3.2 MJERE DISPERZIJE NEGRUPISANIH PODATAKA

- Interval varijacije (razmak varijacije)
- Varijansa i standardna devijacija
- Parametri osnovnog skupa i statistike uzorka

Interval varijacije

Određivanje intervala varijacije za negrupisane podatke

Interval varijacije = Najveća vrijednost – Najmanja vrijednost

Primjer 3-11

U tabeli 3.4 su date ukupne površine u kvadratnim miljama četiri države centralnog juga u SAD.

Odrediti interval varijacije ove serije podataka.

Tabela 3.4

Država	Ukupna površina (u kvadratnim miljama)
Arkanzas	53,182
Lujzijana	49,651
Oklahoma	69,903
Teksas	267,277

Primjer 3-11: Rješenje

Interval varijacije = Najveća vrijednost – Najmanja vrijednost

= 267,277 – 49,651

= **217,626 kvadratnih milja**

Dakle, ukupne površine ove četiri države su raspršene u razmaku od 217,626 kvadratnih milja.

Interval varijacije

Nedostaci

- Interval varijacije, kao i aritmetička sredina, ima nedostatak što na njega utiču ekstremne vrijednosti. Prema tome, interval varijacije nije dobra mjera disperzije za seriju podataka koja sadrži ekstremne vrijednosti.
- Za njegovo izračunavanje koriste se samo dvije vrijednosti: najveća i najmanja. Sve ostale vrijednosti serije podataka se zanemaruju. Stoga, interval varijacija nije zadovoljavajuća mjera disperzije.

Varijansa i standardna devijacija

- Standardna devijacija je najčešće korišćena mjera disperzije.

- Vrijednost standardne devijacije pokazuje koliko blizu su vrijednosti serije podataka grupisane oko aritmetičke sredine.

- Uopšteno, manja vrijednost standardne devijacije serije podataka ukazuje da su vrijednosti te serije raspršene veoma malo oko aritmetičke sredine.

Varijansa i standardna devijacija

- Nasuprot tome, veća vrijednost standardne devijacije serije podataka ukazuje da su vrijednosti te serije raspršene u relativno velikom razmaku oko aritmetičke sredine.

- Standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena varijanse.

Varijansa i standardna devijacija

- Varijansa osnovnog skupa se označava sa σ^2 (čitamo kao sigma na kvadrat), a varijansa uzorka se označava sa s^2.

- U skladu sa tim, standardna devijacija osnovnog skupa se označava sa σ, a standardna devijacija uzorka se označava sa s.

Varijansa i standardna devijacija

Osnovne formule za varijansu i standardnu devijaciju negrupisanih podataka

$$\sigma^2 = \frac{\sum (x-\mu)^2}{N} \quad \text{i} \quad s^2 = \frac{\sum \left(x-\overline{x}\right)^2}{n-1}$$

$$\sigma = \sqrt{\frac{\sum (x-\mu)^2}{N}} \quad \text{i} \quad s = \sqrt{\frac{\sum \left(x-\overline{x}\right)^2}{n-1}}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, σ je standardna devijacija osnovnog skupa, i s je standardna devijacija uzorka.

Tabela 3.5

x	$x - \bar{x}$
82	$82 - 84 = -2$
95	$95 - 84 = +11$
67	$67 - 84 = -17$
92	$92 - 84 = +8$
	$\Sigma(x - \bar{x}) = 0$

Varijansa i standardna devijacija

Radne formule za varijansu i standardnu devijaciju negrupisanih podataka

$$\sigma^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N} \quad \text{i} \quad s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}$$

$$\sigma = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N}} \quad \text{i} \quad s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, σ je standardna devijacija osnovnog skupa, i s je standardna devijacija uzorka.

Primjer 3-12

Do 2009. godine, putnicima aviokompanija se nije naplaćivao prijavljeni prtljag. Međutim, negdje oko 2009. godine, mnoge američke aviokompanije počele su da naplaćuju naknadu za torbe. Prema Zavodu za statistiku transporta, američke aviokompanije su u 2010. godini prikupile više od 3 milijarde dolara prihoda od naknada za prtljag. Sledeća tabela navodi prihode od ovih naknada za šest američkih aviokompanija za 2010. godinu.

Odrediti varijansu i standardnu devijaciju za ove podatke.

Primjer 3-12

Aviokompanija	Prihodi od naknada za prtljag (u milionima dolara)
United	313
Continental	342
American	581
Delta	952
US Airways	514
Air Tran	152

Primjer 3-12: Rješenje

Označimo sa x prihod od naknade prtljaga (u milionima dolara) jedne aviokompanije. Vrijednosti Σx i Σx^2 su izračunate u Tabeli 3.6.

x	x^2
313	97,969
342	116,964
581	337,561
952	906,304
514	264,196
152	23,104
$\Sigma x = 2854$	$\Sigma x^2 = 1{,}746{,}098$

Primjer 3-12: Rješenje

Korak 1. Izračunati Σx

Zbir vrijednosti prve kolone u Tabeli 3.6 daje 2,854.

Korak 2. Naći Σx^2

Rezultati ovog koraka su prikazani u drugoj koloni Tabele 3.6, što daje sumu 1,746,098.

Primjer 3-12: Rješenje

Korak 3. Odrediti varijansu

$$s^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1} = \frac{1,746,098 - \frac{(2,854)^2}{6}}{6-1}$$

$$= \frac{1,746,098 - 1,357,552.667}{5}$$

$$= 77,709.06666$$

Primjer 3-12: Rješenje

Korak 4. Izračunati standardnu devijaciju

Standardna devijacija se računa uzimanjem (pozitivnog) kvadratnog korijena varijanse:

$$s = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{77,709.06666}$$

$$= 278.7634601 = \$278.76 million$$

Dakle, standardna devijacija prihoda od naknada za prtljag ovih šest aviokompanija u 2010. godini je **$278.76 miliona**.

Dvije napomene

1. Vrijednosti varijanse i standardne devijacije nisu nikada negativne.

2. Jedinice mjere varijanse uvijek predstavljaju kvadrirane jedinice mjere originalnih podataka.

Primjer 3-13

Date su bruto zarade (u hiljadama dolara) za svih šest zaposlenih u jednoj manjoj firmi u 2011. godini.

88.50   108.40   65.50   52.50   79.80   54.60

Izračunati varijansu i standardnu devijaciju za ove podatke.

Primjer 3-13: Rješenje

Neka x predstavlja bruto zarade jednog zaposlenog ove firme u 2011. godini. Vrijednosti $\sum x$ i $\sum x^2$ su izračunate u Tabeli 3.7.

x	x^2
88.50	7832.25
108.40	11,750.56
65.50	4290.25
52.50	2756.25
79.80	6368.04
54.60	2981.16
$\Sigma x = 449.30$	$\Sigma x^2 = 35{,}978.51$

Primjer 3-13: Rješenje

$$\sigma^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{N}}{N} = \frac{35{,}978.51 - \frac{(449.30)^2}{6}}{6} = 388.90$$

$$\sigma = \sqrt{388.90} = \$19.721 \text{ hiljada} = \$19{,}721$$

Dakle, standardna devijacija bruto zarada svih šest zaposlenih ove firme u 2011. godini je **$19,721**.

Upozorenje

Primijetimo da $\sum x^2$ nije isto što i $(\sum x)^2$. Vrijednost $\sum x^2$ se dobija sabiranjem kvadriranih vrijednosti x. Vrijednost $(\sum x)^2$ se dobija tako što se kvadrira vrijednost $\sum x$.

Parametri osnovnog skupa i statistike uzorka

- Numerička mjera kao što je aritmetička sredina, medijana, modus, interval varijacije, varijansa, ili standardna devijacija izračunata za podatke osnovnog skupa naziva se parametar skupa, ili jednostavno **parametar**.

- Deskriptivna mjera koja se računa za podatke uzorka naziva se statistika uzorka, ili jednostavno **statistika**.

3.3 ARITMETIČKA SREDINA, VARIJANSA I STANDARDNA DEVIJACIJA ZA GRUPISANE PODATKE

- Aritmetička sredina grupisanih podataka
- Varijansa i standardna devijacija grupisanih podataka

Aritmetička sredina grupisanih podataka

Izračunavanje aritmetičke sredine grupisanih podataka

Aritmetička sredina skupa: $\mu = \frac{\sum mf}{N}$

Aritm. sredina uzorka: $\bar{x} = \frac{\sum mf}{n}$

gdje je m sredina a f frekvencija intervala.

Primjer 3-14

U tabeli 3.8 prikazana je raspodjela frekvencija vremena svakodnevno provedenog u putovanju od kuće do radnog mjesta *svih* 25 zaposlenih u jednom preduzeću (u minutima) .

Izračunati aritmetičku sredinu vremena provedenog u putovanju od kuće do radnog mjesta.

Primjer 3-14

Dnevno vrijeme putovanja na posao (u minutima)	Broj zaposlenih
0 do manje od 10	4
10 do manje od 20	9
20 do manje od 30	6
30 do manje od 40	4
40 do manje od 50	2

Primjer 3-14: Rješenje

Dnevno vrijeme putovanja na posao (u minutima)	f	m	mf
0 do manje od 10	4	5	20
10 do manje od 20	9	15	135
20 do manje od 30	6	25	150
30 do manje od 40	4	35	140
40 do manje od 50	2	45	90
	N=25		$\sum mf$=535

Primjer 3-14: Rješenje

$$\mu = \frac{\sum mf}{N} = \frac{535}{25} = \mathbf{21.40\ minutes}$$

Dakle, zaposleni ovog preduzeća u prosjeku provedu **21.40 minuta** dnevno putujući od kuće do radnog mjesta.

Primjer 3-15

Tabela 3.9 sadrži raspodjelu frekvencija broja porudžbina koje su pristizale svakog dana, tokom proteklih 50 dana, u kancelariju jednog preduzeća koje se bavi porudžbinama poštom.

Izračunati prosjek.

Primjer 3-15

Broj porudžbina	Broj dana
10-12	4
13-15	12
16-18	20
19-21	14

Primjer 3-15: Rješenje

Broj porudžbina	*f*	*m*	*mf*
10-12	4	11	44
13-15	12	14	168
16-18	20	17	340
19-21	14	20	280
	n=50		$\sum mf$=832

Primjer 3-15: Rješenje

$$\bar{x} = \frac{\sum mf}{n} = \frac{832}{50} = \mathbf{16.64\ orders}$$

Znači, ovo preduzeće je u periodu od 50 dana u prosjeku primalo **16.64 porudžbina** dnevno.

Varijansa i standardna devijacija grupisanih podataka

Osnovne formule za varijansu i standardnu devijaciju grupisanih podataka

$$\sigma^2 = \frac{\sum f(m-\mu)^2}{N} \quad \text{i} \quad s^2 = \frac{\sum f(m-\bar{x})^2}{n-1}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, a m sredina intervala. U oba slučaja, standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena iz varijanse.

Varijansa i standardna devijacija grupisanih podataka

Radne formule za varijansu i standardnu devijaciju grupisanih podataka

$$\sigma^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{N}}{N} \quad i \quad s^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1}$$

gdje je σ^2 varijansa osnovnog skupa, s^2 je varijansa uzorka, a m sredina intervala.

Varijansa i standardna devijacija grupisanih podataka

Radne formule za varijansu i standardnu devijaciju grupisanih podataka

Standardna devijacija se dobija uzimanjem pozitivnog kvadratnog korijena varijanse.

Standardna devijacija skupa: $\sigma = \sqrt{\sigma^2}$

Standardna devijacija uzorka: $s = \sqrt{s^2}$

Primjer 3-16

Sledeći podaci, koji su preuzeti iz Tabele 3.8 primjera 3-14, predstavljaju raspodjelu svakodnevnog vremena provedenog u putovanju od kuće do posla (u minutima) za svih 25 zaposlenih jednog preduzeća.

Izračunati varijansu i standardnu devijaciju.

Primjer 3-16

Dnevno vrijeme putovanja na posao (u minutima)	Broj zaposlenih
0 do manje od 10	4
10 do manje od 20	9
20 do manje od 30	6
30 do manje od 40	4
40 do manje od 50	2

Primjer 3-16: Rješenje

Dnevno vrijeme putovanja na posao (u minutima)	f	m	mf	m^2f
0 do manje od 10	4	5	20	100
10 do manje od 20	9	15	135	2025
20 do manje od 30	6	25	150	3750
30 do manje od 40	4	35	140	4900
40 do manje od 50	2	45	90	4050
	N=25		$\sum mf$=535	$\sum m^2f$=14,825

Primjer 3-16: Rješenje

$$\sigma^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{N}}{N} = \frac{14{,}825 - \frac{(535)^2}{25}}{25} = \frac{3376}{25} = 135.04$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{135.04} = 11.62 \text{ minutes}$$

Prema tome, standardna devijacija vremena provedenog putujući na posao ovih zaposlenih je **11.62 minuta**.

Primjer 3-17

Sledeći podaci, preuzeti iz Tabele 3.9 primjera 3-15, predstavljaju raspodjelu frekvencija broja porudžbina koje su pristizale svakoga dana, tokom proteklih 50 dana u kancelariju jednog preduzeća koje se bavi porudžbinama poštom.

Izračunati varijansu i standardnu devijaciju.

Primjer 3-17

Broj porudžbina	f
10-12	4
13-15	12
16-18	20
19-21	14

Primjer 3-17: Rješenje

Broj porudžbina	f	m	mf	m^2f
10-12	4	11	44	484
13-15	12	14	168	2352
16-18	20	17	340	5780
19-21	14	20	280	5600
	n=50		$\sum mf$=832	$\sum m^2f$=14,216

Primjer 3-17: Rješenje

$$s^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1} = \frac{14{,}216 - \frac{(832)^2}{50}}{50-1} = 7.5820$$

$$s = \sqrt{s^2} = \sqrt{7.5820} = 2.75 \text{ orders}$$

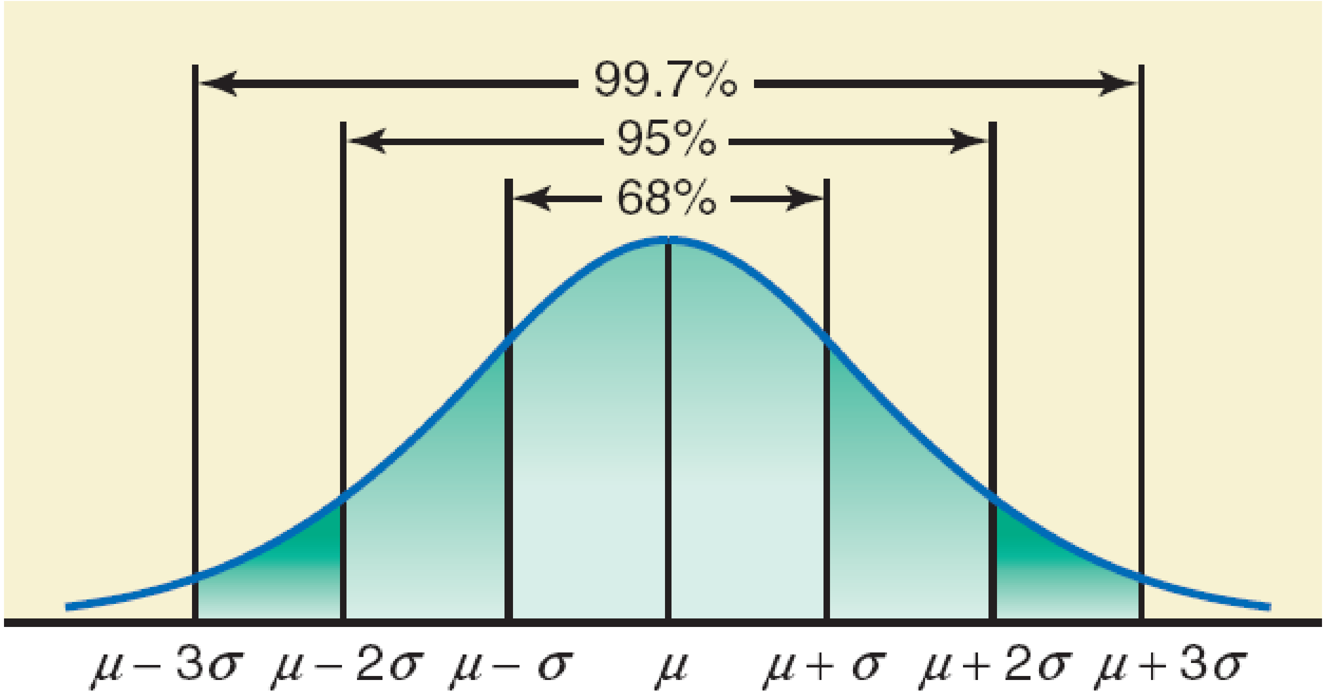
Dakle, standardna devijacija broja porudžbina koje su pristizale u kancelariju ovog preduzeća tokom poslednjih 50 dana je **2.75**.

Empirijsko pravilo

Za raspodjelu u obliku zvona, približno

1. 68% vrijednosti se nalazi u opsegu jedne standardne devijacije od aritmetičke sredine
2. 95% vrijednosti se nalazi u opsegu dvije standardne devijacije od aritmetičke sredine
3. 99.7% vrijednosti se nalazi u opsegu tri standardne devijacije od aritmetičke sredine

Slika 3.9 Ilustracija empirijskog pravila.



POGLAVLJE 4

VJEROVATNOĆA

4.1 EKSPERIMENT, ISHODI I PROSTOR UZORKA

Definicija

Eksperiment je proces, čiji rezultat izvođenja je jedna i samo jedna od mnogih opservacija. Te opservacije se nazivaju ***ishodi*** eksperimenta. Skup svih ishoda jednog eksperimenta se naziva ***prostor uzorka***.

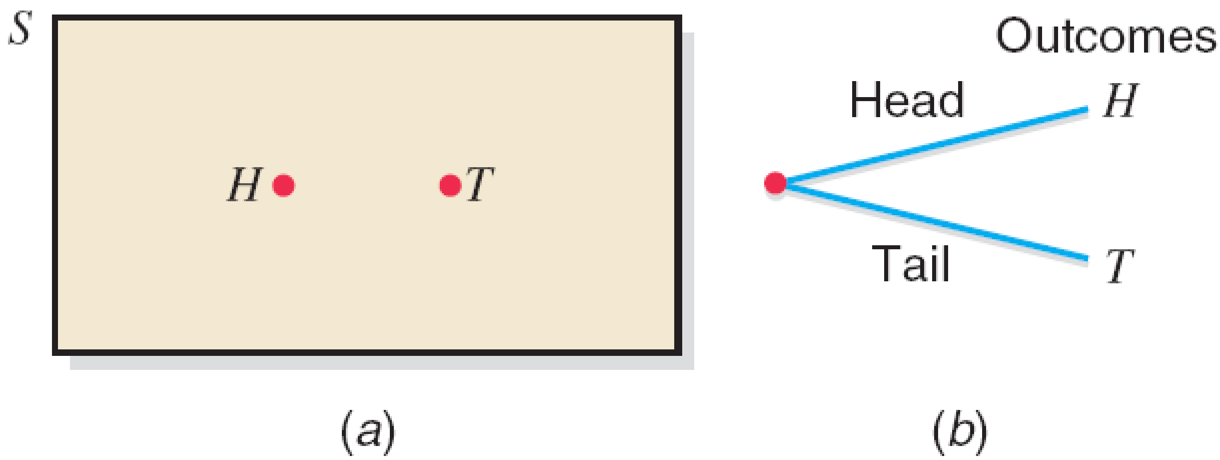
Tabela 4.1 Primjeri eksperimenata, ishoda i prostora uzorka

Eksperiment	Ishodi	Prostor uzorka
Bacanje novčića jednom	Pismo, glava	S = {pismo, glava}
Bacanje kocke jednom	1, 2, 3, 4, 5, 6	S = {1, 2, 3, 4, 5, 6}
Bacanje novčića dva puta	*GG, GP, PG, PP*	S = { *GG, GP, PG, PP*}
Igranje lutrije	dobiti, izgubiti	S = { dobiti, izgubiti}
Polaganje testa	proći, pasti	S = { proći, pasti}
Biranje studenta	muško, žensko	S = { muško, žensko}

Primjer 4-1

Nacrtati Venov dijagram i stablo ishoda za eksperiment bacanja novčića jedanput.

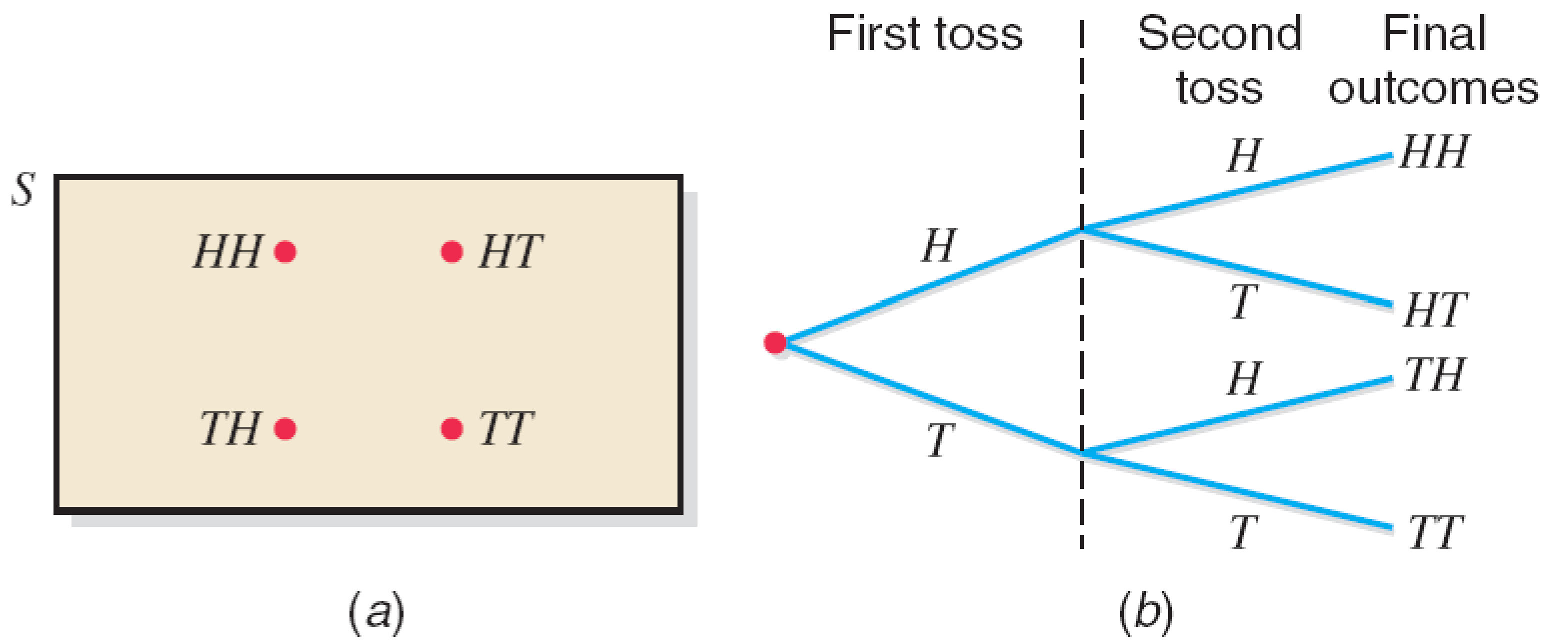
Slika 4.1 (a) Venov dijagram i (b) stablo ishoda za jedno bacanje novčića.



Primjer 4-2

Nacrtati Venov dijagram i stablo ishoda za eksperiment bacanja novčića dva puta.

Slika 4.2 (a) Venov dijagram i (b) stablo ishoda za dva bacanja novčića.



Elementarni i složeni događaji

Definicija

Događaj je skup jednog ili više ishoda eksperimenta.

Događaj koji sadrži jedan i samo jedan od (krajnjih) ishoda nekog eksperimenta se naziva ***elementarni događaj*** i obično obilježava sa E_i.

Složeni događaj je skup više od jednog ishoda nekog eksperimenta.

Primjer 4-3

Razmotrimo primjer o biranju dva zaposlena iz jedne firme i bilježenju svaki put da li je izabrani zaposleni muško ili žensko. Svaki od četiri krajnja ishoda (*MM, MW, WM* i *WW*) za ovaj eksperiment je elementarni događaj. Ta četiri događaja mogu da se obilježe sa E_1, E_2, E_3, i E_4, redom. Dakle,

$$\boldsymbol{E_1} = (MM),\ \ \boldsymbol{E_2} = (MW),\ \ \boldsymbol{E_3} = (WM),\ \text{ i }\ \boldsymbol{E_4} = (WW)$$

Primjer 4-4

Razmotrimo ponovo primjer 4-3 o biranju dva zaposlena iz jedne firme i bilježenju svaki put da li je izabrani zaposleni muško ili žensko. Neka je *A* događaj da je izabran najviše jedan muškarac. Događaj *A* će se realizovati ako je izabran jedan ili nijedan muškarac. Dakle, događaj *A* je

$$A = \{MW, WM, WW\}$$

Pošto događaj *A* sadrži više od jednog ishoda, on je složeni događaj. Venov dijagram na slici 4.3 daje grafički prikaz složenog događaja *A*.

Slika 4.3 Venov dijagram za događaj *A*.

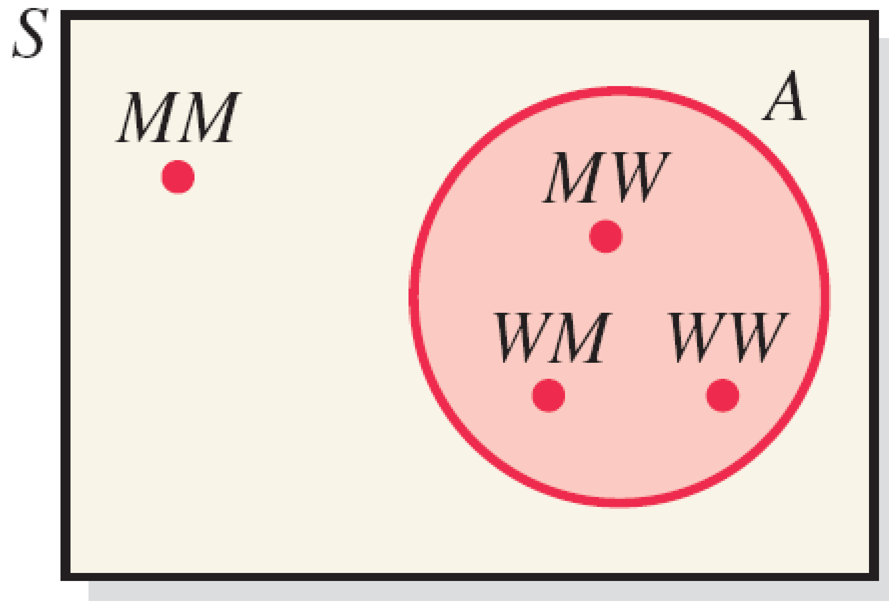


Figure 4.4 Venn diagram for event *A*.