



REPUBLIKA SLOVENIJA
**MINISTRSTVO ZA IZOBRAŽEVANJE,
ZNANOST IN ŠPORT**

Univerza v Ljubljani



EVROPSKA UNIJA
EVROPSKI
SOCIALNI SKLAD

Pogojna verjetnost

Verjetnost in statistika, študijsko leto 2017/18

doc. dr. Boštjan Kuzman

Univerza v Ljubljani, Pedagoška fakulteta

Projekt "IKT v pedagoških študijskih programih UL" delno financirata Republika Slovenija in Evropska unija iz Evropskega socialnega sklada.

1. Definicija pogojne verjetnosti
2. Osnovne lastnosti pogojne verjetnosti
3. Neodvisnost dogodkov
4. Zaporedje neodvisnih poskusov
5. Popolna verjetnost
6. Bayesov izrek o verjetnosti hipotez

Definicija pogojne verjetnosti

Zgled (Dve kocki)

Vržemo dve običajni igralni kocki. Kakšna je verjetnost, da je vsaj na eni padlo sodo število pik, če vemo, da je vsota padlih pik enaka 6?

Zgled (Dve kocki)

Vržemo dve običajni igralni kocki. Kakšna je verjetnost, da je vsaj na eni padlo sodo število pik, če vemo, da je vsota padlih pik enaka 6?

Vseh izidov pri metu dveh kock je 36:

(1,1),	(1,2),	(1,3),	(1,4),	(1,5),	(1,6)
(2,1),	(2,2),	(2,3),	(2,4),	(2,5),	(2,6)
(3,1),	(3,2),	(3,3),	(3,4),	(3,5),	(3,6)
(4,1),	(4,2),	(4,3),	(4,4),	(4,5),	(4,6)
(5,1),	(5,2),	(5,3),	(5,4),	(5,5),	(5,6)
(6,1),	(6,2),	(6,3),	(6,4),	(6,5),	(6,6)

- 27 izidov ima sodo število pik vsaj na eni kocki. $P(A) = 27/36$.
- 5 izidov ima vsoto 6: $P(B) = 5/36$.
- za 2 izida velja oboje: $P(A \cap B) = 2/36$.

Smiselno se zdi reči, da je iskana verjetnost $2/5$, saj podatek o vsoti število možnih izidov skrči iz 36 na 5 izidov, od katerih sta ugodna 2.

Definicija

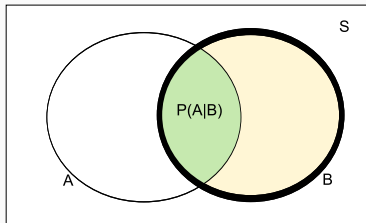
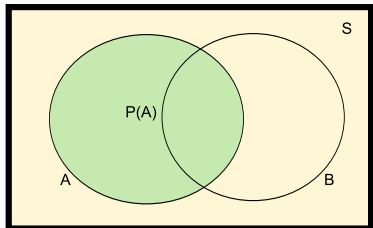
Naj bo (S, P) verjetnostni prostor in naj bosta $A, B \subseteq S$ neka dogodka, tako da je $P(B) > 0$. *Pogojno verjetnost dogodka A pri pogoju B* definiramo kot razmerje

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Definicija

Naj bo (S, P) verjetnostni prostor in naj bosta $A, B \subseteq S$ neka dogodka, tako da je $P(B) > 0$. *Pogojno verjetnost dogodka A pri pogoju B* definiramo kot razmerje

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$



Predstavljamo si lahko, da $P(A)$ meri delež A v celotnem prostoru izidov S , $P(A|B)$ pa delež A znotraj manjše množice B .

Zgled (Met kocke)

Naj bo $A = \{6\}$ dogodek, da je padla šestica, in $B = \{2, 4, 6\}$ dogodek, da je padlo sodo število. Potem je

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{3/6} = \frac{1}{3}$$

in

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/6}{1/6} = 1.$$

Zgled (Več pik zmaga)

Ana in Bine vržeta vsak svojo kocko. Zmaga tisti, ki je vrgel več pik, če je rezultat izenačen, pa met ponovita. Denimo, da je Ana v zadnji igri zmagala. Kakšna je verjetnost, da je vrgla 5 pik?

Zgled (Več pik zmaga)

Ana in Bine vržeta vsak svojo kocko. Zmaga tisti, ki je vrgel več pik, če je rezultat izenačen, pa met ponovita. Denimo, da je Ana v zadnji igri zmagala. Kakšna je verjetnost, da je vrgla 5 pik?

Če izide zapišemo kot pare (a, b) , je možnih izidov $36 - 6 = 30$, saj lahko izpustimo pare z $a = b$.

Za Anino zmago je ugodnih 15 izidov:

$$A = \{(2, 1), (3, 2), (3, 1), \dots, (6, 5), (6, 4), (6, 3), (6, 2), (6, 1)\}.$$

Izidov, pri katerih je Ana vrgla 5 pik, pa je 5:

$$B = \{(5, 6), (5, 4), (5, 3), (5, 2), (5, 1)\}.$$

Ker je $|A \cap B| = 4$, sledi $P(B|A) = 4/15$.

Zgled (Dve dami)

Vi in vaš soigralec iz kompleta 54 kart izbereta vsak 2 karti. Denimo, da ima vaš partner dve dami. Kakšna je potem verjetnost, da imate vi srčevo damo?

Zgled (Dve dami)

Vi in vaš soigralec iz kompleta 54 kart izbereta vsak 2 karti. Denimo, da ima vaš partner dve dami. Kakšna je potem verjetnost, da imate vi srčevo damo?

Označimo $A \equiv$ imate srčevo damo, in $B \equiv$ soigralec ima dve dami. Iščemo $P(A|B) = P(A \cap B)/P(B)$, zato izračunamo

$$P(B) = \frac{\binom{4}{2} \binom{52}{2}}{\binom{54}{2} \binom{52}{2}}, P(A \cap B) = \frac{\binom{1}{1} \binom{51}{1} \binom{3}{2}}{\binom{54}{2} \binom{52}{2}}.$$

Sledi:

$$P(A|B) = \frac{\binom{3}{2} \binom{51}{1}}{\binom{52}{2} \binom{4}{2}} = \frac{1}{52}.$$

Osnovne lastnosti pogojne verjetnosti

Izrek (Pogojna verjetnost je verjetnost)

Naj bo (S, P) verjetnostni prostor in naj bo $B \subseteq S$ dogodek z neničelno verjetnostjo. Potem je preslikava $Q: \mathcal{P}(S) \rightarrow \mathbb{R}$, ki jo definiramo s predpisom $Q(A) = P(A|B)$, verjetnost na S .

Izrek (Pogojna verjetnost je verjetnost)

Naj bo (S, P) verjetnostni prostor in naj bo $B \subseteq S$ dogodek z neničelno verjetnostjo. Potem je preslikava $Q: \mathcal{P}(S) \rightarrow \mathbb{R}$, ki jo definiramo s predpisom $Q(A) = P(A|B)$, verjetnost na S .

Dokaz.

Nenegativnost in normiranost preslikave Q sta očitni. Za dokaz števnice aditivnosti Q naj bodo $A_1, A_2, \dots \subseteq S$ paroma ločeni dogodki. Potem je

$$\begin{aligned} Q\left(\bigcup_{i=1}^{\infty} A_i\right) &= \frac{P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right)}{P(B)} = \frac{\sum_{i=1}^{\infty} P(A_i \cap B)}{P(B)} = \\ &= \sum_{i=1}^{\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) = \sum_{i=1}^{\infty} Q(A_i), \end{aligned}$$

saj so tudi dogodki $A_1 \cap B, A_2 \cap B, \dots$ paroma ločeni in lahko pri tretjem enačaju uporabimo števno aditivnost P . □

Trditev (Osnovne lastnosti pogojne verjetnosti)

Za pogojno verjetnost $P(\cdot | B)$ in vse dogodke $A \subseteq S$ velja:

- (a) $P(A \cap B) = P(A|B)P(B)$ (produktno pravilo).
- (b) $P(A^c|B) = 1 - P(A|B)$.
- (c) Če je $A \subseteq B$, potem je $P(A|B) = \frac{P(A)}{P(B)}$ in $P(B|A) = 1$. Posebej, $P(B|B) = 1$.
- (d) Če je $A \cap B = \emptyset$, je $P(A|B) = 0$.

Zgled (Hokejski turnir)

Hokejsko moštvo se je uvrstilo v polfinale, kjer ima 60% možnosti za zmago. Če zmagajo, bodo dobili v finalu močnega nasprotnika, proti katerem imajo le 40% možnosti. Kakšna je verjetnost, da osvojijo turnir?

Zgled (Hokejski turnir)

Hokejsko moštvo se je uvrstilo v polfinale, kjer ima 60% možnosti za zmago. Če zmagajo, bodo dobili v finalu močnega nasprotnika, proti katerem imajo le 40% možnosti. Kakšna je verjetnost, da osvojijo turnir?

Naj bosta A in B dogodka, da naši hokejisti zmagajo v polfinalu oziroma v finalu. Potem je $A \cap B$ dogodek, da so osvojili turnir, zato po produktnem pravilu velja

$$P(A \cap B) = P(A) \cdot P(B|A) = 0.6 \cdot 0.4 = 0.24.$$

Trditev (Produktno pravilo za več dogodkov)

Naj bodo $A_1, A_2, \dots, A_n \subseteq S$ taki dogodki, da je $P(A_1 \cap \dots \cap A_{n-1}) > 0$.

Potem velja

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

Trditev (Produktno pravilo za več dogodkov)

Naj bodo $A_1, A_2, \dots, A_n \subseteq S$ taki dogodki, da je $P(A_1 \cap \dots \cap A_{n-1}) > 0$.
Potem velja

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

Dokaz.

Desno stran v zgornji enakosti zapišemo kot

$$P(A_1) \cdot \frac{P(A_2 \cap A_1)}{P(A_1)} \cdot \frac{P(A_3 \cap (A_2 \cap A_1))}{P(A_2 \cap A_1)} \cdot \dots \cdot \frac{P(A_n \cap (A_{n-1} \cap \dots \cap A_2 \cap A_1))}{P(A_{n-1} \cap \dots \cap A_2 \cap A_1)}$$

in po krajšanju ostane le člen desno zgoraj.

Uporabili smo, da so imenovalci neničelni, kar sledi iz pogoja $P(A_1 \cap \dots \cap A_{n-1}) > 0$ in monotonosti verjetnostne funkcije P . \square

Zgled (Bonboni)

V vrečki je b belih in c črnih bonbonov. Vsak dan pojemo naključno izbran bonbon. Kakšna je verjetnost, da prva dva dni pojemo bel, druga dva dni pa črn bonbon?

Zgled (Bonboni)

V vrečki je b belih in c črnih bonbonov. Vsak dan pojemo naključno izbran bonbon. Kakšna je verjetnost, da prva dva dni pojemo bel, druga dva dni pa črn bonbon?

Naj A_i pomeni, da i -ti dan pojemo bel bonbon. Zanima nas

$$\begin{aligned} P(A_1 \cap A_2 \cap \bar{A}_3 \cap \bar{A}_4) &= P(A_1)P(A_2|A_1)P(\bar{A}_3|A_1 \cap A_2)P(\bar{A}_4|A_1 \cap A_2 \cap \bar{A}_3) \\ &= \frac{b}{b+c} \cdot \frac{b-1}{b+c-1} \cdot \frac{c}{b+c-2} \cdot \frac{c-1}{b+c-3}. \end{aligned}$$

Neodvisnost dogodkov

Zgled

Vržemo dva kovanca. Na prvem pade grb. Izkušnje kažejo, da to ne vpliva na drugi kovanec. Kako to povedati z enačbo?

Zgled

Vržemo dva kovanca. Na prvem pade grb. Izkušnje kažejo, da to ne vpliva na drugi kovanec. Kako to povedati z enačbo?

Ideja: pogojna verjetnost naj bo enaka "začetni" verjetnosti,

$$P(A|B) = P(A).$$

Toda, če je $P(B) = 0$, potem $P(A|B)$ ni definirana...

Definicija

Dogodka $A, B \subseteq S$ sta *neodvisna*, če velja

$$P(A \cap B) = P(A)P(B).$$

Opomba

Razmisli, da v primeru, ko je $P(A), P(B) > 0$, velja:

- A in B neodvisna $\iff P(A|B) = P(A)$ in $P(B|A) = P(B)$.
- A in B ločena $\implies A$ in B odvisna.

Zgled (Met kocke)

Vržemo običajno kocko in označimo dogodke:

$A \equiv$ pade sodo pik,

$B \equiv$ padeta največ 2 piki,

$C \equiv$ pade 1 pika.

Brez računanja presodi, ali sta A in B neodvisna?

Kaj pa B in C oziroma A in C ? Domneve preveri z računom.

Zgled (Met kocke)

Vržemo običajno kocko in označimo dogodke:

$A \equiv$ pade sodo pik,

$B \equiv$ padeta največ 2 piki,

$C \equiv$ pade 1 pika.

Brez računanja presodi, ali sta A in B neodvisna?

Kaj pa B in C oziroma A in C ? Domneve preveri z računom.

Neodvisna sta A in B , druga dva para pa ne.

Zgled (Spol otrok)

V družini so trije otroci, oba spola sta enako verjetna. Naj bo $A \equiv$ med otroci je največ 1 deklica, in $B \equiv$ med otroci sta zastopana oba spola. Ali sta dogodka A in B neodvisna? Kaj pa, če so otroci štirje?

Zgled (Spol otrok)

V družini so trije otroci, oba spola sta enako verjetna. Naj bo $A \equiv$ med otroci je največ 1 deklica, in $B \equiv$ med otroci sta zastopana oba spola. Ali sta dogodka A in B neodvisna? Kaj pa, če so otroci štirje?

Pri 3 otrocih sta dogodka neodvisna, pri 4 otrocih odvisna!

Trditev

Če sta A in B neodvisna, sta neodvisna tudi A in B^c , in posledično tudi A^c in B^c .

Dokaz.

Naredi sam.



Pri treh ali več dogodkih A_1, A_2, \dots, A_n ločimo dva tipa neodvisnosti:

paroma neodvisni $\iff P(A_i \cap A_j) = P(A_i)P(A_j)$ za vse $i \neq j$.

povsem neodvisni $\iff P(\cap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$ za vse $I \subseteq \{1, \dots, n\}$.

Pri treh ali več dogodkih A_1, A_2, \dots, A_n ločimo dva tipa neodvisnosti:

paroma neodvisni $\iff P(A_i \cap A_j) = P(A_i)P(A_j)$ za vse $i \neq j$.

povsem neodvisni $\iff P(\cap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$ za vse $I \subseteq \{1, \dots, n\}$.

Zgled (5 kovancev)

Vržemo 5 kovancev, A_i je dogodek, da na i -tem pade grb. Potem so dogodki A_1, \dots, A_5 povsem neodvisni (in zato tudi paroma neodvisni).

Zgled (Paroma neodvisno ni povsem neodvisno)

Vržemo dva različno velika kovanca in označimo dogodke:

$A \equiv$ na večjem kovancu pade cifra,

$B \equiv$ na manjšem kovancu pade grb,

$C \equiv$ pade natanko 1 cifra.

Potem ne velja enakost $P(A \cap B \cap C) = P(A)P(B)P(C)$, čeprav so posamezni pari neodvisni.

Zaporedje neodvisnih poskusov

Zaporedje neodvisnih poskusov je zaporedje ponovitev istega poskusa, pri katerem so izidi različnih ponovitev med seboj neodvisni.

Bernoullijev poskus je poskus z dvema možnima izidoma, A in A^c . Verjetnost ugodnega izida običajno označimo s $p = P(A)$ in neugodnega s $q = 1 - p = P(A^c)$.

Bernoullijevo zaporedje je zaporedje neodvisnih ponovitev Bernoullijevega poskusa.

Zgled

Kovanec z verjetnostjo grba p vržemo n -krat. Potem so izidi pri posameznih metih neodvisni.

Trditev

Naj bo $p = P(A)$ verjetnost dogodka A v Bernoullijevem poskusu.

Potem velja:

- (i) Verjetnost, da se je dogodek A v n ponovitvah poskusa zgodil natanko k -krat, je enaka

$$\binom{n}{k} p^k q^{n-k}. \quad (\text{Bernoullijev obrazec})$$

- (ii) Verjetnost, da se je dogodek A zgodil k -tič natanko v n -ti ponovitvi, je enaka

$$\binom{n-1}{k-1} p^k q^{n-k}. \quad (\text{Pascalov obrazec}).$$

- (iii) Posebej, verjetnost, da se je dogodek A prvič zgodil v n -ti ponovitvi, je enaka

$$pq^{n-1}.$$

Dokaz.

Naj bo A_i dogodek, da se je v i -to ponovitvi poskusa zgodil dogodek A . Potem je

$$A_1 \cap A_2 \cap \dots \cap A_k \cap \bar{A}_{k+1} \cap \dots \cap \bar{A}_n$$

dogodek, da se je v n ponovitvah poskusa najprej k -krat zgodil A , nato pa $(n - k)$ -krat \bar{A} . Zaradi neodvisnosti je verjetnost tega dogodka enaka produktu $p^k q^{n-k}$. Toda k ponovitev dogodka A bi se lahko zgodilo tudi z drugačnim vrstnim redom dogodkov A in \bar{A} . Vseh možnih razporeditev je $\binom{n}{k}$, saj izbiramo k od možnih n mest. Ker so načini med seboj nezdružljivi, lahko njihove verjetnosti seštejemo in rezultat do Bernoullijev obrazec.

Pascalov obrazec izpeljemo iz Bernoullijevega, saj vemo, da se je dogodek A k -tič ponovil natanko v n -ti ponovitvi poskusa, pred tem, pa se je ponovil $(k - 1)$ -krat v prvih $(n - 1)$ ponovitvah. Ker je zadnja ponovitev neodvisna od prejšnjih, ustrezni verjetnosti zmnožimo in dobimo željeni rezultat.

Zadnji obrazec pa je seveda le poseben primer prejšnjega za $k = 1$.

Zgled

Verjetnost, da v 5 metih kovanca padeta natanko 2 grba, je enaka

$$\binom{5}{2} p^2 (1-p)^3,$$

kar je približno 31%, če je $p = 1/2$.

Verjetnost, da je drugi grb padel v petem metu, pa je enaka

$$\binom{4}{1} p^2 (1-p)^3$$

oziroma približno 12%, če je $p = 1/2$.

Zgled (Kroglice v žari)

Denimo, da je v žari b belih in c črnih kroglic.

Če iz žare večkrat zapored izberemo naključno kroglico, pogledamo njeno barvo in jo vrnemo v žaro, lahko to situacijo obravnavamo kot zaporedje neodvisnih Bernoullijevih poskusov.

Če kroglic med ponovitvami ne vračamo, pa zaporedni poskusi niso neodvisni. V tem primeru situacijo običajno obravnavamo s pomočjo Izreka o popolni verjetnosti.

Popolna verjetnost

Definicija

Popoln sistem dogodkov je števna družina dogodkov A_1, A_2, \dots , za katero velja:

- Dogodki imajo pozitivno verjetnost: $P(A_i) > 0$ za vsak $i \in I$.
- Dogodki so paroma nezdružljivi: $A_i \cap A_j = \emptyset$ za vse $i \neq j$.
- Unija dogodkov je celoten prostor izidov: $\bigcup A_i = S$.

Z drugimi besedami: popoln sistem dogodkov je števno razbitje prostora izidov na naključne dogodke. V praksi to pomeni, da se pri kateremkoli izidu poskusa zgodi natanko eden od teh dogodkov.

Izrek (Izrek o popolni verjetnosti)

Naj bo (S, P) verjetnostni prostor in A_1, A_2, \dots popoln sistem dogodkov. Potem za poljuben dogodek $B \subseteq S$ velja

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i).$$

Izrek (Izrek o popolni verjetnosti)

Naj bo (S, P) verjetnostni prostor in A_1, A_2, \dots popoln sistem dogodkov. Potem za poljuben dogodek $B \subseteq S$ velja

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i).$$

Dokaz.

Ker je $B = B \cap S = B \cap (\cup_{i \in I} A_i) = \cup_{i \in I} (B \cap A_i)$ števna unija disjunktnih dogodkov A_i , sledi

$$P(B) = P(\cup_{i \in I} (B \cap A_i)) = \sum_{i \in I} P(B \cap A_i) = \sum_{i \in I} P(B|A_i)P(A_i)$$

iz števne aditivnosti verjetnosti P in definicije pogojne verjetnosti. □

Zgled (Krvne skupine)

Porazdelitev krvnih skupin A, B, O, AB v celotni populaciji je naslednja:

O: 44%, A: 42%, B: 10%, AB: 4%.

Denimo, da je verjetnost, da oseba s krvno skupino O zbolí za boleznijo X, enaka 0.05, pri ostalih krvnih skupinah pa 0.15. Potem je verjetnost, da naključno izbrana oseba zbolí, enaka

$$\begin{aligned}P(X) &= P(X|O)P(O) + P(X|A)P(A) + P(X|B)P(B) + P(X|AB)P(AB) \\ &= 0.05 \cdot 0.44 + 0.15 \cdot (0.42 + 0.10 + 0.04) \doteq 0.11.\end{aligned}$$

Zgled (Dvofazni poskus)

Iz žare s 5 modrimi in 3 rdečimi kroglicami naključno izvlečemo kroglico in jo damo v drugo žaro, v kateri so 3 modre in 5 rdečih kroglic. Nato iz druge žare izvlečemo eno kroglico. Kakšna je verjetnost, da je ta kroglica modra?

Zgled (Dvofazni poskus)

Iz žare s 5 modrimi in 3 rdečimi kroglicami naključno izvlečemo kroglico in jo damo v drugo žaro, v kateri so 3 modre in 5 rdečih kroglic. Nato iz druge žare izvlečemo eno kroglico. Kakšna je verjetnost, da je ta kroglica modra?

Dogodka $A_1 \equiv$ "prvič smo izvlekli modro" in $A_2 = \bar{A}_1 \equiv$ "prvič smo izvlekli rdečo kroglico" sestavljata popoln sistem dogodkov. Zato je verjetnost dogodka $B \equiv$ "drugič smo izvlekli modro" enaka

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) = \frac{4}{9} \cdot \frac{5}{8} + \frac{3}{9} \cdot \frac{3}{8} = \frac{29}{72} \doteq 0.40.$$

Zgled (Uganka za snubca)

Kralj je snubcu svoje hčerke postavil naslednjo nalogo: 4 bele in 4 črne kroglice razporedi v dve skrinjici, kakorkoli želiš. Nato bo hčerka naključno izbrala eno skrinjico in v njej naključno kroglico. Če bo kroglica bela, dobiš hčerko za ženo. Če bo črna, pa te obglavim.

Ali lahko snubec sploh kaj vpliva na svojo usodo?

Zgled (Uganka za snubca)

Kralj je snubcu svoje hčerke postavil naslednjo nalogo: 4 bele in 4 črne kroglice razporedi v dve skrinjici, kakorkoli želiš. Nato bo hčerka naključno izbrala eno skrinjico in v njej naključno kroglico. Če bo kroglica bela, dobiš hčerko za ženo. Če bo črna, pa te obglavim.

Ali lahko snubec sploh kaj vpliva na svojo usodo?

Presenetljiv odgovor je, da lahko. Če da v prvo skrinjico 1 belo kroglico, v drugo skrinjico pa vse ostale kroglice, bo verjetnost, da princeska izvleče belo, enaka

$$P(B) = P(B|S_1)P(S_1) + P(B|S_2)P(S_2) = 1 \cdot \frac{1}{2} + \frac{3}{7} \cdot \frac{1}{2} = \frac{5}{7}.$$

Od vseh možnih razporeditev kroglic ima ta najugodnejšo verjetnost za snubca. Bralec naj sam razmisli še o preostalih možnostih.

Bayesov izrek o verjetnosti hipotez

Izrek (Bayesov izrek)

Naj bosta $A, B \subseteq S$ poljubna dogodka z neničelno verjetnostjo. Potem je

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}.$$

Splošneje, če je $\{A_1, A_2, \dots\}$ popoln sistem dogodkov in $B \subseteq S$ poljuben dogodek, potem je

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_i P(B|A_i)P(A_i)} = \frac{P(B \cap A_j)}{\sum_i P(B \cap A_i)}.$$

Izrek (Bayesov izrek)

Naj bosta $A, B \subseteq S$ poljubna dogodka z neničelno verjetnostjo. Potem je

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}.$$

Splošneje, če je $\{A_1, A_2, \dots\}$ popoln sistem dogodkov in $B \subseteq S$ poljuben dogodek, potem je

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_i P(B|A_i)P(A_i)} = \frac{P(B \cap A_j)}{\sum_i P(B \cap A_i)}.$$

Dokaz.

Prva enakost sledi iz definicije pogojne verjetnosti. Za drugo v prvi zamenjaš A z A_j in $P(B)$ z vsoto iz izreka o popolni verjetnosti. \square

Kaj pove Bayesov izrek?

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_i P(B|A_i)P(A_i)} = \frac{P(B \cap A_j)}{\sum_i P(B \cap A_i)}.$$

V opazovanem poskusu je dogodek B odvisen od hipotetičnih dogodkov A_i , ki so se zgodili pred njim. Zato so verjetnosti $P(A_i)$ in $P(B|A_i)$ načeloma znane vnaprej. Z pa izrekom izračunamo verjetnost hipotetičnega dogodka $P(A_j|B)$ na podlagi tega, ali se je zgodil kasnejši dogodek B .

”Bayesov izrek je za verjetnost kot Pitagorov izrek za geometrijo.”
(H. Jeffreys)

Zgled (Tovarne baterij)

Baterije za mobilne telefone izdelujejo v 3 različnih tovarnah. Pri tem je verjetnost okvare 1% v prvi, 2% v drugi in 3% v tretji tovarni, število vseh izdelanih baterij pa je 1000 v prvi, 3000 v drugi in 6000 v tretji tovarni. Izmed 10000 izdelanih telefonov izberemo naključnega. Kakšna je verjetnost, da je bil izdelan v prvi tovarni, če ima okvarjeno baterijo?

Zgled (Tovarne baterij)

Baterije za mobilne telefone izdelujejo v 3 različnih tovarnah. Pri tem je verjetnost okvare 1% v prvi, 2% v drugi in 3% v tretji tovarni, število vseh izdelanih baterij pa je 1000 v prvi, 3000 v drugi in 6000 v tretji tovarni. Izmed 10000 izdelanih telefonov izberemo naključnega. Kakšna je verjetnost, da je bil izdelan v prvi tovarni, če ima okvarjeno baterijo?

Če je A_i dogodek, da je telefon izdelan v i -ti tovarni, in B , da ima okvarjeno baterijo, potem je

$$\begin{aligned} P(A_1|B) &= \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3)} = \\ &= \frac{0.01 \cdot 0.1}{0.01 \cdot 0.1 + 0.02 \cdot 0.3 + 0.03 \cdot 0.6} = 4\% \end{aligned}$$

Zgled (Mamogram)

Z mamogramom preverjamo prisotnost raka na dojki, za katerim zboli približno 1 procent žensk, starih med 40 in 50 let. Test pa ni povsem zanesljiv, in včasih pokaže pozitiven izid, če bolezn ni, ali obratno. Označimo dogodka $B \equiv$ "smo bolni", in $A \equiv$ "test je pozitiven". Znano je, da velja:

Verjetnost pozitivnega testa, če smo bolni, je $P(A|B) = 90\%$,

Verjetnost pozitivnega testa, če nismo bolni, je $P(A|\bar{B}) = 10\%$.

Kakšna je torej verjetnost, da smo zares bolni, če je test pozitiven?

Zgled (Mamogram)

Z mamogramom preverjamo prisotnost raka na dojki, za katerim zboli približno 1 procent žensk, starih med 40 in 50 let. Test pa ni povsem zanesljiv, in včasih pokaže pozitiven izid, če boleznini ni, ali obratno. Označimo dogodka $B \equiv$ "smo bolni", in $A \equiv$ "test je pozitiven". Znano je, da velja:

Verjetnost pozitivnega testa, če smo bolni, je $P(A|B) = 90\%$,

Verjetnost pozitivnega testa, če nismo bolni, je $P(A|\bar{B}) = 10\%$.

Kakšna je torej verjetnost, da smo zares bolni, če je test pozitiven?

Ker je $\{B, \bar{B}\}$ PSD, lahko po Bayesu izračunamo

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.1 \cdot 0.99} \doteq$$

Zgled (Mamogram)

Z mamogramom preverjamo prisotnost raka na dojki, za katerim zboli približno 1 procent žensk, starih med 40 in 50 let. Test pa ni povsem zanesljiv, in včasih pokaže pozitiven izid, če boleznini ni, ali obratno. Označimo dogodka $B \equiv$ "smo bolni", in $A \equiv$ "test je pozitiven". Znano je, da velja:

Verjetnost pozitivnega testa, če smo bolni, je $P(A|B) = 90\%$,

Verjetnost pozitivnega testa, če nismo bolni, je $P(A|\bar{B}) = 10\%$.

Kakšna je torej verjetnost, da smo zares bolni, če je test pozitiven?

Ker je $\{B, \bar{B}\}$ PSD, lahko po Bayesu izračunamo

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})} = \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.1 \cdot 0.99} \doteq 8\%.$$

Verjetnost, da smo bolni, če je test pozitiven, je komaj 8%?

Presenetljivi rezultat je posledica tega, da je bolezen razmeroma redka, verjetnost pozitivnega izida pri zdravih ženskah pa razmeroma velika. Če bi sistematično testirali 1000 kandidatk, bi bilo v povprečju bolnih 10, od katerih bi jih 9 dobilo pozitiven izvid testa, med 990 zdravimi pa bi dobili še 99 lažno pozitivnih izidov, torej je resnično bolnih med pozitivnimi izidi ravno $9/(9 + 99) \doteq 8\%$.

Zato je testiranje v primeru pozitivnega izida nujno ponoviti. Zavedati pa se moramo, da gornji izračuni veljajo le ob predpostavki, da je bolezen neodvisno in enakomerno porazdeljena po celotni populaciji. Če testirana oseba spada v katero izmed rizičnih skupin, bo njena bolezen v primeru pozitivnega testa bistveno verjetnejša.

Zgled

V nekem mestu je delež taksijev bele barve enak $x = 15\%$, preostali taksiji pa so rumeni. Ko neki taksi povzroči nesrečo, očividec trdi, da je bil bele barve. Vendar test pokaže, da očividec ponoči barvo taksija pravilno prepozna le v $y = 80\%$ primerov. Kakšna je torej verjetnost, da je bil v nesreči res udeležjen taksi bele barve, če tako trdi očividec?

Zgled

V nekem mestu je delež taksijev bele barve enak $x = 15\%$, preostali taksiji pa so rumeni. Ko neki taksi povzroči nesrečo, očividec trdi, da je bil bele barve. Vendar test pokaže, da očividec ponoči barvo taksija pravilno prepozna le v $y = 80\%$ primerov. Kakšna je torej verjetnost, da je bil v nesreči res udeležen taksi bele barve, če tako trdi očividec?

$A \equiv$ udeleženi taksi je bele barve,

$B \equiv$ očividec trdi, da je udeleženi taksi bele barve,

Potem je $P(A) = 15\%$, $P(B|A) = 80\%$ in po Bayesu sledi

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{xy}{xy + (1-x)(1-y)} \doteq 41\%.$$

Verjetnost je torej manjša od polovice!

Oglejmo si še, kako je v isti nalogi verjetnost $P(A|B)$ odvisna od spremembe parametrov $x, y \in [0, 1]$:

- Kdaj je verjetnost vsaj polovična?

$$P(A|B) \geq 1/2 \iff \frac{xy}{xy + (1-x)(1-y)} \geq 1/2 \iff x + y \geq 1.$$

Oglejmo si še, kako je v isti nalogi verjetnost $P(A|B)$ odvisna od spremembe parametrov $x, y \in [0, 1]$:

- Kdaj je verjetnost vsaj polovična?

$$P(A|B) \geq 1/2 \iff \frac{xy}{xy + (1-x)(1-y)} \geq 1/2 \iff x + y \geq 1.$$

- Za $k = 0, 1, \dots, 9$ lahko izračunamo tudi, kdaj je $P(A|B) \geq k/10$ in to prikažemo s pomočjo interaktivnega modela.