deb, shunday tanlanmaga aytiladiki, bunda olingan ob'ekt tajribadan so'ng (keyingisini olishdan oldin) bosh to'plamga qaytariladi.

7-ta'rif. Takroriy bo'lmagan tanlanma deb, ajratib olingan ob'yekt kuzatishdan so'ng bosh to'plamga qaytarilmaydi.

Odatda, ko'p hollarda, qaytarilmaydigan tasodifiy tanlashdan foydalaniladi.

Tanlanmadagi ma'lumotlar bo'yicha bosh to'plamning bizni qiziqtirayotgan belgisi haqida etarlicha ishonch bilan fikr yuritish uchun tanlanmaning ob'yektlari bosh to'plamni to'g'ri tasvirlashi zarur. Bu talab qisqacha bunday ta'riflanadi: tanlanma *reprezentativ* (vakolatli) bo'lishi kerak. Odatda, tanlanmaning reprezentativligini ta'minlash uchun bosh to'plam har bir elementining tanlanmaga tushish ehtimoli teng deb olinadi.

Amaliyotda tanlanmani ajratib olishda turli usullardan foydalaniladi. Bu usullarni 2 tipga ajratish mumkin:

1. Bosh to'plamni qism to'plamlarga ajratmasdan tanlanma olish, bunda oddiy tasodifiy: a) qaytarilmaydigan; b) qaytariladigan usullardan foydalaniladi.
2. Bosh to'plamni qism to'plamlarga ajratib so'ngra tanlanma olish, bunda bosh to'plam: a) tipik; b) mexanik; v) seriyalab qism to'plamlarga ajratiladi, so'ngra tanlanma ajratib olinadi.

8-ta'rif. Agar bosh to'plamdan ob'yektlar bittadan tasodifiy ravishda olinib tanlanma tanlansa, bu *oddiy tasodifiy* tanlash deyiladi.

Tipik tanlashda bosh to'plamni uning "tipik" xususiyatlarini e'tiborga olgan holda qism to'plamlarga ajratiladi, so'ngra uning qism to'plamlaridan tanlanma ajratib olinadi.

Mexanik tanlash bosh to'plamni mexanik ravishda qism to'plamlarga ajratiladi, so'ngra uning qism to'plamlaridan tanlanma ajratib olinadi.

Seriyali tanlash bosh to'plamni qism to'plamlarga seriyalab ajratiladi, so'ngra uning qism to'plamlaridan tanlanma ajratib olinadi.

Odatda, ko'p hollarla, tanlanma ajratib olishda yuqoridagi usullardan aralash foydalaniladi, ya'ni ko'rsatilgan usullardan birgalikda foydalaniladi. Masalan, bosh to'plamni ba'zan bir xil hajmli seriyalarga ajratiladi, keyin oddiy tasodifiy tanlash bilan ayrim ob'yektlar olinadi.

Bosh to'plamdan tanlanma olingan bo'lsin. Bunda tanlanmaning x_i qiymati n_i $(i=1,2,...)$ marta kuzatilgan va $\sum_i n_i = n$ bo'lsin. Kuzatilgan x_i qiymatlar variantalar, variantalarning ortib yoki kamayib borish tartibida yozilgan ketma-ketligi esa *variatsion qator* deyiladi. Kuzatishlar soni-n_i chastotalar, ularning tanlanma hajmiga nisbati esa $W_i = \frac{n_i}{n}$ -nisbiy chastotalar deyiladi.

2-mashg'ulot

4. Tanlanmaning statistik taqsimoti deb, variantalar va ularga mos chastotalar yoki nisbiy chastotalar ro'yxatiga aytiladi:

$$\begin{matrix} x_i: & x_1 & x_2 & \dots & x_k & \dots \\ n_i: & n_1 & n_2 & \dots & n_k & \dots \end{matrix} \quad \text{yoki} \quad \begin{matrix} x_i: & x_1 & x_2 & \dots & x_k & \dots \\ W_i: & W_1 & W_2 & \dots & W_k & \dots \end{matrix}. \qquad (1)$$

Shunday qilib, taqsimot ehtimollar nazariyasida tasodifiy miqdorning mumkin bo'lgan qiymatlari va ularning ehtimollari orasidagi moslikni, matematik statistikada esa kuzatilgan variantalar va ularning chastotalari yoki nisbiy chastotalari orasidagi moslikni bildiradi.

1-Misol. Hajmi 40 bo'lgan tanlanmaning chastotalari taqsimoti:

$$\begin{matrix} x_i: & 2 & 6 & 12 \\ n_i: & 6 & 20 & 14 \end{matrix}$$

berilgan. Nisbiy chastotalar taqsimotini yozing.

Yechish. Nisbiy chastotalarni topamiz. Buning uchun chastotalarni tanlanma hajmiga bo'lamiz.

$$W_1 = \frac{6}{40} = 0{,}15; \ W_2 = \frac{20}{40} = 0{,}5; \ W_3 = \frac{14}{40} = 0{,}35.$$

U holda, nisbiy chastotalar taqsimoti:

$$\begin{matrix} x_i: & 2 & 6 & 12 \\ n_i: & 0{,}15 & 0{,}5 & 0{,}35 \end{matrix}.$$

Faraz qilamiz, X-son belgining chastotalar statistik taqsimoti ma'lum bo'lsin. Quyidagi belgilashlar kiritamiz: n_x - X belgining x dan kichik qiymatlari kuzatilgan kuzatishlar soni; n-umumiy kuzatishlar soni.

Ma'lumki, $X < x$ hodisaning nisbiy chastotasi: $\frac{n_x}{n}$. Agar x o'zgaradigan bo'lsa, u holda, nisbiy chastota ham o'zgaradi.

Demak, $\frac{n_x}{n}$ nisbiy chastota x ning funksiyasidir.

5. *9-ta'rif.* Taqsimotning *empirik funksiyasi* (tanlanmaning taqsimot funksiyasi) deb har bir x qiymat uchun $X < x$ hodisaning nisbiy chastotasini aniqlaydigan $F_n^*(x)$ funksiyaga aytiladi.

Demak, ta'rifga ko'ra $$F_n^*(x) = \frac{n_x}{n}. \qquad (2)$$

Bu yerda $n_x - x$ dan kichik variantalar soni, n-tanlanma hajmi.

2-Misol. Tanlanmaning quyidagi taqsimoti:

$$\begin{matrix} x_i: & 2 & 6 & 10 \\ n_i: & 12 & 18 & 30 \end{matrix}$$

bo'yicha uning empirik funksiyasini tuzing.

Yechish. Tanlanma hajmini topamiz: $n = 12+18+30=60$

$$F_n^*(x) = \begin{cases} 0{,}2 \ agar \ x < 2, \\ 0{,}3 \ agar \ 2 \le x < 6, \\ 0{,}5 \ agar \ 6 \le x < 10, \\ 1 \ agar \ x \ge 10 \end{cases}$$

Bosh to'plamning $F(x)$-taqsimot funksiyasi nazariy taqsimot funksiyasi deb ataladi. Empirik funksiya $F_n^*(x)$ $X < x$ hodisaning nisbiy chastotasini, nazariy taqsimot funksiya $F(x)$ esa $X < x$ hodisaning ro'y berish ehtimolini aniqlaydi. $F_n^*(x)$ funksiya uchun $F(x)$ funksiyaning barcha xossalari o'rinli. Ya'ni:

1) $F_n^*(x) \in [0;1]$;

2) $F_n^*(x)$-kamaymaydigan funksiya;

3) agar x_1-eng kichik varianta bo'lsa, u holda $x < x_1$ qiymatlar uchun $F_n^*(x) = 0$ qiymatlar uchun

agar x_k-eng katta varianta bo'lsa, u holda $x \geq x_k$ qiymatlar uchun $F_n^*(x) = 1$.

Shunday qilib, tanlanmaning empirik taqsimot funksiyasi bosh to'plam nazariy taqsimot funksiyasini baholash uchun xizmat qiladi.

Haqiqatan ham, Bernulli teoremasiga asosan, $\lim\limits_{n\to\infty} P\left(\left|F(x) - F_n^*(x)\right| < \varepsilon\right) = 1$. Demak, tanlanmaning empirik taqsimot funksiyasidan bosh to'plam nazariy (integral) funksiyasining taxminiy ko'rinishi sifatida foydalanish mumkin.

6. Ko'rgazmalilik uchun statistik taqsimotning turli grafiklari chiziladi, masalan, poligon va gistogramma.

Chastotalar poligonini yasash uchun Dekart koordinatalar sistemasida kesmalari (x_i, n_i) $(i = 1, 2, ...)$ nuqtalarni tutashtiruvchi siniq chiziq hosil qilish kerak. *Nisbiy chastotalar poligonini* yasash uchun esa Dekart koordinatalar sistemasida kesmalari (x_i, W_i) $(i = 1, 2, ...)$ nuqtalarni tutashtiruvchi siniq chiziq hosil qilish kerak bo'ladi. Chastotalar va nisbiy chastotalar poligonini diskret tasodifiy miqdorlarning grafik usulda berilishi deb ham tushunish mumkin.

Agar kuzatilayotgan belgi uzluksiz bo'lsa, u holda uni grafik usulda tasvirlash uchun gistogramma yasash maqsadga muvofiqdir, buning uchun belgining kuzatiladigan qiymatlarini o'z ichiga olgan intervalni uzunligi o'zgarmas-h bo'lgan bir nechta qismiy intervallarga bo'linadi va har bir i-qismiy interval uchun n_i-ya'ni i-intervaldagi variantalar chastotalarining yig'indisi topiladi. So'ngra, Dekart koordinatalar sistemasida *chastotalar gistogrammasi,* asoslari h uzunlikdagi intervallar, balandliklari esa $\frac{n_i}{h}$ nisbatlarga (chastota zichligi) teng bo'lgan to'g'ri to'rtburchaklardan iborat pog'onaviy figura, yoki *nisbiy chastotalar gistogrammasi* asoslari h uzunlikdagi intervallar, balandliklari esa $\frac{W_i}{h}$ nisbatga (nisbiy chastota zichligi) teng bo'lgan to'g'ri to'rtburchaklardan iborat pog'onaviy figura, yasaladi.

O'z-o'zini tekshirish uchun savollar.

1. Matematik statistika vazifalarini ayting.
2. Tanlanma olishning qanday usullari bor?
3. Tanlanmaning reprezentativligi nimadan iborat?
4. Tanlanmaning statistik taqsimoti ta'rifini bering.
5. Empirik taqsimot funksiya ta'rifini keltiring.
6. Poligon va gistogramma qanday quriladi?

Mustaqil yechish uchun masalalar.

1. Quyidagi tanlanma berilgan: 2, 1, 3, 3, 4, 4, 3, 3, 3, 2, 3, 1, 1, 2, 3, 3, 4, 2, 2, 3.
a) variatsion qatorni tuzing;
b) chastotalar jadvalini tuzing;
v) nisbiy chastotalar poligonini chizing.

2. Korxona ishchilaridan tavakkaliga 20 tasi tanlanib, ularning tarif razryadlari xaqida quyidagi ma'lumotlar olingan: 1, 2, 4, 6, 3, 4, 4, 2, 6, 3, 5, 3, 3, 1, 5, 4, 2, 5, 4, 3.
Shu ma'lumotlarga asoslangan holda:
a) tanlanmaning statistik taqsimotini tuzing va chastotalar poligonini yasang;
b) empirik funksiyani tuzing.

3. Tanlanma

x_i	4	5	7	12
n_i	5	2	3	10

chastotalar taqsimoti ko'rinishida berilgan. Nisbiy chastotalar taqsimotini toping.

4. Chastotalar poligonini yasang.

x_i	15	20	25	30	10
n_i	10	15	30	20	25

5. Tanlanmaning quyidagi berilgan taqsimoti bo'yicha chastotalar gistogrammasini yasang.

Intervallar ro'yxati	Qismiy intervallar	Qismiy intervallardagi variantalar chastotalarining yig'indisi
i	$x_i - x_{i+1}$	n_i
1	2-5	6
2	5-8	10
3	8-11	4
4	11-14	5
		$n = \sum n_i = 25$

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.
5. Крамер Н.Ш. Теория вероятностей и математическая статистика. Учебник М.,2001.

9-ma'ruza. Taqsimot parametrlarining statistik baholari. Baholarga qo'yiladigan talablar.

Tayanch iboralar. *Statistik baho, siljimagan baho, effektiv baho, asosli baho, tanlanma o'rtachasi, bosh to'plam o'rtachasi, tanlanma dispersiyasi, "tuzatilgan" dispersiya, bosh to'plam dispersiyasi.*

Reja.

9.1. Taqsimot parametrlarining statistik baholari.
9.2. Baholarga qo'yiladigan talablar.
9.3. Variatsion qatorning ba'zi xarakteristikalari.

Ma'lumki, matematik statistika masalaridan biri tanlanmadan asosida bosh to'plam taqsimot funksiyasining xarakteristikalari hisoblangan noma'lum parametrlar uchun statistik baholar o'rnatish edi. Bu masala qanday hal qilinishini ko'rib chiqamiz.

Faraz qilamiz, bosh to'plamning son belgisini o'rganish talab qilinayotgan va belgining taqsimot funksiyasi nazariy mulohazalar asosida aniqlangan bo'lsin. Bu taqsimotni aniqlaydigan noma'lum parametrlarni baholash masalasini ko'rib chiqaylik. Masalan, bosh belgi, to'g'rirog'i o'rganilayotgan belgi bosh to'plamda normal taqsimlanganligi oldindan ma'lum bo'lsa, u holda matematik kutilmani va o'rtacha kvadratik chetlanishni baholash, ya'ni taqribiy hisoblash zarur, chunki bu ikki parametr normal taqsimotni to'liq aniqlaydi, agar belgi Puasson taqsimotiga ega deyishga asos bo'lsa, u holda bu taqsimotni aniqlaydigan $\lambda > 0$ parametrni baholash, ya'ni taqribiy hisoblash zarur.

Odatda, tadqiqotchi ixtiyorida tanlanma asosida olingan ma'lumotlar, masalan, tanlanma son belgisini n marta kuzatish natijasida olingan $x_1, x_2, ..., x_n$ qiymatlar bo'ladi. Demak, baholanayotgan belgining bahosi xuddi shu ma'lumotlar orqali ifodalaniladi.

Tanlanmadagi $x_1, x_2, ..., x_n$ qiymatlarni erkli $X_1, X_2, ..., X_n$-tasodifiy miqdorlar deb qarab, nazariy taqsimot noma'lum parametrining statistik bahosini topish uchun kuzatilayotgan tasodifiy miqdorlar orqali shunday funksiya topish kerakki, u baholanayotgan parametrning taqribiy qiymatini bersin. Masalan, normal taqsimotning matematik kutilishini baholash uchun ushbu

$$\overline{X} = \frac{X_1 + X_2 + ... + X_n}{n}$$

funksiya xizmat qiladi.

Shunday qilib, nazariy taqsimot noma'lum parametrning *statistik bahosi* deb kuzatilgan tasodifiy miqdorlardan tuzilgan funksiyaga aytiladi.

Statistik baho baholanayotgan parametrning yaxshi bahosi bo'lishi uchun u ma'lum bir talablarni qanoatlantirishi lozim. Quyida mana shu talablarni ko'rib chiqamiz:

Bosh to'plam $F(x)$-nazariy taqsimot funksiyasining θ parametri noma'lum bo'lib uning statistik bahosi θ^* bo'lsin. Bosh to'plamdan olingan n hajmli tanlanma bo'yicha θ_1^* bahoni topamiz. Tajribani takrorlaymiz, ya'ni bosh to'plamdan yana n hajmli tanlanma olib θ_2^* bahoni topamiz. Tajribani ko'p marta takrorlab, $\theta_1^*, \theta_2^*, ..., \theta_n^*$ sonlar ketma-ketliginini hosil qilamiz, umuman olganda, $\theta_1^*, \theta_2^*, ..., \theta_n^*$ sonlar har xil bo'ladi. U holda θ^* bahoni tasodifiy miqdor, $\theta_1^*, \theta_2^*, ..., \theta_n^*$ sonlarni esa uning mumkin bo'lgan qiymatlari sifatida qarash mumkin.

θ^* tasodifiy miqdorning $M(\theta^*)$-matematik kutilishini hisoblaymiz. $M(\theta^*)$ va θ noma'lum parametr qiymatlarini taqqoslasak ular orasida:

1) $M(\theta^*) < \theta$;

2) $M(\theta^*) = \theta$;

3) $M(\theta^*) > \theta$.

munosabatlardan biri albatta o'rinli bo'ladi. Matematik kutilmasi baholanayotgan parametrga teng bo'lmagan statistik bahoni ishlatish sistematik xatolarga olib keladi. Shu sababli, θ^* bahoning matematik kutilmasi baholanayotgan parametrga teng bo'lishini talab qilish tabiiy holdir.

Demak, $M(\theta^*) = \theta$ talabga rioya qilish sistematik xatolardan saqlaydi.

1-ta'rif. Agar bosh to'plamdan ixtiyoriy hajmli tanlanma olinganda ham θ^* bahoning matematik kutilmasi baholanayotgan θ parametrga teng, ya'ni $M(\theta^*) = \theta$, bo'lsa, u holda θ^* baho *siljimagan baho* deb ataladi, aks holda θ^* siljigan baho deyiladi.

2-ta'rif. Agar θ^* baho va θ noma'lum parametrlar uchun $\lim\limits_{n\to\infty} M(\theta^*) = \theta$ munosabat o'rinli bo'lsa, u holda θ^* baho *asimptotik siljimagan baho* deb ataladi.

Ammo shuni ham ta'kidlash keraki, siljimagan baho har doim ham baholanayotgan parametrga yaxshi yaqinlashadi deb hisoblash xato bo'ladi.

θ_l^* ni θ ning tarqibiy qiymati sifatida qabul qilib, katta xatoga yo'l qo'ygan bo'lar edik. Shu sababli, statistik baholarga effektivlik talabi qo'yiladi.

3-ta'rif. Agar $\theta_l^* \in \theta^*$ bahoning dispersiyasi eng kichik, ya'ni $\inf\limits_{\theta_i^*} D(\theta_i^*) = \theta_l^*$ bo'lsa, u holda θ_l^* *effektiv baho* deb ataladi.

Umuman olganda, effektiv baho mavjud bo'lmasligi ham mumkin.

4-ta'rif. Agar θ^*, $\theta_i^* (i = 1, 2, ...)$ baholar va θ noma'lum parametrlar uchun $\lim\limits_{n\to\infty} \dfrac{M\left((\theta^*(x_1, x_2, ..., x_n) - \theta)^2\right)}{\inf\limits_{\theta_i^*} M\left((\theta_i^*(x_1, x_2, ..., x_n) - \theta)^2\right)} = 1$ munosabat o'rinli bo'lsa, u holda θ^* baho *asimptotik effektiv baho* deb ataladi.

Juda katta hajmli (*n* etarlicha katta bo'lganida) tanlanmalar qaralganda statistik baholarga asoslilik talabi qo'yiladi.

5-ta'rif. Asosli baho deb baholanayotgan parametrga $n \to \infty$ da ehtimol bo'yicha yaqinlashadigan θ^* bahoga aytiladi, ya'ni $\lim\limits_{n\to\infty} P\left\{\left|\theta^* - \theta\right| > \varepsilon\right\} = 0$, bu yerda $\varepsilon > 0$–yetarli darajada kichik son.

Agar bahoning dispersiyasi $n \to \infty$ da nolga intilsa, u holda bunday baho asosli ham bo'ladi.

Agar N hajmli bosh to'plamning mumkin bo'lgan $x_1, x_2, ..., x_N$ -qiymatlari turli bo'lsa, $\overline{x_B}$ *-bosh to'plam o'rtachasi*

$$\overline{x_B} = \frac{x_1 + x_2 + ... + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (1)$$

formula bilan topiladi; agar N hajmli bosh to'plamning mumkin bo'lgan $x_1, x_2, ..., x_k$ - qiymatlari mos ravishda $N_1, N_2, ..., N_k$ chastotalarga ega bo'lib, $N_1 + N_2 + ... + N_k = N$ bo'lsa:

$$\overline{x_B} = \frac{x_1 N_1 + x_2 N_2 + ... + x_k N_k}{N} = \frac{1}{N}\sum_{i=1}^{k} x_i N_i \,. \qquad (2)$$

Bosh to'plamning kuzatilayotgan X belgisini tasodifiy miqdor sifatida qarasak, uning matematik kutilmasi uchun $M(X) = \overline{x_B}$ tenglik o'rinli bo'ladi.

Agar n hajmli tanlanmaning mumkin bo'lgan $x_1, x_2, ..., x_n$ -qiymatlari turli bo'lsa, $\overline{x_T}$ -*tanlanma o'rtacha*

$$\overline{x_T} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (3)$$

formula bilan topiladi; agar n hajmli tanlanmaning mumkin bo'lgan $x_1, x_2, ..., x_k$ - qiymatlari mos ravishda $n_1, n_2, ..., n_k$ chastotalarga ega bo'lib, $n_1 + n_2 + ... + n_k = n$ bo'lsa:

$$\overline{x_T} = \frac{x_1 n_1 + x_2 n_2 + ... + x_k n_k}{n} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i \,. \qquad (4)$$

Bosh to'plam o'rtachasi-*M(X)* ning statistik bahosi sifatida

$$\overline{x_T} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ yoki } \overline{x_T} = \frac{x_1 n_1 + x_2 n_2 + ... + x_k n_k}{n} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i$$

-tanlanma o'rtacha qabul qilinadi.

Ma'lumki, katta sonlar qonuniga (Chebishev teoremasi)asosan ixtiyoriy kichik $\varepsilon > 0$ son uchun

$$\lim_{n\to\infty} P\left(\left|\overline{x_T} - M(\overline{x_T})\right| < \varepsilon\right) = \lim_{n\to\infty} P\left(\left|\overline{x_T} - a\right| < \varepsilon\right) = 1,$$

ya'ni n ortishi bilan $\overline{x_T}$ -tanlanma o'rtachasi bosh to'plam matematik kutilmasiga ehtimol bo'yicha yaqinlashadi. Bundan esa, $\overline{x_T}$ baho a uchun asosli baho bo'lishi kelib chiqadi.

Agar bosh to'plamdan ancha katta hajmli bir nechta tanlanmalar olinib har birining tanlanma o'rtachalari topiladigan bo'lsa, ular o'zaro taqriban teng bo'ladi. Bu tanlanma o'rtachaning *turg'unlik xossasi* deyiladi.

1-м*isol.* Tanlanmaning

$$\begin{array}{llll} x_i: & 4 & 8 & 11 \\ n_i: & 5 & 10 & 5 \end{array}$$

statistik taqsimoti bo'yicha bosh to'plam matematik kutilmasining siljimagan bahosini toping.

Y*echish.*

(4) formuladan foydalanamiz. U holda $\overline{x_T} = \frac{4\cdot 5 + 8\cdot 10 + 11\cdot 5}{20} = \frac{155}{20} = 7,75$.

Agar N hajmli bosh to'plamning mumkin bo'lgan $x_1, x_2, ..., x_N$ -qiymatlari turli bo'lsa, *bosh to'plam dispersiyasi*

$$D(X) \equiv D_B = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x_B})^2 \qquad (5)$$

formula bilan topiladi; agar N hajmli bosh to'plamning mumkin bo'lgan $x_1, x_2, ..., x_k$ -qiymatlari mos ravishda $N_1, N_2, ..., N_k$ chastotalarga ega bo'lib, $N_1 + N_2 + ... + N_k = N$ bo'lsa:

$$D_B = \frac{1}{N}\sum_{i=1}^{k} N_i (x_i - \overline{x_B})^2 . \qquad (6)$$

Bosh to'plam o'rtacha kvadratik chetlanishi

$$\sigma(X) \equiv \sigma_B = \sqrt{D_B} \qquad (7)$$

formula bilan aniqlanadi.

Agar n hajmli tanlanmaning mumkin bo'lgan $x_1, x_2, ..., x_n$ -qiymatlari turli bo'lsa, *tanlanma dispersiyasi*

$$D(x) \equiv D_T = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x_T})^2 \qquad (8)$$

formula bilan topiladi; agar n hajmli tanlanmaning mumkin bo'lgan $x_1, x_2, ..., x_k$ -qiymatlari mos ravishda $n_1, n_2, ..., n_k$ chastotalarga ega bo'lib, $n_1 + n_2 + ... + n_k = n$ bo'lsa:

$$D_T = \frac{1}{n}\sum_{i=1}^{n} n_i (x_i - \overline{x_T})^2 . \qquad (9)$$

Misol.

2. Tanlanmaning

x_i :	4	8	11
n_i :	5	10	5

statistik taqsimoti bo'yicha uning dispersiyasini toping.

Yechish.

(4) formuladan foydalansak: $\overline{x_T} = 7,75$. Dispersiyani hisoblash uchun (9) formuladan foydalanamiz. U holda

$$D_T = \frac{5\cdot(4-7,75)^2 + 10\cdot(8-7,75)^2 + 5\cdot(11-7,75)^2}{20} = \frac{70,3125 + 0,625 + 70,3125}{20} = 7,0625 .$$

Dispersiyani hisoblashda (5), (6), (8), (9) formulalar noqulay, shu sababli, dispersiya va matematik kutilmalarning xossalaridan foydalanib, dispersiyani hisoblash uchun qulay bo'lgan quyidagi formulani keltirib chiqarish mumkin:

$$D = \overline{x^2} - \left(\overline{x}\right)^2, \qquad \overline{x} = \frac{\sum_i n_i x_i}{n}, \overline{x^2} = \frac{\sum_i n_i x_i^2}{n} . \qquad (10)$$

Bosh to'plam dispersiyasi uchun baho sifatida tanlanma dispersiyasi $D_T = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x_T})^2$ qanday baho bo'lishini ko'rib chiqamiz. Qulaylik uchun $m = M(X)$, $\sigma_1^2 = D_B$ belgilashlar kiritib olamiz.

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x_T}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left[x_i - m - (\overline{x_T} - m)\right]^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 -$$

$$-\frac{2}{n}(\overline{x_T} - m)^2\sum_{i=1}^{n}(x_i - m) + \frac{n}{n}(\overline{x_T} - m)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 -$$

$$-\frac{2}{n}(\overline{x_T} - m)(\overline{x_T} - m)n + (\overline{x_T} - m)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 - (\overline{x_T} - m)^2.$$

Agar $M(\overline{x_T}-m)^2 = D(\overline{x_T}) = \frac{1}{n}\sigma_1^2$ belgilashni e'tiborga olsak,

$$M(\sigma^2) = \frac{1}{n} M\left(\sum_{i=1}^{n}(x_i - m)^2\right) - M(\overline{x_T}-m)^2 = \sigma_1^2 - \frac{1}{n}\sigma_1^2 = \frac{n-1}{n}\sigma_1^2 .$$

Demak, tanlanma dispersiyasi-D_T bosh to'plam dispersiyasi D_B uchun siljimagan baho bo'lolmas ekan, shu sababli, bosh to'plam dispersiyasi uchun siljimagan statistik baho sifatida

$$s^2 = \frac{n}{n-1} D_T \qquad (11)$$

-"tuzatilgan" dispersiya olinadi.

Bosh to'plam o'rtacha kvadratik chetlanishining bahosi sifatida $s = \sqrt{\frac{n}{n-1} D_T}$ - "tuzatilgan" o'rtacha kvadratik chetlanish olinadi.

Shuni alohida ta'kidlash keraki, s siljimagan baho bo'la olmaydi, shuning uchun uni "tuzatilgan" o'rtacha kvadratik chetlanish deb ataymiz.

1-eslatma. $D_T = \frac{1}{n}\sum_{i=1}^{n} n_i (x_i - \overline{x_T})^2$ va $s^2 = \frac{1}{n-1}\sum_{i=1}^{n} n_i (x_i - \overline{x_T})^2$ formulalar maxrajlari bilan farqlanadi. U holda n ning katta qiymatlarida tanlanma dispersiyasi va "tuzatilgan" dispersiyalarning farqi juda kam bo'ladi. Shu sababli, "tuzatilgan" dispersiyadan $n < 30$ hajmli tanlanmalarda foydalanish tavsiya etiladi.

2-eslatma. Agar tanlanmaning variatsion qatorida x_i-variantalarning qiymatlari katta sonlardan iborat bo'lsa, u holda x_i variantadan $u_i = \frac{x_i - c_1}{c_2}$ -shartli variantaga o'tish orqali u_i-variantalari kichik sonlardan iborat yangi variatsion qator hosil qilinadi, so'ngra yangi tanlanma uchun $\overline{u_T}$ va $D_T(u)$ lar topiladi. Oldingi tanlanmaning $\overline{x_T}, D_T(x)$ xarakteristikalarini topish uchun $\overline{x_T} = c_2\overline{u_T} + c_1$ va $D_T(x) = c_2^2 D_T(u)$ formulalardan foydalaniladi.

Matematik statistika va uning tatbiqlarida variatsion qatorning tanlanma o'rtachasi va tanlanma dispersiyasidan tashqari boshka xarakteristikalari ham ishlatiladi. Shulardan ba'zilarini keltiramiz.

Eng katta chastotaga ega bo'lgan varianta *moda* deb ataladi va M_0 kabi belgilanadi.

Mediana deb, variatsion qator variantalarini son jihatidan teng ikki qismga ajratadigan variantaga aytiladi va M_e kabi belgilanadi. Variantalar sonining juft yoki toqligiga qarab, mediana quyidagicha aniqlanadi.

$$M_e = \begin{cases} x_{k+1}, & agar \quad n = 2k+1, \\ \dfrac{x_k + x_{k+1}}{2}, & agar \quad n = 2k. \end{cases}$$

Variatsiya qulochi R deb eng katta va eng kichik variantalar ayirmasiga aytiladi:

$$R = x_{\max} - x_{\min} .$$

Variatsiya qulochi variatsion qator tarqoqligining eng sodda xarakteristikasi bo'lib xizmat qiladi.

Variatsion qator tarqoqligining yana bir xarakteristikasi sifatida *o'rtacha absolyut chetlanish* θ ham ishlatiladi.

$$\theta = \frac{\sum_i n_i \left| x_i - \overline{x_T} \right|}{n}$$

Variatsiya koeffitsienti V deb tanlanma o'rtacha kvadratik chetlanishining tanlanma o'rtachasiga nisbatini foizlardagi ifodasiga aytiladi:

$$V = \frac{\sigma_T}{\overline{x_T}} \cdot 100\% .$$

Variatsiya koeffitsienti ikkita yoki undan ortiq variatsion qatorlarning tarqoqliklarini taqqoslash uchun xizmat qiladi: variatsion qatorlardan variatsiya koeffitsienti katta bo'lgani ko'proq tarqoqlikka ega bo'ladi.

Misol. Quyidagi

$$\begin{array}{lcccc} x_i: & 1 & 3 & 6 & 16 \\ n_i: & 4 & 10 & 5 & 1 \end{array}$$

tanlanma uchun M_0, M_e, R, θ, V -xarakteristikalarni hisoblaymiz.

$M_0 = M_e = 3; \quad R = 15;$

$$\overline{x_T} = \frac{4\cdot 1 + 10\cdot 3 + 5\cdot 6 + 1\cdot 16}{4+10+5+1} = \frac{20}{20} = 4, \theta = \frac{4\cdot|1-4| + 10\cdot|3-4| + 5\cdot|6-4| + 1\cdot|16-4|}{20} = 2,2$$

$$\sigma_T = \sqrt{\frac{4\cdot(1-4)^2 + 10\cdot(3-4)^2 + 5\cdot(6-4)^2 + 1\cdot(16-4)^2}{20}} \approx 3,24, V = \frac{3,24}{4}\cdot 100\% = 80,1\%$$

O'z-o'zini tekshirish uchun savollar.

1. Statistik baho ta'rifini bering.
2. Siljimagan, asosli va effektiv baholar ta'riflarini keltiring.
3. $\overline{x_T}$ -bosh to'plam matematik kutilmasi uchun siljimagan va asosli baho bo'lishini tushuntiring.
4. D_T - tanlanma dispersiyasi bosh to'plam dispersiyasi uchun siljigan baho bo'lishini tushuntiring.
5. Variatsion qatorning xarakteristikalarini ta'riflang.

Mustaqil yechish uchun masalalar.

1. Bosh to'plamdan $n = 50$ hajmdagi tanlanma ajratilgan. Quyidagi

x_i	2	5	7	10
n_i	16	12	8	14

taqsimot bo'yicha bosh to'plam o'rtachasining siljimagan bahosini toping.

2. Guruhdagi 40 ta talabaning yozma ishlari baholarining chastotalari jadvali berilgan.

x_i	2	3	4	5
n_i	3	8	25	4

Tanlanma o'rtachasi va tanlanma dispertsiyasini toping.

3. $n=26$ hajmli tanlanma bo'yicha tanlanma dispersiyasining $D_T = 3$ bahosi topilgan. Bosh to'plam dispersiyasining siljimagan bahosini toping.

4. n=10 hajmli tanlanmaning berilgan taqsimoti bo'yicha tanlanma dipersiyasini toping.

x_i	102	104	108
n_i	2	3	5

5. Ushbu n=100 hajmli tanlanmaning berilgan taqsimoti bo'yicha tanlanma dispersiyasini toping.

x_i	156	160	164	168	172	176	180
n_i	10	14	26	28	12	8	2

10-ma'ruza. Nuqtaviy va intervalli baholar. Ishonchli intervallari.

Tayaich iboralar. *Nuqtaviy baho, intervalli baho, bahoning ishonchliligi, bahoning aniqligi, ishonchlilik intervali,*

*Reja***.**

10.1. Nuqtaviy va intervalli baholar.
10.2. Ishonchli ehtimol va ishonchli interval.
10.3. Normal taqsimotning noma'lum parametrlari uchun intervalli baholar.

Faraz qilaylik, bosh to'plam X belgisining taqsimot funksiyasi $F(x,\theta)$ bo'lib, θ noma'lum parametr bo'lsin. Bosh to'plamdan olingan tanlanmaning kuzatilgan qiymatlari $x_1, x_2, ..., x_n$ bo'lsin.

1-ta'rif. Tanlanmadan tuzilgan ixtiyoriy $L(x_1, x_2, ..., x_n)$ funksiyaga *statistika* deyiladi.

Nuqtaviy baholashda taqsimot funksiyaning noma'lum θ parametri uchun shunday $L(x_1, x_2, ..., x_n)$ statistika qidiriladiki, $L(x_1, x_2, ..., x_n)$ ni θ parametr uchun taqribiy qiymat deb olinadi. Bu holda $L(x_1, x_2, ..., x_n)$ statistika θ parametrning bahosi deyiladi.

2-ta'rif. Agar noma'lum parametr bitta $\tilde{\theta}$ son bilan baholansa, u holda bu baho nuqtaviy baho deyiladi.

Yuqorida tanishgan statistik baholar: tanlanma o'rtachasi, tanlanma "tuzatilgan" dispersiyasi, moda, mediana, variasiya qulochi va boshqalar nuqtaviy baho hisoblanadi.

Tajribalar soni juda katta bo'lsa, nuqtaviy bahoning qiymati odatda noma'lum parametrga yaqin bo'ladi. Ammo, kuzatishlar soni kam bo'lsa, $\tilde{\theta}$ nuqtaviy baho va θ parametr orasidagi farq sezilarli darajada bo'lishi mumkin. Bunday hollarda θ parametrni baholash uchun *intervalli baholardan* foydalanish maqsadga muvofiq hisoblanadi.

3-ta'rif. Ikkita son (interval chetlari) bilan aniqlanadigan baho intervalli baho deb ataladi.

Intervalli bahoda bahoning aniqliligi va ishonchliligi tushunchalarini kiritishimiz kerak bo'ladi. Buni quyida ko'rib chiqamiz.

Tanlanma ma'lumotlari asosida topilgan $\tilde{\theta}$-statistik xarakteristika θ parametrning bahosi bo'lsin. θ ni o'zgarmas son deb faraz qilamiz. Ma'lumki, $\tilde{\theta}$ ning aniqligi yuqori bo'lgan sari $|\theta-\tilde{\theta}|$ ning qiymati kamayib boradi, ya'ni $|\theta-\tilde{\theta}|<\delta$ $(\delta>0)$ tengsizlikda δ qancha kichik bo'lsa, baho shuncha aniq bo'ladi. Shu sababli, δ *bahoning aniqligi* deb ataladi.

$|\theta-\tilde{\theta}|<\delta$ tengsizlikning bajarilish ehtimoli γ θ parametrning $\tilde{\theta}$ baho bo'yicha *ishonchliligi*(*ishonchlilik ehtimoli*) deyiladi. Bu yerda, $P\left(|\theta-\tilde{\theta}|<\delta\right)=\gamma$. Ko'p hollarda, ishonchlilik oldindan beriladi. Masalan, 0,95; 0,99; 0,999 va hokazo.

$P\left(\left|\theta-\tilde{\theta}\right|<\delta\right)=\gamma$ ehtimollikni quyidagicha yozib olamiz:

$$P\left(\tilde{\theta}-\delta<\theta<\tilde{\theta}+\delta\right)=\gamma . \qquad (1)$$

Bu munosabatni quyidagicha tushunish kerak: $\left(\tilde{\theta}-\delta,\tilde{\theta}+\delta\right)$ interval θ noma'lum parametrni o'z ichiga olish (qoplash) ehtimoli γ ga teng.

$\left(\tilde{\theta}-\delta,\tilde{\theta}+\delta\right)$ interval noma'lum parametrni berilgan γ ishonchlilik bilan qoplovchi *ishonchlilik intervali* deb ataladi.

1-eslatma. $\left(\tilde{\theta}-\delta,\tilde{\theta}+\delta\right)$ interval tasodifiy chetki nuqtalarga ega, chunki turli tanlanmalar uchun $\tilde{\theta}$ ning qiymatlari turlicha bo'ladi. Shu sababli, tanlanma o'zgarsa $\left(\tilde{\theta}-\delta,\tilde{\theta}+\delta\right)$ intervalning chetki nuqtalari ham o'zgaradi.

Ishonchlilik intervallarni topish qanday amalga oshirilishi bilan normal taqsimot qonuniga bo'ysinuvchi tasodifiy miqdorlar misolida tanishib chiqamiz.

Bosh to'plamning X belgisi normal taqsimlangan bo'lsin. Ma'lumki, bu taqsimotni ikkita parametr: a va σ aniqlaydi. Faraz qilamiz ulardan biri, σ-o'rtacha kvadratik chetlanish, ma'lum ikkinchisi, a-matematik kutilma, noma'lum bo'lsin. Bu taqsimotning matematik kutilmasi a uchun ishonch intervalini γ ishonch bilan δ aniqlikda topamiz.

Tanlanma o'rtachasi $\overline{x_T}$ ni $\overline{X_T}$ tasodifiy miqdor sifatida qaraymiz. X belgi normal taqsimlanganligi sababli tanlanma o'rtacha ham normal taqsimlangan bo'ladi. shu bilan birga, $\overline{X_T}$ ning parametrlari quyidagicha: $M(\overline{X_T})=a;$ $\quad D(\overline{X_T})=\frac{\sigma^2}{n}$.

$P\left(\left|\overline{X_T}-a\right|<\delta\right)=\gamma$ munosabat o'rinli bo'lsin.U holda $P\left(\left|X-a\right|<\delta\right)=2\Phi\left(\frac{\delta}{\sigma}\right)$ formuladan foydalanib, X ni $\overline{X_T}$ bilan σ ni esa $\sigma(\overline{X_T})=\sigma/\sqrt{n}$ bilan almashtirsak quyidagi munosabatni hosil qilamiz:

$$P\left(\left|\overline{X_T}-a\right|<\delta\right)=2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)=2\Phi(t), \qquad (2)$$

bu erda $t=\frac{\delta\sqrt{n}}{\sigma}$. Bundan $\delta=\frac{t\sigma}{\sqrt{n}}$ bo'ladi. U holda (2) quyidagi ko'rinishni oladi:

$$P\left(\left|\overline{X_T}-a\right|<\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t), \text{ yoki } P\left(\overline{X_T}-\frac{t\sigma}{\sqrt{n}}<a<\overline{X_T}+\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t). \quad (3)$$

Shunday qilib, ishonch intervali $\overline{X_T}$ ni $\overline{x_T}$ ga almashtirganimizdan so'ng $\left(\overline{x_T}-\frac{t\sigma}{\sqrt{n}};\overline{x_T}+\frac{t\sigma}{\sqrt{n}}\right)$ dan iborat bo'ladi. Bundan $\left(\overline{x_T}-\frac{t\sigma}{\sqrt{n}};\overline{x_T}+\frac{t\sigma}{\sqrt{n}}\right)$ tasodifiy interval a parametrni $\gamma=2\Phi(t)$ ehtimol bilan $\frac{t\sigma}{\sqrt{n}}$ aniqlikda qoplashi kelib chiqadi.

(3) dan quyidagi xulosalarni chiqaramiz: tanlanma hajmining ortishi baholash aniqligi oshishiga olib keladi; agar γ ishonchlilik orttirilsa, t parametr ortadi va bu esa baholash aniqligi kamayishiga olib keladi.

1-Misol. X tasodifiy miqdor normal taqisimlangan bo'lib uning o'rtacha kvadratik chetlanishi $\sigma=3$. Tanlanma hajmi $n=36$ va bahoning ishonchliligi $\gamma=0,95$ bo'lsin. Noma'lum parametr a-matematik kutilmaning $\overline{x_T}$-tanlanma o'rtacha bo'yicha ishonchlilik intervallarini toping.

Yechish. Jadvaldan foydalanib t ni topamiz, ya'ni $2\Phi(t)=0,95 \Rightarrow \Phi(t)=0,475 \Rightarrow t=1,96$. Bahoning aniqligi: $\delta=\dfrac{t\sigma}{\sqrt{n}}=\dfrac{1,96\cdot 3}{\sqrt{36}}=0,98$.

U holda ishonchlilik intervali: $(\overline{x_T}-0,98;\ \overline{x_T}+0,98)$.

Berilgan $\gamma=0,95$ ishonchlilikni quyidagicha tushunish kerak: agar yetarlicha ko'p sondagi tanlanmalar olingan bo'lsa, u holda ularning 95%i shunday ishonchli intervallarni aniqlaydiki, bu intervallar parametrni haqiqatan ham o'z ichiga oladi; 5% hollardagina parametr interval chegarasidan tashqarida yotishi mumkin.

2-eslatma. Agar matematik kutilmani oldindan berilgan δ aniqlik va γ ishonchlilik bilan baholash talab qilinsa, u holda bu aniqlikni beradigan tanlanmaning minimal hajmi

$$n=\frac{t^2\sigma^2}{\delta^2} \qquad (4)$$

formuladan topiladi.

Bosh to'plamning X belgisi normal taqsimlangan va uning a-matematik kutilmasini $\overline{x_T}$-tanlanma o'rtachasi orqali baholashda σ-o'rtacha kvadratik chetlanish noma'lum bo'lsin. U holda

$$\overline{x_T}-t(\gamma,n)\frac{s}{\sqrt{n}}<a<\overline{x_T}+t(\gamma,n)\frac{s}{\sqrt{n}}. \qquad (7)$$

interval a uchun ishonch intervali bo'lib xizmat qiladi. Bu erda s - "tuzatilgan" o'rtacha kvadratik chetlanish; $t(\gamma,n)$ esa berilgan n va γ bo'yicha maxsus jadvaldan topiladi.Bunday jadvallar ehtimollar nazariyasi va matematik statistikaga oid adabiyotlarda beriladi.

2-Misol. Bosh to'plamdan $n=10$ hajmli tanlanma olingan va quyidagi statistik taqsimot tuzilgan:

x_i:	-2	1	2	3	4	5
n_i:	2	1	2	2	2	1

Bosh to'plamning X belgisi normal taqsimlangan bo'lsa, uning a-matematik kutilmasi uchun $\overline{x_T}$ bo'yicha $\gamma=0,95$ ishonchlilik bilan ishonchli intervalni toping.

Yechish. Tanlanma o'rtachasini va "tuzatilgan" o'rtacha kvadratik chetlanishni mos ravishda quyidagi formulalardan topamiz:

$$\overline{x_T}=\frac{\sum_{i=1}^{6}n_ix_i}{n}, \qquad s=\sqrt{\frac{\sum_{i=1}^{6}n_i(x_i-\overline{x_T})^2}{n-1}}.$$

U holda: $\overline{x_T}=2$, $s=2,4$. Jadvaldan $\gamma=0,95$ va $n=10$ larga mos $t(\gamma,n)=2,26$ ni topamiz. Topilganlarni (7) ifodaga qo'yib: $0,3<a<3,7$ ishonchli intervalni hosil qilamiz. Bu interval noma'lum a-matematik kutilmani $\gamma=0,95$ ishonchlilik bilan qoplaydi.

Bosh to'plamning o'rganilatgan X son belgisi normal taqsimlangan bo'lsin. Uning σ-o'rtacha kvadratik chetlanishi uchun tanlanma ma'lumotlari bo'yicha γ ehtimol bilan ishonch intervali topish talab qilinsin.

Ma'lumki, tanlanmaning s^2- "tuzatilgan" dispersiyasi σ^2-bosh to'plam dispersiyasi uchun siljimagan bahodir. Shu sababli, σ-paremetrni s orqali baholaymiz. Buning uchun

$$P\left(|\sigma - s| < \delta\right) = \gamma \text{, yoki } P\left(s-\delta < \sigma < s+\delta\right) = \gamma$$

munosabat bajarilishini talab qilamiz. Tayyor jadvaldan foydalanish uchun $s-\delta<\sigma<s+\delta$ qo'sh tengsizlikni teng kuchli

$$s\left(1-\frac{\delta}{s}\right) < \sigma < s\left(1+\frac{\delta}{s}\right)$$

tengsizlik bilan almashtiramiz. $q = \delta/s$ belgilashdan so'ng

$$s(1-q) < \sigma < s(1+q), \qquad (q<1) \tag{8}$$

ishonch intervalini hosil qilamiz. Agar $q>1$ bo'lsa ishonch intervali quyidagi ko'rinishda bo'ladi:

$$0<\sigma<s(1+q). \tag{9}$$

bu erda q-n va γ bo'yicha maxsus jadvaldan topiladi.

3-Misol. Bosh to'plamning X son belgisi normal taqsimlangan va $n=50$ hajmli tanlanmaning "tuzatilgan" dispersiyasi: $s=1{,}5$ bo'lsin. σ-noma'lum parametrni $\gamma=0{,}95$ ishonchlilik bilan qoplaydigan ishonch intervalini toping.

Yechish. Jadvaldan $n=50$ va $\gamma=0{,}95$ qiymatlarga mos $q=0{,}21$ ni topamiz. Bu erda $q<1$ bo'lgani uchun (8) tengsizlikdan foydalanib, $1{,}185<\sigma<1{,}815$ ishonchlilik intervalini topamiz.

O'z-o'zini tekshirish uchun savollar.

1. Statistik baho ta'rifini bering.
2. Ishonchlilik ehtimoli va ishonchli interval tushunchalarini ta'riflang.
3. Normal taqsimlangan bosh to'plam matematik kutilishi uchun ishonchli intervallarni keltiring.
4. Ishonchli intervallarda qatnashuvchi parametrlarni izohlang.

Mustaqil yechish uchun masalalar.

1. Bosh to'plamning normal taqsimlangan X son belgisining noma'lum a-matematik kutilmasini $\gamma=0{,}95$ ishonchlilik bilan qoplaydigan ishonch intervalini toping. Bunda o'rtacha kvadratik chetlanish $\sigma=4$, tanlanma o'rtacha $\overline{x_T}=10{,}2$ va tanlanma hajmi $n=16$ deb olinsin.

2. 10 ta erkli o'lchashlar natijasida sterjen uzunligi (mm) uchun quyidagi ma'lumotlar olingan: 23, 24, 23, 25, 25, 26, 26, 25, 24, 25. O'lchash xatoligini normal taqsimlangan deb faraz qilib, sterjen uzunligining matematik kutilmasi uchun $\gamma=0{,}95$ ishonchlilik bilan ishonch intervalini toping.

3. Bosh to'plamning normal taqsimlangan X belgisining a-matematik kutilmasini tanlanma o'rtacha bo'yicha $\gamma=0{,}925$ ishonchlilik va $\delta=0{,}2$ aniqlik bilan baholash uchun tanlanmaning minimal hajmini toping. O'rtacha kvadratik chetlanishni $\sigma=1{,}5$ deb oling.

4. Bosh to'plamdan $n=12$ hajmli tanlanma olingan:

x_i	-0,5	-0,4	-0,2	0	0,2	0,6	0,8	1	1,2	1,5
n_i	1	2	1	1	1	1	1	1	2	1

Bosh to'plamning normal taqsimlangan X belgisining a-matematik kutilmasini $\gamma = 0,95$ ishonchlilik bilan qoplaydigan ishonch oralig'ini toping.

5. Bosh to'plamning X-son belgisi normal taqsimlangan. n hajmli tanlanma bo'yicha "tuzatilgan" o'rtacha kvadratik chetlanishi s topilgan.

a) bosh to'plam o'rtacha kvadratik chetlanishi σ ni;

b) bosh to'plam dispersiyasini $\gamma = 0,99$ ishonchlilik bilan qoplaydigan ishonch oralig'ini toping, bunda $n = 10$, $s = 5,1$.

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.
5. Крамер Н.Ш. Теория вероятностей и математическая статистика. Учебник М.,2001.

IV. Statistik gipotezalarni tekshirish.

11-ma'ruza. Statistik gipotezalar. Statistik kriteriy.

Tayanch iboralar. Statistik gipoteza, oddiy gipoteza, murakkab gipoteza, statistik kriteriy, kuzatiladigan qiymat, kritik nuqtalar, muhimlilik darajasi, kriteriy quvvati, Normal taqsimot, normalangan normal taqsimot, gipoteza, kritik soha, kritik nuqtalar, kriteriy, Laplas funksiyasi.

Reja.

10.1. Statistik gipotazalar.
10.2. I va II tur xatoliklar.
10.3. Statistik kriteriy.
10.4. Ikki tomonli, o'ng va chap tomonli kritik sohalarni aniqlash va unda asosiy gipotezani tekshirish.

Amaliyotda, texnikada va iqtisodiyotda ko'pincha tasodifiylik bilan bog'liq bo'lgan biror faktni aniqlashtirish uchun statistik usul bilan tekshirish mumkin bo'lgan gipotezalarga tayanib ish ko'riladi.

Ma'lumki, har qanday ilmiy asoslangan farazni gipoteza deb aytishimiz mumkin, ammo har qanday gipotezani stastistik gipoteza deb ayta olmaymiz, chunki uning alohida ajralib turadigan xususiyatlari bor, bu xususiyatlarni alohida ta'kidlash uchun biz quyidagi ta'rifni keltiramiz.

1-ta'rif. Statistik gipoteza deb, kuzatilayotgan tasodifiy miqdorning taqsimot qonuni, yoki agar tasodifiy miqdor bo'ysinadigan taqsimot qonunning ko'rinishi ma'lum bo'lsa, u holda bu taqsimot qonunning noma'lum parametrlari haqidagi gipotezaga aytiladi.

Masalan, quyidagi gipotezalar statistik gipotezalarga misol bo'la oladi:

1. Bir xil ishlab chiqarish sharoitlarida bir xil ishni bajarayotgan ishchilarning mehnat unumdorligi normal taqsimot qonun bo'yicha taqsimlangan;
2. Parallel ishlayotgan stanoklarda tayyorlanayotgan bir xil turdagi detallarning o'rtacha o'lchamlari bir-biriga teng;
3. Normal taqsimot qonuniga bo'ysinuvchi ikki to'plamning dispersiyalari o'zaro teng;

1-gipotezada taqsimotning ko'rinishi haqida, 2 va 3-gipotezalarda esa parametrlar haqida faraz qilingan.

«Ertaga yomg'ir yog'adi», «Bu yil mo'l hosil olamiz» kabi gipotezalar statistik gipotezalar bo'la olmaydi, chunki ularda na taqsimot qonunining ko'rinishi haqida, na uning parametrlari haqida so'z boradi.

Oldinga surilgan gipoteza tanlanma natijalarga asoslanib tekshiriladi va natijada yoki qabul qilinishi yoki rad qilinishi mumkin.

Asosiy (yoki nolinchi) gipoteza deb ilgari surilgan H_0-gipotezaga, konkurent (yoki alternativ) gipoteza deb, asosiy gipotezaga zid bo'lgan H_1-gipotezaga aytiladi.

Masalan, "X tasodifiy miqdor Puasson taqsimot qonuniga bo'ysunadi" gipotezasi surilgan bo'lsin.

Bu holda:

$H_0: \quad P(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \quad (\lambda>0, k=0,1,2,3,...); \quad H_1: P(X=k)\neq\frac{\lambda^k e^{-\lambda}}{k!}$.

Faqat bitta da'voni o'z ichiga olgan gipoteza *oddiy gipoteza*; bittadan ortiq sondagi da'volarni o'z ichiga olgan gipoteza esa *murakkab gipoteza* deyiladi.

Masalan, agar X tasodifiy miqdor ko'rsatkichli taqsimot:

$$F(x)=\begin{cases} 0, & agar \quad x<0, \\ 1-e^{-\lambda x}, & agar \quad x\geq 0. \end{cases}$$

qonuniga bo'ysinib, uning λ parametri noma'lum bo'lsin. U holda $H_0: \quad \lambda=2,5$ gipoteza oddiy gipoteza; $H_1: \quad \lambda>2,5$ gipoteza esa murakkab gipotezadir.

Ilgari surilgan gipoteza to'g'ri yoki noto'g'ri bo'lishi mumkin, shu sababli uni tekshirib ko'riladi va so'ngra xulosa chiqariladi. Gipotezani tekshirish natijasida ikki turdagi xatolikka yo'l qo'yilishi mumkin.

Agar to'g'ri gipoteza rad etilsa, qilingan xatolikni I tur xatolik, agar noto'g'ri gipoteza qabul qilinsa qilingan xatolik II tur xatolik deb ataladi. Bu xatoliklarni jadvalda quyidagicha tasvirlash mumkin.

H_0-gipoteza	to'g'ri	noto'g'ri
Rad qilindi	I tur xatolik	To'g'ri qaror
Qabul qilindi	To'g'ri qaror	II tur xatolik

Amaliyotda I va II tur xatoliklarning oqibatlari har xil bo'lishi mumkin. Masalan, agar samolyotga "uchishga ruxsat berilsin" degan to'g'ri qaror rad etilgan bo'lsa, u holda bu I tur xatolik bo'lib, bunday xatolik moddiy zararga olib kelishi mumkin; agar samolyotning nosozligiga qaramasdan "uchishga ruxsat berilsin" degan noto'g'ri qaror qabul qilinsa, u holda bu II tur xatolik bo'lib, bunday xatolik halokatga olib kelishi mumkin.

Albatta, I tur xatolik II tur xatolikga qaraganda og'irroq oqibatlarga olib keladigan misollar ham keltirish mumkin.

To'g'ri qarorni ikki holda qabul qilish mumkin:

1) agar ilgari surilgan gipoteza haqiqatan ham to'g'ri bo'lsa, gipoteza qabul qilinadi;

2) agar ilgari surilgan gipoteza haqiqatan ham noto'g'ri bo'lsa, gipoteza qabul qilinmaydi.

I tur xatolikka yo'l qo'yish ehtimoli α bilan belgilanadi; u *muhimlilik (qiymatdorlik) darajasi* deb ataladi. Ko'p hollarda: $\alpha=0,05; \quad \alpha=0,01$.

Biz ma'lum taqsimot qonuniga bo'ysinuvchi belgining noma'lum parametrlari haqida ilgari surilgan gipoteza statistik usulda qanday tekshirilishini ko'rib chiqamiz.

Asosiy gipoteza ilgari surilgandan so'ng, uning to'g'ri yoki noto'g'ri ekanligini tekshirib ko'rish kerak bo'ladi. Shu maqsadda maxsus tanlangan, aniq, yoki taxminiy taqsimoti ma'lum bo'lgan tasodifiy miqdor ishlatiladi. Bu tasodifiy miqdorni K bilan belgilaymiz.

2-ta'rif. Statistik kriteriy (yoki oddiygina kriteriy) deb, asosiy gipotezani tekshirish uchun xizmat qiladigan K-tasodifiy miqdorga aytiladi.

Masalan, agar normal taqsimot qonuniga ega X va Y bosh to'plamlarning dispersiyalari tengligi haqidagi gipoteza tekshirilayotgan bo'lsa, u holda K kriteriy sifatida "tuzatilgan" tanlanma dispersiyalar nisbati olinadi:

$$K = \frac{s_x^2}{s_y^2} \quad . \qquad (1)$$

Turli tajribalarda dispersiyalar har xil, oldindan ma'lum bo'lmagan qiymatlar qabul qilganligi uchun K tasodifiy miqdor bo'lib, u Fisher-Snedekor qonuni bo'yicha taqsimlangan.

Gipotezani tekshirish uchun kriteriyga kirgan miqdorlarning xususiy qiymatlari tanlanma bo'yicha hisoblanadi va shunday qilib kriteriyning kuzatiladigan (xususiy) qiymati hosil qilinadi.

K_{kuzat}*-kuzatiladigan qiymat* deb, statistik kriteriyning tanlanmalar bo'yicha hisoblangan qiymatiga aytiladi. Masalan, ikkita tanlanma asosida topilgan dispersiyalar: $s_1^2 = 20$ va $s_2^2 = 5$ bo'lsa, u holda

$$K_{kuzat} = F = \frac{s_1^2}{s_2^2} = \frac{20}{5} = 4$$

Tanlangan K kriteriyning mumkin bo'lgan barcha qiymatlari to'plami kesishmaydigan ikkita qism to'plamlarga ajratiladi:

$$K = K^- \cup K^+, K^- \cap K^+ = \varnothing .$$

Ulardan biri H_0-asosiy gipoteza rad qilinadigan, ikkinchisi esa asosiy gipoteza qabul qilinadigan qiymatlarini o'z ichiga oladi.

3-ta'rif. Kritik soha deb, kriteriyning H_0-asosiy gipotezani rad qiladigan qiymatlar to'plamiga aytiladi.

4-ta'rif. Gipotezani qabul qilinish sohasi deb, kriteriyning asosiy gipotezani qabul qiladigan qiymatlar to'plamiga aytiladi.

Statistik gipotezalarni tekshirishning asosiy printsiplari E. Neyman, K. Pirson va boshqa matematiklar tomonidan ishlab chiqilgan bo'lib, bu printsipni quyidagicha ta'riflash mumkin: agar kriteriyning kuzatiladigan qiymati kritik sohaga tegishli bo'lsa, asosiy gipoteza rad qilinadi, agar kriteriyning kuzatilayotgan qiymati gipotezaning qabul qilinish sohasiga tegishli bo'lsa, asosiy gipoteza qabul qilinadi.

Kriteriy bir o'lchovli tasodifiy miqdor bo'lgani uchun uning mumkin bo'lgan barcha qiymatlari to'plami biror intervaldan iborat bo'ladi. Shu sababli, kritik soha va gipotezaning qabul qilinish sohasi ham intervaldan iborat bo'ladi, demak, ularni ajratib turuvchi nuqtalar to'g'risida gapirish mumkin.

5-ta'rif. Kritik nuqtalar deb, kritik sohani gipotezaning qabul qilinish sohasidan ajratib turuvchi nuqtalarga aytiladi.

Agar kritik soha $K > k_{kr}$ tengsizlik bilan aniqlansa, u holda uni *o'ng tomonli* kritik soha, tengsizlik aksincha bo'lsa *chap tomonli* kritik soha deyiladi. Agar kritik soha $K < k_{kr}', K > k_{kr}'$ tengsizliklar bilan aniqlansa, u holda uni ikki tamonli kritik soha deyiladi.

Chap tomonli va ikki tomonli kritik sohalarni aniqlash o'ng tomonli kritik sohani topishga o'xshash bo'lganligi sababli biz faqat o'ng tomonli kritik sohani topish bilan tanishib chiqamiz.

Kritik sohani topish uchun kritik nuqtani aniqlash etarli. Bu nuqtani aniqlash uchun esa α ning qiymati berilishi kerak. So'ngra, quyidagi talabga asoslanib, k_{kr} nuqta topiladi: H_0-asosiy gipoteza o'rinli bo'lishi shartida tanlangan K kriteriyning k_{kr} nuqtadan katta bo'lishi ehtimoli α-muhimlilik darajasiga teng bo'lsin:

$$P(K > k_{kr}) = \alpha . \qquad (2)$$

Har bir kriteriy uchun (2) shartni qanoatlantiruvchi kritik nuqtalarni topish jadvallari mavjud.

Kritik nuqta topilgandan so'ng, $x_1, x_2, ..., x_n$ tanlanma ma'lumotlari bo'yicha kriteriyning kuzatish qiymati topiladi. Bunda agar $K > k_{kr}$ bo'lsa, u holda H_0 asosiy gipoteza rad qilinadi; agar $K < k_{kr}$ bo'lsa, u holda gipotezani rad qilishga asos yo'q deyiladi.

1-eslatma. H_0 gipoteza qabul qilingan bo'lsin. Shu bilan bu gipoteza isbotlandi deyish xato bo'ladi. Aslida "kuzatish natijalari H_0 gipotezaga mos keladi va demak, uni rad qilishga asos yo'q" deyish to'g'riroq bo'ladi.

Amalda gipotezani katta ishonch bilan qabul qilish uchun boshqa statistik usullar bilan tekshiriladi yoki tanlanma hajmi orttirilib tajriba takrorlanadi. Gipotezani qabul qilishdan ko'ra ko'proq uni rad qilishga harakat qilinadi. Haqiqatan, ma'lumki biror umumiy da'voni rad qilish bu uchun bu da'voga zid bo'lgan bitta misolni keltirish kifoya. Shu sababli *kriteriy quvvati* tushunchasi kritiladi.

6-ta'rif. Konkurent gipoteza to'g'ri bo'lganda kriteriyning kritik sohada bo'lish ehtimoli kriteriy quvvati deb ataladi.

Agar II tur xatolikka yo'l qo'yish ehtimoli β bo'lsa, u holda kriteriy quvvati $1-\beta$ ga teng bo'ladi. Bundan ko'rinadiki, quvvat qancha katta bo'lsa II tur xatolikka yo'l qo'yish ehtimoli shuncha kam bo'ladi. Yuqoridagi ta'riflardan ko'rinib turibdiki, α ning kamayishi β ning o'sishiga olib keladi, va aksincha. Masalan, $\alpha = 0$ bo'lsa, u holda barcha gipotezalar qabul qilinadi, jumladan noto'g'rilari ham. Shu sababli, ikkala parametrni bir paytda kamaytirib bo'lmaydi. I tur va II tur xatoliklarni kamaytirishning yagona yo'li tanlanma hajmini oshirishdir.

Statistik gepotezani tekshirish qanday amalga oshirilishini quyidagi misolda ko'rib chiqamiz.

Normal taqsimlangan ikki bosh to'plamning dispersiyalarni taqqoslash. Dispersiyalar haqidagi gipotezalar, ayniqsa texnikada muhim ahamiyatga ega, chunki tarqoqlik xarakteristikasi bo'lgan dispersiya mashina va uskunalarning, o'lchov asboblarining, texnologik protsesslarning aniqligini baholashda juda muhim ko'rsatkich hisoblanadi.

Normal taqsimlangan bosh to'plam dispersiyalarining tengligi haqida gipoteza ilgari surilsa kriteriy sifatida $F = \frac{S_x^2}{S_y^2}$ kattalik olinishni aytib o'tgan edik. Bunda F tasodifiy miqdor bo'ysinadigan Fisher-Snedekor taqsimotining erkinlik darajalari quyidagicha aniqlanadi: $k_1 = n_1 - 1, \quad k_2 = n_2 - 1$, bu erda n_1-hisoblanganda qiymati katta bo'lgan "tuzatilgan" dispersiyaga mos tanlanmaning hajmi, n_2-hisoblanganda qiymati kichik bo'lgan "tuzatilgan" dispersiyaga mos tanlanmaning hajmi. Kritik nuqta $k_{kr} = F_{kr}(\alpha; k_1, k_2)$ tenglik bilan jadvaldan aniqlanadi.

Misol. Normal taqsimlangan X va Y bosh to'plamlardan olingan $n_1 = 11$ va $n_2 = 14$ hajmli ikkita erkli tanlanma bo'yicha "tuzatilgan" dispersiyalar: $s_x^2 = 0,76$, $s_x^2 = 0,38$ topilgan. $\alpha = 0,05$ muhimlilik darajasida quyidagi gipotezani tekshiring: $H_0: \quad D(X) = D(Y)$; $H_1: \quad D(X) > D(Y)$.

Yechish. Gipotezani tekshirish uchun $F = \frac{s_x^2}{s_y^2}$ kriteriyni tanlaymiz. U holda

$$K_{kuzat} = F = \frac{s_1^2}{s_2^2} = \frac{0,76}{0,38} = 2.$$

Fisher-Snedekor taqsimotining kritik nuqtalar jadvalidan $\alpha = 0,05$, $k_1 = n_1 - 1 = 10$, $k_2 = n_2 - 1 = 13$ bo'yicha $k_{kr} = F_{kr}(0,05; 10,13) = 2,67$ kritik nuqtani topamiz. 2<2,67 , ya'ni $K_{kuzat} < k_{kr}$ bo'lgani uchun gipotezani rad qilishga asos yo'q.

O'z-o'zini tekshirish uchun savollar.

1. Statistik kriteriyni tushuntiring.
2. Ikki tomonli kritik soha qanday aniqlanadi?
3. Chap va o'ng tomonli kritik soha qanday aniqlanadi?
4. Statistik gipoteza ta'rifini bering. Misollar keltiring.
5. Statistik gipotezalarning turlarini ayting.
6. Birinchi va ikkinchi tur xatoliklar nimalardan iborat?
7. Muhimlilik darajasi va kriteriy quvvatini tushuntiring.
8. Kritik soha va kritik nuqta tushunchalarini ayting.

Mustaqil ishlash uchun misollar.

1. Normal taqsimlangan X va Y bosh to'plamlardan hajmlari mos ravishda n_1 va n_2 bo'lgan ikkita erkli tanlanma ajratib olingan va ularning "tuzatilgan" dispersiyalari: s_x^2 va s_y^2 topilgan. α muhimlilik darajasida $H_0: \quad D(X) = D(Y)$ asosiy gipotezani tekshiring, bunda H_1: $D(X) > D(Y)$:

a) $n_1 = 10, \quad n_2 = 15, \quad s_x^2 = 3,6, \quad s_y^2 = 2,4, \quad \alpha = 0,05$;

b) $n_1 = 13, \quad n_2 = 18, \quad s_x^2 = 3,6, \quad s_y^2 = 4,8, \quad \alpha = 0,01$;

v) $n_1 = 16, \quad n_2 = 12, \quad s_x^2 = 0,72, \quad s_y^2 = 2,4, \quad \alpha = 0,05$;

s) $n_1 = 14, \quad n_2 = 10, \quad s_x^2 = 1,6, \quad s_y^2 = 2,4, \quad \alpha = 0,01$.

2. Normal taqsimlangan ikkita bosh to'plamdan hajmlari mos ravishda $n = 30$ va $m = 20$ bo'lgan erkli tanlanmalar olinib $\bar{x} = 600, \bar{y} = 550$ o'rtachalari topilgan. Bosh to'plam dispersiyalari: $D(X) = 120, \quad D(Y) = 100$. $\alpha = 0,05$ -muhimlilik darajasida $H_0: \quad M(X) = M(Y)$, $H_1: \quad M(X) \neq M(Y)$ gipotezalarni tekshiring.

3. Normal taqsimlangan ikkita bosh to'plamdan hajmlari mos ravishda $n=40$ va $m=50$ bo'lgan erkli tanlanmalar olinib $\overline{x}=1050, \overline{y}=1275$ o'rtachalari topilgan. Bosh to'plam dispersiyalari: $D(X)=120,\quad D(Y)=100$. $\alpha=0,01$-muhimlilik darajasida $H_0:\quad M(X)=M(Y)$, $H_1:\quad M(X)>M(Y)$ gipotezalarni tekshiring.

4. Normal taqsimlangan ikkita bosh to'plamdan hajmlari mos ravishda $n=50$ va $m=50$ bo'lgan erkli tanlanmalar olinib $\overline{x}=1050, \overline{y}=875$ o'rtachalari topilgan. Bosh to'plam dispersiyalari: $D(X)=120,\quad D(Y)=100$. $\alpha=0,01$-muhimlilik darajasida $H_0:\quad M(X)=M(Y)$, $H_1:\quad M(X)<M(Y)$ gipotezalarni tekshiring.

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.
5. Sirajdinov S.X., Mamatov M.M. Ehtimollar nazariyasi va matematik statistika. Toshkent-1980.

V. Regression tahlil elementlari.

12-ma'ruza. Funksional, statistik va korrelyatsion bog'lanishlar. Korrelyatsion jadval. Korrelyatsiya nazariyasining ikki asosiy masalasi.

Tayanch iboralar. *Funksional bog'lanish, statistik bog'lanish, korrelyatsion bog'lanish, korrelyatsion panjara, shartli o'rtacha, tanlanma regressiyasi, tanlanma regressiya tenglamasi.*

Reja.

12.1. Funksional, statistik va korrelyatsion bog'lanishlar.
12.2. Shartli o'rtacha.
12.3. Korrelyatsion jadval.
12.4. Regressiya tanlanma tenglamasi va tanlanma chizig'i.
12.5. Korrelyatsiya nazariyasining ikki asosiy masalasi.

Kundalik faoliyatimizdagi ko'pgina amaliy masalalarda, tajribalarda o'rganilayotgan Y belgining (tasodifiy miqdorning) bitta yoki bir nechta boshqa belgilarga (tasodifiy miqdorlarga) bog'liqligini aniqlash va baholash talab qilinadi. Avvalam bor Y belgining bitta X tasodifiy miqdorga bog'liqligini o'rganamiz.

Ikki belgi funksional bog'lanish bilan, yoki statistik bog'lanish bilan bog'langan, yoki umuman erkli bo'lishi mumkin.

1-ta'rif. Agar X belgining har bir mumkin bo'lgan qiymatiga Y belgining bitta mumkin bo'lgan qiymati mos kelsa, u holda Y X belging funksiyasi deyiladi:

$$Y = f(X).$$

Misollar.

1. X diskret tasodifiy miqdorning taqsimoti:

X	2	3
p	0,6	0,4

$Y = X^2$ funksiyaning taqsimoti topilsin.

Y*echish.* Y ning mumkin bo'lgan qiymatlarini topamiz: $y_1 = 4$ $y_2 = 9$. U holda Y ning taqsimoti:

Y	4	9
p	0,6	0,4

2. X uzluksiz tasodifiy miqdor normal taqsimlangan bo'lib, $M(X) = a = 2$ va $\sigma(X) = 0,5$ bo'lsa, $Y = 3X + 1$ chiziqli funksiyaning zichlik funksiyasini toping.

Y*echish.* Y ning sonli xarakteristikalarini topamiz:

$$M(Y) = 3\cdot 2 + 1 = 7, \quad \sigma(Y) = 3\cdot 0,5 = 1,5.$$

U holda Y ning zichlik funksiyasi: $g(y) = \dfrac{1}{1,5\sqrt{2\pi}} e^{-\frac{(y-7)^2}{2\cdot(1,5)^2}}$.

Funksional bog'lanishlar aniq va tabiiy fanlar: matematika, fizika, ximiya va boshqa fanlarda ayniqsa yaqqol kuzatiladi.

Masalan, termometrdagi simob ustunining balandligi X havo harorati Y haqida aniq va bir qiymatli ma'lumot beradi; aylana radiusi R va uning uzunligi C orasida $C = 2\pi R$ geometriyadan ma'lum bo'lgan formula bilan aniqlangan funksional bog'lanish mavjuddir.

Iqtisodiy jarayonlarda, umuman jamiyatning boshqa sohalarida tasodifiy belgilar orasida qat'iy funksional bog'lanish kamdan-kam uchraydi. Buning asosiy sabablaridan biri belgilarga ta'sir etuvchi faktorlarning xilma-xilligi va tasodifiyligidir. Bu holatda belgilar orasidagi moslik statistik bog'lanish bo'lishi mumkin.

2-ta'rif. Agar miqdorlardan birining o'zgarishi ikkinchi miqdor taqsimotining o'zgarishiga olib kelsa, u holda bu ikki miqdor orasidagi bog'lanishga statistik bog'lanish deyiladi.

Masalan, agar $Y(Z_1, Z_2, V_1, V_2)$ va $X(Z_1, Z_2, U_1, U_2)$ (Z_i, V_i, U_i -tasodifiy faktorlar) lar berilgan bo'lsin. Bu holda Y va X lar orasidagi bog'lanish statistik bog'lanish deyiladi, chunki ularning har biri bog'liq bo'lgan tasodifiy faktorlar ichida umumiylari: Z_1, Z_2 va umumiy bo'lmaganlari: $V_{i,}$ U_i $(i=1,2)$ bor.

Statistik bog'lanishni matematik ifodalash murakkab, shu sababli uning xususiy hollaridan biri hisoblangan korrelyatsion bog'lanish bilan tanishib chiqamiz.

3-ta'rif. Agar bir-biriga statistik bog'lanishda bo'lgan ikki miqdordan birining o'zgarishi ikkinchi miqdor o'rtacha qiymatining o'zgarishiga olib kelsa, u holda bunday statistik bog'lanish *korrelyatsion bog'lanish* deb ataladi.

Bir-biri bilan korrelyatsion bog'lanishda bo'lgan tasodifiy miqdorlarga misollar keltiramiz.

1. Mehnat unumdorligi X va jami ishlab chiqarilgan mahsulot Y;
2. Yig'ib olingan hosil miqdori Y va ishlatilgan o'g'itlar miqdori X;
3. Jami mahsulot miqdori X va korxonaning ish haqi fondi Y;
4. Sarflangan kapital mablag'lar X va shu mablag'lardan olingan sof foyda Y;
5. Korxonaning texnika bilan qurollanganlik darajasi X va mehnat unumdorligi ko'rsatkichi Y.

Yuqoridagi ta'rifdan ko'rinib turibdiki, korrelyatsion bog'lanishni matematik ifodalash, ya'ni $y = f(x)$ ko'rinishda yozish, uchun *shartli o'rtacha* tushunchasini kiritishimiz kerak.

4-ta'rif. $X = x$ qiymatga mos keluvchi Y ning kuzatilgan qiymatlarining arifmetik o'rtachasini $\overline{y_x}$ -shartli o'rtacha deb ataymiz.

Xuddi shunday usulda $\overline{x_y}$ -shartli o'rtacha tushunchasi ham aniqlanadi.

5-ta'rif. $Y = y$ qiymatga mos keluvchi X ning kuzatilgan qiymatlari arifmetik o'rtachasini $\overline{x_y}$ -shartli o'rtacha deb ataymiz.

3-*Misol.* X miqdorning $x_1 = 5$ qiymatiga Y miqdorning $y_1 = 6,\quad y_2 = 7,\quad y_3 = 8$ qiymatlari mos keldi. $\overline{y_{x_1}} = ?$

Yechish. $\overline{y_{x_1}} = \dfrac{y_1 + y_2 + y_3}{3} = \dfrac{6+7+8}{3} = 7$.

Agar X va Y tasodifiy miqdorlar (belgilar) ustida kuzatishlar o'tkazilgan bo'lib, kuzatishlar natijalari mos ravishda $(x_1, y_1),\ (x_2, y_2),\ \ldots,\ (x_n, y_n)$ lardan iborat bo'lsa, u holda X va Y orasidagi bog'lanishni (munosabatni) ushbu jadval ko'rinishida ifodalash mumkin.

X	x_1	x_2	...	x_n
Y	y_1	y_2	...	y_n

Agar yuqoridagi jadvalda x_i va y_i lar turli qiymatlarini qabul qilsa, u holda shartli o'rtacha tushunchasidan foydalanmaymiz.

Agar kuzatishlar soni ko'p, ya'ni x_i qiymat m_{x_i} marta, y_j qiymat m_{y_j} marta, (x_i, y_j) juftliklar $m_{x_i y_i}$ marta takrorlanishi mumkin bo'lsa, u holda yuqoridagi jadval o'rniga *korrelyatsion jadval* yoki *korrelyatsion panjara* deb ataluvchi jadval hosil bo'ladi. $m_{x_i}, m_{y_j}, m_{x_i y_j}$ lar mos ravishda $x_i, y_j, (x_i, y_j)$ larning chastotalari deyiladi. $m_{x_i y_j} = m_{ij}$ belgilash kiritib quyidagi jadvalni hosil qilamiz. Bu erda $\sum_i m_{ij} = m_{y_j}$, $\sum_j m_{ij} = m_{x_i}$, $\sum_i m_{x_i} = \sum_j m_{y_j} = n$.

Y / X	y_1	y_2	$\dots$	y_l	m_x
x_1	m_{11}	m_{12}	$\dots$	m_{1l}	m_{x_1}
x_2	m_{21}	m_{22}	$\dots$	m_{2l}	m_{x_2}
. . .	. . .	. . .	. . .	. . .	. . .
x_k	m_{k1}	m_{k2}	$\dots$	m_{kl}	m_{x_k}
m_y	m_{y_1}	m_{y_2}	$\dots$	m_{y_l}	n

Bu holatda shartli o'rtacha tushunchasidan foydalanishimiz zarur.

Korrelyatsion panjarada shartli o'rtacha topilishiga doir misol ko'rib chiqamiz.

Misol.

5. Berilgan jadvaldan foydalanib, tanlanma shartli o'rtacha-$\overline{y_x}$ ni toping.

X / Y	3	4	6	7	8	n_y
8	5	3	-	-	-	8
12	3	4	5	4	2	18
15	-	3	3	6	2	14
n_x	8	10	8	10	4	$n=30$

Yechish. Hisoblashlarni quyidagi jadvalga joylashtiramiz:

X / Y	3	4	6	7	8	n_y
8	5	3	-	-	-	8
12	3	4	5	4	2	18
15	-	3	3	6	2	14

n_x	8	10	8	10	4	$n=30$
$\overline{y_x}$	9,5	11,7	13,125	13,8	13,5	

Belgilar orasidagi korrelyatsion munosabatlar (bog'lanishlar) to'g'ri, teskari, to'g'ri chiziqli va egri chiziqli bo'lishi mumkin. Masalan, to'g'ri korrelyatsion bog'lanishda belgilardan birining ortishi (kamayishi) boshqasining o'rtachasi ortishiga (kamayishiga) olib keladi, teskari bog'lanishda esa aksincha va hakozo.

Masalan, daraxtning yoshi X ortib borishi bilan daraxtdagi xalqalar soni Y ortib boradi, havoning harorati X pasayishi bilan nafas olish tezligi Y kamayadi va h.k.

Y ning X ga korrelyatsion bog'liqligi deb, $\overline{y_x}$ shartli o'rtachaning x ga funksional bog'lanishiga aytiladi: $\overline{y_x} = f(x)$. Bu tenglama Y ning X ga *regressiya tanlanma tenglamasi* (ba'zida Y ning X ga regressiya tenglamasi), $f(x)$ funksiya esa Y ning X ga tanlanma regressiyasi (ba'zida regressiya funksiyasi) deb ataladi. Bu tenglama grafigi esa Y ning X ga *regressiya tanlama chizig'i* (ba'zida Y ning X ga regressiya chizig'i) deyiladi.

X ning Y ga regressiya tanlama tenglamasi va regressiya tanlama chizig'i ham yuqoridagiga o'xshash aniqlanadi: $\overline{x_y} = \varphi(y)$.

Korrelyatsiya nazariyasi belgilar orasidagi bog'lanishni o'rganish jarayonida asosan quyidagi ikki masalani hal qiladi.

1-masala. Belgilar orasidagi korrelyatsion bog'lanish formasini aniqlash, ya'ni regressiya funksiyasining ko'rinishini (chiziqli, chiziqsiz va h.k.) topish.

Agar $f(x)$ va $\varphi(y)$ regressiya funksiyalarining ikkalasi ham chiziqli bo'lsa, u holda X va Y belgilar orasidagi korrelyatsion bog'lanish *chiziqli,* aks holda esa *chiziqsiz* deyiladi.

2-masala. Korrelyatsion bog'lanish zichligini (kuchini) aniqlash.

Y belgining X belgiga korrelyatsion bog'lanishiining zichligi $X = x$ qiymatga mos Y ning mumkin bo'lgan qiymatlari $\overline{y_x}$ -shartli o'rtacha atrofida tarqoqligi darajasini baholaydi. Agar tarqoqlik katta bo'lsa, Y belgi X belgiga kuchsiz bog'langanligidan yoki ular orasida bog'liqlik yo'qligidan darak beradi. Aksincha, kichik tarqoqlik belgilar orasida ancha kuchli (zich) bog'liqlik borligini ko'rsatadi.

O'z-o'zini tekshirish uchun savollar.

1. Belgilar orasida qanday bog'lanishlar bo'lishi mumkin?
2. Korrelyatsion bog'lanish ta'rifini bering.
3. Korrelyatsion jadval qanday tuziladi?
4. Shartli o'rtacha qiymat ta'rifini bering.
5. Regressiya tanlanma tenglamasi, regressiya funksiyasi va regressiya tanlanma chizig'i ta'riflarini bering.
6. Korrelyatsiya nazariyasi ikki asosiy masalasini ayting.

Mustaqil yechish uchun masalalar.

1. Berilgan jadvaldan foydalanib, tanlanma shartli o'rtacha- $\overline{x_y}$ ni toping.

X / Y	4	4,5	5	5,5	6
8	5	3	-	-	-
10	2	4	5	4	3
13	-	1	1	2	2

2. Berilgan jadvaldan foydalanib, tanlanma shartli o’rtacha- $\overline{y_x}$ ni toping.

X / Y	3	3,5	4	4,5	5
7	5	3	-	-	-
9	2	3	5	3	1
13	-	1	1	2	2

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O’quv qo’llanma . T. 2008 y.
5. Крамер Н.Ш. Теория вероятностей и математическая статистика. Учебник М.,2001.

13 - ma'ruzalar. To'g'ri chiziqli regressiya tanlanma tenglamasi. Tanlanma korreliyatsiya koeffitsient.

Tayanch iboralar. *Regressiya, to'g'ri chiziqli regressiya, eng kichik kvadratlar usuli, to'g'ri chiziqli regressiya tanlanma koeffitsienti, burchak koeffitsien, korrelyatsion bog'lanish zichligi, tanlanma korrelyatsiya koeffitsienti.*

Reja.

1. To'g'ri chiziqli regressiya tanlanma tenglamasi va uning parametrlarini topish.
2. Korrelyatsion bog'lanish zichligi.
3. Korrelyatsiya tanlanma koeffitsienti va uning xossalari.

1. Ma'lumki, korrelyatsiion bog'langan X va Y belgilarning regressiya tanlanma tenglamasi

$$\overline{y_x} = f(x), \text{ yoki } \overline{x_y} = \varphi(y)$$

ko'rinishda yozilib, agar $f(x)$ va $\varphi(y)$ regressiya funksiyalarining ikkalasi ham chiziqli bo'lsa, u holda X va Y belgilar orasidagi korrelyatsion bog'lanish chiziqli deb atalar edi. Biz mana shu chiziqli korrelyatsion bog'lanishni atroflicha o'rganib chiqamiz.

Buning uchun (X,Y) juftlikning sonli belgilari sistemasini o'rganamiz. Bunda ikki: 1) ma'lumotlar guruhlanmagan; 2) ma'lumotlar guruhlangan hollarni alohida-alohida qarashimiz kerak bo'ladi.

1) Tanlanma ustida o'tkazilgan n ta erkli tajriba natijasida olingan ma'lumotlardan $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ sonlar juftligi ketma-ketligini hosil qilingan bo'lib, bu ma'lumotlarni guruhlash shart bo'lmasin, ya'ni X belgining turli x qiymatlari va ularga mos Y belgining y qiymatlari bir martadan kuzatilgan bo'lsin. Bunday holatda shartli o'rtacha tushunchasidan foydalanish shart emas. Shuning uchun izlanayotgan

$$\overline{y_x} = kx + b \qquad (1)$$

tanlanma regressiya to'g'ri chizig'i tenglamasini quyidagicha yozishimiz mumkin

$$y = kx + b. \qquad (2)$$

Bu tenglamadagi burchak koeffitsientni ρ_{yx} bilan belgilab, uni Y ning X ga *regressiya tanlanma koeffitsienti* deb ataymiz. Shunday qilib, Y ning X ga to'g'ri chiziqli regressiya tanlanma tenglamasini

$$Y = \rho_{yx} x + b \qquad (3)$$

ko'rinishda izlaymiz.

Bu tenglamadagi noma'lum ρ_{yx} va b koeffitsientlarni shunday tanlashimiz keraki, natijada kuzatish ma'lumotlari bo'yicha topilgan $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ nuqtalarni xOy tekislikka joylashtirganimizda bu nuqtalar mumkin qadar (3) to'g'ri chiziq yaqin atrofida yotsin. Bunday talabni bajarishdan oldin $Y_i - y_i \quad (i = 1, 2, ..., n)$ ifoda bilan aniqlanadigan chetlanish tushunchasini kiritib olamiz, bu erda Y_i-(3) tenglamadan kuzatilgan x_i qiymatga mos keluvchi ordinata; y_i esa x_i ga mos kuzatilgan ordinata. Noma'lum ρ_{yx} va b koeffitsientlarni shunday tanlaymizki, chetlanishlar kvadratlarining yig'indisi eng kichik, ya'ni $\min\sum_{i=1}^{n}\left(Y_i-y_i\right)^2$, bo'lsin (noma'lum ρ_{yx} va b koeffitsientlarni topishning bu usuli *eng kichik kvadratlar* usuli deb ataladi).

Har bir chetlanish noma'lum ρ_{yx} va b koeffitsientlarga bog'liq bo'lgani uchun chetlanishlari kvadratlari yig'indisining funksiyasi F ham bu koeffitsientlarga bog'liq bo'ladi:

$$F(\rho_{yx},b)=\sum_{i=1}^{n}(Y_i-y_i)^2=\sum_{i=1}^{n}(\rho_{yx}x_i+b-y_i)^2 .$$

Bu funksiyaning minimumini topish uchun noma'lum parametrlar bo'yicha F ning xususiy hosilalarni hisoblab nolga tenglashtiramiz (hozircha ρ_{xy} o'rniga ρ yozib turamiz):

$$\frac{\partial F}{\partial\rho}=2\sum_{i=1}^{n}(\rho+b-y_i)x_i=0,\qquad \frac{\partial F}{\partial b}=2\sum_{i=1}^{n}(\rho+b-y_i)=0.$$

Elementar almashtirishlar bajarib ρ va b ga nisbatan quyidagi tenglamalar sistemasini olamiz:

$$\begin{cases}\rho\sum\limits_{i=1}^{n}x_i^2+b\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}x_iy_i\\ \rho\sum\limits_{i=1}^{n}x_i+nb=\sum\limits_{i=}^{n}y_i\end{cases}\qquad(4)$$

Bu sistemani yechib izlanayotgan parametrlarni topamiz (ixchamlik uchun i indekslarni tushirib qoldiramiz):

$$\rho_{yx}=\frac{n\sum xy-\sum x\sum y}{n\sum x^2-\left(\sum x\right)^2},\quad b=\frac{n\sum x^2\sum y-\sum x\sum xy}{n\sum x^2-\left(\sum x\right)^2}\qquad(5)$$

Xuddi shu usulda X ning Y ga regressiya to'g'ri chiziqli tanlanma tenglamasini topish mumkin.

$$\overline{x_y}=\rho_{xy}y+c\,.\qquad(6)$$

1-misol. Hajmi $n=5$ bo'lgan tanlanmaning taqsimoti

x_i	1	1,5	3	4,5	5
y_i	1,25	1,4	1,5	1,75	2,25

bo'yicha Y ning X ga regressiya to'g'ri chiziqli tanlanma tenglamasini toping.

Yechish. Ma'lumotlar asosida quyidagi jadvalni tuzamiz:

x_i	y_i	x_i^2	x_iy_i
1	1,25	1	1,25
1,5	1,4	2,25	2,1
3	1,5	9	4,5
4,5	1,75	20,25	4,875
5	2,25	25	11,25
$\sum_i=15$	$\sum_i=8,15$	$\sum_i=57,5$	$\sum_i=26,975$

Jadvaldagi hisoblangan qiymatlarni (5) formulaga qo'ysak:

$$\rho_{yx}=\frac{n\sum x_i y_i-\sum x_i\sum y_i}{n\sum x_i^2-\left(\sum x_i\right)^2}=\frac{5\cdot 26,975-15\cdot 8,15}{5\cdot 57,5-15^2}=0,202,$$

$$b=\frac{n\sum x_i^2\sum y_i-\sum x_i\sum x_i y_i}{n\sum x_i^2-\left(\sum x_i\right)^2}=\frac{5\cdot 57,5\cdot 8,15-15\cdot 26,975}{5\cdot 57,5-15^2}=1,024.$$

U holda regressiya tanlanma tenglamasi: $\overline{y_x}=0,202x+1,024$.

2) Faraz qilamiz, kuzatish natijasida olingan ma'lumotlar ko'p sonli (kamida 50 ta kuzatish o'tkazilishi kerak), ya'ni guruhlanadigan, bo'lib X belgining x qiymatiga va mos Y belgining y qiymati bir necha martadan kuzatilgan bo'lsin, ya'ni ma'lumotlar ichida takrorlanadiganlari ham bor, u holda ular korrelyatsion jadval ko'rinishida beriladi.

Quyidagi (soddalik uchun i indekslarni tushirib qoldiramiz):

$$\overline{x}=\frac{1}{n}\sum x\Rightarrow\sum x=n\overline{x},\quad \overline{y}=\frac{1}{n}\sum y\Rightarrow\sum y=n\overline{y},\quad \overline{x^2}=\frac{1}{n}\sum x^2\Rightarrow\sum x^2=n\overline{x^2},$$

$\sum xy=n_{xy}xy$ ((x,y) juftlik n_{xy} marta kuzatilishi hisobga olingan)

ayniyatlardan foydalanib, (4) tenglamalar sistemasini quyidagicha yozib olamiz:

$$\begin{cases}\left(n\overline{x^2}\right)\rho_{yx}+\left(n\overline{x}\right)b=\sum n_{xy}xy,\\ \left(\overline{x}\right)\rho_{yx}+b=\overline{y}.\end{cases}\qquad(7)$$

Bu sistemani p_{yx} va b ga nisbatan yechib, izlanayotgan regressiya tanlama tenglamasini topamiz:

$$\overline{y_x}=\rho_{yx}x+b.\qquad(8)$$

Ammo (7) sistemaning yechimini topishdagi ba'zi bir hisoblashlarni yengillashtirish maqsadida (8) tenglamani $\overline{y}$ uchun ham yozib:

$$\overline{y}=\rho_{yx}\overline{x}+b,\qquad(9)$$

chunki $(\overline{x},\overline{y})$ nuqta ham (8) tenglamaning yechimi bo'ladi, (8) va (9) tenglamalardan tenglamalar sistemasi hosil qilamiz va yangi sistemadan

$$\overline{y_x}-\overline{y}=\rho_{yx}(x-\overline{x})\qquad(10)$$

regressiya tanlama tenglamasini hosil qilamiz.

(7) sistemadan regressiya koeffitsientini topamiz:

$$\rho_{yx}=\frac{\sum n_{xy}xy-n\overline{x}\,\overline{y}}{n\left(\overline{x^2}-\left(\overline{x}\right)^2\right)}=\frac{\sum n_{xy}xy-n\overline{x}\,\overline{y}}{n\sigma_x^2}$$ (ma'lumotlar guruhlanmasa $n_{xy}=1$).

1-eslatma. Agar (x_i,y_i) ma'lumotlarda katta sonlar qatnashsa x_i, y_j variantalardan mos ravishda $u_i=\dfrac{x_i-c_1}{h_1}$, $v_j=\dfrac{y_j-c_2}{h_2}$ shartli variantalarga o'tib hisoblashlarni ancha yengillashtirish mumkin.

Ma'lumki, korrelyatsiya nazariyasining asosiy masalalaridan biri korrelyatsion bog'lanish zichligini (kuchini) aniqlashdir.

Y belgining X belgiga korrelyatsion bog'lanish zichligi Y ning $X = x$ ga mos qiymatlarining $\overline{y_x}$ shartli o'rtacha qiymat atrofida tarqoqligi bo'yicha baholanadi. Agar tarqoqlik katta bo'lsa, u holda Y ning X ga kuchsiz bog'langanligini yoki umuman bog'lanmaganligini bildiradi. Tarqoqlikning kamligi esa ular orasida ancha kuchli bog'lanish borligini ko'rsatadi.

Y va X belgilar orasidagi korrelyatsion bog'lanish zichligini xarakterlovchi kattaliklar: korrelyatsiya tanlanma koeffitsienti va tanlanma korrelyatsion nisbatlar bilan tanishib chiqamiz. Bu ikki kattalikning vazifalari bir-biriga o'xshasa ham turli shakldagi masalalarni hal qiladi. Shu sababli, bu ikki kattalikni alohida-alohida o'rganamiz.

3. ***Korrelyatsiya tanlanma koeffitsienti*** belgilar orasidagi chiziqli bog'lanish zichligini aniqlab beradi. Uning formulasi keltirib chiqarish uchun Y ning X to'g'ri chiziqli regressiya tanlanma tenglamasini

$$\overline{y_x} - \overline{y} = \rho_{yx}(x - \overline{x}) \qquad (10)$$

parametri ρ_{yx} ning

$$\rho_{yx} = \frac{\sum n_{xy}xy - n\overline{x}\,\overline{y}}{n\sigma_x^2} \qquad (11)$$

ifodasining ko'rinishini o'zgartiramiz. Buning uchun (11) tenglikning ikkala tomonini ham $\frac{\sigma_x}{\sigma_y}$ nisbatga ko'paytiramiz. U holda:

$$\frac{\sigma_x}{\sigma_y}\rho_{yx} = \frac{\sum n_{xy}xy - n\overline{x}\,\overline{y}}{n\sigma_x\sigma_y}$$ (agar ma'lumotlar guruhlanmasa: $n_{xy} = 1$).

Hosil bo'lgan tenglikning o'ng tomonini r_T bilan belgilaymiz va uni tanlanma korrelyatsiya koeffitsienti deb ataymiz:

$$r_T = \frac{\sum xy - n\overline{x}\,\overline{y}}{n\sigma_x\sigma_y}$$ (ma'lumotlar guruhlanmasa), (12)

yoki

$$r_T = \frac{\sum n_{xy}xy - n\overline{x}\,\overline{y}}{n\sigma_x\sigma_y}$$ (ma'lumotlar guruhlansa). (13)

Bu yerda x, y lar mos ravishda X va Y belgilarning kuzatilgan qiymatlari; n_{xy} - kuzatilgan (x, y) juftlikning chastotasi; n - tanlanma hajmi; $\overline{x}$, $\overline{y}$ -mos tanlanma o'rtachalar; σ_x, σ_y - tanlanma o'rtacha kvadratik chetlanishlari.

r_T - tanlanma korrelyatsiya koeffitsienti bosh to'plam r-korrelyatsiya koeffitsientining bahosi hisoblanadi, shuning uchun Y va X kattaliklarning son belgilari orasidagi chiziqli bog'liqligining o'lchovi hisoblanadi.

Agar tanlanma etarlicha katta hajmga ega va reprezentativ bo'lsa, u holda belgilar orasidagi zichlik haqida tanlanma ma'lumotlari bo'yicha olingan xulosa ma'lum darajada bosh to'plamga ham tarqatilishi mumkin. Masalan, normal qonun bo'yicha taqsimlangan bosh to'plam korrelyatsiya koeffitsientini baholash uchun $(n \geq 50)$

$$r_T-3\frac{1-r_T^2}{\sqrt{n}}\le r_B\le r_T+3\frac{1+r_T^2}{\sqrt{n}}$$

formuladan foydalanish mumkin.

Tanlanma korrelyatsiya koeffitsienti uchun quyidagi xossalar o'rinli:

1-xossa. Tanlanma korrelyatsiya koeffitsientining absolyut qiymati birdan ortmaydi, ya'ni $|r_T|\le 1$, yoki $-1\le r_T\le 1$.

2-xossa. Tanlanma korrelyatsiya koeffitsientining absolyut qiymati ortsa, belgilar orasidagi chiziqli korrelyatsion bog'lanish zichligi ortadi.

3-xossa. Agar $|r_T|=1$ bo'lsa, u holda kuzatilayotgan belgilar chiziqli funksional bog'langan bo'ladi.

4-xossa. Agar $r_T=0$ bo'lib, regressiya tanlanma chiziqlari to'g'ri chiziqlardan iborat bo'lsa, u holda X va Y belgilar orasidagi bog'lanish chiziqli korrelyatsion bog'lanish bo'lmaydi.

1-eslatma. Agar $r_T=0$ bo'lsa, u holda o'rganilayotgan belgilar chiziqsiz korrelyatsion bog'lanishda (masalan, parabolik, ko'rsatkichli va h.k.) va hattoki, funksional bog'lanishda bo'lishi mumkin.

Yuqorida keltirilgan xossalardan tanlanma korrelyatsiya koeffitsientining ma'nosi kelib chiqadi: tanlanma korrelyatsiya koeffitsienti tanlanmada son belgilar orasidagi chiziqli korrelyatsion bog'lanish zichligini xarakterlaydi: $|r_T|$ kattalik 1 ga qancha yaqin bo'lsa, chiziqli korrelyatsion bog'lanish shuncha kuchli; $|r_T|$ kattalik 0 ga qancha yaqin bo'lsa, chiziqli korrelyatsion bog'lanish shuncha kuchsiz.

2-eslatma. Tanlanma korrelyatsiya koeffitsientining ishorasi regressiya koeffitsientlarining ishoralari bilan bir xil bo'ladi, bu quyidagi formulalardan kelib chiqadi:

$$\rho_{yx}=r_T\frac{\sigma_y}{\sigma_x},\qquad \rho_{xy}=r_T\frac{\sigma_x}{\sigma_y}. \qquad (14)$$

3-eslatma. Tanlanma korrelyatsiya koeffitsienti tanlanma regressiya koeffitsientlarining geometrik o'rtacha qiymatiga teng: $r_T=\pm\sqrt{p_{yx}p_{xy}}$.

Haqiqatan ham (14) dan:

$$\rho_{yx}\rho_{xy}=r_T^2\Rightarrow r_T=\pm\sqrt{p_{yx}p_{xy}}\ .$$

Ildiz oldidagi ishora regressiya koeffitsientlari ishoralari bilan bir xil qilib olinishi lozim.

2-misol. Cho'chqa bolasining og'irligi Y (kg.) va yoshi X (haftalarda) orasidagi bog'lanish quyidagi jadval bilan berilgan.

x	0	1	2	3	4	5	6	7	8
y	1,3	2,5	3,9	5,2	5,3	7,5	9,0	10,8	13,1

Shu ma'lumotlar bo'yicha tanlanma korrelyatsiya koeffitsientini toping.

Yechish. $r_T=\dfrac{\sum x_iy_i-n\overline{x}\overline{y}}{n\sigma_x\sigma_y}$ formulada zarur hisoblashlarni bajarsak, $r_T=0{,}98$ ekanligini topamiz. Bundan esa cho'chqa bolasining og'irligi va yoshi orasidagi bog'lanish kuchli, degan xulosaga kelamiz.

4-eslatma. Tanlanma korrelyatsiya koeffitsientini hisoblashni soddalashtirish uchun shartli variantaga o'tish mumkin (bunda r_T ning qiymati o'zgarmaydi).

O'z-o'zini tekshirish uchun savollar.

1. Eng kichik kvadratlar usulining mohiyatini tushuntiring.
2. Belgilar orasidagi to'g'ri chiziqli regressiya tanlanma tenglamalarini keltiring.
3. To'g'ri chiziqli regressiya tanlama tenglamasidagi parametrlarni izohlang.
4. Tanlanma korrelyatsiya koeffitsienti xossalarini keltiring.

MUSTAQIL YECHISH UCHUN MASALALAR.

1. Tanlanmaning quyidagi jadvali yordamida Y ning X ga to'g'ri chiziqli regressiya tanlanma tenglamasini tuzing.

X	10	2	7	5
Y	8	2	6	4

2. Shahardagi 10 ta oziq-ovqat magazini bo'yicha bir oylik tovar ayirboshlash hajmi- X va shu davr mobaynidagi muomala xarajatlari-Y hajmi o'rganilgan. X ning Y ga bog'liqligi ifodalovchi to'g'ri chiziqli regressiya tanlanma tenglamasini toping.

X (mln. Sum)	200	300	320	410	304	500	540	600	650	700
Y (mln. sum)	20	27	30	36	38	44	50	56	58	60

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.

14 - ma'ruza. Tanlanma korrelyatsion nisbat. Egri chiziqli korrelyatsiya.

Tayanch iboralar. *Korrelyatsiya, egri chiziqli korrelyatsiya, to'plamli korrelyatsiya, tanlanma korrelyatsion nisbat, shartli varianta.*

Reja.

16.1. Tanlanma korrelyatsion nisbat va uning xossalari.

16.2. Egri chiziqli korrelyatsiya.

Kuzatilayotgan (yoki biz o'rganmoqchi bo'lgan) X va Y belgilar orasidagi chiziqli korrelyatsion bog'lanish zichligini baholash uchun r_T-korrelyatsiya tanlanma koeffitsienti xizmat qilsa, chiziqsiz yoki umuman ixtiyoriy ko'rinishdagi korrelyatsion bog'lanishning zichligini qanday baholash mumkin degan savol bo'lishi tabiiydir. Umumiy holda korrelyatsion bog'lanishning zichligini aniqlash uchun *tanlanma korrelyatsion nisbat* deb ataluvchi xarakteristika ishlatiladi. Bu xarakteristika bilan tanishib chiqishdan oldin tanlanma korrelyatsion nisbatni kiritish bilan bog'liq bo'lgan ba'zi tushunchalarni keltirib o'tamiz.

1-ta'rif. Bosh to'plamning biror bir guruhsiga tegishli belgilarning arifmetik o'rtachasi guruh o'rtachasi deb ataladi.

Guruh o'rtachasini ba'zi hollarda shartli o'rtacha deb ham yuritish mumkin. Yuqorida foydalanilgan shartli o'rtacha tushunchasida bu holat yuz bergan.

Guruh o'rtachasi va guruhlar hajmi ma'lum bo'lsa umumiy to'plam o'rtachasini (bosh to'plam o'rtachasi) topish mumkin.

1-m*isol.* Ikki guruhdan tashkil topgan to'plam o'rtachasini toping:

Guruh	birinchi		ikkinchi	
Belgining qiymatlari	1	6	1	5
Chastota	10	15	20	30
HAJM	10+15		20+30	

Yechish. Guruh o'rtachalarini topamiz:

$$\overline{x}_1 = \frac{10\cdot 1 + 15\cdot 6}{25} = 4, \qquad \overline{x}_2 = \frac{20\cdot 1 + 30\cdot 5}{50} = 3{,}4.$$

Guruh o'rtachalari bo'yicha umumiy o'rtachani topamiz:

$$\overline{x} = \frac{25\cdot 4 + 50\cdot 3{,}4}{25+50} = 3{,}6.$$

2-ta'rif. Guruhga tegishli belgilarning guruh o'rtachasiga nisbatan dispersiyasi guruh dispersiyasi deb ataladi:

$$D_{гр}(X_j) = \frac{\sum n_i\left(x_i - \overline{x}_j\right)^2}{N_j}, \qquad (1)$$

bu yerda n_i - x_i qiymatning chastotasi; j-guruh nomeri; $\overline{x}_j$ - j guruhning guruh o'rtachasi; $N_j = \sum n_i$ - j guruh hajmi.

2-misol. Ikki guruhdan tashkil topgan to'plamning guruh dispersiyasini toping:

Guruh	birinchi		ikkinchi	
Belgining qiymatlari	1	6	1	5
Chastota	10	15	20	30
HAJM	10+15		20+30	

Yechish. 1-misoldan ma'lumki, $\overline{x_1}=4$, $\overline{x_2}=3,4$. Endi guruh dispersiyalarini topamiz:

$$D_{гр}(X_1)=\frac{10\cdot(1-4)^2+15\cdot(6-4)^2}{25}=6,$$

$$D_{гр}(X_2)=\frac{20\cdot(1-3,4)^2+30\cdot(5-3,4)^2}{50}=\frac{115,2+76,8}{50}=3,84.$$

3-ta'rif. Guruh dispersiyalarining guruhlar hajmi bo'yicha olingan arifmetik o'rtachasi guruhlar ichki dispersiyasi deb ataladi:

$$\overline{D_{гр}}=\frac{\sum N_j D_{гр}(X_j)}{n},$$

bu erda, N_j - j guruh hajmi; $n=\sum N_j$ -umumiy to'plam hajmi.

3-misol. 2-misolda guruhlar ichki dispersiyasini topsak:

$$\overline{D_{гр}}=\frac{25\cdot 6+50\cdot 3,84}{75}=4,56.$$

4-ta'rif. Guruh o'rtachalarining umumiy to'plam o'rtachasiga (bosh to'plam o'rtachasi) nisbatan dispersiyasi guruhlararo dispersiya deb ataladi:

$$D_{гр}\left(\overline{x_j}\right)=\frac{\sum N_j\left(\overline{x_j}-\overline{x}\right)^2}{n},$$

bu erda $\overline{x_j}$ - j guruhning guruh o'rtachasi; N_j - j guruh hajmi; $\overline{x}$ -umumiy o'rtacha; $n=\sum N_j$ -umumiy to'plam hajmi.

4-misol. 2-misolda guruhlararo dispersiyani topsak:

$$D_{гр}\left(\overline{x_j}\right)=\frac{25\cdot(4-3,6)^2+50\cdot(3,4-3,6)^2}{75}=\frac{4+2}{75}=0,08.$$

Endi bu tushunchalardan foydalanib tanlanma korrelyatsion nisbat tushunchasini aniqlaymiz.

5-ta'rif. Y ning X ga tanlanma korrelyatsion nisbati deb,

$$\eta_{yx}=\frac{\sigma_{\overline{y_x}}}{\sigma_y} \qquad (2)$$

nisbat bilan aniqlanuvchi kattalikka aytiladi.

Bu yerda $\sigma_{\overline{y_x}}=\sqrt{\dfrac{\sum n_x\left(\overline{y_x}-\overline{y}\right)^2}{n}}$ -shartli yoki guruhlararo o'rtacha kvadratik chetlanish; $\sigma_y=\sqrt{\dfrac{\sum n_y\left(y-\overline{y}\right)^2}{n}}$ -o'rtacha kvadratik chetlanish; n - tanlanma hajmi; n_x

n_x - X belgining x qiymati chastotasi; n_y - Y belgining y qiymati chastotasi; $\overline{y}$ - Y belgining umumiy o'rtachasi; $\overline{y_x}$ - Y belgining $X=x$ ga mos shartli o'rtachasi (x guruhning guruh o'rtachasi).

X ning Y ga tanlanma korrelyatsion nisbati ham shu kabi aniqlanadi:

$$\eta_{xy} = \frac{\sigma_{\overline{x_y}}}{\sigma_x} \qquad (3)$$

5-misol. $n=50$ hajmli quyidagi korrelyatsion jadval bo'yicha Y belgining X belgiga korrelyatsion nisbati η_{yx} ni toping.

$Y \backslash X$	10	20	30	n_y
15	4	28	6	38
25	6	-	6	12
n_x	10	28	12	$n=50$
$\overline{y_x}$	21	15	20	

Yechish. $\overline{y}$-umumiy o'rtachani topamiz:

$$\overline{y} = \frac{\sum n_i y_i}{n} = \frac{38\cdot 15+12\cdot 25}{50} = \frac{870}{50} = 17,4.$$

σ_y-o'rtacha kvadratik chetlanishni topamiz:

$$\sigma_y = \sqrt{\frac{\sum n_y\left(y-\overline{y}\right)^2}{n}} = \sqrt{\frac{38\cdot(15-17,4)^2+12\cdot(25-17,4)^2}{50}} = 4,27.$$

$\sigma_{\overline{y_x}}$-shartli o'rtachaning o'rtacha kvadratik chetlanishni (yoki guruhlararo o'rtacha kvadratik chetlanish) topamiz:

$$\sigma_{\overline{y_x}} = \sqrt{\frac{\sum n_x\left(\overline{y_x}-\overline{y}\right)}{n}} = \sqrt{\frac{10(21-17,4)^2+28(15-17,4)^2+12(20-17,4)^2}{50}} = 2,73.$$

Topilganlarni (2) formulaga qo'ysak: $\eta_{yx} = \frac{\sigma_{\overline{y_x}}}{\sigma_y} = \frac{2,73}{4,27} = 0,64.$

Tanlanma korrelyatsion nisbat uchun quyidagi xossalar o'rinli. η_{yx} va η_{xy} kattaliklar uchun aniqlangan xossalar bir xil bo'lganligi sababli tanlanma korrelyatsion nisbat xossalarini η kattalik uchun sanab o'tamiz.

1-xossa. Tanlama korrelyatsion nisbat quyidagi qo'sh tengsizlikni qanoatlantiradi: $0 \le \eta \le 1$.

2-xossa. Agar $\eta=1$ bo'lsa, belgilar funksional bog'lanishda, ya'ni $Y=f(X)$, bo'ladi.

3-xossa. Tanlanma korrelyatsion nisbat tanlanma korrelyatsiya koeffitsientining absolyut qiymatidan kichik emas: $\eta \ge |r_T|$.

4-xossa. Agar $\eta = |r_T|$ bo'lsa, belgilar orasida chiziqli bog'lanish bo'ladi.

5-xossa. Agar $\eta=0$ bo'lsa, belgilar korrelyatsion bog'lanishda bo'lmaydi.

Tanlanma korrelyatsion nisbatning afzalligi uning istalgan korrelyatsion bog'lanish, shu jumladan, chiziqli bog'lanish zichligining ham o'lchovi bo'lib xizmat qilishidadir. Shu bilan birga tanlanma korrelyatsion nisbat kamchilikka ham ega: u bog'lanish shakli haqida hech qanday ma'lumot bermaydi.

Agar X va Y belgilar orasidagi korrelyatsion bog'lanish o'rganilayotgan bo'lib, $\overline{y_x} = f(x)$ yoki $\overline{x_y} = \varphi(y)$ regressiya grafiklari egri chiziq bilan tasvirlanadigan bo'lsa, u holda korrelyatsiya *egri chiziqli* deyiladi.

Egri chiziqli korrelyatsiya nazariyasida ham chiziqli korrelyatsiya nazariyasi kabi masalalar, ya'ni korrelyatsion bog'lanish shakli va zichligini aniqlash bilan shug'ullaniladi. Egri chiziqli korrelyatsiyada Y ning X ga regressiya funksiyalari ko'rinishiga quyidagilar misol bo'lishi mumkin:

$\overline{y_x} = ax^2 + bx + c$ (ikkinchi tartibli parabolik korrelyatsiya);

$\overline{y_x} = ax^3 + bx^2 + cx + d$ (uchinchi tartibli parabolik korrelyatsiya);

$\overline{y_x} = \dfrac{a}{x} + b$ (giperbolik korrelyatsiya);

$\overline{y_x} = ae^{bx}$ (ko'rsatkichli korrelyatsiya) va h.k.

Regressiya funksiyasining ko'rinishini aniqlash uchun Dekart koordinatalar sistemasida $(x; \overline{y_x})$ nuqtalarning o'rni topiladi va ularning joylashishiga qarab regressiya funksiyasining taxminiy ko'rinishi haqida gipoteza qilinadi; o'rganilayotgan masalaning mohiyatidan kelib chiqqan holda oxirgi xulosa qabul qilinadi.

Belgilar orasidagi korrelyatsion bog'lanishni ifodalovchi regressiya funksiyalarining noma'lum parametrlarni aniqlash yoki statistik baholash masalalari ham muhim hisoblanadi.

Regressiya funksiyasining noma'lum parametrlari ham eng kichik kvadratlar usuli yordamida baholanadi. Egri chiziqli korrelyatsiya zichligini baholashda tanlanma korrelyatsion nisbatdan foydalanamiz.

n marta kuzatish ma'lumotlari asosida belgilar orasidagi korrelyatsion bog'lanish egri chiziqli korrelyatsiyaning sodda hollaridan biri ikkinchi tartibli parabolik korrelyatsiya deb hisoblaymiz va bu ko'rinishdagi korrelyatsiyaning noma'lum parametrlarini tanlanma ma'lumotlari yordamida baholaymiz.. Aniqlik uchun Y ning X ga regressiya tanlanma tenglamasini qaraymiz. Bunda regressiya tanlanma tenglamasi

$$\overline{y_x} = ax^2 + bx + c \qquad (4)$$

ko'rinishda bo'lib, a, b, c noma'lum parametrlarni tanlanma ma'lumotlari bo'yicha topish kerak bo'ladi. Noma'lum koeffitsientlarni $\nu_i = y_i - (ax_i^2 + bx_i + c), \quad i = 1, 2, 3, ..., n$ chetlanishlar kvadratlarining yig'indisi eng kichik bo'ladigan qilib, tanlaymiz. Shu maqsadda, quyidagi funksiyani kiritamiz:

$$F(a,b,c) = \sum_{i=1}^{n} \nu_i^2 = \sum_{i=1}^{n} \left(y_i - \left(ax_i^2 + bx_i + c\right)\right)^2$$

Bu funksiyani ekstremumga tekshirib va tegishli almashtirishlardan so'ng quyidagi sistemani hosil qilamiz.

$$\begin{cases} a\sum_{i=1}^{n} n_{x_i} x_i^4 + b\sum_{i=1}^{n} n_{x_i} x_i^3 + c\sum_{i=1}^{n} n_{x_i} x_i^2 = \sum_{i=1}^{n} n_{x_i} x_i^2 \overline{y_{x_i}}, \\ a\sum_{i=1}^{n} n_{x_i} x_i^3 + b\sum_{i=1}^{n} n_{x_i} x_i^2 + c\sum_{i=1}^{n} n_{x_i} x_i = \sum_{i=1}^{n} n_{x_i} x_i \overline{y_{x_i}}, \\ a\sum_{i=1}^{n} n_{x_i} x_i^2 + b\sum_{i=1}^{n} n_{x_i} x_i + cn = \sum_{i=1}^{n} n_{x_i} \overline{y_{x_i}}. \end{cases} \qquad (5)$$

Kuzatish natijalari-(x_i, y_i) juftliklardan foydalanib a,b,c larga nisbatan tenglamalar sistemasi hosil qilamiz va undan a,b,c noma'lum parametrlar topiladi.

6-misol. Korrelyatsiya jadvali ma'lumotlari asosida $\overline{y_x} = ax^2 + bx + c$ ko'rinishdagi Y ning X ga regressiya tanlama tenglamasini toping.

Y \ X	1	1,1	1,2	n_y
6	8	2	-	10
7	-	30	-	30
7,5	-	1	9	10
n_x	8	33	9	$n = 50$

Y*echish.* Korrelyatsion jadval ma'lumotlari asosida quyidagt jadvalni tuzamiz.

x	n_x	$\overline{y_x}$	$n_x x$	$n_x x^2$	$n_x x^3$	$n_x x^4$	$n_x \overline{y_x}$
1	8	6	8	8	8	8	48
1,1	33	6,73	36,3	39,93	43,93	48,32	222,09
1,2	9	7,5	10,8	12,96	15,55	18,66	67,50
Σ	50	-	55,1	60,89	67,48	74,98	337,59

Bu jadvalning Σ qatoridagi sonlarni (2) ga qo'yib quyidagi tenglamalar sistemasini hosil qilamiz:

$$\begin{cases} 74,98a + 67,48b + 60,89c = 413,93, \\ 67,48a + 60,89b + 55,10c = 373,30, \\ 60,89a + 55,10b + 50c = 337,59. \end{cases}$$

Bu sistemadan $a = 1,94,\quad b = 2,98,\quad c = 1,10$ yechimlarni topamiz. U holda regressiya tenglamasi

$$\overline{y_x} = 1,94x^2 + 2,98x + 1,10$$

ko'rinishda bo'ladi. Tekshirish uchun tenglama bo'yicha hisoblangan $\overline{y_x}$ ning qiymatlari bilan jadval bo'yicha topilgan $\overline{y_x}$ ning qiymatlarini taqqoslash mumkin.

Yuqorida keltirilgan boshqa turdaga egri chiziqli regressiya tenglamalarining koeffitsiyetlarini topishda ham eng kichik kvadratlar usulidan foydalanish mumkin, ammo ba'zi hollarda oldin ma'lum bir almashtirishlarni amalga oshirish zarur.

Masalan, $y = ax^b \quad (a > 0, b > 0)$ regressiya tenglamasidagi noma'lum a, b koeffitsiyentlarni topishda avvalam bor bu tenglamani $\ln y = \ln a + b \ln x$ ko'rinishda yozib olamiz, so'ngra $u = \ln x, \quad z = \ln y$ belgilashlar yordamida $z = bu + \ln a$ chiziqli funksiyani hosil qilamiz.

O'z-o'zini tekshirish uchun savollar.

1. Tanlanma korrelyatsiya koeffitsienti va regressiya tanlanma koeffitsienti orasida qanday munosabat bor?
2. Tanlanma korrelyatsion nisbat nima uchun xizmat qiladi? Uning xossalarini keltiring.
3. Egri chiziqli korrelyatsiya nima?
4. Egri chiziqli korrelyatsiyadagi regressiya funksiyalariga misollar keltiring.

Mustaqil yechish uchun masalalar.

1. $n = 50$ hajmli quyidagi korrelyatsion jadval bo'yicha Y belgining X belgiga tanlanma korrelyatsion nisbati η_{yx} ni toping.

X / Y	10	20	30	n_y
15	4	28	6	38
25	6	-	6	12
n_x	10	28	12	n=50
y_x	21	15	20	

2. Jadvaldagi ma'lumotlar asosida $\overline{y_x} = ax^2 + bx + c$ regressiya tanlanma tenglamasini toping.

X / Y	0	4	6	7	10	n_y
7	19	1	1	-	-	21
13	2	14	-	-	-	16
40	-	3	22	2	-	27
80	-	-	-	15	-	15
200	-	-	-	-	21	21
n_x	21	18	23	17	21	$n = 100$

3. Jadvalda keltirilgan ma'lumotlar asosida $\overline{x_y} = ay^2 + by + c$ regressiya tanlanma tenglamasini va η_{yx} tanlanma korrelyatsion nisbatni aniqlang.

X / Y	6	30	50	n_y
1	15	-	-	15
2	1	14	-	15
3	-	2	18	20
4	16	16	18	50
n_x	32	32	36	$n = 100$

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.
5. Крамер Н.Ш. Теория вероятностей и математическая статистика. Учебник М.,2001.

15 - ma’ruza. To’plamiy korrelyatsiya. Umumiy va xususiy tanlanma korrelyatsiya koeffitsienti.

Tayanch iboralar. *To’plamiy korrelyatsiya. **X**ususiy tanlanma korrelyatsiya koeffitsienti, umumiy tanlanma korrelyatsiya koeffitsienti.*

Reja:

17.1. Umumiy tanlanma korrelyatsiya koeffitsienti.
17.2. Xususiy tanlanma korrelyatsiya koeffitsienti.

Ba’zi amaliy masalalarda ikkita emas, balki ikkitadan ko’proq belgilar orasidagi bog’lanishni o’rganish zaruriyati tug’iladi. Bu holda belgilar orasidagi korrelyatsion bog’lanish to’plamiy (ko’plik) korrelyatsiya deb ataladi.

To’plamli korrelyatsiyaning eng sodda holi bo’lgan uchta belgi orasidagi chiziqli korrelyatsiyani qaraymiz. Bu holda X, Y va Z belgilar orasidagi korrelyatsion munosabat

$$z = ax + by + cz \qquad (6)$$

tenglama ko’rinishida ifodalanadi. Bunda quyidagi:

1. Kuzatish ma’lumotlari bo’yicha regressiyaning a, b, c koeffitsientlarni topish, ya’ni $z = ax + by + cz$ tanlanma tenglamani topish;
2. Z belgi bilan ikkala Y va Z belgilar orasidagi bog’lanish zichligini baholash;
3. Y fiksirlanganda (o’zgarmaganda) Z va X orasidagi, X fiksirlanganda Z va Y bog’lanish zichligini topish masalalarini hal qilish zarur.

Birinchi masala eng kichik kvadratlar usuli bilan hal qilinadi. Analitik geometriyadan ma’lumki, (3) chiziqli bog’lanish tenglamasini:

$$z - \bar{z} = a\left(x - \bar{x}\right) + b\left(y - \bar{y}\right) \qquad (7)$$

ko’rinishda yozib olish mumkin. Bu ko’rinishda esa 1-masalani hal qilish osonroq..

Ba’zi elementar hisolashlardan so’ng a va b koeffitsientlar uchun quyidagi formulalarni topamiz:

$$a = \frac{r_{xz} - r_{yz} r_{xy}}{1 - r_{xy}^2} \frac{\sigma_z}{\sigma_x}, \qquad b = \frac{r_{yz} - r_{yx} r_{zx}}{1 - r_{xy}^2} \frac{\sigma_z}{\sigma_y}. \qquad (8)$$

Bunda r_{xz}, r_{yz}, r_{xy} – mos ravishda X va Z, Y va Z, X va Y belgilar orasidagi korrelyatsiya koeffitsientlari; $\sigma_x, \sigma_y, \sigma_z$ – o’rtacha kvadratik chetlanishlar.

Z belgining X va Y belgilar bilan bog’liqliq zichligi quyidagi:

$$R = \sqrt{\frac{r_{xz}^2 - 2r_{xy} r_{xz} r_{yz} + r_{yz}^2}{1 - r_{xy}^2}}, \quad 0 \le R \le 1 \qquad (9)$$

korrelyatsiya umumiy tanlanma koeffitsienti bilan baholanadi.

Shuningdek, Y fiksirlanganda (o’zgarmaganda) Z va X orasidagi, X fiksirlanganda Z va Y bog’lanish zichligi mos ravishda:

$$r_{xz(y)} = \frac{r_{xz} - r_{xy} r_{yz}}{\sqrt{\left(1 - r_{xy}^2\right)\left(1 - r_{yz}^2\right)}}, \qquad (10)$$

$$r_{yz(x)} = \frac{r_{yz} - r_{xy} r_{xz}}{\sqrt{\left(1 - r_{xy}^2\right)\left(1 - r_{xz}^2\right)}} \qquad (11)$$

korrelyatsiya xususiy tanlanma koeffitsientlari bilan baholanadi.

Tabiatda turli-tuman jarayonlarni o'rganishda, tasodifiy jarayonlarning o'zaro bog'liqlik qonunlarini ochishda, hamda umuman prognozlash masalalarida korrelyatsion va regression analizning xulosalari katta ahamiyatga egadir. Xususan, iqtisodiy jarayonlarni tadqiq etishda turli iqtisodiy ko'rsatkichlarning bir-biriga bog'liqligini aniqlash va shu asosda muhim xulosalar chiqarishda korrelyatsiya nazariyasining elementlari muvaffaqiyatli tatbiq etib kelinmoqda.

O'z-o'zini tekshirish uchun savollar.

1. To'plamiy korrelyatsiyani tushuntiring.
2. Tanlanma to'lamiy korrelyatsiya koeffitsienti va xususiy korrelyatsiya koeffitsientlari nimani xarakterlaydi?

Adabiyotlar.

1. Бабажанов Ш.Ш. Материалы для самостоятельных работ по теории вероятностей и математической статистике. Учебное пособие. Т., 2006.
2. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие. М.: Высшая школа,1998. 479с.
3. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. Учебное пособие.. М.: Инфра-М,1997.
4. Mamurov E.N., Adirov T.X. Ehtimollar nazariyasi va matematik statistika. O'quv qo'llanma . T. 2008 y.
5. Крамер Н.Ш. Теория вероятностей и математическая статистика. Учебник М.,2001.