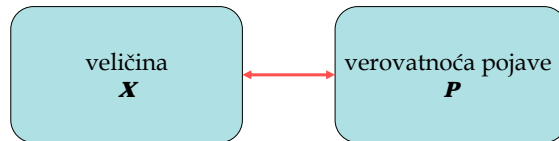


Statistička analiza u hidrologiji

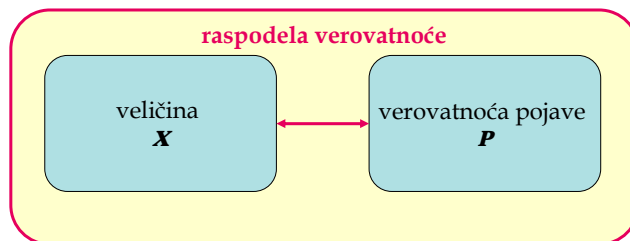
■ Uvod

- Statistička analiza se primenjuje na podatke osmatranja hidroloških veličina (najčešće: protoka i kiša)
- Cilj: opisivanje veze između veličine i verovatnoće njene pojave



Statistička analiza u hidrologiji

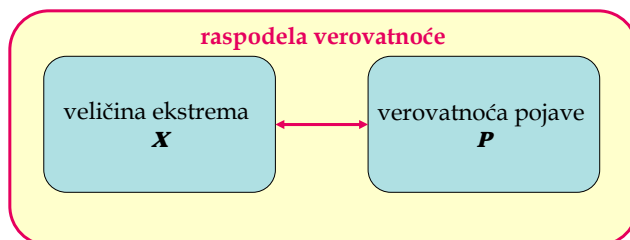
- Veza između veličine i verovatnoće = raspodela verovatnoće



Statistička analiza u hidrologiji

■ Cilj statističke analize:

- pronaći raspodelu verovatnoće ("model") koja dovoljno dobro opisuje vezu X - P u osmotrenom nizu podataka
- uz pomoć odabrane raspodele, odrediti:
 - verovatnoću pojave zadatog ekstrema, $P(X)$
 - veličinu ekstrema zadate verovatnoće pojave, $X(P)$



Statistička analiza u hidrologiji

■ Rezultati statističke analize koriste se za:

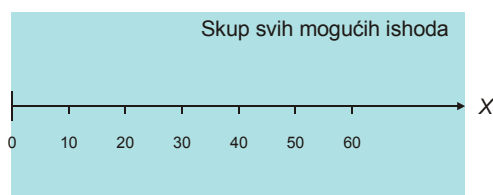
- projektovanje objekata i sistema za zaštitu od poplava
 - analiza maksimalnih protoka, kiša (analiza velikih voda)
- analizu raspoloživih količina vode za potrebe svih vidova korišćenja voda (vososnabdevanje, hidroenergetika, navodnjavanje)
 - analiza srednjih godišnjih protoka (analiza srednjih voda)
- analizu dugotrajnih sušnih perioda za potrebe vodosnabdevanja ili poljoprivrede
 - analiza minimalnih protoka, maksimalnih beskišnih perioda (analiza malih voda)
- analize kvaliteta voda i garantovanih ekoloških protoka
 - analiza minimalnih protoka (analiza malih voda)

Osnovni pojmovi iz verovatnoće

- Slučajna promenljiva
 - veličina koja se ponaša po nekom zakonu verovatnoće, tj. uzima određene vrednosti sa nekom verovatnoćom
- Ishodi ili realizacije
 - vrednosti koje uzima slučajna promenljiva
- Skup svih mogućih ishoda
 - oblast definisanosti slučajne promenljive
- Slučajni događaj
 - podskup skupa svih mogućih ishoda

Osnovni pojmovi

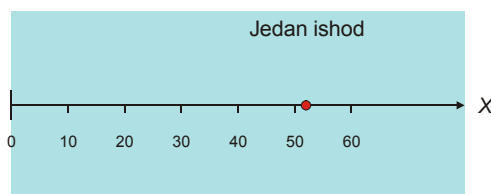
- Primer:
 - Visina kiše kao slučajna promenljiva X
 - Skup svih mogućih ishoda: $0 \leq X < \infty$



Osnovni pojmovi

■ Primer:

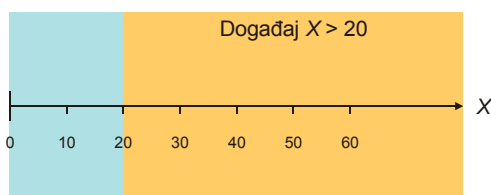
- Visina kiše kao slučajna promenljiva X
- Skup svih mogućih ishoda: $0 \leq X < \infty$
- Ishodi ili realizacije (osmatranja): $X = 52 \text{ mm}$



Osnovni pojmovi

■ Primer:

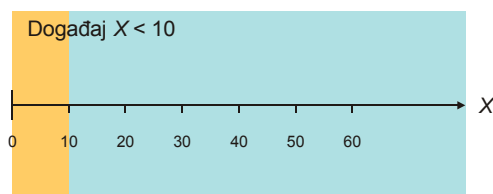
- Visina kiše kao slučajna promenljiva X
- Skup svih mogućih ishoda: $0 \leq X < \infty$
- Ishodi ili realizacije (osmatranja): $X = 52 \text{ mm}$
- Slučajni događaj:
 $X > 20 \text{ mm}$



Osnovni pojmovi

■ Primer:

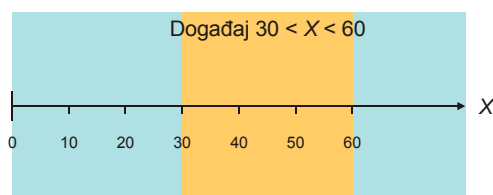
- Visina kiše kao slučajna promenljiva X
- Skup svih mogućih ishoda: $0 \leq X < \infty$
- Ishodi ili realizacije (osmatranja): $X = 52 \text{ mm}$
- Slučajni događaj:
 $X > 20 \text{ mm}, X \leq 10 \text{ mm}$



Osnovni pojmovi

■ Primer:

- Visina kiše kao slučajna promenljiva X
- Skup svih mogućih ishoda: $0 \leq X < \infty$
- Ishodi ili realizacije (osmatranja): $X = 52 \text{ mm}$
- Slučajni događaj:
 $X > 20 \text{ mm}, X \leq 10 \text{ mm}, 30 \leq X \leq 60 \text{ mm}$



Osnovni pojmovi

■ Slučajne promenljive:

- *prekidne ili diskretne*: skup svih mogućih ishoda = skup celih brojeva
 - broj dana u godini sa kišom većom od 10 mm
 - broj dana u godini sa temperaturom ispod 0°C
 - broj talasa velikih voda u godini sa maksimalnim protokom većim od neke vrednosti
- *neprekidne ili kontinualne*: skup svih mogućih ishoda = skup realnih brojeva
 - protok
 - visina kiše
 - nivo vode
 - zapremine talasa velikih voda
 - nivo podzemnih voda

Zakon raspodele verovatnoće

- Ishodi ili realizacije i događaji se dešavaju sa određenom verovatnoćom, prema RASPODELI VEROVATNOĆE

Zakon raspodele verovatnoće: diskretna slučajna promenljiva

$$X : \begin{pmatrix} x_1 & x_2 & x_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}$$

$$p_i = P\{X = x_i\}$$

$$p_1 + p_2 + p_3 + \dots = \sum_i p_i = 1$$

Zakon raspodele verovatnoće: diskretna slučajna promenljiva

■ Primer:

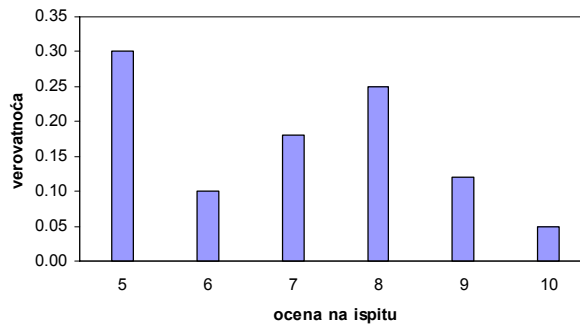
• bacanje novčića $X : \begin{pmatrix} P & G \\ 0.5 & 0.5 \end{pmatrix}$

• bacanje kocke $X : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$

• ocena na ispitu $X : \begin{pmatrix} 5 & 6 & 7 & 8 & 9 & 10 \\ 0.30 & 0.10 & 0.18 & 0.25 & 0.12 & 0.05 \end{pmatrix}$

Zakon raspodele verovatnoće: diskretna slučajna promenljiva

- Grafički prikaz
 - ocena na ispitu



Zakon raspodele verovatnoće: diskretna slučajna promenljiva

- Primeri događaja
 - ocena na ispitu

- verovatnoća da se padne ispit: $P\{X = 5\} = 0.30$

- verovatnoća da se položi ispit:

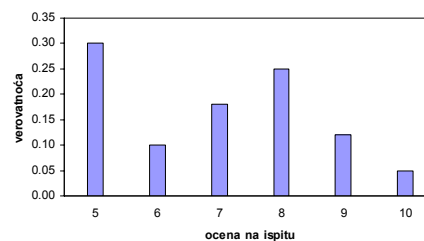
$$\begin{aligned} P\{X > 5\} &= P\{X \geq 6\} = P\{X = 6 \text{ ili } X = 7 \text{ ili } X = 8 \text{ ili } X = 9 \text{ ili } X = 10\} = \\ &= P\{X = 6\} + P\{X = 7\} + P\{X = 8\} + P\{X = 9\} + P\{X = 10\} = \\ &= 0.10 + 0.18 + 0.25 + 0.12 + 0.05 = 0.70 \end{aligned}$$

ili

$$P\{X \neq 5\} = 1 - P\{X = 5\} = 1 - 0.30 = 0.70$$

- verovatnoća za odličnu ocenu:

$$\begin{aligned} P\{X \geq 9\} &= P\{X = 9\} + P\{X = 10\} \\ &= 0.12 + 0.05 = 0.17 \end{aligned}$$



Zakon raspodele verovatnoće: kontinualna slučajna promenljiva

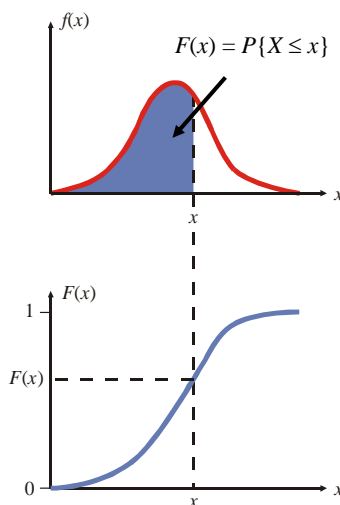
- Raspodela verovatnoće za kontinualnu slučajnu promenljivu

- funkcija gustine verovatnoće $f(x)$

$$\int_{-\infty}^{\infty} f(u) du = 1$$

- funkcija raspodele verovatnoće $F(x)$

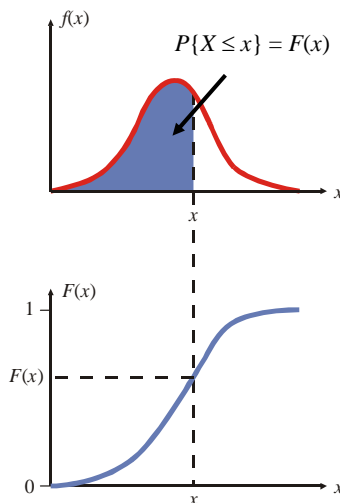
$$F(x) = P\{X \leq x\} = \int_{-\infty}^x f(u) du$$



Zakon raspodele verovatnoće: kontinualna slučajna promenljiva

- Verovatnoće događaja

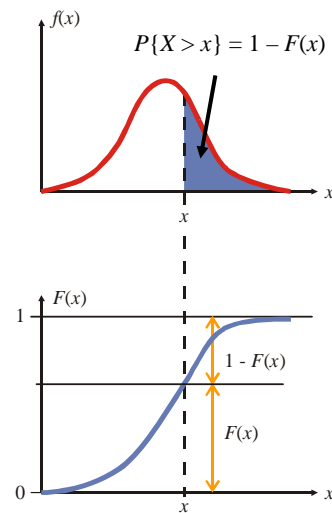
$$P\{X \leq x\} = F(x)$$



Zakon raspodele verovatnoće: kontinualna slučajna promenljiva

- Verovatnoće događaja

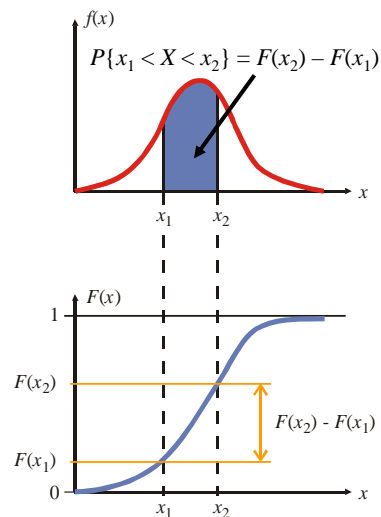
$$P\{X > x\} = 1 - P\{X \leq x\} = 1 - F(x)$$



Zakon raspodele verovatnoće: kontinualna slučajna promenljiva

- Verovatnoće događaja

$$\begin{aligned} P\{x_1 < X < x_2\} &= \\ &= 1 - P\{X < x_1\} - P\{X > x_2\} = \\ &= 1 - P\{X > x_2\} - P\{X < x_1\} = \\ &= P\{X < x_2\} - P\{X < x_1\} \\ &= F(x_2) - F(x_1) \end{aligned}$$



Zakon raspodele verovatnoće: kontinualna slučajna promenljiva

- Primeri događaja:
 - eksponencijalna raspodela:

$$f(x) = e^{-x}, \quad x \geq 0$$

$$F(x) = P\{X \leq x\} = \int_0^x e^{-u} du = -e^{-u} \Big|_0^x = 1 - e^{-x}$$

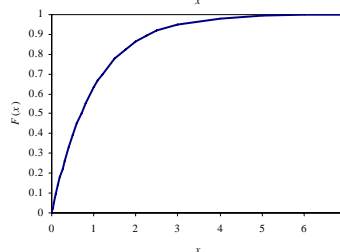
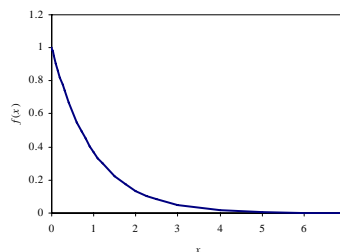
$$P\{X \leq 1\} = F(1) = 1 - e^{-1} = 1 - 0.368 = 0.632$$

$$P\{1 < X < 2\} = F(2) - F(1) = 1 - e^{-2} - 1 + e^{-1} = 0.368 - 0.135 = 0.233$$

$$P\{2 < X < 3\} = F(3) - F(2) = 1 - e^{-3} - 1 + e^{-2} = 0.135 - 0.050 = 0.085$$

$$P\{X > 3\} = 1 - F(3) = 1 - 1 + e^{-3} = 0.050$$

$$P\{X \leq 1\} + P\{1 < X < 2\} + P\{2 < X < 3\} + P\{X > 3\} = 0.632 + 0.233 + 0.085 + 0.050 = 1.000$$

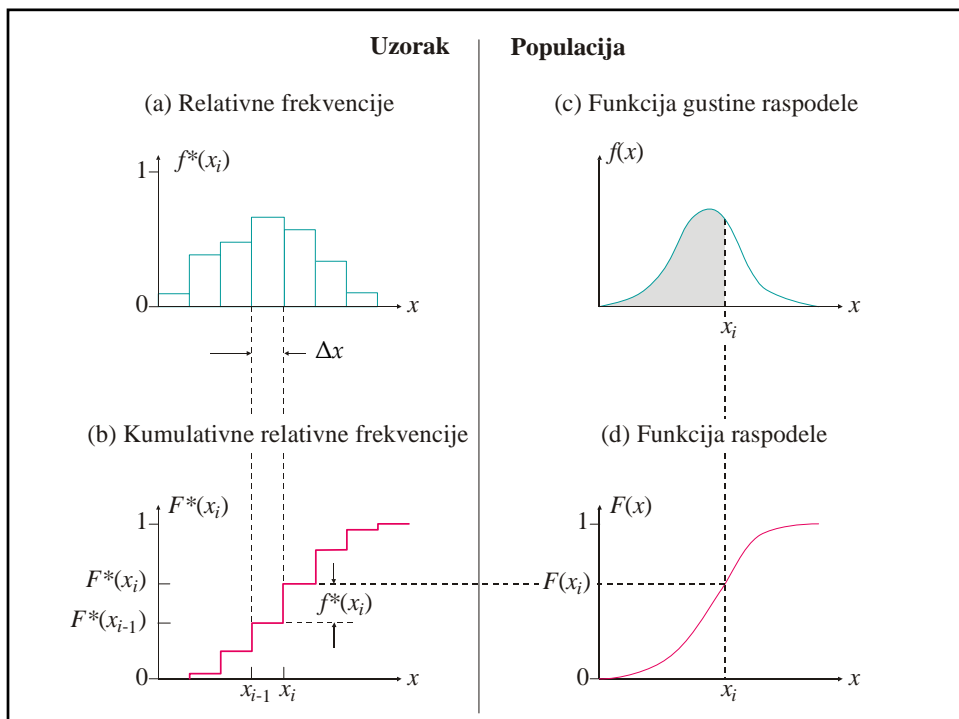


Populacija i uzorak

- uzorak je deo populacije
- raspodela populacije može biti poznata ili nepoznata (u hidrologiji: nepoznata)
- na osnovu osobina uzorka zaključujemo o populaciji

Populacija i uzorak

Populacija	Uzorak
verovatnoća	učestalost (frekvencija), empirijska verovatnoća
-	apsolutna frekvencija: broj podataka u klasi f
$P\{a \leq X < b\} = F(b) - F(a)$	relativna frekvencija: broj podataka u klasi u odnosu na ukupan broj podataka $f^* = f/N$
funkcija raspodele: $P\{X \leq x\} = F(x)$	kumulativna relativna frekvencija: broj podataka $\leq x$ u odnosu na ukupan broj podataka $F^* = \sum f^*$
	empirijska funkcija raspodele



Empirijska raspodela verovatnoće, $F_e(x)$

- “Pandan” funkciji raspodele $F(x)$ tj. verovatnoći $P\{X \leq x\}$ koji se određuje na osnovu uzorka

- kao kumulativna relativna frekvencija:

$$F_e(x_k) = P\{X \leq x_k\} = \frac{k}{N} = \frac{\text{broj podataka} \leq x_k}{\text{broj podataka u nizu}}$$

x_k - k -ti podatak u nizu uređenom u rastući redosled

- primer ($N = 51$):

$$F_e(x_5) = P\{X \leq x_5\} = P\{X \leq 3360\} = \frac{5}{51} = 0.098$$

k	x_k
1	2680
2	2996
3	3190
4	3310
5	3360
6	...

Empirijska raspodela

- Kumulativna relativna frekvencija kao empirijska raspodela

$N = 40$

k	x_k	k / N
1	$x_1 = x_{\min}$	1/40
2	x_2	2/40
3	x_3	3/40
4	x_4	4/40
5	x_5	5/40
...		
38	x_{38}	38/40
39	x_{39}	39/40
40	$x_{40} = x_{\max}$	40/40 = 1

$$P\{X \leq x_{\max}\} = \frac{N}{N} = 1$$



$$P\{X > x_{\max}\} = 1 - P\{X \leq x_{\max}\} = 0$$

sigurno će X biti manje od x_{\max}
tj. nemoguće da X bude veće od x_{\max}



Empirijska raspodela

- "Korekcija" kumulativne relativne frekvencije kao empirijska raspodela

$N = 40$

k	x_k	$(k-1)/N$
1	$x_1 = x_{\min}$	$0/40 = 0$
2	x_2	$1/40$
3	x_3	$2/40$
4	x_4	$3/40$
5	x_5	$4/40$
...		
38	x_{38}	$37/40$
39	x_{39}	$38/40$
40	$x_{40} = x_{\max}$	$39/40$

$$P\{X \leq x_{\min}\} = \frac{0}{N} = 0$$



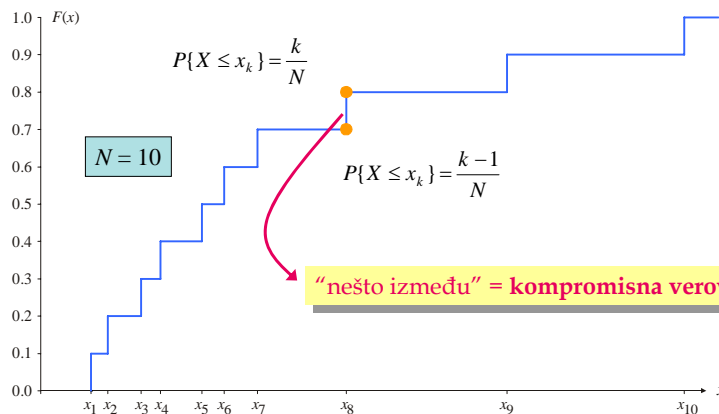
$$P\{X > x_{\min}\} = 1 - P\{X \leq x_{\min}\} = 1$$

nemoguće je da X bude manje od x_{\min}
tj. sigurno će X biti veće od x_{\min}



Empirijska raspodela

- Kumulativna relativna frekvencija kao empirijska raspodela



Empirijska raspodela

- Kompromisna verovatnoća po Hejzenu kao empirijska raspodela

$N = 40$

k	x_k	$(k - 0.5) / N$
1	$x_1 = x_{\min}$	0.5/40
2	x_2	1.5/40
3	x_3	2.5/40
4	x_4	3.5/40
5	x_5	4.5/40
...		
38	x_{38}	37.5/40
39	x_{39}	38.5/40
40	$x_{40} = x_{\max}$	39.5/40

$$P\{X \leq x_k\} = \frac{k - 0.5}{N}$$

$$P\{X \leq x_{\min}\} = \frac{0.5}{N} = \frac{0.5}{40} = 0.0125$$

$$P\{X \leq x_{\max}\} = \frac{N - 1}{N} = \frac{39.5}{40} = 0.9875$$

$$P\{X > x_{\max}\} = 1 - P\{X \leq x_{\max}\} = 0.0125$$

Empirijska raspodela

- Kompromisna verovatnoća po Vejbulu kao empirijska raspodela

$N = 40$

k	x_k	$k / (N + 1)$
1	$x_1 = x_{\min}$	1/41
2	x_2	2/41
3	x_3	3/41
4	x_4	4/41
5	x_5	5/41
...		
38	x_{38}	38/41
39	x_{39}	39/41
40	$x_{40} = x_{\max}$	40/41

$$P\{X \leq x_k\} = \frac{k}{N + 1}$$

$$P\{X \leq x_{\min}\} = \frac{1}{N + 1} = \frac{1}{41} = 0.0244$$

$$P\{X \leq x_{\max}\} = \frac{N}{N + 1} = \frac{40}{41} = 0.9756$$

$$P\{X > x_{\max}\} = 1 - P\{X \leq x_{\max}\} = 0.0244$$

Osobine raspodela verovatnoće

■ Momenti raspodele

- momenti oko koordinatnog početka

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

- momenti oko sredine

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

Osobine raspodela verovatnoće

■ Mere centralne tendencije

- srednja vrednost
 - težište gustine raspodele:

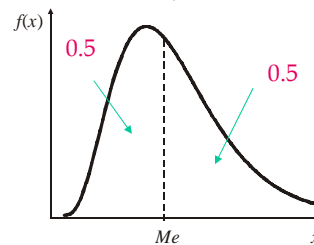
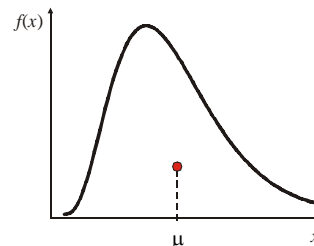
$$\mu'_1 = \mu = \int_{-\infty}^{\infty} x f(x) dx$$

- iz uzorka:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

- medijana:

$$F(Me) = \int_{-\infty}^{Me} f(x) dx = \int_{Me}^{\infty} f(x) dx = 0.5$$



Osobine raspodela verovatnoće

■ Mere odstupanja od srednje vrednosti

- disperzija (varijansa):

$$\mu_2 = \sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

- iz uzorka:

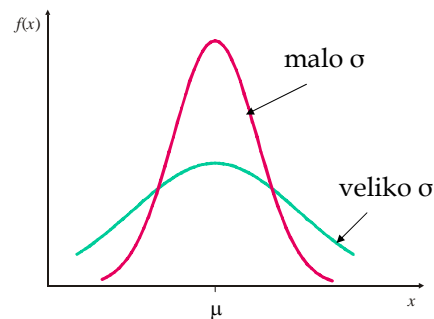
$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$$

- standardna devijacija:

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

- koeficijent varijacije:

$$C_v = \frac{\sigma}{\mu} \quad c_v = \frac{S}{\bar{x}}$$



Osobine raspodela verovatnoće

■ Asimetrija raspodele

- treći momenat:

$$\mu_3 = \int_{-\infty}^{\infty} (x - \mu)^3 f(x) dx$$

- koeficijent asimetrije:

$$C_s = \frac{\mu_3}{\sigma^3}$$

- iz uzorka:

$$c_s = \frac{N}{(N-1)(N-2)} \frac{1}{S^3} \sum_{i=1}^N (x_i - \bar{x})^3$$

