Bayesova formula s primjerima

Peulić, Ana

Sveučilište J.J. Strossmayera u Osijeku
Fakultet primijenjene matematike i informatike
Sveučilišni prijediplomski studij matematike

Ana Peulić

Bayesova formula s primjerima

Završni rad

Osijek, 2023.

Sveučilište J.J. Strossmayera u Osijeku
Fakultet primijenjene matematike i informatike
Sveučilišni prijediplomski studij matematike

Ana Peulić

Bayesova formula s primjerima

Završni rad

Mentor: doc. dr. sc. Ivan Papić

Osijek, 2023.

Sažetak

Tema ovog rada je Bayesova formula. Kroz rad prolazimo i uvjetnu vjerojatnost budući da je formula zasnovana na istoj. Navest ćemo primjene i proći kroz nekolicinu primjera pokazujući na koje sve načine Bayesova formula može biti upotrebljena. Vjerojatnost nekog događaja uz danu realizaciju nekog drugog događaja, može se računati na mnoge načine. Primjetit ćemo da nas navedena formula, od naizgled složenog računa, dovodi do računa minimalnog napora.

Ključne riječi

vjerojatnost, uvjetna vjerojatnost, formula potpune vjerojatnosti, Bayesova formula

Bayes' theorem with examples

Abstract

The topic of this paper is Bayes' theorem. Through paper, we also go through the conditional probability since the formula is based on the same. We will list applications and go through several examples showing in what ways the Bayes' theorem can be used. The probability of an event with the specified condition, can be calculated in many ways. We will notice that the given formula from a seemingly complex calculation leads to an calculation with minimal effort.

Key words

probability, conditional probability, formula of total probability, Bayes' theorem

Sadržaj

Uvod

Ovaj je rad baziran na pojmu vjerojatnosti. Kada bismo preveli na jezik shvatljiv većini ljudi, vjerojatnost bismo definirali kao šansu pojave nekoga događaja.
U prvom poglavlju proći ćemo kroz definicije u svrhu što bolje usvojenosti i lakše interpretacije pojmova. Usvajanje definicija često je svedeno na čitanje i pamćenje istih. Dubljom analizom informacija postiže se kvalitetnija obrada rezultata što i jest svrha ovoga rada. Osim pojmova vezanih uz vjerojatnost, obradit ćemo i uvjetnu vjerojatnost. Dokazat ćemo da je upravo uvjetna vjerojatnost ključna za razumijevanje Bayesove formule.
U drugom, ujedno i zadnjem poglavlju, prelazimo na glavnu temu, a to je Bayesova formula i njezine primjene. Proći ćemo kroz razne primjere i vidjeti u kojim sve segmentima čovjekova djelovanja dano gradivo može biti primijenjeno.

1 Osnovni pojmovi iz teorije vjerojatnosti

1.1 Pojmovi

Zahvaljujući engleskom matematičaru i statističaru, Thomasu Bayesu (1702.-1761.), utemeljena je Bayesova formula na način kakvu ju danas poznajemo. Kako bi nam postupci izračunavanja bili što jasniji, definirajmo pojmove pomoću kojih ćemo dolaziti do nama bitnih zaključaka.

Definicija 1. ***Slučajan pokus*** *je pokus čiji ishod nije jednoznačno određen uvjetima u kojima se pokus izvodi.*

Definicija 2. *Svaki ishod jednog izvođenja slučajnog pokusa naziva se* ***elementarni događaj*** ω.

Definicija 3. ***Skup*** *ili* ***prostor elementarnog događaja*** Ω *je skup svih mogućih ishoda slučajnog pokusa.*

Napomena 1. *Elementi skupa* Ω *su elementarni događaji* ω.

Definicija 4. *Neka je* $\Omega \neq \emptyset$. *Familija* $\mathcal{F}$ *podskupova skupa* Ω *jest* ***σ-algebra*** *skupova na* Ω *ukoliko vrijedi:*

1. $\emptyset \in \mathcal{F}$
2. *Ako je* $A \in \mathcal{F}$, *onda je i* $A^c \in \mathcal{F}$.
3. *Ako je dana prebrojiva familija skupova* $(A_i, i \in I) \subseteq \mathcal{F}, I \subseteq \mathbb{N}$, *onda* $\mathcal{F}$ *sadrži i njihovu uniju* $\bigcup_{i \in I} A_i \in \mathcal{F}$.

Definicija 5. *Elemente odabrane* σ*-algebre na* Ω *zovemo* ***događajima***.

Definicija 6. ***Suprotan događaj*** *događaju* $A \subseteq \Omega$ *jest skup* $A^c = \{\omega \in \Omega : \omega \notin A\}$.

Definicija 7. ***Relativna frekvencija*** *događaja* A *definira se kao kvocijent frekvencije* n_A *(broj ponavljanja realizacije* A *tijekom* n *izvođenja pokusa) i broja izvođenja pokusa* n.

Definicija 8. *Neka je* $\Omega \neq \emptyset$ *skup elementarnih događaja te* $\mathcal{F}$ σ*-algebra događaja na tom skupu.* ***Vjerojatnost na skupu*** Ω *definira se preslikavanjem* $P : \mathcal{F} \to \mathbb{R}$ *ukoliko vrijedi:*

(A1) $P(A) \geq 0, \forall A \in \mathcal{F}$ ***(nenegativnost vjerojatnosti)***

(A2) $P(\Omega) = 1$ ***(normiranost vjerojatnosti)***

(A3) *Ako je* $(A_i, i \in I) \subseteq \mathcal{F}$ *prebrojiva familija, gdje je* $A_i \cap A_j = \emptyset$, *za* $i \neq j$, *onda vrijedi:* $P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ ***(σ-aditivnost vjerojatnosti)***.

Definicija 9. *Uređena trojka* $(\Omega, \mathcal{F}, P)$ *naziva se* ***vjerojatnosni prostor****, pri čemu je* $\Omega \neq \emptyset$ *prostor elementarnih događaja,* $\mathcal{F}$ σ*-algebra na* Ω, $P : \mathcal{F} \to \mathbb{R}$ *vjerojatnost.*

1.2 Uvjetna vjerojatnost

Proučavamo vjerojatnost događaja uz uvjet da je prethodno došlo do realizacije nekog drugog događaja. Dakle, vjerojatnost realizacije događaja A, ukoliko je prethodno došlo do realizacije događaja B, definiramo na sljedeći način.

Definicija 10. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor te* $A, B \in \mathcal{F}$ *događaji iz* Ω*, pri čemu je* $P(B) > 0$*.* ***Uvjetna vjerojatnost događaja A uz dani događaj*** *B jest preslikavanje* $P_B : \mathcal{F} \to [0, 1]$ *definirano na sljedeći način:*

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

Koristimo i oznaku $P(A|B) = P_B(A)$.

Napomena 2. *Uočimo,* $P(A^c|B) = 1 - P(A|B)$.

Prirodno je zapitati se što ako jedan događaj ne zavisi o drugome. Odnosno, što ako realizacija događaja B ne utječe na vjerojatnost događaja A? Pretpostavljamo da je tada:

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

Dakle, vjerojatnost događaja A ostaje nepromijenjena bez obzira dolazi li do realizacije događaja B ili pak ne. To govori samo jedno, a to je da događaj A ne zavisi o događaju B, tj. da su navedeni događaji nezavisni jedno o drugome.

Definicija 11. *Neka je* $(\Omega, \mathcal{F}, P)$ *vjerojatnosni prostor. Dva su događaja* $A, B \in \mathcal{F}$ ***nezavisna*** *ako vrijedi:*

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

Pretpostavili smo da je događaj A nezavisan o događaju B pa slijedeći definiciju dolazimo do zaključka:

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

Uz uvjetnu vjerojatnost, pojavljuje se još jedna jako bitna formula. Radi se o formuli potpune vjerojatnosti do koje dolazimo na sljedeći način.

Definicija 12. *Familija događaja* $(H_i, i \in I), I \subseteq \mathbb{N}$ *u vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$ *čini* ***potpun sustav događaja*** *ukoliko vrijedi:*

1. $H_i \neq \emptyset$, *za sve* $i \in I$
2. $H_i \cap H_j = \emptyset$, *za sve* $i \neq j$, $i, j \in I$
3. $\bigcup_{i \in I} H_i = \Omega$.

Teorem 1 (Formula potpune vjerojatnosti)**.** *([1], str. 36.) Neka je* $(H_i, i \in I)$ *potpun sustav događaja u vjerojatnosnom prostoru* $(\Omega, \mathcal{F}, P)$*. Tada za svaki događaj* $A \in \mathcal{F}$ *vrijedi:*

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

Dokaz. Uočimo da se događaj $A \in \mathcal{F}$ može prikazati na sljedeći način:

$$A = \bigcup_{i \in I} (A \cap H_i).$$

Također, uočimo da je $\{A \cap H_i,\ i \in I\},\ I \subseteq \mathbb{N}$, familija disjkunktnih skupova. Primjenom svojstva *(A3)* dobivamo:

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

Iz formule uvjetne vjerojatnosti poznato je:

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

Izražavanjem brojnika te sumiranjem obje strane dolazimo do izraza:

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

$\square$

Dakle, potrebno je izabrati familiju događaja koja tvori potpun sustav događaja, a potom kroz navedenu formulu dolazimo do tražene vjerojatnosti.

2 Bayesova formula

2.1 Definicija

Prethodno je poglavlje zaključeno potpunim sustavom događaja. Što su zapravo elementi navedenoga skupa? Primjetimo da su elementi potpunog sutava događaja međusobno disjunktni događaji čija je unija jednaka sigurnom događaju. Te elemente nazivamo hipotezama. Dolazimo do formule koja je bazirana na vjerojatnosti hipoteza ukoliko znamo da je došlo do realizacije događaja A.

Teorem 2 (Bayesova formula)**.** *([1], str. 39.) Neka je $(H_i, i \in I)$, $I \subseteq \mathbb{N}$ potpun sustav događaja u vjerojatnosnom prostoru $(\Omega, \mathcal{F}, P)$ te neka je A bilo koji događaj iz $\mathcal{F}$, pri čemu je $P(A) > 0$. Tada za svaki $i \in I$ vrijedi sljedeće:*

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

Dokaz. Prema definiciji uvjetne vjerojatnosti imamo:

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

Znamo da je $P(H_i \cap A) = P(A \cap H_i)$.

Pogledajmo vjerojatnost događaja A ukoliko je došlo do realizacije H_i.

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

Izrazimo iz jednakosti $P(A \cap H_i)$, odnosno $P(H_i \cap A)$, ubacimo u gornji izraz te dobivamo:

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

□

Primjer 1. *Marta je pozvana na tematsku zabavu pod nazivom: "Crno-bijeli svijet." Prijateljice su joj odlučile pomoći oko oblačenja. Petra joj je donijela tri bijele i tri crne haljine, Dora četiri bijele i pet crnih, dok je Lorena donijela samo dvije crne haljine. Kolika je vjerojatnost da je Marta obukla Loreninu haljinu, ako znamo da se na kraju ipak odlučila za crnu?*

Rješenje:

Neka su H_1, H_2, H_3 događaji koji čine potpun sustav događaja, gdje je:

$$\begin{aligned} H_1 &= \{\text{haljina je Petrina}\} \\ H_2 &= \{\text{haljina je Dorina}\} \\ H_3 &= \{\text{haljina je Lorenina}\}. \end{aligned}$$

Odredimo vjerojatnost događaja $H_i,\ i = 1, 2, 3.$
Ukupno ima 17 haljina, od čega je šest Petrinih, devet Dorinih i dvije Lorenine. Stoga je:

$$\begin{aligned} P(H_1) &= \frac{6}{17} \\ P(H_2) &= \frac{9}{17} \\ P(H_3) &= \frac{2}{17}. \end{aligned}$$

Možemo uočiti da su navedeni događaji disjunkti i pozitivne vjerojatnosti te da u uniji daju skup Ω.
Označimo s A događaj čiju vjerojatnost tražimo, tj.

$$A = \{\text{Marta je obukla crnu haljinu}\}.$$

Vjerojatnost događaja A računamo formulom potpune vjerojatnosti (FPV):

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

Izračunali smo vjerojatnost svakog od događaja H_i te sada preostaje odrediti vjerojatnost događaja A uz uvjet da je došlo do realizacije događaja $H_i,\ za\ i = 1, 2, 3.$

$$\begin{aligned} P(A|H_1) &= \frac{3}{6} = \frac{1}{2} \text{ jer je Petra donijela 6 haljina od čega su 3 crne} \\ P(A|H_2) &= \frac{5}{9} \text{ jer je Dora donijela 9 haljina od čega je 5 crnih} \\ P(A|H_3) &= \frac{2}{2} = 1 \text{ jer je Lorena donijela 2 haljine i obje su crne} \end{aligned}$$

Uvrstimo u FPV:

$$P(A) = \frac{6}{17} \cdot \frac{1}{2} + \frac{9}{17} \cdot \frac{5}{9} + \frac{2}{17} \cdot 1 = 0.5882.$$

Konačnu, traženu vjerojatnost događaja (da je Marta obukla crnu haljinu ukoliko znamo da je haljina Lorenina), računamo Bayesovom formulom:

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

Uvrstimo tražene podatke te dobijemo:

$$P(H_3|A) = \frac{P(H_3) \cdot P(A|H_3)}{P(A)} = \frac{\frac{2}{17} \cdot 1}{0.5882} = 0.2.$$

Zaključujemo, vjerojatnost da je Marta obukla Loreninu haljinu ako znamo da je obukla crnu, iznosi 20%.

2.2 Primjeri korištenja Bayesove formule

Kao što je navedeno na početku, vjerojatnost i statistika primijenjuju se u raznim područjima naše svakodnevice, većinom nesvijesno. Često se susrećemo sa izazovima te prije poduzimanja prvog poteza, procijenimo rizike i zatim razmatramo buduće i korisne postupke. Često se zapitamo kolika je uopće šansa za uspjehom ukoliko se odlučimo za pojedini potez. Razmišljajući na taj način procjenjujemo vjerojatnosti uspjeha/neuspjeha.

2.2.1 Primjeri u medicini

Ima li smisla promatrati učinak određenog lijeka ako se osoba nad kojom vršimo promatranje nije striktno držala propisane terapije? Kolika je vjerojatnost da ćemo promatranjem više takvih osoba dobiti točne podatke o tom lijeku? Odlučimo li pokrenuti određeno istraživanje, koliko ćemo se namučiti u pronalasku prave osobe za to ispitivanje? Nadahnuti tim pitanjima, prolazimo kroz sljedeći primjer.

Primjer 2. *Za kliničko istraživanje odabrana je osoba cijepljena protiv COVID-19. U vrijeme cijepljenja odabrane osobe, u KBC Osijek na raspolaganju je bilo tri vrste cjepiva: Pfizer, Johnson & Johnson te Moderna. Udio Pfizer-a tada je bio 60%, Johnson & Johnson 25%, a preostali udio bio je Moderna. Tadašnja cjepiva Pfizer i Moderna primala su se u dvije doze, dok je Johnson&Johnson samo u jednoj. Pretpostavimo da je osoba primila sve potrebne*

doze.
1. Kolika je vjerojatnost da je odabrana osoba cijepljena cjepivom Moderna?
2. Kolika je vjerojatnost da je odabrana osoba cijepljena cjepivom Moderna, ako znamo da je odabrana osoba primila dvije doze?

Rješenje:

Označimo s H_1, H_2, H_3 događaje koji čine potpun sustav događaja.

$$H_1 = \{\text{osoba je primila cjepivo Pfizer}\}$$
$$H_2 = \{\text{osoba je primila cjepivo Johnson \& Johnson}\}$$
$$H_3 = \{\text{osoba je primila cjepivo Moderna}\}$$

Također, označimo sa A događaj:

$$A = \{\text{osoba je primila 2 doze}\}.$$

Odredimo vjerojatnosti hipoteza.

$$P(H_1) = 0.6 \text{ jer je udio cjepiva Pfizer bio } 60\%.$$
$$P(H_2) = 0.25 \text{ jer je udio cjepiva Johnson \& Johnson bio } 25\%.$$
$$P(H_3) = 0.15 \text{ jer je udio cjepiva Moderna bio } 15\%.$$

Dakle, vjerojatnost da je odabrana osoba cijepljena cjepivom moderna iznosti 15%.

S druge strane,

$$P(A|H_1) = 1 \text{ jer se cjepivo Pfizer primao u dvije doze}$$
$$P(A|H_2) = 0 \text{ jer se cjepivo Johnson \& Johnson primao u jednoj dozi}$$
$$P(A|H_3) = 1 \text{ jer se cjepivo Moderna primao u dvije doze.}$$

Budući da sada imamo sve potrebne vrijednosi, primijenimo Bayesovu formulu:

$$\begin{aligned} P(H_3|A) &= \frac{P(H_3) \cdot P(A|H_3)}{P(H1) \cdot P(A|H1) + P(H2) \cdot P(A|H2) + P(H3) \cdot P(A|H3)} \\ &= \frac{0.15 \cdot 1}{0.6 \cdot 1 + 0.25 \cdot 0 + 0.15 \cdot 1} = \frac{0.15}{0.75} = 0.2. \end{aligned}$$

Dakle, ako znamo da je odabrana osoba primila dvije doze, vjerojatnost da je cijepljena cjepivom Moderna iznosi 20%.

2.2.2 Primjeri u poljoprivredi

Poljoprivreda je nerijetko vezana za rizike. Čovjek ne može u potpunosti sam upravljati svojim urodom. Velik je broj čimbenika o kojima ovisi urod određene kulture, poput dovoljne količine kiše, topline, temeljne obrade tla i slično. Kada je sve gotovo, ostaje samo pitanje jesu li usjevi dovoljno dobri za daljnju proizvodnju hrane. Pogledajmo sljedeći primjer.

Primjer 3. *Ratari su krenuli u žetvu i svoje plodove voze na silos. Tamo se provode dodatne provjere kako bi se ustanovilo jesu li zrna pšenice još uvijek vlažna ili su dovoljno suha za otkup. Ukoliko nije dovoljno suha, pšenica se vraća proizvođaču. Trećina je dovežene pšenice sorte Kraljica, polovina je sorte Grendor, a šestina je sorte Srpanjka. Stopa vlažnosti sorte Kraljica iznosi 17%, stopa vlažnosti sorte Grendor 12%, dok stopa vlažnosti sorte Srpanjka iznosi 15%.*
1. Ako je komisija uzela slučajan uzorak bilo kojeg od proizvođača, kolika je vjerojatnost uzimanja sorte Grendor?
2. Ako je dodatnim provjerama ustanovljeno da je slučajan uzorak pšenice izrazito vlažan, kolika je vjerojatnost da se radi o sorti Grendor?

Rješenje:

Potpun sustav događaja čine događaji:

$$H_1 = \{\text{uzorak je sorte Kraljica}\}$$
$$H_2 = \{\text{uzorak je sorte Grendor}\}$$
$$H_3 = \{\text{uzorak je sorte Srpanjka}\}.$$

Označimo s A događaj:

$$A = \{\text{uzorak je vlažan}\}.$$

Iz zadanih podataka možemo odrediti vjerojatnost hipoteza.

$$\begin{aligned} P(H_1) &= \frac{1}{3} \text{ jer je trećina pšenice sorte Kraljica} \\ P(H_2) &= \frac{1}{2} \text{ jer je polovina pšenice sorte Grendor} \\ P(H_3) &= \frac{1}{6} \text{ jer je šestina sorte Srpanjka} \end{aligned}$$

Iz određenih vjerojatnosti uočimo, vjerojatnost da je slučajno izabrana sorta Grendor jest 50%.

Kako je:

$$\begin{aligned} P(A|H_1) &= 0.17 \\ P(A|H_2) &= 0.12 \\ P(A|H_3) &= 0.15, \end{aligned}$$

imamo:

$$\begin{aligned} P(H_2|A) &= \frac{P(H_2) \cdot P(A|H_2)}{P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + P(H_3) \cdot P(A|H_3)} \\ &= \frac{\frac{1}{2} \cdot 0.12}{\frac{1}{3} \cdot 0.17 + \frac{1}{2} \cdot 0.12 + \frac{1}{6} \cdot 0.15} = 0.4235. \end{aligned}$$

Zaključujemo, ako je dodatnim istraživanjima potvrđena velika prisutnost vlage, vjerojatnost da je riječ o sorti Grendor iznosi 42.35%.

2.2.3 Primjeri u policiji

Tragajući za počiniteljem određenog kaznenog djela, policajci se često moraju osvrnuti na ljude oko sebe i pažljivo pratiti tragove. Uvijek je prisutna sumnja u istinitost izjava od strane svjedoka, ali što veći uzorak ispitanika trebao bi rezultirati pravednijom presudom. Kolika je vjerojatnost da svjedoci govore istinu? Kolika je vjerojatnost da se, prema riječima svjedoka, radi o opisanoj osobi, ilustrirat ćemo u primjeru koji slijedi.

Primjer 4. *Tijekom noći došlo je do provale u stan. Nakon izlaska policije na teren, 25% svjedoka reklo je kako misle da je riječ o muškoj osobi, dok 75% svjedoka tvrdi da su vidjeli sumnjivu osobu ženskog spola. 15% ispitanika koji tvrde da je riječ o provalniku, uvjereni su kako je osoba viša od 180cm, dok iz grupe onih ljudi koji tvrde da je riječ o provalnici, njih 90% sigurno je da je osoba niža od 180cm.*
1. Kolika je vjerojatnost da je osoba niža od 180cm?
2. Ako je provjerama kamera s okolnog područja ustanovljeno da je sumnjiva osoba ipak niža od 180cm, kolika je vjerojatnost da se radi o osobi ženskog spola?

Rješenje:

Dva su događaja koja čine potpun sustav događaja:

$$H_1 = \{\text{osoba je muškog spola}\}$$
$$H_2 = \{\text{osoba je ženskog spola}\}.$$

Primijetimo:

$$P(H_1) = 0.25 \text{ jer je 25\% svjedoka vidjelo provalnika}$$
$$P(H_2) = 0.75 \text{ jer je 75\% svjedoka vidjelo provalnicu.}$$

Označimo događaje:

$$A = \{\text{osoba je visoka 180cm ili više}\}$$
$$A^c = \{\text{osoba je niža od 180cm}\}.$$

Prema zadanim podacima:

$$P(A|H_1) = 0.15$$
$$P(A^c|H_2) = 0.9.$$

Prema definiciji uvjetne vjerojatnosti suprotnog događaja:

$$P(A^c|H_1) = 1 - P(A|H_1) = 1 - 0.15 = 0.85$$
$$P(A|H_2) = 1 - P(A^c|H_2) = 1 - 0.9 = 0.1.$$

Tražimo $P(A^c)$.
Prema formuli potpune vjerojatnosti:

$$\begin{aligned} P(A^c) &= P(H_1) \cdot P(A^c|H_1) + P(H_2) \cdot P(A^c|H_2) \\ &= 0.25 \cdot 0.85 + 0.75 \cdot 0.9 = 0.8875. \end{aligned}$$

Dakle, vjerojatnost da je provalnik/provalnica niži/niža od 180cm iznosi 88.75%.

Pronađimo $P(H_2|A^c)$.
Upotrijebimo Bayesov teorem.

$$\begin{aligned} P(H_2|A^c) &= \frac{P(H_2) \cdot P(A^c|H_2)}{P(A^c)} \\ &= \frac{0.75 \cdot 0.90}{0.8875} = 0.7606 \end{aligned}$$

Dakle, ukoliko je ustanovljeno da je osoba niža od 180cm, vjerojatnost da se radi o osobi ženskog spola iznosi 76.06%.

2.2.4 Primjeri u prehrambenoj industriji

Tijekom proizvodnje, posebice namirnica, nužna je dobra kvaliteta specifičnih sastojaka koji tvore gotov proizvod. Čak i najmanja pogreška pri proizvodnji može uzrokovati zdravstvene probleme i uzrokovati opasne simptome uključujući i smrtonosne. Razumno je razmišljati da to nije u cilju poizvođača, međutim, kada se propust dogodi, nužno je sanirati napravljenu štetu i spriječiti nastanak još veće, globalnije. Kada u medijima čujemo o povlačenju proizvoda kojeg imamo u kućanstvu, nadamo se da nije riječ o onome u našem hladnjaku ili pak onome kojeg smo već konzumirali. Kolike su šanse da se radi baš o tom proizvodu? Kolika je šansa da smo u sebe unijeli kancerogenu tvar? Ilustrirat ćemo kroz sljedeći primjer.

Primjer 5. *U tvornici čokolade došlo je do propusta pri proizvodnji tijekom jednog dana. Smatra se da je u čokoladama utvrđen etilen oksid te je čokolade navedenoga datuma potrebno baciti. 40% je mliječnih čokolada, dok tamna i bijela zauzimaju podjednak udio u proizvodnji. Ustanovilo se da je 45% mliječnih, 60% tamnih i 65% bijelih čokolada toga dana pravilno proizvedeno.*

1. Kolika je vjerojatnost da ćemo, nakon provjere datuma otisnutog na pakiranju, čokoladu baciti?
2. Kolika je vjerojatnost da smo kupili mliječnu čokoladu, ako znamo da smo ju kasnije bacili?

Rješenje:

Događaji koji čine potpun sustav događaja:

$$\begin{aligned} H_1 &= \{\text{čokolada je mliječna}\} \\ H_2 &= \{\text{čokolada je tamna}\} \\ H_3 &= \{\text{čokolada je bijela}\}. \end{aligned}$$

Koliko iznose njihove vjerojatnosti? Iz uvjeta zadatka vidimo da je u proizvodnji udio mliječnih čokolada jednak 40%, dok 60% čine tamne i bijele čokolade. Budući da je njihov udio u proizvodnji podjednak, udio tamne iznosi 30%, a također toliko iznosi i udio bijele čokolade. Sada je:

$$\begin{aligned} P(H_1) &= 0.4 \\ P(H_2) &= 0.3 \\ P(H_3) &= 0.3. \end{aligned}$$

Neka su:

$$\begin{aligned} A &= \{\text{čokolada se smije konzumirati}\} \\ A^c &= \{\text{čokoladu je potrebno baciti}\}. \end{aligned}$$

Odredimo vjerojatnost događaja A^c uz danu realizaciju H_i, *za* $i = 1, 2, 3$.

$P(A|H_1) = 0.45$ jer je 45% mliječnih čokolada tog dana pravilno proizvedeno
$P(A|H_2) = 0.6$ jer je 60% mliječnih čokolada tog dana pravilno proizvedeno
$P(A|H_3) = 0.65$ jer je 65% mliječnih čokolada tog dana pravilno proizvedeno

Uvrstimo u formulu:

$$\begin{aligned} P(A) &= P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + P(H_3) \cdot P(A|H_3) \\ &= 0.4 \cdot 0.45 + 0.3 \cdot 0.6 + 0.3 \cdot 0.65 = 0.56. \end{aligned}$$

Prema vjerojatnosti suprotnog događaja:

$$P(A^c) = 1 - P(A) = 1 - 0.56 = 0.44.$$

Vjerojatnost da smo bacili čokoladu iznosi 0.44%.

Odredimo $P(H_1|A^c)$.
Prema Bayesovom teoremu:

$$\begin{aligned} P(H_1|A^c) &= \frac{P(H_1) \cdot P(A^c|H_1)}{P(A^c)} \\ &= \frac{P(H_1) \cdot (1 - P(A|H_1))}{P(A^c)} \\ &= \frac{0.4 \cdot (1 - 0.55)}{0.44} = 0.41. \end{aligned}$$

Vjerojatnost da smo uzeli mliječnu čokoladu koju smo kasnije bacili iznosi 41%.

2.2.5 Primjeri u automobilskoj industriji

U prethodnom odlomku, uveli smo se u svijet proizvodnje i nastanka propusta tijekom iste. Posjedovanje automobila neophodno je za uklapanje u užurbani svijet koji se odvija oko nas. Pri kupnji vozila važan nam je sigurnosni sustav koji ono pruža. Automobilskih je nesreća sve više i ljudi često stradavaju u njima. Kroz primjer ćemo proći kroz selekciju automobila koji su spremni za prodaju.

Primjer 6. *U tvornici automobila proizvodeno je 67% crvenih automobila, dok su preostali automobili sivi. U svrhu testiranja vozila, odvoze se automobili jedan po jedan.*
1. Kolika je vjerojatnost da je je prvo izabran sivi automobil?
2. Tijekom prve vožnje ustanovilo se da automobil ima grešku na motoru. Kasnije se saznalo da 30% crvenih automobila ima isti problem, dok je kod sivih automobila pojavnost tog kvara trostruko manja. Kolika je sada vjerojatnost da je izabran sivi automobil?

Rješenje:

Događaji koji čine potpun sustav događaja:

$$H_1 = \{\text{automobil je crvene boje}\}$$
$$H_2 = \{\text{automobil je sive boje}\},$$

pri čemu su

$$P(H_1) = 0.67 \text{ jer } 67\% \text{ čine crveni automobili}$$
$$P(H_2) = 0.33 \text{ jer } 33\% \text{ čine sivi automobili.}$$

Označimo događaje:

$$I = \{\text{ima grešku na motoru}\}$$
$$N = \{\text{nema grešku na motoru}\}.$$

Već smo naveli da je $P(H_2) = 0.33$, tj. vjerojatnost da je prvo izabran sivi automobil je 33%.

Moramo odrediti $P(H_2|I)$.
Budući da nam je događaj S jedna od hipoteza potpunog sustava događaja, koristimo Bayesov teorem.

$$P(H_2|I) = \frac{P(H_2) \cdot P(I|H_2)}{P(I)}$$

Iz zadatka vidimo:

$$P(I|H_1) = 0.3 \text{ jer } 30\% \text{ crvenih automobila ima grešku na motoru}$$
$$P(I|H_2) = 0.1 \text{ jer } 10\% \text{ sivih automobila ima grešku na motoru.}$$

Uvrstimo u formulu:

$$\begin{aligned} P(H_2|I) &= \frac{P(H_2) \cdot P(I|H_2)}{P(H_1) \cdot P(I|H_1) + P(H_2) \cdot P(I|H_2)} \\ &= \frac{0.33 \cdot 0.1}{0.67 \cdot 0.3 + 0.33 \cdot 0.1} = 0.1410. \end{aligned}$$

Dakle, ako automobil ima grešku na motoru, vjerojatnost da je to sivi automobil iznosi 14.10%.

2.2.6 Primjeri u sportu

Napetost tijekom gledanja utakmice dodatno pojačava i ozljeda nekog od igrača. Naravno da bismo voljeli da nesreća uopće nema, ali kada dođe do izvanrednih situacija voljeli bismo da je tim za koji mi navijamo siguran i zaštičen. Kolika je šansa da je upravo naš najdraži igrač ozlijeđen? Kolika je vjerojatnost da će se suprotne strane tribina sukobiti? Sve su to problemi koje želimo minimizirati provodeći posebnu proceduru par dana prije i na sam dan igranja. Prođimo kroz primjer koji predstavlja jedan od navedenih problema.

Primjer 7. *Tijekom nogometne utakmice "Bijelih" i "Plavih" došlo je do nereda na tribini. Nekolicina navijača počela je bacati baklje na teren te je jedna osoba ozlijeđena. Među ekipom "Bijeli", 70% ima tamnu kosu, dok su ostali svijetlokosi. S druge strane, 90% "Plavih" je svijetlokoso, a ostatak tamnokoso kao i sudci koji su sudili taj dan.*
1) Kolika je vjerojatnost da je tijekom nereda ozlijeđen sudac?
2) Kolika je vjerojatnost da je ozlijeđena osoba igrač "Plavih", ako smo od silnog dima koji je nastao, uspjeli vidjeti da je tamnokosoj osobi pružena pomoć?

Rješenje:

Neka su:

$$\begin{aligned} H_1 &= \{\text{igrač je iz tima "Plavi"}\} \\ H_2 &= \{\text{igrač je iz tima "Bijeli"}\} \\ H_3 &= \{\text{osoba je sudac}\}. \end{aligned}$$

Znamo da je broj igrača svakog tima jednak 11 (uključujući i vratara) te da po utakmici imamo 4 sudca od kojih je jedan neposredno izvan terena. Prema tome, vjerojatnosti od $H_i, za\ i = 1, 2, 3$ računamo na sljedeći način:

$P(H_1) = 11/25 = 0.44$ jer je 11 igrača "Bijelih" od ukupno 25 osoba na terenu
$P(H_2) = 11/25 = 0.44$ jer je 11 igrača "Plavih" od ukupno 25 osoba na terenu
$P(H_3) = 3/25 = 0.12$ jer je 3 sudca od ukupno 25 osoba na terenu.

Stoga, vjerojatnost da je ozlijeđena osoba sudac iznosi 12%.

Označimo sada događaje:

$$\begin{aligned} A &= \{\text{osoba je tamnokosa}\} \\ A^c &= \{\text{osoba je svijetlokosa}\}. \end{aligned}$$

Zanima nas koliko iznosi $P(H_1|A)$, a to dobijemo pomoću iduće formule:

$$P(H_1|A) = \frac{P(H_1) \cdot P(A|H_1)}{P(A)}.$$

Potrebne su nam uvjetne vjerojatnosti:

$$\begin{aligned} &P(A|H_1) = 0.1 \text{ jer udio tamnokosih u timu "Plavi" iznosi } 10\% \\ &P(A|H_2) = 0.7 \text{ jer udio tamnokosih u timu "Bijeli" iznosi } 70\% \\ &P(A|H_3) = 1 \text{ jer su svi sudci tamnokosi.} \end{aligned}$$

Računamo vjerojatnost događaja A, odnosno vjerojatnost da je ozlijeđena osoba tamnokosa.

$$\begin{aligned} P(A) &= P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + P(H_3) \cdot P(A|H_3) \\ &= 0.44 \cdot 0.1 + 0.44 \cdot 0.7 + 0.12 \cdot 1 = 0.472 \end{aligned}$$

Uvrstimo dobivene podatke u gornji izraz za $P(H_1|A)$:

$$\begin{aligned} P(H_1|A) &= \frac{P(H_1) \cdot P(A|H_1)}{P(A)} \\ &= \frac{0.44 \cdot 0.1}{0.472} = 0.0932. \end{aligned}$$

Prema tome, vjerojatnost da je ozlijeđena osoba igrač iz tima "Plavi", ukoliko znamo da je ozlijeđena osoba tamnokosa, iznosi 9.32%.

Literatura

[1] M. Benšić, N. Šuvak, Uvod u vjerojatnost i statistiku, Grafika d.o.o., Osijek 2014.

[2] M. F. Triola, Bayes' Theorem, online materijal http://faculty.washington.edu/tamre/BayesTheorem.pdf (pristupljeno 12.9.2023.)