2.4. BAYESOVA FORMULA.

FORMULA POTPUNE VJEROJATNOSTI. BAYESOVA FORMULA.

BERNOULLIJEVA SHEMA.

2.4.1. POTPUNI SUSTAV DOGAĐAJA

- Neka je $(\mathbf{\Omega}, \mathcal{F}, P)$ vjerojatnosni prostor. (Možemo pretpostaviti da je $\mathcal{F} = \mathcal{P}(\mathbf{\Omega})$.)
- Potpuni sustav događaja je *konačan* skup $\boldsymbol{A} = \{A_1, \ldots, A_n\} \subseteq \mathcal{F}$ takav da istodobno vrijedi:

$$\mathbf{1.)}\ (i \neq j) \Rightarrow (A_i \cdot A_j = \emptyset);$$

$$\mathbf{2.)}\ \sum_{i=1}^{n} A_i = \Omega;$$

$$\mathbf{3.)} \sum_{i=1}^{n} P(A_i) = 1.$$

2.4.2. PRIMJER 1.

- Promatramo slučajni pokus *izvlačenje prvoga broja u jednom kolu igre* LOTO 7/35.
- Svih 35 elementarnih događaja (navedite ih!) tvore jedan potpuni sustav događaja.
- Takav je sustav tehnički malo zahtjevnije ispisati, pa zbog toga tražimo "manje" potpune sustave događaja, tj. skupove A sa što manjim kardinalitetom.

2.4.3. FORMULA POTPUNE VJEROJATNOSTI

- Neka su $(\mathbf{\Omega}, \mathcal{F}, P)$ vjerojatnosni prostor, $\mathcal{A} = \{A_1, \ldots, A_n\}$ potpuni sustav događaja iz $\mathcal{F}$ i $B \in \mathcal{F}$.
- Promotrimo presjeke događaja B sa *svakim* elementom sustava $\mathcal{A}$. To su događaji $B_1 := B \cdot A_1, \ldots, B_n := B \cdot A_n$.
- Događaji $B_1, \ldots, B_n$ su međusobno isključivi (presjek im je prazan skup) jer bi u suprotnom neka dva elementa sustava $\mathcal{A}$ imala neprazan presjek, što je nemoguće prema definiciji potpunoga sustava događaja.
- Unija svih događaja $B_1, \ldots, B_n$ je događaj B.
- Tako smo događaj B rastavili na *konačnu* uniju međusobno isključivih događaja. (Ta unija je konačna jer je $\mathcal{A}$ konačan skup prema definiciji potpunoga sustava događaja.)

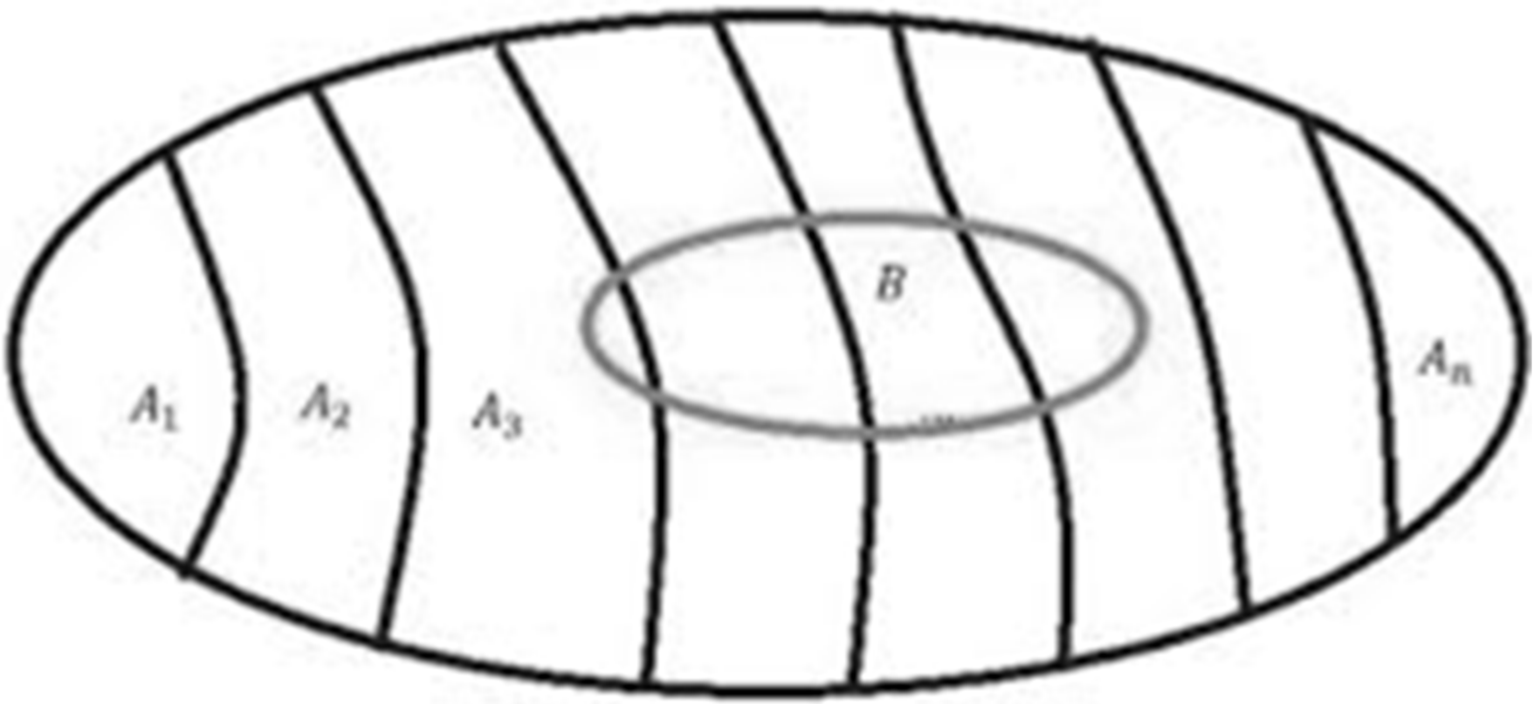
2.4.3. FORMULA POTPUNE VJEROJATNOSTI

- Tako je $P(B) = \sum_{i=1}^{n} P(B \cdot A_i)$.
- Prema formuli za uvjetnu vjerojatnost je:
- $$P(B \cdot A_i) = P(A_i) \cdot P(B \mid A_i)$$
- Uvrštavanjem u izraz za $P(B)$ dobivamo tzv. **formulu potpune vjerojatnosti**:

$$P(B) = \sum_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)$$

- Formulu potpune vjerojatnosti podesno je primijeniti u slučajevima kad elementi potpunoga sustava povlače neki događaj iz algebre događaja (i kad se mogu izračunati ili kad su unaprijed zadane sve pripadne uvjetne vjerojatnosti).

2.4.3. FORMULA POTPUNE VJEROJATNOSTI



2.4.4. BAYESOVA FORMULA

- Dosad smo razmatrali određivanje vjerojatnosti događaja $A \in \mathcal{F}$ u slučaju kad je taj događaj bio posljedica svake hipoteze iz zadanoga potpunoga sustava događaja.
- Sada ćemo promotriti obrnut problem: kako odrediti vjerojatnost da je $A \in \mathcal{F}$ uzrokovao neku hipotezu iz zadanoga potpunoga sustava događaja?
- Neka su $(\mathbf{\Omega}, \mathcal{F}, P)$ vjerojatnosni prostor, $\mathcal{A} = \{A_1, \ldots, A_n\}$ potpuni sustav događaja iz $\mathcal{F}$ i $B \in \mathcal{F}$.
- Prema formuli uvjetne vjerojatnosti vrijedi:

$$P(A_i \mid B) = \frac{P(A_i \cdot B)}{P(B)} = \frac{P(A_i) \cdot P(B \mid A_i)}{P(B)}, \text{ za svaki } i \in [n].$$

2.4.4. BAYESOVA FORMULA

- Prema formuli potpune vjerojatnosti vrijedi:

$$P(B)=\sum_{i=1}^{n} P(A_i)\cdot P(B\mid A_i)$$

- Uvrstimo li ovu jednakost u prethodnu, dobivamo tzv. Bayesovu formulu:

$$P(A_i\mid B)=\frac{P(A_i\cdot B)}{P(B)}=\frac{P(A_i)\cdot P(B\mid A_i)}{\sum_{j=1}^{n} P(A_j)\cdot P(B\mid A_j)},\ \text{za svaki } i\in[n].$$

2.4.5. NAPOMENA

- Elemente potpunoga sustava događaja nazivamo **hipoteze**. Dakle, skup $\boldsymbol{\Omega}$ možemo jednoznačno zadati i navođenjem svih hipoteza.
- **Ponovimo:** formula potpune vjerojatnosti opisuje računanje vjerojatnosti događaja A ako znamo vjerojatnost svake hipoteze i (uvjetnu) vjerojatnost događaja A ako se dogodi svaka hipoteza.
- Bayesova formula opisuje računanje uvjetne vjerojatnosti *hipoteze* uz uvjet da se dogodio neki događaj iz algebre događaja.
- Za njezinu primjenu najprije moramo primijeniti formulu potpune vjerojatnosti. (Obrat ne vrijedi.)

2.4.6. BERNOULLIJEVA SHEMA

- Pretpostavimo da neki slučajan pokus s *točno dva* moguća ishoda (npr. ispitivanje je li neki uređaj ispravan) slučajno i nezavisno ponavljamo ukupno n puta. (*Nezavisno* znači da ishod i-te izvedbe pokusa ne ovisi ni o jednoj drugoj izvedbi toga pokusa.)
- Odaberimo točno jedan ishod pokusa i nazovimo ga *uspjeh*. Preostali ishod pokusa nazovimo *neuspjeh*.
- Neka je p vjerojatnost pojave *uspjeha* u jednom slučajnom pokusu. Tada je vjerojatnost pojave *neuspjeha* u istom slučajnom pokusu jednaka $q := 1 - p$.
- Ovako dobivenu shemu nazivamo Bernoullijeva shema.

2.4.6. BERNOULLIJEVA SHEMA

- Vjerojatnost da se u *n* izvedbi slučajnoga pokusa dogodi točno *k uspjeha* jednaka je
- $$p_k = \binom{n}{k} \cdot p^k \cdot q^{n-k} = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$
- Ovu je vjerojatnost za velike vrijednosti varijable *n* vrlo teško izračunati, pa ćemo u drugom dijelu predmeta izvesti formule kojima ćemo aproksimirati te vjerojatnosti (tzv. *granični teoremi u Bernoullijevoj shemi*).
- Zadatke u kojima se primjenjuje Bernoullijeva shema rješavat ćemo u okviru nastavne cjeline o binomnoj razdiobi.

1. Trgovina *Sve po malo kuna* iz Konjskoga Brda nabavlja robu od točno dvaju dobavljača: *Mućak trade* d.o.o. iz Krivoga Puta i *Šalabajzer commerce* d.o.o. iz Lukova Šugarja. Pritom *Mućak trade* d.o.o. dobavlja 40% sve nabavljene robe. Vjerojatnosti da dobavljači (u danom poretku) neće pravovremeno dostaviti robu iznose redom 20% i 10%. Izračunajte vjerojatnost da slučajno izabrana pošiljka robe neće biti pravovremeno dostavljena.

Rješenje: Označimo:

$M :=$ {izabranu pošiljku isporučuje *Mućak trade*},
$\check{S} :=$ {izabranu pošiljku isporučuje *Šalabajzer commerce*},
$A :=$ {izabrana pošiljka nije pravovremeno dostavljena}.

Lako se provjeri da je skup

$$S = \{M, \check{S}\}$$

potpuni sustav događaja. Iz podataka u zadatku slijedi:

$$\begin{aligned} P(M) &= 40\% = \frac{2}{5}, \\ P(\check{S}) &= 1 - P(M) = \\ &= 1 - \frac{2}{5} = \\ &= \frac{3}{5}, \\ P(A \mid M) &= 20\% = \frac{1}{5}, \\ P(A \mid \check{S}) &= 10\% = \frac{1}{10}. \end{aligned}$$

Tražimo $P(A)$. Primijenimo formulu potpune vjerojatnosti, pa dobijemo:

$$\begin{aligned} P(A) &= P(M) \cdot P(A \mid M) + P(\check{S}) \cdot P(A \mid \check{S}) = \\ &= \frac{2}{5} \cdot \frac{1}{5} + \frac{3}{5} \cdot \frac{1}{10} = \\ &= \frac{7}{50} = 0.14. \end{aligned}$$

2. U doigravanje za ulazak u 1. županijsku rukometnu ligu ulaze RK „*Cicvara*" iz Bedekovčine i točno jedna od sljedećih triju momčadi: RK „*Kajgana*" iz Špičkovine, RK „*Ćevap*" iz Luga i RK „*Majburger*" iz Đurmanca. Svaki od navedenih triju protivnika RK "*Cicvara*" je jednakovjerojatan, tj. među njima nema "najvjerojatnijega" drugoga sudionika doigravanja. Statistike dosadašnjih međusobnih ogleda RK "*Cicvara*" sa svakom od njih zadane su u tablici 1.

Momčad	*Ukupan broj susreta s RK „Cicvara"*	*Pobjeda RK „Cicvara"*	*Pobjeda momčadi*	*Broj remija*
RK „*Kajgana*"	11	5	5	1
RK „*Ćevap*"	12	6	6	0
RK „*Majburger*"	13	7	4	2

Tablica 1.

Isplati li se kladiti da će prva utakmica doigravanja završiti pobjedom RK "*Cicvara*"? (Klađenje na neki događaj isplati se ako je vjerojatnost toga događaja strogo veća od 0.5.) Objasnite svoj odgovor.

Rješenje: Označimo:

K := {u prvoj utakmici doigravanja igra RK „*Kajgana*"},
$Ć$:= {u prvoj utakmici doigravanja igra RK „*Ćevap*"},
M := {u prvoj utakmici doigravanja igra RK „*Majburger*"}.

Lako se provjeri da je

$$\mathcal{A} = \left\{Ć, K, M\right\}$$

potpuni sustav događaja. Zbog ravnopravnosti svih mogućih protivnika RK „*Cicvara*" vrijede jednakosti:

$$P(Ć) = P(K) = P(M) = \frac{1}{3}.$$

Nadalje, neka je

B = {u prvoj utakmici doigravanja pobijedila je RK „*Cicvara*"}.

Iz tablice 1. proizlazi:

$$P(B \mid K) = \frac{5}{11},$$
$$P(B \mid Ć) = \frac{6}{12} = \frac{1}{2},$$

$$P(B \mid M) = \frac{7}{13}.$$

Primjenom formule potpune vjerojatnosti slijedi da je tražena vjerojatnost jednaka:

$$P(B) = P(K) \cdot P(B \mid K) + P(Ć) \cdot P(B \mid Ć) + P(M) \cdot P(B \mid M) =$$
$$= \frac{1}{3} \cdot \frac{5}{11} + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{7}{13} = \frac{427}{858} < \frac{429}{858} = \frac{1}{2}.$$

Dakle, ne isplati se kladiti u pobjedu RK „*Cicvara*“.

3. Na izvanrednom ispitnom roku moguće je polagati točno jedan ispit. Poznato je da 60% studenata bira *Osnove elektrotehnike*, 25% studenata *Matematiku*, a svi ostali studenti *Fiziku*. Prolaznost iz *Osnova elektrotehnike* iznosi 20%, iz *Matematike* 32%, a iz *Fizike* 40%. Na slučajan način odabran je točno jedan student koji je polagao ispit na izvanrednom ispitnom roku.

a) Izračunajte vjerojatnost da je izabrani student položio svoj izabrani ispit.

b) Ako znamo da je izabrani student položio svoj izabrani ispit, izračunajte vjerojatnost da je taj student položio *Matematiku*.

Rješenje: Neka su:

H_1 = {izabrani predmet je *Osnove elektrotehnike*},
H_2 = {izabrani predmet je *Matematika*},
H_3 = {izabrani predmet je *Fizika*},
A = {student je položio izabrani predmet}.

Iz podataka u zadatku proizlaze sljedeće vjerojatnosti:

$$\begin{aligned} P(H_1) &= 60\%, & P(A \mid H_1) &= 20\%, \\ P(H_2) &= 25\%, & P(A \mid H_2) &= 32\%, \\ P(H_3) = 1-\big(P(H_1)+P(H_2)\big) = 1-(60\%+25\%) &= 15\%, & P(A \mid H_3) &= 40\%. \end{aligned}$$

a) Treba izračunati vjerojatnost $P(A)$. Primjenom formule potpune vjerojatnosti odmah dobivamo:

$$\begin{aligned} P(A) &= \sum_{i=1}^{3} P(H_i)\cdot P(A \mid H_i) = \\ &= 0.6\cdot 0.2 + 0.25\cdot 0.32 + 0.15\cdot 0.4 = \\ &= 0.12 + 0.08 + 0.06 = \\ &= 0.26 = 26\%. \end{aligned}$$

b) Treba izračunati vjerojatnost $P(H_2 \mid A)$. Primjenom Bayesove formule odmah dobivamo:

$$\begin{aligned} P(H_2 \mid A) &= \frac{P(H_2)\cdot P(A \mid H_2)}{P(A)} = \\ &= \frac{0.25\cdot 0.32}{0.26} = \\ &= \frac{0.08}{0.26} = \frac{4}{13} \approx 0.31 = 31\%. \end{aligned}$$

4. (*problem Montyja Halla*[1]) Natjecatelj u TV-showu treba odabrati točno jedna od triju vrata. Iza točno jednih vrata nalazi se slika novoga automobila, a iza svake od preostalih dviju vrata slika koze. Cilj natjecatelja je odabrati vrata iza kojih se nalazi slika automobila jer tada (i samo tada!) osvaja automobil. Voditelj TV-showa unaprijed zna koja slika se nalazi iza svakih vrata. Nakon što natjecatelj izabere točno jedna vrata, voditelj showa otvara točno jedna od preostalih vrata iza kojih se nalazi slika koze, te nudi natjecatelju promjenu izbora vrata. (Ako može birati vrata iza kojih je slika koze, voditelj će to učiniti na slučajan način.) Isplati li se natjecatelju promijeniti prvotni izbor? Objasnite svoj odgovor.

Rješenje: Bez smanjenja općenitosti možemo pretpostaviti da su vrata označena brojevima 1, 2 i 3, da je natjecatelj izabrao vrata označena brojem 1 i da je voditelj potom otvorio vrata označena brojem 2. (Inače možemo prenumerirati vrata tako da vrijedi ova pretpostavka.)

Neka je

$$A_i = \{\text{slika automobila se nalazi iza vrata označena brojem } i\},\ \forall i \in [3].$$

Tada je $\{A_1, A_2, A_3\}$ potpuni sustav događaja. Očito je

$$P(A_i) = \frac{1}{3},\ \forall i \in [3].$$

Neka je

$$K := \{\text{voditelj je otvorio vrata označena brojem } 2\}.$$

Očito je tada

$$P(A_2 \mid K) = 0$$

jer ako voditelj otvori vrata označena brojem 2, onda iza tih vrata ne može biti slika automobila.

Dakle, trebamo usporediti vjerojatnosti $P(A_1 \mid K)$ i $P(A_3 \mid K)$. U tu svrhu odredimo:

- $P(K \mid A_1)$. Znamo da se slika automobila nalazi iza vrata br. 1. Zbog toga voditelj na slučajan način bira hoće li otvoriti vrata br. 2 ili vrata br. 3, pa je $P(K \mid A_1) = 0.5$.

[1] Monty Hall (1921. – 2017.), američki televizijski voditelj, najpoznatiji po showu *Let's make a Deal*.

- $P(K \mid A_2)$. Ako je iza drugih vrata slika automobila, onda, prema zahtjevu zadatka, voditelj ne smije otvoriti ta vrata. Zbog toga je $P(K \mid A_2) = 0$.

- $P(K \mid A_3)$. Slika automobila se nalazi iza vrata br. 3., a natjecatelj je izabrao vrata br. 1. Zbog toga voditelj ne smije otvoriti ni vrata br. 3 (jer bi otkrio da se iza tih vrata nalazi slika automobila), ni vrata br. 1. Dakle, u ovom slučaju voditelj obavezno mora otvoriti vrata br. 2. Zbog toga je $P(K \mid A_3) = 1$.

Primjenom Bayesove formule dobivamo:

$$\begin{aligned}
P(A_1 \mid K) &= \frac{P(A_1)\cdot P(K \mid A_1)}{P(A_1)\cdot P(K \mid A_1) + P(A_2)\cdot P(K \mid A_2) + P(A_3)\cdot P(K \mid A_3)} = \\
&= \left(\text{zbog } P(A_1) = P(A_2) = P(A_3) = \frac{1}{3} \text{ i } P(K \mid A_2) = 0\right) = \\
&= \frac{P(K \mid A_1)}{P(K \mid A_1) + P(K \mid A_3)} = \\
&= \frac{0.5}{0.5+1} = \\
&= \frac{0.5}{1.5} = \frac{1}{3}, \\
P(A_3 \mid K) &= \frac{P(K \mid A_3)}{P(K \mid A_1) + P(K \mid A_3)} = \\
&= \frac{1}{0.5+1} = \\
&= \frac{1}{1.5} = \\
&= \frac{2}{3} > \frac{1}{3} = P(A_1 \mid K).
\end{aligned}$$

Dakle, isplati se promijeniti prvotni izbor.

Napomena: „Prirodan“ zaključak da se ne isplati promijeniti izbor zbog navodne jednakosti $P(A_1 \mid K) = P(A_3 \mid K) = \frac{1}{2}$ posljedica je **netočne** pretpostavke da voditelj otvara vrata *nezavisno o slici koja se nalazi iza tih vrata*. Preciznije, ovaj slučaj nastupa ako pretpostavimo da voditelj na slučajan način bira vrata različita od vrata koja je izabrao natjecatelj i otvara ih *bez obzira je li iza tih vrata slika automobila ili slika koze*. Dokažite!

5. Čokolada *Slatkica* proizvodi se na ukupno pet strojeva marke A, tri stroja marke B i dva stroja marke C. Svaki stroj marke A proizvede 3% „lom"-čokolade, svaki stroj marke B 2% „lom"-čokolade, a svaki stroj marke C 1% „lom"-čokolade. Na slučajan način biramo jednu proizvedenu čokoladu.

a) Izračunajte vjerojatnost da je izabrana čokolada „lom".

b) Ako znamo da je izabrana čokolada „lom", izračunajte vjerojatnost da je ona proizvedena na stroju marke A. Iskažite dobiveni rezultat u postotcima.

Rješenje: Primijetimo da je ukupan broj strojeva jednak $5+3+2=10$. Neka su:

$$\begin{aligned}
H_1 &= \{\text{izabrana čokolada je proizvedena na stroju } A\},\\
H_2 &= \{\text{izabrana čokolada je proizvedena na stroju } B\},\\
H_3 &= \{\text{izabrana čokolada je proizvedena na stroju } C\},\\
D &= \{\text{izabrana čokolada je "lom"}\}.
\end{aligned}$$

Iz zadanih podataka slijedi:

$$\begin{aligned}
P(H_1) &= \frac{5}{10} = 0.5,\\
P(H_2) &= \frac{3}{10} = 0.3,\\
P(H_3) &= \frac{2}{10} = 0.2,\\
P(D \mid H_1) &= 3\% = 0.03,\\
P(D \mid H_2) &= 2\% = 0.02,\\
P(D \mid H_3) &= 1\% = 0.01.
\end{aligned}$$

a) Tražimo $P(D)$. Primjenom formule potpune vjerojatnosti dobivamo:

$$\begin{aligned}
P(D) &= \sum_{i=1}^{3} P(H_i) \cdot P(D \mid H_i) =\\
&= 0.5 \cdot 0.03 + 0.3 \cdot 0.02 + 0.2 \cdot 0.01 = 0.23 = 23\%.
\end{aligned}$$

b) Tražimo $P(H_1 \mid D)$. Primjenom Bayesove formule dobivamo:

$$\begin{aligned}
P(H_1 \mid D) &= \frac{P(H_1) \cdot P(D \mid H_1)}{P(D)} =\\
&= \frac{0.5 \cdot 0.03}{0.23} = \frac{15}{23} \approx 0.6522 = 65.22\%.
\end{aligned}$$

6. Neki električni stroj se polovicu vremena nalazi u normalnom režimu rada, petinu vremena u otežanom režimu rada, a preostali dio vremena u mirovanju. Vjerojatnost kvara stroja u normalnom režimu jednaka je 2%, u otežanom režimu 6%, a u mirovanju 4‰.

   a) Izračunajte vjerojatnost da je taj stroj pokvaren.

   b) Ako je ustanovljeno da je taj stroj pokvaren, izračunajte vjerojatnost da se kvar dogodio u otežanom režimu rada.

Rješenje: Neka su:

$$\begin{aligned}
H_1 &= \{\text{stroj se nalazi u normalnom režimu rada}\},\\
H_2 &= \{\text{stroj se nalazi u otežanom režimu rada}\},\\
H_3 &= \{\text{stroj se nalazi u mirovanju}\},\\
A &= \{\text{stroj se pokvario}\}.
\end{aligned}$$

Iz zadanih podataka slijedi:

$$\begin{aligned}
P(H_1) &= \frac{1}{2} = 0.5,\\
P(H_2) &= \frac{1}{5} = 0.2,\\
P(H_3) &= 1 - \left(\frac{1}{2} + \frac{1}{5}\right) =\\
&= \frac{3}{10} = 0.3,\\
P(A \mid H_1) &= 2\% = 0.02,\\
P(A \mid H_2) &= 6\% = 0.06,\\
P(A \mid H_3) &= 4‰ = 0.004.
\end{aligned}$$

a) Tražimo $P(A)$. Primjenom formule potpune vjerojatnosti dobivamo:

$$\begin{aligned}
P(A) &= \sum_{i=1}^{3} P(H_i) \cdot P(A \mid H_i) =\\
&= 0.5 \cdot 0.02 + 0.2 \cdot 0.06 + 0.3 \cdot 0.004 =\\
&= 0.0232 = 2.32\%.
\end{aligned}$$

b) Tražimo $P(H_2 \mid A)$. Primjenom Bayesove formule dobivamo:

$$
\begin{aligned}
P(H_2 \mid A) &= \frac{P(H_2) \cdot P(A \mid H_2)}{P(A)} = \\
&= \frac{0.2 \cdot 0.06}{0.0232} = \\
&= \frac{15}{29} \approx 0.5172 = 51.72\%.
\end{aligned}
$$

7. Nekim binarnim kanalom prenose se podaci zapisani (samo) pomoću znakova 0 i 1. Vjerojatnost da je poslani znak jednak 1 iznosi 55%, a vjerojatnost da je poslani znak pogrešno primljen (i interpretiran) iznosi 20%. Ako je primljen (i interpretiran) znak 1, izračunajte vjerojatnost da je poslani znak jednak 0.

Rješenje: Neka su:

$N = \{\text{poslani znak jednak je } 0\}$,
$J = \{\text{poslani znak jednak je } 1\}$,
$B = \{\text{primljeni (i interpretirani) znak je } 1\}$.

Lako se provjeri da je skup $\mathcal{A} = \{J, N\}$ potpun sustav događaja.

Iz zadanih podataka slijedi:

$$\begin{aligned} &P(J) = 55\% = 0.55, \\ &P(N) = 1 - P(J) = 1 - 0.55 = 0.45, \\ &P(B \mid J) = 80\% = 0.8, \\ &P(B \mid N) = 20\% = 0.2. \end{aligned}$$

Tražimo vjerojatnost $P(N \mid B)$. Primjenom Bayesove formule dobijemo:

$$\begin{aligned} P(N \mid B) &= \frac{P(N) \cdot P(B \mid N)}{P(J) \cdot P(B \mid J) + P(N) \cdot P(B \mid N)} = \\ &= \frac{0.45 \cdot 0.2}{0.55 \cdot 0.8 + 0.45 \cdot 0.2} = \\ &= \frac{9}{53} \end{aligned}$$