FRI, Univerza v Ljubljani

Verjetnostni račun in statistika

Aleksandar Jurišić

Ljubljana, 1. oktober 2007

Predstavitev

Aleksandar Jurišić

FRI, Jadranska 21, soba 5

e-pošta: **ajurisic@valjhun.fmf.uni-lj.si**

WWW: **http://lkrv.fri.uni-lj.si/∼ajurisic**

asistenti:

dr. Oliver Dragičević (oliver.dragicevic@fmf.uni-lj.si),
mag. Gregor Šega (gregor.sega@fmf.uni-lj.si) in
univ. dip. ing. Boris Cergol (boris.cergol@fri.uni-lj.si)

Verjetnost in statistika [20116]

Cilj predmeta: Predstaviti osnove teorije verjetnosti in njeno uporabo v statistiki, predstaviti osnove statistike.

Kratka vsebina: Definicija verjetnosti, slučajne spremenljivke in vektorji, diskretne in zvezne porazdelitve, matematično upanje, disperzija in višji momenti, karakteristične funkcije, zaporedja slučajnih spremenljivk in slučajni procesi, osnovna naloga statistike, ocenjevanje parametrov, testiranje statističnih hipotez, analiza variance, kovariance in linearne regresije.

Obveznosti študenta

Uspešno opravljena *kolokvija*. Če ni šlo, je potrebno opraviti *pisni izpit* iz reševanja nalog.

Ko ima enkrat pozitivno oceno iz reševanja nalog, mora v tekočem letu opraviti še *izpit iz teorije* (lahko je pisni ali ustni - odvisno od števila prijavljenih).

Viri

Pri predavanjih se bomo pretežno opirali na naslednje knjige:

W. Mendenhall in T. Sincich, Statistics for engineering and the sciences, 4th edition, Prentice Hall, 1995.

D. S. Moore (Purdue University), *Statistika: znanost o podatkih* (prevedeno v slovenščino leta 2007).

Milan Hladnik: *Verjetnost in statistika*. Založba FE in FRI, Ljubljana 2002.

Anuška Ferligoj: *Osnove statistike na prosojnicah*. Samozaložba, Ljubljana 1995.

Obstaja obilna literatura v knjižnicah.

Gradiva bodo dosegljiva preko internetne učilnice (moodle).

Pri delu z dejanskimi podatki se bomo v glavnem naslonili na prosti statistični program **R**.

Program je prosto dostopen na:

`http://www.r-project.org/`

proti koncu semestra pa morda tudi Minitab.

Zamenjalna šifra

Tomaž Pisanski, Skrivnostno sporočilo
Presek V/1, 1977/78, str. 40-42.

```
YHW?HD+CVODHVTHVO-!JVG:CDCYJ(JV/-V?HV(
-T?HVW-4YC4(?-DJV/-(?S-VO3CWC%J(-V4-DC
V!CW-?CVNJDJVD-?+-VO3CWC%J(-VQW-DQ-VJ+
V?HVDWHN-V3C:CODCV!H+?-DJVD-?+CV3JO-YC
```

(črko Č smo zamenjali s C, črko Ć pa z D)

Imamo $26! = 40329146112665635584000000$
možnosti z direktnim preizkušanjem,
zato v članku dobimo naslednje nasvete:

(0) Relativna frekvenca črk in presledkov v slovenščini: presledek 173,

E	A	I	O	N	R	S	L	J	T	V	D	
89	84	74	73	57	44	43	39	37	37	33	30	
K	M	P	U	Z	B	G	"C	H	"S	C	"Z	F
29	27	26	18	17	15	12	12	9	9	6	6	1

(1) Na začetku besed so najpogostejše črke
N, S, K, T, J, L.

(2) Najpogostejše končnice pa so
E, A, I, O, U, R, N.

(3) Ugotovi, kateri znaki zagotovo predstavljajo samoglasnike in kateri soglasnike.

(4) V vsaki besedi je vsaj en samoglasnik ali samoglasniški R.

(5) V vsaki besedi z dvema črkama je ena črka samoglasnik, druga pa soglasnik.

(6) detektivska sreča

(0)	V	-	C	D	J	?	H	W	O	(	+	3
	23	19	16	12	11	10	9	7	6	6	5	4

Y	4	!	/	Q	:	%	T	N	S	G
4	3	3	2	2	2	2	2	2	1	1

Zaključek `V --> ' '` (drugi znaki z visoko frekvenco ne morejo biti).

Dve besedi se ponovita: `03CWC%J(-`,
opazimo pa tudi eno sklanjatev:
`D-?+-` ter `D-?+C`.

Torej nadaljujemo z naslednjim tekstom:

```
YHW?HD+C ODH TH O-!J G:CDCYJ(J /- ?H
(-T?H W-4YD4(?-DJ /-(?S- 03CWC%J(- 4-DC
!CW-?C NJDJ D-?+- 03CWC%J(- QW-DQ- J+
?H DWHN- 3C:C0DC !H+?-DJ D-?+C 3J0-YC
```

(3) Kanditati za samoglasnike e,a,i,o so znaki z visokimi frekvancami. Vzamemo:

$$\{\text{e,a,i,o}\} = \{\text{-,C,J,H}\}$$

(saj D izključi -,H,J,C in ? izključi -,H,C,
znaki -,C,J,H pa se ne izključujejo)

Razporeditev teh znakov kot samoglasnikov izgleda prav verjetna. To potrdi tudi gostota končnic, gostota parov je namreč:

AV	CV	HV	JV	VO	?H	-D	DC	JM	W-	DJ	UC	CW	-?	VD
7	5	5	5	4	4	4	3	3	3	3	3	3	3	3

(5) Preučimo besede z dvema črkama:

```
    Samoglasnik na koncu
1)  da ga na pa ta za (ha ja la)
2)  "ce je le me ne se "se te ve "ze (he)
3)  bi ji ki mi ni si ti vi
4)  bo do (ho) jo ko no po so to
5)  ju mu tu (bu)
6)  r"z rt
```

```
   Samoglasnik na za"cetku
1) ar as                 (ah aj au)
2) en ep                  (ej eh)
3) in iz ig
4) on ob od os on         (oh oj)
5) uk up u"s ud um ur (uh ut)
```

in opazujemo besedi: /- ?H

ter besedi: J+ ?H.

J+ ima najmanj možnosti, + pa verjetno ni črka n, zato nam ostane samo še:

```
J+ ?H          DWHN-
/- ?H
iz te         (ne gre zaradi: D-?+C)
ob ta(e,o)    (ne gre zaradi: D-?+C)
od te         (ne gre zaradi: D-?+C)
```

tako da bo potrebno nekaj spremeniti in preizkusiti še naslednje:
on bo; on jo; in so; in se; in je; in ta; en je; od tu ...

(6) Če nam po dolgem premisleku ne uspe najti rdeče niti, bo morda potrebno iskati napako s prijatelji (tudi računalniški program z metodo lokalne optimizacije ni zmogel problema zaradi premajhne dolžine tajnopisa, vsekakor pa bi bilo problem mogoče rešiti s pomočjo elektronskega slovarja).

Tudi psihološki pristop pomaga, je svetoval Martin

Juvan in naloga je bila rešena (poskusite sami!).

Podobna naloga je v angleščini dosti lažja, saj je v tem jeziku veliko členov THE, A in AN, vendar pa zato običajno najprej izpustimo presledke iz teksta, ki ga želimo spraviti v tajnopis. V angleščini imajo seveda črke drugačno gostoto kot v slovenščini. Razdelimo jih v naslednjih pet skupin:

1. E, z verjetnostjo okoli 0.120,
2. T, A, O, I, N, S, H, R, vse z verjetnostjo med 0.06 in 0.09,
3. D, L, obe z verjetnostjo okoli 0.04,
4. C, U, M, W, F, G, Y, P, B, vse z verjetnostjo med 0.015 in 0.028,
5. V, K, J, X, Q, Z, vse z verjetnostjo manjšo od 0.01.

Najbolj pogosti pari so (v padajočem zaporedju): TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI in OF,

Najbolj pogoste trojice pa so (v padajočem zaporedju): THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR in DTH.

Začetki verjetnosti

Ahil in Ajaks kockata, amfora, okrog 530 pr.n.š, Eksekias, Vatikan

Naključnost so poznale že stare kulture: Egipčani, Grki, ...a je niso poskušale razumeti – razlagale so jo kot voljo bogov.

Leta 1662 je plemič Chevalier de Mere zastavil matematiku Pascalu vprašanje zakaj določene stave prinašajo dobiček druge pa ne. Pascal si je o tem začel dopisovati s Fermatom in iz tega so nastali začetki verjetnostnega računa.

Prvo tovrstno razpravo je napisal že leta 1545 italijanski kockar in matematik Cardano, a ni bila širše znana.

Tudi leta 1662 je anglež John Graunt sestavil na osnovi podatkov prve zavarovalniške tabele.

... začetki

Leta 1713 je Jakob Bernoulli objavil svojo *Umetnost ugibanja* s katero je verjetnostni račun postal resna in splošno uporabna veda. Njegov pomen je še utrdil Laplace, ko je pokazal njegov pomen pri analizi astronomskih podatkov (1812).

Leta 1865 je avstrijski menih Gregor Mendel uporabil verjetnostno analizo pri razlagi dednosti v genetiki.

V 20. stoletju se je uporaba verjetnostnih pristopov razširila skoraj na vsa področja.

Poskusi, dogodki in verjetnost

Verjetnostni račun obravnava zakonitosti, ki se pokažejo v velikih množicah enakih ali vsaj zelo podobnih pojavov. Predmet verjetnostnega računa je torej empirične narave in njegovi osnovni pojmi so povzeti iz izkušnje. Osnovni pojmi v verjetnostnem računu so: poskus, dogodek in verjetnost dogodka.

Poskus je realizacija neke množice skupaj nastopajočih dejstev (kompleksa pogojev). Poskus je torej vsako dejanje, ki ga opravimo v natanko določenih pogojih.

Primeri:

- met igralne kocke,
- iz kupa 20 igralnih kart izberemo eno karto.

Dogodki

Pojav, ki v množico skupaj nastopajočih dejstev ne spada in se lahko v posameznem poskusu zgodi ali pa ne, imenujemo *dogodek*.

Primeri:

- v poskusu meta igralne kocke je na primer dogodek, da vržemo 6 pik;
- v poskusu, da vlečemo igralno karto iz kupa 20 kart, je dogodek, da izvlečemo rdečo barvo.

Za poskuse bomo privzeli, da jih lahko neomejeno velikokrat ponovimo. Dogodki se bodo nanašali na isti poskus.

Poskuse označujemo z velikimi črkami iz konca abecede, npr. X, Y, X_1. Dogodke pa označujemo z velikimi črkami iz začetka abecede, npr. A, C, E_1.

Vrste dogodkov

Dogodek je lahko:

- **gotov** dogodek – G: ob vsaki ponovitvi poskusa se zgodi.

  Primer: dogodek, da vržemo 1, 2, 3, 4, 5, ali 6 pik pri metu igralne kocke;

- **nemogoč** dogodek – N: nikoli se ne zgodi.

  Primer: dogodek, da vržemo 7 pik pri metu igralne kocke;

- **slučajen** dogodek: včasih se zgodi, včasih ne.

  Primer: dogodek, da vržemo 6 pik pri metu igralne kocke.

Računanje z dogodki

Dogodek A je *poddogodek* ali **način** dogodka B, kar zapišemo $A \subset B$, če se vsakič, ko se zgodi dogodek A, zagotovo zgodi tudi dogodek B.

Primer: Pri metu kocke je dogodek A, da pade šest pik, način dogodka B, da pade sodo število pik.

Če je dogodek A način dogodka B in sočasno dogodek B način dogodka A, sta dogodka *enaka*: $A \subset B \wedge B \subset A \iff A = B$.

Vsota dogodkov A in B, označimo jo z $A \cup B$ ali $A + B$, se zgodi, če se zgodi **vsaj** eden od dogodkov A in B.

Primer: Vsota dogodka A, da vržemo sodo število pik, in dogodka B, da vržemo liho število pik, je gotov dogodek.

Velja: $A \cup B = B \cup A$; $A \cup N = A$; $A \cup G = G$; $A \cup A = A$

$$B \subset A \iff A \cup B = A\,;\; A \cup (B \cup C) = (A \cup B) \cup C$$

...Računanje z dogodki

Produkt dogodkov A in B, označimo ga z $A \cap B$ ali AB, se zgodi, če se zgodita A in B **hkrati**.

Primer: Produkt dogodka A, da vržemo sodo število pik, in dogodka B, da vržemo liho število pik, je nemogoč dogodek.

Velja: $A \cap B = B \cap A$; $A \cap N = N$; $A \cap G = A$; $A \cap A = A$

$$B \subset A \iff A \cap B = B \;;\; A \cap (B \cap C) = (A \cap B) \cap C$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \;;\; A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Dogodku A **nasproten** dogodek $\overline{A}$ imenujemo negacijo dogodka A.

Primer: Nasproten dogodek dogodku, da vržemo sodo število pik, je dogodek, da vržemo liho število pik.

Velja: $A \cap \overline{A} = N$; $A \cup \overline{A} = G$; $\overline{N} = G$; $\overline{\overline{A}} = A$

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \;;\; \overline{A \cap B} = \overline{A} \cup \overline{B}$$

...Računanje z dogodki

Dogodka A in B sta **nezdružljiva**, če se ne moreta zgoditi hkrati, njun produkt je torej nemogoč dogodek, $A \cap B = N$.

Primer: Dogodka, A – da pri metu kocke pade sodo število pik in B – da pade liho število pik, sta nezdružljiva.

Poljuben dogodek in njegov nasprotni dogodek sta vedno nezdružljiva. Ob vsaki ponovitvi poskusa se zagotovo zgodi eden od njiju, zato je njuna vosta gotov dogodek: $A \cap \overline{A} = N \ \wedge \ A \cup \overline{A} = G$.

Če lahko dogodek A izrazimo kot vsoto nezdružljivih in mogočih dogodkov, rečemo, da je A *sestavljen* dogodek. Dogodek, ki ni sestavljen, imenujemo *osnoven* ali *elementaren* dogodek.

Primer: Pri metu kocke je šest osnovnih dogodkov: E_1, da pade 1 pika, E_2, da padeta 2 piki, ... , E_6, da pade 6 pik. Dogodek, da pade sodo število pik je sestavljen dogodek iz treh osnovnih dogodkov (E_2, E_4 in E_6).

...Računanje z dogodki

Množico dogodkov $S = \{A_1, A_2, \ldots, A_n\}$ imenujemo **popoln sistem dogodkov**, če se v vsaki ponovitvi poskusa zgodi natanko eden od dogodkov iz množice S.

To pomeni, da so vsi mogoči

$$A_i \neq N,$$

paroma nezdružljivi

$$A_i \cap A_j = N \quad i \neq j$$

in njihova vsota je gotov dogodek

$$A_1 \cup A_2 \cup \ldots \cup A_n = G.$$

Primer: Popoln sistem dogodkov pri metu kocke sestavljajo na primer osnovni dogodki ali pa tudi dva dogodka: dogodek, da vržem sodo število pik, in dogodek, da vržem liho število pik.

Verjetnost

Opišimo najpreprostejšo verjetnostno zakonitost. Denimo, da smo n–krat ponovili dan poskus in da se je k–krat zgodil dogodek A. Ponovitve poskusa, v katerih se A zgodi, imenujemo ugodne za dogodek A, število

$$f(A) = \frac{k}{n}$$

pa je *relativna frekvenca* (pogostost) dogodka A v opravljenih poskusih.

Statistični zakon, ki ga kaže izkušnja, je:

> *Če poskus X dolgo ponavljamo, se relativna frekvenca slučajnega dogodka ustali in sicer skoraj zmeraj toliko bolj, kolikor več ponovitev poskusa napravimo.*

Statistična definicija verjetnost

To temeljno zakonitost so empirično preverjali na več načinov. Najbolj znan je poskus s kovanci, kjer so določali relativno frekvenco grba ($f(A)$):

- Buffon je v 4040 metih dobil $f(A) = 0,5069$,
- Pearson je v 12000 metih dobil $f(A) = 0,5016$,
- Pearson je v 24000 metih dobil $f(A) = 0,5005$.

Ti poskusi kažejo, da se relativna frekvenca grba pri metih kovanca običajno ustali blizu 0.5. Ker tudi drugi poskusi kažejo, da je ustalitev relativne frekvence v dovolj velikem številu ponovitev poskusa splošna zakonitost, je smiselna naslednja *statistična definicija verjetnosti*:

> *Verjetnost dogodka A v danem poskusu je število $P(A)$, pri katerem se navadno ustali relativna frekvenca dogodka A v velikem številu ponovitev tega poskusa.*

Osnovne lastnosti verjetnosti

1. Ker je relativna frekvenca vedno nenegativna, je verjetnost $P(A) \geq 0$.
2. $P(G) = 1$, $P(N) = 0$ in $A \subset B \Rightarrow P(A) \leq P(B)$
3. Naj bosta dogodka A in B nezdružljiva. Tedaj velja

$$P(A \cup B) = P(A) + P(B)$$

Klasična definicija verjetnosti

Pri določitvi verjetnosti si pri nekaterih poskusih in dogodkih lahko pomagamo s *klasično definicijo verjetnosti*:

> *Vzemimo, da so dogodki iz popolnega sistema dogodkov* $\{E_1, E_2, \ldots, E_s\}$ *enako verjetni:*
> $P(E_1) = P(E_2) = \ldots = P(E_s) = p$.
> *Tedaj je* $P(E_i) = 1/s \quad i = 1, \ldots, s$.
> *Če je nek dogodek* A *sestavljen iz* r *dogodkov iz tega popolnega sistema dogodkov, potem je njegova verjetnost* $P(A) = r/s$.

Primer: Izračunajmo verjetnost dogodka A, da pri metu kocke padejo manj kot 3 pike.
Popolni sistem enako verjetnih dogodkov sestavlja 6 dogodkov.
Od teh sta le dva ugodna za dogodek A (1 in 2 piki).

Zato je verjetnost dogodka A enaka $\dfrac{2}{6} = \dfrac{1}{3}$.

Geometrijska verjetnost

V primerih, ko lahko osnovne dogodke predstavimo kot ‘enakovredne’ točke na delu premice (ravnine ali prostora),
določimo verjetnost sestavljenega dogodka kot razmerje dolžin (ploščin, prostornin) dela, ki ustreza ugodnim izidom,
in dela, ki ustreza vsem možnim izidom.

Primer:

Še dve lastnosti verjetnosti

4. Za dogodka A in B velja:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Primer: Denimo, da je verjetnost, da študent naredi izpit iz Sociologije $P(S) = 2/3$. Verjetnost, da naredi izpit iz Politologije je $P(P) = 5/9$. Če je verjetnost, da naredi vsaj enega od obeh izpitov $P(S \cup P) = 4/5$, kolikšna je verjetnost, da naredi oba izpita?

$$P(S \cap P) = P(S) + P(P) - P(S \cup P) =$$

$$= \frac{2}{3} + \frac{5}{9} - \frac{4}{5} = 0,42$$

Za dogodke A, B in C velja:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$

Kako lahko to pravilo posplošimo še na več dogodkov?

Namig: Pravilo o vključitvi in izključitvi za množice $A_1, A_2, \ldots, A_n$:

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{i=1}^{n} |A_i| - \sum_{1 \le i_1 < i_2 \le n}^{n} |A_{i_1} \cap A_{i_2}|$$

$$+ \sum_{1 \le i_1 < i_2 < i_3 \le n}^{n} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| - \cdots + (-1)^{n-1} |A_1 \cap A_2 \cap \ldots \cap A_n|$$

...Še dve lastnosti verjetnosti

5. $P(\overline{A}) = 1 - P(A)$

Primer: Iz kupa 32 kart slučajno povlečemo 3 karte. Kolikšna je verjetnost, da je med tremi kartami vsaj en as (dogodek A)?

Pomagamo si z nasprotnim dogodkom. Nasprotni dogodek $\overline{A}$ dogodka A je, da med tremi kartami ni asa. Njegova verjetnost po klasični definiciji verjetnosti je določena s kvocientom števila vseh ugodnih dogodkov v popolnem sistemu dogodkov s številom vseh dogodkov v tem sistemu dogodkov. Vseh dogodkov v popolnem sistemu dogodkov je $\binom{32}{3}$, ugodni pa so tisti, kjer zbiramo med ne–asi, t.j. $\binom{28}{3}$. Torej je

$$P(\overline{A}) = \frac{\binom{28}{3}}{\binom{32}{3}} = 0,66\ ;\ P(A) = 1 - P(\overline{A}) = 1 - 0,66 = 0,34.$$

...Še dve lastnosti verjetnosti

Posledica. Če so dogodki A_i, $i \in I$ paroma nezdružljivi, velja

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

Velja tudi za neskončne množice dogodkov.

Aksiomi Kolmogorova

Dogodek predstavimo z množico zanj ugodnih izidov; gotov dogodek G ustreza univerzalni množici; nemogoč dogodek pa prazni množici.

Neprazna družina dogodkov $\mathcal{D}$ je *algebra*, če velja:

- $A \in \mathcal{D} \Rightarrow \overline{A} \in \mathcal{D}$
- $A, B \in \mathcal{D} \Rightarrow A \cup B \in \mathcal{D}$

Pri neskončnih množicah dogodkov moramo drugo zahtevo posplošiti

- $A_i \in \mathcal{D}, i \in I \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{D}$

Dobljeni strukturi rečemo *σ-algebra*.

. . . Aksiomi Kolmogorova

Naj bo $\mathcal{D}$ σ-algebra v G. *Verjetnost na G* je preslikava $P : \mathcal{D} \to \mathbb{R}$ z lastnostmi:

1. $P(A) \geq 0$
2. $P(G) = 1$
3. Če so dogodki A_i, $i \in I$ paroma nezdružljivi, je

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

Trojica $(G, \mathcal{D}, P)$ določa *verjetnostni prostor*.

Iz teh treh aksiomov lahko izpeljemo vse ostale lastnosti verjetnosti (Hladnik, str. 12).

Pogojna verjetnost

Opazujemo dogodek A ob poskusu X, ki je realizacija kompleksa pogojev K. Verjetnost dogodka A je tedaj $P(A)$.

Kompleksu pogojev K pridružimo mogoč dogodek B, tj. $P(B) > 0$. Realizacija tega kompleksa pogojev $K' = K \cap B$ je poskus X' in verjetnost dogodka A v tem poskusu je $P_B(A)$, ki se z verjetnostjo $P(A)$ ujema ali pa ne.

Pravimo, da je poskus X' poskus X s pogojem B in verjetnost $P_B(A)$ **pogojna verjetnost** dogodka A glede na dogodek B, kar zapišemo takole:

$$P_B(A) = P(A/B)$$

Pogojna verjetnost $P(A/B)$ v poskusu X' je verjetnost dogodka A v poskusu X s pogojem B.

Pogosto pogojno verjetnost pišejo tudi $P(A/B)$.

…Pogojna verjetnost

Denimo, da smo n–krat ponovili poskus X in da se je ob tem k_B–krat zgodil dogodek B. To pomeni, da smo v n ponovitvah poskusa X napravili k_B–krat poskus X'. Dogodek A se je zgodil ob poskusu X' le, če se je zgodil tudi B, t.j. $A \cap B$. Denimo, da se je dogodek $A \cap B$ zgodil ob ponovitvi poskusa $k_{A\cap B}$–krat. Potem je relativna frekvenca dogodka A v opravljenih ponovitvah poskusa X':

$$f_B(A) = f(A/B) = \frac{k_{A\cap B}}{k_B} = \frac{k_{A\cap B}/n}{k_B/n} = \frac{f(A\cap B)}{f(B)}$$

oziroma

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

Pogojna verjetnost P_B ima prav take lastnosti kot brezpogojna. Trojica $(B, \mathcal{D}_B, P_B)$, $\mathcal{D}_B = \{A\cap B \mid A \in \mathcal{D}\}$ je zopet verjetnostni prostor.

...Pogojna verjetnost

Primer: Denimo, da je v nekem naselju 900 polnoletnih prebivalcev. Zanima nas struktura prebivalcev po spolu (M – moški, Ž – ženski spol) in po zaposlenosti (Z – zaposlen(a), N – nezaposlen(a)). Podatke po obeh spremenljivkah uredimo v dvorazsežno frekvenčno porazdelitev, ki jo imenujemo tudi *kontingenčna tabela*:

$spol \setminus zap.$	Z	N	
M	460	40	500
Ž	240	160	400
	700	200	900

…Pogojna verjetnost

Poglejmo, kolikšna je verjetnost, da bo slučajno izbrana oseba moški pri pogoju, da je zaposlena.

$$P(Z) = \frac{700}{900} \quad , \quad P(M \cap Z) = \frac{460}{900}$$

$$P(M/Z) = \frac{P(M \cap Z)}{P(Z)} = \frac{460 \cdot 900}{900 \cdot 700} = \frac{460}{700}$$

ali neposredno iz kontingenčne tabele

$$P(M/Z) = \frac{460}{700}.$$

...Pogojna verjetnost

Iz formule za pogojno verjetnost sledi:

$$P(A \cap B) = P(B)P(A/B),$$

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

Torej velja:

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

Dogodka A in B sta **neodvisna**, če velja

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

Zato za neodvisna dogodka A in B velja $P(A \cap B) = P(A) \cdot P(B)$.

Za nezdružljiva dogodka A in B velja $P(A/B) = 0$.

...Pogojna verjetnost

Primer: Iz posode, v kateri imamo 8 belih in 2 rdeči krogli, dvakrat na slepo izberemo po eno kroglo. Kolikšna je verjetnost dogodka, da je prva krogla bela (B_1) in druga rdeča (R_2).

1. Če po prvem izbiranju izvlečeno kroglo ne vrnemo v posodo (odvisnost), je:

$$P(B_1 \cap R_2) = P(B_1) \cdot P(R_2/B_1) =$$

$$= \frac{8}{10} \cdot \frac{2}{9} = 0,18$$

2. Če po prvem izbiranju izvlečeno kroglo vrnemo v posodo (neodvisnost), je:

$$P(B_1 \cap R_2) = P(B_1) \cdot P(R_2/B_1) =$$

$$= P(B_1) \cdot P(R_2) = \frac{8}{10} \cdot \frac{2}{10} = 0,16.$$

. . . Pogojna verjetnost

Dogodka A in B sta neodvisna, če je $P(A/B) = P(A/\overline{B})$.

$$P(A \cap B \cap C) = P(A) \cdot P(B/A) \cdot P(C/(A \cap B))$$

(Tudi to pravilo lahko posplošimo naprej.)

Dogodki A_i, $i \in I$ so *neodvisni*, če je $P(A_j) = P(A_j / \bigcap_{i=1}^{j-1} A_i)$, $j \in I$.

Za neodvisne dogodke A_i, $i \in I$ velja

$$P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i).$$

Obrazec za razbitja in večstopenjski poskusi

Naj bo H_i, $i \in I$ *razbitje* gotovega dogodka: $\bigcup_{i\in I} H_i = G$ in dogodki so paroma nezdružljivi $H_i \cap H_j = N, i \neq j$.

Zanima nas verjetnost dogodka A, če poznamo verjetnost $P(H_i)$, in pogojno verjetnost $P(A/H_i)$ za $i \in I$:

$$A = A \cap (H_1 \cup H_2 \cdots H_n) = (A \cap H_1) \cup \cdots \cup (A \cap H_n).$$

Ker so tudi dogodki $A \cap H_i$ paroma nezdružljivi, velja:

$$P(A) = \sum_{i\in I} P(A \cap H_i) = \sum_{i\in I} P(H_i) P(A/H_i).$$

Na stvar lahko pogledamo tudi kot na večstopenjski poskus: v prvem koraku se zgodi natanko eden od dogodkov H_i, ki ga imenujemo hipoteza (hipoteze sestavljajo popolen sistem dogodkov). Šele izidi na prejšnjih stopnjah določajo, kako bo potekal poskus na naslednji stopnji. Omejimo se na poskus z dvema stopnjama.

Naj bo A eden izmed mogočih dogodkov na drugi stopnji. Včasih nas zanima po uspešnem izhodu tudi druge stopnje, verjetnost tega, da se je na prvi stopnji zgodil dogodek H_i.

Odgovor dobimo iz zgornjega obrazca in mu pravimo *Bayesov obrazec*:

$$P(H_k/A) = \frac{P(H_k) \cdot P(A/H_k)}{\sum_{i \in I} P(H_i) \cdot P(A/H_i)}$$

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	0
H2	0	1	1	,8*,4*,6	=0,192	2	0	0
H3	1	0	1	,2*,6*,6	=0,072	2	0	0
H4	1	1	0	,2*,4*,4	=0,032	2	0	0
H5	1	0	0	,2*,6*,4	=0,048	1	1	0,048
H6	0	1	0	,8*,4*,4	=0,128	1	1	0,128
H7	0	0	1	,8*,6*,6	=0,288	1	1	0,288
H8	0	0	0	,8*,6*,4	=0,192	0	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:

$$\begin{aligned}
P(\text{nobena krogla ni zadela}) &= 0,192 = P(N'),\\
P(\text{dve krogli sta zadeli}) = 0,192 + 0,072 + 0,032 &= 0,296 = P(D),\\
P(\text{tri krogle so zadele}) = 0,048 + 0,128 + 0,288 &= 0,048 = P(T).
\end{aligned}$$

Vsota teh verjetnosti je seveda enaka 1.

Končno uporabimo Bayesov obrazec:

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

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

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

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).