4 Kertaus: Tilastot ja todennäköisyys

4.1 Kurssin keskeiset asiat

1.

Viimeisen muuttujan arvon 4 summafrekvenssi on **25**, joten havaintoyksiköiden lukumäärä on 25. Lasketaan puuttuvat frekvenssit taulukkoon:

- summafrekvenssi (*sf*) saadaan summaamalla siihenastiset frekvenssit
- suhteellinen frekvenssi (*f* %) saadaan vertaamalla muuttujan frekvenssiä havaintoyksiköiden kokonaismäärään, joka on 25, ja ilmoittamalla osamäärä prosentteina
- suhteellinen summafrekvenssi (*sf* %) saadaan summaamalla siihenastiset suhteelliset frekvenssit (tai vertaamalla muuttujan summafrekvenssiä havaintoyksiköiden kokonaismäärään)

x	f	$f\,\%$	sf	$sf\,\%$
1	**5**	$\frac{5}{25} \cdot 100 = 20$	5	20
2	$15 - 5 = 10$	$\frac{10}{25} \cdot 100 = 40$	**15**	$20 + 40 = 60$
3	$18 - 15 = 3$	$\frac{3}{25} \cdot 100 = 12$	$25 - 7 = 18$	$60 + 12 = 72$
4	**7**	$\frac{7}{25} \cdot 100 = 28$	**25**	$72 + 28 = 100$

2.

a) ”Ei kertaakaan” viittaa muuttujan arvoon nolla, jonka frekvenssi on 8. Kyselyyn osallistuneista kahdeksan ei käynyt elokuvissa edellisen kuukauden aikana kertaakaan.

b) ”Korkeintaan kaksi kertaa” viittaa muuttujan arvoihin 0–2. Korkeintaan kaksi kertaa elokuvissa käyneiden prosenttiosuus nähdään muuttujan arvon 2 suhteellisesta summafrekvenssistä, joka on 78 %.

c) ”Ainakin kolme kertaa” viittaa muuttujan arvoihin 3–5, joiden suhteellisten frekvenssien summa on

$$6\ \% + 12\ \% + 4\ \% = 22\ \%.$$

Ainakin kolme kertaa elokuvissa käyneitä oli siis 22 %.

Tämä voidaan laskea myös kertymien avulla. Kaikkien kyselyyn osallistuneiden määrä prosentteina on 100 %. Korkeintaan kaksi kertaa elokuvissa käyneitä oli 78 %, joten ainakin kolme kertaa käyneitä oli $100\ \% - 78\ \% = 22\ \%$.

d) Moodi on se muuttujan arvo, jonka frekvenssi on suurin, eli Mo = 1 (kerta).
Mediaani on se muuttujan arvo, jonka suhteellinen summafrekvenssi ensimmäisen kerran ylittää arvon 50 %, eli Md = 1 (kerta).

3.

a) Muuttujan arvon 170 cm suhteellinen summafrekvenssi on kuvaajan perusteella 75 %, joten 75 % opiskelijoista on pituudeltaan korkeintaan 170 cm.

b) Muuttujan arvon 160 cm suhteellinen summafrekvenssi on kuvaajan perusteella 50 %, joten opiskelijoiden mediaanipituus on 160 cm.

c) Muuttujan arvon 155 cm suhteellinen summafrekvenssi on kuvaajan perusteella 25 %, joten 25 % opiskelijoista on pituudeltaan korkeintaan 155 cm. Opiskelijoita on yhteensä 220, joten korkeintaan 155 senttisiä opiskelijoita on

$$0{,}25 \cdot 220 = 55 \text{ (opiskelijaa).}$$

4.

a) Moodi on se muuttujan arvo, jonka frekvenssi on suurin.
Sadepäivien lukumäärä 4 esiintyy yleisimmin, joten sadepäivien moodi on 4 päivää.

b) Asetetaan sadepäivien lukumäärät suuruusjärjestykseen:

2 2 **3** **4** 4 4

Mediaani on suuruusjärjestykseen asetetun aineiston keskimmäinen luku. Havaintoyksiköitä on nyt yhteensä kuusi, eli parillinen määrä, joten keskimmäistä lukua ei ole. Tällöin mediaani on kahden keskimmäisen luvun keskiarvo:

$$\mathrm{Md} = \frac{3+4}{2} = \frac{7}{2} = 3{,}5.$$

Sadepäivien mediaani on siis 3,5 päivää.

c) Lasketaan sadepäivien keskiarvo:

$$\bar{x} = \frac{4+3+4+2+4+2}{6} = \frac{19}{6} = 3{,}16\ldots \approx 3{,}2.$$

Sadepäivien keskiarvo on siis 3,2 päivää.

5.

a) Tutkimukseen osallistuneiden lukumäärä saadaan summaamalla frekvenssit:

$$15 + 55 + 45 + 30 + 5 = 150$$

Tutkimukseen osallistui 150 lukiolaista.

b) Moodi on se muuttujan arvo, jonka frekvenssi on suurin, eli Mo = 1 (kpl/päivä).

Mediaanin määrittämiseksi täydennetään frekvenssitaulukkoon summafrekvenssisarake:

Uutissivustojen lkm	***f***	***sf***
0	15	15
1	55	70
2	45	115
3	30	145
4	5	150

Mediaani on se muuttujan arvo, jonka alapuolelle jää puolet (50 %) havaintoyksiköistä. Havaintoyksiköitä on yhteensä 150 ja näistä puolet on $\frac{150}{2} = 75$ kappaletta.

Muuttujan arvon 2 summafrekvenssi on ensimmäinen, joka ylittää arvon 75, joten Md = 2 (kpl/päivä).

c) Lasketaan seurattujen uutissivustojen keskiarvo:

$$\bar{x} = \frac{15 \cdot 0 + 55 \cdot 1 + 45 \cdot 2 + 30 \cdot 3 + 5 \cdot 4}{150} = \frac{255}{150} = 1{,}7.$$

Seurattujen uutissivustojen keskiarvo on siis 1,7 (kpl/päivä).

6.

a) ”Kaksi tietokonetta” vastaa muuttujan arvoa 2, jonka frekvenssi on 60. Kaksi tietokonetta oli siis 60 kotitaloudessa.

b) ”Vähintään kolme” viittaa muuttujan arvoihin 3–5. Lasketaan frekvenssien summa:

$$120 + 40 + 10 = 170$$

Vähintään kolme tietokonetta oli 170 kotitaloudessa.

c) Lasketaan kaikkien muuttujan arvojen frekvenssien summa:

$$10 + 20 + 60 + 120 + 40 + 10 = 260$$

Tutkimukseen osallistui 260 kotitaloutta.

d) Moodi on se muuttujan arvo, jonka frekvenssi on suurin, eli Mo = 3 (tietokonetta).

Lasketaan tietokoneiden lukumäärän keskiarvo:

$$\bar{x} = \frac{10 \cdot 0 + 20 \cdot 1 + 60 \cdot 2 + 120 \cdot 3 + 40 \cdot 4 + 10 \cdot 5}{260} = \frac{710}{260}$$

$$= 2{,}73 \ldots \approx 2{,}7$$

Tutkimukseen osallistuneissa kotitalouksissa oli keskimäärin 2,7 tietokonetta.

7.

Merkitään Maxin viimeisellä viidennellä heitolla saamaa pistemäärää kirjaimella *x*.

Kaikkien heittojen pistemäärien summa on tällöin
$8 + 5 + 9 + 8 + x = x + 30$ ja pistemäärien keskiarvo on $\frac{x+30}{5}$.

Tiedetään, että keskiarvo on 7 joten saadaan yhtälö:

$$\frac{x + 30}{5} = 7 \quad | \cdot 5$$

$$x + 30 = 35$$

$$x = 35 - 30 = 5.$$

Max heitti viimeisellä tikalla tuloksen 5.

8.

a) Sophien päätelmä on oikea.

Perustelu: Arvosanojen mediaani on suuruusjärjestykseen asetettujen arvosanojen keskimmäinen luku. Sophien sai toisesta jäljellä olevista välikokeesta arvosanan $x < 7$ ja toisesta arvosanan $y > 7$. Jälkimmäinen arvosana voi olla $y = 8$ tai $y > 8$, joten suuruusjärjestykseen asetetut arvosanat ovat

x 7 **7** **8** 8 9 tai x 7 **7** **8** y 9 tai x 7 **7** **8** 9 y

Kaksi keskimmäistä lukua ovat kaikissa tapauksissa 7 ja 8, joten arvosanojen mediaani on $\frac{7+8}{2} = \frac{15}{2} = 7{,}5$ eli yli seitsemän.

b) Merkitään Sophien kahden viimeisen välikokeen arvosanaa kirjaimella *x.*

Kaikkien kurssien arvosanojen summa on tällöin $7 + 7 + 8 + 9 + x + x = 2x + 31$ ja arvosanojen keskiarvo on $\frac{2x+31}{6}$.

Tiedetään, että keskiarvo on 8,5 joten saadaan yhtälö:

$$\frac{2x + 31}{6} = 8{,}5 \quad | \cdot 6$$

$$2x + 31 = 51$$

$$2x = 20$$

$$x = \frac{20}{2} = 10.$$

Kahden viimeisen välikokeen arvosana oli siis 10.

9.

a) Moodi on se muuttuja arvo, jonka frekvenssi on suurin.

$\mathrm{Mo_A} = 7,$

$\mathrm{Mo_B} = 2$ ja $7,$

$\mathrm{Mo_C} = 4,$

joten suurin moodi on muuttujalla A.

b) Päätellään jakaumien keskikohdat, eli muuttujien mediaanit (ne muuttujan arvot, joiden alapuolelle jää puolet havainnoista):

$\mathrm{Md_A} = 7,$

$\mathrm{Md_B} = 3{,}5$

$\mathrm{Md_C} = 4,$

joten suurin mediaani on muuttujalla A. Täten suurin keskiarvokin on muuttujalla A.

c) Keskihajonta kuvaa muuttujan arvojen hajaantumista keskiarvon ympärille. Keskihajonta on suurin silloin, kun suuri osa muuttujan arvoista on kaukana keskiarvosta. Muuttujan B arvoista suurimmat frekvenssit ovat pienimmillä (1 ja 2) sekä suurimmilla (6 ja 7) arvoilla kun taas lähellä keskiarvoa olevilla muuttujan arvoilla on pieni frekvenssi. Täten suurin keskihajonta on muuttujalla B.

10.

a) Mansikkamäärän keskiarvo on $\bar{x} = 25$ (kpl) ja keskihajonta on $s = 3{,}5$ (kpl), joten mansikkamäärän $x = 29$ (kpl) normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{29 - 25}{3{,}5} = 1{,}142 \ldots \approx 1{,}1.$$

b) Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos se on yli kahden keskihajonnan päässä keskiarvosta. Tällöin normitetun arvon itseisarvo on suurempi kuin 2.

Asiakkaan rasiassa oli 29 mansikkaa. Normitetun arvon itseisarvo on $|z| = 1{,}142 \ldots < 2$, joten määrä ei poikkea merkitsevästi keskiarvosta.

c) Lasketaan ne mansikkamäärät, jotka poikkeavat tasan kahden keskihajonnan verran keskiarvosta $\bar{x} = 25$ (kpl).

Keskiarvosta kaksi keskihajontaa alaspäin oleva lukumäärä on:

$$\bar{x} - 2s = 25 - 2 \cdot 3{,}5 = 25 - 7 = 18 \text{ (kpl)}.$$

Keskiarvosta kaksi keskihajontaa ylöspäin oleva lukumäärä on:

$$\bar{x} + 2s = 25 + 2 \cdot 3{,}5 = 25 + 7 = 32 \text{ (kpl)}.$$

Kun rasian mansikkamäärä on välillä 18–32 mansikkaa, niin määrä ei poikkea merkitsevästi keskiarvosta.

11.

Todennäköisyys saadaan laskemalla tapahtumalle suotuisien alkeistapauksien osuus kaikista alkeistapauksista.

Alkeistapauksia, eli taideleirille osallistujia on yhteensä $4 + 5 + 3 = 12$ (osallistujaa).

a) Kuvanveistäjiä on viisi. Kysytty todennäköisyys on

$$P(\text{"kuvanveistäjä"}) = \frac{5}{12}.$$

b) Kuvanveistäjiä ja valokuvaajia on yhteensä $5 + 3 = 8$. Kysytty todennäköisyys on

$$P(\text{"kuvanveistäjä tai valokuvaaja"}) = \frac{8}{12}^{(4} = \frac{2}{3}.$$

12.

a) Juoksijoita on yhteensä $5 + 23 + 16 + 8 + 2 = 54$. Vähintään 2500 metriä juosseiden lukumäärä on $16 + 8 + 2 = 26$ (kappaletta), joten kysytty todennäköisyys on

$$P(\text{"juoksi vähintään 2500 m"}) = \frac{26^{(2}}{54} = \frac{13}{27}.$$

b) Vähintään 2000 metriä juosseita on yhteensä $23 + 16 + 8 + 2 = 49$. Näistä juoksijoista vielä ainakin 1000 metriä juoksevat kuuluvat vähintään 3000 metriä juokseviin, joiden lukumäärä on $8 + 2 = 10$ (juoksijaa).

Kysytty todennäköisyys on:

$$P(\text{"2000 m juossut juoksee vielä ainakin 1000 m"}) = \frac{10}{49}.$$

13.

Valitaan yksi numero 50 päänumeron joukosta. Alkeistapauksia on yhteensä 50.

a) Valittu numero on korkeintaan 10, jos se on välillä 1–10. Suotuisia alkeistapauksia on kymmenen. Kysytty todennäköisyys on

$$P(\text{"numero on korkeintaan 10"}) = \frac{10}{50} = \frac{1}{5}.$$

b) Valittu numero on vähintään 20, jos se on välillä 20–50. Suotuisia alkeistapauksia on $50 - 19 = 31$. Kysytty todennäköisyys on

$$P(\text{"numero on vähintään 20"}) = \frac{31}{50}.$$

c) Välillä 1–50 puolet numeroista on parillisia. Suotuisia alkeistapauksia on 25. Kysytty todennäköisyys on

$$P(\text{"numero on parillinen"}) = \frac{25}{50} = \frac{1}{2}.$$

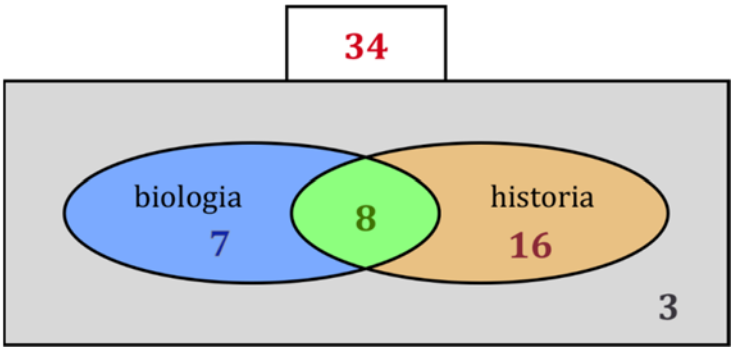
d) Valittu numero on parillinen ja korkeintaan 10, jos se joku numeroista 2, 4, 6, 8 ja 10. Suotuisia alkeistapauksia on viisi. Kysytty todennäköisyys on

$$P(\text{"numero on parillinen ja korkeintaan 10"}) = \frac{5}{50} = \frac{1}{10}.$$

14.

Ratkaistaan tehtävä Venn-diagrammin avulla. Alkeistapauksia eli opiskelijoita on yhteensä 34.

- biologiaa ja historiaa opiskelevia on 8
- pelkkää biologiaa opiskelevia on $15 - 8 = 7$
- pelkkää historiaa opiskelevia on $24 - 8 = 16$
- niitä opiskelijoita, jotka eivät opiskele biologiaa eivätkä historiaa, on $34 - 7 - 8 - 16 = 3$



a) Tapahtumalle ”opiskelee joko biologiaa tai historiaa (tai molempia)” suotuisia alkeistapauksia on yhteensä $7 + 8 + 16 = 31$. Kysytty todennäköisyys on

$$P(\text{"opiskelee biologiaa tai historiaa"}) = \frac{31}{34}.$$

b) Tapahtumalle ”ei opiskele biologiaa eikä historiaa” suotuisia alkeistapauksia on kolme. Kysytty todennäköisyys on

$$P(\text{"ei opiskele biologiaa eikä historiaa"}) = \frac{3}{34}.$$

c) Tarkastellaan historiaa opiskelevia, joten alkeistapauksia on 24. Tapahtumalle ”opiskelee myös biologiaa” suotuisia alkeistapauksia on kahdeksan. Kysytty todennäköisyys on

$$P(\text{"historiaa opiskeleva opiskelee myös biologiaa"}) = \frac{8}{24}^{(8} = \frac{1}{3}.$$

15.

Kahden nopan heitossa alkeistapaukset ovat kahden silmäluvun pareja. Alkeistapauksia on yhteensä $6 \cdot 6 = 36$.

a) Havainnollistetaan alkeistapauksia taulukon avulla. Lasketaan taulukkoon kutakin silmälukuparia vastaava summa.

2. noppa	**6**	7	8	9	10	11	12
	5	6	7	8	9	10	11
	4	5	6	7	8	9	10
	3	4	5	6	7	8	9
	2	3	4	5	6	7	8
	1	2	3	4	5	6	7
		1	**2**	**3**	**4**	**5**	**6**
		1. noppa					

Tapaukselle ”silmälukujen summa on vähintään kahdeksan” suotuisia alkeistapauksia on yhteensä 15. Kysytty todennäköisyys on

$$P(\text{"summa on vähintään 8"}) = \frac{15^{(3}}{36} = \frac{5}{12}.$$

b) Taulukkoon on laskettu kutakin alkeistapausta vastaavan silmälukuparin **summa** ja **tulo**.

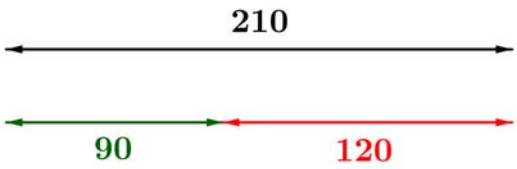
2. noppa 6	7 6	8 12	9 18	10 24	11 30	12 36
5	6 5	7 10	8 15	9 20	10 25	11 30
4	5 4	6 8	7 12	8 16	9 20	10 24
3	4 3	5 6	6 9	7 12	8 15	9 18
2	3 2	4 4	5 6	6 8	7 10	8 12
1	2 1	3 2	4 3	5 4	6 5	7 6
	1	2	3	4	5	6
	1. noppa					

Tapaukselle ”silmälukujen summa on suurempi kuin niiden tulo” suotuisia alkeistapauksia on yhteensä 11. Kysytty todennäköisyys on

$$P(\text{"summa > tulo"}) = \frac{11}{36}.$$

16.

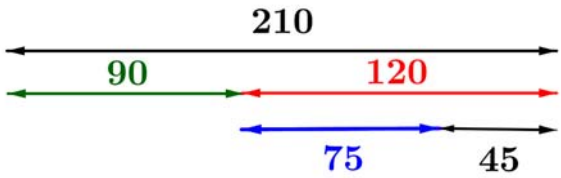
Liikennevalot toimivat 90 s + 120 s = 210 s mittaisissa ajanjaksoissa. Havainnollistetaan yksittäistä 210 sekunnin ajanjaksoa janalla, jonka pituus on 210 yksikköä. Tähän janaan sisältyy 90 yksikön mittainen osuus, jonka liikennevalo palaa vihreällä sekä 120 yksikön mittainen osuus, jonka liikennevalo palaa punaisella.



a) Liikennevaloihin saapuva jalankulkija voi ylittää kadun pysähtymättä, jos hän tulee valoihin niiden 90 sekunnin aikana, jolloin valot ovat vihreällä (oletetaan, ettei jalankulkija lähde ylittämään katua punaisen valon palaessa). Kysytty todennäköisyys on

$$P(\text{"voi ylittää kadun pysähtymättä"}) = \frac{90}{210} \overset{(30}{} = \frac{3}{7}.$$

b) Liikennevaloihin saapuva jalankulkija joutuu odottamaan valojen vaihtumista yli 45 min, jos valot ovat juuri ehtineet vaihtua vihreästä puneiseksi ja seuraavaan valojen vaihtumiseen on yli 45 sekuntia.



Kysytty todennäköisyys on

$$P(\text{"joutuu odottamaan yli 45 s"}) = \frac{120-45}{210} = \frac{75}{210}^{(15} = \frac{5}{14}.$$

17.

a) Ensimmäisessä maljakossa on yhteensä $6 + 4 = 10$ tulppaania, joista kuusi on valkoisia. Todennäköisyys, että ensimmäisestä maljakosta nostettu tulppaani on valkoinen, on

$$P(\text{"1. tulppaani on valkoinen"}) = \frac{6}{10}^{(2} = \frac{3}{5}.$$

Toisessa maljakossa on yhteensä $3 + 7 = 10$ tulppaania, joista kolme on valkoisia. Todennäköisyys, että toisesta maljakosta nostettu tulppaani on valkoinen, on

$$P(\text{"2. tulppaani on valkoinen"}) = \frac{3}{10}.$$

Kertolaskusäännön mukaan

$$P(\text{"1. tulppaani on valkoinen } \textbf{ja} \text{ 2. tulppaani on valkoinen"}) = \frac{3}{5} \cdot \frac{3}{10} = \frac{9}{50}.$$

b) Lasketaan todennäköisyys vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla. Tapahtuman $A =$”ainakin yksi tulppaani on keltainen” vastatapahtuma on $\bar{A} =$”molemmat tulppaanit ovat valkoisia”.

Vastatapahtuman todennäköisyys laskettiin kohdassa a. Kysytty todennäköisyys on

$$
\begin{aligned}
&P(\text{"ainakin yksi tulppani on keltainen"}) \\
&= 1 - P(\text{"molemmat tulppaanit ovat valkoisia"}) \\
&= 1 - \frac{9}{50} \\
&= \frac{41}{50}.
\end{aligned}
$$

18.

a) Esimerkiksi:

- nopan heitto
  $A =$"saadaan 1. heitolla silmäluku 6" ja $B =$"saadaan 2. heitolla silmäluku 6"

- koulumatkan varrella on kahdet liikennevalot, joiden toiminta ei riipu toisistaan
  $A =$"1. valo on punaisella" ja $B =$"2. valo on punaisella"

b) Esimerkiksi:

- kahden kortin nosto korttipakasta, kun ensin nostettua korttia ei palauteta pakkaan
  $A =$"saadaan 1. nostolla ässä" ja $B =$"saadaan 2. nostolla ässä"

- hedelmäkorissa on viisi punaista ja neljä vihreää omenaa, nostetaan yksi omena ja sitten toinen omena palauttamatta ensin nostettua takaisin koriin
  $A =$"ensin nostettu omena on punainen" ja $B =$"toiseksi nostettu omena on vihreä"

19.

Todennäköisyydet ovat

$$P(\text{"sinisilmäinen"}) = \frac{2}{3} \quad \text{ja} \quad P(\text{"ei-sinisilmäinen"}) = 1 - \frac{2}{3} = \frac{1}{3}.$$

Koulun opiskelijoiden lukumäärä on suuri, joten yksittäisen opiskelijan valitseminen ei oleellisesti muuta seuraavan todennäköisyyttä olla sinisilmäinen.

a) Kertolaskusäännön mukaan

$P(\text{"kaikki sinisilmäisiä"})$

$= P(\text{"1. sinisilmäinen } \textbf{ja} \text{ 2. sinisilmäinen } \textbf{ja} \text{ … } \textbf{ja} \text{ 5. sinisilmäinen"})$

$= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \left(\frac{2}{3}\right)^5$

$= 0{,}1316 \ldots \approx 0{,}132.$

Huomaa, että vastaus pyydettiin kolmen desimaalin tarkkuudella.

b) Tapahtuman ”ainakin yksi sinisilmäinen” vastatapahtuma on ”ei yhtään sinisilmäistä”. Vastatapahtuman todennäköisyys on kertolaskusäännön mukaan

$$\begin{aligned} P(\text{"kaikki ei-sinisilmäisiä"}) &= \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \\ &= \left(\frac{1}{3}\right)^5 \\ &= 0{,}00411 \ldots. \end{aligned}$$

Kysytty todennäköisyys on

$$\begin{aligned} &P(\text{"ainakin yksi sinisilmäinen"}) \\ &= 1 - P(\text{"kaikki ei-sinisilmäisiä"}) \\ &= 1 - 0{,}00411 \ldots \\ &= 0{,}9958 \ldots \\ &\approx 0{,}996. \end{aligned}$$

20.

Pelaajia on kentällä yhteensä viisi. Näistä vaihtoon joutuvat kaksi pelaajaa voidaan valita $\binom{5}{2} = 10$ eri tavalla. Alkeistapauksia on yhteensä 10.

a) Tapahtumalle ”Henna ja Maria joutuvat vaihtoon” suotuisia alkeistapahtumia on ainoastaan yksi. Kysytty todennäköisyys on

$$P(\text{"Henna ja Maria vaihtoon"}) = \frac{1}{10}.$$

b) Jos kumpikaan ei joudu vaihtoon, niin vaihtoon joutuvat valitaan $5 - 2 = 3$ pelaajan joukosta. Näistä kaksi voidaan valita $\binom{3}{2} = 3$ eri tavalla. Tapahtumalle ”kumpikaan ei joudu vaihtoon” suotuisia alkeistapahtumia on kolme. Kysytty todennäköisyys on

$$P(\text{"ei kumpikaan vaihtoon"}) = \frac{3}{10}.$$

c) Tapahtuman ”ainakin toinen vaihtoon” vastatapahtuma on ”ei kumpikaan vaihtoon”. Vastatapahtuman todennäköisyys laskettiin kohdassa b. Kysytty todennäköisyys on

$$P(\text{"ainakin toinen vaihtoon"}) = 1 - \frac{3}{10} = \frac{7}{10}.$$

21.

a) Tomaatteja on yhteensä 20. Kolme tomaattia voidaan valita

$$\binom{20}{3} = 1140$$

eri tavalla.

Alkeistapauksia on yhteensä 1140.

Tuoreita tomaatteja on $20 - 6 = 14$. Valitaan tuoreista tomaateista kolme:

$$\binom{14}{3} = 364.$$

Tapahtumalle ”kolme tuoretta tomaattia” suotuisia alkeistapauksia on siis 364. Kysytty todennäköisyys on

$$P(\text{"kolme tuoretta"}) = \frac{364}{1140} = 0{,}319\ldots \approx 0{,}32.$$

b) Valitaan kuudesta pilaantuneesta tomaatista kaksi ja 14 tuoreesta tomaatista yksi. Kysytty todennäköisyys on

$$P(\text{"2 pilaantunutta } \textbf{ja} \text{ 1 tuore"})$$

$$= \frac{\overbrace{\binom{6}{2}}^{\text{2 pilantunutta}} \cdot \overbrace{\binom{14}{1}}^{\text{1 tuore}}}{\binom{20}{3}} = \frac{15 \cdot 14}{1140} = \frac{210}{1140}$$

$$= 0{,}184\ldots \approx 0{,}18.$$

22.

a) Pakassa on yhteensä 52 korttia. Kaksi korttia voidaan valita

$$\binom{52}{2} = 1326$$

eri tavalla.

Alkeistapauksia on siis yhteensä 1326.

Kuvakortteja on yhteensä $4 \cdot 3 = 12$ korttia. Kuvakorteista voidaan valita kaksi

$$\binom{12}{2} = 66$$

eri tavalla.

Tapahtumalle ”kaksi kuvakorttia” suotuisia alkeistapauksia on siis 66. Kysytty todennäköisyys on

$$P(\text{"kaksi kuvakorttia"}) = \frac{66}{1326} = \frac{11}{221} = 0{,}0497\ldots \approx 0{,}050.$$

b) Tapahtuma ”ensimmäinen kortti on pata ja toinen kortti on kuvakortti” sisältää kaksi erillistä tilannetta: ensimmäinen kortti voi olla

- pata ja kuvakortti: 3 vaihtoehtoa **tai**
- pata ja numerokortti: patoja on yhteensä 13, joten numeropatoja on $13 - 3 = 10$ vaihtoehtoa.

Toiselle kortille (kuvakortille) suotuisien korttien lukumäärä pakassa riippuu ensiksi valitusta padasta:

- Jos ensimmäinen kortti on pata ja numerokortti, niin kaikki kuvakortit ovat vielä pakassa ja toiselle kortille suotuisia kortteja on 12 vaihtoehtoa. Kysytty todennäköisyys saadaan kertolaskusäännön mukaan:

$$P(\text{"1. kortti on numeropata } \textbf{ja} \text{ 2. kortti on kuvakortti"})$$
$$= \frac{10}{52} \cdot \frac{12}{51} = \frac{120}{2652}.$$

- Jos ensimmäinen kortti on pata ja kuvakortti, niin toiselle kortille suotuisia kortteja on $12 - 1 = 11$ vaihtoehtoa. Kysytty todennäköisyys saadaan kertolaskusäännön mukaan:

$$P(\text{"1. kortti on kuvapata } \textbf{ja} \text{ 2. kortti on kuvakortti"})$$
$$= \frac{3}{52} \cdot \frac{11}{51} = \frac{33}{2652}.$$

Kokonaistodennäköisyys saadaan yhteenlaskusäännön mukaan:

$P(\text{"ensimmäinen kortti on pata ja toinen kortti on kuvakortti"})$

$$= \frac{120}{2652} + \frac{33}{2652} = \frac{153}{2652}$$

$$= \frac{3}{52} = 0{,}0576 \ldots \approx 0{,}058.$$

23.

$P(\text{"poutaa"}) = 0{,}85$, joten $P(\text{"sataa"}) = 1 - 0{,}85 = 0{,}15$.

a) Kertolaskusäännön mukaan

$$P(\text{"1. päivänä poutaa } \textbf{ja} \text{ 2. päivänä poutaa"}) = 0{,}85 \cdot 0{,}85$$
$$= 0{,}7225 \approx 0{,}72.$$

b) Lasketaan todennäköisyys vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla. Tapahtuman ”ainakin yksi poutapäivä” vastatapahtuma on ”molempina päivinä sataa”.

Vastatapahtuman todennäköisyys on kertolaskusäännön mukaan

$$P(\text{"1. päivänä sataa } \textbf{ja} \text{ 2. päivänä sataa"}) = 0{,}15 \cdot 0{,}1 = 0{,}0225$$

Kysytty todennäköisyys on

$$P(\text{"ainakin yksi poutapäivä"})$$
$$= 1 - P(\text{"molempina päivinä sataa"})$$
$$= 1 - 0{,}0225$$
$$= 0{,}9775 \approx 0{,}98.$$

Todennäköisyys voidaan laskea myös yhteenlaskusäännöllä. Tapahtuma ”ainakin yksi poutapäivä” voi tarkoittaa seuraavia tilanteita:

- 1. päivänä poutaa ja 2. päivänä sataa
- 1. päivänä sataa ja 2. päivänä poutaa
- molempina päivinä poutaa.

Kysytty todennäköisyys on

$$
\begin{aligned}
&P(\text{"ainakin yksi poutapäivä"}) \\
&= P(\text{"1. poutaa } \textbf{ja} \text{ 2. sataa"}) + P(\text{"1. sataa } \textbf{ja} \text{ 2. poutaa"}) \\
&\qquad + P(\text{"1. poutaa } \textbf{ja} \text{ 2. poutaa"}) \\
&= 0{,}85 \cdot 0{,}15 + 0{,}15 \cdot 0{,}85 + 0{,}85 \cdot 0{,}85 \\
&= 0{,}9775 \approx 0{,}98.
\end{aligned}
$$

c) Tapahtuma ”toinen on poutapäivä ja toinen ei” voi tarkoittaa seuraavia tilanteita:

- 1. päivänä poutaa ja 2. päivänä sataa
- 1. päivänä sataa ja 2. päivänä poutaa

Kysytty todennäköisyys on yhteenlaskusäännön mukaan

$$
\begin{aligned}
&P(\text{"ainakin yksi poutapäivä"}) \\
&= P(\text{"1. poutaa } \textbf{ja} \text{ 2. sataa"}) + P(\text{"1. sataa } \textbf{ja} \text{ 2. poutaa"}) \\
&= 0{,}85 \cdot 0{,}15 + 0{,}15 \cdot 0{,}85 \\
&= 0{,}255 \approx 0{,}26.
\end{aligned}
$$

24.

a) Suomi pelaa finaalissa Saksaa vastaan silloin, jos välierissä Saksa voittaa Senegalin ja Suomi voittaa Singaporen. Todennäköisyydet ovat:

Saksa voittaa Senegalin	0,65
Suomi voittaa Singaporen	$1 - 0{,}50 = 0{,}50$

Kertolaskusäännön mukaan

$$P(\text{"Saksa voittaa Senegalin } \textbf{ja} \text{ Suomi voittaa Singaporen"})$$
$$= 0{,}65 \cdot 0{,}50 = 0{,}325.$$

b) Suomen voittotodennäköisyydet yksittäisessä pelissä ovat:

- Suomi voittaa Saksan **0**
- Suomi voittaa Senegalin $1 - 0{,}60 = \mathbf{0{,}40}$
- Suomi voittaa Singaporen $1 - 0{,}50 = \mathbf{0{,}50}$

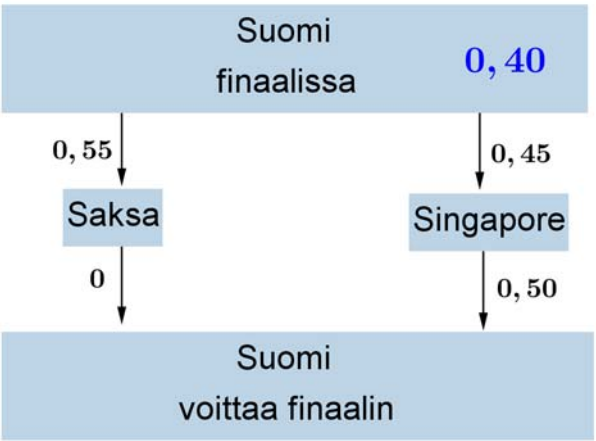
Mahdollisia Suomen välieräpareja on kolme:

Suomen välieräpari:		**Suomen voittotodennäköisyys**
1	Suomi – Saksa	**0**
2	Suomi – Senegal	$\mathbf{0{,}40}$
3	Suomi – Singapore	$\mathbf{0{,}50}$

Lasketaan todennäköisyydet Suomen voitolle koko kilpailussa näissä kolmessa eri tilanteessa:

1 Mikäli Suomen välierävastustaja on Saksa, Suomella ei ole mahdollisuutta päästä finaaliin, joten Suomen todennäköisyys voittaa koko kilpailu on <u>nolla</u>.

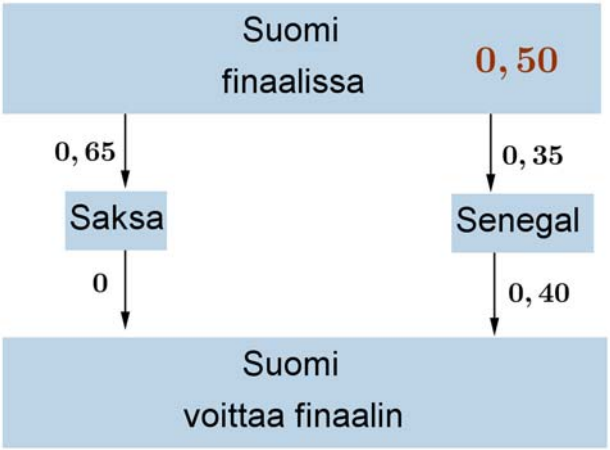
2 Mikäli Suomen välierävastustaja on Senegal, Suomen todennäköisyys päästä finaaliin on **0,40**. Finaalissa vastustajana on joko Saksa (todennäköisyydellä 0,55) tai Singapore (todennäköisyydellä 0,45). Puukaavio havainnollistaa Suomen voittotodennäköisyyksiä näissä kahdessa eri tilanteessa.



Tässä tilanteessa Suomi voittaa todennäköisyydellä

$$\mathbf{0,40} \cdot 0{,}55 \cdot 0 + \mathbf{0,40} \cdot 0{,}45 \cdot 0{,}50 = \underline{0{,}09.}$$

3 Mikäli Suomen välierävastustaja on Singapore, Suomen todennäköisyys päästä finaaliin on **0, 50**. Finaalissa vastustajana on joko Saksa (todennäköisyydellä 0,65) tai Senegal (todennäköisyydellä 0,35). Puukaavio havainnollistaa Suomen voittotodennäköisyyksiä näissä kahdessa eri tilanteessa.



Tässä tilanteessa Suomi voittaa todennäköisyydellä

$$\mathbf{0,50} \cdot 0{,}65 \cdot 0 + \mathbf{0,50} \cdot 0{,}35 \cdot 0{,}40 = \underline{0{,}07.}$$

Kun verrataan todennäköisyyksiä keskenään, voidaan todeta, että Suomen todennäköisyys voittaa koko kilpailu on suurin tilanteessa **2** eli välieräpareilla Suomi – Senegal ja Saksa – Singapore. Todennäköisyys voittoon on tällöin 0,09 = 9 %.

25.

Musiikkitapahtumaan menijöitä on yhteensä $6 + 4 = 10$. Sisälle pääsevät kolme henkilöä voidaan valita $\binom{10}{3} = 120$ eri tavalla.

Tyttöjä on yhteensä neljä. Näistä kolme voidaan valita $\binom{4}{3} = 4$ eri tavalla.

Kysytty todennäköisyys on

$$P(\text{"kolme tyttöä"}) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4}{120} \quad {}^{(4} = \frac{1}{30} = 0{,}0333\ldots \approx 0{,}033.$$

26.

Erilaisia setelijonoja on yhteensä kuusi:

(5,5,10,10) (5,10,5,10) (5,10,10,5) (10,5,5,10)

(10,5,10,5) (10,10,5,5)

Jokainen jono on yhtä mahdollinen: alkeistapauksia on yhteensä kuusi.

Tapahtumalle ”jokaiselle oikea vaihtoraha” suotuisia tapauksia on kaksi: (5,5,10,10) ja (5,10,5,10), joten kysytty todennäköisyys on

$$P(\text{"kaikille oikea vaihtoraha"}) = \frac{2^{(2}}{6} = \frac{1}{3}.$$

27.

Todennäköisyydet ovat seuraavat:

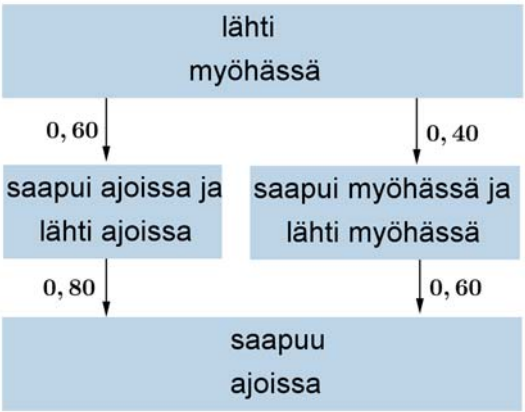
- jos bussi lähtee pysäkiltä myöhässä, niin se saapuu seuraavalle pysäkille
  - ajoissa todennäköisyydellä 0,60
  - myöhässä todennäköisyydellä $1 - 0{,}60 = 0{,}40$.

- jos bussi lähtee pysäkiltä ajoissa, niin se saapuu seuraavalle pysäkille
  - ajoissa todennäköisyydellä 0,80
  - myöhässä todennäköisyydellä $1 - 0{,}80 = 0{,}20$.

- jos bussi saapuu pysäkille ajoissa, se myös lähtee pysäkiltä ajoissa

- jos bussi saapuu pysäkille myöhässä, se myös lähtee pysäkiltä myöhässä

Bussi lähti kaksi pysäkkiä aiemmin myöhässä. Piirretään puukaavio havainnollistamaan eri tilanteita ja todennäköisyyksiä.



Kysytty todennäköisyys on yhteenlaskusäännön mukaan

$$P(\text{"saapuu ajoissa"}) = 0{,}60 \cdot 0{,}80 + 0{,}40 \cdot 0{,}60 = 0{,}72.$$

28.

a) Kolminumeroisia lukuja ovat luvut 100–999. Niitä on yhteensä $999 - 99 = 900$.

Lukumäärä voidaan laskea myös tuloperiaatteella. Kolminumeroisessa luvussa ensimmäinen numero on välillä 1–9 (9 vaihtoehtoa) ja kaksi muuta numeroa välillä 0–9 (10 vaihtoehtoa).

Erilaisia kolmenumeroisia lukuja on tuloperiaatteen mukaan

$$9 \cdot 10 \cdot 10 = 900 \text{ kappaletta.}$$

b) Kolminumeroisia lukuja on yhteensä 900 ja luvuista puolet on parillisia. Parillisia kolminumeroisia lukuja on siis yhteensä $\frac{900}{2} = 450$.

Lukumäärä voidaan laskea myös tuloperiaatteella. Kolminumeroisessa parillisessa luvussa ensimmäinen numero on välillä 1–9 (9 vaihtoehtoa), toinen numeroa välillä 0–9 (10 vaihtoehtoa) ja kolmas numero on joku numeroista 0, 2, 4, 6, 8 (viisi vaihtoehtoa).

Erilaisia kolminumeroisia parillisia lukuja on tuloperiaatteen mukaan

$$9 \cdot 10 \cdot 5 = 450 \text{ kappaletta.}$$

c) Alkeistapauksia eli kolminumeroisia lukuja on yhteensä 900. Tapahtumalle ”luku on 100” suotuisia alkeistapauksia on ainoastaan yksi. Kysytty todennäköisyys on

$$P(\text{"kolminumeroinen luku on luku 100"}) = \frac{1}{900}.$$

29.

a) Romaaneja on yhteensä viisi. Erilaisia viiden romaanin järjestyksiä on tuloperiaatteen mukaan yhteensä $5! = 120$.

b) Kaksi suomalaista dekkaria voidaan järjestää jonon alkuun 2! = 2 eri tavalla. Tämän jälkeen loput $5 - 2 = 3$ romaania voidaan järjestää jonon loppuun 3! = 6 eri tavalla. Tuloperiaatteen mukaan erilaisia järjestyksiä on yhteensä $2 \cdot 6 = 12$.

c) Viidestä kirjasta voidaan valita kaksi $\binom{5}{2} = 10$ eri tavalla. Suomalaisia dekkareita on kaksi, joten tapahtumalle ”valitut kaksi kirjaa ovat suomalaisia dekkareita” suotuisia alkeistapauksia on ainoastaan yksi. Kysytty todennäköisyys on

$$P(\text{"valitaan suomalaiset dekkarit"}) = \frac{1}{10}.$$

Todennäköisyys voidaan laskea myös kertolaskusäännöllä:

$$P(\text{"valitaan suomalaiset dekkarit"})$$

$$= P\begin{pmatrix}\text{"1.valittu kirja on suomalainen dekkari } \textbf{ja} \\ \text{2. valittu kirja on suomalainen dekkari"}\end{pmatrix}$$

$$= \frac{2}{5} \cdot \frac{1}{4} = \frac{2}{20}^{(2} = \frac{1}{10}.$$

4.2 Matemaattisia malleja

30.

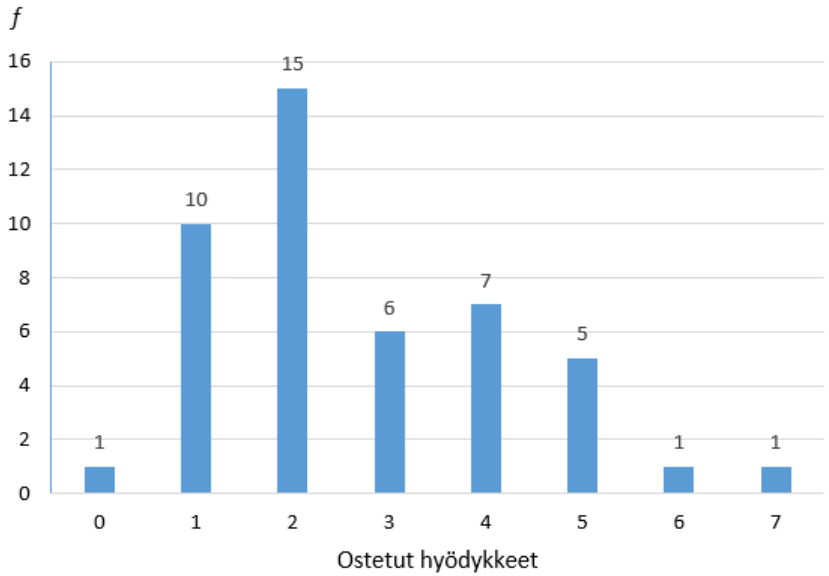
a) Laaditaan frekvenssijakaumat tilasto-ohjelmalla.

	A	B	C	D	E
1	**Tavara, palvelu tai sisältö (kpl)**	**f**	**sf**	**f%**	**sf%**
2	0	1	1	2	2
3	1	10	11	22	24
4	2	15	26	33	57
5	3	6	32	13	70
6	4	7	39	15	85
7	5	5	44	11	96
8	6	1	45	2	98
9	7	1	46	2	100

b) Suurin frekvenssi (*f*) on muuttujan arvolla 2, joten Mo = 2 (kappaletta).

Mediaani on se muuttujan arvo, jonka suhteellinen summafrekvenssi (*sf* %) ensimmäisen kerran ylittää 50 %, joten Md = 2 (kappaletta).

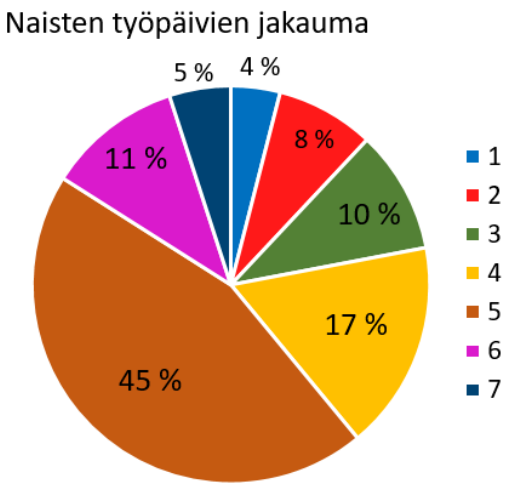
c) Piirretään pylväsdiagrammi tilasto-ohjelmalla. Pylvään korkeus määräytyy frekvenssin mukaan.



31.

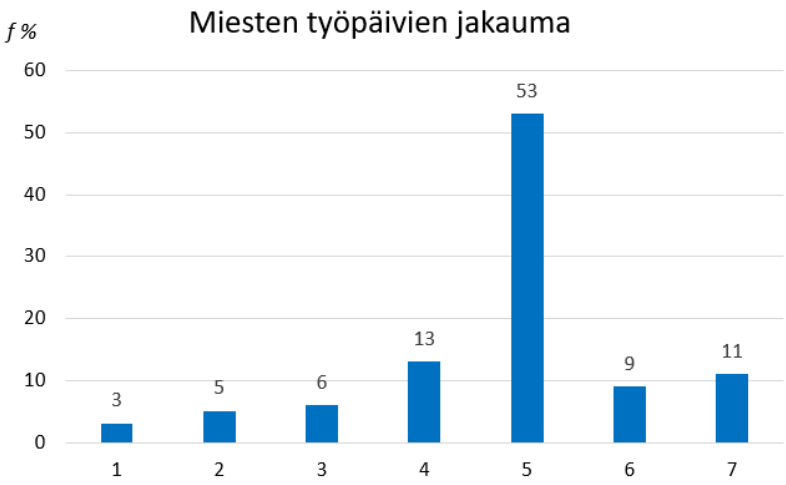
a) Tallennetaan naisten työpäivien prosenttijakauma taulukkolaskentaohjelmaan. Piirretään jakaumasta sektoridiagrammi.

A	B
Lukumäärä	**Naiset (f%)**
1	4
2	8
3	10
4	17
5	45
6	11
7	5



b) Tallennetaan miesten työpäivien prosenttijakauma taulukkolaskentaohjelmaan. Piirretään jakaumasta pylväsdiagrammi. Pylvään korkeus määräytyy prosenttiosuuden mukaan.

A	B
Lukumäärä	**Miehet (f%)**
1	3
2	5
3	6
4	13
5	53
6	9
7	11



c) Moodi on yleisin muuttujan arvo, joten sekä naisten että miesten työpäivien moodi on Mo = 5 (työpäivää viikossa).

Mediaani on se muuttujan arvo, jonka suhteellinen summafrekvenssi ensimmäisen kerran ylittää arvon 50 %. Mediaani voidaan määrittää esimerkiksi täydentämällä taulukkolaskentaohjelman taulukkoon työpäivien suhteelliset summafrekvenssisarakkeet:

A	B	C
Lukumäärä	**Naiset (f%)**	**Naiset (sf%)**
1	4	4
2	8	12
3	10	22
4	17	39
5	45	84
6	11	95
7	5	100

A	B	C
Lukumäärä	**Miehet (f%)**	**Miehet (sf%)**
1	3	3
2	5	8
3	6	14
4	13	27
5	53	80
6	9	89
7	11	100

Tästä nähdään, että sekä naisten että miesten työpäivien mediaani on Md = 5 (työpäivää viikossa).

32.

a) Käytetään ratkaisussa taulukkolaskentaohjelmaa tai laskinohjelmiston taulukkolaskentasovellusta. Aloitetaan laskemalla kunkin luokan luokkakeskus. Luokkakeskus saadaan laskemalla luokan ala- ja ylärajan keskiarvo:

- soluun D2 kirjoitetaan laskukaava $= (B2 + C2)/2$
- kopioidaan kaava sarakkeen D seuraaville riveille
- tallennetaan kunkin luokan frekvenssi sarakkeeseen E

	A	B	C	D	E
1	**Pistemäärä**	**alaraja**	**yläraja**	**luokkakeskus**	**f**
2	25-39	25	39	32	6
3	40-54	40	54	47	25
4	55-69	55	69	62	22
5	70-84	70	84	77	31
6	85-99	85	99	92	13
7	100-114	100	114	107	4

Määritetään tunnusluvut luokkakeskusten ja frekvenssien perusteella tilastotoimintojen avulla: pistemäärän

- moodi on $\text{Mo} = 77$ (pistettä)
- mediaani on $\text{Md} = 62$ (pistettä)
- keskiarvo on $\bar{x} = 66{,}752 \ldots \approx 67$ (pistettä)
- keskihajonta on $s_n = 18{,}752 \ldots \approx 19$ (pistettä)

b) Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos se on yli kahden keskihajonnan päässä keskiarvosta $\bar{x} = 66{,}752\ldots$

Kaisan pistemäärä 88 pistettä on yli keskiarvon. Lasketaan keskiarvosta kaksi keskihajontaa ylöspäin oleva pistemäärä ja verrataan sitä Kaisan pistemäärään. Käytetään laskussa tarkkoja arvoja: tehdään lasku taulukkolaskennan soluviittauksilla tai laskinsovelluksessa kopioimalla tunnuslukujen tarkat arvot laskukenttään.

$$\bar{x} + 2s_n = 66{,}752\ldots + 2 \cdot 18{,}752\ldots = 104{,}258\ldots > 88.$$

Kaisan pistemäärä on siis alle kahden keskihajonnan päässä keskiarvosta, joten se ei poikkea merkitsevästi keskiarvosta.

Tämä voidaan laskea myös normittamalla Kaisan pistemäärä $x = 88$. Pistemäärä poikkeaa merkitsevästi keskiarvosta, jos normitetun arvon itseisarvo on suurempi kuin kaksi.

Normitettu arvo on

$$z = \frac{x - \bar{x}}{s_n} = \frac{88 - 66{,}752\ldots}{18{,}752\ldots} = 1{,}133\ldots$$

Normitetun arvon itseisarvo on $|z| = 1{,}133\ldots < 2$, joten Kaisan pistemäärä ei poikkea merkitsevästi keskiarvosta.

33.

a) Käytetään ratkaisussa taulukkolaskentaohjelmaa. Aloitetaan laskemalla kunkin luokan luokkakeskus:

- soluun D2 kirjoitetaan laskukaava $= (B2 + C2)/2$
- kopioidaan kaava sarakkeen D seuraaville riveille
- muutetaan sarakkeen D solujen pyöristystarkkuus yhden desimaalin tarkkuuteen

	A	B	C	D
1	**Ratkaisut (lkm)**	**alaraja**	**yläraja**	**luokkakeskus**
2	0-9	0	9	4,5
3	10-19	10	19	14,5
4	20-29	20	29	24,5
5	30-39	30	39	34,5
6	40-49	40	49	44,5
7	50-59	50	59	54,5
8	60-69	60	69	64,5
9	70-79	70	79	74,5
10	80-89	80	89	84,5
11	90-99	90	99	94,5
12	100-109	100	109	104,5

Kirjoitetaan kunkin luokan frekvenssi sarakkeeseen E. Lasketaan summafrekvenssit sarakkeeseen F ja suhteelliset summafrekvenssit sarakkeeseen G.

	A	B	C	D	E	F	G
1	**Ratkaisut (lkm)**	**alaraja**	**yläraja**	**luokkakeskus**	**f**	**sf**	**sf%**
2	0-9	0	9	4,5	1	1	2,6
3	10-19	10	19	14,5	0	1	2,6
4	20-29	20	29	24,5	2	3	7,9
5	30-39	30	39	34,5	1	4	10,5
6	40-49	40	49	44,5	1	5	13,2
7	50-59	50	59	54,5	3	8	21,1
8	60-69	60	69	64,5	4	12	31,6
9	70-79	70	79	74,5	7	19	50,0
10	80-89	80	89	84,5	9	28	73,7
11	90-99	90	99	94,5	8	36	94,7
12	100-109	100	109	104,5	2	38	100,0

Moodi on sen luokan luokkakeskus, jonka frekvenssi on suurin: ratkaisujen lukumäärän moodi on Mo = 84,5 (tehtävää).

Mediaani on sen luokan luokkakeskus, jonka suhteellinen summafrekvenssi ensimmäisen kerran ylittää arvon 50 %: ratkaisujen lukumäärän mediaani on Md = 84,5 (tehtävää).

Keskiarvo ja keskihajonta määritetään luokkakeskusten ja frekvenssien perustella tilastotoimintojen avulla: ratkaisujen lukumäärän

- keskiarvo on $\bar{x} = 73{,}710 \ldots \approx 73{,}7$ (tehtävää)
- keskihajonta on $s_n = 22{,}870 \ldots \approx 22{,}9$ (tehtävää)

b) Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos se on yli kahden keskihajonnan päässä keskiarvosta $\bar{x} = 73{,}710\ldots$ Normitetun arvon itseisarvo on tällöin suurempi kuin kaksi.

Normitetaan Eliaksen laskemien tehtävien lukumäärä $x = 107$. Käytetään laskussa tarkkoja arvoja eli tehdään lasku taulukkolaskennan soluviittauksilla tai laskinsovelluksessa kopioimalla tunnuslukujen tarkat arvot laskukenttään.

Normitettu arvo on

$$z = \frac{x - \bar{x}}{s_n} = \frac{107 - 73{,}710\ldots}{22{,}870\ldots} = 1{,}455\ldots$$

Normitetun arvon itseisarvo on $|z| = 1{,}455\ldots < 2$, joten Eliaksen tekemien tehtävien lukumäärä ei poikkea merkitsevästi keskiarvosta.

34.

a) Kirjoitetaan katsojaluvut taulukkolaskentaohjelman tai laskinsovelluksen taulukkoon.

	A	B
1	Jokerit	9173
2	HIFK	8266
3	Kärpät	5821
4	TPS	5534
5	Tappara	5359
6	Ilves	5177

Määritetään tunnusluvut tilastotoimintojen avulla:

- keskiarvo on $\bar{x} = 6555$ (katsojaa)
- keskihajonta on $s_n = 1564{,}818\ldots \approx 1565$ (katsojaa)

b) Lasketaan keskiarvosta $\bar{x} = 6555$ keskihajonnan $s_n = 1564{,}818\ldots$ verran alas- ja ylöspäin poikkeavat katsojamäärät:

- keskiarvosta keskihajonnan verran alaspäin oleva katsojaluku on:

  $\bar{x} - s_n = 6555 - 1564{,}818\ldots = 4990{,}181\ldots \approx \mathbf{4990}$.

- keskiarvosta keskihajonnan verran ylöspäin oleva katsojaluku on:

  $\bar{x} + s_n = 6555 + 1564{,}818\ldots = 8119{,}818\ldots \approx \mathbf{8120}$.

Verrataan joukkueiden katsojalukuja laskettuihin rajoihin: Kärppien, TPS:n, Tapparan ja Ilveksen katsojaluvut jäävät rajojen 4990 – 8120 väliin kun taas Jokerien ja HIFK:n katsojamäärät ylittävät ylärajan 8120.

Jokerien ja HIFK:n katsojamäärät poikkeavat yli keskihajonnan verran keskiarvosta.

35.

a) Kirjoitetaan pituudet taulukkolaskentaohjelman tai laskinsovelluksen taulukkoon ja määritetään tunnusluvut tilastotoimintojen avulla. Vaahteran pituuden

- keskiarvo on $\bar{x} = 14{,}228\ldots \approx 14{,}2$ (metriä)
- keskihajonta on $s_n = 3{,}004\ldots \approx 3{,}0$ (metriä)

b) Aineiston pisin vaahtera on pituudeltaan 19,8 metriä. Lasketaan pituuden $x = 19{,}8$ normitettu arvo:

$$z = \frac{x - \bar{x}}{s_n} = \frac{19{,}8 - 14{,}228\ldots}{3{,}004\ldots\ldots} = 1{,}854\ldots$$

Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos normitetun arvon itseisarvo on suurempi kuin kaksi. Nyt normitetun arvon itseisarvo on $|z| = 1{,}854\ldots < 2$, joten aineiston pisin vaahtera ei poikkea merkitsevästi keskiarvosta.

c) Kahdessa eri jakaumassa (vaahteran pituus ja pihlajan pituus) normitetut arvot vastaavat toisiaan.

Pisimmän vaahteran normitettu arvo laskettiin kohdassa b. Normitettu arvo on $z = 1{,}854\ldots$

Merkitään kysyttyä pihlajan pituutta kirjaimella *x.* Normitettu arvo on

$$\frac{x - \bar{x}}{s} = \frac{x - 6{,}4}{0{,}67}.$$

Merkitään normitetut arvot yhtä suuriksi:

$$\frac{x - 6{,}4}{0{,}67} = 1{,}854\ldots$$

Ratkaistaan yhtälöstä tuntematon *x.* Tehdään se tarkoilla arvoilla kopioimalla taulukkolaskentaohjelmassa tai laskinsovelluksessa laskettu tarkka arvo yhtälönratkaisun laskukenttään.

Yhtälön ratkaisu on $x = 7{,}642\ldots \approx 7{,}6$ (metriä). Pisintä vaahteraa vastaavan pihlajan pituus on siis 7,6 metriä.

36.

Lasketaan arvosanojen keskiarvo painottamalla kutakin arvosanaa sen prosenttiosuuden ilmaisemalla prosenttikertoimella:

$$\begin{aligned}\bar{x} &= 0{,}013 \cdot 4 + 0{,}098 \cdot 5 + 0{,}158 \cdot 6 + 0{,}203 \cdot 7 + 0{,}233 \cdot 8 \\ &\qquad + 0{,}234 \cdot 9 + 0{,}061 \cdot 10 \\ &= 7{,}491 \approx 7{,}5.\end{aligned}$$

Arvosanojen keskiarvo on 7,5.

37.

Kirjoitetaan arvosanat ja niiden frekvenssit taulukkolaskentaohjelman tai laskinsovelluksen taulukkoon. Määritetään tunnusluvut tilastotoimintojen avulla:

- keskiarvo on $\bar{x} = 4{,}860\ldots \approx 4{,}86$,
- keskihajonta on $s_n = 1{,}249\ldots \approx 1{,}25$.

	A	B
1	Arvosana	Frekvenssi
2	7	7
3	6	20
4	5	30
5	4	16
6	3	9
7	2	4
8	0	0

38.

a) Kirjoitetaan maksuhäiriö-merkintöjen lukumäärät taulukkolaskentaohjelman tai laskinsovelluksen taulukkoon (kuvassa sarake B).

	A	B
1	**Vuosi**	**Maksuhäiriöt (lkm)**
2	2007	309300
3	2008	292500
4	2009	305400
5	2010	318700
6	2011	327500
7	2012	348000
8	2013	360000
9	2014	366700
10	2015	369900
11	2016	373100

Määritetään merkintöjen lukumäärän keskiarvo tilastotoimintojen avulla:

- keskiarvo on $\bar{x} = 337\ 110$

Maksuhäiriömerkintöjä oli kyseisellä aikavälillä keskimäärin 337 110 merkintää per vuosi.

b) Vaihteluvälin pituus lasketaan vähentämällä suurimmasta muuttujan arvosta pienin muuttujan arvo. Vaihteluvälin pituuden laskemiseksi pitää siis määrittää maksuhäiriömerkintöjen pienin ja suurin arvo. Useimmat tilastotoiminnot kertovat suurimman ja pienimmän arvon samassa yhteydessä muiden tilastollisten tunnuslukujen kanssa.

Maksuhäiriömerkintöjen

- pienin arvo on 292 500,
- suurin arvo on 373 100.

Maksuhäiriömerkintöjen vaihteluvälin pituus on

$$373\ 100 - 292\ 500 = 80\ 600$$

c) Lasketaan aineiston pienimmän arvon 292 500 normitettu arvo. Laskua varten tarvitaan keskiarvo sekä keskihajonta:

- keskiarvo on $\bar{x} = 337\ 110$
- keskihajonta on $s_n = 28\ 440{,}866\ldots$

Pienimmän arvon $x = 292\ 500$ normitettu arvo on

$$z = \frac{x - \bar{x}}{s_n} = \frac{292\ 500 - 337\ 110}{28\ 440{,}866\ldots} = -1{,}568\ldots$$

Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos normitetun arvon itseisarvo on suurempi kuin kaksi. Nyt normitetun arvon itseisarvo on $|z| = 1{,}568\ldots < 2$, joten aineiston pienin arvo ei poikkea merkitsevästi keskiarvosta.

39.

Lasketaan pituuksien x ja leveyksien y keskiarvot taulukkolaskentaohjelmalla:

- pituuksien keskiarvo on $\bar{x} = x_0 = 83{,}212\ldots \approx 83{,}2$ (cm)
- leveyksien keskiarvo on $\bar{y} = y_0 = 54{,}012\ldots \approx 54{,}0$ (cm)

Näiden perusteella keskimääräisen levyn pinta-ala on

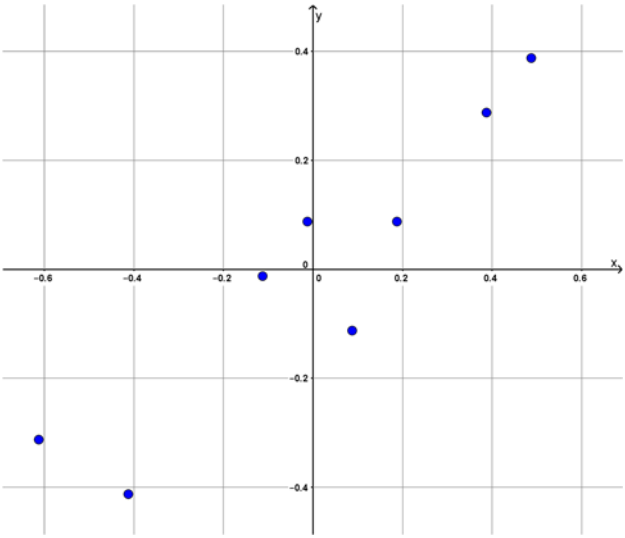
$$A = x_0 \cdot y_0 = 83{,}212\ldots \cdot 54{,}012\ldots$$
$$= 4\,494{,}515\ldots \approx 4490\ (\mathrm{cm}^2).$$

Esitetään mittaustulokset (x, y)-koordinaatistossa, jonka origo on muuttujien keskiarvojen määräämässä pisteessä (x_0, y_0).

Tätä varten lasketaan mittaustulosten poikkeamat keskiarvosta, eli erotukset $x - x_0$ ja $y - y_0$.

	A	B	C	D
1	**x**	**y**	**x-x0**	**y-y0**
2	83.4	54.1	0.1875	0.0875
3	83.6	54.3	0.3875	0.2875
4	82.8	53.6	-0.4125	-0.4125
5	82.6	53.7	-0.6125	-0.3125
6	83.1	54	-0.1125	-0.0125
7	83.7	54.4	0.4875	0.3875
8	83.2	54.1	-0.0125	0.0875
9	83.3	53.9	0.0875	-0.1125

Esitetään poikkeamien (kuvassa sarakkeet C ja D) määräämät pisteet koordinaatistossa.



40.

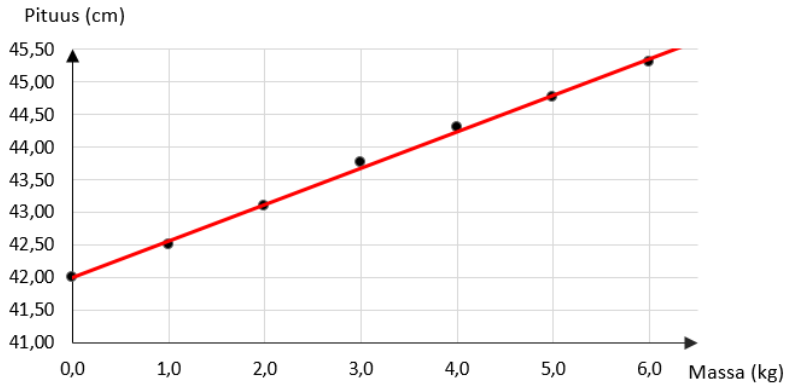
a) Kirjoitetaan punnusten massat ja niitä vastaavat kierrejousen pituudet taulukkoon ja piirretään taulukon avulla sirontakuvaaja: selittävän muuttujan (punnuksen massan) arvot ovat x-akselilla ja selitettävän muuttujan (jousen pituuden) arvot ovat y-akselilla.

	A	B
1	**Massa (kg)**	**Pituus (cm)**
2	0,0	42,00
3	1,0	42,50
4	2,0	43,10
5	3,0	43,75
6	4,0	44,30
7	5,0	44,76
8	6,0	45,30

Sovitetaan pistejoukkoon suora. Riippuvuutta kuvaavan suoran yhtälö on

$$y = 0{,}55785\ldots x + 41{,}999\ldots \approx 0{,}5579x + 42{,}00.$$

Huomaa, että kerroin ja vakio pyydettiin neljän merkitsevän numeron tarkkuudella.



b) Punnuksen massa x ratkaistaan sijoittamalla pituus $y = 44{,}00$ regressiosuoran yhtälöön:

$$44{,}00 = 0{,}5579x + 42{,}00$$

Yhtälö voidaan ratkaista symbolisen laskennan ohjelmalla: $x = 3{,}584 \ldots \approx 3{,}6$ (kg). Punnuksen massa on siis 3,6 kiloa.

c) Jousen pituus y lasketaan sijoittamalla massa $x = 2{,}5$ regressiosuoran yhtälöön:

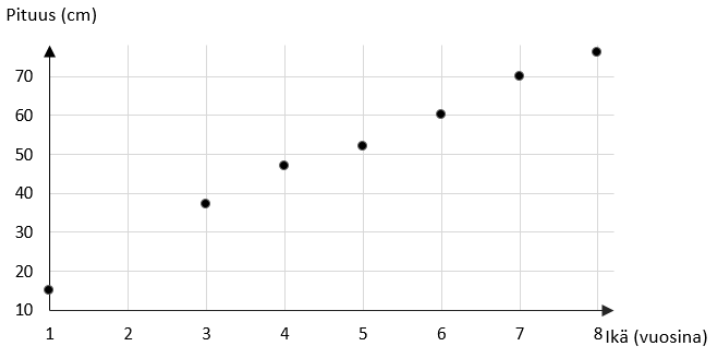
$$y = 0{,}5579 \cdot 2{,}5 + 42{,}00 = 43{,}39475 \approx 43{,}39 \text{ (cm)}$$

Jousen pituus on siis 43,39 cm.

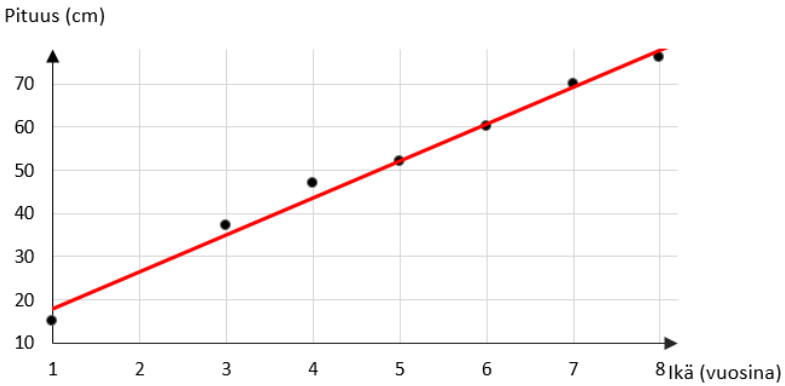
41.

a) Kirjoitetaan iät ja niitä vastaavat pituudet taulukkoon ja piirretään taulukon avulla sirontakuvaaja: selittävän muuttujan (iän) arvot ovat *x*-akselilla ja selitettävän muuttujan (pituuden) arvot ovat *y*-akselilla.

	A	B
1	**Ikä (vuosina)**	**Pituus (**
2	1	15
3	3	37
4	4	47
5	5	52
6	6	60
7	7	70
8	8	76



Havaintoarvot näyttävät asettuvan suoralle, joten lineaarinen malli sopii parhaiten kuvaamaan hauen pituuden riippuvuutta sen iästä. Sovitetaan pistejoukkoon suora.



Riippuvuutta kuvaavan suoran yhtälö on

$$y = 8{,}549\ldots x + 9{,}475\ldots \approx 8{,}55x + 9{,}48.$$

2-vuotiaan hauen pituus y lasketaan sijoittamalla ikä $x = 2$ regressiosuoran yhtälöön:

$$y = 8{,}55 \cdot 2 + 9{,}48 = 26{,}58 \approx 27 \text{ (cm)}.$$

Mallin mukaan hauki on 2-vuotiaana pituudeltaan 27 cm.

12-vuotiaan hauen pituus y lasketaan sijoittamalla ikä $x = 12$ regressiosuoran yhtälöön:

$$y = 8{,}55 \cdot 12 + 9{,}48 = 112{,}08 \approx 112 \text{ (cm)}.$$

Mallin mukaan hauki on 12-vuotiaana pituudeltaan 112 cm.

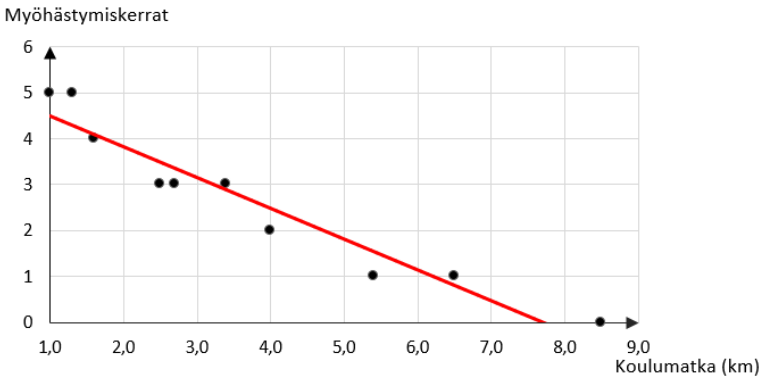
42.

a) Kirjoitetaan koulumatkat ja niitä vastaavat myöhästymiskerrat taulukkoon ja piirretään taulukon avulla sirontakuvaaja: selittävän muuttujan (koulumatkan pituuden) arvot ovat x-akselilla ja selitettävän muuttujan (myöhästymiskertojen) arvot ovat y-akselilla.

	A	B
1	**Koulumatka (km)**	**Myöhästymis-kerrat**
2	2,7	3
3	1,6	4
4	4,0	2
5	1,3	5
6	1,0	5
7	6,5	1
8	8,5	0
9	2,5	3
10	5,4	1
11	3,4	3

Sovitetaan pistejoukkoon suora. Riippuvuutta kuvaavan suoran yhtälö on

$$y = -0{,}669\ldots x + 5{,}169\ldots \approx -0{,}67x + 5{,}2.$$



b) Korrelaatiokertoimen arvo on

$$r = -0{,}9610 \ldots \approx -0{,}96.$$

Korrelaatiokertoimen itseisarvo on $|r| = 0{,}96 > 0{,}8$, joten lineaarinen riippuvuus on voimakas.

43.

Nappeja on yhteensä $7 + 5 = 12$.

a) Valitaan 12 napista kolme: valinta voidaan tehdä

$$\binom{12}{3} = 220$$

eri tavalla.

b) Valitaan seitsemästä pienestä napista kaksi: valinta voidaan tehdä

$$\binom{7}{2} = 21$$

eri tavalla.

Valitaan viidestä isosta napista yksi: valinta voidaan tehdä

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

eri tavalla.

Kaksi pientä nappia ja yksi iso nappi voidaan tuloperiaatteen mukaan valita

$$\binom{7}{2} \cdot \binom{5}{1} = 21 \cdot 5 = 105$$

eri tavalla.

Kysytty todennäköisyys on

$$\begin{aligned} P(\text{"kaksi pientä ja yksi iso nappi"}) &= \frac{105}{220} = \frac{21}{44} \\ &= 0{,}477\ldots \approx 0{,}48. \end{aligned}$$

c) Tapahtuman $A =$ ”ainakin yksi koristenappi” vastatapahtuma on $\bar{A} =$ ”ei yhtään koristenappia”. Lasketaan kysytty todennäköisyys säännön $P(A) = 1 - P(\bar{A})$ avulla.

Todennäköisyys, että pussista otettu nappi on koristenappi, on $28\ \% = 0{,}28$. Todennäköisyys, että nappi ei ole koristenappi on silloin $100\ \% - 28\ \% = 72\ \% = 0{,}72$.

Jos pussista otetaan n nappia, niin vastatapahtuman todennäköisyys on kertolaskusäännön mukaan $P(\bar{A}) = 0{,}72^n$. Kysytyn tapahtuman todennäköisyys on

$$P(A) = 1 - P(\bar{A}) = 1 - 0{,}72^n.$$

Halutaan, että tämä todennäköisyys on vähintään 0,95:

$$1 - 0{,}72^n \geq 0{,}95,$$

josta ratkaistaan tuntematon $n \geq 9{,}119\ \ldots$

Pussista pitää ottaa vähintään 10 nappia.

44.

Kirjaimia on yhteensä 10. Sanassa on viisi vokaalia (kirjaimet Y, I, O ja A) ja viisi konsonanttia (kirjaimet L, P ja S). Kolme kirjainta voidaan valita yhteensä

$$\binom{10}{3} = 120$$

eri tavalla.

a) Todennäköisyys, että ensin otettu kirjain on vokaali on $\frac{5}{10} = \frac{1}{2}$.

b) Valitaan viidestä vokaalista yksi ja viidestä konsonantista kaksi.

Kysytty todennäköisyys on

$$P(\text{"yksi vokaali } \textbf{ja} \text{ kaksi konsonanttia"})$$

$$= \frac{\overbrace{\binom{5}{1}}^{\text{1 vokaali}} \cdot \overbrace{\binom{5}{2}}^{\text{2 konsonanttia}}}{\binom{10}{3}}$$

$$= \frac{5 \cdot 10}{120} = \frac{50}{120}$$

$$= \frac{5}{12} = 0{,}416\ldots \approx 0{,}42.$$

c) Kirjaimia I ja L on kaksi ja O-kirjaimia on yksi.

Kysytty todennäköisyys on

$$P(\text{"ILO"})$$

$$= \frac{\overbrace{\binom{2}{1}}^{\text{I-kirjain}} \cdot \overbrace{\binom{2}{1}}^{\text{L-kirjain}} \cdot \overbrace{\binom{1}{1}}^{\text{O-kirjain}}}{\binom{10}{3}} = \frac{2 \cdot 2 \cdot 1}{120} = \frac{4}{120}$$

$$= \frac{1}{30} = 0{,}0333\ldots \approx 0{,}033.$$

45.

Palloja on yhteensä $5 + 10 = 15$.

Tapahtuman $A =$"ainakin yksi punainen pallo" vastatapahtuma on $\bar{A} =$"kaikki viisi palloa ovat mustia". Lasketaan kysytty todennäköisyys säännön $P(A) = 1 - P(\bar{A})$ avulla.

Vastatapahtuman todennäköisyys on

$$P(\bar{A}) = P(\text{"viisi mustaa palloa"}) = \frac{\binom{10}{5}}{\binom{15}{5}} = \frac{252}{3003} = \frac{12}{143}.$$

Kysytty todennäköisyys on silloin

$$\begin{aligned} &P(\text{"ainakin yksi punainen pallo"}) \\ &= 1 - P(\text{"viisi mustaa palloa"}) \\ &= 1 - \frac{12}{143} \\ &= \frac{131}{143} = 0{,}916\ldots \approx 0{,}92. \end{aligned}$$

Tapahtuma $A =$"viisi samanväristä palloa" sisältää kaksi tapausta:

- kaikki viisi palloa ovat punaisia: tämän tapauksen todennäköisyys on

$$P(\text{"viisi punaista palloa"}) = \frac{\binom{5}{5}}{\binom{15}{5}} = \frac{1}{3003}$$

- kaikki viisi palloa ovat mustia: tämän tapauksen todennäköisyys laskettiin yllä, ja todennäköisyys on
$$P(\text{"viisi mustaa palloa"}) = \frac{252}{3003}$$

Kysytty kokonaistodennäköisyys on yhteenlaskusäännön nojalla

$$P(\text{"viisi samanväristä palloa"})$$
$$= P(\text{"viisi punaista palloa"}) + P(\text{"viisi mustaa palloa"})$$
$$= \frac{1}{3003} + \frac{252}{3003} = \frac{253}{3003}$$
$$= \frac{23}{273} = 0{,}0842\ldots \approx 0{,}084.$$

46.

a) Erilaisia istumajärjestyksiä on tuloperiaatteen mukaan yhteensä

$$30! = 2{,}652\ldots \cdot 10^{32} \approx 2{,}65 \cdot 10^{32}.$$

b) Valitaan 30 pulpetista kolme, jotka jäävät tyhjäksi. Valinta voidaan tehdä

$$\binom{30}{3} = 4060$$

eri tavalla.

Erilaisia 27 oppilaan istumajärjestyksiä on tuloperiaatteen mukaan yhteensä

$$\begin{aligned}\underbrace{30 \cdot 29 \cdot 28 \cdot \ldots \cdot 4}_{\text{27 tulontekijää}} &= \frac{30!}{3 \cdot 2} \\ &= \frac{30!}{6} \\ &= 4{,}420\ldots \cdot 10^{31} \approx 4{,}42 \cdot 10^{31}.\end{aligned}$$

c) Tietokone käsittelee 10^{12} järjestystä sekunnissa. Vuorokaudessa on $24 \cdot 60 \cdot 60 = 86\,400$ sekuntia, joten vuorokaudessa tietokone käsittelee $86\,400 \cdot 10^{12} = 864 \cdot 10^{14}$ järjestystä. Vuodessa on 365,25 vuorokautta, joten yhdessä vuodessa tietokone käsittelee

$$365{,}25 \cdot 864 \cdot 10^{14} = 315\,576 \cdot 10^{14}$$

järjestystä.

a-kohdan erilaisia istumajärjestyksiä on yhteensä 30! Lasketaan, kuinka monta vuotta tietokoneella menee näiden läpikäymiseen kun se käsittelee $315\,576 \cdot 10^{14}$ järjestystä vuodessa:

$$\frac{30!}{315\,576 \cdot 10^{14}} = 8{,}405\ldots \cdot 10^{12} \text{ (vuotta).}$$

Tietokoneelta menee järjestysten läpikäymiseen noin $8{,}4 \cdot 10^{12}$ vuotta.

47.

a) Ovikoodin neljä keskenään erisuurta paritonta numeroa valitaan lukujen 1, 3, 5, 7 ja 9 joukosta. Ensimmäisen numeron valintaan on viisi vaihtoehtoa, toisen valintaan on neljä vaihtoehtoa, kolmannen valintaan kolme ja viimeisen valintaan kaksi vaihtoehtoa.

Tuloperiaatteen mukaan vaihtoehtoja on yhteensä

$$5 \cdot 4 \cdot 3 \cdot 2 = 120.$$

b) Numero 9 ei ole käytössä, joten ensimmäisen numeron valintaan on neljä vaihtoehtoa: 1, 3, 5 ja 7.

Koodissa ei esiinny ”vierekkäisiä” parittomia numeroita, joten numero 1 ei kelpaa ensimmäiseksi numeroksi. Jos nimittäin ensimmäiseksi numeroksi valitaan 1, niin toisen numeron pitää olla 5 tai 7: mahdolliset koodit ovat tällöin 1<u>53</u>7, 1<u>57</u>3, 17<u>35</u> ja 1<u>75</u>3, joista yksikään ei kelpaa koska jokaisessa esiintyy <u>vierekkäisiä</u> parittomia numeroita.

Samalla tavalla päätellään, ettei numero 7 kelpaa ensimmäiseksi numeroksi.

Ensimmäisen numeron on siis oltava joko 3 tai 5. Jos ensimmäinen numero on 3, niin toisen numeron pitää olla 7: mahdolliset koodit ovat tällöin 3715 ja 3<u>75</u>1, joista vain **3715** kelpaa. Jos ensimmäinen numero on 5, niin toisen numeron on oltava 1: mahdolliset koodit ovat tällöin 5<u>13</u>7 ja 5173, joista vain **5173** kelpaa.

Professorin riittää siis kokeilla kaksi koodia: 3715 ja 5173.

48.

Värisävyn ilmoitetaan kuudella merkillä, ja käytössä on 16 merkkiä. Jokaisen merkin valintaan on siis 16 vaihtoehtoa. Tuloperiaatteen mukaan erilaisia koodeja on yhteensä

$$\overbrace{16 \cdot 16}^{\text{punainen}} \cdot \overbrace{16 \cdot 16}^{\text{vihreä}} \cdot \overbrace{16 \cdot 16}^{\text{sininen}} = 16^2 \cdot 16^2 \cdot 16^2 = 16^6$$
$$= 16\,777\,216.$$

Erilaisia värisävyjä on siis yhteensä 16 777 216.

49.

Tapahtuman ”ainakin yksi vasenkätinen” vastatapahtuma on ”ei yhtään vasenkätistä”, eli ”kaikki oikeakätisiä”.

Todennäköisyys, että satunnaisesti valittu henkilö on vasenkätinen on 10 %, joten todennäköisyys että valittu henkilö on oikeakätinen on $100\ \% - 10\ \% = 90\ \%$.

Lasketaan kysytty todennäköisyys vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla.

Merkitään ryhmään tulevien henkilöiden lukumäärää kirjaimella *n.* Kertolaskusäännön mukaan vastatapahtuman todennäköisyys on

$$P(\text{"kaikki oikeakätisiä"}) = 0{,}90^n,$$

joten

$$P(\text{"ainakin yksi vasenkätinen"})$$
$$= 1 - P(\text{"kaikki oikeakätisiä"})$$
$$= 1 - 0{,}90^n.$$

Halutaan, että tämä todennäköisyys on vähintään 0,8:

$$1 - 0{,}90^n \geq 0{,}8,$$

josta ratkaistaan tuntematon $n \geq 15{,}275\ \ldots$

Ryhmässä tulee olla vähintään 16 ihmistä.

50.

Tulos ”vähintään 8 pistettä” sisältää tilanteet ”8 pistettä” ja ”9 pistettä” ja ”10 pistettä”. Lasketaan kunkin tilanteen todennäköisyys.

Tuloksen ”10 pistettä” saa arvaamalla kaikki 10 vastausta oikein. Oikean arvauksen todennäköisyys on jokaisessa tehtävässä $\frac{1}{2}$. Kertolaskusäännön mukaan

$$P(\text{"10 pistettä"}) = \left(\frac{1}{2}\right)^{10}.$$

Yksi väärä arvaus johtaa pistemäärään $9 - 1 = 8$ pistettä. Tulosta ”9 pistettä” ei siis voi saavuttaa millään tavalla, joten $P(\text{"9 pistettä"}) = 0$

Tuloksen ”8 pistettä” saa arvaamalla yhden vastauksen väärin ja 9 vastausta oikein. Väärä arvaus voi osua mihin hyvänsä kymmenestä kysymyksestä: ensimmäiseen kysymykseen tai toiseen kysymykseen tai kolmanteen jne. Tapahtuma ”yksi väärin” sisältää siis **10** eri tilannetta.

Tuloksen ”8 pistettä” todennäköisyys on kerto- ja yhteenlaskusääntöjen mukaan

$$P(\text{"8 pistettä"}) = \mathbf{10} \cdot \overbrace{\frac{1}{2}}^{\text{yksi väärä}} \cdot \overbrace{\left(\frac{1}{2}\right)^{9}}^{\text{9 oikein}} = 10 \cdot \left(\frac{1}{2}\right)^{10}.$$

Kysytty kokonaistodennäköisyys saadaan yhteenlaskusäännöllä:

$$P(\text{"vähintään 8 oikein"})$$
$$= P(\text{"8 oikein"}) + P(\text{"9 oikein"}) + P(\text{"10 oikein"})$$
$$= 10 \cdot \left(\frac{1}{2}\right)^{10} + 0 + \left(\frac{1}{2}\right)^{10}$$
$$= 0{,}0107 \ldots \approx 0{,}011 = 1{,}1\ \%.$$

51.

Tehtävä voidaan ratkaista monella eri tavalla: esitetään kaksi ratkaisua.

Ratkaisu 1: Merkitään punaisten pallojen lukumäärää kirjaimella *x.* Valkoisia palloja on yhteensä *n* kappaletta, joten palloja on yhteensä $x + n$ kappaletta. Punaisten pallojen osuus on silloin

$$\frac{x}{x+n}.$$

Todennäköisyys sille, että valitaan punainen pallo, on 0,4. Punaisia palloja on siis 40 % kaikista palloista, joten saadaan yhtälö:

$$\frac{x}{x+n} = 0{,}40,$$

mistä ratkaistaan $x = \frac{2}{3}n$.

Ratkaisu 2: Punaisen pallon todennäköisyys on 0,4 joten punaisia palloja on 40 % kaikista palloista. Valkoisia palloja on tällöin $100\ \% - 40\ \% = 60\ \%$.

Merkitään punaisten pallojen lukumäärää kirjaimella *x.* Valkoisia palloja on *n* kappaletta. Verrataan punaisten pallojen määrää valkoisten pallojen määrään:

$$\frac{x}{n} = \frac{40}{60},$$

josta ratkaistaan $x = \frac{2}{3}n$.

Kun valkoisia palloja on *n* kappaletta, niin punaisia palloja on $\frac{2}{3}n$ kappaletta.

4.3 Monivalintatehtävät

1.

Asetetaan arvosanat suuruusjärjestykseen: 7 7 $\underline{7\ 8}$ 8 9

Moodi on keskimmäisin arvosana.

Nyt kun keskimmäisiä arvosanoja on kaksi, on moodi näiden keskiarvo, eli $\text{Mo} = \frac{7+8}{2} = 7{,}5$.

Oikea vaihtoehto on siis b.

2.

Merkitään isän ikää kirjaimella *x.* Perheenjäsenten ikien keskiarvo on silloin

$$\frac{14 + 16 + 48 + x}{4}.$$

Tiedetään, että keskiarvo on 32 vuotta, joten saadaan yhtälö:

$$\frac{14 + 16 + 48 + x}{4} = 32 \quad | \cdot 4$$
$$78 + x = 128$$
$$x = 128 - 78 = 50$$

Isän ikä on siis 50. Oikea vaihtoehto on siis a.

Tässä isän ikä ratkaistiin yhtälön avulla. Voit myös laskea perheenjäsenten ikien keskiarvon annetuilla vaihtoehdoilla ja huomata, että vaihtoehto a antaa oikean keskiarvon, sillä

$$\frac{14 + 16 + 48 + 50}{4} = \frac{128}{4} = 32.$$

3.

Moodi on se muuttujan arvo, jonka frekvenssi on suurin.

Arvolla $x = 3$ on suurin frekvenssi, joten moodi on 3.

Oikea vaihtoehto on b.

4.

Muuttujan arvo poikkeaa merkitsevästi keskiarvosta, jos se on yli kahden keskihajonnan päässä keskiarvosta. Tällöin normitetun arvon itseisarvo on suurempi kuin 2.

Normitetaan annetut vaihtoehdot: keskiarvo on $\bar{x} = 10$ ja keskihajonta on $s = 1{,}5$.

a-kohdan arvon $x = 7{,}5$ normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{7{,}5 - 10}{1{,}5} = -1{,}666\ldots$$

Normitetun arvon itseisarvo on $|z| = 1{,}666\ldots < 2$, joten a-kohdan arvo ei poikkea merkitsevästi keskiarvosta.

b-kohdan arvon $x = 11{,}5$ normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{11{,}5 - 10}{1{,}5} = 1.$$

Normitetun arvon itseisarvo on $|z| = 1 < 2$, joten b-kohdan arvo ei poikkea merkitsevästi keskiarvosta.

c-kohdan arvon $x = 14{,}5$ normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{14{,}5 - 10}{1{,}5} = 3.$$

Normitetun arvon itseisarvo on $|z| = 3 > 2$, joten c-kohdan arvo poikkeaa merkitsevästi keskiarvosta.

Oikea vaihtoehto on siis c.

Tehtävä voidaan ratkaista myös laskemalla tasan kahden keskihajonnan verran keskiarvosta alas- ja ylöspäin olevat arvot:

- keskiarvosta kaksi keskihajontaa alaspäin oleva arvo on:

  $\bar{x} - 2s = 10 - 2 \cdot 1{,}5 = 10 - 3 = 7.$

- keskiarvosta kaksi keskihajontaa ylöspäin oleva lukumäärä on:

  $\bar{x} + 2s = 10 + 2 \cdot 1{,}5 = 10 + 3 = 13.$

Arvot, jotka ovat välillä 7–13 eivät poikkea merkitsevästi keskiarvosta, joten ainoastaan c-kohdan arvo poikkeaa.

5.

Tapahtuman $A =$”ainakin kerran vähintään 3” todennäköisyys kannattaa laskea vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla.

Vastatapahtuman $\bar{A} =$”molemmilla kerroilla 1 tai 2” todennäköisyys on kertolaskusäännön nojalla

$$P(\text{"molemmilla kerroilla 1 tai 2"}) = \frac{2}{6} \cdot \frac{2}{6} = \frac{4}{36}^{(4} = \frac{1}{9}.$$

Kysytty todennäköisyys on

$$\begin{aligned} &P(\text{"ainakin kerran vähintään 3"}) \\ &= 1 - P(\text{"molemmilla kerroilla 1 tai 2"}) \\ &= 1 - \frac{1}{9} \\ &= \frac{8}{9}. \end{aligned}$$

Oikea vaihtoehto on b.

6.

Kertolaskusäännön mukaan

$$P(\text{"kuningas ja ässä"}) = \frac{4}{52} \cdot \frac{4}{51} = \frac{16}{2652} \overset{(4}{=} \frac{4}{663}.$$

Tämä ilmoittaa todennäköisyyden sille, että ensimmäinen kortti oli kuningas ja toinen kortti oli ässä. Suotuisa on myös tapaus, jossa kortit on nostettu toisessa järjestyksessä:

$$P(\text{"ässä ja kuningas"}) = \frac{4}{52} \cdot \frac{4}{51} = \frac{16}{2652} \overset{(4}{=} \frac{4}{663}.$$

Kysytty todennäköisyys on

$$\begin{aligned} P(\text{"kuningas ja ässä"}) + P(\text{"ässä ja kuningas"}) &= \frac{4}{663} + \frac{4}{663} \\ &= \frac{8}{663}. \end{aligned}$$

Oikea vaihtoehto on siis b.

Todennäköisyys voidaan laskea myös kombinaatioiden avulla:

$$\begin{aligned} P(\text{"valitaan kuningas ja ässä"}) &= \frac{\binom{4}{1} \cdot \binom{4}{1}}{\binom{52}{2}} = \frac{4 \cdot 4}{1326} = \frac{16}{1326} \overset{(2}{} \\ &= \frac{8}{663}. \end{aligned}$$

7.

Valitaan kahdestakymmenestä kolme:

$$\binom{20}{3} = 1140.$$

Oikea vaihtoehto on a.

8.

Kertolaskusäännön mukaan

$$P(\text{"kolme sinistä palloa"}) = \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{3}{8} = \frac{60}{720}^{(60} = \frac{1}{12}.$$

Oikea vaihtoehto on c.

9.

Kertolaskusäännön mukaan

$$P(\text{"kaikki maaliin"}) = 0{,}2 \cdot 0{,}2 \cdot 0{,}2 = 0{,}008.$$

Oikea vaihtoehto on b.

10.

Suotuisia jonoja ovat P P T T T T ja T P P T T T ja T T P P T T ja T T T P P T sekä T T T T P P.

Tuloperiaatteen mukaan jonoja on yhteensä

$$5 \cdot 2! \cdot 4! = 5 \cdot 2 \cdot 24 = 240.$$

Oikea vaihtoehto on c.

11.

Mediaani on se muuttujan arvo, jonka suhteellinen summafrekvenssi (sf %) ensimmäisen kerran ylittää arvon 50 %.

Mediaani on siis $x = 15$.

Oikea vaihtoehto on c.

12.

Muuttujan x arvon kasvaessa myös muuttujan y arvo kasvaa. Regressiosuora on siis nouseva suora, jonka kulmakerroin on positiivinen luku. Annetuista vaihtoehdoista ainoastaan suoran $y = 2{,}1 + 1{,}8x$ kulmakerroin $k = 1{,}8$ on positiivinen luku.

Oikea vaihtoehto on b.

13.

Muuttujan *x* arvon kasvaessa muuttujan *y* arvo pienenee, joten korrelaatiokerroin on negatiivinen luku. Muuttujien välinen korrelaatio on voimakas, joten korrelaatiokertoimen itseisarvo on lähellä lukua 1.

Oikea vaihtoehto on a.

14.

Tapahtuman $A =$”ainakin yksi oikein” todennäköisyys kannattaa laskea vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla.

Vastatapahtuman $\bar{A} =$”kaikki väärin” todennäköisyys on kertolaskusäännön nojalla

$$P(\text{"kaikki väärin"}) = \underbrace{\frac{2}{3} \cdot \frac{2}{3} \cdot \ldots \cdot \frac{2}{3}}_{11\text{ kpl}} = \left(\frac{2}{3}\right)^{11}.$$

Kysytty todennäköisyys on

$$P(\text{"ainakin yksi oikein"}) = 1 - \left(\frac{2}{3}\right)^{11}.$$

Oikea vaihtoehto on a.

15.

Kertolaskusäännön mukaan oikean järjestyksen todennäköisyys on

$$P(\text{"oikea järjestys"}) = \frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{1} = \frac{1}{24}.$$

Oikea vaihtoehto on a.

16.

Pituudet ovat pienimmästä suurimpaan:

162 165 170 173 176 178

Keskihajonta kuvaa pituuksien keskimääräistä hajaantumista keskiarvon ympärille. Oikea vaihtoehto voidaan päätellä esimerkiksi seuraavasti:

Lasketaan pituuksien keskiarvo:

$$\frac{162 + 165 + 170 + 173 + 176 + 178}{6} = \frac{1024}{6}$$

$$= 170{,}666 \approx 171 \text{ (cm)}$$

Lasketaan pituuksien *suurin* poikkeama keskiarvosta: tämä on nyt pienimmän pituuden poikkeama keskiarvosta, eli luku

$$171 - 162 = 9 \text{ (cm)}.$$

Pituuksien keskimääräinen poikkeama on korkeintaan suurimman poikkeaman suuruinen, eli keskihajonta on korkeintaan 9 cm.

Oikea vaihtoehto on b.

17.

Kysytty todennäköisyys on yhteenlaskusäännön nojalla

$$
\begin{aligned}
&P(\text{"vain toinen läpäisee kokeen"})\\
&= P(\text{"Rolf läpäisee ja Benjamin ei"})\\
&\qquad\qquad + P(\text{"Benjamin läpäisee ja Rolf ei"})\\
&= 0{,}95 \cdot (1 - 0{,}88) + 0{,}88 \cdot (1 - 0{,}95)\\
&= 0{,}95 \cdot 0{,}12 + 0{,}88 \cdot 0{,}05\\
&= 0{,}158.
\end{aligned}
$$

Oikea vaihtoehto on a.

18.

Kuuden kilpailijan järjestyksiä on tuloperiaatteen mukaan yhteensä $6! = 720$ kappaletta.

Suotuisia järjestyksiä ovat

S S M M M M ja M S S M M M ja M M S S M M ja M M M S S M sekä M M M M S S,

missä S = suomalainen ja M = muunmaalainen.

Tuloperiaatteen mukaan suotuisia järjestyksiä on yhteensä

$$5 \cdot 2! \cdot 4! = 5 \cdot 2 \cdot 24 = 240.$$

Kysytty todennäköisyys on

$$P(\text{"suomalaiset peräkkäin"}) = \frac{240}{720} = 0{,}333\ldots \approx 0{,}33.$$

Oikea vaihtoehto on c.

19.

Moodi on se muuttujan arvo, jonka frekvenssi on suurin, joten moodi on Mo = 4.

Oikea vaihtoehto on a.

20.

Mediaani on se muuttujan arvo, jonka suhteellinen summafrekvenssi ensimmäisen kerran ylittää arvon 50 %, joten mediaani on Md = 3.

Oikea vaihtoehto on c.

21.

Ratkaistaan tehtävä taulukon avulla. Sijoitetaan ensin tehtävässä annetut tiedot taulukkoon.

	Sinisilmäinen	Ei-sinisilmäinen	**Yhteensä**
Tummahiuksinen	3		8
Ei-tummahiuksinen			
Yhteensä	15		**28**

Lasketaan taulukosta puuttuvat tiedot annettujen tietojen avulla.

	Sinisilmäinen	Ei-sinisilmäinen	**Yhteensä**
Tummahiuksinen	3	$8 - 3 = 5$	8
Ei-tummahiuksinen	$15 - 3 = 12$	$13 - 5 = 8$ $20 - 12 = 8$	$28 - 8 = 20$
Yhteensä	15	$28 - 15 = 13$	**28**

Tapaukselle ”ei-sinisilmäinen eikä tummahiuksinen” suotuisia alkeistapauksia on yhteensä 8 (taulukon tummennettu solu).

Kysytty todennäköisyys on $\frac{8}{28}^{(4} = \frac{2}{7}$.

Oikea vaihtoehto on b.

22.

Tapahtuman A = ”korkeintaan kaksi kertaa 5” todennäköisyys kannattaa laskea vastatapahtuman ja säännön $P(A) = 1 - P(\bar{A})$ avulla.

Vastatapahtuman $\bar{A} =$ ”kolme kertaa 5” todennäköisyys on kertolaskusäännön mukaan

$$P(\bar{A}) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}.$$

Kysytty todennäköisyys on

$$P(\text{"korkeintaan kaksi kertaa 5"})$$

$$= 1 - P(\text{"kolme kertaa 5"})$$

$$= 1 - \frac{1}{216} = \frac{215}{126}.$$

Oikea vaihtoehto on b.

23.

Sämpylöitä on yhteensä $7 + 8 = 15$ ja niistä valitaan 3. Valinta voidaan tehdä $\binom{15}{3}$ eri tavalla.

Valitaan 2 vehnäsämpylää ja 1 ruissämpylä:

$$P(\text{"2 vehnäsämpylää ja 1 ruissämpylä"}) = \frac{\overbrace{\binom{8}{2}}^{\text{2 vehnää}} \cdot \overbrace{\binom{7}{1}}^{\text{1 ruis}}}{\binom{15}{3}}$$

Oikea vaihtoehto on c.

24.

Havaintoyksiköiden lukumäärä saadaan summaamalla frekvenssit.

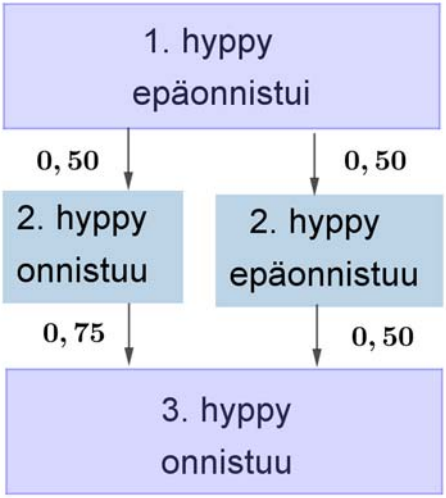
Muuttujan arvon $x = 3$ frekvenssi on $9 - 5 = 4$.

Frekvenssien summa on $3 + 2 + 4 + 3 + 1 + 4 = 17$.

Oikea vaihtoehto on a.

25.

Piirretään puukaavio havainnollistamaan eri tilanteita ja todennäköisyyksiä.



Kysytty todennäköisyys on

$$P(\text{"3. hyppy onnistuu"}) = 0{,}50 \cdot 0{,}75 + 0{,}50 \cdot 0{,}50 = 0{,}625.$$

Oikea vaihtoehto on a.

26.

Maapallon pinta-ala on $A = 4\pi r^2 = 4\pi \cdot 6371^2 \text{ km}^2$.

Tapahtumalle ”osuu Australiaan” suotuisan osan pinta-ala on $A_1 = 7\,692\,000 \text{ km}^2$.

Kysytty todennäköisyys on

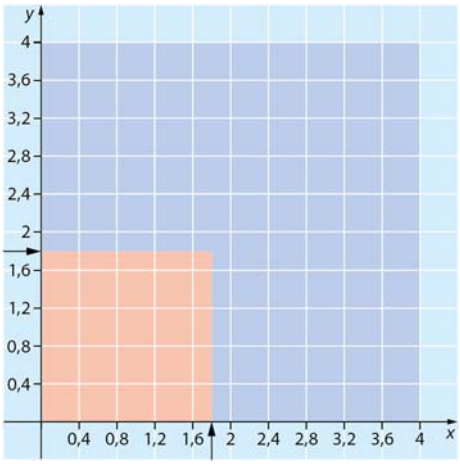
$$\frac{A_1}{A} = \frac{7\,692\,000 \text{ km}^2}{4\pi \cdot 6371^2 \text{ km}^2} = 0{,}01508 \ldots \approx 1{,}51\ \%.$$

Oikea vaihtoehto on b.

27.

Kaikkia alkeistapauksia, eli eri mahdollisuuksia valita kaksi reaalilukua x ja y väliltä $[0, 4]$ kuvaa neliö, jonka pinta-ala on $A = 4 \cdot 4 = 16$.

Tapahtumalle ”molemmat luvut ovat pienempiä kuin 1,8” suotuisia alkeistapauksia kuvaa kuvioon piirretty pienempi punainen neliö, jonka pinta-ala on $A_1 = 1{,}8 \cdot 1{,}8 = 3{,}24$.



Kysytty todennäköisyys saadaan pinta-alojen avulla.

$$P(\text{"molemmat luvut pienempiä kuin 1,8"}) = \frac{A_1}{A} = \frac{3{,}24}{16}$$
$$= 0{,}2025 \approx 0{,}20.$$

Oikea vaihtoehto on c.

28.

Arvosanoja on yhteensä $4 + 3 + 6 + 5 + 3 + 1 = 22$.

Arvosanojen keskiarvo on

$$\frac{4 \cdot 5 + 3 \cdot 6 + 6 \cdot 7 + 5 \cdot 8 + 3 \cdot 9 + 1 \cdot 10}{22} = \frac{157}{22}$$

$$= 7{,}136 \ldots \approx 7{,}14.$$

Oikea vaihtoehto on b.

29.

Muuttujan arvoa $x = 10{,}0$ vastaava normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{10{,}0 - \bar{x}}{0{,}40}.$$

Tiedetään, että normitettu arvo on 1,20, joten saadaan yhtälö:

$$\frac{10{,}0 - \bar{x}}{0{,}40} = 1{,}20 \quad | \cdot 0{,}40$$

$$10{,}0 - \bar{x} = 0{,}48$$

$$\bar{x} = 10{,}0 - 0{,}48 = 9{,}52.$$

Muuttujan arvojen keskiarvo on siis $\bar{x} = 9{,}52$.

Oikea vaihtoehto on a.

30.

Viiden osan järjestyksiä on yhteensä $5! = 120$.

Oikeita järjestyksiä on yksi. Oikean järjestyksen todennäköisyys on $\frac{1}{120}$.

Oikea vaihtoehto on b.

31.

Tehtävä voidaan ratkaista Venn-diagrammin tai taulukon avulla. Ratkaistaan tehtävä taulukon avulla.

Sijoitetaan ensin tehtävässä annetut tiedot taulukkoon.
Kaikkien työntekijöiden määrä prosentteina on 100 %.

	Käyttää valkoista paitaa	Ei käytä valkoista paitaa	**Yhteensä**
Käyttää solmiota			36
Ei käytä solmiota			
Yhteensä	60		**100**

Merkitään sekä valkoista paitaa että solmiota käyttävien prosenttiosuutta kirjaimella x. Täydennetään taulukkoon tehtävän ratkaisuun tarvittavat puuttuvat tiedot annettujen tietojen avulla.

	Käyttää valkoista paitaa	Ei käytä valkoista paitaa	**Yhteensä**
Käyttää solmiota	x	$36 - x$	36
Ei käytä solmiota	$60 - x$		
Yhteensä	60		**100**

Tapaukselle ”valkoinen paita tai solmio (tai molemmat)” suotuisien alkeistapauksien prosenttiosuus on yhteensä

$$x + (36 - x) + (60 - x) = 96 - x.$$

Tiedetään, että tämä prosenttiosuus on $0{,}80 = 80\ \%$, joten saadaan yhtälö

$$96 - x = 80,$$

mistä ratkaistaan $x = 96 - 80 = 16$. Valkoista paitaa tai solmiota (tai molempia) käyttää siis 16 % työntekijöistä.

Oikea vaihtoehto on c.

32.

Jokaisessa monivalintatehtävässä on kolme vaihtoehtoa, joista yksi on oikein ja kaksi väärin. Väärän vastauksen todennäköisyys on jokaisen tehtävän kohdalla

$$P(\text{"väärin"}) = \frac{2}{3}.$$

Merkitään ”laskettavien” tehtävien lukumäärää kirjaimella *x.*

Muodostetaan tapahtuman ”ainakin yksi oikea vastaus” todennäköisyys vastatapahtuman avulla.

Vastatapahtuma on ”kaikki vastaukset ovat väärin”.

Kertolaskusäännön mukaan

$$P(\text{”kaikki } x \text{ vastausta on väärin”}) = \underbrace{\frac{2}{3} \cdot \frac{2}{3} \cdot \ldots \cdot \frac{2}{3}}_{x \text{ kpl}} = \left(\frac{2}{3}\right)^x.$$

Tapahtuman ”ainakin yksi oikea vastaus” todennäköisyys on tällöin

$$P(\text{”ainakin yksi oikein”})$$
$$= 1 - P(\text{”kaikki väärin”})$$
$$= 1 - \left(\frac{2}{3}\right)^x.$$

Halutaan, että todennäköisyys tälle tapahtumalle on vähintään 0,95.

Jos ”laskettavien” tehtävien lukumäärä on $x = 6$, niin todennäköisyys ainakin yhdelle oikealle vastaukselle on

$$1 - \left(\frac{2}{3}\right)^6 = 0{,}912 \ldots < 0{,}95.$$

Kuusi tehtävää ei siis riitä.

Jos ”laskettavien” tehtävien lukumäärä on $x = 7$, niin todennäköisyys ainakin yhdelle oikealle vastaukselle on

$$1 - \left(\frac{2}{3}\right)^7 = 0{,}941 \ldots < 0{,}95.$$

Seitsemän tehtävää ei siis riitä.

Jos ”laskettavien” tehtävien lukumäärä on $x = 8$, niin todennäköisyys ainakin yhdelle oikealle vastaukselle on

$$1 - \left(\frac{2}{3}\right)^8 = 0{,}960 \ldots > 0{,}95.$$

Kahdeksan tehtävää siis riittää.

Oikea vaihtoehto on c.

33.

Merkitään yksittäisen arvan voittotodennäköisyyttä kirjaimella *p.* Niklas ostaa kaksi arpaa. Todennäköisyys, että molemmat ovat voittoarpoja on kertolaskusäännön mukaan

$$P(\text{”voitto ja voitto”}) = p \cdot p = p^2.$$

Tiedetään, että tämä todennäköisyys on 0,42, joten saadaan yhtälö:

$$p^2 = 0{,}42.$$

Kysytty todennäköisyys ratkeaa neliöjuuren avulla:

$$p = \sqrt{0{,}42} = 0{,}648\ldots \approx 65\ \%.$$

Oikea vaihtoehto on a.

34.

Kortteja on yhteensä 20. Pelaaja nostaa kaksi korttia. Nosto voidaan tehdä $\binom{20}{2} = 190$ eri tavalla.

Tapahtumalle ”vain toinen korteista on hyökkäyskortti” suotuisa on tilanne, jossa toinen on hyökkäyskortti ja toinen ei.

Kysytty todennäköisyys on

$$P(\text{"1 hyökkäyskortti ja 1 puolustuskortti"})$$

$$= \frac{\overbrace{\binom{12}{1}}^{\text{1 hyökkäys}} \cdot \overbrace{\binom{8}{1}}^{\text{1 puolustus}}}{\binom{20}{2}} = \frac{12 \cdot 8}{190} = \frac{96}{190} = 0{,}5052\ldots \approx 0{,}505.$$

Oikea vaihtoehto on a.

35.

Jokaisen merkin kohdalla oikean arvauksen todennäköisyys on $\frac{1}{2}$.

Todennäköisyys, että kaikki neljä merkkiä arvataan oikein, on kertolaskusäännön mukaan

$$P(\text{"neljä merkkiä oikein"}) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}.$$

Oikea vaihtoehto on c.

Todennäköisyys voidaan laskea myös tuloperiaatteella:

Jokaisen merkin kohdalla on kaksi vaihtoehtoa: 0 tai 1.
Tuloperiaatteen mukaan neljän merkin valitsemiseen on yhteensä $2 \cdot 2 \cdot 2 \cdot 2 = 2^4 = 16$ eri vaihtoehtoa.

Näistä vaihtoehdoista vain yksi on oikea: 1111. Kysytty todennäköisyys on

$$P(\text{"oikein arvattu"}) = \frac{1}{16}.$$