

6.2. Funkcija izvodnica momenata

Glavne poruke

- FI momenata ima slična svojstva kao i FI vjerojatnosti, ali nije ograničena na slučajne varijable s vrijednostima u $\mathbb{N}_0$.
- Ipak, FI momenata **nije** definirana za varijable koje imaju "pretežak rep". Primjerice, čim je $\mathbb{E}[|X|^k] = \infty$ za neki k , X nema FI momenata.
- FI momenata se općenito koriste za dobivanje ograda na vjerojatnosti oblika $\mathbb{P}(X > u)$ za velike $u > 0$, ali time se nećemo baviti.

Def. 6.12 (FI momenata)

Neka je X (bilo kakva) slučajna varijabla. Ako postoji $t_0 > 0$ za koji vrijedi

$$\mathbb{E}[e^{tX}] < \infty, \quad \forall |t| \leq t_0, \quad (6.9)$$

funkcija $M : [-t_0, t_0] \rightarrow [0, \infty)$ definirana sa

$$M(t) := \mathbb{E}[e^{tX}], \quad (6.10)$$

zove se **funkcija izvodnica momenata** slučajne varijable X .

Napomena.

- U nastavku ćemo često pisati $M_X := M$ kako bi naglasili da se radi o FI momenata za slučajnu varijablu X .
- Ako ne vrijedi (6.9), FI momenata nije definirana, te se rezultati koje ćemo izvesti ne mogu koristiti za takve slučajne varijable.

Napomena 6.13 (o uvjetu (6.9))

- Ako je $t > 0$ (odnosno $t < 0$), uvjet $\mathbb{E}[e^{tX}] < \infty$ će vrijediti ako vjerojatnosti $\mathbb{P}(X > u)$ (odnosno $\mathbb{P}(X < -u)$) **nisu prevelike** za velike $u > 0$.
- Gornji uvjet je ekvivalentan je uvjetu

$$\mathbb{E}[e^{t_0|X|}] < \infty. \quad (6.11)$$

- To slijedi jer za sve $|t| \leq t_0$ vrijedi

$$e^{tX} \leq e^{|tX|} \leq e^{t_0|X|} \leq \max\{e^{t_0X}, e^{-t_0X}\} \leq e^{t_0X} + e^{-t_0X},$$

pa specijalno i

$$\mathbb{E}[e^{tX}] \leq \mathbb{E}[e^{t_0|X|}] \leq \mathbb{E}[e^{t_0X}] + \mathbb{E}[e^{-t_0X}].$$

Teorem 6.14 (Svojstva FI momenata)

Pretpostavimo da postoji $t_0 > 0$ takav da vrijedi (6.9).

- (a) Za sve $k \in \mathbb{N}_0$ je $\mathbb{E}[|X|^k] < \infty$, pa specijalno postoji **k -ti moment od X** definiran s $\mu_k := \mathbb{E}[X^k]$.
- (b) Za sve $|t| \leq t_0$ vrijedi¹

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k.$$

- (c) Za sve $k \in \mathbb{N}_0$ vrijedi $M^{(k)}(0) = \mu_k$.

¹Zato se M_X zove funkcija izvodnica **momenata**.

Dokaz.

Dokaz je baziran na jednakosti $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, $x \in \mathbb{R}$.

$$(a) \left[\mathbb{E}[|X|^k] < \infty, \forall k \right]$$

$$e^{t_0|X|} = \sum_{n=0}^{\infty} \frac{t_0^n |X|^n}{n!} \geq \frac{t_0^k}{k!} |X|^k, \quad \forall k \in \mathbb{N}_0. \quad (6.12)$$

Monotonost očekivanja povlači da je

$$\frac{t_0^k}{k!} \mathbb{E}[|X|^k] \leq \mathbb{E}[e^{t_0|X|}] \stackrel{(6.11)}{<} \infty, \quad \forall k \in \mathbb{N}_0.$$

Tvrđnja sada slijedi budući da je $t_0 > 0$.

Dokaz.

$$(b) \left[M_X(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k, \forall |t| \leq t_0 \right]$$

Za $|t| \leq t_0$,

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] = \mathbb{E} \left[\sum_{n=0}^{\infty} \frac{t^n X^n}{n!} \right] \\ &= \sum_{n=0}^{\infty} \mathbb{E} \left[\frac{t^n X^n}{n!} \right] = \sum_{n=0}^{\infty} \frac{\mu_n}{n!} t^n. \end{aligned}$$

Napomena.* Zamjena $\mathbb{E}$ i $\sum_{n=0}^{\infty}$ u trećoj jednakosti opravdana je verzijom tzv. Fubinijevog teorema (vidi bilješke). Dovoljan uvjet ovdje je zadovoljen samo za $|t| \leq t_0$ jer je samo tada

$$\mathbb{E} \left[\sum_{n=0}^{\infty} \frac{|t|^n |X|^n}{n!} \right] = \mathbb{E}[e^{|t| \cdot |X|}] \leq \mathbb{E}[e^{t_0 \cdot |X|}] \stackrel{(6.11)}{<} \infty$$

Dokaz.

$$(c) \left[M^{(k)}(0) = \mu_k \right]$$

DZ – slijedi jer

$$M_X(t) = \sum_{k=0}^{\infty} \frac{\mu_k}{k!} t^k$$

deriviramo "član po član".

Primjer 6.15 ($\text{Exp}(\lambda)$ razdioba)

- Za $X \sim \text{Exp}(\lambda)$, tj. $f_X(t) = \lambda e^{-\lambda t}$, $t > 0$, vrijedi

$$M_X(t) = [\text{raspis na ploči}] = \frac{\lambda}{\lambda - t}, \quad t < \lambda := t_0.$$

[Formalno bi trebali uzeti $t_0 := \lambda - \epsilon > 0$ neki mali ϵ tako da vrijedi (6.9).]

- Specijalno, iz $M'_X(t) = \frac{\lambda}{(\lambda - t)^2}$ slijedi da je

$$\mathbb{E}[X] = M'_X(0) = \frac{1}{\lambda}.$$

Daljnjim deriviranjem možemo dobiti i $\mathbb{E}[X^k]$ za $k = 2, 3, \dots$, ali pogledajmo puno elegantniji pristup.

Primjer 6.15 ($\text{Exp}(\lambda)$ razdioba – momenti)

- Ideja: Znamo da je $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n]$, za $t < t_0$, pa ako nađemo niz $(a_n)_{n \geq 0}$ za koji je $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} a_n$ (za t -ove na proizvoljno maloj okolini nule), nužno je $\mathbb{E}[X^n] = a_n, \forall n \geq 0$.
- Ako je $E \sim \text{Exp}(1)$ i $|t| < 1$,

$$\begin{aligned} M_E(t) &= [t < 1 = \lambda] = \frac{\lambda}{\lambda - t} \Big|_{\lambda=1} = \frac{1}{1-t} \\ &= [|t| < 1] = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot n!. \end{aligned}$$

$\rightsquigarrow \mathbb{E}[E^n] = n!$, za sve $n \geq 0$.

- Općenito, budući da je $X := \frac{E}{\lambda} \sim \text{Exp}(\lambda)$,

$$\mathbb{E}[X^n] = \frac{\mathbb{E}[E^n]}{\lambda^n} = \frac{n!}{\lambda^n}, \quad \forall n \in \mathbb{N}_0.$$

Primjer 6.16 ($N(\mu, \sigma^2)$)

Za $Z \sim N(0, 1)$ imamo

$$\begin{aligned} M_Z(t) &= \mathbb{E}[e^{tZ}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \cdot e^{tz} dz \\ &= [\pm \frac{t^2}{2}] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2tz + t^2) + \frac{t^2}{2}} dz \\ &= e^{\frac{t^2}{2}} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}} dz}_{= f_{N(t,1)}(z)} = e^{\frac{t^2}{2}}, \quad \forall t \in \mathbb{R}. \end{aligned}$$

to jest

$$M_{N(0,1)}(t) = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R} \tag{6.15}$$

Primjer 6.16 ($N(\mu, \sigma^2)$)

Za $X \sim N(\mu, \sigma^2)$,

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] = [X \sim \sigma Z + \mu] = \mathbb{E}[e^{t\sigma Z + t\mu}] = e^{t\mu} \mathbb{E}[e^{t\sigma Z}] \\ &= e^{t\mu} M_Z(t\sigma) = [M_Z(t) = e^{\frac{t^2}{2}}] = e^{\mu t + \frac{\sigma^2}{2} t^2}, \quad \forall t \in \mathbb{R}. \end{aligned} \quad (6.16)$$

Primjer 6.17 (FI ne postoji uvijek!)

- Ako za slučajnu varijablu X postoji $k \geq 0$ t.d. je $\mathbb{E}[|X|^k] = \infty$, tj. X **nema** k -ti moment, po prethodnom teoremu ne postoji $t_0 > 0$ t.d. vrijedi (6.9), tj. **ne postoji** M_X .
- Intuitivno, to su varijable koje imaju "**pretežak rep**" u smislu da su vjerojatnosti $\mathbb{P}(|X| > u)$ (tj. $\mathbb{P}(X > u)$ ili $\mathbb{P}(X < -u)$) za velike $u > 0$ **relativno velike**.
- npr. to je slučaj za $X \sim \text{Pareto}(\alpha)$ za proizvoljan $\alpha > 0$ ($X \geq 1$ te $\mathbb{P}(X > u) = u^{-\alpha}$, $u \geq 1$) – u Primjeru 5.28 smo pokazali da je $\mathbb{E}[X^\alpha] = \infty$, pa je i $\mathbb{E}[X^k] = \infty$ za sve $k \geq \alpha$.

NAPOMENA*.

- Ako za X ne postoji M_X , možemo koristiti tzv. **karakterističnu funkciju** $t \mapsto \mathbb{E}[e^{itX}] \in \mathbb{C}$ koja **uvijek** postoji.
- Karakteristična funkcija ima slična svojstva kao i FI momenata.
- Ipak, ako postoji M_X možemo dobiti eksponencijalne ocjene na vjerojatnosti oblika $\mathbb{P}(X > a)$; vidi vježbe (Chernoffova ograda), te tzv. **teoriju velikih devijacija** (engl. *large deviations*).

Ključna svojstva (isto kao kod FI vjerojatnosti)

Tm. 6.18.(DZ) Ako su slučajne varijable X i Y **nezavisne** te postoji $t_0 > 0$ tako da su M_X i M_Y dobro definirane na $[-t_0, t_0]$, M_{X+Y} je također dobro definirana na $[-t_0, t_0]$ te vrijedi

$$M_{X+Y}(t) = M_X(t)M_Y(t), \quad |t| \leq t_0. \quad (6.17)$$

Nap. 6.19 (FI momenata jedinstveno određuje distribuciju) Može se pokazati da ako vrijedi $M_X = M_Y$ na nekom otvorenom intervalu oko 0, tada slučajne varijable X i Y imaju istu distribuciju (bez dokaza).

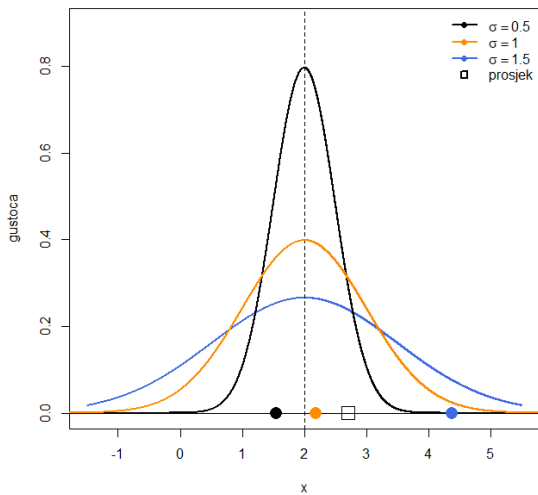
Pr. 6.20 (zbroy nezavisnih normalnih je opet normalna)(DZ) Koristeći (6.17) i prethodnu napomenu lako se pokaže da za $X \sim N(\mu_1, \sigma_1^2)$ i $Y \sim N(\mu_2, \sigma_2^2)$ **nezavisne** vrijedi

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Pr. 6.21 (Crowdsourcing)

- Pretpostavimo da su $X_i \sim N(\mu, \sigma_i^2)$, $i = 1, \dots, n$ nezavisne, za neko zajedničko očekivanje $\mu \in \mathbb{R}$, te (moguće različite) varijance $\sigma_1, \dots, \sigma_n > 0$.
- **Cilj:** na temelju realizacija $X_1 = x_1, \dots, X_n = x_n$, želimo procijeniti μ .
- **Motivacija:** Imamo proizvod čiju (nepoznatu) kvalitetu želimo procijeniti. Parametar μ predstavlja kvalitetu proizvoda, a X_i predstavlja ocjenu koji i -ti korisnik daje tom proizvodu – veći σ_i odgovara korisniku s manjom ekspertizom.

Pr. 6.21 (Crowdsourcing, $n = 3$)



Pr. 6.21 (Crowdsourcing)

- Osnovni procjenitelj koji možemo koristiti je $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, a njegovu "grešku" možemo definirati kao

$$\text{err}(\overline{X}_n) := \mathbb{E}[|\overline{X}_n - \mu|] .$$

- Po prethodnom primjeru znamo da $S_n := \sum_{i=1}^n X_i$ ima normalnu razdiobu, pa nadalje slijedi da i $\overline{X}_n - \mu = \frac{1}{n} \cdot S_n + \mu$ ima normalnu razdiobu, i to s parametrima

$$\mathbb{E}[\overline{X}_n - \mu] = [\mathbb{E}[X_i] = \mu, \forall i] = \frac{1}{n} \cdot n\mu - \mu = 0 ,$$

$$\text{Var}(\overline{X}_n - \mu) = \text{Var}(\overline{X}_n) = [\text{nezavisnost}] = \frac{\sum_{i=1}^n \sigma_i^2}{n^2} =: c_n^2 .$$

- Specijalno, ako je $Z \sim N(0, 1)$, imamo $c_n |Z| \sim |\overline{X}_n - \mu|$, pa odmah slijedi

$$\text{err}(\overline{X}_n) = c_n \mathbb{E}[|Z|] = \sqrt{\frac{2}{\pi}} \cdot \frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n}$$

Pr. 6.21 (Crowdsourcing)

$$\text{err}(\bar{X}_n) = \sqrt{\frac{2}{\pi}} \cdot \frac{\sqrt{\sum_{i=1}^n \sigma_i^2}}{n}$$

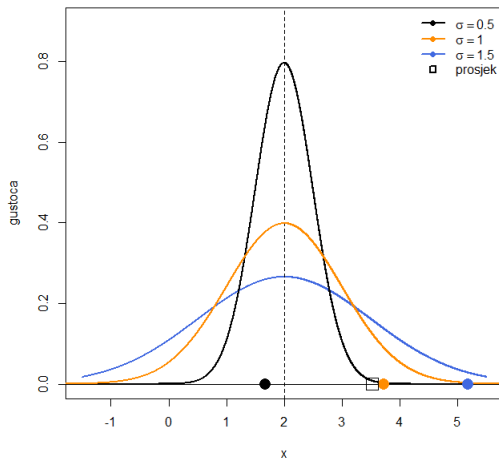
- Ako je $\sigma_i = \sigma > 0$, $\forall i$, onda je $\text{err}(\bar{X}_n) = \sqrt{\frac{2}{\pi}} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow 0$, kada $n \rightarrow \infty$.
- Na primjer, ako dopustimo da varijance σ_i ovise o n , tj. $\sigma_i = \sigma_i^{(n)}$, te ako za neki $\epsilon > 0$ vrijedi

$$\max_{i=1, \dots, n} \sigma_i^{(n)} = n^{1+\epsilon}, \quad \forall n \geq 1,$$

tj. **barem jedna varijanca je dovoljno velika** imamo

$$\text{err}(\bar{X}_n) \geq \sqrt{\frac{2}{\pi}} \frac{n^{1+\epsilon}}{n} = \sqrt{\frac{2}{\pi}} n^{\epsilon} \longrightarrow +\infty, \quad n \rightarrow \infty.$$

Pr. 6.21 (Crowdsourcing)



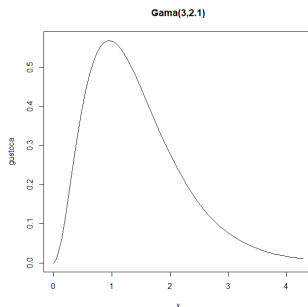
Za alternativan procjenitelj kada su σ_i poznate vidjeti vježbe. Što ako su σ_i nepoznate?

Primjer 6.22 (Gama razdioba)

Kažemo da slučajna varijabla X ima **gama** razdiobu s parametrima $\alpha, \lambda > 0$ (oznaka $X \sim \text{Gama}(\alpha, \lambda)$) ako je neprekidna s gustoćom

$$f(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \quad t > 0,$$

uz $f(t) = 0$ za $t \leq 0$, gdje je $\Gamma(\alpha) := \int_0^\infty y^{\alpha-1} e^{-y} dy$ tzv. gama funkcija (vrijedi $\Gamma(n) = (n-1)!$ za sve $n \in \mathbb{N}$).



Primjer 6.22 (Gama razdioba) – svojstva

$$f_{\text{Gama}(\alpha, \lambda)}(t) = c \cdot t^{\alpha-1} e^{-\lambda t}, \quad t > 0$$

- očito je $\text{Gama}(1, \lambda) = \text{Exp}(\lambda)$;
- $X \sim \text{Gama}(\alpha, 1)$ povlači da je $\frac{X}{\lambda} \sim \text{Gama}(\alpha, \lambda)$, za sve $\lambda > 0$ (DZ);
- $X \sim \text{Gama}(\alpha, \lambda)$ povlači $M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha$, za $t < \lambda$ (DZ).
- koristi se u Bayesovskoj statistici kao tzv. apriorna distribucija.

Prop. 6.23 (zbroj nezavisnih $\text{Exp}(\lambda)$)

Ako su $X_1, \dots, X_n$ njd takve da je $X_i \sim \text{Exp}(\lambda)$, vrijedi

$$T_n := X_1 + \dots + X_n \sim \text{Gama}(n, \lambda).$$

[U Poissonovom procesu s intenzitetom $\lambda > 0$, gornji T_n je upravo vrijeme kada se dogodio n -ti događaj.]

Dokaz.

$$\begin{aligned} M_{T_n}(t) &= [X_1, \dots, X_n \text{ njd}] = (M_{X_1}(t))^n = \left(\frac{\lambda}{\lambda - t} \right)^n \\ &= M_{\text{Gama}(n, \lambda)}(t), \quad t < \lambda. \end{aligned}$$