Esimerkki 1. Heitetään yhtä noppaa kerran, alkeistapaukset (siis pisteluvut) ovat

$$\Omega = \{1,2,3,4,5,6\}.$$

Tapahtuma

$$A = \text{"pisteluku on vähintään 5"} = \{5,6\},$$

ja sen todennäköisyys on

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{2}{6} = \frac{1}{3} \approx 0.333.$$

Esimerkki 2. Jatkoa esimerkkiin 1, liittyy todennäköisyyden ominaisuuksiin, *Tuominen, s. 11*. Nyt

$$A^c = \text{"pisteluku on enintään 4"} = \{1,2,3,4\},$$

joten $P(A^c) = \frac{4}{6} = \frac{2}{3} = 1 - \frac{1}{3} = 1 - P(A)$.

Merkitään

$$B = \text{"pisteluku ei ole 1 eikä 6"} = \{2,3,4,5\},$$

jolloin $P(B) = \frac{4}{6} = \frac{2}{3}$, $A \cup B = \{5,6\} \cup \{2,3,4,5\} = \{2,3,4,5,6\}$ ja $P(A \cup B) = \frac{5}{6}$. Toisaalta saadaan

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{3} + \frac{2}{3} - \frac{1}{6} = \frac{5}{6},$$

sillä $A \cap B = \{5,6\} \cap \{2,3,4,5\} = \{5\}$ ja näin ollen $P(A \cap B) = \frac{1}{6}$. Saadaan myös $A \setminus B = \{5,6\} \setminus \{2,3,4,5\} = \{6\}$, joten

$$P(A \setminus B) = \frac{1}{6} = \frac{1}{3} - \frac{1}{6} = P(A) - P(A \cap B).$$

Merkitään sitten

$$C = \text{"pisteluku ei ole 1"} = \{2,3,4,5,6\},$$

jolloin $A \subset C$, $P(C) = \frac{5}{6}$ ja $P(A) \leqslant P(C)$.

Esimerkki 3. Todennäköisyyden tulkintaa: *frekvenssitulkinta*. Heitetään noppaa n kertaa. Merkitään

$$F_n(A) = \text{"A:n esiintymiskertojen lukumäärä } n \text{ heitossa" ja}$$

$$f_n(A) = \frac{F_n(A)}{n} = \text{"suhteellinen frekvenssi"}.$$

n	$F_n(A)$	$f_n(A)$
1 000	308	0.308
5 000	1701	0.340
20 000	6743	0.337

Ilmeisesti toistettaessa n kertaa (n "iso") on $f_n(A) = P(A)$. Vaihtoehtoisesti: heitetään n noppaa kerran, jolloin tulos on sama.

Esimerkki 4. Klassisen todennäköisyyden laajennus: *geometrinen todennäköisyys*. Perusjoukko Ω on geometrinen kuvio (jana, tasoalue, 3-ulotteinen kappale,...).

Tikkataulu, numerot 1–10, säde 10 cm.

$$A = \text{"tulee kymppi"},\ m(A) = \pi(1\text{ cm})^2 \text{ ja } m(\Omega) = \pi(10\text{ cm})^2.$$

Näin ollen

$$P(A) = \frac{m(A)}{m(\Omega)} = \frac{\pi(1\text{ cm})^2}{\pi(10\text{ cm})^2} = \frac{1}{100} = 0.01.$$

Tässä oletetaan, että tikka osuu (eikä siis mene ohi) satunnaiseen kohtaan taulua! Muutenkaan oletukset eivät aivan vastaa todellisuutta, hyvä – ehkä huonompikin – heittäjä osuu taulun keskiosaan todennäköisemmin kuin sen laidoille.

Esimerkki 5. Ruletti: 18 punaista, 18 mustaa ja 1 vihreä ruutu.

$\Omega =$ "kaikki 37 ruutua",
$n(\Omega) = 37$,
$A =$ "tulee punainen",
$n(A) = 18$,
$P(A) = \frac{18}{37} < \frac{19}{37} = P(A^c)$.

Tähän huomioon perustuu se, että pitkällä aikavälillä ruletissa pelin pitäjä voittaa aina, jos pelataan pelkästään väreillä punainen ja musta, joissa pelin pitäjä maksaa voittona pelaajan panoksen suuruisen summan.

Esimerkki 6. Olkoon $\Omega = \{0,1\}$.

(i) $\mathcal{F}_1 = \{\emptyset, \{0\}, \{0,1\}\}$ ei ole σ-algebra, sillä $\{0\}^c = \{1\} \notin \mathcal{F}_1$. Näin ollen ehto ($\sigma$A$_2$) ei toteudu. Toteutuuko (σA$_3$)?

(ii) $\mathcal{F}_2 = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$ on σ-algebra.

(iii) Onko $\mathcal{F}_3 = \{\emptyset, \{0,1\}\}$ σ-algebra?

Esimerkki 7. *Painotetun nopan heitto*. Perusjoukko on $\Omega = \{1,2,3,4,5,6\}$ ja tapahtumien joukko on $\mathcal{F} = \mathcal{P}(\Omega)$; siis aivan samat kuin tavallisen symmetrisen nopan tapauksessakin. Nyt kuitenkin noppaa on painotettu niin, että

$$P(6) = \frac{1}{2} \quad \text{ja} \quad P(i) = \frac{1}{10}, \text{ kun } i = 1,2,3,4,5.$$

Tällöin $(\Omega, \mathcal{F}, P)$ on todennäköisyysavaruus. Havaitaan, että

$$\sum_{i=1}^{6} P(i) = P(1) + P(2) + \cdots + P(6) = 1.$$

Tapahtuman

$$A = \text{"ei saada kuutosta yhden nopan heitossa"} = \{1,2,3,4,5\}$$

todennäköisyys on

$$P(A) = \sum_{i=1}^{5} P(i) = P(1) + \cdots + P(5) = 5 \cdot \frac{1}{10} = \frac{1}{2} = 0.5.$$

Esimerkki 8. Valitaan luku t umpimähkään suljetulta väliltä $I = [0,1]$. Tällöin

$$P(\{t \in [\tfrac{1}{2}, 1]\}) = \frac{\frac{1}{2}}{1} = \frac{1}{2},$$
$$P(\{t = k\}) = 0, \text{ kun } k \in I.$$

Esimerkki 9. Oletetaan, että herra X valitsee viidestä työpaikasta yhden tai jää ilman työtä. Laske tn, että herra X valitsee ensimmäisen tai toisen työpaikan.

Ratkaisu. Ensimmäinen pyrkimys on päästä ”sanoista eroon”:

$$A_i = \text{”herra X valitsee } i\text{:nnen työpaikan”}, \; i = 1, 2, \ldots, 5,$$
$$A_0 = \text{”herra X jää ilman työtä”}.$$

Lasketaan siis $P(A_1 \cup A_2)$. Todetaan ensin, että eri A_i:t ($i = 0, 1, \ldots, 5$) eivät voi sattua yhtä aikaa; siis $A_i \cap A_j = \emptyset$, $i \neq j$; sanotaan, että A_i:t ovat *erillisiä*. Näin ollen

$$P(A_1 \cup A_2) = P(A_1) + P(A_2).$$

Esimerkki 10. Olkoon A ja B tn-avaruuden $(\Omega, \mathcal{F}, P)$ tapahtumia, joille $P(A) = 0.5$, $P(B) = 0.6$ sekä $P(A \cap B) = 0.3$. Lausu seuraavat tapahtumat sanallisesti ja laske niiden todennäköisyydet: a) $A \cup B$, b) $A \setminus B$, c) B^c ja d) $(A \cup B) \setminus (A \cap B)$.

Ratkaisu. a) Sanallisesti

$$\begin{aligned} A \cup B &= \text{”joko } A \text{ tai } B \text{ tai molemmat tapahtuvat”} \\ &= \text{”tapahtumista } A \text{ ja } B \text{ ainakin toinen sattuu”}. \end{aligned}$$

Tämän todennäköisyys on yhteenlaskukaavan mukaan

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 - 0.6 - 0.3 = 0.8.$$

b) Sanallisesti

$$A \setminus B = \text{”}A \text{ tapahtuu, mutta } B \text{ ei tapahdu”},$$

jonka todennäköisyys on Lauseen 1.3.2 kohdan (vi) mukaan

$$P(A \setminus B) = P(A) - P(A \cap B) = 0.5 - 0.3 = 0.2.$$

c) Sanallisesti

$$B^c = \text{”}B \text{ ei tapahdu”},$$

ja sen todennäköisyys on Lauseen 1.3.2 kohdan (ii) mukaan

$$P(B^c) = 1 - P(B) = 1 - 0.6 = 0.4.$$

d) Sanallisesti

$$(A \cup B) \setminus (A \cap B) = \text{”tapahtumista } A \text{ ja } B \text{ sattuu täsmälleen yksi”},$$

ja sen todennäköisyys on Lauseen 1.3.2 kohdan (vi) ja yhteenlaskukaavan mukaan

$$\begin{aligned} P((A \cup B) \setminus (A \cap B)) &= P(A \cup B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) - P(A \cap B) \\ &= P(A) + P(B) - 2P(A \cap B) \\ &= 0.5 + 0.6 - 2 \cdot 0.3 = 0.5. \end{aligned}$$

Esimerkki 11. Herrat M, T ja K pelaavat. Oletetaan, että herra M on 2 kertaa niin taitava kuin herra K ja herra T on kolme kertaa niin taitava kuin herra K. Laske tn, että herra M tai herra K voittaa.

Ratkaisu. Merkitään

$\omega_{\mathrm{M}} =$ "herra M voittaa", $\omega_{\mathrm{T}} =$ "herra T voittaa" ja $\omega_{\mathrm{K}} =$ "herra K voittaa",

jolloin $\Omega = \{\omega_{\mathrm{M}}, \omega_{\mathrm{T}}, \omega_{\mathrm{K}}\}$. Nyt

$$\begin{aligned} p_{\mathrm{M}} &= P(\omega_{\mathrm{M}}) = P(\text{"herra M voittaa"}), \\ p_{\mathrm{T}} &= P(\omega_{\mathrm{T}}) = P(\text{"herra T voittaa"}), \\ p_{\mathrm{K}} &= P(\omega_{\mathrm{K}}) = P(\text{"herra K voittaa"}). \end{aligned}$$

Tällöin $p_{\mathrm{M}} = 2 \cdot p_{\mathrm{K}}$ ja $p_{\mathrm{T}} = 3 \cdot p_{\mathrm{K}}$, joten saadaan aksiooman (TN_2) perusteella

$$p_{\mathrm{M}} + p_{\mathrm{T}} + p_{\mathrm{K}} = 2 \cdot p_{\mathrm{K}} + 3 \cdot p_{\mathrm{K}} + p_{\mathrm{K}} = 6 \cdot p_{\mathrm{K}} = 1.$$

Näin ollen

$$p_{\mathrm{K}} = \frac{1}{6},$$

ja kysytty tn on siis

$$P(\text{"herra M tai herra K voittaa"}) = p_{\mathrm{M}} + p_{\mathrm{K}} = 2 \cdot \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}.$$

Esimerkki 12. *Symmetrinen tn-avaruus*; yhden nopan heitto symmetrisellä nopalla. Tällöin $\Omega = \{1,2,3,4,5,6\}$, ja jos $A =$ "pisteluku vähintään 4" $= \{4,5,6\}$, niin

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = \frac{1}{2}.$$

Esimerkki 13. Numeroituva (ääretön) tn-avaruus ei voi olla symmetrinen: Merkitään $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}$.

(i) Jos

$$p_i = P(\omega_i) = \varepsilon > 0 \text{ kaikilla } i = 1,2,3,\ldots,$$

niin

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} \varepsilon = \infty \neq 1.$$

(ii) Jos

$$p_i = P(\omega_i) = 0 \text{ kaikilla } i = 1,2,3,\ldots,$$

niin

$$\sum_{i=1}^{\infty} p_i = \sum_{i=1}^{\infty} 0 = 0 \neq 1.$$

Näin ollen aksiooma (TN_2) ei ole voimassa.

Esimerkki 14. Kuinka moneen eri järjestykseen kolme henkilöä M, T ja K voidaan asettaa?

Ratkaisu. Merkitään $E = \{\mathrm{M}, \mathrm{T}, \mathrm{K}\}$. Mahdolliset järjestykset ovat

(M,T,K), (M,K,T), (T,M,K), (T,K,M), (K,M,T) ja (K,T,M).

Näin ollen eri järjestyksiä on

$$6 = 3 \cdot 2 \cdot 1 = 3!,$$

ja sanotaan, että E:llä on 6 *permutaatiota*.

Kuinka moneen eri järjestykseen voidaan 10 henkilöä asettaa?

Esimerkki 15. Valitaan umpimähkään yksi 2-numeroinen luku. Kuinka monella tavalla tämä on mahdollista? Entä mikä on todennäköisyys, että luvussa esiintyy numero 1?

Ratkaisu. Ensimmäisen numeron on oltava jokin luvuista $1, \ldots, 9$, joten se voidaan valita 9 tavalla. Toinen luku voi olla mikä tahansa numero $0, \ldots, 9$, joten se voidaan valita 10 tavalla. Siis yhteensä saamme $9 \cdot 10 = 90$ mahdollisuutta.

Numero 1 esiintyy kaikissa luvuissa $10, 11, \ldots, 19$, joita on 10 kpl. Tämän lisäksi numero 1 esiintyy luvuissa $21, 31, 41, \ldots, 91$, joita on 8 kpl. Näin ollen 2-numeroisia lukuja, joissa numero 1 esiintyy, on yhteensä $10 + 8 = 18$ kpl. Saadaan siis

$$P(\text{"umpimähkään valitussa 2-numeroisessa luvussa esiintyy numero 1"})$$
$$= \frac{18}{90} = \frac{1}{5} = 0.2.$$

Esimerkki 16. Herra K on ostamassa uutta autoa. Hän voi valita kolmesta automerkistä ja neljästä väristä. Lisäksi hän voi joko ottaa tai olla ottamatta autoonsa erikoisvanteet. Kuinka monta eri vaihtoehtoa herra K:lla on valita uusi autonsa?

Ratkaisu. Vaihtoehtoja on yhteensä $3 \cdot 4 \cdot 2 = 24$. On oleellista, että ”eri vaiheissa” vaihtoehtojen lkm:t ovat riippumattomia. Tämä tarkoittaa sitä, että auton väri ei riipu auton merkistä jne.

Esimerkki 17. Jatkoa *Tuomisen* esimerkkiin 1.5.6. Mikä on todennäköisyys, että umpimähkään täytetyssä veikkaussarakkeessa on 12 oikein?

Ratkaisu. Olkoon

$A_i =$ "i. kohde on väärin ja kaikki muut oikein" $i = 1, 2, \ldots, 13$.

$A =$ "12 oikein".

Tällöin

$$A = \bigcup_{i=1}^{13} A_i \quad \text{ja} \quad A_i \cap A_j = \emptyset \quad \text{kun } i \neq j.$$

Koska $n(A_i) = 2$ ("väärä" kohde voidaan valita täytettäväksi kahdella tavalla, kaikki muut vain yhdellä; tuloperiaate), $i = 1, 2, \ldots, 13$, niin

$$n(A) = \sum_{i=1}^{13} n(A_i) = 13 \cdot 2 = 26$$

ja

$$P(A) = \frac{n(A)}{n(\Omega)} = 26 \cdot 3^{-13} \approx 1.63 \cdot 10^{-5}.$$

Esimerkki 18. Herra K on jälleen ostamassa uutta autoa. Hän voi valita kolmesta automerkistä, joista ensimmäisessä on 3, toisessa 2 ja kolmannessa 10 värivaihtoehtoa. Lisäksi ensimmäiseen merkkiin hän voi joko ottaa tai olla ottamatta erikoisvanteet. Kuinka monta eri vaihtoehtoa herra K:lla on valita uusi autonsa?

Ratkaisu. Ensimmäinen automerkki voidaan valita $3 \cdot 2 = 6$ tavalla (tuloperiaate), toinen 2 tavalla ja kolmas 10 tavalla. Näin ollen erilaisia vaihtoehtoja on $6 + 2 + 10 = 18$ *(summaperiaate).*

Esimerkki 19. Lotto. $E = \{1,2,3,\dots,39\}$. Valitaan 7 numeroa.

(i) 7-variaatioiden lukumäärä on

$$(39)_7 = 39 \cdot 38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 \cdot 33 \approx 7.75 \cdot 10^{10}.$$

Tällöin

$$(1,2,10,20,25,30,35) \neq (2,1,10,20,25,30,35).$$

Saadaan siis ”liikaa vaihtoehtoja”.

(ii) 7-kombinaatioiden lukumäärä on

$$\binom{39}{7} = \frac{39!}{7! \cdot 32!} = 15\,380\,937.$$

Tällöin

$$\{1,2,10,20,25,30,35\} = \{2,1,10,20,25,30,35\},$$

ja havaitaan, että kombinaatiot ovat ”oikea” malli kuvaamaan mahdollisia lottorivejä. Todetaan lisäksi, että

$$P(\text{”7 oikein”}) = \frac{1}{15\,380\,937} \approx 6.5 \cdot 10^{-8}.$$

Esimerkki 20. Korttipakka (52 korttia), nostetaan 3 korttia. Mikä on tn, että saadaan pari?

Ratkaisu. Merkitään

$$A_i = \text{"saadaan } i\text{-pari"}, \; i = 1, \dots 13 \quad \text{ja} \quad A = \text{"saadaan pari"}.$$

Tällöin

$$A = \bigcup_{i=1}^{13} A_i.$$

Lasketaan kysytty todennäköisyys ensin käyttämällä *variaatioita*. Merkitään siis $E_v =$ "3-variaatiot", jolloin

$$n(E_v) = (52)_3 = 52 \cdot 51 \cdot 50 = 132\,600.$$

Määrätään sitten $n(A_i)$, $i = 1, \dots, 13$; ajatellaan, että kolme korttia "sijoitetaan" 3-jonoon $(\sqcup \sqcup \sqcup)$.

(i) Ensimmäinen kortti, jonka arvo on i, voidaan sijoittaa kolmella tavalla jonoon.
(ii) Toinen kortti, jonka arvo on i, voidaan sijoittaa kahdella tavalla jonoon.
(iii) Kaksi korttia, joiden arvo on i, voidaan valita neljästä mahdollisesta kortista $\binom{4}{2} = 6$ tavalla.
(iv) Kolmanteen, vielä vapaana olevaan 3-jonon paikkaan, voidaan valita kortti 48 tavalla (kaikki kortit, joiden arvo ei ole i).

Näin ollen $n(A_i) = 3 \cdot 2 \cdot 6 \cdot 48 = 1728$ ja kysytty tn on

$$P(A) = \sum_{i=1}^{13} P(A_i) = 13 \cdot \frac{1728}{132\,600} \approx 0.169.$$

Ratkaistaan sitten sama tehtävä käyttämällä *kombinaatioita*. Merkitään siis nyt $E_k =$ "3-kombinaatiot", jolloin

$$n(E_k) = \binom{52}{3} = \frac{(52)_3}{3!} = 22\,100.$$

Määrättäessä nyt $n(A_i)$:ta ajatellaan, että kolme korttia valitaan pakasta "ilman järjestystä". Tällöin

(i) i-pari voidaan valita neljästä mahdollisesta kortista $\binom{4}{2} = 6$ tavalla.
(ii) Kolmas kortti voidaan valita 48 tavalla (kaikki kortit, joiden arvo ei ole i).

Näin ollen $n(A_i) = 6 \cdot 48 = 288$ ja kysytty tn on

$$P(A) = \sum_{i=1}^{13} P(A_i) = 13 \cdot \frac{288}{22\,100} \approx 0.169.$$

Kumpi tapa (variaatiot vai kombinaatiot) on mielestäsi tilanteeseen paremmin sopiva?

Esimerkki 21. 5 palloa, joista 3 valkoista ja 2 mustaa. Laske tn, että nostettaessa 2 palloa *ilman takaisinpanoa* saadaan 1 valkoinen ja 1 musta.

Ratkaisu. Nyt $N = 5$, $K = 3$, $N - K = 2$ ja $n = 2$. Merkitään

$$A = \text{"saadaan 1 valkoinen ja 1 musta"},$$

jolloin

$$P(A) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = \frac{\frac{3!}{1!2!} \cdot \frac{2!}{1!1!}}{\frac{5!}{2!3!}} = \frac{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1}{1 \cdot 2 \cdot 1 \cdot 1 \cdot 1 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{3}{5} = 0.6.$$

Esimerkki 22. Jatkoa edelliseen esimerkkiin. Laske tn, että nostettaessa 2 palloa *takaisinpanolla* (pallot siis palautetaan noston jälkeen) saadaan 1 valkoinen ja 1 musta.

Ratkaisu. Nyt

$$P(A) = \binom{2}{1}\frac{3^1 \cdot 2^1}{5^2} = \frac{2!}{1!1!} \cdot \frac{3 \cdot 2}{25} = \frac{2 \cdot 1 \cdot 3 \cdot 2}{25} = \frac{12}{25} = 0.48.$$

Esimerkki 23. Korttipakka, $N = 52$. Nostetaan kolme korttia, mikä on tn, että saadaan täsmälleen yksi ässä? a) otp, b) oitp.

Ratkaisu. Merkitään

$$A_i = \text{"}i\text{ ässää"} \quad \text{ja} \quad K = \text{"ässien lukumäärä pakassa"} = 4.$$

a)

$$P(A_1) = \binom{3}{1}\frac{4^1 \cdot 48^2}{52^3} = \frac{3!}{1!2!} \cdot \frac{4 \cdot 48^2}{52^3} = \frac{3 \cdot 4 \cdot 48^2}{52^3} \approx 0.197.$$

b)

$$\begin{aligned} P(A_1) &= \frac{\binom{4}{1}\binom{48}{2}}{\binom{52}{3}} = \frac{\frac{4!}{1!3!} \cdot \frac{48!}{2!46!}}{\frac{52!}{3!49!}} = \frac{4!48!49!3!}{3!2!46!52!} \\ &= \frac{4 \cdot 3 \cdot 48 \cdot 47}{52 \cdot 51 \cdot 50} \approx 0.204. \end{aligned}$$

Esimerkki 24. *Tuominen, 1: 33.* Laatikossa on 6 punaista ja 9 valkoista palloa. Kokeessa laatikosta nostetaan 3 palloa ilman takaisinpanoa. Laske tn:t tapahtumille

$$\begin{aligned}
A &= \text{"ainakin yksi on punainen"}, \\
B &= \text{"kaikki ovat punaisia"}, \\
C &= \text{"pallot ovat keskenään samanvärisiä"}, \\
D &= \text{"nostossa on sekä valkoisia että punaisia palloja"}.
\end{aligned}$$

Ratkaisu. Nyt $N = 6+9 = 15$, $K =$ "punaisten lukumäärä" $= 6$ ja $n = 3$. Merkitään

$$Q_k = \text{"otoksessa } k \text{ punaista palloa"}, \; k = 0,1,2,3.$$

Kyseessä on oitp, joten

$$\begin{aligned}
P(A) &= 1 - P(A^c) = 1 - P(Q_0) = 1 - \frac{\binom{6}{0}\binom{9}{3}}{\binom{15}{3}} = 1 - \frac{1 \cdot \frac{9!}{3!6!}}{\frac{15!}{3!12!}} \\
&= 1 - \frac{9!3!12!}{3!6!15!} = 1 - \frac{9 \cdot 8 \cdot 7}{15 \cdot 14 \cdot 13} = 1 - \frac{24}{130} = \frac{53}{65} \approx 0.815. \\
P(B) &= P(Q_3) = \frac{\binom{6}{3}\binom{9}{0}}{\binom{15}{3}} = \frac{\frac{6!}{3!3!} \cdot 1}{\frac{15!}{3!12!}} = \frac{6!3!12!}{3!3!15!} \\
&= \frac{6 \cdot 5 \cdot 4}{15 \cdot 14 \cdot 13} = \frac{4}{91} \approx 0.0439. \\
P(C) &= P(Q_0 \cup Q_3) = P(Q_0) + P(Q_3) - P(Q_0 \cap Q_3) \\
&= 1 - P(A) + P(B) - P(\emptyset) = 1 - \frac{53}{65} + \frac{4}{91} - 0 \\
&= \frac{1352}{5915} \approx 0.229. \\
P(D) &= 1 - P(D^c) = 1 - P(C) = 1 - \frac{1352}{5915} = \frac{4563}{5915} \approx 0.771.
\end{aligned}$$

Esimerkki 25. *Tuominen, 1: 43.* n palloa sijoitetaan umpimähkään k lokeroon. Millä tn:llä lokerossa 1 on tasan i palloa $(i = 0,1,\ldots,n)$?

Ratkaisu. Kyseessä on otp; valitaan n kertaa lokero k:sta mahdollisesta. Merkitään

$$A_i = \text{"lokero 1 valitaan } i \text{ kertaa"}, \; i = 0,1,\ldots,n.$$

Näin ollen

$$P(A_i) = \binom{n}{i}\frac{1^i(k-1)^{n-i}}{k^n} = \binom{n}{i}\frac{(k-1)^{n-i}}{k^n}.$$

Esimerkki 26. Jatkoa esimerkkiin 19. Lotossa arvotaan 7 varsinaista numeroa ja 3 lisänumeroa. Näin ollen ”täysin vääriä” numeroita on $39-(7+3)=29$ ja erilaisia rivejä saadaan

$$\frac{39!}{7!\,3!\,29!}=7.63\cdot 10^{10}.$$

Tätä lukua kutsutaan *multinomikertoimeksi*.

Esimerkki 27. Erään viikon aikana sattui 7 liikenneonnettomuutta. Millä todennäköisyydellä ne sattuivat kaikki eri päivinä?

Ratkaisu. Tulkitsemme viikonpäivät *lokeroiksi* ($k=7$) ja onnettomuudet *palloiksi* ($n=7$). Nyt $n_1=n_2=\cdots=n_7=1$, jolloin siis $n_1+n_2+\cdots+n_7=1+1+\cdots+1=7=n$ (tämä merkitsee siis sitä, että jokaisena viikonpäivänä sattuu yksi onnettomuus). Onnettomuuksien ”sijoitusmahdollisuuksien” lukumäärä on $k^n=7^7=823\,543$. Suotuisia tapauksia on $n!=7!=5040$ (onnettomuuksien tapahtumapäivien järjestysten lukumäärä). Kun merkitään

$$A=\text{”onnettomuudet sattuivat eri päivinä”},$$

niin saadaan

$$P(A)=\frac{n!}{n_1!\,n_2!\cdots n_7!}\cdot\frac{1}{k^n}=\frac{7!}{7^7}=\frac{5040}{823\,543}\approx 0.00612.$$

Esimerkki 28. *Otanta äärettömästä populaatiosta.* $N=5\,000\,000$ (Suomen väestö), $K=500\,000$ (Helsingissä asuvat) ja $n=3$ (otoskoko). Merkitään

$$H=\text{”otokseen osuu 1 Helsingissä asuva henkilö”}.$$

Yritetään laskea ensin oitp (jollainen esim. puhelinhaastattelu käytännössä on – eihän samalle henkilölle soiteta samassa kyselyssä kahta kertaa). Tällöin saadaan

$$\begin{aligned}P(H)&=\frac{\binom{500\,000}{1}\binom{4\,500\,000}{2}}{\binom{5\,000\,000}{3}}=\frac{500\,000\cdot\frac{4\,500\,000!}{2!\,4\,499\,998!}}{\frac{5\,000\,000!}{3!\,4\,999\,997!}}\\&=\frac{1\,500\,000\cdot 4\,500\,000!}{4\,499\,998\cdot 5\,000\,000!}\approx ?.\end{aligned}$$

Nyt kuitenkin

$$\begin{aligned}&n=3\ll 5\,000\,000=N,\\&n=3\ll 500\,000=K\text{ ja}\\&n=3\ll 4\,500\,000=5\,000\,000-500\,000=N-K,\end{aligned}$$

joten voidaan käyttää otp:

$$P(H)=\binom{3}{1}\frac{500\,000^1\cdot 4\,500\,000^2}{5\,000\,000^3}=\frac{3\cdot 500\,000\cdot 4\,500\,000^2}{5\,000\,000^3}\approx 0.135.$$

Esimerkki 29. Erään kirjan sivulla on 30 riviä, kullakin 60 merkkiä. Tällä sivulla on 10 painovirhettä. Millä tn:llä ne kaikki ovat samalla rivillä.

Ratkaisu. Lokeroina ovat nyt sivulla olevat $30 \cdot 60 = 1800$ merkkiä, *palloina* 10 painovirhettä. Koska kuhunkin *lokeroon* mahtuu vain yksi *pallo*, niin painovirheet voivat sijaita sivulla $\binom{1800}{10}$ eri tavalla. Pidämme näitä symmetrisinä alkeistapauksina. Suotuisien alkeistapausten lukumäärä päätellään seuraavasti: rivi voidaan valita 30 tavalla, jonka jälkeen kyseessä on 10 *pallon* sijoittaminen 60 *lokeroon*, kuhunkin *lokeroon* korkeintaan yksi pallo, eri tapoja on siis $\binom{60}{10}$. Kysytty tn on siis

$$\frac{30 \cdot \binom{60}{10}}{\binom{1800}{10}} \approx 30 \cdot \frac{60^{10}}{1800^{10}} = 30^{-9} \approx 5.1 \cdot 10^{-14}.$$

Esimerkki 30. *Ehdollinen todennäköisyys.* Nopanheitto.

$$A = \text{"saadaan kuutonen"} = \{6\} \subset \Omega = \{1, 2, \ldots, 6\},$$
$$B = \text{"saadaan parillinen silmäluku"} = \{2, 4, 6\}.$$

Nyt

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{1}{6}.$$

Tutkitaan sitten tapahtuman

$$C = \text{"saadaan kuutonen, kun tiedetään, että saadaan parillinen"}$$

todennäköisyyttä. Tällöin

$$P(C) = P(A \,|\, B) = \frac{1}{3} = \frac{n(A)}{n(B)}.$$

Merkitään nyt

$$D = \text{"saadaan viitonen" ja}$$
$$E = \text{"saadaan viitonen, kun tiedetään, että saadaan parillinen"}$$

Tällöin

$$P(E) = P(D \,|\, B) = 0 \neq \frac{1}{3} = \frac{n(D)}{n(B)}.$$

Pelastus:

$$P(E) = P(D \,|\, B) = 0 = \frac{0}{3} = \frac{n(D \cap B)}{n(B)}.$$

Esimerkki 31. *Tu 1: 50*. Osoita: Jos $P(A) = P(B) = \frac{2}{3}$, niin $P(A\,|\,B) \geqslant \frac{1}{2}$.

Todistus. Koska

$$P(A\,|\,B) \geqslant \frac{1}{2} \Leftrightarrow \frac{P(A\cap B)}{P(B)} \geqslant \frac{1}{2} \Leftrightarrow 2\cdot P(A\cap B) \geqslant P(B) = \frac{2}{3} \Leftrightarrow P(A\cap B) \geqslant \frac{1}{3},$$

niin riittää todistaa viimeinen epäyhtälö. Yhteenlaskukaavasta seuraa

$$\frac{4}{3} - P(A\cap B) = P(A) + P(B) - P(A\cap B) = P(A\cup B) \leqslant 1,$$

joten

$$P(A\cap B) \geqslant \frac{4}{3} - 1 = \frac{1}{3}. \qquad \square$$

Esimerkki 32. *Tuominen, 1: 45*. Korttipakasta vedetään 5 korttia (ilman takaisinpanoa). Laske tn, että mukana on ainakin yksi ässä ehdolla, että kaikkien arvo on vähintään 10.

Ratkaisu. Merkitään

A_k = "nostetuissa korteissa on täsmälleen k korttia $\geqslant 10$",

B = "mukana nostetuissa korteissa ainakin yksi ässä", ja

C = "kaikkien nostettujen korttien arvo $\geqslant 10$".

Nyt A_k:lle suotuisten alkeistapausten lkm on

$$n(A_k) = \binom{K}{k}\binom{N-K}{n-k},$$

missä $N = 52$, $n = 5$ ja K = "$\geqslant 10$ korttien lkm" $= 20$. Tällöin $C = A_5$, joten

$$n(C) = \binom{20}{5}\binom{32}{0} = \frac{20!}{5!\,15!}\cdot 1 = 15\,504$$

ja $B\cap C = C\setminus(C\setminus B)$, missä

$C\setminus B$ = "nostetut kortit ovat joukossa kympit, jätkät, rouvat ja kuninkaat".

Yhteensä siis

$$n(C\setminus B) = \binom{16}{5} = \frac{16!}{5!\,11!} = 4\,368,$$

joten

$$n(B\cap C) = n(C\setminus(C\setminus B)) = n(C) - n(C\setminus B) = 15\,504 - 4\,368 = 11\,136.$$

Nyt saadaan kysytty ehdollinen tn

$$P(B\,|\,C) = \frac{P(B\cap C)}{P(C)} = \frac{\frac{n(B\cap C)}{n(\Omega)}}{\frac{n(C)}{n(\Omega)}} = \frac{n(B\cap C)}{n(C)} = \frac{11\,136}{15\,504} \approx 0{,}718.$$

Esimerkki 33. Nopanheitto. $A = \{5,6\}$, $B = \{2,4,6\}$. Tällöin
$$P(A) = \frac{2}{6} = \frac{1}{3}$$
ja
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$
Ehto B ei siis "auta", se "ei lisää tietoa". Näin ollen A ja B ovat *riippumattomia*, merkitään $A \perp\!\!\!\perp B$.

Esimerkki 34. Korttipakan 52 kortista vedetään yksi kortti. Olkoon
$$A = \text{"kortti on ässä"},$$
$$B = \text{"kortti on pata"}.$$
Onko $A \perp\!\!\!\perp B$?

Ratkaisu. Nyt
$$P(A) = \frac{4}{52} = \frac{1}{13}, \quad P(B) = \frac{13}{52} \quad \text{ja} \quad P(A \cap B) = \frac{1}{52},$$
joten
$$P(A \cap B) = \frac{1}{52} = \frac{1}{13} \cdot \frac{13}{52} = P(A)P(B).$$
Siis riippumattomuuden määritelmän perusteella $A \perp\!\!\!\perp B$.

Esimerkki 35. Tarkastelemme kolmilapsisia perheitä. Pitämällä tyttö- ja poikalasten syntymiä yhtä todennäköisinä ja kunkin perheen lasten sukupuolia riippumattomina voimme kuvailla kolmilapsisia perheitä symmetrisellä tn-avaruudella, jossa on seuraavat 8 alkeistapausta:

(p,p,p) (p,p,t) (p,t,p) (t,p,p) (p,t,t) (t,p,t) (t,t,p) (t,t,t),

missä p = poika ja t = tyttö. Olkoon
$$A = \text{"perheessä on sekä tyttöjä että poikia"},$$
$$B = \text{"perheessä on enintään yksi tyttö"}.$$
Ovatko A ja B riippumattomia?

Ratkaisu. Nyt
$$P(A) = \frac{6}{8} = \frac{3}{4}, \quad P(B) = \frac{4}{8} = \frac{1}{2} \quad \text{ja} \quad P(A \cap B) = \frac{3}{8},$$
joten
$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{3}{8}}{\frac{3}{4}} = \frac{1}{2} = P(B).$$
Siis Lauseen 1.8.3 perusteella $A \perp\!\!\!\perp B$.

Esimerkki 36. *Tuominen, 1: 62.* Anna esimerkki tapahtumista A, B ja C, jotka eivät ole riippumattomia, vaikka

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

Ratkaisu. Heitetään lanttia kerran. Merkitään

$$A=\text{"tulee klaava"},\quad B=\text{"tulee kruuna"}\quad\text{ja}\quad C=\text{"ei tule kumpikaan"}.$$

Nyt $A\cap B\cap C=\emptyset$ ja $P(C)=0$, joten

$$P(A\cap B\cap C)=P(\emptyset)=0=P(A)P(B)P(C),$$

mutta

$$P(A\cap B)=P(\emptyset)=0\neq\frac{1}{4}=\frac{1}{2}\cdot\frac{1}{2}=P(A)P(B).$$

Siis A ja B eivät ole riippumattomia, joten A, B ja C eivät ole riippumattomia.

Esimerkki 37. *Toistokoe:* Vakioveikkaus. Merkitään

$$A_i=\text{"veikkaaja onnistuu } i\text{. rivissä"},\quad i=1,2,\dots,13,$$
$$B_k=\text{"veikkaajalla on } k \text{ oikein"},\quad i=0,1,\dots,13.$$

Tällöin

$$P(A_i)=p=\frac{1}{3},\quad\text{ja}\quad P(A_i^c)=1-p=q=\frac{2}{3},$$
$$P(B_k)=\binom{13}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{13-k},$$
$$P(B_{13})=\binom{13}{13}\left(\frac{1}{3}\right)^{13}\left(\frac{2}{3}\right)^{0}=\left(\frac{1}{3}\right)^{13}\approx 3.30\cdot 10^{-15}.$$

Pohdittavaa. Mitä vikaa mallissa on ajatellen todellista tilannetta vakioveikkauksessa?

Esimerkki 38. *Tuominen, 1: 71.* Laske n-kertaisessa toistokokeessa tn, että A sattuu i. kerralla ehdolla, että A sattuu k kertaa.

Ratkaisu. Merkitään

$$A_k=\text{"}A\text{ sattuu } k \text{ kertaa"}\quad\text{ja}\quad B_i=\text{"}A\text{ sattuu } i\text{. kerralla"}.$$

Tällöin

$$P(A_k\,|\,B_i)=\frac{P(A_k\cap B_i)}{P(A_k)}=\frac{\binom{n-1}{k-1}p^kq^{n-k}}{\binom{n}{k}p^kq^{n-k}}$$
$$=\frac{(n-1)!}{(k-1)!\,(n-k)!}\cdot\frac{k!\,(n-k)!}{n!}=\frac{k}{n},\quad k=0,1,\dots,n.$$

Esimerkki 39. On todettu, että tietty sairaus on 0.5 %:lla populaation jäsenistä. Diagnosointimenetelmä toteaa sairaan sairaaksi tn:llä 0.95 ja terveen sairaaksi tn:llä 0.01. Mikä on tn, että diagnosoitava henkilö todetaan sairaaksi?

Ratkaisu. Merkitään

$$T = \text{"henkilö todetaan sairaaksi"} \quad \text{ja} \quad S = \text{"henkilö on sairas"}$$

Nyt ositus on S ja S^c eli terveet ja sairaat. Kokonaistodennäköisyyden kaavan perusteella kysytty tn on

$$\begin{aligned} P(T) &= P(S)P(T \mid S) + P(S^c)P(T \mid S^c) \\ &= 0.005 \cdot 0.95 + 0.995 \cdot 0.01 \approx 0.015. \end{aligned}$$

Esimerkki 40. Kolme konetta valmistaa samanlaisia tuotteita. Ensimmäisen koneen päivätuotanto on 6 000, toisen 1 000 ja kolmennen 3 000 kpl. On havaittu tietyn vian esiintyvän 1. koneen valmistamista 10 %:lla, 2. koneen valmistamista 8 %:lla ja 3. koneen valmistamista 15 %:lla. Varastoitaessa päivätuotanto kolmen eri koneen valmistamat tuotteet menevät sekaisin. Mikä on tn, että poimittaessa varastosta umpimähkään yksi tuote siinä on mainittu vika?

Ratkaisu. Kuvailemme koetta symmetrisellä tn-avaruudella, jonka perusjoukon muodostavat päivätuotannon 10 000 tuotetta. Merkitään

$$\begin{aligned} A_i &= \text{"tuote on lähtöisin } i\text{. koneesta"}, \quad i = 1,2,3, \\ B &= \text{"tuotteessa on mainittu vika"}. \end{aligned}$$

Nyt saadaan tn:t

$$\begin{aligned} P(A_1) &= \frac{6\,000}{10\,000} = 0.6, \\ P(A_2) &= \frac{1\,000}{10\,000} = 0.1, \\ P(A_3) &= \frac{3\,000}{10\,000} = 0.3, \end{aligned}$$

sekä ehdolliset tn:t

$$P(B \mid A_1) = 0.10, \quad P(B \mid A_2) = 0.08 \quad \text{ja} \quad P(B \mid A_3) = 0.15.$$

Näin ollen kokonaistodennäköisyyden kaavaa käyttäen saamme

$$\begin{aligned} P(B) &= P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + P(A_3)P(B \mid A_3) \\ &= 0.06 \cdot 0.10 + 0.1 \cdot 0.08 + 0.3 \cdot 0.15 \approx 0.11. \end{aligned}$$

Esimerkki 41. *Ehdon kääntäminen – Bayesin kaava*. Jatkoa esimerkkiin 39. Oletetaan, että potilas diagnosoidaan sairaaksi. Mikä on tällä ehdolla tn, että hän todella on sairas? Entä terve?

Ratkaisu. Nyt kysytään siis ehdollista todennäköisyyttä

$$P(S \mid T) = \frac{P(S \cap T)}{P(T)}.$$

Kertolaskukaavan (L.1.7.4) perusteella

$$P(S \cap T) = P(S)P(T \mid S)$$

ja kokonaistodennäköisyyden kaavan perusteella

$$P(T) = P(S)P(T \mid S) + P(S^c)P(T \mid S^c).$$

Näin ollen saadaan

$$\begin{aligned} P(S \mid T) &= \frac{P(S)P(T \mid S)}{P(S)P(T \mid S) + P(S^c)P(T \mid S^c)} \\ &= \frac{0.005 \cdot 0.95}{0.005 \cdot 0.95 + 0.995 \cdot 0.01} \approx 0.323. \end{aligned}$$

Tästä saadaan, että

$$P(S^c \mid T) = 1 - P(S \mid T) \approx 0.677.$$

Siis valtaosan sairaiksi diagnosoiduista muodostavat terveet!

Esimerkki 42. *Tuominen 1:84*. Hajamielinen herra unohtaa sateenvarjonsa kaupassa käydessään tn:lla $\frac{1}{4}$. Eräänä päivänä hän on käynyt neljässä kaupassa, ja huomannut kotiin palattuaan sateenvarjonsa unohtuneen. Laske tn, että sateenvarjo on i. kaupassa, $i = 1, 2, 3, 4$.

Ratkaisu. Merkitään

$$\begin{aligned} A_i &= \text{"varjo on unohtunut kauppaan } i\text{"}, \quad i = 1, 2, 3, 4, \\ B &= \text{"varjo on unohtunut"}. \end{aligned}$$

Nyt siis

$$\begin{aligned} P(A_1) &= \frac{1}{4} = \frac{64}{256}, \quad P(A_2) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16} = \frac{48}{256}, \\ P(A_3) &= \left(\frac{3}{4}\right)^2 \cdot \frac{1}{4} = \frac{9}{64} = \frac{36}{256}, \quad P(A_4) = \left(\frac{3}{4}\right)^3 \cdot \frac{1}{4} = \frac{27}{256}. \end{aligned}$$

Koska $A_i \subset B$ kaikilla i, niin $P(B \mid A_i) = 1$ kaikilla i. Nyt saadaan

$$\sum_{k=1}^{4} P(A_k) = \frac{64}{256} + \frac{48}{256} + \frac{36}{256} + \frac{27}{256} = \frac{175}{256},$$

joten Bayesin kaavan perusteella

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{k=1}^{4} P(A_k)P(B \mid A_k)} = \frac{P(A_i)}{\sum_{k=1}^{4} P(A_k)} = \frac{256}{175} P(A_i)$$

kaikilla $i = 1, 2, 3, 4$. Saadaan siis esim. $P(A_1 \mid B) \approx 0.366$.

Esimerkki 43. Toistokokeiden yhteydessä käytimme kokeen kuvailuun tapahtumia

$$B_k = \text{"A:n esiintymiskertojen lukumäärä on } k\text{"}, \quad k = 0, 1, 2, \ldots.$$

Merkitsemme nyt

$$X = \text{"A:n esiintymiskertojem lukumäärä"}.$$

Tätä merkintätapaa käyttäen tapahtumat B_k ovat lausuttavissa muodossa

$$B_k = (X = k), \quad k = 0, 1, 2, \ldots.$$

Tällä lähestymistavalla on nyt se etu, että koko satunnaiskoe on voitu kuvailla yhden suureen, *satunnaismuuttujan* X, avulla. Satunnaismuuttujan määrittelyn kannalta on ratkaisevaa todeta:

Satunnaismuuttujan arvo on täysin määrätty, kun satunnaiskokeen tulos tunnetaan.

Esimerkki 44. Symmetristä rahaa heitetään kaksi kertaa. Käytämme kokeen kuvailuun tn-avaruutta $(\Omega, \mathcal{P}(\Omega), P)$, jonka perusjoukko Ω koostuu seuraavista neljästä alkeistapauksesta:

$$a = (\text{kr,kr}), \quad b = (\text{kr,kl}), \quad c = (\text{kl,kr}), \quad d = (\text{kl,kl}).$$

Olkoon

$$X = \text{"kruunien lukumäärä"}.$$

Tällöin X on satunnaismuuttuja; se on kuvaus $\Omega \to \mathbb{R}$, jolle

$$X(a) = 2, \quad X(b) = 1, \quad X(c) = 1, \quad X(d) = 0.$$

Vastaavasti kuvaus $Y : \Omega \to \mathbb{R}$,

$$Y(a) = 2, \quad Y(b) = 0, \quad Y(c) = 0, \quad Y(d) = 2,$$

on satunnaismuuttuja tässä avaruudessa, sanallisesti se voitaisiin ilmaista esimerkiksi:

$$Y = \text{"kruunien ja klaavojen lukumäärien ero"}.$$

Esimerkki 45. Heitetään kolikkoa, kunnes tulee ensimmäinen kruuna. Merkitään

$$X = \text{"sen heiton järjestysluku, jolloin tulee ensimmäinen kruuna".}$$

Tällöin X on diskreetti satunnaismuuttuja, jonka arvojoukko on $\{1, 2, 3, \ldots\}$ ja vastaavina pistetodennäköisyyksinä

$$
\begin{aligned}
p_1 &= P(X = 1) = \frac{1}{2}, \\
p_2 &= P(X = 2) = \frac{1}{2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^2 = \frac{1}{4}, \\
p_3 &= P(X = 3) = \left(\frac{1}{2}\right)^2 \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^3 = \frac{1}{8}, \\
&\vdots \\
p_k &= P(X = k) = \left(\frac{1}{2}\right)^{k-1} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^k.
\end{aligned}
$$

Esimerkiksi

$$
\begin{aligned}
&P(\text{"ensimmäinen kruuna tulee kymmenennellä heitolla"}) \\
&= p_{10} = P(X = 10) = \left(\frac{1}{2}\right)^{10}.
\end{aligned}
$$

Esimerkki 46. *Jatkoa esimerkkiin 45.* Määrätään X:n kertymäfunktio. Se on määritelmän mukaan funktio $F : \mathbb{R} \to \mathbb{R}$, jolle

$$F(x) = \sum_{k: x_k \leqslant x} p_k, \quad x \in \mathbb{R}.$$

Saamme siis

$$
F(x) = \begin{cases}
0, & x < 1, \\
\frac{1}{2} & 1 \leqslant x < 2, \\
\frac{1}{2} + \frac{1}{4} = \frac{3}{4}, & 2 \leqslant x < 3, \\
\frac{3}{4} + \frac{1}{8} = \frac{7}{8}, & 3 \leqslant x < 4, \\
& \vdots \\
\frac{2^k - 1}{2^k}, & k \leqslant x < k + 1, \\
& \vdots
\end{cases}
$$

Havaitsemme, että $F(x) \to 1$, kun $x \to \infty$. Kuitenkaan F ei koskaan saavuta 1:stä, miten tulkitset tämän?

Esimerkki 47. *Tuominen 2:18.* Lentoyhtiö tietää kokemuksesta, että keskimäärin 5 % paikan varanneista jää saapumatta koneeseen. Siksi yhtiö myykin 257 lippua koneeseen, johon mahtuu 250 matkustajaa. Millä tn:llä jokainen koneeseen saapuva saa paikan?

Ratkaisu. Merkitään

$$\begin{aligned} n &= \text{"myytyjen lippujen lukumäärä"} = 257, \\ A &= \text{"lipun ostanut matkustaja jää saapumatta koneeseen"}, \\ p &= P(A) = 0.05, \quad (q = 1 - p = 1 - 0.05 = 0.95) \\ X &= \text{"saapumatta jääneiden matkustajien lukumäärä"}. \end{aligned}$$

Kyseessä on toistokoe, jossa "toistoina" ovat myydyt liput, ja näin ollen X on binomijakautunut,

$$X \sim \mathrm{Bin}(n, p) = \mathrm{Bin}(257, 0.05).$$

Saadaan siis kysytty tn

$$\begin{aligned} &P(\text{"jokainen koneeseen saapuva saa paikan"}) \\ &= P(X \geqslant 7) = 1 - P(X \leqslant 6) \\ &= 1 - \left[P(X=0) + P(X=1) + P(X=2) + \cdots + P(X=6)\right] \\ &= 1 - \left[\binom{257}{0} 0.05^0 \cdot 0.95^{257} + \binom{257}{1} 0.05^1 \cdot 0.95^{256} + \cdots \right. \\ &\quad \left. + \binom{257}{6} 0.05^6 \cdot 0.95^{251}\right] \approx 0.975. \end{aligned}$$

Esimerkki 48. *Tuominen 2:19.* Jatkoa edelliseen esimerkkiin 47: Arvioi kysyttä tn:tä korvaamalla poisjäävien lkm:n jakauma sopivalla Poisson-jakaumalla.

Ratkaisu. Tilanteeseen sopiva Poisson-jakautunut sm on

$$X \sim \mathrm{Poisson}(np) = \mathrm{Poisson}(257 \cdot 0.05) = \mathrm{Poisson}(12.85).$$

Näin ollen saadaan kysytty arvio

$$\begin{aligned} &P(\text{"jokainen koneeseen saapuva saa paikan"}) \\ &= P(X \geqslant 7) = 1 - P(X \leqslant 6) \\ &= 1 - \left[P(X=0) + P(X=1) + P(X=2) + \cdots + P(X=6)\right] \\ &= 1 - e^{-12.85}\left[\frac{12.85^0}{0!} + \frac{12.85^1}{1!} + \frac{12.85^2}{2!} + \cdots + \frac{12.85^6}{6!}\right] \\ &\approx 0.972. \end{aligned}$$

Esimerkki 49. Jatkoa esimerkkiin 47. Olemme johtaneet (*Tuominen*, **Huomautus 2.1.10**) ptn:ien numeeriseksi laskemiseksi kaavan

$$p_{k+1} = \frac{n-k}{k+1} \cdot \frac{p}{q} \cdot p_k, \quad k = 0, \dots, n-1,\ p_k = P(X = k).$$

Koska nyt $\frac{p}{q} = \frac{0.05}{0.95} = \frac{1}{19}$, niin saadaan

$$\begin{aligned}
p_0 &= 0.95^{257} \approx 1.88 \cdot 10^{-6},\\
p_1 &= \frac{257}{1} \cdot \frac{1}{19} \cdot p_0 = \frac{257}{19} \cdot 0.95^{257} \approx 2.55 \cdot 10^{-5},\\
p_2 &= \frac{256}{2} \cdot \frac{1}{19} \cdot p_1 = \frac{128}{19} \cdot p_1 \approx 1.71 \cdot 10^{-4},\\
p_3 &= \frac{255}{3} \cdot \frac{1}{19} \cdot p_2 \approx 7.68 \cdot 10^{-4},\\
p_4 &= \frac{254}{4} \cdot \frac{1}{19} \cdot p_3 \approx 2.57 \cdot 10^{-3},\\
p_5 &= \frac{253}{5} \cdot \frac{1}{19} \cdot p_4 \approx 6.83 \cdot 10^{-3},\\
p_6 &= \frac{252}{6} \cdot \frac{1}{19} \cdot p_5 \approx 0.0151,\\
p_7 &= \frac{251}{7} \cdot \frac{1}{19} \cdot p_6 \approx 0.0285,\\
p_8 &= \frac{250}{8} \cdot \frac{1}{19} \cdot p_7 \approx 0.0469,\\
p_9 &= \frac{249}{9} \cdot \frac{1}{19} \cdot p_8 \approx 0.0683,\\
p_{10} &= \frac{248}{10} \cdot \frac{1}{19} \cdot p_9 \approx 0.0891,\\
p_{11} &= \frac{247}{11} \cdot \frac{1}{19} \cdot p_{10} \approx 0.105,\\
p_{12} &= \frac{246}{12} \cdot \frac{1}{19} \cdot p_{11} \approx 0.114,\\
p_{13} &= \frac{245}{13} \cdot \frac{1}{19} \cdot p_{12} \approx 0.113,\\
p_{14} &= \frac{244}{14} \cdot \frac{1}{19} \cdot p_{13} \approx 0.103,\\
&\vdots
\end{aligned}$$

Havaitsemme, että jakauman *moodi* on $k = 12$. Voimme tarkistaa tämän havainnon kaavalla

$$k = [np + p] = [257 \cdot 0.05 + 0.05] = [12.9] = 12.$$

Huom. Merkintä $[x]$ tarkoittaa x:n *kokonaisosaa*, se on siis suurin kokonaisluku, joka $\leqslant x$.

Esimerkki 50. Jatkoa esimerkkiin 24. Kyseessä on oitp ja näin ollen *hypergeometrinen jakauma*. Merkitään

$$X = \text{"punaisten pallojen lukumäärä otoksessa"},$$

jolloin

$$X \sim \mathrm{Hyperg}(N,K,n) = \mathrm{Hyperg}(15,6,3).$$

Saadaan siis

$$\begin{aligned}
P(A) &= P(X \geqslant 1) = 1 - P(X < 1) = 1 - P(X = 0) = \ldots \approx 0.815, \\
P(B) &= P(X = 3) = \ldots \approx 0.0439, \\
P(C) &= P((X = 0) \cup (X = 3)) = P(X = 0) + P(X = 3) = \ldots \approx 0.229, \\
P(D) &= P((X = 1) \cup (X = 2)) = P(X = 1) + P(X = 2) = \ldots \approx 0.771.
\end{aligned}$$

Esimerkki 51. Laitteella on eri käyttökerroilla edellisistä riippumatta sama tn $p = 0.05$ vioittua (laitteen käyttö ei siis "heikennä sen kuntoa"). Laske tn, että laite kestää vähintään 20 kertaa.

Ratkaisu. Olkoon

$$X = \text{"käyttökertojen lukumäärä ennen vioittumista"}.$$

Tällöin

$$X \sim \mathrm{Geom}(0.05) \quad \text{ja} \quad q = 1 - p = 0.95,$$

joten

$$\begin{aligned}
P(X \geqslant 20) &= 1 - P(X \leqslant 19) \\
&= 1 - \big[P(X = 0) + P(X = 1) + P(X = 2) + \cdots + P(X = 19)\big] \\
&= 1 - (0.05 \cdot 0.95^0 + 0.05 \cdot 0.95^1 + \cdots + 0.05 \cdot 0.95^{19}) \\
&= 1 - 0.05 \sum_{k=0}^{19} 0.95^k = 1 - 0.05 \cdot \frac{1 - 0.95^{20}}{1 - 0.95} \\
&= 1 - (1 - 0.95^{20}) = 0.95^{20} \approx 0.368.
\end{aligned}$$

Esimerkki 52. Jatkoa edelliseen esimerkkiin 51. Kysytty tn on kätevintä laskea X:n kertymäfunktion F_X avulla, saadaan siis

$$\begin{aligned}
P(X \geqslant 20) &= 1 - P(X \leqslant 19) = 1 - F_X(19) = 1 - (1 - 0.95^{20}) \\
&= 0.95^{20} \approx 0.368.
\end{aligned}$$

Esimerkki 53. Jatkoa esimerkkiin 51. Laske tn, että laite kestää vähintään 30 kertaa, kun tiedetään, että laite on jo kestänyt 10 kertaa.

Ratkaisu. Geometrisen jakauman *muistinmenetysominaisuuden* perusteella

$$P(X \geqslant 30 \,|\, X \geqslant 10) = P(X \geqslant 20) \approx 0.368.$$

Esimerkki 54. *Tuominen 2:16.* Suoritetaan riippumattomia kokeita siihen asti kunnes saadaan ensimmäinen positiivinen tulos; kullakin kerralla positiivisen tuloksen tn on 0.05. Kuvatkoon sm X suorituskertojen lukumäärää. Määritä

a) X:n arvojoukko,
b) X:n ptnf,
c) X:n kf,
d) X:n odotusarvo,
e) ”todennäköisin” suorituskertojen lkm (ns. moodi).

Ratkaisu. Nyt siis

$$X = \text{"kerta, jolloin saadaan ensimmäinen positiivinen tulos"}$$
$$p = \text{"positiivisen tuloksen tn"} = 0.05.$$

Sm X ei ole suoraan jakautunut geometrisesti, mutta sen sijaan

$$X - 1 \sim \mathrm{Geom}(0.05).$$

a) $X \in \{1, 2, 3, \ldots\} = \mathbb{N}_+$.

b) X:n ptnf on

$$f_X(k) = P(X = k) = P(X - 1 = k - 1) = 0.05 \cdot 0.95^{k-1},$$

missä $k \in \mathbb{N}_+$.

c) X:n kf on

$$F_X(k) = P(X \leqslant k) = \sum_{i=1}^{k} f_X(k) = \sum_{i=1}^{k} 0.05 \cdot 0.95^{i-1} = 0.05 \sum_{i=0}^{k-1} 0.95^i$$
$$= 0.05 \cdot \frac{1 - 0.95^{k-1+1}}{1 - 0.95} = 1 - 0.95^k, \quad k \in \mathbb{N}_+.$$

d) X:n odotusarvo on

$$EX = \sum_{k=1}^{\infty} k \cdot f_X(k) = \sum_{k=1}^{\infty} k \cdot 0.05 \cdot 0.95^{k-1} = 0.05 \sum_{k=1}^{\infty} k \cdot 0.95^{k-1}$$
$$= 0.05 \cdot \frac{1}{(1 - 0.95)^2} = \frac{0.05}{0.05^2} = \frac{1}{0.05} = 20,$$

sillä

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k x^{k-1}, \quad \text{kun } |x| < 1,$$

katso *Tuominen*, **Lause 2.2.3** (iii), todistus.

e) Nyt

$$f_X(k) = 0.05 \cdot 0.95^{k-1} > 0.05 \cdot 0.95^k = f_X(k+1)$$

kaikilla $k \in \mathbb{N}_+$, siis

$$f_X(1) > f_X(2) > \cdots > f_X(k) > f_X(k+1) > \cdots,$$

joten ”todennäköisin” suorituskertojen lukumäärä (moodi) on 1.

Esimerkki 55. Jatkoa esimerkkiin 48. Olemme johtaneet (*Tuominen*, **Huomautus 2.1.18**) ptn:ien numeeriseksi laskemiseksi kaavan

$$p_{k+1} = \frac{\lambda}{k+1} \cdot p_k, \quad k = 0, 1, 2, \ldots.$$

Koska nyt $\lambda = 12.85$, niin saadaan

$$\begin{aligned}
p_0 &= e^{-12.85} \approx 2.63 \cdot 10^{-6}, \\
p_1 &= \frac{12.85}{1} \cdot p_0 \approx 3.37 \cdot 10^{-5}, \\
p_2 &= \frac{12.85}{2} \cdot p_1 \approx 2.17 \cdot 10^{-4}, \\
p_3 &= \frac{12.85}{3} \cdot p_2 \approx 9.29 \cdot 10^{-4}, \\
p_4 &= \frac{12.85}{4} \cdot p_3 \approx 2.98 \cdot 10^{-3}, \\
p_5 &= \frac{12.85}{5} \cdot p_4 \approx 7.66 \cdot 10^{-3}, \\
p_6 &= \frac{12.85}{6} \cdot p_5 \approx 0.0164, \\
p_7 &= \frac{12.85}{7} \cdot p_6 \approx 0.0301, \\
p_8 &= \frac{12.85}{8} \cdot p_7 \approx 0.0484, \\
p_9 &= \frac{12.85}{9} \cdot p_8 \approx 0.0691, \\
p_{10} &= \frac{12.85}{10} \cdot p_9 \approx 0.0888, \\
p_{11} &= \frac{12.85}{11} \cdot p_{10} \approx 0.103, \\
p_{12} &= \frac{12.85}{12} \cdot p_{11} \approx 0.111, \\
p_{13} &= \frac{12.85}{13} \cdot p_{12} \approx 0.110, \\
p_{14} &= \frac{12.85}{14} \cdot p_{13} \approx 0.101, \\
&\vdots
\end{aligned}$$

Vertaamalla ptn:ien numeerisia arvoja esimerkissä 49 binomijakaumalla laskettuihin vastaaviin arvoihin havaitsemme, että Poisson-jakauma antaa varsin hyvän approksimaation binomijakaumalle (ainakin tässä tapauksessa). Lisäksi havaitsemme yllä olevista arvoista, että jakauman *moodi* on $k = 12$. Voimme tarkistaa tämän havainnon Poisson-jakaumalle kaavalla

$$k = [\lambda] = [12.85] = 12.$$

Esimerkki 56. Jatkoa esimerkkeihin 24 ja 50. Koska otoskoko $n = 3$, niin $p_k = P(X = k) = 0$, kun $k \geqslant 4$. Näin ollen X:n odotusarvo on

$$\begin{aligned}
EX &= \sum_{k=0}^{\infty} = \sum_{k=0}^{3} = \sum_{k=0}^{3} k \cdot \frac{\binom{6}{3}\binom{9}{3-k}}{\binom{15}{3}} \\
&= 0 \cdot \frac{\binom{6}{0}\binom{9}{3}}{\binom{15}{3}} + 1 \cdot \frac{\binom{6}{1}\binom{9}{2}}{\binom{15}{3}} + 2 \cdot \frac{\binom{6}{2}\binom{9}{1}}{\binom{15}{3}} + 3 \cdot \frac{\binom{6}{3}\binom{9}{0}}{\binom{15}{3}} \\
&= \frac{\binom{6}{1}\binom{9}{2} + 2 \cdot \binom{6}{2}\binom{9}{1} + 3 \cdot \binom{6}{3}\binom{9}{0}}{\binom{15}{3}} \\
&= \frac{6 \cdot \frac{9!}{2!7!} + 2 \cdot \frac{6!}{2!4!} \cdot 9 + 3 \cdot \frac{6!}{3!3!}}{\frac{15!}{3!12!}} = \frac{\frac{6 \cdot 9 \cdot 8}{2} + 2 \cdot \frac{6 \cdot 5}{2} \cdot 9 + 3 \cdot \frac{6 \cdot 5 \cdot 4}{3 \cdot 2}}{\frac{15 \cdot 14 \cdot 13}{3 \cdot 2}} \\
&= \frac{3 \cdot 9 \cdot 8 + 6 \cdot 5 \cdot 9 + 6 \cdot 5 \cdot 2}{5 \cdot 7 \cdot 13} = \frac{546}{455} = \frac{6}{5} = 1.2.
\end{aligned}$$

Esimerkki 57. Jatkoa edelliseen esimerkkiin. Koska $X \sim \text{Hyperg}(15, 6, 3)$, niin Lauseen 2.2.3 (ii) perusteella saamme X:n odotusarvon suoraan kaavaan sijoittamalla

$$EX = 3 \cdot \frac{6}{15} = \frac{6}{5} = 1.2.$$

Esimerkki 58. (Pietarin paradoksi) (*Nicolas Bernoulli*). Eräässä peliluolassa pelataan seuraavin säännöin:

- Pelaaja saa luolan pitäjältä S €.
- Pelaaja heittää kolikkoa. Jos tulee kruuna, peli päättyy, pelaaja maksaa takaisin 1 € ja saa pitää loput. Jos taas tulee klaava, niin pelaajan on jatkettava kolikkoa heittämällä kunnes tulee ensimmäinen kruuna.
- Jos ensimmäinen kruuna tulee n. kerralla, niin peli päättyy ja pelaaja maksaa luolan pitäjälle 2^{n-1} €.

Olkoon X pelaajan voittoa kuvaava satunnaismuuttuja. Pyrimme selvittämään, että mikä tulee S:n olla, jotta peli olisi reilu, so. $EX = 0$. Nyt

$$\begin{aligned}P(X = S - 2^{n-1}) &= P(\text{"1. kruuna } n\text{. heitolla''})\\ &= P((\underbrace{\text{kl,kl,kl,}\ldots\text{,kl}}_{n-1}\text{,kr})) = \left(\frac{1}{2}\right)^n,\end{aligned}$$

joten

$$\begin{aligned}\sum_{n=1}^{\infty} x_n \cdot P(X = x_n) &= \sum_{n=1}^{\infty} (S - 2^{n-1}) \cdot P(X = S - 2^{n-1})\\ &= \sum_{n=1}^{\infty} (S - 2^{n-1}) \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty} \left(\frac{S}{2^n} - \frac{1}{2}\right) = -\infty.\end{aligned}$$

Siis EX **ei ole olemassa!** Edelleen tarkastelumme mukaan, olipa S miten suuri hyvänsä, peli on "keskimääräisessä mielessä" pelaajalle äärettömän tappiollinen (ja luolan pitäjälle äärettömän voitollinen). Tätä tulosta on aikanaan pidetty varsin yllättävänä, jopa järjenvastaisena. Eräs tapa selvittää "paradoksi" on todeta, että käytännössä pelaajan maksama summa on rajoitettu, ja muuttaa sääntöjä seuraavasti:

- Olkoon maksimisumma, jonka pelaaja voi joutua maksamaan luolan pitäjälle, esimerkiksi 10 000 €.
- Koska
$$2^{13} < 10\,000 < 2^{14}.$$
niin säännöt ovat kuten yllä, kun $n < 15$.
- Jos taas 14 ensimmäisellä kerralla tulee pelkkiä klaavoja, niin peli päättyy ja pelaaja maksaa luolan pitäjälle 10 000 €.

Näillä säännöillä peli on reilu, sillä

$$\begin{aligned}EX &= \sum_{n=1}^{\infty} x_n \cdot P(X = x_n) = \sum_{n=1}^{14} (S - 2^{n-1}) \left(\frac{1}{2}\right)^n + \sum_{n=15}^{\infty} (S - 10\,000) \left(\frac{1}{2}\right)^n\\ &= S\left(1 - \left(\frac{1}{2}\right)^{14}\right) - 7 + (S - 10\,000)\left(\frac{1}{2}\right)^{14} = S - 7 - \frac{10\,000}{2^{14}} = 0\\ \Leftrightarrow S &= 7 + \frac{10\,000}{2^{14}} \approx 7.61.\end{aligned}$$

Esimerkki 59. Määrää vakio c siten, että funktio $f:\mathbb{R}\to\mathbb{R}$,

$$f(x)=\frac{c}{1+x^2},\quad x\in\mathbb{R},$$

on tiheysfunktio.

Ratkaisu. Havaitsemme, että **Huomautuksen 2.3.2** ehdot ovat tässä tapauksessa yhtäpitäviä ehtojen

(i) $c\geqslant 0$

(ii) $c\int_{-\infty}^{\infty}\frac{1}{1+x^2}\,dx=c\Big/_{-\infty}^{\infty}\arctan x=c\cdot\left(\frac{\pi}{2}-\left(-\frac{\pi}{2}\right)\right)=c\cdot\pi=1$

kanssa. Näemme, että (i) ja (ii) toteutuvat jos ja vain jos $c=\frac{1}{\pi}$.

Esimerkki 60. Jatkoa edelliseen esimerkkiin 59. Jos X on satunnaismuuttuja, jonka tiheysfunktio on f, mitkä ovat todennäköisyydet $P(X>0)$ ja $P(|X|\geqslant 1)$?

Ratkaisu.

$$\begin{aligned}P(X>0)&=\int_0^{\infty}f(x)\,dx=\frac{1}{\pi}\int_0^{\infty}\frac{1}{1+x^2}\,dx==\frac{1}{\pi}\Big/_{0}^{\infty}\arctan x\\&=\frac{1}{\pi}\left(\frac{\pi}{2}-0\right)=\frac{1}{2}.\\P(|X|\geqslant 1)&=1-P(|X|<1)=1-P(-1<X<1)\\&=1-\int_{-1}^{1}f(x)\,dx=1-\frac{1}{\pi}\int_{-1}^{1}\frac{1}{1+x^2}\,dx\\&=1-\frac{1}{\pi}\Big/_{-1}^{1}\arctan x=1-\frac{1}{\pi}(\arctan 1-\arctan(-1))\\&=1-\frac{1}{\pi}\left(\frac{\pi}{4}+\frac{\pi}{4}\right)=1-\frac{1}{2}=\frac{1}{2}.\end{aligned}$$

Esimerkki 61. Jatkoa edellisiin esimerkkeihin 59 ja 60. Mikä on X:n kertymäfunktio F_X? Määritä $P(X=0)$.

Ratkaisu.

$$\begin{aligned}F_X(x)&=P(X\leqslant x)=\int_{-\infty}^{x}f(x)\,dx=\frac{1}{\pi}\int_{-\infty}^{x}\frac{1}{1+x^2}\,dx\\&=\frac{1}{\pi}\Big/_{-\infty}^{x}\arctan x=\frac{1}{\pi}\left(\arctan x+\frac{\pi}{2}\right).\end{aligned}$$

Kertymäfunktio F_X on integraalifunktiona kaikkialla jatkuva, joten

$$P(X=0)=F_X(0)-F_X(0-)=F_X(0)-\lim_{x\to 0-}F_X(x)=0.$$

Esimerkki 62. Kuvatkoon jatkuva sm X radion ikää, joka on eksponenttijakautunut parametrina $\frac{1}{10}$. Tämä merkitsee sitä, että radion iän odotusarvo on $EX = 10$ (yksikkönä vuosi). Tällöin

$$\begin{aligned}
&P(\text{"radio kestää vähintään 10 vuotta"}) = P(X \geqslant 10)\\
&\quad = \int_{10}^{\infty} \frac{1}{10} e^{-\frac{x}{10}}\,dx = e^{-1} \approx 0.368.\\
&P(\text{"radio kestää vähintään 30 vuotta, kun tiedetään, että se on}\\
&\quad \text{kestänyt jo 20 vuotta"}) = P(X \geqslant 30 \,|\, X > 20)\\
&\quad = P(X \geqslant 10) = e^{-1} \approx 0.368.
\end{aligned}$$

Esimerkki 63. Herra K valvoo työkseen konetta, jonka häiriötön toiminta-aika on $\mathrm{Exp}\left(\frac{1}{20}\right)$ jakautunut sm (yksikkönä tunti). Jos koneeseen tulee häiriö, se on heti pysäytettävä; muuten se räjähtää. Työpäivä on kello 8–16, johon sisältyy luvallinen ruokatunti 11–12. Ruokatunnin aikana kukaan ei valvo konetta.

Perjantaina Herra K ei palaakaan ruokatunnilta töihin. Johtaja saapuu kello 16 ja havaitsee koneen räjähtäneen. Laske tn, että syy on herra K:n lintsauksessa.

Ratkaisu. Merkitään

$$X = \text{"koneen häiriötön toiminta-aika"},$$

jolloin $X \sim \mathrm{Exp}\left(\frac{1}{20}\right)$. Kysytty tn on siis

$$\begin{aligned}
P(12 \leqslant X \leqslant 16 \,|\, 11 \leqslant X \leqslant 16) &= \frac{P((12 \leqslant X \leqslant 16) \cap (11 \leqslant X \leqslant 16))}{P(11 \leqslant X \leqslant 16)}\\
= \frac{P(12 \leqslant X \leqslant 16)}{P(11 \leqslant X \leqslant 16)} &= \frac{F_X(16) - F_X(12)}{F_X(16) - F_X(11)} = \frac{1 - e^{-\frac{16}{20}} - \left(1 - e^{-\frac{12}{20}}\right)}{1 - e^{-\frac{16}{20}} - \left(1 - e^{-\frac{11}{20}}\right)}\\
= \frac{-e^{-\frac{16}{20}} + e^{-\frac{12}{20}}}{-e^{-\frac{16}{20}} + e^{-\frac{11}{20}}} &\approx 0.780.
\end{aligned}$$

Esimerkki 64. Olkoon $X \sim N(0, 1)$. Tällöin

$$\begin{aligned} P(X \leqslant 1) &= \Phi(1) \approx 0.841345, \\ P(X > 2) &= 1 - P(X \leqslant 2) = 1 - \Phi(2) \\ &\approx 1 - 0.977250 = 0.022750, \\ P(1 < X \leqslant 2) &= P(X \leqslant 2) - P(X \leqslant 1) = \Phi(2) - \Phi(1) \\ &\approx 0.977250 - 0.841345 = 0.135905. \end{aligned}$$

Esimerkki 65. *Tuominen 2:38.* Eräässä väestössä miespuolisen henkilön pituus on sm X, joka noudattaa normaalijakaumaa; $\mu = 178$, $\sigma = 5$ (yksikkönä cm). Määritä

a) $P(168 \leqslant X \leqslant 188)$,
b) $P(X > 188)$,
c) $P(X < 193 \,|\, X > 188)$.

Ratkaisu.

a)

$$\begin{aligned} P(168 \leqslant X \leqslant 188) &= P\left(\frac{168 - 178}{5} \leqslant \frac{X - 178}{5} \leqslant \frac{188 - 178}{5}\right) \\ &= P\left(-2 \leqslant \frac{X - 178}{5} \leqslant 2\right) = \Phi(2) - \Phi(-2) \\ &= \Phi(2) - (1 - \Phi(2)) = 2\Phi(2) - 1 \\ &\approx 2 \cdot 0.977250 - 1 \approx 0.955. \end{aligned}$$

b)

$$\begin{aligned} P(X > 188) &= 1 - P(X \leqslant 188) = 1 - P\left(\frac{X - 178}{5} \leqslant \frac{188 - 178}{5}\right) \\ &= 1 - P\left(\frac{X - 178}{5} \leqslant 2\right) = 1 - \Phi(-2) \approx 1 - 0.977250 \\ &\approx 0.0228. \end{aligned}$$

c)

$$\begin{aligned} P(188 < X < 193) &= P\left(\frac{188 - 178}{5} \leqslant \frac{X - 178}{5} \leqslant \frac{193 - 178}{5}\right) \\ &= P\left(2 \leqslant \frac{X - 178}{5} \leqslant 3\right) = \Phi(3) - \Phi(2) \\ &\approx 0.998650 - 0.977250 = 0.021400, \end{aligned}$$

joten kysyttu ehdollinen tn on

$$\begin{aligned} P(X < 193 \,|\, X > 188) &= \frac{P(188 < X < 193)}{P(X > 188)} \\ &\approx \frac{0.021400}{0.02275} \approx 0.941. \end{aligned}$$

Esimerkki 66. Olkoon $X \sim \mathrm{Tas}(0,10)$. Tällöin

$$f_X(x) = \begin{cases} \frac{1}{10}, & 0 < x < 10, \\ 0, & \text{muulloin.} \end{cases}$$

Näin ollen X:n odotusarvo on

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_0^{10} \frac{1}{10}\, x\, dx = \frac{1}{10} \Big/_{0}^{10} \frac{1}{2} x^2 \\ &= \frac{1}{20}(100 - 0) = \frac{100}{20} = 5. \end{aligned}$$

Esimerkki 67. Jatkoa edelliseen esimerkkiin 66. X:n odotusarvo saadaan helposti **Lauseen 2.4.3** kohdan (i) perusteella

$$EX = \frac{0+10}{2} = \frac{10}{2} = 5.$$

Esimerkki 68. Jatkoa esimerkkiin 63. Koska koneen häiriötöntä toiminta-aikaa kuvaava sm $X \sim \mathrm{Exp}\left(\frac{1}{20}\right)$, niin X:n odotusarvo on **Lauseen 2.4.3** kohdan (ii) mukaan

$$EX = \frac{1}{\frac{1}{20}} = 20 \quad \text{(tuntia).}$$

Esimerkki 69. Normaalijakautuneen sm:n odotusarvo on **Lauseen 2.4.3** kohdan (iii) mukaan suoraan jakauman ensimmäinen parametri. Esimerkissä 64 käsiteltiin standardinormaalijakaumaa $X \sim N(0,1)$, joten X:n odotusarvo on $EX = \mu = 0$. Esimerkissä 65 käsiteltiin yleistä normaalijakaumaa $X \sim N(178, 5^2)$, joten X:n odotusarvo on $EX = \mu = 178$ (cm).

Esimerkki 70. *Tuominen 2:44*. Millainen on satunnaismuuttujan $\mathbf{1}_A$ kertymäfunktio, kun $A \in \mathcal{F}$, $P(A) = p$?

Ratkaisu. Nyt $P(A^c) = 1 - p$. *Tuominen*, esimerkin 2.5.2 mukaan kysytty kf on

$$F(x) = P(\mathbf{1}_A \leqslant x) = \begin{cases} P(\emptyset), & x < 0, \\ P(A^c), & 0 \leqslant x < 1, \\ P(\Omega), & x \geqslant 1, \end{cases} = \begin{cases} 0, & x < 0, \\ 1 - p, & 0 \leqslant x < 1, \\ 1, & x \geqslant 1. \end{cases}$$

Esimerkki 71. *Tuominen 2:69*. Osoita, että $X + Y$ on sm olettaen, että X ja Y ovat sm:ia samassa tn-avaruudessa $(\Omega, \mathcal{F}, P)$.

Todistus. Olkoon $x \in \mathbb{R}$ mielivaltainen. Osoittaaksemme, että $X + Y$ on sm, tutkimme joukkoa

$$(X + Y > x) = \{\omega \in \Omega : X(\omega) + Y(\omega) > x\}.$$

Rationaalilukujen joukko on numeroituva; olettakaamme, että $q_1, q_2, \ldots$ on jokin rationaalilukujen numerointi. Väitämme, että

$$(X + Y > x) = \bigcup_{i=1}^{\infty} \{\omega \in \Omega : X(\omega) > q_i\} \cap \{\omega \in \Omega : Y(\omega) > x - q_i\}. \tag{1}$$

Jos ω on yhtälön (1) oikealla puolella olevan joukon alkio, on olemassa rationaaliluku q_k siten, että

$$X(\omega) > q_k \quad \text{ja} \quad Y(\omega) > x - q_k,$$

josta seuraa, että

$$X(\omega) + Y(\omega) > x$$

eli ω on myös (1):ssa vasemmalla puolella olevan joukon alkio. Jos kääntäen ω kuuluu joukkoon $(X + Y > x$, niin

$$X(\omega) + Y(\omega) > x \quad \text{eli} \quad X(\omega) > x - Y(\omega).$$

Koska rationaalilukujen joukko on tiheä $\mathbb{R}$:ssä (so. kahden reaaliluvun välissä on aina rationaaliluku), on olemassa rationaaliluku q_k siten, että

$$X(\omega) > q_k > x - Y(\omega)$$

eli

$$X(\omega) > q_k \quad \text{ja} \quad Y(\omega) > x - q_k.$$

Siis ω kuuluu myös yhtälön (1) oikealla puolella olevaan joukkoon. Yhtälöstä (1) seuraa nyt, että $(X + Y > x) \in \mathcal{F}$, sillä

$$\{\omega \in \Omega : X(\omega) > q_i\} \in \mathcal{F} \quad \text{ja} \quad \{\omega \in \Omega : Y(\omega) > x - q_i\} \in \mathcal{F} \quad \text{kaikilla } i$$

ja tapahtumien numeroituva yhdiste on edelleen tapahtuma. Komplementtiin siirtymällä saamme

$$(X + Y \leqslant x) \in \mathcal{F}. \qquad \square$$

Esimerkki 72. Eksponenttijakautuneet satunnaismuuttujat X ja Y kuvaavat vahvistimen ja CD-soittimen kestoa (siis ikää) parametreinaan $\frac{1}{500}$ ja $\frac{1}{100}$ (yksikkönä tunti). Laske todennäköisyydet tapahtumille

A = "sekä vahvistin että CD-soitin kestävät yli 100 tuntia",
B = "ainakin toinen laitteista kestää yli 100 tuntia",

kun oletetaan, että $X \perp\!\!\!\perp Y$.

Ratkaisu. Merkitään

$$\begin{aligned} X &= \text{"vahvistimen ikä"} \sim \operatorname{Exp}\left(\tfrac{1}{500}\right), \\ Y &= \text{"CD-soittimen ikä"} \sim \operatorname{Exp}\left(\tfrac{1}{100}\right). \end{aligned}$$

Koska $X \perp\!\!\!\perp Y$, niin saadaan

$$\begin{aligned} P(A) &= P((X > 100) \cap (Y > 100)) = P(X > 100)\, P(Y > 100) \\ &= (1 - P(X \leqslant 100))\,(1 - P(Y \leqslant 100)) \\ &= \left(1 - \left(1 - e^{-\frac{1}{500}\cdot 100}\right)\right)\left(1 - \left(1 - e^{-\frac{1}{100}\cdot 100}\right)\right) \\ &= e^{-\frac{1}{5}} \cdot e^{-1} = e^{-\frac{6}{5}} \approx 0.301, \\ P(B) &= P((X > 100) \cup (Y > 100)) \\ &= P(X > 100) + P(Y > 100) - P(X > 100)\, P(Y > 100) \\ &= (1 - P(X \leqslant 100)) + (1 - P(Y \leqslant 100)) - P(X > 100)\, P(Y > 100) \\ &= e^{-\frac{1}{5}} + e^{-1} - e^{-\frac{6}{5}} \approx 0.885. \end{aligned}$$

Esimerkki 73. Oletetaan, että huoneessa on 5 lamppua ja että lamppujen toiminta on toisistaan riippumatonta. Jokaisen lampun kestoa (siis ikää) kuvaava satunnaismuuttuja on eksponenttijakautunut parametrinaan $\frac{1}{150}$ (yksikkönä tunti). Laske tn, että huoneessa tulee täysi pimeys 100 tunnissa.

Ratkaisu. Merkitään

$$X_i = \text{"}i\text{:nnen lampun ikä"} \sim \operatorname{Exp}\left(\tfrac{1}{150}\right), \quad i = 1, \ldots, 5.$$

Koska X_i:t ovat toisistaan riippumattomia, niin saadaan

$$\begin{aligned} &P(\text{"huoneessa tulee täysi pimeys 100 tunnissa"}) \\ &\quad = P(\max\{X_1, X_2, X_3, X_4, X_5\} \leqslant 100) = P\Big(\bigcap_{i=1}^{5} (X_i \leqslant 100)\Big) \\ &\quad = \prod_{i=1}^{5} F_{X_i}(100) = \left(1 - e^{-\frac{100}{150}}\right)^5 = \left(1 - e^{-\frac{2}{3}}\right)^5 \approx 0.027. \end{aligned}$$

Esimerkki 74. Olkoon satunnaismuuttujalla X jatkuva jakauma, tiheysfunktiona f_X. Johda satunnaismuuttujan $Y = |X|$ tiheysfunktio. Sovella tulosta tapauksiin a) $X \sim \mathrm{Tas}(-1,1)$, b) $X \sim N(0,1)$.

Ratkaisu. Selvästi

$$F_Y(y) = P(|X| \leqslant y) = 0, \quad \text{kun } y \leqslant 0,$$

joten

$$f_Y(y) = 0, \quad \text{kun } y \leqslant 0.$$

Jos $y > 0$, niin

$$F_Y(y) = P(|X| \leqslant y) = P(-y \leqslant X \leqslant y) = F_X(y) - F_X(-y),$$

joten

$$f_Y(y) = F'_Y(y) = F'_X(y) - F'_X(-y) = f_X(y) + f_X(-y).$$

Näin ollen

$$f_Y(y) = \begin{cases} 0, & y \leqslant 0, \\ f_X(y) + f_X(-y), & y > 0. \end{cases}$$

a) Jos $X \sim \mathrm{Tas}(-1,1)$, niin

$$f_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < 1, \\ 0, & \text{muulloin.} \end{cases}$$

Näin ollen

$$f_Y(y) = \begin{cases} \frac{1}{2} + \frac{1}{2} = 1, & 0 < y < 1, \\ 0, & \text{muulloin.} \end{cases}$$

Siis $Y \sim \mathrm{Tas}(0,1)$.

b) Jos $X \sim N(0,1)$, niin

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Näin ollen

$$f_Y(y) = \begin{cases} 0, & y \leqslant 0, \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{(-y)^2}{2}}, & y > 0, \end{cases} = \begin{cases} 0, & y \leqslant 0, \\ \sqrt{\frac{2}{\pi}}\, e^{-\frac{y^2}{2}}, & y > 0. \end{cases}$$

Esimerkki 75. Olkoon X:n jakauma diskreetin ja jatkuvan "sekoitus". Jos $D \subset \mathbb{R}$ on X:n epäjatkuvuuspisteiden joukko, niin

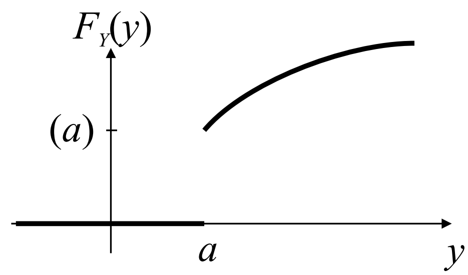
$$F_X(x) = \int_{-\infty}^{x} f_1(y)\,dy + \sum_{y \leqslant x,\, y \in D} f_2(y),$$

missä f_1 on X:n tiheysfunktio joukossa $\mathbb{R} \setminus D$ ja f_2 pistetodennäköisyysfunktio joukossa D. Tällöin

$$EX = \int_{-\infty}^{\infty} x\, f_1(x)\,dx + \sum_{x \in D} x\, f_2(x).$$

Sovellus. Olkoon nyt $X \sim N(0,1)$ ja $Y = \max(X, a)$, missä $a \in \mathbb{R}$. Tällöin

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P((X \leqslant y) \cap (a \leqslant y)) \\ &= \begin{cases} 0, & y < a, \\ \int_{-\infty}^{y} \varphi(t)\,dt, & y \geqslant a, \end{cases} = \begin{cases} 0, & y < a, \\ \int_{-\infty}^{a} \varphi(t)\,dt + \int_{a}^{y} \varphi(t)\,dt, & y \geqslant a. \end{cases} \end{aligned}$$



Pisteessä a on "hyppy"

$$F_Y(a) - \lim_{y \to a-} F_Y(y) = F_Y(a) - 0 = \Phi(a).$$

Siis

$$f_2(y) = \begin{cases} \Phi(a), & y = a, \\ 0, & \text{muulloin,} \end{cases}$$

$$f_1(y) = \begin{cases} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2}, & y > a, \\ 0, & \text{muulloin.} \end{cases}$$

Näin ollen

$$\begin{aligned} EY &= \int_{-\infty}^{\infty} y\, f_1(y)\,dy + \sum_{y \in a} y\, f_2(y) = \int_{a}^{\infty} y \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2}\,dy + a\,\Phi(a) \\ &= -\frac{1}{\sqrt{2\pi}} \int_{a}^{\infty} -y\, e^{-\frac{1}{2}y^2}\,dy + a\,\Phi(a) = -\frac{1}{\sqrt{2\pi}} \Big/_{a}^{\infty} e^{-\frac{1}{2}y^2} + a\,\Phi(a) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}a^2} + a\,\Phi(a). \end{aligned}$$

Esimerkki 76. *Tuominen 3:1.* Laske EX^2 ja $E\sin(2\pi X)$, kun $X \sim \mathrm{Tas}(0,1)$.

Ratkaisu. Koska $X \sim \mathrm{Tas}(0,1)$, niin

$$f_X(x) = \begin{cases} 1, & \text{kun } 0 < x < 1, \\ 0, & \text{muulloin.} \end{cases}$$

Lauseen 3.1.8 (ii) mukaan

$$EX^2 = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx = \int_0^1 x^2\, dx = \Big/_{\!0}^{1} \frac{1}{3}x^3 = \frac{1}{3} - 0 = \frac{1}{3},$$

ja

$$\begin{aligned} E\sin(2\pi X) &= \int_{-\infty}^{\infty} \sin(2\pi x)\, f_X(x)\, dx = \int_0^1 \sin(2\pi x)\, dx \\ &= \frac{1}{2\pi}\int_0^1 2\pi \cdot \sin(2\pi x)\, dx = \frac{1}{2\pi} \Big/_{\!0}^{1} -\cos(2\pi x) \\ &= \frac{1}{2\pi}(-1-(-1)) = 0. \end{aligned}$$

Esimerkki 77. Olkoon X ja Y riippumattomia satunnaismuuttujia, joille $EX = EY = \mu$ ja $D^2X = D^2Y = \sigma^2$. Tällöin

$$D^2X = EX^2 - (EX)^2,$$

joten

$$\sigma^2 = EX^2 - \mu^2.$$

Näin ollen

$$EX^2 = \sigma^2 + \mu^2,$$

ja vastaavasti $EY^2 = \sigma^2 + \mu^2$. Koska $X \perp\!\!\!\perp Y$, niin saadaan

$$\begin{aligned} D^2(XY) &= E(XY)^2 - (E(XY))^2 = E(X^2\,Y^2) - (EX\,EY)^2 \\ &= EX^2\,EY^2 - (EX)^2\,(EY)^2 = (\sigma^2+\mu^2)^2 - \mu^2\cdot\mu^2 \\ &= (\sigma^2+\mu^2)^2 - \mu^4. \end{aligned}$$

Huomaa, että jos $\mu, \sigma \neq 0$, niin

$$D^2(XY) = (\sigma^2+\mu^2)^2 - \mu^4 \neq \sigma^4 = \sigma^2\cdot\sigma^2 = D^2X\,D^2Y.$$

Esimerkki 78. (Steinerin sääntö) Jos X:llä on varianssi, niin kaikilla $c \in \mathbb{R}$ on voimassa

$$D^2X \leqslant E(X-c)^2.$$

Todistus.

$$\begin{aligned}
E(X-c)^2 &= E\big((X-EX)+(EX-c)\big)^2 \\
&= E\big((X-EX)^2 + 2(X-EX)(EX-c) + (EX-c)^2\big) \\
&= E(X-EX)^2 + 2(EX-c)\,E(X-EX) + (EX-c)^2 \\
&= D^2X + (EX-c)^2 \geqslant D^2X.
\end{aligned}$$

Tästä nähdään lisäksi, että $E(X-c)^2$ saa aidon absoluuttisen minimiarvon, kun $c = EX$. $\square$

Esimerkki 79. *Tuominen 3:10.* Suorakulmion sivujen todelliset pituudet ovat a ja b. Mittauksessa tapahtuu virhe siten, että mittaustulokset ovat

$$a+X,\ \ b+Y,$$

missä $X \perp\!\!\!\perp Y$, X ja Y ovat kumpikin jakaumaltaan $\mathrm{Tas}(-1,1)$, $a,b > 1$. Olkoon Z mittausvirheiden aiheuttama pinta-alan virhe. Laske EZ ja D^2Z.

Ratkaisu. Koska $X, Y \sim \mathrm{Tas}(-1,1)$, niin Lauseen 2.4.3 (i) ja 3.2.7 (vi) mukaan

$$EX = EY = \frac{-1+1}{2} = 0$$

ja

$$D^2X = D^2Y = \frac{(1-(-1))^2}{12} = \frac{1}{3}.$$

Tästä seuraa, että

$$EX^2 = EX^2 - 0^2 = EX^2 - (EX)^2 = D^2X = \frac{1}{3},$$

ja samoin saadaan, että $EY^2 = \frac{1}{3}$. Lisäksi saadaan, että pinta-alan virhe on

$$Z = (a+X)(b+Y) - ab = ab + bX + aY + XY - ab = XY + bX + aY.$$

Koska $X \perp\!\!\!\perp Y$, niin näin ollen

$$\begin{aligned}
EZ &= E(XY + bX + aY) = EX\,EY + bEX + aEY \\
&= 0\cdot 0 + b\cdot 0 + a\cdot 0 = 0
\end{aligned}$$

ja

$$\begin{aligned}
D^2Z &= EZ^2 - (EZ)^2 = E(XY + bX + aY)^2 - 0^2 \\
&= E\big(X^2Y^2 + b^2X^2 + a^2Y^2 + 2(bX^2Y + aXY^2 + abXY)\big) \\
&= EX^2EY^2 + b^2EX^2 + a^2EY^2 \\
&\quad + 2(bEX^2EY + aEXEY^2 + abEXEY) \\
&= \frac{1}{3}\cdot\frac{1}{3} + b^2\cdot\frac{1}{3} + a^2\cdot\frac{1}{3} + 2\cdot 0 = \frac{1}{9}(3a^2 + 3b^2 + 1).
\end{aligned}$$

Esimerkki 80. Vuotuisen viljasadon tutkimiseksi koko maan viljelty ala on jaettu suureen määrään keskenään samanlaisiksi (esim. pinta-alan, sääolosuhteiden ja maaperän suhteen) oletettuja alueita. Oletamme, että viljasato on satunnaismuuttuja X, jolla on odotusarvo μ ja varianssi σ^2. Arvot μ ja σ^2 ovat tuntemattomia, tehtävänä on arvioida μ:ta otoksen avulla. Kuinka suuri on otoksen oltava, jotta otoskeskiarvon poikkeama μ:sta olisi enintään 0.01 (yksikkönä tonni) todennäköisyydellä $1-\alpha$, kun aikaisempien tilastojen perusteella voidaan arvioida $\sigma \leqslant 0.2$.

Ratkaisu. Merkitsemme otoskeskiavoa $\overline{X}_n$:llä, jolloin Lauseen 3.5.1 todistuksessa esiintyneen arvion perusteella ($\varepsilon = 0.01$)

$$P(|\overline{X}_n - \mu| < 0.01) \geqslant 1 - \frac{\sigma^2}{n \cdot 0.01^2} \geqslant 1 - \frac{0.2^2}{n \cdot 0.01^2} = 1 - \frac{400}{n} \geqslant 1 - \alpha,$$

joka on yhtäpitävää sen kanssa, että $n \geqslant \frac{400}{\alpha}$. Kiinnittämällä eri *varmuustasoja* saamme esimerkiksi

$$\alpha = 0.05,\ n \geqslant 8\,000; \quad \alpha = 0.01,\ n \geqslant 40\,000; \quad \alpha = 0.001,\ n \geqslant 400\,000.$$

Esimerkki 81. *Tuominen 3:25.* Sm X saa arvot 1, 2, 3, kunkin tn:llä $\frac{1}{3}$. a) Määritä X:n tngf, b) laske EX, EX^2 ja EX^3.

Ratkaisu. a) Koska $p_k = P(X = k) = \frac{1}{3}$ kaikilla $k = 1, 2, 3$, niin

$$G(t) = \sum_{k=1}^{3} p_k\, t^k = \frac{1}{3}(t + t^2 + t^3).$$

b)

$$G'(t) = \frac{1}{3}(1 + 2t + 3t^2), \text{ joten } EX = G'(1) = \frac{1}{3} + \frac{2}{3} + \frac{3}{3} = 2.$$

Koska

$$G''(t) = \frac{1}{3}(2 + 6t),$$

niin

$$G''(1) = \frac{2}{3} + 2 = \frac{8}{3} = E(X(X-1)) = E(X^2 - X) = EX^2 - EX = EX^2 - 2,$$

ja saadaan, että

$$EX^2 = \frac{8}{3} + 2 = \frac{14}{3} = 4\frac{2}{3} \approx 4.67.$$

EX^3:n laskemiseksi toteamme ensin, että

$$G'''(t) = \frac{1}{3} \cdot 6 = 2, \text{ joten } G'''(1) = 2.$$

Derivoimalla vielä kerran Lauseen 3.6.4 (ii) todistuksessa saadaan

$$\begin{aligned} G'''(1) &= E(X(X-1)(X-2)) = E(X^3 - 3X^2 + 2X) = EX^3 - 3EX^2 + 2EX \\ &= EX^3 - 3 \cdot \frac{14}{3} + 2 \cdot 2 = EX^3 - 10 = 2, \end{aligned}$$

joten

$$EX^3 = 10 + 2 = 12.$$

Esimerkki 82. *Tuominen 3:31.* Kioskiin saapuvien asiakkaiden lkm muodostaa Poisson-prosessin, intensiteettinä 10 (yksikkönä tunti). Oletamme, että 15 minuutin aikana on saapunut 2 asiakasta. Mikä on tällä ehdolla tn, että

a) molemmat saapuivat ensimmäisen 5 minuutin aikana,
b) ainakin toinen saapui viimeisen 5 minuutin aikana?

Ratkaisu. Miksi voimme olettaa, että kioskiin saapuvien asiakkaiden lukumäärä muodostaa Poisson-prosessin?

- Asiakkaat saapuvat kioskiin satunnaisina ajankohtina.
- Asiakkaiden saapumisajankohdat ovat toisistaan riippumattomia.
- Asiakkaita saapuu keskimäärin yhtä paljon yhtä pitkillä aikaväleillä.

Pätevätkö nämä ehdot mielestäsi riittävän hyvin kioskeihin yleensä?

Olkoon X saapuvien asiakkaiden muodostama Poisson-prosessi, $X(\Delta) \sim \text{Poisson}(\lambda t)$, missä $\lambda = \frac{1}{6}$ (aika t minuuteissa). Tällöin siis $X(\Delta_1) \perp\!\!\!\perp X(\Delta_2)$, kun aikavälit Δ_1 ja Δ_2 ovat erillisiä.

a) Kysytty ehdollinen tn on riippumattomuusoletuksen perusteella

$$\begin{aligned}
&P(X([0,5]) = 2 \,|\, X([0,15]) = 2) \\
&= \frac{P((X([0,5]) = 2) \cap (X([0,15]) = 2))}{P(X([0,15]) = 2)} \\
&= \frac{P((X([0,5]) = 2) \cap (X(]5,15]) = 0))}{P(X([0,15]) = 2)} \\
&= \frac{P(X([0,5]) = 2) \cdot P(X(]5,15]) = 0)}{P(X([0,15]) = 2)} \\
&= \frac{e^{-\frac{5}{6}} \cdot \frac{\left(\frac{5}{6}\right)^2}{2} \cdot e^{-\frac{10}{6}}}{e^{-\frac{15}{6}} \cdot \frac{\left(\frac{15}{6}\right)^2}{2}} = \frac{1}{9} \cdot e^0 = \frac{1}{9} \approx 0.111.
\end{aligned}$$

b) Vastaavasti saadaan ehdollinen tn

$$\begin{aligned}
&P(X(]10,15]) \geqslant 1 \,|\, X([0,15]) = 2) \\
&= \frac{P((X(]10,15]) \geqslant 1) \cap (X([0,15]) = 2))}{P(X([0,15]) = 2)} \\
&= \frac{P((X([0,10]) = 0) \cap (X(]10,15]) = 2)}{P(X([0,15]) = 2)} \\
&\quad + \frac{P((X([0,10]) = 1) \cap (X(]10,15]) = 1)}{P(X([0,15]) = 2)} \\
&= \frac{P(X([0,10]) = 0) \cdot P(X(]10,15]) = 2)}{P(X([0,15]) = 2)} \\
&\quad + \frac{P(X([0,10]) = 1) \cdot P(X(]10,15]) = 1)}{P(X([0,15]) = 2)} \\
&= \frac{e^{-\frac{10}{6}} \cdot e^{-\frac{5}{6}} \cdot \frac{\left(\frac{5}{6}\right)^2}{2} + e^{-\frac{10}{6}} \cdot \frac{10}{6} \cdot e^{-\frac{5}{6}} \cdot \frac{5}{6}}{e^{-\frac{15}{6}} \cdot \frac{\left(\frac{15}{6}\right)^2}{2}} = \frac{5}{9} \approx 0.556.
\end{aligned}$$

Esimerkki 83. Pariston kestoajan odotusarvo on $\frac{1}{2}$ ja hajonta $\frac{1}{2}$ (yksikkönä vuosi). Pariston kuluttua loppuun se vaihdetaan heti uuteen; eri paristot oletetaan riippumattomiksi. Kuinka monta paristoa vähintään on oltava, jotta ne riittäisivät neljäksi vuodeksi vähintään todennäköisyydellä 0.95

a) ilman jakaumaoletusta,

b) olettaen kestoaika eksponenttijakautuneeksi?

Ratkaisu. Olkoon paristojen lukumäärä n ja

$$X_i = \text{"}i.\text{ pariston kestoaika"} \quad i = 1, \ldots, n,$$

jolloin $EX_i = \frac{1}{2}$, $D^2X_i = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Ehto on nyt lausuttavissa muodossa

$$P(S_n > 4) \geqslant 0.95,$$

missä $S_n = X_1 + \ldots + X_n$ ja $ES_n = \frac{n}{2}$ sekä $D^2S_n = \frac{n}{4}$.

a) Normaaliapproksimaatiota käyttäen

$$\begin{aligned} P(S_n > 4) &= P\left(\frac{S_n - \frac{1}{2}}{\frac{\sqrt{n}}{2}} > \frac{4 - \frac{n}{2}}{\frac{\sqrt{n}}{2}}\right) \approx 1 - \Phi\left(\frac{4 - \frac{n}{2}}{\frac{\sqrt{n}}{2}}\right) \\ &= \Phi\left(\frac{\frac{n}{2} - 4}{\frac{\sqrt{n}}{2}}\right) \geqslant 0.95 \\ \Leftrightarrow \quad & \frac{\frac{n}{2} - 4}{\frac{\sqrt{n}}{2}} \geqslant 1.645. \end{aligned}$$

Tästä saadaan ratkaisuksi $n = 15$.

b) Jos $X_i \sim \text{Exp}(2)$, $i = 1, \ldots, n$, niin Lauseen 3.7.6 mukaan $S_n \sim \text{Gamma}(n, 2)$, sillä paristojen kestoajat oletettiin riippumattomiksi. Tällöin Lauseen 3.7.4 (iii) mukaan $4S_n \sim \text{Gamma}(n, \frac{1}{2})$, joten Huomautuksen 3.7.8 mukaan $4S_n \sim \chi^2_{2n}$. Taulukosta näemme, että

$$P(4S_n > 16) \geqslant 0.95 \Leftrightarrow n \geqslant 14.$$

Esimerkki 84. Kuvatkoon X kruunien lukumäärää heitettäessä symmetristä rahaa 10 kertaa. Pyrimme määräämään normaaliapproksimaation avulla tapahtuman $(X > 7)$ todennäköisyydelle likiarvon. Koska X:n arvojoukko on $\{0, 1, \ldots, 10\}$, tapahtuma $(X > 7)$ on lausuttavissa seuraavissa ekvivalenteissa muodoissa

$$\begin{aligned}(X > 7) &= (X \geqslant 8) = (7 < X \leqslant 10) = (8 \leqslant X <\leqslant 10)\\ &= (7 < X < 11) = (8 \leqslant X < 11).\end{aligned}$$

Koska $X \sim \mathrm{Bin}(n, p) = \mathrm{Bin}(10, \frac{1}{2})$, niin se on likimain jakutunut

$$N(np, npq) = N(10 \cdot \tfrac{1}{2}, 10 \cdot \tfrac{1}{2} \cdot \tfrac{1}{2}) = N(5, \tfrac{5}{2}).$$

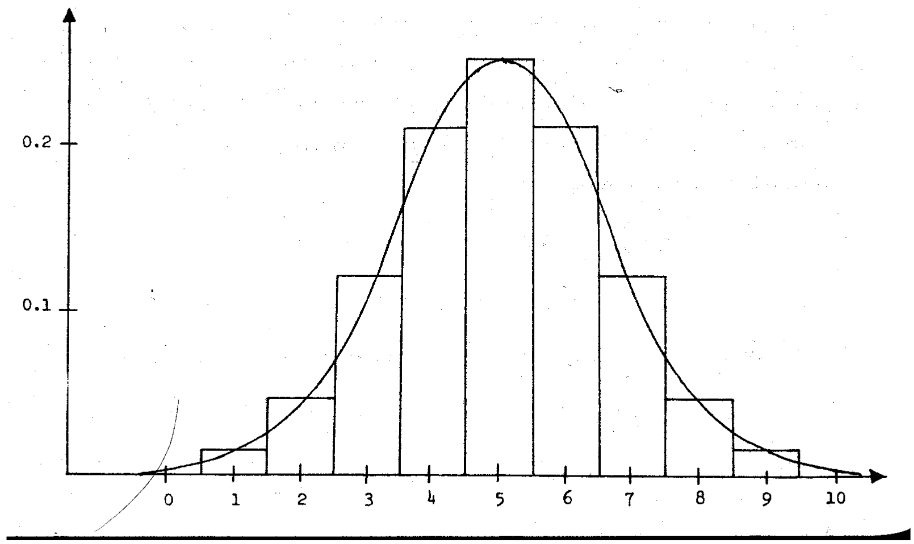
Näin ollen saamme kaikille em. tapahtumille eri likiarvon:

$$\begin{aligned}
P(X > 7) &= 1 - P(X \leqslant 7) = 1 - P\left(\frac{X - np}{\sqrt{npq}} \leqslant \frac{7 - np}{\sqrt{npq}}\right)\\
&\approx 1 - \Phi\left(\frac{7-5}{\sqrt{2.5}}\right) \approx 1 - \Phi(1.26) \approx 0.1038,\\
P(X \geqslant 8) &= 1 - P(X < 8) = 1 - P\left(\frac{X-5}{\sqrt{2.5}} < \frac{8-5}{\sqrt{2.5}}\right)\\
&\approx 1 - \Phi(1.90) \approx 0.0287,\\
P(7 < X \leqslant 10) &= P\left(\frac{7-5}{\sqrt{2.5}} < \frac{X-5}{\sqrt{2.5}} \leqslant \frac{10-5}{\sqrt{2.5}}\right)\\
&\approx \Phi\left(\frac{5}{\sqrt{2.5}}\right) - \Phi\left(\frac{2}{\sqrt{2.5}}\right) \approx \Phi(3.16) - \Phi(1.26)\\
&\approx 0.1030,\\
P(8 \leqslant X <\leqslant 10) &= \ldots \approx \Phi(3.16) - \Phi(1.90) \approx 0.0280,\\
P(7 < X < 11) &= \ldots \approx \Phi(3.80) - \Phi(1.26) \approx 0.1038,\\
P(8 \leqslant X < 11) &= \ldots \approx \Phi(3.80) - \Phi(1.90) \approx 0.0286.
\end{aligned}$$

Vertailun vuoksi voimme laskea myös tarkan arvon

$$\begin{aligned}P(X > 7) &= P(X = 8) + P(X = 9) + P(X = 10)\\ &= \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right) \cdot \left(\frac{1}{2}\right)^{10} = \frac{7}{128} \approx 0.0547.\end{aligned}$$

Vaikeus johtuu siitä, että diskreettiä satunnaismuuttujaa on pyritty approksimoimaan jatkuvalla jakaumalla. Havainnollistamme tilanteen seuraavalla kuviolla:



Todennäköisyyden $P(k_1 \leqslant X \leqslant k_2)$ tarkka arvo geometrisesti merkitsee vastaavien suorakulmioiden yhteenlaskettua pinta-alaa, kun taas approksimaatio merkitsee tiheysfunktion rajoittaman kuvion pinta-alaa. Paras yhteensopivuus saadaan yleensä korvaamalla väli $[k_1, k_2]$ välillä $[k_1 - \frac{1}{2}, k_2 + \frac{1}{2}]$ eli tekemällä ns. *jatkuvuuskorjaus*:

$$P(k_1 \leqslant X \leqslant k_2) \approx \int_{k_1-\frac{1}{2}}^{k_2+\frac{1}{2}} f(x)\,dx = F(k_2 + \tfrac{1}{2}) - F(k_1 - \tfrac{1}{2}),$$

missä f ja F ovat approksimoiva tiheysfunktio ja kertymäfunktio. Binomijakauman tapauksessa saamme siis seuraavan kaavan (jota ei ole tarkoitettu opeteltavaksi ulkoa): Jos $X \sim \mathrm{Bin}(n, p)$, niin

$$P(k_1 \leqslant X \leqslant k_2) \approx \Phi\left(\frac{k_2 + \frac{1}{2} - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - \frac{1}{2} - np}{\sqrt{npq}}\right).$$

Ratkaisemme alkuperäisen tehtävämme nyt jatkuvuuskorjauksella: Kirjoitamme

$$(X > 7) = (8 \leqslant X \leqslant 10),$$

teemme jatkuvuuskorjauksen

$$(X > 7) = (7.5 \leqslant X \leqslant 10.5),$$

ja käytämme normaaliapproksimaatiota

$$\begin{aligned} P(X > 7) = P(7.5 \leqslant X \leqslant 10.5) &= P\left(\frac{7.5-5}{\sqrt{2.5}} \leqslant \frac{X-5}{\sqrt{2.5}} \leqslant \frac{10.5-5}{\sqrt{2.5}}\right) \\ &\approx \Phi\left(\frac{5.5}{\sqrt{2.5}}\right) - \Phi\left(\frac{2.5}{\sqrt{2.5}}\right) \approx \Phi(3.48) - \Phi(1.58) \approx 0.0568. \end{aligned}$$

Huomaamme, että saimme näinkin pienellä n:n arvolla ($n = 10$) varsin tyydyttävän likiarvon.

Esimerkki 85. Olkoon funktio $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} x^2 + \frac{xy}{3}, & 0 < x < 1,\ 0 < y < 2, \\ 0, & \text{muulloin.} \end{cases}$$

a) Osoita, että f määrittelee jatkuvan jakauman tiheysfunktion.
b) Oletetaan, että f on satunnaismuuttujaparin (X, Y) tiheysfunktio. Määritä $P(X + Y \leqslant 1)$.

Ratkaisu. a) On osoitettava (*Tuominen*, Huomautus 5.3.2), että

(i) $f(x,y) \geqslant 0$ kaikilla $(x,y) \in \mathbb{R}$,
(ii) f on integroituva yli $\mathbb{R}^2$:n ja

$$\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1.$$

Ehto (i) ja f:n integroituvuus ovat selviä. Lisäksi

$$\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = \int_0^1 \int_0^2 \left(x^2 + \frac{xy}{3}\right) dx\,dy = \int_0^1 \left(\Big/_{\!0}^{2} x^2 y + \frac{xy^2}{6} \right) dx$$

$$= \int_0^1 \left(2x^2 + \frac{2}{3}x\right) dx = \Big/_{\!0}^{1} \left(\frac{2}{3}x^3 + \frac{1}{3}x\right) = \frac{2}{3} + \frac{1}{3} = 1. \qquad \square$$

b) Nyt saadaan

$$P(X + Y \leqslant 1) = P(Y \leqslant 1 - X) = \int_0^1 \int_0^{1-x} \left(x^2 + \frac{xy}{3}\right) dx\,dy$$

$$= \int_0^1 \left(\Big/_{\!0}^{1-x} x^2 y + \frac{xy^2}{6} \right) dx = \int_0^1 \left(x^2(1-x) + \frac{x(1-x)^2}{6}\right) dx$$

$$= \int_0^1 \left(-\frac{5}{6}x^3 + \frac{2}{3}x^2 + \frac{1}{6}x\right) dx = \Big/_{\!0}^{1} \left(-\frac{5}{24}x^4 + \frac{2}{9}x^3 + \frac{1}{12}x^2\right)$$

$$= -\frac{5}{24} + \frac{2}{9} + \frac{1}{12} = -\frac{15}{72} + \frac{16}{72} + \frac{6}{72} = \frac{7}{72} \approx 0.097.$$

Esimerkki 86. Olkoon $X \sim \text{Gamma}(r, \lambda)$ ja olkoon $Y \sim \text{Exp}(x)$ ehdolla $(X = x)$. Tiedämme siis, että

$$f_X(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, \quad x > 0,$$

ja kun $x > 0$, niin

$$f_Y(y \mid X = x) = xe^{-xy}, \quad y > 0.$$

Näistä saamme (X, Y):n tiheysfunktioksi

$$\begin{aligned} f(x, y) &= f_X(x) \cdot f_Y(y \mid X = x) \\ &= \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^r e^{-(\lambda+y)x}, & x > 0 \text{ ja } y > 0, \\ 0, & \text{muulloin.} \end{cases} \end{aligned}$$

Nyt voimme johtaa Y:n reunatiheysfunktion

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x, y) = \int_0^{\infty} \frac{\lambda^r}{\Gamma(r)} x^r e^{-(\lambda+y)x} \, dx \\ &= \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(r+1)}{(\lambda+y)^{r+1}} = \frac{r\lambda^r}{(\lambda+y)^{r+1}}, \quad y > 0. \end{aligned}$$

Edelleen voimme johtaa X:n ehdollisen jakauman ehdolla $(Y = y)$, kun $y > 0$,

$$f_X(x \mid Y = y) = \frac{f(x, y)}{f_Y(y)} = \frac{(\lambda+y)^{r+1}}{\Gamma(r+1)} x^r e^{-(\lambda+y)x}, \quad x > 0.$$

Siis ehdollinen jakauma on $\text{Gamma}(r+1, \lambda+y)$. Ehdolliset odotusarvot ovat (gamma- ja eksponenttijakauman ominaisuudet!)

$$E(X \mid Y = y) = \frac{r+1}{\lambda+y}, \quad x > 0$$

ja

$$E(Y \mid X = x) = \frac{1}{x}, \quad y > 0.$$

Lisäoletuksella $r > 1$ saamme Y:n odotusarvon

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} f_X(x) \, E(Y \mid X = x) \, dx = \int_0^{\infty} \frac{1}{x} \cdot \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \, dx \\ &= \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(r-1)}{\lambda^{r-1}} = \frac{\lambda}{r-1}. \end{aligned}$$

Esimerkki 87. Olkoon satunnaismuuttujaparin (X, Y) jakauma tasainen $\mathbb{R}^2$:n avoimessa yksikkökiekossa

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}.$$

Tällöin $m(A) = \pi \cdot 1^2 = \pi$, joten

$$f(x, y) = \begin{cases} \frac{1}{\pi}, & \text{kun } (x, y) \in A, \\ 0, & \text{muulloin.} \end{cases}$$

Johdamme satunnaismuuttujien reunatiheysfunktiot:

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\, dy \Big/_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{y}{\pi} = \frac{2\sqrt{1-x^2}}{\pi}, \quad -1 < x < 1,$$

ja vastaavasti

$$f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \frac{1}{\pi}\, dx = \frac{2\sqrt{1-y^2}}{\pi}, \quad -1 < y < 1.$$

Huomaamme, että satunnaismuuttujat X ja Y eivät ole riippumattomia, sillä

$$f(x, y) \neq f_X(x)\, f_Y(y).$$

Ehdolliset tiheysfunktiot: Kun $-1 < y < 1$, niin

$$f_X(x \mid Y = y) = \frac{1}{2\sqrt{1-y^2}}, \quad -\sqrt{1-y^2} < x < \sqrt{1-y^2},$$

ja kun $-1 < x < 1$, niin

$$f_Y(y \mid X = x) = \frac{1}{2\sqrt{1-x^2}}, \quad -\sqrt{1-x^2} < y < \sqrt{1-x^2}.$$

Ehdolliset jakaumat ovat siis tasaisia, mikä ei liene yllätys!

Totesimme, että X ja Y eivät ole riippumattomia, mutta sen sijaan ne ovat korreloimattomia:

$$EX = \frac{1}{\pi} \iint_A x\, dx\, dy = 0 = \frac{1}{\pi} \iint_A y\, dx\, dy = EY$$

ja

$$E(XY) = \frac{1}{\pi} \iint_A xy\, dx\, dy = 0,$$

joten $\mathrm{Cov}(X, Y) = E(XY) - EX\, EY = 0$ ja näin ollen

$$\mathrm{Corr}(X, Y) = \frac{\mathrm{Cov}(X, Y)}{DX\, DY} = 0.$$