TILASTOMATEMATIIKKA

Harjoitusviikon 1 tehtävien ratkaisut, kevät 2024

Harjoituksen teemoja ovat:

(i) Joukko-opin peruslaskutoimitukset

(ii) Satunnaiskoe, otosavaruus ja tapahtuma

(iii) Todennäköisyyden peruslaskutoimitukset

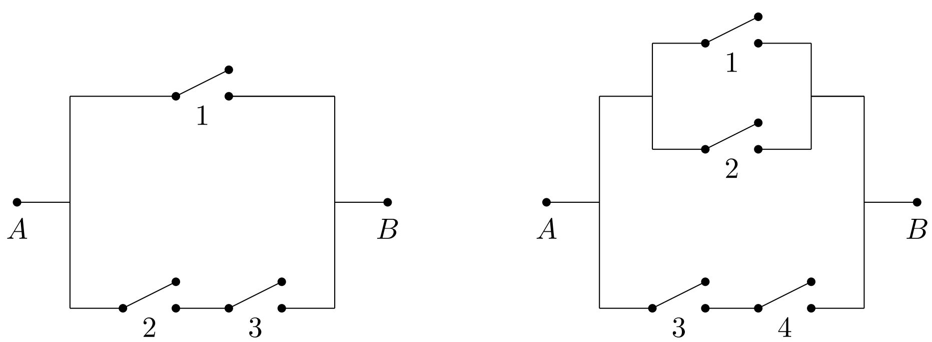
(iv) Tapahtumien riippumattomuus

(v) Ehdollinen todennäköisyys

(vi) Kokonaistodennäköisyys ja Bayesin kaava

Tehtävät 1-4 käsittelevät ensisijaisesti teemoja (i)-(iv), joita käsitellään pääasiassa alkuviikon harjoituksissa. Teemat (v) ja (vi) on lähtökohtaisesti tarkoitettu käsiteltäväksi loppuviikon harjoituksissa. Myös nopeampi eteneminen on mahdollista, jos kokee hallitsevansa alkuviikon asiat. Alkuviikon harjoituksissa suositellaan laskettavaksi **tehtävät 1 ja 2** ja loppuviiikon harjoituksissa **tehtävät 5 ja 6**.

1. Tarkastellaan oheisessa kuvassa olevaa kahta virtapiiriä, jotka koostuvat rinnan tai sarjaan kytketyistä kytkimistä. Sähkö pääsee kulkemaan kytkimen läpi vain, jos se on kiinni. Sitä varten olkoon K_i tapahtuma ”virtapiirin kytkin i on kiinni”.



Kuva 1: Eräitä virtapiirejä

a) Esitä molemmille piireille joukko-opillisesti tapahtumien K_i avulla tapahtumat

$$V = \text{”A:sta virtaa sähkö pisteeseen B”},$$
$$E = \text{”A:sta ei virtaa sähkö pisteeseen B”}.$$

Piirrä Vennin diagrammi vasemman puoleisen piirin tapahtumalle V.

b) Oletetaan, että kytkimet ovat kiinni toisista riippumatta $90\,\%$ todennäköisyydellä. Millä todennäköisyydellä virtapiirit toimivat, kun toimiminen tarkoittaa sitä, että pisteestä A virtaa sähkö pisteeseen B?

Ratkaisu:

a) Tarkastellaan ensin vasemmanpuoleista piiriä. Sähkö voi virrata yläkautta, alakautta tai molempia reittejä pitkin. Sähkö pääsee kulkemaan yläkautta täsmälleen silloin, kun kytkin 1 on kiinni, eli K_1 tapahtuu. Vastaavasti sähkö pääsee virtaamaan alakautta täsmälleen silloin, kun molemmat kytkimet 2 ja 3 ovat kiinni, eli $K_2 \cap K_3$ tapahtuu.

Niinpä joukko-opillisesti $V = K_1 \cup (K_2 \cap K_3)$. Koska E on V:n komplementti, on De Morganin kaavojen mukaan

$$E = \overline{K_1 \cup (K_2 \cap K_3)} = \overline{K_1} \cap (\overline{K_2 \cap K_3}) = \overline{K_1} \cap (\overline{K_2} \cup \overline{K_3}).$$

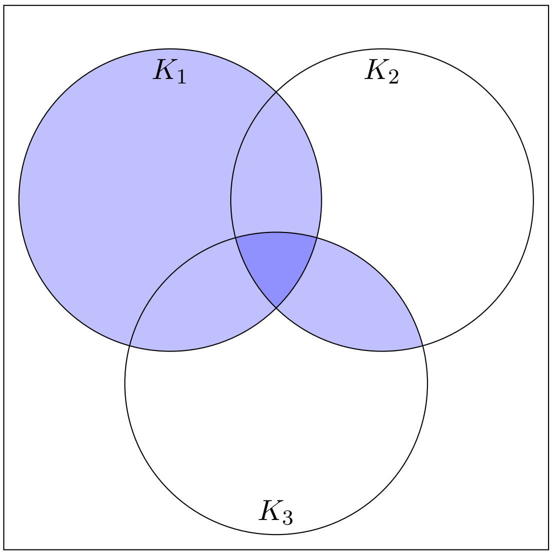
Vastaavasti oikeanpuoleiselle piirille saadaan

$$V = (K_1 \cup K_2) \cup (K_3 \cap K_4)$$

ja

$$E = \overline{V} = (\overline{K_1 \cup K_2}) \cap (\overline{K_3 \cap K_4}) = \overline{K_1} \cap \overline{K_2} \cap (\overline{K_3} \cup \overline{K_4}).$$

Kuvaan 2 on piirretty vasemmanpuoleisen piirin tapahtuma $V = K_1 \cup (K_2 \cap K_3)$.



Kuva 2: Vennin diagrammi tehtävän 1 tapahtumalle $V = K_1 \cup (K_2 \cap K_3)$ vasemmanpuoleisessa piirissä

b)Kysytään todennäköisyyttä $\mathbb{P}(V)$ molemmille piireille. Todennäköisyyden laskukaavan $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ ja tapahtumien K_i riippumattomuuden mukaan vasemmanpuoleiselle piirille

$$\begin{aligned}
\mathbb{P}(V) &= \mathbb{P}(K_1) + \mathbb{P}(K_2 \cap K_3) - \mathbb{P}(K_1 \cap K_2 \cap K_3) \\
&= \mathbb{P}(K_1) + \mathbb{P}(K_2)\mathbb{P}(K_2) - \mathbb{P}(K_1)\mathbb{P}(K_2)\mathbb{P}(K_3) \\
&= 0.9 + 0.9^2 - 0.9^3 = 0.981.
\end{aligned}$$

Kuten kuvan 2 Vennin diagrammista nähdään, tulee leikkaus $K_1 \cap K_2 \cap K_3$ kahteen kertaan, kun joukkojen K_1 ja $K_2 \cap K_3$ todennäköisyydet lasketaan yhteen. Siksi sen todennäköisyys täytyy lopuksi vähentää. Saman asian toki kertoo myös todennäköisyyden laskukaava joukkoyhdisteelle, jota yllä käytettiin.

Vastaavasti oikeanpuoleiselle piirille saadaan

$$\begin{aligned}
\mathbb{P}(V) &= \mathbb{P}(K_1 \cup K_2) + \mathbb{P}(K_3 \cap K_4) - \mathbb{P}((K_1 \cup K_2) \cap (K_3 \cap K_4)) \\
&= \mathbb{P}(K_1 \cup K_2) + \mathbb{P}(K_3 \cap K_4) - \mathbb{P}((K_1 \cap K_3 \cap K_4) \cup (K_2 \cap K_3 \cap K_4)) \\
&= \mathbb{P}(K_1) + \mathbb{P}(K_2) - \mathbb{P}(K_1 \cap K_2) + \mathbb{P}(K_3 \cap K_4) \\
&\quad - (\mathbb{P}(K_1 \cap K_3 \cap K_4) + \mathbb{P}(K_2 \cap K_3 \cap K_4) - \mathbb{P}(K_1 \cap K_2 \cap K_3 \cap K_4)) \\
&= 0.9 + 0.9 - 0.9^2 + 0.9^2 - (0.9^3 + 0.9^3 - 0.9^4) = 2 \cdot 0.9 + 0.9^4 - 2 \cdot 0.9^3 \\
&= 0.9981.
\end{aligned}$$

2. Oletetaan, että Lauri Markkanen pussittaa vapaaheiton 88 % todennäköisyydellä ja että hän saa 3 vapaaheittoa.

a) Määrää kolmen vapaaheiton tuloksia kuvastava otosavaruus, kun meitä kiinnostaa, pussittaako Markkanen pallon vai ei.
Ratkaisu: Merkitään

$$I = \text{"Markkasen heitto menee koriin"},$$
$$O = \text{"Heitto ei mene koriin"},$$

jolloin Markkasen kolmen vapaaheiton onnistumisia voidaan kuvastaa merkkijonoilla XYZ, missä $X, Y, Z \in \{I, O\}$. Koska yksittäisessä heitossa on vain 2 eri lopputulosta, on otosavaruudessa $2^3 = 8$ alkiota, jotka on helppo luetella. Tällä tavalla mallinnettuna otosavaruus on

$$S = \{OOO, OOI, OIO, IOO, OII, IOI, IIO, III\}.$$

Tämä ei ole suinkaan ainoa mahdollisuus otosavaruuden valitsemiseen. Tärkeintä on, että otosavaruus on selkeä ja informatiivinen.

b) Millä todennäköisyydellä Markkanen pussittaa kaikki vapaaheitot?
Ratkaisu: Merkitään tapahtumaa

$$B = \text{"Markkanen onnistuu kaikissa heitoissa"}.$$

Kun otosavaruus on mallinnettu selkeästi, voidaan sieltä poimia B:lle suotuisat alkeistapahtumat. Joukko-opillisesti

$$B = \{III\}.$$

Oletetaan, että heittojen lopputulokset ovat toisistaan riippumattomia, jolloin todennäköisyydeksi tulee

$$\mathbb{P}(B) = 0.88^3 \approx 68\,\%.$$

c) Millä todennäköisyydellä Markkanen pussittaa ainakin yhden vapaaheiton?
Ratkaisu: Merkitään tapahtumaa

$$C = \text{"Markkanen pussittaa ainakin yhden vapaaheiton"}.$$

Jälleen otosavaruudesta on helppo poimia suotuisat alkeistapahtumat ja laskea niiden todennäköisyyksien summa. Tässä todennäköisyys voidaan laskea helpommin komplementin avulla, sillä joukko-opillisesti

$$\overline{C} = \{OOO\}.$$

Oletetaan jälleen, että heittojen lopputulokset ovat toisistaan riippumattomia, jolloin samalla tavalla kuin b)-kohdassa saadaan

$$\mathbb{P}(C) = 1 - \mathbb{P}(\overline{C}) = 1 - (1 - 0.88)^3 \approx 99.8\%.$$

Kolmas tapa on tunnistaa, että tilanteeseen sopii binomijakauma, mutta siitä enemmän tulevilla viikoilla.

d) Millä todennäköisyydellä Markkanen pussittaa kaksi vapaaheittoa, jos hän on heittänyt ensimmäisen heiton ohi?
Ratkaisu: Koska tiedetään, että ensimmäinen heitto on mennyt ohi, niin tässä perusjoukkona (otosavaruutena) on

$$D = \{OOO, OOI, OIO, OII\}$$

ja kysytään alkeistapahtuman $\{OII\}$ todennäköisyyttä joukossa D. Jos O ja I olisivat yhtä todennäköisiä, saataisiin klassisella todennäköisyysmallilla todennäköisyydeksi 1/4. Nyt näin ei ole, joten emme voi menetellä edellä kuvatulla tavalla. Luontevin tapa ratkaista tehtävä on ehdollisen todennäköisyyden käyttäminen, josta enemmän loppuviikolla. Lasketaan tähän kuitenkin mallin vuoksi kysytty todennäköisyys $\mathbb{P}(\{OII\}|D)$, joksi saadaan

$$\begin{aligned}\mathbb{P}(\{OII\}|D) &= \frac{\mathbb{P}(\{OII\} \cap D)}{\mathbb{P}(D)} = \frac{\mathbb{P}(\{OII\})}{\mathbb{P}(D)} = \frac{0.12 \cdot 0.88^2}{0.12^3 + 2 \cdot 0.12^2 \cdot 0.88 + 0.12 \cdot 0.88^2} \\ &\approx 77\,\%.\end{aligned}$$

Koska emme periaatteessa tässä vaiheessa tiedä ehdollista todennäköisyyttä, ratkaistaan tämä vielä niin, ettei ehdollista todennäköisyyttä tarvita. Riippumattomuuden nojalla ensimmäisen heiton lopputuloksella ei ole merkitystä jäljellä olevien heittojen onnistumiseen. Meitä kiinnostaa ainoastaan, että onnistuuko Markkanen kahdessa jäljellä olevassa heitossa vai ei. Niinpä otosavaruudeksi voidaan ottaa kahden vapaaheiton lopputulosten joukko $\{OO, OI, IO, II\}$. Kysytään todennäköisyyttä $\mathbb{P}(\{II\}) = 0.88^2 \approx 77\,\%$.

3. Oletetaan, että 100 satunnaisesti valitun teekkarin joukossa 54 opiskelee fysiikkaa, 69 opiskelee matematiikkaa sekä 35 opiskelee matematiikka ja fysiikkaa. Merkitse seuraavat tapahtumat joukko-opillisesti ja määrää todennäköisyys, että opiskelija

   a) opiskelee matematiikkaa tai fysiikkaa.

   b) ei opiskele kumpaakaan näistä aineista.

   c) opiskelee matematiikkaa, muttei fysiikkaa.

Ratkaisu: Merkitään tapahtumat A="teekkari opiskelee fysiikkaa" ja B="teekkari opiskelee matematiikkaa". Otosavaruus S sisältää 100 teekkaria ($\#S = 100$) ja heistä 54 opiskelee fysiikkaa ($\#A = 54$), joten satunnaisesti valittu teekkari opiskelee fysiikkaa todennäköisyydellä $\mathbb{P}(A) = 54/100 = 0,54$. Vastaavasti matematiikan osalta $\#B = 69$ ja $\mathbb{P}(B) = 69/100 = 0,69$. Molempia opiskelee 35 ($\#(A \cap B) = 35$).

a) "opiskelee matematiikkaa tai fysiikkaa": $A \cup B$.

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0,54 + 0,69 - 0,35 = 0,88 = 88\%$$

b) "ei opiskele kumpaakaan": $\overline{A \cup B}$.

$$\mathbb{P}(\overline{A \cup B}) = 1 - \mathbb{P}(A \cup B) = 1 - 0,88 = 0,12 = 12\%$$

c) "opiskelee matematiikkaa, muttei fysiikkaa": $B \cap \overline{A} = B \setminus (A \cap B)$

$$\mathbb{P}(B \cap \overline{A}) = \frac{69 - 35}{100} = \frac{34}{100} = 0,34 = 34\%$$

4. Olkoot A ja B tapahtumia, joille $\mathbb{P}(A) = 0.3$, $\mathbb{P}(B) = 0.5$ ja $\mathbb{P}(A \cap B) = 0.24$.

   a) Laske todennäköisyydet $\mathbb{P}(A \cup B), \mathbb{P}(\overline{A} \cup B), \mathbb{P}(A \setminus B)$ ja $\mathbb{P}(\overline{A} \cup \overline{B})$.

   b) Ovatko tapahtumat A ja B riippumattomia? Entäpä A ja $\overline{B}$?

Ratkaisu: Käytetään hyväksi kaavakokoelmasta löytyviä todennäköisyyden peruskaavoja. Joukkoja kannattaa havainnollistaa myös Vennin diagrammien avulla.

a) Kaavasta

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \tag{1}$$

saadaan

$$\mathbb{P}(A \cup B) = 0.3 + 0.5 - 0.24 = 0.56.$$

Sijoittamalla kaavaan (1) A:n paikalle $\overline{A}$ saadaan

$$\begin{aligned}\mathbb{P}(\overline{A} \cup B) &= \mathbb{P}(\overline{A}) + \mathbb{P}(B) - \mathbb{P}(\overline{A} \cap B) = 1 - \mathbb{P}(A) + \mathbb{P}(B) - (\mathbb{P}(B) - \mathbb{P}(A \cap B)) \\ &= 1 - \mathbb{P}(A) + \mathbb{P}(A \cap B) \\ &= 0.7 + 0.24 = 0.94.\end{aligned}$$

Joukkoerotuksen $A \setminus B$ todennäköisyydeksi saadaan

$$\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = 0.06.$$

De Morganin kaavalla saadaan

$$\mathbb{P}(\overline{A} \cup \overline{B}) = 1 - \mathbb{P}(\overline{\overline{A} \cup \overline{B}}) = 1 - \mathbb{P}(A \cap B) = 0.76.$$

b) Tapahtumat A ja B eivät ole riippumattomia, sillä

$$\mathbb{P}(A \cap B) = 0.24 \neq 0.15 = 0.3 \cdot 0.5 = \mathbb{P}(A)\mathbb{P}(B).$$

Myöskään A ja $\overline{B}$ eivät ole riippumattomia, sillä

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A)\mathbb{P}(\overline{B}) \Leftrightarrow \mathbb{P}(A) - \mathbb{P}(A \cap B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) \Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Sama asia voidaan toki todeta myös laskemalla

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B) = 0.06$$

ja

$$\mathbb{P}(A)\mathbb{P}(\overline{B}) = 0.3 \cdot 0.5 = 0.15.$$

5. Kaksi tehdasta A ja B valmistaa radioita. Jokainen tehtaalla A valmistettu radio on viallinen todennäköisyydellä 0.05, kun taas tehtaalla B valmistettu radio on viallinen todennäköisyydellä 0.01. Ostetaan 2 radiota, jotka molemmat ovat yhtä todennäköisesti peräisin joko tehtaalta A tai tehtaalta B.

   a) Millä todennäköisyydellä molemmat ovat viallisia?

   b) Jos ensimmäinen radio, jonka testaat, on viallinen, niin millä todennäköisyydellä myös toinen on viallinen?

   c) Jos molemmat radiot todetaan viallisiksi, niin millä todennäköisyydellä ne olivat peräisin tehtaalta A?

Ratkaisu: Merkitään V_i = ”radio i on viallinen” ja E_i = ”radio i on ehjä”, $i = 1, 2$. Merkitään lisäksi A = ”radio on peräisin tehtaalta A”, jolloin $\overline{A}$ = ”radio on peräisin tehtaalta B”. Oletetaan, että radioiden viallisuus on toisistaan riippumatonta.

a) Kysytään todennäköisyyttä $\mathbb{P}(V_1 \cap V_2)$. Kokonaistodennäköisyyden mukaan

$$\mathbb{P}(V_1 \cap V_2) = \mathbb{P}(V_1 \cap V_2|A)\mathbb{P}(A) + \mathbb{P}(V_1 \cap V_2|\overline{A})\mathbb{P}(\overline{A}) = 0.05^2 \cdot \tfrac{1}{2} + 0.01^2 \cdot \tfrac{1}{2} = 0.0013.$$

b) Tiedetään, että V_1 toteutuu, ja kysytään ehdollista todennäköisyyttä $\mathbb{P}(V_2|V_1)$. Ehdollisen todennäköisyyden määritelmän mukaan

$$\mathbb{P}(V_2|V_1) = \frac{\mathbb{P}(V_1 \cap V_2)}{\mathbb{P}(V_1)}.$$

Osoittaja laskettiin jo a)-kohdassa, joten riittää laskea nimittäjä. Samalla tavalla kuin a)-kohdassa saadaan kokonaistodennäköisyyden mukaan

$$\mathbb{P}(V_1) = \mathbb{P}(V_1|A)\mathbb{P}(A) + \mathbb{P}(V_1|\overline{A})\mathbb{P}(\overline{A}) = 0.05 \cdot \tfrac{1}{2} + 0.01 \cdot \tfrac{1}{2} = 0.03.$$

Edellisen perusteella

$$\mathbb{P}(V_2|V_1) = \frac{0.0013}{0.03} = \frac{13}{300} \approx 4.3\,\%.$$

c) Kysytään todennäköisyyttä $\mathbb{P}(A|V_1 \cap V_2)$. Bayesin kaavan ja edellä lasketun mukaan

$$\mathbb{P}(A|V_1 \cap V_2) = \frac{\mathbb{P}(V_1 \cap V_2|A)\mathbb{P}(A)}{\mathbb{P}(V_1 \cap V_2)} = \frac{0.05^2 \cdot \frac{1}{2}}{0.0013} \approx 96\,\%.$$

Kannattaa piirtää tilannetta havainnollistava puukaavio, jossa ensimmäinen haarautuminen tapahtuu radioita valmistavan tehtaan suhteen.

6. Vakuutusyhtiö uskoo, että ihmiset voidaan jakaa kahteen ryhmään: tapaturma-alttiisiin ja ihmisiin, jotka eivät ole tapaturma-alttiita. Yhtiön mukaan tapaturma-altiille henkilölle sattuu jokin vahinko yhden vuoden aikana 40 % todennäköisyydellä, kun taas vahingon todennäköisyys on 20 % henkilölle, joka ei ole tapaturma-altis. Oletetaan, että 30 % ihmisistä on tapaturma-alttiita.

a) Millä todennäköisyydellä uudelle asiakkaalle sattuu vahinko vuoden sisällä?

b) Jos uudelle asiakkaalle sattuu vahinko vuoden sisällä, niin millä todennäköisyydellä hän on tapaturma-altis?

c) Ovatko onnettomuuden sattuminen ja tapaturma-alttius riippumattomia?

Ratkaisu: Merkitään

$$\begin{aligned} T &= \text{”asiakas on tapaturma-altis”}, \\ V &= \text{”asiakkaalle sattuu vahinko vuoden sisällä”}. \end{aligned}$$

Tiedetään todennäköisyydet

$$\begin{aligned} \mathbb{P}(T) &= 0.30 \Rightarrow \mathbb{P}(\overline{T}) = 0.70 \\ \mathbb{P}(V|T) &= 0.40, \\ \mathbb{P}(V|\overline{T}) &= 0.20. \end{aligned}$$

a) Kysytään todennäköisyyttä $\mathbb{P}(V)$, joka kokonaistodennäköisyyden mukaan on

$$\mathbb{P}(V) = \mathbb{P}(V|T)\mathbb{P}(T) + \mathbb{P}(V|\overline{T})\mathbb{P}(\overline{T}) = 0.40 \cdot 0.30 + 0.20 \cdot 0.70 = 26\,\%.$$

b) Kysytään todennäköisyyttä $\mathbb{P}(T|V)$, joka Bayesin kaavan ja a)-kohdan mukaan on

$$\mathbb{P}(T|V) = \frac{\mathbb{P}(T \cap V)}{\mathbb{P}(V)} = \frac{\mathbb{P}(V|T)\mathbb{P}(T)}{\mathbb{P}(V)} = \frac{0.40 \cdot 0.30}{0.26} \approx 46\,\%.$$

c) Kysytään tapahtumien V ja T riippumattomuutta. Tapahtumat *eivät ole riippumattomat*, sillä esimerkiksi $\mathbb{P}(V|T) = 0.40$, mutta a)-kohdan mukaan $\mathbb{P}(V) = 0.26$. Sama asia voidaan toki todeta laskemalla

$$\mathbb{P}(V \cap T) = \mathbb{P}(V|T)\mathbb{P}(T) = 0.40 \cdot 0.30 = 0.12 \neq 0.26 \cdot 0.30 = \mathbb{P}(V)\mathbb{P}(T).$$

Edelleen sama asia olisi voitu todeta ehdoista $\mathbb{P}(V|T) \neq \mathbb{P}(V|\overline{T})$ ja $0 < \mathbb{P}(T) < 1$.

7. Pumpun venttiilin toimintaa valvotaan automaattisella hälytysjärjestelmällä. Todennäköisyys sille, että järjestelmä hälyttää, kun venttiili ei toimi, on 0.98. Todennäköisyys sille, että järjestelmä ei hälytä, kun venttiili toimii, on 0.985. Todennäköisyys sille, että venttiili ei toimi, on 0.00001. Määrää todennäköisyys sille, että venttiili ei toimi, kun järjestelmä hälyttää.

Ratkaisu:
Merkitään

$$\begin{aligned} H &= \text{"järjestelmä hälyttää"}, \\ V &= \text{"venttiili toimii"}. \end{aligned}$$

Tiedetään todennäköisyydet $\mathbb{P}(H|\overline{V}) = 0.98$, $\mathbb{P}(\overline{H}|V) = 0.985$ ja $\mathbb{P}(\overline{V}) = 0.00001$. Kysytään todennäköisyyttä

$$\mathbb{P}(\overline{V}|H) = \frac{\mathbb{P}(\overline{V} \cap H)}{\mathbb{P}(H)}.$$

Lasketaan ensin todennäköisyys, että järjestelmä hälyttää. Kokonaistodennäköisyyden kaavan mukaan

$$\begin{aligned} \mathbb{P}(H) = \mathbb{P}(H|V)\mathbb{P}(V) + \mathbb{P}(H|\overline{V})\mathbb{P}(\overline{V}) &= (1 - \mathbb{P}(\overline{H}|V))\mathbb{P}(V) + \mathbb{P}(H|\overline{V})\mathbb{P}(\overline{V}) \\ &= 0.015 \cdot 0.99999 + 0.98 \cdot 0.00001 \approx 0.015. \end{aligned}$$

Kysytty todennäköisyys saadaan Bayesin kaavalla

$$\mathbb{P}(\overline{V}|H) = \frac{\mathbb{P}(H|\overline{V})\mathbb{P}(\overline{V})}{\mathbb{P}(H)} \approx 6.53 \cdot 10^{-4}.$$

Todennäköisyyttä kannattaa havainnollistaa myös puukaavion avulla. Tulos voi tuntua yllättävältä, vaikka järjestelmä toimii todennäköisyyksien mielessä kohtuullisen hyvin.

8. Erästä tietoliikennelinjaa pitkin pystyttiin lähettämään kahta erilaista merkkijonoa: 000 ja 111. Tiedettiin, että jonon 000 lähettämistodennäköisyys oli 0.3. Häiriöiden vuoksi kummankin jonon yksittäinen bitti vastaanotetaan oikein todennäköisyydellä 0.6. Lisäksi tiedetään, että bitit siirtyvät linjalla toisistaan riippumattomasti. Kumpi on todennäköisemmin lähetetty merkkijono, jos vastaanotetaan merkkijono 010?

Ratkaisu: Merkitään

$$\begin{aligned} L &= \text{"lähetetään 000"}, \\ V &= \text{"vastaanotaan 010"}, \\ A_i &= \text{"}i\text{:s bitti säilyy samana"}, \quad i = 1, 2, 3. \end{aligned}$$

Lasketaan todennäköisyys

$$\mathbb{P}(L|V) \overset{\text{Bayes}}{=} \frac{\mathbb{P}(V|L)\mathbb{P}(L)}{\mathbb{P}(V)}.$$

Kokonaistodennäköisyyden kaavan mukaan

$$\begin{aligned}\mathbb{P}(V) &= \mathbb{P}(V|L)\mathbb{P}(L) + \mathbb{P}(V|\overline{L})\mathbb{P}(\overline{L}) = \mathbb{P}(A_1 \cap \overline{A_2} \cap A_3)\mathbb{P}(L) + \mathbb{P}(\overline{A_1} \cap A_2 \cap \overline{A_3})\mathbb{P}(\overline{L}) \\ &= 0.6 \cdot 0.4 \cdot 0.6 \cdot 0.3 + 0.4 \cdot 0.6 \cdot 0.4 \cdot 0.7 \\ &= 0.1104.\end{aligned}$$

Näin ollen

$$\mathbb{P}(L|V) = \frac{0.6 \cdot 0.4 \cdot 0.6 \cdot 0.3}{0.1104} \approx 39\%,$$

eli todennäköisemmin lähetettiin 111.

9. Tunnetusti geneettinen perimä vaikuttaa useisiin sairauksiin, joten on tärkeää pystyä ennustamaan lasten riskiä saada sairautta aiheuttava perimä. Oletetaan yksinkertaisuuden vuoksi, että sairautta voidaan ennustaa yhden geeniparin avulla, jossa toinen geeni periytyy äidiltä ja toinen isältä. Geenillä on 2 muotoa: A, joka on dominoiva, ja a, joka on resessiivinen.

   Oletetaan, että henkilöllä on riski sairastua tautiin vain, jos geeniperimä on aa, ja että kaikki geeniparit ovat yhtä todennäköisiä. Oletetaan, että perheen isällä on taudin aiheuttava geneettinen riski ja että perheen äidillä ei ole riskiä. Jos perheeseen syntyy kaksi lasta, niin millä todennäköisyydellä

   a) kummallakaan lapsella ei ole riskiä sairastua tautiin, jos äidillä on AA?

   b) kummallakaan lapsella ei ole riskiä sairastua tautiin?

   c) äidillä on AA, jos kummallakaan lapsella ei ole riskiä sairastua tautiin?

Ratkaisu: Käytetään hieman epätäsmällisiä merkintöjä, mutta joista toivottavasti hahmottuu hyvin, mistä on kyse. Tiedetään, että isän perimä on aa ja että äidin perimä on jokin vaihtoehdoista aA, Aa tai AA. Merkitään Ä_{xy}, jos äidin perimä on xy, ja

$$E = \text{"äidillä ei ole riskiä sairastua tautiin"}.$$

Koska tiedetään, että E sattuu ja että kaikki geenivaihtoehdot ovat yhtä todennäköisiä, niin

$$\mathbb{P}_E(\text{Ä}_{AA}) = \mathbb{P}(\text{Ä}_{AA}|E) = \frac{\mathbb{P}(\text{Ä}_{AA} \cap E)}{\mathbb{P}(E)} = \frac{1/4}{3/4} = 1/3,$$

missä alaindeksi E tarkoittaa, että todennäköisyys lasketaan tapahtuman E suhteen. Ilman alaindeksiä varustettu $\mathbb{P}$ tarkoittaa, että todennäköisyysfunktio on määritelty kaikkien perimävaihtoehtojen joukossa.
Vastaavasti saadaan

$$\mathbb{P}_E(\text{Ä}_{Aa}) = \mathbb{P}_E(\text{Ä}_{aA}) = \frac{1}{3}.$$

Lapsen perimän hahmottamiseksi kannattaa laatia perimätaulukko

Äiti\Isä	a	a
A	Aa	Aa
a	aa	aa

josta näkyy, että todennäköisyys sille, että lapsella ei ole riskiä sairastua tautiin, on $1/2$, kun tiedetään, että äidin perimä on Aa (lapsi saa molemmilta vanhemmiltaan yhden geenin, jotka muodostavat lapsen geeniparin). Jos merkitään

$$R = \text{"lapsella ei ole riskiä sairastua tautiin"},$$

niin esimerkiksi

$$\mathbb{P}_E(R|\text{Ä}_{Aa}) = \frac{1}{2}.$$

Vastaavalla tavalla voidaan päätellä

$$\mathbb{P}_E(R|\text{Ä}_{AA}) = 1 \quad \text{ja} \quad \mathbb{P}_E(R|\text{Ä}_{aA}) = \frac{1}{2}.$$

Käytetään tapahtumalle ”kummallakaan lapsella ei ole riskiä sairastua tautiin” merkintää RR.

a) Kysytään todennäköisyyttä $\mathbb{P}_E(RR|\text{Ä}_{AA})$. Yllä olevan perusteella

$$\mathbb{P}_E(RR|\text{Ä}_{AA}) = 1,$$

sillä äidiltä tulee varmasti geeni A, joka takaa sen, ettei sairastumismahdollisuutta ole kummallakaan lapsella.

b) Kysytään todennäköisyyttä $\mathbb{P}_E(RR)$. Kokonaistodennäköisyyden mukaan

$$\begin{aligned}\mathbb{P}_E(RR) &= \mathbb{P}_E(RR|\text{Ä}_{AA})\mathbb{P}_E(\text{Ä}_{AA}) + \mathbb{P}_E(RR|\text{Ä}_{Aa})\mathbb{P}_E(\text{Ä}_{Aa}) + \mathbb{P}_E(RR|\text{Ä}_{aA})\mathbb{P}_E(\text{Ä}_{aA}) \\ &= 1 \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{2}.\end{aligned}$$

c) Kysytään todennäköisyyttä $\mathbb{P}_E(\text{Ä}_{AA}|RR)$. Bayesin kaavan mukaan

$$\mathbb{P}_E(\text{Ä}_{AA}|RR) = \frac{\mathbb{P}_E(\text{Ä}_{AA} \cap RR)}{\mathbb{P}_E(RR)} = \frac{\mathbb{P}_E(\text{Ä}_{AA})}{\mathbb{P}_E(RR)} \overset{\text{b)}}{=} \frac{1/3}{1/2} = \frac{2}{3}.$$

Todennäköisyyksiä kannattaa hahmotella myös puukaavion avulla.
Tässä tehtävässä alaindeksi E saattaa sekottaa merkinnällisesti, mutta hyvä puoli on siinä, että tulee myös merkinnällä kerrottua mikä on otosavaruus, jossa todennäköisyys lasketaan. Todennäköisyys riippuu aina valitusta otosavaruudesta, joten epäselvissä tapauksissa otosavaruus on syytä täsmentää.