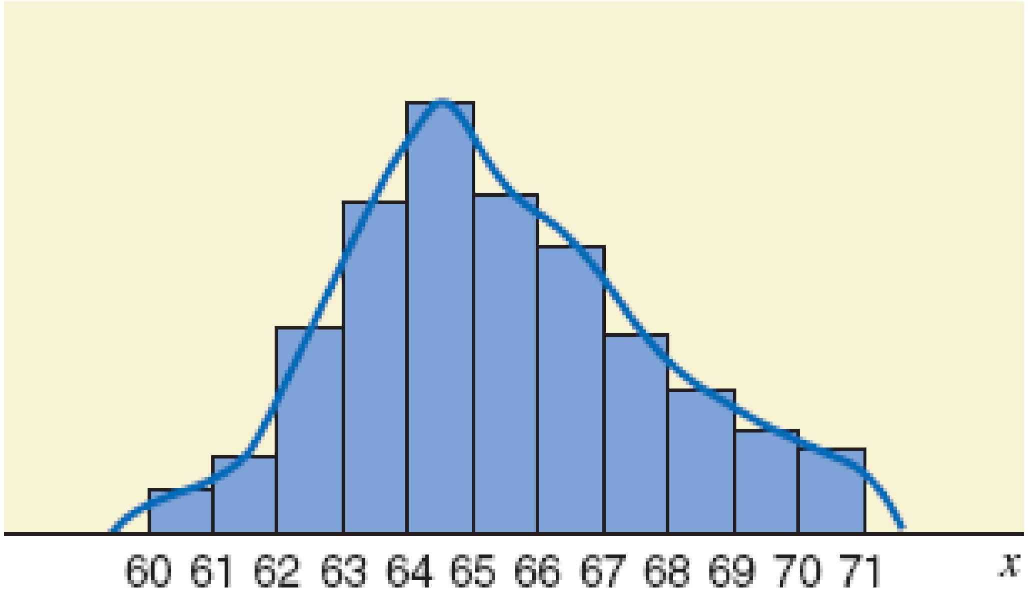
POGLAVLJE 6

NEPREKIDNE SLUČAJNE PROMJENLJIVE I NORMALNA RASPODJELA

6.1 NEPREKIDNA RASPODJELA VJEROVATNOĆA

Visina studentkinja (u inčima) x	f	Relativna frekvencija
60 do 61	90	0.018
61 do 62	170	0.034
62 do 63	460	0.092
63 do 64	750	0.150
64 do 65	970	0.194
65 do 66	760	0.152
66 do 67	640	0.128
67 do 68	440	0.088
68 do 69	320	0.064
69 do 70	220	0.044
70 do 71	180	0.036
	N=5000	Zbir = 1

Slika 6.2 Glatki poligon raspodjele vjerovatnoća za visine.

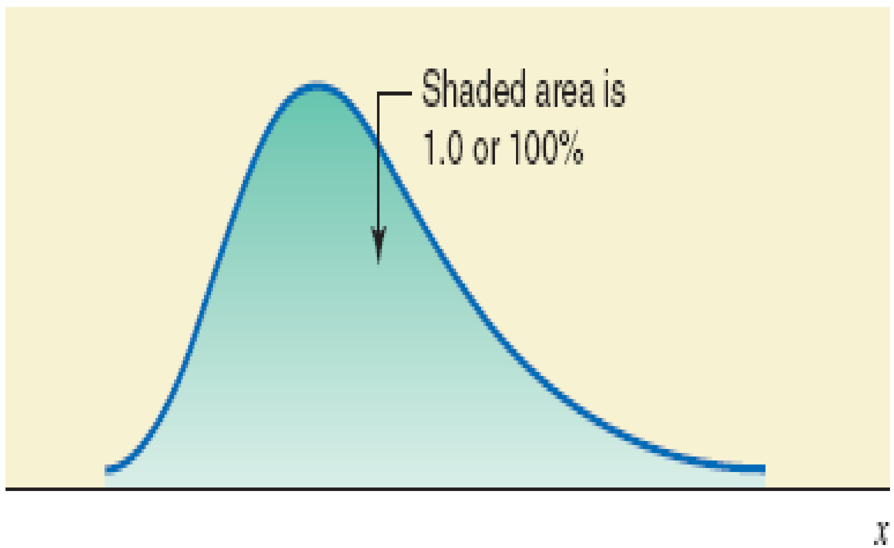


NEPREKIDNA RASPODJELA VJEROVATNOĆA

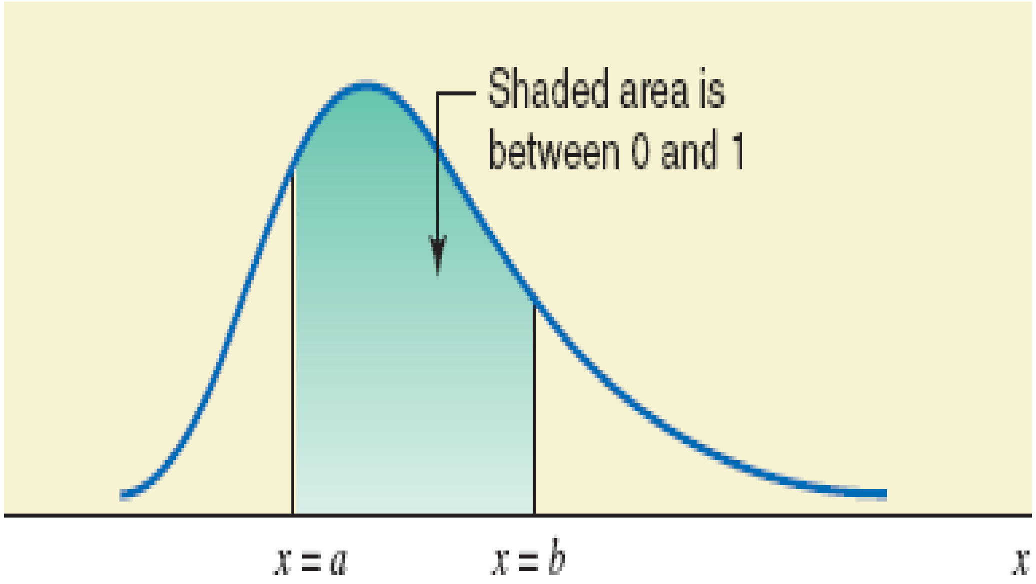
Dvije karakteristike

1. Vjerovatnoća da x uzme vrijednost iz bilo kog intervala se nalazi u intervalu od 0 do 1
2. Ukupna vjerovatnoća svih (međusobno disjunktnih) intervala u okviru kojih x može da uzme vrijednost je 1.0

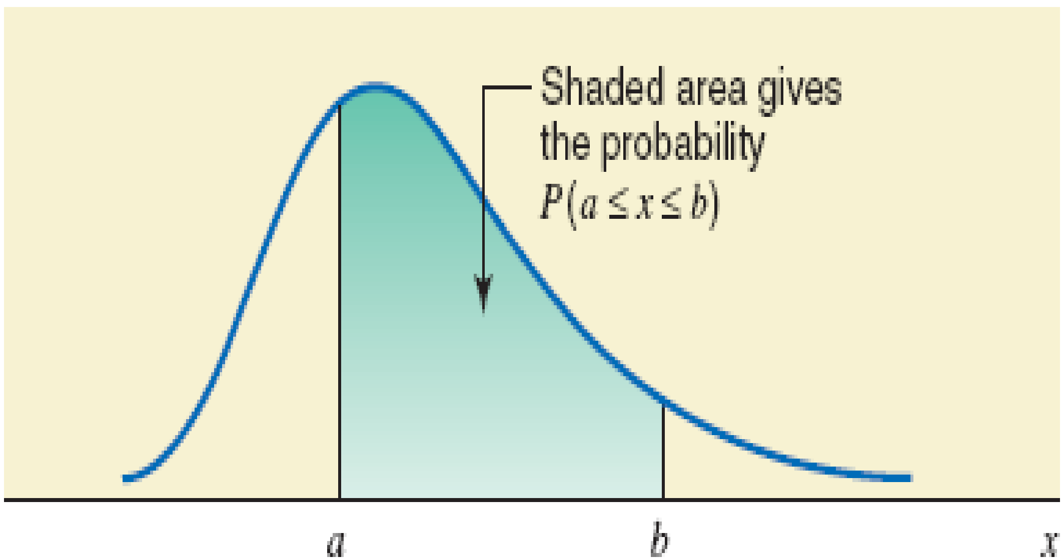
Slika 6.4 Ukupna površina ispod krive raspodjele vjerovatnoća.



Slika 6.3 Površina ispod krive između dvije tačke.



Slika 6.5 Površina ispod krive kao vjerovatnoća.



Slika 6.6 Vjerovatnoća da se x nađe u intervalu od 65 do 68.

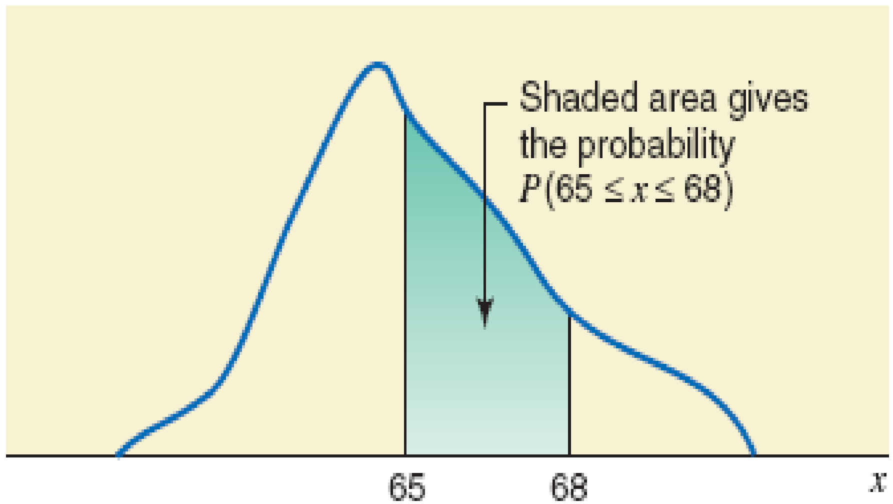


Figure 6.6 Probability that x lies in the interval 65 to 68.

Slika 6.7 Vjerovatnoća da *x* uzme jednu vrijednost je nula.

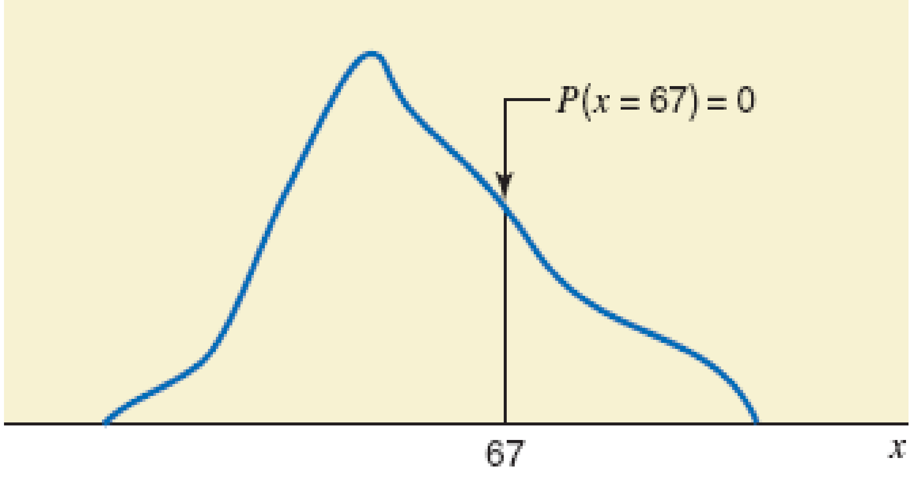


Figure 6.7 The probability of a single value of *x* is zero.

Slika 6.8 Vjerovatnoća za "od 65 do 68" i "između 65 i 68".

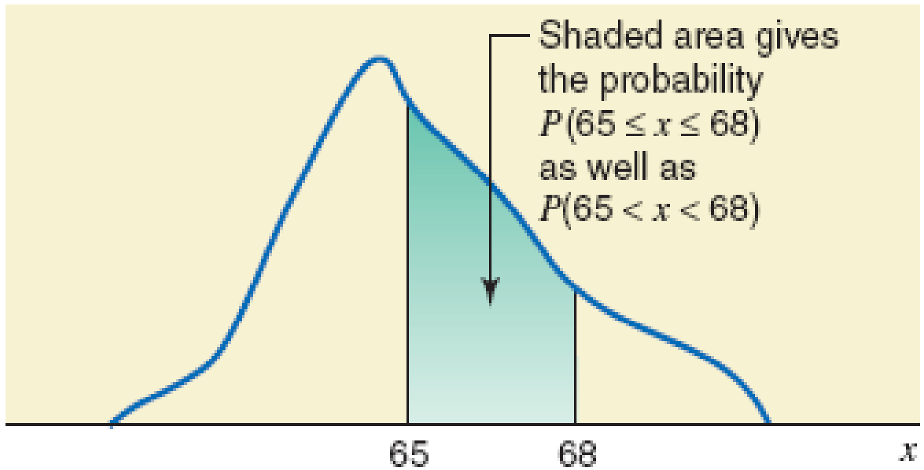


Figure 6.8 Probability "from 65 to 68" and "between 65 and 68."

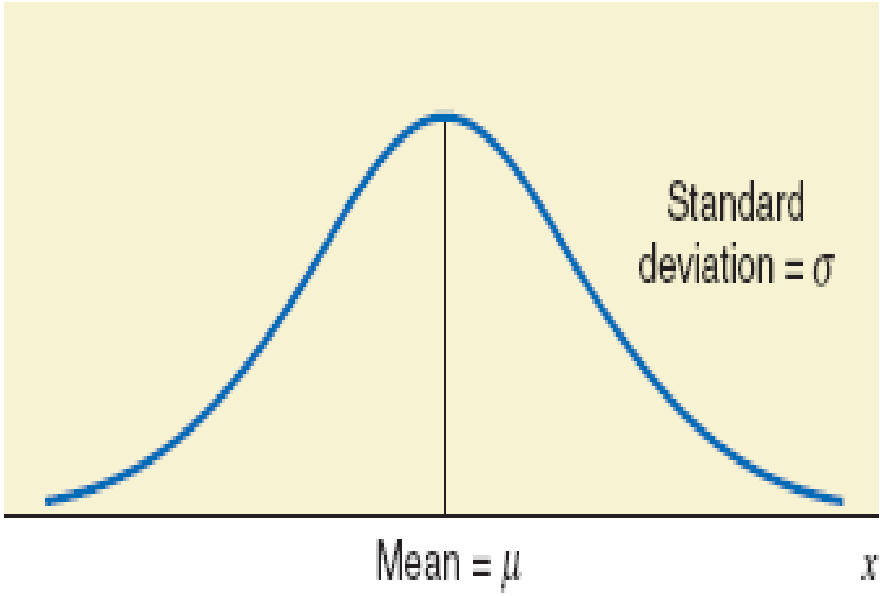
6.2 NORMALNA RASPODJELA

Normalna raspodjela vjerovatnoća

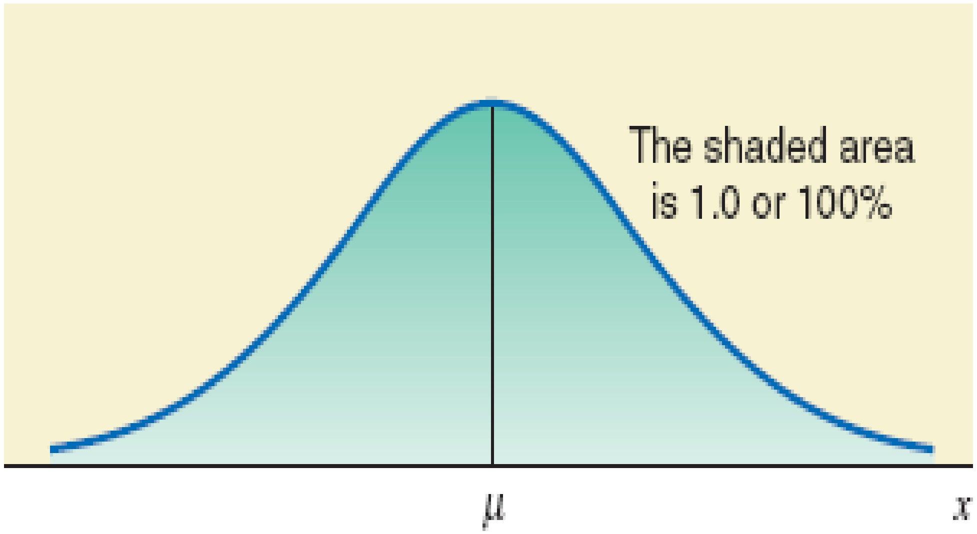
Normalna raspodjela vjerovatnoća, kada se grafički prikaže, daje krivu u obliku zvona takvu da:

1. Ukupna površina ispod krive je 1.0.
2. Kriva je simetrična u odnosu na aritmetičku sredinu.
3. Dva kraja krive se protežu u beskonačnost.

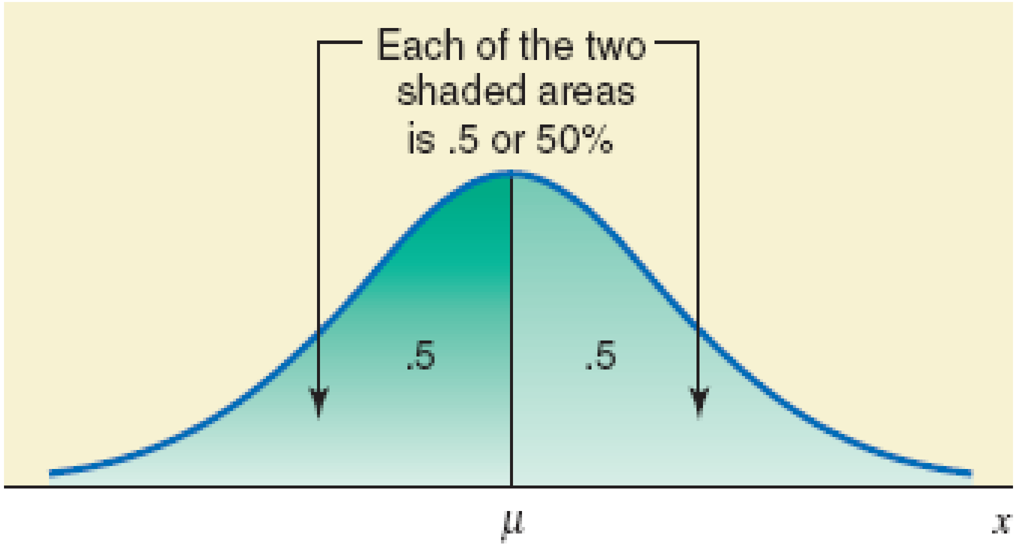
Slika 6.11 Normalna raspodjela sa aritmetičkom sredinom μ i standardnom devijacijom σ.



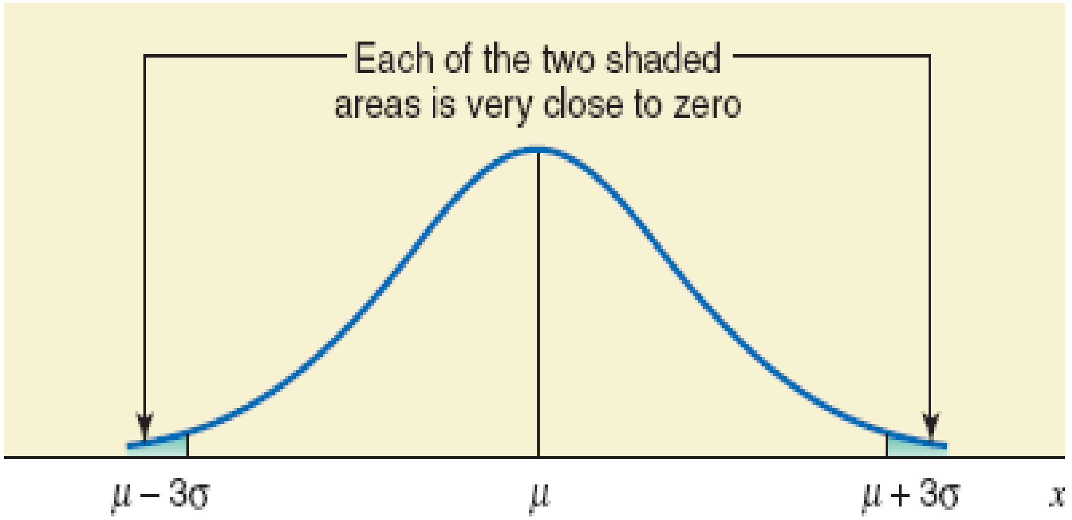
Slika 6.12 Ukupna površina ispod normalne krive.



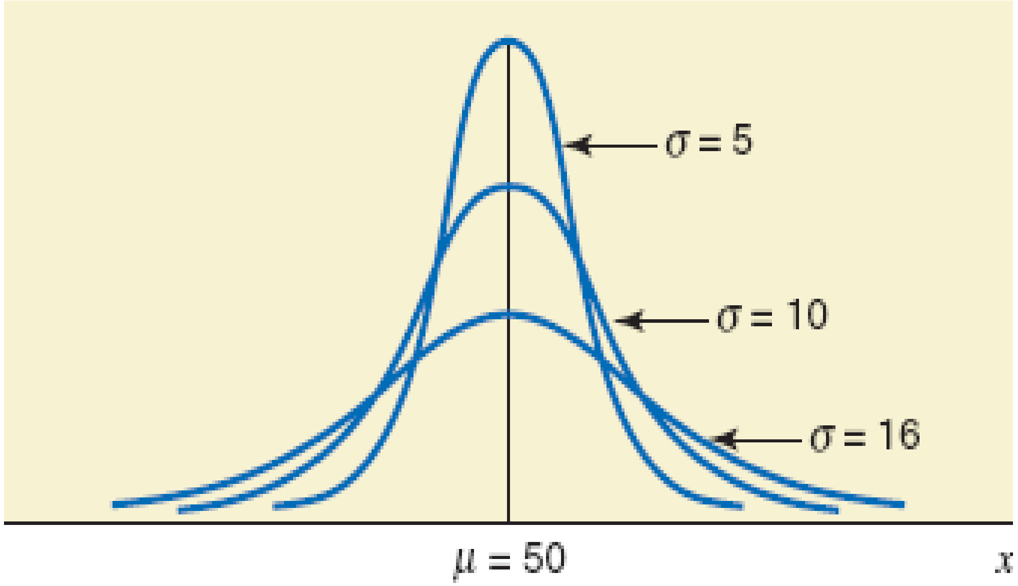
Slika 6.13 Normalna kriva je simetrična u odnosu na aritmetičku sredinu.



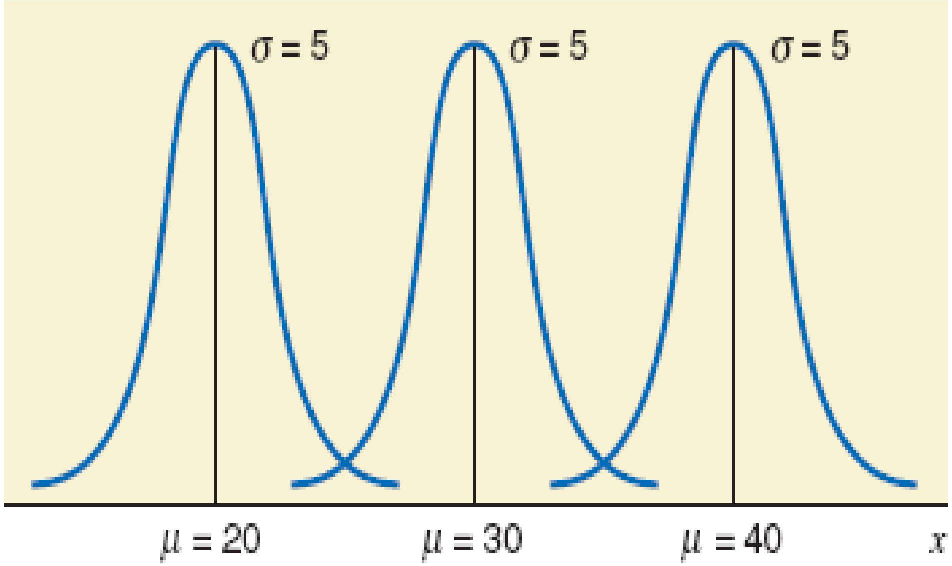
Slika 6.14 Površine ispod normalne krive izvan $\mu \pm 3\sigma$.



Slika 6.15 Tri krive normalne raspodjele sa istom aritmetičkom sredinom ali različitim standardnim devijacijama.



Slika 6.16 Tri krive normalne raspodjele sa različitim aritmetičkim sredinama ali sa istom standardnom devijacijom.

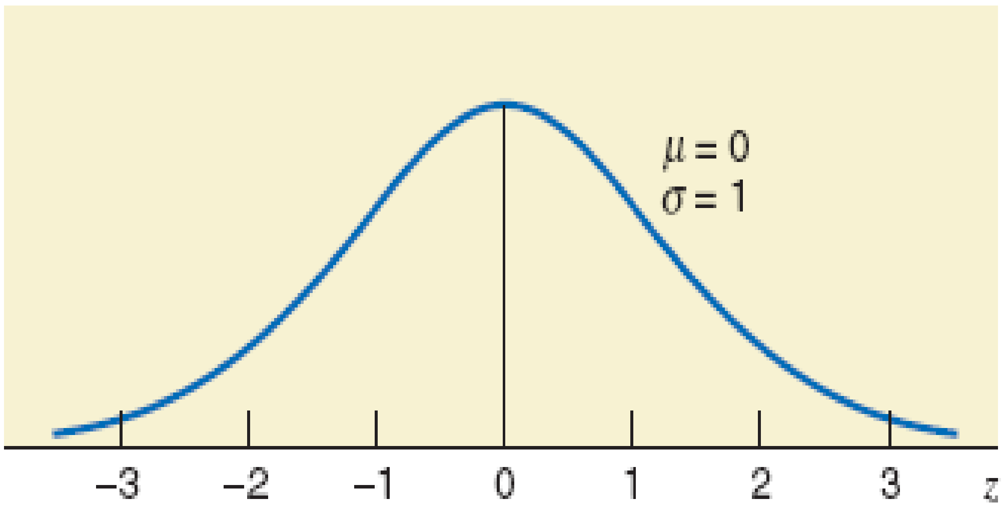


6.3 STANDARDIZOVANA NORMALNA RASPODJELA

Definicija

Normalna raspodjela sa $\mu = 0$ i $\sigma = 1$ se naziva ***standardizovana normalna raspodjela***.

Slika 6.17 Kriva standardizovane normalne raspodjele.



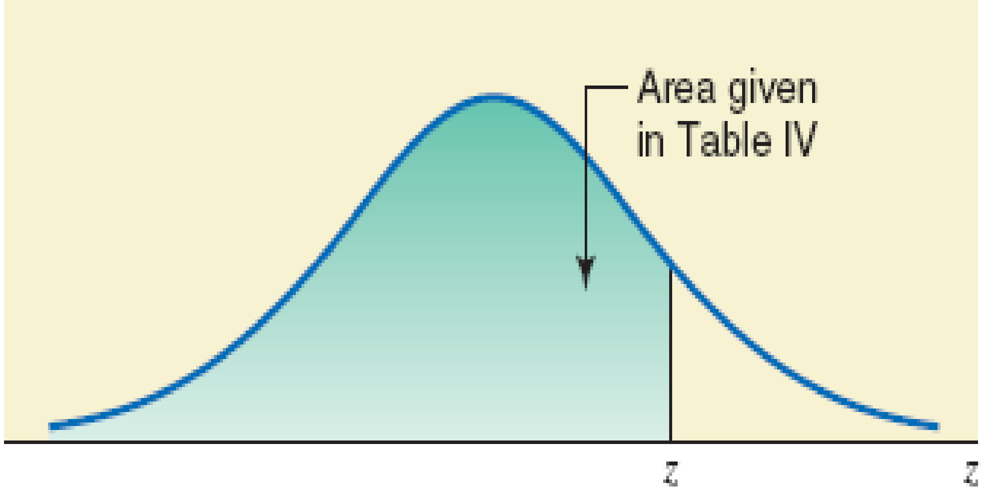
STANDARDIZOVANA NORMALNA RASPODJELA

z vrijednosti ili *z* rezultati

Definicija

Vrijednosti na horizontalnoj osi kod standardizovane normalne krive se označavaju sa *z* i nazivaju se ***z* vrijednostima ili *z* rezultatima**. Određena vrijednost *z* predstavlja rastojanje između aritmetičke sredine i tačke koja je prikazana sa *z* iskazano u standardnim devijacijama.

Slika 6.18 Površina ispod standardizovane normalne krive.



Primjer 6-1

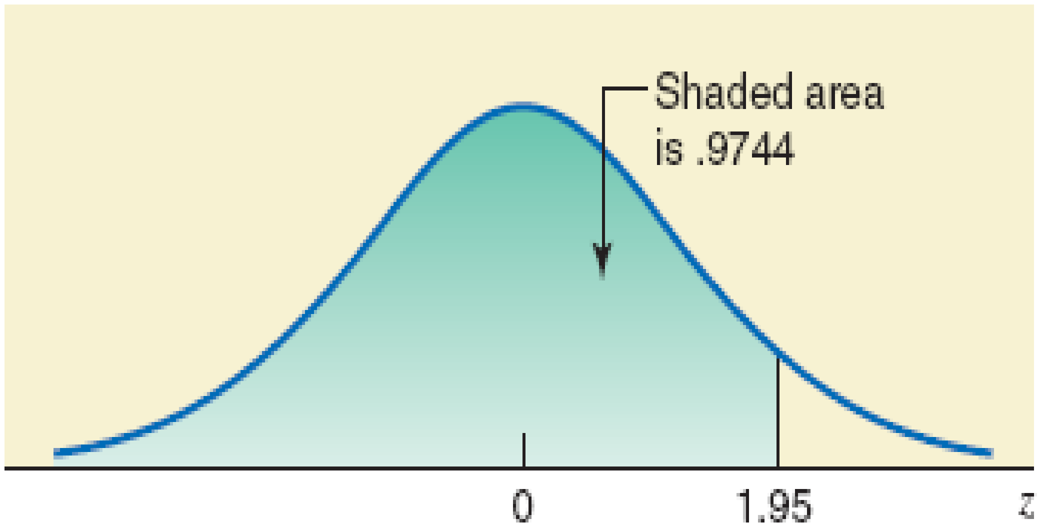
Odrediti površinu ispod standardizovane normalne krive ulijevo od z = 1.95.

Tabela 6.2 Površina ispod standardizovane normalne krive ulijevo od $z = 1.95$

z	**.00**	**.01**	**. . .**	**.05**	**. . .**	**.09**
−3.4	.0003	.0003	. . .	.0003	. . .	.0002
−3.3	.0005	.0005	. . .	.0004	. . .	.0003
−3.2	.0007	.0007	. . .	.0006	. . .	.0005
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
1.9	.9713	.9719	. . .	.9744	. . .	.9767
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
.	.	.	. . .	.	. . .	.
3.4	.9997	.9997	. . .	.9997	. . .	.9998

Required area

Slika 6.19 Površina ulijevo od $z = 1.95$.



Primjer 6-2

Odrediti površinu ispod standardizovane normalne krive od z = -2.17 do z = 0.

Primjer 6-2: Rješenje

- Da bismo odredili površinu od z=-2.17 do z =0 , prvo određujemo površinu ulijevo od z=0 i površinu ulijevo od z=-2.17 u Tablici IV. Kao što je prikazano u Tabeli 6.3, ove dvije površine su 0.5 i 0.0150, respektivno. Zatim oduzimamo 0.0150 od 0.5 da bismo odredili traženu površinu.
- Površina od -2.17 do 0 = $P(-2.17 \leq z \leq 0)$

= 0.5000 - 0.0150 = **0.4850**

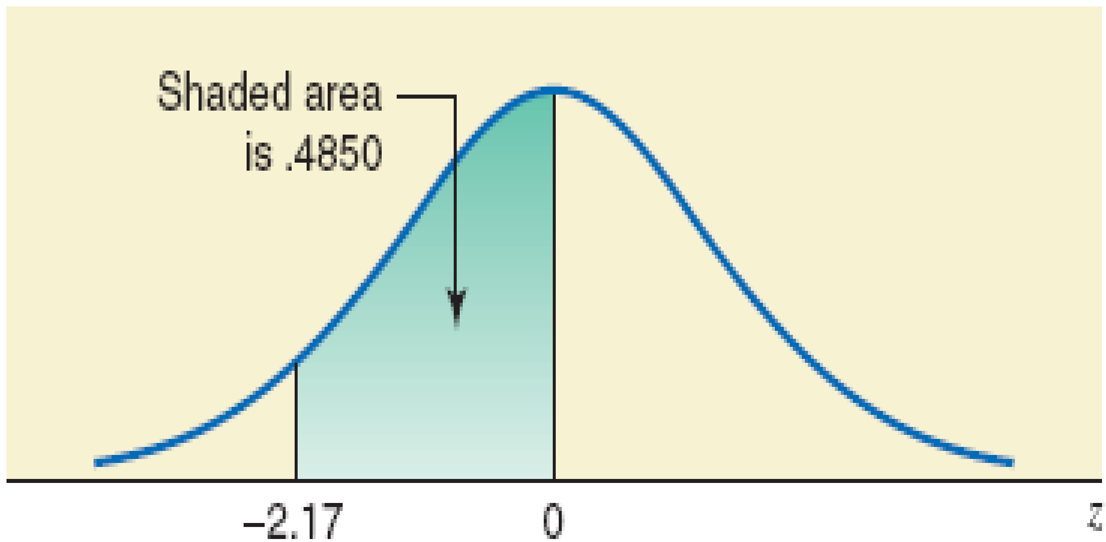
Tabela 6.3 Površina ispod standardizovane normalne krive

z	00	.01	...	.07	...	.09
−3.4	.0003	.0003	...	.0003	...	.0002
−3.3	.0005	.0005	...	.0004	...	.0003
−3.2	.0007	.0007	...	.0005	...	.0005
.	.	.	...	.	...	.
.	.	.	...	.	...	.
−2.1	.0179	.0174	...	.0150	...	.0143
.	.	.	...	.	...	.
.	.	.	...	.	...	.
0.0	.5000	.5040	...	.5279	...	.5359
.	.	.	...	.	...	.
.	.	.	...	.	...	.
3.4	.9997	.9997	...	.9997	...	.9998

Area to the left of $z = 0$

Area to the left of $z = -2.17$

Slika 6.20 Površina od z = -2.17 do z = 0.



Primjer 6-3

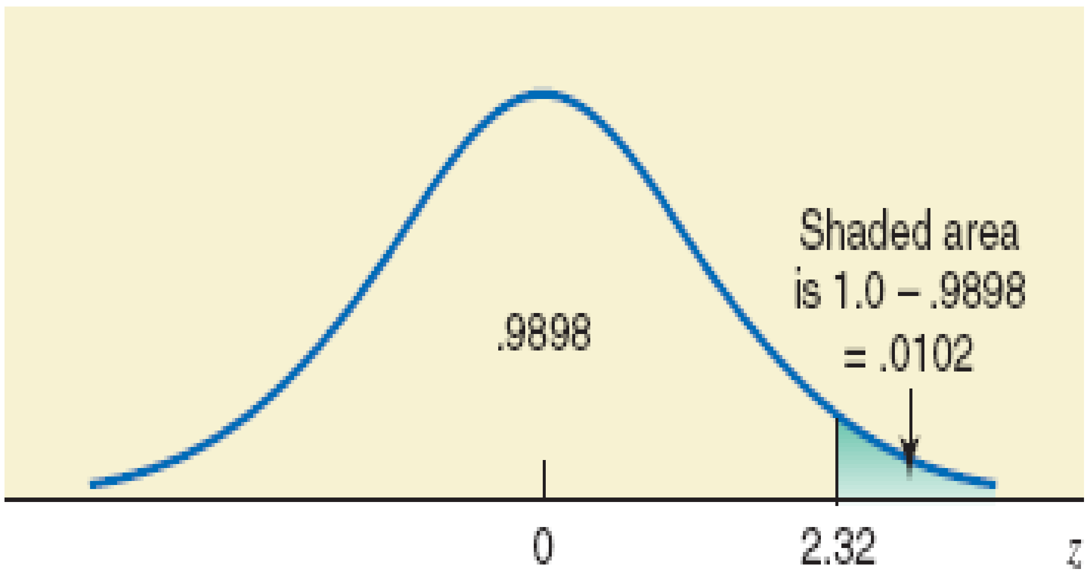
Odrediti sledeće površine ispod standardizovane normalne krive.

a) Površina udesno od $z = 2.32$
b) Površina ulijevo od $z = -1.54$

Primjer 6-3: Rješenje

a) Da bismo odredili površinu udesno od $z=2.32$, prvo nalazimo površinu ulijevo od $z=2.32$. Onda oduzimamo ovu površinu od 1.0, što je ukupna površina ispod krive. Tražena površina je $1.0 - 0.9898 = 0.0102$.

Slika 6.21 Površina udesno od $z = 2.32$.

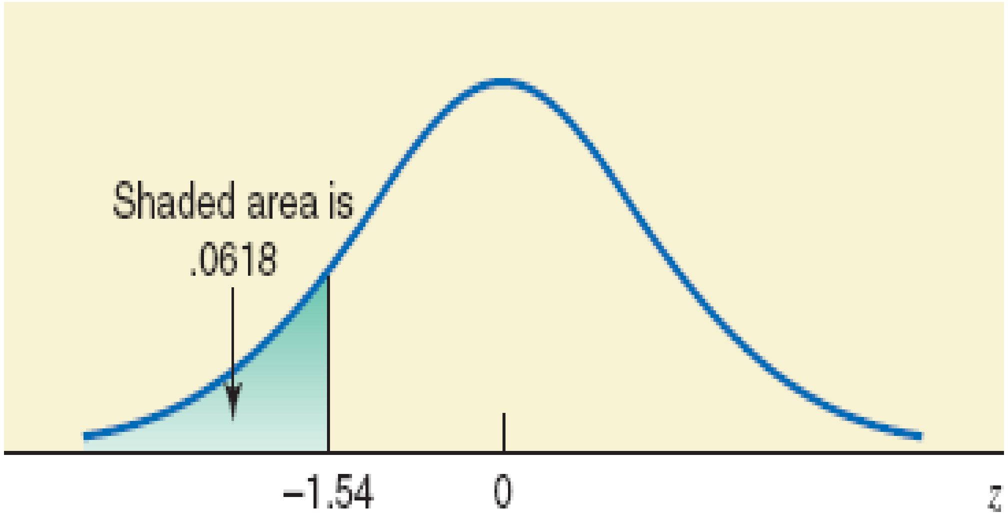


Primjer 6-3: Rješenje

b) Da bismo odredili površinu ispod standardizovane normalne krive ulijevo od $z=-1.54$, U Tablici IV nalazimo površinu koja odgovara za -1.5 iz z kolone i za 0.04 iz prvog reda iz zaglavlja tablice. Ova površina je 0.0618. Površina ulijevo od -1.54

$$= P\,(z < -1.54) = 0.0618$$

Slika 6.22 Površina ulijevo od z = -1.54.



6.4 STANDARDIZACIJA NORMALNE RASPODJELE

Transformisanje vrijednosti x u vrijednost z

Kod normalne slučajne promjenljive x, određena (realizovana) vrijednost x može da se transformiše u odgovarajuću vrijednost z korišćenjem formule

$$z = \frac{x - \mu}{\sigma}$$

gdje su μ i σ aritmetička sredina i standardna devijacija normalne raspodjele za x, respektivno.

POGLAVLJE 7

UZORAČKA RASPODJELA

7.1 RASPODJELA OSNOVNOG SKUPA I UZORAČKA RASPODJELA

- Raspodjela osnovnog skupa
- Uzoračka raspodjela

Raspodjela osnovnog skupa

Definicija

Raspodjela osnovnog skupa je raspodjela vjerovatnoća slučajne promjenljive X u osnovnom skupu.

Raspodjela osnovnog skupa

- Pretpostavimo da na kursu iz napredne statistike ima samo pet studenata i da su dati njihovi rezultati na kraju semestra

  70 78 80 80 95

- Neka je x slučajna promjenljiva koja se odnosi na broj poena studenata

Tabela 7.1 Raspodjela frekvencija i relativnih frekvencija osnovnog skupa

x	f	Relative Frequency
70	1	1/5 = .20
78	1	1/5 = .20
80	2	2/5 = .40
95	1	1/5 = .20
	$N = 5$	Sum = 1.00

Tabela 7.2 Raspodjela vjerovatnoća osnovnog skupa

x	$P(x)$
70	.20
78	.20
80	.40
95	.20
	$\Sigma P(x) = 1.00$

Uzoračka raspodjela

Definicija

Raspodjela vjerovatnoća statistike $\bar{x}$ se naziva uzoračkom raspodjelom. Ona predstavlja skup parova različitih vrijednosti koje može uzeti statistika $\bar{x}$ i odgovarajućih vjerovatnoća.

Uopšteno govoreći, raspodjela vjerovatnoća bilo koje statistike uzorka se naziva ***uzoračkom raspodjelom***.

Uzoračka raspodjela

- Vratimo se na podatke iz Tabele 7.1 o rezultatima koje je pet studenata ostvarilo na kraju semestra iz napredne statistike
- Posmatrajmo sve uzorke od po tri studenta koje je moguće izabrati bez ponavljanja iz tog osnovnog skupa.
- Ukupan broj mogućih uzoraka je

$$ {}_5C_3 = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 10 $$

Uzoračka raspodjela

- Obilježimo rezultate studenata slovima A, B, C, D i E tako da je
  - A = 70,   B = 78,   C = 80,   D = 80,   E = 95
- Prema tome, 10 mogućih uzoraka od po tri rezultata su
  - ABC,   ABD,   ABE,   ACD,   ACE,
    ADE,   BCD,   BCE,   BDE,   CDE

Tabela 7.3 Svi mogući uzorci od po tri elemenata i njihove aritmetičke sredine

Uzorak	Rezultati iz uzorka	$\bar{x}$
ABC	70, 78, 80	76.00
ABD	70, 78, 80	76.00
ABE	70, 78, 95	81.00
ACD	70, 80, 80	76.67
ACE	70, 80, 95	81.67
ADE	70, 80, 95	81.67
BCD	78, 80, 80	79.33
BCE	78, 80, 95	84.33
BDE	78, 80, 95	84.33
CDE	80, 80, 95	85.00

Tabela 7.4 Raspodjela frekvencija i relativnih frekvencija statistike $\bar{x}$ za uzorke od po 3 el.

$\bar{x}$	f	Relative Frequency
76.00	2	2/10 = .20
76.67	1	1/10 = .10
79.33	1	1/10 = .10
81.00	1	1/10 = .10
81.67	2	2/10 = .20
84.33	2	2/10 = .20
85.00	1	1/10 = .10
	$\Sigma f = 10$	Sum = 1.00

Tabela 7.5 Uzoračka raspodjela statistike $\bar{x}$ za uzorke od po tri elementa

$\bar{x}$	$P(\bar{x})$
76.00	.20
76.67	.10
79.33	.10
81.00	.10
81.67	.20
84.33	.20
85.00	.10
	$\Sigma P(\bar{x}) = 1.00$

7.2 SLUČAJNE (UZORAČKE) I NESLUČAJNE (SISTEMATSKE) GREŠKE

Definicija

Slučajna (uzoračka) greška je razlika između vrijednosti statistike uzorka i vrijednosti parametra posmatranog osnovnog skupa. Npr. za aritmetičku sredinu je

$$\text{Slučajna greška} = \bar{x} - \mu$$

pod pretpostavkom da je riječ o slučajnom uzorku i da nije napravljena nijedna od neslučajnih grešaka.

SLUČAJNE (UZORAČKE) I NESLUČAJNE (SISTEMATSKE) GREŠKE

Definicija

Greške koje nastaju prilikom prikupljanja, bilježenja podataka i njihovog unošenja u tabele se nazivaju ***neslučajnim (sistematskim) greškama***.

Razlozi nastanka neslučajnih grešaka

1. Ako uzorak nije slučajan (i samim tim je nereprezentativan), rezultati na osnovu uzorka mogu se značajno razlikovati od rezultata popisa.
2. Pitanja mogu biti tako formulisana da nisu razumljiva svim ispitanicima iz uzorka ili iz osnovnog skupa.
3. Ispitanici mogu namjerno da daju netačne informacije kao odgovore na neka osjetljiva pitanja.
4. Anketar može jednostavno napraviti grešku i prilikom evidentiranja podataka ili njihovog unosa u bazu podataka.

Primjer 7-1

Vratimo se ponovo na podatke iz Tabele 7.1. Pretpostavimo da smo iz osnovnog skupa izabrali samo jedan uzorak koji sadrži ostvareni broj poena tri studenata 70, 80 i 95. Odrediti slučajnu grešku.

Primjer 7-1: Rješenje

$$\mu = \frac{70 + 78 + 80 + 80 + 95}{5} = 80.60$$

$$\bar{x} = \frac{70 + 80 + 95}{3} = 81.67$$

$$\textit{Slučajna greška} = \bar{x} - \mu = 81.67 - 80.60 = 1{,}07$$

Odnosno, prosječan broj poena u uzorku studenata je za 1,07 veći od prosjeka u osnovnom skupu.

SLUČAJNE (UZORAČKE) I NESLUČAJNE (SISTEMATSKE) GREŠKE

- Sada pretpostavimo da je broj poena drugog studenta u uzorku unijet greškom, odnosno da je umjesto 80 unijet broj 82.
- Tada bi aritmetička sredina uzorka bila

$$\bar{x} = \frac{70 + 82 + 95}{3} = 82.33$$

SLUČAJNE (UZORAČKE) I NESLUČAJNE (SISTEMATSKE) GREŠKE

- Razlika između aritmetičke sredine uzorka i osnovnog skupa bi bila tada

$$\bar{x} - \mu = 82.33 - 80.60 = 1.73$$

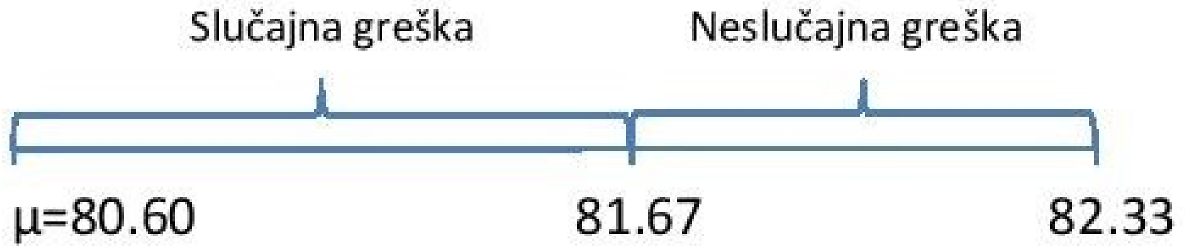
- Ova razlika ne predstavlja samo slučajnu grešku.
  - Samo 1.07 od ove razlike se odnosi na slučajnu grešku.

SLUČAJNE (UZORAČKE) I NESLUČAJNE (SISTEMATSKE) GREŠKE

- Ostatak predstavlja neslučajnu grešku.
  - Jednaka je 1.73 – 1.07 = 0.66
  - Nastala je prilikom pogrešnog unošenja drugog podatka uzorka
- Takođe,

$$Neslučajna\ greška = Netačna\ \bar{x} - Tačna\ \bar{x} = 82.33 - 81.67 = 0.66$$

Slika 7.1 Slučajna i neslučajna greška.



7.3 ARITMETIČKA SREDINA I STANDARDNA DEVIJACIJA STATISTIKE $\bar{x}$

Definicija

Aritmetička sredina i standardna greška od uzoračke raspodjele statistike $\bar{x}$ se nazivaju ***aritmetička sredina i standardna greška*** statistike $\bar{x}$ i označavaju se sa $\mu_{\bar{x}}$ i $\sigma_{\bar{x}}$, respektivno.

ARITMETIČKA SREDINA I STANDARDNA GREŠKA STATISTIKE $\bar{x}$

Aritmetička sredina statistike $\bar{x}$

Aritmetička sredina statistike $\bar{x}$ je uvijek jednaka aritmetičkoj sredini osnovnog skupa. Prema tome,

$$\mu_{\bar{x}} = \mu$$

ARITMETIČKA SREDINA I STANDARDNA GREŠKA STATISTIKE $\bar{x}$

Standardna greška statistike $\bar{x}$

Standardna greška statistike $\bar{x}$ je

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

gdje σ je standardna devijacija osnovnog skupa, a n veličina uzorka. Ova formula se primjenjuje pod uslovom da je $n/N \leq 0.05$, gdje je N veličina osnovnog skupa.

ARITMETIČKA SREDINA I STANDARDNA GREŠKA STATISTIKE $\bar{x}$

Ako uslov $n/N \leq 0.05$ nije zadovoljen, koristimo sledeću formulu pri računanju $\sigma_{\bar{x}}$:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

gdje se faktor $\sqrt{\frac{N-n}{N-1}}$ naziva popravnim faktorom za konačne skupove.

Dva važna zapažanja

1. Disperzija statistike uzorka $\bar{x}$ je manja od disperzije odgovarajuće promjenljive X u osnovnom skupu, odnosno
$$\sigma_{\bar{x}} < \sigma_x$$
2. Standardna greška $\sigma_{\bar{x}}$ se smanjuje sa povećavanjem uzorka.

7.4 OBLIK UZORAČKE RASPODJELE $\overline{x}$

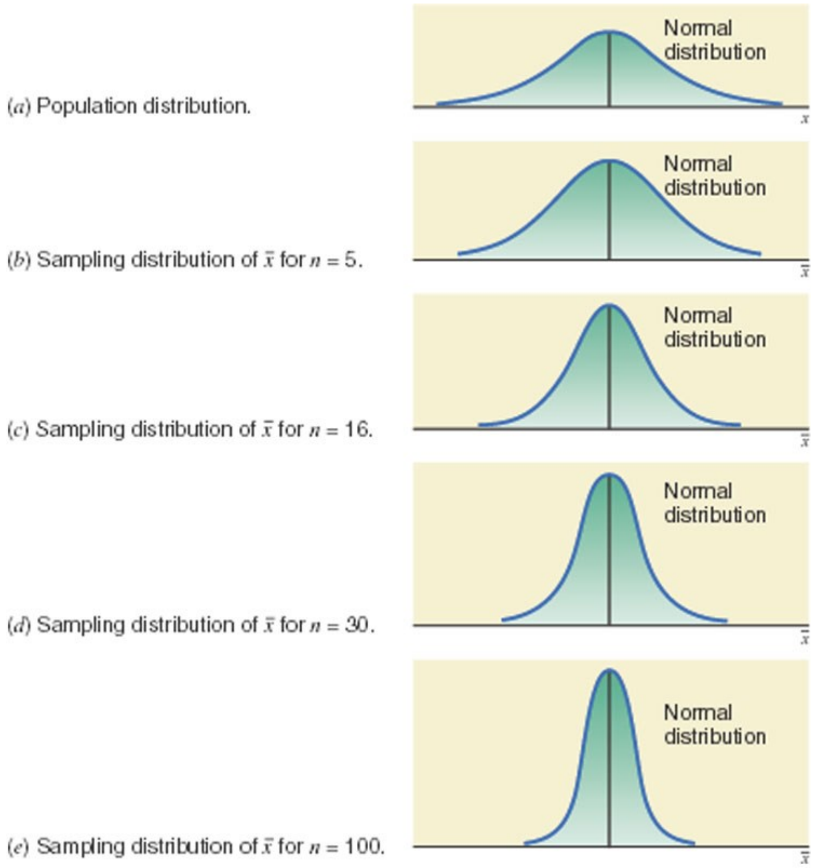
- Osnovni skup iz kojeg su izabrani uzorci ima normalnu raspodjelu.
- Osnovni skup iz kojeg su izabrani uzorci nema normalnu raspodjelu.

Uzorci iz osnovnog skupa sa normalnom raspodjelom

Kada osnovni skup iz kojeg su izabrani uzorci ima normalnu raspodjelu sa aritmetičkom sredinom μ i standardnom devijacijom σ, tada će uzoračka raspodjela statistike $\bar{x}$, takođe imati normalnu raspodjelu sa sledećom aritmetičkom sredinom i standardnom devijacijom, bez obzira na veličinu uzorka:

$$\mu_{\bar{x}} = \mu \quad i \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Slika 7.2 Raspodjela osnovnog skupa i uzoračka raspodjela statistike $\bar{x}$.



Uzorci iz osnovnog skupa koji nemaju normalnu raspodjelu

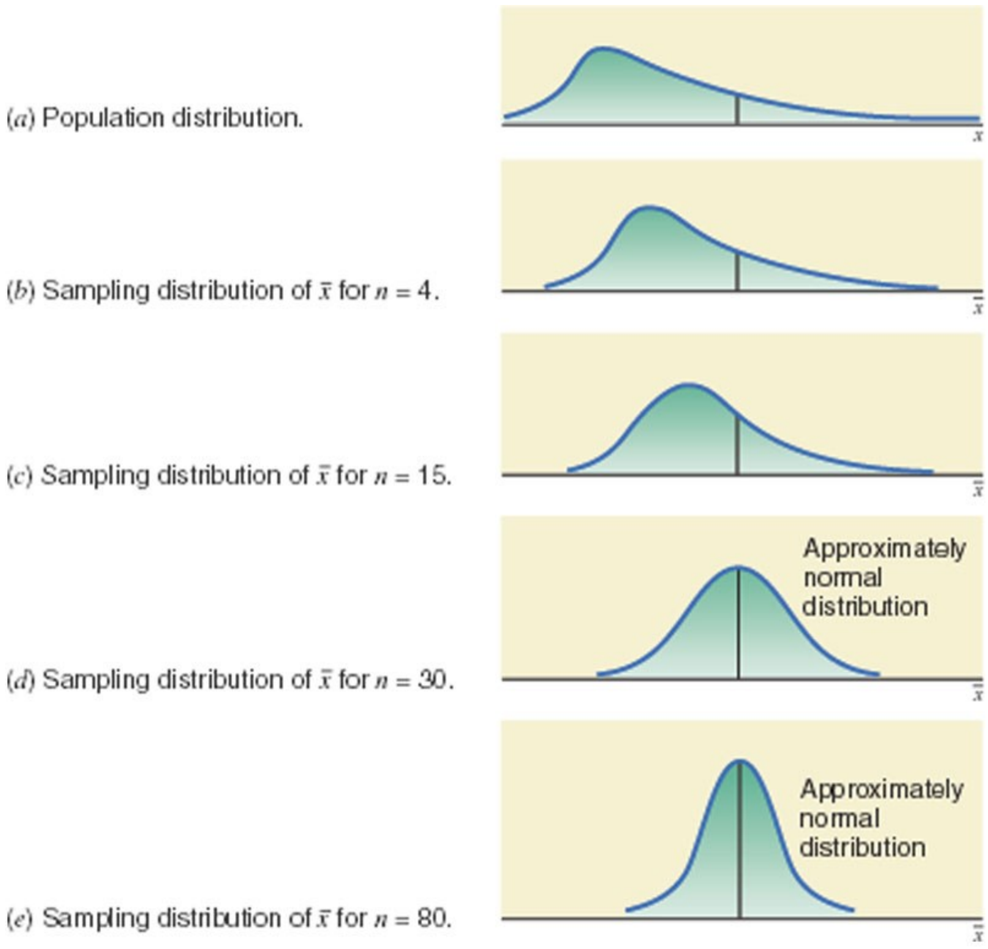
Centralna granična teorema

Prema ***centralnoj graničnoj teoremi***, za veliki uzorak, uzoračka raspodjela statistike $\overline{x}$ je približno normalna , bez obzira na oblik raspodjele osnovnog skupa. Aritmetička sredina i standardna devijacija uzoračke raspodjele statistike $\overline{x}$ su

$$\mu_{\bar{x}} = \mu \quad i \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Uzorci se najčešće smatraju velikim kada je $n \geq 30$.

Slika 7.6 Raspodjela osnovnog skupa i uzoračka raspodjela statistike $\bar{x}$.



Standardizovana promjenljiva z za $\bar{x}$

Standardizovana normalna promjenljiva z za promjenljivu $\bar{x}$ računa se kao

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

7.6 PROPORCIJA OSNOVNOG SKUPA I UZORKA

Proporcija osnovnog skupa i proporcija uzorka, označavaju se sa p i $\hat{p}$, respektivno, i računaju se

$$p = \frac{X}{N} \quad i \quad \hat{p} = \frac{x}{n}$$

PROPORCIJA OSNOVNOG SKUPA I UZORKA

gdje je

- N = ukupan broj elemenata u osnovnom skupu
- n = ukupan broj elemenata u uzorku
- X = broj elemenata u osnovnom skupu sa određenom karakteristikom
- x = broj elemenata u uzorku sa određenom karakteristikom

Primjer 7-7

Pretpostavimo da u jednom gradu živi 789654 porodica, od kojih 563282 ima kuću u svom vlasništvu. Izabran je slučajan uzorak od 240 porodica, od kojih 158 posjeduje kuću. Odrediti proporciju porodica koje posjeduju kuću osnovnom skupu i u uzorku.

Primjer 7-7: Rješenje

$$p = \frac{X}{N} = \frac{563{,}282}{789{,}654} = .71$$

$$\hat{p} = \frac{x}{n} = \frac{158}{240} = .66$$

7.7 ARITMETIČKA SREDINA, STANDARDNA DEVIJACIJA I OBLIK UZORAČKE RASPODJELE STATISTIKE $\hat{p}$

- Uzoračka raspodjela statistike $\hat{p}$
- Aritmetička sredina i standardna devijacija statistike $\hat{p}$
- Oblik uzoračke raspodjele $\hat{p}$

Uzoračka raspodjela uzoračke proporcije $\hat{p}$

Definicija

Raspodjela vjerovatnoća proporcije uzorka, $\hat{p}$, naziva se ***uzoračkom raspodjelom proporcije*** i predstavlja parove vrijednosti koje može uzeti statistika $\hat{p}$ i odgovarajućih vjerovatnoća.

Aritmetička sredina i standardna devijacija statistike $\hat{p}$

Aritmetička sredina proporcija uzoraka

Aritmetička sredina proporcija uzoraka, $\hat{p}$, označava se sa $\mu_{\hat{p}}$ i jednaka je proporciji osnovnog skupa, p. Dakle,

$$\mu_{\hat{p}} = p$$

Aritmetička sredina i standardna devijacija statistike $\hat{p}$

Standardna devijacija proporcija uzoraka

Standardna devijacija proporcija uzoraka, $\hat{p}$, označava se sa $\sigma_{\hat{p}}$ i dobija na osnovu formule

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

gdje je p proporcija osnovnog skupa, $q = 1 - p$, a n je veličina uzorka. Ova formula se koristi kada je $n/N \leq 0.05$, gdje je N veličina osnovnog skupa.

Aritmetička sredina i standardna devijacija statistike $\hat{p}$

Ako je $n/N > 0.05$, tada $\sigma_{\hat{p}}$ se računa kao:

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$$

gdje se faktor $\sqrt{\frac{N-n}{N-1}}$ naziva popravnim faktorom za konačne skupove.

Oblik uzoračke raspodjele statistike $\hat{p}$

Centralna granična teorema za proporciju uzorka

Prema centralnoj graničnoj teoremi, uzoračka raspodjela statistike $\hat{p}$ je približno normalna ako se radi o velikim uzorcima. Kada je riječ o proporciji, u praksi se smatra da su uzorci dovoljno veliki ako je np i nq veće od 5 – odnosno, ako važi

$$np > 5 \quad \text{i} \quad nq > 5$$

Standardizovana normalna promjenljiva z za promjenljivu $\hat{p}$

Standardizovana normalna promjenljiva z za promjenljivu $\hat{p}$ računa se

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}}$$