1

ELEMENTARNI POJMOVI TEORIJE VEROVATNOĆA I STATISTIKE

Zorica Mladenović, profesor
zorima@eunet.rs,
http://avs.ekof.bg.ac.rs

2

Struktura

- Slučajna promenljiva i raspodela verovatnoće
  - Prekidna
  - Neprekidna
- Očekivana vrednost i varijansa
- Linearna transformacija slučajne promenljive
- Važne teorijske raspodele
  - Normalna raspodela

Slučajna promenljiva i raspodela verovatnoće

- Polazni pojam teorije verovatnoće: prostor elementarnih događaja
- To je skup svih mogućih ishoda nekog eksperimenta
- Primeri:
  1. Novčić se baca jednom: $\Omega=\{\Pi,\Gamma\}$
  2. Novčić se baca dva puta: $\Omega=\{\Pi\Pi,\Pi\Gamma,\Gamma\Pi,\Gamma\Gamma\}$
  3. Kocka za igru se baca jednom: $\Omega=\{1,2,3,4,5,6\}$
  4. Kocka za igru se baca dva puta:

$$\Omega=\left\{\begin{matrix}11,12,13,14,15,16,\\21,22,23,24,25,26,\\31,32,33,34,35,36,\\41,42,43,44,45,46,\\51,52,53,54,55,56,\\61,62,63,64,65,66\end{matrix}\right\}$$

Prekidna slučajna promenljiva i raspodela verovatnoće

- Slučajna promenljiva: funkcija kojom se prostor elementarnih događaja preslikava u skup realnih brojeva.
- Na skupu sa konačno mnogo elemenata definiše se prekidna slučajna promenljiva.
- Primer 1:

  Novčić se baca tri puta

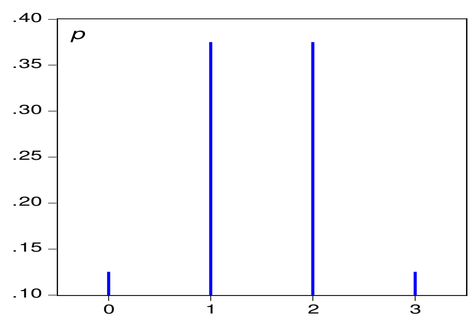
$$\Omega=\{\Pi\Pi\Pi,\Pi\Pi\Gamma,\Pi\Gamma\Pi,\Pi\Gamma\Gamma,\Gamma\Pi\Pi,\Gamma\Pi\Gamma,\Gamma\Gamma\Pi,\Gamma\Gamma\Gamma\}$$

Neka je X slučajna promenljiva koja pokazuje **broj pojavljivanja pisma** u sva tri bacanja:

$$X:\begin{pmatrix}0 & 1 & 2 & 3\\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8}\end{pmatrix}$$

Prekidna slučajna promenljiva i raspodela verovatnoće II

- Parovi vrednosti: *(0, 1/8), (1, 3/8), (2, 3/8) i (3, 1/8)* predstavljaju raspodelu verovatnoće slučajne promenljive *X*.
- Grafički prikaz raspodele: histogram



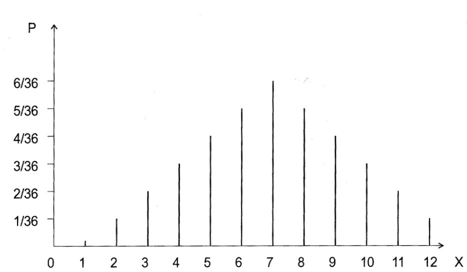
Prekidna slučajna promenljiva i raspodela verovatnoće III

- Primer 2:

  Kocka za igru se baca dva puta

  Neka je *X* slučajna promenljiva koja pokazuje zbir dobijenih brojeva iz oba bacanja

$$X : \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{pmatrix}$$



7

Prekidna slučajna promenljiva i raspodela verovatnoće IV

Opšta definicija

Određena je raspodela verovatnoće prekidne slučajne promenljive X ako su poznate moguće konkretne vrednosti: $x_1, x_2, \ldots, x_m$ i njihove verovatnoće: $p_1, p_2, \ldots, p_m$.

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ p_1 & p_2 & \ldots & p_m \end{pmatrix}$$

Uslovi

$1.\ p_1, p_2, \ldots, p_m \geq 0.$

$2.\ p_1 + p_2 + \ldots + p_m = 1.$

8

Očekivana vrednost i varijansa prekidne slučajne promenljive - osnovna numerička obeležja raspodele

- Očekivana vrednost: **broj** oko koga se grupišu vrednosti slučajne promenljive

$$E(X) = x_1 p_1 + x_2 p_2 + \ldots + x_m p_m.$$

- Varijansa: **broj** kojim se meri prosek kvadrata odstupanja pojedinačnih vrednosti od očekivane vrednosti

$$v(X) = E(X - E(X))^2$$
$$= (x_1 - E(X))^2 p_1 + \ldots + (x_m - E(X))^2 p_m.$$

- Kvadratni koren iz varijanse: standardno odstupanje (devijacija).

Očekivana vrednost i varijansa prekidne slučajne promenljive – primer 1

$$X:\begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$$

$$\begin{aligned} E(X) &= x_1p_1 + x_2p_2 + x_3p_3 + x_4p_4 \\ &= 0\cdot\frac{1}{8} + 1\cdot\frac{3}{8} + 2\cdot\frac{3}{8} + 3\cdot\frac{1}{8} \\ &= 1.5. \end{aligned}$$

$$\begin{aligned} v(X) = \sigma^2(X) &= (x_1 - E(X))^2 p_1 + \dots + (x_4 - E(X))^2 p_4 \\ &= (0-1.5)^2\frac{1}{8} + (1-1.5)^2\frac{3}{8} + (2-1.5)^2\frac{3}{8} + (3-1.5)^2\frac{1}{8} \\ &= 0.75 \end{aligned}$$

$$\sigma(X) = \sqrt{0.75} = 0.87.$$

Očekivana vrednost i varijansa prekidne slučajne promenljive – primer 2

$$X:\begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{pmatrix}$$

$$\begin{aligned} E(X) &= x_1p_1 + x_2p_2 + \dots + x_{11}p_{11} \\ &= 2\cdot\frac{1}{36} + 3\cdot\frac{2}{36} + \dots + 11\cdot\frac{2}{36} + 12\cdot\frac{1}{36} \\ &= 7 \end{aligned}$$

$$\begin{aligned} E\left(X^2\right) &= x_1^2p_1 + x_2^2p_2 + \dots + x_{11}^2p_{11} \\ &= 4\cdot\frac{1}{36} + 9\cdot\frac{2}{36} + \dots + 121\cdot\frac{2}{36} + 144\cdot\frac{1}{36} \\ &= 54.8333 \end{aligned}$$

$$v(X) = E(X^2) - E(X)^2 = 54.8333 - 49 = 5.8333.$$

$$\sigma(X) = \sqrt{5.8333} = 2.42.$$

Linearna transformacija slučajne promenljive

Neka je:

$$X - dato\ , E(\ X\ ), v(\ X\ )$$

Tada je:

$$Y = \alpha_0 + \alpha X\ , \alpha_0\ , \alpha = const.$$

$$E(\ Y\ ) = \alpha_0 + \alpha E(\ X\ )$$

$$v(\ X\ ) = \alpha^2 v(\ X\ )$$

Važna posledica:

$$\left.\begin{array}{l} X\ , E(\ X\ ) = \mu, v(X) = \sigma^2 \\ z = \dfrac{X - \mu}{\sqrt{v(\ X\ )}} = \dfrac{X - \mu}{\sigma} \end{array}\right\} \Rightarrow \begin{cases} E(\ z\ ) = 0, \\ v(\ z\ ) = 1. \end{cases}$$

Sa z označena je standardizovana slučajna promenljiva

Primer 1.1

Definiše se igra sa sledećim pravilima:

- Početni ulog: 2 dolara
- Novčić se baca tri puta i prebrojava se broj pojavljivanja pisma u tri bacanja (X)
- Dobitak se određuje u dolarskom iznosu X^2-X
- Koliki je očekivani ukupni dobitak?

Primer 1.1

X	0	1	2	3
p	1/8	3/8	3/8	1/8
Slučajna promenljiva ukupnog dobitka:				
$Y=(X^2-X)-2$	-2	-2	0	4
p	1/8	3/8	3/8	1/8
Očekivani ukupni dobitak je u stvari gubitak: (1/8)*(-2-6+4)=-0.5.				

Vežba 1

Data je raspodela verovatnoće prekidne slučajne promenljive:

X	2	3	4	5	6	7
p	2/30	5/30	8/30	8/30	5/30	2/30

a) Prikazati grafički raspodelu verovatnoće slučajne promenljive X.

b) $P(X \leq 4) = ?$

c) $E(X) = ?, v(X) = ?$

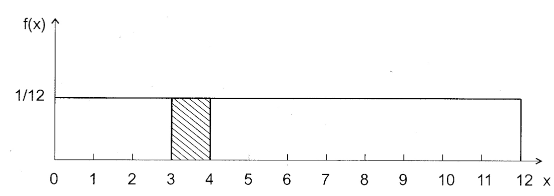
d) $E(2X - 3) = ?, v(2X - 3) = ?$

Neprekidna slučajna promenljiva i funkcija gustine

- Definiše se na skupu sa beskonačno mnogo elemenata iz poznatog intervala.
- Primer 3:
  - *Eksperiment*: posmatramo kretanje kazaljki na satu.
  - *Slučajna promenljiva*: vreme koje pokazuju kazaljke kada se zaustave na slučaj.
  - *Interval mogućih vrednosti*: (0, 12).
  - Verovatnoća da će kazaljke pokazati vreme 3:25:36 je nula.
  - Verovatnoća da će kazaljke pokazati vreme u intervalu 3:00 -4:00 je 1/12.

Neprekidna slučajna promenljiva i funkcija gustine II

- Raspodela verovatnoće neprekidne slučajne promenljive naziva se funkcija gustine, *f(x)*.
- Grafički prikaz iz primera 3:



Poruka: kod neprekidne slučajne promenljive ima smisla govoriti o verovatnoći u intervalu, što je površina ispod krive u datim granicama intervala.

17

Neprekidna slučajna promenljiva i funkcija gustine III

Opšta definicija
Neka je neprekidna slučajna promenljiva X definisana na intervalu (a,b).
Raspodela verovatnoće X naziva se funkcija gustine, $f(x)$, i definiše se na sledeći način:

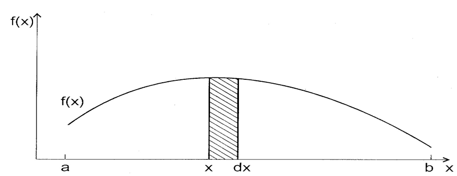
$$P(x < X \leq x + dx) = f(x)dx, \ \ a \leq x \leq b$$

Pri tome:
1. $f(x)$ ne uzima negativne vrednosti i
2. ukupna površina ispod krive $f(x)$ u granicama definisanja slučajne promenljive je jedan.

18

Neprekidna slučajna promenljiva i funkcija gustine IV

Grafički prikaz:



Očekivana vrednost i varijansa kod neprekidne slučajne promenljive:

$$E(X) = \int_a^b f(x) x dx$$

$$v(X) = \int_a^b f(x)(x - E(x))^2 dx$$

19

Važne neprekidne raspodele

- Ravnomerna (uniformna) raspodela
- Normalna raspodela
- Hi-kvadrat raspodela
- t – raspodela
- F – raspodela.

20

Ravnomerna raspodela

- Slučajna promenljiva *X* uzima sve vrednosti u intervalu *[a,b], a<b*.
- Funkcija gustine ravnomerne raspodele:

$$f(x) = \frac{1}{b-a},$$

$$E(X) = \frac{a+b}{2}, v(X) = \frac{(b-a)^2}{12}$$

- Primer 3. je specijalni slučaj ove raspodele za *a=0 i b=12*.

21

Normalna raspodela

- Slučajna promenljiva X uzima sve vrednosti na realnoj pravoj.
- Oznaka:

$$X : N(\mu, \sigma^2)$$

$$\mu - \textit{srednja vrednost}, E(X) = \mu$$

$$\sigma^2 - \textit{varijansa}, v(X) = \sigma^2 = E(X-\mu)^2$$

- Raspodela je u potpunosti određena parametrima očekivane (srednje) vrednosti i varijanse.
- Alternativni naziv: Gausova raspodela

22

Normalna raspodela II

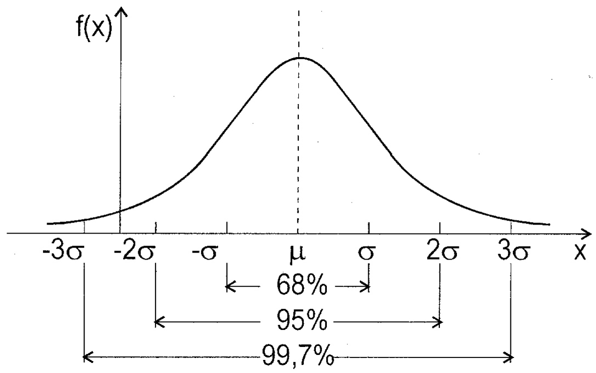
- Funkcija gustine normalne raspodele:

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

$$E(X) = \mu,\ v(X) = \sigma^2$$

23

Funkcija gustine normalne raspodele: grafički prikaz



24

Normalna raspodela u poznatoj formi

25

Značaj normalne raspodele

- Mnoge pojave u prirodi i društvu poseduju normalnu raspodelu (visina pojedinaca, koeficijent inteligencije, itd.).
- Mnoge teorijske raspodele se definišu prema normalnoj raspodeli i poprimaju njena svojstva pod određenim uslovima.
- Relevantnost primene zasniva se na rezultatu *centralne granične teoreme*: zbirno dejstvo (prosek) velikog broja nezavisnih slučajnih faktora poseduje normalnu raspodelu (opisna definicija).

26

Standardizovana normalna raspodela

- Oznaka *z*:

$$\left.\begin{array}{l} X : N(\mu,\sigma^2) \\ z = \dfrac{X-\mu}{\sigma} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} z : N(0,1) \\ E(z)=0, v(z)=1. \end{array}\right.$$

- Kada od slučajne promenljive *X* oduzmemo srednju vrednost i razliku podelimo sa standardnom devijacijom tada dobijamo novu slučajnu promenljivu *z*, čija je srednja vrednost 0, i varijansa 1.
- Ukoliko je polazna raspodela normalna, onda se ona postupkom standardizacije sl. promenljive ne menja.

27

Standardizovana normalna raspodela II

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

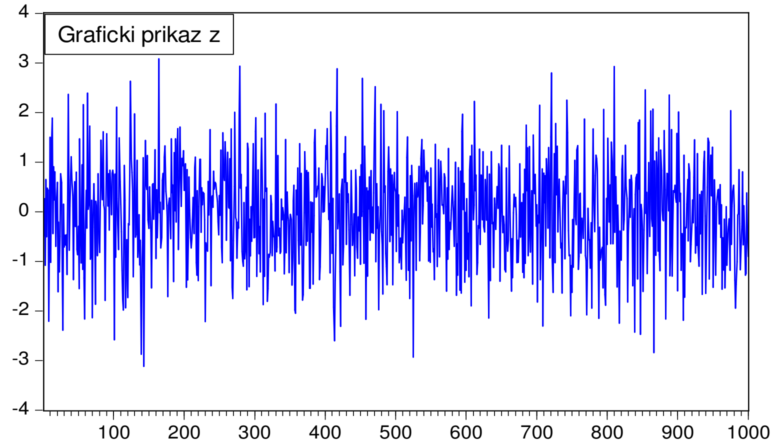
$$P(-1 \le z \le 1) \approx 0.68$$
$$P(-2 \le z \le 2) \approx 0.95$$
$$P(-3 \le z \le 3) \approx 0.99$$

28

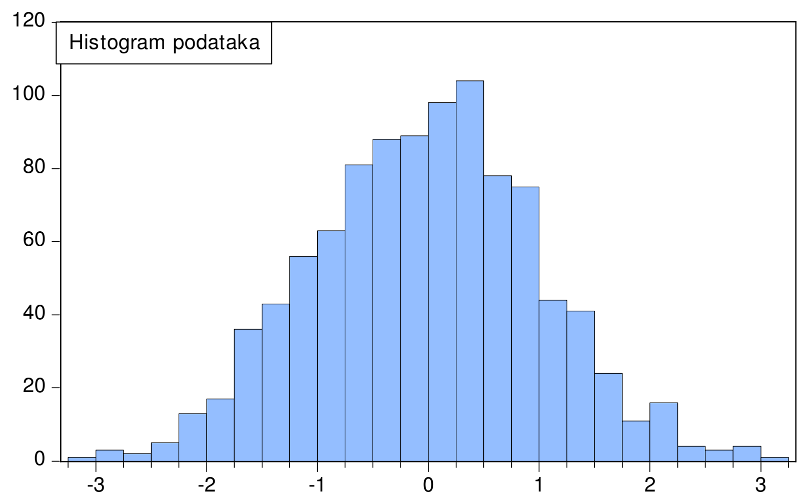
Standardizovana normalna raspodela
na osnovu generisanih podataka:
Primer iz programskog paketa EVIEWS I

- Generisani podaci: z
- Uzorak: 1000
- Ocena sred.vrednosti: -0.0218
- Maksimalna vrednost: 3.0813
- Minimalna vrednost: -3.1151
- Ocena stand. devijacije: 1.0145

Standardizovana normalna raspodela
na osnovu generisanih podataka:
Primer iz programskog paketa EVIEWS II



Standardizovana normalna raspodela
na osnovu generisanih podataka:
Primer iz programskog paketa EVIEWS III



Podinterval		Učestalost
-3.250000	-3.000000	1.000000
-3.000000	-2.750000	3.000000
-2.750000	-2.500000	2.000000
-2.500000	-2.250000	5.000000
-2.250000	-2.000000	13.00000
-2.000000	-1.750000	17.00000
-1.750000	-1.500000	36.00000
-1.500000	-1.250000	43.00000
-1.250000	-1.000000	56.00000
-1.000000	-0.750000	63.00000
-0.750000	-0.500000	81.00000
-0.500000	-0.250000	88.00000
-0.250000	0.000000	89.00000
0.000000	0.250000	98.00000
0.250000	0.500000	104.0000
0.500000	0.750000	78.00000
0.750000	1.000000	75.00000
1.000000	1.250000	44.00000
1.250000	1.500000	41.00000
1.500000	1.750000	24.00000
1.750000	2.000000	11.00000
2.000000	2.250000	16.00000
2.250000	2.500000	4.000000
2.500000	2.750000	3.000000
2.750000	3.000000	4.000000
3.000000	3.250000	1.000000

Standardizovana normalna raspodela
na osnovu generisanih podataka:
Van intervala (-2,2) je oko 5% (50) podataka