THE PERSPECTIVE OF NEW ERA SCIENCES
INTERNATIONAL SCIENTIFIC JOURNAL
www.global-science-review.com
editor@global-science-review.com

STATISTIK MA’LUMOTLARNING TAQSIMOTI

Ilmiy rahbar: Shakirova Nigora Akhralovna
Toshkent davlat iqtisodiyot universiteti
Statistika kafedrasi PhD, dotsenti
E-mail: shakirova0130@gmail.com
ORCID: 0009-0006-5025-4764
Moyliyev Diyor Donyorvich
Toshkent davlat iqtisodiyot universiteti, “Moliya” fakulteti
MMT-82/25 guruh talabasi
moyliyevdiyor17@gmail.com
ORCID: 0009-0001-2220-2338

ANNOTATSIYA

Maqolada statistik ma’lumotlarning taqsimoti, diskret va uzluksiz tasodifiy miqdorlar hamda binomial, Puasson va normal taqsimotlarning matematik asoslari ko‘rib chiqiladi. Normal taqsimotni tavsiflash uchun zichlik funksiyasi va 68–95–99,7 foizlik empirik qoida keltiriladi. Shuningdek, gistogramma, taqsimot poligoni, Q–Q grafigi, asimmetriya va ekssess ko‘rsatkichlari hamda Shapiro–Wilk testi orqali normallikni baholash yondashuvlari izohlanadi. Ish nazariy-tahliliy xarakterga ega: dastlabki hujjatda qayta hisoblash mumkin bo‘lgan xom kuzatuvlar to‘plami keltirilmagani sabab, chizmalar metodik tushuntirish uchun ishlatiladi va empirik natija sifatida talqin qilinmaydi. Maqolada normal taqsimot haqidagi farazni birgina test yoki grafik asosida qat’iy tasdiqlash mumkin emasligi ta’kidlanadi.

Kalit so‘zlar: statistik taqsimot, diskret o‘zgaruvchi, uzluksiz o‘zgaruvchi, normal taqsimot, ehtimollik zichligi, Shapiro–Wilk testi, Q–Q grafigi.

АННОТАЦИЯ

В статье рассматриваются распределение статистических данных, дискретные и непрерывные случайные величины, а также математические

основы биномиального, распределения Пуассона и нормального распределения. Приведены функция плотности нормального распределения и эмпирическое правило 68–95–99,7 %. Описаны гистограмма, полигон частот, Q–Q график, коэффициенты асимметрии и эксцесса, а также критерий Шапиро–Уилка. Работа носит теоретико-аналитический характер: исходный документ не содержит набора первичных наблюдений, достаточного для независимого пересчёта, поэтому графики используются как иллюстрации методики, а не как доказательство эмпирических результатов. Подчёркивается, что предположение о нормальности нельзя окончательно подтвердить только одним тестом или графиком.

Ключевые слова: статистическое распределение, дискретная переменная, непрерывная переменная, нормальное распределение, плотность вероятности, критерий Шапиро–Уилка, Q–Q график.

ABSTRACT

This article reviews statistical distributions, discrete and continuous random variables, and the mathematical foundations of the binomial, Poisson, and normal distributions. It presents the normal density function and the 68–95–99.7 empirical rule. Histograms, frequency polygons, Q–Q plots, skewness and kurtosis measures, and the Shapiro–Wilk test are discussed as tools for assessing normality. The study is theoretical and analytical: the supplied manuscript does not include raw observations that would allow the reported sample statistics to be reproduced. Therefore, the figures are used to illustrate methods rather than to claim empirical findings. The article emphasizes that normality should not be treated as proven by a single test or plot.

Keywords: statistical distribution, discrete variable, continuous variable, normal distribution, probability density, Shapiro–Wilk test, Q–Q plot.

1. KIRISH

Zamonaviy iqtisodiyot, moliya, tibbiyot, ta’lim va sanoatda qarorlar qabul qilish statistik ma’lumotlarni tizimli tahlil qilishni talab etadi. Ma’lumotlarning qanday taqsimlanganini tushunish o‘rtacha qiymat va tarqalishni tavsiflash, ehtimollarni hisoblash, tavakkalchilikni baholash hamda statistik xulosalarning ishonchliligini tekshirish uchun zarur. Shu sababli tasodifiy miqdorning diskret yoki uzluksiz ekanini aniqlash va unga mos ehtimollik modelini tanlash tahlilning dastlabki bosqichi hisoblanadi.

Diskret tasodifiy miqdorlar sanaladigan alohida qiymatlarni qabul qiladi; uzluksiz miqdorlar esa ma’lum oraliqda istalgan qiymatni olishi mumkin. Binomial va Puasson taqsimotlari ko‘plab sanash jarayonlarini ifodalashda qo‘llanadi, normal taqsimot esa o‘lchash xatolari va ko‘plab statistik modellar nazariyasida muhim o‘rin tutadi. Normal taqsimotning amaliy tavsifi va standart og‘ishlar oralig‘ida kuzatuvlarning kutiladigan ulushi 68–95–99,7 qoidasi orqali ifodalanadi.[1]

Biroq normal taqsimotni barcha ma’lumotlarga avtomatik ravishda tatbiq etish to‘g‘ri emas. Gistogramma va Q–Q grafigi shaklni ko‘rishga yordam beradi, Shapiro–Wilk kabi testlar esa statistik dalil beradi. Shapiro va Wilk normallikni tekshirish mezonini taklif qilgan; keyingi solishtirma tadqiqotlarda turli normallik testlarining quvvati va qo‘llanish shartlari baholangan.[2]

Tadqiqotning maqsadi — statistik taqsimotlarning asosiy turlarini tushuntirish, ularning matematik formulalarini to‘g‘ri ifodalash va normal taqsimotni baholashda vizual hamda statistik usullardan foydalanish tamoyillarini tizimlashtirish. Maqsadga erishish uchun o‘zgaruvchilar tasnifi yoritiladi, taqsimot funksiyalari keltiriladi, normallik diagnostikasi usullari izohlanadi va mavjud materialning cheklovlari ko‘rsatiladi.

2. MATERIALLAR VA METODLAR

[1] OpenStax. Introductory Statistics 2e: The Standard Normal Distribution. Normal taqsimot uchun empirik 68–95–99,7 qoidasi va uning talqini.

[2] Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. DOI: 10.1093/biomet/52.3-4.591; Razali, N. M., & Wah, Y. B. (2011). Power comparisons of normality tests. Journal of Statistical Modeling and Analytics, 2(1), 21–33.

Tadqiqot nazariy-tahliliy xarakterda bajarildi. Asosiy usullar: matematik formulalarni tahlil qilish, diskret va uzluksiz taqsimotlarni taqqoslash, statistik diagnostika usullarini tavsiflash hamda normallik haqidagi xulosalarni metodologik jihatdan baholash. Nazariy ta'riflar normallik testi va normal ehtimollik grafigi bo'yicha ilmiy manbalar bilan solishtirildi.3

2.1. Tasodifiy miqdorlarni tasniflash

Diskret tasodifiy miqdor X uchun ehtimolliklar yig'indisi birga teng bo'ladi:

$$\sum_i P(X = x_i) = 1$$

Bu yerda x_i — X qabul qilishi mumkin bo'lgan qiymatlar, $P(X = x_i)$ esa mos ehtimollikdir. Uzluksiz tasodifiy miqdor uchun ehtimollik zichligi f(x) ishlatiladi; [a, b] oraliqqa tushish ehtimoli zichlik funksiyasining shu oraliq bo'yicha integraliga teng:

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

2.2. Taqsimot modellarining formulalari

Binomial taqsimot mustaqil, ikki natijali n ta sinovdagi muvaffaqiyatlar sonini tavsiflaydi. Bu yerda p — bitta sinovda muvaffaqiyat ehtimoli, k — muvaffaqiyatlar soni:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

Puasson taqsimoti belgilangan vaqt yoki makon oralig'ida sodir bo'ladigan hodisalar sonini, λ parametri esa shu oraliq uchun kutiladigan hodisalar sonini ifodalaydi:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots$$

Normal taqsimot ikki parametr bilan belgilanadi: μ — o'rtacha qiymat, σ — standart og'ish (σ > 0). Uning zichlik funksiyasi quyidagicha:

[3] NIST/SEMATECH e-Handbook of Statistical Methods. Normal Probability Plot. https://www.itl.nist.gov/div898/handbook/eda/section3/normprpl.htm

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Normal taqsimot simmetrik bo‘lib, egri chiziqning markazi μ nuqtasida joylashadi. Nazariy normal taqsimot uchun taxminan 68,27 % qiymat μ ± σ, 95,45 % qiymat μ ± 2σ va 99,73 % qiymat μ ± 3σ oralig‘iga to‘g‘ri keladi.4

2.3. Normallikni baholash mezonlari

Vizual tekshiruvda gistogramma taqsimotning shakli, bir nechta cho‘qqi borligi va ehtimoliy chet qiymatlarni ko‘rishga yordam beradi. Q–Q grafigida kuzatilgan tartiblangan qiymatlar normal taqsimotning nazariy kvantillari bilan solishtiriladi: nuqtalar to‘g‘ri chiziqqa yaqin joylashsa, normallik farazi bilan moslik kuchayadi; tizimli egilish yoki dumlardagi og‘ishlar esa farq bo‘lishi mumkinligini ko‘rsatadi.5

Shapiro–Wilk mezoni tartiblangan x_(i) kuzatuvlari va a_i koeffitsiyentlari orqali hisoblanadi:

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Bu yerda n — tanlanma hajmi, $\bar{x}$ — tanlanma o‘rtachasi, x_(i) — tartiblangan kuzatuvlar, a_i — test uchun belgilangan koeffitsiyentlar. Odatda nol gipoteza ma’lumotlar normal taqsimotdan olinganini bildiradi. $p > 0{,}05$ bo‘lsa, nol gipotezani rad etish uchun yetarli dalil topilmagan deb talqin qilinadi; bu holat normal taqsimot isbotlandi degani emas. $p \leq 0{,}05$ bo‘lsa, tanlanma normalligiga qarshi statistik dalil bor deb qaraladi.

[4] Empirik qoida nazariy normal taqsimotga tegishli: taxminan 68,27 %, 95,45 % va 99,73 % qiymatlar mos ravishda μ ± σ, μ ± 2σ va μ ± 3σ oraliqlariga tushadi.

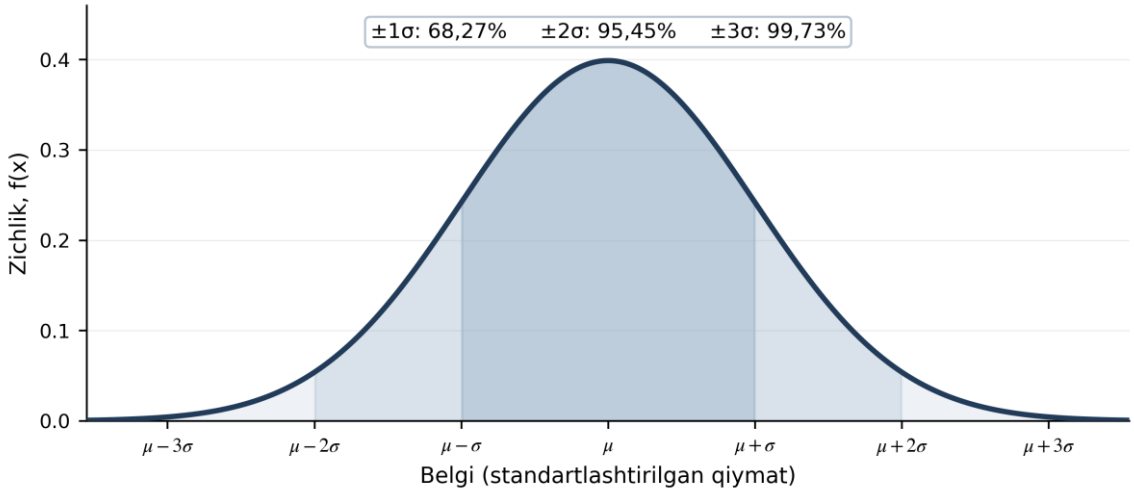
[5] NIST/SEMATECH e-Handbook of Statistical Methods, “Normal Probability Plot”: Q–Q/normal ehtimollik grafigidagi nuqtalarning chiziqqa yaqinligi taqsimotga moslikni vizual baholashga yordam beradi.

Taqsimotning simmetriyasi asimmetriya (skewness), cho‘qqi va dumlar xususiyati esa ekssess (kurtosis) orqali tavsiflanishi mumkin. Normal taqsimotda nazariy asimmetriya 0, ortiqcha ekssess (excess kurtosis) 0 ga teng; xom ekssess ko‘rsatkichidan foydalanilganda normal taqsimot qiymati 3 bo‘ladi. Hisobotda qaysi ta’rif ishlatilgani aniq ko‘rsatilishi lozim.

3. NATIJALAR

Nazariy tahlil uchta asosiy natijani berdi. Birinchidan, diskret va uzluksiz o‘zgaruvchilar uchun ehtimollikni ifodalash usullari farq qiladi: diskret holatda ehtimolliklar yig‘indisi, uzluksiz holatda esa ehtimollik zichligi integrali qo‘llanadi. Ikkinchidan, binomial va Puasson modellarida hodisalar soni, normal modelda esa uzluksiz ko‘rsatkichlarning o‘rtacha qiymat atrofidagi tarqalishi ifodalanadi. Uchinchidan, normallikni tekshirishda grafik va statistik mezonlarni birgalikda qo‘llash maqsadga muvofiq.

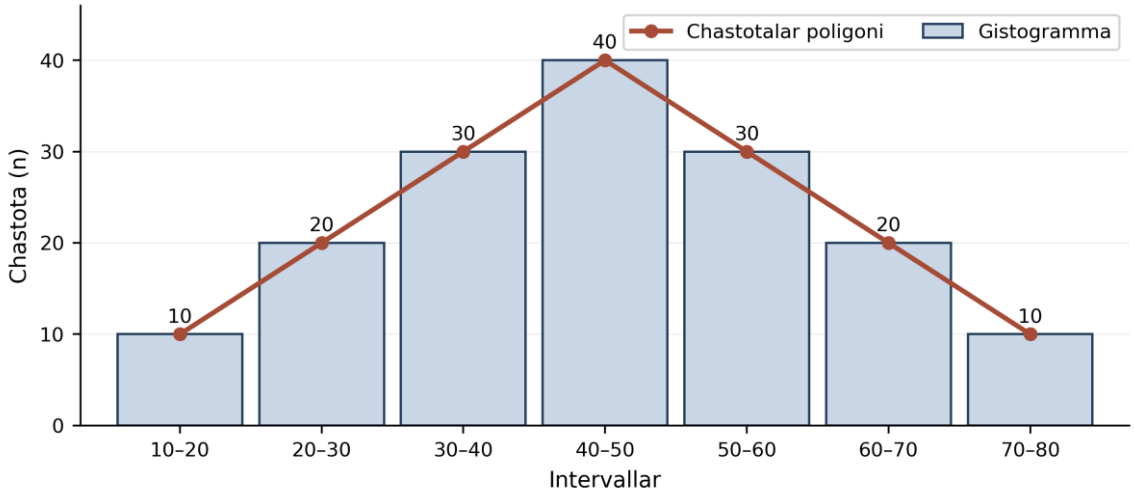
3.1. Normal taqsimot va 3σ qoidasi



1-rasm. Normal taqsimot zichligi va 68–95–99,7 empirik qoidasi.

1-rasm nazariy normal taqsimotning simmetrik ko‘rinishini ko‘rsatadi. Egri chiziq μ nuqtasida eng yuqori qiymatga ega bo‘ladi; μ dan uzoqlashgan sari zichlik kamayadi. Rasmda berilgan foizlar nazariy normal taqsimot uchun ma’lum ehtimolliklarni anglatadi, ular ushbu maqola uchun o‘lchangan ma’lumotlar ulushi emas.

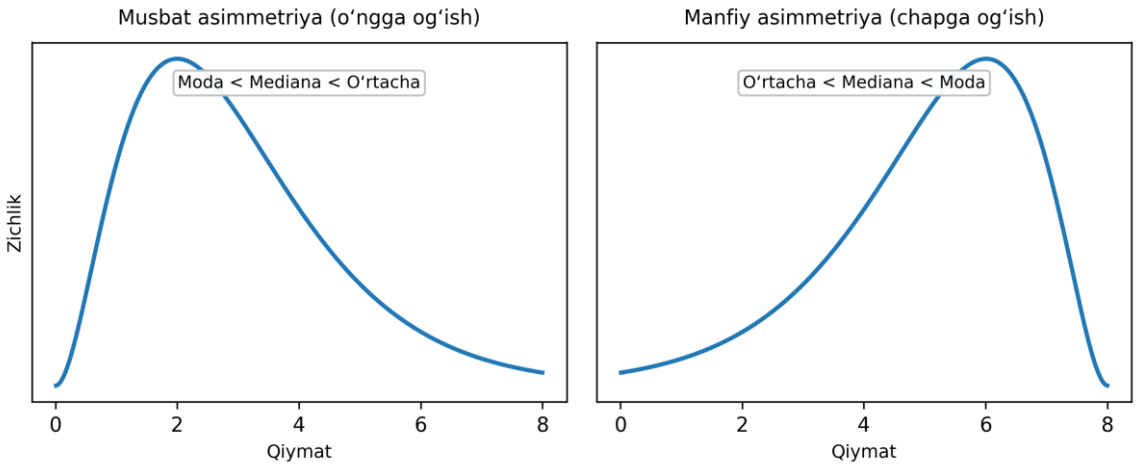
3.2. Gistogramma va chastotalar poligoni



2-rasm. Intervallar bo‘yicha gistogramma va chastotalar poligoni (shartli, metodik misol).

2-rasmda intervallar bo‘yicha chastotalarni tasvirlash usuli ko‘rsatilgan. Ustunlar intervaldagi kuzatuvlar sonini, poligon esa shu intervallar chastotalarining o‘zgarishini vizual taqqoslashga xizmat qiladi. Rasm metodik misol sifatida keltirilgan: asl hujjatda xom kuzatuvlar jadvali mavjud emasligi sababli ustunlar haqiqiy tanlanmaga tegishli deb qabul qilinmasligi kerak. Empirik tahlilda ustun balandliklari tekshirilgan ma’lumotlardan hisoblanishi zarur.

3.3. Asimmetriya turlari



3-rasm. Musbat va manfiy asimmetriyaning sxematik ko‘rinishi.

Musbat asimmetriyada o‘ng dum nisbatan uzun bo‘lib, ko‘p hollarda o‘rtacha qiymat mediana va modadan katta bo‘ladi. Manfiy asimmetriyada chap dum uzunlashadi va ko‘p hollarda o‘rtacha qiymat mediana hamda modadan kichik bo‘ladi. Bular umumiy yo‘nalishlarni ifodalovchi qoidalar bo‘lib, taqsimotning aniq shakli va multimodallik mavjudligi ham hisobga olinishi kerak.

Dastlabki nusxada moda Mo = 45,2 va asimmetriya As = +0,42 qiymatlari keltirilgan. Biroq tanlanma hajmi, kuzatuvlar ro‘yxati va hisoblash tartibi berilmaganligi uchun ushbu qiymatlarni mustaqil qayta hisoblash va tasdiqlash mumkin emas. Shu sababli ular ushbu qayta ishlangan variantda ishonchli empirik natija sifatida talqin qilinmaydi.

4. MUHOKAMA

Olingan nazariy natijalar statistik taqsimot turini ma’lumotlarning tabiatiga qarab tanlash muhimligini ko‘rsatadi. Sanaladigan hodisalar soni uchun binomial yoki Puasson modeli mos kelishi mumkin; uzluksiz o‘lchovlar uchun zichlik funksiyasi va gistogramma kabi vositalar qo‘llanadi. Normal taqsimot ko‘plab parametrik usullarda muhim bo‘lsa-da, uni avtomatik ravishda barcha ko‘rsatkichlarga tatbiq etish mumkin emas.

Q–Q grafigi normallikdan og‘ishning shaklini ko‘rishga yordam beradi, Shapiro–Wilk testi esa formal statistik gipotezani tekshiradi. Bu ikki usul bir-birini to‘ldiradi: katta tanlanmalarda juda kichik og‘ishlar ham testda ahamiyatli chiqishi, kichik tanlanmalarda esa testning og‘ishni aniqlash quvvati cheklangan bo‘lishi mumkin. Shu bois p-qiymat bilan bir qatorda grafiklar, kuzatuvlar konteksti va kerak bo‘lganda boshqa ko‘rsatkichlar ham tahlil qilinadi.6

Muhim metodologik cheklov shuki, qayta ishlash uchun berilgan maqolada to‘liq dastlabki ma’lumotlar, tanlanma hajmi, dispersiya hisob-kitoblari yoki

[6] Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. Journal of Statistical Modeling and Analytics, 2(1), 21–33.

Shapiro–Wilk testining W va p qiymatlari keltirilmagan. Shuning uchun bu versiyada test natijasi e’lon qilinmaydi va ma’lumotlarning normal taqsimlanganligi haqida empirik tasdiq berilmaydi. Grafiklar aniqlik va o‘qiluvchanlik uchun qayta chizilgan, chastotalar grafigi esa shartli misol sifatida yorliqlangan.

Natijani empirik maqola darajasiga olib chiqish uchun dastlabki ma’lumotlar jadvalini, tanlanma hajmini, o‘rtacha qiymat va standart og‘ishni, asimmetriya/ekssess hisoblarini hamda Shapiro–Wilk testining W va p qiymatlarini kiritish lozim. Shunda grafiklar real ma’lumotlardan qayta qurilib, xulosalar natijalar bilan izchil bog‘lanadi.

5. XULOSA VA TAVSIYALAR

Maqolada statistik taqsimotlarning asosiy tushunchalari va normal taqsimotni baholashda foydalaniladigan vositalar tartibga solindi. Diskret taqsimotlarda ehtimolliklar yig‘indisi birga teng bo‘lishi, uzluksiz taqsimotlarda esa oraliq ehtimoli zichlik funksiyasining integrali orqali topilishi ko‘rsatildi. Binomial, Puasson va normal taqsimot formulalari aniq va o‘qiladigan shaklda berildi.

Normal taqsimotga moslikni baholashda 68–95–99,7 qoidasi nazariy tayanch bo‘lib xizmat qiladi, ammo u alohida tanlanma bo‘yicha o‘lchangan natijani almashtirmaydi. Gistogramma, Q–Q grafigi, asimmetriya va ekssess ko‘rsatkichlari hamda Shapiro–Wilk testi kompleks ravishda qo‘llanishi kerak. $p > 0,05$ natijasi normallikni isbotlamaydi; u faqat tanlangan ahamiyatlilik darajasida nol gipotezani rad etishga yetarli asos yo‘qligini bildiradi.

Amaliy tavsiya: har qanday empirik xulosa oldidan tanlanma hajmi va xom ma’lumotlar aniq berilsin, grafiklar shu ma’lumotlardan qurilsin va test statistikasi bilan p-qiymat hisobotga kiritilsin. Ma’lumotlar taqsimotining xususiyatlari tekshirilgandan keyingina t-test, ANOVA yoki korrelyatsion tahlil kabi parametrik usullarni qo‘llash to‘g‘risida qaror qabul qilish lozim.

FOYDALANILGAN ADABIYOTLAR

1. Gauss, C. F. (1809). Theoria motus corporum coelestium in sectionibus conicis solem ambientium. Hamburg.
2. Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591
3. Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson–Darling tests. Journal of Statistical Modeling and Analytics, 2(1), 21–33.
4. D'Agostino, R. B., & Stephens, M. A. (Eds.). (1986). Goodness-of-fit techniques. Marcel Dekker.
5. Seber, G. A. F., & Lee, A. J. (2003). Linear Regression Analysis (2nd ed.). John Wiley & Sons.
6. OpenStax. (2023). Introductory Statistics 2e: The Standard Normal Distribution. https://openstax.org/books/introductory-statistics-2e/pages/6-1-the-standard-normal-distribution (murojaat sanasi: 09.10.2026).
7. NIST/SEMATECH. (2012). e-Handbook of Statistical Methods: Normal Probability Plot. https://www.itl.nist.gov/div898/handbook/eda/section3/normprpl.htm (murojaat sanasi: 09.10.2026).