

DISKRETNE SLUČAJNE PROMENLJIVE

Zakon raspodele

- **Diskretne** slučajne promenljive su one koje imaju **konačan ili prebrojiv** skup vrednosti.
- Diskretne slučajne promenljive su određene svojim vrednostima, i njihovim verovatnoćama.
- Te informacije su sadržane u **zakonu raspodele**.

$$X: \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p(x_1) & p(x_2) & \dots & p(x_n) \end{pmatrix} \begin{matrix} \leftarrow \text{vrednosti} \\ \leftarrow \text{verovatnoće} \end{matrix}$$

- Šematski prikaz zakona raspodele slučajne promenljive X se sastoji od dve vrste:
 1. Prva vrsta sadrži sve vrednosti iz skupa vrednosti R_X
 2. Druga vrsta sadrži verovatnoću one vrednosti ispod koje je napisana.

Zakon raspodele

U svakom šematskom prikazu zakona raspodele važe sledeća pravila:

1. Vrednosti iz skupa vrednosti R_X se navode u rastućem poretku (tj. od najmanje do najveće, bez ponavljanja)

$$x_1 < x_2 < \dots < x_n < \dots$$

2. Verovatnoće svih vrednosti su pozitivne ($p(x_i) > 0$) i u zbiru iznose 1 (zato što su obuhvaćeni svi mogući ishodi verovatnosnog eksperimenta)

$$p(x_1) + p(x_2) + \dots + p(x_n) + \dots = 1$$

Zakon raspodele – primer

Primer: Verovatnosni eksperiment može da ima ishode 0, 2, 1 i -1. Poznato je da se 0 javlja u 50% slučajeva, 2 u 30% slučajeva, i da se vrednosti 1 i -1 javljaju jednakoverovatno. Napisati zakon raspodele slučajne promenljive X koja predstavlja ishode ovog eksperimenta.

- Sortirane vrednosti: -1, 0, 1, 2.
- Verovatnoće: $P(X=0)=0.5$, $P(X=2)=0.3$, $P(X=1)=P(X=-1)=k$.
- Iz osobine zakona raspodele $p(-1)+p(0)+p(1)+p(2) = 1$ izračunavamo $0.5+0.3+k+k=1 \Rightarrow k=0.1$
- Zakon raspodele:

$$X: \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.5 & 0.1 & 0.3 \end{pmatrix}$$

Određivanje verovatnoća

- Zakon raspodele slučajne promenljive X sadrži **potpunu informaciju** o X , tj. o verovatnosnom eksperimentu koji je opisan putem X .
- **Verovatnoću događaja A** nad skupom elementarnih ishoda $\Omega = R_X$ možemo pročitati iz zakona raspodele.
- Postupak: u zakonu raspodele identifikujemo sve vrednosti x_i koje su sadržane u događaju A , i saberemo njihove verovatnoće:

$$P(A) = \sum_{x_i \in A} p(x_i)$$

Određivanje verovatnoća – primer

Primer: Iz eksperimenta iz prethodnog primera izračunati verovatnoće sledećih događaja: A – ishod je pozitivan broj, B – ishod je manji od $4/3$, C – ishod je neparan broj, D – ishod nije manji od 3, E – ishod je broj koji udvostručen i potom uvećan za 3 pripada intervalu $[3, 6]$.

Zakon raspodele $X: \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.5 & 0.1 & 0.3 \end{pmatrix}$

- $P(A) = P(X > 0) = p(1) + p(2) = 0.1 + 0.3 = 0.4$
- $P(B) = P(X < 4/3) = p(-1) + p(0) + p(1) = 0.7$
- $P(C) = P(X \text{ je neparan}) = p(-1) + p(1) = 0.2$
- $P(D) = P(X \geq 3) = 0$
- $P(E) = P(3 \leq 2X + 3 \leq 6) = P(0 \leq 2X \leq 3) = P(0 \leq X \leq 3/2) = p(0) + p(1) = 0.6$

Nastaviće se...

Funkcija raspodele

- Pojam slučajne promenljive X omogućava prelazak sa skupa elementarnih događaja Ω na skup realnih brojeva $\mathbb{R}$ koji ima bogatiju strukturu.
- Zbog toga uvodimo **funkciju raspodele** $F_X(x)$ slučajne promenljive X

$$F_X(x) = P(X \leq x), \quad x \in \mathbb{R}$$

- Funkcija raspodele je definisana preko verovatnoće, što znači da je možemo pročitati iz zakona raspodele.

Funkcija raspodele - osobine

- Funkcija raspodele $F_X(x)$ slučajne promenljive X zadovoljava sledeće osobine:

1. $F_X(x)$ je neopadajuća funkcija
2. $F_X(-\infty) = 0, F_X(\infty) = 1$
3. $P(a \leq X < b) = F_X(b) - F_X(a)$
4. $P(X \geq a) = 1 - F_X(a)$
5. $F_X(x)$ je neprekidna sa leve strane

- dokazi svih osobina slede iz definicije $F_X(x) = P(X < x)$

Funkcija raspodele – primer

Primer: Slučajna promenljiva X je data zakonom raspodele

$$X: \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.5 & 0.1 & 0.3 \end{pmatrix}$$

Odrediti $F_X(0)$, $F_X(0.2)$, $F_X(\sqrt{2})$ i napisati funkciju $F_X(x)$.

- $F_X(0) = P(X < 0) = p(-1) = 0.1$
- $F_X(0.2) = P(X < 0.2) = p(-1) + p(0) = 0.6$
- $F_X(\sqrt{2}) = P(X < \sqrt{2}) = p(-1) + p(0) + p(1) = 0.7$

$$F_X(x) = \begin{cases} 0, & x \leq -1 \\ 0.1, & -1 < x \leq 0 \\ 0.6, & 0 < x \leq 1 \\ 0.7, & 1 < x \leq 2 \\ 1, & x > 2 \end{cases}$$

Numeričke karakteristike

- Slučajnu promenljivu X opisuju i određene numeričke karakteristike
- Najznačajnije među njima su **matematičko očekivanje $E(X)$** i **disperzija $D(X)$** , ukoliko postoje.

$$E(X) = \sum_i x_i p(x_i)$$

$$D(X) = E(X^2) - E^2(X)$$

- Više toga o numeričkim karakteristikama, njihovom značenju i osobinama ćemo raditi kasnije.

Numeričke karakteristike– primer

Primer: Slučajna promenljiva X je data zakonom raspodele

$$X: \begin{pmatrix} -1 & 0 & 1 & 2 \\ 0.1 & 0.5 & 0.1 & 0.3 \end{pmatrix}$$

Odrediti matematičko očekivanje i disperziju sl. pr. X .

- matematičko očekivanje:

$$E(X) = (-1)*0.1 + 0*0.5 + 1*0.1 + 2*0.3 = 0.6$$

- Disperzija:

$$E(X^2) = (-1)^2*0.1 + 0^2*0.5 + 1^2*0.1 + 2^2*0.3 = 1.4$$

$$D(X) = 1.4 - 0.6^2 = 1.04$$

NEKE ZNAČAJNE DISKRETNE SLUČAJNE PROMENLJIVE

Raspodele „sa imenom“

- Uočeno je da postoje klase verovatnosnih eksperimenata koje se mogu opisati sličnim (tipskim) slučajnim promenljivama
- Takve slučajne promenljive su dobile ime i posebno su proučavane
- U njima figurišu parametri, čije vrednosti određujemo iz konkretnih primera
- Kada odredimo ime i parametre, određen je i zakon raspodele (a samim tim i sve ostalo...) u vezi te slučajne promenljive
- Diskretne raspodele „sa imenom“: geometrijska, binomna, Poasonova,...

Geometrijska raspodela $G(p)$

Kada se javlja?

- Posmatrajmo verovatnosni eksperiment koji ima samo dva moguća ishoda: **uspeh** i **neuspeh**.
- Verovatnoća uspeha je **p**, a verovatnoća neuspeha **q** ($q=1-p$).
- Eksperiment ponavljamo do prve pojave uspeha (tj. „dok ne uspe“) i potom prestajemo.

Slučajna promenljiva **X** koja **predstavlja broj ponavljanja** ovakvog eksperimenta ima geometrijsku $G(p)$ raspodelu.

- Parametar p , $0 < p < 1$, odgovara verovatnoći uspeha!
- Primeri: bacanje kockice dok ne padne 6; gađanje mete dok ne bude pogodoena; kontrola proizvoda do nailaska na prvi neispravan...

Geometrijska raspodela $G(p)$

Kako izgleda?

- Da bi odredili skup vrednosti, pitamo se koliko puta možemo da ponavljamo eksperiment do prvog uspeha? Najmanje jednom, a najviše nije unapred određeno...

$$\mathbf{R}_X = \mathbb{N} = \{1, 2, \dots, k, \dots\}$$

- Verovatnoća $p(k)$ znači da se prvi uspeh desio u k -tom ponavljanju, i da mu je prethodilo $k-1$ neuspeha

$$p(k) = q^{k-1}p, \quad k \in \mathbf{R}_X$$

- Može se pokazati da za ovakav zakon raspodele, matematičko očekivanje i disperzija kod geometrijske raspodele iznose:

$$\mathbf{E}(X) = \frac{1}{p} \quad \mathbf{D}(X) = \frac{q}{p^2}$$

Geometrijska raspodela - primer

Primer: Bacamo kockicu za igru dok ne padne šestica. Odrediti raspodelu sl. pr. X koja predstavlja broj bacanja. Izračunati verovatnoće (a) da će šestica pasti u trećem bacanju; (b) da će biti potrebno više od tri bacanja.

- Raspodela je geometrijska, verovatnoća uspeha pri jednom bacanju je $1/6$, dakle $p=1/6$, $q=5/6$, i X : **$G(1/6)$** .

a) $P(X=3) = p(3) = (5/6)^2(1/6) = 0.116$

b) $P(X>3) = p(4) + p(5) + \dots$

ovaj beskonačni zbir možemo izračunati pomoću geometrijskog reda, ili možemo iskoristiti suprotni događaj:

$$P(X>3) = 1 - P(X \leq 3) = 1 - (p(1) + p(2) + p(3)) = 1 - ((5/6)^0(1/6) + (5/6)^1(1/6) + (5/6)^2(1/6)) = 1 - 0.421 = 0.579$$

Binomna raspodela $\mathcal{B}(n,p)$

Kada se javlja?

- Posmatrajmo opet verovatnosni eksperiment koji ima samo dva moguća ishoda: uspeh i neuspeh.
- Verovatnoća uspeha je p , a verovatnoća neuspeha q ($q=1-p$).
- **Eksperiment ponavljamo n puta**, i zanima nas koliko je puta (od tih n pokušaja) uspeo.

Slučajna promenljiva **X** koja **predstavlja broj uspeha** (pri n ponavljanja) eksperimenta ima binomnu $\mathcal{B}(n,p)$ raspodelu.

- Parametri: $n, n \in \mathbb{N}$ – broj ponavljanja,
 $p, 0 < p < 1$ - verovatnoća uspeha
- Primeri: koliko puta će pasti šestica pri 10 bacanja, koliko puta će strelac pogoditi metu pri 100 gađanja, koliko će biti neispravnih proizvoda u seriji od 5000 proizvedenih...

Binomna raspodela $\mathcal{B}(n,p)$

Kako izgleda?

- Da bi odredili skup vrednosti, pitamo se koliko puta može da uspe eksperiment koji ponavljamo n puta? Najmanje nijednom, a najviše svih n puta.

$$\mathbf{R}_X = \{0,1,\dots,k,\dots,n\}$$

- Verovatnoća $p(k)$ znači da se uspeh desio nekih k puta, i da se neuspeh desio preostalih $n-k$ puta. Postoji $\binom{n}{k}$ kombinacija za k uspeha u n pokušaja, pa je:

$$p(k) = \binom{n}{k} p^k q^{n-k}, \quad k \in \mathbf{R}_X$$

- Može se pokazati da za ovakav zakon raspodele, matematičko očekivanje i disperzija kod binomne raspodele iznose:

$$\mathbf{E}(X) = np \quad \mathbf{D}(X) = npq$$

Binomna raspodela - primer

Primer: Bacamo kockicu za igru 10 puta. Odrediti raspodelu sl. pr. X koja predstavlja broj palih šestica. Izračunati verovatnoće (a) šestica će pasti 3 puta; (b) šestica će pasti bar tri puta. (c) Koliki je očekivani broj palih šestica?

- Raspodela je binomna, verovatnoća uspeha pri jednom bacanju je $1/6$ a bacanja ima 10, dakle $n=10$, $p=1/6$, $q=5/6$, $\mathbf{X: \mathcal{B}(10, 1/6)}$.

a) $P(X=3) = p(3) = \binom{10}{3} (1/6)^3 (5/6)^7$

b) $P(X \geq 3) = p(3) + \dots + p(10) = \sum_{k=3}^{10} \binom{10}{k} (1/6)^k (5/6)^{10-k}$

c) $E(X) = 10 * 1/6 = 1.67$

Napomena: očekivana vrednost $E(X)$ nekad ne može da se realizuje (nemoguće je da šestica padne 1.67 puta) tj. očekivana vrednost i najverovatniji ishod slučajne promenljive nisu isti pojmovi.

Binomna vs geometrijska raspodela

- Naglasicemo **razlike** između slučajnih promenljivih koje imaju binomnu, i onih koje imaju geometrijsku raspodelu.

	Broj ponavljanja	Broj uspeha
Binomna $\mathcal{B}(n,p)$	n	?
Geometrijska $\mathcal{G}(p)$	?	1

Aproksimacije binomne raspodele

- Pri velikom broju ponavljanja eksperimenta, izračunavanje verovatnoće događaja vezanih za binomnu raspodelu može biti računski komplikovano.
- Npr. ako je verovatnoća da je neki proizvod neispravan 0.08, i posmatramo seriju od 5000 proizvoda, verovatnoća da među njima bude manje od 350 neispravnih je

$$\sum_{k=0}^{349} \binom{5000}{k} 0.08^k 0.92^{5000-k}$$

- postoje dve aproksimacije binomne raspodele, koje omogućavaju približno izračunavanje ovakvih verovatnoća:
 1. Aproksimacija Poasonovom raspodelom
 2. Izračunavanje pomoću Gausove raspodele (Moavr-Laplasova teorema)

Poasonova raspodela $\mathcal{P}(\lambda)$

Kada se javlja?

- Poasonova raspodela je uvedena kao granični slučaj binomne raspodele
- broj ponavljanja $n \rightarrow \infty$, verovatnoća uspeha $p \rightarrow 0$, ali je proizvod ove dve vrednosti konstantan $np = \lambda, \lambda > 0$.
- Utvrđeno je da slučajne promenljive sa Poasonovom raspodelom opisuju pojave kod kojih nas zanima **broj uspešnih realizacija u određenom vremenskom intervalu** (ukoliko su nezavisne jedna od druge i sa konstantnom stopom)
- Primenuje se u matematičkom modelovanju redova čekanja, retkih događaja, itd...
- Primeri: broj klijenata na šalteru banke tokom 1h; broj saobraćajnih nesreća u nekoj opštini tokom jednog dana; broj meteorita većih od 1m koji udare u Zemlju tokom jedne godine; broj vojnika u pruskoj vojsci stradali od šutanja konja u toku jedne godine (prva dokumentovana primena), ...

Poasonova raspodela $\mathcal{P}(\lambda)$

Kako izgleda?

- Da bi odredili skup vrednosti, pitamo se koliko puta može da se realizuje neki događaj u fiksnom vremenskom intervalu? Najmanje nijednom, a najveći broj nije unapred određen.

$$\mathbf{R}_X = \{0, 1, \dots, k, \dots\}$$

- Verovatnoća $p(k)$ je izvedena kao granična vrednost verovatnoće kod binomne raspodele ($e \approx 2.72$ je Ojlerov broj).

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \in \mathbf{R}_X, \quad \lambda > 0$$

- Može se pokazati da i matematičko očekivanje i disperzija kod Poasonove raspodele imaju vrednost λ

$$\mathbf{E}(X) = \lambda$$

$$\mathbf{D}(X) = \lambda$$

Poasonova raspodela - primer

Primer: Utvrđeno je da neki call centar u proseku dobija 7 poziva u minuti, i da broj poziva u minuti ima Poasonovu raspodelu. Izračunati verovatnoće (a) u toku jednog minuta će biti 5 poziva; (b) u toku jednog minuta će biti bar 1 poziv.

- Raspodela je Poasonova, prosečan tj. očekivani broj poziva odgovara parametru λ pa je raspodela slučajne promenljive X koja predstavlja broj poziva u minuti **$X: \mathcal{P}(7)$** .

$$\text{a)} \quad P(X=5) = p(5) = \frac{7^5}{5!} e^{-7} = 0.1277$$

$$\text{b)} \quad P(X \geq 1) = 1 - p(0) = 1 - \frac{7^0}{0!} e^{-7} = 0.9991$$

Poasonova raspodela – aproksimacija binomne

Iz same konstrukcije Poasonove raspodele kao graničnog slučaja binomne kad $n \rightarrow \infty$, $p \rightarrow 0$ vidimo da

$$\mathcal{B}(n,p) \rightarrow \mathcal{P}(np) \leftarrow \lambda = np$$

- aproksimacija je tim bolja ukoliko je n veće i p manje.
- npr. $\mathcal{B}(5000, 0.08) \rightarrow \mathcal{P}(400)$, pa je iz primera o neispravnim proizvodima

$$P(X < 350) \approx \sum_{k=0}^{349} \frac{400^k}{k!} e^{-400}$$

- Aproksimirani izraz je nešto manje računski kompleksan, ali se u praksi i dalje mnogo češće koristi aproksimacija normalnom raspodelom tj. Moavr-Laplasova teorema.