Todennäköisyyslaskennan ja tilastotieteen peruskurssi

Esimerkkikokoelma 1

Aiheet:

Joukko-opin peruskäsitteet
Todennäköisyyslaskennan peruskäsitteet
Todennäköisyyslaskennan peruslaskusäännöt
Klassinen todennäköisyys ja kombinatoriikka
Todennäköisyyden aksioomat
Kokonaistodennäköisyyden ja Bayesin kaavat

Avainsanat:

Alkeistapahtuma
Alkio
Bayesin kaava
Binomikaava
Binomikerroin
Ehdollinen todennäköisyys
Ehtotapahtuma
Empiirinen todennäköisyys
Frekvenssi
Frekvenssitulkinta
Jono
Joukko
Kertolaskuperiaate
Kertoma
Klassinen todennäköisyys
Koetoisto
Kokonaistodennäköisyyden kaava
Kombinaatio
Kombinatoriikka
Komplementti
Komplementtitapahtuma
Leikkaus
Lukumääräfunktio
Mahdoton tapahtuma
Mitta
Osajono
Osajoukko
Otanta
Otanta palauttaen
Otanta palauttamatta
Otosavaruus
Permutaatio
Perusjoukko
Pistevieraus
Riippumattomuus
Sattuma
Satunnaisilmiö
Satunnaiskoe
Satunnaisotanta
Suhteellinen frekvenssi
Suhteellinen osuus
Suotuisa alkeistapahtuma
Tapahtuma
Todennäköisyys
Toisensa poissulkevuus
Tyhjä joukko
Tulosääntö
Unioni
Variaatio
Varma tapahtuma
Yhdiste
Yhteenlaskuperiaate
Yhteenlaskusääntö

Joukko-oppia

Joukko ja sen alkiot

Joukko voidaan määritellä luettelemalla sen alkiot tai kertomalla sääntö, jonka avulla voidaan päätelllä, kuuluuko valittu alkio joukkoon vai ei. Matematiikassa joukko määritellään usein antamalla ehto, jonka joukon alkioiden on toteutettava. Jos alkio x kuuluu joukkoon A, niin merkitsemme $x \in A$. Vastaavasti, jos x ei kuulu joukkoon A, niin merkitsemme $x \notin A$. Jos A on niiden perusjoukon S alkioiden x joukko, jotka toteuttavat ehdon $P(x)$ eli joille lause $P(x)$ on tosi, niin merkitsemme

$$A = \{x \in S \mid P(x)\}.$$

Jos joukko A on äärellinen ja sen alkiot ovat $a_1, a_2, \ldots, a_n$, niin merkitsemme $A = \{a_1, a_2, \ldots, a_n\}$.

Osajoukko ja tyhjä joukko

Jos jokainen joukon A alkio on myös joukon B alkio, on joukko A joukon B *osajoukko* ja merkitsemme $A \subset B$. Joukko on *tyhjä*, jos siinä ei ole yhtään alkiota. Merkitsemme tyhjää joukkoa symbolilla $\varnothing$. Tyhjä joukko on kaikkien joukkojen osajoukko.

Joukkojen perusoperaatiot

Olkoot A ja B perusjoukon S osajoukkoja. Joukkojen A ja B *yhdiste* $A \cup B$ on niiden perusjoukon alkioiden joukko, jotka kuuluvat joukkoon A tai joukkoon B (tai molempiin):

$$A \cup B = \{x \mid x \in A \ \mathit{tai}\ x \in B\}.$$

Joukkojen A ja B *leikkaus* $A \cap B$ on niiden alkioiden joukko, jotka kuuluvat joukkoon A ja joukkoon B:

$$A \cap B = \{x \in S \mid x \in A \ \mathit{ja}\ x \in B\}.$$

Jos $A \cap B = \varnothing$, niin A ja B ovat *erilliset*.

Joukon A *komplementti* A^c on niiden alkioiden joukko, jotka eivät kuulu joukkoon A:

$$A^c = \{x \in S \mid x \notin A\}.$$

Joukkojen A ja B *erotus* $A \backslash B$ on niiden alkioiden joukko, jotka kuuluvat joukkoon A, mutta eivät kuulu joukkoon B:

$$A \backslash B = \{x \mid x \in A \ \mathit{ja}\ x \notin B\}.$$

Perusjoukon S osajoukkojen $A_1, A_2, \ldots$ *yhdiste* on

$$\bigcup_{i=1}^{\infty} A = \{x \in S \mid \mathit{on\ olemassa}\ i \text{ siten, että } x \in A_i\}$$

ja *leikkaus* on

$$\bigcap_{i=1}^{\infty} A_i = \{x \in S \mid x \in A_i \ \mathit{kaikilla}\ i\}.$$

Todennäköisyyden käsite

Satunnaisilmiön matemaattinen mallintamien

Satunaisilmiö voi realisoitua usealla eri tavalla. Satunnaisilmiön realisaatio voi olla valitun havainnoitsijan näkökulmasta jossain määrin tiedossa, mutta yleensä ilmiön realisaatiota ei täysin tarkkaan tunneta tai voida ennustaa. *Stokastiikka* on tieteenala, jossa satunnaisilmiöitä mallinnetaan ja analysoidaan matemaattisin keinoin. Satunnaisilmiön matemaattista mallia kutsutaan usein *stokastiseksi malliksi*.

Satunnaisilmiön stokastinen malli rakentuu seuraaviin käsitteisiin:

(i) *Realisaatiot* eli *otospisteet* ovat satunnaisilmiön mahdollisia toteumia.

(ii) *Otosavaruus* on kaikkien mahdollisten realisaatioiden joukko.

(iii) *Tapahtumat* ovat otosavaruuden osajoukkoja. Tapahtuma *A sattuu*, jos satunnaisilmiön realisaatio $s \in A$.

Realisaatiota s vastaava *alkeistapahtuma* on yhden alkion joukko $\{s\}$.

Stokastiikan ja joukko-opin peruskäsitteet vastaavat seuraavalla tavalla toisiaan:

Otosavaruus	↔	Perusjoukko
Realisaatio	↔	Perusjoukon alkio
Tapahtuma	↔	Perusjoukon osajoukko

Todennäköisyys

Todennäköisyysjakauma eli *todennäköisyysmitta* on kuvaus, joka liittää jokaiseen otosavaruuden tapahtumaan $A \subset S$ reaaliluvun Pr(A), ja toteuttaa ehdot:

(i) Varman tapahtuman todennäköisyys on $\Pr(S) = 1$.

(ii) Jokaiselle tapahtumalle pätee $0 \le \Pr(A) \le 1$.

(iii) Mille tahansa äärelliselle tai äärettömälle jonolle $A_1, A_2, \ldots$ toisensa poissulkevia tapahtumia (eli $A_i \cap A_j = \varnothing$ kun $i \neq j$) pätee

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

Symmetrinen todennäköisyys ja diskreetti tasajakauma

Jos äärellisen otosavaruuden $S = \{s_1, s_2, \ldots, s_n\}$ kaikki realisaatiot ovat *symmetrisiä* eli yhtä todennäköisiä, niin tapahtuman $A \subset S$ todennäköisyys saadaan kaavasta

$$\Pr(A) = \frac{n(A)}{n(S)},$$

missä

$n(A)$ = tapahtumaan A sisältyvien realisaatioiden lukumäärä

ja

$n(S)$ = kaikkien mahdollisten realisaatioiden lukumäärä.

Yllä määritelty todennäköisyysmitta $A \to n(A)/n(S)$ on äärellisen joukon S *tasajakauma*.

Todennäköisyyslaskennan peruslaskusäännöt

Tapahtumien yhdisteleminen

Stokastiikan ja joukko-opin *perusoperaatiot* vastaavat seuraavalla tavalla toisiaan:

"A *ei* satu"

$\leftrightarrow$ Joukon A *komplementti* A^c

"A tai B sattuu"

$\leftrightarrow$ Joukkojen A ja B *yhdiste* $A \cup B$

"A ja B sattuvat"

$\leftrightarrow$ Joukkojen A ja B *leikkaus* $A \cap B$

"A sattuu mutta B ei"

$\leftrightarrow$ Joukkojen A ja B *erotus* $A \setminus B$

Vastakohdan todennäköisyys

Tapahtuman A *vastakohdan*

$$A^c = \{x \in S \mid x \notin A\}$$

todennäköisyys on

$$\Pr(A^c) = 1 - \Pr(A).$$

Yleinen yhteenlaskusääntö

Yleisen yhteenlaskusäännön mukaan

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

Toisensa poissulkevien tapahtumien yhteenlaskusääntö

Tapahtumat A ja B *poissulkevat toisensa*, jos ne eivät voi sattua samanaikaisesti, eli $A \cap B = \varnothing$. Siten toisensa poissulkevat tapahtumat ovat joukkoina erilliset. Jos tapahtumat A ja B ovat toisensa poissulkevia, niin pätee *toisensa poissulkevien tapahtumien yhteenlaskusääntö*

$$\Pr(A \cup B) = \Pr(A) + \Pr(B).$$

Tapahtumat A_1, A_2, … *poissulkevat toisensa*, jos vain yksi niistä voi sattua. Tällöin $A_i \cap A_j = \varnothing$ aina kun $i \neq j$. Tällöin pätee *toisensa poissulkevien tapahtumien yhteenlaskusääntö*

$$\Pr(A_1 \cup A_2 \cup \cdots \cup A_k) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_k).$$

Ehdollinen todennäköisyys

Tapahtuman A *ehdollinen todennäköisyys* tapahtuman B sattuessa määritellään kaavalla

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \quad \text{kun } \Pr(\mathrm{B}) > 0,$$

missä $\Pr(A \cap B)$ on todennäköisyys, että A ja B sattuvat. Kun $\Pr(B)=0$, jätetään $\Pr(A|B)$ määrittelemättä.

Yleinen tulosääntö

Yleisen tulosäännön mukaan

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B).$$

Monen tapahtuman yleisen tulosäännön mukaan

$$\begin{aligned}&\Pr(A_1 \cap A_2 \cap \cdots \cap A_k)\\ &\quad = \Pr(A_1) \times \Pr(A_2 | A_1) \times \Pr(A_3 | A_1 \cap A_2) \times \cdots \times \Pr(A_k | A_1 \cap A_2 \cap \cdots \cap A_{k-1})\end{aligned}$$

Tilastollinen riippumattomuus ja riippumattomien tapahtumien tulosääntö

Tapahtumat A ja B ovat tilastollisesti *riippumattomat*, jos

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

pätee. Yo. riippumattomien tapahtumien tulosääntö on yhtäpitävä sen kanssa, että

$$\Pr(A \mid B) = \Pr(A).$$

Tarkastellaan tapahtumia $A_1, A_2, \ldots , A_k$. Jos tapahtumat $A_1, A_2, \ldots , A_k$ ovat *riippumattomat*, niin pätee *riippumattomien tapahtumien tulosäännön* yleistys

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_k) = \Pr(A_1)\Pr(A_2)\Pr(A_3)\cdots\Pr(A_k).$$

Satunnaisotanta ja tulosääntö

Yksinkertaisessa satunnaisotannassa äärellisestä perusjoukosta S poimitaan umpimähkään alkioita järjestettyyn listaan tai järjestämättömään osajoukkoon yksi kerrallaan. Poiminnan tulosta kutsutaan *otokseksi* ja arvonnassa käytettyä menetelmää *otantamenetelmäksi*.

(i) Jos otanta tehdään *palauttamatta,* eli poimittua alkiota ei palauteta takaisin perusjoukkoon, ovat yksittäisten poimintojen tulokset keskenään tilastollisesti riippuvia, ja poimintatodennäköisyyksiä määrättäessä on sovellettava yleistä tulosääntöä.

(ii) Jos otanta tehdään *palauttaen,* eli poimittu alkio palautetaan aina takaisin perusjoukkoon, ovat yksittäisten poimintojen tulokset toisistaan riippumattomia, ja poimintatodennäköisyyksiä määrättäessä voidaan soveltaa riippumattomien tapahtumien tulosääntöä.

Kombinatoriikkaa

Satunnaisilmiön realisaatioiden lukumäärien laskeminen on usein vaikea tehtävä ja apuna tarvitaan *kombinatoriikaksi* kutsuttua matematiikan osa-aluetta. Tässä esitettävät kombinatoriikan kaavat voidaan johtaa käyttäen kahta perusperiaatetta, *kertolaskuperiaatetta* ja *yhteenlaskuperiaatetta*.

Kertolaskuperiaate

Oletetaan, että operaatio M voidaan suorittaa m eri tavalla ja operaatio N voidaan suorittaa n eri tavalla. Oletetaan lisäksi, että operaatiot M ja N voidaan suorittaa toisistaan riippumatta. Tällöin *yhdistetty operaatio*

”Suoritetaan operaatio M *ja* operaatio N”

voidaan suorittaa $m \times n$ eri tavalla.

Yhteenlaskuperiaate

Oletetaan, että operaatio M voidaan suorittaa m eri tavalla ja operaatio N voidaan suorittaa n eri tavalla. Oletetaan lisäksi, että operaatiot M ja N ovat *toisensa poissulkevia*. Tällöin *yhdistetty operaatio*

”Suoritetaan operaatio M *tai* operaatio N”

voidaan suorittaa $(m + n)$ eri tavalla.

Järjestettyjen listojen lukumäärä

Joukosta S, jossa on n alkiota, voidaan poimia (palauttamatta) k alkiota *järjestettyyn listaan*

$$\frac{n!}{(n-k)!} = n \times (n-1) \times \cdots \times (n-k+1)$$

tavalla. Näin muodostettuja listoja kutsutaan joukon S *k-permutaatioiksi*.

Joukon, jossa on n alkiota, kaikki alkiot voidaan järjestää listaan $n! = n\times(n-1)\times \cdots \times 2\times 1$ tavalla.

Tietyn kokoisten osajoukkojen lukumäärä

Joukosta S, jossa on n alkiota, voidaan valita k:n alkion (järjestämätön) osajoukko *binomikertoimen*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

ilmaisemalla määrällä tapoja. Tällä tavoin muodostettuja osajoukkoja kutsutaan yleisesti joukon S *k-kombinaatioiksi*.

Kaikkien osajoukkojen lukumäärä

Joukon S, joka sisältää n alkiota, kaikkien (järjestämättömien) osajoukkojen lukumäärä on 2^n. *Binomikaavan*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

avulla tämä lukumäärä voidaan esittää muodossa

$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n-1} + \binom{n}{n}.$$

Ositusten lukumäärä

Joukko S, joka sisältää n alkiota, ositetaan erillisiin osajoukkoihin $A_1,\ldots,A_k$ niin, että joukossa A_i on n_i alkiota, missä positiiviset kokonaisluvut n_i toteuttavat ehdon

$$n_1 + n_2 + \cdots + n_k = n.$$

Kuinka monella eri tavalla tällainen ositus voidaan tehdä? Vastauksen antaa *multinomikerroin*

$$\binom{n}{n_1\ n_2 \cdots n_k} = \frac{n!}{n_1!n_2!\cdots n_k!}.$$

Huomaa, että jos $k = 2$, saadaan binomikerroin

$$\binom{n}{n_1\ n_2} = \frac{n!}{n_1!n_2!} = \binom{n}{n_1} = \binom{n}{n_2},$$

missä $n_1 + n_2 = n$.

Kokonaistodennäköisyyden ja Bayesin kaavat

Ositus

Joukot $B_1, B_2, \ldots, B_n$ muodostavat joukon S *osituksen*, jos ne ovat toisensa poisulkevat ja niiden yhdiste on S, eli pätee

(i) $B_i \cap B_j = \varnothing,\ i \neq j,$

(ii) $S = B_1 \cup B_2 \cup \cdots \cup B_n.$

Kokonaistodennäköisyyden kaava

Oletetaan, että joukot $B_1, B_2, \ldots, B_n$ muodostavat otosavaruuden S osituksen. Tällöin pätee *kokonaistodennäköisyyden kaava*

$$\Pr(A) = \sum_{i=1}^{n} \Pr(B_i)\Pr(A \mid B_i).$$

Bayesin kaava

Oletetaan, että joukot $B_1, B_2, \ldots, B_n$ muodostavat otosavaruuden S osituksen. Tällöin pätee *Bayesin kaava*

$$\Pr(B_i \mid A) = \frac{\Pr(B_i)\Pr(A \mid B_i)}{\sum_{i=1}^{n} \Pr(B_i)\Pr(A \mid B_i)},\ i = 1,2,\ldots,n.$$

Esimerkki 1.1.

Symmetristä noppaa heitettäessä jokaisella silmäluvulla on sama todennäköisyys. Jos symmetristä noppaa heitetään kaksi kertaa, kaikkien mahdollisten realisaatioiden muodostama otosavaruus on

$$S = \{(x, y) \mid x = 1, 2, 3, 4, 5, 6\ ; y = 1, 2, 3, 4, 5, 6\},$$

missä x on ensimmäisen heiton ja y toisen heiton tulos.

Otosavaruutta S voidaan kuvata seuraavalla taulukolla:

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

Oletetaan, että 1. heiton tulos ei vaikuta siihen, mitä tulee tulokseksi 2. heitosta. Tällöin voimme olettaa, että jokaisella silmälukuparilla (x, y) on sama todennäköisyys realisoitua.

Määritellään tapahtumat

$$A = \{(x, y) \in S \mid y = 3\},$$

$$B = \{(x, y) \in S \mid x \geq 5\},$$

$$C = \{(x, y) \in S \mid x + y = 6\},$$

$$D = \{(x, y) \in S \mid x - y = 3\},$$

$$E = \{(x, y) \in S \mid y - x \geq 3\}.$$

Merkitse otosavaruutta S kuvaavaan taulukkoon seuraavat tapahtumat:

(a) A, B, C, D, E.

(b) $A \cup C$ = Tapahtumien A ja C yhdiste.

(c) $B \cap E$ = Tapahtumien B ja E leikkaus.

Esimerkki 1.1. – Mitä opimme?

Esimerkissä 1.1. tarkastellaan *tapahtumien määrittelemistä* ja *havainnollistamista* sekä joukko-opin perusoperaatioita *yhdiste* ja *leikkaus*.

Esimerkki 1.1. – Ratkaisu:

Esimerkin 1.1. tapahtumat on varjostettu otosavaruutta S kuvaaviin taulukoihin alla.

(a) $A = \{(x, y) \in S \mid y = 3\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

$B = \{(x, y) \in S \mid x \geq 5\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

$C = \{(x, y) \in S \mid x + y = 6\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

$D = \{(x, y) \in S \mid x - y = 3\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

$E = \{(x, y) \in S \mid y - x \geq 3\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

(b) $A \cup C = \{(x, y) \in S \mid (x, y) \in A \text{ tai } (x, y) \in C\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

(c) $B \cap E = \{(x, y) \in S \mid (x, y) \in B \text{ ja } (x, y) \in E\} = \varnothing$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

Esimerkki 1.2.

Esimerkki 1.2. on jatkoa esimerkille 1.1. Merkitse otosavaruutta S kuvaavaan taulukkoon seuraavat tapahtumat:

(a) D^c = Tapahtuman D komplementti.

(b) $B \setminus C$ = Tapahtumien B ja C erotus.

(c) $C \setminus B$ = Tapahtumien C ja B erotus.

Esimerkki 1.2. – Mitä opimme?

Esimerkki 1.2. on jatkoa esimerkille 1.1. ja siinä tarkastellaan *tapahtumien määrittelemistä* ja *havainnollistamista* sekä joukko-opin perusoperaatioita *komplementti* ja *erotus*.

Esimerkki 1.2. – Ratkaisu:

Esimerkissä 1.2. määritellyt joukot on merkitty alla varjostettuina otosavaruutta S kuvaaviin taulukoihin.

(a) $D^c = \{(x, y) \in S \mid (x, y) \notin D\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

(b) $B \setminus C \;= \{(x, y) \in S \mid (x, y) \in B \text{ ja } (x, y) \notin C\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

(c) $C \setminus B \;= \{(x, y) \in S \mid (x, y) \in C \text{ ja } (x, y) \notin B\}$

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

Esimerkki 1.3.

Esimerkki 1.3. on jatkoa esimerkille 1.1. Määrää todennäköisyydet esimerkin 1.1. kohdissa (a) – (c) määritellyille tapahtumille.

Esimerkki 1.3. – Mitä opimme?

Esimerkki 1.3. on jatkoa esimerkille 1.1. ja siinä tarkastellaan *todennäköisyyslaskennan* ja *joukko-opin peruskäsiteiden* ja *-operaatioiden vastaavuutta* sekä *symmetrisen todennäköisyyden* käsitettä.

Esimerkki 1.3. – Ratkaisu:

Jos symmetristä noppaa heitetään yhden kerran, jokaisella silmäluvulla on sama todennäköisyys realisoitua. Jos siis x on nopanheiton tulos, niin symmetrisen nopan tapauksessa

$$\Pr(\{x\}) = 1/6, \quad x = 1, 2, 3, 4, 5, 6.$$

Jos symmetristä noppaa heitetään kaksi kertaa, on järkevää ajatella, että jokaisella silmälukujen parilla (x,y), missä

x = tulos 1. nopan heitosta

y = tulos 2. nopan heitosta

on sama todennäköisyys realisoitua. Tämä johtuu siitä, että 1. heiton tulos ei vaikuta 2. heiton tulokseen. Siten voimme pitää silmälukujen pareja (x, y) ovat symmetrisinä, jolloin

$$\Pr(\{(x, y)\}) = 1/36 , \quad x = 1, 2, 3, 4, 5, 6; \quad y = 1, 2, 3, 4, 5, 6.$$

Symmetrian perusteella tapahtuman M todennäköisyys $\Pr(M)$ saadaan diskreetistä tasajakaumasta

$$\Pr(M) = \frac{n(M)}{n(S)},$$

missä

$n(M)$ = tapahtumaan M sisältyvien realisaatioiden lukumäärä,

$n(S)$ = kaikkien mahdollisten realisaatioiden lukumäärä.

Esimerkin 1.1. otosavaruus on

$$S = \{(x, y) \mid x = 1, 2, 3, 4, 5, 6 \text{ ja } y = 1, 2, 3, 4, 5, 6\},$$

missä x = 1. heiton tulos ja y = 2. heiton tulos.

Otosavaruudessa S on 36 alkiota, joten

$$n(S) = 36.$$

(a) Koska $n(A) = 6$, niin

$$\Pr(A) = n(A)/n(S) = 6/36 = 1/6.$$

Koska $n(B) = 12$, niin

$$\Pr(B) = n(B)/n(S) = 12/36 = 1/3.$$

Koska $n(C) = 5$, niin

$\Pr(C) = n(C)/n(S) = 5/36.$

Koska $n(D) = 3$, niin

$\Pr(D) = n(D)/n(S) = 3/36 = 1/12.$

Koska $n(E) = 6$, niin

$\Pr(E) = n(E)/n(S) = 6/36 = 1/6.$

(b) Koska $n(A\cup C) = 10$, niin

$\Pr(A\cup C) = n(A\cup C)/n(S) = 10/36 = 5/18.$

Sama tulos saadaan johdettua myös *yleisen yhteenlaskusäännön*

$\Pr(A\cup C) = \Pr(A) + \Pr(C) - \Pr(A\cap C).$

avulla: Kohdan (a) mukaan

$\Pr(A) = 6/36.$

$\Pr(C) = 5/36.$

ja koska

$A\cap C = \{(3,3)\},$

niin

$\Pr(A\cap C) = 1/36.$

Siten

$\Pr(A\cup C) = \Pr(A) + \Pr(C) - \Pr(A\cap C) = 6/36 + 5/36 - 1/36 = 10/36.$

(c) Koska $B \cap E = \varnothing$, niin

$\Pr(B \cap E) = 0.$

Esimerkki 1.4.

Esimerkki 1.4. on jatkoa esimerkille 1.2. Määrää todennäköisyydet esimerkin 1.2. kohdissa (a) – (c) määritellyille tapahtumille.

Esimerkki 1.4. – Mitä opimme?

Esimerkki 1.4. on jatkoa esimerkille 1.2. ja siinä tarkastellaan *todennäköisyyslaskennan* ja *joukko-opin peruskäsiteiden* ja *-operaatioiden vastaavuutta* sekä *symmetrisen todennäköisyyden* käsitettä; ks. myös esimerkkiä 1.3.

Esimerkki 1.4. – Ratkaisu:

Symmetrian perusteella tapahtuman M todennäköisyys $\Pr(M)$ saadaan diskreetistä tasajakaumasta

$$\Pr(M) = \frac{n(M)}{n(S)}$$

missä

$n(M)$ = tapahtumaan M sisältyvien realisaatioiden lukumäärä,

$n(S)$ = kaikkien mahdollisten realisaatioiden lukumäärä.

Esimerkin 1.1. otosavaruus on

$$S = \{(x, y) \mid x = 1, 2, 3, 4, 5, 6 \text{ ja } y = 1, 2, 3, 4, 5, 6\}$$

jossa x = 1. heiton tulos ja y = 2. heiton tulos.

Otosavaruudessa S on 36 alkiota, joten

$$n(S) = 36.$$

(a) Koska $n(D^c) = 33$, niin

$$\Pr(D^c) = n(D^c)/n(S) = 33/36 = 11/12$$

Sama tulos saadaan johdettua myös *vastakohdan todennäköisyyden* kaavan

$$\Pr(D^c) = 1 - \Pr(D)$$

avulla: Koska

$$\Pr(D) = 3/36 = 1/12$$

niin

$$\Pr(D^c) = 1 - \Pr(D) = 1 - 1/12 = 11/12$$

(b) Koska $n(B \setminus C) = 11$, niin

$$\Pr(B \setminus C) = 11/36$$

(c) Koska $n(C \setminus B) = 4$, niin

$$\Pr(C \setminus B) = 4/36 = 1/9$$

Esimerkki 1.5.

Esimerkki 1.5. on jatkoa esimerkille 1.1. Tarkastellaan 1. ja 2. nopanheiton silmälukujen erotusta

$$z = x - y$$

missä x = 1. heiton tulos ja y = 2. heiton tulos. Määritellään lisäksi tapahtumat

A = {1. nopalla saadaan 5},

B = {2. nopalla saadaan 5},

C = {Erotus on 4}.

(a) Määrää silmälukujen erotuksen $z = x - y$ otosavaruus.

(b) Määrää ehdollinen todennäköisyys

$$\Pr(A \mid B)$$

ja vertaa sitä tapahtuman A todennäköisyyteen. Ovatko tapahtumat A ja B riippumattomat?

(c) Määrää ehdollinen todennäköisyys

$$\Pr(C \mid A)$$

ja vertaa sitä tapahtuman C todennäköisyyteen. Ovatko tapahtumat C ja A riippumattomat?

(d) Määrää ehdollinen todennäköisyys

$$\Pr(C \mid B)$$

ja vertaa sitä tapahtuman C todennäköisyyteen. Ovatko tapahtumat C ja B riippumattomat?

Esimerkki 1.5. – Mitä opimme?

Esimerkki 1.5. on jatkoa esimerkille 1.1. ja siinä tarkastellaan *ehdollisen todennäköisyyden* ja *tilastollisen riippumattomuuden* käsitteitä.

Esimerkki 1.5. – Ratkaisu:

1. nopanheiton tulokseen x liittyvä otosavaruus: $\{1, 2, 3, 4, 5, 6\}$

2. nopanheiton tulokseen y liittyvä otosavaruus: $\{1, 2, 3, 4, 5, 6\}$

Muodostetaan silmälukujen x ja y mahdollisia erotuksia $z = x - y$ kuvaava aputaulukko:

$z = x - y$		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	0	1	2	3	4	5
	2	–1	0	1	2	3	4
	3	–2	–1	0	1	2	3
	4	–3	–2	–1	0	1	2
	5	–4	–3	–2	–1	0	1
	6	–5	–4	–3	–2	–1	0

(a) Aputaulukosta nähdään, että kahden nopanheiton silmälukujen erotuksen

$$z = x - y$$

otosavaruus on

$$\{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}.$$

Symmetrian perusteella (ks. esimerkki 1.3.) aputaulukosta saadaan erotuksille $z = x - y$ seuraavat todennäköisyydet:

z	–5	–4	–3	–2	–1	0	1	2	3	4	5
Pr	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

(b) Jos käytämme apuna esimerkin 1.1. otosavaruutta

$$S = \{(x, y) \mid x = 1, 2, 3, 4, 5, 6 \text{ ja } y = 1, 2, 3, 4, 5, 6\}$$

kuvaavaa taulukkoa

(x, y)		1. heiton tulos x					
		1	2	3	4	5	6
2. heiton tulos y	1	(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
	2	(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
	3	(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
	4	(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
	5	(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
	6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

niin diskreetin tasajakauman avulla on helppo nähdä, että

$$\Pr(A) = \Pr(\{\text{1. nopalla saadaan 5}\}) = 6/36 = 1/6$$

$$\Pr(B) = \Pr(\{\text{2. nopalla saadaan 5}\}) = 6/36 = 1/6$$

Koska

$$A \cap B = \{\text{1. nopalla saadaan 5 } \textit{ja} \text{ 2. nopalla saadaan 5}\}$$

niin *diskreetin tasajakauman määritelmän* nojalla

$\Pr(A \cap B) = 1/36.$

Siten *ehdollisen todennäköisyyden määritelmän* perusteella nähdään, että

$\Pr(A|B) = \Pr(A \cap B)/\Pr(B) = (1/36)/(1/6) = 1/6$

Koska

$\Pr(A|B) = \Pr(A) = 1/6$

tapahtumat A ja B *ovat riippumattomia*.

Tehtävän voi ratkaista myös käyttämällä apuna ennen (a)-kohdan ratkaisua esitettyä aputaulukkoa rajoittumalla tarkastelemaan niitä soluja, joissa 2. nopalla saadaan 5.

Näitä soluja on 6 ja täsmälleen yksi niistä vastaa sitä, että 1. nopalla on saatu 5. Siten suoraan diskreetin tasajakauman määritelmästä nähdään, että

$\Pr(A|B) = 1/6.$

Huomautus:

Vaikka tehtävä voidaan ratkaista käyttämällä yo. aputaulukkoa, on syytä oppia käyttämään ehdollisen todennäköisyyden kaavaa. Realisaatioiden taulukointi ja niiden lukumäärien laskeminen on hankalaa, jos otosavaruus on iso, ja se on jopa mahdotonta, kun otosavaruus on ääretön.

(c) Kohdassa (c) kysytään *ehdollista todennäköisyyttä* tapahtumalle

$C = \{(x, y) \in S \mid z = x - y = 4\},$

kun tapahtuma

$A = \{x = 5\}$

on sattunut.

Kohdan (b) otosavaruutta S kuvaavasta taulukosta on helppo nähdä seuraavaa: Jos A on sattunut, jäljellä on 6 vaihtoehtoa, joista erotus $z = x - y$ voi saada arvon 4 vain yhdellä tavalla, jos y saanut arvon 1. Siten

$\Pr(C|A) = 1/6.$

Sama tulos saadaan myös *ehdollisen todennäköisyyden määritelmän* perusteella:

$\Pr(C|A) = \Pr(C \cap A)/\Pr(A) = (1/36)/(1/6) = 1/6.$

Koska

$\Pr(C) = 2/36 \neq \Pr(C|A),$

tapahtumat C ja A *eivät ole riippumattomia*.

(d) Kohdassa (d) kysytään *ehdollista todennäköisyyttä* tapahtumalle

$$C = \{(x, y) \in S \mid z = x - y = 4\},$$

kun tapahtuma

$$B = \{y = 4\}$$

on sattunut.

Kohdan (b) otosavaruutta S kuvaavasta taulukosta on helppo nähdä seuraavaa: Jos B on sattunut, erotus $z = x - y$ ei voi saada arvoa 4. Siten

$$\Pr(C|B) = 0.$$

Sama tulos saadaan myös *ehdollisen todennäköisyyden määritelmän* perusteella:

$$\Pr(C|B) = \Pr(C \cap B)/\Pr(B) = (0/36)/(1/6) = 0/6 = 0.$$

Esimerkki 1.6.

Olkoot $\Pr(A) = 0.2$ ja $\Pr(B) = 0.6$. Määrää tapahtuman $A \cup B$ todennäköisyys, kun

(a) $\Pr(A \cap B) = 0.1$.

(b) A ja B ovat toisensa poissulkevia.

Esimerkki 1.6. – Mitä opimme?

Esimerkissä 1.6. tarkastellaan *yleistä yhteenlaskusääntöä* ja *poissulkevien tapahtumien yhteenlaskusääntöä*.

Esimerkki 1.6. – Ratkaisu:

Yleisen yhteenlaskusäännön mukaan aina pätee

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

(a) Tiedämme, että $\Pr(A) = 0.2$ ja $\Pr(B) = 0.6$. Näin ollen

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.2 + 0.6 - 0.1 = 0.7.$$

(b) Tiedämme, että $\Pr(A) = 0.2$ ja $\Pr(B) = 0.6$. Koska mahdottoman tapahtuman todennäköisyys on nolla, havaitaan että

$$\Pr(A \cap B) = \Pr(\varnothing) = 0.$$

Näin siis

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.2 + 0.6 - 0 = 0.8.$$

Esimerkki 1.7.

Olkoot $\Pr(A) = 0.2$ ja $\Pr(B) = 0.6$ kuten esimerkissä 1.6. Määrää tapahtuman $A \cup B$ todennäköisyys, kun

(a) A ja B ovat riippumattomia.

(b) $\Pr(A \mid B) = 0.1$.

Esimerkki 1.7. – Mitä opimme?

Esimerkissä 1.7. tarkastellaan *yleistä yhteenlaskusääntöä*, *ehdollisen todennäköisyyden* ja *riippumattomuuden* käsitteitä sekä *yleistä tulosääntöä* ja *riippumattomien tapahtumien tulosääntöä*.

Esimerkki 1.7. – Ratkaisu:

Yleisen yhteenlaskusäännön mukaan aina pätee

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

(a) Tiedämme, että $\Pr(A) = 0.2$ ja $\Pr(B) = 0.6$. Jos A ja B ovat riippumattomia, niin riippumattomien tapahtumien tulosäännön mukaan

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

Siten

$$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \Pr(A) + \Pr(B) - \Pr(A)\Pr(B) \\ &= 0.2 + 0.6 - 0.2 \times 0.6 = 0.68. \end{aligned}$$

(b) Tiedämme, että $\Pr(A) = 0.2$, $\Pr(B) = 0.6$ ja $\Pr(A|B) = 0.1$. Yleisen tulosäännön mukaan aina pätee

$$\Pr(A \cap B) = \Pr(A|B)\Pr(B).$$

Siten

$$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \Pr(A) + \Pr(B) - \Pr(A|B)\Pr(B) \\ &= 0.2 + 0.6 - 0.1 \times 0.6 = 0.74. \end{aligned}$$

Esimerkki 1.8.

Olkoot $\Pr(A) = 0.5$ ja $\Pr(B) = 0.6$. Yritä määrätä tapahtuman $A \cup B$ todennäköisyys, kun

(a) $\Pr(A \cap B) = 0.1$.

(b) A ja B ovat toisensa poissulkevia.

(c) A ja B ovat riippumattomia.

(d) $\Pr(A|B) = 0.1$.

Milloin tämä on mahdollista?

Esimerkki 1.8. – Mitä opimme?

Esimerkissä 1.8. tarkastellaan *yleistä yhteenlaskusääntöä*, *toisensa poissulkevuuden* ja *riippumattomuuden* käsitteitä sekä *ehdollisen todennäköisyyden* määritelmää; ks. myös esimerkkejä 1.6. ja 1.7.

Esimerkki 1.8. – Ratkaisu:

Yleisen yhteenlaskusäännön mukaan aina pätee

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

(a) Tiedämme, että $\Pr(A) = 0.5$ ja $\Pr(B) = 0.6$. Jos $\Pr(A \cap B) = 0.1$, niin

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.5 + 0.6 - 0.1 = 1.$$

(b) Tiedämme, että $\Pr(A) = 0.5$ ja $\Pr(B) = 0.6$. Jos A ja B ovat toisensa poissulkevia, niin $A \cap B = \varnothing$. Tällöin

$$\Pr(A \cap B) = \Pr(\varnothing) = 0,$$

koska mahdottoman tapahtuman todennäköisyys on nolla. Siten

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) = 0.5 + 0.6 - 0 = 1.1 > 1,$$

mikä on mahdotonta. Annetut tiedot ovat siten ristiriitaisia.

(c) Tiedämme, että $\Pr(A) = 0.5$ ja $\Pr(B) = 0.6$. Jos A ja B ovat riippumattomia, niin riippumattomien tapahtuman tulosäännön mukaan

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

Siten

$$\begin{aligned}\Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \Pr(A) + \Pr(B) - \Pr(A)\Pr(B) \\ &= 0.5 + 0.6 - 0.5 \times 0.6 = 0.8.\end{aligned}$$

(d) Tiedämme, että $\Pr(A) = 0.5$ ja $\Pr(B) = 0.6$. Yleisen tulosäännön mukaan

$$\Pr(A \cap B) = \Pr(A \mid B)\Pr(B).$$

Siten

$$\begin{aligned}\Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ &= \Pr(A) + \Pr(B) - \Pr(A \mid B)\Pr(B) \\ &= 0.5 + 0.6 - 0.1 \times 0.6 = 1.04 > 1,\end{aligned}$$

mikä on mahdotonta. Annetut tiedot ovat siten ristiriitaisia.

Esimerkki 1.9.

Uurnassa on 8 keltaista ja 4 sinistä palloa.

(a) Poimitaan uurnasta satunnaisesti kolme palloa *palauttaen*.

Tällöin uurnasta nostetaan palloja yksi pallo kerrallaan ja jokainen nostettu pallo palautetaan ennen seuraavan pallon nostoa takaisin uurnaan. Mikä on todennäköisyys, että havaitset kolme sinistä palloa?

(b) Poimitaan uurnasta satunnaisesti kolme palloa *palauttamatta*.
Tällöin nostettuja palloja ei palauteta takaisin uurnaan. Mikä on todennäköisyys, että havaitset kolme sinistä palloa?

(c) Poimitaan uurnasta satunnaisesti kolme palloa *palauttamatta*. Mikä on todennäköisyys, että viimeisenä poimittu pallo on sininen, jos kaksi edellistä ovat olleet keltaisia?

Ohje: Sovella (a)-kohdassa riippumattomien tapahtumien tulosääntöä ja (b)-kohdassa yleistä tulosääntöä.

Esimerkki 1.9. – Mitä opimme?

Esimerkissä 1.9. havainnollistetaan *ehdollisen todennäköisyyden* ja *riippumattomuuden* käsitteitä sekä *yleisen tulosäännön* ja *riippumattomien tapahtumien tulosäännön* soveltamista *satunnaisotantaan*.

Yksinkertaisessa satunnaisotannassa *palauttaen* voi sama objekti voi tulla poimituksi useita kertoja.

Yksinkertaisessa satunnaisotannassa *palauttamatta* voi sama objekti voi tulla poimituksi vain kerran.

Esimerkki 1.9. – Ratkaisu:

Poimitaan kolme palloa uurnasta, jossa on 8 keltaista ja 4 sinistä palloa.

Merkitään

$A_i = \{i.\ \text{pallo on sininen}\}$,

$A_i^c = \{i.\ \text{pallo on keltainen}\}$.

(a) Kysytty todennäköisyys on

$$\Pr(A_1 \cap A_2 \cap A_3).$$

Koska pallojen poiminta tapahtuu palauttaen, niin tapahtumat A_1, A_2 ja A_3 ovat riippumattomia. Siten *riippumattomien tapahtumien tulosäännön* perusteella

$$\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1)\Pr(A_2)\Pr(A_3) = (4/12)^3 = 1/27 = 0.037037,$$

koska poiminnan jokaisessa vaiheessa uurnassa on 12 palloa, joista 4 on sinistä.

(b) Kysytty todennäköisyys on

$$\Pr(A_1 \cap A_2 \cap A_3).$$

Koska poiminta tapahtuu *palauttamatta*, niin tapahtumat A_1, A_2 ja A_3 ovat tilastollisesti riippuvia. Tässä tapauksessa riippumattomien tapahtumien tulosääntöä ei voi käyttää. *Yleisen tulosäännön* perusteella

$$\Pr(A_1 \cap A_2 \cap A_3)$$
$$= \Pr(A_1)\Pr(A_2 | A_1)\Pr(A_3 | A_1 \cap A_2)$$
$$= (4/12) \times (3/11) \times (2/10) = 1/55 = 0.018182.$$

Laskutoimituksen perustelu:

(i) Ensimmäistä palloa poimittaessa uurnassa on 12 palloa, joista 4 on sinistä.

(ii) Jos ensimmäisenä poimittu pallo oli sininen, toista palloa poimittaessa uurnassa on jäljellä 11 palloa, joista 3 on sinistä.

(iii) Jos ensimmäisenä ja toisena poimitut pallot olivat sinisiä, kolmatta palloa poimittaessa uurnassa on jäljellä 10 palloa, joista 2 on sinistä.

(c) Kysytty todennäköisyys on

$$\Pr(A_3 \mid A_1^c \cap A_2^c)$$

Jos uurnassa on aluksi 8 keltaista ja 4 sinistä palloa, ja 2 keltaisista palloista otetaan pois, niin jäljelle jää 6 keltaista ja 4 sinistä palloa. Siten todennäköisyys saada sininen pallo kolmantena on

$$\Pr(A_3 \mid A_1^c \cap A_2^c) = 4/10 = 0.4.$$

Esimerkki 1.10.

Erään liikeyrityksen johto on harkinnut sulautumista toiseen yritykseen. Johto on selvittänyt osakkaiden mielipiteet sulautumisesta ja luokitellut osakkaat kolmeen luokkaan omistettujen osakkeiden lukumäärän ja mielipiteen mukaan. Luokittelun tuloksena saadut *frekvenssit* on annettu alla olevassa taulukossa. Määrää todennäköisyydet seuraaville tapahtumille:

(a) Satunnaisesti valitulla osakkaalla on alle 200 osaketta.

(b) Satunnaisesti valittu osakas on sulautumista vastaan.

(c) Satunnaisesti valitulla osakkaalla on 200 – 1000 osaketta *ja* hän on sulautumisen puolesta.

(d) Satunnaisesti valitulla osakkaalla on 200 – 1000 osaketta *tai* hän on sulautumisen puolesta.

(e) Satunnaisesti valittu osakas on sulautumista vastaan *ehdolla* että hänellä on alle 200 osaketta. Kysymys: Ovatko se, että osakkaalla on alle 200 osaketta ja se, että osakas on sulautumista vastaan tapahtumina riippumattomia?

(f) Satunnaisesti valitulla osakkaalla on 200 osaketta tai enemmän *ehdolla* että hän on sulautumisen puolesta.

Osakkaat		Mielipide sulautumisesta		
		Puolesta	Ei mielipidettä	Vastaan
Osakkeiden lukumäärä	< 200	38	9	29
	200 – 1000	30	7	42
	1000 <	32	4	59

Esimerkki 1.10. – Mitä opimme?

Esimerkissä 1.10. havainnollistetaan *ehdollisen todennäköisyyden* ja *riippumattomuuden* käsitteitä.

Esimerkki 1.10. – Ratkaisu:

Määrätään ensin taulukon solujen rivi- ja sarakesummat sekä kokonaissumma:

Osakkaiden lukumäärä		Mielipide sulautumisesta			Yhteensä
		Puolesta	Ei mielipidettä	Vastaan	
Osakkeiden lukumäärä	< 200	38	9	29	76
	200 – 1000	30	7	42	79
	1000 <	32	4	59	95
Yhteensä		100	20	130	250

Vastaavat todennäköisyydet saadaan jakamalla kaikki taulukon luvut osakkaiden kokonaislukumäärällä 250. Tuloksena saadaan seuraava taulukko:

Osakkaiden lukumäärä		Mielipide sulautumisesta			Yhteensä
		Puolesta	Ei mielipidettä	Vastaan	
Osakkeiden lukumäärä	< 200	0.152	0.036	0.116	0.304
	200 – 1000	0.120	0.028	0.168	0.316
	1000 <	0.128	0.016	0.236	0.380
Yhteensä		0.400	0.080	0.520	1

Esimerkiksi:

Pr(Osakkaalla 200 – 1000 osaketta) = 0.316

Pr(Osakas on sulautumisen puolesta *ja* osakkaalla on yli 1000 osaketta) = 0.128

Edellä määrätystä todennäköisyyksien taulukosta saadaan:

(a) Pr(Osakkaalla on alle 200 osaketta) = 0.304

(b) Pr(Osakas on sulautumista vastaan) = 0.52

(c) Pr(Osakkaalla on 200 – 1000 osaketta *ja* osakas on sulautumisen puolesta)
= 0.120

(d) *Yleisen yhteenlaskusäännön* nojalla:

Pr(Osakkaalla on 200 – 1000 osaketta *tai* osakas on sulautumisen puolesta)

= Pr(Osakkaalla on 200 – 1000 osaketta) + Pr(Osakas on sulautumisen puolesta)
– Pr(Osakkaalla on 200 – 1000 osaketta *ja* osakas on sulautumisen puolesta)

= 0.316 + 0.400 – 0.120

= 0.596

(e) Käyttämällä *ehdollisen todennäköisyyden* määritelmää saadaan:

Pr(Osakas on sulautumista vastaan *ehdolla*, *että* osakkaalla on alle 200 osaketta)

= Pr(Osakas on sulautumista vastaan *ja* osakkaalla on alle 200 osaketta)
/Pr(Osakkaalla on alle 200 osaketta)

= 0.116/0.304

= 0.382.

(b)-kohdan mukaan

Pr(Osakas on sulautumista vastaan) = 0.52.

Koska

Pr(Osakas on sulautumista vastaan *ehdolla*,
että osakkaalla on alle 200 osaketta) = 0.382

≠ Pr(Osakas on sulautumista vastaan) = 0.52,

tapahtumat ”Osakas on sulautumista vastaan” ja ”Osakkaalla on alle 200 osaketta” *eivät ole riippumattomia*. Siten tieto siitä, että ehtotapahtuma

” Osakkaalla on alle 200 osaketta”

on sattunut sisältää *informaatiota*, jota voidaan käyttää hyväksi, kun määräämme tapahtuman

”Osakas on sulautumista vastaan”

todennäköisyyttä.

(f) Käyttämällä *ehdollisen todennäköisyyden määritelmää* saadaan:

Pr(Osakkaalla on 200 osaketta tai enemmän *ehdolla*,
että osakas on sulautumisen puolesta)

= Pr(Osakkaalla on 200 osaketta tai enemmän *ja* osakas on sulautumisen puolesta)
/Pr(Osakas on sulautumisen puolella)

= (0.120 + 0.128)/0.400

= 0.620.

Huomaa, että

Pr(Osakkaalla on 200 osaketta tai enemmän *ehdolla*,
että osakas on sulautumisen puolesta) = 0.620

≠ Pr(Osakkaalla on 200 osaketta tai enemmän) = 0.696.

Siten tieto siitä, että ehtotapahtuma

”Osakas on sulautumisen puolesta”

on sattunut sisältää *informaatiota*, jota voidaan käyttää hyväksi, kun määräämme tapahtuman

”Osakkaalla on 200 osaketta tai enemmän”

todennäköisyyttä.

Esimerkki 1.11.

Potilaan ikä saattaa vaikuttaa siihen millaista hoitoa hän saa. Eräässä USA:ssa tehdyssä tutkimuksessa verrattiin eri-ikäisten naisten pääsemistä mammografiaan (röntgentutkimus rintasyövän toteamiseksi), kun heidän rinnoissaan oli havaittu kyhmyjä. Tulokset on annettu taulukossa alla. Taulukon solut ovat todennäköisyyksiä, että kumpikin tapahtumista sattuu; esim. 0.321 on todennäköisyys, että potilas on alle 65-vuotias ja hänelle on tehty mammografia.

Todennäköisyys		Mammografia tehty	Mammografiaa ei ole tehty
Ikä	alle 65	0.321	0.124
	65 tai yli	0.365	0.190

(a) Määrää seuraavien tapahtumien todennäköisyydet:

A = {potilas on alle 65-vuotias}

B = {potilas on 65-vuotias tai yli}

C = {potilaalle on tehty mammografia}

D = {potilaalle ei ole tehty mammografiaa}

(b) Ovatko tapahtumat B ja C riippumattomia?

(c) Määrää todennäköisyys sille, että potilaalle on tehty mammografia, jos hän on ollut alle 65-vuotias ja vertaa sitä todennäköisyyteen, että potilaalle on tehty mammografia, jos hän on ollut 65-vuotias tai yli.

Esimerkki 1.11. – Mitä opimme?

Esimerkissä 1.11. havainnollistetaan *ehdollisen todennäköisyyden* käsitettä.

Esimerkki 1.11. – Ratkaisu:

(a) Kysytyt todennäköisyydet saadaan laskemalla tehtävän taulukosta ns. *reuna-todennäköisyydet* eli rivi- ja sarakesummat:

Todennäköisyys		Mammografia tehty	Mammografiaa ei ole tehty	Summa
Ikä	alle 65	0.321	0.124	$\Pr(A) = 0.445$
	65 tai yli	0.365	0.190	$\Pr(B) = 0.555$
Summa		$\Pr(C) = 0.686$	$\Pr(D) = 0.314$	1

(b) Jos tapahtumat B ja C ovat riippumattomia, niin *riippumattomien tapahtumien tulosäännöstä* seuraa, että

$$\Pr(B\cap C) = \Pr(B)\Pr(C).$$

Taulukosta saadaan

$$\Pr(B\cap C) = 0.365,$$

$$\Pr(B)\Pr(C) = 0.381\ .$$

Koska

$$\Pr(B\cap C) \neq \Pr(B)\Pr(C),$$

tapahtumat B ja C eivät ole riippumattomia.

(c) Ehdollisen todennäköisyyden määritelmän mukaan

$$\Pr(C|A) = \Pr(C\cap A)/\Pr(A) = 0.321/0.445 = 0.721,$$

$$\Pr(C|B) = \Pr(C\cap B)/\Pr(B) = 0.365/0.555 = 0.657.$$

Tämän perusteella nuoremmilla potilailla todennäköisyys päästä mammografiaan on jonkin verran suurempi kuin vanhemmilla potilailla.

Esimerkki 2.1.

Tarkastellaan kirjainten a, e, i, k, l, m, p (7 kpl) muodostamaa joukkoa $S = \{a, e, i, k, l, m, p\}$.

(a) Kuinka monella eri tavalla joukon S kirjaimet voidaan järjestää jonoon?

(b) Kuinka monta erilaista 3:n alkion osajonoa joukon S kirjaimista voidaan muodostaa?

(c) Kuinka monta erilaista 3:n alkion osajoukkoa joukon S kirjaimista voidaan muodostaa?

Esimerkki 2.1. – Mitä opimme?

Esimerkissä 2.1. tarkastellaan esimerkkejä *kombinatoriikan perusongelmista* sekä niiden *ratkaisemista kombinatoriikan perusperiaatteiden avulla*.

Esimerkki 2.1. – Ratkaisu:

(a) Joukossa $S = \{a, e, i, k, l, m, p\}$ on

$$n(S) = 7$$

alkiota. Siten joukon S alkiot voidaan järjestää

$$7! = 7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1 = 5040$$

eri tavalla.

Tulos voidaan perustella käyttämällä ns. *lokeromallia*: Koska joukossa S on 7 alkiota, muodostetaan lokerikko, jossa on 7 lokeroa. Ideana on täyttää lokerikko joukon S alkioilla vaiheittain

Alla olevan taulukon varjostetut solut kuvaavat ko. lokerikkoa. Jokaiseen lokeroon on merkitty luvulla n kuinka monella tavalla lokeron täyttö voidaan tehdä.

Lokeron nro	1	2	3	4	5	6	7
n	7	6	5	4	3	2	1

1. lokero: Lokero voidaan täyttää 7:llä eri tavalla kirjaimilla a, e, i, k, l, m, p

2. lokero: Lokero voidaan täyttää 6:lla eri tavalla jäljelle jääneillä kirjaimilla, koska 1 kirjaimista on käytetty.

3. lokero: Lokero voidaan täyttää 5:llä eri tavalla jäljelle jääneillä kirjaimilla, koska 2 kirjaimista on käytetty.

4. lokero: Lokero voidaan täyttää 4:llä eri tavalla jäljelle jääneillä kirjaimilla, koska 3 kirjaimista on käytetty.

5. lokero: Lokero voidaan täyttää 3:lla eri tavalla jäljelle jääneillä kirjaimilla, koska 4 kirjaimista on käytetty.

6. lokero: Lokero voidaan täyttää 2:lla eri tavalla jäljelle jääneellä kirjaimella, koska 5 kirjaimista on käytetty

7. lokero: Lokero voidaan täyttää 1:llä eri tavalla jäljelle jääneellä kirjaimella, koska 6 kirjaimista on käytetty.

Koska jokainen täyttöoperaatio voidaan tehdä riippumatta aikaisemmin suoritetuista täyttöoperaatioista, niin kombinatoriikan *kertolaskuperiaatteen* mukaan koko lokerikko voidaan täyttää

$$7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1 = 7! = 5040$$

eri tavalla.

(b) Joukon $S = \{a, e, i, k, l, m, p\}$ alkioista voidaan muodostaa

$$\mathrm{P}(7,3) = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{4\cdot3\cdot2\cdot1} = 7\cdot6\cdot5 = 210$$

erilaista 3:n alkion järjestettyä listaa, jossa mikään alkio ei toistu.

Tulos voidaan perustella käyttämällä ns. *lokeromallia*: Muodostetaan lokerikko, jossa on 3 lokeroa. Ideana on täyttää lokerikko joukon S alkioilla vaiheittain.

Alla olevan taulukon varjostetut solut kuvaavat ko. lokerikkoa. Jokaiseen lokeroon on merkitty luvulla n kuinka monella tavalla lokeron täyttö voidaan tehdä.

Lokeron nro	1	2	3
n	7	6	5

1. lokero: Lokero voidaan täyttää 7:llä eri tavalla kirjaimilla a, e, i, k, l, m, p
2. lokero: Lokero voidaan täyttää 6:lla eri tavalla jäljelle jääneillä kirjaimilla, koska 1 kirjaimista on käytetty.
3. lokero: Lokero voidaan täyttää 5:llä eri tavalla jäljelle jääneillä kirjaimilla, koska 2 kirjaimista on käytetty.

Koska jokainen täyttöoperaatio voidaan tehdä riippumatta aikaisemmin suoritetuista täyttöoperaatioista, niin kombinatoriikan *kertolaskuperiaatteen* mukaan koko lokerikko voidaan täyttää

$$7\cdot 6\cdot 5 = \frac{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{4\cdot 3\cdot 2\cdot 1} = \frac{7!}{4!} = \frac{7!}{(7-3)!} = 210$$

eri tavalla. Tulos seuraa myös (a)-kohdassa esitetystä tarkastelusta pysäyttämällä lokeroiden täyttö sen jälkeen, kun 3. lokero on saatu täytetyksi.

(c) Joukon $S = \{a, e, i, k, l, m, p\}$ alkioista voidaan muodostaa

$$\binom{7}{3} = \frac{7!}{3!(7-3)!} = \frac{7!}{3!4!} = \frac{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1\cdot 4\cdot 3\cdot 2\cdot 1} = 35$$

erilaista 3:n alkion *osajoukkoa* eli *kombinaatiota*.

Tulos voidaan perustella seuraavalla tavalla: Olkoon joukon S alkioiden 3:n alkion *osajoukkojen* lukumäärä x, jossa x on vielä toistaiseksi tuntematon luku. (b)-kohdan mukaan joukon S alkioiden 3:n alkion järjestettyjen listojen lukumäärä on

$$\mathrm{P}(7,3) = \frac{7!}{(7-3)!}.$$

Joukon S alkioiden 3:n alkion järjestetyt listat, joissa mikään alkio ei toistu, voidaan muodostaa kahdessa vaiheessa:

(i) Valitaan joukon S alkioiden joukosta 3:n alkion järjestämätön *osajoukko*. Tämä voidaan tehdä x eri tavalla, jossa x on vielä toistaiseksi tuntematon luku.

(ii) Järjestetään kohdassa (i) valitut 3 alkiota *jonoksi*. Tämä voidaan tehdä (a)-kohdan mukaan 3! eri tavalla.

Koska operaatiot (i) ja (ii) voidaan suorittaa toisistaan riippumatta, niin joukon S alkioiden 3-permutaatiot voidaan muodostaa kombinatoriikan *kertolasku-periaatteen* mukaan

$$x \cdot 3!$$

eri tavalla.

Olemme määränneet joukon S alkioiden 3-permutaatioiden lukumäärän kahdella eri tavalla ja saamme siten x:n ratkaisemiseksi yhtälön

$$\mathrm{P}(7,3) = \frac{7!}{(7-3)!} = x \cdot 3!.$$

Siten

$$x = \frac{\mathrm{P}(7,3)}{3!} = \frac{7!}{3!(7-3)!} = \binom{7}{3}.$$

Esimerkki 2.2.

Olkoot tietokoneen salasanat muotoa XXXXX, jossa X on jokin vokaaleista a, e, i, o, u (5 kpl). Laske erilaisten sanojen lukumäärä, kun sanojen muodostamista rajoittavat seuraavat ehdot:

(a) Kaikkien kirjainten on oltava salasanassa erilaisia.

(b) Salasanassa on oltava yksi ”pari” eli täsmälleen kaksi samaa vokaalia (esim. eaioe).

(c) Salasanassa on oltava ”kolmoset” eli täsmälleen kolme samaa vokaalia (esim. aaoea).

(d) Salasanassa on oltava ”täyskäsi” eli ”kolmoset” ja ”pari” (esim. ioioo).

Esimerkki 2.2. – Mitä opimme?

Esimerkissä 2.2. tarkastellaan *esimerkkejä kombinatoriikan perusongelmista* sekä niiden *ratkaisemista kombinatoriikan perusperiaatteiden avulla*. Ks. myös esimerkkiä 2.1.

Esimerkki 2.2. – Ratkaisu:

Kaikki muodostettavat salasanat ovat muotoa XXXXX, jossa

X = a, e, i, o, u (5 kpl)

Kohdat (a) – (d) eroavat toisistaan siten, että salasanojen muodostamista rajoittavat niissä erilaiset ehdot. Sovellamme tunnusten muodostamisessa ns. *lokeromallia*: Salasanaa asetetaan vastaamaan *lokerikko*, jossa on 5 lokeroa.

(a) Kaikkien kirjainten on oltava salasanassa erilaisia.

Täytetään lokerot vaiheittain: Koska samaa vokaalia ei saa käyttää kuin kerran, k. lokero voidaan täyttää

$$5 - k + 1 \ , \ k = 1, 2, 3, 4, 5$$

eri tavalla. Koska jokainen lokerontäyttöoperaatio voidaan suorittaa riippumatta aikaisemmista operaatioista, kombinatoriikan *kertolaskuperiaatteen* nojalla salasanojen kokonaislukumäärä on

$$5\times4\times3\times2\times1 = 120$$

(b) Salasanassa on oltava yksi ”pari” eli täsmälleen kaksi samaa vokaalia.

5 lokeroa voidaan täyttää kahden saman vokaalin muodostamalla parilla *binomikertoimen*

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{2\cdot1\cdot3\cdot2\cdot1} = 10$$

ilmaisemalla lukumäärällä eri tapoja ja vokaali pariin voidaan valita 5:llä eri tavalla.

Koska kolmeen jäljellä olevaan lokeroon jokaiseen on valittava eri vokaali, voidaan loput vokaaleista valita $4\times3\times2$ eri tavalla. Siten vokaalit voidaan valita lokeroihin $5\times4\times3\times2$ eri tavalla.

Kombinatoriikan *kertolaskuperiaatteen* nojalla salasanojen kokonaislukumäärä on

$$\binom{5}{2}\times5\times4\times3\times2 = 10\times5\times4\times3\times2 = 1200$$

(c) Salasanassa on oltava ”kolmoset” eli täsmälleen kolme samaa vokaalia.

5 lokeroa voidaan täyttää kolmen saman vokaalin muodostamilla kolmosilla *binomikertoimen*

$$\binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5!}{3!2!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot2\cdot1} = 10$$

ilmaisemalla lukumäärällä eri tapoja ja vokaali voidaan valita kolmosiin 5:llä eri tavalla.

Koska kahteen jäljellä olevaan lokeroon jokaiseen on valittava eri vokaali, voidaan muut vokaalit valita 4×3 eri tavalla. Siten vokaalit voidaan valita lokeroihin $5\times4\times3$ eri tavalla.

Kombinatoriikan *kertolaskuperiaatteen* nojalla salasanojen kokonaislukumäärä on

$$\binom{5}{3}\times 5\times 4\times 3 = 10\times 5\times 4\times 3 = 600$$

(d) Salasanassa on oltava ”täyskäsi” eli ”kolmoset” ja ”pari”.

5 lokeroa voidaan täyttää kahden saman vokaalin muodostamalla parilla *binomi-kertoimen*

$$\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{5!}{2!3!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{2\cdot 1\cdot 3\cdot 2\cdot 1} = 10$$

ilmaisemalla lukumäärällä eri tapoja. Sen jälkeen paikat kolmosille on määrätty. Pariin voidaan valita vokaali 5:llä eri tavalla ja sen jälkeen kolmosiin 4:llä eri tavalla. Siten vokaalit voidaan valita lokeroihin 5×4 eri tavalla.

Kombinatoriikan *kertolaskuperiaatteen* nojalla salasanojen lukumäärä on

$$\binom{5}{2}\times 5\times 4 = 10\times 5\times 4 = 200$$

Esimerkki 2.3.

Tehtävä 2.3. on jatkoa tehtävälle 2.2. Oletetaan, että salasanat arvotaan kaikkien mahdollisten salasanojen joukosta. Laske seuraavien tapahtumien todennäköisyydet:

(a) Kaikki kirjaimet ovat salasanassa erilaisia.

(b) Salasanassa on yksi ”pari” eli täsmälleen kaksi samaa vokaalia (esim. eaioe).

(c) Salasanassa on ”kolmoset” eli täsmälleen kolme samaa vokaalia (esim. aaoea).

(d) Salasanassa on ”täyskäsi” eli ”kolmoset” ja ”pari” (esim. ioioo).

Esimerkki 2.3. – Mitä opimme?

Esimerkki 2.3. on jatkoa esimerkille 2.2. ja siinä tarkastellaan *kombinatoriikan soveltamista tapahtumien* (*klassisen*) *todennäköisyyden määräämiseen*. Ks. myös esimerkkejä 2.1., 1.3. ja 1.4.

Esimerkki 2.3. – Ratkaisu:

Kaikki muodostettavat salasanat ovat muotoa XXXXX, jossa

X = a, e, i, o, u (5 kpl)

Käytetään *klassisen todennäköisyyden määritelmää*: Tapahtuman A klassinen todennäköisyys on

$$\Pr(A) = k / n = p$$

jossa

k = tapahtumalle A *suotuisien* tulosvaihtoehtojen lukumäärä

n = kaikkien mahdollisten tulosvaihtoehtojen lukumäärä ja *kaikki* tulosvaihtoehdot on oletettu yhtä todennäköisiksi.

Kaikkien mahdollisten salasanojen lukumäärä n on kombinatoriikan *kertolaskuperiaatteen* nojalla

$$n = 5\times5\times5\times5\times5 = 5^5 = 3125$$

(a) Kaikki kirjaimet ovat salasanassa erilaisia.

Tehtävän 2.3. mukaan

$$k = 120$$

Siten

$$p = k/n = 120/3125 = 0.0384$$

(b) Salasanassa on yksi ”pari” eli täsmälleen kaksi samaa vokaalia (esim. eaioe).

Tehtävän 2.3. mukaan

$$k = 1200$$

Siten

$$p = k/n = 1200/3125 = 0.384$$

(c) Salasanassa on ”kolmoset” eli täsmälleen kolme samaa vokaalia (esim. aaoea).

Tehtävän 2.3. mukaan

$$k = 600$$

Siten

$$p = k/n = 600/3125 = 0.192$$

(d) Salasanassa on ”täyskäsi” eli ”kolmoset” ja ”pari” (esim. ioioo).

Tehtävän 2.3. mukaan

$$k = 200$$

Siten

$$p = k/n = 200/3125 = 0.064$$

Esimerkki 2.5.

Erässä tv-vastaanottimia on 25 vastaanotinta, joista 5 on viallista.

(a) Kuinka monella eri tavalla vastaanotinten joukosta voidaan poimia 5 vastaanotinta niin, että mukaan tulee täsmälleen 1 viallinen vastaanotin, jos poiminta tehdään *palauttamatta*?

(b) Mikä on todennäköisyys, että poimittaessa vastaanotinten joukosta umpimähkään 5 vastaanotinta mukaan tulee täsmälleen 1 viallinen vastaanotin, jos poiminta tehdään *palauttamatta*?

Esimerkki 2.5. – Mitä opimme?

Esimerkissä 2.5. tarkastellaan *kombinatoriikan soveltamista tapahtumien toden-näköisyyden määräämiseen otannassa palauttamatta*. Ks. myös esimerkkejä 2.1. ja 1.9.

Esimerkki 2.5. – Ratkaisu:

(a) Tehtävänä on valita 4 vastaanotinta 20:n ehjän vastaanottimen joukosta ja 1 vastaanotin 5:n viallisen vastaanottimen joukosta ja laskea niiden tapojen lukumäärä, jolla tämä voidaan tehdä.

4 vastaanotinta voidaan valita 20:n ehjän joukosta *binomikertoimen*

$$\binom{20}{4}$$

ilmaisemalla lukumäärällä eri tapoja.

1 vastaanotin voidaan valita 5:n viallisen joukosta *binomikertoimen*

$$\binom{5}{1}$$

ilmaisemalla lukumäärällä eri tapoja.

Nämä valinnat voidaan tehdä toisistaan riippumatta, joten *kombinatoriikan kertolaskuperiaatteen* mukaan valintojen kokonaislukumääräksi saadaan

$$\binom{20}{4}\binom{5}{1} = \frac{20!}{4!(20-4)!}\cdot\frac{5!}{1!(5-1)!} = \frac{20!}{4!16!}\cdot\frac{5!}{1!4!} = \frac{20\cdot19\cdot18\cdot17}{4\cdot3\cdot2}\cdot\frac{5}{1} = 24225$$

(b) Käytetään *klassisen todennäköisyyden määritelmää*: Tapahtuman A klassinen todennäköisyys on

$$\Pr(A) = k/n$$

jossa

k = tapahtumalle A *suotuisien* tulosvaihtoehtojen lukumäärä

n = kaikkien mahdollisten tulosvaihtoehtojen lukumäärä

ja *kaikki* tulosvaihtoehdot on oletettu yhtä todennäköisiksi.

Kaikkien tulosvaihtoehtojen lukumäärä: 5 vastaanotinta voidaan poimia 20 vastaan-ottimen joukosta *binomikertoimen*

$$\binom{25}{5} = \frac{25!}{5!(25-5)!} = \frac{25!}{5!20!} = \frac{25\cdot24\cdot23\cdot22\cdot21}{5\cdot4\cdot3\cdot2} = 53130$$

ilmaisemalla lukumäärällä eri tapoja.

Tapahtumalle *A suotuisien* tulosvaihtoehtojen lukumäärä: (a)-kohdan mukaan 4 vastaanotinta voidaan valita 20:n ehjän vastaanottimen joukosta ja 1 vastaanotin 5:n viallisen vastaanottimen joukosta tulon

$$\binom{20}{4}\binom{5}{1} = 24225$$

ilmaisemalla lukumäärällä eri tapoja.

Siten todennäköisyys valita 5 vastaanotinta satunnaisesti 25:n vastaanottimen joukosta ja saada 4 vastaanotinta 20:n ehjän vastaanottimen joukosta ja 1 vastaanotin 5:n viallisen vastaanottimen joukosta on

$$\frac{\binom{20}{4}\binom{5}{1}}{\binom{25}{5}} = \frac{\frac{20!}{4!16!}\cdot\frac{5!}{1!4!}}{\frac{25!}{5!20!}} = \frac{20\cdot19\cdot18\cdot17}{24\cdot23\cdot22\cdot21} = 0.455957$$

Esimerkki 2.6.

Eräässä tehtaassa on 3 valmistuslinjaa, joilla tehdään samanlaisia CD-soittimia. Linja A valmistaa soittimista 30 %, linja B 25 % ja linja C 45 %. A:n valmistamista soittimista keskimäärin 2 %, B:n valmistamista soittimista 3 % ja C:n valmistamista soittimista 4 % on osoittautunut viallisiksi.

Valitaan satunnaisesti yksi soitin tarkistusta varten.

(a) Mikä on todennäköisyys, että soitin on viallinen?

(b) Mikä on todennäköisyys, että soitin on tehty linjalla A, jos se on viallinen?

Esimerkki 2.6. – Mitä opimme?

Esimerkissä 2.6. tarkastellaan *kokonaistodennäköisyyden* ja *Bayesin kaavojen* soveltamista.

Tehtävä 2.6. – Ratkaisu:

Määritellään seuraavat tapahtumat:

$$A = \{\text{Soittimen on valmistanut linja A}\}$$

$$B = \{\text{Soittimen on valmistanut linja B}\}$$

$$C = \{\text{Soittimen on valmistanut linja C}\}$$

Tehtävän asettelun mukaan seuraavat *todennäköisyydet* tunnetaan:

$$\Pr(A) = 0.30$$

$$\Pr(B) = 0.25$$

$$\Pr(C) = 0.45$$

Määritellään tapahtuma

$$V = \{\text{Soitin on viallinen}\}$$

Tehtävän asettelun mukaan myös seuraavat *ehdolliset todennäköisyydet* tunnetaan:

$$\Pr(V|A) = 0.02$$
$$\Pr(V|B) = 0.03$$
$$\Pr(V|C) = 0.04$$

(a) Tehtävänä on määrätä $\Pr(V)$.

Kokonaistodennäköisyyden kaavan mukaan:

$$\begin{aligned}\Pr(V) &= \Pr(A)\Pr(V|A) + \Pr(B)\Pr(V|B) + \Pr(C)\Pr(V|C) \\ &= 0.30\times0.02 + 0.25\times0.03 + 0.45\times0.04 = 0.0315\end{aligned}$$

(b) Tehtävänä on määrätä $\Pr(A|V)$.

Bayesin kaavan mukaan:

$$\Pr(A|V) = \frac{\Pr(A\cap V)}{\Pr(V)} = \frac{\Pr(A)\Pr(V|A)}{\Pr(A)\Pr(V|A) + \Pr(B)\Pr(V|B) + \Pr(C)\Pr(V|C)}$$
$$= \frac{0.30\times0.02}{0.0315} \approx 0.190$$

Esimerkki 2.7.

Huumeiden käytön paljastamiseen tarkoitetun pikatestin luotettavuudesta on käytössä seuraavat tiedot: Henkilö, joka käyttää huumeita tulee oikein luokitelluksi huumeiden käyttäjäksi todennäköisyydellä 0.98. Toisaalta henkilö, joka *ei käytä huumeita* tulee virheellisesti luokitelluksi huumeiden käyttäjäksi todennäköisyydellä 0.05.

Oletetaan, että testiä sovelletaan ihmisjoukkoon, jossa 1 % ihmisistä käyttää huumeita.

(a) Mikä on todennäköisyys, että satunnaisesti valittu henkilö tulee pikatestissä luokitelluksi huumeiden käyttäjäksi?

(b) Mikä on todennäköisyys, että pikatestissä huumeiden käyttäjäksi luokiteltu henkilö ei käytä huumeita?

Esimerkki 2.7. – Mitä opimme?

Esimerkissä 2.7. tarkastellaan *kokonaistodennäköisyyden* ja *Bayesin kaavojen* soveltamista.

Esimerkki 2.7. – Ratkaisu:

Merkitään tehtävän tapahtumavaihtoehtoja seuraavalla tavalla:

H = “Testi luokittelee henkilön huumeiden käyttäjäksi”

K = “Henkilö käyttää huumeita”

E = “Henkilö ei käytä huumeita”

Tehtävän asettelun mukaan seuraavat todennäköisyydet tunnetaan:

$$\Pr(H|K) = 0.98$$

$\Pr(H|E) = 0.05$

$\Pr(K) = 0.01$

Komplementtitapahtuman todennäköisyyden kaavan mukaan

$\Pr(E) = 1 - \Pr(K) = 0.99$

Tehtävässä kysytään todennäköisyyksiä $\Pr(H)$ ja $\Pr(E|H)$.

(a) *Kokonaistodennäköisyyden kaavan* mukaan:

$$\Pr(H) = \Pr(E)\Pr(H \mid E) + \Pr(K)\Pr(H \mid K) = 0.99 \times 0.05 + 0.01 \times 0.98 = 0.0593$$

(b) Bayesin kaavan mukaan:

$$\begin{aligned} \Pr(E \mid H) &= \frac{\Pr(E)\Pr(H \mid E)}{\Pr(E)\Pr(H \mid E) + \Pr(K)\Pr(H \mid K)} \\ &= \frac{0.99 \times 0.05}{0.99 \times 0.05 + 0.01 \times 0.98} = 0.834739 \end{aligned}$$

Kommentti:

(a)-kohdan mukaan pikatesti luokittelee ihmisistä, joista 1 % käyttää huumeita, melkein 6 % huumeiden käyttäjiksi. (b)-kohdan mukaan niistä, jotka pikatesti on luokitellut huumeiden käyttäjiksi, yli 83 % ei kuitenkaan käytä huumeita. Kysymys: Kuinka reilua olisi käyttää tällaista huumetestiä työpaikoilla?

Huomautus:

Tehtävässä ei puhuta mistään todellisuudessa sovelletusta huumetestistä.