4. DISKRETNE SLUČAJNE VARIJABLE

4.1. SLUČAJNE VARIJABLE. FUNKCIJA RAZDIOBE SLUČAJNE VARIJABLE. ČEBIŠEVLJEV POUČAK.

4.1.1. POJAM SLUČAJNE VARIJABLE

- Pri svakoj izvedbi nekoga slučajnog pokusa ostvaruje se točno jedan elementaran događaj.
- Svrha pokusa može biti i *mjerenje* numeričke veličine čije vrijednosti ovise o ishodu slučajnoga pokusa.
- Npr. u *izvlačenju prvoga broja* u igri LOTO 7/35 možemo *svakom* od 35 elementarnih događaja pridružiti pripadni izvučeni broj.
- Analogno možemo učiniti i u slučajnim pokusima *izvlačenje prvih dvaju brojeva*, ..., *izvlačenje svih sedam brojeva* (uz oprez: prostori elementarnih događaja u svakom od tih pokusa su međusobno različiti!)
- Preslikavanje koje *svakom* elementarnom događaju pridružuje neki realan broj pripada u kategoriju tzv. *slučajnih varijabli.*

4.1.2. DISKRETNA SLUČAJNA VARIJABLA

- Neka je S bilo koji *konačan* ili *prebrojiv* podskup skupa $\mathbb{R}$.
- *Podsjetnik*: Skup S je prebrojiv ako postoji bijekcija $f : S \to \mathbb{N}$.
- Promatramo slučajni pokus čiji je vjerojatnosni prostor $(\mathbf{\Omega}, \mathcal{F}, P)$.
- **Diskretna slučajna varijabla** je svako preslikavanje $X : \mathbf{\Omega} \to S$ takvo da je za svaki $s \in S$ skup $A_s := \{\omega \in \Omega\colon X(\omega) = s\} \in \mathcal{F}$.
- Zahtjev $A_s := \{\omega \in \Omega\colon X(\omega) = s\} \in \mathcal{F}$ sigurno vrijedi ako je $\mathcal{F} = \mathcal{P}(\mathbf{\Omega})$, ali u općem slučaju ne mora biti ispunjen.

4.1.3. PRIMJER 1.

- Neka je $S = [35] := \{1, 2, ..., 35\}$.
- Promatramo slučajni pokus *izvlačenje prvoga broja u igri* LOTO 7/35.
- Skup svih elementarnih događaja toga slučajnoga pokusa je $\mathbf{\Omega} = \{\omega_k : k \in [35]\}$, gdje je $\omega_k := \{\text{izvučen je broj } k\}$.
- Uz pretpostavku nenamještenoga izvlačenja, vjerojatnost *svakoga* od 35 elementarnih događaja iznosi

$$p_k = p = \frac{1}{35}.$$

- Preslikavanje $X : \mathbf{\Omega} \to S$ definirano pravilom:
  - $X(\omega_k) = k$,
- je diskretna slučajna varijabla. Njezina prirodna domena je skup $\mathbf{\Omega}$. Njezina slika je skup S.
- Na skupu $\mathbf{\Omega}$ moguće je zadati još neke diskretne slučajne varijable, npr. :
- $X(\omega_k)$ = zbroj svih znamenaka broja k,
- $X(\omega_k)$ = umnožak svih znamenaka broja k itd.
- Te diskretne slučajne varijable imaju *istu prirodnu domenu*, *ali različite slike* (i različita pravila).

4.1.4. PRIMJER 2.

- Neka je $S = \{0, 1, 2\}$.
- Promatramo slučajni pokus *nogometna utakmica između NK Akumulator i NK Uskok* (u 1. kolu međužupanijske lige).
- Skup svih elementarnih događaja je $\mathbf{\Omega} = \{\{neodlučeno\}, \{pobjeda\ NK\ Akumulator\}, \{pobjeda\ NK\ Uskok\}\}$.
- Uz pretpostavke da su obje momčadi podjednake "jakosti" i da utakmica nije namještena, sva tri elementarna događaja su jednako vjerojatna. Vjerojatnost svakoga od njih iznosi
  - $$p = \frac{1}{3}.$$
- Preslikavanje $X : \mathbf{\Omega} \to S$ definirano s:
- $X(\{neodlučeno\}) = 0$;
- $X(\{pobjeda\ NK\ Akumulator\}) = 1$;
- $X(\{pobjeda\ NK\ Uskok\}) = 2$;
- je diskretna slučajna varijabla.
- **Zadatak**: Sami zadajte barem još dvije diskretne slučajne varijable na skupu $\mathbf{\Omega}$.

4.1.5. PRIMJER 3.

- Neka je $S = \{0, 1\}$.
- Promatramo slučajni pokus *rođenje djeteta*.
- Skup svih elementarnih događaja je $\mathbf{\Omega} = \{\{$rođen je dječak$\}, \{$rođena je djevojčica$\}\}$.
- Uz pretpostavku da tijekom razdoblja trudnoće ultrazvukom nije određen djetetov spol, *oba* elementarna događaja su jednako vjerojatna. Vjerojatnost svakoga od njih iznosi
  - $p = \dfrac{1}{2}$.
- Svakom pojedinom elementarnom događaju prirodno je pridružiti spol rođenoga djeteta, tj.
- $\{$rođen je dječak$\} \mapsto$ dječak,
- $\{$rođena je djevojčica$\} \mapsto$ djevojčica.
- Ako *kodiramo* spol djeteta šiframa:
- dječak $\leftrightarrow 0$, djevojčica $\leftrightarrow 1$,
- gornje preslikavanje postaje diskretna slučajna varijabla $X : \mathbf{\Omega} \to S$.
- Bez provedenoga kodiranja gornje preslikavanje ne bismo svrstali u kategoriju slučajnih varijabli. Naime, iako je skup $A = \{$dječak, djevojčica$\}$ konačan, on nije podskup skupa realnih brojeva.
- Zbog ovakvih slučajnih pokusa, u definiciji slučajne varijable tražimo da je S podskup skupa $\mathbb{R}$, a ne bilo kakav konačan ili prebrojiv skup.

4.1.6. FUNKCIJA VJEROJATNOSTI

- Neka je S konačan ili prebrojiv podskup skupa $\mathbb{R}$.
- Neka je $(\Omega, \mathcal{F}, P)$ vjerojatnosni prostor i neka je $X : \Omega \to S$ diskretna slučajna varijabla.
- U definiciji slučajne varijable X navodi se da za svaki $s \in S$ skup $A_s := \{\omega \in \Omega \colon X(\omega) = s\}$ treba biti element algebre događaja $\mathcal{F}$.
- Zbog toga za svaki $s \in S$ *postoji jedinstveni* $p_s \in [0, 1]$ takav da je $p_s = P(A_s)$. p_s je, dakle, vjerojatnost događaja A_s.
- Zbog toga ima smisla promatrati preslikavanje koje svakom $s \in S$ pridružuje broj p_s. Kratko pišemo:
- $p_s = P(A_s) =: P(X = s)$
- i čitamo: "pe es je vjerojatnost da slučajna varijabla iks poprimi vrijednost es".

4.1.6. FUNKCIJA VJEROJATNOSTI

- Formalno, *funkcija vjerojatnosti diskretne slučajne varijable* je realna funkcija $f : S \to [0, 1]$ definirana pravilom:

$$f(s) = p_s.$$

- Za različite vrijednosti $s \in S$ događaji A_s su međusobno isključivi, a mora se dogoditi točno jedan od njih. Zbog toga vrijedi jednakost:

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

4.1.7. TABLICA RAZDIOBE

- Nakon što svakoj vrijednosti $s \in S$ pridružimo odgovarajuću vjerojatnost $p_s \in [0, 1]$, dobiveni rezultat pregledno možemo zapisati ovako:

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

- Gornji zapis nazivamo **tablica razdiobe** slučajne varijable X.
- Umjesto termina „tablica razdiobe" još se koriste termini **zakon razdiobe** i **distribucija**.

4.1.8. FUNKCIJA RAZDIOBE VJEROJATNOSTI

- Ako znamo tablicu razdiobe neke slučajne varijable, onda ima smisla definirati realnu funkciju $F : S \to [0, 1]$ s:

$$F(s) := P(X \leq s) = \sum_{s_i \leq s} P(X = s_i) = \sum_{s_i \leq s} p_i$$

- Funkcija F naziva se **funkcija razdiobe vjerojatnosti** (u literaturi često i kraće: *funkcija razdiobe*) slučajne varijable X.
- Prema definiciji, interpretacija funkcije F je:
- *$F(s)$ je* ***vjerojatnost*** *da slučajna varijabla X poprimi vrijednost jednaku ili manju od s.*

4.1.9. NAPOMENE

- **1.** Vrijede sljedeće “jezične ekvivalencije”:
- *barem* $\leftrightarrow$ *najmanje (jednako)* $\leftrightarrow$ *ne manje od* $\leftrightarrow$ *jednako ili veće od*
- *najviše (jednako)* $\leftrightarrow$ *ne veće od* $\leftrightarrow$ *jednako ili manje od.*
- **2.** Iz svojstava funkcije vjerojatnosti i definicije funkcije razdiobe vjerojatnosti slijedi **(oprez**: obratite pozornost na znakove nejednakosti):

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

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

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

4.1.10. OČEKIVANJE SLUČAJNE VARIJABLE

- Prema definiciji vjerojatnosti, tablicu razdiobe slučajne varijable X možemo protumačiti i ovako:
- *Tablica razdiobe slučajne varijable X određena je modalitetima slučajne varijable X* (to su svi elementi skupa S) *i njihovim relativnim frekvencijama* (za svaki $s \in S$ to je vjerojatnost p_s).
- Tako se tablica razdiobe može "poistovjetiti" s tablicom relativnih frekvencija iz opisne statistike, a funkcija razdiobe vjerojatnosti s kumulativnim nizom relativnih frekvencija „manje od".

4.1.10. OČEKIVANJE SLUČAJNE VARIJABLE

- U opisnoj statistici naučili smo pojmove *aritmetička sredina*, *varijanca* i *standardna devijacija*.
- Analogone tih pojmova primijenit ćemo na diskretne slučajne varijable.
- *Težinska* ili *vagana* aritmetička sredina svih vrijednosti slučajne varijable naziva se **matematičko očekivanje** ili, kraće, **očekivanje**.
- *Očekivanje* diskretne slučajne varijable X označava se s $E(X)$ i računa prema formuli:

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

4.1.10. OČEKIVANJE SLUČAJNE VARIJABLE

- **Oprez**: Ako je S beskonačan skup, očekivanje se definira kao zbroj odgovarajućega reda, pa općenito ne mora postojati (postojanje ovisi o tome je li dotični red konvergentan ili divergentan).
- Uzmemo li *bilo koje* a, $b \in \mathbb{R}$ i definiramo li slučajnu varijablu Y s

$$Y := a \cdot X + b \sim \begin{pmatrix} a \cdot s_1 + b & a \cdot s_2 + b & \ldots & a \cdot s_n + b & \ldots \\ p_1 & p_2 & \ldots & p_n & \ldots \end{pmatrix},$$

- tada nije teško pokazati da vrijedi jednakost:
- $E(a \cdot X + b) = a \cdot E(X) + b.$

4.1.11. NAPOMENA

- Ako je f bilo koja realna funkcija i X diskretna slučajna varijabla, tada je $Y := f\ (X)$ također diskretna slučajna varijabla.
- Njezina slika je skup $A = \{f\ (s)\ :\ s \in\ S\}$, a tablica razdiobe:

$$Y = f(X) \sim \begin{pmatrix} f(s_1) & f(s_2) & \ldots & f(s_n) & \ldots \\ p_1 & p_2 & \ldots & p_n & \ldots \end{pmatrix}$$

- Kad god je to moguće, tablicu razdiobe slučajne varijable Y treba "reducirati" tako da među vrijednostima $f\ (s_1),\ldots,\ f\ (s_n)$ nema međusobno jednakih brojeva.

4.1.12. VARIJANCA I STANDARDNA DEVIJACIJA SLUČAJNE VARIJABLE

- Analogno kao i za *grupirane* podatke u opisnoj statistici, definiraju se *varijanca* i *standardna devijacija* slučajne varijable X
- Pritom u definicijskim formulama "ulogu" aritmetičke sredine "igra" očekivanje slučajne varijable
- Definicijske formule su:

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

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

- Interpretacije su analogne onima iz opisne statistike:
- **Varijanca** diskretne slučajne varijable je prosječno kvadratno odstupanje vrijednosti slučajne varijable od očekivanja te varijable.
- **Standardna devijacija** diskretne slučajne varijable je prosječno odstupanje vrijednosti slučajne varijable od očekivanja te varijable.

4.1.13. NAPOMENA

- Može se pokazati da vrijede sljedeće jednakosti:

$$\mathbf{1}.\ V(X) = E\left(X^2\right) - \left(E(X)\right)^2.$$

$$\mathbf{2}.\ V(X) = E\left(\left(X - E(X)\right)^2\right).$$

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

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

4.1.14. ČEBIŠEVLJEV TEOREM

- **Podsjetnik**: *Čebiševljev poučak* u opisnoj statistici govorio je o "raspršenosti" elemenata statističkoga niza oko aritmetičke sredine toga niza.
- U „terminologiji" diskretnih slučajnih varijabli aritmetičku sredinu zamjenjujemo matematičkim očekivanjem. Tako dobivamo:
- **Teorem 1.** (*Čebiševljev teorem za diskretne slučajne varijable*) Neka je X diskretna slučajna varijabla s očekivanjem E i standardnom devijacijom σ. Tada za svaki $\alpha > 0$ vrijedi nejednakost:

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