REPETITORIJ VJEROJATNOSTI I STATISTIKE ZA STUDENTE ELEKTROTEHNIKE

pripremio: mr.sc. Bojan Kovačić, viši predavač

nerecenzirana autorizirana verzija

Sadržaj

PREDGOVOR....4
1. OSNOVE KOMBINATORIKE....5
1.1. Permutacije i kombinacije. Binomni teorem....5

2. OSNOVE DISKRETNE TEORIJE VJEROJATNOSTI....7
2.1. Operacije s događajima....7

2.2 Neke posebne relacije među događajima....7

2.3. Korisni identiteti s događajima....7

2.4. Vjerojatnost. Vjerojatnosni prostor....8

2.5. Neka korisna svojstva vjerojatnosti....8

2.6. Klasičan vjerojatnosni prostor....8

2.7. Geometrijska vjerojatnost....9

2.8. Uvjetna vjerojatnost....9

2.9. Nezavisnost događaja....9

2.10. Potpuni sustav događaja. Formula potpune vjerojatnosti. Bayesova formula....10

2.11. Bernoullijeva shema....11

3. OSNOVE OPISNE (DESKRIPTIVNE) STATISTIKE....12
3.1. Modaliteti i njihove frekvencije....12

3.2. Srednje vrijednosti....13

3.3. Mjere raspršenja....15

4. DISKRETNE SLUČAJNE VARIJABLE....21
4.1. Diskretna jednolika slučajna varijabla i diskretna jednolika razdioba....23

4.2. Binomna slučajna varijabla i binomna razdioba....23

4.3. Poissonova slučajna varijabla i Poissonova razdioba....24

4.4. Geometrijska slučajna varijabla i geometrijska razdioba....25

4.5. Hipergeometrijska slučajna varijabla i hipergeometrijska razdioba....26

5. NEPREKIDNE (KONTINUIRANE) SLUČAJNE VARIJABLE 27
5.1. Neprekidna jednolika slučajna varijabla i neprekidna jednolika razdioba 28
5.2. Eksponencijalna slučajna varijabla i eksponencijalna razdioba 29
5.3. Normalna slučajna varijabla i normalna razdioba 30
5.4. Čebiševljeva nejednakost za neprekidne slučajne varijable. Pravilo $3\cdot\sigma$ 33
5.5. Granični teoremi u Bernoullijevoj shemi 33
6. DODATAK 34
6.1. Pregled osnovnih pojmova teorije skupova 34
6.2. Pregled formula iz diferencijalnoga i integralnoga računa 35
6.3. Pregled nekih MATLAB-ovih funkcija koje se koriste u vjerojatnosti i statistici 36
6.4. Pregled nekih funkcija MS Excel-a koje se koriste u vjerojatnosti i statistici 38
POPIS KORIŠTENIH OZNAKA 40
POPIS TABLICA 42
KAZALO POJMOVA 43
LITERATURA 48

PREDGOVOR

Ovaj nastavni materijal namijenjen je studentima 2. godine stručnoga studija elektrotehnike na Tehničkom veleučilištu u Zagrebu kao pomoć prigodom polaganja pisanog dijela ispita iz predmeta *Vjerojatnost i statistika*.

Tekst je pisan tako da redoslijed tematskih cjelina odgovara redoslijedu obrade tih cjelina u navedenom predmetu. Radi boljeg razumijevanja obrađene materije, osim matematičkih formula i postupaka, čije se poznavanje provjerava na ispitu, navedene su i teorijske činjenice. To ni u kojem slučaju ne znači da ovaj materijal može zamijeniti nastavne materijale prema kojima se održavaju predavanja i auditorne vježbe, nego mu je osnovna svrha poslužiti kao koristan podsjetnik na definicije, svojstva i formule koji se na ispitu možda zaborave.

Ugodna mi je dužnost zahvaliti svima koji su na bilo koji način pomogli u nastajanju ovog nastavnoga materijala. Tu ponajprije mislim na dekana Tehničkoga veleučilišta u Zagrebu mr. sc. Gorana Malčića, višega predavača, prodekana za nastavu i studente Borisa Uremovića, dipl. ing. građ., višega predavača, i na pročelnika Elektrotehničkoga odjela Tehničkog veleučilišta u Zagrebu mr. sc. Krunoslava Martinčića, višega predavača.

Posebno zahvaljujem kolegama Mandi Orlić Bachler i Luki Marohniću na korisnim primjedbama i prijedlozima, te svim studentima koji su svojim pitanjima na nastavi i konzultacijama izravno utjecali na sadržaj i kvalitetu ove verzije teksta.

Pokude za sve „preživjele” nenamjerne pogreške, kojih u tekstu nesumnjivo ima i nakon višestrukih korektura, preuzimam isključivo na sebe. Unaprijed zahvaljujem svima koji me obavijeste o svakoj uočenoj pogreški ili nekom drugom propustu.

Svim korisnicima ovoga nastavnoga materijala želim uspješno korištenje.

U Zagrebu, ožujka 2021.

Bojan Kovačić

1. OSNOVE KOMBINATORIKE

Teorem 1. (**pravilo jednakosti** (**bijekcije**)) Neka su S i T konačni skupovi. Tada je $\operatorname{card}(S)=\operatorname{card}(T)$ ako i samo ako postoji bijekcija među skupovima S i T.

Teorem 2. (**princip zbroja**) Neka su $n\in\mathbb{N}$ i $S_1,\ldots,S_n$ konačni, u parovima disjunktni skupovi (tj. vrijedi ekvivalencija $\left(S_i\cap S_j=\varnothing\right)\Leftrightarrow(i\neq j)$). Tada je skup $S:=\bigcup_{i=1}^{n}S_i$ konačan i vrijedi:

$$\operatorname{card}(S)=\sum_{i=1}^{n}\operatorname{card}(S_i).$$

Teorem 3. (**teorem o uzastopnom prebrojavanju**) Neka su $n\in\mathbb{N}$ i $S_1,\ldots,S_n$ konačni skupovi. Tada je Kartezijev umnožak $\prod_{k=1}^{n}S_k:=S_1\times\ldots\times S_n$ konačan skup i vrijedi:

$$\operatorname{card}\left(\prod_{k=1}^{n}S_k\right)=\prod_{k=1}^{n}\operatorname{card}(S_k).$$

1.1. Permutacije i kombinacije. Binomni teorem.

Neka su $n,r\in\mathbb{N}$ i $k\in\{1,2,\ldots,n\}$. Neka je $S=\{a_1,\ldots,a_n\}$ bilo koji n-člani skup.

Permutacija (**bez ponavljanja**) skupa S je bilo koja bijekcija toga skupa u samoga sebe. Ekvivalentno, permutacija skupa S je svaka uređena n-torka koja sadrži sve elemente toga skupa. Ukupan broj svih međusobno različitih permutacija skupa S jednak je:

$$P_n=n!.$$

Napomena: Dogovorno se uzima 0! = 1.

Pretpostavimo da smo od svih elemenata skupa S formirali strukturu A tako da se u A element a_1 pojavljuje točno k_1 puta, element a_2 pojavljuje točno k_2 puta, ..., element a_n pojavljuje točno k_n puta, pri čemu je $k_i\in\mathbb{N}\cup\{0\}$, za svaki $i=1,\ldots,n$. Neka je $m:=\sum_{i=1}^{n}k_i$. Kažemo da smo na taj način formirali ***m*-člani multiskup *A* na *S***.

Permutacija multiskupa A je svaka uređena m-torka koja sadrži sve elemente toga multiskupa. (Za tu permutaciju kažemo i da je **permutacija s ponavljanjem** skupa S.) Ukupan broj svih različitih permutacija multiskupa A jednak je:

$$P_m^{k_1,\dots,k_n} := \binom{m}{k_1,\dots,k_n} := \frac{m!}{\prod_{i=1}^{n}(k_i!)} = \frac{\left(\sum_{i=1}^{n} k_i\right)!}{\prod_{i=1}^{n}(k_i!)}.$$

Broj $\binom{m}{k_1,\dots,k_n}$ naziva se **multinomni koeficijent**.

***k*–permutacija bez ponavljanja** (ili kraće: ***k*-permutacija**) skupa S je bilo koja uređena k–torka međusobno različitih elemenata toga skupa. Ukupan broj svih različitih k–permutacija skupa S jednak je:

$$P(n,k) = \prod_{i=0}^{k-1}(n-i) = \frac{n!}{(n-k)!}.$$

***k*–permutacija s ponavljanjem** skupa S je bilo koja uređena k–torka u kojoj se na svakoj poziciji (komponenti) može pojaviti svaki element toga skupa. Ukupan broj svih različitih k–permutacija s ponavljanjem skupa S jednak je:

$$\overline{P}(n,k) = n^k.$$

***k*–kombinacija bez ponavljanja** (ili kraće: ***k*-kombinacija**) skupa S je bilo koji k–člani podskup toga skupa. Ukupan broj svih različitih k–kombinacija skupa S jednak je:

$$C(n,k) = \binom{n}{k} := \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k!}.$$

Broj $\binom{n}{k}$ naziva se **binomni koeficijent** i čita: „en povrh ka“.

***k*-kombinacija s ponavljanjem** skupa S je bilo koji k-člani multiskup na S. Ukupan broj svih različitih k–kombinacija s ponavljanjem skupa S jednak je:

$$\left(\binom{n}{k}\right) := \overline{C}(n,k) = \binom{n+k-1}{k}.$$

Teorem 4. (*neka svojstva binomnih koeficijenata i* **binomni teorem**) Za sve $k,n\in\mathbb{N}$ takve da je $k\le n$ i za sve $x,y\in\mathbb{C}$ vrijede sljedeće jednakosti:

1. $\binom{n}{0}=\binom{n}{n}=1,\ \binom{n}{1}=n,\ \binom{n}{2}=\frac{n\cdot(n-1)}{2},\dots,\ \binom{n}{k}=\frac{n!}{k!\cdot(n-k)!}.$

2. $\binom{n}{k}=\binom{n}{n-k}=\frac{n}{k}\cdot\binom{n-1}{k-1}=\binom{n-1}{k-1}+\binom{n-1}{k}.$

3. (binomna formula) $(x+y)^n=\sum_{i=0}^{n}\binom{n}{i}\cdot x^{n-i}\cdot y^i.$

2. OSNOVE DISKRETNE TEORIJE VJEROJATNOSTI

Slučajni pokus je svaki pokus čiji ishod nije unaprijed zadan. Svaki mogući ishod takvoga pokusa naziva se **elementarni događaj** i uobičajeno označava s ω. Skup svih elementarnih događaja naziva se **prostor elementarnih događaja** i označava s Ω. Kao ishod slučajnoga pokusa javlja se točno jedan element skupa Ω.

Algebra događaja je familija $\mathcal{F}$ podskupova skupa Ω koja ima sljedeća svojstva:

A1. $\varnothing,\ \Omega \in \mathcal{F}$.

A2. $(A, B \in \mathcal{F}) \Rightarrow (A \cup B \in \mathcal{F})$.

A3. $(A \in \mathcal{F}) \Rightarrow \left(A^C := \Omega \setminus A \in \mathcal{F}\right)$.

Elementi algebre $\mathcal{F}$ nazivaju se **događaji**. Prazan skup $\varnothing$ naziva se **nemoguć događaj**, a skup Ω (promatran kao element algebre $\mathcal{F}$) **siguran događaj**.

2.1. Operacije s događajima

1. Zbroj događaja: $A + B := A \cup B$

Interpretacija: $A + B$ se ostvari ako se ostvari barem jedan od događaja A i B.

2. Umnožak događaja: $A \cdot B := A \cap B$

Interpretacija: $A \cdot B$ se ostvari ako se istovremeno ostvare događaj A i događaj B.

3. Razlika događaja: $A - B := A \setminus B$

Interpretacija: $A - B$ se ostvari ako se ostvari događaj A, ali ne i događaj B.

2.2 Neke posebne relacije među događajima

1. Događaj A **povlači** događaj B ako vrijedi $A \subset B$.

2. Događaji A i B su **jednaki** ako vrijedi $A = B$.

3. Događaji A i B su **disjunktni** ili **međusobno isključivi** ako vrijedi $A \cdot B = \varnothing$.

4. Događaj A^C naziva se **događaj suprotan događaju** A. On se ostvaruje ako i samo ako se A ne ostvaruje.

2.3. Korisni identiteti s događajima

Za bilo koje događaje $A,\ B,\ C \in \mathcal{F}$ vrijede sljedeće jednakosti.

1. $A - B = A \cdot B^C$.

2. $\left(A^C\right)^C = A.$

3. (*prva de Morganova formula*) $(A+B)^C = A^C \cdot B^C.$

4. (*druga de Morganova formula*) $(A \cdot B)^C = A^C + B^C.$

5. $A \cdot (B+C) = A \cdot B + A \cdot C.$

6. $(A+B) \cdot (A+C) = A + B \cdot C.$

2.4. Vjerojatnost. Vjerojatnosni prostor

Vjerojatnost na algebri $\mathcal{F}$ je svako preslikavanje $P : \mathcal{F} \to [0, 1]$ sa svojstvima:

P1. $P(\emptyset) = 0,\ P(\Omega) = 1.$

P2. (monotonost) $(A \subset B) \Rightarrow \left(P(A) \le P(B)\right).$

P3. (aditivnost) $(A \cdot B = \emptyset) \Rightarrow \left(P(A+B) = P(A) + P(B)\right).$

Uređena trojka $(\Omega, \mathcal{F}, P)$ naziva se **vjerojatnosni prostor**. Ovisno o broju elemenata skupa Ω, taj prostor može biti *konačan* ili *beskonačan*.

2.5. Neka korisna svojstva vjerojatnosti

Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor. Za bilo koje $A, B, C \in \mathcal{F}$ vrijede sljedeće jednakosti.

1. $P(A+B) = P(A) + P(B) - P(A \cdot B).$

2. $P(A+B) = P(A) + P(B \cdot A^C).$

3. $P(A) = P(A \cdot B^C) + P(A \cdot B).$

4. $P(A^C) = 1 - P(A).$

5. $P(A+B+C) = P(A) + P(B) + P(C) - P(A \cdot B) - P(A \cdot C) - P(B \cdot C) + P(A \cdot B \cdot C).$

2.6. Klasičan vjerojatnosni prostor

Klasičan vjerojatnosni prostor je svaki *konačan* vjerojatnosni prostor $(\Omega, \mathcal{F}, P)$ u kojemu vrijedi:

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{\text{ukupan broj povoljnih ishoda za } A}{\text{ukupan broj mogućih ishoda}},\ \forall A \in \mathcal{F}..$$

Ako je $\Omega = \{\omega_1, ..., \omega_n\}$, onda vrijede sljedeće jednakosti.

1. $P(\omega_i) = \dfrac{1}{\text{card}(\Omega)} = \dfrac{1}{n},\ \forall i = 1, ..., n.$

2. $\sum_{i=1}^{n} P(\omega_i) = 1.$

2.7. Geometrijska vjerojatnost

Ako su Ω i $A \subseteq \Omega$ *ograničeni i izmjerivi* podskupovi skupa $\mathbb{R}^n$, za neki $n \in \mathbb{N}$, onda vjerojatnost događaja A definiramo formulom:

$$P(A) = \frac{m(A)}{m(\Omega)},$$

gdje je m tzv. *Lebesqueova mjera* (dalje u tekstu: *mjera*) na skupu $\mathbb{R}^n$. Takvu vjerojatnost nazivamo **geometrijska vjerojatnost**.

Za $n = 1$ mjera bilo kojega intervala jednaka je razlici gornje i donje granice toga intervala.

Za $n = 2$ mjera skupa jednaka je površini toga skupa.

Za $n = 3$ mjera skupa jednaka je volumenu toga skupa.

Posebno, vrijede sljedeće tvrdnje:

1. Mjera bilo kojega *jednočlanoga* skupa jednaka je nuli.

2. Mjera bilo koje ravninske krivulje u $\mathbb{R}^2$ (npr. pravac, kružnica, elipsa itd.) jednaka je nuli.

3. Mjera bilo koje ravninske plohe u $\mathbb{R}^3$ (npr. sfera, plašt kocke itd.) jednaka je nuli.

2.8. Uvjetna vjerojatnost

Neka su $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i $A, B \in \mathcal{F}$.

Vjerojatnost događaja A **uz uvjet da se dogodio događaj** B (tzv. **uvjetna vjerojatnost**) definirana je formulom:

$$P(A \mid B) := \frac{P(A \cdot B)}{P(B)}.$$

Općenito, vrijedi formula:

$$P\left(\prod_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} P\left(A_k \mid \prod_{i=1}^{k-1} A_i\right), \ \forall n \in \mathbb{N}.$$

2.9. Nezavisnost događaja

Neka su $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i $A, B \in \mathcal{F}$. Kažemo da su događaji A i B **nezavisni** ako vrijedi jednakost:

$$P(A \cdot B) = P(A) \cdot P(B).$$

Općenito, za n događaja $A_1,\dots,A_n \in \mathcal{F}$ kažemo da su **nezavisni** ako za *svaki* neprazan skup $I \subseteq \{1,\dots,n\}$ vrijedi jednakost:

$$P\left(\prod_{i\in I} A_i\right) = \prod_{i\in I} P(A_i).$$

Npr. tri događaja A_1, A_2 i A_3 su nezavisna ako vrijede sljedeće četiri jednakosti:

$$\begin{cases} P(A\cdot B) = P(A)\cdot P(B), \\ P(A\cdot C) = P(A)\cdot P(C), \\ P(B\cdot C) = P(B)\cdot P(C), \\ P(A\cdot B\cdot C) = P(A)\cdot P(B)\cdot P(C). \end{cases}$$

2.10. Potpuni sustav događaja. Formula potpune vjerojatnosti. Bayesova formula

Neka je $(\Omega,\mathcal{F},P)$ vjerojatnosni prostor. Kažemo da je *konačan* skup događaja $A=\{A_1,\dots,A_n\}\subseteq\mathcal{F}$ **potpuni sustav događaja** ako vrijedi:

1. $A_i\cdot A_j=\varnothing$, za sve $i\neq j$.

2. $\sum_{i=1}^{n} A_i = \Omega$.

3. $\sum_{i=1}^{n} P(A_i) = 1$.

Elemente skupa A nazivamo **hipoteze**.

Ako su $A=\{A_1,\dots,A_n\}$ potpuni sustav događaja i $B\in\mathcal{F}$, onda vrijedi **formula potpune vjerojatnosti**:

$$P(B) = \sum_{i=1}^{n} P(A_i)\cdot P(B\mid A_i).$$

Za primjenu ove formule treba znati vjerojatnost svake hipoteze i odgovarajuće uvjetne vjerojatnosti događaja B.

Obratno, ako treba odrediti vjerojatnost ostvaraja svake pojedine hipoteze uz uvjet da se dogodio događaj B, onda primjenjujemo **Bayesovu formulu**:

$$P(A_i\mid B) = \frac{P(A_i\cdot B)}{P(B)} = \frac{P(A_i)\cdot P(B\mid A_i)}{\sum_{j=1}^{n} P(A_j)\cdot P(B\mid A_j)},\ \forall i=1,\dots,n.$$

2.11. Bernoullijeva shema.

Neka je $n \in \mathbb{N}$ proizvoljan, ali fiksiran.

Bernoullijeva shema je vjerojatnosni model n-terostrukoga ponavljanja slučajnoga pokusa koji ima točno dva moguća ishoda: *uspjeh* i *neuspjeh*. Pritom je vjerojatnost pojavljivanja *svakoga* ishoda u *svakom* pojedinom pokusu stalna.

Neka je p vjerojatnost pojavljivanja uspjeha u *svakom* slučajnom pokusu. Tada je vjerojatnost p_k da se u svih n pokusa dogodi točno k uspjeha jednaka:

$$p_k = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}, \ \forall k = 0,1,\ldots,n.$$

Za velike n ove se vjerojatnosti mogu računati pomoću odgovarajućih graničnih teorema (vidjeti stranice 24. i 33.).

3. OSNOVE OPISNE (DESKRIPTIVNE) STATISTIKE

3.1. Modaliteti i njihove frekvencije

Neka je $y_1, y_2, ..., y_n$ konačan (ne)numerički niz podataka dobivenih mjerenjem određenoga statističkoga obilježja. Svaki element toga niza naziva se **modalitet** statističkoga obilježja. Broj $n \in \mathbb{N}$ naziva se **duljina statističkoga niza**.

Pretpostavimo da se u nizu $y_1, y_2, ..., y_n$ pojavljuje točno k različitih modaliteta. Označimo te modalitete s $x_1, ..., x_k$. Sljedeće definicije vrijede za svaki $i = 1, ..., k$.

Apsolutna frekvencija f_i **modaliteta** x_i jednaka je ukupnom broju pojavljivanja dotičnoga modaliteta u zadanom nizu.

Relativna frekvencija r_i **modaliteta** x_i iskazuje se u postotcima i definira formulom:

$$r_i = \frac{f_i}{n} \cdot 100 \ [\%].$$

Pretpostavimo da je niz $y_1, y_2, ..., y_n$ dobiven mjerenjem kvantitativnoga ili kvalitativnoga redoslijednoga obilježja. Tada su svaka dva različita modaliteta međusobno usporediva. Zbog toga možemo pretpostaviti da su svi međusobno različiti modaliteti $x_1, ..., x_k$ označeni tako da vrijedi:

$$x_1 \le x_2 \le ... \le x_k.$$

Apsolutna frekvencija „veće od“ $f_i^>$ **modaliteta** x_i jednaka je zbroju apsolutnih frekvencija svih modaliteta jednakih ili većih (boljih) od x_i:

$$f_i^> = \sum_{x_j \ge x_i} f_j.$$

Apsolutna frekvencija „manje od“ $f_i^<$ **modaliteta** x_i jednaka je zbroju apsolutnih frekvencija svih modaliteta jednakih ili manjih (slabijih) od x_i:

$$f_i^< = \sum_{x_j \le x_i} f_j.$$

Apsolutne frekvencije „veće od“/„manje od“ tvore **kumulativne nizove apsolutnih frekvencija**.

Relativna frekvencija „veće od“ $r_i^>$ **modaliteta** x_i iskazuje se u postotcima i definira formulom:

$$r_i^> = \frac{f_i^>}{n} \cdot 100 \ [\%].$$

Relativna frekvencija „manje od“ $r_i^<$ **modaliteta** x_i iskazuje se u postotcima i definira formulom:

$$r_i^< = \frac{f_i^<}{n} \cdot 100\ [\%].$$

Relativne frekvencije „veće od“ i „manje od“ tvore **kumulativne nizove relativnih frekvencija.**

Ako je broj različitih modaliteta relativno velik, podatke je pogodno grupirati u (*ne*)*prave zatvorene razrede*. **Zatvoreni razred** je svaki interval oblika $[a,b\rangle \subset \mathbb{R}$. Ako zatvorene razrede možemo poredati u niz $[a,b\rangle$, $[b,c\rangle$, $[c,d\rangle$, ..., onda takve razrede nazivamo **pravi razredi**. Ako između bilo kojih dvaju uzastopnih razreda postoji razmak, razrede nazivamo **nepravima** ili **nominalnima**. Ne istaknemo li drugačije, dalje u tekstu pod pojmom *razred* podrazumijevamo pravi razred, dok pod pojmom *grupirani podaci* podrazumijevamo podatke grupirane u prave razrede.

Neka je $[a,b\rangle$ zatvoreni razred. Broj a nazivamo **donja granica razreda**, a broj b **gornja granica razreda**. Broj $h := b - a$ nazivamo **širina razreda** ili **razredna širina,** a broj $s := \frac{a+b}{2}$ **sredina razreda** ili **razredna sredina**.

Osim tablično, modalitete i njihove apsolutne/relativne frekvencije podesno je prikazati i grafički npr. strukturnim krugom, jednostavnim stupcima, strukturnim stupcima i sl.

Poligon apsolutnih frekvencija je otvorena poligonalna crta dobivena spajanjem uređenih parova točaka (x_i, f_i) i (x_{i+1}, f_{i+1}) u ravnini, za svaki $i = 1,...,k-1$. Analogno se definira **poligon relativnih frekvencija.**

Histogram je površinski grafikon kojega tvori konačan niz zatvorenih pravokutnika. Duljine osnovicâ tih pravokutnika odgovaraju širinama razreda. *Površine* tih pravokutnika jednake su apsolutnim/relativnim frekvencijama dotičnih razreda.

3.2. Srednje vrijednosti

Ne istaknemo li drugačije, pod pojmom *niz* podrazumijevamo niz numeričkih podataka (numerički niz).

Niz je pogodno opisivati tzv. **srednjim vrijednostima**. Srednje vrijednosti dijele se na **potpune** (izračunavaju se korištenjem svih elemenata niza) i **položajne** (izračunavaju se korištenjem položaja elemenata unutar niza).

Potpune srednje vrijednosti su **aritmetička sredina**, **geometrijska sredina** i **harmonijska sredina**. Formule za izračun tih vrijednosti iz negrupiranih podataka,

odnosno njihovu procjenu iz grupiranih podataka, navedene su u tablici 3. (vidjeti stranicu 18.)

U *položajne* srednje vrijednosti ubrajamo **mod** i **medijan**.

Mod je modalitet s najvećom apsolutnom/relativnom frekvencijom. On je jedina srednja vrijednost koju je moguće određivati i za numerička i za nenumerička obilježja.

Zavisno o broju modova, razdioba može biti **unimodalna** (ima jedinstven mod), **bimodalna** (ima točno dva različita moda) i **multimodalna** (ima barem tri različita moda).

Ako su podaci grupirani u $m \in \mathbb{N}$ razreda, za *procjenu* moda najprije treba izračunati **korigiranu apsolutnu/relativnu frekvenciju** svakoga razreda. Ta frekvencija dobije se dijeljenjem originalne apsolutne/relativne frekvencije razreda i originalne razredne širine:

$$\begin{cases} f_i^{corr} = \dfrac{f_i}{s_i}, \\ r_i^{corr} = \dfrac{r_i}{s_i}, \end{cases} \text{ za svaki } i = 1, \ldots, m.$$

Potom se odredi razred s najvećom korigiranom apsolutnom/relativnom frekvencijom. Taj razred naziva se **modalni razred**. Tada se mod procjenjuje prema formuli:

$$Mo = a + \frac{f - f^{\#}}{2 \cdot f - \left(f^{\#} + f^{\#\#}\right)} \cdot h,$$

gdje su Mo mod, a donja granica modalnoga razreda, f korigirana frekvencija modalnoga razreda, $f^{\#}$ korigirana frekvencija razreda koji neposredno prethodi modalnom razredu, $f^{\#\#}$ korigirana frekvencija razreda koji neposredno slijedi iza modalnoga razreda i h razredna širina modalnoga razreda.

Napomene: **1.)** Ako su podaci grupirani u razrede *jednakih širina*, nije potrebno računati korigirane frekvencije. U tom slučaju su korigirane frekvencije jednake originalnim frekvencijama razreda.

2.) Ako modalni razred nema prethodnika, onda je $f^{\#} = 0$. Analogno, ako modalni razred nema sljedbenika, onda je $f^{\#\#} = 0$.

Percentili su položajne vrijednosti koje dijele niz na 100 jednakobrojnih dijelova. Pritom ***k*–ti percentil** P_k ima svojstvo da ukupno $k\%$ članova niza nije veće od P_k, odnosno da $(100-k)\%$ članova niza nije manje od P_k.

Formule za određivanje percentila iz negrupiranih podataka, odnosno njihovu procjenu iz grupiranih podataka, navedene su u tablici 4. (vidjeti stranicu 18.)

Za $k \in \{10, 20, ..., 90\}$ govorimo o **decilima**. **Prvi decil** je 10. percentil, **drugi decil** je 20. percentil itd.

Za $k \in \{25, 50, 75\}$ govorimo o **kvartilima**. **Prvi** ili **donji kvartil** je 25. percentil, **drugi kvartil** ili **medijan** je 50. percentil, a **treći** ili **gornji kvartil** je 75. percentil.

Formule za određivanje svih triju kvartila iz negrupiranih podataka, odnosno njihovu procjenu iz grupiranih podataka, navedene su u tablici 5. (vidjeti stranicu 19.)

3.3. Mjere raspršenja

Reprezentativnost pojedine srednje vrijednosti iskazuje se pomoću odgovarajućih **mjera raspršenja** (**disperzije**). Mjere raspršenja dijele se na **apsolutne** i **relativne**.

Apsolutne mjere raspršenja su **raspon varijacije**, **interpercentil**, **srednje apsolutno odstupanje**, **varijanca** (**disperzija**) i **standardna devijacija** (**standardno odstupanje**).

Relativne mjere raspršenja su **koeficijent kvartilne devijacije** i **koeficijent varijacije**.

Raspon varijacije (oznaka: R) je razlika najveće i najmanje vrijednosti u nizu.

Interpercentil je razlika bilo kojih dvaju percentila P_k i P_l, pri čemu je nužno $k \geq l$. On označava raspon varijacije središnjih $(k-l)\%$ članova niza. Za $k = 75$ i $l = 25$ dobiva se **interkvartil** I_q. On označava raspon varijacije središnje polovice članova niza.

Odstupanje Δy_i elementa y_i niza $y_1, ..., y_n$ **od** njihove **aritmetičke sredine** A definirano je pravilom:

$$\Delta y_i := y_i - A, \ \forall i = 1, 2, ..., n.$$

Za ta odstupanja vrijedi jednakost:

$$\sum_{i=1}^{n} \Delta y_i = 0.$$

Srednje apsolutno odstupanje (oznaka: MAD) niza $y_1, ..., y_n$ definira se kao aritmetička sredina svih elemenata niza $|\Delta y_1|, ..., |\Delta y_n|$. Ono se obično (neprecizno) interpretira kao prosjek svih apsolutnih vrijednosti odstupanja elemenata niza od njihove aritmetičke sredine.

Varijanca ili **srednje kvadratno odstupanje** (oznaka: σ^2) niza $y_1, ..., y_n$ definira se kao aritmetička sredina svih elemenata niza $(\Delta y_1)^2, ..., (\Delta y_n)^2$. Ono se obično

(neprecizno) interpretira kao prosjek kvadrata svih odstupanja elemenata niza od njihove aritmetičke sredine.

Standardno odstupanje ili **standardna devijacija** (oznaka: σ) niza $y_1,...,y_n$ definira se kao drugi korijen iz varijance. Obično se (netočno) interpretira kao prosječno odstupanje elemenata niza od njihove aritmetičke sredine.

Formule za izračun srednjega apsolutnoga odstupanja, varijance i standardne devijacije iz negrupiranih podataka, odnosno njihovu procjenu iz grupiranih podataka, navedene su u tablici 6. (vidjeti stranicu 20.)

Koeficijent kvartilne devijacije V_q jednak je omjeru interkvartila I_q i zbroja prvoga i trećega kvartila:

$$V_q = \frac{Q_3 - Q_1}{Q_3 + Q_1}.$$

Taj koeficijent se interpretira kao intenzitet varijabiliteta središnje polovice statističkoga niza. Koristi se u slučajevima kad je medijan reprezentativnija srednja vrijednost od aritmetičke sredine. Jedan od kriterija za interpretaciju intenziteta varijabiliteta iskazanoga koeficijentom kvartilne devijacije naveden je u tablici 2.

Tablica 1. Skala intenziteta varijabiliteta koeficijenta kvartilne devijacije

V_q	0.0 – 0.1	0.1 – 0.2	0.2 – 0.3	0.3 – 0.5	0.5 – 1.0
interpretacija	vrlo slab	relativno slab	umjeren	relativno jak	jak

Koeficijent varijacije V jednak je količniku standardne devijacije i aritmetičke sredine iskazanom u postotcima:

$$V = \frac{\sigma}{A} \cdot 100 \ [\%].$$

On predstavlja relativno prosječno odstupanje vrijednosti numeričkoga niza od aritmetičke sredine niza. Primjenjuje se u slučajevima kad je aritmetička sredina dobar reprezentant numeričkoga niza. Jedan od kriterija za interpretaciju varijabiliteta niza iskazanoga koeficijentom varijacije naveden je u tablici 3.

Tablica 2. Skala intenziteta varijabiliteta koeficijenta varijacije

V[%]	0 – 10	10 – 30	30 – 50	50 – 70	≥ 70
interpretacija	vrlo slab	relativno slab	umjeren	relativno jak	jak

Teorem 6. (**Čebiševljevo pravilo u opisnoj statistici**) Neka je (y_n) konačan niz vrijednosti kvantitativnoga obilježja. Neka su $\overline{y}$ i σ redom pripadna aritmetička sredina, odnosno standardna devijacija. Tada se za svaki prirodan broj $k > 1$ u segmentu $\left[\overline{y} - k \cdot \sigma,\ \overline{y} + k \cdot \sigma\right]$ nalazi najmanje $100 \cdot \left(1 - \frac{1}{k^2}\right)$ % članova niza (y_n).

Korolar 1. (posljedice Čebiševljeva pravila)

a) U segmentu $\left[\overline{y} - 2 \cdot \sigma,\ \overline{y} + 2 \cdot \sigma\right]$ nalazi se najmanje 75% članova niza (y_n).

b) U segmentu $\left[\overline{y} - 3 \cdot \sigma,\ \overline{y} + 3 \cdot \sigma\right]$ nalazi najmanje 89% članova niza (y_n).

Za usporedbu raznorodnih numeričkih nizova koristi se **standardizirano obilježje**. Neka je (y_n) niz vrijednosti kvantitativnoga obilježja. Neka su $\overline{y}$ i σ redom pripadna aritmetička sredina, odnosno standardna devijacija. Tada je **niz standardiziranih vrijednosti** toga obilježja niz (z_n) čiji je opći član dan pravilom:

$$z_k = \frac{y_k - \overline{y}}{\sigma}.$$

Standardizacija vrijednosti numeričkoga obilježja se obično provodi kod izračunavanja vrijednosti funkcije razdiobe vjerojatnosti normalne slučajne varijable (vidjeti stranicu 30.).

Tablica 3. Određivanje potpunih srednjih vrijednosti

srednja vrijednost	*izračun iz negrupiranih podataka*	*procjena iz grupiranih podataka*
aritmetička sredina	$A=\frac{\sum_{i=1}^{n} y_i}{n}=\frac{y_1+y_2+\ldots+y_n}{n}$	$A=\frac{f_1\cdot s_1+\ldots+f_m\cdot s_m}{n}=\frac{\sum_{i=1}^{m} f_i\cdot s_i}{n}$
geometrijska sredina	$G=\sqrt[n]{y_1\cdot y_2\cdot\ldots\ldots\cdot y_n}$	$G=\sqrt[n]{\prod_{i=1}^{m}(s_i)^{f_i}}=\sqrt[n]{(s_1)^{f_1}\cdot\ldots\cdot(s_m)^{f_m}}$
harmonijska sredina	$H=\frac{n}{\frac{1}{y_1}+\frac{1}{y_2}+\ldots+\frac{1}{y_n}}=\frac{n}{\sum_{i=1}^{n}\frac{1}{y_i}}$	$H=\frac{n}{\frac{f_1}{s_1}+\frac{f_2}{s_2}+\ldots+\frac{f_m}{s_m}}=\frac{n}{\sum_{i=1}^{m}\frac{f_i}{s_i}}$

Napomena: Prilikom grupiranja u razrede pretpostavlja se da su podaci grupirani u točno m pravih razreda. Veličine f_i i s_i označavaju apsolutnu frekvenciju, odnosno sredinu i-toga razreda, za svaki $i=1,\ldots,m$. Pritom vrijedi jednakost:

$$\sum_{i=1}^{m} f_i = n.$$

Tablica 4. Određivanje percentila iz (ne)grupiranih podataka

percentil	*određivanje iz negrupiranih podataka*	*procjena iz grupiranih podataka*
P_k	$P_k=\begin{cases} y_{\left\lceil\frac{k\cdot n}{100}\right\rceil}, \text{ ako } n \text{ nije djeljiv sa 100,} \\ \frac{1}{2}\cdot\left(y_{\frac{k\cdot n}{100}}+y_{\frac{k\cdot n}{100}+1}\right), \text{ ako je } n \text{ djeljiv sa 100.}\end{cases}$	$P_k=a+\frac{\frac{k}{100}\cdot n-f_{g-1}^{<}}{f_g}\cdot h$

Napomena: Prilikom grupiranja u razrede pretpostavlja se da su podaci grupirani u točno m uzlazno sortiranih razreda. Tada najprije treba formirati kumulativni niz apsolutnih frekvencija „manje od“. Potom treba odrediti prvi član toga niza koji obuhvaća ukupno $\left\lceil\frac{k}{100}\cdot n\right\rceil$ podataka. Tom članu odgovara razred $[a,b\rangle$. Broj $g\in\mathbb{N}$ je redni broj toga razreda u uzlaznom poretku svih m razreda, dok je $h:=b-a$ pripadna razredna širina.

Za $k \in \{25, 50, 75\}$ iz tablice 4. dobiva se tablica 5.

Tablica 5. Određivanje kvartila iz (ne)grupiranih podataka

kvartil	*određivanje iz negrupiranih podataka*	*procjena iz grupiranih podataka*
1. (donji)	$Q_1 = \begin{cases} y_{\lceil \frac{n}{4} \rceil}, \text{ ako } n \text{ nije djeljiv s 4;} \\ \frac{1}{2} \cdot (y_{\frac{n}{4}} + y_{\frac{n}{4}+1}), \text{ ako je } n \text{ djeljiv s 4.} \end{cases}$	$Q_1 = a + \frac{\frac{1}{4} \cdot n - f_{g-1}^{<}}{f_g} \cdot h$
2. (medijan)	$Me = Q_2 = \begin{cases} y_{\frac{n+1}{2}}, \text{ ako je } n \text{ neparan;} \\ \frac{1}{2} \cdot (y_{\frac{n}{2}} + y_{\frac{n}{2}+1}), \text{ ako je } n \text{ paran.} \end{cases}$	$Me = a + \frac{\frac{1}{2} \cdot n - f_{g-1}^{<}}{f_g} \cdot h$
3. (gornji)	$Q_3 = \begin{cases} y_{\lceil \frac{3}{4} \cdot n \rceil}, \text{ ako } n \text{ nije djeljiv s 4;} \\ \frac{1}{2} \cdot (y_{\frac{3}{4} \cdot n} + y_{\frac{3}{4} \cdot n+1}), \text{ ako je } n \text{ djeljiv s 4.} \end{cases}$	$Q_3 = a + \frac{\frac{3}{4} \cdot n - f_{g-1}^{<}}{f_g} \cdot h$

Napomena: Prilikom grupiranja u razrede pretpostavlja se da su podaci grupirani u točno m uzlazno sortiranih razreda. Tada najprije treba formirati kumulativni niz apsolutnih frekvencija „manje od“. Potom treba odrediti prvi član toga niza koji obuhvaća ukupno $\left\lceil \frac{n}{4} \right\rceil$ podataka (u slučaju donjega kvartila), $\left\lceil \frac{n}{2} \right\rceil$ podataka (u slučaju medijana), odnosno $\left\lceil \frac{3}{4} \cdot n \right\rceil$ podataka (u slučaju gornjega kvartila). Tom članu odgovara razred $[a, b\rangle$. Broj $g \in \mathbb{N}$ je redni broj toga razreda u uzlaznom poretku svih m razreda, dok je $h := b - a$ pripadna razredna širina.

Tablica 6. Određivanje mjera raspršenja (disperzije)

mjera raspršenja	*izračun iz negrupiranih podataka*	*procjena iz grupiranih podataka*
srednje apsolutno odstupanje	$MAD = \frac{1}{n} \cdot \sum_{k=1}^{n} \lvert y_k - A \rvert$	$MAD = \frac{1}{n} \cdot \sum_{i=1}^{m} f_i \cdot \lvert s_i - A \rvert$
varijanca	$\sigma^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - A)^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} y_i^2 - A^2$	$\sigma^2 = \frac{1}{n} \cdot \sum_{i=1}^{m} f_i \cdot (s_i - A)^2 = \frac{1}{n} \cdot \sum_{i=1}^{m} f_i \cdot s_i^2 - A^2$
standardna devijacija	$\sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (y_i - A)^2} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} y_i^2 - A^2}$	$\sigma = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{m} f_i \cdot (s_i - A)^2} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{m} f_i \cdot s_i^2 - A^2}$

Napomena: Prilikom grupiranja u razrede pretpostavlja se da su podaci grupirani u točno m razreda. Veličine f_i i s_i označavaju redom apsolutnu frekvenciju, odnosno razrednu sredinu i-toga razreda, za svaki $i = 1, ..., m$. Pritom vrijedi jednakost:

$$\sum_{i=1}^{m} f_i = n.$$

U praksi se vrlo često podaci prikazuju tablično tako da se u tablici navedu svi modaliteti x_i i njihove apsolutne frekvencije f_i, za $i = 1, ..., k$. U tom slučaju za određivanje mjera iz tablica 3. i 6. treba primijeniti formule iz trećega stupca dotične tablice (tj. računati kao da su podaci grupirani u razrede). Pritom u svakoj pojedinoj formuli:

- ukupan broj razreda m treba zamijeniti ukupnim brojem modaliteta k,
- razrednu sredinu s_i treba zamijeniti modalitetom x_i.

Napomena: Položajne mjere iz tablica 4. i 5. nije moguće određivati na analogan način. Za određivanje tih mjera u ovome je slučaju najpodesnije razgrupirati podatke, odnosno formirati originalni niz negrupiranih podataka.

4. DISKRETNE SLUČAJNE VARIJABLE

Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor. Neka je S *konačan* ili *prebrojiv* podskup skupa $\mathbb{R}$. (Najčešće uzimamo da je $S \subseteq \mathbb{Z}$.)

Diskretna slučajna varijabla je svako preslikavanje $X : \Omega \to S$ takvo da za svaki $s \in S$ vrijedi relacija:

$$A_s := \{\omega \in \Omega : X(\omega) = s\} \subseteq \mathcal{F}.$$

Kažemo da je skup S **slika** diskretne slučajne varijable X. Pišemo: $R(X) = S$.

Ne istaknemo li drugačije, dalje u tekstu ovoga poglavlja pretpostavljamo da je X diskretna slučajna varijabla čija je slika skup S.

Svakom $s \in S$ jednoznačno je pridružen broj $p_s := P(A_s)$. Kratko pišemo:

$$p_s = P(X = s).$$

Zbog toga je svaka varijabla X jednoznačno određena svojom **tablicom razdiobe (distribucije):**

$$X \sim \begin{pmatrix} s_1 & s_2 & \dots & s_n & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}.$$

Neka su $A \subseteq S$ i $g : A \to \mathbb{R}$. Tada je $g(X)$ također slučajna varijabla. Njezina slika je skup $S_1 := \{g(s) : s \in S\}$, a tablica razdiobe

$$g(X) \sim \begin{pmatrix} g(s_1) & g(s_2) & \dots & g(s_n) & \dots \\ p_1 & p_2 & \dots & p_n & \dots \end{pmatrix}.$$

Funkcija vjerojatnosti varijable X je funkcija $f : S \to [0, 1]$ definirana pravilom:

$$f(s) = p_s.$$

Osnovno svojstvo funkcije vjerojatnosti je:

$$\sum_{s \in S} f(s) = 1.$$

Funkcija razdiobe vjerojatnosti varijable X je funkcija $F : S \to [0, 1]$ definirana pravilom:

$$F(s) = P(X \le s) = \sum_{x \le s} f(x),$$

gdje je f funkcija vjerojatnosti varijable X.

Neka korisna svojstva funkcije razdiobe vjerojatnosti

1. $P(X < s) = F(s) - p_s$.

2. $P(X \geq s) = 1 + p_s - F(s)$.

3. $P(X > s) = 1 - F(s)$.

Matematičko očekivanje (ili kraće **očekivanje**) $E(X)$ varijable X definirano je pravilom:

$$E(X) = \sum_{s \in S} s \cdot p_s.$$

Ono se obično interpretira kao analogon *težinske* aritmetičke sredine.

Varijanca $\mathrm{Var}(X)$ i **standardna devijacija** $\sigma(X)$ varijable X definirane su pravilima:

$$\begin{cases} \mathrm{Var}(X) = E\left(\left(X - E(X)\right)^2\right), \\ \sigma(X) = \sqrt{\mathrm{Var}(X)}. \end{cases}$$

Njihove interpretacije su potpuno analogne interpretacijama varijance, odnosno standardne devijacije u opisnoj statistici. Tako se varijanca slučajne varijable obično interpretira kao prosjek kvadrata odstupanja vrijednosti slučajne varijable od njezina očekivanja, a standardna devijacija slučajne varijable kao prosjek odstupanja vrijednosti slučajne varijable od njezina očekivanja.

Za bilo koju varijablu X, te $a, b \in \mathbb{R}$ takve da je $a \neq 0$ vrijede sljedeća svojstva:

1. $\mathrm{Var}(X) = \left(\sum_{s \in S} s^2 \cdot p_s\right) - \left(E(X)\right)^2$.

2. $\mathrm{Var}(X) = E(X^2) - \left(E(X)\right)^2$.

3. $\mathrm{Var}(a \cdot X + b) = a^2 \cdot \mathrm{Var}(X)$.

4. $\sigma(a \cdot X + b) = |a| \cdot \sigma(X)$.

Teorem 7. (**Čebiševljev teorem za diskretne slučajne varijable**) Neka je X diskretna slučajna varijabla s očekivanjem $E(X)$ i standardnom devijacijom σ. Tada za svaki $\alpha > 0$ vrijedi nejednakost:

$$P\left(X \in \left[E(X) - \alpha \cdot \sigma,\ E(X) + \alpha \cdot \sigma\right]\right) > 1 - \frac{1}{\alpha^2}.$$

4.1. Diskretna jednolika slučajna varijabla i diskretna jednolika razdioba

Kažemo da je varijabla X **jednolika** diskretna **slučajna varijabla** ako postoji konačan skup $S=\{s_1,...,s_n\}$ takav da je $R(X)=S$ i ako vrijedi:

$$P(X=k)=\frac{1}{n},\ \forall k\in S.$$

Pišemo: $X\sim U(n)$ i kažemo da X ima **jednoliku razdiobu**.

Očekivanje, **varijanca** i **standardna devijacija** varijable $X\sim U(n)$ su redom dani formulama:

$$\begin{cases} E(X)=\dfrac{n+1}{2}, \\ \mathrm{Var}(X)=\dfrac{n^2-1}{12}, \\ \sigma(X)=\dfrac{\sqrt{3\cdot(n^2-1)}}{6}. \end{cases}$$

4.2. Binomna slučajna varijabla i binomna razdioba

Neka su $n\in\mathbb{N}$ proizvoljan, ali fiksiran, te X slučajna varijabla koja označava ukupan broj uspjeha u n–terostrukom ponavljanju slučajnoga pokusa modeliranoga Bernoullijevom shemom (vidjeti stranicu 11.) Tada kažemo da je X **binomna slučajna varijabla s parametrima *n* i *p***.

Pišemo: $X\sim B(n,p)$ i kažemo da varijabla X ima **binomnu razdiobu**.

Slika varijable X je skup $\{0,1,...,n\}$. **Funkcija vjerojatnosti** varijable X je dana pravilom:

$$P(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k},\ \forall k=0,...,n.$$

Očekivanje, **varijanca** i **standardna devijacija** varijable $X\sim B(n,p)$ su redom dani formulama:

$$\begin{cases} E(X)=n\cdot p, \\ \mathrm{Var}(X)=n\cdot p\cdot(1-p), \\ \sigma(X)=\sqrt{n\cdot p\cdot(1-p)}. \end{cases}$$

4.3. Poissonova slučajna varijabla i Poissonova razdioba

Poissonova slučajna varijabla je prikladna za opis slučajnih pokusa u kojima vrijede sljedeće pretpostavke (tzv. **Poissonov model**):

1. Događaji su međusobno nezavisni.

2. Vjerojatnost da će se neki događaj dogoditi barem dvaput u dovoljno malom prostornom ili vremenskom intervalu je praktički zanemariva.

3. Vjerojatnost događaja je ista u dovoljno malim međusobno jednakim prostornim ili vremenskim intervalima.

Kažemo da je varijabla X **Poissonova slučajna varijabla s parametrom** $\lambda > 0$ ako je $R(X) = \mathbb{N} \cup \{0\}$ i ako je pripadna **funkcija vjerojatnosti** dana pravilom:

$$P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \text{ za } k \in \mathbb{N} \cup \{0\}.$$

Pišemo: $X \sim Po(\lambda)$ i kažemo da X ima **Poissonovu razdiobu s parametrom** λ.

Očekivanje, **varijanca i standardna devijacija** varijable $X \sim Po(\lambda)$ su redom dani formulama:

$$\begin{cases} E(X) = \mathrm{Var}(X) = \lambda, \\ \sigma(X) = \sqrt{\lambda}. \end{cases}$$

Za velike $n \in \mathbb{N}$ i male $p \in \langle 0,1 \rangle$ binomnu slučajnu varijablu $X \sim B(n,p)$ možemo dobro aproksimirati Poissonovom slučajnom varijablom s parametrom $\lambda = n \cdot p$. Preciznije, vrijedi sljedeća tvrdnja.

Teorem 8. (Poissonov teorem) Neka je $X \sim B(n,p)$. Ako istovremeno $n \to +\infty$, $p \to 0$ i $n \cdot p \to \lambda = \text{const.}$, onda vrijedi aproksimacija:

$$\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \approx \frac{\lambda^k}{k!} \cdot e^{-\lambda}.$$

Ovaj teorem se često koristi u sljedećem obliku.

Teorem 9. Neka su $n \in \mathbb{N}$ takav da je $n \geq 20$. Neka su $p \in \left\langle 0, \frac{10}{n} \right\rangle$ i $X \sim B(n,p)$. Tada vrijedi aproksimacija:

$$P(X = k) \approx \frac{(n \cdot p)^k}{k!} \cdot e^{-n \cdot p}.$$

4.4. Geometrijska slučajna varijabla i geometrijska razdioba

Pretpostavimo da neki slučajni pokus koji ima točno dva ishoda (*uspjeh* i *neuspjeh*) izvodimo sve dok se prvi put ne dogodi uspjeh. Neka je p vjerojatnost pojave uspjeha u *svakom* slučajnom pokusu. Neka je X slučajna varijabla koja označava ukupan broj ponavljanja pokusa sve do pojave uspjeha (uračunavajući i pokus čiji je ishod uspjeh).

Varijablu X nazivamo **geometrijska slučajna varijabla s parametrom *p***. Pišemo: $X \sim G(p)$ i kažemo da X ima **geometrijsku razdiobu**.

Slika varijable X je skup $\mathbb{N}$. Njezina **funkcija vjerojatnosti** je dana pravilom:

$$P(X = k) = (1-p)^{k-1} \cdot p, \ \forall k \in \mathbb{N}.$$

Pripadna **funkcija razdiobe vjerojatnosti** $F : \mathbb{N} \to [0,1]$ dana je pravilom:

$$F(k) = 1 - (1-p)^k .$$

Iz definicije funkcije F slijedi:

$$\left.\begin{aligned} P(X \le k) &= 1-(1-p)^k, \\ P(X > k) &= (1-p)^k, \end{aligned}\right\} \forall k \in \mathbb{N}.$$

Očekivanje, **varijanca** i **standardna devijacija** varijable $X \sim G(p)$ su redom dani formulama:

$$\begin{cases} E(X) = \dfrac{1}{p}, \\ \mathrm{Var}(X) = \dfrac{1-p}{p^2}, \\ \sigma(X) = \dfrac{\sqrt{1-p}}{p}. \end{cases}$$

Geometrijska razdioba može se karakterizirati sljedećim osnovnim svojstvom.

Teorem 10. Neka je X slučajna varijabla čija je slika skup $\mathbb{N}$. Tada je X geometrijska slučajna varijabla ako i samo ako za sve $k, l \in \mathbb{N}$ vrijedi:

$$P(X = k + l \mid X > k) = P(X = m).$$

Kažemo da geometrijska slučajna varijabla ima **svojstvo „zaboravljanja“**.

4.5. Hipergeometrijska slučajna varijabla i hipergeometrijska razdioba

Pretpostavimo da u skupu od $N \in \mathbb{N}$ elemenata njih točno $M \leq N$ ima obilježje (svojstvo) O (npr. „škart“ element i sl.). Iz promatranoga skupa odaberemo uzorak od $n \leq N$ elemenata. Neka je X slučajna varijabla koja označava ukupan broj elemenata iz odabranoga uzorka koji imaju obilježje O.

Varijablu X nazivamo **hipergeometrijska slučajna varijabla s parametrima N, M** i **n**. Pišemo: $X \sim H(N, M, n)$ i kažemo da X ima **hipergeometrijsku razdiobu.**

Slika varijable $X \sim H(N, M, n)$ je skup $S = \{k \in \mathbb{Z} : \max\{0, n+M-N\} \leq k \leq \min\{M, n\}\}$.

Pripadna **funkcija vjerojatnosti** je dana pravilom:

$$P(X=k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}, \ \forall k \in S.$$

Očekivanje, **varijanca** i **standardna devijacija** varijable $X \sim H(M, m, n)$ su redom dani formulama:

$$\begin{cases} E(X) = \dfrac{M}{N} \cdot n, \\ \mathrm{Var}(X) = \dfrac{M \cdot (N-M) \cdot (N-n) \cdot n}{N^2 \cdot (N-1)}, \\ \sigma(X) = \dfrac{1}{N} \cdot \sqrt{\dfrac{M \cdot (N-M) \cdot (N-n) \cdot n}{N-1}}. \end{cases}$$

Napomena: U slučaju velikoga broja elemenata osnovnoga skupa (N) i relativno maloga broja elemenata uzorka (M) hipergeometrijska slučajna varijabla $X_1 \sim H(M, m, n)$ može se dobro aproksimirati binomnom slučajnom varijablom $X_2 \sim B\left(n, \dfrac{M}{N}\right)$.

5. NEPREKIDNE (KONTINUIRANE) SLUČAJNE VARIJABLE

Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i neka je $S \subseteq \mathbb{R}$ neprebrojiv skup. (S je najčešće otvoreni/poluotvoreni/poluzatvoreni/zatvoreni interval, unija takvih intervala i sl.) Svaku slučajnu varijablu $X : \Omega \to S$ takvu da je $R(X) = S$ nazivamo **neprekidna** (**kontinuirana**) **slučajna varijabla**.

Ne istaknemo li drugačije, dalje u tekstu u ovoga poglavlja pretpostavljamo da je X neprekidna slučajna varijabla.

Za svaku varijablu X vrijedi:

$$P(X = s) = 0, \ \forall s \in S.$$

Varijabla X zadaje se svojom **funkcijom gustoće vjerojatnosti**. Preciznije, funkcija gustoće vjerojatnosti varijable X je funkcija $f : \mathbb{R} \to [0,1]$ koja ima sljedeća svojstva:

1. $f(x) \geq 0, \ \forall x \in \mathbb{R}.$

2. $P(a \leq X \leq b) = \int_a^b f(x) \cdot \mathrm{d}x.$

3. $\int_{-\infty}^{+\infty} f(x) \cdot \mathrm{d}x = 1.$

Funkcija f određuje vjerojatnost da varijabla X poprimi proizvoljnu vrijednost iz nekoga intervala.

Funkcija razdiobe vjerojatnosti varijable X je funkcija $F : \mathbb{R} \to [0,1]$ definirana pravilom:

$$F(x) = \int_{-\infty}^{x} f(t) \cdot \mathrm{d}t.$$

Funkcija F određuje vjerojatnost da varijabla X poprimi vrijednost ne veću od x. Preciznije, vrijedi jednakost:

$$F(x) = P(X \leq x) = P(X < x).$$

Funkcija F ima sljedeća svojstva:

1. $\lim_{x \to +\infty} F(x) = 1.$

2. $\lim_{x \to -\infty} F(x) = 0.$

3. $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = F(b) - F(a).$

4. $F'(x) = f(x), \ \forall x \in \mathbb{R}.$

Očekivanje, **varijanca** i **standardna devijacija** varijable X definirani su redom formulama:

$$\begin{cases} E(X) = \int\limits_{-\infty}^{+\infty} x \cdot f(x) \cdot \mathrm{d}x, \\ \mathrm{Var}(X) = \int\limits_{-\infty}^{+\infty} \left(x - E(X)\right)^2 \cdot f(x) \cdot \mathrm{d}x, \\ \sigma(X) = \sqrt{\mathrm{Var}(X)}. \end{cases}$$

Neka korisna svojstva očekivanja, varijance i standardne devijacije varijable X su iskazana sljedećim tvrdnjama.

1. $\mathrm{Var}(X) = E\left(\left(X - E(X)\right)^2\right) = E\left(X^2\right) - \left(E(X)\right)^2 = \left(\int\limits_{-\infty}^{+\infty} x^2 \cdot f(x) \cdot \mathrm{d}x\right) - \left(E(X)\right)^2.$

2. Neka su $A \subseteq S$, $g : A \to \mathbb{R}$ i $Y = g(X)$. Tada je Y neprekidna slučajna varijabla čije je očekivanje dano pravilom:

$$E(Y) = \int\limits_{-\infty}^{+\infty} g(x) \cdot f(x) \cdot \mathrm{d}x.$$

3. Neka su $a, b \in \mathbb{R}$ takvi da je $a \neq 0$. Neka je $Y = a \cdot X + b$. Tada je Y neprekidna slučajna varijabla čiji su očekivanje, varijanca i standardna devijacija redom dani izrazima:

$$\begin{cases} E(Y) = a \cdot E(X) + b, \\ \mathrm{Var}(Y) = a^2 \cdot \mathrm{Var}(X), \\ \sigma(Y) = |a| \cdot \sigma(X). \end{cases}$$

5.1. Neprekidna jednolika slučajna varijabla i neprekidna jednolika razdioba

Neka je X neprekidna slučajna varijabla takva da je $R(X) = [a,b]$, za neke $a, b \in \mathbb{R}$, $a < b$. Kažemo da je X **neprekidna jednolika slučajna varijabla** ako je njezina funkcija gustoće vjerojatnosti $f : \mathbb{R} \to [0,1]$ dana pravilom:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{za } x \in [a,b]; \\ 0, & \text{inače.} \end{cases}$$

Pišemo: $X \sim U(a, b)$ i kažemo da X ima **neprekidnu jednoliku razdiobu s parametrima a i b**.

Pripadna funkcija razdiobe vjerojatnosti $F:\mathbb{R}\to[0,1]$ je dana pravilom:

$$F(x)=\begin{cases}0, \text{ za } x<a;\\ \dfrac{x-a}{b-a}, \text{ za } x\in[a,b];\\ 1, \text{ za } x>b.\end{cases}$$

Očekivanje, **varijanca** i **standardna devijacija** varijable X su redom dani formulama:

$$\begin{cases}E(X)=\dfrac{a+b}{2},\\ \mathrm{Var}(X)=\dfrac{(b-a)^2}{12},\\ \sigma(X)=\dfrac{\sqrt{3}}{6}\cdot(b-a).\end{cases}$$

5.2. Eksponencijalna slučajna varijabla i eksponencijalna razdioba

Neka je X neprekidna slučajna varijabla takva da je $R(X)=[0,+\infty\rangle$. Neka je $a>0$. Kažemo da je X **eksponencijalna slučajna varijabla** ako je njezina funkcija gustoće vjerojatnosti $f:\mathbb{R}\to[0,1]$ dana pravilom:

$$f(x)=\begin{cases}a\cdot e^{-a\cdot x}, \text{ za } x>0;\\ 0, \text{ za } x\le 0.\end{cases}$$

Pišemo: $X\sim Ex(a)$ i kažemo da X ima **eksponencijalnu razdiobu.** Pripadna **funkcija razdiobe vjerojatnosti** $F:\mathbb{R}\to[0,1]$ je dana pravilom:

$$F(x)=\begin{cases}1-e^{-a\cdot x}, \text{ za } x>0,\\ 0, \text{ za } x\le 0.\end{cases}$$

Očekivanje, **standardna devijacija i varijanca** varijable X su redom dani izrazima:

$$\begin{cases}E(X)=\sigma(X)=\dfrac{1}{a},\\ \mathrm{Var}(X)=\dfrac{1}{a^2},\end{cases}$$

Može se pokazati da varijabla X ima tzv. **svojstvo „zaboravljanja“**, tj. da za sve $s,t>0$ vrijedi:

$$P(X>s+t\mid X>t)=P(X>s).$$

5.3. Normalna slučajna varijabla i normalna razdioba

Neka je X neprekidna slučajna varijabla takva da je $R(X)=\mathbb{R}$. Neka su $\mu\in\mathbb{R}$ i $\sigma>0$ konstante. Kažemo da je X **normalna slučajna varijabla** ako je njezina funkcija gustoće vjerojatnosti $f:\mathbb{R}\to[0,1]$ dana formulom:

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

Pišemo: $X\sim N\left(\mu,\sigma^2\right)$ i kažemo da X ima **normalnu razdiobu s parametrima μ i σ^2**. Pripadnu funkciju razdiobe vjerojatnosti F nije moguće zapisati analitički (pomoću zatvorene formule), pa se njezine vrijednosti određuju približno numeričkim metodama.

Očekivanje, **varijanca** i **standardna devijacija** varijable $X\sim N\left(\mu,\sigma^2\right)$ su redom dani izrazima:

$$\begin{cases} E(X)=\mu, \\ \mathrm{Var}(X)=\sigma^2, \\ \sigma(X)=\sigma. \end{cases}$$

Posebno, za $\mu=0$ i $\sigma=1$ dobivamo **standardnu** ili **jediničnu normalnu slučajnu varijablu.** Pišemo: $X\sim N(0,\ 1)$ i kažemo da X ima **standardnu** ili **jediničnu normalnu razdiobu**. Pripadna funkcija gustoće te varijable je funkcija $f:\mathbb{R}\to[0,1]$ definirana formulom:

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

Njezina funkcija razdiobe vjerojatnosti označava se s F^*. Vrijednosti te funkcije prikazane su u tablici 7. (vidjeti stranicu 32.)

Računanje vrijednosti funkcije razdiobe vjerojatnosti varijable $X\sim N\left(\mu,\sigma^2\right)$ svodi se na računanje vrijednosti funkcije razdiobe vjerojatnosti varijable $Y=\dfrac{X-\mu}{\sigma}\sim N(0,1)$. Varijablu Y nazivamo **standardizirana normalna varijabla.**

Neka korisna svojstva normalne slučajne varijable

Neka su F i F^* redom funkcija razdiobe vjerojatnosti varijable $X \sim N\left(\mu, \sigma^2\right)$, odnosno varijable $Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$. Tada vrijede sljedeće tvrdnje.

1. $P(X \geq 2 \cdot \mu - x) = P(X \leq x),\ \forall x \in \mathbb{R}.$

2. $P(X \leq 2 \cdot \mu - x) = P(X \geq x),\ \forall x \in \mathbb{R}.$

3. Ako su $a,\ b \in \mathbb{R}$ takvi da je $P(X \leq a) \leq F^*(b)$, onda vrijedi: $\mu \geq a - \sigma \cdot b.$

4. Ako su $a,\ b \in \mathbb{R}$ takvi da je $P(X \leq a) \geq F^*(b)$, onda vrijedi: $\mu \leq a - \sigma \cdot b.$

5. Ako su $a,\ b \in \mathbb{R}$ takvi da je $P(X \geq a) \leq F^*(b)$, onda vrijedi: $\mu \leq a + \sigma \cdot b.$

6. Ako su $a,\ b \in \mathbb{R}$ takvi da je $P(X \geq a) \geq F^*(b)$, onda vrijedi: $\mu \geq a + \sigma \cdot b.$

7. $F(x) = F^*\left(\dfrac{x-\mu}{\sigma}\right).$

8. $P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = F^*\left(\dfrac{b-\mu}{\sigma}\right) - F^*\left(\dfrac{a-\mu}{\sigma}\right).$

9. $F^*(-x) = 1 - F^*(x),\ \forall x \geq 0.$

Tablica 7. Tablica vrijednosti funkcije razdiobe vjerojatnosti F^* varijable $X \sim N(0,\ 1)$

x	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0,55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900

Uputa za primjenu tablice:

Znamenku jedinica i znamenku desetinki treba pronaći u recima, dok znamenku stotinki treba pronaći u stupcima. Npr. $F^*(1.23)$ očita se na presjeku retka **1.2** i stupca **0.03**:

$$F^*(1.23) = 0.89065.$$

Vrijednosti funkcije F^* za $x \in [-3, 0\rangle$ dobiju se korištenjem svojstva:

$$F^*(-x) = 1 - F^*(x),\ \forall x \in [0, 3].$$

Npr.

$$F^*(-1.23) = 1 - F^*(1.23) = 1 - 0.89065 = 0.10935.$$

5.4. Čebiševljeva nejednakost za neprekidne slučajne varijable. Pravilo $3\cdot\sigma$

Teorem 11. (**Čebiševljev teorem za neprekidne slučajne varijable**) Neka je X neprekidna slučajna varijabla čije je očekivanje $E(X)<+\infty$. Tada za svaki $\varepsilon>0$ vrijedi nejednakost:

$$P\left(\left|X-E(X)\right|\geq\varepsilon\right)\leq\frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Ova ocjena može se poboljšati.

Teorem 12. Neka su $X\sim N\left(\mu,\ \sigma^2\right)$ i $k\in\mathbb{N}$. Tada vrijedi:

$$P\left(\left|X-\mu\right|<k\cdot\sigma\right)=2\cdot F^*(k)-1.$$

Korolar 2. (**pravilo** $3\cdot\sigma$) Neka je $X\sim N\left(\mu,\sigma^2\right)$. Tada se u segmentu $\left[\mu-3\cdot\sigma,\ \mu+3\cdot\sigma\right]$ nalazi ukupno 99.73% svih vrijednosti varijable X.

5.5. Granični teoremi u Bernoullijevoj shemi

Uz određene uvjete binomnu slučajnu varijablu možemo aproksimirati normalnom slučajnom varijablom. Preciznije, vrijede sljedeći **granični teoremi**.

Teorem 13. (**lokalni de Moivrè – Laplaceov teorem**) Neka su $p\in\langle 0,\ 1\rangle$ i $(X_n)_{n\in\mathbb{N}}$ niz binomnih slučajnih varijabli $X_n\sim B(n,p)$. Tada za dovoljno velike $n\in\mathbb{N}$ vrijedi:

$$P(X_n=k)\approx\frac{1}{\sqrt{2\cdot\pi\cdot n\cdot p\cdot(1-p)}}\cdot e^{\frac{-(k-n\cdot p)^2}{2\cdot n\cdot p\cdot(1-p)}},\ \forall k\in\mathbb{N}.$$

Teorem 14. (**integralni de Moivrè – Laplaceov teorem**) Neka su $p\in\langle 0,\ 1\rangle$ i $(X_n)_{n\in\mathbb{N}}$ niz binomnih slučajnih varijabli $X_n\sim B(n,p)$. Tada za sve $a,\ b\in\mathbb{R}$ takve da je $a\leq b$ vrijedi:

$$P(a\leq X_n\leq b)\approx F^*\left(\frac{b-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right)-F^*\left(\frac{a-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right).$$

Napomena: Umjesto aproksimacije u teoremu 14. često se koristi bolja aproksimacija:

$$P(a\leq X_n\leq b)\approx F^*\left(\frac{b+\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right)-F^*\left(\frac{a-\frac{1}{2}-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right).$$

6. DODATAK

6.1. Pregled osnovnih pojmova teorije skupova

Kažemo da je skup A **podskup** skupa B i pišemo $A \subseteq B$ ako $(x \in A) \Rightarrow (x \in B)$, tj. ako svaki element skupa A pripada skupu B.

Kažemo da je skup A **pravi podskup** skupa B i pišemo $A \subset B$ ako je $A \subseteq B$ i ako postoji barem jedan $x \in B$ takav da $x \notin A$.

Neka je A bilo koji skup. Tada skup $\mathcal{P}(A) := \{X : X \subseteq A\}$ nazivamo **partitivni skup** skupa A. ($\mathcal{P}(A)$ je *skup svih podskupova* skupa A.)

Kažemo da su skupovi A i B **jednaki** i pišemo $A = B$ ako je $A \subseteq B$ i $B \subseteq A$. U suprotnom, kažemo da su skupovi A i B **različiti** i pišemo $A \neq B$.

Osnovne operacije sa skupovima:

- **presjek skupova**: $A \cap B := \{x : (x \in A) \wedge (x \in B)\}$ - skup kojega tvore elementi koji pripadaju skupu A i skupu B
- **unija skupova**: $A \cup B := \{x : (x \in A) \vee (x \in B)\}$ - skup kojega tvore elementi koji pripadaju skupu A ili skupu B (može i obama);
- **razlika skupova:** $A \setminus B := \{x : (x \in A) \wedge (x \notin B)\}$ - skup kojega tvore elementi koji pripadaju skupu A, ali ne pripadaju skupu B;
- **simetrična razlika skupova**: $A \Delta B := (A \setminus B) \cup (B \setminus A)$ - skup kojega tvore elementi koji pripadaju ili skupu A ili skupu B, ali ne i obama.
- **komplement skupa**: Ako su S *univerzalni* skup i $A \subseteq S$, onda je $A^C := \{x : (x \in S) \wedge (x \notin A)\}$ skup kojega tvore svi elementi skupa S koji ne pripadaju skupu A.

Kažemo da su skupovi A i B **disjunktni** ako je $A \cap B = \varnothing$, tj. ako nemaju nijedan zajednički element.

Kažemo da je *beskonačan* skup A **prebrojiv** ako postoji bijekcija $f : \mathbb{N} \to A$. Npr. skupovi $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, svi njihovi *beskonačni* podskupovi i unije takvih podskupova su prebrojivi.

Kažemo da je *beskonačan* skup A **neprebrojiv** ako postoji bijekcija $f : \mathbb{R} \to A$. Npr. skup iracionalnih brojeva, svi „karakteristični“ intervali (otvoreni, poluotvoreni, poluzatvoreni i zatvoreni) i unije takvih intervala su neprebrojivi.

6.2. Pregled formula iz diferencijalnoga i integralnoga računa

L'Hôpital–Bernoullijevo pravilo: $\lim\limits_{x\to c}\left(\dfrac{f(x)}{g(x)}\right)=\lim\limits_{x\to c}\left(\dfrac{f'(x)}{g'(x)}\right)$

Tablica 8. Tablica nekih standardnih antiderivacija

f	*standardna antiderivacija F*
$a,\ a\in\mathbb{R}$	$a\cdot x$
$x^n,\ n\neq -1$	$\dfrac{1}{n+1}\cdot x^{n+1}$
$a^x,\ a>0$	$\dfrac{1}{\ln a}\cdot a^x$
$e^{a\cdot x},\ a\neq 0$	$\dfrac{1}{a}\cdot e^{a\cdot x}$
$\sin(a\cdot x),\ a\neq 0$	$\left(-\dfrac{1}{a}\right)\cdot\cos(a\cdot x)$
$\cos(a\cdot x),\ a\neq 0$	$\dfrac{1}{a}\cdot\sin(a\cdot x)$

Formula za djelomičnu (parcijalnu) integraciju: $\int u\cdot \mathrm{d}v = u\cdot v-\int v\cdot \mathrm{d}u$

Neka osnovna svojstva integrala:

$$\int\big(f(x)\pm g(x)\big)\cdot \mathrm{d}x=\int f(x)\cdot \mathrm{d}x\pm\int g(x)\cdot \mathrm{d}x,$$
$$\int a\cdot f(x)\cdot \mathrm{d}x=a\cdot\int f(x)\cdot \mathrm{d}x.$$

Newton-Leibnizova formula:

$\int\limits_a^b f(x)\cdot \mathrm{d}x=F(b)-F(a)$, gdje je F standardna antiderivacija funkcije f.

Nepravi integrali:

$$\int\limits_{-\infty}^{b} f(x)\cdot \mathrm{d}x=\lim_{a\to-\infty}\left(\int\limits_a^b f(x)\cdot \mathrm{d}x\right),$$
$$\int\limits_{a}^{+\infty} f(x)\cdot \mathrm{d}x=\lim_{b\to+\infty}\left(\int\limits_a^b f(x)\cdot \mathrm{d}x\right),$$

6.3. Pregled nekih MATLAB-ovih funkcija koje se koriste u vjerojatnosti i statistici

Funkcija	*Značenje*
bar	crtanje jednostavnih stupaca
barh	crtanje jednostavnih redaka
betacdf	funkcija razdiobe vjerojatnosti beta slučajne varijable
betainv	inverz funkcije razdiobe vjerojatnosti beta slučajne varijable
betapdf	funkcija gustoće beta slučajne varijable
betastat	očekivanje i varijanca beta slučajne varijable
binocdf	funkcija razdiobe vjerojatnosti binomne slučajne varijable
binoinv	inverz funkcije razdiobe vjerojatnosti binomne slučajne varijable
binopdf	funkcija gustoće binomne slučajne varijable
binostat	očekivanje i varijanca binomne slučajne varijable
cdf	funkcija razdiobe vjerojatnosti određene slučajne varijable
expcdf	funkcija razdiobe vjerojatnosti eksponencijalne slučajne varijable
expinv	inverz funkcije razdiobe vjerojatnosti eksponencijalne slučajne varijable
exppdf	funkcija gustoće eksponencijalne slučajne varijable
expstat	očekivanje i varijanca eksponencijalne slučajne varijable
factorial	funkcija $n!$
geocdf	funkcija razdiobe vjerojatnosti geometrijske slučajne varijable
geoinv	inverz funkcije razdiobe vjerojatnosti geometrijske slučajne varijable
geomean	geometrijska sredina niza podataka
geopdf	funkcija gustoće geometrijske slučajne varijable
geostat	očekivanje i varijanca geometrijske slučajne varijable
harmmean	harmonijska sredina niza podataka
hist	crtanje histograma (negrupirani podaci)
histc	crtanje histograma (podaci grupirani u razrede)
hygecdf	funkcija razdiobe vjerojatnosti hipergeometrijske slučajne varijable
hygeinv	inverz funkcije razdiobe vjerojatnosti hipergeometrijske slučajne varijable
hygepdf	funkcija gustoće hipergeometrijske slučajne varijable
hygestat	očekivanje i varijanca hipergeometrijske slučajne varijable
icdf	inverz funkcije razdiobe vjerojatnosti određene slučajne varijable
iqr	interkvartil niza podataka
kurtosis	koeficijent zaobljenosti niza podataka

Funkcija	*Značenje*
mad	srednje apsolutno odstupanje niza podataka
max	najveći element niza podataka
mean	aritmetička sredina niza podataka; očekivanje neke slučajne varijable
median	medijan niza podataka
min	najmanji element niza podataka
mode	mod niza podataka
moment	središnji momenti niza podataka
nchoosek	binomni koeficijent $\binom{n}{k}$
normcdf	funkcija razdiobe vjerojatnosti normalne slučajne varijable
norminv	inverz funkcije razdiobe vjerojatnosti normalne slučajne varijable
normpdf	funkcija gustoće normalne slučajne varijable
normstat	očekivanje i varijanca normalne slučajne varijable
pdf	funkcija gustoće vjerojatnosti određene slučajne varijable
perms	ispis svih permutacija nekoga skupa
poisscdf	vrijednosti vjerojatnosti Poissonove slučajne varijable
poissoninv	inverz funkcije razdiobe vjerojatnosti Poissonove slučajne varijable
poisspdf	funkcija gustoće vjerojatnosti Poissonove slučajne varijable
poisstat	očekivanje i varijanca Poissonove slučajne varijable
prctile	percentili niza podataka
quantile	kvantili niza podataka
range	raspon varijacije niza podataka
skewness	koeficijent asimetrije niza podataka
std	standardna devijacija niza podataka ili neke slučajne varijable
tabulate	tablično grupiranje podataka
tiedrank	rang modaliteta
unidcdf	funkcija razdiobe vjerojatnosti diskretne jednolike slučajne varijable
unidcdf	funkcija razdiobe vjerojatnosti neprekidne jednolike slučajne varijable
unidinv	inverz funkcije razdiobe diskretne/neprekidne jednolike slučajne varijable
unidpdf	funkcija gustoće vjerojatnosti diskretne/neprekidne jednolike slučajne varijable
unidstat	očekivanje i varijanca diskretne jednolike slučajne varijable
var	varijanca niza podataka ili neke slučajne varijable
zscore	standardizirana vrijednost

6.4. Pregled nekih funkcija MS Excel-a koje se koriste u vjerojatnosti i statistici

Funkcija	*Značenje*
AVEDEV	srednje apsolutno odstupanje niza negrupiranih podataka
AVERAGE	aritmetička sredina negrupiranoga niza podataka
BETADIST	funkcija razdiobe vjerojatnosti beta slučajne varijable
BINOMDIST	funkcija razdiobe vjerojatnosti binomne slučajne varijable
COMBIN	binomni koeficijent
EXPONDIST	funkcija razdiobe vjerojatnosti eksponencijalne slučajne varijable
FACT	funkcija $n!$
FACTDOUBLE	funkcija $n!!$
GEOMEAN	geometrijska sredina negrupiranoga niza podataka
HARMEAN	harmonijska sredina negrupiranoga niza podataka
HYPGEOMDIST	funkcija razdiobe vjerojatnosti hipergeometrijske slučajne varijable
KURT	koeficijent zaobljenosti negrupiranoga niza podataka
MAX	najveća vrijednost u negrupiranom nizu podataka
MEDIAN	medijan negrupiranoga niza podataka
MIN	najmanja vrijednost u negrupiranom nizu podataka
MODE	mod negrupiranoga niza podataka
MULTINOMIAL	multinomni koeficijent
NORMDIST	funkcija razdiobe vjerojatnosti normalne slučajne varijable
NORMINV	inverz funkcije razdiobe vjerojatnosti normalne slučajne varijable
NORMSDIST	funkcija razdiobe vjerojatnosti standardne normalne slučajne varijable
NORMSINV	inverz funkcije razdiobe vjerojatnosti standardne normalne slučajne varijable
PERCENTILE	percentili negrupiranoga niza podataka
PERMUT	ukupan broj k-permutacija n-članoga skupa
POISSON	funkcija razdiobe vjerojatnosti Poissonove slučajne varijable
PROB	vjerojatnost da neki broj pripada određenom intervalu
QUARTILE	kvartili negrupiranoga niza podataka
RANK	rang modaliteta
SKEW	koeficijent asimetrije negrupiranoga niza podataka
STANDARDIZE	standardizirana vrijednost
STDEV	standardna devijacija uzorka podataka
STDEVP	standardna devijacija cijele populacije podataka

Funkcija	*Značenje*
VAR	varijanca uzorka podataka
VARP	varijanca cijele populacije podataka

POPIS KORIŠTENIH OZNAKA

Oznaka	*Značenje*
$:=$	definira se
$\mathbb{N}$	skup prirodnih brojeva
$\mathbb{Z}$	skup cijelih brojeva
$\mathbb{R}$	skup realnih brojeva
$\mathbb{C}$	skup kompleksnih brojeva
$\lvert\ \rvert$	apsolutna vrijednost
$\varnothing$	prazan skup
card	kardinalni broj skupa
$\subseteq$	podskup skupa
$\subset$	pravi podskup skupa
$\cap$	presjek skupova
$\cup$	unija skupova
$\setminus$	razlika skupova
^{c}	komplement skupa
$\langle a,b\rangle$	otvoreni interval, tj. $\{x\in\mathbb{R}: a<x<b\}$
$\langle a,b]$	poluotvoreni interval, tj. $\{x\in\mathbb{R}: a<x\le b\}$
$[a,b\rangle$	poluzatvoreni interval, tj. $\{x\in\mathbb{R}: a\le x<b\}$
$[a,b]$	segment, tj. $\{x\in\mathbb{R}: a\le x\le b\}$
$\sum_{i=1}^{n} x_i$	zbroj $x_1+\ldots+x_n$
$\prod_{i=1}^{n} x_i$	umnožak $x_1\cdot\ldots\cdot x_n$
$\lceil x\rceil$	najmanji cijeli broj jednak ili veći od x
$\times$	Kartezijev umnožak skupova
$n!$	umnožak prvih n prirodnih brojeva
$P(n,\ k)$	broj k-permutacija n-članoga skupa
$\binom{\sum_{i=1}^{n} k_i}{k_1,\ldots,k_n}$	multinomni koeficijent $\dfrac{\left(\sum_{i=1}^{n} k_i\right)!}{\prod_{i=1}^{n}(k_i!)}$
$\binom{n}{k}$	binomni koeficijent $\dfrac{n\cdot\ldots\cdot(n-k+1)}{k!}$
$\overline{C}(n,k)$	broj k-kombinacija s ponavljanjem n-članoga skupa

Oznaka	*Značenje*
f_i	apsolutna frekvencija modaliteta x_i
$f_i^<$	apsolutna frekvencija „manje od“ modaliteta x_i
$f_i^>$	apsolutna frekvencija „veće od“ modaliteta x_i
r_i	relativna frekvencija modaliteta x_i
$r_i^<$	relativna frekvencija „manje od“ modaliteta x_i
$r_i^>$	relativna frekvencija „veće od“ modaliteta x_i
A	aritmetička sredina numeričkoga niza
Me	medijan niza
Mo	mod niza
Q_1	1. ili donji kvartil
Q_3	3. ili gornji kvartil
R	raspon varijacije niza
I_q	Interkvartil
V_q	koeficijent kvartilne devijacije
MAD	srednje apsolutno odstupanje
σ^2	varijanca (disperzija) numeričkoga niza
σ	standardna devijacija
V	koeficijent varijacije
E	(matematičko) očekivanje slučajne varijable
Var	varijanca slučajne varijable
D	prirodna domena funkcije
F^*	funkcija gustoće standardne normalne razdiobe

POPIS TABLICA

1. Tablica 1. Skala intenziteta varijabiliteta koeficijenta kvartilne devijacije 16
2. Tablica 2. Skala intenziteta varijabiliteta koeficijenta varijacije 16
3. Tablica 3. Određivanje potpunih srednjih vrijednosti 18
4. Tablica 4. Određivanje percentila iz (ne)grupiranih podataka 18
5. Tablica 5. Određivanje kvartila iz (ne)grupiranih podataka 19
6. Tablica 6. Određivanje mjera raspršenja (disperzije) 20
7. Tablica 7. Tablica vrijednosti funkcije razdiobe vjerojatnosti standardne normalne slučajne varijable 33
8. Tablica 8. Tablica nekih standardnih antiderivacija 35

KAZALO POJMOVA

algebra događaja 7
apsolutna frekvencija 12
- "manje od" 12
- "veće od" 12
- korigirana 14
- kumulativni niz 12

Bernoullijeva shema 11
binomna razdioba 23
binomna slučajna varijabla 23
- funkcija vjerojatnosti 23
- očekivanje 23
- slika 23
- standardna devijacija 23
- varijanca 23

binomni koeficijent 6
Čebiševljeva nejednakost 33
Čebiševljevo pravilo u opisnoj statistici 17
de Morganove formule 8
decili 14
diskretna jednolika slučajna varijabla 23
- očekivanje 23
- standardna devijacija 23
- varijanca 23

diskretna slučajna varijabla 21
- funkcija razdiobe vjerojatnosti 21
- funkcija vjerojatnosti 21
- očekivanje 22
- standardna devijacija 22
- tablica razdiobe 21
- varijanca 22

događaj 7
- elementarni 7
- nemoguć 7
- siguran 7
- suprotan 7

događaji	
- disjunktni	7
- nezavisni	9, 10
- operacije s događajima	7
duljina statističkoga niza	12
eksponencijalna slučajna varijabla	29
- funkcija gustoće	29
- funkcija razdiobe vjerojatnosti	29
- očekivanje	29
- standardna devijacija	29
- varijanca	29
formula	
- Bayesova	10
- potpune vjerojatnosti	10
geometrijska slučajna varijabla	25
- funkcija razdiobe vjerojatnosti	25
- funkcija vjerojatnosti	25
- očekivanje	25
- slika	25
- standardna devijacija	25
- varijanca	25
hipoteze	10
hipergeometrijska slučajna varijabla	26
- funkcija razdiobe vjerojatnosti	26
- funkcija vjerojatnosti	26
- očekivanje	26
- slika	26
- standardna devijacija	26
- varijanca	26
histogram	13
k-kombinacija	
- bez ponavljanja	6
- s ponavljanjem	6
k-permutacija	
- bez ponavljanja	6
- s ponavljanjem	6

kvartil
- donji 15, 19
- gornji 15, 19

medijan 15, 19
mjere raspršenja 15
- apsolutne 15
- relativne 15

mod 14
multinomni koeficijent 6
multiskup 5
modalitet obilježja 12
neprekidna jednolika slučajna varijabla 28
- funkcija gustoće 28
- funkcija razdiobe vjerojatnosti 28, 29
- očekivanje 29
- standardna devijacija 29
- varijanca 29

neprekidna slučajna varijabla 27
- funkcija gustoće vjerojatnosti 27
- funkcija razdiobe vjerojatnosti 27
- očekivanje 28
- standardna devijacija 28
- varijanca 28

normalna slučajna varijabla 30
- funkcija gustoće 30
- očekivanje 30
- standardizirana 30
- standardna devijacija 30
- varijanca 30

odstupanje od aritmetičke sredine 15
percentili 14, 18
permutacija
- bez ponavljanja 5
- multiskupa 5
- s ponavljanjem 5

Poissonov model 24

Poissonova slučajna varijabla 24
- očekivanje 24
- standardna devijacija 24
- varijanca 24

poligon apsolutnih/relativnih frekvencija 13
potpuni sustav događaja 10
teorem
- binomni 6
- Čebiševljev za diskretne slučajne varijable 22
- Čebiševljev za neprekidne slučajne varijable 33
- integralni de Moivrè-Laplaceov 33
- lokalni de Moivrè-Laplaceov 33
- o uzastopnom prebrojavanju 5
- Poissonov 24

pravilo $3\cdot\sigma$ 33
pravilo jednakosti (bijekcije) 5
princip zbroja 5
prostor elementarnih događaja 7
raspon varijacije 15
razdioba
- bimodalna 14
- binomna 23
- diskretna jednolika 23
- eksponencijalna 29
- geometrijska 25
- hipergeometrijska 26
- multimodalna 14
- neprekidna jednolika 28, 29
- normalna 30
- Poissonova 24
- standardna normalna 30
- unimodalna 14

relativna frekvencija 12
- "manje od" 13
- "veće od" 12
- korigirana 14

- kumulativni niz 13

slučajna varijabla

- diskretna 21
- neprekidna 27

slučajni pokus 7

sredina

- aritmetička 13, 18
- geometrijska 13, 18
- harmonijska 13, 18

srednje apsolutno odstupanje 15, 20

srednje vrijednosti 13

- položajne 13
- potpune 13

standardizirano obilježje 17

standardna devijacija numeričkoga niza 16, 20

standardna (jedinična) normalna slučajna varijabla 30

- funkcija gustoće 30
- očekivanje 30
- standardna devijacija 30
- varijanca 30

svojstvo „zaboravljanja“ 25, 29

varijanca numeričkoga niza 15, 20

vjerojatnosni prostor 8

- klasičan 8

vjerojatnost 8

- definicija 8
- geometrijska 9
- uvjetna 9

zatvoreni razred 13

- donja granica 13
- gornja granica 13
- nepravi 13
- pravi 13
- sredina 13
- širina 13

LITERATURA

1. N. Sarapa: *Teorija vjerojatnosti*, Školska knjiga, Zagreb, 2002.

2. M. Benšić, N. Šuvak: *Uvod u vjerojatnost i statistiku*, Odjel za matematiku Sveučilišta u Osijeku, Osijek, 2014.

3. N. Elezović: *Diskretna vjerojatnost*, Element, Zagreb, 2007.

4. N. Elezović: *Slučajne varijable*, Element, Zagreb, 2007.

5. Ž. Pauše: *Uvod u matematičku statistiku*, Školska knjiga, Zagreb, 1993.

6. S. Suljagić: *Vjerojatnost i statistika*, interna skripta, Tehničko veleučilište u Zagrebu, Zagreb, 2003.

7. D. Veljan: *Kombinatorna i diskretna matematika*, Algoritam, Zagreb, 2001.