OSNOVNI POJMI IZ VERJETNOSTI IN STATISTIKE

Verjetnost in verjetnostni račun

Osnovni pojmi
Poskus je vsako dejanje, ki poteka v skaldu z množico pogojev (je realizaciaja kompleksa pogojev).

Dogodek je vsak pojav, ki se pri danem poskusu lahko zgodi in ne spada v množico pogojev. Dogodke označujemo z velikimi tiskanimi črkami: A, B, C...

Dogodek je *gotov*, če se zgodi pri vsaki ponovitvi poskusa. Oznaka: G
Verjetnost gotovega dogodka je 1. $P(G) = 1$

Dogodek je *negotov*, če se ne zgodi pri nobeni ponovitvi poskusa. Oznaka: N.
Verjetnost negotovega dogodka je 0. $P(N) = 0$

Dogodek je *slučajen*, če se pri nekaterih ponovitvah poskusa zgodi pri drugih pa ne.

Definiciji verjetnosti
Ločimo dve definiciji:

a) Statistična definicija verjetnosti:
Pri tem poskus ponovimo n-krat, od tega se k-krat zgodi dogodek A. Če se pri večanju n, relativna frekvenca $f = \frac{k}{n}$ približuje k neki vrednosti, lahko to smatramo kot verjetnost dogodka A. Zapišemo: $P(A) = f$.

Zgled:
Met kovanca: A ... pade cifra

n	**A**	**f**
10	6	$6/10$
20	11	$11/20$
100	51	$51/100$

b) Klasična definicija verjetnosti:
Če pri poskusu nastopi lahko n enakovrednih situacij in je za dogodek A ugodnih r situacij, vzamemo razmerje $\frac{r}{n}$ za verjetnost dogodka A. Verjenost označimo s P(A).

Zgled:
Met kocke A ... vžemo sodo pik:
$n = 6$
$r = 3$

$$P(A) = \frac{3}{6} = \frac{1}{2}$$

Lastnosti verjetnosti in operacije
Vsota dogodkov $A+B$ je dogodek, da se od dogodkov A in B zgodi vsaj eden. Označujemo tudi $A \cup B$.
Produkt dogodkov $A \cdot B$ je dogodek, da se zgodita dogodka A in B hkrati. Označujemo tudi $A \cap B$.

Zgled:
Poskus: met kocke
A ... pade vsaj 5
B ... pade sodo pik
$A+B=\{2,4,5,6\}$
$A \cdot B=\{6\}$

Za vsoto in produkt veljajo naslednje lastnosti:
- $A+B=B+A$
- $A \cdot B=B \cdot A$
- $(A+B)+C=A+(B+C)$
- $A \cdot (B \cdot C)=(A \cdot B) \cdot C$
- $(A+B) \cdot C=A \cdot C+B \cdot C$

1. $A+A=A$

- $A \cdot A=A$

Negacija $\overline{A}$ dogodka A; velja:
$A+\overline{A}=G$
$A \cdot \overline{A}=N$
Verjetnost negacije je enaka: $P\left(\overline{A}\right)=1-P(A)$

Verjetost pri sestavljenih dogodkih
Zanima nas verjetnost pri vsoti dveh dogodkov $A+B$. Če naredimo nek poskus n-krat, pri čemer se dogodek A zgodi k_1-krat, dogodek B zgodi k_2-krat in dogodek AB zgodi k_3-krat, lahko zapišemo relativne frekvence:
$$f(A)=\frac{k_1}{n}$$
$$f(B)=\frac{k_2}{n}$$
$$f(AB)=\frac{k_3}{n}$$
Če pogledamo relativno frekvenco za vsoto, ji moramo odšteti del, ki večkrat ponovi: $f(A+B)=\frac{k_1+k_2-k_3}{n}=\frac{k_1}{n}+\frac{k_2}{n}-\frac{k_3}{n}=f(A)+f(B)-f(AB)$.
Iz tega sledi za verjetnost vsote naslednji izraz: $P(A+B)=P(A)+P(B)-P(AB)$.

Zgled:
Izbiramo naravna števila:
A ... število deljivo s 2
B ... deljivo s 3
A + B ... deljivo vsaj z enim:

$$P(A+B)=P(A)+P(B)-P(AB)=\frac{1}{2}+\frac{1}{3}-\frac{1}{6}=\frac{2}{3}$$

Če sta dogodka A in B nezdružljiva, velja da je verjetnost za vsoto $P(A+B)=P(A)+P(B)$, saj je verjetnost produkta nič, ker sta je rezultat negotov dogodek.

Zgled:
Poskus: met kocke
A ... pade sodo pik
B ... pade 3

$$P(A+B)=P(A)+P(B)=\frac{3}{6}+\frac{1}{6}=\frac{2}{3}$$

Podobno lahko izračunamo verjetnost za vsoto treh dogodkov:

$$P(A+B+C)=P(A+B)+P(C)-P((A+B)\cdot C)=$$
$$=P(A)+P(B)-P(A\cdot B)+P(C)-P((A\cdot C+B\cdot C)=$$
$$=P(A)+P(B)-P(A\cdot B)+P(C)-\left[P(A\cdot C)+P(B\cdot C)-P(A\cdot C\cdot B\cdot C)\right]=$$
$$=P(A)+P(B)+P(C)-P(A\cdot B)-P(A\cdot C)-P(B\cdot C)+P(A\cdot B\cdot C)$$

$$\underline{\underline{P(A+B+C)=P(A)+P(B)+P(C)-P(A\cdot B)-P(A\cdot C)-P(B\cdot C)+P(A\cdot B\cdot C)}}$$

Pogojna verjetnost
Pri pogojni verjetnosti nas zanima verjetnost dogodka A, pri pogoju, da se je dogodek B zgodil. Označimo s $P(A/B)$. Pri tem predpostavimo, da je verjetnost D mogoč dogodek, to pomeni $P(B)>0$.

Če naredimo nek poskus n-krat, pri čemer se dogodek A zgodi k_1-krat, dogodek B zgodi k_2-krat in dogodek AB zgodi k_3-krat, lahko zapišemo relativne frekvence:

$$f(A)=\frac{k_1}{n}$$
$$f(B)=\frac{k_2}{n}$$
$$f(AB)=\frac{k_3}{n}$$

Če pogledamo relativno frekvenco za frekvenco dogodka A pri pogoju B, dobimo:

$$f(A/B)=\frac{k_3}{k_2}=\frac{\frac{k_3}{n}}{\frac{k_2}{n}}=\frac{f(AB)}{f(B)}.$$

Če gledamo relativne frekvenc v limiti, ko se približuje določenemu številu, dobimo izraz za pogojno verjetnost: $P(A/B)=\frac{P(AB)}{P(B)}$. Velja tudi obratno: $P(B/A)=\frac{P(AB)}{P(A)}$.

Zgled:
Tovarna izdeluje izdelke naslednje kvalitete: 40 % izdelkov je I. kvalitete, 50 % izdelkov je II. kvalitete, 10 % pa je škarta. Zanima nas verjetnost izbire izdelka I. kvalitete, pri pogoju, da je uporaben.
A ... I.kvaliteta
U ... da je uporaben
$$P(A/U)=\frac{P(A\cdot U)}{P(U)}=\frac{0,40}{0,90}=\frac{4}{9}$$

Če velja, da je pogojna verjetnost enaka verjetnosti dogodka $P(A/B)=P(A)$, sta dogodka A in B neodvisna. Iz zveze za pogojno verjetnost potem sledi:
$$\begin{aligned}P(A\cdot B)&=P(B)\cdot P(A/B)\\&=P(A)\cdot P(B/A)\end{aligned}.$$

Zgled:
Imamo posodo z 6 belimi in 9 črnimi kroglami. Kolikšna je verjetnost, da bomo v dveh zaporednih poskusih potegnili beli kroglo, če pri tem krogel ne vračmo.
A ... prvič belo
B ... drugič belo
$$P(A\cdot B)=P(A)\cdot P(B/A)=\frac{6}{15}\cdot\frac{5}{14}=\frac{1}{7}$$

Če sta dogodka A in B neodvisna, velja $P(A/B)=P(A)$, iz tega sledi, da $P(A\cdot B)=P(A)\cdot P(B)$. Velja pa tudi obratno. Namreč, č eta zveza velja, sledi:
$$\left.\begin{aligned}P(A\cdot B)=P(B)\cdot P(A/B)\\P(A)\cdot P(B)=P(B)\cdot P(A/B)\end{aligned}\right\}\Rightarrow P(A)=P(A/B)$$

Za tri dogodke A, B, C velja: $P(A\cdot B\cdot C)=P(A)\cdot P(B/A)\cdot P(C/A\cdot B)$. To pomeni, da če se zgodi A, se lahko B zgodi pri pogoju, da se je prej zgodil A, dogodek C pa se bo zgodil, če sta se prej zgodila A in B.

Relejni poskusi
Naj poskus poteka v večih stopnjah. Na prvi stopnji se lahko zgodi natanko eden od dogodkov $H_1,H_2,...,H_n$, ki so med seboj nezdružljivi. To pomeni, da velja: $H_1+H_2+...+H_n=G; H_i\cdot H_j=N; i\neq j$. Na drugi stopnji se lahko zgodi dogodek A.
$$\begin{aligned}&P(A)=P(AG)=P(A)\cdot\left(H_1+H_2+...+H_n\right)=P\left(A\cdot H_1+A\cdot H_2+...+A\cdot H_n\right)=\\&=P\left(A\cdot H_1\right)+P\left(A\cdot H_2\right)+...+P\left(A\cdot H_n\right)=\\&=P\left(H_1\right)P\left(A/H_1\right)+P\left(H_2\right)P\left(A/H_2\right)+...+P\left(H_n\right)P\left(A/H_n\right)\end{aligned}$$

Krajši zapis: $P(A)=\sum_{i=1}^{n}P\left(H_i\right)P\left(A/H_i\right)$. Pri tem dobimo obrazec za popolno verjetnost.

Lahko pa iščemo tudi obratno: $P(A/H_k)=P(A)\cdot P(H_k/A)\overset{\text{ali}}{=}P(H_k)\cdot P(A/H_k)$. Iz tega sledi: $P(H_k/A)=\dfrac{P(H_k)P(A/H_k)}{P(A)}$. To je Bayesov obrazec.

Zgled:

1. Imamo tri posode s belimi in črnimi kroglicami. Nato vržemo kocko, pri čemer vzamemo kroglo iz prve posode, če pade 1, 2, 3 pike, iz druge, če pade 4 ali 5 pik, če pa pade 6 pik, potem vzamemo kroglo iz tretje posede. Kolikšna je verjetnost, da bomo izbrali belo kroglo?

A ... bela krogla na koncu
H_1 ... v I. posodo
H_2 ... v II. posodo
H_3 ... v III. Posodo

$$P(H_1)=\frac{3}{6};\ P(H_2)=\frac{2}{6};\ P(H_3)=\frac{1}{6}$$

$$P(A)=P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3)=$$

$$=\frac{3}{6}\cdot\frac{2}{6}+\frac{2}{6}\cdot\frac{3}{12}+\frac{1}{6}\cdot\frac{2}{8}=\frac{7}{24}$$

Kolikšna pa je verjetnost, je bela krogla iz tretje posode?

$$P(H_3/A)=\frac{P(H_3)P(A/H_3)}{P(A)}=\frac{\frac{1}{6}\cdot\frac{2}{8}}{\frac{7}{24}}=\frac{1}{7}$$

2. Nek izdelek izdelujejo tri tovarne v razmerju:

Tovarna A	90%	I. kvalitete
	10%	II. kvalitete
Tovarna B	80%	I. kvalitete
	20%	II. kvalitete
Tovarna C	70%	I. kvalitete
	30%	II. kvalitete

Kupimo 1 izdelek I. kvalitete. Kolikšna je verjetnost, da je izdelek iz tovarne A.
A ... da je izdelek I. kvalitete

$$P(H_1)=\frac{1}{3};\ P(H_2)=\frac{1}{3};\ P(H_3)=\frac{1}{3}$$

$$P(A/H_1)=0{,}90$$

$$P(A/H_2)=0{,}80$$

$$P(A/H_3)=0{,}70$$

$$P(H_1/A)=\frac{P(H_1)P(A/H_1)}{P(A)}=\frac{\frac{1}{3}\cdot\frac{9}{10}}{\frac{1}{3}\cdot\frac{9}{10}+\frac{1}{3}\cdot\frac{8}{10}+\frac{1}{3}\cdot\frac{7}{10}}=\frac{3}{8}$$

Zaporedje neodvisnih poskusov (Bernoulijevo zaporedje)
Poskus velikokrat ponavljamo in vsakič gledamo ali se dogodek A zgodi ali ne pri danem poskusu. Pri tem ni vpliva nekega koraka na njegovega naslednika.

Ponovimo poskus n-krat, verjetnost za dogodek A naj bo $P(A)=p$ in negacija dogodka A naj bo $P\left(\overline{A}\right)=q=1-p$. Sedaj pa nas zanima, kolikšna je verjetnost, da se je od n neodvisnih poskusov dogodek A zgodil k- krat, kjer velja $0 \le k \le n$. Izračunamo jo lahko po formuli: $P_n(k)=\binom{n}{k}\cdot p^k \cdot q^{n-k}$, kar preberemo verjetnost, da se od n ponovitev dogodek A zgodi k-krat.
Pri tem ločimo dve možnosti:

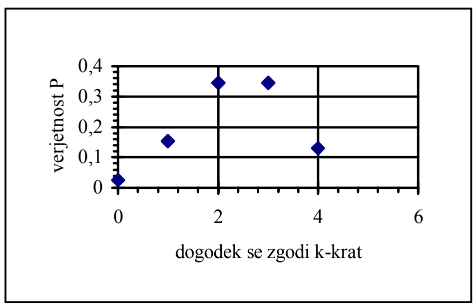
a) Če je $k_0 = np - q$ celo število, je največja verjetnost dosežena pri k_0 in k_0+1.
b) Če $k_0 = np - q$ ni celo število, označimo s k_0 celi del tega števila, kar označimo s $k_0 = [np-q]$. Potem je največja verjetnost dosežena pri številu k_0+1, ki je enak: $k_0+1=[np-q]+1$.

Zgleda:

a) Primer, ko je celo število

$$\left.\begin{array}{l} n=4 \\ p=0,6 \end{array}\right\} \Rightarrow P_n(k)=\binom{4}{k}\cdot(0,6)^k\cdot(0,4)^{4-k} \quad np-q=4\cdot\frac{6}{10}-\frac{4}{10}=2$$

k	$P_n(k)$
0	0,0256
1	0,1536
2	**0,3456**
3	**0,3456**
4	0,1296

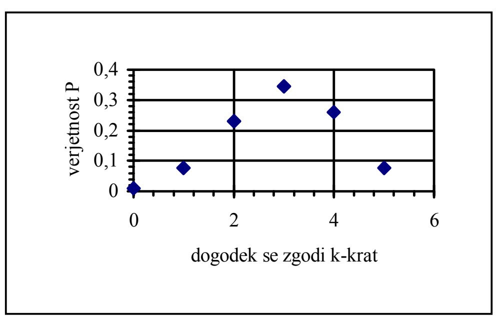


b) Primer, ko ni celo število

$$\left.\begin{array}{l} n=5 \\ p=0,6 \end{array}\right\} \Rightarrow P_n(k)=\binom{5}{k}\cdot(0,6)^k\cdot(0,4)^{5-k} \quad np-q=5\cdot\frac{6}{10}-\frac{4}{10}=2,6$$

$$[np-q]=2 \Rightarrow k_0+1=2+1=3$$

k	$P_n(k)$
0	0,0102
1	0,0768
2	0,2304
3	**0,3456**
4	0,2592
5	0,0777



Veliko število ponovitev n

V primerih ko pa imamo zelo veliko ponovitev poskusa, to pomeni velik n, pa vzamemo približke, saj je računanje binomskih simbolov in rokovanje z velikimi potencami bolj zamudno.

Pri tem aproksimiramo s pomočjo dveh aproksimacij, in sicer:

a) La-Placeova lokalna aproksimacija

La-Placeovo lokalno aproksimacijo dobimo, tako da si pomagamo s standarno Gaussovo krivuljo. Ta ima obliko $\varphi(x)=\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{1}{2}x^2}$, medtem ko je oblika splošne Gaussove krivulje enaka: $\varphi(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}$. Verjetnost k elementov izmad n elementov lahko potem zapišemo kot: $P_n(k)\doteq\frac{1}{\sqrt{2\pi npq}}\cdot e^{-\frac{(k-np)^2}{2npq}}$. Če uvedemo še $x_k=\frac{k-np}{\sqrt{npq}}$, lahko izraz za verjetnost zapišemo kot: $P_n(k)=\frac{1}{\sqrt{npq}}\varphi(x_k)$, kjer so vrednosti $\varphi(x)$ tabelirane. Ta približek je dober za velike n in za primer, ko p ni preblizu 0 ali 1. Če pa je blizu 0 ali 1, pa uporabimo Poissonovo aproksimacijo.

Primer:

Verjetnost, da se izdelek pokvari, je 0,05. kolikšna je verjetnost, da se od 1000 izdelkov pokvari 50.

$$\left.\begin{aligned}n&=1000\\k&=50\\p&=0,05\end{aligned}\right\}\Rightarrow P_{1000}(50)=\binom{1000}{50}\cdot(0,05)^{50}\cdot(0,95)^{50}\doteq\frac{1}{6,89}\varphi(0)=\frac{0,3989}{6,89}=0,058$$

$$k-np=50-\frac{1000}{100}5=0$$

$$\sqrt{npq}=\sqrt{1000\cdot\frac{5}{100}\cdot\frac{95}{100}}=\sqrt{47,5}=6,89$$

b) Poissonova lokalna aproksimacija

Ločimo dve možnosti:

- če je p blizu 0, vzamemo približek: $P_n(k)=\frac{(np)^k\cdot e^{-np}}{k!}$
- če je p blizu 1, vzamemo približek: $P_n(k)=\frac{(nq)^{n-k}\cdot e^{-nq}}{(n-k)!}$.

Primer:

Verjetnost, da pride do okvare dela je 0,02. Naprava je sestavljena iz 50 delov. Kolikšna je verjetnost, da se okvarita dva dela.

$$n=50;\ k=2;\ \ p=0,02\Rightarrow P_{50}(2)=\binom{50}{2}\cdot(0,02)^2\cdot(0,98)^{48}=0,1859$$

s Poissonovim približkom: $\doteq\frac{1^2e^{-1}}{2!}=\frac{1}{2e}=0,1839$

Po La Placeu $\doteq 0,237$

Približek za verjetnost Bernoullijevega zaporedja

Približek za verjetnost Bernoullijevega zaporedja, kjer velja $a \le k \le b$, lahko izračunamo po:

- točni formuli: $P(a \le k \le b) = \sum_{k=a}^{b} \binom{n}{k} \cdot p^k \cdot q^{n-k}$;
- po približku: $P(a \le k \le b) \doteq \frac{1}{\sqrt{2\pi npq}} \cdot \int_a^b e^{-\frac{(x-np)^2}{2npq}} dx$.

Zadnjo enačbo pa lahko poenostavimo z uvedbo nove spremenljivke t, kjer dobimo naslednje:

$P(a \le k \le b) \doteq \frac{1}{\sqrt{2\pi}} \cdot \int_\alpha^\beta e^{-\frac{1}{2}t^2} dt = \Phi(\beta) - \Phi(\alpha)$, pri čemer sta $\alpha = \frac{a-np}{\sqrt{npq}}$ in $\beta = \frac{b-np}{\sqrt{npq}}$.

Vrednosti funkcije $\Phi(x)$ so tabelirane, veljata pa naslednji lastnosti:

- $\Phi(\infty) = \frac{1}{2}$
- $\Phi(-x) = -\Phi(x)$.

Primer:

Verjetnost, da dobimo pri križanju pšenice AB spet A, je 0,25. kolikšna je verjetnost, da bomo pri 4800 križanjih dobili sorto A manj kot 1230.

Točen izračun: $P(0 \le k \le 1230) = \sum_{k=0}^{1230} \binom{4800}{k} \cdot (0,25)^k \cdot (0,75)^{4800-k}$

Približek: $P(a \le k \le b) \doteq \Phi\left(\frac{b-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-np}{\sqrt{npq}}\right) = \Phi\left(\frac{30}{30}\right) - \Phi\left(-\frac{1200}{30}\right) = \Phi(1) + \Phi(40) =$.

$= 0,3143 + 0,5 = 0,8413$

V primeru, ko je p blizu 0 ali 1 zopet uporabimo za izračun verjetnosti vsoto približkov Poissonove aproksimacije:

1) če je p blizu 0, vzamemo približek: $P(a \le k \le b) = \sum_{k=a}^{b} \frac{(np)^k \cdot e^{-np}}{k!}$

- če je p blizu 1, vzamemo približek: $P(a \le k \le b) = \sum_{k=a}^{b} \frac{(nq)^{n-k} \cdot e^{-nq}}{(n-k)!}$.

Sedaj pa nas zanima se relativna frekvenca dogodka A, če imamo podano veliko število poskusov n in verjetnost dogodka A označeno z p. Relativna frekvenca je enaka: $f(A) = \frac{k}{n}$.

$$P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = P\left(-\varepsilon \le \frac{k}{n} - p \le \varepsilon\right) = P\left(\underbrace{+np - \varepsilon n}_{a} \le k \le \underbrace{n\varepsilon + np}_{b}\right) \doteq$$

$$\doteq \Phi\left(\frac{n\varepsilon \cancel{+np} \cancel{-np}}{\sqrt{npq}}\right) - \Phi\left(\frac{-\varepsilon n \cancel{+np} \cancel{-np}}{\sqrt{npq}}\right) = 2\Phi\left(\frac{\varepsilon n}{\sqrt{npq}}\right)$$

Pri velikih n velja za razliko prave verjetnosti posameznega dogodka p in relativno frekvenco naslednji izraz: $P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) \doteq 2\Phi\left(\frac{\varepsilon n}{\sqrt{npq}}\right)$.

Slučajne spremenljivke

Slučanja spremenljivka je takšna spremenljivka, ki zavzema določeno vrednost, odvisno od slučaja. Označujemo jih z velikimi tiskanimi črkami: X, Y, ...

Pri slučajnih spreenljivkah je pomembna zaloga vrednosti verjetnosti. Ločimo dve vrsti spremenljivk, od tega pa je odvisna tudi zaloga vrednosti:

a) pri diskretnih spremenljivkah je zaloga vrednosti posamezne točke
b) pri nediskretnih spremenljivkah pa je zaloga vrednosti zvezna funkcija.

Zapis slučajnih spremenljivk:

1. Diskretne slučanje spremenljivke:

   Podajamo jih z verjetnostno shemo:

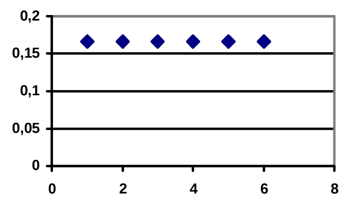
   $X:\begin{pmatrix} x_1 & x_2 & \ldots & x_n & \ldots \\ p_1 & p_2 & \ldots & p_n & \ldots \end{pmatrix}$, kjer je $P(x = x_1) = p_1$ oziroma na splošno: $P(x = x_k) = p_k$, kjer je k 0, 1, 2, ... pri tem pa velja, da je vsota vseh verjetnosti enaka 1: $\sum_{k=1}^{n} p_k = 1$. Lahko pa jih podamo tudi grafično.

   Zgled:

   a. met kocke

      X .. število pik

      $X:\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{pmatrix}$



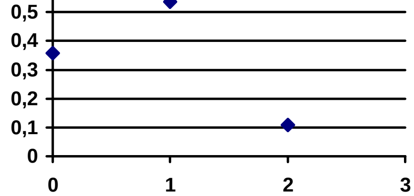
   b. met kovanca

      $X = \begin{cases} 1, \text{ če pade cifra} \\ 2, \text{ če pade glava} \end{cases} \quad X:\begin{pmatrix} 1 & 2 \\ 0,5 & 0,5 \end{pmatrix}$

   c. posoda s 5 belimi in 3 črnimi kroglami, potegnemo 2

      X ... število belih

      $X:\begin{pmatrix} 0 & 1 & 2 \\ \frac{5}{14} & \frac{30}{56} & \frac{3}{28} \end{pmatrix}$



Zapis Bernoullijevo porazdeljene slučajne spremenljivke

Splošno: $P(X=k)=\binom{n}{k}p^k q^{n-k}; k=0,1,2,3...$

Zapis Poissonovo porazdeljene slučajne spremenljivke

Splošno

$X:\begin{pmatrix} 0 & 1 & ... & k & ... \\ p_1 & p_2 & ... & p_k & ... \end{pmatrix}$, kjer je $p_k=\dfrac{\lambda^k \cdot e^{-\lambda}}{k!}; \lambda>0, k=0,1,2,3...$

Dokaz, da je vsota vseh verjetnosti enaka 1:

$$\sum p_k=\sum \frac{\lambda^k \cdot e^{-\lambda}}{k!}=e^{-\lambda}\left(1+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+...\right)=e^{-\lambda}\cdot e^{\lambda}=1$$

2. Nediskretne slučanje spremenljivke
   Tu podajamo verjetnosti s tako znano porazdelitveno funkcijo: $F(x)=P(X<x)$; za vsak x. Ta zapis se da uporabiti tudi za diskretne spremenljivke.

   Lastnosti porazdelitvene funkcije:
   - Če je $x_1<x_2$, potem $F(x_1)\le F(x_2)$
   - $F(-\infty)=0$, lahko se zgodi tudi že prej
   - $F(\infty)=1$, lahko se zgodi tudi že prej
   - $P\left(x_1\le x<x_2\right)=F(x_2)-F(x_1)$

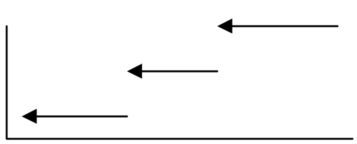
   - $P(X=x_1)=F(x_1+0)-F(x_1)$, prvi člen je pri zvzni funkciji enak nič, ni pa nujno, da je pri nezvezni funkciji.

Absolutno zvezne slučajne spremenljivke

Slučajna spremenljivka X ima tako porazdelitveno funkcijo, če obstaja funkcija p(t), da velja: $F(x)=\int_{-\infty}^{x} p(t)dt;\ x\in\mathbb{R}$. Funkcija p(t) se imenuje gostota verjetnosti.

Lastnosti:
- $F'(X)=p(x)$
- $\int_{-\infty}^{\infty} p(t)dt=1$
- $P(X=x_1)=0$
- F(X) je kot funkcija zvezna.

Pri diskretnik spremenljivkah, ne moremo računati porazdelitvene funkcije s integralom, ker le-ta ni zvezna.



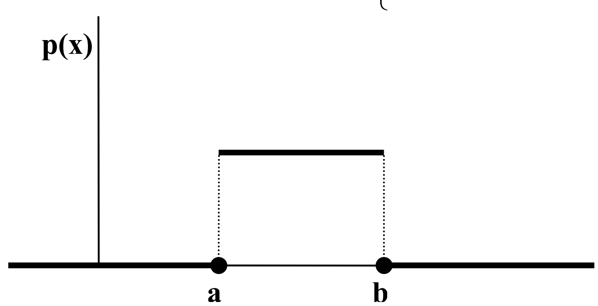
Interval, na katerem je gostota verjetnosti $p(t)$ zvezna, ga omejuje: $P(x_1 \le X \le x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} p(t)dt = p(\xi)(x_2 - x_1)$, kjer je interval $[x_1, x_2]$ majhen in kjer ξ leži znotraj intervala. Vrednost $p(\xi)$ se imenuje gostota verjetnosti, saj pove verjetnost na dolžino intervala: $p(\xi) = \dfrac{P(x_1 \le X \le x_2)}{x_2 - x_1}$.

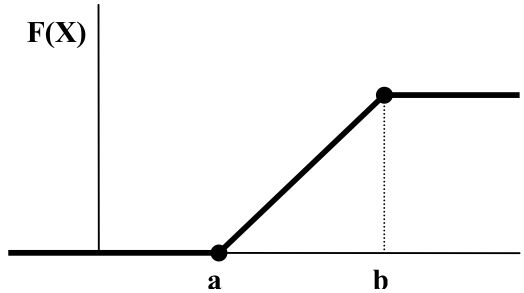
Primeri zvezno porazdeljenih slučajnih spremenljivk:

1) Enakomerna porazdelitev

Gostota verjetnosti je: $p(X) = \begin{cases} 0, & x < a, x > b \\ \dfrac{1}{b-a}; & a \le x \le b \end{cases}$

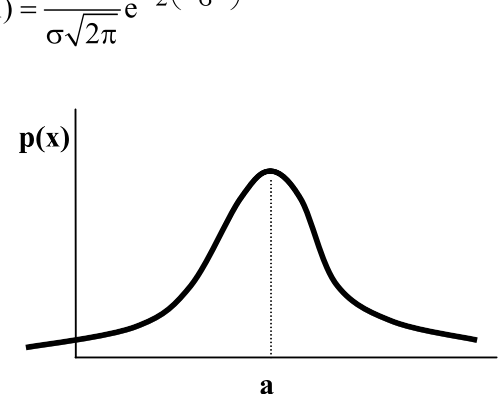
Porazdelitvena funkcija: $F(X) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \ge b \end{cases}$





2) Normalna ali Gaussova porazdelitev

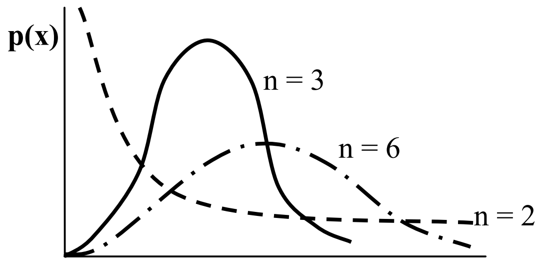
$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}$$



Velikokrat jo zapišemo kot $N(a, \sigma)$.

3) Porazdelitev hi kvadrat $\chi^2(n)$

$$p(x)=\begin{cases} A_n\cdot x^{\frac{n}{2}-1}\cdot e^{-\frac{x}{2}} & x>0; \\ 0 & \text{sicer}\end{cases} \qquad A_n=\frac{1}{2^{\frac{n}{2}}\cdot\Gamma\left(\frac{n}{2}\right)}$$



4) Studentova porazdelitev

$$p(x)=B_n\cdot\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}} \quad ;B_n=\frac{1}{\sqrt{n}\cdot B\left(\frac{n}{2},\frac{1}{2}\right)}$$

Slučajne n-terice ali slučajni vektorji

To so sheme slučajnih spremenljivk. Podamo jih lahko v tabelah. Tako lahko za n = 2, to pmeni, da imamo slučajno dvojico, zapišemo tako tabelo:

$X \setminus Y$	y_1	y_2	$\cdots$	y_m
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}

Če je $P(X=x_i;Y=y_i)=P(X=x_i)\cdot P(Y=y_i)$ za vse možnosti, rečemo, da sta slučajni spremenljivki X in Y neodvisni.

Matematično upanje slučajne spremenljivke

1. Če je X diskretna slučajna spremenljivka

   V tem primeru izračunamo matematično upanje kot vsoto produktov $E(X)=\sum_{i=1}^{n}x_ip_i$.

   Verjetnost lahko zamenjamo tudi z relativnimi frekvencami.

   Matematično upanje je mera za centralno tendenco, okrog katere se spremenljivka najbolj verjetno nahaja. Pri meritvah je to neko povprečje.

   Zgled:
   Poissonova porazdelitev:

$$p(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$

$$E(X)=\sum_{k=0}^{\infty}k\frac{\lambda^k e^{-\lambda}}{k!}=e^{-\lambda}\cdot\sum_{k=0}^{\infty}k\frac{\lambda\cdot\lambda^{k-1}}{(k-1)!}=e^{-\lambda}\cdot\lambda\left(\frac{1}{1}+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+...\right)=e^{-\lambda}\cdot\lambda\cdot e^{\lambda}=\lambda$$

2. Če je X absolutno zvezna slučajno spremenljivka
   Za absolutno zvezne slučajne spremenljivke lahko matematično upanje izračunamo kot $E(X) = \int_{-\infty}^{\infty} x \cdot p(x)dx$.
   Primeri:
   a) Enakomerna porazdelitev slučajne spremenljivke X
   $$E(X) = \int_a^b x \cdot \frac{1}{b-a}dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{a+b}{2}$$
   b) Porazdelitev z gostoto: $p(X) = \begin{cases} 0, x < 0 \\ e^{-x}; x \geq 0 \end{cases}$
   $$E(X) = \int_0^{\infty} x \cdot e^{-x}dx = \Gamma(2) = 1$$
   c) Normalna porazdelitev $N(a,\sigma)$
   $$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}$$
   $$E(X) = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} dx = a$$

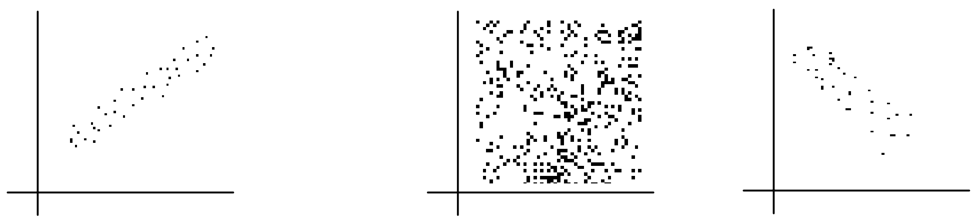
Disperzija
Disprezija je mera za razpršenost slučajne spremenljivke oziroma pri tem gledamo razpršenost spremenljivk okoli matematičnega upanja.
1. Za diskretno spremenljivko
   $$D(X) = \sum_k \left(x_k - E(X)\right)^2 pk$$
2. Za zvezno spremenljivko
   $$D(X) = \int_{-\infty}^{\infty} \left(x_k - E(X)\right)^2 p(x)dx$$

Povezava med matematičnim upanjem in disperzijo je: $D(X) = E(X^2) - E(X)^2$.
Dokaz:
$$D(X) = \int_{-\infty}^{\infty} \left(x^2 - 2xE(X) + E(X)^2\right)p(x)dx =$$
$$= \underbrace{\int_{-\infty}^{\infty} \left(x^2\right)p(x)dx}_{=E(X^2)} - \left(2E(X)\right)\underbrace{\int_{-\infty}^{\infty} xp(x)dx}_{=E(X)} + \left(E(X)^2\right)\underbrace{\int_{-\infty}^{\infty} p(x)dx}_{=1} =$$
$$= E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2$$

Disperzijo pa lahko povežemo s standardno deviacijo ali odklonom: $\sigma(X) = \sqrt{D(X)}$.

Nekaj lastnosti matematičnega upanja in disperzije
1) $E(X+Y) = E(X) + E(Y)$
2) $E(cX) = cE(X); \quad c = const$
3) $E(c) = c$

4) Če sta spremenljivki X in Y neodvisni, velja: $E(XY) = E(X) \cdot E(Y)$. Za odvisne ne velja. Prav tako ne velja obratno, čeprav se lahko zgodi, da je to res, ampak sta spremenljivki X in Y odvisni.
Kadar velja, da je $E(XY) = E(X) \cdot E(Y)$, rečemo, da sta spremenljivki X in Y nekorelirani. To pomeni da iz neodvisnosti sledi nekoreliranost: $\text{neodvisnost} \Rightarrow \text{nekoreliranost}$. Korelacija je razlika leve od desne strani: $K(X,Y) = E(XY) - E(X) \cdot E(Y)$. Če je $K(X,Y) = 0$, potem sta X,Y nekorelirani.
Definiramo lahko tudi korelacijki koeficient, ki se giblje na intervalu od -1 do 1: $r(XY) = \dfrac{K(XY)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} = \dfrac{K(XY)}{\sigma(X) \cdot \sigma(Y)}$.



r = 0,9 | r = 0 | r = -0,8

5) $D(cX) = c^2 \cdot D(X)$

6) $D(c) = 0$

7) $D(X+Y) = D(X) + D(Y) + 2K(XY)$

Dokaz:

$$D(X+Y) = E\left((X+Y)^2\right) - \left(E(X+Y)\right)^2 = E\left(\left(X^2 + 2XY + Y^2\right)\right) - \left(E(X) + E(Y)\right)^2 =$$

$$= E\left(X^2\right) + E(2XY) + E\left(Y^2\right) - \left[E(X)^2 + E(Y)^2 + 2E(XY)\right] =$$

$$= D(X) + D(Y) + 2K(XY)$$

S pomočjo Bernoullijevega zaporedja, Gaussove normalne porazdelitve in zgornjih lastnosti, lahko izpeljemo La Place-ovo aproksimacijo:

$$X_i : \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}$$

$$X = X_1 + X_2 + ... + X_n$$

$$E(X_i) = 1 \cdot p + 0 \cdot q = p \Rightarrow E(X) = \sum_{i=1}^{n} E(X_i) = np$$

$$E(X^2{}_i) = 1 \cdot p + 0 \cdot q = p \Rightarrow E(X^2) = \sum_{i=1}^{n} E(X^2{}_i) = np$$

$$D(X_i) = E(X^2{}_i) - E(X_i)^2 = p - p^2 = p(1-p) = pq \Rightarrow D(X) = \sum_{i=1}^{n} D(X_i) = npq$$

Če vzamemo normalno porazdelitev z isto upognjenostjo in disperzijo, dobimo:

$$N(a,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} = \frac{1}{\sqrt{npq}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-np}{\sqrt{npq}}\right)^2} \doteq p_k$$

ewe