9 STATISTIKA I VJEROJATNOST

SAŽETAK:

Cilj ovog poglavlja je upoznati studente pomorskih studija s osnovnim pojmovima teorije vjerojatnosti i matematičke statistike te njihovim primjenama.
Prvi dio posvećen je teoriji vjerojatnosti s obzirom da je ona temelj za razumijevanje modela potrebnih za statističke analize. Započinje osnovnim pojmom teorije vjerojatnosti. Nakon toga definiraju se pojmovi uvjetne vjerojatnosti, Bayesovo pravilo i nezavisnost događaja. U nastavku se uvodi pojam slučajnih varijabli (diskretnih i neprekidnih) i njihova najosnovnija obilježja (funkcije vjerojatnosti, gustoće i raspodjele, očekivanje, varijanca i standardna devijacija). Predstavljeni su primjeri binomne, Poissonove i normalne raspodjele i njihova primjena.

Drugi dio je posvećen statistici i napisano je s ciljem razumijevanja koncepata koji se koriste u statističkim analizama. Započinje pregledom i diskusijom osnovnih pojmova deskriptivne statistike. U nastavku se uvodi inferencijalna statistika testiranje statističkih hipoteza i metoda regresijske analize. Prikazana je primjena Poissonove distribucije u teoriji pouzdanosti. Ovo poglavlje završava skupom zadataka za vježbu, individualne aktivnosti i učenje.

CILJ: **Upoznavanje s osnovnim pojmovima teorije vjerojatnosti i statistike te stjecanje vještina kako bi u praksi mogli obraditi i protumačiti statističke podatke poput AIS ili satelitskih podataka, ili podataka prikupljeni mjerenjem, testom ili anketom i kako bi uspješno primijenili elemente statističkog zaključivanja u stručnim kolegijima npr. *Dijagnostika kvarova, Teorija pouzdanosti u održavanju brodskih sustava, Upravljanje kvalitetom u pomorstvu, Upravljanje rizicima…***

Ishodi učenja:

1. Računati vjerojatnost s klasičnim definicijama koristeći osnovne principe brojanja
2. Računati ukupnu i uvjetnu vjerojatnost
3. Provjeriti neovisnost događaja kroz primjene
4. Procijeniti karakteristična svojstva distribucije promatranih podataka
5. Koristiti i interpretirati osnovne statističke pokazatelje
6. Primijeniti osnovne metode deskriptivne i inferencijalne statistike na konkretnom primjeru
7. Provesti jednostavnu statističku analizu prikupljenih podataka te pravilno interpretirati dobivene rezultate.
8. Izračunati parametre regresijskoga modela
9. Odabrati statističku metodologiju primjerenu istraživanom problem u pomorstvu

Co-funded by the Erasmus+ Programme of the European Union

Potrebna predznanja iz matematike:

Elementarna matematika, aritmetika, teorija skupova, diferencijalni i integralni račun realnih funkcija te geometrije ravnine i prostora.

Povezanost s rješavanjem problema iz pomorske domene:

Razlog zašto u pomorskim problemima primjenjujemo probabilističke i statističke metode je taj što su ti problemi često nepredvidivi. Kretanje brodova u moru je djelomično nasumično. Djelomično znači da, naravno, možemo upravljati brodom i reagirati na stvarno stanje, ali u mnogim iznimno teškim situacijama to nije dovoljno. Ne možemo dovoljno rano predvidjeti mnoga potencijalno rizična stanja. Jer more ne možemo opisati na deterministički način. Morski valovi nisu deterministički. Ne postoje "jednadžbe mora" koje mogu opisati kako će se valovi ponašati na nekom mjestu i u vrijeme, čak i ako poznajemo trenutnu situaciju tamo.
Drugi razlog zašto su vjerojatnost i statistika toliko važni u pomorstvu je teorija pouzdanosti koju primjenjujemo na brodsku opremu. Brodovi obično provode duge dane u moru, stoga moramo biti sigurni da će svi mehanički, električni i elektronički uređaji ispravno raditi.

Co-funded by the
Erasmus+ Programme
of the European Union

Sadržaj

Co-funded by the
Erasmus+ Programme
of the European Union

9.1 OSNOVNI PRINCIPI PREBROJAVANJA

9.1.1 Princip umnoška (princip produkta)

Ako skup A sadrži *n* elemenata, skup B sadrži *m* elemenata, tada postoji ***n·m*** različitih uređenih parova (*a*, *b*) pri čemu element a pripada skupu A, a element b pripada skupu B.

Primjer 9.1

Koliko ima dvoznamenkastih parnih prirodnih brojeva?

Rješenje:

Dvoznamenkasti parni prirodni brojevi mogu se predstaviti kao uređeni parovi (a, b), gdje je $a \in \{1,2,3,4,5,6,7,8,9\} = A$ i $b \in \{0,2,4,6,8\} = B$. Broj elemenata u skupu A je n=9, a broj elemenata u skupu B je m=5. Dakle, ukupan broj dvoznamenkastih parnih prirodnih brojeva je

$$9 \cdot 5 = 45.$$

VAŽNO Budući se uređena trojka (a, b, c) mogu predstaviti kao uređen par (a, b, c), možemo proširiti princip množenja na slučaj tri skupa.

9.1.2 Princip umnoška (tri i više skupa)

Ako skup A sadrži n elemenata, skup B sadrži m elemenata, a skup C sadrži p elemenata, tada postoji $n \cdot\ m \cdot p$ različitih uređenih trojki (a, b, c), gdje a pripada A, b pripada B i c pripadaju C. Slično pravilo vrijedi za četiri, pet i više skupova.

Primjer 9.2

Koliko ima troznamenkastih neparnih prirodnih brojeva?

Rješenje:

Troznamenkasti neparni prirodni brojevi mogu se predstaviti kao uređene trojke (a, b, c), gdje je $a \in \{1,2,3,4,5,6,7,8,9\} = A, b \in \{0,1,2,3,4,5,6,7,8,9\} = B$ i $c \in \{1,3,5,7,9\} = C$. Broj elemenata u skupu A je $n = 9$, broj elemenata u skupu B je $m = 10$ i broj elemenata u skupu C je $p = 5$. Dakle, ukupan broj troznamenkastih neparnih prirodnih brojeva je

$$9 \cdot 10 \cdot 5 = 450.$$

Princip zbroja

Ako skup A sadrži n elemenata, skup B sadrži m elemenata i A∩B sadrži l elemenata, tada A∪B sadrži $n + m - l$ elemenata.

Primjer 9.3

Alice je kupila u trgovini: jabuke, šljive, banane, borovnice i naranče. Njezin suprug John kupio je, u drugoj trgovini, kupine, ogrozd, grožđe, šljive, kruške, grejp i banane. Koliko će različitih vrsta voća pojesti danas kod kuće?

Co-funded by the Erasmus+ Programme of the European Union

Rješenje:

Ako je

$n = 5,\ m = 7,\ l = 2.$

Onda , Alice i njen suprug imaju

$$m + n - l = 5 + 7 - 2 = 10 \text{ različitih vrsta voća.}$$

Definicija: Permutacija

Permutacija je svaki raspored elemenata konačnog skupa u određenom redoslijedu.

Pri permutiranju svih elemenata skupa od n elemenata redoslijed elemenata je važan.

Broj permutacija

Broj svih permutacija skupa od $\boldsymbol{n}$ elementa je

$$\boldsymbol{P_n = n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1}$$

$\boldsymbol{n}$! *se zove faktorijel od* $\boldsymbol{n}$.

Primjer 9.4

Na koliko načina možemo posložiti pet različitih knjiga na policu.

Rješenje:

Potrebno je izračunati broj svih permutacija skupa od 5 elemenata. Taj broj je jednak umnošku prvih 5 prirodnih brojeva, odnosno

$$P_5 = 5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120.$$

Zadatak 9.1

Na koliko različitih načina sedam različitih brodova može pristati u luci?

Rješenje: $7! = 5040$.

9.1.3 Varijacija bez ponavljanja k-tog razreda u n-članom skupu

Definicija: Varijacija

Varijacija bez ponavljanja ***k***-tog razreda u ***n***-članom skupu svaka je uređena ***k***-torka različitih elemenata danog skupa $(k \leq n)$.

Broj varijacija bez ponavljanja r-tog razreda od n

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

Primjer 9.5

Na koliko različitih načina 5 osoba može sjesti na 11 stolica?

Rješenje:

Neka je N broj uređenih 5-orki (permutacija) uzetih iz skupa od 11 elemenata. Onda je ,

$n = 11,\ k = 5$ i

$$N = P(11,5) = \frac{11!}{(11-5)!} = \frac{11!}{6!} = 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 = 55440.$$

Primjer 9.6

U luci je devet (9) slobodnih vezova. Na koliko različitih načina se šest (6) brodova može rasporediti na slobodne vezove?

Rješenje:

Neka je N broj permutacija od 6 elemenata iz skupa od 9 članova. Onda je,

$n = 9,\ k = 6$ i

$$N = P(9,6) = \frac{9!}{(9-6)!} = \frac{9!}{3!} = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 60480.$$

9.1.4 Varijacije s ponavljanjem

Različite uređeni nizovi duljine k koji se mogu formirati iz skupa od n elemenata, gdje se elementi mogu ponavljati, nazivaju se *varijacijama s ponavljanjima*.

Broj **permutacija s ponavljanjem** skupa od n elemenata je

$$n^k.$$

U varijacijama s ponavljanjem, kao i u varijacijama bez ponavljanja, važan je poredak elemenata (uređene *k*-torke).

Primjer 9.7

Koliko riječi od 3 slova se može napraviti od slova a, b, c, d?

Rješenje:

Treba izračunati broj N različitih nizova (riječi) duljine 3 uzetih iz skupa od 4 elementa $\{a, b, c, d\}$, gdje je dopušteno ponavljanje slova. Tako je ,

$$n = 4,\ k = 3 \qquad \text{i} \qquad N = 4^3 = 64.$$

Primjer 9.8

Koliko riječi od 4 slova se može napraviti od slova a, b, c?

Rješenje:

Co-funded by the Erasmus+ Programme of the European Union

Treba izračunati broj N različitih nizova (riječi) od 4 slova uzetih iz 6-članog skupa elementa $\{a, b, c\}$, gdje je dopušteno ponavljanje slova. Tako je ,

$$n = 3,\ k = 4 \qquad \text{i} \qquad N = 3^4 = 81$$

Remark

U nekim zadacima formulu za broj svih permutacija s ponavljanjima nije jednostavno primijeniti. Lakše je koristiti princip množenja.

Primjer 9.9

Na koliko načina pet brodova iz zemlje A može pristati u tri različite luke u zemlji B?

Rješenje:

1. način:

- Neka su luke u zemlji B označene s h1, h2, h3.
- Svaki brod ima tri mogućnosti izbora: pristati u h1, u h2 ili u h3.
- Koristeći princip množenja, vidimo da je:
  - dva broda imaju 3·3=9 mogućnosti izbora,
  - tri broda imaju 3·3·3=27 mogućnosti izbora,
  - četiri broda imaju 3·3·3 ·3=81 mogućnosti izbora.
- Konačno, pet brodova ima

$$N = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5 = 243$$

mogućnosti izbora.

2. način:

Imamo nizove duljine $k = 5$ (pet brodova) koji se sastoje od brojeva iz skupa od 3 elementa $\{1,2,3\}$ (luke) $n = 3$. Ponavljanja su dopuštena jer dva ili više brodova mogu izabrati istu luku. Tako,

$$N = 3^5 = 243.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.1.5 Kombinacije

Definicija: Kombinacije bez ponavljanja

Kombinacija k –-tog razreda od n elemenata $(k \leq n)$ je svaki podskup od k – elemenata n -članog skupa.

Ukupan broj kombinacija bez ponavljanja k -tog razreda od n elemenata jednak je .

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!\,(n-k)!},$$

gdje je $\binom{n}{k}$ *binomni koeficijent*

Primjer 9.10

Anne, Beth, Charlie i Donald pokušavaju osvojiti dvije karte za Bahame.

a) Koje su sve kombinacije dobitnika?

b) Koji je broj svih mogućnosti za odabir dobitnika?

c) Ako su izabrani Anne i Charlie, je li važno tko je dobio prvu, a tko drugu kartu?

Rješenje:

a) Sve kombinacije dobitnika su:

$$\{A,B\},\{A,C\},\{A,D\},\{B,C\},\{B,D\},\{C,D\},$$

gdje A označava Anne, B Beth itd.

b) Broj svih mogućnosti za odabir dobitnika između 4 kandidata je:

$$N = C(4,2) = \binom{4}{2} = \frac{4!}{2!\,(4-2)!} = \frac{4!}{2!\,2!} = \frac{4\cdot 3}{2\cdot 1} = 6.$$

c) Ne, nije važno tko je dobio prvu, a tko drugu kartu. Vrijedi $\{A,C\} = \{C,A\}$.

Remark

Vrijede sljedeće formule za binomni koeficijent:

- $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n(n-1)\cdots(k+1)}{(n-k)!}$
- $\binom{n}{k} = \binom{n}{n-k}$
- $\binom{n}{0} = \binom{n}{n} = 1$
- $\binom{n}{1} = \binom{n}{n-1} = n$

Co-funded by the
Erasmus+ Programme
of the European Union

Zadatak 9.2

1. Izračunati $\binom{4000}{3998}$.

2. Koliko podskupova od 2 elementa iz skupa od četiri tisuće elemenata se može napraviti?

Rješenje:

1. $7\ 998\ 000$

2. $7\ 998\ 000$

Primjer 9.11

Koliko različitih povjerenstava od 4 člana se može sastaviti u klubu od 10 članova?

Rješenje:

Povjerenstvo je 4 –člani podskup od 10 –članog kluba. Dakle, broj N različitih povjerenstava je jednak $C(10,4)$. Vrijedi

$$N = C(10,4) = \binom{10}{4} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4!} = 210.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.2 DOGAĐAJ. VJEROJATNOST

Računanje vjerojatnosti je vezano uz neki ***pokus***, eksperiment ili simulaciju. Promatramo rezultat pokusa ili ***ishod***.

Primjer 9.12

Pretpostavimo da izvedimo pokus bacanja dva različita novčića. Neka H predstavlja glavu, a T pismo. Tada je svaki mogući ishod bacanja novčića element skupa

$$\Omega = \{(H,H),(H,T),(T,H),(T,T)\}.$$

Definicija: Elementarni događaj. Prostor elementarnih događaja.

Za pokus kod kojeg ne znamo unaprijed koji će se od mogućih ishoda pojaviti, kažemo da je ***slučajni pokus***, a njegove ishode zovemo ***elementarni događaji***.
Prostor elementarnih događaja *je* skup Ω koji kao elemente sadrži sve što se može realizirati kada se izvede pokus (ishodi pokusa) tj. skup svih elementarnih događaja slučajnog pokusa. Svako element skupa Ω naziva se elementarni događaj.

Jedina pretpostavka na prostor elementarnih događaja jest da je neprazan i da zaista sadrži sve što se može realizirati u pokusu.

Definicija: Događaj

Događaj je svaki podskup prostora elementarnih događaja.

Napomena

Prostor elementarnih događaja Ω je događaj (*siguran događaj*), svaki element prostora elementarnog događaja je događaj, prazni skup $\emptyset$ je događaj (*nemogući događaj*).

Primjer 9.13

Svako slovo riječi *leopard* napiše se na posebnu karticu i karte se miješaju. Navedite prostor elementarnih događaja za ishod izvlačenja jedne karte.

Odgovor: Prostor elementarnih događaja je $\Omega = \{l,e,o,p,a,r,d\}$ (također $\Omega = \{e,p,l,a,d,o,r\}$ jer redoslijed nije bitan).

Sada pretpostavimo da nas zanima je li izvučeno slovo samoglasnik. Izvlačenje samoglasnika nazivamo događaj A. Događaj A je pojava bilo kojeg od ishoda e,o,a. Vrijedi da je događaj $A = \{e,o,a\}$ podskup prostora Ω.

Slučajni događaj je takav događaj koji se može, ali ne mora realizirati, tj. realizira se uz određenu vjerojatnost.

Co-funded by the Erasmus+ Programme of the European Union

Definicija: Vjerojatnost (Klasičan pristup)

Neka je Ω konačni prostor slučajnih događaja pokusa s n mogućih ishoda, i neka su svi elementarni događaji jednako mogući. Ako je događaj A podskup od Ω (A⊂Ω) događaj takav da A sadrži k elemenata, tada je vjerojatnost događaja A, koja se označava $P(A)$, dana izrazom

$$P(A) = \frac{k}{n}. 0 \leq P \leq 1.$$

Za siguran događaj: 1=⇒ P (obrat ne vrijedi).

Za nemoguć događaj: 0=⇒ P (obrat ne vrijedi).

Primjer 9.14

Pretpostavimo da izvodimo pokus bacanja dvije različite šesterostrane kocke. Koja je vjerojatnost da ukupan ishod bacanja bude deset?

Rješenje:

Prostor elementarnih događaja za bacanje 2 kocke je Ω={(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),...,(5,6),(6,6)}. Broj elementarnih događaja je $n = 6 \cdot 6 = 36.$ Neka A označava događaj da je ukupan ishod (zbroj ishoda bacanja kockica) deset. Vrijedi,

$$A = \{(6,4),(5,5),(4,6)\}$$

Kako je $k = 3$ slijedi da je tražena vjerojatnost

$$P(A) = \frac{3}{36} = \frac{1}{12}.$$

Vjerojatnost da se slučajni događaj A ne realizira jednaka je omjeru broja nepovoljnih ishoda i svih mogućih ishoda:

$$P(\bar{A}) = \frac{n-k}{n} = 1 - P(A), 0 \leq P \leq 1.$$

gdje je:

n - broj svih mogućih ishoda

k - broj svih povoljnih ishoda za događaj A

$(n-k)$ - broj svih nepovoljnih ishoda za događaj A.

.

Neka je Ω prostor slučajnih događaja konačan ili beskonačan. *Familiju* svih događaja označavamo s F i nazivamo je *algebrom događaja*, a pojedine elemente te familije događaja zovemo *jednostavno događajima.* Ako je Ω konačan, svi njegovi podskupovi su događaji. Ako Ω nije konačan, smatramo događaje kao veliku beskonačnu familiju podskupova od Ω. Na odabranoj familiji događaja definirat ćemo vjerojatnost aksiomatskom definicijom.

Co-funded by the
Erasmus+ Programme
of the European Union

Definicija: Vjerojatnost (Aksiomatski pristup)

Vjerojatnost događaja A je realan broj $P(A)$ koji zadovoljava sljedeća tri uvjeta:

1. $P(A) \geq 0$
2. $P(\Omega) = 1$
3. Za svaku prebrojivu familiju $A_i, A_2, \ldots \subseteq \mathrm{F}$ međusobno isključivih događaja (oni događaji koji ne mogu nastupiti istodobno, odnosno čiji su pripadni skupovi elementarnih događaja disjunktni tj. Ai ∩ Aj = ∅ čim je $i \neq j$,), vjerojatnost realizacije jednog ili drugog događaja jednaka je zbroju vjerojatnosti realizacije svakog pojedinog događaja. Vrijedi
$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \cdots .$$

Uvjete 1. - 3. nazivamo aksiomima vjerojatnosti.

Svojstva vjerojatnosti

1. $P(\emptyset) = 0$ (vjerojatnost nemogućeg događaja);
2. $0 \leq P(A) \leq 1$
3. $P(A \cup B) = P(A) + P(B)$, ako su A i B međusobno isključivi;
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, (za događaje koji se međusobno ne isključuju vjerojatnost da nastupi barem jedan od događaja A ili B)
5. $A \subset B \Rightarrow P(A) \leq P(B)$
6. $P(\bar{A}) = 1 - P(A)$
7. $P(A \setminus B) = P(A) - P(A \cap B)$
8. $B \subset A \Rightarrow P(A \setminus B) = P(A) - P(B)$
9. $$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Teorem o množenju vjerojatnosti (multiplikacijski teorem)

Ako se događaji A i B međusobno ne isključuju, tj. vjerojatnost realizacije jednog ne zavisi o vjerojatnosti realizacije drugog događaja (nezavisni događaji), tada je vjerojatnost istovremene realizacije događaja A i B jednaka umnošku vjerojatnosti realizacije događaja A i događaja B:
$$P(A \cap B) = P(A) \cdot P(B).$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.3 GEOMETRIJSKA VJEROJATNOST

Govorit ćemo o geometrijskoj interpretaciji zadavanja vjerojatnosti koja je vrlo primjenjiva u rješavanju različitih zadataka.

Promatrat ćemo jedino one neprebrojive skupove elementarnih događaja Ω koji imaju konačnu geometrijsku mjeru **m(Ω)** kao npr. duljina, površina, obujam.

Definicija: Geometrijska vjerojatnost

Vjerojatnost događaja A ⊆ Ω, tj. vjerojatnost da je slučajno odabrana točka x iz A, jednaka je omjeru mjere od A i mjere od Ω:

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

Primjer 9.15

U kvadrat je upisana kružnica. Točka se nasumično odabire iz površine kvadrata. Izračunajte vjerojatnost da se nalazi unutar kruga.

Rješenje:

Neka je Ω kvadrat i A je krug radijusa $r > 0$ upisan unutar Ω. Vrijedi

$$m(A) = \pi r^2, \qquad m(\Omega) = 4r^2.$$

$$P(A) = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}$$

Primjer 9.16

Dva broda imaju vrijeme dolaska između 14 i 15 sati. Oba broda moraju pristati na isti vez. Nakon pristajanja, svakom brodu treba 10 minuta da obnovi zalihe i napusti dok. Kolika je vjerojatnost da brodovi neće morati čekati na vez?

Rješenje:

Kao jedinicu vremena postavili smo 10 minuta. Vremenski interval između 14 i 15 sati označimo s $[0,6]$. Neka je x vrijeme kada će prvi brod stići, a y vrijeme kada će drugi brod stići. Tada vrijedi $x, y \in [0,6]$ i $|x - y| \geq 1$.

Neka je $\Omega = [0,6] \times [0,6]$ prostor slučajnih događaja. Površina Ω je μ(Ω)=36.

Co-funded by the
Erasmus+ Programme
of the European Union

Događaj da oba broda neće morati čekati može se predstaviti skupom

$$A = \{(x,y) \in \Omega : |x-y| \geq 1\}.$$

Budući je

$$A = \{(x,y) \in \Omega : y \leq x-1 \text{ or } y \geq x+1\},$$

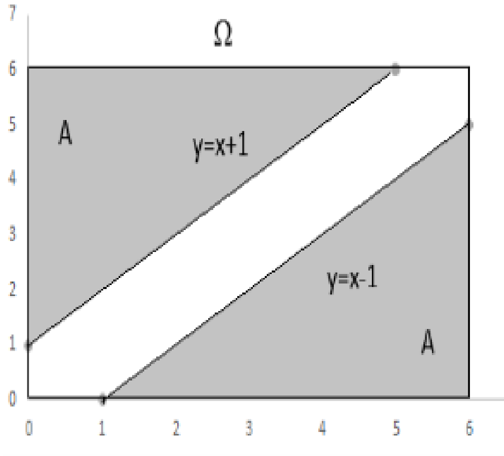
Površina skupa A je

$$\mu(A) = \frac{1}{2}\cdot 5 \cdot 5 + \frac{1}{2}\cdot 5 \cdot 5 = 25.$$

Primjenom formule za geometrijsku vjerojatnost dobije se

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{25}{36} \approx 0.6944.$$

Postoji 69.44% šanse da brodovi ne moraju čekati na vez.



Slika 9.1 . ilustracija primjera 9.16

Co-funded by the
Erasmus+ Programme
of the European Union

9.4 UVJETNA VJEROJATNOST. BAYESOVA FORMULA

Definicija: Nezavisni događaji

Dva događaja A i B se nazivaju *nezavisni* (statistički nezavisni) ako

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

Napomena

Za fizičke nezavisne eksperimente se podrazumijeva da su statistički nezavisni događaji.

Primjer 9.17

Pretpostavimo da su dvije kuglice nasumice izvučene iz vrećice koja sadrži 4 crvene i 3 crne kuglice. Prije drugog izvlačenja prethodno izvučena kuglica vraća se u vrećicu i sve se dobro promiješa. Kolika je vjerojatnost da su obje odabrane kuglice crvene?

Rješenje:

Neka je A događaj da je u prvom pokušaju izvučena crvena kuglica i B događaj izvlačenja crvene kuglice u drugom pokušaju. Prostor elementarnih događaja Ω tog slučajnog pokusa sastoji se od svih uređenih parova (x, y), pri čemu i x i y označavaju boju izvučenih kuglica iz skupa od 7 kuglica (4 crvene, 3 crne). Dakle, u svakom izvlačenju moguće je 7 ishoda.

Princip umnoška (princip produkta) povlači da se prostor elementarnih događaja (skup mogućih ishoda) Ω sastoji $7 \cdot 7 = 49$ elemenata. Kako $A \cap B$ sadrži sve uređene parove oblika (crvena, crvena), onda je $4 \cdot 4 = 16$ elemenata u skupu $A \cap B$. Vjerojatnost izvlačenja crvene kuglice u svakom pojedinom izvlačenju uvijek je ista i iznosi $\frac{4}{7}$. Stoga, vjerojatnost da su u dva pokušaja izvučene crvene kuglice, možemo izračunati kao

$$P(A \cap B) = \frac{16}{49} = \frac{4}{7} \cdot \frac{4}{7} = P(A)P(B).$$

Dakle, događaji A i B su **nezavisni**.

Primjer 9.18

Pretpostavimo da se nasumice izvlače dvije kuglice iz vrećice koja sadrži 4 crvene i 3 crne kuglice. Prva izvučena kuglica se ne vraća u vrećicu prije izvlačenja druge kuglice. Kolika je vjerojatnost da su obje odabrane kuglice crvene?

Rješenje:

Svaka od 7 kuglica može biti izvučena u prvom pokušaju. Nakon prvog izvlačenja, u vrećici je ostalo 6 kuglica i svaka od njih može biti izvučena u drugom izvlačenju.

Stoga se prostor elementarnih događaja (skup mogućih ishoda) Ω sastoji od $7 \cdot 6 = 42$ elemenata. Neka je A događaj da je u prvom pokušaju izvučena crvena kuglica i B događaj izvlačenja crvene kuglice u drugom pokušaju.

Co-funded by the Erasmus+ Programme of the European Union

Skup A sadrži $4 \cdot 6 = 24$ elementa i skup B sadrži $6 \cdot 4 = 24$ elementa. To povlači

$$P(A) = P(B) = \frac{24}{42} = \frac{4}{7}.$$

Skup $A \cap B$ sadrži $4 \cdot 3 = 12$ elemenata. Zbog toga je

$$P(A \cap B) = \frac{12}{42} = \frac{2}{7} \neq \frac{4}{7} \cdot \frac{4}{7} = P(A)P(B).$$

Slijedi, događaji A i B su nezavisni.

Napomena

Kako je u Primjeru 9.18 $P(A \cap B) = \frac{12}{42} = \frac{4}{7} \cdot \frac{3}{6}$ i $P(A) = \frac{4}{7}$ onda se drugi faktor u umnošku $\frac{3}{6}$ može protumačiti kao vjerojatnost da je druga izvučena kuglica crvene boje uz uvjet da je prva izvučena kuglica bila crvene boje. Ovo se označava kao $P(B|A)$.

Posljednja napomena sugerira opći zakon

Formula uvjetne vjerojatnosti

Neka $P(A)$ označava vjerojatnost događaja A, a $P(B|A)$ označava vjerojatnost događaja B nakon što se događaj A dogodio. Ako je $P(A \cap B)$ vjerojatnost da se pojave A i B, tada

$$P(A \cap B) = P(A)P(B|A).$$

Ako je $P(A) > 0$ imamo formulu uvjetne vjerojatnosti

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Ili ako je $P(B) > 0$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Napomena

U primjeru 9.18 $P(A|B) = \frac{3}{6} = \frac{1}{2}$ označava vjerojatnost događaja A nakon što se događaj B dogodio. U primjeru je vjerojatnost da je prva izvučena kuglica bila crvena pod uvjetom da je druga izvučena kuglica crvena.

P(A|B) se naziva "uvjetna vjerojatnost za A nakon što se dogodio događaj B", ili "vjerojatnost za A pod uvjetom B".

NAPOMENA:

Ako su A i B nezavisni događaji, onda je $P(A|B) = P(A)$ uz uvjet $P(B) > 0$.

Zadatak 9.3

Pokazati da ako su A i B nezavisni događaji, onda su A i $\bar{B}$ nezavisni događaji.

Rješenje:

$$\begin{aligned} P(A \cap \bar{B}) &= P\big(A \cap (\Omega \setminus B)\big) = P(A \setminus B) = P(A) - P(A \cap B) = P(A) - P(A)P(B) \\ &= P(A)\big(1 - P(B)\big) = P(A)P(\bar{B}) \end{aligned}$$

Zadatak 9.4

Pokazati da ako su A i B nezavisni događaji, onda su
$\bar{A}$ i $\bar{B}$ nezavisni događaji.

Rješenje:

$$
\begin{aligned}
P(\bar{A} \cap \bar{B}) &= P\big((\Omega \setminus A) \cap (\Omega \setminus B)\big) = P\big(\Omega \setminus (A \cup B)\big) = 1 - P(A \cup B) \\
&= 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A) - P(B) + P(A)P(B) \\
&= \big(1 - P(A)\big)\big(1 - P(B)\big) = P(\bar{A})P(\bar{B}).
\end{aligned}
$$

BAYESOVA FORMULA

Pretpostavimo da je $P(B) > 0$. Onda je

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Potpun sustav događaja čine događaji $E_1, \dots, E_n \subset \Omega$ ako vrijedi sljedeće:

1) $\bigcup_{i=1}^{n} E_i = \Omega$
2) $E_i \cap E_j = \emptyset$, za $i \neq j$ ($E_1, \dots, E_n$ su međusobno disjunktni, tj. događaji se međusobno isključuju).

Potpuna vjerojatnost

Neka je $A \subset \Omega$. Tada je

$$A = \Omega \cap A = \left(\bigcup_{i=1}^{n} E_i\right) \cap A = \bigcup_{i=1}^{n}(E_i \cap A) = (E_1 \cap A) \cup (E_2 \cap A) \cup \dots \cup (E_n \cap A).$$

Događaji $(E_i \cap A)$ i $\big(E_j \cap A\big)$ se isključuju za $i \neq j$ te je

Onda vrijedi formula potpune vjerojatnosti:

$$P(A) = P(A|E_1)P(E_1) + \dots + P(A|E_n)P(E_n).$$

Primjer 9.19

Tri identične zdjele označene su 1, 2, 3. Prva posuda sadrži 3 crvene i 3 plave kuglice. Druga posuda sadrži 4 crvene i 2 plave kuglice. Treća posuda sadrži 1 crvenu i 5 plavih. Prvo se nasumično odabire zdjela, a zatim se iz zdjele nasumično bira kuglica.
Kolika je vjerojatnost da je odabrana plava kuglica?

Rješenje:

E_i – događaj da je odabrana zdjela i. Vrijedi $P(E_i) = \frac{1}{3}$.
A – događaj da je odabrana plava kuglica. Vrijedi

$$P(A|E_1) = \frac{1}{2}, \qquad P(A|E_2) = \frac{1}{3}, \qquad P(A|E_3) = \frac{5}{6}$$

i

$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + P(A|E_3)P(E_3) = \frac{1}{2}\cdot\frac{1}{3} + \frac{1}{3}\cdot\frac{1}{3} + \frac{5}{6}\cdot\frac{1}{3} = \frac{5}{9}.$$

Ako je odabrana plava kuglica, koja je vjerojatnost da je odabrana zdjela 2?

Rješenje:

Primjenom Baysove formule vrijedi:

$$P(E_2|A) = \frac{P(A|E_2)P(E_2)}{P(A)} = \frac{\frac{1}{3}\cdot\frac{1}{3}}{\frac{5}{9}} = \frac{1}{5}.$$

Primjer 9.20

Određena bolest ima stopu incidencije od 1%. Pretpostavimo da je za neki dijagnostički test lažno negativna stopa 10%, a lažno pozitivna stopa 2%. Izračunajte vjerojatnost da osoba, nasumično odabrana iz populacije, koja je pozitivna na testu, zaista ima tu bolest.

Rješenje:

Definirajte događaje za dotičnu osobu:

D: ima bolest, T: test ukazuje na bolest (test je pozitivan)

Komplementarni događaji su:

$\bar{D}$: nema bolest, $\bar{T}$: test ne ukazuje na bolest (test je negativan).

Budući da bolest ima stopu učestalosti od 1%, imamo

$$P(D) = 0.01$$

i za komplementarni događaj

$$P(\bar{D}) = 1 - P(D) = 0.99.$$

Budući da bolest ima lažno negativnu stopu od 10%, imamo

$$P(\bar{T}|D) = 0.1.$$

Budući je lažno pozitivna stopa 2%, imamo

$$P(T|D) = 0.02$$

Računamo $P(D|T)$.

Primjenjujemo Baysovu formulu:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)},$$

Gdje je

$$P(T|D) = 1 - P(\bar{T}|D) = 1 - 0.1 = 0.9.$$

Po formuli potpune vjerojatnosti

$$P(T) = P(T|D)P(D) + P(T|\bar{D})P(\bar{D}) = 0.9 \cdot 0.01 + 0.02 \cdot 0.99 = 0.028.$$

Konačno,

Co-funded by the
Erasmus+ Programme
of the European Union

$$P(D|T) = \frac{0.9 \cdot 0.01}{0.028} = 0.3125.$$

Samo 31,25% nasumično odabranih ljudi čiji je test pozitivan zapravo ima neku bolest.
Rezultat se čini iznenađujućim. Ovdje je kritična varijabla niska učestalost bolesti u populaciji.

Zadatak 9.5

U gornjem primjeru izračunajte vjerojatnost da osoba, nasumično odabrana iz populacije, čiji je test negativan, zapravo nema tu bolest.

Rješenje:
0.99897

Co-funded by the
Erasmus+ Programme
of the European Union

9.5 SLUČAJNA VARIJABLA: DISKRETNA i NEPREKIDNA SLUČAJNA VARIJABLA

Defincija: slučajna varijabla

Neka je Ω prostor elementarnih događaja. Svaka realna funkcija $X: \Omega \to \mathbb{R}$ zove se *slučajna varijabla.*

Napomena:

$X(\omega)$ je vrijednost funkcije X za događaj ω iz Ω.

Primjer 9.21

Pretpostavimo da bacamo novčić. Tada je prostor elementarnih događaja Ω={H,T}, gdje H predstavlja glavu, a T pismo bacanja novčića. Slučajnu varijablu X:Ω→R definiramo pomoću $X(H) = 1$ i $X(T) = 0$.

Primjer 9.22

Pretpostavimo da bacamo kocku. Tada je prostor elementarnih događaja $\Omega = \{1,2,3,4,5,6\}$,, gdje brojevi predstavljaju brojeve na kockicama. Slučajnu varijablu $X: \Omega \to \mathbb{R}$ definiramo pomoću

$$X(1) = 1,\ X(2) = 2,\ X(3) = 3,\ X(4) = 4,\ X(5) = 5,\ X(6) = 6.$$

Jednostavno kažemo da je X broj koji se okrenuo pri bacanju pravilno izrađene igraće kockice.

Napomena

Ako je novčić koji bacamo u primjeru 9.21 pravilno izrađen, vjerojatnost da dobijemo glavu je 1/2, a vjerojatnost da dobijemo pismo je 1/2. Pišemo:

$$P(H) = P(\{\omega \in \Omega: X(\omega) = 1\}) = \frac{1}{2}$$

i

$$P(T) = P(\{\omega \in \Omega: X(\omega) = 0\}) = \frac{1}{2}.$$

Jednostavnije pišemo $P(X = x) = P(\{\omega \in \Omega: X(\omega) = x\})$.

Slično shvaćamo notacije: $P(X < x),\ P(X > x),\ P(X \le x),\ P(X \ge x)$.

Definicija: Funkcija kumulativne distribucije (razdiobe)

Funkcija kumulativne distribucije (funkcije distribucije) slučajne varijable $\boldsymbol{X}$ definirana je izrazom $F(x) = P(X \le x)$ za svaki realni broj x.

Svojstva funkcije kumulativne distribucije

1. Neka je F funkcija distribucije neke slučajne varijable. Za svaki realni broj $x\ vrijedi$:
   $0 \le F(x) \le 1$
2. $\lim_{x \to -\infty} F(x) = 0,\ \lim_{x \to \infty} F(x) = 1$

Co-funded by the Erasmus+ Programme of the European Union

3. $\lim_{x \to a^+} F(x) = F(a)$
4. Funkcija distribucije F slučajne varijable monotono je rastuća funkcija, tj.
   $x_1 < x_2 \Rightarrow F(x_1)$

Slučajna varijabla može biti:

- diskretna ili diskontinuirana slučajna varijabla
- Neprekidna slučajna varijabla.

9.5.1 Diskretna slučajna varijabla

Definicija: diskretna slučajna varijabla

Slučajna varijabla se naziva *diskretnom* ako s pozitivnom vjerojatnošću poprima ili konačno ili prebrojivo mnogo vrijednosti s određenom vjerojatnošću, tj. postoji konačan niz $x_1, \dots, x_n$ takav da je

(1)

$$P(X = x_i) = p_i, \ \sum_{i=1}^{n} p_i = 1, \ \ p_1, \dots, p_n \text{ su pozitivni brojevi}$$

ili postoji beskonačni niz $x_1, x_2, \dots$takav da

(2)

$$P(X = x_i) = p_i, \ \sum_{i=1}^{\infty} p_i = 1, \ \ p_1, p_2, \dots \text{ su pozitivni brojevi .}$$

Napomena

Kada kažemo da je zadana slučajna varijabla X, pod tim podrazumijevamo da su zadane sve njene vrijednosti i pripadne vjerojatnosti. Skup svih parova

$$\big(x_i, p(x_i)\big), \ \ i = 1,2, \cdots$$

tvori *distribuciju (razdiobu) slučajne varijable X*.

Funkcija vjerojatnosti diskretne slučajne varijable je jednoznačno preslikavanje skupa x_iu skup $p(x_i)$, a opisuje kako se vjerojatnost ishoda mijenja s njegovom brojčanom vrijednošću.

Definicija: Funkcija vjerojatnosti slučajne varijable

Za diskretnu slučajnu varijablu X koja može poprimiti vrijednosti x_1, x_2, … funkcija vjerojatnosti je funkcija za koju vrijedi formula (1) odnosno (2).

Štoviše, (1) se može zapisati u obliku tablice

Co-funded by the
Erasmus+ Programme
of the European Union

$X = x_i$	x_1	x_2	…	x_n
$P\ (X = x_i)$	p_1	p_2	…	p_n

Ili jednostavnije

x_i	x_1	x_2	…	x_n
p_i	p_1	p_2	…	p_n

Napomena: U primjeru 9.21 ako je novčić pravilno izrađen, onda vrijedi

x_i	0	1
p_i	0.5	0.5

Parametri distribucije:

Definicija: Očekivanje slučajne varijable

Očekivana vrijednost (*očekivanje ili srednja vrijednost)* diskretne slučajne varijable X definira se kao

$$E(X) = \sum_i x_i P(X = x_i).$$

Linearnost očekivanja

Ako je $E(X) < \infty$ i $E(Y) < \infty$, onda za bilo koje realne brojeve a, b vrijedi

$$E(aX + bY) = aE(X) + bE(Y).$$

Očekivanje funkcije od X.

Neka je $Y = g(X)$, gdje je g funkcija diskretne slučajne varijable X. Onda je Y diskretna slučajna varijabla s očekivanjem

$$E(Y) = \sum_i g(x_i) P(X = x_i).$$

Definicija: Varijanca, standardna devijacija

Varijanca slučajne diskretne varijable X definirana je izrazom

$$V(X) = E\big(X - E(X)\big)^2 = E(X^2) - E(X)^2.$$

Standardna devijacija od X is definirana kao

$$\sigma(X) = \sqrt{V(X)}.$$

Primjer 9.23

Slučajna varijabla X ima sljedeću distribuciju:

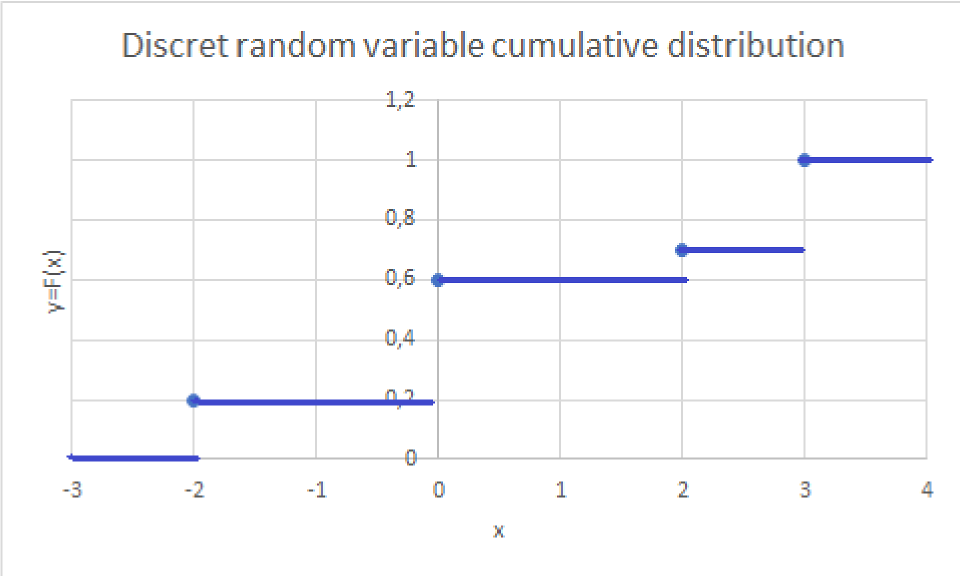
$X = x_i$	-2	0	2	3
$P(X = x_i)$	0.2	0.4	0.1	0.3

Co-funded by the
Erasmus+ Programme
of the European Union

Izračunati: $E(X)$, $E(X^2)$, $V(X)$ i $\sigma(X)$. Nacrtati graf funkcije kumulativne distribucije.

Rješenje:

$$E(X) = (-2)\cdot(0.2) + 0\cdot(0.4) + 2\cdot(0.1) + 3\cdot(0.3) = 0.7$$
$$E(X^2) = (-2)^2\cdot(0.2) + 0^2\cdot(0.4) + 2^2\cdot(0.1) + 3^2\cdot(0.3) = 3.9$$
$$V(X) = 3.9 - (0.7)^2 = 3.41$$
$$\sigma(X) = \sqrt{3.41} \approx 1.8466$$



Slika 9.2. Ilustracija za primjer 9.23

Napomena

Slučajna varijabla $Y = X^2$ u prethodnom primjeru ima sljedeću distribuciju:

$Y = y_i$	0	4	9
$P(Y = y_i)$	0.4	0.3	0.3

$$P(Y=0) = P(X^2=0) = P(X=0) = 0.4$$

$$P(Y=4) = P(X^2=4) = P(X=2 \text{ or } X=-2) = P(X=-2) + P(X=2) = 0.2 + 0.1 = 0.3.$$

VAŽNO! Događaji $\{X=-2\}$ i $\{X=2\}$ su disjunktni, stoga je vjerojatnost njihove sume jednaka sumi njihovih pojedinačnih vjerojatnosti.

$$P(Y=9) = P(X^2=9) = P(X=3) = 0.3$$

Vrijedi da je $P(X=-3) = 0$.

Korištenjem distribucije $Y = X^2$ moguće je izračunati $E(X^2)$ kako slijedi,

$$E(X^2) = E(Y) = 0\cdot 0.4 + 4\cdot 0.3 + 9\cdot 0.3 = 3.9.$$

Co-funded by the Erasmus+ Programme of the European Union

Primjer 9.24

Pretpostavimo da bacamo dvije kocke. Slučajnu varijablu X definiramo kao zbroj dobivenih brojeva. Naći distribuciju vjerojatnosti X. Izračunati očekivanu vrijednost, varijancu i standardnu devijaciju slučajne varijable X.

Rješenje:

Prostor elementarnih događaja je

$$\Omega = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6),\\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5),(4,6),\\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5)(6,6)\}$$

Broj elemenata od Ω je $\mathbf{36}$.

Primjenom klasične definicije vjerojatnosti imamo

$$P(X=2) = P(\{(1,1)\}) = \frac{1}{36}$$

$$P(X=3) = P(\{(1,2),(2,1)\}) = \frac{2}{36} = \frac{1}{18}$$

$$P(X=4) = P(\{(1,3),(2,2),(3,1)\}) = \frac{3}{36} = \frac{1}{12}$$

$$P(X=5) = P(\{(1,4),(2,3),(3,2),(4,1)\}) = \frac{4}{36} = \frac{1}{9}$$

$$P(X=6) = P(\{(1,5),(2,4),(3,3),(4,2),(5,1)\}) = \frac{5}{36}$$

$$P(X=7) = P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \frac{6}{36} = \frac{1}{6}$$

$$P(X=8) = P(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = \frac{5}{36}$$

$$P(X=9) = P(\{(3,6),(4,5),(5,4),(6,3)\}) = \frac{4}{36} = \frac{1}{9}$$

$$P(X=10) = P(\{(4,6),(5,5),(6,4)\}) = \frac{3}{36} = \frac{1}{12}$$

$$P(X=11) = P(\{(5,6),(6,5)\}) = \frac{2}{36} = \frac{1}{18}$$

$$P(X=12) = P(\{(6,6)\}) = \frac{1}{36}.$$

Funkcija vjerojatnosti slučajne varijable X prikazana je sljedećom tablicom:

$X = x_i$	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Očekivanje $E(X)$ od X se računa kao:

$$E(X) = 2\cdot\frac{1}{36} + 3\cdot\frac{1}{18} + 4\cdot\frac{1}{12} + 5\cdot\frac{1}{9} + 6\cdot\frac{5}{36} + 7\cdot\frac{1}{6} + 8\cdot\frac{5}{36} + 9\cdot\frac{1}{9} + 10\cdot\frac{1}{12} + 11\cdot\frac{1}{18} + 12\cdot\frac{1}{36} = 7$$

Varijanca od X se računa u dva koraka. Prvo se računa očekivanje od X^2.

Co-funded by the
Erasmus+ Programme
of the European Union

$$E(X^2) = 2^2 \cdot \frac{1}{36} + 3^3 \cdot \frac{1}{18} + 4^2 \cdot \frac{1}{12} + 5^2 \cdot \frac{1}{9} + 6^2 \cdot \frac{5}{36} + 7^2 \cdot \frac{1}{6} + 8^2 \cdot \frac{5}{36} + 9^2 \cdot \frac{1}{9} + 10^2 \cdot \frac{1}{12} + 11^2 \cdot \frac{1}{18} + 12^2 \cdot \frac{1}{36} = \frac{329}{6} \approx 54.83$$

Nakon toga se računa varijanca

$V(X) = E(X^2) - E(X)^2 \approx 54.83 - 7^2 = 5.83.$

Standardna devijacija je

$\sigma(X) = \sqrt{V(X)} \approx 2.42.$

9.5.2 Kontinuirana slučajna varijabla

Definicija: Neprekidna slučajna varijabla

Slučajna varijabla Xse naziva ***kontinuirana (neprekidna)***, ako postoji neprekidna ne-negativna funkcija f takva da za svaki realni broj a vrijedi

$$P(X \leq a) = \int_{-\infty}^{a} f(x)dx.$$

Funkcija f se naziva *funkciju gustoće vjerojatnosti* (*funkcija gustoće*) slučajne X.

Neprekidna slučajna varijabla može poprimiti neprebrojivo (beskonačno) mnogo vrijednosti. Zato se za kontinuiranu slučajnu varijablu ne računa vjerojatnost u određenoj točki, nego nad određenim intervalom.

Funkcija vjerojatnosti ili funkcija gustoće vjerojatnosti kontinuirane slučajne varijable $\boldsymbol{X}$ ima svojstva:

1) $f(x) \geq 0$
2) $\int_{-\infty}^{\infty} f(x)dx = 1.$ (površina ispod krivulje funkcije vjerojatnosti je 1)
3) Vjerojatnost da kontinuirana slučajna varijabla poprimi vrijednost unutar segmenta $a \leq X \leq b$ je dana sa sljedećim integralom:

$$F(x) = P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

Funkcija distribucije kontinuirane slučajne varijable $\boldsymbol{X}$ je funkcija koja daje vjerojatnost da će slučajna varijabla $\boldsymbol{X}$ poprimiti vrijednost jednaku ili manju od nekog realnog broja $\boldsymbol{x}$**.**

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

Funkcija distribucije kontinuirane slučajne varijable i vjerojatnosti

1 $P(X \geq a) = P(X > a) = 1 - F(a)$

2 $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = F(b) - F(a)$

3 $P(X = a) = 0$

Co-funded by the Erasmus+ Programme of the European Union

Primjer 9.25

Dana je funkcija gustoće vjerojatnosti slučajne varijable X,

$$f(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & inače. \end{cases}$$

Pronađite kumulativnu funkciju distribucije ***F*** slučajne varijable X.

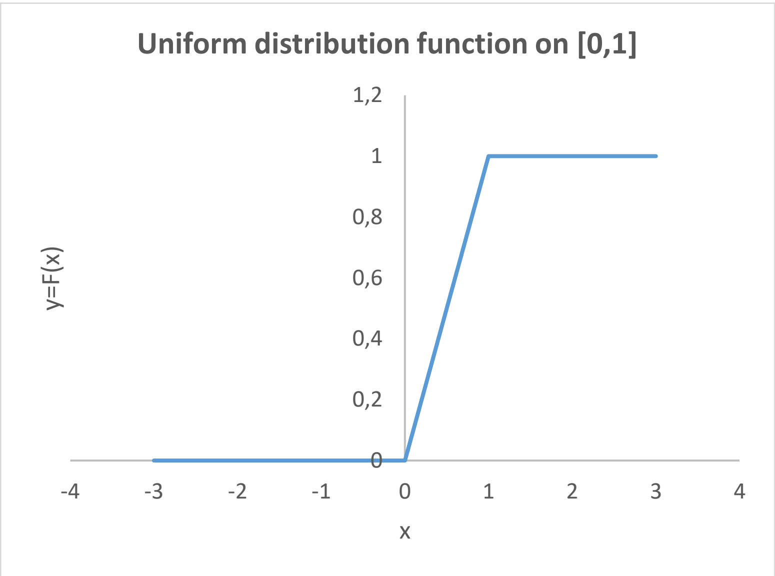
Rješenje:

Koristit ćemo se formulom za kumulativnu funkciju distribucije. Moramo razmotriti tri slučaja:

1. $x < 0$: $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0$
2. $0 \leq x \leq 1$: $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0\, dt + \int_{0}^{x} 1\, dt = 0 + x = x$
3. $x > 1$: $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{0} 0dt + \int_{0}^{1} 1dt + \int_{1}^{x} 0dt = 0 + 1 + 0 = 1$

Konačno, funkcija distribucije je dana sa

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 1 \\ 1, & x \geq 1. \end{cases}$$



Slika 9.3. *Ilustracija primjera 9.27*

9.5.3 Funkcija kumulativne distribucije uniformne slučajne varijable

Funkcija kumulativne distribucije uniformne slučajne varijable $\boldsymbol{X}$ dana je izrazom

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{1}{b-a}, & a \leq x \leq b. \\ 1, & x > b \end{cases}$$

Co-funded by the
Erasmus+ Programme
of the European Union

Zadatak 9.6

Nađite kumulativnu funkciju distribucije uniformne slučajne varijable X na intervalu [1,4].

Rješenje:

Kako je $a = 1,\ b = 4$, vrijedi $b - a = 4 - 1 = 3$ i

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{3}, & x \in [1,4] \\ 1, & x > 4. \end{cases}$$

Parametri distribucije:

Definicija: Očekivanje

Očekivanje slučajne kontinuirane varijable X (ako integral konvergira) definira se kao

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

gdje je $\boldsymbol{f}$ funkcija gustoće od $\boldsymbol{X}$.

Slično kao i za diskretnu slučajnu varijablu, vrijedi linearnost očekivanja $E(X)$.
Ako je $E(X) < \infty$ i $E(y) < \infty$, onda za bilo koja dva realna broja a, b vrijedi

$$E(aX + bY) = aE(X) + bE(Y).$$

Očekivanje funkcije od X

Neka je $Y = g(X)$, pri čemu je g kontinuirana funkcija od kontinuirane slučajne varijable X s pridruženom funkcijom gustoće vjerojatnosti $f(x)$. Y je kontinuirana slučajna varijabla i vrijedi

$$E(Y) = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Slično kao i za diskretnu slučajnu varijablu vrijedi sljedeća definicija:

Definicija: Varijanca

Varijanca *(ili drugi središnji moment) kontinuirane slučajne varijable X s pridruženom funkcijom gustoće vjerojatnosti f(x)*

$$\sigma^2 = V(X) = E\big(X - E(X)\big)^2 = E(X^2) - E(X)^2.$$

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

Standardna devijacija *od X*

$$\sigma(X) = \sqrt{V(X)}.$$

Co-funded by the
Erasmus+ Programme
of the European Union

Primjer 9.26

Izračunati $E(X)$, $E(X^2)$, $V(X)$, $\sigma(X)$ za uniformo slučajne brojeve na intervalu $[0,1]$.

Rješenje:

Gustoća je zadana kao

$f(x) = 1$ za $x \in [0,1]$, inače je $f(x) = 0$.

Dakle vrijedi,

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x dx = \frac{1}{2}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^1 x^2 dx = \frac{1}{3}$$

$$V(X) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

$$\sigma(X) = \sqrt{\frac{1}{12}} \approx 0.2887.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.6 DISTRIBUCIJE DISKRETNE SLUČAJNE VARIJABLE: BINOMNA I POISSONOVA DISTRIBUCIJA

Definicija: Bernoullijevi pokus

Bernoullijev pokus je neovisni pokus za koji vrijedi sljedeća pravila:

1. Pokus se ponavlja određeni broj puta (n puta).
2. Svako ponavljanje ima samo dva moguća ishoda, *„povoljni ishod"* (engl. success (S)) i „nepovoljni ishod" (engl. failure (F)). Vjerojatnosti tih ishoda se ne mijenjaju tijekom pokusa.
3. Vjerojatnost koja odgovara povoljnom ishodu pokusa se označava s p, vjerojatnost koja odgovara nepovoljnom ishodu pokusa se označava s q.
4. p i q moraju biti neke nenegativne vrijednosti te mora vrijediti: $p + q = 1,\ q = 1 - p$.
5. Pokušaji su nezavisni (ishod prethodnog pokušaja ne utječe na ishod sljedećeg pokušaja).

Zanima nas slučajna varijabla X gdje X označava broj povoljnih ishoda. Važno je imati na umu da su moguće vrijednosti od X 0,1,2,...,n.
Naš sljedeći cilj je izračunati distribuciju vjerojatnosti za slučajnu varijablu X, pri čemu X računa broj uspjeha u Bernoullijevom eksperimentu s n pokušaja.

Definicija: BINOMNA DISTRIBUCIJA

Ako je X broj povoljnih ishoda u Bernoullijevom pokusu s n nezavisnih pokušaja, pri čemu je vjerojatnost povoljnog ishoda p u svakom pokušaju i vjerojatnost nepovoljnog ishoda je $q = 1 - p$, onda je vjerojatnost nastupanja povoljnog ishoda k puta u n Bernoullijevih pokusa dana izrazom:

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0,1,2, \dots, n.$$

$\binom{n}{k}$ predstavlja broj načina na koji se dana kombinacija uspjeha i neuspjeha može ostvariti
n – broj Bernoullijevih pokusa
k – broj pokusa s uspješnim ishodom
$n - k$ – broj pokusa s neuspješnim ishodom.

Primjer 9.27

Košarkaš izvodi 4 nezavisna slobodna bacanja. Svaki put je vjerojatnost osvajanja boda 0,6, a svaki postignuti bod se bilježi. Ovdje je svako slobodno bacanje pokušaj i pretpostavlja se da su pokušaji neovisni. Svaki pokušaj ima dva ishoda: bod (uspjeh) ili bez bodova (neuspjeh). Vjerojatnost uspjeha je $p = 0{,}6$, a vjerojatnost neuspjeha je $q = 1 - p = 0{,}4$. Zanima nas varijabla X koja broji broj uspjeha u 4 pokušaja. Ovo je primjer Bernoullijevog eksperimenta s n=4 pokušaja.

Co-funded by the Erasmus+ Programme of the European Union

Rješenje:

$n = 4, \quad p = 0.6, \quad q = 0.4$

$P(X = 0) = \binom{4}{0} \cdot (0.6)^0 \cdot (0.4)^4 = (0.4)^4 = 0.0256$

$P(X = 1) = \binom{4}{1} \cdot (0.6)^1 \cdot (0.4)^3 = 4 \cdot 0.6 \cdot (0.4)^3 = 0.1536$

$P(X = 2) = \binom{4}{2} \cdot (0.6)^2 \cdot (0.4)^2 = 6 \cdot (0.6)^2 \cdot (0.4)^2 = 0.3456$

$P(X = 3) = \binom{4}{3} \cdot (0.6)^3 \cdot (0.4)^1 = 4 \cdot (0.6)^3 \cdot 0.4 = 0.3456$

$P(X = 4) = \binom{4}{4} \cdot (0.6)^4 \cdot (0.4)^0 = (0.6)^4 = 0.1296$

Konačno

x_i	0	1	2	3	4
p_i	0.0256	0.1536	0.3456	0.3456	0.1296

Definicija: Poissonova distribucija

Poissonova distribucija s parametrom $\lambda > 0$ zadana je izrazom

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \qquad k = 0,1,2,3, \dots$$

Glavno svojstvo Poissonove distribucije je da ako X slijedi Poissonovu distribuciju s parametrom $\lambda > 0$ onda je $E(X) = V(X) = \lambda$.

Napomena

Poissonova distribucija je distribucija vjerojatnosti za broj događaja koji se događaju nasumično i neovisno u fiksnom intervalu vremena (ili prostora). Ako je srednja vrijednost broja događaja po intervalu λ tada je vjerojatnost opažanja k događaja u ovom intervalu dana Poissonovom distribucijom.

Primjer 9.28

Pretpostavimo da je prosječni broj nekih događaja u fiksnom vremenskom intervalu 2. Odredite vjerojatnost da ćemo danas promatrati pet događaja u tom vremenskom intervalu.

Rješenje:

Slučajna varijabla X slijedi Poissonovu distribuciju sa srednjom vrijednošću 2, pronađite $P(X = 5)$. Budući je $\lambda = 2$, slijedi

$$P(X = 5) = e^{-2} \frac{2^5}{5!} \approx 0.03609.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.7 NORMALNA DISTRIBUCIJA

Definicija: Normalna distribucija

Normalna (Gaussian) distribucija je raspodjela vjerojatnosti kontinuirane (neprekidne) varijable pri čemu je njezina funkcija gustoće dana sljedećim izrazom

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

Označava se s $N(\mu, \sigma)$, gdje je μ srednja vrijednost i σ je standardna devijacija.

Definicija: Jedinična ili standarna normalna distribucija

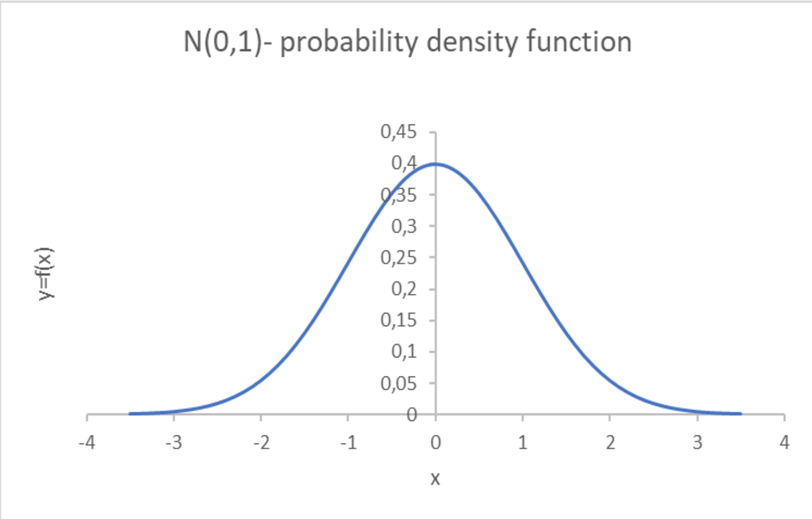
Jedinična ili standardna normalna distribucija je normalna (Gaussian) distribucija s parametrima:

srednja vrijednost $\mu = 0$ i varijanca $\sigma^2 = 1$.

Gustoća standardne normalne distribucije dana je izrazom

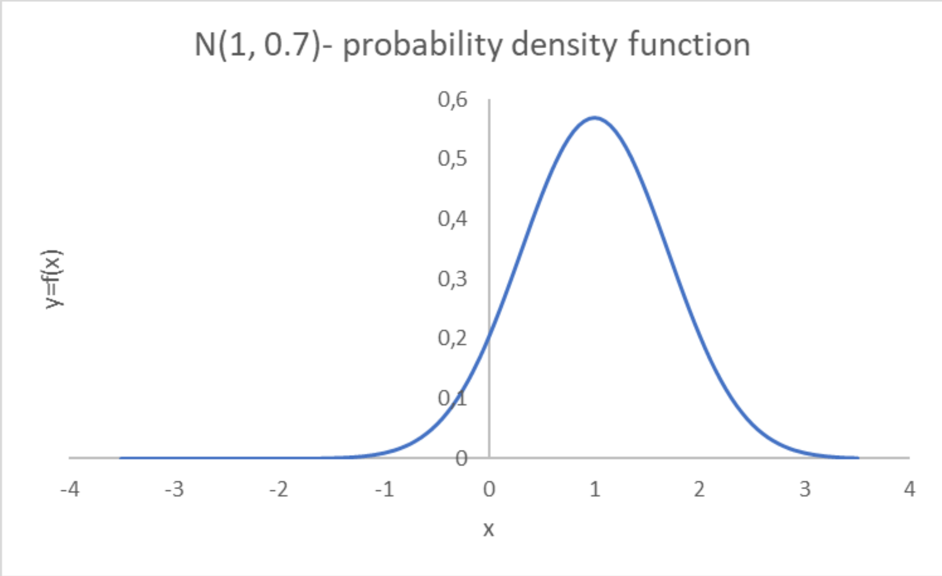
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad -\infty < x < \infty.$$

Označava se s $N(0,1)$.



Slika 9.4: *Standardna normalna distribucija*

Co-funded by the Erasmus+ Programme of the European Union



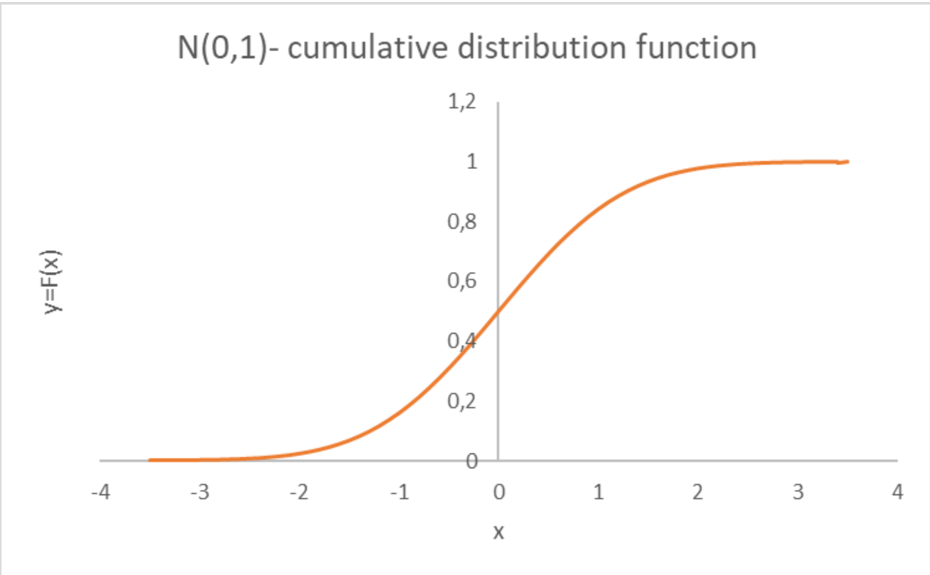
Slika 9.5: *Normalna distribucija* $\mu = 1, \sigma = 0.7$

Napomena:
Graf funkcije gustoće normalne distribucije zove se **Gaussova krivulja**.

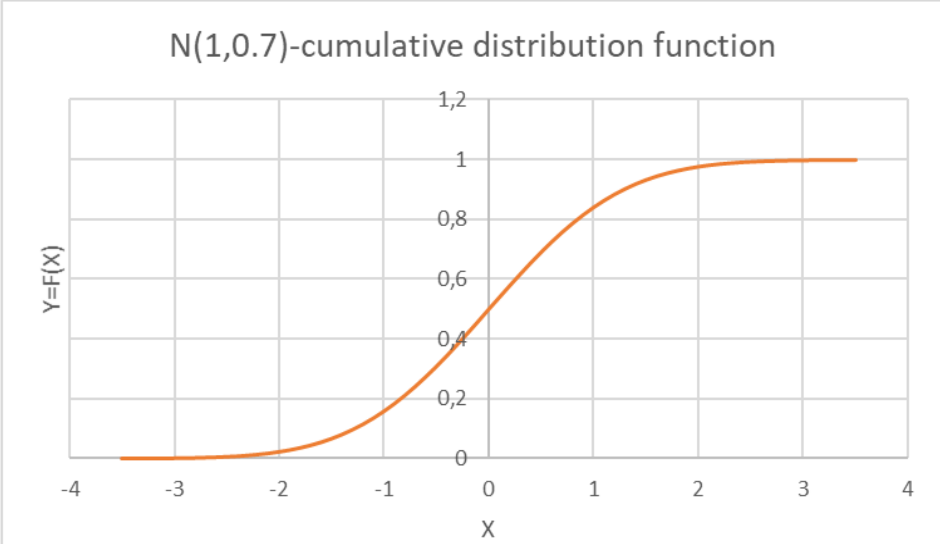
Kumulativna funkcija distribucije standardne normalne slučajne varijable definirana je sljedećim izrazom

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \qquad -\infty < x < \infty.$$

Ovaj integral nije elementaran.



Slika 9.6. *Funkcija distribucije standardne normalne distribucije*



Slika 9.7. *Funkcija kumulativne distribucije standardne normalne distribucije* $\mu = 1, \sigma = 0.7$

Primjer 9.29

Varijabla X ima normalnu distribuciju sa srednjom vrijednošću od 20 i standardnom devijacijom od 3. Jedan rezultat je slučajno odabran.

1. Kolika je vjerojatnost da je iznad 15?
2. Kolika je vjerojatnost da je u intervalu $[18,22]$??

Rješenje:

1. Varijabla X ima normalnu distribuciju $N(20,3)$. Moramo izračunati $P(X > 15)$. Neka je F funkcija kumulativne distribucije od X.

$$P(X > 15) = 1 - P(X \leq 15) = 1 - F(15) \approx 1 - 0.0478 = 0.9522$$

$F(15)$ određujemo računanjem NORM.DIST(15;20;3;1) u Excelu ili NORMDIST(15;20;3;1) u OpenOffice Calc.

2. $P(18 \leq X \leq 22) = F(22) - F(18) \approx 0.7475 - 0.2525 = 0.4950$

Zadatak 9.7

Pretpostavimo da je duljina haringe normalno raspoređena sa srednjom vrijednošću od 8 inča i standardnom devijacijom od 1,5 inča. Ribarski brod može uloviti 10 tona haringi dnevno. Procjena:

1. Koliko je riba dulje od 11 inča?

2. Koliko biste očekivali da će biti niži od 6 inča?

3. Koliko ih ima duljinu između 8 i 10 inča?

Rješenje:

1. $\approx 0.0228 \cdot 10\ t = 0.228\ t$
2. $\approx 0.0912 \cdot 10\ t = 0.912\ t$
3. $\approx 0.4088 \cdot 10\ t = 4.088\ t$

Co-funded by the
Erasmus+ Programme
of the European Union

9.8 STATISTIKA: OSNOVNI ELEMENTI I DEFINICIJE

Statistika je znanost o prikupljanju, analizi, prezentiranju i tumačenju podataka.
Statističko istraživanje fokusirano je na skup objekata, tj. ***jedinki*** (ljudi, životinja, brodova, marina, luka, država, gradova, poduzeća itd.) i skup odabranih veličina (karakteristika) koje se na njima promatraju (poput njihove dužine, težine, brzine, cijene ili starosti). Veličine koje se na jedinkama promatraju zovemo ***varijablama***. ***Populacija*** uključuje sve jedinke koje se žele obuhvatiti istraživanjem, tj. o kojima se želi zaključivati. ***Uzorak*** se sastoji od jednog opažanja izvučenog iz populacije (malog dijela populacije).
Podaci moraju biti prikupljeni na uzorku koji je reprezentativan za opisanu populaciju. ***Reprezentativan uzorak*** mora odražavati populaciju tj. u njemu trebaju biti zastupljene sve tipične karakteristike populacije. Najčešći je način odabira jedinki iz populacije u reprezentativan uzorak tzv. ***slučajni uzorak***, tj. takav izbor u kojemu svaka jedinka ima jednaku šansu biti izabrana u uzorak. Izbor jedinki u reprezentativan uzorak iz populacije može se postići pomoću ***generatora slučajnih brojeva***.

Za dani slučajni uzorak podataka:

$$x_1, x_2, \dots, x_n$$

definiramo nekoliko brojeva koji karakteriziraju ***određeni uzorak***.

Postoje dvije glavne grane statistike: deskriptivna i inferencijalna. Deskriptivna statistika se koristi da kaže nešto o skupu informacija koje su prikupljene iz malog dijela. Inferencijalna statistika koristi se za predviđanja ili usporedbi o većoj skupini (populaciji) korištenjem informacija prikupljenih o malom dijelu te populacije.

U statističkim istraživanjima razlikujemo nekoliko osnovnih tipova varijabli koje se međusobno razlikuju po svojstvima vrijednosti koje mogu poprimiti. Osnovni tipovi varijabli su:

- Kvalitativne varijable
- Numeričke varijable
- Ordinarne varijable.

Karakteristika je ***kvalitativnih*** varijabli da njihove vrijednosti nisu, po svojim svojstvima korištenim u istraživanju, realni brojevi. Primjeri takve varijable su spol osobe (m ili ž), krvne grupe (A, B, AB, 0), tip broda (jahta, kruzer, katamaran, trajekt, tanker)…

Numeričke varijable prirodno primaju vrijednosti iz skupa realnih brojeva. Tipični su primjeri numeričkih varijabli brzina broda, snaga motora, količina emisije CO2…

Napomena

Kategorije kvalitativnih varijabli mogu se izraziti brojevima što ih ne čini numeričkim varijablama. Primjerice, spol osobe jedna je kvalitativna varijabla. Kategoriju "ženski spol" možemo označiti s "1", a kategoriju "muški spol" s "2", što može biti korisno prilikom obrade podataka. Time smo kategorijama

Co-funded by the Erasmus+ Programme of the European Union

kvalitativne varijable pridružili numeričke vrijednosti, ali samu varijablu nismo učinili numeričkom po njezinim svojstvima.

Među numeričkim varijablama razlikujemo diskretne i kontinuirane varijable.
Diskretni podaci sastoje se od cijelih brojeva i obično opisuju određeni broj objekata. Na primjer, u istraživanju se može opisati broj vezova u različitim marinama ili broj uplova i isplova u lukama. Diskretne varijable mogu poprimiti samo konačno ili prebrojivo mnogo vrijednosti.

Kontinuirani (***mjereni podaci***), za razliku od diskretnih podataka, mogu poprimiti bilo koju stvarnu vrijednost iz skupa realnih brojeva ili nekog intervala. Na primjer, količina vremena koju je grupa djece provela igrajući računalne igrice bio bi mjeren podatak, jer bi mogli provesti bilo koji broj sati igrajući se.

Za numeričke varijable možemo definirati numeričke karakteristike koje imaju logičnu interpretaciju i mogu se iskoristiti s ciljem prikazivanja skupa podataka. Ovdje ćemo definirati neke od najčešće korištenih numeričkih karakteristika skupa podataka.

Definicija: Srednja vrijednost uzorka

Srednja vrijednost uzorka definira se izrazom

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

Definicija: Varijanca uzorka

Varijanca uzorka se definira izrazom

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Napomena:
Uzorak varijance je mjera kako su stavke u uzorku raspoređene oko njihove srednje vrijednosti.

Definicija: Standardna devijacija uzorka

Standardne devijacija uzorka je drugi korijen od varijance.

$$s = \sqrt{s^2}.$$

Primjer 9.30

Uzorak:

$$3,\ 3,\ 3,\ 3,\ 3$$

srednja vrijednost:

$$\bar{x} = \frac{1}{5}(3+3+3+3+3) = 3$$

varijanca:

Co-funded by the Erasmus+ Programme of the European Union

$$s^2 = \frac{1}{5}[(3-3)^2 + (3-3)^2 + (3-3)^2 + (3-3)^2 + (3-3)^2] = 0$$

standardna devijacija:

$$s = \sqrt{0^2} = 0$$

Nema disperzije podataka oko srednje vrijednosti.

Primjer 9.31

Uzorak:

$$3,\ 2,\ 4,\ 3,\ 3$$

srednja vrijednost:

$$\bar{x} = \frac{1}{5}(3+2+4+3+3) = 3$$

varijanca:

$$s^2 = \frac{1}{5}[(3-3)^2 + (2-3)^2 + (4-3)^2 + (3-3)^2 + (3-3)^2] = \frac{2}{5} = 0.4$$

standardna devijacija:

$$s = \sqrt{\frac{2}{5}} = 0.63$$

Postoji disperzija podataka oko srednje vrijednosti.

Primjer 9.32

Uzorak:

$$5,\ 2,\ 4,\ 2,\ 2$$

srednja vrijednost:

$$\bar{x} = \frac{1}{5}(5+2+4+2+2) = 3$$

varijanca:

$$s^2 = \frac{1}{5}[(5-3)^2 + (2-3)^2 + (4-3)^2 + (2-3)^2 + (2-3)^2] = \frac{8}{5} = 1.6$$

Standardna devijacija

$$s = \sqrt{\frac{8}{5}} = 1.26$$

Postoji značajna disperzija podataka oko srednje vrijednosti.

Napomena

Standardna devijacija i varijanca su pozitivne osim ako svi elementi u uzorku nisu isti. U tom slučaju su uvijek jednaki nuli.

Co-funded by the
Erasmus+ Programme
of the European Union

Definicija: Procjena varijance

Procjena varijance je zadana izrazom

$$\hat{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Standardna devijacija od procjene varijance se odredi kao

$$\hat{s} = \sqrt{\hat{s}^2}.$$

Primjer 9.33

Uzorak:

$$5,\ 2,\ 4,\ 2,\ 2$$

srednja vrijednost:

$$\bar{x} = \frac{1}{5}(5+2+4+2+2) = 3$$

Procjena varijance:

$$\hat{s}^2 = \frac{1}{4}[(5-3)^2 + (2-3)^2 + (4-3)^2 + (2-3)^2 + (2-3)^2] = \frac{8}{4} = 2$$

Standardna devijacija od procjene varijance

$$\hat{s} = \sqrt{2} = 1.41$$

9.9 TESTIRANJE STATISTIČKIH HIPOTEZA

Statistička hipoteza slutnja je koja se odnosi na statistički model. Cilj je testiranja hipoteze odlučiti o istinitosti ili neistinitosti te slutnje.

Definicija: nul i alternativna -hipoteza

- *Nul hipoteza je* činjenica (slutnja) koja se testira eksperimentom. Označava se s H_0.
- *Alternativna hipoteza* je činjenica koja je kontradiktorna H_0 i obuhvaća sve one distribucije koje nisu sadržane u H_0. Označava se s H_a.

Postoje dva načina donošenja odluke: odbacivanje H_0 ili prihvaćanje H_0.

Definicija: pogreška I. tipa. pogreška II. tipa:

- ***pogreška I. tipa:*** pojavljuje se kada odbacimo H_0 iako je H_0 istinita (pogrešna odluka).
- ***pogreška II. tipa:*** pojavljuje se kada ne odbacimo H_0 iako je H_0 neistinita.

Vjerojatnost ***pogreške I. tipa*** označava se s α. Vjerojatnost ***pogreške II. tipa*** označava se s β.
NAPOMENA:
Pogreška I. tipa i ***pogreška II. tipa*** su **pogrešne odluke.**
Preostale dvije odluke su **ispravne**:

- Ne odbaciti H_0 kada je, zapravo, H_0 istina (ispravna odluka).
- Odbaciti H_0 kada je, zapravo, H_0 lažna (ispravna odluka).

	H_0 **je istinita**	H_0 je lažna
H_0 je prihvaćena (tj. nije odbačena)	ispravno	pogreška tipa II.
H_0 je odbačena	pogreška tipa I.	ispravno

Uvest ćemo dva testa u kojima želimo provjeriti hipotezu da je srednja vrijednost populacije $\boldsymbol{\mu_0}$.. U prvom (Model I) pretpostavljamo da je poznata standardna devijacija populacije $\boldsymbol{\sigma}$.
U drugom testu (Model II) ne znamo vrijednost $\boldsymbol{\sigma}$. U oba slučaja pretpostavljamo da populacija slijedi normalnu distribuciju $\boldsymbol{N(\mu, \sigma)}$.

<u>Model I testa je sljedeći.</u>
Pretpostavimo da populacija slijedi normalnu distribuciju $\boldsymbol{N\ (\mu, \sigma)}$ i da je poznata standardna devijacija populacije $\boldsymbol{\sigma}$. Nulta hipoteza je $\boldsymbol{H_0}: \boldsymbol{\mu} = \boldsymbol{\mu_0}$, gdje je $\boldsymbol{\mu_0}$ predviđena vrijednost srednje vrijednosti za cijelu populaciju. Iz nasumičnog uzorka **x** želimo provjeriti hipotezu $\boldsymbol{H_0}: \boldsymbol{\mu} = \boldsymbol{\mu_0}$ nasuprot alternativne hipoteze $\boldsymbol{H_1}: \boldsymbol{\mu} \neq \boldsymbol{\mu_0}$. Pretpostavljamo neku unaprijed zamišljenu ili unaprijed postavljenu "razinu značajnosti" $\boldsymbol{\alpha}$ testa. Unaprijed postavljena vrijednost $\boldsymbol{\alpha}$ je vjerojatnost da ćemo odbiti $\boldsymbol{H_0}$, ali je zapravo $\boldsymbol{H_0}$ istinita, tj. vjerojatnost *pogreške tipa I*. Ako $\boldsymbol{\alpha}$ nije zadana, prihvaćeni standard je postaviti $\boldsymbol{\alpha} = \mathbf{0.05}$..

Nultu hipotezu H_0provjeravamo na sljedeći način:

Co-funded by the
Erasmus+ Programme
of the European Union

Prvo, izračunavamo srednju vrijednost x̄ od uzorka x. Zatim izračunavamo vrijednost slučajne varijable u sa standardnom normalnom distribucijom $N(0,1)$ pomoću formule:

$$u = \frac{\bar{x} - \mu_0}{\sigma}\sqrt{n}.$$

Iz tablice standardne normalne distribucije $N(0,1)$ ili koristeći kalkulator inverzne normalne distribucije (pogledajte na primjer:
http://onlinestatbook.com/2/calculators/inverse_normal_dist.html),
ili pomoću neke popularne proračunske tablice kao što je Excel, OpenOffice Calc, nalazimo kritičnu vrijednost u takvu da je

$$P(|u| \geq u_\alpha) = \alpha.$$

Skup vrijednosti u definirani nejednadžbom $|u| \geq u_\alpha$ je ***kritično područje ovog testa***, tj. ako dobijemo takvu vrijednost u da $|u| \geq u_\alpha$ hipoteza H_0 mora biti odbačena. Ako je $|u| < u_\alpha$, ne možemo odbaciti hipotezu H_0.
Gore opisani test se ponekad naziva ***dvostranim*** testom budući da imamo dvostrano kritično područje, tj.

$$|u| \geq u_\alpha \quad \text{tj.} \quad u \in (-\infty, -u_\alpha] \cup [u_\alpha, \infty).$$

To se događa kad god imamo alternativnu hipotezu oblika $H_a\colon \mu \neq \mu_0$.
Međutim, u nekim slučajevima moramo provesti ***jednostrane*** testove.
Kada je alternativna hipoteza oblika $H_a\colon \mu < \mu_0$, imamo lijevo kritično područje, tj. $u \leq u_\alpha$.. U ovom slučaju izvodimo u_α tako da

$$P(u \leq u_\alpha) = \alpha.$$

Kada je alternativna hipoteza oblika $H_a\colon \mu > \mu_0$, imamo desno kritično područje, tj. $u \geq u_\alpha$.. U ovom slučaju izvodimo u_α tako da

$$P(u \geq u_\alpha) = \alpha.$$

Primjer 9.34

U 81 slučajno odabranih proizvodnih kompanija, računa se cijena materijala utrošenog u proizvodnju jednog određenog proizvoda. Srednja vrijednost ovih veličina je $\bar{x} = 540$ € i $s = 150$ €. Uz razinu značajnosti $\alpha = 0.05$, testirati hipotezu da je prosječna cijena materijala za proizvodnju takvog proizvoda 600 €.

Rješenje:

Nulta hipoteza: Prosječna cijena potrošenog materijala za proizvodnju proizvoda je 600 € tj.

$$H_0\colon \mu_0 = 600.$$

Budući je veliki uzorak, može se pretpostaviti da je $s = \sigma = 150$. Primjenom Modela I, računa se

$$u = \frac{540 - 600}{150}\sqrt{81} = -3.6.$$

Vrijedi

$$P(|u| \geq u_\alpha) = 0.05$$

za $u_\alpha = 1.96$ (pogledajte tablicu, proračunsku tablicu ili neki kalkulator inverzne normalne distribucije, npr. (opcija: izvan jer je test dvostran).

Vidimo to

$$|-3.6| = 3.6 \geq 1.96 = u_\alpha.$$

Stoga se na razini značajnosti $\alpha = 0.05$ hipoteza H_0 mora odbaciti u korist alternativne hipoteze da prosječni materijalni trošak proizvodnje razmatranog proizvoda nije jednak 600 €.

Model II testa je sljedeći.

Pretpostavimo da populacija slijedi normalnu distribuciju $\boldsymbol{N(\mu, \sigma)}$ i da su i srednja vrijednost $\boldsymbol{\mu}$ i standardna devijacija populacije $\boldsymbol{\sigma}$ nepoznata. Mali uzorak je nasumično odabran iz populacije. Nulta hipoteza je $\boldsymbol{H_0: \mu = \mu_0}$. Iz slučajnog uzorka $\boldsymbol{x}$ želimo provjeriti istinitost hipoteze $\boldsymbol{H_0}$ u odnosu na alternativnu hipotezu $\boldsymbol{H_1: \mu \neq \mu_0}$. Pretpostavljamo neku "razinu značajnosti" α testa.

Nultu hipotezu $\boldsymbol{H_0}$ provjeravamo na sljedeći način:

Prvo izračunavamo srednje vrijednosti $\bar{x}$ i standardne devijacije s ili procjenitelj standardne devijacije $\hat{s}$. Zatim izračunavamo vrijednost slučajne varijable t pomoću formule:

$$t = \frac{\bar{x} - \mu_0}{s}\sqrt{n-1} = \frac{\bar{x} - \mu_0}{\hat{s}}\sqrt{n}.$$

Pod uvjetom da H_0 sadrži slučajnu varijablu $\boldsymbol{t}$ koja slijedi Studentovu t-distribuciju s $n - 11$ stupnjeva slobode. Iz tablice ***t-distribucije*** ili pomoću kalkulatora inverzne normalne distribucije (pogledajte na primjer: https://www.statology.org/inverse-t-distribution-calculator) ili pomoću neke proračunske tablice kao što su Excel, OpenOffice Calc, nalazimo takvu kritičnu vrijednost t_α da

$$P(|t| \geq t_\alpha) = \alpha.$$

Skup vrijednosti t definiran nejednakošću $|t| \geq t_\alpha$ kritično je područje ovog testa, tj. ako iz formule dobijemo takvu vrijednost $\boldsymbol{t}$ da je $|t| \geq t_\alpha$, hipoteza H_0 mora biti odbačena u korist H_1.

Ako je $|t| < t_\alpha$ ne može se odbaciti hipoteza H_0.

Gore opisani test je naravno dvostran (vidi Model I) budući da imamo dvostrano kritično područje, tj.

$$|t| \geq t_\alpha$$

Ovo se događa kad god imamo alternativnu hipotezu oblika $H_1: \mu \neq \mu_0$.

Co-funded by the
Erasmus+ Programme
of the European Union

Međutim, u nekim slučajevima moramo provesti jednostrane testove.
Kada je alternativna hipoteza oblika H_1: $\mu < \mu_0$, imamo lijevo kritično područje, tj. $t \leq t_\alpha$. U ovom slučaju izvodimo t_α tako da

$$P(t \leq t_\alpha) = \alpha.$$

Kada je alternativna hipoteza oblika $H_1: \mu > \mu_0$, imamo desnostrano kritično područje, tj. $t \geq t_\alpha$. U ovom slučaju izvodimo t_α tako da

$$P(t \geq t_\alpha) = \alpha.$$

Primjer 9.35

Stroj izrađuje metalne ploče određenih dimenzija, čija je nominalna debljina **0.04 mm**. Nakon slučajnog odabira i mjerenja **25** ploča, njihova je prosječna debljina bila
$\bar{x}$=0.037 mm i $\hat{s}$=0.005 mm. Može li se zaključiti da su proizvedene ploče tanje od **0.04 mm**? Pretpostavimo razinu značajnosti **α=0.01**.

Rješenje:

Nulta hipoteza H_0 je da je prosječni trošak materijala za proizvodnju proizvoda **600 €,** tj.

$$H_0: \mu_0 \geq 0.04 \text{ and } H_1: \mu_0 < 0.04.$$

Izračunaj

$$t = \frac{0.037 - 0.04}{0.005}\sqrt{25} = -3$$

Koristimo ***t***-distribuciju s **24** stupnja slobode i test je lijevostran.
Iz tablice ***t***-distribucije ili pomoću kalkulatora inverzne normalne distribucije (pogledajte na primjer: https://www.statology.org/inverse-t-distribution-calculator) ili pomoću neke popularne proračunske tablice kao što su Excel, OpenOffice Calc, dobivamo $t_\alpha = -2.492$.
Imamo $t = -3 < -2.492 = t_\alpha$.
Stoga ***t*** pripada kritičnom području. Moramo odbiti H_0 u korist H_1. Proizvedene ploče su tanje od **0.04 mm**.

Co-funded by the
Erasmus+ Programme
of the European Union

9.10 REGRESIJA, KORELACIJA

Pretpostavim da imamo dva uzorka podataka, različitih statističkih karakteristika, iz određene populacije.

$$x:\ \ x_1, x_2, \dots, x_n \qquad y:\ \ y_1, y_2, \dots, y_n$$

Zanima nas sljedeće pitanje: *Postoji li neka veza između uzoraka populacije*?

Definicija: Kovarijanca

Kovarijanca od x, y se definira kao

$$cov(x, y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

Definicija: Pearsonov koeficijent korelacije

Pearsonov koeficijent korelacije se definira kao

$$r = \frac{cov(x, y)}{s(x)s(y)}$$

gdje je $s(x)$ standardna devijacija od x i $s(y)$ je standardna devijacija od y.

<u>Napomena:</u>

Koeficijent r ima vrijednost između -1 i 1.

Vrijednost 1 srednja vrijednost znači da postoji ukupna pozitivna linearna korelacija između x i y, 0 da ne postoji linearna korelacija između x i y, a -1 da postoji ukupna negativna linearna korelacija između x i y.

Primjer 9.36

Za

$$x:\ \ 3, 3, 4, 5, 5 \qquad y:\ \ 5, 7, 6, 4, 8$$

$$\bar{x} = 4, \qquad \bar{y} = 6$$

$$cov(x, y) =$$

$$\frac{1}{5}\big((3-4)(5-6) + (3-4)(7-6) + (4-4)(6-6) + (5-4)(4-6) + (5-4)(8-6)\big)$$

$$= \frac{1}{5}(1 - 1 + 0 - 2 + 2) = 0$$

$$r = \frac{0}{s(x)s(y)} = 0.$$

Primjer 9.37

Za

$$x:\ \ 3, 3, 4, 5, 5 \qquad y:\ \ 5, 8, 6, 6, 10$$

$$\bar{x} = 4, \qquad \bar{y} = 7$$

$$cov(x, y) =$$

$$\frac{1}{5}\big((3-4)(5-7)+(3-4)(8-7)+(4-4)(6-7)+(5-4)(6-7)$$
$$+(5-4)(10-7)\big)=\frac{1}{5}(2-1+0-1+3)=\frac{3}{5}$$
$$s^2(x)=$$
$$\frac{1}{5}((3-4)^2+(3-4)^2+(4-4)^2+(5-4)^2+(5-4)^2)$$
$$=\frac{1}{5}(1+1+0+1+1)=\frac{4}{5}$$
$$s(x)=\sqrt{\frac{4}{5}}=\frac{2}{\sqrt{5}}.$$
$$s^2(y)=$$
$$\frac{1}{5}((5-7)^2+(8-7)^2+(6-7)^2+(6-7)^2+(10-7)^2)$$
$$=\frac{1}{5}(4+1+1+1+9)=\frac{16}{5}$$
$$s(y)=\sqrt{\frac{16}{5}}=\frac{4}{\sqrt{5}}.$$
$$r=\frac{\frac{3}{5}}{\frac{2}{\sqrt{5}}\cdot\frac{4}{\sqrt{5}}}=\frac{3}{8}=0.375.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.10.1 Linearna regresija

Pretpostavimo da imamo dva uzorka:

$$x:\ \ x_1, x_2, \dots, x_n, \qquad y:\ \ y_1, y_2, \dots, y_n.$$

Primjenom ***linearne regresije*** modelira se veza između dvije varijable x i y.

Jedna od varijabli (x) naziva se ***neovisna*** varijabla. Druga varijabla (y) naziva se ***ovisna*** (***zavisna)*** varijabla.
Veza između x i y modelira se primjenom linearne funkcije

$$y = ax + b$$

čiji se nepoznati koeficijenti a, b procijene iz podataka.

- Koeficijenti a, b najčešće se određuju metodom ***najmanjih kvadrata*** koja minimizira vrijednosti kvadrata udaljenosti između opaženih podataka i regresijske krivulje (pravca).

Moramo riješiti sljedeći problem

$$\sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min, \ \ a, b = ?$$

Moguće je pokazati da takav problem minimalizacije uvijek ima rješenja a, b dobivena pomoću formula:

$$a = \frac{cov(x, y)}{s^2(x)}, \ \ b = \bar{y} - a\bar{x},$$

gdje je

- $cov(x, y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$ kovarijanca od x, y,
- $s^2(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ varijacija uzorka od x i
- $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ su srednja vrijednost od x i y, respektivno.

Primjer 9.38

Zadani su uzorci podataka

$$x:\ \ 3, 5, 2, 2, 1, 4, 6, 1 \qquad y:\ \ 3, 4, 3, 4, 2, 5, 4, 3.$$

Vrijedi

$$cov(x, y) = 1, \ \ s^2(x) = 3, \ \ \bar{x} = 3, \ \ \bar{y} = \frac{7}{2}$$

$$a = \frac{cov(x, y)}{s^2(x)} = \frac{1}{3}, \qquad b = \bar{y} - a\bar{x} = \frac{7}{2} - \frac{1}{3} \cdot 3 = \frac{5}{2}.$$

Linearna regresija je $y = \frac{1}{3}x + \frac{5}{2}$. Na sljedećoj slici prikazan je njen graf.

Co-funded by the
Erasmus+ Programme
of the European Union

Slika 9.8. *Regresijska funkcija* $y = \frac{1}{3}x - \frac{5}{2}$

Zadatak 9.8

Odrediti linearnu regresiju za podatke

$$x:\ \ 1, 2, 4 \qquad y:\ \ 0, 2, 1.$$

Rješenje:

$$y = \frac{3}{14}x + \frac{1}{2}.$$

Co-funded by the
Erasmus+ Programme
of the European Union

9.11 PRIMJENA

9.11.1 Pouzdanost - primjena Poissonove distribucije

Vjerojatnost pojave k kvarova u jedinici vremena zadana je kao

$$f(k) = e^{-\lambda}\frac{\lambda^k}{k!},$$

gdje je λ prosječni broj kvarova (očekivani broj kvarova) u jedinici vremena.

Funkcija

$$f(k;\lambda,t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

je vjerojatnost pojave k kvarova u trenutku t.

Potrebno je uočiti da vjerojatnost od 0 kvarova u trenutku t je zadana kao

$$f(0;\lambda,t) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

Definicija: Funkcija pouzdanosti

Funkcija pouzdanosti $R(t)$ je vjerojatnost od nula kvarova u vremenskom intervalu duljine t. Definirana je kao

$$R(t) = \begin{cases} 1, & t < 0 \\ e^{-\lambda t}, & t \geq 0 \end{cases}.$$

Vjerojatnost pojave najviše m kvarova u vremenskom intervalu duljine $t > 0$ zadana je kao

$$R(t,m) = \sum_{k=0}^{m} e^{-\lambda t}\frac{(\lambda t)^k}{k!}.$$

Zadatak 9.9

Neki složeni sustav ima prosječnu stopu kvarova tranzistora $\lambda = 0.002$ po satu. Potrebno je odrediti pouzdanost za 30 dana ako broj kvarova tranzistora ne prelazi 1.

Rješenje:

$\lambda = 0.002$
$t = 30 \cdot 24 = 720$
$m \leq 1$
$\lambda t = 1.44$

Moramo izračunati $R(720,1)$.
$R(720,1) = e^{-1.44} + e^{-1.44} \cdot 1.44 \approx 0.58$

Co-funded by the Erasmus+ Programme of the European Union

9.12 VJEŽBE

1. Koliko ima dvoznamenkastih neparnih prirodnih brojeva većih od 30?
2. Koliko ima troznamenkastih parnih prirodnih brojeva manjih od 500?
3. Na koliko načina možemo rasporediti slova riječi MONTANA?
4. U zemlji A postoji sedam luka. Na koliko načina tamo mogu pristati četiri različita broda?
5. Na koliko se načina u četiri kutije može sakriti šest novčića?
6. Šest točaka leži na kružnici. Koliko se upisanih trokuta može nacrtati s tim točkama kao vrhovima?
7. Ako su slova algebre riječi postavljena nasumično u nizu, kolika je vjerojatnost da će dva uzastopna slova biti a.
8. Ako su slova riječi oko postavljena nasumično u nizu, kolika je vjerojatnost da će tri uzastopna slova biti samoglasnici.
9. Par želi imati troje djece. Pretpostavljamo da je jednako vjerojatan spol djeteta. Navedite vjerojatnosti da:
   a) par ima barem jednog dječaka
   b) nema djevojčice starije od dječaka
   c) par ima točno jednu djevojčicu.
10. Dva broda A i B imaju vrijeme dolaska između 13 i 17 sati. Oba broda moraju pristati na isti vez. Nakon pristajanja, svakom brodu treba 30 minuta da obnovi zalihe i napusti dok. Kolika je vjerojatnost
    a) da brodovi neće morati čekati na vez
    b) da brod A neće morati čekati na vez
    c) da će brod B morati čekati na vez.
11. Pretpostavimo da su dvije kuglice izvučene bez zamjene (prva kuglica nije zamijenjena prije izvlačenja druge) nasumce iz vrećice koja sadrži 4 crvene i 3 crne kuglice. Neka je A događaj izvlačenja crvene kuglice prvi put, a B događaj izvlačenja crvene kuglice drugi put. Jesu li događaji A i B neovisni?
12. Tri identične zdjele označene su $1, 2, 3$. Prva posuda sadrži 3 crvene i 3 plave kuglice. Druga posuda sadrži 4 crvene i 2 plave kuglice. Treća posuda sadrži 1 crvenu i 5 plavih kuglica. Nasumično se odabire zdjela, a iz nje se nasumično bira kuglica. a) Koja je vjerojatnost da je odabrana kuglica plava? b) S obzirom da je odabran a kuglica crvena, koja je vjerojatnost da je odabrana zdjela 2?
13. Pretpostavimo da je vjerojatnost da će John riješiti određeni problem $\frac{2}{3}$, da će ga Mary riješiti $\frac{3}{4}$, a da će ga Bill riješiti $\frac{1}{2}$. Kolika je vjerojatnost
    a) da će bar jedna osoba to riješiti
    b) da će Mary i Bill to riješiti, ali John neće
    c) da će John i Mary to riješiti, ali Bill neće
    d) da će to riješiti barem dvoje ljudi.

Co-funded by the
Erasmus+ Programme
of the European Union

14. Kutija sadrži tri kovanice: dvije obične kovanice i jednu lažnu kovanicu s dvije glave $(P(H) = 1)$. Nasumično odaberete novčić i bacite ga.
    a) Kolika je vjerojatnost da će pasti jedan na drugoga?
    b) Nasumično odaberete novčić i bacite ga i dobijete glave. Koja je vjerojatnost da je to novčić s dvije glave?

15. Bacite pravi novčić tri puta:
    a) Kolika je vjerojatnost tri glave?
    b) Koja je vjerojatnost da promatrate točno jednu glavu?
    c) S obzirom da ste promatrali barem jednu glavu, koja je vjerojatnost da promatrate najmanje dvije glave?

16. Artikle koji prolaze kroz inspekcijsku liniju vizualno pregledavaju dva uzastopna inspektora. Kada proizvod s nedostatkom prođe kroz liniju za inspekciju, vjerojatnost da će ga prvi inspektor dobiti je 0,1. Drugi inspektor će propustiti pet od deset neispravnih predmeta koji prođu pored prvog inspektora. Koja je vjerojatnost da neispravan artikl dobiju oba inspektora?

17. Na ispitu se zadaju dva problema s zaključivanjem, 1 i 2. 35% učenika riješilo je zadatak 1, a 15% učenika riješilo je oba zadatka. Koliko je učenika koji će riješili drugi zadatak ako su riješili prvi?

18. U skupini od 50 ljudi, od 35 koji puše 20 je muškaraca. Od 15 osoba koje ne puše, 10 je žena. Koja je vjerojatnost da ako je slučajno odabrana osoba muškarac da je onda i pušač?

19. Određena bolest ima stopu incidencije od 2%. Pretpostavimo da je za neki dijagnostički test lažno negativna stopa 1%, a lažno pozitivna stopa 1%. Izračunajte vjerojatnost da će osoba, nasumično odabrana iz populacije:
    a) tko je pozitivan na testu zapravo ima bolest.
    b) tko je negativan, zapravo nije bolestan

20. Distribucija vjerojatnosti za slučajnu varijablu X prikazana je u tablici.

x_i	-3	-1	0	2	3
p_i	0.1	0.4	0.2	0.2	0.1

Pronađite srednju vrijednost, varijancu i standardnu devijaciju X..

21. Pet od 100 muškaraca i dvije od 1000 žena su daltonisti. Iz grupe od istog broja muškaraca i žena nasumično je uzeta jedna osoba. Utvrđeno je da je ova osoba slijepa za boje. Koja je vjerojatnost da je to bio muškarac?

22. Tri identične zdjele označene su $1, 2, 3$. Prva posuda sadrži 3 crvene i 4 plave i 3 crne kuglice. Druga posuda sadrži 6 crvenih, 2 plave i 2 crne kuglice. Treća posuda sadrži 2 crvene i 5 plavih i 3 crne kuglice. Nasumično se odabire zdjela, a iz nje se nasumično bira kuglica.

    a) Koja je vjerojatnost da je odabrana kuglica crna?
    b) S obzirom da je odabrana crna kuglica, koja je vjerojatnost da je odabrana zdjela 3?
    c) S obzirom da je odabrana plava kuglica, koja je vjerojatnost da je odabrana zdjela 1?

Co-funded by the
Erasmus+ Programme
of the European Union

23. Tri identične zdjele označene su 1,2,3. Prva posuda sadrži 3 crvene i 4 plave i 3 crne kuglice. Druga posuda sadrži 6 crvenih, 2 plave i 2 crne kuglice. Treća posuda sadrži 2 crvene i 5 plavih i 3 crne kuglice. Nasumično je odabrana zdjela, a iz zdjele su nasumično odabrane dvije kuglice bez zamjene.
    a) Koja je vjerojatnost da su obje odabrane kuglice plave?
    b) S obzirom da su obje odabrane kuglice plave, koja je vjerojatnost da je odabrana zdjela 2?
    c) S obzirom da su odabrane plava i crna kuglica, koja je vjerojatnost da je odabrana zdjela 2?
24. Za nadzornika u tvornici rade tri muškarca i tri žene. Želi se izabrati dva radnika za poseban posao. Odluči se nasumično odabrati dva radnika. Neka Y označava broj žena u izboru. Pronađite distribuciju vjerojatnosti za Y.
25. Svaka od tri kuglice se nasumično stavlja u jednu od tri zdjelice. Nađite distribuciju vjerojatnosti za Y = broj praznih zdjelica.
26. Novčić se baca tri puta. Neka je Y jednako broju promatranih glava.
    a) Izračunajte vjerojatnosti povezane s Y = 0,1,2 i 3.
    b) Konstruirajte tablicu distribucije vjerojatnosti.
    c) Pronađite očekivanu vrijednost i standardnu devijaciju Y.
27. Osiguravajuće društvo izdaje jednogodišnju policu od $\$2000$ kojom se osigurava od slučaja A koji se povijesno događa jednom od svakih 100 vlasnika police. Koliko bi tvrtka trebala naplatiti policu ako zahtijeva da očekivani profit po polici bude $\$75$?
28. Košarkaš izvodi 4 nezavisna slobodna bacanja s vjerojatnošću od 0,7 da će postići koš pri svakom šutu. Neka je Y = broj koševa koje postigne. Odredite distribuciju vjerojatnosti za slučajnu varijablu Y. Odredite vjerojatnost da on postigne najmanje 3 koša.
29. Pretpostavimo da radio sadrži šest tranzistora, od kojih su dva neispravna. Nasumično se odaberu tri tranzistora, izvade iz radija i pregledaju. Neka je X jednak broju opaženih nedostataka. Pronađite distribuciju vjerojatnosti za X.
30. Pretpostavimo da su dvije kuglice izvučene sa zamjenom (prva je kuglica zamijenjena prije izvlačenja druge) nasumce iz vrećice koja sadrži 5 crvenih i 3 crne kuglice. Neka je X jednako broju izvučenih crnih kuglica. Nađite distribuciju vjerojatnosti za X.
31. Pretpostavimo da su dvije kuglice izvučene bez zamjene nasumce iz vrećice koja sadrži 5 crvenih i 3 crne kuglice. Neka je X jednako broju izvučenih crnih kuglica. Nađite distribuciju vjerojatnosti za X
32. Distribucija vjerojatnosti za slučajnu varijablu X dana je u tablici.

x_i	0	2	4	5	6
p_i	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{12}$

    a) Pronađite kumulativnu funkciju distribucije od X. Skicirajte graf.

Co-funded by the Erasmus+ Programme of the European Union

b) Izračunajte: : i) $P(2 \leq X \leq 4)$ ii) $P(0 < X < 4)$ iii) $P(X > 1)$.

33. Nogometaš izvodi tri neovisna jedanaesterca s vjerojatnošću $\frac{4}{5}$ da postigne gol pri svakom udarcu. Neka X bude broj golova koje postigne. Pronađite distribuciju vjerojatnosti za slučajnu varijablu X. Pronađite očekivani broj golova.

34. Varijabla X ima normalnu distribuciju sa srednjom vrijednošću od 10 i standardnom devijacijom od 2. Jedan rezultat je slučajno odabran. Kolika je vjerojatnost da je između 11 i 12?

35. Dubina mora izmjerena je u 5 neovisnih pokusa, a rezultati su bili (u metrima): $862, 870, 876, 866, 871.$ Znajući da je distribucija mjerenja normalna sa standardnom devijacijom od 5 m, a razina značajnosti α=0.05, potvrdite hipotezu da je prosječna dubina mora u tom području $870m$..

36. U biokemijskom eksperimentu izmjeren je životni vijek pojedinih organizama. Raspodjela tog vremena može se pretpostaviti normalnom. Obavljeno je 8 mjerenja, a rezultati su (u satima): 4,7; 5.3; 4,0; 3,8; 6.2; 5,5; 4,5; 6.0. Uz pretpostavku razine značajnosti α=0.05, ocijenite hipotezu da je prosječni životni vijek ovih organizama 4.0 sata.

37. Nađite funkciju linearne regresije za podatke: $x = -1, 1, 2, \quad y = 1, -2, 2.$

38. Neki složeni sustav ima prosječnu stopu kvarova $\lambda = 0.005$ kvarova lampe po satu. Kolika je pouzdanost za razdoblje od 60 dana ako broj kvarova lampe ne može biti veći od 2?

Rješenjes:

1. 35
2. 200
3. 1260
4. 720
5. 4096
6. 20
7. $\frac{2}{7}$
8. $\frac{1}{10}$
9. a) $\frac{7}{8}$ b) $\frac{1}{2}$ c) $\frac{3}{8}$
10. a) ≈ 0.7656 b) ≈ 0.8828 c) ≈ 0.1172
11. not
12. a) $\frac{5}{9}$ b) $\frac{1}{2}$
13. a) $\frac{23}{24}$ b) $\frac{1}{8}$ c) $\frac{1}{4}$ d) $\frac{17}{24}$
14. a) $\frac{2}{3}$ b) $\frac{1}{2}$
15. a) $\frac{1}{8}$ b) $\frac{3}{8}$ c) $\frac{4}{7}$
16. 0.05
17. 42.8%

Co-funded by the
Erasmus+ Programme
of the European Union

18. $\frac{4}{5}$
19. a) ≈ 0.6689 b) ≈ 0.9998
20. $0,\ \ 3\ ,\ \ \sqrt{3}$
21. $\frac{25}{26}$
22. a) $\frac{4}{15}$ b) $\frac{3}{8}$ c) $\frac{4}{11}$
23. a) $\frac{17}{135}$ b) $\frac{1}{17}$ c) $\frac{4}{393}$
24. $P(Y=0)=\frac{1}{5},\ \ P(Y=1)=\frac{3}{5},\ \ P(Y=2)=\frac{1}{5}.$
25. $P(Y=0)=\frac{2}{9},\ \ P(Y=1)=\frac{2}{3},\ \ P(Y=2)=\frac{1}{9}.$
26. a) $P(Y=0)=\frac{1}{8},\ P(Y=1)=\frac{3}{8},\ P(Y=2)=\frac{3}{8},\ P(Y=3)=\frac{1}{8}$ c) $E(Y)=1.5,$ $V(Y)=0.75.$
27. \$95
28. $P(Y=0)=0.0081, P(Y=1)=0.0756, P(Y=2)=0.2646, P(Y=3)=0.4116,$ $P(Y=4)=0.2401,\ P(Y\geq 3)=0.6517.$
29. $P(X=0)=\frac{1}{5},\ \ P(X=1)=\frac{3}{5},\ P(X=2)=\frac{1}{5}.$
30. $P(X=0)=\frac{25}{64},\ P(X=1)=\frac{15}{32},\ \ P(X=2)=\frac{9}{64}.$
31. $P(X=0)=\frac{5}{14},\ \ P(X=1)=\frac{15}{28},\ \ P(X=2)=\frac{3}{28}.$
32. a) $F(x)=0\ \ x<0,\ \ F(x)=\frac{1}{3}\ \ 0\leq x<2,\ \ F(x)=\frac{5}{12}\ \ 2\leq x<4,\ \ F(x)=\frac{3}{4}\ \ 4\leq x<5,\ \ F(x)=\frac{11}{12}\ \ 5\leq x<6,\ F(x)=1\ \ x\geq 6.$ b) i) $\frac{5}{12}$ ii) $\frac{1}{12}$ iii) $\frac{2}{3}$.
33. $P(X=0)=\frac{1}{125},\ \ P(X=1)=\frac{12}{125},\ \ P(X=2)=\frac{48}{125},\ \ P(X=3)=\frac{64}{125}.$ $E(X)=2.4.$
34. 0.1499
35. $|u|=0.447<1.96=u_{\alpha},\ H_0$ ne može se odbaciti
36. $|t|=3.17>2.365=t_{\alpha},\ H_0$ mora se odbaciti
37. $y\ \cong 0.0714x+0.2857$
38. ≈ 0.0255