

**TANLAMA TAQSIMOTINI TAHLIL QILISH: STATISTIKA QATORLARI,
GRAFIK TAQDIMOT – GISTOGRAMMA VA POLIGON, EMPIRIK
TAQSIMOT FUNKSIYASINING NAZARIY ASOSLARI
VA AMALIY JIHALTLARI**

Esanov Erkin Abduraxmanovich

Toshkent davlat texnika universiteti, Oliy

Matematika kafedrası katta o'qituvchisi

Fayzullayev Sayfulloh Xayrullo o'g'li

Toshkent davlat texnika universiteti, Hayot

faoliyati xavfsizligi kafedrası talabasi

Annotatsiya.

Ushbu maqolada tanlama taqsimotini tahlil qilishning nazariy va amaliy jihatlari batafsil ko'rib chiqilgan bo'lib, statistik qatorlar, gistogramma va poligon kabi grafik taqdimot usullari hamda empirik taqsimot funksiyasining asoslari yoritilgan. Maqola tanlama taqsimotining matematik modellarini va ularning real hayotdagi qo'llanilishini tahlil qiladi, shu bilan birga amaliy misollar orqali statistik ma'lumotlarni vizualizatsiya qilish imkoniyatlarini ko'rsatadi. Tadqiqot natijalari empirik taqsimot funksiyasining amaliy jihatlari ta'kidlaydi, bu esa statistik tahlilni osonlashtiradi va qaror qabul qilish jarayonini yaxshilaydi.

Kalit so'zlar: tanlama taqsimoti, statistik qatorlar, gistogramma, poligon, empirik taqsimot funksiyasi, statistik tahlil.

Аннотация.

В данной статье подробно рассмотрены теоретические и практические аспекты анализа выборочного распределения, включая статистические ряды, графическое представление — гистограмму и полигон, а также основы эмпирической функции распределения. Статья анализирует математические модели выборочного распределения и их применение в реальной жизни, с практическими примерами визуализации статистических данных. Результаты исследования подчеркивают практические аспекты эмпирической функции распределения, что упрощает статистический анализ и улучшает процесс принятия решений.

Ключевые слова: выборочное распределение, статистические ряды, гистограмма, полигон, эмпирическая функция распределения, статистический анализ.

Annotation.

This article provides a detailed examination of the theoretical and practical aspects of sample distribution analysis, including statistical series, graphical

presentations such as histograms and polygons, and the foundations of the empirical distribution function. The article analyzes mathematical models of sample distribution and their real-life applications, with practical examples of statistical data visualization. The research findings emphasize the practical aspects of the empirical distribution function, which simplifies statistical analysis and improves decision-making processes.

Keywords: sample distribution, statistical series, histogram, polygon, empirical distribution function, statistical analysis.

Statistika fanida tanlama taqsimoti tahlili muhim o'rin tutadi, chunki u katta ma'lumotlar to'plamidan olingan namunalarni tushunish va bashorat qilishga imkon beradi. Tanlama taqsimoti – bu ma'lumotlar to'plamining statistik xususiyatlarini ifodalovchi taqsimot bo'lib, u nazariy taqsimotlardan (masalan, normal taqsimot) farqli o'laroq, real ma'lumotlar asosida quriladi. Ushbu mavzu dolzarbligi shundaki, zamonaviy sanoat, tibbiyot va iqtisodiyotda ma'lumotlar tahlili ekologik monitoringdan tortib, moliyaviy bashoratlargacha qo'llaniladi. Masalan, global ma'lumotlar hajmi 2023 yilda 120 zettabaytga yetgan bo'lib, bu tanlama tahlilining zarurligini oshiradi. Maqolaning maqsadi tanlama taqsimotini tahlil qilish usullarini – statistik qatorlar, gistogramma, poligon va empirik taqsimot funksiyasini – nazariy va amaliy jihatdan ko'rib chiqishdir. Gipoteza: Empirik taqsimot funksiyasi amaliy tahlilda nazariy taqsimotlardan ustunroq bo'lib, ma'lumotlarning real dinamikasini yaxshiroq aks ettiradi.

Nazariy Asoslar. Tanlama taqsimotining nazariy asoslari statistik qatorlar, grafik tahlil metodlari va empirik taqsimot funksiyasining (ETF) matematik xususiyatlariga tayanadi. Statistik qatorlar tanlama elementlarini tartibga solish va ularni chastotalar bo'yicha guruhlash orqali ma'lumotlarning ichki tuzilishini aniqlash imkonini beradi. Ular odatda **diskret** (aniq qiymatlar bo'yicha tuzilgan) va **uzluksiz** (oraliqlar bo'yicha guruhlangan) ko'rinishga ajratiladi. Diskret qatorlar ma'lumotlarning aniq qiymatlar atrofida qanday to'planganini ko'rsatsa, uzluksiz qatorlar tanlamaning umumiy zichlik tuzilmasini baholashda qo'l keladi. Grafik taqdimot usullari ichida **gistogramma** va **chastota poligoni** tanlama taqsimotining shaklini aniqlashda eng ko'p qo'llaniladigan vositalardandir. Gistogramma chastota yoki chastota zichligini to'rtburchaklar orqali tasvirlab, taqsimotning simmetriya darajasi, markaziy tendensiyalari va potensial moda nuqtalarini vizual baholash imkonini beradi. Chastota poligoni esa gistogrammaning yuqori nuqtalarini chiziq orqali birlashtirish yo'li bilan olinadi va taqsimotning yanada silliq, uzluksiz shaklini ko'rsatadi. Ushbu grafiklar taqsimotning cho'ziqligi, konsentratsiya darajasi va ekstremal qiymatlar haqida tezkor xulosalar chiqarish imkonini beradi. Empirik taqsimot funksiyasi (ETF yoki EDF) esa tanlama asosida qurilgan kumulyativ taqsimot

funksiyasi bo'lib, nazariy taqsimotni amaliy kuzatuvlar asosida baholashda muhim vosita hisoblanadi. Matematik jihatdan 1-rasmda u:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x)$$

1-rasm.

ko'rinishda aniqlanadi, bu yerda $I(\cdot)$ — indikator funksiyasi, n — tanlama hajmi. ETF har bir x qiymatida tanlamaning qanchalik ulushi ushbu qiymatdan kichik yoki teng ekanini ko'rsatadi. Nazariy statistikada ETF ning xossalari Kolmogorovning asimptotik teoremlari bilan asoslanadi, xususan, Kolmogorov–Smirnov mezon empirik va nazariy taqsimotlar o'rtasidagi farqning maksimal og'ishiga asoslanib, ularni moslik darajasi bo'yicha baholaydi. Amaliy statistikada ETF tanlama taqsimotining kumulyativ xususiyatlarini o'rganish, ekstremal qiymatlarni aniqlash, ma'lumotlarda anomaliya mavjudligini tekshirish hamda nazariy modellarni tanlash jarayonida keng qo'llanadi. Shuningdek, u parametrik bo'lmagan metodlarning asosiy elementi bo'lib, kichik hajmli tanlamalar bilan ishlashda yuqori ishonchlilikka ega.

1-jadval. Nazariy Asoslar Jadvali.

Nazariy komponent	Tavsifi	Amaliy qo'llanilishi
Statistik qatorlar	Tanlama elementlarini tartibga solish va chastotalar bo'yicha guruhlash orqali ma'lumotlarning ichki tuzilishini aniqlash. Diskret va uzluksiz qatorlarga ajratiladi.	Ma'lumotlarning tarqalishini baholash, tanlamaning zichlik tuzilmasini ko'rsatish.
Gistogramma	Chastota yoki chastota zichligini to'rtburchaklar orqali tasvirlash.	Taqsimotning simmetriya darajasi, markaziy tendensiyalar va moda nuqtalarini vizual baholash.
Chastota poligoni	Gistogramma yuqori nuqtalarini chiziq orqali birlashtirish orqali olinadi.	Taqsimotning silliq, uzluksiz shaklini ko'rsatish, cho'ziqlik, konsentratsiya va ekstremal qiymatlarni tezkor baholash.
Empirik taqsimot	Tanlama asosida qurilgan kumulyativ taqsimot funksiyasi.	Tanlama taqsimotining kumulyativ xususiyatlarini o'rganish, ekstremal qiymatlarni

funksiyasi (ETF)		aniqlash, anomalionalarni tekshirish, nazariy modellarni tanlash, parametrik bo'lmagan metodlarda qo'llash.
-----------------------------	--	--

Amaliy Jihatlar. Tanlama taqsimotini tahlil qilishning amaliy jihatlarini statistik qatorlar, gistogramma, poligon va empirik taqsimot funksiyasining (EDF) real ma'lumotlar bilan ishlashda qo'llanilishida namoyon bo'ladi. Ushbu usullar ma'lumotlarning ichki tuzilishini vizual va matematik jihatdan o'rganishga imkon berib, qaror qabul qilish jarayonini soddalashtiradi va ma'lumotlarning ishonchligini oshiradi. Amaliyotda tanlama taqsimoti tahlili tibbiyot, iqtisodiyot, muhandislik va ijtimoiy fanlarda keng qo'llaniladi, chunki u katta ma'lumotlar to'plamidan (big data) muhim mintiqalarni ajratib olishga yordam beradi. Masalan, global ma'lumotlar hajmining o'sishi (2023 yilda 120 zettabaytga yetgan) tanlama tahlilining zarurligini yanada kuchaytirmoqda, chunki to'liq ma'lumotlar to'plamini qayta ishlash qimmat va vaqt talab qiladi. Amaliy misol sifatida, 100 ta talabaning bo'yi (sm) haqidagi ma'lumotlar tanlamasini ko'rib chiqaylik. Tanlama 150-200 sm oraliqda bo'lib, o'rtacha qiymati 170 sm, dispersiyasi 100 sm² va standart og'ishi 10 sm ni tashkil etadi. Bu tanlamani tahlil qilish uchun avvalo statistik qatorlar quramiz: ma'lumotlarni 5 ta oraliqqa (150-160, 160-170, 170-180, 180-190, 190-200) bo'lib, chastotalarni hisoblaymiz. Masalan, chastotalar quyidagicha bo'lishi mumkin: 150-160 sm – 15 ta, 160-170 sm – 30 ta, 170-180 sm – 25 ta, 180-190 sm – 20 ta, 190-200 sm – 10 ta. Bu qatorlar orqali ma'lumotlarni 10 ta oraliqqa bo'lish va gistogramma qurish mumkin, bu yerda to'rtburchaklar balandligi chastotani, kengligi esa oraliqni ifodalaydi. Gistogramma taqsimotning simmetriyasi va cho'ziqligini ko'rsatib, masalan, normal taqsimotga yaqinligini vizual baholashga yordam beradi – bu holda taqsimot o'ngga cho'zilgan bo'lishi mumkin, chunki yuqori qiymatlar kamroq.

Gistogrammadan keyin chastota poligonini quramiz: gistogrammaning yuqori nuqtalarini chiziq orqali bog'lab, taqsimotning yumshoq shaklini hosil qilamiz. Poligon taqsimotning modasi (eng yuqori chastotali oraliq, masalan, 160-170 sm) va ekstremal qiymatlarini aniqroq ko'rsatadi, bu ma'lumotlarning tarqalishini tushunishga yordam beradi. EDF orqali taqsimot normal taqsimotga yaqinligini baholaymiz: EDF $F_n(x)$ funksiyasi orqali kumulyativ chastotalarni hisoblab, nazariy normal taqsimot bilan taqqoslaymiz. Kolmogorov testi D-statistikasi 0.12 bo'lib, p-qiymati 0.05 dan yuqori bo'lsa (masalan, 0.08), gipoteza tasdiqlanadi, ya'ni tanlama normal taqsimotga mos keladi. Bu misol tibbiyotda bemorlar ma'lumotlarini tahlil qilishda qo'llaniladi, masalan, kasalliklar taqsimotini modellashtirishda, chunki EDF bemorlarning kasallik darajasini kumulyativ baholashga imkon beradi va kasallik tarqalishini oldindan ko'rishga yordam beradi. Boshqa amaliy sohada, masalan, iqtisodiyotda, aksiyalar

bahosi tanlamasini tahlil qilishda gistogramma va EDF qo'llaniladi: aksiyalar bahosining o'rtacha 100 USD, dispersiya 400 USD² bo'lgan tanlamada poligon taqsimotning chapga cho'zilganligini ko'rsatishi mumkin, bu esa investorlar uchun risklarni baholashga yordam beradi.

Muhandislikda esa materiallarning chidamliligini tahlil qilishda EDF materialning buzilish ehtimolini prognozlashga yordam beradi, masalan, ko'mir konlarida qatlamlarning siljishini modellashtirishda. Ushbu usullar katta ma'lumotlar (big data) bilan ishlashda muhim, chunki ular ma'lumotlarning ichki qonuniyatlarini tezda aniqlaydi va qaror qabul qilishni tezlashtiradi. Amaliyotda EDF va grafiklar R yoki Python (matplotlib kutubxonasi) orqali quriladi, bu esa tahlilni avtomatlashtiradi va xatolarni kamaytiradi. Quydagi jadval (tanlama ma'lumotlarining statistik tahliliga misol).

2-jadval. Tanlama ma'lumotlarining statistik tahliliga misol.

Statistik parametr	Qiymati	Izoh
O'rtacha	170 sm	Taqsimot markazi
Mediana	168 sm	O'rtacha qiymat
Moda	170 sm	Eng ko'p uchraydigan qiymat
Dispersiya	100 sm ²	Taqsimot tarqalishi
Standart og'ish	10 sm	O'rtachadan og'ish

Xulosa.

Tanlama taqsimotini tahlil qilish statistik qatorlar, gistogramma, chastota poligoni va empirik taqsimot funksiyasi (EDF) yordamida ma'lumotlarni samarali va tizimli tushunishga imkon beradi. Statistik qatorlar tanlamaning ichki tuzilishini ochib beradi, gistogramma va poligon esa taqsimotning vizual shaklini aniqlashga yordam beradi, shu bilan birga simmetriya, markaziy tendensiya va ekstremal qiymatlarni baholash imkonini beradi. Empirik taqsimot funksiyasi nazariy taqsimot bilan taqqoslash orqali ma'lumotlarning mosligini aniqlashda muhim vosita hisoblanadi. Amaliy jihatdan ushbu metodlar tibbiyot, iqtisodiyot va muhandislik sohalarida keng qo'llaniladi, masalan, bemorlar yoki ishlab chiqarish parametrlarining taqsimotini aniqlash va nazariy modellarga moslashda. Kelajakda sun'iy intellekt (AI) va avtomatlashtirilgan tizimlar bilan integratsiya ushbu tahlilning aniqligi va tezkorligini sezilarli darajada oshirishi kutilmoqda. Shu bilan birga, parametric va non-parametric metodlarni birlashtirish orqali katta hajmli va murakkab ma'lumotlar bilan ishlash imkoniyatlari kengayadi.

Adabiyotlar ro'yxati.

1. Khan, M. A. (2022). *Empirical Distribution Function in Statistics*. Journal of Statistical Research.

2. Montgomery, D. C. (2020). *Introduction to Statistical Quality Control*. Wiley.
3. Devore, J. L. (2015). *Probability and Statistics for Engineering and the Sciences*. Cengage Learning.
4. Kolmogorov, A. N. (1933). *Foundations of the Theory of Probability*. Chelsea Publishing.
5. Smirnov, N. V. (1948). *Table for Estimating the Goodness of Fit of Empirical Distributions*. Annals of Mathematical Statistics.
6. Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489.
7. Särndal, C. E., Swensson, B., & Wretman, J. (2003). *Model Assisted Survey Sampling*. Springer.
8. Hogg, R. V., McKean, J., & Craig, A. T. (2018). *Introduction to Mathematical Statistics*. Pearson.
9. Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill.
10. Everitt, B. S., & Skrondal, A. (2010). *The Cambridge Dictionary of Statistics*. Cambridge University Press.
11. Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences*. Pearson.
12. Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
13. Sheskin, D. J. (2020). *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press.
14. Zar, J. H. (2010). *Biostatistical Analysis*. Pearson.
15. Rice, J. A. (2006). *Mathematical Statistics and Data Analysis*. Cengage Learning.