Statistika i osnovna mjerenja

Kontinuirane raspodjele

M. Makek

2017/2018

Kontinuirana slučajna varijabla

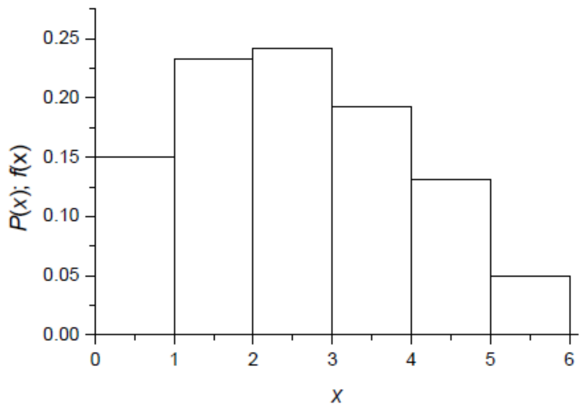
Definicija: Slučajna varijabla ***X*** je **kontinuirana** ako skup njenih vrijednosti čini interval brojeva, tj. ako je za ***A*** i ***B*** **(*A<B*)** svaki ***x* Є [*A,B*]** moguć.

- Kontinuirana slučajna varijabla će biti reprezentacija opservabli poput mase, duljine, vremena, itd.
- Većina karakterističnih veličina (očekivanje, varijanca, idr.) kontinuirane slučajne varijable može se dobiti iz veličina diskretne slučajne varijable uz prelazak sa ***sume*** na ***integral***

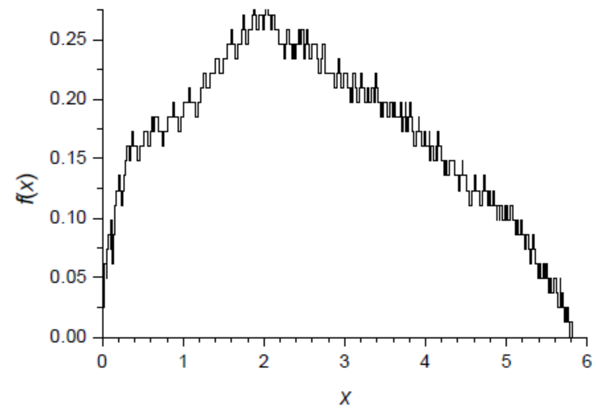
Kontinuirana slučajna varijabla

- Primjer: mjerimo dubinu jezera na ~10000 mjesta

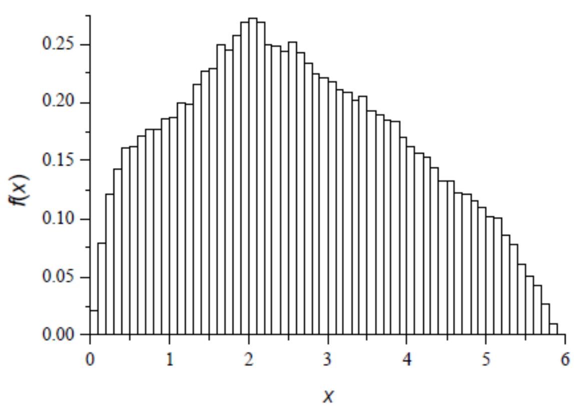
zaokruženo na 1 m



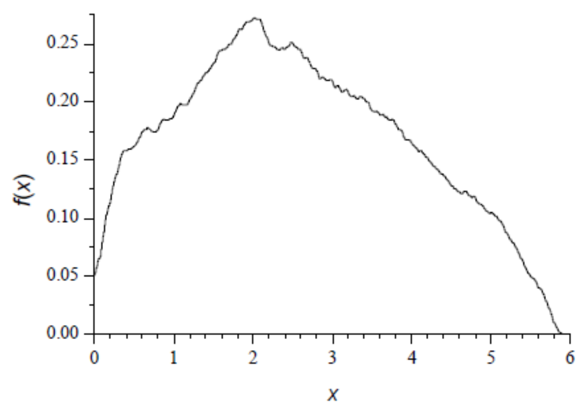
zaokruženo na 1 cm



zaokruženo na 1 dm



kontinuirana raspodjela



Prijelaz s diskretne na kontinuiranu: $p(x) \to f(x) \cdot dx$ $\quad \sum_{x \in D} \to \int_{-\infty}^{\infty}$

Funkcija gustoće vjerojatnosti

Definicija: ***funkcija gustoće vjerojatnosti*** kontinuirane slučajne varijable ***X*** je funkcija ***f(x)*** takva da za bilo koja dva broja $a \leq b$ vrijedi:

$$\boxed{P(a \leq X \leq b) = \int_a^b f(x)dx}$$

Iz definicije slijedi da je:
$P(X=x_0) = 0$, za svaki x_0!

→ *Vjerojatnost da X poprimi vrijednost u intervalu [a,b] dana je površinom ispod funkcije gustoće vjerojatnosti u tom intervalu*

- Pri tome *f(x)* zadovoljava svojstva:
  1. $f(x) \geq 0, \qquad \forall x$
  2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Funkcija raspodjele

Funkcija raspodjele (kumulativna funkcija distribucije) diskretne slučajne varijable ***X*** definirana je:

$$\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy}$$

- Općenito vrijedi:

$\boxed{P(a \leq X \leq b) = F(b) - F(a)}$ → distribucijska funkcija ima važnu ulogu kod kontuiruiranih sl. var. jer pomoću nje jednostavno dobivamo vjerojatnost

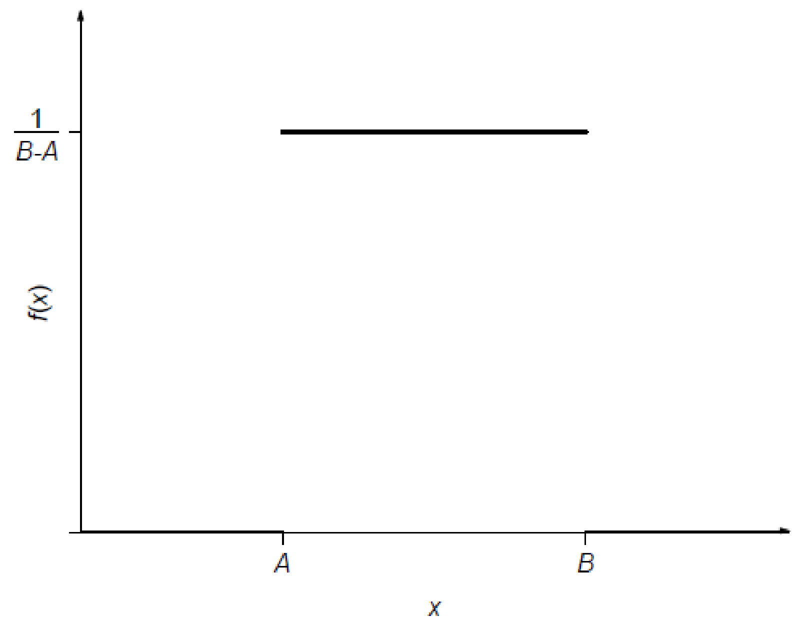
- Iz distribucijske funkcije jednostavno dobivamo funkciju gustoće vjerojatnosti iz:

$$f(x) = \frac{dF(x)}{dx}$$

Funkcija gustoće vjerojatnosti

Primjer: ***uniformna ili pravokutna raspodjela*** u intervalu [A,B]

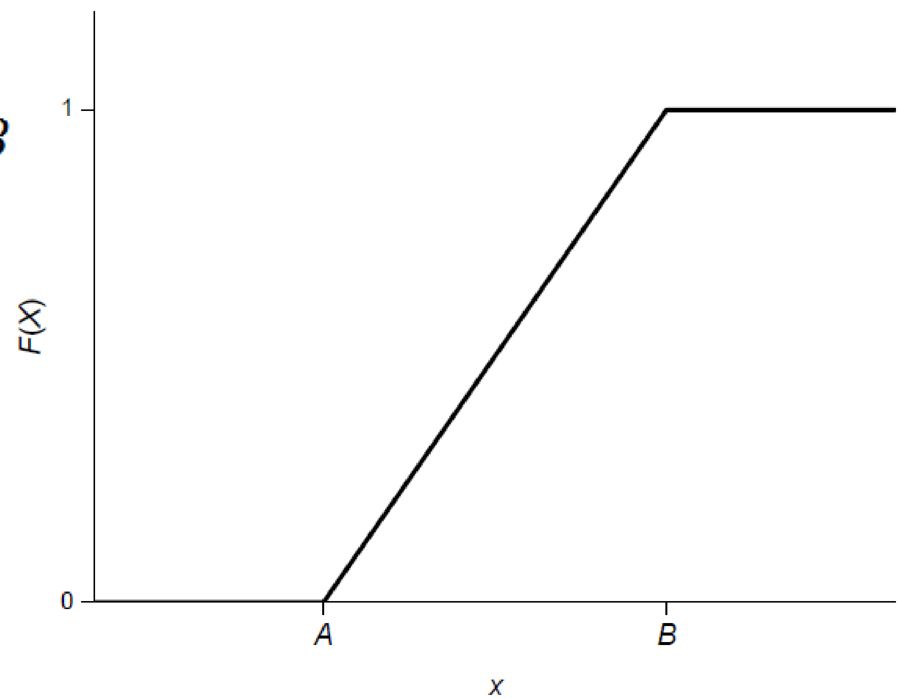
$$f(x) = \begin{cases} \dfrac{1}{B-A} & A \leq x \leq B \\ 0 & inače \end{cases}$$



Funkcija raspodjele

- Primjer: ***uniformna ili pravokutna raspodjela*** u intervalu [A,B]

$$F(x) = \begin{cases} 0, & x < A \\ \dfrac{x - A}{B - A}, & A \leq x \leq B \\ 1, & x > B \end{cases}$$



Očekivanje i varijanca kontinuirane slučajne varijable

Očekivanje kontinuirane slučajne varijable ***X*** definira se:

$$\boxed{\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx}$$

Varijanca kontinuirane slučajne varijable ***X*** je:

$$\boxed{V(X) = \sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx = E[(X-\mu)^2]}$$

Može se pokazati da i ovdje vrijedi:

$$V(X) = E(X^2) - [E(X)]^2$$

Očekivanje i varijanca kontinuirane slučajne varijable

- Primjer: ***uniformna ili pravokutna raspodjela*** u intervalu [A,B]

$$f(x) = \begin{cases} \dfrac{1}{B-A} & A \leq x \leq B \\ 0 & inače \end{cases}$$

- Očekivanje: $E(X) = \dfrac{A+B}{2}$

- Varijanca: $V(X) = \dfrac{(A-B)^2}{12}$

Funkcija izvodnica i momenti

Definicije:

- funkcija izvodnica:

$$g(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

- pomoćni moment:

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

- centralni moment:

$$M_r = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx$$

Očekivanje i varijanca *funkcije* kontinuirane slučajne varijable

- Za ***funkciju h(X)*** kontinuirane slučajne varijable ***X*** s gustoćom vjerojatnosti ***f(x)***, očekivanje je:

$$E\big(h(X)\big) = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x) f(x) dx$$

- Za ***funkciju h(X)*** kontinuirane slučajne varijable ***X*** s gustoćom vjerojatnosti ***f(x)***, varijanca je:

$$V\big(h(X)\big) = E[h(X)^2] - \big[E\big(h(X)\big)\big]^2$$

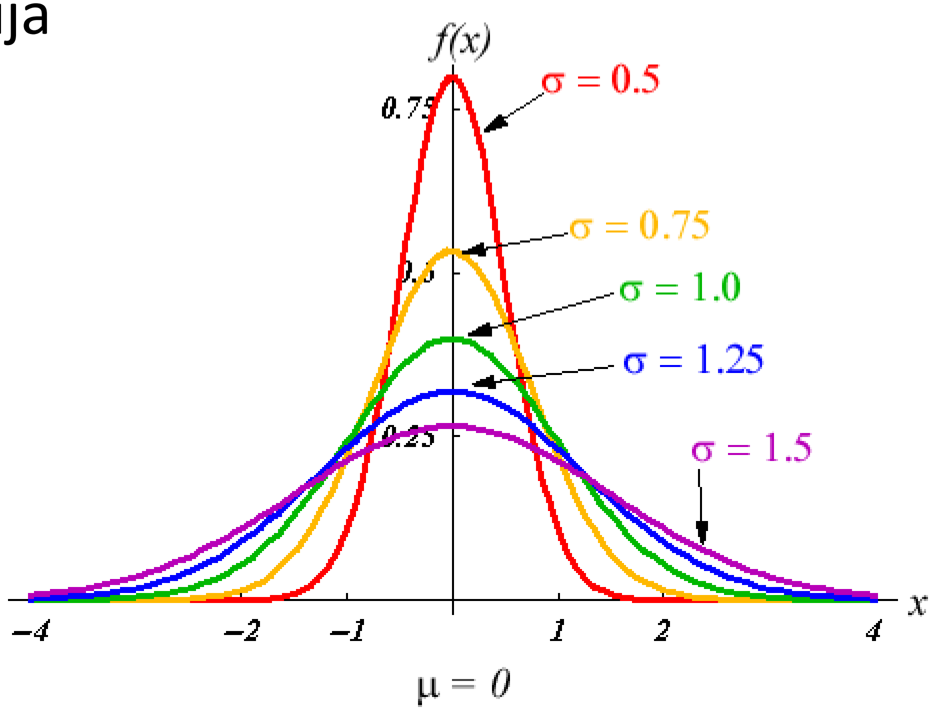
- <u>Specijalno</u>: linearna funkcija h(X)=aX+b
  - Očekivanje: **E(aX+b)=a E(X)+b**
  - Varijanca: **V(aX+b)=a^2 V(X)**

Normalna (Gaussova) raspodjela

- Slučajna varijabla ***X*** ima normalnu raspodjelu ako je njena funkcija gustoće vjerojatnosti:

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

gdje x može biti iz $(-\infty, \infty)$



- Slučajnu varijablu koja ima normalnu raspodjelu označavamo sa: $X \sim N(\mu, \sigma^2)$

Normalna (Gaussova) raspodjela

Normiranost raspodjele slijedi iz definicije:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

uz supstituciju: $u = x - \mu$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{u^2}{2\sigma^2}} du = 2\frac{1}{\sigma\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{u^2}{2\sigma^2}} du = \frac{2}{\sigma\sqrt{2\pi}} \frac{\sqrt{\pi}}{\frac{2}{\sqrt{2}\sigma}} = 1$$

Slijedi: $\boxed{\int_{-\infty}^{\infty} f(x) dx = 1}$

Iz tablica:

$$\int_{0}^{\infty} e^{-a^2x^2} dx = \frac{\sqrt{\pi}}{2a}, a > 0$$

Normalna (Gaussova) raspodjela

Očekivanje – prema definiciji:

$$E(X) = \int_{-\infty}^{\infty} x\, f(x) dx = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

- Uz supstituciju: $u = \dfrac{x-\mu}{\sigma}$
- Slijedi: $\boxed{E(X) = \mu}$

Varijanca – prema definiciji se može pokazati da slijedi:

$$\boxed{V(X) = \sigma^2}$$

Normalna (Gaussova) raspodjela

Standardizirana raspodjela

- Uvodimo substituciju: $u = \frac{x-\mu}{\sigma}$
- Funkcija gustoće vjerojatnosti standardne normalne raspodjele: $$\boxed{\varphi(u) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}}$$

→ Normalna raspodjela s očekivanjem 0 i varijancom 1 ***U~N(0,1)***

- Veza s "običnom" Gaussovom raspodjelom: $$f(x) = \frac{1}{\sigma}\,\varphi(u)$$
- Kumulativna funkcija standardizirane normalne raspodjele: $$\Phi(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$$ → Vrijednosti dane u tablicama!

Normalna (Gaussova) raspodjela

Vjerojatnost da x poprimi vrijednost u intervalu (x_1, x_2) je po definiciji:

$$P(x_1 < x < x_2) = \frac{1}{\sqrt{2\pi}\sigma}\int_{x_1}^{x_2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$$

- Prelaskom na standardiziranu raspodjelu, tj. uvođenjem supstitucije $u = \dfrac{x-\mu}{\sigma}$ vjerojatnost postaje:

$$P(x_1 < x < x_2) = \frac{1}{\sqrt{2\pi}}\int_{u_1}^{u_2} e^{-\frac{u^2}{2}}du$$

- Koristimo distribucijsku funkciju:

$$\boxed{P(x_1 < x < x_2) = \Phi(u_2) - \Phi(u_1)}$$

Normalna (Gaussova) raspodjela

Vjerojatnost – primjer

- Uzmimo normalnu raspodjelu μ=50 i σ=2
- Kolika je vjerojatnost da ***X*** poprimi vrijednost između 46 i 53

$$P(46 < X < 53) = ?$$

- Uz substituciju $u = \dfrac{x-\mu}{\sigma}$:

$$\begin{aligned} P(46 < X < 53) &= P(-2 < U < 1{,}5) \\ &= \Phi(1{,}5) - \Phi(-2) \end{aligned}$$

- Tablice: http://www.stat.ufl.edu/~athienit/Tables/Ztable.pdf

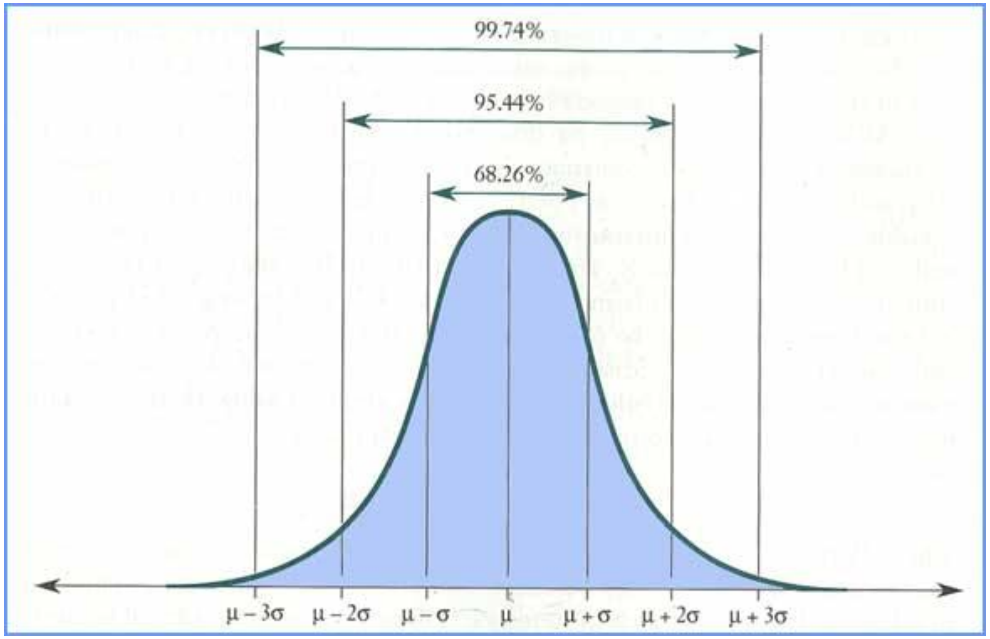
→ P = 0,9332-0,0228=0,9104

Normalna (Gaussova) raspodjela

- Vjerojatnost da X poprimi vrijednost:

Interval	P(x)
$(\mu-\sigma, \mu+\sigma)$	0,683
$(\mu-2\sigma, \mu+2\sigma)$	0,954
$(\mu-3\sigma, \mu+3\sigma)$	0,997

- Simetričnost α_3=0
- Spljoštenost α_4=3



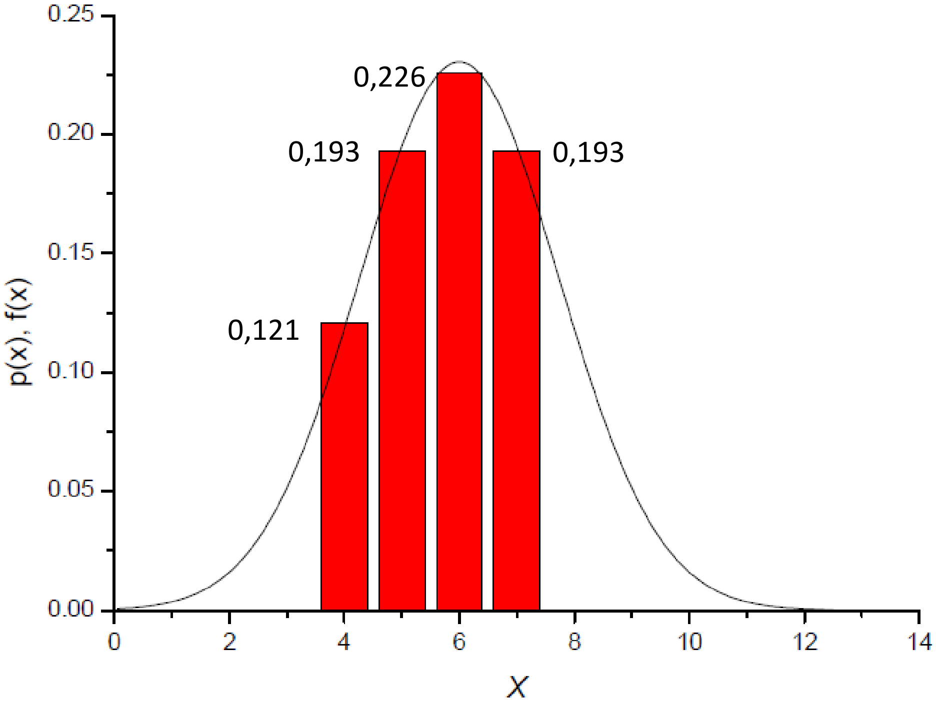
Značaj normalne raspodjele

1. Objašnjava najveći broj statističih opažanja
2. Opisuje raspodjelu slučajne pogreške mjerenja
3. Općenito opisuje pojave gdje postoji velik broj slučajnih utjecaja → pokazuje sa da je ukupna raspodjela veličine podvrgnute tim utjecajima, Gaussova raspodjela

Aproksimacija binomne raspodjele Gaussovom

Primjer: koja je vjerojatnost da u 12 bacanja novčića ispadne najmanje 4 a najviše 7 glava?

- Binomna raspodjela:
  - n=12
  - p=1/2
  - → $P(4 \leq X \leq 7)$ = 0,733
- Aproksimacija Gaussovom:
  - 1/13 < 1/2 < 12/13
  - npq=3<9
  - μ = 6, σ^2 = 3
  - → $P(3.5 < X < 7.5)$ = 0,733



Aproksimacija binomne raspodjele Gaussovom

- Pokazuje se da kod binomne raspodjele:
  - koeficijent asimetrije $\alpha_3 \rightarrow 0$ kad $n \rightarrow$ oo $\alpha_3 = \frac{q-p}{\sqrt{npq}}$
  - koeficijent spljoštenosti $\alpha_4 \rightarrow 3$ kad $n \rightarrow$ oo $\alpha_4 = \frac{1-6pq}{npq} + 3$

  → to su upravo svojstva normalne raspodjele

  → Binomna teži u normalnu kad n → oo
- U praksi se uglavnom pokazuje da je aproksimacija Gaussovom raspodjelom dobra ako su zadovoljeni uvjeti:

1. $\frac{1}{n+1} < p < \frac{n}{n+1}$

2. $npq > 9$

Aproksimacija Poissonove raspodjele Gaussovom

- Ako X ima Poissonovu raspodjelu $X \sim P(\lambda)$ sa velikim λ ($\lambda>20$) može se dobro aproksimirati Gaussovom $X \sim N(\lambda,\lambda)$

<u>Primjer</u>: Geigerov brojač mjeri raspade sa srednjom brzinom od 25 s^{-1}

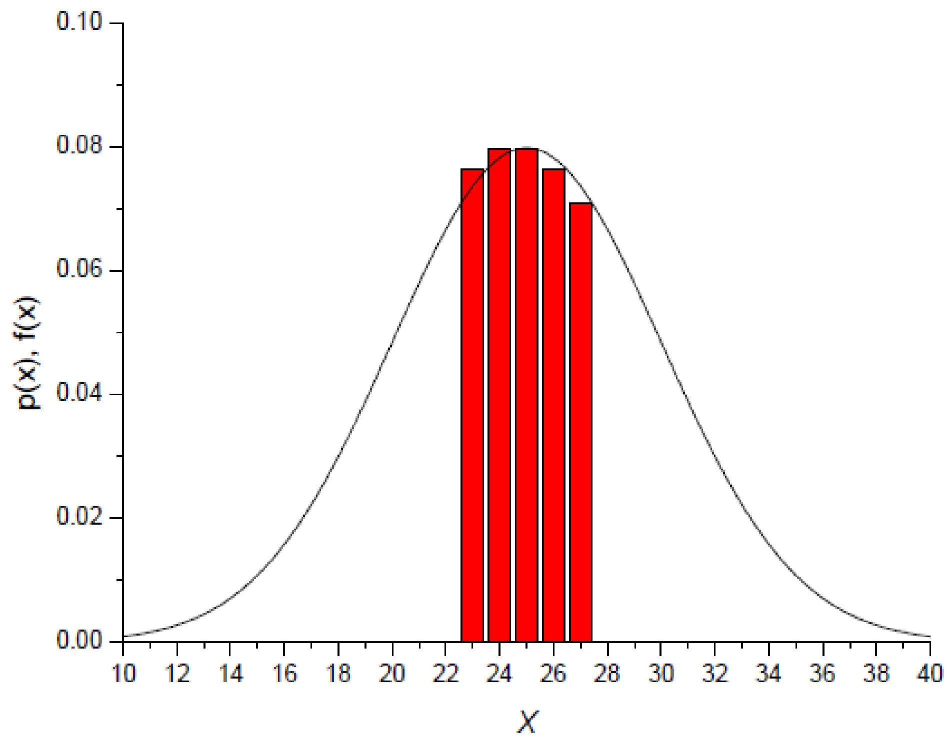
Koja je vjerojatnost da opazimo najmanje 23 i najviše 27 raspada/s ?

Poisson $X \sim P(25)$:

$P(23 \leq X \leq 27)$ = 0,3826

Gauss $X \sim (25,25)$:

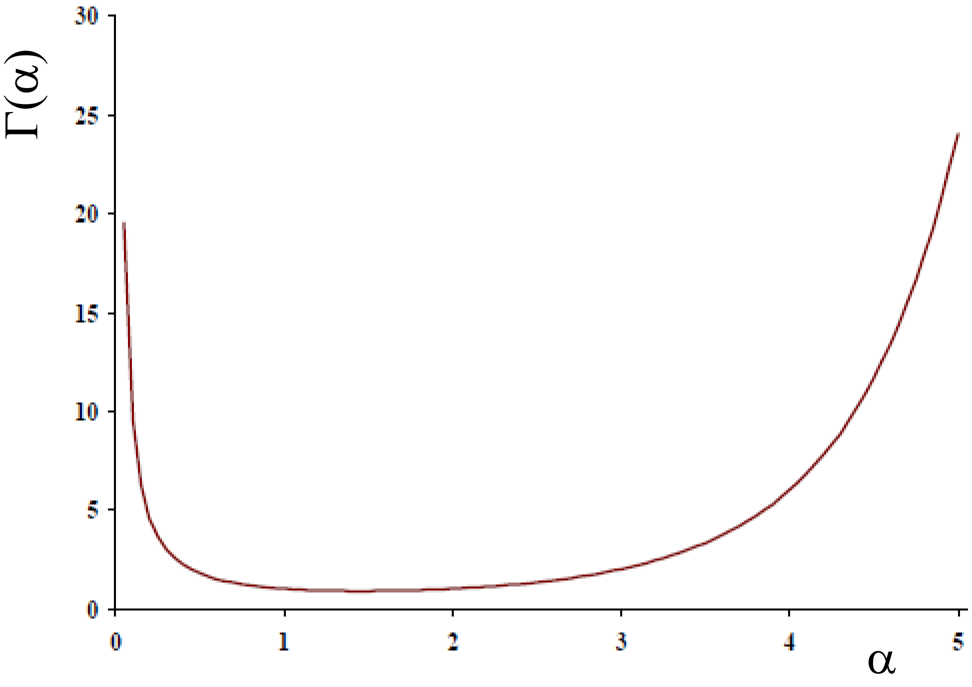
$P(22.5 < X < 27.5)$ = 0,3830



Obitelj gama raspodjela

Definiramo Γ-funkciju: $$\boxed{\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \qquad \alpha > 0}$$

- Svojstva:
  1. $\forall \alpha > 1 \rightarrow \Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ (rekurzija)
  2. $\forall n \in \mathbb{N} \rightarrow \Gamma(n) = (n-1)!$ **(poopćenje faktorijela!)**
  3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

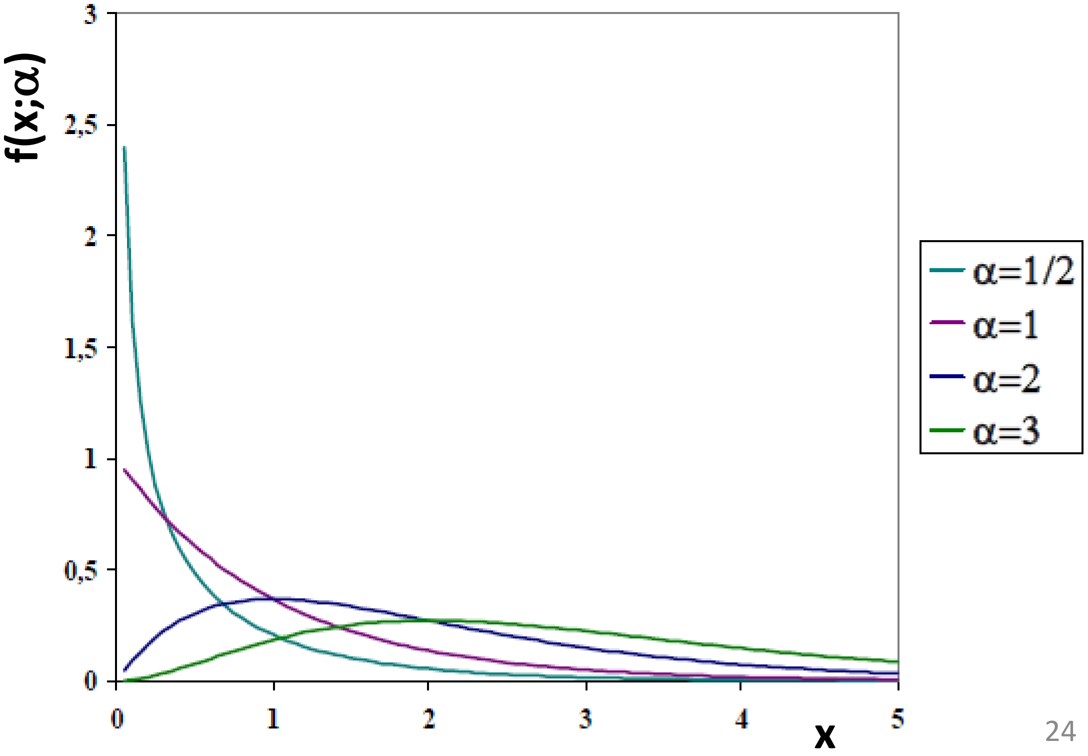


Standardna Γ-raspodjela

- Funkcija gustoće vjerojatnosti Γ raspodjele slučajne varijable ***X*** je:

$$f(x;\alpha)=\boxed{\begin{cases}\dfrac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}, & x\geq 0\\ 0, & x<0\end{cases}}$$

- gdje je parametar raspodjele $\alpha>0$



Standardna Γ-raspodjela

- Svojstva – osnovni uvjeti raspodjele vjerojatnosti:

1. $f(x;\alpha) \geq 0, \qquad \forall x$
2. $\int_{-\infty}^{\infty} f(x;\alpha)dx = \frac{1}{\Gamma(\alpha)} \int_{0}^{\infty} x^{\alpha-1}e^{-x}dx = \frac{\Gamma(\alpha)}{\Gamma(\alpha)} = 1$

- Očekivanje: $E(X)=\alpha$
- Varijanca: $V(X)=\alpha$

Općenita Γ-raspodjela

- Funkcija gustoće vjerojatnosti:

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{x^{\alpha-1} e^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

gdje su α, β >0

- Očekivanje: $E(X)=\alpha\beta$

- Varijanca: $V(X)=\alpha\beta^2$

Za β=1 općenita → standardna Γ-raspodjela