

Биостатистика и анализа података

Марија Цупарић

Други час



Друге корисне мере су мере расејања које показују расутост унутар популације

- ★ распон популације
- ★ дисперзија популације σ^2
- ★ стандардно одступање (девијација) популације σ
- ★ међуквартилно растојање



- ★ Распон је разлика између највеће и најмање вредности X —а када се посматра цела популација.
- ★ Узорачки распон је разлика између највећег и најмањег елемента узорка, тј. $R = X_{(n)} - X_{(1)}$, где су $X_{(1)}$ и $X_{(n)}$ прва и последња статистика поретка.
- ★ Ова величина није нешто посебно добра мера, јер њеним познавањем не добијамо неке посебно корисне информације о популацији.



```
diff(range(x))
```



Пример. Резултати студената на испиту у два семестра

	први семестар	други семестар
обим узорка	23	26
средњи број поена $\bar{x}$	75	75
медијана број поена	75	75
распон	50 (од 50 до 100)	50 (од 50 до 100)

Стварна расподела поена

први семестар	други семестар
50 50 50 50 50 50	50
60 60	65 65
70 70	70 70 70
75	74 74 74 74
80 80	75 75 75 75 75 75
85 85 85	76 76 76 76
100 100 100 100 100 100	80 80 80
	85 85
	100

- ★ Средње квадратно одступање случајне величине X од свог очекивања μ .
- ★ Природна оцена дисперзије је узорачка дисперзија

$$\overline{S}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X})^2$$

или поправљена узорачка дисперзија

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})^2.$$



`var(x) #daje popravljenu uzoracku disperziju`

Рачунски једноставнија формула:

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n X_i^2 - \frac{n}{n-1} \overline{X}^2.$$

Пример. Подаци о дужини трајања телефонских разговора:
 10 20 6 12 15 8 4 9 3 13

Тада је $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 10$. Према томе

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{(10-10)^2 + \dots + (12-10)^2}{9}$$

или, како је $\sum_{i=1}^n x_i^2 = 1244$,

$$s_n^2 = \frac{1}{9} \cdot 1244 - \frac{10}{9} \cdot 10^2 = 27.11.$$



- ★ Квадратни корен из дисперзије
- ★ Процењујемо га узорачким стандардним одступањем, тј. $s = \sqrt{s^2}$

R `sd(x)`

Пример (наставак).

$$s = \sqrt{27.11} = 5.21.$$



Међуквартилно растојање

- ★ Разлика између трећег и првог квантила.

први квантил Q_1 - вредност од које је $1/4$ популације мање, а $3/4$ веће

други квантил Q_2 - вредност од које је $2/4$ популације мање

трећи квантил Q_3 - вредност од које је $3/4$ популације мање, а $1/4$ веће

- ★ Узорачки α -квантил се добија тако што се у варијационом низу пронађе онај елемент од којег је $\alpha \cdot 100\%$ елемената мање (или једнако).
- ★ Ако је $\alpha = 0.5$ онда је одговарајући квантил медијана, ако је $\alpha = 0.25$ онда је први квантил, а ако је $\alpha = 0.75$ онда је трећи квантил.

$$IQR = q_3 - q_1$$

R

```
quantile(x)
IQR(x)
```


Идентификација аутлајера

- ★ Аутлајер је члан узорка који се не уклапа у постојећи статистички модел, односно одступа од осталих чланова узорка.
- ★ Треба испитати који је њихов узрок и какав је њихов утицај на модел.
- ★ Ако је њихов узрок грешка, онда их треба избацити из узорка.
- ★ Најједноставнији начин за детекцију аутлајера је кутијаста дијаграм (boxplot).



Конструкција куријастог дијаграма:

- (1) нацртати хоризонталну референтну скалу и уцртати први q_1 и трећи q_3 узорачки квартил;
- (2) уцртати $f_1 = q_1 - 1.5/IQR$ и $f_3 = q_3 + 1.5/IQR$;
- (3) уцртати тачке a_1 (прва вредност из узорка већа од f_1) и a_3 (последња вредност из узорка мања од f_3);
- (4) уцртати вредности $F_1 = q_1 - 2 \cdot 1.5 \cdot IQR$ и $F_3 = q_3 + 2 \cdot 1.5 \cdot IQR$;
- (5) нацртати правоугаоник са крајњим тачкама q_1 и q_3 , унутар којег је медијана означена линијом;
- (6) тачке a_1 и a_3 повезати са правоугаоником;
- (7) ако има тачака из узорка које се налазе између F_1 и f_1 , односно f_3 и F_3 треба их уцртати као празне кружиће (благитлајери), а ако има тачака из узорка које се налазе пре F_1 , односно после F_3 треба их уцртати као пуне кружиће (правитлајери).

Ако се медијана налази на средини правоугаоника, онда је расподела симетрична у односу на њу.



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

$$(2) IQR = q_3 - q_1 = 47 - 32 = 15$$

$$f_1 = q_1 - 1.5 IQR = 32 - 1.5 \cdot 15 = 9.5$$

$$f_3 = q_3 + 1.5 IQR = 47 + 1.5 \cdot 15 = 69.5$$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

$$(2) IQR = q_3 - q_1 = 47 - 32 = 15$$

$$f_1 = q_1 - 1.5 \cdot IQR = 32 - 1.5 \cdot 15 = 9.5$$

$$f_3 = q_3 + 1.5 \cdot IQR = 47 + 1.5 \cdot 15 = 69.5$$

$$(3) a_1 = 12, a_3 = 61$$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

(2) $IQR = q_3 - q_1 = 47 - 32 = 15$

$$f_1 = q_1 - 1.5 \cdot IQR = 32 - 1.5 \cdot 15 = 9.5$$

$$f_3 = q_3 + 1.5 \cdot IQR = 47 + 1.5 \cdot 15 = 69.5$$

(3) $a_1 = 12, a_3 = 61$

(4) $F_1 = q_1 - 2 \cdot 1.5 \cdot IQR = 32 - 3 \cdot 15 = -13,$

$$F_3 = q_3 + 2 \cdot 1.5 \cdot IQR = 47 + 3 \cdot 15 = 92$$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

(2) $IQR = q_3 - q_1 = 47 - 32 = 15$

$$f_1 = q_1 - 1.5 \cdot IQR = 32 - 1.5 \cdot 15 = 9.5$$

$$f_3 = q_3 + 1.5 \cdot IQR = 47 + 1.5 \cdot 15 = 69.5$$

(3) $a_1 = 12, a_3 = 61$

(4) $F_1 = q_1 - 2 \cdot 1.5 \cdot IQR = 32 - 3 \cdot 15 = -13,$

$$F_3 = q_3 + 2 \cdot 1.5 \cdot IQR = 47 + 3 \cdot 15 = 92$$

(5) $m_e = x_{(11)} = 40$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

(2) $IQR = q_3 - q_1 = 47 - 32 = 15$

$$f_1 = q_1 - 1.5 \cdot IQR = 32 - 1.5 \cdot 15 = 9.5$$

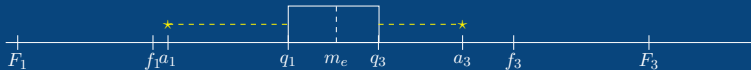
$$f_3 = q_3 + 1.5 \cdot IQR = 47 + 1.5 \cdot 15 = 69.5$$

(3) $a_1 = 12, a_3 = 61$

(4) $F_1 = q_1 - 2 \cdot 1.5 \cdot IQR = 32 - 3 \cdot 15 = -13,$

$$F_3 = q_3 + 2 \cdot 1.5 \cdot IQR = 47 + 3 \cdot 15 = 92$$

(5) $m_e = x_{(11)} = 40$



Пример. Проучавана је дужина хоспитализације након повреде главе и добијени су резултати: 8, 12, 20, 27, 30, 32, 35, 36, 40, 40, 40, 40, 41, 42, 45, 47, 50, 52, 61, 89, 108. Нацртати кутијасти дијаграм ових података.

- (1) Први узорачки квартил је на позицији $\frac{\text{положај од } m_e + 1}{2} = 6$ од почетка а трећи узорачки квартил је на позицији 6 од краја низа. Односно

$$q_1 = x_{(6)} = 32, \quad q_3 = x_{(16)} = 47$$

(2) $IQR = q_3 - q_1 = 47 - 32 = 15$

$$f_1 = q_1 - 1.5 \cdot IQR = 32 - 1.5 \cdot 15 = 9.5$$

$$f_3 = q_3 + 1.5 \cdot IQR = 47 + 1.5 \cdot 15 = 69.5$$

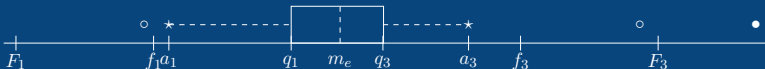
(3) $a_1 = 12, a_3 = 61$

(4) $F_1 = q_1 - 2 \cdot 1.5 \cdot IQR = 32 - 3 \cdot 15 = -13,$

$$F_3 = q_3 + 2 \cdot 1.5 \cdot IQR = 47 + 3 \cdot 15 = 92$$

(5) $m_e = x_{(11)} = 40$

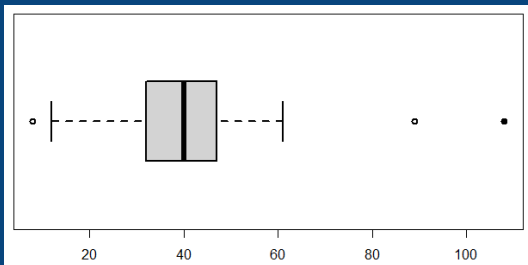




```
x<-c(8,12,20,27,30,32,35,36,40,40,40,40,41,42,45,47,50,52,61,89,
108)
```



```
b <- boxplot(x,horizontal = TRUE, lwd=2)
points(x[21],order(x[21]),pch=19,col="black",lwd=2)
b$stats
```



	[,1]
[1,]	12
[2,]	32
[3,]	40
[4,]	47
[5,]	61

← a_1
← q_1
← m_e
← q_3
← a_3



```
summary(x)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
8.00	32.00	40.00	42.62	47.00	108.00

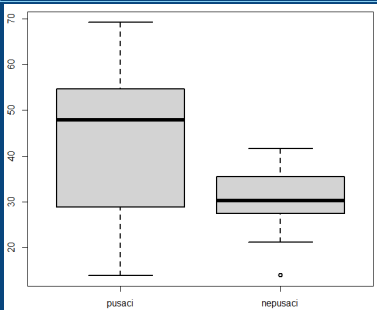
Напомена: Формуле за први и трећи узорачки квартил се не поклапају увек са формулама које користи функција `quantile()`

n	k	q_1	m_2	q_3
$2k + 1$	непаран	$\frac{X_{(\frac{k+1}{2})} + X_{(\frac{k+3}{2})}}{2}$	$X_{(k+1)}$	$\frac{X_{(\frac{3k+1}{2})} + X_{(\frac{3k+3}{2})}}{2}$
$2k + 1$	паран	$X_{(\frac{k}{2}+1)}$	$X_{(k+1)}$	$X_{(\frac{3k+2}{2})}$
$2k$	непаран	$X_{(\frac{k+1}{2})}$	$\frac{X_{(k)} + X_{(k+1)}}{2}$	$X_{(\frac{3k+1}{2})}$
		$\frac{3X_{(\frac{k}{2})} + X_{(\frac{k}{2}+1)}}{4}$		$\frac{X_{(\frac{3k}{2})} + 3X_{(\frac{3k+2}{2})}}{4}$
$2k$	паран	$\frac{X_{(\frac{k}{2})} + X_{(\frac{k}{2}+1)}}{2}$	$\frac{X_{(k)} + X_{(k+1)}}{2}$	$\frac{X_{(\frac{3k}{2})} + X_{(\frac{3k+2}{2})}}{2}$
		$\frac{X_{(\frac{k}{2})} + 3X_{(\frac{k}{2}+1)}}{4}$		$\frac{3X_{(\frac{3k}{2})} + X_{(\frac{3k+2}{2})}}{4}$

Пример. Проучаван је утицај пушења на спавање. Подаци представљају време у минутима које је било потребно да испитаници заспе. Упоредити резултате графички.



```
x<-c(69.3, 56.0, 22.1, 47.6, 53.2, 48.1, 52.7, 34.4, 60.2, 43.8, 23.2,
     13.8)
y<-c(28.6, 25.1, 26.4, 34.9, 29.8, 28.4, 38.5, 30.2, 30.6, 31.8, 41.6,
     21.1, 36.0, 37.9, 13.9)
b1<-boxplot(x,y, names=c("pusaci", "nepusaci"))
b1$stats
```



	[,1]	[,2]
[1,]	13.80	21.10
[2,]	28.80	27.40
[3,]	47.85	30.20
[4,]	54.60	35.45
[5,]	69.30	41.60

← a_1
← q_1
← m_e
← q_3
← a_3

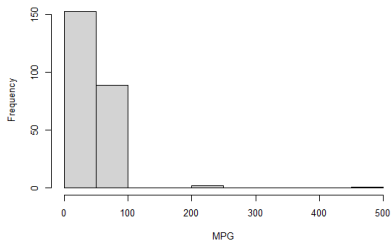
Пример. Размотримо базу података `topgear` у оквиру пакета `crmReg` и утицај аутлајера на променљиву `MPG` (потрошња горива).



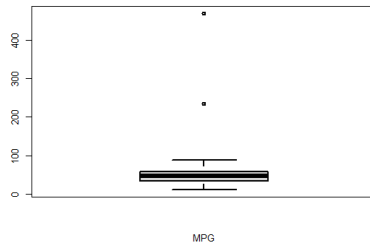
```
hist(topgear$MPG,main="Petrošnja goriva", xlab="MPG")  
boxplot(topgear$MPG,main="Petrošnja goriva", xlab="MPG",lwd = 2)  
hist(sort(topgear$MPG)[1:242],main="Petrošnja goriva bez autlajera",  
      xlab="MPG")  
boxplot(sort(topgear$MPG)[1:242],main="Petrošnja goriva bez  
      autlajera", xlab="MPG",lwd = 2)
```



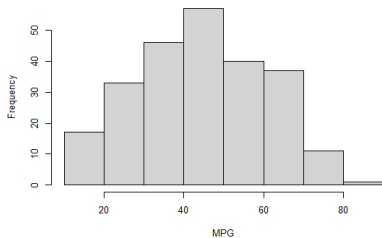
Potrošnja goriva



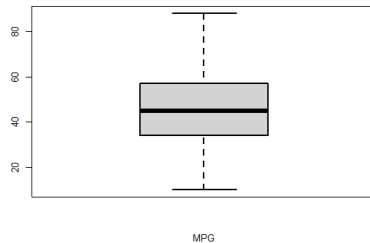
Potrošnja goriva



Potrošnja goriva bez autajera



Potrošnja goriva bez autajera



Да ли знате нешто о вероватноћи?

- ★ „40% је шансе за кишу”
- ★ Шта тај проценат значи?

И без формалног знања вероватноће, већина људи интуитивно зна како да интерпретира ове бројеве.

- ★ број близу 0 или близу 0%, догађај о којем је реч не очекијемо да ће се десити
- ★ бројеви близу 1 или 100% показују да се очекује да ће се догађај десити
- ★ Шта треба радити ако је број 0.5 или 50%?

Догађаји који имају вероватноћу близу 0 не значи да се неће десити, само значи да су такви догађаји ретки. Слично, они догађаји који имају вероватноћу близу 1 не значи да ће се сигурно десити, већ да се такви догађаји често дешавају.



Главно питање је како се догађајима додељују вероватноће.

Случајни експеримент подразумева комплексе услова који се могу неограничено много пута понављати, али који не доводе увек до једнозначно одређеног исхода

- ★ од интереса су експерименти чији се исходи не могу предвидети са сигурношћу
- ★ бацање новчића, дељење карата у покеру, давање дозе лека,...

Да би се анализирали овакви експерименти треба размотрити две ствари: могућности и вероватноће. Не можемо говорити о томе шта је вероватно ако не знамо шта је могуће.

Ω : Сваком случајном експерименту придружујемо скуп свих могућих исхода које зовемо **простор елементарних догађаја (простор исхода)**

елементарни исходи су елементи простора исхода.



Пример

- а) Истовремено се бацају новчић и коцкица. Скуп елементарних исхода је

$$\Omega = \{П1, П2, П3, П4, П5, П6, Г1, Г2, Г3, Г4, Г5, Г6\}.$$

- б) Баца се новчић два пута. Скуп елементарних исхода је

$$\Omega = \{ПП, ПГ, ГП, ГГ\}.$$

- в) Новчић се баца четири пута и региструје се број колико пута се појавило писмо. Скуп елементарних исхода је

$$\Omega = \{0, 1, 2, 3, 4\}.$$



Пример Извлачи се једна карта из стандардог шпила од 52 карте (без џокера). Потенцијални скупови елементарних исхода:

$$\Omega_1 = \{\text{црвена, црна}\}$$

$$\Omega_2 = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$$

$$\Omega_3 = \{A\clubsuit, A\diamondsuit, A\heartsuit, A\spadesuit, \dots, K\clubsuit, K\diamondsuit, K\heartsuit, K\spadesuit\} \text{ (свака карта понаособ)}$$

$$\Omega_4 = \{\text{слика (краљ, дама, жандар), није слика}\}$$

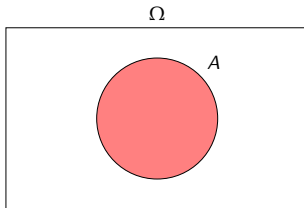
$$\Omega_5 = \{\text{слика, карта с бројем}\}$$

$$\Omega_6 = \{\text{слика, ас, није слика}\}$$

$\Omega_1, \Omega_2, \Omega_3, \Omega_4$ јесу скупови исхода; Ω_5 није (нема исхода који одговара асу); Ω_6 није (асу одговара више од једног исхода).



Нека је Ω простор исхода неког експеримента. Неке од подскупова скупа Ω називамо **догађајима**.



Кажемо да се догађај A реализује ако и само ако се оствари неки исход ω који припада подскупу A .

- ★ Ω сигуран догађај - увек се реализује
- ★ $\emptyset$ немогућ догађај - никад се не реализује

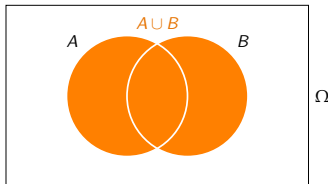
Пример. Новчић се баца четири пута. Догађај A је догађај да се писмо појавило већи број пута од главе.

$$A = \{\text{nnnn, nnnг, nngn, ngnn, gnnn}\}.$$

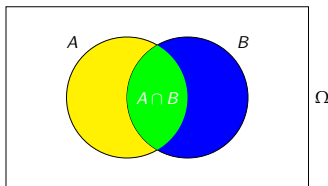


Ако су A и B догађаји, тада

$A \cup B = \{\omega \in \Omega \mid \omega \in A \vee \omega \in B\}$ - догађај $A \cup B$ се реализује ако и само ако се реализују бар један од догађаја A и B



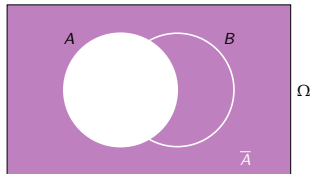
$A \cap B = AB = \{\omega \in \Omega \mid \omega \in A \wedge \omega \in B\}$ - догађај AB се реализује ако и само ако се реализују и A и B



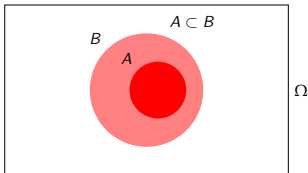
Пресек и унија се једноставно могу проширити на n догађаја.



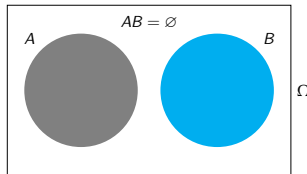
$\bar{A} = \{\omega \in \Omega \mid \omega \notin A\}$ - догађај $\bar{A}$ је супротан догађају A ако и само ако када се реализује $\bar{A}$ не реализује се A



$A \subset B: \omega \in A \implies \omega \in B$ - релација импликације, односно A имплицира B ако и само ако када се реализује догађај A реализује се и догађај B



$AB = \emptyset$ - A и B су дисјунктни (узајамно се искључују), односно догађаји A и B се не могу истовремено реализовати



Пример Коцкица се баца два пута. Одредити догађаје:

- ★ A - збир палих бројева је мањи од 6
 $A = \{11, 12, 13, 14, 21, 22, 23, 31, 32, 41\}$
- ★ B - пали су исти бројеви $B = \{11, 22, 33, 44, 55, 66\}$
- ★ Пали су бројеви чији је збир мањи од 6 или су пала два иста броја
 $A \cup B = \{11, 12, 13, 14, 21, 22, 23, 31, 32, 33, 41, 44, 55, 66\}$
- ★ Пали су бројеви чији је збир мањи од 6 и пала су два иста броја
 $AB = \{11, 22\}$
- ★ Пали су бројеви чији је збир бар 6
 $\bar{A} =$
 $\{15, 16, 24, 25, 26, 33, 34, 35, 36, 42, 43, 44, 45, 46, 51, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66\}$
- ★ Пали су бројеви чији је збир бар 6 и пала су два иста броја
 $\bar{A}B = \{33, 44, 55, 66\}$

Да ли су догађаји A и B дисјунктни?

Нису, јер $AB \neq \emptyset$.



Како догађајима доделити вероватноћу?

- ▷ субјективно
- ▷ класично (математички)
- ▷ статистички

Субјективни начин дефинисања вероватноће:

- ★ Лично мишљење уз доступне информације.
- ★ Он се често користи код догађаја који се нису десили раније и који се неће десити поново.
 - ★ која је вероватноћа да неки тим победи у утакмици (свака утакмица се одвија под релативно другачијим околностима)

Предност овог метода је што се увек може користити. Међутим, мана је што зависи од тачности и потпуности доступних информација као и способности појединца да доноси закључке. Овај приступ се назива и бајесов-ким.



Дефиниција: Нека се изводи експеримент у коме је сваки од његових исхода једнако вероватан. Нека је $n(A)$ број начина на које се може догодити догађај A , а n укупан број исхода експеримента. Тада је

$$P(A) = \frac{n(A)}{n}.$$

Пример Кутија садржи 5 артикала од којих су 2 дефектна. Артикли се узимају из кутије на случајан начин. Колика је вероватноћа да је трећи извучени артикал последњи дефектни артикал који је извучен?

Тада $\Omega = \{ddiii, didii, diidi, diiid, iddii, ididi, idiid, iiddi, iidid, iiidd\}$, а $A = \{didii, iddii\}$. Према томе

$$P(A) = \frac{2}{10} = \frac{1}{5}.$$



Дефиниција: Вероватноћа догађаја A се рачуна по формули

$$P(A) = \frac{\text{број експеримената у којима се догађај } A \text{ реализовао}}{\text{укупан број експеримената}}$$

Пример. Извођење експеримента бацања коцкице.



```
> eksperiment<-sample(c(1,2,3,4,5,6),1000,replace=TRUE)
> table(eksperiment)
```

```
eksperiment
 1    2    3    4    5    6 
161 166 153 173 163 184
```

Статистичка вероватноћа добијања шестице је

$$\frac{184}{1000} = 0.184$$

Класична вероватноћа добијања шестице је

$$\frac{1}{6} = 0.167$$

Мана овог метода је да се подразумева да се експеримент изводи велики број пута под практично истим условима, што у реалности није могуће. Али може се користити за рачунање вероватноћа компликованих догађаја који се могу симулирати (Монте Карло метод).

Класична дефиниција - неједнаке вероватноће исхода

Дефиниција: Нека се изводи експеримент чији могући исходи имају вероватноће редом $p_1, \dots, p_n$. Вероватноћа догађаја A једнака је збиру вероватноћа исхода који реализују догађај A .

Пример Баца се кутија шибица. Вероватноће добијања њених шест страна су: по $\frac{1}{20}$ за две странице најмање површине, по $\frac{1}{10}$ за две странице „средње“ површине, и по $\frac{7}{20}$ за две странице највеће површине. Вероватноћа догађаја A да шибица не падне на страницу највеће површине је $P(A) = 2 \cdot \frac{1}{20} + 2 \cdot \frac{1}{10} = \frac{3}{10}$.



```
> sibice<-sample(c("a1","a2","b1","b2","c1","c2"),10000,  
  replace=TRUE, prob=c(0.05,0.05,0.1,0.1,0.35,0.35))  
> table(sibice)
```

```
sibice  
 a1  a2  b1  b2  c1  c2  
521 496 975 978 3543 3487
```

$$P(A) \approx \frac{521+496+975+978}{10000} = \frac{2970}{10000} = 0.297$$