

Zgled. Trije lovci so hkrati ustrelili na divjega prašiča in ga ubili. Ko so prišli do njega, so našli v njem eno samo kroglo. Kolikšne so verjetnosti, da je vepra ubil

(a) prvi,

(b) drugi,

(b) tretji

lovec, če poznamo njihove verjetnosti, da zadanejo: 0, 2; 0, 4 in 0, 6?

Na ta način jim namreč lahko pomagamo pri pošteni delitvi plena (kajti ne smemo pozabiti, da imajo vsi v rokah nevarno orožje).

Sestavimo popolen sistem dogodkov in uporabimo dejstvo,
da so lovci med seboj neodvisni, torej

$$P(A * B * C) = P(A) * P(B) * P(C).$$

To nam zna pomagati pri računanju verjetnosti hipotez.

	.2	.4	.6				
	prvi	drugi	tretji	P (H_i)	st.kr.	P (E/H_i)	P (E*H_i)
H1	1	1	1	, 2*, 4*, 6	=0, 048	3	0
H2	0	1	1	, 8*, 4*, 6	=0, 192	2	0
H3	1	0	1	, 2*, 6*, 6	=0, 072	2	0
H4	1	1	0	, 2*, 4*, 4	=0, 032	2	0
H5	1	0	0	, 2*, 6*, 4	=0, 048	1	0, 048
H6	0	1	0	, 8*, 4*, 4	=0, 128	1	0, 128
H7	0	0	1	, 8*, 6*, 6	=0, 288	1	0, 288
H8	0	0	0	, 8*, 6*, 4	=0, 192	0	0
vsota					=1, 000		0, 464

$$P(\text{ena krogla je zadela}) = 0,048 + 0,128 + 0,288 = 0,464 = P(E).$$

Ostale verjetnosti računamo za preiskus:

$$P(\text{nobena krogla ni zadela}) = 0,192 = P(N'),$$

$$P(\text{dve krogli sta zadeli}) = 0,192 + 0,072 + 0,032 = 0,296 = P(D),$$

$$P(\text{tri krogle so zadele}) = 0,048 + 0,128 + 0,288 = 0,464 = P(T).$$

Vsota teh verjetnosti je seveda enaka 1.

Končno uporabimo Bayesov obrazec:

$$P(H_5/E) = \frac{P(H_5 * E)}{P(E)} = \frac{0,048}{0,464} = 0,103 = P(\text{prvi je zadel}),$$

$$P(H_6/E) = \frac{P(H_6 * E)}{P(E)} = \frac{0,128}{0,464} = 0,276 = P(\text{drugi je zadel}),$$

$$P(H_7/E) = \frac{P(H_7 * E)}{P(E)} = \frac{0,288}{0,464} = 0,621 = P(\text{tretji je zadel}).$$

Tudi vsota teh verjetnosti pa je enaka 1.

Delitev plena se opravi v razmerju 10,3 : 27,6 : 62,1

(in ne 2 : 4 : 6 oziroma 16,6 : 33,3 : 50,

kot bi kdo utegnil na hitro pomisliti).

Bernoullijevo zaporedje neodvisnih poskusov

O zaporedju neodvisnih poskusov

$$X_1, X_2, \dots, X_n, \dots$$

govorimo tedaj, ko so verjetnosti izidov v enem poskusu neodvisne od tega, kaj se zgodi v drugih poskusih.

*Zaporedje neodvisnih poskusov se imenuje **Bernoullijevo zaporedje**, če se more zgoditi v vsakem poskusu iz zaporedja neodvisnih poskusov le dogodek A z verjetnostjo $P(A) = p$ ali dogodek $\bar{A}$ z verjetnostjo $P(\bar{A}) = 1 - P(A) = 1 - p = q$.*

Primer: Primer Bernoullijevega zaporedja poskusov je met kocke, kjer ob vsaki ponovitvi poskusa pade šestica (dogodek A) z verjetnostjo $P(A) = p = 1/6$ ali ne pade šestica (dogodek $\bar{A}$) z verjetnostjo $P(\bar{A}) = q = 5/6$.

...Bernoullijevo zaporedje neodvisnih poskusov

V Bernoullijevem zaporedju neodvisnih poskusov nas zanima, kolikšna je verjetnost, da se v n zaporednih poskusih zgodi dogodek A natanko k -krat. To se lahko zgodi na primer tako, da se najprej zgodi k -krat dogodek A in nato v preostalih $(n - k)$ poskusih zgodi nasprotni dogodek $\bar{A}$:

$$P\left(\bigcap_{i=1}^k (X_i = A) \cap \bigcap_{i=k+1}^n (X_i = \bar{A})\right) = \prod_{i=1}^k P(A) \cdot \prod_{i=k+1}^n P(\bar{A}) = p^k \cdot q^{n-k}$$

Dogodek $P_n(k)$, da se dogodek A v n zaporednih poskusih zgodi natanko k -krat, se lahko zgodi tudi na druge načine in sicer je teh toliko, na kolikor načinov lahko izberemo k poskusov iz n poskusov. Teh je $\binom{n}{k}$. Ker so ti načini nezdružljivi med seboj, je verjetnost dogodka $P_n(k)$ enaka

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

Tej zvezi pravimo *Bernoullijev obrazec*.

...Bernoullijevo zaporedje neodvisnih poskusov

Primer: Iz posode, v kateri imamo 8 belih in 2 rdeči krogli, na slepo izberemo po eno kroglo in po izbiranju izvlečeno kroglo vrnemo v posodo. Kolikšna je verjetnost, da v petih poskusih izberemo 3–krat belo kroglo?

Dogodek A je, da izvlečem belo kroglo. Potem je

$$p = P(A) = \frac{8}{10} = 0,8$$

$$q = 1 - p = 1 - 0,8 = 0,2$$

Verjetnost, da v petih poskusih izberemo 3–krat belo kroglo, je:

$$P_5(3) = \binom{5}{3} 0,8^3 (1 - 0,8)^{5-3} = 0,205$$

Računanje $P_n(k)$

Uporaba rekurzije: $P_n(0) = q^n$

$$P_n(k) = \frac{(n - k + 1)p}{kq} P_n(k - 1), \quad \text{za } k = 1, \dots$$

Stirlingov obrazec:

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

Poissonov obrazec: za p blizu 0

$$P_n(k) \approx \frac{(np)^k e^{-np}}{k!}$$

Laplaceov točkovni obrazec:

$$P_n(k) \approx \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k - np)^2}{2npq}}$$

Računanje $P_n(k)$

Program R: Vrednost $P_n(k)$ dobimo z ukazom

```
dbinom( $k$ , size= $n$ , prob= $p$ )
```

```
> dbinom(50, size=1000, prob=0.05)
```

```
[1] 0.05778798
```

Izpeljava rekurzivne zveze

$$\begin{aligned}\frac{P_n(k)}{P_n(k-1)} &= \frac{\binom{n}{k} p^k q^{n-k}}{\binom{n}{k-1} p^{k-1} q^{n-k+1}} = \\ &= \frac{n! (k-1)! (n-k+1)! p}{k! (n-k)! n! q} = \frac{(n-k+1)p}{kq}\end{aligned}$$

Torej je res

$$P_n(k) = \frac{(n-k+1)p}{kq} P_n(k-1), \quad \text{za } k = 1, \dots$$

Laplaceov intervalski obrazec

Zanima nas, kolikšna je verjetnost $P_n(k_1, k_2)$, da se v Bernoullijevem zaporedju neodvisnih poskusov v n zaporednih poskusih zgodi dogodek A vsaj k_1 -krat in manj kot k_2 -krat. Označimo

$$x_k = \frac{k - np}{\sqrt{npq}} \quad \text{in} \quad \Delta x_k = x_{k+1} - x_k = \frac{1}{\sqrt{npq}}.$$

Tedaj je, če upoštevamo Laplaceov točkovni obrazec,

$$P_n(k_1, k_2) = \sum_{k=k_1}^{k_2-1} P_n(k) = \frac{1}{\sqrt{2\pi}} \sum_{k=k_1}^{k_2-1} e^{-\frac{1}{2}x_k^2} \Delta x_k$$

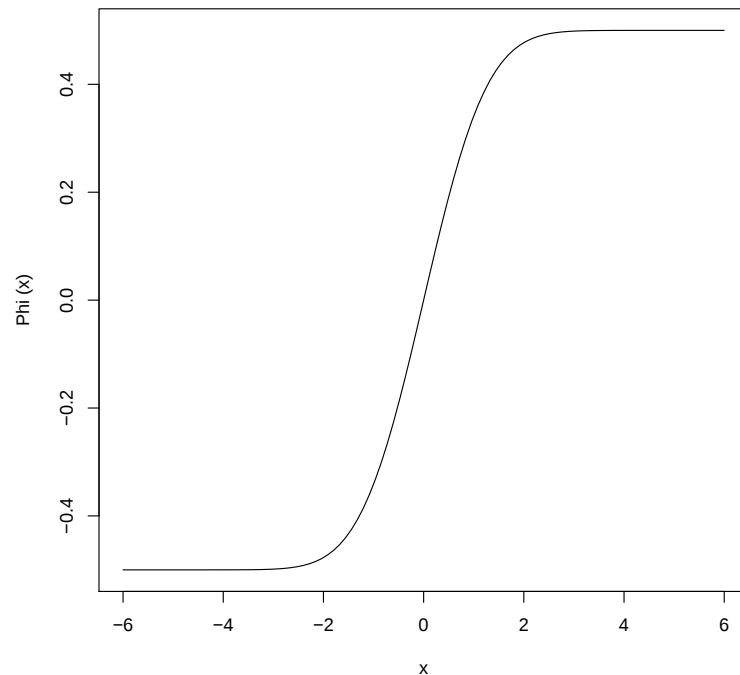
Za (zelo) velike n lahko vsoto zamenjamo z integralom

$$P_n(k_1, k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_{k_1}}^{x_{k_2}} e^{-\frac{1}{2}x^2} dx$$

Funkcija napake $\Phi(x)$

Funkcija napake imenujemo funkcijo

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}t^2} dt$$



```
> Phi <- function(x) {pnorm(x)-0.5}
> curve(Phi,-6.6)
```

Funkcija napake je liha, zvezno odvedljiva, strogo naraščajoča funkcija. $\Phi(-\infty) = -\frac{1}{2}$, $\Phi(0) = 0$, $\Phi(\infty) = \frac{1}{2}$ in $P_n(k_1, k_2) \approx \Phi(x_{k_2}) - \Phi(x_{k_1})$.

Vrednosti funkcije napake najdemo v tabelah ali pa je vgrajena v statističnih programih.

```
> x2 <- (50 - 1000*0.05)/sqrt(1000*0.05*0.95)
> x1 <- (0 - 1000*0.05)/sqrt(1000*0.05*0.95)
> pnorm(x2)-pnorm(x1)
[1] 0.5
```

Bernoullijev zakon velikih števil

IZREK 1 (J. Bernoulli, 1713) *Naj bo k frekvenca dogodka A v n neodvisnih ponovitvah danega poskusa, v katerem ima dogodek A verjetnost p . Tedaj za vsak $\varepsilon > 0$ velja*

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{k}{n} - p\right| < \varepsilon\right) = 1.$$

Ta izrek opravičuje statistično definicijo verjetnosti.

Slučajne spremenljivke in porazdelitve

Denimo, da imamo poskus, katerega izidi so števila (npr. pri metu kocke so izidi števila pik). Se pravi, da je poskusom prirejena neka količina, ki more imeti različne vrednosti. Torej je spremenljivka. Katero od mogočih vrednosti zavzame v določeni ponovitvi poskusa, je odvisno od slučaja. Zato ji rečemo **slučajna spremenljivka**.

Da je slučajna spremenljivka znana, je potrebno vedeti

1. kakšne vrednosti more imeti (*zaloge vrednosti*) in
2. kolikšna je verjetnost vsake izmed možnih vrednosti ali intervala vrednosti.

Predpis, ki določa te verjetnosti, imenujemo **porazdelitveni zakon**.

...Slučajne spremenljivke

Slučajne spremenljivke označujemo z velikimi tiskanimi črkami iz konca abecede, vrednosti spremenljivke pa z enakimi malimi črkami. Tako je npr. $(X = x_i)$ dogodek, da slučajna spremenljivka X zavzame vrednost x_i .

Porazdelitveni zakon slučajne spremenljivke X je poznan, če je mogoče za vsako realno število x določiti verjetnost

$$F(x) = P(X < x)$$

$F(x)$ imenujemo **porazdelitvena funkcija**.

Najpogosteje uporabljamo naslednji vrsti slučajnih spremenljivk:

1. **diskretna** slučajna spremenljivka, pri kateri je zaloga vrednosti neka števna množica;
2. **zvezna** slučajna spremenljivka, ki lahko zavzame vsako realno število znotraj določenega intervala.

Lastnosti porazdelitvene funkcije

1. Funkcija F je definirana na vsem $\mathbb{R}$ in velja $0 \leq F(x) \leq 1, x \in \mathbb{R}$
2. Funkcija F je naraščajoča $x_1 < x_2 \implies F(x_1) \leq F(x_2)$
3. $F(-\infty) = 0$ in $F(\infty) = 1$
4. Funkcija je v vsaki točki zvezna od leve $F(x-) = F(x)$
5. Funkcija ima lahko v nekaterih točkah skok. Vseh skokov je največ števno mnogo.
6. $P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$
7. $P(x_1 < X < x_2) = F(x_2) - F(x_1+)$
8. $P(X \geq x) = 1 - F(x)$
9. $P(X = x) = F(x+) - F(x)$

Diskretne slučajne spremenljivke

Zaloga vrednosti diskretne slučajne spremenljivke X je števna množica $\{x_1, x_2, \dots, x_m, \dots\}$. Dogodki

$$X = x_k \quad k = 1, 2, \dots$$

sestavljajo popoln sistem dogodkov. Označimo verjetnost posameznega dogodka s

$$P(X = x_i) = p_i$$

Vsota verjetnosti vseh dogodkov je enaka 1:

$$p_1 + p_2 + \dots + p_m + \dots = 1$$

Verjetnostna tabela

Verjetnostna tabela prikazuje diskretno slučajno spremenljivko s tabelo tako, da so v prvi vrstici zapisane vse vrednosti x_i , pod njimi pa so pripisane pripadajoče verjetnosti:

$$X : \begin{pmatrix} x_1 & x_2 & \cdots & x_m & \cdots \\ p_1 & p_2 & \cdots & p_m & \cdots \end{pmatrix}$$

Porazdelitvena funkcija je v tem primeru

$$F(x_k) = P(X < x_k) = \sum_{i=1}^{k-1} p_i$$

Enakomerna diskretna porazdelitev

Končna diskretna slučajna spremenljivka se porazdeljuje **enakomerno**, če so vse njene vrednosti enako verjetne. Primer take slučajne spremenljivke je število pik pri metu kocke

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

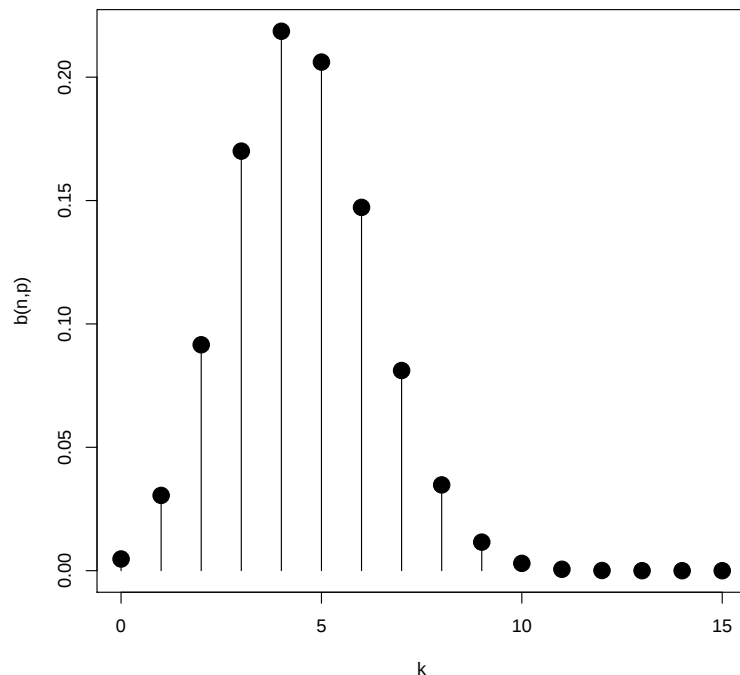
Binomska porazdelitev

Binomska porazdelitev ima zalogo vrednosti $\{0, 1, 2, \dots, n\}$ in verjetnosti, ki jih računamo po Bernoullijevem obrazcu:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$k = 0, 1, 2, \dots, n$. Binomska porazdelitev je natanko določena z dvema podatoma – parametroma: n in p . Če se slučajna spremenljivka X porazdeljuje binomsko s parametroma n in p , zapišemo:

$$X : B(n, p)$$



```
> h <- dbinom(0:15,size=15,prob=0.3)
> plot(0:15,h,type='h',xlab='k',ylab='b(n,p)')
> points(0:15,h,pch=16,cex=2)
```