# *KRIVULJE RASPODJELE*

- Krivulje raspodjele predstavljaju zakon vjerojatnosti pojave neke hidrološke veličine.
- Za slučajnu varijablu X kažemo da je poznata ako znamo zakon njene raspodjele.
- Razlikujemo dvije vrste slučajnih varijabli: diskretnu i kontinuiranu

- Za diskretnu varijablu skup mogućih vrijednosti $R_x$ je diskretan skup:

$$\left\{ x_1, x_2 \ldots x_n \right\}$$

- Pripadne vjerojatnosti

$$p_i = P\left(X = x_i\right) \text{za } \forall \text{ i} = 1,2,\ldots\text{n}$$

- Razdioba slučajne varijable *X* može se zadati i u slijedećem obliku:

$$X : \begin{pmatrix} x_1, x_2, \ldots x_n \\ \\ p_1, p_2, \ldots p_n \end{pmatrix}$$

- Nužan uvjet

$$p_i \geq 0, \quad \sum_i p_i = 1$$

# *Kontinuirana slučajna varijabla*

- Vjerojatnost se promatra kao mogućnost ostvarivanja vrijednosti nekog intervala

$$x, x + \Delta x$$

- Zakon razdiobe slučajne varijable $X$ je zadan ukoliko je poznata tzv. ***funkcija gustoće vjerojatnosti f(x)***

$$f(x) : x \in R$$

$$p(a \le x \le b) = \int_a^b f(x) \cdot dx$$

Funkcija gustoće je nenegativna funkcija, a vjerojatnost pojave sigurnog događaja je definirana putem slijedećeg integrala:

$$\int_{-\infty}^{+\infty} f(x)\,dx = 1$$

# Vjerojatnost kontinuirane varijable

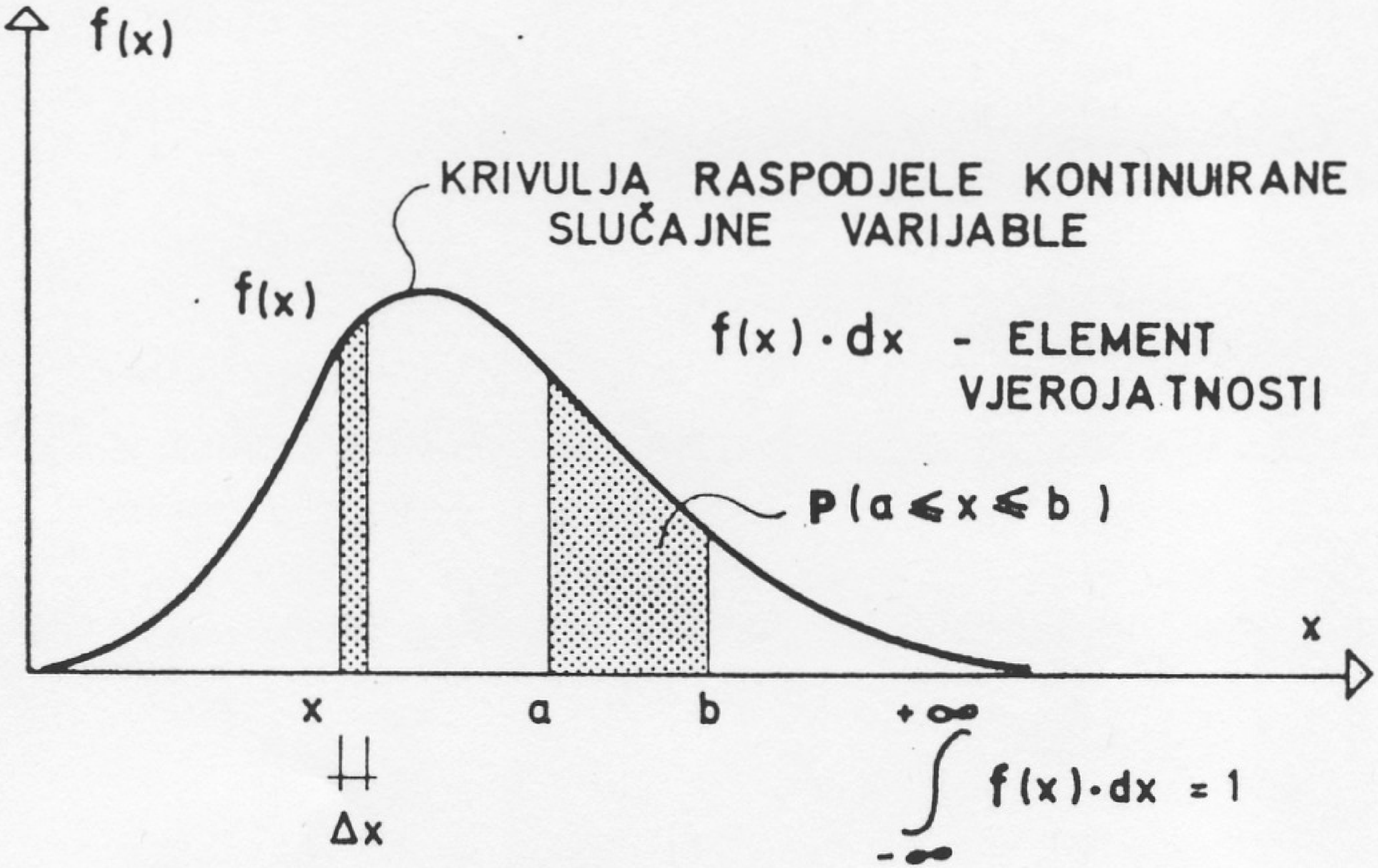
$$P(x_o \leq X \leq x_o + \Delta x) = \int_{x_o}^{x_o + \Delta x} f(x) \cdot dx$$

# Funkcija distribucije ili funkcija razdiobe slučajne varijable X

$$F(x) = P(X \le x)$$

$$F(x) = \int_{-\infty}^{x} f(x)dx$$

- Razdioba slučajne varijable je poznata ukoliko je poznata ili funkcija gustoće ili funkcija razdiobe.



f(x)
KRIVULJA RASPODJELE KONTINUIRANE SLUČAJNE VARIJABLE
f(x)
f(x)·dx - ELEMENT VJEROJATNOSTI
P(a ≤ x ≤ b)
x
x
a
b
+∞
Δx
$\int_{-\infty}^{+\infty} f(x)\cdot dx = 1$
-∞


SLIKA 4

- Hidrološki podaci su statističke varijable, a mogu biti i diskretne i kontinuirane.
- Dnevni, mjesečni protok je primjer diskretne varijable, a otjecanje prikazano putem hidrograma predstavlja kontinuiranu varijablu.
- Promatrat će se vjerojatnost pojave neke hidrološke veličine.

***Krivuljom raspodjele*** se dobiju značajne informacije o empirijskom nizu na temelju samo nekoliko parametara.

Normalna i lognormalna raspodjela su raspodjele određene na temelju 2 statistička parametra, srednje vrijednosti i standardne devijacije.

- Da bismo došli do pojma vjerojatnosti pojave neke hidrološke veličine potrebno je na temelju mjerenih podataka odrediti mjerodavnu funkciju raspodjele.
- Grafički prikaz funkcije raspodjele je krivulja raspodjele.

- Da bi se odredila vjerojatnost pojave slučajne varijable X na temelju hidrološkog niza potrebno je izvršiti statističku analizu koja podrazumijeva:
- proračun **osnovnih statističkih parametara** uzorka i određivanje **empirijske raspodjele**,
- izbor **teoretske funkcije raspodjele** i ocjenu parametara,
- testiranje prilagodbe **empirijske i teoretske** raspodjele.

- Osnovni statistički parametri uzorka:
- Aritimetička sredina
- Varijanca
- Koeficijent varijacije
- Koeficijent asimetrije

$$\bar{x} = \frac{1}{N}\sum_{I=1}^{N} x_i$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2$$

$$c_v = \frac{\sigma}{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2}{N\bar{x}^2}}$$

$$c_s = \frac{\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^3}{N\sigma}$$

- Podatke opažanja je potrebno najprije poredati u opadajući niz:

$$x_i \leq x_{i+1}$$

- Funkcija raspodjele se definira kao vjerojatnost pojave da slučajna varijabla poprimi određenu promatranu vrijednost ili vrijednost veću od promatrane.

$$F(X) = P(X \geq x)$$

- Empirijska razdioba se aproksimira izrazom:

$$P(X \geq x_m) = \frac{m - 0{,}5}{N}$$

- Teoretska funkcija raspodjele je dana analitičkim izrazom:

$$P(X \geq x) = F(x, \alpha, \beta, \gamma, ...)$$

- čiji oblik ovisi o parametrima
- Primjena teoretskih funkcija raspodjele pruža mogućnost da se ekstrapoliraju teoretske funkcije raspodjele u područja malih vjerojatnosti.

- Krivulje raspodjele se u hidrologiji koriste za određivanje vjerojatnosti pojave hidroloških fenomena.
- Oblici krivulja raspodjele mogu biti različiti, a kao primjer će se obrađivati normalna i lognormalna raspodjela

# NORMALNA RASPODJELA

- Normalna raspodjela je najčešće opisivana i najčešće korištena
- Normalna raspodjela je dvoparametarska i u cijelosti je definirana poznavanjem aritimetičke sredine i standardne devijacije

$$(\overline{x}, \sigma)$$

# Funkcija gustoće normalne raspodjele

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

- Krivulja normalne raspodjele je simetrična pa je njen koeficijent asimetrije jednak nuli.
- Koeficijent asimetrije:

$$r_3 = \frac{\mu_3}{\sigma^3} = n^{1/2} \frac{\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^3}{\left(\sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2\right)^{3/2}}$$

# Koeficijent spljoštenosti

$$r_4 = \frac{\mu_4}{\sigma^4} = n \frac{\sum_{i=1}^{n} \left( x_i - \overline{x} \right)}{\left( \sum_{i=1}^{n} \left( x_i - \overline{x} \right)^2 \right)^2}$$

- Za $r_4>3$, krivulja je s visokim vrhom, a za $r_4<3$, krivulja je s plitkim vrhom raspodjele

# LOGARITAMSKO NORMALNA RASPODJELA

$$f(x)=\frac{1}{\sigma'\sqrt{2\pi}}e^{-\frac{\left(\ln x-\overline{x}\right)}{}}$$

# NORMALNA RASPODJELA

$$Q_p = \overline{Q} + z_p \sigma$$

| *T(god)* | *P (%)* | $z_p$ | $Q_p$ |
|---|---|---|---|
| 100 | 0,01 | 3,715 | |
| 10 | 0,1 | 3,090 | |
| 1 | 1 | 2,326 | |
| 0,5 | 2 | 2,054 | |
| 0,25 | 4 | 1,752 | |
| 0,1 | 10 | 1,281 | |
| 0,05 | 20 | 0,842 | |

# LOGNORMALNA RASPODJELA

$$\ln Q_p = \overline{Q}' + z_p \sigma'$$

$$\overline{Q}' = \ln \frac{\overline{Q}^2}{\sqrt{\sigma^2 + \overline{Q}^2}}$$

$$\sigma' = \sqrt{\ln\left(1 + \frac{\sigma^2}{\overline{Q}^2}\right)}$$

# LOGNORMALNA RASPODJELA

$$\ln Q_p = \overline{Q}' + z_p \sigma'$$

$$\overline{Q}' = \ln \frac{\overline{Q}^2}{\sqrt{\sigma^2 + \overline{Q}^2}}$$

$$\sigma' = \sqrt{\ln\left(1 + \frac{\sigma^2}{\overline{Q}^2}\right)}$$

| *T(god)* | *P (%)* | $z_p$ | $Q_p$ |
|---|---|---|---|
| 100 | 0,01 | 3,715 | |
| 10 | 0,1 | 3,090 | |
| 1 | 1 | 2,326 | |
| 0,5 | 2 | 2,054 | |
| 0,25 | 4 | 1,752 | |
| 0,1 | 10 | 1,281 | |
| 0,05 | 20 | 0,842 | |

- Dobivene vrijednosti se grafički mogu prikazati na pravcu normalne raspodjele.
- Kad bi se dobivene vrijednosti crtale u dekadskom mjerilu, tj. običnom koordinatnom sustavu taj prikaz ne bi bio pravac, već krivulja.
- Radi jednostavnije usporedbe dviju raspodjela (empirijske i teoretske), one se grafički prikazuju na tzv. papiru vjerojatnosti, a pravac koji odgovara funkciji raspodjele se u statistici zove Henry-ev pravac.

- Pravac prolazi kroz dvije karakteristične točke:

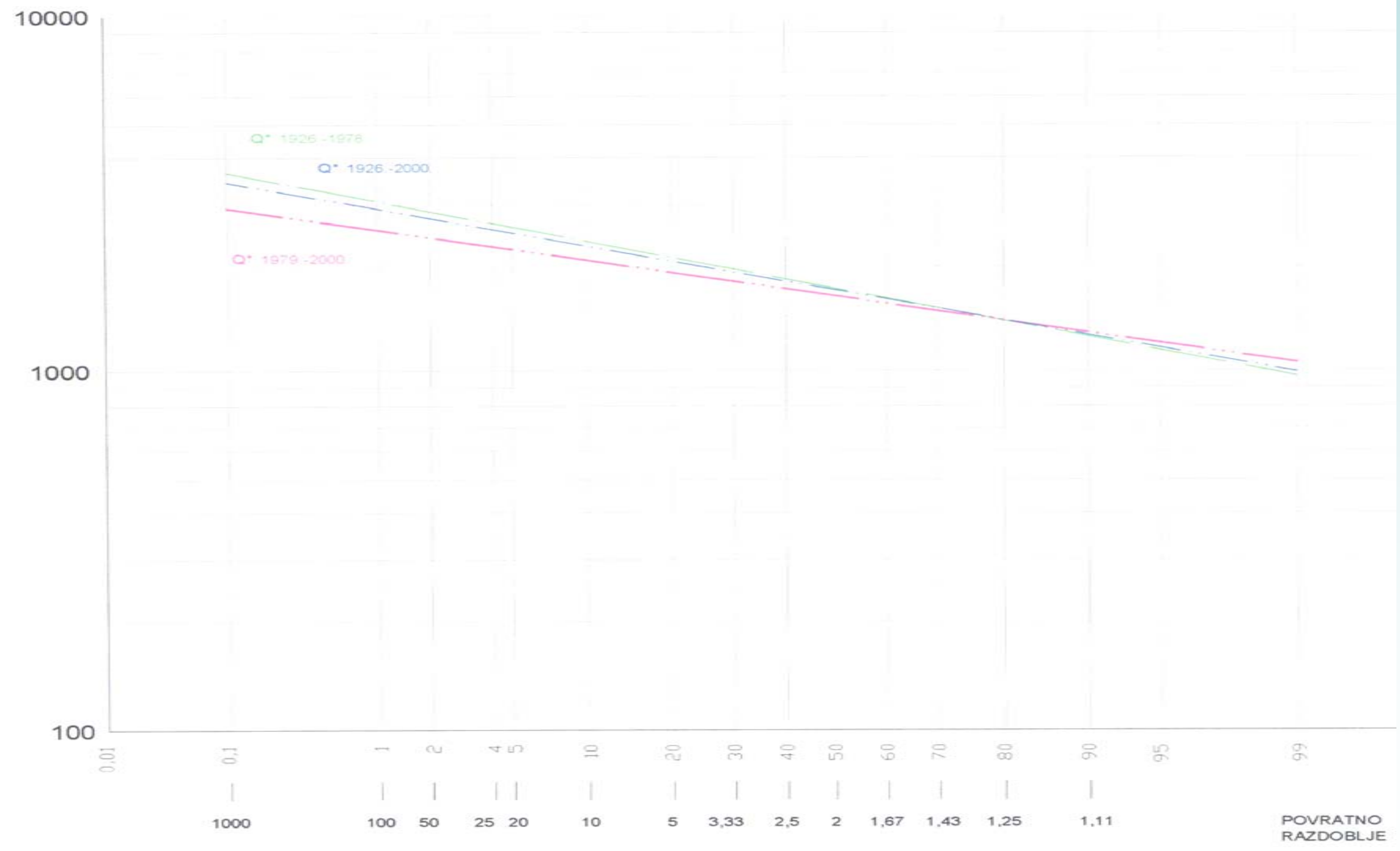
$$\left(\overline{x};\, 0,5\right)$$

$$(\overline{x}+\sigma;0,84134\,)$$

- Te dvije točke mogu poslužiti kao kontrola prilagodbe.

# LOG-NORMALNE KRIVULJE

## RASPODJELE ZA TRI NIZA MAKSIMALNIH GODIŠNJIH PROTOKA SAVE KOD ZAGREBA RAZLIČITE DULJINE TRAJANJA



$$\chi^2 - test$$

Kod određivanja krivulja raspodjele postavlja se pitanje prilagodbe teoretske i empirijske raspodjele tj. pitanje koliko dobro neka raspodjela prati raspodjelu empirijskih podataka.

Odgovor na to se daje putem gore navedenog testa.

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(f_i - f_{ti}\right)^2}{f_{ti}}$$

Ako razlike između empirijskih i teoretskih frekvencija nisu prevelike tj. imaju slučajan karakter tada i pripadna vrijednost testa neće biti velika. Ukoliko su razlike prevelike, tada svakako nisu slučajne te treba odbaciti pretpostavku da se varijabla ravna po navedenoj raspodjeli.

- Ukoliko je

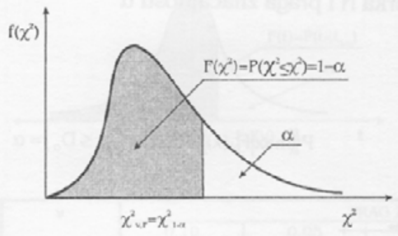
$$\chi^2 \leq \chi^2_{dop}$$

$\chi^2_{dop}$ ovisi o broju stupnjeva slobode

Broj stupnjeva slobode k = broj razreda r - 2 - 1

$k = r - 2 - 1$

# $\chi^2 - raspodjela$




$$f(\chi^2)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}(\chi^2)^{\frac{\nu-2}{2}}\cdot e^{-\frac{\chi^2}{2}}$$

$$\nu>0 \quad 0\le\chi^2\le+\infty$$

| $\nu$ | Vrednosti $\chi^2_{1-\alpha}$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0.50 | 0.25 | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| 1 | 0.455 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 1.386 | 2.773 | 4.605 | 5.991 | 7.378 | 9.210 | 10.60 |
| 3 | 2.366 | 4.108 | 6.251 | 7.815 | 9.348 | 11.34 | 12.84 |
| 4 | 3.357 | 5.386 | 7.779 | 9.488 | 11.14 | 13.28 | 14.86 |
| 5 | 4.351 | 6.626 | 9.236 | 11.07 | 12.83 | 15.09 | 16.75 |
| 6 | 5.348 | 7.841 | 10.64 | 12.59 | 14.45 | 16.81 | 18.55 |
| 7 | 6.346 | 9.037 | 12.02 | 14.07 | 16.01 | 18.48 | 20.28 |
| 8 | 7.344 | 10.22 | 13.36 | 15.51 | 17.53 | 20.09 | 21.96 |
| 9 | 8.343 | 11.39 | 14.68 | 16.92 | 19.02 | 21.67 | 23.59 |
| 10 | 9.342 | 12.55 | 15.99 | 18.31 | 20.48 | 23.21 | 25.19 |
| 11 | 10.34 | 13.70 | 17.28 | 19.68 | 21.92 | 24.72 | 26.78 |
| 12 | 11.34 | 14.85 | 18.55 | 21.03 | 23.34 | 26.22 | 28.30 |
| 13 | 12.34 | 15.98 | 19.81 | 22.36 | 24.74 | 27.69 | 29.82 |
| 14 | 13.34 | 17.12 | 21.06 | 23.68 | 26.12 | 29.14 | 32.32 |
| 15 | 14.34 | 18.25 | 22.31 | 25.00 | 27.49 | 30.58 | 32.80 |
| 16 | 15.34 | 19.37 | 23.54 | 26.30 | 28.85 | 32.00 | 34.27 |
| 17 | 16.34 | 20.49 | 24.77 | 27.59 | 30.19 | 33.41 | 35.72 |
| 18 | 17.34 | 21.60 | 25.99 | 28.87 | 31.53 | 34.81 | 37.48 |
| 19 | 18.34 | 22.72 | 27.20 | 30.14 | 32.85 | 36.19 | 38.58 |
| 20 | 19.34 | 23.83 | 28.41 | 31.41 | 34.17 | 37.57 | 40.00 |
| 21 | 20.34 | 24.93 | 29.62 | 32.67 | 35.48 | 38.93 | 41.40 |
| 22 | 21.34 | 26.04 | 30.81 | 33.92 | 36.78 | 40.29 | 42.80 |
| 23 | 22.34 | 27.14 | 32.01 | 35.17 | 38.08 | 41.64 | 44.18 |
| 24 | 23.34 | 28.24 | 33.20 | 36.42 | 39.36 | 42.98 | 45.56 |
| 25 | 24.34 | 29.34 | 34.38 | 37.65 | 40.65 | 44.31 | 46.93 |
| 26 | 25.34 | 30.43 | 35.56 | 38.89 | 41.92 | 45.64 | 48.29 |
| 27 | 26.34 | 31.53 | 36.74 | 40.11 | 43.19 | 46.96 | 49.64 |
| 28 | 27.34 | 32.62 | 37.92 | 41.34 | 44.46 | 48.28 | 50.99 |
| 29 | 28.34 | 33.71 | 39.09 | 42.56 | 45.72 | 49.59 | 52.34 |
| 30 | 29.34 | 34.80 | 40.26 | 43.97 | 46.98 | 50.89 | 53.67 |

- Općenito, izraz

$$P\left\{\chi^2 > \chi^2_{1-\alpha}\right\} = \alpha$$

- predstavlja vjerojatnost odbacivanja hipoteze da se raspodjela ravna po određenoj promatranoj raspodjeli

# PRIMJENA TESTA

- U primjenama ta vjerojatnost odbacivanja hipoteze iznosi 5%, a rjeđe 1%.

$$\alpha = 0{,}05$$

# PRILAGOĐAVANJE EMPIRIJSKIM PODACIMA

- Pretpostavimo da je N empirijskih podataka grupirano u razrede od kojih svaki ima širinu $\Delta$.
- Neka je poznata srednja vrijednost i varijanca navedenog skupa.
- Da bismo empirijskoj raspodjeli prilagodili normalnu raspodjelu, aproksimirat ćemo očekivanje $\mu$ aritimetičkom sredinom, a varijancu varijancom empirijskih podataka.

- Teoretske frekvencije računat ćemo za svaki razred, služeći se funkcijom vjerojatnosti:

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

- Neka je $x_i$ sredina i-tog razreda kojem pripada empirijska vjerojatnost $f_i$.
- Vjerojatnost da varijabla primi vrijednost iz i-tog razreda jednaka je površini ispod krivulje vjerojatnosti nad dotičnim razredom.

- Ta je vjerojatnost približno jednaka:

$$P_i \approx \Delta \cdot f(x_i)$$

- Uvedu li se supstitucijski izrazi:

$$u = \frac{x-\mu}{\sigma}$$

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

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

$\varphi(u)$

Funkcija standardne ili jedinične normalne raspodjele čije se vrijednosti za različite u mogu naći tabelarno ili se dobiju proračunom. Te vrijednosti služe za proračun vrijednosti f(x) pri čemu vrijedi da je:

$$f(x) = \frac{1}{\sigma}\varphi(u)$$

- Ako se vratimo na vjerojatnost da varijabla primi vrijednost iz i-tog razreda slijedi:

$$P_i \approx \Delta \cdot f(x_i) = \frac{\Delta}{\sigma} \cdot \varphi(u_i)$$

$$f_{ti} = N \cdot P_i$$

- Teoretska ili očekivana frekvencija $f_{ti}$ u i-tom razredu:

$$f_{ti} = \frac{N \cdot \Delta}{\sigma} \cdot \varphi(u_i)$$

$$x_i \to u_i \to \varphi(u_i) \to f_{ti}$$

- Empirijske frekvencije $f_i$ i teoretske frekvencije $f_{ti}$ neće se podudarati. Ako se varijabla na koju se odnose empirijski podaci pokoravaju zakonu normalne raspodjele razlike neće biti prevelike, imat će slučajan karakter.

| RAZRED | $f_{ei}$ | $x_i$ | $u_i$ | $\varphi(u_i)$ | $f_{ti}$ | $(f_{ei}-f_{ti})^2/f_{ti}$ |
|---|---|---|---|---|---|---|
| 0-22 | | 11 | | | | |
| 22-44 | | 34 | | | | |
| 44-66 | | 7 | | | | |

$$u_i = \frac{Q_i - \overline{Q}}{\sigma}$$

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

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

$$\boxed{P_i \approx \Delta \cdot f(x_i) = \frac{\Delta}{\sigma} \cdot \varphi(u_i)}$$

$$f_{ti} = N \cdot P_i$$

$$f_{ti} = \frac{N \cdot \Delta}{\sigma} \cdot \varphi(u_i)$$

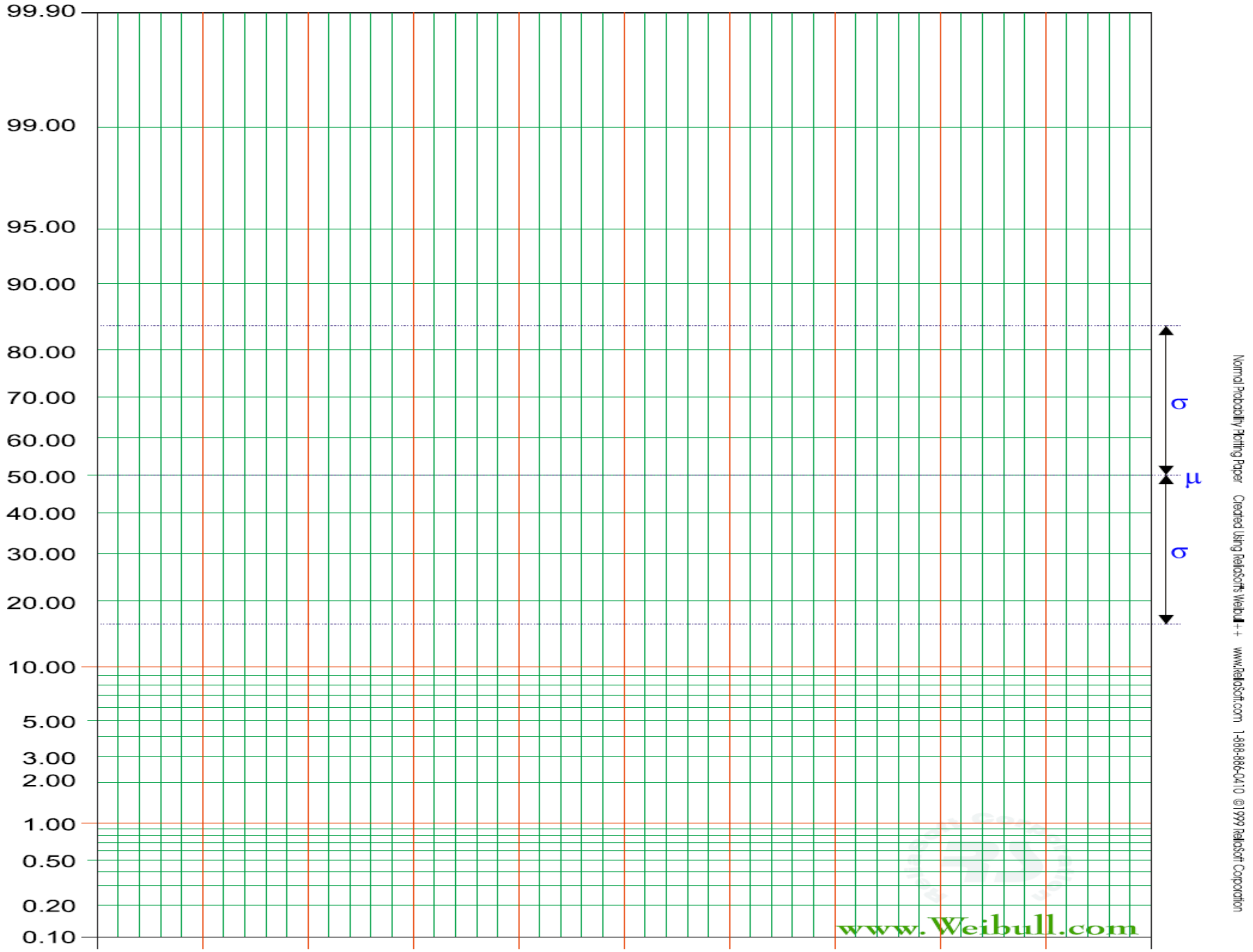
$$x_i \to u_i \to \varphi(u_i) \to f_{ti}$$

# Kontrola točnosti

$$\sum_i f_{ei} \approx \sum_i f_{ti}$$

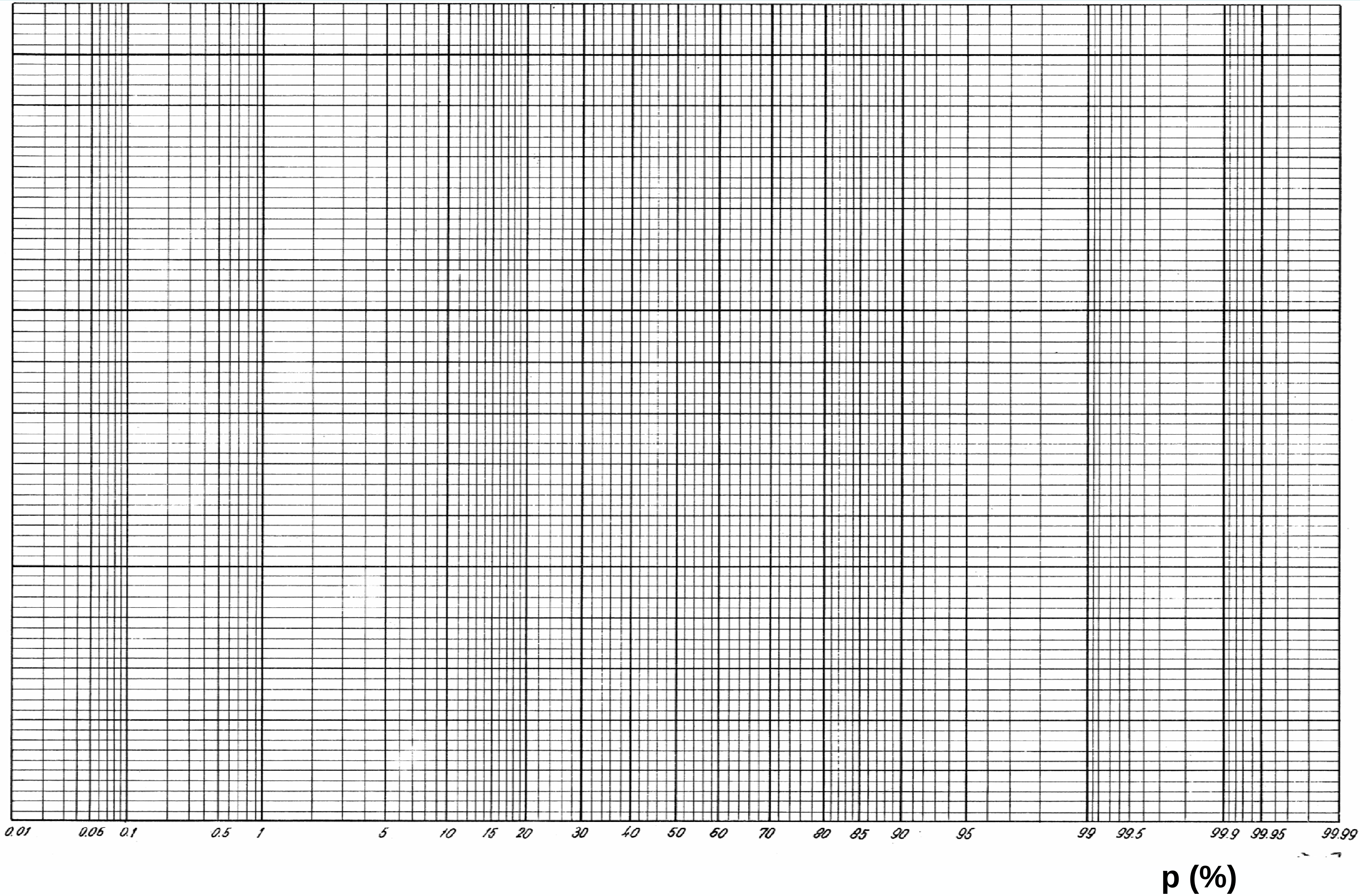
# LOGNORMALNA RASPODJELA

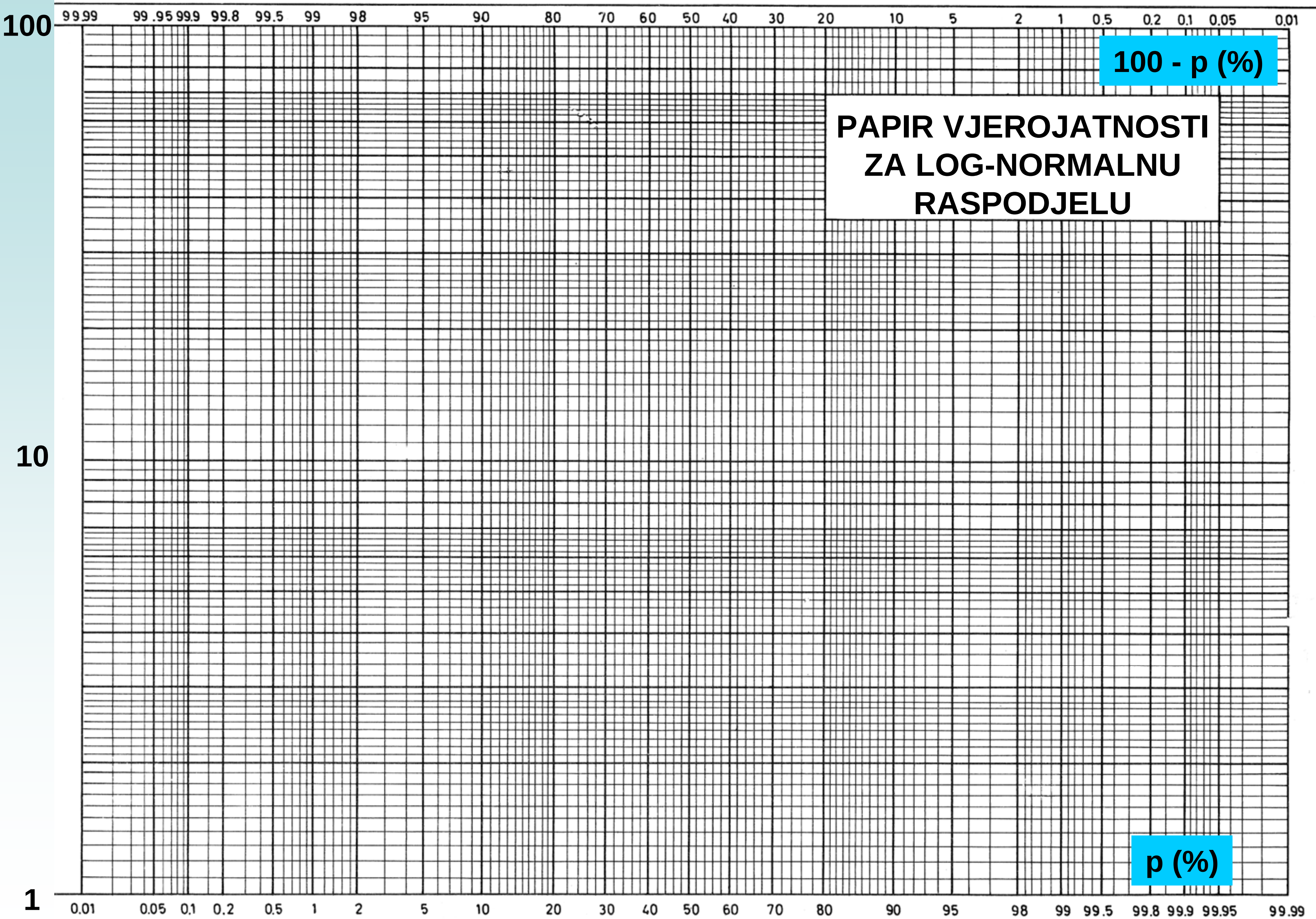
| RAZREDI | GORNJA GRANICA RAZREDA | $p_i(\%)$ | $\Delta p_i = p_i - p_{i-1}$ | $f_{ti} = \Delta p_i * n$ | $f_{ei}$ | $(f_{ei} - f_{ti})^2 / f_{ti}$ |
|---|---|---|---|---|---|---|
| 420-520 | 420 | | | | | |
| 520-620 | 520 | | | | | |
| | | | | | | |
| | | | | | | |



99.90
99.00
95.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
5.00
3.00
2.00
1.00
0.50
0.20
0.10
σ
μ
σ
www.Weibull.com
Normal Probability Plotting Paper Created Using ReliaSoft's Weibull++ www.ReliaSoft.com 1-888-886-0410 ©1999 ReliaSoft Corporation

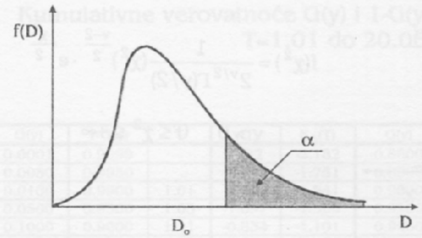
# PAPIR VJEROJATNOSTI ZA NORMALNU RASPODJELU




PAPIR VJEROJATNOSTI ZA LOG-NORMALNU RASPODJELU
100 - p (%)
p (%)
100
10
1
99.99 99 .95 99.9 99.8 99.5 99 98 95 90 80 70 60 50 40 30 20 10 5 2 1 0.5 0.2 0.1 0.05 0.01
0.01 0.05 0.1 0.2 0.5 1 2 5 10 20 30 40 50 60 70 80 90 95 98 99 99.5 99.8 99.9 99.95 99.99

Test Kolmogorov-Smirnova

Granične vrijednosti $D_o$ u funkciji veličine uzorka N i praga značajnosti $\alpha$




$$P\{(\max|F_N(x)| - F(x)) = D_N \leq D_o\} = \alpha$$

| OBIM UZORAKA (N) | PRAG ZNAČAJNOSTI α | | | | |
|---|---|---|---|---|---|
| | 0.20 | 0.15 | 0.10 | 0.05 | 0.01 |
| 1 | 0.900 | 0.925 | 0.950 | 0.975 | 0.995 |
| 2 | .684 | .726 | .776 | .842 | .929 |
| 3 | .565 | .597 | .642 | .708 | .829 |
| 4 | .494 | .525 | .564 | .624 | .734 |
| 5 | .446 | .474 | .510 | .563 | .669 |
| 6 | .410 | .436 | .470 | .521 | .618 |
| 7 | .381 | .405 | .438 | .486 | .577 |
| 8 | .358 | .381 | .411 | .457 | .543 |
| 9 | .339 | .360 | .388 | .432 | .514 |
| 10 | .322 | .342 | .368 | .409 | .486 |
| 11 | .307 | .326 | .352 | .391 | .468 |
| 12 | .295 | .313 | .338 | .375 | .450 |
| 13 | .284 | .302 | .325 | .361 | .433 |
| 14 | .274 | .292 | .314 | .349 | .418 |
| 15 | .266 | .283 | .304 | .338 | .404 |
| 16 | .258 | .274 | .295 | .328 | .691 |
| 17 | .250 | .266 | .286 | .318 | .380 |
| 18 | .244 | .259 | .278 | .309 | .370 |
| 19 | .237 | .252 | .272 | .301 | .361 |
| 20 | .231 | .246 | .264 | .294 | .352 |
| 25 | .210 | .220 | .240 | .264 | .320 |
| 30 | .190 | .200 | .220 | .242 | .290 |
| 35 | .180 | .190 | .210 | .230 | .270 |
| 40 | .170 | .180 | .190 | .190 | .230 |
| 50 | .150 | .160 | .170 | .190 | .230 |
| 60 | .140 | .150 | .160 | .170 | .210 |
| 70 | .130 | .140 | .150 | .160 | .190 |
| 80 | .120 | 0.130 | .140 | .150 | .180 |
| 90 | .110 | .120 | .130 | .140 | .170 |
| 100 | .100 | .110 | .120 | .140 | .160 |
| ASIMPTOTSKA FORMULA | $\frac{1.07}{\sqrt{N}}$ | $\frac{1.14}{\sqrt{N}}$ | $\frac{1.22}{\sqrt{N}}$ | $\frac{1.36}{\sqrt{N}}$ | $\frac{1.63}{\sqrt{N}}$ |

# Test Kolmogorova

- Za ocjenu prilagodbe empirijske raspodjele teoretskom uzima se najveća apsolutna razlika između te dvije funkcije raspodjele:

$$D_n = \max\left|F_n(x) - F(x)\right|$$

$$D_n \leq D_o$$

$$P(D_n \geq D_o) = \alpha$$

- $D_{dop}$ su različite ovisno o pragu značajnosti $\alpha$ i opsegu uzorka n, te su dane tabelarno.