4. lekcija SLUČAJNE VARIJABLE

Uvod: Pojam slučajne varijable i razdiobe vjerojatnosti

Sjetimo se da su ishodi ili elementarni događaji upravo oni konkretni događaji koji mogu biti rezultati razmatranog pokusa. Nakon provođenja pokusa, tj. pojavljivanja nekog ishoda, u pravilu nešto prebrojavamo ili mjerimo (kraće, nešto registriramo). Rezultat je neki broj, ovisno o tomu koji se ishod pojavio. Matematički se ta aktivnost može predočiti funkcijom koja svakom ishodu pridružuje neki broj.
Vrijednosti te funkcije ovise o ishodu koji se dogodio. Dakle ta je funkcija definirana na skupu svih ishoda u pokusu i slučajno (ovisno o ishodu koji se dogodio) postiže određene vrijednosti. Zato se ta funkcija zove **slučajnom varijablom**. Dakle:
Slučajna varijabla je funkcija koja svakom ishodu pridružuje neki realni broj.
Kraće:
Slučajna varijabla je funkcija X: S $\rightarrow$ R, gdje je S skup ishoda u nekom pokusu
Opišimo četiri tipična slučaja pojavljivanja slučajne varijable.

I. U nekom pokusu uočimo događaj A kojemu je vjerojatnost p. Pokus izvodimo nezavisno n puta. Slučajna varijabla X registrira koliko se puta pojavio događaj A.

II. U fiksiranom vremenskom intervalu u svakom se trenutku može, a ne mora, dogoditi događaj A (na primjer, u svakom trenutku na neku adresu može doći poruka). Slučajna varijabla X registrira broj pojavljivanja događaja A unutar tog vremenskog intervala (broj poruka na toj adresi).

III. U svakom trenutku može se, ali ne mora, dogoditi događaj A (na primjer, može stići poruka na neku adresu). Slučajna varijabla X registrira vrijeme između dvaju uzastopnih pojavljivanja događaja A (izmmeđu dviju uzastopnih poruka).

IV. Slučajna varijabla X registrira pogrješku pri mjerenju neke veličine (ili mjeri neku masu, temperaturu koja se razvuje u nekom pokusu i sl.)

To će redom biti tipični primjeri **binomne, Poissonove, eksponencjalne, normalne slučajne varijable**. Prve dvije su **diskretne**, a posljednje dvije **kontinuirane (neprekinute)**.

Diskretne slučajne varijable.

Binomna razdioba

Razmotrimo primjer I. Imamo p(A) = p, a tada suprotni događaj događaja *A* ima vjerojatnost q=1-p. Vidimo da X može postići vrijednosti 0,1,2,…,n (događaj *A* može se dogoditi ili nikako ili jednom ili dvaput itd. a najviše *n* puta). Dakle, skup vrijeednosti od X je R(X) = {0,1,2,…,n}.
Također:
p(X=i) = p(*i* puta se dogodio događaj *A*, a *n-i* puta se nije dogodio *A*)

$$= \binom{n}{i} \text{p(prvih } i \text{ puta dogodio se } A\text{, a ostalih } n\text{-}i \text{ puta nije se dogodio } A)$$

$$= \binom{n}{i} p^i(1-p)^{n-i} .$$

Broj $\binom{n}{i}$ koji se tu pojavljuje kao faktor može se shvatiti kao broj načina na koliko možemo od *n* mjesta izabrati *i* mjesta za događaj *A*. Treba uočiti da je taj faktor za i=0, odnosno za i=n jednak 1. Naime, $\binom{n}{0} = \binom{n}{n} = 1$.

To je bio **tipični primjer binomne razdiobe**. Treba uočiti da se ta razdioba opisuje pomoću **prirodnog broja** *n* (broj nezavisnih izvođenja pokusa) i **realnog broja** *p* između 0 i1 (vjerojatnost pojavljivanja događaja *A* u jednom izvođenju pokusa). Ti brojevi nazivaju se **parametrima razdiobe** (treba uočiti da je *n* **diskretan**, a *p* **kontinuiran** parametar).

Definicija 1. Slučajna varijabla X distribuirana je prema binomnom zakonu s parametrima *n* i *p*, ako je

R(X) = {0,1,…,n} i p(X=i) = $\binom{n}{i} p^i(1-p)^{n-i}$, $i \in$ R(X).

Ako je tako pišemo X~B(n,p).

Uočite da X ovisi o dvama parametrima: diskretnom parametru n i kontinuiranom parametru p.
Jedno od važnih svojstava slučajne varijable X je njena očekivana vrijednost ili očekivanje E(X). Intuitivno, to je prosječna vrijednost koju ta varijabla postiže. Preciznu definiciju navest ćemo poslije.
Pokušajmo problem očekivanja binomne razdiobe riješiti zdravorazumski. Pođimo od posebnog primjera kada je
p=1/6 i n= 120.
Možemo smatrati da je X slučajna varijabla koja broji koliko je bilo šestica pri 120 bacanja kocke. Jasno je da očekujemo 20 šestica (jednako kao i petica, četvorka itd.).
Dakle
E(X) = 20 = 120:6 = 120 ·1/6

Analognim zaključivanjem dobili bismo:

Ako je $X \sim B(n,p)$, onda je $E(X) = np$.

Primjer 1. Bacamo kocku 6 puta.
a) Koliko šestica očekujemo da ćemo dobiti?
b) Koliko puta treba baciti kocku pa da vjerojatnost da bude bar jedna šestica bude veća od 0.5?

Ako slučajna varijabla registrira broj šestica, onda je $X=B(6,1/6)$.

a) Očekivani broj šestica upravo je očekivanje slučajne varijable X, dakle $E(X)=6\cdot 1/6=1$.
b) Pretpostavimo da kocku treba baciti *n* puta. Zadatak se može zapisati kao: $p(X\geq 1) \geq 0.5$.

Međutim:

$$p(X\geq 1)=1-p(X=0)$$
$$=1-(5/6)^n.$$

Dakle, mora biti:
$1-(5/6)^n \geq 0.5$
odakle se dobije $n\geq 4$.

Primjer 2. Bacamo novčić 3 puta. Slučajna varijabla X registrira koliko se puta pojavio P (pismo). Zapišimo X.

X je binomna slučajna varijabla, odnosno X ima binomnu razdiobu s paramaetrima $n=3$ i $p= 1/2$ (dakle i $q=1/2$). Zato je $R(X) = \{0,1,2,3\}$. A vjerojatnosti su:

$$p(X=0)=(\frac{1}{2})^3 = \frac{1}{8};\ p(X=1)=\binom{3}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^2 = \frac{3}{8} = p(X=2)\ \text{ i }\ p(X=3)=\frac{1}{8}.$$

Razdiobu vjerojatnosti kraće zapisujemo kao

$p(X=0) = 1/8$
$p(X=1) = 3/8$
$p(X=2) = 3/8$
$p(X =3) = 1/8$

To se kraće zapisuje kao:

0	1	2	3
1/8	3/8	3/8	1/8

U gornjem su redu tablice vrijednosti koje slučajna varijabla postiže, a u donjem dijelu vjerojatnosti s kojima se te vrijednosti postižu. Treba uočiti da je zbroj brojeva u drugom redu jednak 1. To je zato što je ukupna vjerojatnost 1 (događaji [X=0], [X=1], [X=2],[X=3] međusobno se isključuju i zbroj im je sigurni događaj).
Kažemo da je tom tablicom zadana slučajna varijabla X.

Točnije, tom je tablicom zadana **razdioba vjerojatnosti** slučajne varijable X, tj. zadane su vjerojatnosti s kojima slučajna varijabla X postiže pojedine vrijednosti.

Općenito, s brojevima $p(X=i) = \binom{n}{i} p^i(1-p)^{n-i}$, $i=0,1,\ldots,n$ zadana je razdioba vjerojatnosti binomne slučajne varijable $X \sim B(n,p)$ pa se govori o binomnoj razdiobi.

Poissonova razdioba.

Razmotrimo sad tipični primjer II. Podsjetimo se:

II. U fiksiranom vremenskom intervalu u svakom se trenutku može, a ne mora, dogoditi događaj A (na primjer, u svakom trenutku na neku adresu može doći poruka). Slučajna varijabla X registrira broj pojavljivanja događaja A unutar tog vremenskog intervala (broj poruka na toj adresi).

Intuitivno je jasno da treba postojati parametar prema kojemu će se razlikovati različite adrese, odnosno razdioba vjerojatnosti na njima. Najjednostavniji takav parametar je prosječan broj poruka a na adresi (to će biti očekivanje slučajne varijable X).
Binomna razdioba primjer je diskretne razdiobe s konačno mnogo vrijednosti.
Poissonova razdioba također je diskretna, međutim prima beskonačno mnogo (prebrojivo) vrijednosti. Točnije, $R(X) = \{0,1,2,3,\ldots\}$.
U proširenim lekcijama ima izvod za vjerojatnost da bude točno odredjeni broj poruka.
Pokazuje se da u idealnim uvjetima vrijedi $p(X=i) = e^{-a}a^i/i!$. To je osnova za definiranje Poissonove razdiobe.

Definicija 2. Kažemo da je diskretna slučajna varijabla X distribuirana prema Poissonovu zakonu s parametrom $a>0$, ako je:
1. $R(X) = \{0,1,2,\ldots\}$
2. $p(X=i) = e^{-a}a^i/i!$, $i=0,1,2,3,\ldots$

Ako je tako pišemo $X \sim P(a)$.

Uočite da ta razdioba ovisi samo o jednom parametru (za razliku od binomne razdiobe koja ovisi o dvama parametrima).

Zadatak 1. Pokažite da je $p(X=0)+p(X=1)+p(X=2)+\ldots = 1$.

Uputa: Koristite se formulom za razvoj u red eksponencijalne funkcije.

Poissonova se razdioba pojavljuje na primjer kod slučajnih varijabla koje broje broj poziva u jedinici vremena na nekoj telefonskoj centrali, broj ulaza na neku adresu, broj kvarova na nekom složenom uređaju i sl. Primjenu Poissonove razdiobe pokazujemo na nekoliko primjera.

Primjer 3. Prosječan broj poziva u minuti na nekoj telefonskoj centrali je 8. Odredite vjerojatnost da na toj centrali u nekoj minuti bude:

a) najviše 8 poziva,
b) između 5 i 10 poziva,
c) barem 5 poziva

Prije rješavanja pokušajte procijeniti ove vjerojatnosti odoka.
Da bismo riješili zadatak, **moramo prihvatiti jedan dogovor**, a to je da se **broj poziva ponaša prema Poissonovu zakonu**. Ako je tako, onda je slučajna varijabla X koja registrira broj poziva u minuti na toj centrali, Poissonova s parametrom a=8 (naime, parametar *a* je očekivanje, a očekivanje odgovara prosječnom broju poziva, odnosno prosječnoj vrijednosti što je postiže ta slučajna varijabla). Sad se pitanja mogu formulirati ovako:

a) $p(X\leq 8)=p_0+p_1+\ldots+p_8=e^{-8}(1+8/1!+8^2/2!+\ldots+8^8/8!)=0.5925$ (na 4 decimalna mjesta)
c) $p(5<X<10) = p_6+p_7+p_8+p_9 = e^{-8}(8^6/6!+\ldots+8^9/9!)= 0.4168$
d) $p(X>4)=1-p(X\leq 4)=1-p_0-p_1-p_2-p_3-p_4=0.9004$.

Zadatak 2. Predpostavimo da netko dobije prosječno 3 poruke na sat.
(a) Koliko je puta vjerojatnije da tijekom sata dođe jedna poruka od toga da ne dođe ni jedna?
(b) Koji je najvjerojatnij broj poruka u jednom satu i koja mu je vjerojatnost.
(c) Kontroliranjem u 1000 sati dobiveni su rezultati o broju poruka. U koliko odprilike od tih 1000 sati neće biti ni jednog poziva, u koliko će biti točno jedan poziv itd.

Napomena.

1. Iako je vjerojatnosti Poissonove razdiobe lakše računati nego binomne i ovdje je često dobro koristiti rekurzivnu formulu koju je lako izvesti:
   $p_{i+1}/p_i = a/(i+1)$.
2. Može se pokazati da Poissonova razdioba nastaje kao limes binomnih razdioba kod kojih parametar *n* teži u beskonačnost, a umnožak *np* je stalan broj. Tada taj niz razdioba (tj. njihovih vrijednosti i pripadnih vjerojatnosti) teži upravo k razdiobi P(np). Drugim riječima, ako želimo tako dobiti razdiobu P(a), onda parametar *n*

redom uzima vrijednosti 2,3,4,…, a , parametar *p* redom uzima vrijednosti a/2, a/3, a/4,….
Zato se nekada Poissonova razdioba koristi za aproksimaciju binomne (jer za dosta veliki *n* i ne prevelik *np*, vjerojatnosti možemo računati kao kod Poissonove razdiobe).

Definicija diskretne slučajne varijable

Na osnovi razmatranih primjera uvodimo sljedeću definiciju.

Diskretna slučajna varijabla jest slučajna varijabla kojoj je skup vrijednosti konačan ili beskonačan prebrojiv.
Drugim riječima slučajna varijabla X je diskretna ako je
$R(X) = \{x_1,x_2,x_3,\ldots,x_n\}$ ili $R(X) = \{x_1,x_2,x_3,\ldots\}$

Događaji $[X=x_i]$ čine potpun skup događaja. To znači da se oni međusobno isključuju i da je suma njihovih vjerojatnosti 1:
$\sum p(X=x_i) = 1.$

Ako označimo:
$p_i = p(X=x_i)$,
onda se razdioba vjerojatnosti slučajne diskretne varijable X može zapisati kao:

x1	x2	x3	xn
p1	p2	p3	pn

ako je skup vrijednosti konačan, odnosno kao:

x1	x2	x3…..
p1	p2	p3 ….

ako je skup vrijednosti beskonačan prebrojiv.

Navedime primjere dviju diskretnih razdioba koje nisu ni binomne ni Poissonove.

Primjer 4. (Jednolika diskretna razdioba). To je razdioba koja prima konačan skup vrijednosti, a svaku vrijednost s jednakom vjerojatnošću. Opći oblik te razdiobe je:

x_1	x_2	………	x_n
1/n	1/n	………	1/n

Na primjer, u pokusu bacanja kocke jedan put, slučajna varijabla X koja registrira rezultat je jednoliko distribuirana. Njena je razdioba:

1 2 3 4 5 6

1/6 1/6 1/6 1/6 1/6 1/6

Primjer 5. Bacamo kocku dok se ne pojavi broj 6. Slučajna varijabla X registrira broj bacanja. Odredimo razdiobu vjerojatnosti od X.

To je pokus s beskonačno mnogo, ali prebrojivo ishoda. Slučajna varijabla X primjer je slučajne varijable koja poprima beskonačno mnogo (ali prebrojivo) vrijednosti:
$R(X) = \mathbf{N} = \{1,2,3,4, \dots\}$.
Izračunajmo nekoliko početnih vjerojatnosti:

p(X=1) = p(prvi put 6) = 1/6
p(X=2) = p(prvi put nije 6, drugi put 6) = 5/61/6
p(X=3) = p(prva dva puta nije 6, treći put 6) = $(5/6)^2 1/6$
p(X=4) = p(prva tri puta nije 6, četvrti put 6) = $(5/6)^3 1/6$.

Zato je razdioba vjerojatnosti slučajne varijable X:

1	2	3	4 ……………	n……….
1/6	5/61/6	$(5/6)^2 1/6$	$(5/6)^3 1/6$……	$(5/6)^{n-1} 1/6$…..

Koristeći se formulom za sumu beskonačnog reda možemo provjeriti da je zbroj vjerojatnosti zaista jednak 1.

$$1/6 + 5/61/6 + (5/6)^2 1/6 + (5/6)^3 1/6 + \dots = 1/6(1 + 5/6 + (5/6)^2 + (5/6)^3 + \dots)$$
$$= 1/6 \frac{1}{1-\frac{5}{6}}$$
$$= 1.$$

Očekivanje i varijanca diskretne slučajne varijable

Sad ćemo matematički precizno definirati očekivanje diskretne slučajne varijable prema uzoru na aritmetičku sredinu uzorka, i varijancu prema uzoru na varijancu uzorka. Razdioba slučajne varijable može se grafički predočiti točkama pravca (vrijednosti slučajne varijable) nad kojima su podignuti štapovi kojma su visine pripadne vjerojatnosti.
Grafički prikazi diskretnih slučajnih varijabla podsjećaju na **sustav materijalnih čestica na pravcu**.
Podsjetimo se:
sustav materijalnih čestica na pravcu čine čestice masa $m_1,m_2,\dots,m_n$ smještene u točkama pravca s koordinatama $x_1,x_2,\dots,x_n$.
To se kraće može zapisati kao $(x_1,m_1),(x_2,m_2),\dots(x_n,m_n)$.

Pri razmatranju sustava čestica uobičajeno je definirati **ukupnu masu**

$m := m_1+m_2+\ldots+m_n$
i **relativne mase**, brojeve:
$m_1/m, m_2/m,\ldots,m_n/m$.
Treba uočiti da je zbroj relativnih masa 1:

$$m_1/m+m_2/m+\ldots+m_n/m = (m_1+m_2+\ldots+m_n)/m = m/m = 1.$$

Zato je tablicom

x_1	x_2	x_n
m_1/m	m_2/m	m_n/m

zadana razdioba vjerojatnosti.
Obratno, svaka razdioba vjerojatnosti diskretne slučajne varijable zadaje neki sustav čestica na pravcu (kojemu je ukupna masa 1).
Pri toj korespondenciji, vjerojatnosti $p_i=p(X=x_i)$ korespondiraju relativnim masama mi/m.

Dvije temeljne karakteristike (značajke) sustava čestica jesu njegovo **težište** i **moment inercije oko težišta**. Sjetimo se koordinate težišta:

$$x_T =(x_1m_1+x_2m_2+\ldots+x_nm_n)/m = x_1m_1/m+x_2m_2/m+\ldots+x_nm_n/m.$$

Analogno težištu sustava čestica (odnosno aritmetičkoj sredini uzorka) definira se očekivanje slučajne varijable X kojoj je razdioba vjerojatnosti zadana tablicom

x_1	x_2	x_3	x_n
p_1	p_2	p_3	p_n

$$E(X) := x_1p_1+x_2p_2+\ldots+x_np_n$$

Ako diskretna slučajna varijabla postiže beskonačno mnogo, ali prebrojivo vrijednosti, očekivanje se definira analogno:

$$E(X) := x_1p_1+x_2p_2+x_3p_3+\ldots$$

To je **beskonačni red** i može se dogoditi da **divergira** (ili da **konvergira**, ali ne **konvergira apsolutno**). Tada slučajna varijabla **nema očekivanja** (odnosno pripadni sustav čestica nema težište).

Primjer 6. Pokažimo: ako je $X\sim P(a)$, onda je $E(X) = a$.

Uvrštavajući u formulu:
$E(X) = x_0p_0 + x_1p_1+\ldots+x_ip_i+\ldots$
$x_i=i$, $p_i = e^{-a}a^i/i!$, dobit ćemo:
$E(X) = e^{-a}(0\cdot a^0/0! + 1\cdot a^1/1! +2\cdot a^2/2!+\ldots+i\cdot a^i/i!+\ldots)$

$= e^{-a} \cdot a(1+a^1/1!+a^2/2!+\dots)$
$= e^{-a} \cdot a \cdot e^a$
$= a.$

Ako je X~P(a), onda je E(X) = a.

Primjer 7. Bacamo kocku dok se ne pojavi 6. Slučajna varijabla X registrira broj bacanja. Izračunajmo E(X).

Ta slučajna varijabla postiže beskonačno mnogo, ali prebrojivo vrijednosti; vidjeli smo da ima sljedeću razdiobu:

1	2	3	4	…………….n……….
1/6	5/61/6	$(5/6)^2 1/6$	$(5/6)^3 1/6$	……$(5/6)^{n-1} 1/6$…

Prije računanja pokušajmo predvidjeti očekivanje. Intuitivno, zbog ravnopravnosti brojeva 1,2,3,4,5,6 očekujemo da će se u 6 bacanja pojaviti jednom pojaviti 6 (slično je za ostale brojeve). Provjerimo to predviđanje računom. Koristit ćemo se formulom:

$1+2x+3x^2+4x^3+\dots= 1/(1-x)^2, \quad -1<x<1.$

Ta se formula dobije deriviranjem formule za zbroj geometrijskog reda:

$1+x+x^2+x^3+x^4+\dots = 1/1-x, \quad -1<x<1.$

$E(X) = 1\cdot 1/6+ 2\cdot 5/61/6+3\cdot(5/6)^2 1/6+ 4\cdot (5/6)^3 1/6+\dots+n\cdot (5/6)^{n-1} 1/6+\dots$
$=1/6(1+2\cdot 5/6+3\cdot(5/6)^2 +4\cdot(5/6)^3+\dots+n\cdot(5/6)^{n-1}+\dots)$
$=1/6\cdot 1/(1-5/6)^2$
$= 6$ (kako smo i predvidjeli).

Varijanca slučajne varijable.

Vratimo se na sustav čestica na pravcu. Sjetimo se momenta inercije oko težišta tog sustava.
$I_T = (x_1-x_T)^2 m_1+(x_2-x_T)^2 m_2+\dots+(x_n-x_T)^2 m_n.$
Prema uzoru na tu formulu (odnosno na varijancu uzorka), definiramo **disperziju** iliti **varijancu** V(X) diskretne slučajne varijable X (odnosno razdiobe vjerojatnosti slučajne vjerojatnosti):

$V(X) := (x_1-E(X))^2 p_1+(x_2-E(X))^2 p_2+\dots+(x_n-E(X))^2 p_n.$

Ako X postiže beskonačno mnogo (ali prebrojivo) vrijednosti, taj je red beskonačan i može biti divergentan (tada razdioba vjerojatnosti nema varijancu, odnosno pripadni sustav čestica nema moment inercije).
Treba uočiti da je varijanca V(X) pozitivna i da može biti 0 ako i samo ako postiže samo jednu vrijednost (s vjerojatnošću 1).

Za računanje varijance katkad je jednostavnija formula:

$V(X) = (x_1^2p_1+x_2^2p_2+\ldots+x_n^2p_n) - E^2(X)$

(slično je ako je R(X) beskonačan, prebrojiv skup). Tu je $E^2(X)$ kraća oznaka za $((E(X))^2$.

Zadatak 3. Izvedite gornju formulu.

Primjer 8. Bacamo novčić 2 puta. Slučajna varijabla X registrira koliko puta se pojavio P. Izračunajmo varijancu od X.

Vidjeli smo da X ima sljedeću razdiobu vjerojatnosti:

0	1	2	3
1/8	3/8	3/8	1/8

i da je E(X) = 1.5.

$V(X) = (0-1.5)^2 1/8+(1-1.5)^2 3/8+(2-1.5)^2 3/8+(3-1.5)^2 1/8$
$= 2.25\cdot 1/8+0.25\cdot 3/8+0.25\cdot 3/8+2.25\cdot 1/8$
$= 0.75$

To ćemo izračunati i pomoću druge formule.

$V(X) = 0^2\cdot 1/8+1^2\cdot 3/8+2^2\cdot 3/8+3^2\cdot 1/8 - 1.5^2$
$= 3-2.25=0.75.$

Moment inercije je **mjera disperzije masa u odnosu na težište**; što je masa koncentriranija uz težište, moment inercije je manji i obratno. Analogno tome, varijanca slučajne varijable je mjera disperzije (raspršenosti) vjerojatnosti u odnosu na očekivanje; što je vjerojatnost koncentriranija uz očekivanje, varijanca je manja i obratno.

Primjer 9. Izračunajmo očekivanje i varijancu diskretne jednolike razdiobe.

Ako zadržimo oznake kao do sada, dobit ćemo:

$E(X)= (x_1+x_2\ldots+x_n)/n,$

$V(X)=(x_1^2+x_2^2+\ldots+x_n^2)/n-((x_1+x_2\ldots+x_n)/n)^2$

Očekivanje i varijanca binomne i Poissonove razdiobe.

Očekivanje smo već računali. Može se pokazati da vrijedi.

Ako je X~B(n,p), onda je E(X) = np i V(X)=npq.

Ako je X~P(a), onda je E(X) = a i V(X)=a.