Curs 7: Distribuții continue clasice: uniformă, exponențială, normală

Conf.dr. Maria Jivulescu

Departamentul de Matematică
UPT

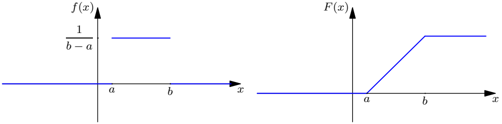
Cuprins

Distribuția uniformă, exponențială, normală:

- Definiție. Densitate de probabilitate.
- Funcția de repartiție.
- Medie. Dispersie

Distribuția uniformă

- Parametrii: a, b
- Domeniu de valori(adică mulțimea de valori posibile unde $f_X(x) > 0$): $[a, b]$
- Notație: Unif$[a,])$ sau uniform[a,b];
- Model: valorile din domeniu au probabilitate egală de apariție;
- den. de prob. $f(x) = \begin{cases} \frac{1}{b-a} & \text{dacă } x \in [a, b] \\ 0 & \text{dacă } x \notin (a, b) \end{cases}$
- funcție de repartiție $F(x) = \begin{cases} 0, & \text{dacă} \quad x < a \\ \frac{x-a}{b-a} & \text{dacă} \quad a \leq x < b \\ 1 & \text{dacă} \quad x \geq b \end{cases}$



https://mathlets.org/mathlets/probability-distributions/

Distribuția exponențială

- Parametru: θ;
- Domeniu de valori $[0, \infty)$;
- Notație:Exp(θ);
- Den de prob. $f(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ \frac{1}{\theta} e^{-x/\theta} & \text{dacă } x \geq 0 \end{cases}, \quad \theta > 0$
- Funcția de repartiție: $F(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ 1 - e^{-x/\theta} & \text{dacă } x \geq 0 \end{cases},$
- $M(X) = \theta$, $\sigma^2(X) = \theta^2$
- Model: timpul de așteptare pentru un proces continuu
- Lipsa de memorie: probabilitatea de a aștepta încă t minute nu este afectată de aceea de a fi așteptat deja s minute fără eveniment.

$$P(X > s + t / X > s) = P(X > t)$$

.

Distribuția normală

Definiție

O variabilă aleatoare continua, X, ce are densitatea de probabilitate

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-m)^2}{2\sigma^2}, \quad x \in \mathbb{R},$$

se numește variabilă aleatoare normal distribuită, de parametri m, σ, unde $m \in \mathbb{R}$, $\sigma > 0$ (notăm cu $N(m, \sigma)$).

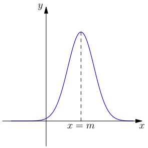


Figure: Graficul densității de probabilitate pentru $X \sim N(m = 2, \sigma = 0.75)$.

Distribuția normală

- Parametrii: m, σ;
- Domeniu de valori $(-\infty, \infty)$;
- Notație: $N(m, \sigma)$ sau $N(m, \sigma^2)$
- Den de prob: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-m)^2}{2\sigma^2}, \quad x \in \mathbb{R}$,
- Funcția de repartiție: nu are formulă (valorile se pot determina folosind tabele sau software matematice: `pnorm` in R)
- Dacă $X \sim N(m, \sigma^2)$, atunci $M(X) = m, \sigma^2(X) = \sigma^2$.
- Model: măsurarea erorilor, înălțime, media big data;

Distribuită normal standard

- Parametrii: $m = 0, \sigma = 1$;
- Den de prob: $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \; t \in \mathbb{R}.$;
- φ este pară;
- graficul său se numește clopotul lui Gauss;
- Notație $Z \sim N(0,1)$;
- Funcția de repartiție: $\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$
- $M(Z) = 0, \sigma^2(Z) = 1$

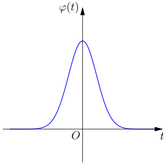


Figure: Densitatea de probabilitate a unei variabile aleatoare $Z \sim N(0,1)$.

Distribuită normal standard

Funcția de repartiție a variabilei aleatoare $Z \sim N(0,1)$ se notează cu Φ:

$$\Phi(x) = P(Z \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

Proprietăți:

- $\Phi(-x) = 1 - \Phi(x),\ \ x \in \mathbb{R}$
- Valoarea funcției Φ într-un punct x se poate calcula doar prin metode aproximative, dar există tabele pentru aceste valori;

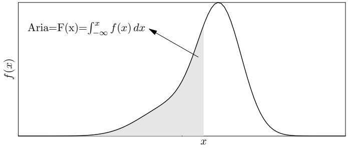


Figure: Ilustrarea semnificației geometrice a valorii funcției de repartiție într-un punct.

Calculul probabilităților: regula 3-σ

De exemplu, dacă $X \sim N(\mu, \sigma^2)$, atunci

$$P(\mu - \sigma \leq X \leq \mu + \sigma) \approx 0.68,$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95,$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.99.$$

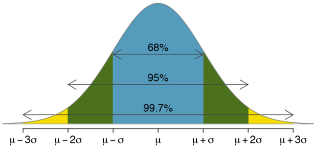


Figure: Regula 3-σ a unei variabile aleatoare $X \sim N(\mu, \sigma)$.

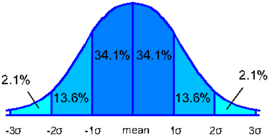


Figure: Regula 3-σ a unei variabile aleatoare $Z \sim N(0,1)$.

De exemplu, dacă $Z \sim N(0,1)$, atunci

$$P(-1 \leq Z \leq 1) \approx 0.68,$$

$$P(-2 \leq Z \leq 2) \approx 0.95,$$

$$P(-3 \leq Z \leq 3) \approx 0.99.$$

De exemplu, se poate calcula $\phi(1)$. Intr-adevăr,
$\phi(1) = P(Z \leq 1) \approx 0.68 + 0.16 = 0.84$.

Definiție mediană

Mediana unei variabile aleatore X este valoarea x pentru care $P(X \leq x) = P(X \geq x) = 0.5$.
Adică, acea valoare x pentru care $F(x) = 0.5$.

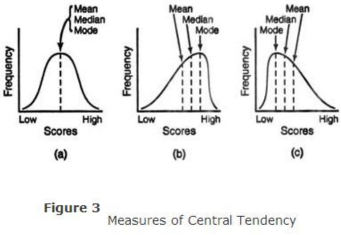


Figure: Media, mediana și modulul unei variabile aleatoare.

Cvantila distribuției normale standard

In general, α–cvantila este unicul număr real x_α, unde $\alpha \in (0,1)$, pentru care $F(x_\alpha) = P(X \leq x_\alpha) = \alpha$.
Echivalent, $x_\alpha = F^{-1}(\alpha)$, unde $F : \mathbb{R} \to (0,1)$, strict crescătoare, inversa $F^{-1} : (0,1) \to \mathbb{R}$.
Evident, mediana este $x_{0.5}$.
Observație: Pentru $Z \sim N(0,1)$, notăm z_α.

Propoziție

Între cvantilele $z_{1-\alpha}$ și z_α ale unei variabile aleatoare $Z \sim N(0,1)$, are loc:

$$z_{1-\alpha} = -z_\alpha \; \alpha \in (0,1).$$

Avem, $z_{0.5} = 0$, deci $P(Z \leq 0) = 0.5$.
`https://www.calculator.net/z-score-calculator.html?c1raw=0&c1mean=0&c1sd=1&calctype=zscore&x=0&y=0`

Binemeritată pauză

Pauză!

Dupa pauză... Transformarea variabilelor aleatoare

Transformarea variabilelor aleatoare

Problemă: Dacă X este o v.a. continuă și $Y = g(X)$.
Cum se pot determina densitatea de prob f_Y (p.d.f.) și funcția de repartiție F_Y (c.d.f.) ale lui Y (cunoscând f_X și F_X)?

Exemplu: Fie $X \sim Unif[0,1]$ și $Y = e^X$. Să se determine F_Y, f_y și $M(Y)$.

Soluție: Se știe că $F_X(x) = \begin{cases} 0, & \text{dacă} \quad x < 0 \\ x, & \text{dacă} \quad 0 \leq x < 1 \\ 1, & \text{dacă} \quad x \geq 1 \end{cases}$

Care este domeniul de valori al lui Y, i.e. D_Y?
Evident, $D_Y = [1, e]$ (e^x este o funcție crescătoare).
Prin definiție avem că $F_Y(y) = P(Y \leq y)$.

- $F_Y(y) = P(Y \leq y) = 0$, pentru $y < 1$;
- $F_Y(y) = P(Y \leq y) = 1$, pentru $y > e$;
- $F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq ln(y)) = F_X(ln(y)) = ln(y)$, pentru $1 \leq y \leq e$;

Deci, $F_Y(y) = \begin{cases} 0, & \text{dacă} \quad y < 1 \\ In(y), & \text{dacă} \quad 1 \leq y < e \\ 1, & \text{dacă} \quad y \geq e \end{cases}$

Pentru a obține p.d.f. f_Y, vom deriva funcția F_Y (pentru ca F_Y este continuă). Avem:

$f_Y(y) = F_Y'(y) = \begin{cases} \frac{1}{y}, & \text{dacă} \quad 1 \leq y \leq e \\ 0, & \text{altfel} \end{cases}$

Pentru a determina $M(Y)$, aplicăm LOTUS:

$$M(Y) = M(e^X) = \int_{-\infty}^{\infty} e^x f_X(x)dx = \int_0^1 e^x dx = e - 1.$$

Sau

$$M(Y) = \int_{-\infty}^{\infty} y f_Y(y)dy = \int_1^e y\frac{1}{y}dy = e - 1.$$

Proprietate

Fie $X \sim N(m, \sigma^2)$. Atunci, $Z = \frac{X-m}{\sigma}$ are proprietatea că $Z \sim (0, 1)$.

Intr-adevăr, se poate demonstra că

$$f_X(x) = \frac{1}{\sigma} \varphi_Z \left(\frac{x-m}{\sigma} \right), \quad F_X(x) = \Phi \left(\frac{x-m}{\sigma} \right)$$

Cum putem folosi această informație?

Fie v.a $X \sim N(m, \sigma^2)$. Avem:

$P(a \leq X \leq b) = P(a < X \leq b) = F_X(b) - F_X(a) = ?$

$P(a \leq X \leq b) = P(\frac{a-m}{\sigma} \leq Z = \frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}) = \Phi \left(\frac{b-m}{\sigma} \right) - \Phi \left(\frac{a-m}{\sigma} \right)$

Exemplu:

Fie $X \sim N(1, 0.4^2)$. Să se calculeze: $P(0.75 < X \leq 1.3)$ și să se determine $x \in \mathbb{R}$ astfel încât $P(X \leq x) = 0.95$?

Rezolvare: $P(0.75 < X \leq 1.3) = F_X(1.3) - F_X(0.75) =?$
Vom standardiza variabila X, transformând-o în variabila $Z = \frac{X-m}{\sigma}$.
Avem: $P(0.75 < X \leq 1.3) = P(\frac{0.75-1}{0.4} < Z \leq \frac{1.3-1}{0.4}) =$
$\Phi((1.3-1)/0.4) - \Phi((0.75-1)/0.4) = \Phi(0.75) - \Phi(-0.625) =$
$\Phi(0.75) - 1 + \Phi(0.625)$.
Avem de calculat $P(X \leq x) = F_X(x) = \Phi((x-1)/0.4) = 0.95$.
Deci $(x-1)/0.4 = \Phi^{-1}(0.95) = z_{0.95}$. Din tabele se știe că 0.95–cvantila distribuției normale standard este $z_{0.95} = 1.64$.
Deci $x = 1 + 0.4 \cdot 1.64$.