Slučajne varijable
Materijali za nastavu iz Statistike

Kristina Krulić Himmelreich i Ksenija Smoljak

2012/13

Slučajna varijabla

Slučajna varijabla je funkcija X koja elementarnim događajima pridružuje brojeve. Dakle,

$$X : \Omega \to \mathbb{R}.$$

Slučajne varijable dijelimo na:

1. Diskretne slučajne varijable
2. Neprekidne (kontinuirane) slučajne varijable

Diskretne slučajne varijable

Označimo s R_X skup svih različitih vrijednosti koje slučajna varijabla X može poprimiti. Kažemo da je zadan **zakon razdiobe** ili **distribucija** slučajne varijable X ako je zadan skup $R_X = \{a_1, a_2, a_3, \dots\}$, te niz brojeva $p_i \geq 0$ tako da
1) $p_i = P(X = a_i)$
2) $\sum_{i=1}^{\infty} p_i = 1$
Zakon razdiobe zapisujemo u obliku tablice:

$$X \sim \begin{pmatrix} a_1 & a_2 & a_3 & \dots \\ p_1 & p_2 & p_3 & \dots \end{pmatrix}$$

Budući je skup svih vrijednosti koje slučajna varijabla može poprimiti $R_X = \{a_1, a_2, a_3, \dots\}$ diskretan (konačan ili prebrojiv) skup, kažemo da je X **diskretna slučajna varijabla**.

Diskretne slučajne varijable

Neka je $X : \Omega \to \mathbb{R}$ slučajna varijabla. **Funkcija vjerojatnosti** od X je funkcija $f : \ R_X \to [0, 1]$ definirana s

$$f(a_i) := P(X = a_i) = p_i$$

Funkcija distribucije slučajne varijable X je funkcija $F : \mathbb{R} \to [0, 1]$ definirana s

$$F(x) := P(X \leq x), \quad x \in \mathbb{R}.$$

Vrijedi

$$F(x) = \sum_{a_i \leq x} p(a_i).$$

Matematičko očekivanje diskretne slučajne varijable je broj $\mathsf{E}[X]$ definiran s

$$\mathsf{E}[X] = \sum_{a_i \in R_X} a_i \cdot f(a_i)$$

Broj

$$\mathsf{V}[X] := \mathsf{E}[X^2] - (\mathsf{E}[X])^2$$

zove se **varijanca** diskretne slučajne varijable X.
Standardna devijacija slučajne varijable X je broj

$$\sigma_X := +\sqrt{\mathsf{V}[X]}$$

Zadaci

Zadatak

Pokažite da je funkcija f dana tablicom funkcija vjerojatnosti neke slučajne varijable X. Izračunajte $E[X]$ i $V[X]$.

x_i	1	2	3	4	5
$f(x_i)$	$\frac{1}{12}$	$\frac{5}{24}$	$\frac{1}{6}$	$\frac{3}{8}$	$\frac{1}{6}$

Zadatak

Odredite konstantu a tako da funkcija f dana tablicom bude funkcija vjerojatnosti slučajne varijable X. Izračunajte $E[X]$ i $V[X]$.

x_i	-1	0	2	3	4
$f(x_i)$	$\frac{a}{2}$	$\frac{a}{10}$	$\frac{a}{5}$	$\frac{a}{4}$	$\frac{a}{20}$

Binomna razdioba

Osnovna svojstva koja opisuju binomnu razdiobu (distribuciju):

1. Izvodimo pokus koji ima dva moguća ishoda. Jedan ishod ćemo zvati "uspjeh" a drugi "neuspjeh".
2. Vjerojatnost "uspjeha" jednaka je p. Vjerojatnost "neuspjeha" je tada $q = 1 - p$.
4. Pokus ponavljamo n puta. Pokusi su međusobno nezavisni.
5. Binomna slučajna varijabla broji broj "uspjeha" k u tih n pokusa.

Binomna razdioba

Slučajna varijabla X ima **binomnu razdiobu** ili **distribuciju** s parametrima n i p ako je X poprima vrijednosti iz skupa $\{0, 1, 2, \dots, n\}$ s vjerojatnostima

$$P(X = k) = \binom{n}{k} p^k \cdot q^{n-k}, \qquad 0 \leq k \leq n,$$

gdje je $q = 1 - p$. Slučajnu varijablu X koja ima binomnu razdiobu označavamo s:

$$X \sim B(n, p)$$

Očekivanje binomne razdiobe: $\quad \mathrm{E}[X] = np$
Varijanca binomne razdiobe: $\quad \mathrm{V}[X] = npq$

Zadaci

Zadatak

Baca se ispravna kocka 10 puta. Nađite vjerojatnost da se

(a) 5 puta pojavi broj 6,

(b) barem jednom pojavi broj 6,

(c) 7 puta pojavi neparan broj.

(d) Koji je očekivani broj pojavljivanja broja većeg od 4?

Zadatak

U skladištu je 1000 proizvoda među kojima je 50 neispravnih. Izvlačimo na sreću 5 proizvoda s vraćanjem. Izračunajte vjerojatnost da je od tih 5 proizvoda

(a) točno 1 neispravan,

(b) najviše 1 neispravan

(c) barem 1 neispravan.

Hipergeometrijska razdioba

Osnovna svojstva koja opisuju hipergeometrijsku razdiobu:

1. U skupu od N elemenata njih M ima neko svojstvo, dok preostalih $N - M$ nema to svojstvo
2. Pokus se sastoji od slučajnog izvlačenja, bez vraćanja, n elemenata iz skupa od N elemenata
3. Hipergeometrijska slučajna varijabla broji broj "uspjeha" (odnosno elemenata koji imaju neko svojstvo) k u izvlačenju ukupno n elemenata.

Hipergeometrijska razdioba

Slučajna varijabla X ima **hipergeometrijsku razdiobu** ili **distribuciju** u oznaci $X \sim Hg(M, N, n)$ ako je funkcija gustoće te slučajne varijable zadana s:

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \quad \max\left(0, n - (N - M)\right) \leq k \leq \min\left(M, n\right)$$

Očekivanje hipergeometrijske razdiobe: $\quad \mathrm{E}\left[X\right] = \dfrac{nM}{N}$

Napomena: Za dovoljno velike n vrijedi $Hg(M, N, n) \sim B(n, p)$.

Zadatak

Riješite prethodni zadatak uz uvjet da se proizvodi ne vraćaju.

Zadaci

Zadatak

Odredite očekivani broj dječaka u obitelji s 8 djece pod pretpostavkom da je spol djeteta jednakovjerojatan. Kolika je vjerojatnost da će se ostvariti očekivani broj dječaka?

Zadatak

Vjerojatnost da je proizvod proizveden u tvornici neispravan je 0.02. Pošiljka od 10000 proizvoda poslana je u prodaju. Nađite očekivanje i standardnu devijaciju broja neispravnih proizvoda.

Zadatak

Vjerojatnost da strijelac pogodi cilj je 0.25. Strijelac 7 puta gađa cilj. Izračunajte vjerojatnost da pogodi cilj barem dvaput.

Poissonova razdioba

Slučajna varijabla X ima **Poissonovu razdiobu** ili **distribuciju** s parametrom $\lambda > 0$ ako je funkcija gustoće te slučajne varijable zadana formulom:

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

Slučajnu varijablu X koja ima Poissonovu razdiobu označavamo s:

$$X \sim P(\lambda)$$

Očekivanje Poissonove razdiobe: $\quad \mathrm{E}[X] = \lambda$
Varijanca Poissonove razdiobe: $\quad \mathrm{V}[X] = \lambda$
Binomna razdioba $B(n, p)$ može se **aproksimirati** Poissonovom razdiobom $P(np)$, to jest $\lambda = np$. Aproksimacija je to bolja što je parametar n veći, a parametar p manji.

Zadaci

Zadatak

Vjerojatnost da je proizvod neispravan iznosi 1%. Iz skladišta uzimamo paket od 100 proizvoda. Kolika je vjerojatnost da od tih 100 proizvoda

(a) je 5 neispravnih, (b) broj neispravnih nije veći od 10?

Zadatak

Ako je vjerojatnost da je jedna osoba alergična na pelud 0.002, odredite vjerojatnost da su od 4000 osoba na pelud

(a) alergične 4 osobe (b) alergične više od 2 osobe.

Zadatak

Pretpostavimo da je 2% proizvoda neke tvornice neispravno. Nađite vjerojatnost da u uzorku od 100 proizvoda postoje 3 neispravna.

Neprekidne slučajne varijable

Za slučajnu varijablu X kažemo da je **neprekidna** ako vrijedi sljedeće:

(i) R_X je interval ili unija intervala u $\mathbb{R}$

(ii) postoji nenegativna funkcija $f : \mathbb{R} \to \mathbb{R}$ tako da za svaka dva broja a, b $(a < b)$ vrijedi

$$P(a \leq X \leq b) = \int_a^b f(t)dt$$

Funkciju f zovemo **funkcija gustoće** od X. Vjerojatnost da vrijednost slučajne varijable X upadne u interval $[a, b]$ jednaka je dakle površini ispod grafa funkcije gustoće na tom intervalu.

Funkcija distribucije

Funkcija distribucije F od X definirana je s:

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

Vrijedi:

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

Svojstva neprekidne slučajne varijable:

(1) Za svaki broj $a \in \mathbb{R}$ je $P(X = a) = 0$

(2)
$$\int_{-\infty}^{\infty} f(t)dt = P(-\infty < X < \infty) = 1$$

što znači da je ukupna površina ispod grafa funkcije gustoće jednaka 1.

Matematičko očekivanje i varijanca od X

Matematičko očekivanje od X definirano je s:

$$E[X] = \int_{-\infty}^{\infty} t \cdot f(t) dt,$$

a za **varijancu** vrijedi relacija kao i kod diskretnih slučajnih varijabli

$$V[X] = E[X^2] - (E[X])^2$$

gdje je sada

$$E[X^2] = \int_{-\infty}^{\infty} t^2 \cdot f(t) dt.$$

Zadaci

Zadatak

Dokažite da je funkcija

(a)

$$f(x) = \begin{cases} \frac{2}{9}x - \frac{4}{9} & \text{za } x \in [2,5] \\ 0 & \text{za } x \notin [2,5] \end{cases}$$

(b)

$$f(x) = \begin{cases} \cos x & \text{za } x \in [0, \frac{\pi}{2}] \\ 0 & \text{za } x \notin [0, \frac{\pi}{2}] \end{cases}$$

funkcija gustoće vjerojatnosti neke slučajne varijable X. Nacrtajte graf funkcije f te izračunajte $E[X]$.

Normalna razdioba

Kažemo da je slučajna varijabla X normalno distribuirana, ako je ona kontinuirana, $R_x = \mathbb{R}$, i ako je funkcija vjerojatnosti dana formulom

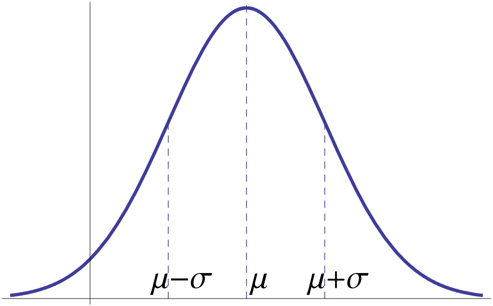
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

gdje su μ, σ $(\sigma > 0)$ proizvoljne konstante.

Pišemo: $X \sim N(\mu, \sigma^2)$

$E[X] = \mu,\ V[X] = \sigma^2$

Gaussova krivulja



Krivulja ovog tipa zove se Gaussova zvonolika krivulja, pa se normalna razdioba nekad zove i Gaussova razdioba.

Standardna normalna razdioba

Posebno je važna normalna razdioba s parametrima $\mu = 0$ i $\sigma^2 = 1$, tj. normalna razdioba $N(0,1)$. Ta razdioba se naziva **standardna normalna razdioba**, pišemo $X \sim N(0,1)$.
Standardna normalna funkcija gustoće vjerojatnosti

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

Standardna normalna funkcija distribucije

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}} t^2 dt$$

Vrijednosti od $\Phi(z)$ su tabelirane.

Veza između $N(\mu, \sigma^2)$ i $N(0, 1)$ distribucije

Ako X ima $N(\mu, \sigma^2)$ distribuciju, onda $Z = \frac{X-\mu}{\sigma}$ ima $N(0, 1)$ distribuciju.

$$P(a \leq X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Napomena: Pomoću normalne razdiobe može se aproksimirati binomna razdioba ako je n velik, a p nije blizu broja 0 ili 1 (u praksi $np \geq 10$).

$$X \sim B(n, p) \text{ tada je } Z = \frac{X - np}{\sqrt{npq}} \approx N(0, 1)$$

Primjer

Pretpostavimo da je visina studenata nekog sveučilišta slučajna varijabla, normalno distribuirana, $X \sim N(172, 7.5^2)$. Nađite vjerojatnost da je:

a) visina slučajno odabranog studenta između 175cm i 185cm
b) slučajno odabrani student niži od 170cm
c) slučajno odabrani student viši od 160cm
d) visina slučajno odabranog studenta između 165cm i 180cm
e) Koliki je očekivani broj studenata viših od 195cm, ako ima ukupno 3000 studenata?

Zadatak

U jednom skladištu od 1000 proizvoda je 20% proizvoda prve klase. Zbog kontrole kvalitete 100 puta je uziman 1 proizvod uz vraćanje. Nađite vjerojatnost da je broj proizvoda prve klase:

a) manji od 16

b) najmanje 30

Zadaci

Zadatak

Neka je X slučajna varijabla da standardnom normalnom distribucijom $N(0,1)$. Nađite:

a) $P(X \leq 1.42)$

b) $P(-1.37 \leq X \leq 2.01)$

c) $P(X \geq 1.13)$

d) $P(|X| \leq 0.5)$

e) $P(0.65 \leq |X|)$

Zadatak

Neka je X slučajna varijabla sa standardnom normalnom distribucijom $N(0,1)$. Odredite t ako je:

a) $P(0 \leq X \leq t) = 0.4236$

b) $P(X \leq t) = 0.7967$

c) $P(t \leq X \leq 2) = 0.1000$

Zadaci

Zadatak

Ako je $X \sim N(8, 4^2)$ skicirajte graf funkcije gustoće i odredite t tako da vrijedi:

a) $P(X \leq t) = 0.14$

b) $P(X \geq t) = 0.975$

c) $P(|X - 8| \leq t) = 0.9$

Zadatak

Pretpostavimo da je temperatura T u mjesecu lipnju normalno distribuirana s očekivanjem 20^oC i standardnom devijacijom 3^oC. Neđite vjerojatnost da je temperatura 15. lipnja između 21^oC i 27^oC.

Zadatak

Pretpostavimo da je na populaciji od 800 stanovnika slučajna varijabla visina V normalno distribuirana s očekivanjem 170cm i standardnom devijacijom 12.7. Nađite broj stanovnika s visinom:

a) između 165.1cm i 177.8cm

b) većom ili jednakom 182.88cm

Zadatak

Prema izvješću Hrvatskog auto kluba očekivano vrijeme između poziva i dolaska do mjesta nesreće je 25 minuta. Pretpostavimo da se radi o normalno distribuiranoj slučajnoj varijabli sa standardnom devijacijom 4.5 minute. Ako slučajno izaberemo 80 poziva, na koliko njih će biti reagirano u roku manjem 15 minuta?