1.1 Tilastotieteen peruskäsitteitä

1. Muodostetaan taulukon perusteella suhteellinen frekvenssijakauma.

Lehti	Levikki	*f* %
Helsingin Sanomat	365 994	$\frac{365994}{2695202} = 0{,}13579... \approx 13{,}6\,\%$
Ilta-Sanomat	143 321	$\frac{143321}{2695202} = 0{,}05317... \approx 5{,}3\,\%$
Aamulehti	130 081	$\frac{130081}{2695202} = 0{,}04826... \approx 4{,}8\,\%$
Turun Sanomat	103 314	$\frac{103314}{2695202} = 0{,}0383... \approx 3{,}8\,\%$
Iltalehti	102 124	$\frac{102124}{2695202} = 0{,}03789... \approx 3{,}8\,\%$
Maaseudun tulevaisuus	83 259	$\frac{83259}{2695202} = 0{,}03089... \approx 3{,}1\,\%$

a) Suhteellisen frekvenssin mukaan Helsingin sanomien prosenttiosuus koko levikistä on 13,6 %.

b) Kolmen suurimman lehden kappalemääräinen levikki on
365 994 + 143 321 + 130 081 = 639 396 lehteä.

c) Lasketaan ensin neljän suurimman lehden kappalemääräinen levikki. Tämä saadaan lisäämällä b-kohdan levikkiin Turun Sanomien levikki 103 314.

Neljän suurimman levikki on 639 396 + 103 314 = 742 710 lehteä.

Tämä prosentteina on

$$\frac{742710}{2695202} = 0{,}27556... \approx 27{,}6\,\%$$

Vastaus: a) 13,6 % b) 639 396 kappaletta c) 27,6 %

2. Muodostetaan taulukon perusteella frekvenssijakaumat.

Oppilaitostyyppi	***f***	***f* %**
peruskoulut	595 727	$\frac{595727}{1165436} = 0{,}511... \approx 51\%$
lukiot	128 642	$\frac{128642}{1165436} = 0{,}1103... \approx 11\%$
ammatillinen koulutus	160 115	$\frac{160115}{1165436} = 0{,}1373... \approx 14\%$
ammattikorkea-koulutus	118 013	$\frac{118013}{1165436} = 0{,}10126... \approx 10\%$
yliopistokoulutus	162 939	$\frac{162939}{1165436} = 0{,}1398... \approx 14\%$
yhteensä	1 165 436	

a) Peruskoulussa tai lukiossa opiskelee 51,1…% + 11,03…% ≈ 62 % opiskelijoista.

b) Yliopistoissa opiskelee

$$\frac{162939}{1165436} = 0{,}1398... \approx 14\% \text{ opiskelijoista.}$$

Muualla kuin yliopistossa opiskelee tällöin

100 % – 13,98… % = 86,019…≈ 86 % opiskelijoista.

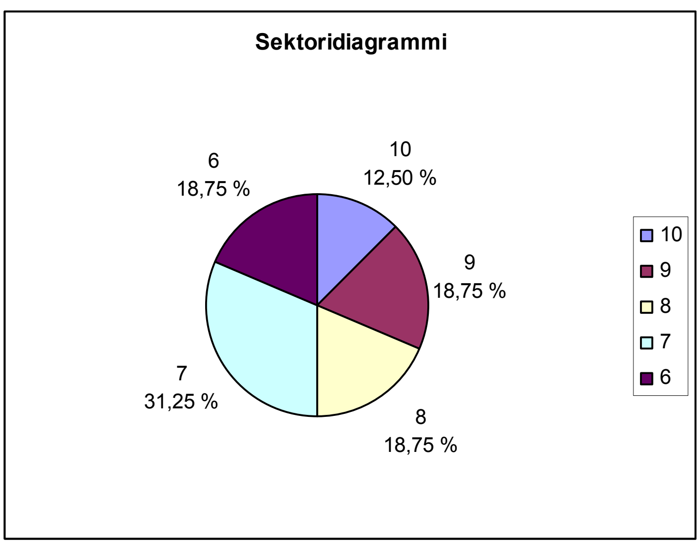
Vastaus: a) 62 % b) 86 %

3. **a)** Muodostetaan frekvenssijakaumat.

Arvosana	***f***	***f* %**
10	2	$\frac{2}{16}=0{,}125=12{,}5\%$
9	3	$\frac{3}{16}=0{,}1875=18{,}75\%$
8	3	$\frac{3}{16}=0{,}1875=18{,}75\%$
7	5	$\frac{5}{16}=0{,}3125=31{,}25\%$
6	3	$\frac{3}{16}=0{,}1875=18{,}75\%$
Yhteensä	16	

Moodi on muuttujan arvo, jolla on suurin frekvenssi. Eniten on arvosanoja 7, joten Mo = 7.

b)

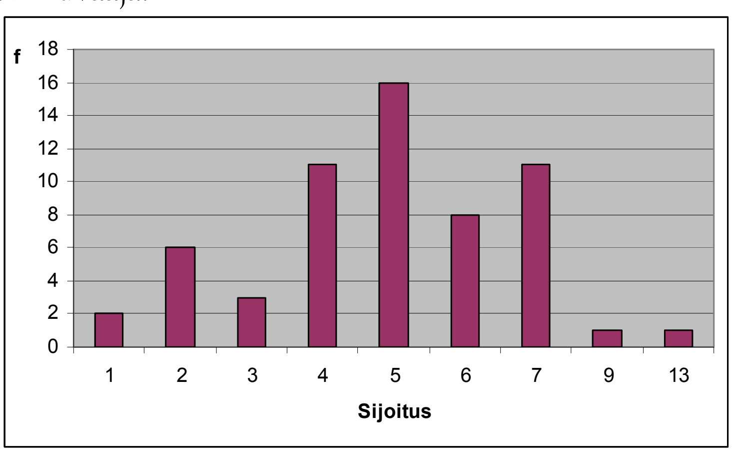


4. **a)** Muodostetaan taulukon perusteella frekvenssijakaumat.

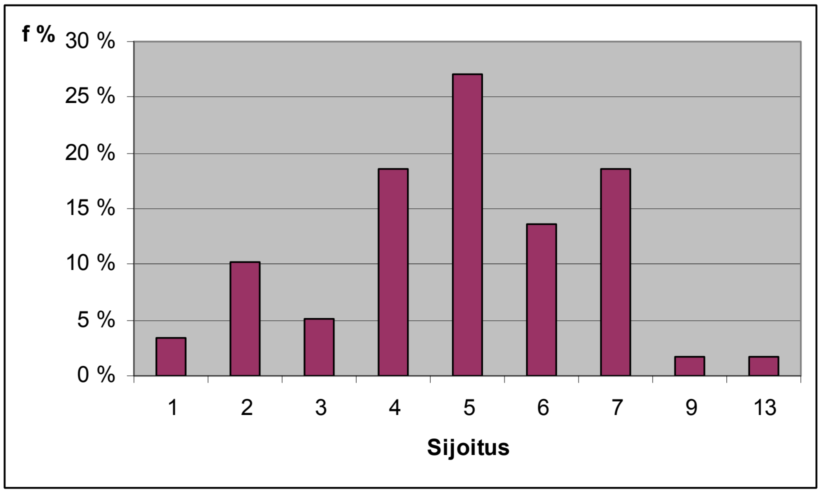
Sijoitus	***f***	***f* %**
1	2	$\frac{2}{59} = 0{,}03389... \approx 3{,}4\%$
2	6	$\frac{6}{59} = 0{,}10169... \approx 10{,}2\%$
3	3	$\frac{3}{59} = 0{,}05084... \approx 5{,}1\%$
4	11	$\frac{11}{59} = 0{,}18644... \approx 18{,}6\%$
5	16	$\frac{16}{59} = 0{,}27118... \approx 27{,}1\%$
6	8	$\frac{8}{59} = 0{,}135593... \approx 13{,}6\%$
7	11	$\frac{11}{59} = 0{,}18644... \approx 18{,}6\%$
9	1	$\frac{1}{59} = 0{,}016694... \approx 1{,}7\%$
13	1	$\frac{1}{59} = 0{,}016694... \approx 1{,}7\%$
Yhteensä	59	

b) Kuvataan frekvenssijakaumaa pylväsdiagrammilla. Pylväsdiagrammissa pylväät ovat erilliset, koska kuvattava muuttuja on diskreetti.

Frekvenssin kuvaaja:



Suhteellisen frekvenssin kuvaaja:



5. **a)** Sektoridiagrammin perusteella asunnon osuus kotitalouksien varallisuudesta on 67 %. Tällöin asunnon arvo on

$$0{,}67 \cdot 248\ 000\,€ = 166\ 160\,€$$

b) Talletusten osuus varallisuudesta on 9 %. Tällöin perheen talletusten määrä on

$$0{,}09 \cdot 248\ 000\,€ = 22\ 320\,€$$

c) Arvopaperisijoitusten osuus varallisuudesta on 6 %. Arvopaperisijoitusten määrä on

$$0{,}06 \cdot 248\,000\,€ = 14\,880\,€$$

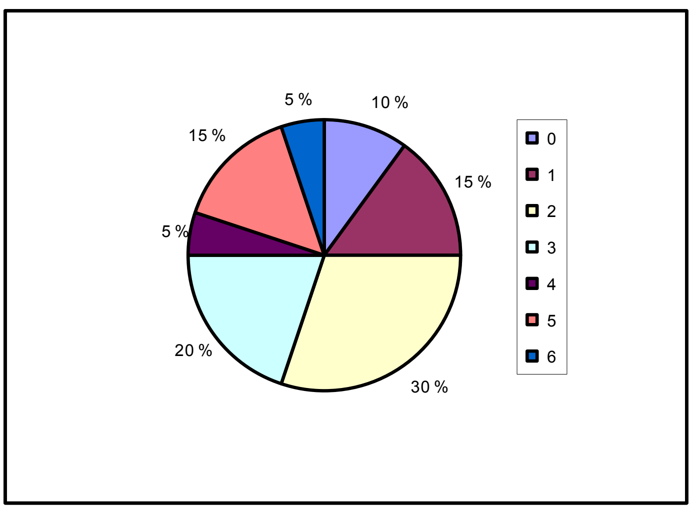
Vastaus: a) 166160€ b) 22 320 € c) 14 880€

6. **a)** Muodostetaan frekvenssijakauma ja suhteellinen frekvenssijakauma.

Päivät/viikko	***f***	***f* %**
0	2	$\frac{2}{20}=0{,}1=10\%$
1	3	$\frac{3}{20}=0{,}15=15\%$
2	6	30 %
3	4	20 %
4	1	5 %
5	3	15 %
6	1	5 %
Yhteensä	20	

b) Moodi on se muuttujan arvo, jolla on suurin frekvenssi. Kuusi työntekijää harrasti liikuntaa kahtena päivänä viikossa, joten Mo = 2.

c) Sektoridiagrammilla havainnollistetaan aineiston suhteellista frekvenssiä.



7. **a)** Määritetään ensin todelliset luokkarajat.

Paino (kg)	**Todellinen alaraja (kg)**	**Todellinen yläraja (kg)**
50–54	49,5	54,5
55–59	54,5	59,5
60–64	59,5	64,5

Lasketaan todellisten luokkarajojen avulla luokkakeskus ja luokkavälin pituus.

Paino (kg)	**Luokkakeskus**	**Luokkavälin pituus**
50–54	$\frac{49,5+54,5}{2}=52$	$54,5-49,5=5$
55–59	$\frac{54,5+59,5}{2}=57$	$59,5-54,5=5$
60–64	$\frac{59,5+64,5}{2}=62$	$64,5-59,5=5$

b) Määritetään ensin todelliset luokkarajat.

5 km juoksuaika (min)	**Todellinen alaraja (min)**	**Todellinen yläraja (min)**
24–29	23,5	29,5
30–35	29,5	35,5
36–41	35,5	41,5

Lasketaan todellisten luokkarajojen avulla luokkakeskus ja luokkavälin pituus.

5 km juoksuaika (min)	**Luokka-keskus**	**Luokkavälin pituus**
24–29	$\frac{23,5+29,5}{2}=26,5$	$29,5-23,5=6$
30–35	$\frac{29,5+35,5}{2}=32,5$	$35,5-29,5=6$
36–41	$\frac{35,5+41,5}{2}=38,5$	$41,5-35,5=6$

8. **a)** Luokitellaan aineisto kuuteen luokkaan ja määritetään luokkien todelliset luokkarajat.

Esimerkiksi:

Korkeusero (m)	***f***	**Todellinen alaraja (m)**	**Todellinen yläraja (m)**
50–99	2	49,5	99,5
100–149	4	99,5	149,5
150–199	6	149,5	199,5
200–249	2	199,5	249,5
250–299	1	249,5	299,5
300–349	1	299,5	349,5

b) Lasketaan luokkien luokkakeskukset.

Korkeusero (m)	**Luokkakeskus (m)**
50–99	$\frac{49,5+99,5}{2}=74,5$
100–149	$\frac{99,5+149,5}{2}=124,5$
150–199	$\frac{149,5+199,5}{2}=174,5$
200–249	$\frac{199,5+249,5}{2}=224,5$
250–299	$\frac{249,5+299,5}{2}=274,5$
300–349	$\frac{299,5+349,5}{2}=324,5$

c) Lasketaan luokkavälin pituus.

Korkeusero (m)	**Luokkavälin pituus (m)**
50–99	99,5 – 49,5= 50
100–149	149,5 – 99,5 = 50
150–199	199,5 – 149,5 = 50
200–249	249,5 – 199,5 = 50
250–299	199,5 – 249,5 = 50
300–349	349,5 – 299,5 = 50

Luokkavälit ovat tasaiset. Luokkavälin pituus on siis 50 m.

9. Luokitellaan aineisto kuuteen luokkaan.

Esimerkiksi:

Luokka (km/h)	***f***
61–70	1
71–80	2
81–90	3
91–100	2
101–110	4
111–120	2

Luokkakeskuksien laskemista varten määritetään luokkien todelliset luokkarajat.

Luokka	**Todellinen alaraja (km/h)**	**Todellinen yläraja (km/h)**	**Luokkakeskus (km/h)**
61–70	60,5	70,5	$\frac{60{,}5+70{,}5}{2}=65{,}5$
71–80	70,5	80,5	$\frac{70{,}5+80{,}5}{2}=75{,}5$
81–90	80,5	90,5	$\frac{80{,}5+90{,}5}{2}=85{,}5$
91–100	90,5	100,5	$\frac{90{,}5+100{,}5}{2}=95{,}5$
101–110	100,5	110,5	$\frac{100{,}5+110{,}5}{2}=105{,}5$
111–120	110,5	120,5	$\frac{110{,}5+120{,}5}{2}=115{,}5$

10. Muodostetaan taulukon perusteella frekvenssijakauma ja lasketaan luokkakeskukset. Luokitellaan aineisto neljään luokkaan.

Esimerkiksi:

Sijoitus	***f***	**Luokkakeskus**
1–3	9	$\frac{1+3}{2}=2$
4–6	6	$\frac{4+6}{2}=5$
7–9	3	$\frac{7+9}{2}=8$
10–12	2	$\frac{10-12}{2}=11$

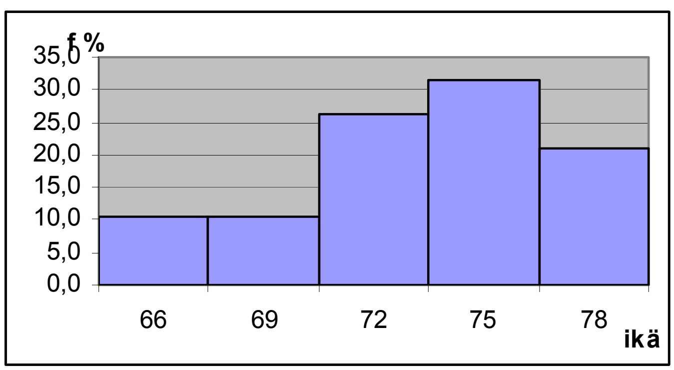
Sijoitukset ovat aina kokonaislukuja ja tarkkoja arvoja. Tästä syystä taulukossa ilmoitetut luokkarajat ovat todellisia luokkarajoja eikä pyöristettyjä likiarvoja.

Sijoitus	**Todellinen alaraja**	**Todellinen yläraja**
1–3	1	3
4–6	4	6
7–9	7	9
10–12	10	12

11. Luokitellaan aineisto tasavälisesti siten, että alimpana luokkana on 65–67 (vuotta). Muodostetaan frekvenssijakaumat ja määritetään luokkakeskukset.

Elinikä (vuotta)	***f***	***f* %**	**Luokkakeskus (vuotta)**
65–67	2	$\frac{2}{19} = 0{,}10526... \approx 10{,}5\%$	$\frac{64{,}5 + 67{,}5}{2} = 66$
68–70	2	10,5 %	$\frac{67{,}5 + 70{,}5}{2} = 69$
71–73	5	$\frac{5}{19} = 0{,}2631... \approx 26{,}3\%$	$\frac{70{,}5 + 73{,}5}{2} = 72$
74–76	6	31,6 %	75
77–79	4	21,1 %	78

Havainnollistetaan aineiston suhteellista frekvenssiä. Kuvataan tätä histogrammilla, koska ikä on jatkuva muuttuja. Piirretään pylväät keskenään yhtä leveiksi ja merkitään pylvään keskikohdalle luokkakeskus.



12. Aineiston inflaatioprosentit luokitellaan siten, että ensimmäisenä luokkana on 0,1–1,0 (%). Muodostetaan frekvenssijakauma ja lasketaan luokkakeskukset.

Inflaatioprosentti (%)	***f***	**Luokkakeskus (%)**
0,1–1,0	3	$\frac{0{,}05+1{,}05}{2}=0{,}55$
1,1–2,0	1	$\frac{1{,}05+2{,}05}{2}=1{,}55$
2,1–3,0	13	2,55
3,1–4,0	10	3,55
4,1–5,0	1	4,55
5,1–6,0	2	5,55

Moodi on sen luokan luokkakeskus, jolla on suurin frekvenssi. Eniten Euroopan valtioita (13 kpl) kuuluu inflaatioprosenttiluokkaan 2,1–3,0 (%). Siis Mo = 2,55.

13. Luokitellaan aineisto annetun taulukon mukaisesti siten, että esimerkiksi luokkaan 56–60 (%) kuuluu ne kunnat, joissa äänestysprosentti oli vähintään 55,5 % ja alle 60,5 %. Kun huomioidaan jokaisessa luokassa todelliset ala- ja ylärajat näin, niin saadaan taulukko:

Äänestysprosentti (%)	***f***	***f* %**	**Luokkakeskus (%)**
56–60	6	$\frac{6}{40}=0{,}15=15\%$	$\frac{55{,}5+60{,}5}{2}=58$
61–65	9	$\frac{9}{40}=0{,}225=22{,}5\%$	$\frac{60{,}5+65{,}5}{2}=63$
66–70	16	40,0 %	68
71–75	7	17,5 %	73
76–80	2	5,0 %	78
Yhteensä	40		

14. Arpoja on kaikkiaan 20 000 kpl.

a) Näistä 12 000 kpl eivät tuo voittoa, joten niiden suhteellinen frekvenssi on

$$\frac{12\,000}{20\,000}\cdot 100\,\% = 60\,\%$$

Siis P(ei voittoa) = 0,6.

b) Päävoitto on 5000 €, ja tällaisia arpoja on 100 kpl, joten näiden suhteellinen frekvenssi on

$$\frac{100}{20\,000}\cdot 100\,\% = 0{,}5\,\%.$$

Siis $P(\text{päävoitto}) = \dfrac{100}{20\,000} = 0{,}005$.

Vastaus: a) 0,6 b) 0,005

15. Laboratoriossa tutkittiin 255 + 87 + 102 + 62 = 506 (kpl) hiiriä.

a) Monivärisiä hiiriä oli 62 kpl, joten niiden suhteellinen frekvenssi on

$$\frac{62}{506}\cdot 100\% = 12{,}252...\%.$$

Siis $P(\text{hiiri on monivärinen}) = 0{,}12252... \approx 0{,}123$.

b) Yksivärisiä hiiriä oli 506 – 62 = 444 (kpl), joten niiden suhteellinen frekvenssi on

$$\frac{444}{506} \cdot 100\ \% = 87{,}747...\ \%.$$

Siis $P(\text{hiiri on yksivärinen}) = 0{,}87747... \approx 0{,}877$.

Vastaus: a) 0,123 b) 0,877

16. Heinäkuussa on päiviä 31 kpl, joten viiden vuoden aikana heinäkuun päiviä on $5 \cdot 31 = 155$ (kpl). Viiden vuoden aikana oli heinäkuisia sadepäiviä 9 + 7 + 7 + 10 + 8 = 41 (kpl).

Lasketaan sadepäivien suhteellinen frekvenssi

$$\frac{41}{155} \cdot 100\ \% = 26{,}451...\ \%.$$

Siis P(sataa) = 0,26451… ≈ 0,265.

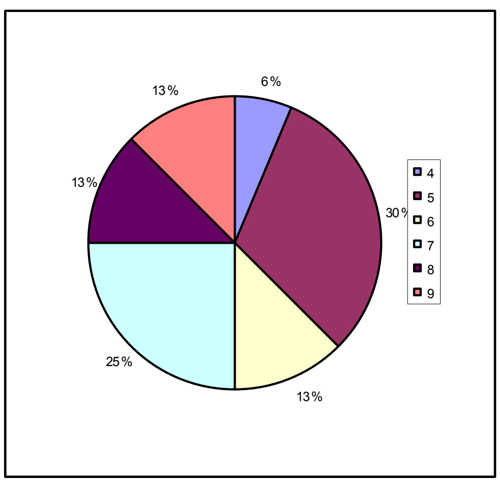
Vastaus: 0,265

17. **a)** Muodostetaan frekvenssijakaumat ja määritetään moodi.

Lyönnit (lkm)	***f***	***f* %**
4	1	$\frac{1}{16} = 0{,}0625 = 6{,}25\%$
5	5	$\frac{5}{16} = 0{,}3125 = 31{,}25\%$
6	2	12,5 %
7	4	25 %
8	2	12,5 %
9	2	12,5 %
Yhteensä	16	

Moodi on se muuttujan arvo, jolla on suurin frekvenssi. Tyypillisimmin reikä pelataan viidellä lyönnillä eli Mo = 5.

b) Suhteellisia frekvenssejä havainnollistetaan sektoridiagrammilla.



18. Luokitellaan aineisto esimerkiksi neljään luokkaan ja laaditaan frekvenssijakaumat.

Esimerkiksi:

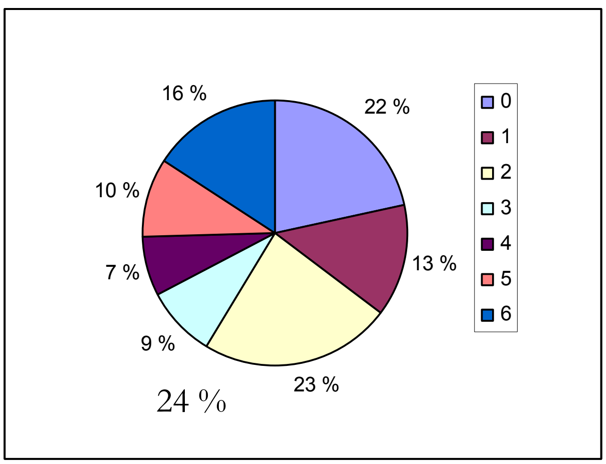
Lauman leijonat (lkm)	***f***	***f* %**
5–8	7	$\frac{7}{22} = 0{,}31818.. \approx 32\%$
9–12	7	$\frac{7}{22} = 0{,}31818.. \approx 32\%$
13–16	5	$\frac{5}{22} = 0{,}22727.. \approx 23\%$
17–20	3	$\frac{3}{22} = 0{,}13636.. \approx 14\%$
Yhteensä	22	

Taulukon luokkarajat ovat todelliset luokkarajat, koska leijonien lukumäärä on aina kokonaisluku (pyöristyksiä ja mittaustarkkuutta ei tarvitse ottaa huomioon).
Tyypillisimmin (seitsemässä laumassa) leijonia oli 5–8 tai 9–12 kpl. Näitä kutsutaan moodiluokiksi. Näin ollen

$$\mathrm{Mo} = \frac{5+8}{2} = 6{,}5 \text{ tai } \mathrm{Mo} = \frac{9+12}{2} = 10{,}5$$

19. Laaditaan tehtävästä 8 saadun pistemäärän frekvenssijakaumat.

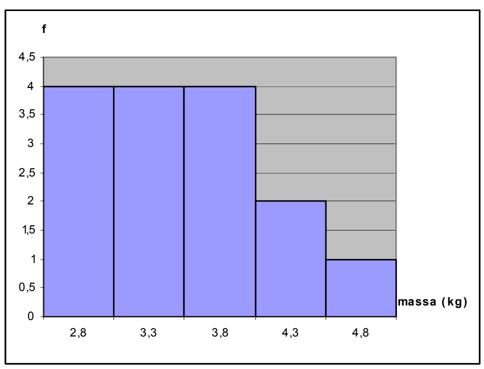
Pistemäärä	***f***	***f* %**
0	2025	$\frac{2025}{9300} = 0{,}21774... \approx 22\%$
1	1245	$\frac{1245}{9300} = 0{,}13387... \approx 13\%$
2	2186	24 %
3	792	8,5 %
4	673	7,2 %
5	908	9,8 %
6	1471	16 %
Yhteensä	9300	



20. **a)** Luokitellaan aineisto viiteen luokkaan, laaditaan frekvenssijakaumat ja lasketaan samaan taulukkoon luokkakeskukset.

Massa (kg)	***f***	***f* %**	**Luokkakeskus (kg)**
2,6–3,0	4	$\frac{4}{15} = 0{,}2666... \approx 27\%$	$\frac{2{,}55 + 3{,}05}{2} = 2{,}8$
3,1–3,5	4	27 %	$\frac{3{,}05 + 3{,}55}{2} = 3{,}3$
3,6–4,0	4	27 %	3,8
4,1–4,5	2	$\frac{2}{15} = 0{,}1333... \approx 13\%$	4,3
4,6–5,0	1	7 %	4,8
Yhteensä	15		

b) Massa on jatkuva muuttuja, joten havainnollistetaan esimerkiksi frekvenssejä histogrammilla. Pylvään keskikohdalle merkitään luokkakeskus.



21. Veripalvelun tilaston mukaan suomalaisia on

$$2{,}01 + 0{,}81 + 1{,}52 + 0{,}38 + 0{,}27 + 0{,}11 + 0{,}27 + 0{,}05 = 5{,}42 \text{ (milj.)}$$

a) Lasketaan veriryhmän AB suhteellinen frekvenssi.

$$\frac{0{,}38 + 0{,}05}{5{,}42} \cdot 100\% = 7{,}9335...\%$$

Siis $P(\mathrm{AB}) = 0{,}079335\ldots \approx 0{,}079$

b) Lasketaan Rhesus + ominaisuuden suhteellinen frekvenssi.

$$\frac{2{,}01 + 0{,}81 + 1{,}52 + 0{,}38}{5{,}42} \cdot 100\% = 87{,}084...\%$$

Siis $P(\text{Rhesus+}) = 0{,}87084\ldots \approx 0{,}87$.

Vastaus: a) 0,079 b) 0,87

1.2 Summafrekvenssit

22. Muodostetaan taulukon perusteella kaikki frekvenssijakaumat.

Myöhästymis-kerrat	*f*	*f* %	*sf*	*sf* %
0	7	$\frac{7}{26} = 0{,}269... \approx 27\,\%$	7	$\frac{7}{26} = 0{,}269... \approx 27\,\%$
1	8	$\frac{8}{26} = 0{,}307... \approx 31\,\%$	15	$\frac{15}{26} = 0{,}576... \approx 58\,\%$
2	6	$\frac{6}{26} = 0{,}230... \approx 23\,\%$	21	$\frac{21}{26} = 0{,}807... \approx 81\,\%$
3	2	$\frac{2}{26} = 0{,}0769... \approx 8\,\%$	23	$\frac{23}{26} = 0{,}884... \approx 88\,\%$
4	2	$\frac{2}{26} = 0{,}076... \approx 8\,\%$	25	$\frac{25}{26} = 0{,}961... \approx 96\,\%$
5	1	$\frac{1}{26} = 0{,}038... \approx 4\,\%$	26	$\frac{26}{26} = 1 = 100\,\%$

a) Suhteellinen frekvenssi 0 myöhästymiskerralle on 27 %.

b) Summafrekvenssi 1 myöhästymiskerran kohdalla on 15. Tämä tarkoittaa, että 15 opiskelijaa myöhästyi kerran tai ei ollenkaan. Koska opiskelijoita oli yhteensä 26 kappaletta, vähintään kahtena aamuna myöhästyi

$26 - 15 = 11$ opiskelijaa.

c) Suhteellinen summafrekvenssi 3 myöhästymiskerran kohdalla on 88 %. Opiskelijoista myöhästyi enintään kolmena aamuna (0, 1, 2 tai 3 aamuna) 88 %.

Vastaus: a) 27% b) 11 c) 88%

23. **a)** Lasketaan summafrekvenssit ja suhteelliset summafrekvenssit.

Tietokone (lkm)	***f***	***sf***	***sf*** %
0	8	8	$\frac{8}{861} = 0{,}009291... \approx 1\%$
1	191	8 + 191 = 199	$\frac{199}{861} = 0{,}23111... \approx 23\%$
2	462	199 + 462 = 661	$\frac{661}{861} = 0{,}7677... \approx 77\%$
3	136	797	93 %
4	58	855	99 %
5	6	861	100 %

b) Suhteellinen summafrekvenssi kohdassa tietokoneita 3 kpl on 93 % eli korkeintaan kolme tietokonetta (0, 1, 2 tai 3 kpl) omistaa 93 % kotitalouksista.

c) Vähintään neljä konetta on, jos kotitaloudessa on 4 tai 5 kpl tietokoneita. Lasketaan prosenttiosuus suhteellisen summafrekvenssin avulla

$$100\ \% - 93\ \% = 7\ \%.$$

Vastaus: b) 93 % c) 7 %

24. Muodostetaan äänimäärien summafrekvenssit sekä suhteelliset summafrekvenssit.

Äänimäärä	***f***	***sf***	***sf* %**
0	1	1	$\frac{1}{59} = 0{,}01694... \approx 1{,}7\%$
2	4	1 + 4 = 5	$\frac{5}{59} = 0{,}0847... \approx 8{,}5\%$
3	7	5 + 7 = 12	$\frac{12}{59} = 0{,}2033... \approx 20\%$
4	15	12 + 15 = 27	46 %
5	18	27 + 18 = 45	76 %
6	12	57	97 %
7	2	59	100 %

a) Opiskelijoista sai enintään 4 ääntä (0, 2, 3 tai 4 ääntä) 27, joka on prosentteina 46 %.

b) Alle neljä ääntä (0, 2 tai 3 ääntä) sai 12 opiskelijaa, joka on prosentteina 20 %.

c) Suhteellisen summafrekvenssin mukaan ainakin kaksi ääntä (2, 3, 4, 5, 6 tai 7 ääntä) sai

$$100\ \% - 1{,}694\ldots\ \% = 98{,}30\ldots\% \approx 98\ \%.$$

Vastaus: a) 46 % b) 20 % c) 98 %

25. Muodostetaan aineiston avulla frekvenssijakaumat.

Luokka (cm)	***f***	***f* %**	***sf***	***sf* %**
160–164	6	$\frac{6}{18}=0{,}333...\approx 33\%$	6	$\frac{6}{18}=0{,}333...\approx 33\%$
165–169	5	$\frac{5}{18}=0{,}277...\approx 28\%$	6 + 5 = 11	$\frac{11}{18}=0{,}611...\approx 61\%$
170–174	3	$\frac{3}{18}=0{,}166...\approx 17\%$	11 + 3 = 14	$\frac{14}{18}=0{,}777...\approx 78\%$
175–179	3	17 %	17	94 %
180–184	1	6 %	18	100 %

a) Summafrekvenssin mukaan alle 170 cm pituisia on 11 oppilasta.

b) Suhteellisen summafrekvenssin mukaan pituudeltaan 180 cm tai enemmän on 100 % – 94 % = 6 %.

c) Suhteellisen summafrekvenssin mukaan pituudeltaan 165 cm tai enemmän on 100 % – 33 % = 67 %.

d) Summafrekvenssin mukaan alle 170 cm pituisia on 14 opiskelijaa.

e) Frekvenssin mukaan vähintään 170 cm, mutta alle 180 cm pituisia opiskelijoita on 3 + 3 = 6.

f) Suhteellisen frekvenssin mukaan vähintään 165 cm, mutta alle 180 cm pituisia on 28 % + 17 % + 17 % = 62 %.

g) Kaikki oppilaat ovat vähintään 160 cm pitkiä eli 100 % oppilaista on yli 160 cm.

Vastaus: a) 11 b) 6 % c) 67 % d) 14 e) 6 f) 62 % g) 18 (100 %)

26. Muodostetaan summafrekvenssi- ja suhteellinen summafrekvenssijakauma.

Aika (min)	***f***	***sf***	***sf* %**
0–59	2	2	$\frac{2}{76} = 0{,}0263... \approx 3\%$
60–119	3	2 + 3 = 5	$\frac{5}{76} = 0{,}0657... \approx 7\%$
120–179	7	5 + 7 = 12	$\frac{12}{76} = 0{,}1578... \approx 16\%$
180–239	28	40	53 %
240–299	25	65	86 %
300–360	11	76	100 %

a) Suhteellisen summafrekvenssin mukaan alle kaksi tuntia (120 min) viipyi 7 % opiskelijoista.

b) Suhteellisen summafrekvenssin mukaan yli viisi tuntia (300 min) viipyi

100 % – 86 % = 14 % opiskelijoista.

Vastaus. a) 7 %   b) 14 %

27. Täydennetään taulukko rivi kerrallaan.

1. rivi

Tiedetään, että frekvenssi *f* on 2, joten ensimmäisen rivin kohdalla myös summafrekvenssi *sf* on 2. Myös suhteellinen frekvenssi *f* % ja suhteellinen summafrekvenssi *sf* % ovat ensimmäisellä rivillä yhtä suuret.

f	***f* %**	***sf***	***sf* %**
2	10 %	2	10 %

Ensimmäisen rivin tietojen avulla voidaan laskea havaintoyksiköiden kokonaismäärä. Merkitään tätä havaintoyksiköiden määrää kirjaimella x.

Koska frekvenssin arvo 2 on havaintoyksiköiden kokonaismäärästä x 10 %, saadaan yhtälö

$$\frac{2}{x} = 0{,}1$$

$$\frac{2}{x} = \frac{1}{10} \quad |\text{kerrotaan ristiin}$$

$$x = 20$$

Havaintoyksiköiden kokonaismäärä on siis 20.

2. rivi

Tiedetään, että summafrekvenssi on 6. Tällöin rivin frekvenssi on 6 – 2 = 4.

f	f %	sf	sf %
2	10 %	2	10 %
4	$\frac{4}{20}=0{,}2=20\,\%$	6	$\frac{6}{20}=0{,}3=30\,\%$

3. rivi

Tiedetään, että suhteellinen summafrekvenssi on 80 %. Merkitään rivin summafrekvenssiä kirjaimella y.

$$\frac{y}{20}=0{,}8 \quad |\cdot 20$$
$$y=16$$

Rivin summafrekvenssi on siis 16.

f	f %	sf	sf %
2	10 %	2	10 %
4	20 %	6	30 %
16 – 6 = 10	$\frac{10}{20}=0{,}5=50\,\%$	16	80 %

4. rivi

Viimeisellä rivillä summafrekvenssi on sama kuin havaintoyksiköiden määrä eli 20. Suhteellinen summafrekvenssi on tällöin 100 %.

28. Muodostetaan frekvenssijakaumien kuvaajien perusteella summafrekvenssijakaumat. Päätellään summafrekvenssijakaumien perusteella oikea kuvaaja.

A

Viikkoraha (€)	***f***	***sf***
3–5	5	5
6–8	10	15
9–11	15	30
12–14	5	35

Tätä vastaava summafrekvenssijakauman kuvaaja on *E*.

B

Viikkoraha (€)	***f***	***sf***
3–5	15	15
6–8	10	25
9–11	6	31
12–14	4	35

Tätä vastaava summafrekvenssijakauman kuvaaja on *F*.

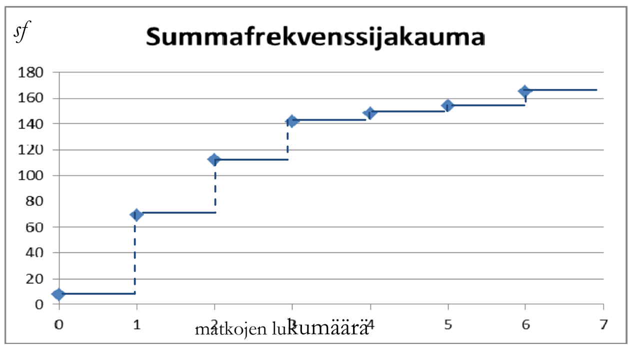
C

Viikkoraha (€)	***f***	***sf***
3–5	4	4
6–8	6	10
9–11	10	20
12–14	15	35

Tätä vastaava summafrekvenssijakauman kuvaaja on *D*.

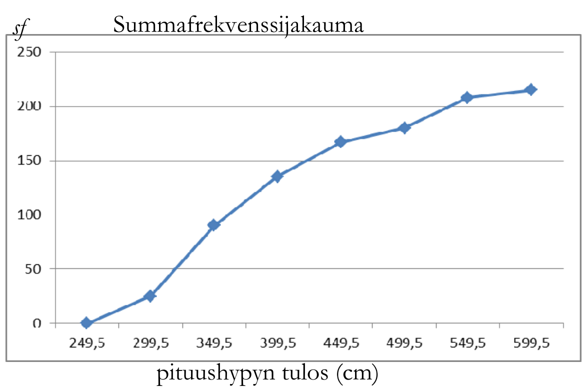
Vastaus: A ja E, B ja F, C ja D

29. **a)** Kun tarkasteltava muuttuja on diskreetti, kuten esimerkiksi lomamatkojen lukumäärä, summafrekvenssejä kuvataan porrasdiagrammilla. Summafrekvenssit merkitään kuvaajaan pisteillä, ja kertymä pysyy aina samana seuraavaan muuttujan arvoon saakka. Vaakaviivat yhdistetään katkoviivalla.



b) Kun tarkasteltava muuttuja on jatkuva, kuten esimerkiksi pituus, summafrekvenssejä kuvataan summakäyrän avulla. Kuvaaja alkaa nollatasolta. Summafrekvenssit merkitään kuvaajaan pisteillä todellisten ylärajojen kohdalle ja pisteet yhdistetään murtoviivalla.

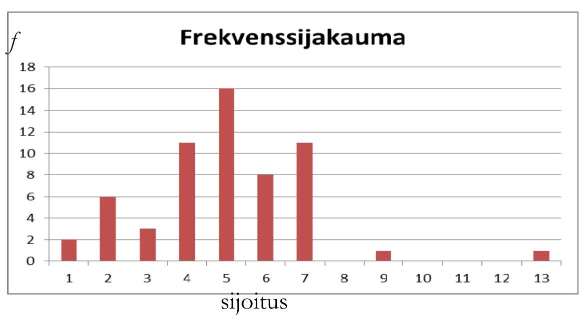
Ensimmäisen luokan alarajalla summafrekvenssi on siis 0, mutta luokan ylärajaan (299,5 cm) mennessä on kertynyt jo 25 havaintoyksikköä jne.



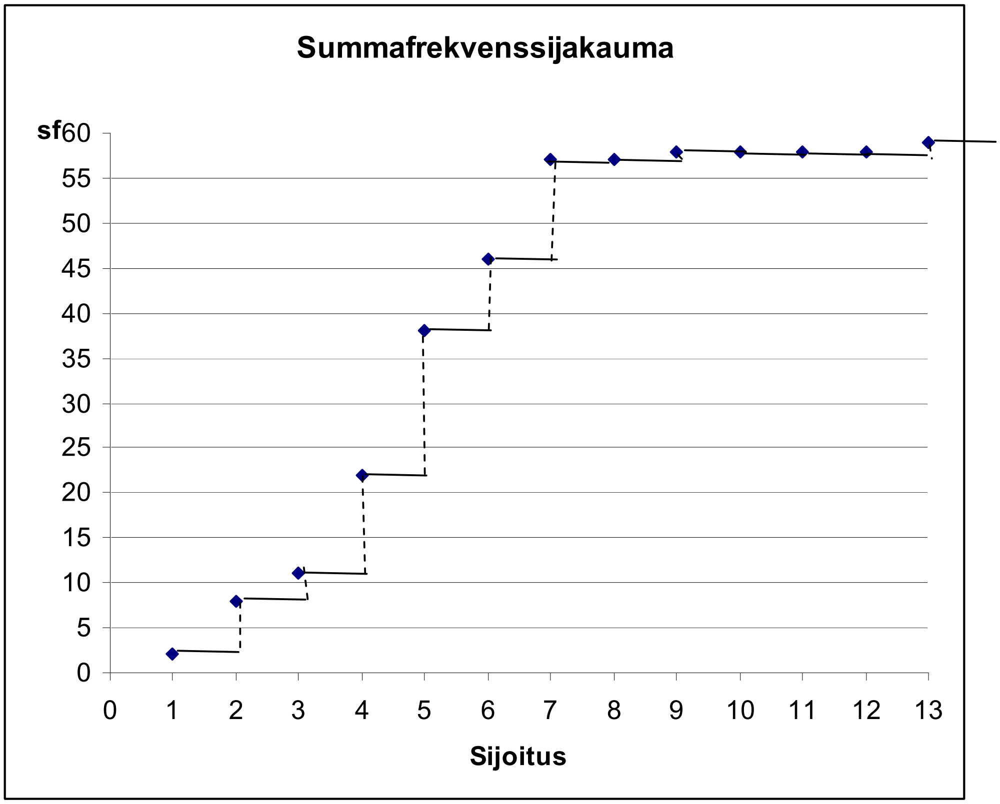
30. Muodostetaan frekvenssi- ja summafrekvenssijakaumat.

Sijoitus	*f*	*sf*
1	2	2
2	6	2 + 6 = 8
3	3	8 + 3 = 11
4	11	11 + 11 = 22
5	16	38
6	8	46
7	11	57
8	0	57
9	1	58
10	0	58
11	0	58
12	0	58
13	1	59

a) Koska kuvattava muuttuja on diskreetti, frekvenssijakaumaa voidaan kuvata pylväsdiagrammilla.



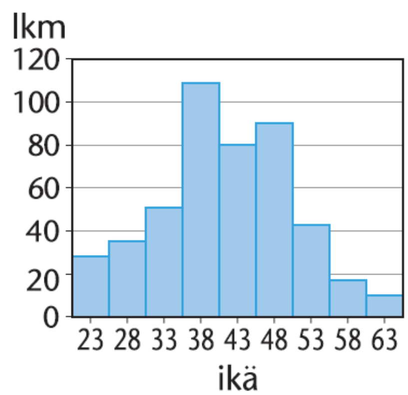
b) Koska kuvattava muuttuja on diskreetti, summafrekvenssijakaumaa kuvataan porrasdiagrammilla. Summafrekvenssit merkitään kuvaajaan pisteillä, ja kertymä pysyy samana aina seuraavaan muuttujan arvoon (sijoitukseen) saakka.



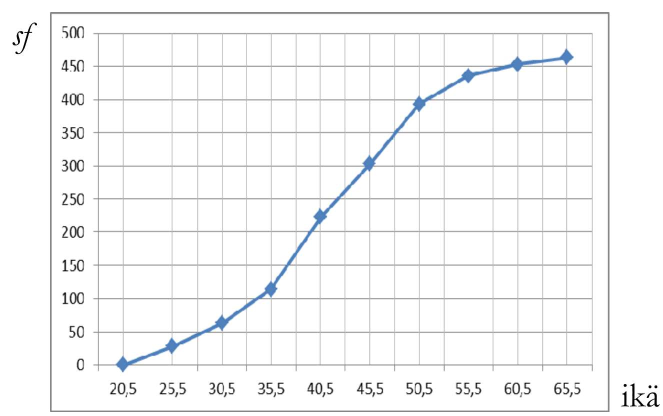
31. Muodostetaan taulukon perusteella sekä summafrekvenssijakauma että suhteellinen summafrekvenssijakauma.

Ikä	**Luokka-keskus**	***f***	***sf***	***sf* %**
21–25	$\frac{20{,}5+25{,}5}{2}=23$	28	28	$\frac{28}{454}=0{,}062\ldots\approx 6\,\%$
26–30	28	35	28 + 35 = 63	$\frac{63}{454}=0{,}139\ldots\approx 14\,\%$
31–35	33	51	63 + 51 = 114	$\frac{114}{454}=0{,}251\ldots\approx 25\,\%$
36–40	38	109	223	$\frac{223}{454}=0{,}491\ldots\approx 49\,\%$
41–45	43	80	303	$\frac{303}{454}=0{,}667\ldots\approx 67\,\%$
46–50	48	90	393	$\frac{393}{454}=0{,}866\ldots\approx 87\,\%$
51–55	53	43	436	$\frac{436}{454}=0{,}960\ldots\approx 96\,\%$
56–60	58	17	453	$\frac{453}{454}=0{,}998\ldots\approx 100\,\%$
61–65	63	1	454	$\frac{454}{454}=100\,\%$

a) Ikä on jatkuva muuttuja, joten piirretään frekvenssijakaumasta histogrammi. Histogrammin voi piirtää myös suhteellisten frekvenssien avulla. Pylväiden keskikohdalle merkitään luokkakeskukset.



b) Piirretään summakäyrä summafrekvenssijakaumasta. Summakäyrän voi piirtää myös suhteellisesta summafrekvenssijakaumasta. Jakauman muoto on tällöin sama. Merkitään summafrekvenssipisteet luokkien todellisten ylärajojen kohdalle.



32. Moodi on se muuttujan arvo, jolla on suurin frekvenssi. Koska arvosanoja 9 on eniten (3 kpl), niin Mo = 9.

Mediaani on suuruusjärjestykseen laitetun aineiston keskimmäinen arvo. Laitetaan kurssiarvosanat suuruusjärjestykseen

7 8 8 9 9 9.

Koska keskimmäisiä arvosanoja on kaksi, mediaani on näiden keskiarvo

$$\frac{8+9}{2} = 8{,}5$$

Siis Md = 8,5.

Vastaus: Mo = 9, Md = 8,5

33. **a)** Moodi on se kuukausi, jolloin oli eniten yöpymisiä eli 13 % kaikista yöpymisistä. Mo = 7 (heinäkuu).

b) Mediaani on se kuukausi, jonka aikana on täyttynyt puolet (50 %) koko vuoden yöpymisistä. Lasketaan kuvaajan pohjalta suhteellinen summafrekvenssijakauma ja katsotaan mediaani siitä.

Kuukausi	***sf* %**
1	12 %
2	12 % + 8 % = 20 %
3	20 % + 8 % = 28 %
4	28 % + 6 % = 34 %
5	40 %
6	49 %
7	62 %
8	74 %
9	80 %
10	85 %
11	91 %
12	100 %

Puolet vuotuisista yöpymisistä ylittyy heinäkuun aikana eli Md = 7 (heinäkuu).

c) Joulukuussa kertyi 10 % kaikista yöpymisistä eli

$$0{,}10 \cdot 4703827 = 470383$$

Vastaus: a) Mo = 7 b) Md = 7 c) 470 383

34. **a)** Tyypillisimmin hiuksia pestään kaksi kertaa viikossa, koska näin tekee 32 havaintoyksikköä. Siis Mo = 2.

Mediaani on se pesujen määrä viikossa, jolloin aineiston suhteellinen summafrekvenssi ylittää ensimmäisen kerran 50 %. Näin tapahtuu, kun hiusten pesukertoja on 3 viikossa. Siis Md = 3.

b) Havaintoaineiston tyypillisin kuukausipalkka kuuluu joko luokkaan 2000–2499 (€) tai 3000–3499 (€). Näihin palkkaluokkiin kuuluu molempiin 42 havaintoyksikköä. Koska palkka on luokiteltu muuttujan arvo, niin moodit ovat näiden luokkien luokkakeskukset.

$$\text{Mo} = \frac{1999{,}5 + 2499{,}5}{2} = 2249{,}5\,(€) \text{ tai } \text{Mo} = \frac{2999{,}5 + 3499{,}5}{2} = 3249{,}5\,(€)$$

Puolet (50 %) koko aineiston havaintoyksiköistä on $\frac{188}{2} = 94$. Katsotaan aineistosta, mihin luokkaan mennessä summafrekvenssi ylittää ensimmäisen kerran 94. Tämä tapahtuu palkkaluokan 2500–2999 (€) kohdalla. Luokitellun aineiston mediaani on luokan luokkakeskus.

$$\text{Md} = \frac{2499{,}5 + 2999{,}5}{2} = 2749{,}5\ (€)$$

Vastaus: a) Mo = 2, Md = 3
b) Mo = 2249,50 € tai Mo = 3249,50 €, Md = 2749,50 €

35. Frekvenssijakauman avulla voidaan määrittää aineiston moodi. Suhteellisen summafrekvenssijakauman avulla voidaan määrittää aineiston mediaani.

Mökkien lukumäärä	***f***	***sf***	***sf*%**
0–99	12	12	$\frac{12}{177} = 0{,}067... \approx 7\%$
100–199	16	12 + 16 = 28	$\frac{28}{177} = 0{,}1581... \approx 16\%$
200–299	23	51	29 %
300–399	25	76	43 %
400–499	23	99	56 %
500–599	17	116	66 %
600–699	6	122	69 %
700–799	25	147	83 %
800–899	13	160	90 %
900–999	17	177	100 %

Moodi on se muuttujan arvo, jolla on suurin frekvenssi. Moodi on tässä aineistossa se yleisin kunnissa oleva mokkien lukumäärä.

Kahdessakymmenessä viidessä kunnassa mökkejä on 300–399 (kpl) tai 700–799 (kpl). Luokitellun aineiston moodit ovat luokkien luokkakeskukset. Taulukon luokkarajat ovat todelliset, koska mökkien määrä on aina kokonaisluku. Moodina on siis

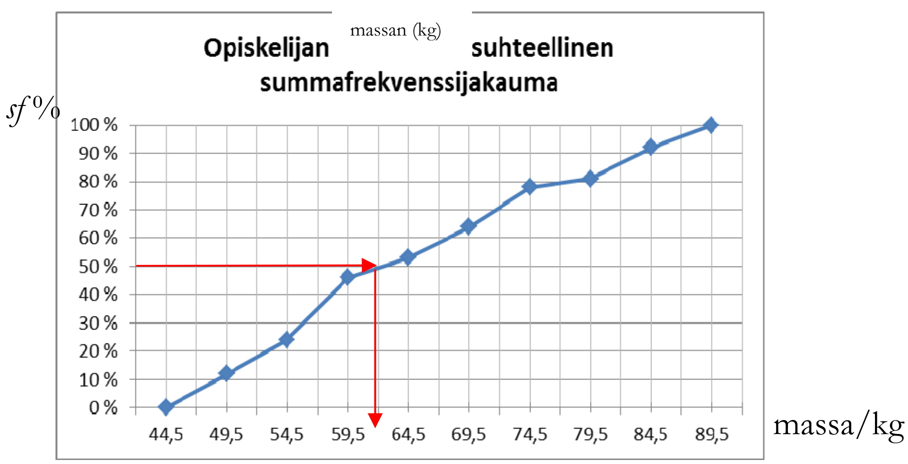
$$\text{Mo} = \frac{300 + 399}{2} = 349{,}5 \ \text{(mökkiä)} \quad \text{tai} \quad \text{Mo} = \frac{700 + 799}{2} = 749{,}5 \ \text{(mökkiä)}.$$

Mediaani on se luokka, jonka suhteellinen summafrekvenssi on ensimmäisenä vähintään 50 %. Tämä ylittyy luokassa 400–499 (mökkiä). Puolet kunnista on siis sellaisia, jossa mökkien lukumäärä on korkeintaan 400–499 (mökkiä). Tämä on mediaaniluokka ja mediaanina käytetään luokan luokkakeskusta

$$\mathrm{Md} = \frac{400+499}{2} = 449{,}5 \text{ (mökkiä).}$$

Vastaus: Mo = 749,5 mökkiä tai Mo = 349,5 mökkiä, tyypillisin mökkien määrä kunnassa. Md = 449,5 mökkiä, puolessa kunnista on korkeintaan tämän verran mökkejä.

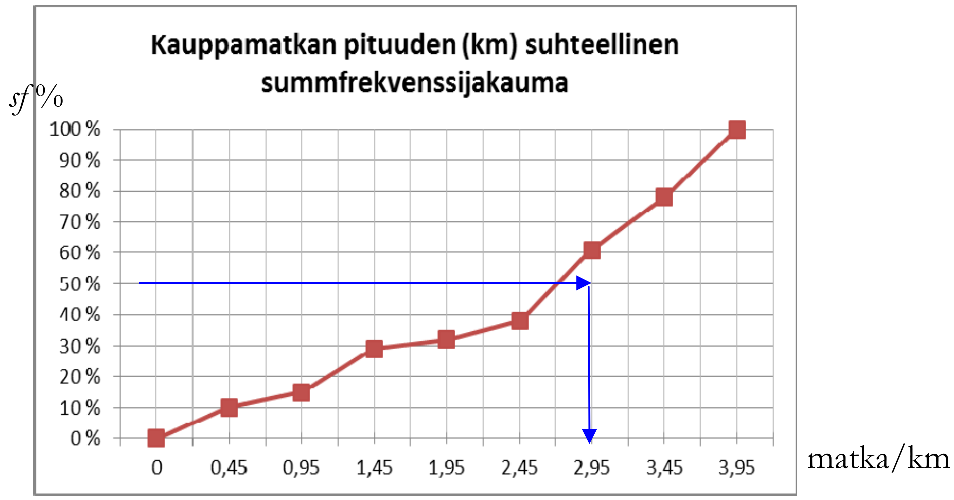
36. a)



Suuntaa antava likiarvo mediaanille luetaan kuvaajasta todellisten ylärajojen 59,5 (kg) ja 64,5 (kg) puolivälistä eli

$$\frac{59{,}5+64{,}5}{2} = 62 \text{ (kg)}$$

b)



Suuntaa antava likiarvo mediaanille luetaan kuvaajasta todellisten ylärajojen 2,45 (km) ja 2,95 (km) puolivälistä eli

$$\frac{2,45+2,95}{2}=2,7\,(\text{km})$$

Vastaus: a) Md = 62 kg b) Md = 2,7 km

37. Muodostetaan taulukon perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Kurssimäärä	***f***	***sf***	***sf* %**
75–77	87	87	$\frac{87}{154} = 0{,}564\ldots \approx$ 56 %
78–80	34	87 + 34 = 121	$\frac{121}{154} = 0{,}785\ldots \approx$ 79 %
81–83	17	121 + 17 = 138	$\frac{138}{154} = 0{,}896\ldots \approx$ 90 %
84–86	9	147	0,954… ≈ 95 %
87–89	5	152	0,987… ≈ 99 %
90–92	2	154	100 %

a) Suhteellisen summafrekvenssijakauman perusteella mediaaniluokka on 75–77 kurssia. Mediaaniluokan luokkakeskus on 76 kursia. (Luokkarajat ovat todellisia rajoja, koska pääsääntöisesti kurssimäärät ovat kokonaislukuja.)

Tässä aineistossa tämä tarkoittaa, että puolet opiskelijoista suorittaa alle 76 kurssia ja puolet yli 76 kurssia.

b) Frekvenssijakauman perustella moodiluokka on 75–77 kurssia.

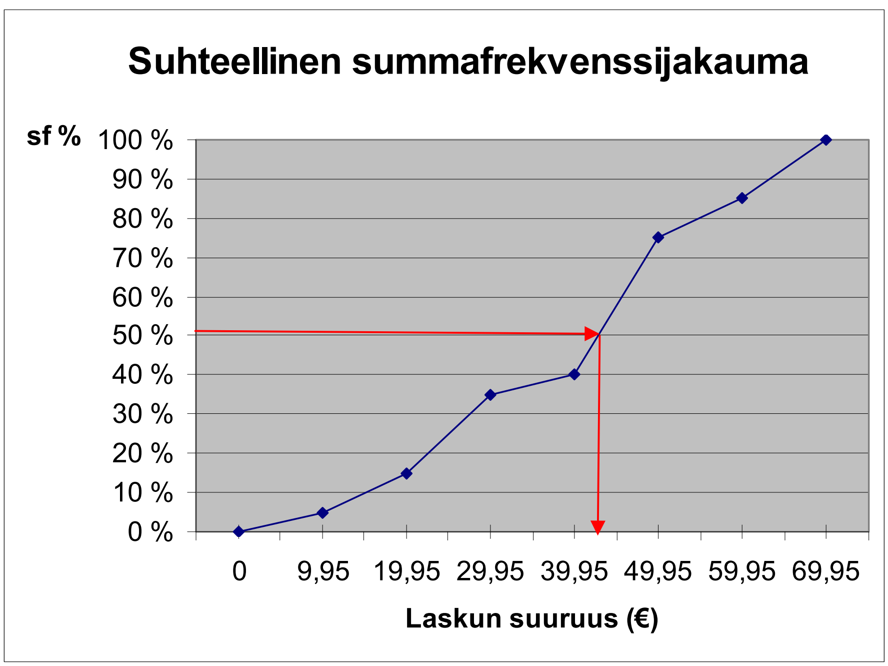
Tässä aineistossa tämä tarkoittaa, että suurin osa opiskelijoista suorittaa 75–77 kurssia.

Vastaus: a) 75–77 kurssia b) 75–77 kurssia

38. Lasketaan luokkien todelliset ylärajat ja piirretään summafrekvenssijakauman kuvaaja.

Puhelinlaskun suuruus (€)	**Todellinen yläraja (€)**	***sf* %**
0–9,9	9,95	5 %
10,0–19,9	19,95	15 %
20,0–29,9	29,95	35 %
30,0–39,9	39,95	40 %
40,0–49,9	49,95	75 %
50,0–59,9	59,95	85 %
60,0–69,9	69,95	100 %

Suhteelliset summafrekvenssipisteet merkitään luokkien todellisten ylärajojen kohdalle ja pisteet yhdistetään murtoviivalla. Jakauman kuvaaja alkaa nollatasolta.



Kuvaajasta voidaan saada mediaanille suuntaa antava likiarvo.

Todellisten ylärajojen puoliväli on

$$\frac{39,95+49,95}{2}=44,95\approx 45\ (€)$$

Mediaani näyttäisi olevan noin

$$\frac{39,95+44,95}{2}=42,45\approx 42\ (€)$$

Vastaus: Md ≈ 42 €

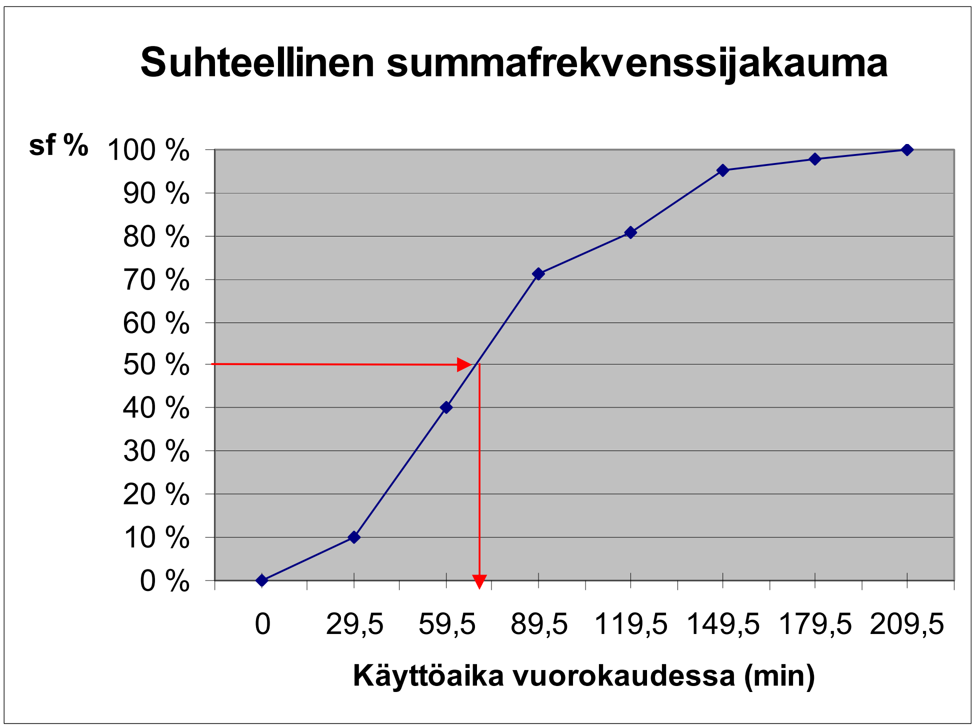
39. Muodostetaan suhteellinen summafrekvenssijakauma.

Tietokoneen käyttöaika vuorokaudessa (min)	***f***	***sf***	***sf* %**
0–29	42	42	$\frac{42}{412}=0{,}1019...\approx 10\%$
30–59	124	42 + 124 = 166	$\frac{166}{412}=0{,}4029...\approx 40\%$
60–89	126	166 + 126 = 292	$\frac{292}{412}=0{,}7087...\approx 71\%$
90–119	41	333	81 %
120–149	60	393	95 %
150–179	12	405	98 %
180–209	7	412	100 %

Luokan 60–89 (min) kohdalla suhteellinen summafrekvenssi ylittää ensimmäisen kerran 50 %. Mediaani on luokan luokkakeskus. Käyttöaika on jatkuva muuttuja, ja luokkakeskus saadaan todellisten luokkarajojen keskiarvona. Siis

$$\text{Md} = \frac{59{,}5 + 89{,}5}{2} = 74{,}5 \ \text{(min)}$$

b) Suhteelliset summafrekvenssipisteet merkitään luokkien todellisten ylärajojen kohdalle ja pisteet yhdistetään murtoviivalla. Jakauman kuvaaja alkaa nollatasolta.



Kuvaajasta voidaan saada mediaanille suuntaa antava likiarvo. Todellisten ylärajojen puoliväli on

$$\frac{59{,}5 + 89{,}5}{2} = 74{,}5 \ \text{(min)}$$

Mediaani näyttäisi olevan hieman tätä pienempi, mutta suurempi kuin $\frac{59{,}5+74{,}5}{2}=67\,(\text{min})$. Esimerkiksi Md ≈ 72 min.

Vastaus: a) Md = 74,5 min b) Md ≈ 72 min

40. Muodostetaan frekvenssijakaumat.

Lyönnit/ peli (lkm)	***f***	***sf***	***f* %**	***sf* %**
40	19	19	$\frac{19}{68}=0{,}279... \approx 28\%$	$\frac{19}{68}=0{,}279... \approx 28\%$
41	2	19 + 2 = 21	$\frac{2}{68}=0{,}0294... \approx 2{,}9\%$	$\frac{21}{68}=0{,}308... \approx 31\%$
42	14	21 + 14 = 35	$\frac{14}{68}=0{,}2058... \approx 21\%$	$\frac{35}{68}=0{,}5147... \approx 51\%$
43	5	40	7 %	59 %
44	9	49	13 %	72 %
45	13	62	19 %	91 %
46	6	68	9 %	100 %

a) Summafrekvenssin mukaan amatööri pelasi <u>alle</u> 45 lyönnin pelejä 49 kpl.

b) Suhteellisen summafrekvenssin mukaan <u>korkeintaan</u> 42 lyönnillä (40, 41 tai 42 lyöntiä) selvittiin 51 % peleistä.

c) Suhteellisen summafrekvenssin mukaan <u>yli</u> 43 lyönnin pelejä oli

100 % – 59 % = 41 %.

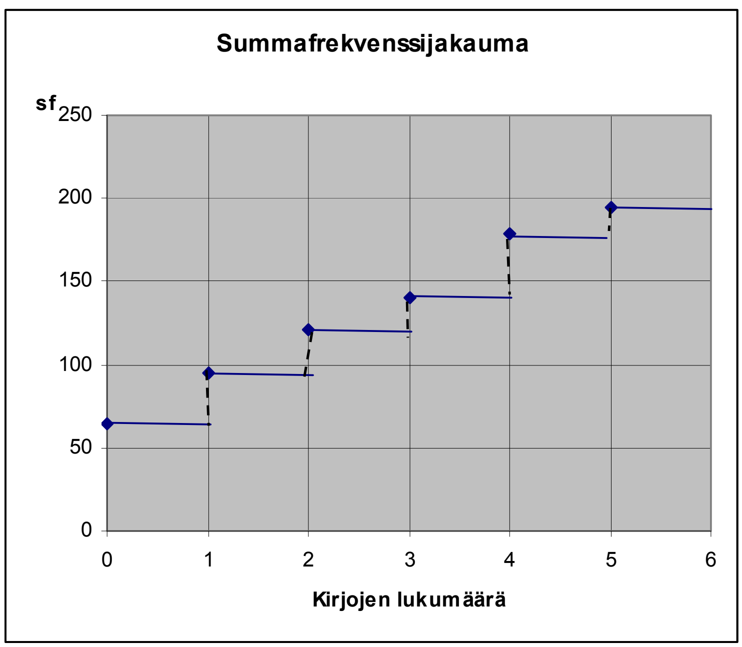
Vastaus: a) 49 peliä b) 51 % c) 41 %

41. Muodostetaan ensin summafrekvenssijakauma.

a)

Kuukaudessa luettujen kirjojen lukumäärä	*f*	*sf*
0	65	65
1	30	65 + 30 = 95
2	26	95 + 26 = 121
3	19	140
4	39	179
5	16	195

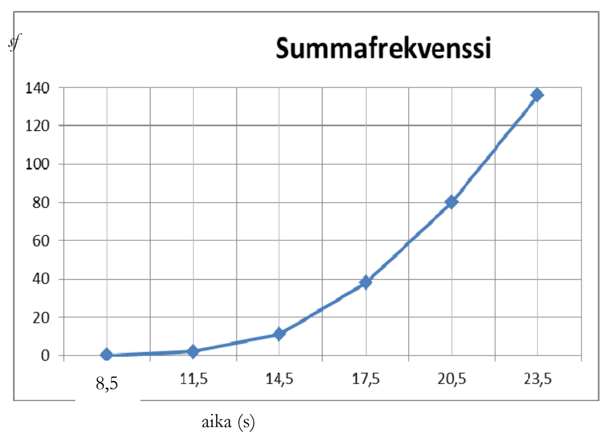
Koska tarkasteltava muuttuja (kirjojen lukumäärä) on diskreetti, summafrekvenssejä kuvataan porrasdiagrammilla. Summafrekvenssit merkitään kuvaan pisteillä, ja kertymä pysyy samana seuraavaan muuttujan arvoon saakka.



b) Muodostetaan ensin summafrekvenssijakauma. Lasketaan kuvaajan piirtämistä varten luokkien todelliset ylärajat samaan taulukkoon.

100 m juoksuun kuluva aika (s)	***f***	***sf***	**Todellinen yläraja (s)**
9–11	2	2	11,5
12–14	9	2 + 9 = 11	14,5
15–17	27	11 + 27 = 38	17,5
18–20	42	80	20,5
21–23	56	136	23,5

Koska muuttuja on jatkuva (aika), summafrekvenssejä kuvataan summakäyrällä. Summafrekvenssipisteet merkitään kuvaajissa todellisten ylärajojen kohdalle ja pisteet yhdistetään murtoviivalla. Jakauman kuvaaja alkaa nollatasolta.



42. Muodostetaan frekvenssijakaumat. Suhteellisesta summafrekvenssijakaumasta nähdään, mihin kellonaikaan mennessä puolen päivän asiakkaista oli käynyt kaupassa.

Aukioloaika (h)	***f***	***sf***	***sf*%**
1 (klo 10)	2	2	$\frac{2}{84} = 0{,}0238... \approx 2{,}4\%$
2 (klo 11)	11	2 + 11 = 13	$\frac{13}{84} = 0{,}1547... \approx 15\%$
3 (klo 12)	6	13 + 6 = 19	$\frac{19}{84} = 0{,}2261... \approx 23\%$
4 (klo 13)	14	33	39 %
5 (klo 14)	15	48	57 %
6 (klo 15)	8	56	67 %
7 (klo 16)	7	63	75 %
8 (klo 17)	17	80	95 %
9 (klo 18)	4	84	100 %

Puolet päivän asiakkaista oli siis käynyt kaupassa klo 14 mennessä. Siis mediaani on klo 14.

Vastaus: Mediaani klo 14; puolet päivän asiakkaista oli käynyt kaupassa

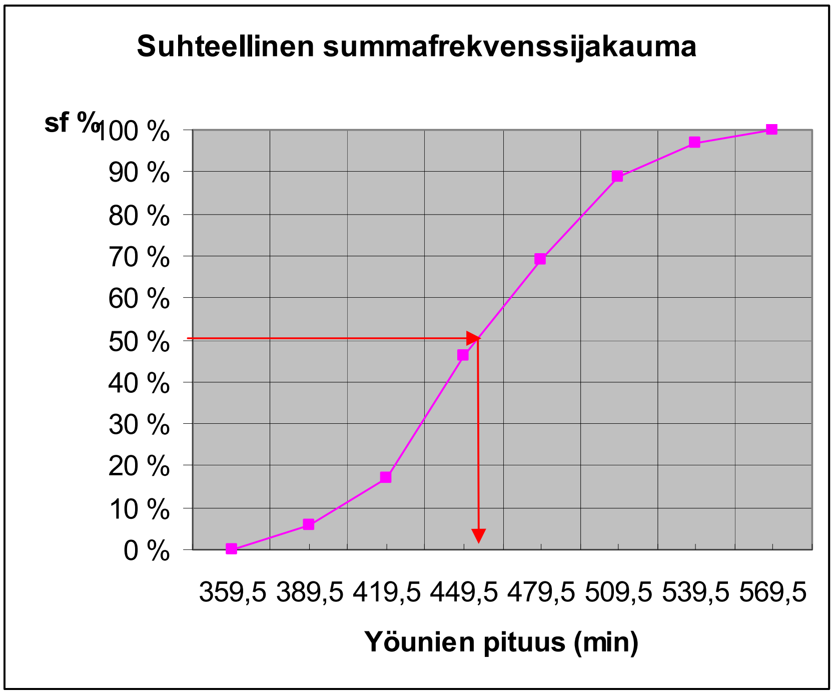
43. **a)** Muodostetaan suhteellinen summafrekvenssijakauma.

Yöunien pituus (min)	***f***	***sf***	***sf %***
360–389	2	2	$\frac{2}{35}=0{,}0571...\approx 6\ \%$
390–419	4	2 + 4 = 6	$\frac{6}{35}=0{,}1714...\approx 17\ \%$
420–449	10	6 + 10 = 16	$\frac{16}{35}=0{,}4571...\approx 46\ \%$
450–479	8	24	69 %
480–509	7	31	89 %
510–539	3	34	97 %
540–569	1	35	100 %

Suhteellisen summafrekvenssin arvo ylittyy ensimmäisen kerran luokassa 450–479 (min). Luokitellun aineiston mediaani on luokan 450–479 (min) luokkakeskus.

$$\text{Md} = \frac{449{,}5 + 479{,}5}{2} = 464{,}5\ \ (\text{min})$$

b) Koska muuttuja on jatkuva (aika), Suhteellisia summafrekvenssejä kuvataan summakäyrällä. Suhteelliset summafrekvenssipisteet merkitään kuvaajissa todellisten ylärajojen kohdalle ja pisteet yhdistetään murtoviivalla. Jakauman kuvaaja alkaa nollatasolta.



Kuvaajasta voidaan saada mediaanille suuntaa antava likiarvo. Todellisten ylärajojen puoliväli on

$$\frac{449{,}5 + 479{,}5}{2} = 464{,}5 \text{ (min)}$$

Mediaani näyttäisi olevan noin

$$\frac{449{,}5 + 464{,}5}{2} = 457 \text{ (min).}$$

Siis Md ≈ 457 min.

Vastaus: a) Md = 464,5 min b) Md ≈ 457 min

1.3 Keskiarvo

44. **a)** Lasketaan ruotsalaisten turistien määrän keskiarvo.

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^{5} x_i}{5} \\ &= \frac{570640 + 581477 + ... + 598851}{5} \\ &= \frac{3022441}{5} \\ &= 604488{,}2 \\ &\approx 604488\end{aligned}$$

b) Lasketaan saksalaisten turistien määrän keskiarvo.

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^{5} x_i}{5} \\ &= \frac{458\,235 + 483\,507 + ... + 513\,703}{5} \\ &= \frac{2\,422\,435}{5} \\ &= 484487\end{aligned}$$

Vastaus: a) 604 488 b) 484 487

45. a) Viisivuotiskauden keskimääräinen työttömyysaste on

$$\bar{x} = \frac{\sum_{i=1}^{5} x_i}{5}$$
$$= \frac{8{,}9 + 8{,}1 + 9{,}8 + 11{,}6 + 11{,}3}{5} = \frac{49{,}7}{5} = 9{,}94 \ \ (\%)$$

(Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

b) Työttömyysaste olisi 2012 ennusteen mukaan 10 % pienempi kuin vuoden 2011 työttömyysaste (11,3 %). Työttömyysaste 2012 olisi siis

$$0{,}9 \cdot 11{,}3 = 10{,}17 \ \ (\%)$$

Keskimääräinen työttömyysaste 2007–2012 olisi siis

$$\bar{x} = \frac{\sum_{i=1}^{6} x_i}{6}$$
$$= \frac{8{,}9 + 8{,}1 + 9{,}8 + 11{,}6 + 11{,}3 + 10{,}17}{6} = 9{,}978... \approx 10{,}0 \ (\%)$$

(Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Vastaus: a) 9,94 % b) 10,0 %

46. Muodostetaan heittojen tulosten frekvenssijakauma ja lasketaan sen avulla keskiarvo.

Heiton tulos	***f***
3	1
4	1
5	2
6	3
7	1
8	3
9	5
10	2

Heittojen keskiarvo on

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^{8} f_i x_i}{18} \\ &= \frac{1\cdot 3 + 1\cdot 4 + 2\cdot 5 + 3\cdot 6 + 1\cdot 7 + 3\cdot 8 + 5\cdot 9 + 2\cdot 10}{6} \\ &= 7{,}27\ldots \approx 7{,}3\end{aligned}$$

Koska 7,3 < 8,1, niin Anssi ei voita Akua.

(Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Vastaus: Anssi ei voita Akua.

47. Muodostetaan frekvenssijakauman perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Poissaolo-tunteja *x*	***f***	***sf***	***sf* %**
0	8	8	$\frac{8}{32} = 0{,}25 = 25\,\%$
1	5	8 + 5 = 13	$\frac{13}{32} = 0{,}406... \approx 41\,\%$
2	6	13 + 6 = 19	0,593… ≈ 59 %
3	2	21	0,656… ≈ 66 %
4	5	26	0,812… ≈ 81 %
5	2	28	0,875 ≈ 88 %
6	2	30	0,937… ≈ 94 %
7	0	30	94 %
8	1	31	0,968… ≈ 97 %
9	1	32	100 %

a) Frekvenssijakauman perusteella moodi on 0 tuntia, koska 0 poissaolotunnin frekvenssi on suurin.

Suhteellisen summafrekvenssijakauman perusteella 50 % ylittyy ensimmäisen kerran poissaolotuntien 2 kohdalla, joten 2 tuntia on mediaani.

b) Poissaolotuntien keskiarvo oli

$$\bar{x} = \frac{\sum_{i=1}^{10} f_i x_i}{32}$$

$$= \frac{8 \cdot 0 + 5 \cdot 1 + 6 \cdot 2 + ... + 1 \cdot 9}{32} = \frac{82}{32} = 2{,}5625 \approx 2{,}6 \quad (\text{h})$$

(Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Vastaus: a) Mo = 0 h, Md = 2 h   b) 2,6 h

48. Lasketaan keskimääräinen kasvuprosentti luokitellun aineiston keskiarvona. Tätä varten tarvitaan luokkien luokkakeskukset. Lasketaan samalla taulukkoon frekvenssin ja luokkakeskuksen tulo.

Keskiarvon voi laskea laskimen tilastotoiminnoilla tai taulukoimalla.

Seuraavassa on esitetty taulukointiratkaisu.

Kasvu-prosentti (%)	f_i	**Luokka-keskus** x_i	$f_i x_i$
0,1–1,0	2	$\frac{0{,}05+1{,}05}{2}=0{,}55$	$2\cdot 0{,}55=1{,}1$
1,1–2,0	5	$\frac{1{,}05+2{,}05}{2}=1{,}55$	$5\cdot 1{,}55=7{,}75$
2,1–3,0	25	2,55	63,75
3,1–4,0	18	3,55	63,9
4,1–5,0	1	4,55	4,55
5,1–6,0	1	5,55	5,55

Tilastossa on valtioita yhteensä 2 + 5 + 25 + 18 + 1 + 1 = 52.

Kasvuprosentin keskiarvo on

$$\bar{x}=\frac{\sum_{i=1}^{6} f_i x_i}{52}$$

$$=\frac{1{,}1+7{,}75+63{,}75+\ldots+5{,}55}{52}=\frac{146{,}6}{52}=2{,}819\ldots\approx 2{,}8\quad(\%)$$

Vastaus: 2,8 %

49. Muodostetaan taulukon perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Ikä	***f***	***sf***	***sf* %**
0–4	291 275	291 275	$\frac{291\,275}{5\,181\,115} = 0{,}056\ldots \approx 6\ \%$
5–14	645 058	291 275 + 645 058 = 936 333	$\frac{936\,333}{5\,181\,115} = 0{,}180\ldots \approx 18\ \%$
15–64	3 467 584	936 333 + 3 467 584 = 4 403 917	85 %
65–74	436 789	4 840 706	93 %
75–84	262 014	5 102 720	98 %
85–104	78 395	5 181 115	100 %

Frekvenssijakauman perusteella moodiluokka on 15–64.

Suhteellisen summafrekvenssijakauman perusteella mediaaniluokka on 15–64.

Keskiarvon laskemiseksi tarvitaan luokkien luokkakeskukset. Lasketaan samalla taulukkoon frekvenssin ja luokkakeskuksen tulo. (Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Ikä	f_i	**Luokka-keskus** x_i	$f_i x_i$
0–4	291 275	$\frac{0+4{,}5}{2} = 2{,}25$	$291\,275 \cdot 2{,}25 = 655\,368{,}75$
5–14	645 058	$\frac{4{,}5+14{,}5}{2} = 9{,}5$	$645\,058 \cdot 9{,}5 = 6\,128\,051$
15–64	3 467 584	39,5	136 969 568
65–74	436 789	69,5	30 356 836
75–84	262 014	79,5	20 830 113
85–104	78 395	94,5	7 408 327,5

Keski-ikä on

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^{6} f_i x_i}{5\,181\,115} \\ &= \frac{655\,368{,}75 + 6\,128\,051 + \ldots + 7\,408\,327{,}5}{5\,181\,115} \\ &= \frac{202\,348\,264{,}3}{5\,181\,115} \\ &= 39{,}054\ldots \\ &\approx 39{,}1 \quad \text{(vuotta)}\end{aligned}$$

Vastaus: keski-ikä 39,1 (vuotta), mediaaniluokka 15–64, moodiluokka 15–64

50. Tarkastellaan veroastetta mediaanin, moodin ja keskiarvon avulla. Luokitellaan aineisto ja muodostetaan tarvittavat frekvenssijakaumat.

Aineisto voidaan luokitella esimerkiksi:

Osuus (%)	***f***	***sf***	***sf* %**
26–30	1	1	$\frac{1}{18} = 0{,}0555\ldots \approx 6\%$
31–35	4	1 + 4 = 5	$\frac{5}{18} = 0{,}277\ldots \approx 28\%$
36–40	3	5 + 3 = 8	$\frac{8}{18} = 0{,}444\ldots \approx 44\%$
41–45	6	14	0,777… ≈ 78 %
46–50	3	17	0,944… ≈ 94 %
51–55	1	18	100 %

Frekvenssijakauman perusteella moodiluokka on luokka 41–45.

Moodina on luokan luokkakeskus $\frac{40{,}5+45{,}5}{2}=43$ (%). Suurimmassa osassa verojen osuus bruttokansantuotteesta on siis 43 %.

Suhteellisen summafrekvenssijakauman perusteella mediaaniluokka on luokka 41–45 (%). Näin ollen mediaani on 43 %. Puolessa OECD maista verojen osuus on siis alle 43 % ja puolessa yli 43 %.

Keskiarvon laskua varten määritetään luokkien luokkakeskukset ja lasketaan frekvenssin ja luokkakeskuksen tulot taulukkoon. (Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Osuus (%)	f_i	**Luokka-keskus** x_i	$f_i x_i$
26–30	1	$\frac{25{,}5+30{,}5}{2}=28$	$1\cdot 28=28$
31–35	4	$\frac{30{,}5+35{,}5}{2}=33$	$4\cdot 33=132$
36–40	3	38	$3\cdot 38=114$
41–45	6	43	258
46–50	3	48	144
51–55	1	53	53

Verojen osuuden (%) keskiarvo on

$$\bar{x}=\frac{\sum_{i=1}^{6} f_i x_i}{18}$$

$$=\frac{28+132+114+258+144+53}{18}=\frac{729}{18}=40{,}5\ \ (\%)$$

OECD-maissa maksetaan siis veroja keskimäärin 40,5 % bruttokansantuotteesta.

Vastaus: Md = 43 %, Mo = 43 %, $\bar{x}=40{,}5$ %

51. Kukkarossa on aluksi seuraavat kolikot:

		
massa:8,50 g	massa: 7,50 g	massa: 7,80 g

a) Kukkaroon lisätään 10 sentin kolikko. Merkitään lisätyn kolikon massaa kirjaimella x (g). Koska kukkarossa olevien kolikoiden keskiarvo $\bar{x} = 6{,}975 \text{ g}$, niin saadaan

$$
\begin{aligned}
\frac{8{,}50 + 7{,}50 + 7{,}80 + x}{4} &= 6{,}975 \quad |\cdot 4 \\
23{,}8 + x &= 27{,}9 \\
x &= 27{,}9 - 23{,}8 \\
x &= 4{,}10
\end{aligned}
$$

Lisätyn kolikon massa on siis 4,10 g.

b) Merkitään yhden lisätyn kolikon massaa kirjaimella x (g), jolloin kahden lisätyn (samanarvoisen) kolikon massa on siis $2x$ (g). Koska kukkarossa olevien kolikoiden keskiarvo $\bar{x} = 7{,}76 \text{ g}$, niin saadaan

$$
\begin{aligned}
\frac{8{,}50 + 7{,}50 + 7{,}80 + 2x}{5} &= 7{,}76 \quad |\cdot 5 \\
23{,}8 + 2x &= 38{,}8 \\
2x &= 15 \quad |: 2 \\
x &= 7{,}50
\end{aligned}
$$

Yhden lisätyn kolikon massa on siis 7,50 g, joten kyseessä on 1 €:n kolikko.
Koska kukkaroon lisättiin kaksi samanarvoista kolikkoa, niin yhteensä lisättiin 2 €.
Kukkarossa on tällöin rahaa

$$2\ € + 1\ € + 0{,}50\ € + 2\cdot 1\ € = 5{,}50\ €$$

Vastaus: a) 4,10 g b) 5,50 €

52. **a)** Merkitään viimeisen kurssin arvosanaa kirjaimella x.

Jotta kurssien keskiarvo pyöristyisi 8:aan, on keskiarvon oltava vähintään 7,5.
Lasketaan millä kurssiarvosanalla keskiarvo on 7,5.

$$\begin{aligned} \frac{8+6+7+8+6+x}{6} &= 7{,}5 \quad |\cdot 6 \\ 35+x &= 45 \\ x &= 10 \end{aligned}$$

Jotta loppuarvosana pyöristyisi 8:aan, on viimeisen kurssin arvosanan oltava 10.

b) Jotta loppuarvosana on 7, on kurssien keskiarvon oltava vähintään 6,5.
Merkitään viimeisen kurssin arvosanaa kirjaimella x.

$$\begin{aligned} \frac{8+6+7+8+6+x}{6} &= 6{,}5 \quad |\cdot 6 \\ 35+x &= 39 \\ x &= 4 \end{aligned}$$

Jotta loppuarvosanaksi tulisi 7, on viimeisestä kurssista saatava vähintään 4.

Vastaus: a) 10 b) 4

53. Merkitään kysyttyä lukua kirjaimella x. Luvun neliö on tällöin x^2.

Koska luvun ja sen neliön keskiarvo on 0,72, saadaan yhtälö

$$\frac{x+x^2}{2}=0{,}72 \quad |\cdot 2$$
$$x+x^2=1{,}44$$
$$x^2+x-1{,}44=0$$

Ratkaistaan yhtälö toisen asteen yhtälön ratkaisukaavalla.

$$x=\frac{-1\pm\sqrt{1^2-4\cdot 1\cdot(-1{,}44)}}{2\cdot 1}$$
$$x=\frac{-1\pm\sqrt{6{,}76}}{2}$$
$$x=\frac{-1\pm 2{,}6}{2}$$
$$x=\frac{-1+2{,}6}{2} \quad \text{tai} \quad x=\frac{-1-2{,}6}{2}$$
$$x=\frac{1{,}6}{2}=0{,}8 \qquad x=\frac{-3{,}6}{2}=-1{,}8$$

Vastaus: Luvut ovat $-1{,}8$ ja $0{,}8$

54. Vuosien 1986 ja 1987 liikevoitot olivat yhtä suuret. Merkitään niitä kirjaimella x.

Koska keskimääräinen liikevoitto vuosina 1985–2000 oli −29 miljoonaa euroa, saadaan yhtälö

$$\frac{22+x+x+74+...+152+219}{16}=-29$$
$$\frac{2x-524}{16}=-29 \quad |\cdot 16$$
$$2x-524=-464$$
$$2x=60 \quad |:2$$
$$x=30$$

Vastaus: Kummankin vuoden liikevoitto on 30 miljoonaa euroa.

55. Eri osasuorituksia painotettiin sen mukaan, mikä oli tehtävän suhteellinen osuus prosentteina.

Elsa sai englannin kurssin loppuarvosanaksi

$$\frac{20\cdot 8{,}5+20\cdot 8+15\cdot 9+15\cdot 7{,}5+30\cdot 9{,}5}{20+20+15+15+30}=\frac{862{,}5}{100}=8{,}625\approx 8{,}6$$

Vastaus: 8,6

56. Polttoaineen keskikulutus laskettiin painottamalla eri olosuhteita sen mukaan, mikä oli kunkin olosuhteen kesto.

Auton keskikulutukseksi (dm³/100 km) saatiin

$$\frac{4\cdot 5{,}80+5\cdot 6{,}90+3\cdot 8{,}10+2\cdot 7{,}20}{4+5+3+2}=\frac{96{,}4}{14}=6{,}885...\approx 6{,}89\ \ (\text{dm}^3/100\text{ km})$$

Vastaus: 6,89 dm^3/100 km

57. Painotettu keskiarvo lasketaan kertomalla suhteellinen atomimassa sen prosenttiosuudella.

$$\begin{aligned}&\frac{0{,}9\cdot 73{,}922475+9{,}0\cdot 75919212+7{,}6\cdot 76{,}919912}{100}+\\&\frac{23{,}5\cdot 77{,}917304+49{,}6\cdot 79{,}916520+9{,}4\cdot 81{,}916698}{100}\\&=78{,}9932746\\&\approx 78{,}993275\end{aligned}$$

Vastaus: 78,993275

58. Muodostetaan taulukon perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma. Määritetään sen jälkeen moodi ja mediaani.

Terälehtiä (lkm), x_i	f_i	***sf***	***sf* %**
4	26	26	$\frac{26}{108} = 0{,}2407... = 24\,\%$
5	32	26 + 32 = 58	$\frac{58}{108} = 0{,}537... \approx 54\,\%$
6	33	58 + 33 = 91	0,842… ≈ 84 %
7	11	102	0,944… ≈ 94 %
8	6	108	100 %

a) Suurin frekvenssi on terälehtien määrällä 6, joten Mo = 6.

Suhteellinen summafrekvenssi ylittää 50 % terälehtien määrän 5 kohdalla, joten Md = 5.

b) Lasketaan terälehtien määrän keskiarvo. (Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

$$\bar{x} = \frac{\sum_{i=1}^{5} f_i x_i}{108}$$

$$= \frac{26\cdot 4 + 32\cdot 5 + 33\cdot 6 + 11\cdot 7 + 6\cdot 6}{108} = \frac{587}{108} = 5{,}435... \approx 5{,}4 \quad (\text{kpl})$$

Vastaus: a) Mo = 6 (kpl) , Md = 5 (kpl) b) 5,4 (kpl)

59. Koska kokeeseen osallistuneiden opiskelijoiden määrää ei tunneta, merkitään sitä kirjaimella *a*. Muodostetaan tämän perusteella lausekkeet eri arvosanojen opiskelijamäärille.

Arvosana x_i	**Prosentti-osuus**	**Opiskelijoiden määrä** f_i
4	1,3 %	$0{,}013a$
5	9,8 %	$0{,}098a$
6	15,8 %	$0{,}158a$
7	20,3 %	$0{,}203a$
8	23,3 %	$0{,}233a$
9	23,4 %	$0{,}234a$
10	6,1 %	$0{,}061a$

Lasketaan arvosanojen keskiarvo luokitellun aineiston keskiarvon kaavalla. Kaikkien havaintojen määrä on sama kuin opiskelijoiden määrä eli *a*.

$$\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{7} f_i x_i}{a} \\
&= \frac{0{,}013a \cdot 4 + 0{,}098a \cdot 5 + \ldots + 0{,}061a \cdot 10}{a} \\
&= \frac{7{,}491\cancel{a}}{\cancel{a}} \\
&= 7{,}491 \\
&\approx 7{,}5
\end{aligned}$$

Vastaus: keskiarvo on 7,5.

60. Luokitellaan aineisto ja muodostetaan tarvittavat frekvenssijakaumat. Pisteet voidaan luokitella esimerkiksi:

Pisteet	***f***	***sf***	***sf* %**
0–9	18	18	$\frac{18}{30} = 0{,}6 = 60\ \%$
10–19	5	18 + 5 = 23	$\frac{23}{30} = 0{,}766 = 77\ \%$
20–29	2	25	0,83… = 83 %
30–39	4	29	0,966… ≈ 97 %
40–49	1	30	100 %

Jakauman mukaan moodiluokka on 0–9 pistettä, koska sen frekvenssi on suurin. Suurin osa pelaajista on siis tehnyt korkeintaan yhdeksän pistettä.

Mediaaniluokka on suhteellisen summafrekvenssijakauman perusteella myös luokka 0–9 pistettä.

Mediaaniluokan luokkakeskus on $\frac{0+9{,}5}{2} = 4{,}75$ pistettä, joten voidaan ajatella, että puolet pelaajista on tehnyt alle 4,75 pistettä ja puolet yli 4,75 pistettä.

Lasketaan keskiarvoa varten luokkien luokkakeskukset. Lasketaan taulukkoon lisäksi frekvenssien ja luokkakeskusten tulot. (Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

	f_i	**Luokka-keskus** x_i	$f_i\, x_i$
0–9	18	$\frac{0+9{,}5}{2} = 4{,}75$	$18 \cdot 4{,}75 = 85{,}5$
10–19	5	$\frac{9{,}5+19{,}5}{2} = 14{,}5$	$5 \cdot 14{,}5 = 72{,}5$
20–29	2	24,5	49
30–39	4	34,5	138
40–49	1	44,5	44,5

Pelaajien pistekeskiarvo on

$$\bar{x} = \frac{\sum_{i=1}^{5} f_i x_i}{30}$$
$$= \frac{85{,}5 + 72{,}5 + 49 + 138 + 44{,}5}{30} = \frac{389{,}5}{30} = 12{,}983... \approx 13$$

Pelaajat tekivät siis keskimäärin 13 pistettä. keskiarvo ei kuitenkaan tässä aineistossa ole paras mahdollinen tunnusluku, koska jakauma on selvästi vino. Mediaani kuvaa aineistoa paremmin.

61. Kymmenen balettitanssijan yhteismassa on noin $10 \cdot 52\ \text{kg} = 520\ \text{kg}$.
Koska viisi jalkapalloilijaa painaa yhteensä 425 kg, kaikkien viidentoista urheilijan yhteismassa on

520 kg + 425 kg = 945 kg

Keskimääräinen paino on tällöin

$$\frac{945\,\text{kg}}{15} = 63\ \text{kg}$$

Vastaus: 63 kg

62. -

1.4 Keskihajonta

63. Keskihajonnan laskemista varten lasketaan ensin keskiarvo. Melun keskiarvo on

$$\overline{x} = \frac{62+44+41+53+72}{5} = \frac{272}{5} = 54{,}4 \text{ (dB)}$$

Keskihajonnan laskemista varten muodostetaan taulukko.

Viikonpäivä	**Melu (dB)** x_i	$(x_i - \overline{x})^2$
ma	62	$(62-54{,}4)^2 = 7{,}6^2$
ti	44	$(44-54{,}4)^2 = (-10{,}4)^2$
ke	41	$(41-54{,}4)^2 = (-13{,}4)^2$
to	53	$(53-54{,}4)^2 = (-1{,}4)^2$
pe	72	$(72-54{,}4)^2 = 17{,}6^2$

Keskihajonta on

$$\begin{aligned} s &= \sqrt{\frac{\sum_{i=1}^{5}(x_i-\overline{x})^2}{n-1}} \\ &= \sqrt{\frac{7{,}6^2+(-10{,}4)^2+(-13{,}4)^2+(-1{,}4)^2+17{,}6^2}{5-1}} \\ &= \sqrt{\frac{657{,}2}{4}} = 12{,}817\ldots \approx 13 \ (dB) \end{aligned}$$

Vastaus: 13 dB

64. **a)** Lasketaan uutuuselokuvien määrän keskiarvo. (Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

$$\bar{x} = \frac{179+150+169+...+176}{10} = \frac{1648}{10} = 164{,}8 \approx 165 \quad (\text{elokuvaa})$$

b) Lasketaan uutuuselokuvien määrän keskihajonta taulukoimalla. (Keskihajonnan voi laskea myös laskimen tilastotoiminnolla.)

Elokuvien lkm. x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
179	179 – 164,8 = 14,2	$(14{,}2)^2 = 201{,}64$
150	150 – 164,8 = –14,8	$(-14{,}8)^2 = 219{,}04$
169	4,2	17,64
163	–1,8	3,24
147	–17,8	316,84
166	1,2	1,44
154	–10,8	116,64
152	–12,8	163,84
192	27,2	739,84
176	11,2	125,44

Keskihajonta

$$\begin{aligned}
s &= \sqrt{\frac{\sum_{i=1}^{10}(x_i - \bar{x})^2}{n-1}} \\
&= \sqrt{\frac{201{,}64 + 219{,}04 + ... + 125{,}44}{10-1}} \\
&= \sqrt{\frac{1905{,}6}{9}} \\
&= 14{,}551... \\
&\approx 15 \quad (\text{elokuvaa})
\end{aligned}$$

Vastaus: a) 165 elokuvaa b) 15 elokuvaa

65. **a)** Lasketaan kummankin hyppääjän hyppyjen keskipituus.

Matti

$$\bar{x}_M = \frac{82+85+78+90+...+79}{12} = \frac{981}{12} = 81,75 \quad \text{(m)}$$

Toni

$$\bar{x}_T = \frac{79+81+81+82+...+78}{12} = \frac{957}{12} = 79,75 \quad \text{(m)}$$

Koska 81,75 m > 79,75 m, Matti hyppää keskimäärin pidempiä hyppyjä kuin Toni.

b) Lasketaan kummakin hyppääjän hyppyjen keskihajonta.

Matti

$$s_M = \sqrt{\frac{(82-81,75)^2+(85-81,75)^2+...+(79-81,75)^2}{12-1}} = \sqrt{\frac{446,25}{11}} = 6,369... \quad \text{(m)}$$

Toni

$$s_T = \sqrt{\frac{(89-79,75)^2+(81-79,75)^2+...+(78-79,75)^2}{12-1}} = \sqrt{\frac{50,25}{11}} = 2,137... \quad \text{(m)}$$

Koska Tonin hyppyjen keskihajonta on pienempi kuin Matin hyppyjen keskihajonta, Toni hyppää tasaisempia hyppyjä.

(Keskiarvot ja -hajonnat voi laskea myös laskimen tilastotoiminnolla.)

Vastaus: a) Matti b) Toni

66. Lasketaan 5 vuoden obligaation keskikorko.

$$\bar{x}_5 = \frac{7{,}93 + 6{,}03 + 4{,}86 + ... + 4{,}54}{7} = \frac{37}{7} = 5{,}285... \approx 5{,}29 \text{ (\%)}$$

Lasketaan keskihajonta taulukoimalla.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
7,93	7,93 – 5,285…= 2,644…	$2{,}644\ldots^2 = 6{,}992\ldots$
6,03	6,03 – 5,285…= 0,744…	$0{,}744\ldots^2 = 0{,}553\ldots$
4,86	–0,425…	0,181…
4,3	–0,985…	0,971…
4,07	–1,215…	1,477…
5,27	–0,0157…	0,00024…
4,54	–0,745…	0,556…

5 vuoden obligaatioiden keskihajonta on siis

$$s_5 = \sqrt{\frac{6{,}992... + 0{,}553... + ... + 0{,}556...}{7-1}} = \sqrt{\frac{10{,}733...}{6}} = 1{,}337... \approx 1{,}34 \text{ (\%)}$$

Lasketaan 10 vuoden obligaation keskikorko

$$\bar{x}_{10} = \frac{8{,}79 + 7{,}08 + 5{,}95 + ... + 5{,}04}{7} = \frac{41{,}87}{7} = 5{,}981... \approx 5{,}98 \text{ (\%)}$$

Lasketaan keskihajonta taulukoimalla.

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
8,79	8,79 – 5,981… = 2,808…	$2{,}808\ldots^2 = 7{,}888\ldots$
7,08	7,08 – 5,981… = 1,0985…	$1{,}0985\ldots^2 = 1{,}206\ldots$
5,95	–0,0314…	0,000987…
4,78	–1,201…	1,443…
4,74	–1,241…	1,541…
5,49	–0,491…	0,241…
5,04	–0,941…	0,886…

10 vuoden obligaatioiden keskihajonta on siis

$$s_{10} = \sqrt{\frac{7{,}888... + 1{,}206... + ... + 0{,}886...}{7-1}} = \sqrt{\frac{13{,}208...}{6}} = 1{,}483... \approx 1{,}48 \ \ (\%)$$

(Keskiarvot ja –hajonnat olisi voinut laskea myös laskimen tilastotoiminnolla.)

10 vuoden obligaation koron keskihajonta on hieman suurempi kuin 5 vuoden obligaation koron keskihajonta. 10 vuoden obligaation korot poikkeavat siis enemmän keskiarvostaan kuin 5 vuoden korot.

Hajontalukujen perusteella 10 vuoden obligaation korot vaihtelevat enemmän kuin 5 vuoden obligaation korot. Toisaalta 10 vuoden obligaation korolla on parempi keskiarvo kuin 5 vuoden obligaation korolla.

67. Keskikorkeuden laskemiseksi määritetään luokkien luokkakeskukset.

Korkeus (m)	***f***	**Luokkakeskus (m)**
0–999	6	$\dfrac{0+999{,}5}{2}=499{,}75$
1000–1999	12	$\dfrac{999{,}5+1999{,}5}{2}=1499{,}5$
2000–2999	9	2499,5
3000–3999	17	3499,5
4000–4999	7	4499,5
5000–5999	6	5499,5
6000–6999	2	6499,5

Taulukkoon on tilastoitu tulivuoria yhteensä
$n = 6 + 12 + 9 + 17 + 7 + 6 + 2 = 59$ (kpl).

Korkeuden keskiarvo on näin ollen

$$\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{7} f_i x_i}{n} \\
&= \frac{6\cdot 499{,}75 + 12\cdot 1499{,}5 + \ldots + 2\cdot 6499{,}5}{59} \\
&= \frac{180\,472}{59} \\
&= 3058{,}84\ldots \\
&\approx 3059 \quad (\mathrm{m})
\end{aligned}$$

Lasketaan keskihajonta taulukoimalla.

$x_i - \overline{x}$	$\left(x_i - \overline{x}\right)^2$
499,75 – 3058,84… = –2559,09…	$(-2559{,}09\ldots)^2$ = 6 548 979,79…
1499,5 – 3058,84… = –1559,34…	$(-1559{,}34\ldots)^2$ = 2 431 564,49…
–559,34…	312 869,57…
440,65…	194 174,66…
1440,65…	2 075 479,74…
2440,65…	5 956 784,83…
3440,65…	11 838 089,91…

Keskihajonta on

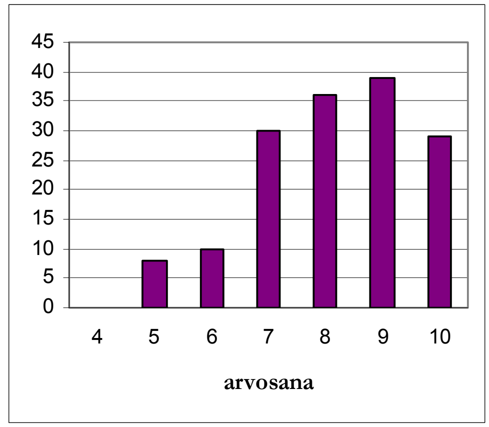
$$
\begin{aligned}
s &= \sqrt{\frac{\sum_{i=1}^{7} f_i (x_i - \overline{x})^2}{n-1}} \\
&= \sqrt{\frac{6 \cdot 6548979{,}79\ldots + 12 \cdot 2431564{,}49\ldots + \ldots + 2 \cdot 11838089{,}91\ldots}{59-1}} \\
&= \sqrt{\frac{148534695{,}25\ldots}{58}} \\
&= 1600{,}29\ldots \\
&\approx 1600 \qquad (\mathrm{m})
\end{aligned}
$$

(Keskiarvon ja -hajonnan olisi voinut laskea myös laskimen tilastotoiminnoilla.)

Vastaus: Keskikorkeus 3059 m ja korkeuden keskihajonta 1600 m.

68. Piirretään pylväsdiagrammi aineistosta.

Arvosana	*f*
4	0
5	8
6	10
7	30
8	36
9	39
10	29
Yhteensä	152



Arvosanojen keskiarvo on

$$\bar{x} = \frac{\sum_{i=1}^{7} f_i x_i}{152}$$
$$= \frac{0 \cdot 4 + 8 \cdot 5 + 10 \cdot 6 + ... + 29 \cdot 10}{152} = \frac{1239}{152} = 8{,}1513... \approx 8{,}15$$

(Keskiarvon voisi laskea myös laskimen tilastotoiminnolla.)

Lasketaan keskihajonta taulukoimalla.
(Keskihajonnan voisi laskea myös laskimen tilastotoiminnolla.)

$x_i - \overline{x}$	$\left(x_i - \overline{x}\right)^2$
4 – 8,151…= –4,151…	$(-4{,}151\ldots)^2 = 17{,}233\ldots$
5 – 8,151…= –3,151…	$(-3{,}151\ldots)^2 = 9{,}930\ldots$
–2,151…	4,628…
–1,151…	1,325…
–0,151…	0,0228…
0,848…	0,720…
1,848…	3,417…

Keskihajonta on

$$
\begin{aligned}
s &= \sqrt{\frac{\sum_{i=1}^{7} f_i (x_i - \overline{x})^2}{152 - 1}} \\
&= \sqrt{\frac{0 \cdot 17{,}233... + 8 \cdot 9{,}930... + ... + 29 \cdot 3{,}417...}{151}} \\
&= \sqrt{\frac{293{,}519...}{151}} \\
&= 1{,}394... \\
&\approx 1{,}39
\end{aligned}
$$

Vastaus: Keskiarvo on 8,15 ja keskihajonta 1,39.

69. Lasketaan aluksi lämpötilojen **keskiarvot** molemmilla paikkakunnilla.

Helsinki

$$
\begin{aligned}
\overline{x}_H &= \frac{17+15+19,2+16,4+18,7+17+20,2+19,1}{8} \\
&= \frac{142,6}{8} \\
&= 17,825 \\
&\approx 17,8 \quad (°\mathrm{C})
\end{aligned}
$$

Sodankylä

$$
\begin{aligned}
\overline{x}_S &= \frac{14,1+12,9+15,9+15,5+14,7+15,5+14,9+15,6}{8} \\
&= \frac{119,1}{8} \\
&= 14,8875 \\
&\approx 14,9 \quad (°\mathrm{C})
\end{aligned}
$$

(Keskiarvon voisi laskea myös laskimen tilastotoiminnolla.)

Keskiarvon perusteella heinäkuun keskilämpötila on korkeampi Helsingissä kuin Sodankylässä (17,8 °C > 14,9 °C).

Määritetään lisäksi lämpötilojen **mediaanit** suuruusjärjestykseen asetetusta aineistosta.

Helsinki

15; 16,4; 17, 17;18,7; 19,1; 19,2; 20,2

Mediaani

$$Md = \frac{17 + 18{,}7}{2} = 17{,}85 \;(°C)$$

Sodankylä

12,9; 14,1; 14,7; 14,9; 15,5; 15,5; 15,6; 15,9

Mediaani

$$Md = \frac{14{,}9 + 15{,}5}{2} = 15{,}2 \quad (°C)$$

Mediaanit ovat kummassakin tapauksessa hyvin lähellä keskiarvoja. Myös mediaanien perusteella heinäkuun keskilämpötila on korkeampi Helsingissä (17,85 °C > 15,2 °C).

Lasketaan vielä kummankin paikkakunnan lämpötilojen **keskihajonnat**.

Helsinki

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
17	17 – 17,825 = –0,825	$(-0{,}825)^2 = 0{,}680625$
15	15 – 17,825 = –2,825	$(-2{,}825)^2 = 7{,}98062$
19,2	1,375	1,890625
16,4	–1,425	2,030625
18,7	0,875	0,765625
17	–0,825	0,680625
20,2	2,375	5,640625
19,1	1,275	1,625625

Keskihajonta

$$\begin{aligned} s_H &= \sqrt{\frac{0{,}680625 + 7{,}98062 + \ldots + 1.625625}{8-1}} \\ &= \sqrt{\frac{21{,}295}{7}} \\ &= 1{,}744\ldots \\ &\approx 1{,}74 \ (°\,\mathrm{C}) \end{aligned}$$

Sodankylä

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
14,1	14,1 – 14,8875 = –0,7875	$(-0{,}7875)^2 = 0{,}6201\ldots$
12,9	12,9 – 14,8875 = –1,9875	$(-1{,}9875)^2 = 3{,}9501\ldots$
15,9	1,0125	1,0251…
15,5	0,6125	0,3751…
14,7	–0,1875	0,0351…
15,5	0,6125	0,3751…
14,9	0,0125	0,000156…
15,6	0,7125	0,5076…

Keskihajonta

$$\begin{aligned} s_S &= \sqrt{\frac{0{,}6201\ldots + 3{,}9501\ldots + \ldots + 0{,}5076}{8-1}} \\ &= \sqrt{\frac{6{,}8887\ldots}{7}} \\ &= 0{,}9920\ldots \\ &\approx 0{,}992 \ (°\,\mathrm{C}) \end{aligned}$$

Heinäkuun lämpötilan keskihajonta on pienempi Sodankylässä
(0,992 °C < 1,74 °C).

Sijaintilukujen perusteella voidaan sanoa, että Helsingissä on heinäkuussa keskimäärin lämpimämpää kuin Sodankylässä.

Hajontalukujen perusteella voidaan sanoa, että Sodankylän keskilämpötilat heinäkuussa poikkeavat toisistaan vähemmän toisistaan kuin Helsingin lämpötilat.

70. Laaditaan aluksi frekvenssijakauma ja lasketaan sen jälkeen keskihajonta ja keskiarvo laskimen tilastotoiminnolla.

Aika (s)	***f***
120	1
125	1
128	2
129	1
130	1
131	1
132	1
133	1
135	2
138	2
142	2
145	1
148	1
150	1
151	1
152	1
155	1
Yhteensä	21

a) Syötetään ajat ja frekvenssit laskimeen. Tilastotoiminnolla saadaan keskihajonnaksi

$s = 9{,}84692361\ldots\ \text{s} \approx 9{,}85\ \text{s}$

b) Tilastotoiminnolla saadaan vastaavasti keskiarvoksi

$$\bar{x} = 137,47619...\ \text{s} \approx 137\ \text{s}.$$

Nopein aika on 120 s. Jos aika on vähintään kahden keskihajonnan päässä keskiarvosta, niin poikkeama on merkittävä.

$$\bar{x} - 2s = 137,476...\ \text{s-}2 \cdot 9,846...\ \text{s} = 117,782...\ \text{s} < 120\ \text{s}.$$

Nopein aika ei siis poikkea merkittävästi keskiarvosta.

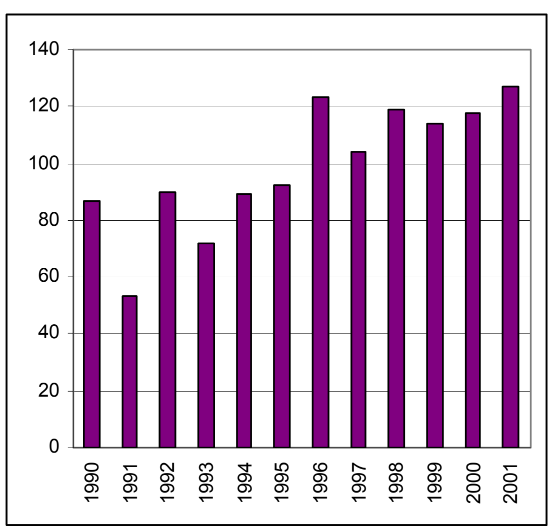
Hitain aika on 155 s. Jos aika on vähintään kahden keskihajonnan päässä keskiarvosta, niin poikkeama on merkittävä.

$$\bar{x} + 2s = 137,476...\ \text{s} + 2 \cdot 9,846...\ \text{s} = 157,17...\ \text{s} > 155\ \text{s}$$

Hitainkaan aika ei siis poikkea merkittävästi keskiarvosta.

Vastaus: a) 9,85 s   b) Eivät poikkea.

71. Piirretään jakaumasta pylväsdiagrammi.



Lasketaan keskiarvo. (Keskiarvo voidaan laskea myös laskimen tilastotoiminnolla.)

$$\bar{x} = \frac{87+53+90+...+127}{12} = \frac{1188}{12} = 99 \quad \text{(poikasta)}$$

Lasketaan hajonta taulukoimalla. (Keskihajonta voidaan laskea myös laskimen tilastotoiminnolla.)

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
87	$87 - 99 = -12$	$(-12)^2 = 144$
53	$53 - 99 = -46$	$(-46)^2 = 2116$
90	–9	81
72	–27	729
89	–10	100
92	–7	49
123	24	576
104	5	25
119	20	400
114	15	225
118	19	361
127	28	784

Keskihajonta

$$\begin{aligned} s &= \sqrt{\frac{144+2116+81+...+784}{12-1}} \\ &= \sqrt{\frac{5590}{11}} \\ &= \sqrt{508{,}18...} \\ &= 22{,}54.... \\ &\approx 22{,}5 \quad \text{(poikasta)} \end{aligned}$$

Jos arvo on vähintään kahden keskihajonnan päässä keskiarvosta, niin arvon poikkeama on merkittävä. Aineiston keskihajonta on s = 22,54… ja keskiarvo $\bar{x} = 99$.

Vuoden 1991 arvo on 53.

$$\bar{x} - 2s = 99 - 2 \cdot 22{,}54... = 53{,}92... > 53$$

Arvo 53 (poikasta) poikkeaa siis enemmän kuin kaksi keskihajontaa, joten poikkeama on merkittävä.

Vuoden 2001 arvo on 127.

$$\bar{x} + 2s = 99 + 2 \cdot 22{,}54... = 144{,}08... > 127$$

Arvo 127 (poikasta) on siis alle kahden keskihajonnan päässä keskiarvosta. Tällöin vuoden 2001 arvo ei poikkea merkittävästi keskiarvosta.

HUOM! Keskihajonnan voi laskea populaation keskihajontana.
Tällöin keskihajonta

$$s = \sqrt{\frac{5590}{12}} = 21{,}583... \approx 21{,}6 \quad \text{(poikasta)}.$$

Vastaus: Jakauman keskiarvo on 99 poikasta ja keskihajonta 22,5 poikasta. Vuoden 1991 arvo poikkeaa merkittävästi keskiarvosta. Vuoden 2001 arvo ei poikkea merkittävästi keskiarvosta.

72. **a)** Luokitellaan aineisto siten, että pienin luokka on 200–249 (g), ja laaditaan frekvenssijakauma.

Massa (g)	***f***	**Luokkakeskus (g)**
200–249	2	$\frac{199{,}5 + 249{,}5}{2} = 224{,}5$
250–299	4	$\frac{249{,}5 + 299{,}5}{2} = 274{,}5$
300–349	3	324,5
350–399	5	374,5
400–449	2	424,5
Yhteensä	16	

b) Keskiarvo

$$\begin{aligned}\bar{x} &= \frac{2 \cdot 224{,}5 + 4 \cdot 274{,}5 + 3 \cdot 324{,}5 + 5 \cdot 374{,}5 + 2 \cdot 424{,}5}{16} \\ &= \frac{5242}{16} \\ &= 327{,}625 \\ &\approx 328 \ \text{(g)}\end{aligned}$$

(Keskiarvo voidaan laskea myös laskimen tilastotoiminnoilla.)

Keskihajonnaksi saadaan laskimen tilastotoiminnolla

$$s = 64{,}46898... \approx 64{,}5 \text{ (g)}$$

c) Jos arvo on vähintään kahden keskihajonnan päässä keskiarvosta, niin arvon poikkeama on merkittävä. Aineiston keskihajonta on $s = 64{,}46898\ldots$ g ja keskiarvo $\bar{x} = 327{,}625$ g.

$$\bar{x} + 2s = 327{,}625\text{ g} + 2 \cdot 64{,}46898...\text{ g} = 457{,}4665...\text{ g} < 520\text{ g}$$

Siis 520 g painava kurkku poikkeaa enemmän kuin kaksi hajontaa keskiarvosta eli poikkeama on merkittävä.

Vastaus: b) $\bar{x} \approx 328$ g, $s \approx 64{,}5$ g c) Poikkeama on merkittävä.

73. Keskihajonnan laskemiseksi tarvitaan keskiarvoa. Tuulen voimakkuuden keskiarvo on

$$\begin{aligned}\bar{x} &= \frac{3{,}4 + 12{,}1 + 7{,}2 + \cdots + 1{,}1 + 1{,}9}{21} \\ &= \frac{79{,}9}{21} \\ &= 3{,}80476...\left(\frac{\text{m}}{\text{s}}\right)\end{aligned}$$

Keskihajonta

$$\begin{aligned}s &= \sqrt{\frac{(3{,}4 - 3{,}80476...)^2 + (12{,}1 - 3{,}80476...)^2 + \cdots (1{,}9 - 3{,}80476...)^2}{21 - 1}} \\ &= 3{,}68815... \\ &\approx 3{,}7\left(\frac{\text{m}}{\text{s}}\right)\end{aligned}$$

Keskihajonta voidaan laskea laskimen tilastotoiminnolla.

Vastaus: $3{,}7\,\frac{\text{m}}{\text{s}}$

74. **a)** Lasketaan pistekeskiarvot molemmille tytöille.

Johanna

$$\begin{aligned}\bar{x}_J &= \frac{5{,}8+5{,}7+5{,}5+5{,}8+5{,}6+5{,}7}{6}\\ &= \frac{34{,}1}{6}\\ &= 5{,}6833...\end{aligned}$$

Elina

$$\begin{aligned}\bar{x}_E &= \frac{5{,}4+5{,}9+5{,}6+5{,}9+5{,}9+5{,}6}{6}\\ &= \frac{34{,}3}{6}\\ &= 5{,}7166...\end{aligned}$$

Koska Elinan keskiarvo on suurempi kuin Johannan keskiarvo, Elinan loppusijoitus on parempi.

b) Lasketaan kummankin tytön pisteiden keskihajonta taulukoimalla.

Johanna

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
5,8	5,8 – 5,68333…= 0,11666…	$(0{,}1166\ldots)^2 = 0{,}013611\ldots$
5,7	5,7 – 5,6833… = 0,01666…	0,0002777…
5,5	5,5 – 5,68333… = –0,18333…	0,0336111…
5,8	5,8 – 5,68333… = 0,11666…	0,0136111…
5,6	5,6 – 5,68333… = –0,08333…	0,0069444…
5,7	5,7 – 5,68333… = 0,01666…	0,0002777…

Keskihajonta

$$
\begin{aligned}
s_J &= \sqrt{\frac{\sum_{i=1}^{6}(x_i-\overline{x})^2}{n-1}} \\
&= \sqrt{\frac{0{,}0136111...+0{,}0002777...+...+0{,}0002777...}{6-1}} \\
&= \sqrt{\frac{0{,}068333...}{5}} \\
&= 0{,}1169...
\end{aligned}
$$

Elina

x_i	$x_i-\overline{x}$	$(x_i-\overline{x})^2$
5,4	5,4 – 5,71666…= –0,31666…	$(–0{,}3166…)^2$ = 0,100277…
5,9	5,9 – 5,71666… = 0,18333…	0,0336111…
5,6	5,6 – 5,71666… = –0,11666…	0,0136111…
5,9	5,9 – 5,71666… = 0,18333…	0,0336111…
5,9	5,9 – 5,71666… = 0,18333…	0,0336111…
5,6	5,6 – 5,71666… = –0,11666…	0,0136111…

Keskihajonta

$$
\begin{aligned}
s_E &= \sqrt{\frac{\sum_{i=1}^{6}(x_i-\overline{x})^2}{n-1}} \\
&= \sqrt{\frac{0{,}1002777...+0{,}0336111...+...+0{,}0136111...}{6-1}} \\
&= \sqrt{\frac{0{,}228333...}{5}} \\
&= 0{,}2136...
\end{aligned}
$$

Koska 0,2136… > 0,1169…, Elinan pisteiden keskihajonta on suurempi.

Keskiarvot ja -hajonnat voi laskea myös laskimen tilastotoiminnoilla.

Vastaus: a) Elinan b) Elinan

75. Koska kokeeseen osallistuneiden opiskelijoiden määrää ei tunneta, merkitään sitä kirjaimella *a*. Muodostetaan tämän perusteella lausekkeet eri arvosanojen opiskelijamäärille.

Arvosana x_i	**Prosentti-osuus**	**Opiskelijoiden määrä** f_i
0	5,80 %	0,0580*a*
1	10,99 %	0,1099*a*
2	17,54 %	0,1754*a*
3	24,78 %	0,2478*a*
4	19,95 %	0,1995*a*
5	15,48 %	0,1548*a*
6	5,46 %	0,0546*a*

Lasketaan arvosanojen keskiarvo luokitellun aineiston keskiarvon kaavalla. Kaikkien havaintojen määrä on sama kuin opiskelijoiden määrä eli *a*.

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^{7} f_i x_i}{a} \\ &= \frac{0,0580a\cdot 0+0,1099a\cdot 1+...+0,0546a\cdot 6}{a} \\ &= \frac{3,1037\cancel{a}}{\cancel{a}} \\ &= 3,1037 \\ &\approx 3,10\end{aligned}$$

Keskihajonta (koko perusjoukko)

$$
\begin{aligned}
s_E &= \sqrt{\frac{\sum_{i=1}^{7} f_i(x_i-\overline{x})^2}{n}} \\
&= \sqrt{\frac{0{,}0580a(0-3{,}1037)^2+0{,}1099a(1-3{,}1037)^2+\ldots+0{,}0546a(6-3{,}1037)^2}{a}} \\
&= \sqrt{\frac{2{,}436346\ldots\cancel{a}}{\cancel{a}}} \\
&= 1{,}56087\ldots \\
&\approx 1{,}56
\end{aligned}
$$

Vastaus: $\overline{x} \approx 3{,}10$, $s \approx 1{,}56$

76. Keskiarvon ja -hajonnan laskemista varten määritetään ensin luokkakeskukset.

Ala (m^2)	***f***	**Luokkakeskus (m^2)**
40–79	6	$\frac{39{,}5+79{,}5}{2}=59{,}5$
80–119	6	99,5
120–159	9	139,5
160–199	10	179,5
200–239	8	219,5
Yhteensä	39	

a) Keskiarvo

$$\begin{aligned}\overline{x} &= \frac{6\cdot 59{,}5+6\cdot 99{,}5+9\cdot 139{,}5+10\cdot 179{,}5+8\cdot 219{,}5}{39}\\ &= \frac{5760{,}5}{39}\\ &= 147{,}7051...\\ &\approx 148\ \ (\mathrm{m}^2)\end{aligned}$$

(Keskiarvon voi laskea myös laskimen tilastotoiminnolla.)

Keskihajonnaksi saadaan laskimen tilastotoiminnolla

$$s = 54{,}42869...\ \mathrm{m}^2 \approx 54{,}4\ \mathrm{m}^2$$

b) Jos arvo on vähintään kahden keskihajonnan päässä keskiarvosta, niin arvon poikkeama on merkittävä. Aineiston keskihajonta on $s = 54{,}42869\ldots\ \mathrm{m}^2$ ja keskiarvo $\overline{x} = 147{,}7051...\ \mathrm{m}^2$, jolloin

$$\overline{x} + 2s = 147{,}7051...\,\mathrm{m}^2 + 2\cdot 54{,}42869...\,\mathrm{m}^2 = 256{,}5625...\,\mathrm{m}^2 > 225\ \mathrm{m}^2$$

Siis 225 m^2 ei poikkea vielä merkittävästi talojen keskikoosta.

Vastaus: a) $\overline{x} \approx 148\ \mathrm{m}^2$, $s \approx 54{,}4\ \mathrm{m}^2$ b) Ei poikkea merkittävästi.

77. Koska lukujen keskiarvo on 16,25, saadaan yhtälö

$$\frac{x+(x+4)+(3x-2)+4x}{4}=16{,}25 \quad |\cdot 4$$
$$9x+2=65$$
$$9x=63 \quad |:9$$
$$x=7$$

Luvut ovat siis

7, $7 + 4 = 11$, $3\cdot 7-2=19$ ja $4\cdot 7=28$.

Lasketaan lukujen keskihajonta taulukoimalla.

x_i	$x_i-\overline{x}$	$(x_i-\overline{x})^2$
7	7 – 16,25 = –9,25	$(-9{,}25)^2 = 85{,}5625$
11	11 – 16,52= –5,25	27,5625
19	19 – 16,25 = 2,75	7,5625
28	28 – 16,25 = 11,75	138,0625

Keskihajonta

$$\begin{aligned} s &= \sqrt{\frac{85{,}5625+27{,}5625+7{,}5625+138{,}0625}{4-1}} \\ &= \sqrt{\frac{258{,}75}{3}} \\ &= \sqrt{86{,}25} \\ &= 9{,}2870... \\ &\approx 9{,}29 \end{aligned}$$

(Lukujen keskihajonnan olisi voinut laskea myös laskimen tilastotoiminnolla.)

HUOM! Tässä keskihajonnan voi laskea myös koko perusjoukkoa vastaavalla keskihajonnalla eli käyttämällä jakajana lukua *n*. Tällöin keskihajonta on 8,0428… ≈ 8,04.

1.5 Normaalijakauma

78. Lasketaan kummankin lemmikin painon normitettu arvo.

Maikin kissa

$$z_M = \frac{4{,}8-4{,}3}{0{,}7} = \frac{0{,}5}{0{,}7} = 0{,}7142...$$

Vilman koira

$$z_V = \frac{12{,}8-10{,}2}{0{,}9} = \frac{2{,}6}{0{,}9} = 2{,}888... \;\; (> 0{,}7142\ldots)$$

Vilman koiran normitettu arvo on suurempi kuin Maikin kissan normitettu arvo, joten koira on suhteessa painavampi.

Vastaus: Vilman koira on suhteessa painavampi.

79. Lasketaan papukaijan ikää 140 vuotta vastaava normitettu arvo.

$$z = \frac{140-92}{9,5} = \frac{48}{9,5} = 5,052...$$

Jotta ihminen olisi suhteessa yhtä vanha, on iän normitettu arvo oltava sama kuin papukaijalla. Merkitään ihmisen ikää kirjaimella x. Saadaan yhtälö

$$\begin{aligned}
\frac{x-82}{4,5} &= \frac{48}{9,5} \\
9,5(x-82) &= 4,5 \cdot 48 \\
9,5x - 779 &= 216 \\
9,5x &= 995 \quad |:9,5 \\
x &= 104,73... \\
x &\approx 105 \quad (\text{vuotta})
\end{aligned}$$

Vastaus: Ihmisen on oltava 105-vuotias.

80. Lasketaan historian koenumeroa (8+ = 8,25) vastaava normitettu arvo.

$$z = \frac{8,25-7,2}{0,8} = \frac{1,05}{0,8} = 1,3125$$

Ruotsin kokeen arvosana (7+ = 7,25) on suhteessa yhtä hyvä kuin historian arvosana. Sen normitettu arvo on siis myös 1,3125.

Merkitään ruotsin kokeen keskiarvoa kirjaimella x. Saadaan yhtälö

$$\begin{aligned}\frac{7{,}25-x}{0{,}8} &= 1{,}3125 \quad \mid \cdot 0{,}8\\ 7{,}25-x &= 1{,}05\\ -x &= -6{,}2 \quad \mid \cdot(-1)\\ x &= 6{,}2\end{aligned}$$

Vastaus: Ruotsin kokeen keskiarvo on 6,2.

81. Muuttujan arvo $x = 35$.

a) $N(35, 4)$, $z = \dfrac{35-35}{4} = \dfrac{0}{4} = 0$

b) $N(39, 4)$, $z = \dfrac{35-39}{4} = \dfrac{-4}{4} = -1$

c) $N(31, 4)$, $z = \dfrac{35-31}{4} = \dfrac{4}{4} = 1$

82. $x \sim N(0{,}1)$, jakauma on normitettu normaalijakauma, joten taulukosta voidaan lukea todennäköisyydet.

a) $P(x < 0{,}25) = 0{,}5987 \approx 0{,}60$

b) $P(x < 1{,}37) = 0{,}9147 \approx 0{,}91$

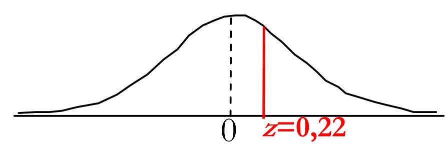
c) $P(x > 1{,}37) = 1 - P(x < 1{,}37) = 1 - 0{,}9147 = 0{,}0853 \approx 0{,}085$

83. $x \sim N(56{,}18)$
Lasketaan, kuinka monen prosentin todennäköisyydellä kokeen pistemäärä on alle 60. Normitetaan aluksi arvo 60 (pistettä).

$$z_{60} = \frac{60-56}{18} = \frac{4}{18} = 0{,}222...$$

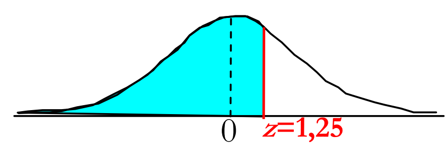
Kysytään siis todennäköisyyttä $P(z < 0{,}22)$. Luetaan taulukosta normitettua arvoa vastaava prosenttiluku eli kertymäfunktion arvo

$$\Phi(0{,}22) = 0{,}5871 = 58{,}71\% \approx 59\%$$



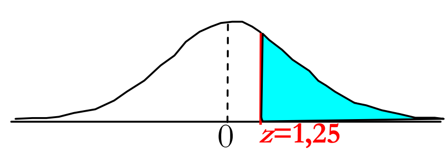
Vastaus: 59 %

84. $x \sim N(20{,}4)$
a) $z_{25} = \dfrac{25-20}{4} = \dfrac{5}{4} = 1{,}25$



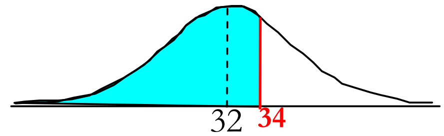
$$P(x \le 25) = P(z \le 1{,}25) = 0{,}8944 \approx 0{,}89$$

b) $P(x \ge 25) = P(z \ge 1{,}25) = 1 - \Phi(1{,}25) = 1 - 0{,}8944 = 0{,}1056 \approx 0{,}11$



85. $x \sim N(32;1{,}7)$

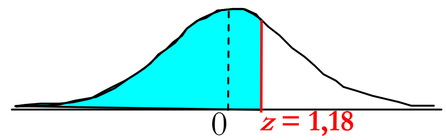
a) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Normitetaan massan arvo 34 kg.

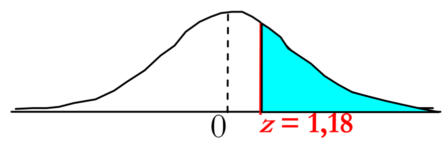
$$z_{34} = \frac{34-32}{1{,}7} = \frac{2}{1{,}7} = 1{,}17647 \approx 1{,}18$$

Kysytty todennäköisyys on siis $P(z \le 1{,}18)$ eli



$$\Phi(1{,}18) = 0{,}8810 \approx 0{,}88$$

b) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



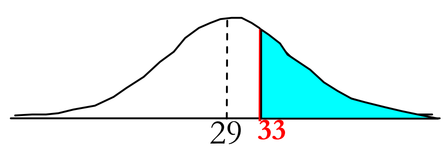
Kysytty todennäköisyys on siis $P(z \ge 1{,}18)$ eli

$$1-\Phi(1{,}18) = 1-0{,}8810 = 0{,}119 \approx 0{,}12$$

Vastaus: a) 0,88 b) 0,12

86. $x \sim N(29;1{,}967)$

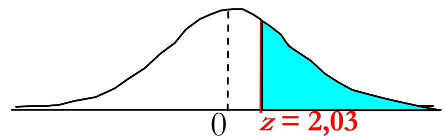
a) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Normitetaan pistemäärä 33.

$$z_{33} = \frac{33-29}{1{,}967} = 2{,}0335 \approx 2{,}03$$

Kysytty todennäköisyys on siis $P(z > 2{,}03)$ eli



$$1 - \Phi(2{,}03) = 1 - 0{,}9788 = 0{,}0212 \approx 0{,}021$$

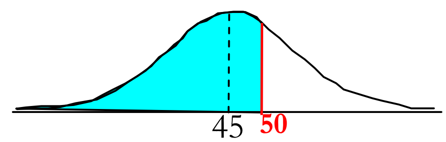
b) Kokeeseen osallistui 233 oppilasta, joten yli 33 pistettä sai

$$0{,}0212 \cdot 233 = 4{,}93... \approx 5 \text{ (oppilasta)}$$

Vastaus: a) 0,021 b) 5 oppilasta

87. $x \sim N(45{,}4)$

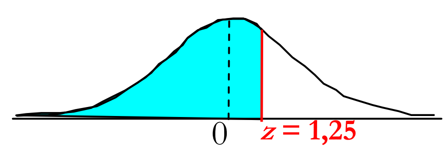
a) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Normitetaan arvo 50.

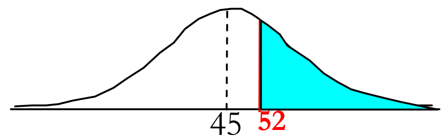
$$z_{50} = \frac{50 - 45}{4} = 1{,}25$$

Kysytty todennäköisyys on siis $P(z < 1{,}25)$ eli



$$\Phi(1{,}25) = 0{,}8944 \approx 0{,}89$$

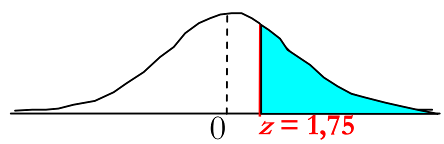
b) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Normitetaan arvo 52.

$$z_{52} = \frac{52-45}{4} = \frac{7}{4} = 1{,}75$$

Kysytty todennäköisyys on siis $P(z > 1{,}75)$ eli



$$1-\Phi(1{,}75) = 1-0{,}9599 = 0{,}0401 \approx 0{,}040$$

Vastaus: a) 0,89 b) 0,040

88. $\mu = 15{,}0\,\text{m} \quad \sigma = 3{,}5\,\text{m}$

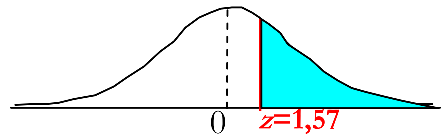
a) Normitetaan pituus 16,7 m.

$$z_{16,7} = \frac{16{,}7-15{,}0}{3{,}5} = \frac{1{,}7}{3{,}5} = 0{,}485... \approx 0{,}49$$

Normitetaan pituus 20,5 m.

$$z_{20,5} = \frac{20{,}5-15{,}0}{3{,}5} = \frac{5{,}5}{3{,}5} = 1{,}571... \approx 1{,}57$$

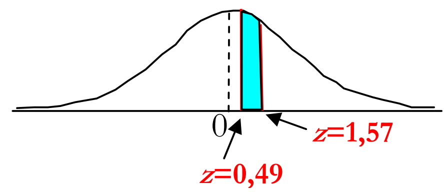
b) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Kysytty todennäköisyys on siis $P(z > 1{,}57)$ eli

$$1 - \Phi(1{,}57) = 1 - 0{,}9418 = 0{,}0582 \approx 0{,}058$$

c) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



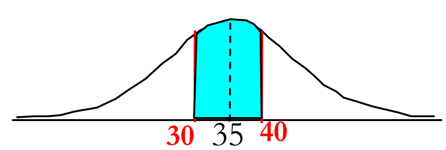
Kysytty todennäköisyys on siis $P(0{,}49 \le z \le 1{,}57)$ eli

$$\Phi(1{,}57) - \Phi(0{,}49) = 0{,}9418 - 0{,}6879 = 0{,}2539 \approx 0{,}25$$

Vastaus: a) $z_{16,7} \approx 0{,}49$, $z_{20,5} \approx 1{,}57$ b) 0,058 c) 0,25

89. $x \sim N(35{,}7)$

Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.

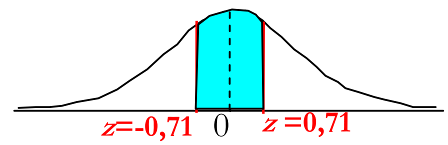


Normitetaan arvot 30 (min) ja 40 (min).

$$z_{30} = \frac{30 - 35}{7} = -0{,}7142... \approx -0{,}71$$

$$z_{40} = \frac{40 - 35}{7} = 0{,}7142... \approx 0{,}71$$

Kysytty todennäköisyys on siis $P(-0{,}71 \le z \le 0{,}71)$ eli $\Phi(0{,}71) - \Phi(-0{,}71)$.



Koska jakauma on symmetrinen, arvon −0,71 vasemmalle puolelle jäävä pinta-ala on yhtä suuri kuin arvon 0,71 oikealle puolelle jäävä pinta-ala.

Arvon −0,71 vasemmalle puolelle jäävä pinta-ala on siis

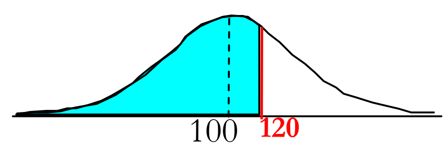
$$\Phi(-0{,}71) = 1 - \Phi(0{,}71) = 1 - 0{,}7611 = 0{,}2389$$

Todennäköisyys, että Luukaksen koulumatka kestää 30−40 min on

$$\Phi(0{,}71) - \Phi(-0{,}71) = 0{,}7611 - 0{,}2389 = 0{,}5222 \approx 0{,}52$$

Vastaus: 0,52

90. $x \sim N(100{,}15)$



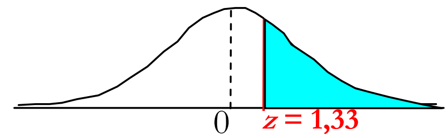
a) Normitetaan arvo 120.

$$z_{120} = \frac{120 - 100}{15} = 1{,}333... \approx 1{,}33$$

Kysytty prosenttiosuus saadaan kertymäfunktion arvona eli

$$\Phi(1{,}33) = 0{,}9082 = 90{,}82\% \approx 90{,}8\%$$

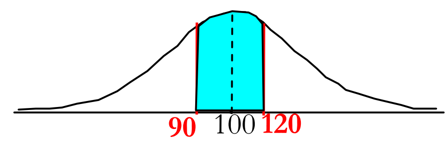
b) Arvon 120 normitettu arvo on $z = 1{,}33$ a-kohdan mukaan.



Kysytty prosenttiosuus on siis

$$1 - \Phi(1{,}33) = 1 - 0{,}9082 = 0{,}0918\% = 9{,}18\% \approx 9{,}2\%$$

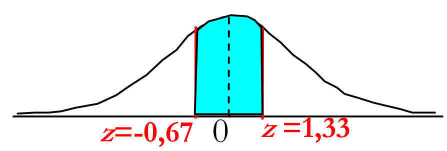
c) Arvon 120 normitettu arvo on $z = 1{,}33$ a-kohdan mukaan.



Normitetaan arvo 90.

$$z_{90} = \frac{90 - 100}{15} = -0{,}666... \approx -0{,}67$$

Kysytty prosenttiosuus on siis



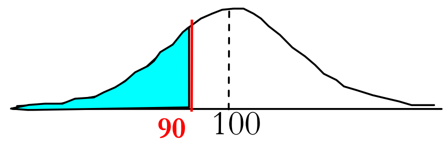
$$\Phi(1{,}33) - \Phi(-0{,}67)$$

Lasketaan ensin prosenttiosuus kohtaan $z = -0{,}67$.

$$\Phi(-0{,}67) = 1 - \Phi(0{,}67) = 1 - 0{,}7486 = 0{,}2514$$

$$\Phi(1{,}33) - \Phi(-0{,}67) = 0{,}9082 - 0{,}2514 = 0{,}6568 \approx 65{,}7\%$$

d) Arvon 90 normitettu arvo on $z = -0{,}67$
c-kohdan mukaan.

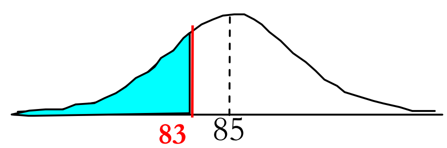


Prosenttiosuus sille, että älykkyysosamäärä on alle 90 on

$$\Phi(-0{,}67) = 1 - \Phi(0{,}67) = 1 - 0{,}7486 = 25{,}14\% \approx 25{,}1\%$$

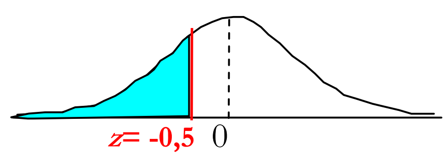
Vastaus: a) 90,8 % b) 9,2 % c) 65,7 % d) 25,1 %

91. $x \sim N(85\,\mathrm{g}, 4\,\mathrm{g})$
Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



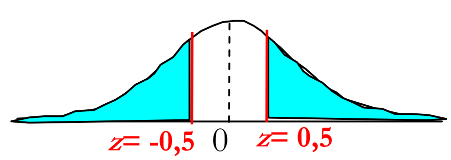
Normitetaan massan arvo 83 g.

$$z_{83} = \frac{83-85}{4} = -\frac{2}{4} = -0{,}5$$



Kysytty todennäköisyys on siis $P(z < -0{,}5)$ eli $\Phi(-0{,}5)$.

Koska jakauma on symmetrinen, kertymäfunktion arvo kohtaan $z = -0{,}5$ on yhtä suuri kuin arvon $z = 0{,}5$ oikealle puolelle jäävä pinta-ala.

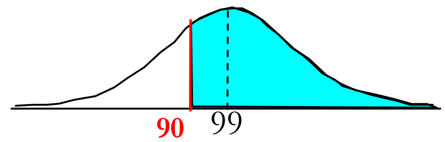


$$\Phi(-0{,}5) = 1 - \Phi(0{,}5) = 1 - 0{,}6915 = 0{,}3085 \approx 0{,}31$$

Vastaus: 0,31

92. $x \sim N(99\,\text{g}, 8\,\text{g})$

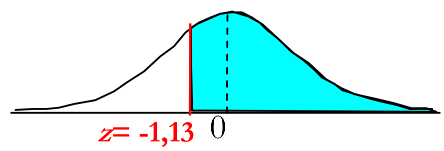
a) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



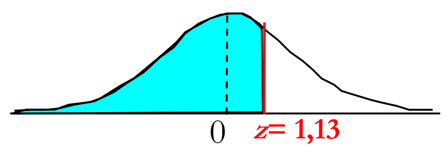
Normitetaan massan arvo 90 g.

$$z_{90} = \frac{90-99}{8} = -\frac{9}{8} = -1{,}125 \approx -1{,}13$$

Kysytty todennäköisyys on normitetun arvon $z = -1{,}13$ oikealle jäävä pinta-ala.



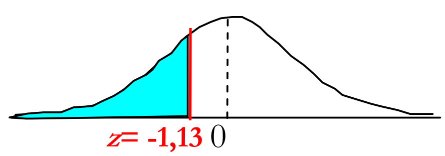
Jakauman symmetrisyydestä johtuen tämä on yhtä suuri kuin normitetun arvon $z = 1{,}13$ vasemmalle jäävä pinta-ala.



Kysytty todennäköisyys on siis $P(z > -1{,}13) = P(z < 1{,}13)$

$$\Phi(1{,}13) = 0{,}8708 \approx 0{,}87$$

b) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.
Arvon 90 normitettu arvo on a-kohdan mukaan $z = -1{,}13$.



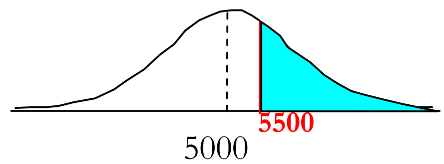
Kysytty todennäköisyys on $P(z < -1{,}13)$ eli

$$\Phi(-1{,}13) = 1 - \Phi(1{,}13) = 1 - 0{,}8708 = 0{,}1292 \approx 0{,}13$$

Vastaus: a) 0,87 b) 0,13

93. $x \sim N(5000\,\text{h}, 8\,\text{h})$

a) Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



Normitetaan kestoaika 5500 h.

$$z_{5500} = \frac{5500-5000}{300} = \frac{500}{300} = \frac{5}{3} = 1{,}666... \approx 1{,}67$$

Kertymäfunktion arvo $\Phi(1{,}67) = 0{,}9525 = 95{,}25\ \%$ kuvaa normitettuun arvoon z = 1,67 mennessä kertynyttä osuutta.

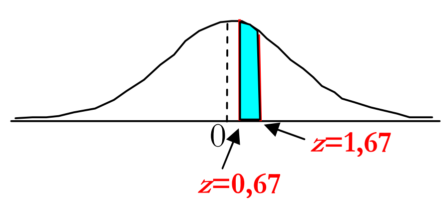
Koska koko pinta-ala on 100 %, saadaan kysytyksi todennäköisyydeksi

100 % – 95,25 % = 4,75 % ≈ 0,048

b) Edellisen kohdan mukaan $z_{5500} = 1{,}666... \approx 1{,}67$. Normitetaan kestoaika 5200 h.

$$z_{5200} = \frac{5200-5000}{300} = \frac{200}{300} = \frac{2}{3} = 0{,}666... \approx 0{,}67$$

Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo.



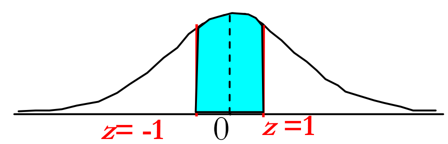
$$\Phi(1{,}67) - \Phi(0{,}67) = 95{,}25\ \% - 74{,}86\ \% = 20{,}39\ \% \approx 0{,}20$$

Vastaus: a) 0,048 b) 0,20

94. $N(17{,}5;\ 2{,}5)$
Normitetaan pituudet 15 cm ja 20 cm.

$$z_{15} = \frac{15 - 17{,}5}{2{,}5} = -1 \qquad z_{20} = \frac{20 - 17{,}5}{2{,}5} = 1$$

Kysytty todennäköisyys on kuvioon merkityn pinta-alan arvo. Tämä saadaan kertymäfunktion arvojen erotuksena



$$\Phi(1) - \Phi(-1)$$

Symmetriasta johtuen arvoon $z = -1$ kertynyt pinta-ala on yhtä suuri kuin arvon $z = 1$ oikealle puolelle jäävä osuus eli

$$\Phi(-1) = 100\% - \Phi(1) = 100\% - 84{,}13\% = 15{,}87\%$$

Todennäköisyys, että hännän pituus on [15 cm, 20 cm] on

$$\Phi(1) - \Phi(-1) = 84{,}13\% - 15{,}87\% = 68{,}26\% \approx 0{,}68$$

Vastaus: 0,68

95. Lakritsinauhojen pituuden keskihajonta on $\sigma = 4{,}0$ cm

a) Olkoon 0,9 m = 90 cm pituisia nauhoja 50 % kaikista nauhoista.

Taulukosta nähdään, että $\Phi(0) = 0{,}50 = 50\,\%$.
Tällöin siis 90 cm pituutta vastaa normitettu arvo $z_{90} = 0$.

$$\begin{aligned} z_{90} &= 0 \\ \frac{90-\mu}{4{,}0} &= 0 \quad \Big|\cdot 4{,}0 \\ 90-\mu &= 0 \\ \mu &= 90 \ \text{(cm)} \end{aligned}$$

b) Olkoon 0,9 m = 90 cm pituisia nauhoja 85 % kaikista nauhoista.

Taulukosta nähdään, että $\Phi(1{,}04) = 0{,}8508 \approx 85\,\%$.
Tällöin siis 90 cm pituutta vastaa normitettu arvo $z_{90} = 1{,}04$.

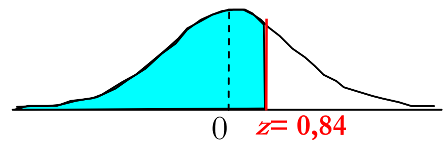
$$\begin{aligned} z_{90} &= 1{,}04 \\ \frac{90-\mu}{4{,}0} &= 1{,}04 \quad \Big|\cdot 4{,}0 \\ 90-\mu &= 4{,}16 \\ -\mu &= 4{,}16-90 \\ \mu &= 85{,}84 \approx 86 \ \text{(cm)} \end{aligned}$$

Vastaus: a) 0,9 m b) 86 cm

96. Mäntyjä, joiden läpimitta yli 26 cm, oli 20 %.
Tällöin mäntyjä, joiden läpimitta oli alle 26 cm oli 100 % – 20 % = 80 %

Normitettu arvo, joka vastaa tätä kertymää, on z = 0,84, koska

$$\Phi(0{,}84) = 0{,}7995 \approx 80\,\%$$



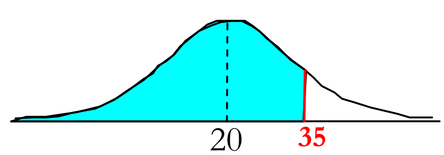
Tiedetään, että läpimitan keskihajonta on $\sigma = 2{,}65\,\text{cm}$. Lasketaan, mikä on keskiarvo μ, jos $z_{26} = 0{,}84$.

$$\frac{26-\mu}{2{,}65} = 0{,}84 \;\Big|\cdot 2{,}65$$
$$26 - \mu = 2{,}226$$
$$-\mu = -23{,}774$$
$$\mu = 23{,}774 \approx 23{,}8 \;\; \text{(cm)}$$

Vastaus: 23,8 cm

97. Hammaslääkärikäyntiin kuluvan ajan keskiarvo on $\mu = 20\,\text{min}$.
Kestoaika on normaalisti jakautunut $N(20,\ \sigma)$.

Koska käynti ei 98 % varmuudella saa ylittää 35 min, etsitään kertymää 98 % vastaava normitettu arvo taulukosta.



$$\Phi(2{,}05) = 0{,}9798 \approx 98\,\%$$

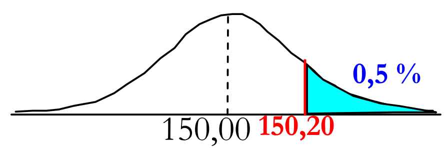
Tällöin siis normitettuun arvoon z = 2,05 on kertynyt 98 %, ja tätä normitettua arvoa vastaa siis 35 min.

$$\frac{35-20}{\sigma}=2,05$$
$$2,05\sigma=15 \,|\cdot 2,05$$
$$\sigma=\frac{15}{2,05}=7,317...\approx 7,3 \text{ (min)}$$

Vastaus: 7,3 min

98. Akselien läpimitan keskiarvo on $\mu=150,00\,\text{mm}$.
Läpimitta on normaalisti jakautunut N(150,00; σ)

Läpimitan 150,20 mm normitettuun arvoon mennessä on kertynyt



100 % – 0,5 % = 99,5 % = 0,995

Akselin läpimitta on siis pienempi tai yhtä suuri kuin 150,20 mm todennäköisyydellä 0,995.

Tätä kertymää vastaava normitettu arvo on taulukon mukaan z = 2,5758, koska $\Phi(2,5758)=0,995$.

$$\frac{150,20-150,00}{\sigma}=2,5758$$
$$2,5758\sigma=0,20 \,|: 2,5758$$
$$\sigma=\frac{0,20}{2,5758}=0,07764...\text{(mm)}$$

Vastaus: Hajonta enintään 0,07 mm

99. Pistemäärä on normaalisti jakautunut $N(30, 10)$.
Laudatur - arvosanan saa enintään 5 % osallistujista. On löydettävä sellainen pistemäärä x, että tätä vastaavaan normitettuun arvoon z mennessä on kertynyt vähintään 95 % osallistujista eli $\Phi(z) = 0{,}95$. Taulukkokirjasta saadaan

$$\Phi(1{,}6449) = 0{,}95$$

Kysyttyä pistemäärää x vastaa siis normitettu arvo $z = 1{,}6449$.

$$\begin{aligned} \frac{x-30}{10} &= 1{,}6449 \;\Big|\cdot 10 \\ x - 30 &= 16{,}449 \\ x &= 46{,}449 \end{aligned}$$

Jotta laudatur- arvosanojen osuus olisi enintään 5 %, niin x pyöristetään ylöspäin eli $x \approx 47$ (pistettä).

Vastaus: 47 pistettä

100. Lasketaan havupuun ($\overline{x} = 18\,\text{m}, s = 1{,}4\,\text{m}$) pituutta 20,2 m vastaava normitettu arvo.

$$z_{20,2} = \frac{20{,}2 - 18}{1{,}4} = \frac{2{,}2}{1{,}4} = 1{,}571...$$

Jotta heinä ($\bar{x} = 13\,\text{cm},\ s = 2{,}3\,\text{cm}$) olisi suhteessa yhtä pitkä, on sen pituuden normitettu arvo oltava yhtä suuri. Merkitään heinän pituutta kirjaimella x.

$$
\begin{aligned}
\frac{x-13}{2{,}3} &= \frac{2{,}2}{1{,}4} \\
1{,}4(x-13) &= 2{,}3 \cdot 2{,}2 \\
1{,}4x - 18{,}2 &= 5{,}06 \\
1{,}4x &= 23{,}26 \quad |:1{,}4 \\
x &= 16{,}614 \approx 17 \quad (\text{cm})
\end{aligned}
$$

Vastaus: Heinän tulisi olla 17 cm pitkä.

101. $x \sim N(0, 1)$
a) $P(x < 0{,}45) = \Phi(0{,}45) = 0{,}6736 \approx 0{,}67$

b) $P(x < -2)$ on symmetriasyistä yhtä suuri kuin normitetun arvon $z = 2$ jälkeinen kertymän osuus. Lasketaan ensin kertymä kohtaan $z = 2$ eli $\Phi(2) = 0{,}9772$. Tällöin saadaan

$$
\begin{aligned}
\Phi(-2) &= 1 - \Phi(2) \\
&= 1 - 0{,}9772 \\
&= 0{,}0228 \approx 0{,}023
\end{aligned}
$$

c) $P(-1 < x < 0{,}98) = \Phi(1) - \Phi(-1)$
Kertymän osuus kohtaan $z = -1$ on symmetriasyistä yhtä suuri kuin normitetun arvon $z = 1$ jälkeinen kertymän osuus eli

$$\Phi(-1) = 1 - \Phi(1) = 1 - 0{,}8413 = 0{,}1587$$

Taulukoitujen arvojen mukaan saadaan myös $\Phi(0{,}98) = 0{,}8365$.

Tällöin saadaan
$$\begin{aligned} P(-1 < x < 0{,}98) &= (0{,}98) - \Phi(-1) \\ &= 0{,}8365 - 0{,}1587 \\ &= 0{,}6778 \approx 0{,}68 \end{aligned}$$

Vastaus: a) 0,67 b) 0,023 c) 0,68

102. $x \sim N(45, 12)$

a) Lasketaan ensin arvoa 50 vastaava normitettu arvo.

$$z_{50} = \frac{50-45}{12} = 0{,}4166... \approx 0{,}42$$

Siis $P(x > 50) = P(z > 0{,}42)$.

Taulukosta saadaan kertymän arvo arvoon $z = 0{,}42$ mennessä eli

$$\Phi(0{,}42) = 0{,}6628$$

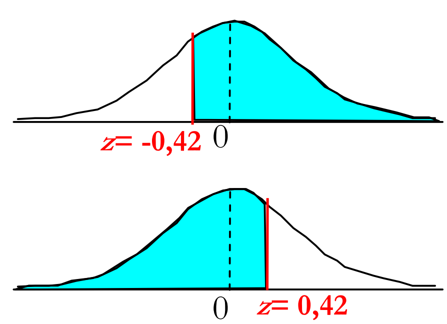
Tämän avulla saadaan

$$P(z > 0{,}42) = 1 - \Phi(0{,}42) = 1 - 0{,}6628 = 0{,}3372 \approx 0{,}34$$

b) Lasketaan ensin arvoa 40 vastaava normitettu arvo.

$$z_{50} = \frac{40-45}{12} = -0{,}4166... \approx -0{,}42$$

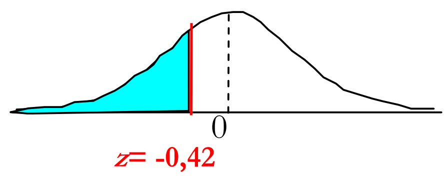
Siis $P(x > 40) = P(z > -0{,}42)$. Tämä tarkoittaa normitetun arvon $z = -0{,}42$ jälkeistä osuutta, joka on symmetrian nojalla yhtä suuri kuin arvoon $z = 0{,}42$ mennessä kertynyt osuus.



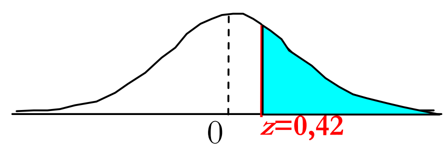
Taulukon avulla saadaan

$$P(z > -0{,}42) = \Phi(0{,}42) = 0{,}6628 \approx 0{,}66$$

c) $P(x < 40) = P(z < -0{,}42) = \Phi(-0{,}42)$



Jakauman symmetrisyydestä johtuen tämä on yhtä suuri kuin normitetun arvon $z = 0{,}42$ jälkeinen osuus.



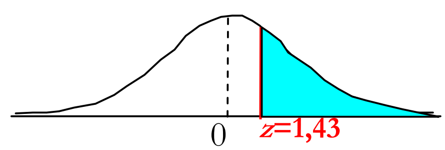
$$\begin{aligned}\text{Siis } \Phi(-0{,}42) &= 1 - \Phi(0{,}42) \\ &= 1 - 0{,}6628 \\ &= 0{,}3372 \approx 0{,}34\end{aligned}$$

Vastaus: a) 0,34 b) 0,66 c) 0,34

103. Pistemäärä x on normaalisti jakautunut $x \sim N(25; 3{,}5)$, joten $\mu = 25$ ja $\sigma = 3{,}5$.

a) Normitetaan pistemäärä 30.

$$z_{30} = \frac{30-25}{3{,}5} = \frac{5}{3{,}5} = 1{,}4285... \approx 1{,}43$$



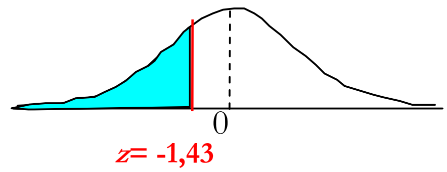
Todennäköisyys, että saadaan yli 30 pistettä on yhtä suuri kuin kertymän osuus normitetun arvon $z = 1{,}43$ jälkeen.

$$\begin{aligned}P(\text{yli 30 pistettä}) &= 1 - \Phi(1{,}43) \\ &= 1 - 0{,}9236 \\ &= 0{,}0764 \approx 0{,}076\end{aligned}$$

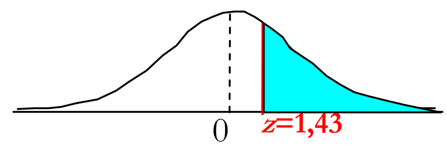
b) Normitetaan pistemäärä 20.

$$z_{20} = \frac{20-25}{3{,}5} = \frac{-5}{3{,}5} = -1{,}4285... \approx -1{,}43$$

Todennäköisyys, että saadaan alle 20 pistettä on yhtä suuri kuin kertymän osuus normitettuun arvoon $z = -1{,}43$ mennessä.



Jakauman symmetrisyydestä johtuen tämä tarkoittaa samaa kuin kertymä arvon $z = 1{,}43$ jälkeen.



$$\begin{aligned} P(\text{alle 20 pistettä}) &= \Phi(-1{,}43) \\ &= 1-\Phi(1{,}43) \\ &= 1-0{,}9236 \\ &= 0{,}0764 \approx 0{,}076 \end{aligned}$$

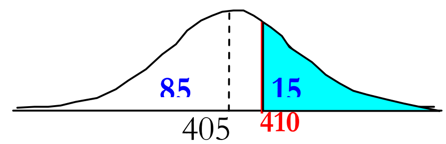
c) Kysytty todennäköisyys on $P(-1{,}43 \le z \le 1{,}43)$. Edellisten kohtien perusteella saadaan

$$\begin{aligned} P(20-30\text{ pistettä}) &= \Phi(1{,}43)-\Phi(-1{,}43) \\ &= 0{,}9236-0{,}0764 \\ &= 0{,}8472 \approx 0{,}85 \end{aligned}$$

Vastaus: a) 0,076 b) 0,076 c) 0,85

104. Purkin massa x noudattaa normaalijakaumaa $x \sim N(405 \text{ g}, \sigma)$.

Jos purkin massa on yli 410 g 15 % todennäköisyydellä, niin silloin massa on alle 410 g todennäköisyydellä 100 % – 15 % = 85 % = 0,85.



Etsitään taulukkokirjasta 85 % = 0,85 kertymää vastaava normitettu arvo, $\Phi(1{,}04) = 0{,}8508 \approx 0{,}85$. Tällöin saadaan

$$z_{410} = \frac{410-405}{\sigma} = 1{,}04$$

$$\frac{5}{\sigma} = 1{,}04$$

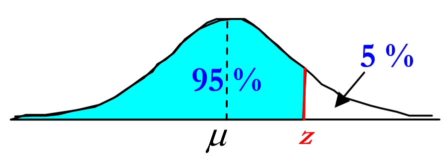
$$1{,}04\sigma = 5 \,|: 1{,}04$$

$$\sigma = \frac{5}{1{,}04} = 4{,}80769... \approx 4{,}8\,(\text{g})$$

Vastaus: 4,8 g

105. Pistemäärä noudattaa normaalijakaumaa, $x \sim N(27{,}36;\, 12{,}23)$.

a) Merkitään alimman pistemäärän (jolla saadaan laudatur) normitettua arvoa z. Jos laudatur- arvosanojen osuudeksi halutaan enintään 5 % kaikista kokelaista, niin kertymän arvo laudaturin alarajalle on $\Phi(z) \geq 0{,}95$.



Etsitään taulukkokirjasta normitettu arvo, jota vastaava kertymäfunktion arvo on suurempi kuin 0,95.

$$\Phi(1{,}65) = 0{,}9505$$

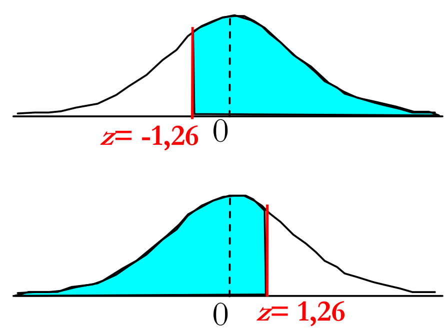
Näin ollen $z = 1{,}65$, jolloin vastaavaksi pistemääräksi saadaan

$$\frac{x-27{,}36}{12{,}23}=1{,}65 \;\Big|\cdot 12{,}23$$
$$x-27{,}36=20{,}1795$$
$$x=47{,}5395\approx 48 \text{ (pistettä)}$$

b) Normitetaan pistemäärä 12.

$$z_{12}=\frac{12-27{,}36}{12{,}23}=-1{,}255...\approx -1{,}26$$

Hyväksyttyjen kokelaiden osuutta kuvaa kertymä kohdan $z_{12} = -1{,}26$ jälkeen.
Jakauman symmetrisyydestä johtuen tämä tarkoittaa samaa kuin kertymä arvoon $z = 1{,}26$



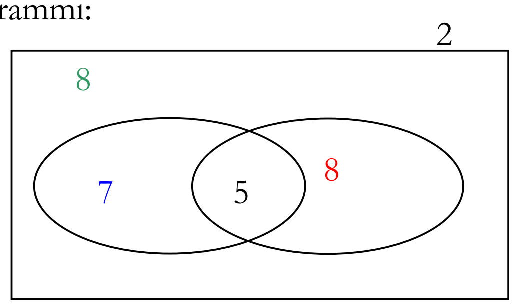
$$\Phi(1{,}26)=0{,}8962=89{,}62\ \%\approx 89{,}6\ \%$$

Vastaus: a) 48 pistettä b) 89,6 %

2.1 Malleja todennäköisyyden määrittämiseksi

106. Lukiolaisia on yhteensä 28 kpl.
Sekä taidemuseossa että eläintarhassa haluaa käydä 5.
Pelkästään taidemuseossa haluaa käydä 12 – 5 =7.
Pelkästään eläintarhassa haluaa käydä 13 – 5 = 8.
Lukiolaisia, jotka käyvät muualla kuin taidemuseossa tai eläintarhassa on
28 – 7 – 5 – 8 = 8.

Laaditaan Venn-diagrammi:



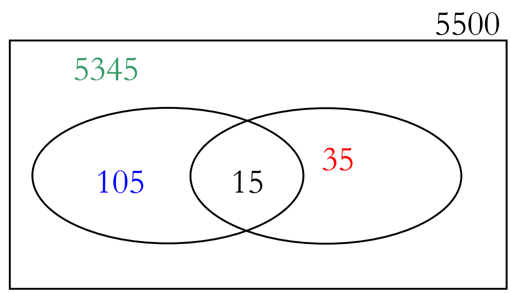
a) $P(\text{käy vain toisessa nähtävyydessä}) = \dfrac{7+8}{28} = \dfrac{15}{28} \approx 0,54$

b) $P(\text{ei käy kummassakaan nähtävyydessä}) = \dfrac{8}{28} \approx 0,29$

Vastaus: a) 0,54 b) 0,29

107. Myymälässä on 5500 tuotetta.
Luomuvihanneksia on 15 kpl.
Pelkästään vihanneksia (ei luomu) on 50 – 15 = 35 (kpl).
Kaikkiaan luomutuotteita on 120 (kpl).
Luomutuotteita, jotka eivät ole vihanneksia on 120 – 15 = 105 (kpl).
Tuotteita, jotka eivät ole vihanneksia eivätkä luomutuotteita on
5500 – 105 – 15 – 35 = 5345 (kpl).

Laaditaan Venn-diagrammi:



a) $P(\text{on luomutuote, muttei vihannes}) = \dfrac{105}{5500} \approx 0{,}019$

b) $P(\text{ei luomutuote, mutta vihannes}) = \dfrac{35}{5500} \approx 0{,}0064$

c) $P(\text{ei luomutuote, eikä vihannes}) = \dfrac{5345}{5500} \approx 0{,}97$

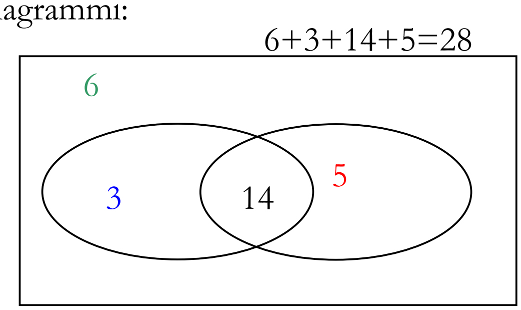
Vastaus: a) 0,019   b) 0,0064   c) 0,97

108. Naisjäsenistä töissä käy 17 (kpl).
Naisjäseniä, joilla on lapsia on 19 (kpl).
Naisia, joilla on lapsia ja jotka käyvät töissä, on 14 (kpl).

- naiset, joilla on lapsia, mutteivät käy töissä: 19 – 14 = 5
- naiset, jotka käyvät töissä, mutta joilla ei ole lapsia: 17 – 14 = 3

Naisia, joilla ei ole lapsia eivätkä käy töissä, on 6 (kpl).

Laaditaan Venn-diagrammi:



Kaikkiaan yhdistyksen naisjäseniä on siis 28 (kpl).

a) $P(\text{ei perheenäiti}) = \dfrac{6+3}{28} = \dfrac{9}{28} \approx 0{,}32$

b) $P(\text{ei töissä}) = \dfrac{6+5}{28} = \dfrac{11}{28} \approx 0{,}39$

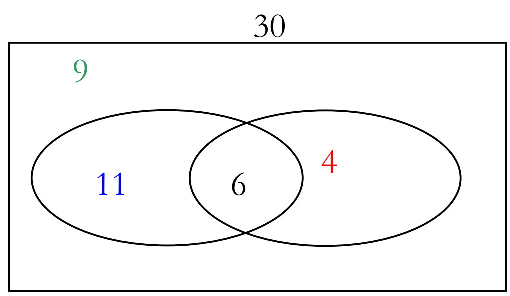
Vastaus: a) 0,32 b) 0,39

109. Ryhmässä on kaikkiaan 30 henkilöä.

- tennistä pelaa 17
- jääpalloa pelaa 10
- molempia pelaa 6

Henkilöitä, jotka eivät harrasta kumpaakaan lajia on: 30 – 10 – 6 – 17 = 9.

Laaditaan Venn-diagrammi:



a) $P(\text{ei harrasta kumpaakaan}) = \dfrac{9}{30} = \dfrac{3}{10}$

b) $P(\text{harrastaa vain toista lajia}) = \dfrac{11+4}{30} = \dfrac{15}{30} = \dfrac{1}{2}$

Vastaus: a) $\dfrac{3}{10}$ b) $\dfrac{1}{2}$

110. Alkeistapaukset (10 kpl) ovat:

ankka, hanhi	hanhi, porsas
ankka, porsas	porsas, nauta
ankka, nauta	nauta, jänis
ankka, jänis	hanhi, nauta
hanhi, jänis	porsas, jänis

a) $P(\text{tarjotaan jänistä}) = \dfrac{4}{10} = 0{,}4$

b) $P(\text{tarjotaan lintua}) = \dfrac{7}{10} = 0{,}7$

c) $P(\text{ei tarjota porsasta}) = \dfrac{6}{10} = 0{,}6$

Vastaus: a) 0,4 b) 0,7 c) 0,6

111. Alkeistapaukset ovat: (kr = kruuna, kl = klaava)

kr, kr, kr	kl, kl, kr
kl, kl, kl	kl, kr, kl
kr, kr, kl	kr, kl, kl
kr, kl, kr	kl, kr, kr

Alkeistapauksia on siis 8 kpl.

a) $P(\text{kolme klaavaa}) = \dfrac{1}{8}$

b) $P(\text{ainakin kaksi klaavaa}) = \dfrac{4}{8} = \dfrac{1}{2}$

Vastaus: a) $\dfrac{1}{8}$ b) $\dfrac{1}{2}$

112. Viisi opiskelijaa: T, H, L, X, Y

T, H, L	H, L, X
T, X, Y	H, L, Y
H, X, Y	T, L, X
L, X, Y	T, L, Y
T, H, X	T, H, Y

Alkeistapauksia on 10 kpl.

$$P(\text{kaikki loppukilpailussa, THL}) = \frac{1}{10} = 0{,}1$$

Vastaus: 0,1

113. Taulukoidaan kaikki mahdolliset tulokset kahden nopan heitossa.

1,6	2,6	3,6	4,6	5,6	6,6
1,5	2,5	3,5	4,5	5,5	6,5
1,4	2,4	3,4	4,4	5,4	6,4
1,3	2,3	3,3	4,3	5,3	6,3
1,2	2,2	3,2	4,2	5,2	6,2
1,1	2,1	3,1	4,1	5,1	6,1

Kaikkia alkeistapauksia on siis 36 kpl.

a) $P(\text{sama silmäluku}) = \frac{6}{36} = \frac{1}{6} \approx 0{,}17$

b) $P(\text{summa vähintään 9}) = \frac{10}{36} = \frac{5}{18} \approx 0{,}28$

c) $P(\text{ainakin toinen vähintään 5}) = \frac{20}{36} = \frac{5}{9} \approx 0{,}56$

Vastaus: a) 0,17 b) 0,28 c) 0,56

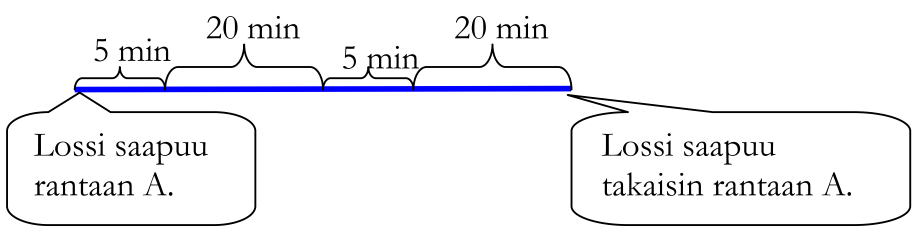
114. Alkeistapauksia on 36 kpl. Pistesummat ovat taulukossa.

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

Silmälukujen summa saa arvot 2 … 12. Taulukon mukaan silmälukujen summa 7 esiintyy useimmin, joten se on todennäköisin.

Vastaus: 7

115.



$$P(\text{suoraan lossiin rannasta A}) = \frac{5}{50} = 0{,}1$$

Vastaus: 0,1

116. Tikkataulun renkaiden (9 kpl) leveys on 2 cm. Kymppi-ympyrän säde on 1 cm. Tikkataulun säde *r* on siis $9 \cdot 2 + 1 = 19$ (cm).

a) $P(\text{saa } 10) = \dfrac{\pi \cdot 1^2}{\pi \cdot 19^2} \approx 0{,}0028$

b) $P(\text{saa vähintään } 8) = \dfrac{\pi \cdot 5^2}{\pi \cdot 19^2} \approx 0{,}069$

c) $P(\text{lukujen 4 ja 5 väliselle kehälle}) = \dfrac{0}{\pi \cdot 19^2} = 0$

Vastaus: a) 0,0028 b) 0,069 c) 0

117. Kattilan pohjan säde r = 7,5 cm. Puurosta syödään 8 cm:n korkuinen kerros, jonka tilavuus on

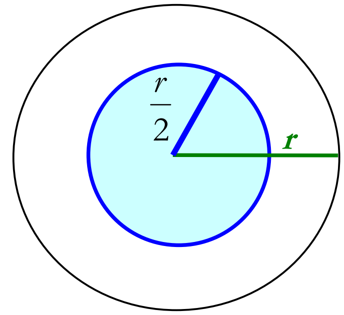
$$V = \pi \cdot 7{,}5^2 \cdot 8 = 1413{,}7...(\text{cm}^3) = 1{,}4137...(\text{dm})$$

Kattilaan jääneen puuron tilavuus on $V = 4{,}5\,\text{l} - 1{,}4137...\,\text{l} = 3{,}086...\,\text{l}$

$$P(\text{manteli kattilassa}) = \frac{3{,}086...\,\text{l}}{4{,}5\,\text{l}} = 0{,}6858... \approx 0{,}69$$

Vastaus: 0,69

118. Merkitään ympyrän sädettä kirjaimella r.



Kivi osuus lähemmäksi keskipistettä, jos se osuus samankeskisen $\frac{r}{2}$-säteisen ympyrän sisään.

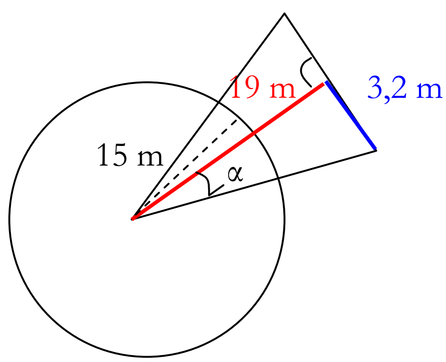
Todennäköisyys siis on

$$\frac{\pi \cdot \left(\frac{r}{2}\right)^2}{\pi \cdot r^2} = \frac{\frac{r^2}{4}}{r^2} = \frac{r^2}{4} \cdot \frac{1}{r^2} = \frac{1}{4} = 0{,}25$$

Vastaus: 0,25

119. Kuvassa on tilanne ylhäältä päin katsottuna (puun kaaduttua).

Koska henkilön etäisyys on pienempi kuin puun korkeus, puu osuus henkilöön, jos hän seisoo kaatuneen puun muodostamassa sektorissa.



Merkitään sektorin keskuskulman puolikasta kirjaimella α.

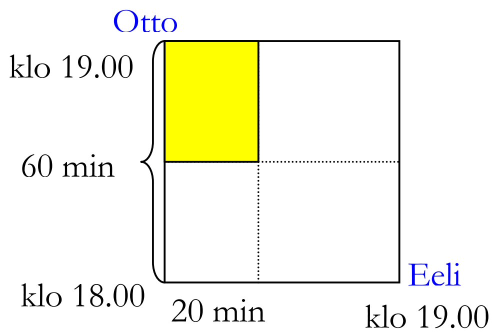
$$\tan \alpha = \frac{3{,}2}{19} = 0{,}168...$$
$$\alpha = 9{,}56...°$$

Keskuskulma on $2\alpha = 19{,}120\ldots° \approx 19{,}1°$.

Todennäköisyys, että henkilö seisoo sektorissa on $\frac{19{,}1°}{360°} \approx 0{,}053$.

Vastaus: 0,053

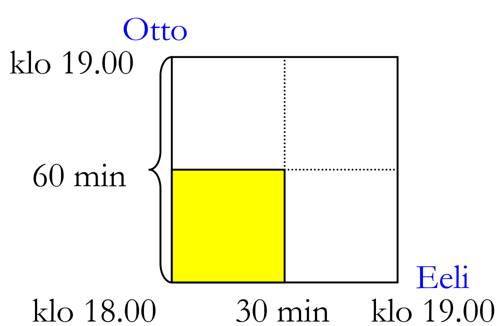
120. a)



Koko neliön pinta-ala on $A = 60^2 = 3600$.
Suotuisan alueen pinta-ala on $A = 20 \cdot 30 = 600$.

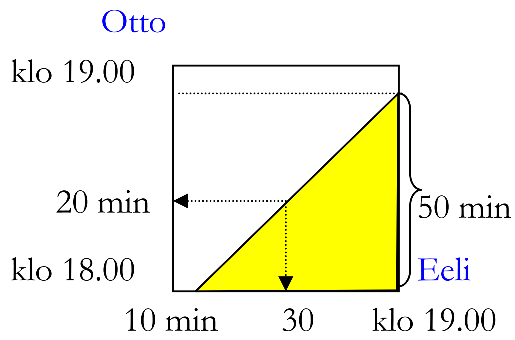
$$P(\text{Eeli ennen 18.20, Otto 18.30 jälkeen}) = \frac{600}{3600} = \frac{1}{6} \approx 0{,}17$$

b)



$$P(\text{molemmat ennen 18.30}) = \frac{30^2}{3600} = \frac{900}{3600} = \frac{1}{4} = 0{,}25$$

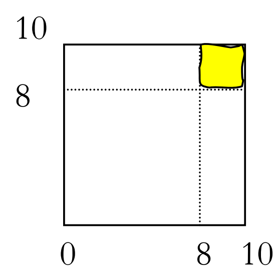
c)



$$P(\text{Otto odottaa yli 10 min}) = \frac{\frac{1}{2} \cdot 50 \cdot 50}{3600} = \frac{1250}{3600} \approx 0{,}35$$

Vastaus: a) 0,17 b) 0,25 c) 0,35

121. Suotuisat alkeistapaukset kuuluvat tummennetun neliön alueelle.



Tummennetun neliön pinta-ala on

$$A = 2^2 = 4$$

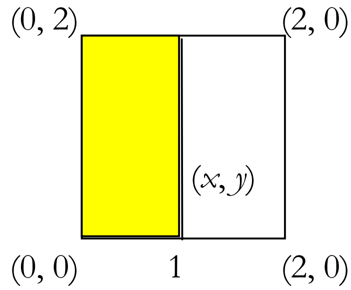
Kaikkia alkeistapauksia kuvaavan ison neliön ala on

$$A = 10^2 = 100$$

$$P(\text{arvotut luvut suurempia kuin 8}) = \frac{4}{100} = 0{,}04$$

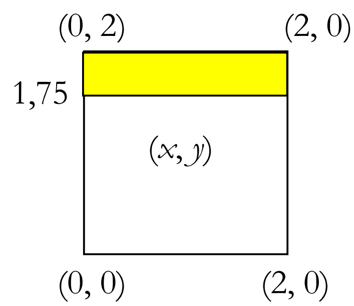
Vastaus: 0,04

122. a)



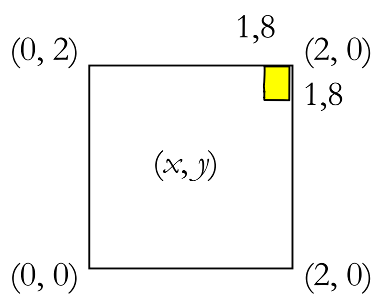
$$P(x\text{-koordinaatti pienempi kuin 1}) = \frac{1 \cdot 2}{2^2} = \frac{2}{4} = \frac{1}{2} = 0{,}5$$

b)



$$P(y\text{-koordinaatti suurempi kuin 1,75}) = \frac{0{,}25 \cdot 2}{2^2} = \frac{0{,}5}{4} = \frac{1}{8} = 0{,}125$$

c)



$$P(x\text{- ja } y\text{-koordinaatit suurempia kuin 1,8}) = \frac{0{,}2^2}{4} = \frac{0{,}04}{4} = 0{,}01$$

Vastaus: a) 0,5 b) 0,125 c) 0,01

123. **a)** $P(\text{alle 200 h}) = \dfrac{245+8}{500} = \dfrac{253}{500} \approx 0{,}51$

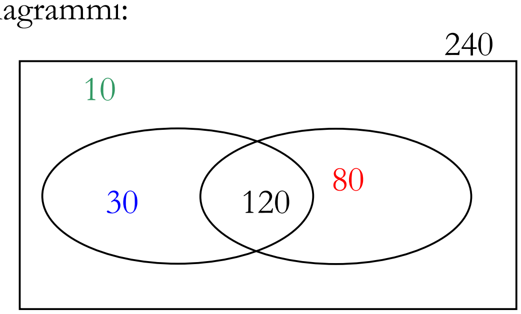
b) $P(\text{kestää vielä 100 h}) = \dfrac{235+12}{500-8} = \dfrac{247}{492} \approx 0{,}50$

Vastaus: a) 0,51 b) 0,50

124. Maatilalla on 240 lintua.

- 200 valkoista lintua
- kanoja 150
- valkoisia kanoja 120
- kanoja, jotka eivät ole valkoisia 150 – 120 = 30
- valkoisia lintuja, jotka eivät ole kanoja 200 – 120 = 80
- lintuja, jotka eivät ole valkoisia eivätkä ole kanoja
  240 – 80 – 120 – 30 = 10

Laaditaan Venn-diagrammi:



a) $P(\text{valkoinen lintu, ei kana}) = \dfrac{80}{240} = \dfrac{1}{3}$

b) $P(\text{kana,ei valkoinen lintu}) = \dfrac{30}{240} = \dfrac{1}{8}$

c) $P(\text{ei kana,ei valkoinen lintu}) = \dfrac{10}{240} = \dfrac{1}{24}$

Vastaus: a) $\dfrac{1}{3}$ b) $\dfrac{1}{8}$ c) $\dfrac{1}{24}$

125. Taidenäyttelyssä on maalauksia 112 (kpl).

Merkitään ei-impressionististen, ranskalaisten maalausten lukumäärää x.
Impressionistisiä on tällöin myös x (kpl).

- näistä ranskalaisia on $\frac{1}{4}x$

Ei-impressionistisiä ja ei-ranskalaisia 80 (kpl)

Koska tauluja on yhteensä 112, niin

$$\begin{aligned} x + x + 80 &= 112 \\ 2x &= 32 \quad |:2 \\ x &= 16 \end{aligned}$$

a) $P(\text{impressionistinen}) = \dfrac{x}{112} = \dfrac{16}{112} = \dfrac{1}{7} \approx 0{,}14$

b) $P(\text{ranskalainen}) = \dfrac{x + \frac{1}{4}x}{112} = \dfrac{\frac{5}{4}x}{112} = \dfrac{20}{112} = \dfrac{5}{28} \approx 0{,}18$

c) $P(\text{ranskalainen ja impressionist.}) = \dfrac{\frac{1}{4}x}{112} = \dfrac{4}{112} = \dfrac{1}{28} \approx 0{,}04$

Vastaus: a) $\dfrac{1}{7} \approx 0{,}14$ b) $\dfrac{5}{28} \approx 0{,}18$ c) $\dfrac{1}{28} \approx 0{,}04$

126. Havainnollistetaan alkeistapauksia koordinaatistossa.

127. Merkitään SDP:n kansanedustajia S_1, S_2
Kokoomuksen kansanedustajia K_1, K_2, K_3

Taulukoidaan kaikki mahdolliset arvonnan tulokset:

S_1, S_2	S_1, K_1	S_1, K_2	S_1, K_3
S_2, K_1	S_2, K_2	S_2, K_3	
K_1, K_2	K_2, K_3	K_1, K_3	

Erilaisia yhdistelmiä on siis 10 kpl.

a) Tapahtumia, joissa kaksi Kokoomuksen kansanedustajaa pääsee hissiin, on 3.

$$P(\text{kaksi Kokoomuksen edustajaa}) = \frac{3}{10}$$

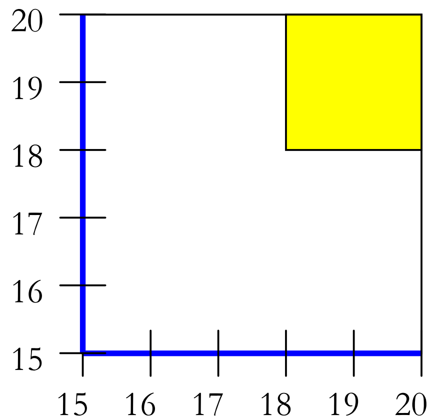
b) Tapahtumia, joissa kaksi SDP:n edustajaa pääsee hissiin, on 1.

$$P(\text{kaksi SDP:n edustajaa}) = \frac{1}{10}$$

Vastaus: a) $\frac{3}{10}$ b) $\frac{1}{10}$

128. Tutkitaan alkeistapauksia geometrisesti.

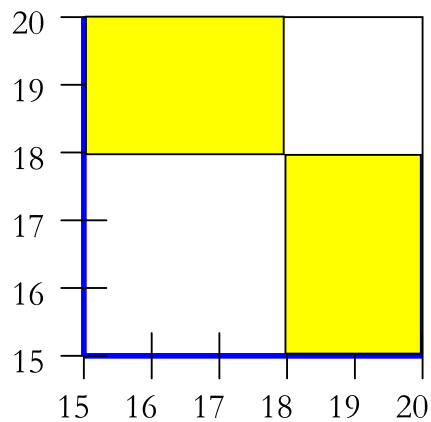
a)



Kaikkia alkeistapauksia kuvaa neliö, jonka ala on $A = 5 \cdot 5 = 25$
Suotuisia tapauksia kuvaa kuvioon muodostuvan, värillisen neliön ala $A = 2 \cdot 2 = 4$

$$P(\text{molemmat täyttäneen 18 v}) = \frac{4}{25}$$

b)



Kaikkia alkeistapauksia kuvaa neliö, jonka ala on $A = 5 \cdot 5 = 25$
Suotuisia tapauksia kuvaa kuvioon muodostuvan, värillisen alueen ala
$A = 2(2 \cdot 3) = 12$

$$P(\text{toinen 18 v, toinen ei}) = \frac{12}{25}$$

Vastaus: a) $\frac{4}{25}$ b) $\frac{12}{25}$

2.2 Komplementti ja kertolaskusääntö

129. P(ukkostaa la) = 0,4
P(ukkostaa su) = 0,3

Kertolaskusäännön mukaan saadaan P(ukkostaa la ja su) = 0,4 · 0,3 = 0,12

Vastaus: 0,12

130. Laatikossa 1 on: 1 keltainen, 3 sinistä ja 2 punaista palloa. Yhteensä 6 palloa.
Laatikossa 2 on: 4 keltainen, 1 sininen, 3 punaista palloa. Yhteensä 8 palloa.

$$P(\text{1. laatikosta punainen pallo}) = \frac{2}{6} = \frac{1}{3}$$

$$P(\text{2. laatikosta punainen pallo}) = \frac{3}{8}$$

$$P(\text{molemmista laatikoista punainen pallo}) = \frac{1}{3} \cdot \frac{3}{8} = \frac{1}{8} = 0{,}125$$

Vastaus: $\frac{1}{8} = 0{,}125$

131. P(Ilari saa häränsilmän) = 0,25
P(Matti saa häränsilmän) = 0,18

a) P(Matti ja Ilari saavat häränsilmän) = 0,18 · 0,25 = 0,045

b) P(Matti ei osu) = 1 − 0,18 = 0,82
P(Matti ei osu kertaakaan) = $0{,}82^3 \approx 0{,}55$

Vastaus: a) 0,045 b) 0,55

132. a) $P(\text{kuutonen}) = \dfrac{1}{6}$

$$P(\text{neljä kertaa kuutonen}) = \left(\frac{1}{6}\right)^4 \approx 0{,}00077$$

b) $P(\text{viisi tai kuusi}) = \dfrac{2}{6}$

$$\begin{aligned} P(\text{kaikilla kerroilla vähintään 5}) &= P(\text{kaikilla kerroilla 5 tai 6}) \\ &= \left(\frac{2}{6}\right)^4 \approx 0{,}012 \end{aligned}$$

Vastaus: a) 0,00077 b) 0,012

133. Verkkopankkiin kirjautuminen: 1) käyttäjätunnus (6 numeroa)
2) salasana (4 numeroa)
3) kertakäyttötunnus (4 numeroa)

Numeroita on käytettävissä 0–9 eli 10 kpl jokaiselle numeropaikalle.

a) Mahdollisia käyttäjätunnuksia on 1 000 000 kpl. Olemassa olevia tunnuksia (asiakkaita) on 600 000.

$$P(\text{osuu olemassa olevaan käyttäjätunnukseen}) = \frac{600\,000}{1\,000\,000} = 0{,}6$$

b) Yhdellä yrityksellä pääsee kirjautumaan verkkopankkiin, jos

arvaa oikein jonkin 600 000:sta **käyttäjätunnuksesta ja**
sen jälkeen arvaa oikein **salasanan ja**
arvaa oikein **kertakäyttötunnuksen**.

$$P(\text{arvaa oikein salasanan}) = \left(\frac{1}{10}\right)^4$$

$$P(\text{arvaa oikein kertakäyttötunnuksen}) = \left(\frac{1}{10}\right)^4$$

Todennäköisyys, että yhdellä yrityksellä pääsee kirjautumaan on

$$P(\text{yhdellä yrityksellä verkkopankkiin}) = 0{,}6 \cdot \left(\frac{1}{10}\right)^4 \cdot \left(\frac{1}{10}\right)^4 = 6 \cdot 10^{-9}$$

Vastaus: a) 0,6 b) $6 \cdot 10^{-9}$

134. Tunnusluvussa on neljä numeroa. Jokaiselle numeropaikalle on 10 vaihtoehtoa (0–9).

a) $P(\text{arvaa oikein yhden numeron}) = \dfrac{1}{10}$

$$P(\text{näppäilee oikein koko tunnusluvun}) = \left(\frac{1}{10}\right)^4 = 0{,}0001$$

b) $P(\text{ei arvaa numeroa oikein}) = \dfrac{9}{10}$

$$\begin{aligned} P(\text{näppäilee ainakin yhden numeron oikein}) &= 1 - P(\text{ei yhtään oikein}) \\ &= 1 - \left(\frac{9}{10}\right)^4 \\ &\approx 0{,}34 \end{aligned}$$

Vastaus: a) 0,0001 b) 0,34

135. P(asuu pk-seudulla) = 0,20
P(asuu muualla) = 1 – 0,20 = 0,80

$$\begin{aligned} P(\text{ainakin yksi kuudesta pk-seudulta}) &= 1 - P(\text{ei yksikään pk-seudulta}) \\ &= 1 - 0{,}80^6 \\ &= 0{,}7378... \\ &\approx 0{,}74 \end{aligned}$$

Vastaus: 0,74

136. $P(\text{paikka vapaana}) = \dfrac{6}{60} = 0{,}1$
P(paikka varattu) = 1 – 0,1 = 0,9

Parkkipaikkoja 10 kappaletta.

a) P(ei yhtään paikkaa vapaana) = $0{,}9^{10}$ = 0,3486… ≈ 0,35

b) Jos autoilija löytää vapaan paikan, niin ainakin yhden paikan on oltava vapaana. Sen komplementti on ”ei yhtään paikkaa vapaana” eli kohdan a todennäköisyys.

$$\begin{aligned} P(\text{ainakin yksi paikka vapaa}) &= 1 - P(\text{ei yhtään paikkaa vapaana}) \\ &= 1 - 0{,}3486\ldots \\ &= 0{,}6513\ldots \\ &\approx 0{,}65 \end{aligned}$$

Vastaus: a) 0,35 b) 0,65

137. Kukkasipulipakkauksessa on 6 sipulia.
P(sipuli itää) = 0,8
P(sipuli ei idä) = 1 – 0,8 = 0,2

a) P(kaikki 6 sipulia itää) = $0{,}8^6 \approx 0{,}26$

b)

$$\begin{aligned} P(\text{ainakin yksi itää}) &= 1 - P(\text{ei yksikään idä}) \\ &= 1 - 0{,}2^6 \\ &= 0{,}999936... \\ &\approx 0{,}99994 \end{aligned}$$

Vastaus: a) 0,26 b) 0,99994

138. P(bitti virheellinen) = 0,00015
P(bitti oikein) = 1 – 0,00015 = 0,99985

a) Jonossa on 16 bittiä.

$$\begin{aligned} P(\text{ainakin yksi virheellinen}) &= 1 - P(\text{kaikki oikein}) \\ &= 1 - 0{,}99985^{16} \\ &\approx 0{,}0024 \end{aligned}$$

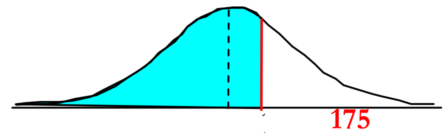
b) Lähetetään 32 kpl bittijonoja, joissa jokaisessa on 16 bittiä.

$$P(\text{yksi bittijono oikein}) = 0{,}99985^{16} = 0{,}9976...$$

$$\begin{aligned} P(\text{ainakin yksi virheellinen jono}) &= 1 - P(\text{kaikki jonot oikein}) \\ &= 1 - 0{,}9976...^{32} \\ &\approx 0{,}074 \end{aligned}$$

Vastaus: a) 0,0024 b) 0,074

139. Pituus $x \sim N(165\text{ cm}, 6\text{ cm})$



Todennäköisyys, että tyttö on korkeintaan 175 cm pitkä, on kuvioon merkityn pinta-alan arvo.

Normitetaan pituus 175 cm.

$$z_{175} = \frac{175-165}{6} = \frac{10}{6} = 1{,}666... \approx 1{,}67$$

Todennäköisyys, että tyttö on alle 175 cm, saadaan kertymäfunktion arvona eli

$$\Phi(1{,}67) = 0{,}9525 \approx 0{,}95$$

$$P(\text{kaikki kolme alle 175 cm}) = 0{,}9525^3 = 0{,}8641... \approx 0{,}86$$

$$\begin{aligned} P(\text{ainakin yksi yli 175cm}) &= 1 - P(\text{kaikki alle 175 cm}) \\ &= 1 - 0{,}9525^3 \\ &= 0{,}13583... \\ &\approx 0{,}14 \end{aligned}$$

140. Korissa on 8 vihreää ja 17 punaista omenaa eli yhteensä 25 omenaa. Korista otetaan neljä omenaa.

a)

$$P(\text{kaikki punaisia}) = \frac{17}{25} \cdot \frac{16}{24} \cdot \frac{15}{23} \cdot \frac{14}{22} = \frac{238}{1265} \approx 0{,}19$$

b)

$$\begin{aligned} P(\text{ainakin yksi vihreä}) &= 1 - P(\text{kaikki punaisia}) \\ &= 1 - \frac{238}{1265} \\ &= 0{,}8118... \\ &\approx 0{,}81 \end{aligned}$$

c)

$$
\begin{aligned}
P(\text{ainakin yksi punainen}) &= 1 - P(\text{kaikki vihreitä}) \\
&= 1 - \left(\frac{8}{25}\cdot\frac{7}{24}\cdot\frac{6}{23}\cdot\frac{5}{22}\right) \\
&= 1 - \frac{7}{1265} \\
&= 0{,}99446... \\
&\approx 0{,}994
\end{aligned}
$$

Vastaus: a) 0,19 b) 0,81 c)0,994

141. Kokoelmassa 12 levyä, joista 5 sinfoniaa, 4 jazzia ja 3 oopperaa.

$$P(\text{kaksi jazz - levyä}) = \frac{4}{12}\cdot\frac{3}{11} = \frac{1}{11} \approx 0{,}091$$

Vastaus: 0,091

142. Pelissä pyritään saamaan kolmella kortilla yhteispistemääräksi 21.

$$P(\text{saadaan } 7) = \frac{4}{52}$$

$$P(\text{saadaan toinen } 7) = \frac{3}{51}$$

$$P(\text{saadaan kolmas } 7) = \frac{2}{50}$$

Todennäköisyys, että saadaan pistemäärä 21 kolmella seiskalla

$$P(\text{saadaan 7 ja 7 ja 7}) = \frac{4}{52}\cdot\frac{3}{51}\cdot\frac{2}{50} \approx 0{,}00018$$

Vastaus: 0,00018

143. a) Pakassa on 12 kuvakorttia, muita 40. Koska kortteja ei palauteta pakkaan, joka nostolla pakassa on yksi kortti vähemmän. Nostetaan viisi korttia.

$$
\begin{aligned}
&P(\text{ainakin yksi on kuvakortti}) \\
&= 1 - P(\text{ei yhtään kuvakorttia}) \\
&= 1 - \frac{40}{52} \cdot \frac{39}{51} \cdot \frac{38}{50} \cdot \frac{37}{49} \cdot \frac{36}{48} \\
&= 0{,}7468... \\
&\approx 0{,}75
\end{aligned}
$$

b) Pakassa on 13 ristiä, joten muita maita on yhteensä 39.

$$
\begin{aligned}
&P(\text{ainakin yksi on risti}) \\
&= 1 - P(\text{ei yhtään ristiä}) \\
&= 1 - \frac{39}{52} \cdot \frac{38}{51} \cdot \frac{37}{50} \cdot \frac{36}{49} \cdot \frac{35}{48} \\
&= 0{,}7784... \\
&\approx 0{,}78
\end{aligned}
$$

Vastaus: a) 0,75 b) 0,78

144. Myynnissä 50 arpaa, joista 30 voittoarpoja. Rasmus ostaa 4 arpaa.

a) $P(\text{kaikki arvat voittoarpoja}) = \frac{30}{50} \cdot \frac{29}{49} \cdot \frac{28}{48} \cdot \frac{27}{47} = 0{,}1189... \approx 0{,}12$

b) $P(\text{mikään ei voittoarpa}) = \frac{20}{50} \cdot \frac{19}{49} \cdot \frac{18}{48} \cdot \frac{17}{47} = 0{,}0210... \approx 0{,}012$

c)
$$
\begin{aligned}
P(\text{ainakin yksi voittoarpa}) &= 1 - P(\text{mikään ei voittoarpa}) \\
&= 1 - 0{,}0210... \\
&= 0{,}978... \\
&\approx 0{,}98
\end{aligned}
$$

d) $$\begin{aligned} P(\text{ainakin yksi sellainen, jolla ei voita}) &= 1 - P(\text{kaikki voittoarpoja}) \\ &= 1 - 0{,}1189... \\ &= 0{,}8811... \\ &\approx 0{,}88 \end{aligned}$$

Vastaus: a) 0,12 b) 0,021 c) 0,98 d) 0,88

145. Merkitään pilaantuneita kirsikoita kirjaimella x.

$$P(\text{I pilaantunut}) = \frac{x}{50}$$

$$P(\text{II pilaantunut}) = \frac{x-1}{49}$$

$$P(\text{molemmat pilaantuneita}) = \frac{3}{175}$$

$$\begin{aligned} \frac{x}{50} \cdot \frac{x-1}{49} &= \frac{3}{175} \\ \frac{x^2 - x}{2450} &= \frac{3}{175} \\ 175x^2 - 175x - 7350 &= 0 \qquad \Big| :175 \\ x^2 - x - 42 &= 0 \\ x &= \frac{1 \pm \sqrt{1 - 4 \cdot 1 \cdot (-42)}}{2} \\ &= \frac{1 \pm \sqrt{169}}{2} = \frac{1 \pm 13}{2} \\ x = \frac{14}{2} = 7 \quad &\text{tai} \quad x = \frac{-12}{2} = -6 \end{aligned}$$

Ratkaisuksi kelpaa vain positiivinen luku, joten $x = 7$.

Vastaus: 7 (kirsikkaa)

146. Merkitään P(arpa voittaa) = x
Molemmat arvat voittavat todennäköisyydellä 0,34 eli

$$P(\text{I ja II voittaa}) = x \cdot x = 0,34$$
$$x^2 = 0,34 \quad \Big|\sqrt{}$$
$$x = \pm\sqrt{0,34}$$
$$x \approx \pm 0,58$$

Vastauksista vain positiivinen kelpaa todennäköisyydeksi.

Vastaus: 0,58

147. Merkitään hävinneiden korttien lukumäärää kirjaimella x.

$$P(\text{I kortti on pata}) = \frac{13-x}{52-x}$$
$$P(\text{II kortti on pata}) = \frac{13-x-1}{52-x-1} = \frac{12-x}{51-x}$$
$$P(\text{saadaan kaksi pataa}) = \frac{13-x}{52-x} \cdot \frac{12-x}{51-x} = \frac{3}{94}$$

$$(13-x)(12-x)94 = (52-x)(51-x)3$$
$$(156-13x-12x+x^2)94 = (2652-52x-51x+x^2)3$$
$$14664-2350x+94x^2 = 7956-309x+3x^2$$
$$91x^2-2041x+6708 = 0$$
$$x = \frac{2041 \pm \sqrt{4165681-4\cdot 91\cdot 6708}}{2\cdot 91}$$
$$= \frac{2041 \pm \sqrt{1723969}}{182}$$
$$= \frac{2041 \pm 1313}{182}$$
$$x = \frac{3354}{182} = 18,428... \quad \text{tai} \quad x = \frac{728}{182} = 4$$

Vain kokonaislukuratkaisu kelpaa vastaukseksi, joten $x = 4$.

Vastaus: Patoja on hävinnyt 4 kpl.

148. $P(\text{vasenkätiset}) = \frac{10}{100} = \frac{1}{10}$

$$P(\text{oikeakätiset}) = 1 - \frac{1}{10} = \frac{9}{10}$$

Merkitään henkilöiden lukumäärää x.

$$P(\text{kaikki oikeakätiset}) = \left(\frac{9}{10}\right)^x$$

$$\begin{aligned} P(\text{ainakin yksi vasenkätinen}) &= 1 - P(\text{kaikki oikeakätiset}) \\ &= 1 - \left(\frac{9}{10}\right)^x \end{aligned}$$

Ryhmässä on ainakin yksi vasenkätinen on todennäköisyydellä

$$1 - \left(\frac{9}{10}\right)^x = 0{,}8$$

Kokeilemalla kokonaislukuja ratkaisuiksi, saadaan henkilömääräksi

$$x = 15 \qquad 1 - \left(\frac{9}{10}\right)^{15} = 0{,}79410.. < 0{,}8$$

$$x = 16 \qquad 1 - \left(\frac{9}{10}\right)^{16} = 0{,}8146.. > 0{,}8$$

Vastaus: 16 henkilöä

149. $P(\text{I valot vihreät}) = \frac{40}{60}$

$P(\text{II valot vihreät}) = \frac{35}{60}$

$P(\text{III valot vihreät}) = \frac{28}{60}$

$$P(\text{I ja II ja III valot vihreät}) = \frac{40}{60} \cdot \frac{35}{60} \cdot \frac{28}{60} \approx 0{,}18$$

Vastaus: 0,18

150. Jokerittomassa pakassa on $4 \cdot 13 = 52$ korttia (neljä maata).
Pieniä kortteja on pakassa $4 \cdot 6 = 24$ (kpl).
Suuria kortteja on pakassa $4 \cdot 6 = 24$ (kpl).
Pelaaja häviää, jos hän nostaa numeron 7, joita on pakassa 4 kpl.

a) Pelaaja lyö vetoa aina suuren kortin puolesta ja ”tuplaa” viisi kertaa.

$$P(\text{suuri kortti}) = \frac{24}{52}$$

$$P(\text{suuri kortti viidessä tuplauksessa}) = \left(\frac{24}{52}\right)^5 = 0{,}0209... \approx 0{,}021$$

b)

$$P(\text{nostetaan 7}) = \frac{4}{52}$$

$$P(\text{nostetaan 7 viisi kertaa}) = \left(\frac{4}{52}\right)^5 = 0{,}00000269... \approx 0{,}0000027$$

Vastaus: a) 0,021 b) 0,0000027

151. $P(\text{säävahinko}) = \dfrac{1}{5}$

$P(\text{hirvivahinko}) = \dfrac{2}{6} = \dfrac{1}{3}$

$$
\begin{aligned}
P(\text{ainakin jompikumpi tuhoista}) &= 1 - P(\text{ei kumpikaan tuhoista}) \\
&= 1 - \frac{4}{5} \cdot \frac{2}{3} \\
&= \frac{7}{15} \\
&= 0,4666... \approx 0,47
\end{aligned}
$$

Vastaus: 0,47

152. $P(\text{siemen itää}) = 0,60$

$P(\text{siemen ei idä}) = 1 - 0,60 = 0,40$

a) Siemeniä istutetaan 3 kpl.

$P(\text{mikään siemen ei idä}) = 0,40^3 = 0,064$

$$
\begin{aligned}
P(\text{ainakin yksi itää}) &= 1 - P(\text{mikään siemen ei idä}) \\
&= 1 - 0,064 \\
&= 0,936
\end{aligned}
$$

b)

$P(\text{jokaisessa viidessä ruukussa ainakin yksi siemen itää})$

$= P(\text{1.ruukussa ainakin yksi itää ja...ja 5. ruukussa ainakin yksi itää})$

$= 0,936^5$

$= 0,7184...$

$\approx 0,718$

Vastaus: a) 0,064; 0,936 b) 0,718

153. Pakassa on 52 korttia, joista 4 on ässiä.

$$P(\text{ässä}) = \frac{4}{52} = \frac{1}{13}$$

a) $P(\text{neljä ässää}) = \left(\frac{1}{13}\right)^4 \approx 0{,}000035$

b) $P(\text{neljä ässää}) = \frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} \cdot \frac{1}{49} \approx 3{,}7 \cdot 10^{-6}$

Vastaus: a) 0,000035 b) $3{,}7 \cdot 10^{-6}$

154. Merkitään oppilaiden määrää kirjaimella x. Tyttöjä on 15 kpl.

$$P(\text{kaksi tyttöä}) = \frac{1}{6}$$

$$\frac{15}{x} \cdot \frac{14}{x-1} = \frac{1}{6}$$

$$\frac{210}{x^2 - x} = \frac{1}{6}$$

$$x^2 - x = 1260$$

$$x^2 - x - 1260 = 0$$

$$x = \frac{1 \pm \sqrt{1 - 4 \cdot 1 \cdot (-1260)}}{2}$$

$$= \frac{1 \pm \sqrt{5041}}{2} = \frac{1 \pm 71}{2}$$

$$x = \frac{72}{2} = 36 \quad \text{tai} \quad x = \frac{-70}{2} = -35$$

Ratkaisuksi kelpaa vain positiivinen luku eli $x = 36$.
Koska tyttöjä on 15 kpl, niin poikia on 36 – 15 = 21 kpl.

Vastaus: 21 poikaa

2.3 Yhteenlaskusääntö

155. Erillisiä tapahtumia: a, b, c, e

156. Korttipakan 52 kortista herttoja on 13, patoja 13 ja ässiä 4.

a)

$$\begin{aligned}
&P(\text{saadaan hertta tai pata})\\
&= P(\text{saadaan hertta}) + P(\text{saadaan pata})\\
&= \frac{13}{52} + \frac{13}{52}\\
&= \frac{1}{2} = 0{,}50
\end{aligned}$$

b)

$$\begin{aligned}
&P(\text{saadaan hertta tai ässä})\\
&= P(\text{saadaan hertta}) + P(\text{saadaan ässä}) - P(\text{saadaan herttaässä})\\
&= \frac{13}{52} + \frac{4}{52} - \frac{1}{52}\\
&= \frac{16}{52}\\
&= \frac{4}{13} \approx 0{,}31
\end{aligned}$$

Vastaus: a) 0,50 b) 0,31

157. $P(\text{veriryhmä O}) = 0{,}33$
$P(\text{Rh-tekijä}) = 0{,}13$
$P(\text{veriryhmä O ja Rh-tekijä}) = 0{,}05$

$$\begin{aligned} P(\text{ryhmä O ja Rh-tekijä}) &= P(\text{ryhmä O}) + P(\text{Rh-tekijä}) - P(\text{ryhmä O ja Rh-tekijä}) \\ &= 0{,}33 + 0{,}13 - 0{,}05 \\ &= 0{,}41 \end{aligned}$$

Vastaus: 0,41

158. $P(\text{lyhytkarvainen}) = 0{,}70$ $\quad$ $P(\text{pitkäkarvainen}) = 0{,}30$

$P(\text{lyhytkarvainen ja musta})$
$= 0{,}25 \cdot 0{,}70$
$= 0{,}175$

$P(\text{pitkäkarvainen ja kääpiö})$
$= 0{,}60 \cdot 0{,}30$
$= 0{,}18$

$P(\text{lyhytkarvainen ja kääpiö})$
$= 0{,}50 \cdot 0{,}70$
$= 0{,}35$

a) Kaikki mustat ovat lyhytkarvaisia, joten

$P(\text{musta tai pitkäkarvainen})$
$= P(\text{musta}) + P(\text{pitkäkarvainen})$
$= 0{,}175 + 0{,}30$
$= 0{,}475$

b)

$P(\text{kääpiökoira})$
$= P(\text{lyhytkarv.kääpiö}) + P(\text{pitkäkarv. kääpiö})$
$= 0{,}35 + 0{,}18$
$= 0{,}53$

c)

$$\begin{aligned}&P(\text{pitkäkarv. tai kääpiökoira})\\&= P(\text{pitkäkarv.}) + P(\text{kääpiö}) - P(\text{pitkäkarv. kääpiö})\\&= 0{,}30 + 0{,}53 - 0{,}18\\&= 0{,}65\end{aligned}$$

Vastaus: a) 0,475 b) 0,53 c) 0,65

159. Ominaisuudet periytyvät toisistaan riippumatta.

$P(\text{ominaisuus A}) = 0{,}15$ $P(\text{ei ominaisuus A}) = 1 - 0{,}15 = 0{,}85$

$P(\text{ominaisuus B}) = 0{,}65$ $P(\text{ei ominaisuus B}) = 1 - 0{,}65 = 0{,}35$

a)

$$P(\text{ominaisuus A, mutta ei B}) = 0{,}15 \cdot 0{,}35 = 0{,}0525$$

b)

$$\begin{aligned}&P(\text{ominaisuus A, mutta ei B tai ominaisuus B, mutta ei A})\\&= 0{,}15 \cdot 0{,}35 + 0{,}65 \cdot 0{,}85 = 0{,}605\end{aligned}$$

Vastaus: a) 0,0525 b) 0,605

160. P (punavihersokea, S) = 0,08
P (ei-punavihersokea, T) = 0,92

Koska kolmen joukossa oltava ainakin kaksi sairasta (S), ovat seuraavat vaihtoehdot mahdollisia:

SST, STS, TSS, SSS

Näistä kolmen ensimmäisen tapahtuman todennäköisyydet ovat samat.

$$\begin{aligned}P(\text{ainakin kaksi punavihersokeaa}) &= P(\text{SST}) + P(\text{STS}) + P(\text{TSS}) + P(\text{SSS})\\&= 3 \cdot 0{,}08^2 \cdot 0{,}92 + 0{,}08^3\\&= 0{,}018176 \approx 0{,}018\end{aligned}$$

Vastaus: 0,018

161. Merkitään klaavaa numerolla 1 ja kruunaa numerolla 0. Tällöin tapahtuma ”saadaan kolmella kolikolla yksi klaava” koostuu kolmesta eri mahdollisuudesta.

$$
\begin{aligned}
&P(\text{saadaan yksi klaava})\\
&= P(1,0,0) + P(0,1,0) + P(0,0,1)\\
&= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\\
&= \frac{3}{8} \approx 0,38
\end{aligned}
$$

Vastaus: 0,38

162. Hissiä odottaa yhteensä 2 + 5 = 7 henkilöä. Näiden joukosta arvotaan kaksi hissiin menijää.

a)

$$
\begin{aligned}
P(\text{molemmat tytöt pääsevät}) &= P(\text{1. on tyttö ja 2. on tyttö})\\
&= \frac{2}{7}\cdot\frac{1}{6}\\
&= \frac{1}{21} \approx 0,048
\end{aligned}
$$

b)

$$
\begin{aligned}
&P(\text{vain toinen on tyttö})\\
&= P(\text{1. on tyttö, 2. on poika}) + P(\text{1. on poika, 2. on tyttö})\\
&= \frac{2}{7}\cdot\frac{5}{6} + \frac{5}{7}\cdot\frac{2}{6}\\
&= \frac{10}{21} \approx 0,48
\end{aligned}
$$

Vastaus: a) 0,048 b) 0,48

163. P(valmistusvika) = 0,12 P(ei vikaa) = 1 – 0,12 = 0,88

P(valmistusvika ja kuljetusvaurio) = 0,12 · 0,50
P(valmistusvika ja ei kuljetusvauriota) = 0,12 · 0,50

P(ei-viallinen maljakko ja kuljetusvaurio) = 0,88 · 0,10
P(ei-viallinen maljakko ja ei kuljetusvaurio) = 0,88 · 0,90

a)

$$\begin{aligned}
&P(\text{maljakossa ainakin toinen vika})\\
&= P(\text{kuljetusvaurio, ei valmistusvikaa}) + P(\text{valmistusvika, ei kuljetusvauriota})\\
&+P(\text{molemmat})\\
&= 0{,}10 \cdot 0{,}88 + 0{,}12 \cdot 0{,}50 + 0{,}12 \cdot 0{,}50\\
&= 0{,}208\\
&\approx 0{,}21
\end{aligned}$$

b)

$$\begin{aligned}
P(\text{ei kumpaakaan vikaa}) &= P(\text{ei valmistusvikaa eikä kuljetusvauriota})\\
&= 0{,}88 \cdot 0{,}90\\
&= 0{,}792\\
&\approx 0{,}79
\end{aligned}$$

Vastaus: a) 0,21 b) 0,79

164. Oletetaan, että valojen toiminta on riippumatonta muista valoista.

Valot	**Vihreä**	**Muu**
1.	0,80	1 – 0,80 = 0,20
2.	0,70	1 – 0,70 = 0,30
3.	0,90	1 – 0,90 = 0,10

a)

$$\begin{aligned}
P(\text{ei pysähdy kertaakaan}) &= P(\text{kaikki ovat vihreitä})\\
&= 0{,}80 \cdot 0{,}70 \cdot 0{,}90\\
&= 0{,}504\\
&\approx 0{,}50
\end{aligned}$$

b)

P(pysähtyy vain kerran)
$= P$(pysähtyy 1. valoissa, ei muissa) $+ P$(pysähtyy 2. valoissa, ei muissa) $+$
P(pysähtyy 3. valoissa, ei muissa)

$$
\begin{aligned}
&= 0{,}20 \cdot 0{,}70 \cdot 0{,}90 + 0{,}80 \cdot 0{,}30 \cdot 0{,}90 + 0{,}80 \cdot 0{,}70 \cdot 0{,}10 \\
&= 0{,}398 \\
&\approx 0{,}40
\end{aligned}
$$

Vastaus: a) 0,50 b) 0,40

165. $P(\text{toimii uusi}) = 0{,}98$ $P(\text{ei toimi uusi}) = 1 - 0{,}98 = 0{,}02$
$P(\text{toimii vanha}) = 0{,}85$ $P(\text{ei toimi vanha}) = 1 - 0{,}85 = 0{,}15$

a)

$$P(\text{molemmat toimivat}) = 0{,}98 \cdot 0{,}85 = 0{,}833$$

b)

$$
\begin{aligned}
P(\text{vain toinen toimii}) &= P(\text{uusi toimii, vanha ei}) + P(\text{uusi ei, vanha toimii}) \\
&= 0{,}98 \cdot 0{,}15 + 0{,}02 \cdot 0{,}85 \\
&= 0{,}164
\end{aligned}
$$

c)

$$P(\text{ei kumpikaan toimi}) = 0{,}02 \cdot 0{,}15 = 0{,}003$$

Vastaus: a) 0,833 b) 0,164 c) 0,003

166. Pussissa on yhteensä 7 + 6 + 8 = 21 karkkia, joista nostetaan umpimähkään kaksi.

a)

$$\begin{aligned} P(\text{molemmat liitulakuja}) &= P(\text{1. on liitulaku ja 2. on liitulaku}) \\ &= \frac{7}{21} \cdot \frac{6}{20} \\ &= \frac{1}{10} \end{aligned}$$

b)

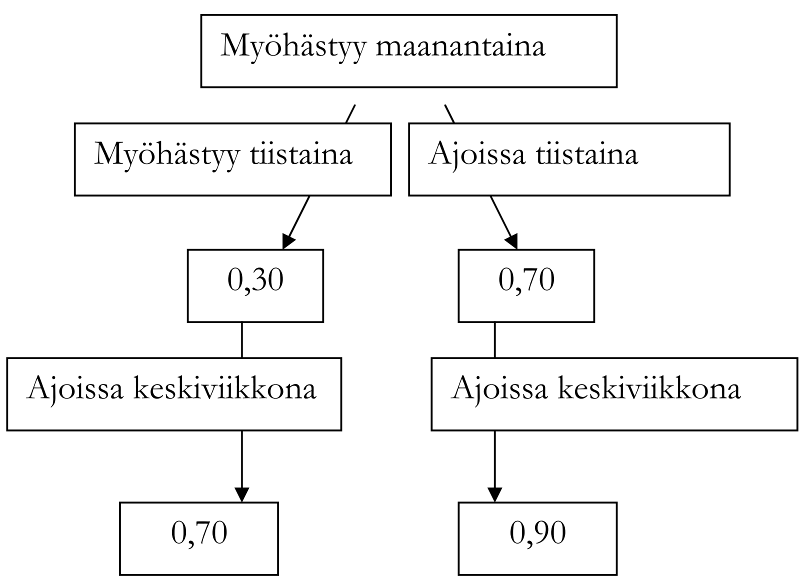
P(ainakin toinen on merirosvoraha)
= *P*(1. on merirosvoraha, 2. joku muu) + *P*(1. joku muu, 2. on merirosvoraha) + *P*(molemmat ovat merirosvorahoja)

$$= \frac{6}{21} \cdot \frac{15}{20} + \frac{15}{21} \cdot \frac{6}{20} + \frac{6}{21} \cdot \frac{5}{20}$$

$$= \frac{1}{2}$$

Vastaus: a) $\frac{1}{10}$ b) $\frac{1}{2}$

167. Tilannetta voi havainnollistaa puumallilla:

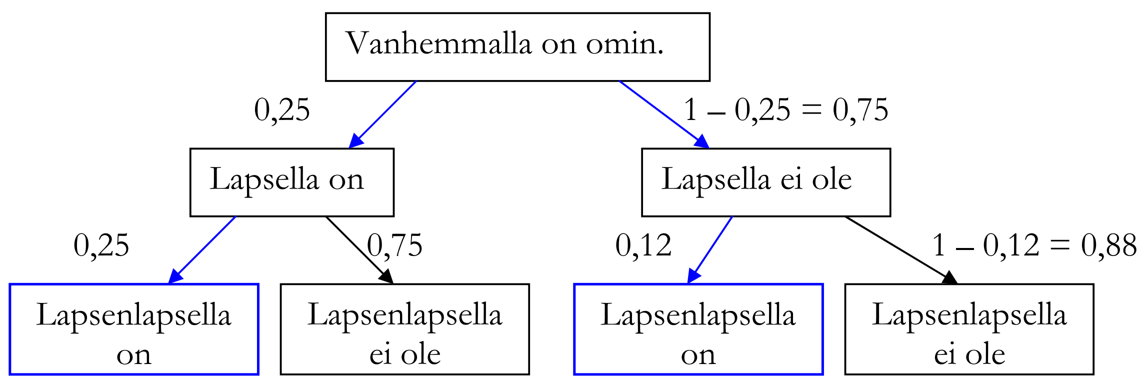


Kaavion molemmat haarat johtavat siihen, että oppilas on ajoissa keskiviikkona. Todennäköisyys on näin ollen

$$\begin{aligned}&P(\text{myöhästyy keskiviikkona})\\&=0{,}30\cdot 0{,}70+0{,}70\cdot 0{,}90\\&=0{,}84\end{aligned}$$

Vastaus: 0,84

168. Tilannetta voi havainnollistaa puumallilla.



$$\begin{aligned}P(\text{lapsenlapsella on ominaisuus})&=0{,}25\cdot 0{,}25+0{,}75\cdot 0{,}12\\&=0{,}1525\\&\approx 0{,}15\end{aligned}$$

Vastaus: 0,15

169. Pakan korteista ei-herttoja on 39. Koska korttia ei noston jälkeen palauteta pakkaan, toisella nostolla pakassa on vain 51 korttia.

P(ainoastaan toinen on hertta)
$= P$(1. on hertta, toinen joku muu) $+ P$(1. on joku muu, toinen on hertta)

$$= \frac{13}{52} \cdot \frac{39}{51} + \frac{39}{52} \cdot \frac{13}{51}$$

$$= \frac{13}{34}$$

$$\approx 0{,}38$$

Vastaus: 0,38

170. $P(\text{vastaus oikein}) = \frac{1}{4}$

$P(\text{vastaus väärin}) = \frac{3}{4}$

Opiskelija arvaa kolme ensimmäistä kohtaa.

a) $P(\text{oikein, väärin, väärin}) = \frac{1}{4} \cdot \left(\frac{3}{4}\right)^2 = \frac{9}{64} \approx 0{,}14$

b) $P(\text{oikein, oikein, väärin}) = \left(\frac{1}{4}\right)^2 \cdot \frac{3}{4} = \frac{3}{64} \approx 0{,}047$

c) $P(\text{oikein, väärin, oikein}) + P(\text{oikein, oikein, väärin}) + P(\text{väärin, oikien, oikein})$

$$= 3 \cdot \left(\frac{1}{4}\right)^2 \cdot \frac{3}{4}$$

$$= \frac{9}{64}$$

$$\approx 0{,}14$$

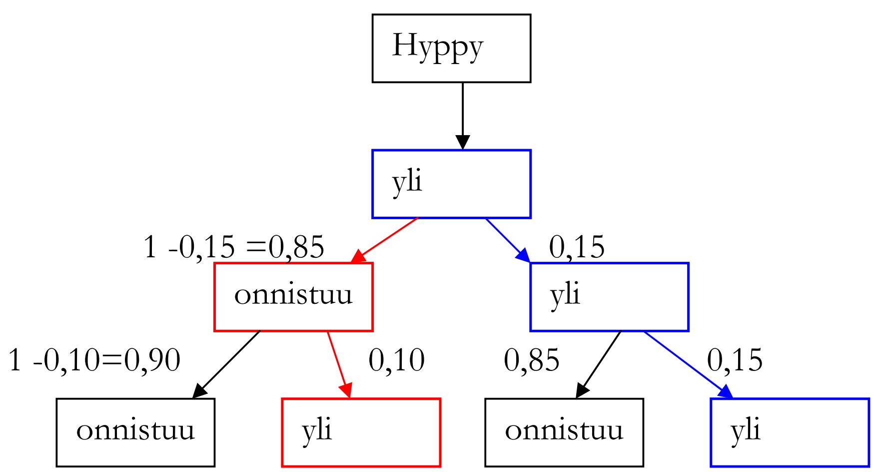
Vastaus: a) 0,14 b) 0,047 c) 0,14

171. Mustia noppia on 3 kpl, valkoisia 5 kpl ja punaisia 4 kpl. Yhteensä noppia on 12 kpl.

P(mustalla kuutonen tai valkoisella ykkönen)
$= P($ musta noppa ja 6$) + P($ valkoinen noppa ja 1$)$
$= \frac{3}{12} \cdot \frac{1}{6} + \frac{5}{12} \cdot \frac{1}{6}$
$= \frac{1}{9}$
$\approx 0{,}11$

Vastaus: 0,11

172. Tilannetta voi havainnollistaa puumallilla.



a) $P(\text{kaksi hyppyä yli}) = 0{,}15^2 = 0{,}0225 \approx 0{,}02$

b) $P(\text{kolmas hyppy yli}) = 0{,}15^2 + 0{,}85 \cdot 0{,}10 = 0{,}1075 \approx 0{,}11$

Vastaus: a) 0,02 b) 0,11

2.4 Tuloperiaate ja kombinaatiot

173. Housuja 4, paitoja 5, kenkiä 3 (kpl)

Tuloperiaatteen mukaan erilaisia asukokonaisuuksia on

$$4 \cdot 5 \cdot 3 = 60$$

Vastaus: 60

174. Kirjaimia on 29 kpl ja numeroita 10 kpl.

Jokainen kirjain voidaan valita 29 tavalla ja vastaavasti jokainen numero 10 tavalla. Rekisterikilvessä on alkuosassa 3 kirjainta ja loppuosassa 3 numeroa. Näin ollen tuloperiaatteen mukaan erilaisia yhdistelmiä on

$$29 \cdot 29 \cdot 29 \cdot 10 \cdot 10 \cdot 10 = 24\,389\,000 \text{ (kpl)}$$

Vastaus: 24 389 000

175. Henkilötunnuksen loppuosassa on kolmen numeron sarja ja sen perässä tarkistusmerkintä (kirjain). Koska kirjaimia on käytössä vain 21, mutta numeroita 10, erilaisia tunnuksia on

$$10 \cdot 10 \cdot 10 \cdot 21 = 21\,000 \text{ (kpl)}$$

Vastaus: 21 000

176. a) Koska kokeen jokaiseen kohtaan on neljä erilaista vastausvaihtoehtoa, erilaisia mahdollisia rivejä on tuloperiaatteen mukaan

$$4^{10} = 1\,048\,576$$

b) Oikeita rivejä on vain yksi, joten

$$P(\text{arvaa kaikki 10 oikein}) = \frac{1}{1\,048\,576} = 9{,}5367...\cdot 10^{-7} \approx 9{,}5\cdot 10^{-7}$$

Vastaus: a) 1 048 576 b) $9{,}5\cdot 10^{-7}$

177. Kutakin merkkiä kohden on käytettävissä 6 paikkaa, joihin voidaan asettaa 1–6 pistettä. Koska paikka voi olla täytetty pisteellä tai ei, niin jokaiselle paikalle on kaksi täyttövaihtoehtoa. Yhteensä erilaisia täyttövaihtoehtoja on

$$2^6 = 64$$

Koska järjestelmässä merkkiin kuuluu vähintään yksi piste, niin täyttövaihtoehdoista pitää vähentää se, jossa kaikki paikat ovat tyhjiä.

Merkkejä on siis 64 – 1 = 63.

Vastaus: 63

178. Punaisia palloja 5 kpl, mustia palloja 10 kpl. Yhteensä palloja on 15 kpl. Poimitaan umpimähkään 5 palloa palauttamatta niitä takaisin.

$$
\begin{aligned}
P(\text{ainakin yksi on punainen}) &= 1 - P(\text{kaikki mustia}) \\
&= 1 - \frac{10}{15}\cdot\frac{9}{14}\cdot\frac{8}{13}\cdot\frac{7}{12}\cdot\frac{6}{11} \\
&= \frac{131}{143} \\
&= 0{,}91608... \\
&\approx 0{,}9161
\end{aligned}
$$

Tapa 2:

$$1 - \frac{\binom{10}{5}}{\binom{15}{5}}$$

$$
\begin{aligned}
P(\text{kaikki samanvärisiä}) &= P(\text{kaikki punaisia}) + P(\text{kaikki mustia}) \\
&= \frac{\binom{5}{5}}{\binom{15}{5}} + \frac{\binom{10}{5}}{\binom{15}{5}} \\
&= 0{,}084249... \\
&\approx 0{,}0842
\end{aligned}
$$

Vastaus: a) 0,9161 b) 0,0842

179. Ensimmäinen kilpailija voidaan valita 30 tavalla, toinen 29 tavalla, kolmas 28 tavalla ja niin edelleen. Erilaisia hyppyjärjestyksiä on siis

$30! = 2{,}6525\ldots\cdot 10^{32} \approx 2{,}7\ 10^{32}$

Vastaus: $2{,}7\ 10^{32}$

180. Seitsemän lasta voi asettua jonoon 7! = 5040 eri tavalla.
Suotuisia järjestyksiä on 2 kpl, joten kysytty todennäköisyys on

$$\frac{2}{5040} = 0{,}0003968... \approx 3{,}97 \cdot 10^{-4}$$

Vastaus: $3{,}97 \cdot 10^{-4}$

181. Kolme tyttöä voi asettua jonoon 3! erilaisella tavalla.

Vastaavasti pojat voivat asettua omaan jonoonsa 3! eri tavalla.
Jos pojat sijoitetaan tyttöjen jälkeen, on tuloperiaatteen mukaan erilaisia jonoja

$$3! \cdot 3! = 36 \ \text{(kpl)}$$

Vastaus: 36

182. Naisia on 2 kpl ja miehiä 4 kpl. Erilaisia jonoja on yhteensä 6! kappaletta.

Naiset A ja B ovat peräkkäin. Kiinnitetään naisten mahdolliset paikat tapaus kerrallaan, jolloin erilaisia järjestyksiä saadaan vain miesten X paikkoja vaihtamalla:

ABXXXX	BAXXXX
XABXXX	XBAXXX
XXABXX	XXBAXX
XXXABX	XXXBAX
XXXXAB	XXXXBA

Jokainen esitetty mahdollisuus sisältää 4! verran erilaisia jonoja (miehet voivat asettua jonoon 4! erilaisella tavalla). Niinpä jonoja, joissa naiset ovat peräkkäin, on

$10 \cdot 4! = 240$ kappaletta

$$P(\text{naiset peräkkäin}) = \frac{10 \cdot 4!}{6!} = \frac{240}{720} = \frac{1}{3} \approx 0{,}33$$

Vastaus: 0,33

183. Viisi maata voidaan umpimähkään asettaa 5! erilaiseen järjestykseen. Olkoot kaksi pienintä maata **1** ja **2**, muita merkitään kirjaimilla A, B ja C. Oletetaan, että oikea järjestys olisi **1** **2** A B C.

Järjestyksiä, joissa ainakin kaksi kolmesta viimeistä maasta on väärässä järjestyksessä, on

1 2 A C B
1 2 B A C
1 2 B C A
1 2 C A B
1 2 C B A

eli 5 kappaletta.

$$P(\text{lopuista ainakin kaksi väärissä paikoissa}) = \frac{5}{5!} = 0{,}041666... \approx 0{,}042$$

Vastaus: 0,042

184. a) $\binom{7}{3} = \frac{7!}{3!4!} = 35$

b) $\binom{13}{6} = \frac{13!}{6!7!} = 1716$

Vastaus: a) 35 b) 1716

185. Tuomariston 5 jäsentä voidaan valita 27 kandidaatin joukosta

$\binom{27}{5} = 80\,730$ eri tavalla.

186. Koska henkilöitä on yhteensä 15, erilaisia 3 hengen pöytäseurueita voidaan muodostaa

$$\binom{15}{3} = 455$$

187. Pakassa on 52 korttia. Pelaajalle jaetaan 5 korttia.

a) Viiden kortin käsiä on

$$\binom{52}{5} = 2\,598\,960$$

b) Punaisia maita on 2 · 13 = 26 (kpl). Näistä viiden kortin käsiä on

$$\binom{26}{5} = 65\,780$$

c) Neljä ässää ja yksi kortti voidaan valita muiden (52 – 4 = 48) joukosta eli käsiä on

$$\binom{4}{4}\binom{48}{1} = 48$$

Vastaus: a) 2 598 960 b) 65 780 c) 48

188. Kuuden henkilön joukosta voidaan arpoa kolme

$$\binom{6}{3} = 20 \text{ tavalla}$$

Arvontatuloksia, joissa Sara, Mimmi ja Harri ovat mukana, on vain yksi.

$$P(\text{Sara,Mimmi ja Harri palkitaan}) = \frac{1}{20} = 0{,}05$$

Vastaus: 0,05

189. Koska aviopareja on 10, henkilöitä on yhteensä 20. Jos kaikki, myös avioparit, kättelisivät toisiaan, kutsuilla tehtäisiin

$$\binom{20}{2} = 190 \text{ kättelyä.}$$

Tästä jää pois kuitenkin aviopariein kättelyt, joita on 10 kappaletta. Kättelyitä tehdään siis 180.

Vastaus: 180

190. Yli 180 cm poikia 10 kpl ja alle 180 cm poikia 8 kpl.

Luokan 18 pojan joukosta voidaan valita neljä

$$\binom{18}{4} = 3060 \text{ tavalla.}$$

Yli 180 cm poikien joukosta voidaan valita kaksi

$$\binom{10}{2} = 45 \text{ tavalla.}$$

Vastaavasti alle 180 cm poikien joukosta voidaan valita kaksi

$$\binom{8}{2} = 28 \text{ tavalla.}$$

Tällöin kokonaisuuksia, joissa on kaksi yli ja kaksi alle 180 cm pitkää poikaa, on tuloperiaatteen mukaan $45 \cdot 28 = 1260$.

$$P(\text{kaksi yli ja kaksi alle 180 cm})$$
$$= \frac{\binom{10}{2} \cdot \binom{8}{2}}{\binom{18}{4}} = \frac{45 \cdot 28}{3060} = 0{,}4117... \approx 0{,}41$$

Vastaus: 0,41

191. Numeroita on 39, joista arvotaan 7.

Erilaisia lottorivejä on $\binom{39}{7} = 15\,380\,937$.

Neljä oikein voidaan valita seitsemän oikean joukosta

$\binom{7}{4} = 35$ eri tavalla.

Loput kolme voidaan valita jäljelle jääneen 32 numeron joukosta

$\binom{32}{3} = 4960$ eri tavalla.

Tuloperiaatteen mukaan erilaisia 4-oikein rivejä on siis olemassa

$$\binom{7}{4} \cdot \binom{32}{3} = 35 \cdot 4960 = 173\,600$$

eli lehtiartikkelin luku oli oikea.

Vastaavalla tavalla voidaan laskea 5-oikein ja 6-oikein –ruudukoiden määrät:

5-oikein: $\binom{7}{5} \cdot \binom{32}{2} = 21 \cdot 496 = 10\,416$

6-oikein: $\binom{7}{6} \cdot \binom{32}{1} = 7 \cdot 32 = 224$

Vastaus: 4 oikein: 173 600, 5-oikein: 10 416, 6-oikein: 224

192. Sieniä opetettiin 78, kurssilainen oppi 49, kokeessa kysyttiin 6.

Kokeen kuusi sientä voidaan valita $\binom{78}{6}$ tavalla.

Jotta kurssilainen tuntisi ne kaikki, ne tulisi valita 49 opitun sienen joukosta.
Tällaisia rivejä on yhteensä $\binom{49}{6}$ erilaista.

$$P(\text{tuntee kaikki sienet kokeessa}) = \frac{\binom{49}{6}}{\binom{78}{6}} = 0{,}054443... \approx 0{,}054$$

Vastaus: 0,054

193. Laatikossa on 10 sinistä ja 7 punaista palloa. Yhteensä 17 palloa.
Laatikosta nostetaan umpimähkään 6 palloa.

a)

$$P(\text{yksi pallo sininen}) = \frac{\binom{10}{1}\binom{7}{5}}{\binom{17}{6}} = 0{,}01696... \approx 0{,}017$$

b)

$$P(\text{kolme punaista}) = \frac{\binom{7}{3}\binom{10}{3}}{\binom{17}{6}} = 0{,}3393... \approx 0{,}34$$

Vastaus: a) 0,017 b) 0,34

194. Housuja 6 kpl ja paitoja 8 kpl. Matkalle otetaan mukaan housuja 3 ja paitoja 3.

a) Matkalla voidaan muodostaa vaatekokonaisuuksia $3 \cdot 3 = 9$ erilaista.

b) Vaatekokonaisuuksia on $\binom{6}{3}\binom{8}{3} = 1120$

Vastaus: a) 9 b) 1120

195. Ryhmässä 3 blondia, 2 brunettea ja 2 tummaverikköä.

a) Seitsemän henkilöä voi asettua jonoon $7! = 5040$ eri tavalla.

b) Jos blondit ovat jonon alussa, niin erilaisia jonoja on

$$3 \cdot 2 \cdot 1 \cdot 4! = 144$$

c) Jos brunetet ovat jonossa ensimmäisenä ja viimeisenä, niin erilaisia jonoja on

$$2 \cdot 5! \cdot 1 = 240$$

d) Merkitään blondia kirjaimella B ja muuta henkilöä kirjaimella M. Kolme blondia voi olla peräkkäin seitsemän henkilön jonossa seuraavasti:

B B B M M M M   M B B B M M M   M M B B B M M

M M M B B B M   M M M M B B B

Tällaisia jonoja on

$$3! \cdot 4! + 4 \cdot 3! \cdot 3! + 4 \cdot 3 \cdot 3! \cdot 2 \cdot 1 + 4 \cdot 3 \cdot 2 \cdot 3! \cdot 1 + 4! \cdot 3! = 720$$

Vastaus: a) 5040 b) 144 c) 240 d) 720

196. Luokassa on 28 oppilasta.

a) Näistä voidaan valita 3 henkilöä

$$\binom{28}{3} = 3276 \text{ eri tavalla.}$$

b) Oppilaista voidaan valita 5 henkilöä

$$\binom{28}{5} = 98\,280 \text{ eri tavalla.}$$

Vastaus: a) 3276 b) 98 280

197. Jokainen yhdeksästä ruudusta voidaan värittää kahdella tavalla eli

$$2^9 = 512 \text{ eri tavalla.}$$

a) Kuvion 2 shakkilautakuvio on yksi väritystapa eli

$$P(\text{saadaan kuvio 2}) = \frac{1}{512} = 0{,}001953... \approx 0{,}002$$

b) Kukin vaakarivi eli kolme ruutua voidaan värittää $2^3 = 8$ eri tavalla.

Koska vaakarivi ei voi olla yksivärinen (sininen tai ruskea), niin yksi vaakarivi voidaan tällöin värittää

$$8 - 2 = 6 \text{ eri tavalla.}$$

Koska vaakarivejä on kolme, niin koko kuvio voidaan värittää

$$6^3 = 216 \text{ eri tavalla.}$$

$$P(\text{mikään vaakarivi ei yksivärinen}) = \frac{216}{512} = 0{,}42187... \approx 0{,}422$$

198. Tuoreet tomaatit 17
Pilaantuneet 5
Yhteensä 22

a) Viisi tomaattia voidaan nostaa laatikosta $\binom{22}{5}$ tavalla.

Tuoreet tomaatit voidaan valita $\binom{17}{5}$ tavalla. Näin ollen

$$P(\text{kaikki 5 ovat tuoreita}) = \frac{\binom{17}{5}}{\binom{22}{5}} = \frac{6188}{26\,334} = 0,23498... \approx 0,23$$

b) Yksi tuore voidaan valita $\binom{17}{1}$ tavalla.

Loput neljä valitaan pilaantuneiden joukosta, ja ne voidaan valita $\binom{5}{4}$ tavalla.

$$P(\text{vain yksi on tuore}) = \frac{\binom{17}{1}\cdot\binom{5}{4}}{\binom{22}{5}} = \frac{85}{26\,334} = 0,003227... \approx 0,0032$$

Vastaus: a) 0,23 b) 0,0032

2.5 Binomitodennäköisyys

199. a) $\binom{7}{3}\cdot 0{,}25^3 \cdot 0{,}75^4 = 35\cdot 0{,}25^3\cdot 0{,}75^4 = 0{,}173... \approx 0{,}17$

b) $\binom{9}{2}\cdot\left(\frac{2}{7}\right)^2\cdot\left(\frac{5}{7}\right)^5 = 36\cdot\left(\frac{2}{7}\right)^2\cdot\left(\frac{5}{7}\right)^5 = 0{,}546... \approx 0{,}55$

Vastaus: a) 0,17 b) 0,55

200. Esimerkiksi: Heitto onnistuu todennäköisyydellä 0,90. Laske todennäköisyys, että 12 heitosta täsmälleen 5 onnistuu.

201. $P(\text{Oona osuu}) = 0{,}70$ $P(\text{Oona ei osu}) = 1 - 0{,}70 = 0{,}30$

$$\begin{aligned} P(\text{24 lyönnistä 15 lyönnillä osuu palloon}) &= \binom{24}{15}\cdot 0{,}70^{15}\cdot 0{,}30^9 \\ &= 0{,}12218... \\ &\approx 0{,}12 \end{aligned}$$

Vastaus: 0,12

202. $P(\text{valkoruskea}) = 0{,}35$
$P(\text{mustavalkoiset}) = 0{,}65$

a) Seitsemän pennun joukosta voidaan valita kaksi mustavalkoista pentua

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

$$P(2 \text{ mustavalkoista, } 5 \text{ valkoruskeaa}) = \binom{7}{2} \cdot 0{,}65^2 \cdot 0{,}35^5 = 0{,}04660... \approx 0{,}047$$

b)

$$P(4 \text{ valkoruskeaa, } 3 \text{ mustavalkoista}) = \binom{7}{4} \cdot 0{,}35^4 \cdot 0{,}65^3 = 0{,}14423... \approx 0{,}14$$

Vastaus: a) 0,047 b) 0,14

203. $P(\text{itää}) = 0{,}88$ $P(\text{ei idä}) = 1 - 0{,}88 = 0{,}12$

a) 50 siemenen joukosta voidaan valita 45 itävää siementä

$\binom{50}{45} = 2\,118\,760$ tavalla.

$$P(\text{täsmälleen } 45 \text{ siemenistä itää}) = \binom{50}{45} \cdot 0{,}88^{45} \cdot 0{,}12^5 = 0{,}16738... \approx 0{,}17$$

b) Tapahtumaan ”kaikki siemenet itävät” ei tarvita binomitodennäköisyyttä, vaan se saadaan laskettua suoraan kertolaskusäännön avulla:

$$P(\text{kaikki itävät}) = 0{,}88^{50} = 0{,}001675... \approx 0{,}0017$$

Vastaus: a) 0,17 b) 0,0017

204. Vastaajista 50 on oltava tyttöjä. Nämä tytöt voidaan valita 100 oppilaan joukosta $\binom{100}{50}$ tavalla.

$$P(\text{puolet tyttöjä, puolet poikia}) = \binom{100}{50} \cdot 0{,}5^{50} \cdot 0{,}5^{50} = 0{,}07958... \approx 0{,}080$$

Vastaus: 0,080

205. P(vasenkätinen) = 0,05 P(oikeakätinen) = 1 – 0,05 = 0,95

Luokan 32 oppilaasta 4 vasenkätistä voidaan valita $\binom{32}{4}$ eri tavalla.

$$P(\text{luokassa 4 vasenkätistä}) = \binom{32}{4} \cdot 0{,}05^{4} \cdot 0{,}95^{28} = 0{,}05345... \approx 0{,}053$$

Vastaus: 0,053

206. P(valmistettu Suomessa) = 0,32
P(ei valmistettu Suomessa) = 1 – 0,32 = 0,68
Asiakas sovittaa yhdeksää vaatetta.

a)

$$\begin{aligned} &P(0,\ 1 \text{ tai } 2 \text{ vaatetta valm. Suomessa}) \\ &= 0{,}68^{9} + \binom{9}{1} \cdot 0{,}32 \cdot 0{,}68^{8} + \binom{9}{2} \cdot 0{,}32^{2} \cdot 0{,}68^{7} \\ &= 0{,}41058... \\ &\approx 0{,}41 \end{aligned}$$

b)

$$P(\text{vähintään 2 vaatetta valm. Suomessa})$$
$$= 1 - P(\text{0 tai 1 valm. Suomessa})$$
$$= 1 - \left[0{,}68^9 + \binom{9}{1}0{,}32 \cdot 0{,}68^8\right]$$
$$= 0{,}8372...$$
$$\approx 0{,}84$$

Vastaus: a) 0,41 b) 0,84

207. $P(\text{veriryhmä A}) = 0{,}37 + 0{,}05 = 0{,}42$
$P(\text{muu veriryhmä }) = 1 - 0{,}42 = 0{,}58$

a) Kymmenestä henkilöstä 4 kuuluu veriryhmään A todennäköisyydellä

$$\binom{10}{4}0{,}42^4 \cdot 0{,}58^6 = 0{,}2487... \approx 0{,}25$$

b)

$$P(\text{vähintään kaksi kuuluu A}) = 1 - P(\text{0 tai 1 kuuluu A})$$
$$= 1 - \left[0{,}58^{10} + \binom{10}{1}0{,}42 \cdot 0{,}58^9\right]$$
$$= 0{,}9644...$$
$$\approx 0{,}96$$

c)

$$P(\text{korkeintaan kaksi kuuluu A}) = P(0 \text{ tai } 1 \text{ tai } 2 \text{ kuuluu A})$$
$$= 0{,}58^{10} + \binom{10}{1}0{,}42 \cdot 0{,}58^{9} + \binom{10}{2}0{,}42^{2} \cdot 0{,}58^{8}$$
$$= 0{,}1371...$$
$$\approx 0{,}14$$

Vastaus: a) 0,25 b) 0,96 c) 0,14

208. P(viallinen) = 0,03
P(virheetön) = 1 – 0,03 = 0,97

a) Kymmenen levyn joukosta voidaan yksi viallinen levy valita

$$\binom{10}{1} = 10 \text{ tavalla.}$$

$$P(10 \text{ joukossa yksi viallinen}) = \binom{10}{1} \cdot 0{,}03 \cdot 0{,}97^{9} = 0{,}2280... \approx 0{,}23$$

b) Tapahtuman ”ainakin kolme viallista kymmenen joukossa” komplementti ”korkeintaan kaksi viallista” on paljon helpompi laskea.

P(10 joukossa ainakin kolme viallista)
= 1 – P(korkeintaan kaksi viallista)
= 1 – [P(0 viallista) + P(1 viallinen) + P(2 viallista)]

$$= 1 - [0{,}97^{10} + \binom{10}{1} \cdot 0{,}03 \cdot 0{,}97^{9} + \binom{10}{2} \cdot 0{,}03^{2} \cdot 0{,}97^{8}]$$
$$= 0{,}002764...$$
$$\approx 0{,}0028$$

Vastaus: a) 0,23 b) 0,0028

209. P (saa ihottuman) = 0,167
P (ei saa ihottumaa) = 1 – 0,167 = 0,833

a) Tallin 14 ponin joukosta 5 ihottuman saavaa ponia voidaan valita

$$\binom{14}{5} = 2002 \text{ tavalla.}$$

$$P(\text{tasan 5 saa ihottuman}) = \binom{14}{5} \cdot 0{,}167^5 \cdot 0{,}833^9 = 0{,}05021... \approx 0{,}050$$

b) Tapahtuman ”ainakin kolme saa ihottuman” komplementti, ”korkeintaan kaksi saa ihottuman”, on helpompi laskea. Se muodostuu kolmesta tapauksesta: ei yksikään saa ihottumaa, tasan yksi saa tai tasan kaksi saa ihottumaa.

P(ainakin 3 saa ihottuman)
= 1 – P(korkeintaan 2 saa ihottuman)
= 1 – [P(ei yksikään saa ihottumaa”) + P(1 saa ihottuman) + P(2 saa ihottuman)]

$$= 1 - [0{,}833^{14} + \binom{14}{1} \cdot 0{,}167 \cdot 0{,}833^{13} + \binom{14}{2} \cdot 0{,}167^2 \cdot 0{,}833^{12}]$$

$$= 0{,}42188...$$

$$\approx 0{,}42$$

Vastaus: 0,42

210. P(virheelliset) = 0,04
P(virheettömät) = 1 – 0,04 = 0,96

a) 21 levyn joukosta voidaan kaksi virheellistä levyä valita $\binom{21}{2}$ tavalla.

$$P(\text{joukossa 2 virheellistä levyä}) = \binom{21}{2} \cdot 0{,}04^2 \cdot 0{,}96^{19} = 0{,}1547... \approx 0{,}15$$

b) Tapahtuman ”ainakin kaksi virheellistä levyä” komplementti ”korkeintaan yksi virheellinen” on helpompi laskea. Se sisältää vain tapahtumat ”ei yhtään virheellistä” ja ”tasan yksi virheellinen”.

P(joukossa ainakin 2 virheellistä levyä)
= 1 – [P(ei yhtään virheellistä levyä) + P(yksi virheellinen levy)]
$= 1 - [0{,}96^{21} + \binom{21}{1} \cdot 0{,}04 \cdot 0{,}96^{20}]$
$= 1 - 0{,}79560...$
$= 0{,}2043...$
$\approx 0{,}20$

Vastaus: a) 0,15 b) 0,20

211. Noppaa heitetään 5 kertaa.
$P(\text{saadaan 6}) = \frac{1}{6}$
$P(\text{saadaan muu kuin 6}) = \frac{5}{6}$

a) $P(\text{kaksi kuutosta}) = \binom{5}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^3 = 0{,}1607... \approx 0{,}16$

b)

$$P(\text{vähintään kaksi kuutosta}) = 1 - P(0 \text{ tai } 1 \text{ kuutonen})$$

$$= 1 - \left[\left(\frac{5}{6}\right)^5 + \binom{5}{1} \cdot \left(\frac{1}{6}\right)^1 \cdot \left(\frac{5}{6}\right)^4 \right] = 0{,}1962... \approx 0{,}20$$

Vastaus: a) 0,16 b) 0,20

212. $P(\text{saadaan kuutonen}) = \frac{1}{6}$

$P(\text{ei saada kuutosta}) = \frac{5}{6}$

Kaksi kuutosta voidaan valita kuuden heiton joukosta $\binom{6}{2} = 15$ tavalla.

$$P(\text{saadaan kaksi kuutosta}) = \binom{6}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^4 = 0{,}200938... \approx 0{,}20$$

Vastaus: 0,20

213. *P*(osuu kymppiin)= 0,08
P(ei osu kymppiin) = 1 – 0,08 = 0,92

a) $P(\text{osuu 1. mutta ei muilla}) = 0{,}08 \cdot 0{,}92^4 = 0{,}05731... \approx 0{,}057$

b) Tapahtuman ”osuu ainakin yhdellä tikalla” komplementti on ”ei osu yhdelläkään tikalla kymppiin”.

P(osuu ainakin yhdellä tikalla kymppiin) = 1 – *P*(ei osu yhdelläkään)
$= 1 - 0{,}92^5$
$= 0{,}3408...$
$\approx 0{,}34$

c) Yksi kymppiin osuva tikka voidaan valita viiden tikan joukosta $\binom{5}{1}$ tavalla.

$$P(\text{osuu yhdellä tikalla kymppiin}) = \binom{5}{1} \cdot 0{,}08 \cdot 0{,}92^4 = 0{,}286557... \approx 0{,}29$$

Vastaus: a) 0,057 b) 0,34 c) 0,29

214. *P*(tallentaja tekee virheen) = 0,049
P(ei tee virhettä) = 1 – 0,049 = 0,951

a) Virhe voi olla mikä tahansa kuudesta numerosta eli erilaisia virhemahdollisuuksia on $\binom{6}{1} = 6$.

$$P(\text{tekee tasan yhden virheen}) = \binom{6}{1} \cdot 0{,}049 \cdot 0{,}951^5 = 0{,}2286... \approx 0{,}23$$

b) Tapahtuma ”enintään yksi virhe” muodostuu kahdesta mahdollisuudesta: ”tallentaja ei tee yhtään virhettä” ja ”tekee tasan yhden virheen”.

$$\begin{aligned}
&P(\text{tekee enintään yhden virheen})\\
&= P(\text{ei tee yhtään virhettä}) + P(\text{tekee yhden virheen})\\
&= 0{,}951^6 + \binom{6}{1}\cdot 0{,}049\cdot 0{,}951^5 = 0{,}9684... \approx 0{,}97
\end{aligned}$$

Vastaus: a) 0,23 b) 0,97

215. P(vaaleasilmäinen) = 0,89
P(muu)= 1 – 0,89 = 0,11

Ryhmästä valitaan satunnaisesti 8 opiskelijaa.

a)

$$\begin{aligned}
&P(6,\ 7 \text{ tai } 8 \text{ opiskelijaa vaaleasilmäisiä})\\
&= \binom{8}{6}0{,}89^6\cdot 0{,}11^2 + \binom{8}{7}0{,}89^7\cdot 0{,}11^1 + 0{,}89^8\\
&= 0{,}95127...\\
&\approx 0{,}95
\end{aligned}$$

b)

$$\begin{aligned}
&P(0,\ 1 \text{ tai } 2 \text{ opiskelijaa vaaleasilmäisiä})\\
&= 0{,}11^8 + \binom{8}{1}0{,}89^1\cdot 0{,}11^7 + \binom{8}{2}0{,}89^2\cdot 0{,}11^6\\
&= 0{,}00004070...\\
&\approx 0{,}000041
\end{aligned}$$

Vastaus: a) 0,95 b) 0,000041

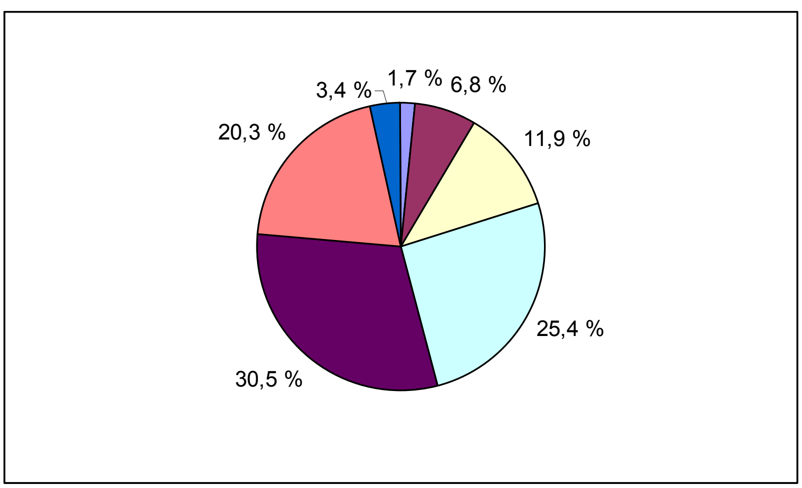
Kertausosa

1. **a)** Muodostetaan taulukon perusteella frekvenssijakaumat.

Äänimäärä	***f***	***f*%**
0	1	$\frac{1}{59} = 0{,}0169... \approx 1{,}7\ \%$
2	4	$\frac{4}{59} = 0{,}0677... \approx 6{,}8\ \%$
3	7	$\frac{7}{59} = 0{,}1186... \approx 11{,}9\ \%$
4	15	$\frac{15}{59} = 0{,}2542... \approx 25{,}4\ \%$
5	18	$\frac{18}{59} = 0{,}3050... \approx 30{,}5\ \%$
6	12	$\frac{12}{59} = 0{,}2033... \approx 20{,}3\ \%$
7	2	$\frac{2}{59} = 0{,}0338... \approx 3{,}4\ \%$

b) Moodi on se muuttujan arvo, jonka frekvenssi on suurin. Mo = 5.

c)



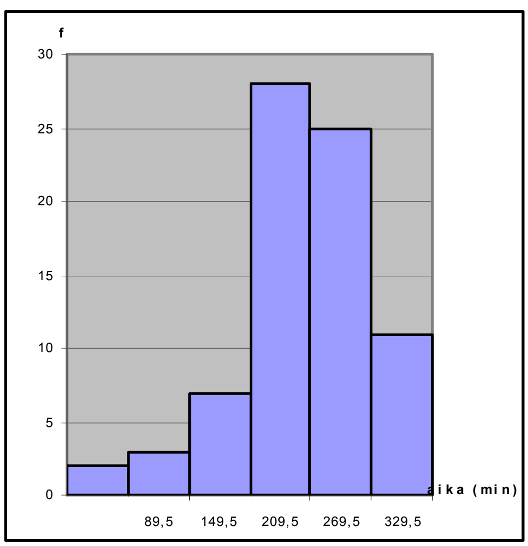
2. **a)** Muodostetaan absoluuttinen frekvenssijakauma ja suhteellinen frekvenssijakauma.

Aika (min)	***f***	***f*%**	**Luokkakeskus (min)**
0–59	2	3 %	$\frac{0+59,5}{2}=29,75$
60–119	3	4 %	89,5
120–179	7	9 %	149,5
180–239	28	37 %	209,5
240–299	25	33 %	269,5
300–360	11	14 %	329,5

b) Moodi on sen luokan luokkakeskus, jonka frekvenssi on suurin.

$$\text{Mo}=\frac{179,5+239,5}{2}=209,5\ \text{(min)}$$

c) Havainnollistetaan frekvenssijakaumaa histogrammilla.



3. Luokitellaan opiskelijoiden pituudet ja muodostetaan frekvenssijakauma ja summafrekvenssijakauma esimerkiksi seuraavasti:

Pituus (cm)	***f***	***f*%**	**Luokkakeskus (cm)**
155–159	1	1,3 %	$\frac{154,5+159,5}{2}=157$
160–164	0	0,0 %	162
165–169	9	11,8 %	167
170-174	5	6,6 %	172
175–179	2	2,6 %	177
180–184	4	5,3 %	182

4. Annettuja arvosanoja on yhteensä 30 kpl.
Lasketaan arvosanan 8 suhteellinen frekvenssi.

$$\frac{6}{30}\cdot 100\% = 20\%$$

$$P(\text{saadaan } 8) = 0{,}2$$

Arvosanan pitää olla vähintään 7 eli 7, 8, 9 tai 10.
Lasketaan näiden arvosanojen suhteelliset frekvenssit yhteen.

$$\frac{8+6+4+3}{30}\cdot 100\% = 70\%$$

$$P(\text{saadaan7, 8, 9 tai 10}) = 0{,}7$$

Arvosanan pitää olla enintään 6 eli 4, 5 tai 6.
Lasketaan näiden arvosanojen suhteelliset frekvenssit yhteen.

$$\frac{2+3+4}{30}\cdot 100\% = 30\%$$

$$P(\text{saadaan 4, 5 tai 6}) = 0{,}3$$

Vastaus: a) 0,2 b) 0,7 c) 0,3

5. **a)** Muodostetaan summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Arvosana	***f***	***sf***	***sf*%**
I	437	437	$\frac{437}{9719}\cdot 100\% = 4,5\ \%$
A	829	437 + 829 = 1266	$\frac{1266}{9719}\cdot 100\% = 13,0\ \%$
B	1863	1266 + 1863 = 3129	32,2 %
C	2388	5517	56,8 %
M	2021	7538	77,6 %
E	1653	9191	94,6 %
L	528	9719	100,0 %

b) Suhteellinen summafrekvenssi arvosanan C kohdalla on 56,8 %, joten näin suuri osuus kokelaista sai arvosanakseen korkeintaan C.

c) Suhteellisen summafrekvenssin mukaan vähintään arvosanan M sai kokelaista 100 % - 56,8 % = 43,2 %.

Vastaus: b) 56,8 % c) 43,2 %

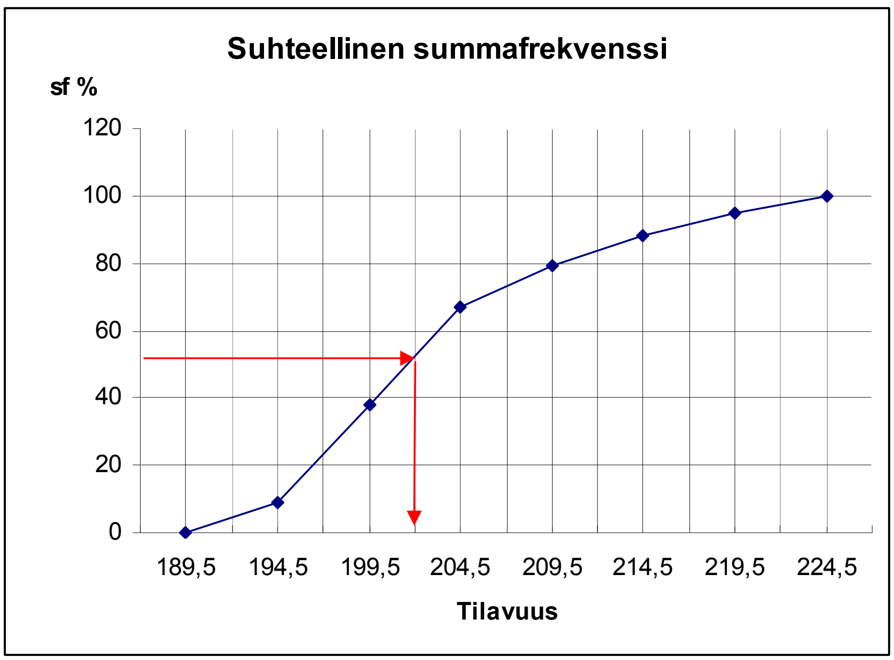
6. Muodostetaan suhteellinen summafrekvenssijakauma.

Ikä	***f***	***sf***	***sf*%**
20–24	64 257	64 257	$\frac{64\ 257}{264\ 622}\cdot 100\ \% = 24,3\ \%$
25–29	56 935	64 257 + 56 935 = 121 192	45,8 %
30–34	48 506	169 698	64,1 %
35–39	43 998	213 696	80,8 %
40–44	50 926	264 622	100,0 %

Mediaaniluokka on se luokka, jossa suhteellinen summafrekvenssi ensimmäisen kerran ylittää 50 %. Mediaaniluokka on 30–34 (vuotta).

7. Muodostetaan suhteellinen summafrekvenssijakauma.

Tilavuus (cm^3)	***f***	**Luokkakeskus (cm^3)**	***sf***	***sf*%**
190–194	240	$\frac{189{,}5+194{,}5}{2}=192$	240	$\frac{240}{2660}\cdot 100\% = 9\%$
195–199	760	197	240 + 720 = 1000	38 %
200–204	780	202	1780	67 %
205–209	320	207	2100	79 %
210–214	250	212	2350	88 %
215–219	170	217	2520	95 %
220–224	140	222	2660	100 %



Likiarvo voidaan lukea suhteellisensummafrekvenssin kuvaajalta.

$$\mathrm{Md} \approx \frac{199{,}5+204{,}5}{2} = 202\ \ (\mathrm{cm}^3)$$

Vastaus: 202 cm^3

8. Keskimääräinen poikueen koko on

$$\bar{x} = \frac{1\cdot 3+2\cdot 4+3\cdot 15+4\cdot 12+5\cdot 33+6\cdot 27+7\cdot 8+8\cdot 2}{104}$$
$$= \frac{503}{104}$$
$$= 4{,}836...$$
$$\approx 4{,}8$$

Vastaus: 4,8 poikasta

9. Frekvenssijakauma on

Ikä	***f***	**Luokkakeskus**
20–24	82 660	22
25–29	98 111	27
30–34	63 251	32
35–39	27 451	37
Yhteensä	271 473	

Keski-ikä on

$$\bar{x} = \frac{22\cdot 82\,660+27\cdot 98\,111+32\cdot 63\,251+37\cdot 27\,451}{271\,473}$$
$$= 27{,}653...$$
$$\approx 27{,}7$$

Vastaus: 27,7 vuotta

10. Osakkeen hinta saadaan painotettuna keskiarvona.

$$\begin{aligned}\bar{x} &= \frac{14{,}15+13{,}00+12{,}80+12{,}61+2\cdot 11{,}81+3\cdot 12{,}12}{9}\\ &= \frac{112{,}54}{9}\\ &= 12{,}504...\\ &\approx 12{,}5\ (€)\end{aligned}$$

Vastaus: 12,5 €

11. Koska lukujen keskiarvo on 8, saadaan yhtälö

$$\begin{aligned}\frac{(2x+1)+x+(x-3)+(-x+5)}{4} &= 8 \quad |\cdot 4\\ 2x+1+x+x-3-x+5 &= 32\\ 3x+3 &= 32\\ 3x &= 29 \quad |:3\\ x &= \frac{29}{3}\end{aligned}$$

Luvut ovat tällöin:

$$2x+1 = 2\cdot\frac{29}{3}+1 = \frac{58}{3}+\frac{3}{3} = \frac{61}{3} = 20\frac{1}{3}$$

$$x = \frac{29}{3} = 9\frac{2}{3}$$

$$x-3 = \frac{29}{3}-3 = \frac{29}{3}-\frac{9}{3} = \frac{20}{3} = 6\frac{2}{3}$$

$$-x+5 = -\frac{29}{3}+5 = -\frac{29}{3}+\frac{15}{3} = -\frac{14}{3} = -4\frac{2}{3}$$

Vastaus: $-4\frac{2}{3}$, $6\frac{2}{3}$, $9\frac{2}{3}$, $20\frac{1}{3}$

12. Lasketaan mittareiden virheet ja niiden keskiarvot.

Mittari *A*

T	$\boldsymbol{T_A}$	$\boldsymbol{T_A - T}$
+10,0	+11,2	+1,2
+6,0	+7,1	+1,1
+2,0	+0,7	-1,3
-2,0	-1,2	+0,8
-6,0	-6,9	-0,9
-10,0	-10,7	-0,7

Virheiden keskiarvo on

$$\begin{aligned}\bar{x}_A &= \frac{1{,}2+1{,}1-1{,}3+0{,}8-0{,}9-0{,}7}{6}\\ &= \frac{0{,}2}{6}\\ &= 0{,}0333...\\ &\approx 0{,}03\end{aligned}$$

Mittari *B*

T	$\boldsymbol{T_B}$	$\boldsymbol{T_B - T}$
+10,0	+10,6	+0,6
+6,0	+6,5	+0,5
+2,0	+2,4	+0,4
-2,0	-1,6	+0,4
-6,0	-5,7	+0,3
-10,0	-9,8	+0,2

Virheiden keskiarvo on

$$\begin{aligned}\bar{x}_B &= \frac{0{,}6+0{,}5+0{,}4+0{,}4+0{,}3+0{,}2}{6}\\ &= \frac{2{,}4}{6}\\ &= 0{,}4\end{aligned}$$

Mittarin *A* virheiden keskiarvo on pienempi kuin mittarin *B* virheiden keskiarvo. Mittari *B* on kuitenkin tarkempi. Se näyttää aina hieman liikaa.

Mittarin *A* virheiden keskiarvo on pienempi, koska virheet ovat erimerkkisiä. Yhteenlaskettuna virheet tällöin kumoavat toisiaan.

Parempi tunnusluku saadaan, jos otetaan huomioon virheiden itseisarvot.

Mittarin *A* virheiden keskiarvo on tällöin

$$\bar{x}_A = \frac{1{,}2 + 1{,}1 + 1{,}3 + 0{,}8 + 0{,}9 + 0{,}7}{6} = \frac{6}{6} = 1$$

13. Frekvenssijakauma:

Arvosana	***f***
4	3
5	6
6	7
7	11
8	10
9	5
10	2
Yhteensä	**44**

Keskiarvo

$$\bar{x} = \frac{3 \cdot 4 + 6 \cdot 5 + \ldots + 2 \cdot 10}{44} = \frac{306}{44} = 6{,}954\ldots \approx 7{,}0$$

Keskihajonta

$$s = \sqrt{\frac{3 \cdot (4 - 6{,}954\ldots)^2 + \ldots + 2(10 - 6{,}954\ldots)^2}{44 - 1}} = 1{,}5693\ldots \approx 1{,}6$$

Vastaus: $\bar{x} \approx 7{,}0$ ja $s \approx 1{,}6$

14. Metsäpinta-alat (ha):

13,4	23,6	45,3	21,1	34,6
70,2	89,4	110,3	46,9	98,7
85,4	12,5	54,9	67,3	87,5

a) Luokitellaan aineisto samankokoisiin luokkiin.

Ala (ha)	***f***	**Luokkakeskus (ha)**
10–29	4	$\frac{9,5+29,5}{2}=19,5$
30–49	3	29,5
50–69	2	59,5
70–89	4	79,5
90–109	1	99,5
110–129	1	119,5
Yhteensä	**15**	

b) Keskiarvo

$$\bar{x}=\frac{4\cdot 19,5+3\cdot 29,5+...+1\cdot 119,5}{15}=54,83...\approx 54,8\text{ (ha)}$$

Keskihajonta

$$s=\sqrt{\frac{4(19,5-54,83...)^2+3(29,5-54,83...)^2+...+(119,5-54,83...)^2}{15-1}}$$
$$=33,352...$$
$$\approx 33,4\text{(ha)}$$

c) $\bar{x}+2s=54,83...+2\cdot 33,352...=121,53...>89$

Arvo, joka poikkeaa keskiarvosta kaksi hajontaa on 121,53…(ha). Tämä on siis suurempi kuin 89 (ha), joten poikkeama ei ole merkittävä.

Vastaus: b) $\bar{x}\approx 54,8\,ha$, $s\approx 33,4$ ha   c) Ei poikkea merkittävästi.

15. Kuukausipalkkojen frekvenssijakauma:

Kuukausipalkka (€)	***f***
2500	2
3000	6
3340	7
3560	10
4010	6
4560	3
4930	1
Yhteensä	35

Lasketaan keskiarvo ja keskihajonta.

$$\bar{x} = \frac{2 \cdot 2500 + 6 \cdot 3000 + ... + 4930}{35} = 3561{,}428... \approx 3561 \text{ (€)}$$

$$s = \sqrt{\frac{2(2500 - 3561{,}428...)^2 + ... + (4930 - 3561{,}428...)^2}{35 - 1}}$$
$$= 557{,}2893...$$
$$\approx 557 \text{ (€)}$$

Yhden keskihajonnan päässä keskiarvosta olevat palkat ovat vähintään

3561,428… - 557,2893…= 3004,14… ≈ 3004 (€)

ja enintään

3561,428… + 557,2893… = 4118,72… ≈ 4119 (€)

b) $\bar{x} + 2s = 3561{,}428... + 2 \cdot 557{,}2893... = 4676{,}007...\text{(€)} > 4560 \text{ (€)}$

Arvo, joka poikkeaa keskiarvosta kaksi hajontaa on 4676,007…(ha). Tämä on siis suurempi kuin 4560 (ha), joten poikkeama ei ole merkittävä

Vastaus: a) 3004 € - 4119 €   b) Ei poikkea merkittävästi.

16. **a)** Normitetaan muuttujan arvo, $z_{70} = \dfrac{70-75}{5} = -1$

Kysytty todennäköisyys on

$$\begin{aligned} P(x < 70) &= \Phi(-1) \\ &= 1 - \Phi(1) \\ &= 1 - 0{,}8413 \\ &= 0{,}1587 \approx 0{,}16 \end{aligned}$$

b) Normitetaan muuttujan arvot, $z_{68} = \dfrac{68-75}{5} = -1{,}4$ ja $z_{85} = \dfrac{85-75}{5} = 2$.

Kertymä normitettuun arvoon z = -1,4 mennessä on

$$\begin{aligned} \Phi(-1{,}4) &= 1 - \Phi(1{,}4) \\ &= 1 - 0{,}9192 \\ &= 0{,}0808 \end{aligned}$$

Kysytty todennäköisyys on

$$\begin{aligned} &P(68 < x < 85) \\ &= \Phi(2) - \Phi(-1{,}4) \\ &= 0{,}9772 - 0{,}0808 \\ &= 0{,}8964 \\ &\approx 0{,}9 \end{aligned}$$

Vastaus: a) 0,16 b) 0,9

17. Kokeen keskiarvo oli 72 pistettä. Pistemäärät noudattavat normaalijakaumaa, joten odotusarvo $\mu = 72$.

Alina sai 82 pistettä kokeesta, jonka pisteiden keskihajonta $\sigma_A = 9{,}2$.
Bertta sai 80 pistettä kokeesta, jonka pisteiden keskihajonta $\sigma_B = 6{,}8$.

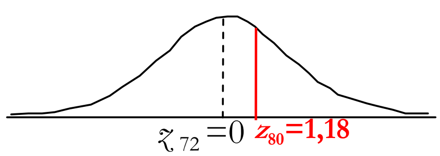
Normitetaan pistemäärät:

$$z_{82} = \frac{82-72}{9{,}2} = \frac{10}{9{,}2} = 1{,}0869... \approx 1{,}09$$

$$z_{80} = \frac{80-72}{6{,}8} = \frac{8}{6{,}8} = 1{,}1764... \approx 1{,}18$$

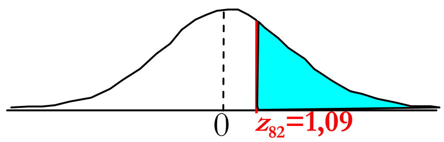
Koska 1,18 > 1,09, Bertta pärjäsi paremmin.

Lukio B:



P(pisteet yli 72, mutta alle 80)
$= \Phi(1{,}18) - \Phi(0)$
$= 0{,}8810 - 0{,}5$
$= 0{,}381$
$\approx 0{,}38 = 38\ \%$

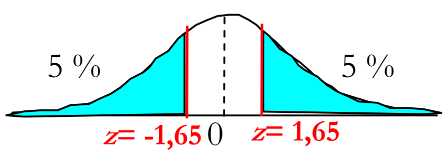
Lukio A:



$P(\text{pisteet yli } 82)$
$= 1 - \Phi(1{,}09)$
$= 1 - 0{,}8621$
$= 0{,}1379$
$\approx 0{,}14 = 14\,\%$

Vastaus: Bertta pärjäsi paremmin. $P(\text{pisteet yli } 82) = 14\,\%$,
$P(\text{pisteet yli } 72\text{, mutta alle } 80) = 38\,\%$

18. $N(25000,\ 2000)$
Tulpan toimintavarmuus on alle 95 %
Etsitään kohta, johon mennessä toimintavarmuus on vähintään 95 %.
$\Phi(1{,}6449) = 0{,}95$, joten $\Phi(-1{,}6449) = 1 - 0{,}95 = 0{,}05 = 5\%$



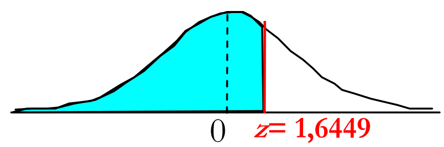
Merkitään ajokilometrejä kirjaimella x.

$$\frac{x - 25\,000}{2000} = -1{,}6449$$
$$x - 25\,000 = -3289{,}8$$
$$x = 21\,710{,}2 \approx 21\,700 \text{ (km)}$$

Vastaus: 21 700 km

19. $N(25, \sigma)$
Haastattelu ei saa 95 % varmuudella ylittää 30 min.
Etsitään normitettu arvo, johon mennessä kertymä on 95 %.

$\Phi(1{,}6449) = 0{,}95 = 95\,\%$

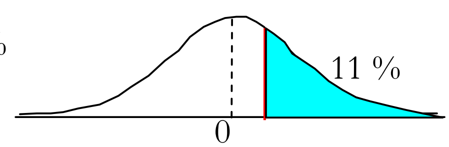


Tällöin siis 30 min normitettu arvo on 1,6449.

$$z_{30} = \frac{30-25}{\sigma} = 1{,}6449$$
$$\frac{5}{\sigma} = 1{,}6449$$
$$\sigma = \frac{5}{1{,}6449}$$
$$= 3{,}030... \approx 3{,}0\,(\text{min})$$

Vastaus: Hajonta korkeintaan 3,0 min

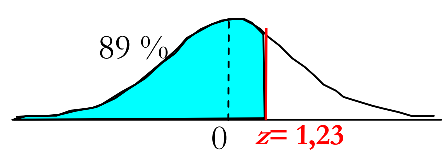
20. $P(\text{vähintään 47 tikkua}) = \dfrac{11}{100} = 0{,}11 = 11\,\%$



100 % - 11 % = 89 %

Etsitään normitettu arvo, johon mennessä on kertynyt 89 % (eli rasiassa enintään 47 tikkua).

$\Phi(1{,}23) = 0{,}8907 \approx 0{,}89$



Merkitään keskimääräistä tikkujen määrää kirjaimella μ.

$$z_{47} = \frac{47-\mu}{4} = 1{,}23 \quad |\cdot 4$$

$$47-\mu = 4{,}92$$

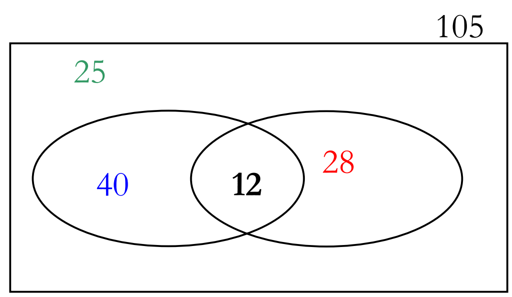
$$-\mu = -42{,}08$$

$$\mu = 42{,}08 \approx 42$$

Vastaus: 42 tikkua

21. Laaditaan Venn-diagrammi.

Tyttöjä on 52, joista **12** harrastaa sählyä.
Sählyä harrastaa yhteensä 40 oppilasta, joten poikia näistä on 40 – 12 = 28.
Poikia, jotka harrastavat muuta kuin sählyä on 105 – 28 – 12 - 40 = 25



a) $P(\text{tyttö, ei sähly}) = \dfrac{40}{105} = 0{,}380... \approx 0{,}38$

b) $P(\text{poika, sähly}) = \dfrac{28}{105} = 0{,}266... \approx 0{,}27$

c) $P(\text{poika, ei sähly}) = \dfrac{25}{105} = 0{,}2380... \approx 0{,}24$

Vastaus: a) 0,38 **b)** 0,27 c) 0,24

22. Alkuruokavaihtoehtoja on 2 kpl.
Pääruokavaihtoehtoja on 2 kpl.
Jälkiruokavaihtoehtoja on 2 kpl.

Eri menu-vaihtoehtoja on siis $2 \cdot 2 \cdot 2 = 8$ (kpl).

$$P(\text{keitto, uunikala, sorbetti}) = \frac{1}{8} = 0{,}125$$

Vastaus: $\frac{1}{8}$

23. Loppukilpailuun voi Nean (N) ja Leevin (L) lisäksi päästä henkilö A, B tai C. Muodostetaan kaikki mahdolliset parit:

NL	NA
NB	NC
LA	LB
LC	AB
AC	BC

Erilaisia pareja on yhteensä 10.

a) Jos valittu pari on NL, NA, NB tai NC, Nea pääsee loppukilpailuun eli

$$P(\text{Nea pääsee}) = \frac{4}{10} = \frac{2}{5}$$

b) Parit, joissa Leevi pääsee, mutta Nea ei, ovat LA, LB ja LC. Näin ollen

$$P(\text{Leevi pääsee, Nea ei}) = \frac{3}{10}$$

Vastaus: a) $\frac{2}{5}$ b) $\frac{3}{10}$

24. **a)** Viikonpäiviä on 7 kpl.

$$P(\text{syntynyt maanantaina}) = \frac{1}{7}$$

b) Kuukausia vuodessa on 12.

$$P(\text{syntynyt tammikuussa}) = \frac{1}{12}$$

c) Suotuisia ovat mainitut 4 tuntia vuorokauden 24 tunnista, joten

$$P(\text{syntynyt klo } 12-16) = \frac{4}{24} = \frac{1}{6}$$

Vastaus: a) $\frac{1}{7}$ b) $\frac{1}{12}$ c) $\frac{1}{6}$

25. Lamppuja on yhteensä 300.

a)Lamppuja, jotka ovat palaneet jo 300 tuntia, on 300 – 12 = 288.
Niistä vielä korkeintaan 300 tuntia palaa 145, joten

$$P(\text{palaa vielä 300 tuntia}) = \frac{145}{288} = 0,5034... \approx 0,50$$

b) 600 tuntia palaneita lamppuja on 300 – 12 – 145 = 143.
Niistä yli 900 tuntia toimii vain 300 – 12 – 145 – 120 = 23, joten

$$P(\text{toimii yli 900 tuntia}) = \frac{23}{143} = 0,1608... \approx 0,16$$

Vastaus: a) 0,50 b) 0,16

26. Suotuisia tapauksia kuvaa Suomen pinta-ala. Alkeistapauksia kuvataan koko maapallon alalla, joka on

$$\begin{aligned} A(\text{pallo}) &= 4\pi r^2 \\ &= 4 \cdot \pi \cdot (6370\ \text{km})^2 \\ &= 509\,904\,363{,}8...\text{km}^2 \end{aligned}$$

$$P(\text{meteoriitti putoaa Suomeen}) = \frac{337\,000\ \text{km}^2}{509\,904\,363{,}8...\text{km}^2} = 0{,}0006609... \approx 0{,}00066$$

Vastaus: 0,00066

27. Kahden nopan heitossa alkeistapauksia on 6 · 6 = 36. Näistä on lihavoitu ne, joissa jälkimmäisellä heitolla saadaan suurempi silmäluku.

(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,3)	**(2,3)**	(3,3)	(4,3)	(5,3)	(6,3)
(1,4)	**(2,4)**	**(3,4)**	(4,4)	(5,4)	(6,4)
(1,5)	**(2,5)**	**(3,5)**	**(4,5)**	(5,5)	(6,5)
(1,6)	**(2,6)**	**(3,6)**	**(4,6)**	**(5,6)**	(6,6)

Tämän tapahtuman todennäköisyys on siis $\frac{15}{36} = \frac{5}{12}$.

Jos ensimmäisen nopanheiton tulos oli 3, kaksi seuraavaa tulosta voivat olla vain(4,5), (4,6) ja (5,6), jotta annettu ehto toteutuisi. Näin ollen tapahtuman ”ensimmäisellä saadaan kolme, toisella tätä suurempi ja kolmannella taas edellistä suurempi silmäluku” todennäköisyys on $\frac{3}{36} = \frac{1}{12}$.

Vastaus:

P(jälkimmäinen suurempi)=$\frac{5}{12}$, P(jälkimmäinen suurempi, jos I oli 3)=$\frac{1}{12}$

28. $P(\text{malaria}) = 0{,}30$, $P(\text{ei sairastu}) = 1 - 0{,}30 = 0{,}70$

a) $P(\text{kukaan ei sairastu}) = 0{,}7^8 \approx 0{,}058$

b) $P(\text{ainakin yksi sairastaa}) = 1 - 0{,}7^8 \approx 0{,}94$

Vastaus: a) 0,058 b) 0,94

29. $P(\text{arpa voittaa}) = \frac{1}{4}$, $P(\text{arpa ei voita}) = \frac{3}{4}$

a) $P(\text{ainakin yksi voittaa}) = 1 - \left(\frac{3}{4}\right)^6 \approx 0{,}82$

b) $P(\text{voittaa korkeintaan viidellä}) = 1 - P(\text{voittaa kuudella})$

$$= 1 - \left(\frac{1}{4}\right)^6 \approx 0{,}9998$$

Vastaus: a) 0,82 b) 0,9998

30. Noppaa heitetään 5 kertaa.

a) $P(\text{saadaan 5 kuutosta}) = \left(\frac{1}{6}\right)^5 = 0{,}0001286... \approx 0{,}00013$

b) Silmälukuja, jotka ovat korkeintaan neljä, on neljä kuudesta.

$$P(\text{saadaan joka heitolla kork. nelonen}) = \left(\frac{4}{6}\right)^5 = 0{,}1316... \approx 0{,}13$$

Vastaus: a) 0,00013 b) 0,13

31. $P(\text{sairastaa}) = \dfrac{1}{3}$, $P(\text{ei sairasta}) = \dfrac{2}{3}$

$$\begin{aligned} P(\text{ainakin yksi sairastaa}) &= 1 - P(\text{ei kukaan sairasta}) \\ &= 1 - \left(\frac{2}{3}\right)^8 \approx 0{,}96 \end{aligned}$$

Vastaus: 0,96

32. P(laskimissa on virhe) = 0,02
P(ei ole virhettä) = 1 – 0,02 = 0,98

$$\begin{aligned} &P(\text{yksikään 26 laskimesta ei ole viallinen}) \\ &= 0{,}98^{26} \\ &= 0{,}5913... \\ &\approx 0{,}59 \end{aligned}$$

Vastaus: 0,59

33. Koska kortteja ei palauteta pakkaan, joka nostolla pakassa on yksi kortti vähemmän.

a) Punaisia maita on pakassa 26, jolloin

$$P(\text{saadaan vain punaisia maita}) = \frac{26}{52} \cdot \frac{25}{51} \cdot \frac{24}{50} \cdot \frac{23}{49} = 0{,}05522... \approx 0{,}055$$

b) Pakassa on 12 kuvakorttia, joten

$$P(\text{saadaan vain kuvakortteja}) = \frac{12}{52} \cdot \frac{11}{51} \cdot \frac{10}{50} \cdot \frac{9}{49} = 0{,}001828... \approx 0{,}0018$$

c) Ensimmäisellä kerralla kelpaa mikä tahansa kortti. Toisella nostolla suljetaan pois 13 korttia suotuisten joukosta, koska yksi maa on jo käytetty. Näin toimitaan myös kahdella seuraavalla nostolla.

$$P(\text{saadaan joka kortilla eri maa}) = \frac{52}{52} \cdot \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49} = 0{,}1054... \approx 0{,}11$$

Vastaus: a) 0,055 b) 0,0018 c) 0,11

34. Olkoon suomalaisia kolikoita x kappaletta. Näin ollen todennäköisyys, että nostetaan molemmilla kerroilla suomalainen kolikko, on

$$\frac{x}{12} \cdot \frac{x-1}{11} = \frac{x^2 - x}{132}$$

Tämän tapahtuman todennäköisyyden tuli tehtävänannon mukaan olla $\frac{5}{33}$. Saadaan siis yhtälö

$$\frac{x^2 - x}{132} = \frac{5}{33} \quad \Big| \cdot 132$$

$$x^2 - x = 20$$

$$x^2 - x - 20 = 0$$

$$x = \frac{-(-1) \pm \sqrt{(-1)^2 - 4 \cdot 1 \cdot (-20)}}{2 \cdot 1}$$

$$x = \frac{1 \pm \sqrt{81}}{2}$$

$$x = \frac{1 \pm 9}{2}$$

$$x = 5 \quad \text{tai} \quad x = -4$$

Koska kolikoiden määrä ei voi olla negatiivinen luku, vain $x = 5$ kelpaa ratkaisuksi. Saksalaisia kolikoita on siis 12 – 5 = 7 kappaletta.

Vastaus: 7 kpl

35. Olkoon todennäköisyys saada ykkönen x.
Tällöin muiden silmälukujen todennäköisyydet ovat $2x$, $3x$, $4x$, $5x$ ja $6x$.

Koska silmälukujen todennäköisyyksien summa on yksi, saadaan yhtälö

$$\begin{aligned} x+2x+3x+4x+5x+6x &= 1 \\ 21x &= 1 \quad |:21 \\ x &= \frac{1}{21} \end{aligned}$$

Silmälukujen 1, 2, 3, 4, 5 ja 6 todennäköisyydet ovat siis

$$\frac{1}{21}, \frac{2}{21}, \frac{3}{21}, \frac{4}{21}, \frac{5}{21} \text{ ja } \frac{6}{21}.$$

Tapahtuman ”kahdella heitolla saadaan kaksi kuutosta” todennäköisyys on näin ollen

$$\frac{6}{21} \cdot \frac{6}{21} = 0{,}0816... \approx 0{,}082$$

36.

Kultakoruja	4
Hopeakoruja	5
Pronssikoruja	3
Yhteensä	12

Nostetaan kaksi korua.

a) $P(\text{molemmat hopeaa}) = \frac{5}{12} \cdot \frac{4}{11} = \frac{5}{33} \approx 0{,}15$

b) P(molemmat samaa metallia)

= P(molemmat kultaa) + P(molemmat hopeaa) + P(molemmat pronssia)

$$=\frac{4}{12}\cdot\frac{3}{11}+\frac{5}{12}\cdot\frac{4}{11}+\frac{3}{12}\cdot\frac{2}{11}$$

$$=\frac{19}{66}\approx 0{,}29$$

Vastaus: a) 0,15 b) 0,29

37.

Tytöt	suomenkieliset	195
	ruotsinkieliset	85
Pojat	suomenkieliset	26
	ruotsinkieliset	52

a) Valitaan umpimähkään yksi opiskelija.

P(suomenkielinen tai poika)

= P(suomenkielinen") + P(poika) - P(suomenkielinen poika)

$$=\frac{221}{358}+\frac{78}{358}-\frac{26}{358}$$

$$=\frac{273}{358}$$

$$=0{,}7625...\approx 0{,}76$$

b) Valitaan kaksi opiskelijaa.

P(ainakin toinen ruotsinkielinen)

= P(1. ruotsinkielinen, 2. ei) + P(1. ei ole ruotsinkielinen, 2. on) + P(molemmat ruotsinkielisiä)

$$=\frac{137}{358}\cdot\frac{221}{357}+\frac{221}{358}\cdot\frac{137}{357}+\frac{137}{358}\cdot\frac{136}{357}$$

$$=0{,}61957...$$

$$\approx 0{,}62$$

Vastaus: a) 0,76 b) 0,62

38. 10 pisteen kortteja pakassa 16 kpl, mukana kuvat ja kympit

11 pisteen kortteja pakassa 4 kpl.

a) P(saadaan 2:lla kortilla summaksi 21)
= P(1. kortti ässä, 2. kuva tai kymppi) + P(1. kortti kuva tai kymppi, 2. on ässä)

$$= \frac{4}{52} \cdot \frac{16}{51} + \frac{16}{52} \cdot \frac{4}{51}$$
$$= 0{,}04826...$$
$$\approx 0{,}048$$

b) Kahden ensimmäisen kortin pistesumma on 10 + 3 = 13. Jos pelaaja nostaa kolmannella kortilla **9**, **10** tai **kuvan**, peli menee metsään. Pakassa on jäljellä näitä kortteja seuraavasti:

9 4 kpl
10 4 kpl
kuvat 11 kpl

$$P(\text{pistesummaksi yli } 21) = \frac{19}{50} = 0{,}38$$

Vastaus: a) 0,048 b) 0,38

39. **a)** Tyttö voidaan valita 6:lla tavalla ja poika samoin eli vaihtoehtoja on $6 \cdot 6 = 36$ (kpl).

b) Poika-tyttö-jonoja, joissa jonon ensimmäinen on poika, on
$6 \cdot 6 \cdot 5 \cdot 5 \cdot 4 \cdot 4 \cdot 3 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = 518\,400$ (kpl)

Vastaus: a) 36 b) 518 400

40. **a)** Kahdestatoista voidaan valita neljä henkilöä $\binom{12}{4} = 495$ eri tavalla.

b) Jos Tupu, Hupu ja Lupu ovat mukana, niin viimeinen pelaaja voidaan valita 12 – 3 = 9 henkilön joukosta.

On siis yhdeksän suotuisaa tapausta.

$$P(\text{samassa pöydässä}) = \frac{9}{495} \approx 0{,}018$$

Vastaus: a) 495 b) 0,018

41. **a)** $P(\text{I kirjain vokaali}) = \frac{5}{10} = 0{,}5$

b) Kolmesta kirjaimesta vain yksi on vokaali.

$$P(\text{vain yksi kirjain vokaali}) = \frac{\binom{5}{1}\binom{5}{2}}{\binom{10}{3}} = \frac{5 \cdot 10}{120} = \frac{5}{12} \approx 0{,}42$$

c) $P(\text{ I, L ja O jossain järjestyksessä}) = \dfrac{\binom{2}{1}\binom{2}{1} \cdot 1}{\binom{10}{3}} = \dfrac{1}{30} \approx 0{,}030$

Vastaus: a) 0,5 b) 0,42 c) 0,030

42. Arvotaan 6 oikeaa numeroa ja 2 lisänumeroa (48 numerosta).

a) $P(\text{neljä oikein}) = \dfrac{\binom{6}{4}\binom{42}{2}}{\binom{48}{6}} \approx 0,0011$

b) $P(\text{viisi ja lisä oikein}) = \dfrac{\binom{6}{5}\binom{2}{1}}{\binom{48}{6}} \approx 9,8 \cdot 10^{-7}$

Vastaus: a) 0,0011 b) $9{,}8 \cdot 10^{-7}$

43. $P(\text{poika}) = 0,513$

$P(\text{tyttö}) = 1 - 0,513 = 0,487$

$P(\text{puolet tyttöjä, puolet poikia}) = \binom{50}{25} \cdot 0,513^{25} \cdot 0,487^{25} \approx 0,11$

Vastaus: 0,11

44. $P(\text{saadaan kuutonen}) = \frac{1}{6}$

$P(\text{ei saada kuutosta}) = \frac{5}{6}$

$$\begin{aligned}
&P(\text{saadaan kuutonen ainakin kahdesti})\\
&= 1 - P(0 \text{ tai } 1 \text{ kertaa kuutonen})\\
&= 1 - \left[\left(\frac{5}{6}\right)^6 + \binom{6}{1}\frac{1}{6}\left(\frac{5}{6}\right)^5\right]\\
&\approx 0{,}26
\end{aligned}$$

Koska todennäköisyys on alle 0,5, niin ei kannata lyödä vetoa.

Vastaus: Ei kannata.

45. $P(\text{pun.vihr. sokea}) = 0{,}08$

$P(\text{ei pun.vihr. sokea}) = 1 - 0{,}08 = 0{,}92$

a) $P(3 \text{ kpl sokeita}) = \binom{8}{3} \cdot 0{,}08^3 \cdot 0{,}92^5 \approx 0{,}019$

b) $P(\text{korkeintaan 2 kpl sokeita})$

$$= 0{,}92^8 + \binom{8}{1} \cdot 0{,}08 \cdot 0{,}92^7 + \binom{8}{2} \cdot 0{,}08^2 \cdot 0{,}92^6 \approx 0{,}98$$

c) $P(\text{vähintään 7 kpl sokeita})$

$$= \binom{8}{7} \cdot 0{,}08^7 \cdot 0{,}92 + 0{,}08^8 \approx 1{,}6 \cdot 10^{-7}$$

Vastaus: a) 0,019 b) 0,98 c) $1{,}6 \cdot 10^{-7}$

Harjoituskokeiden ratkaisut

Harjoituskoe 1

1. Heittotulosten frekvenssijakauma:

Silmäluku	***f***	***sf***
1	9	9
2	7	9 + 7 = 16
3	5	21
4	4	25
5	8	33
6	6	39

a) Koska jokainen opiskelija on heittänyt kerran noppaa, niin heittojen lukumäärän summa on yhtä suuri kuin opiskelijoiden lukumäärä.

$$9 + 7 + 5 + 4 + 8 + 6 = 39$$

b) Moodi on se silmäluku, jonka frekvenssi (heittojen lukumäärä) on suurin.

$$\text{Mo} = 1$$

Mediaani on se silmäluku, jonka kohdalla summafrekvenssi ensimmäisenä ylittää rajan 20, koska $\frac{39}{2} = 19{,}5$.

$$\text{Md} = 3$$

Keskiarvo on

$$\bar{x} = \frac{9 \cdot 1 + 7 \cdot 2 + \ldots + 6 \cdot 6}{39} = 3{,}333\ldots \approx 3{,}33$$

c) Keskihajonta on

$$s = \sqrt{\frac{9(1-3{,}33...)^2 + 7(2-3{,}33...)^2 + ... + 6(6-3{,}33...)^2}{39-1}}$$
$$= 1{,}8400...$$
$$\approx 1{,}84$$

Vastaus: a) 39 b) Mo = 1, Md = 3, $\bar{x} \approx 3{,}33$ c) $s \approx 1{,}84$

2. **a)** Kuusitahkoisessa nopassa kaikkien mahdollisten alkeistapausten määrä on 6.

Ensimmäinen tyttö voi heittää minkä tahansa silmäluvun. Suotuisia alkeistapauksia on siis 6 kappaletta.

Toisen ja kolmannen tytön on saatava juuri tuo sama silmäluku. Kummallakin on siis suotuisia alkeistapauksia 1 kappale.

P(ensimmäinen tyttö heittää minkä tahansa silmäluvun JA toinen heittää saman JA kolmas heittää saman)

$$= \frac{6}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = 0{,}0277... \approx 0{,}03$$

b) Jokaisen pojan todennäköisyys saada heitollaan vähintään silmäluku 6 on sama. Poikien heittämässä nopassa kaikkien alkeistapausten määrä on 12. Näistä suotuisia tapauksia ovat silmäluvut 6, 7, 8, 9, 10, 11 ja 12. Suotuisia alkeistapauksia on siis 7 kappaletta.

P(ensimmäinen poika saa vähintään 6 JA toinen poika saa vähintään 6 JA kolmas poika saa vähintään 6)

$$= \frac{7}{12} \cdot \frac{7}{12} \cdot \frac{7}{12} = \frac{343}{1728} = 0{,}1984... \approx 0{,}20$$

Vastaus: a) 0,03 b) 0,20

3. Kyseessä on toistokoe. Todennäköisyys saada kolikonheitossa kruuna on 0,5.

$$
\begin{aligned}
&P(\text{kymmenellä heitolla saadaan kuusi kruunaa})\\
&=\binom{10}{6}\cdot 0{,}5^{6}\cdot 0{,}5^{10-6}\\
&=210\cdot 0{,}5^{6}\cdot 0{,}5^{4}\\
&=0{,}2050...\\
&\approx 0{,}21
\end{aligned}
$$

Vastaus: 0,21

4. Jokaiseen pinoon tulee $\frac{52}{4}=13$ korttia.

a) Kun jokaisesta neljästä pinosta otetaan yksi kortti, erilaisia neljän kortin ryhmiä saadaan

$$13\cdot 13\cdot 13\cdot 13=28\ 561 \text{ kappaletta.}$$

b) Kun nostettua korttia ei palauteta takaisin pakkaan, korttien määrä vähenee jokaisella kierroksella yhdellä.

Tapauksen ”vähintään yksi ässä” komplementti on ”ei yhtään ässää”. Koska korttipakassa on 4 ässää, on ensimmäisellä nostokerralla komplementin suotuisia tapauksia

52 – 4 = 48 kappaletta.

Toisella nostokerralla suotuisia tapauksia on 47, kolmannella 46 jne.

P(kahdeksalla kortilla ei yhtään ässää)
= P(ensimmäisellä kortilla joku muu kuin ässä JA toisella kortilla joku muu kuin ässä JA…JA kahdeksannella kortilla joku muu kuin ässä)

$$=\frac{\cancel{48}}{52}\cdot\frac{\cancel{47}}{51}\cdot\frac{\cancel{46}}{50}\cdot\frac{\cancel{45}}{49}\cdot\frac{44}{\cancel{48}}\cdot\frac{43}{\cancel{47}}\cdot\frac{42}{\cancel{46}}\cdot\frac{41}{\cancel{45}}$$

$$=\frac{44\cdot 43\cdot 42\cdot 41}{52\cdot 51\cdot 50\cdot 49}$$

$$=\frac{3\,528\,024}{6\,497\,400}$$

$$=0,50143...$$

Kysytty todennäköisyys on siis

P(kahdeksalla kortilla vähintään yksi ässä)
= 1 – P(kahdeksalla kortilla ei yhtään ässää)

$$=1-\frac{3528024}{6497400}$$

$$=1-0,50143...$$

$$=0,4985...$$

$$\approx 0,50$$

Vastaus: a) 28 561 b) 0,50

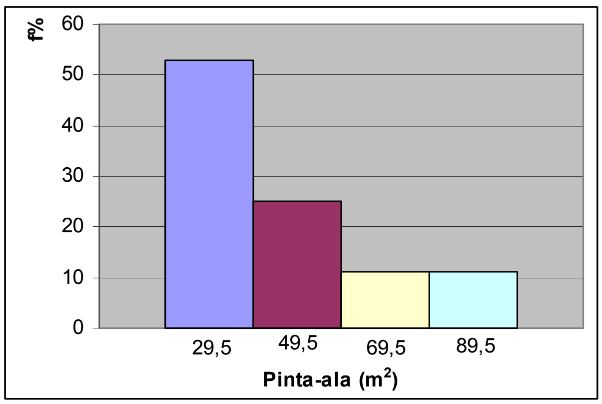
5. **a)** Muodostetaan frekvenssijakaumat.

Pinta-ala (m^2)	***f***	***f %***	***sf***	***sf %***
20–39	133 823	$\frac{133823}{250466}\approx 53\%$	133823	53 %
40–59	61 571	25 %	133823 + 61571 = 195 394	$\frac{195394}{250466}\approx 78\%$
60–79	27 967	11 %	223 361	89 %
80–99	27 105	11 %	250 466	100 %

b) Alle 60 m^2 on suhteellisen summafrekvenssin mukaan 78 % mökeistä.

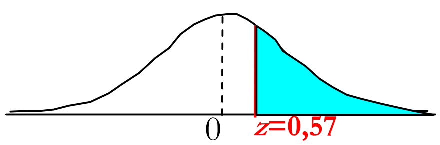
c) Mediaani luokka on se luokka, jossa suhteellinen summafrekvenssi ensimmäisenä ylittää 50 %. Mediaaniluokka on siis 20–39 (m^2).

d) Koska pinta-ala on jatkuva muuttuja, suhteellisen frekvenssin kuvaaja on histogrammi.



6. $\mu = 115\,\mathrm{g},\ \sigma = 5{,}3\,\mathrm{g}$

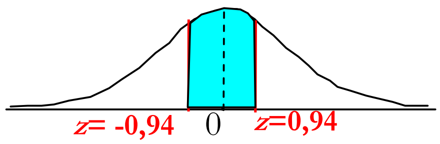
a) $z_{118} = \dfrac{118-115}{5{,}3} = \dfrac{3}{5{,}3} = 0{,}566... \approx 0{,}57$



$$\begin{aligned} P(\text{yli } 118\text{ g}) &= 1-\Phi(0{,}57) \\ &= 1-0{,}7157 \\ &= 0{,}2843 \approx 0{,}28 \end{aligned}$$

b) $z_{110} = \dfrac{110-115}{5,3} = \dfrac{-5}{5,3} = -0,943... \approx -0,94$

$z_{120} = \dfrac{120-115}{5,3} = \dfrac{5}{5,3} = 0,943... \approx 0,94$



P(yli 110 g, mutta alle 120 g)

$= \Phi(0,94) - \Phi(-0,94)$

$= 0,8264 - [1 - 0,8264]$

$= 0,8264 - 0,1736$

$= 0,6528 \approx 0,65$

Vastaus: a) 0,28 b) 0,65

Harjoituskoe 2

1. Määritetään ensin aineiston frekvenssijakaumat.

x	***f***	***f%***	***sf%***
2	4	$\frac{4}{74} \approx 5{,}4\ \%$	5,4 %
3	3	4,1 %	5,4 % + 4,1 % = 9,5 %
4	7	9,5 %	18,9 %
5	17	23,0 %	41,9 %
6	10	13,5 %	55,4 %
7	9	12,2 %	67,6 %
8	17	23,0 %	90,5 %
9	5	6,8 %	97,3 %
10	2	2,7 %	100,0 %
Yhteensä	**74**		

a) Koska suhteellisen summafrekvenssin mukaan enintään 6 pistettä sai 55,4 %, vähintään 7 pistettä sai loput eli

100 % - 55,4 % = 44,6 % ≈ 45 %.

b) Mediaani saadaan suhteellisesta summafrekvenssin avulla. Mediaani sijaitsee kohdassa, jossa 50 % täyttyy. Tämän aineiston mediaani on siis 6.

Moodi on muuttujan arvoista yleisin eli Mo = 5 tai Mo = 8.

Aineiston keskiarvo $\bar{x}$

$$\begin{aligned}\bar{x} &= \frac{2\cdot 4+3\cdot 3+4\cdot 7+\ldots+10\cdot 2}{74}\\ &= 6{,}135\ldots\\ &\approx 6{,}1 \quad \text{(pistettä)}\end{aligned}$$

c) Keskihajonta voidaan laskea laskimen tilastotoimintojen avulla. Aineisto syötetään laskimeen, jolloin (otos)keskihajonnaksi saadaan

$$s = 1{,}98848... \approx 1{,}99 \text{ (pistettä)}$$

d) $\bar{x} + 2s = 6{,}135... + 2 \cdot 1{,}98848... = 10{,}111... > 10$

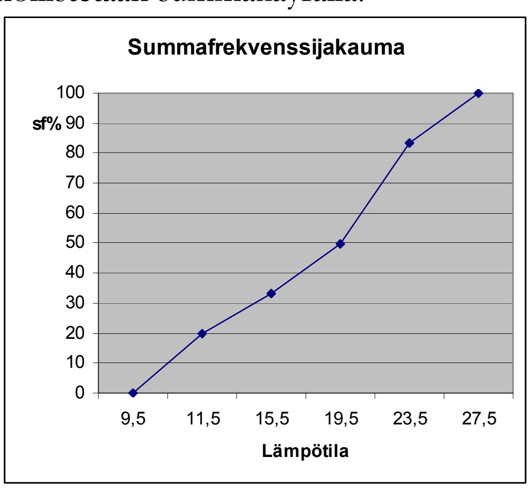
Pistemäärä 10 ei siis ole vähintään kahden keskihajonnan päässä keskiarvosta, joten poikkeama ei ole merkittävä.

Vastaus: a) 45 % b) Mo = 8 tai Mo = 5, Md = 6, $\bar{x} \approx 6{,}1$ c) $s \approx 1{,}99$
d) Ei poikkea

2. Muodostetaan frekvenssijakaumat.

Lämpötila (°C)	***f***	***sf***	***f*%**	***sf*%**
10–13	6	6	$\frac{6}{30} = 20\ \%$	20 %
14–17	4	6 + 4 = 10	13 %	$\frac{10}{30} \approx 33\ \%$
18–21	5	15	17 %	50 %
22–25	10	25	33 %	83 %
26–29	5	30	17 %	100 %

Lämpötila on jatkuva muuttuja, joten suhteellista summafrekvenssiä havainnollistetaan summakäyrällä.



3. Itämistodennäköisyydet siemenille R ja S:

	Itää	**Ei idä**
R	0,9	0,1
S	0,8	0,2

Nostetaan pussista kaksi siementä.

a) Kaksi itävää siementä voidaan nostaa seuraavilla tavoilla (viereen on laskettu kyseisen tapahtuman todennäköisyys):

RR $0{,}4^2 \cdot 0{,}9^2 = 0{,}1296$
RS $0{,}4 \cdot 0{,}9 \cdot 0{,}6 \cdot 0{,}8 = 0{,}1728$
SR $0{,}6 \cdot 0{,}8 \cdot 0{,}4 \cdot 0{,}9 = 0{,}1728$
SS $0{,}6^2 \cdot 0{,}8^2 = 0{,}2304$

Kysytty todennäköisyys on em. todennäköisyyksien summa.

$$\begin{aligned}&P(\text{molemmat itävät})\\&= 0{,}1296 + 2 \cdot 0{,}1728 + 0{,}2304\\&= 0{,}7056\\&\approx 0{,}71\end{aligned}$$

b) Tapahtuman "ainakin toinen itää" komplementti "kumpikaan ei idä" on helpompi laskea. Tämä sisältää seuraavat tapahtumat

RR $0{,}4^2 \cdot 0{,}1^2 = 0{,}0016$
RS $0{,}4 \cdot 0{,}1 \cdot 0{,}6 \cdot 0{,}2 = 0{,}0048$
SR $0{,}6 \cdot 0{,}2 \cdot 0{,}4 \cdot 0{,}1 = 0{,}0048$
SS $0{,}6^2 \cdot 0{,}2^2 = 0{,}0144$

$$\begin{aligned}&P(\text{ainakin toinen itää})\\&= 1 - P(\text{kumpikaan ei idä})\\&= 1 - \left(0{,}0016 + 0{,}0048 \cdot 2 + 0{,}0144\right)\\&= 0{,}9744\\&\approx 0{,}97\end{aligned}$$

Vastaus: a) 0,71 b) 0,97

4. $P(\text{peruuttaa}) = 0{,}07$

$P(\text{pitää paikkansa}) = 1 - 0{,}07 = 0{,}93$

Pekka saa paikan, jos neljä tai enemmän peruuttaa paikkansa.

Vastatapahtuma on: ”ei yhtään tai korkeintaan kolme peruuttaa”.
Lasketaan vastatapahtuman todennäköisyys.

$$\binom{120}{0}\cdot 0{,}07^0 \cdot 0{,}93^{120} + \binom{120}{1}\cdot 0{,}07 \cdot 0{,}93^{119} + \binom{120}{2}\cdot 0{,}07^2 \cdot 0{,}93^{118}$$
$$+\binom{120}{3}\cdot 0{,}07^3 \cdot 0{,}93^{117}$$
$$= 0{,}0281...$$

$$P(\text{Pekka saa paikan}) = 1 - 0{,}0281... = 0{,}9718... \approx 0{,}97$$

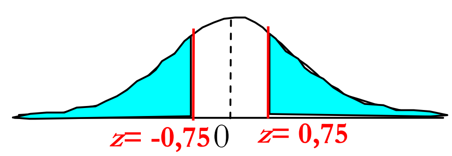
Vastaus: 0,97

5. Pituus on normaalisti jakautunut, $\mu = 160$ ja $\sigma = 20$

a) Pituutta 145 cm vastaa normitettu arvo

$$z_{145} = \frac{145-160}{20} = -0{,}75$$

Koska normaalijakauman kertymäfunktion arvot on taulukoitu vain positiivisille normitetuille arvoille, käytetään hyväksi jakauman symmetriaa: arvoa 0,75 vastaa prosenttiluku 77,34 %. Näin ollen arvoa –0,75 vastaa



(100 – 77,34) % ≈ 23 %

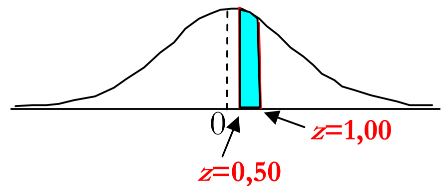
Oppilaista 23 % on siis lyhyempiä kuin 145 cm.

b) Määritetään ensin pituuksia 170 cm ja 180 cm vastaavat normitetut arvot:

$$z_{170} = \frac{170-160}{20} = 0{,}50$$

$$z_{180} = \frac{180-160}{20} = 1{,}00$$

Arvoa 0,50 vastaa prosenttiluku 69,15 %.
Arvoa 1,00 vastaa prosenttiluku 84,13 %.



Näiden pituuksien väliin jää siis

(84,13 – 69,15) % ≈ 15 % opiskelijoista.

Vastaus: a) 23 % b) 15 %

6. **a)** Koska Arthur istuu kiinteällä paikalla, muut asettuvat ”jonoon” kuninkaan paikasta alkaen. Erilaisia jonoja voidaan muodostaa

5! = 120 erilaista

b) Viiden ritarin joukosta voidaan muodostaa erilaisia pareja

$$\binom{5}{2} = 10 \text{ kappaletta}$$

c) Kandidaatteja ovat Lancelot (L) ja ritarit A, B, C ja D. Lancelot pääsee mukaan seuraavissa kokoonpanoissa:

LA, LB, LC ja LD

Koska erilaisia pareja oli 10, niin

$$P(\text{Lancelot pääsee turnajaisiin}) = \frac{4}{10} = 0{,}4$$

Vastaus: a) 120 b) 10 c) 0,4

Harjoituskoe 3

1. Määritetään ensin luokkien luokkakeskukset.

Pituus (cm)	**Luokkakeskus x_i (cm)**	f_i
155–159	$\frac{154,5+159,5}{2}=157$	7
160–164	$\frac{159,5+164,5}{2}=162$	6
165–169	167	6
170–174	172	8
175–179	177	1
180–184	182	2

Keskiarvo

$$\bar{x}=\frac{7\cdot 157+6\cdot 162+...+2\cdot 182}{30}=\frac{4990}{30}=166,333...\approx 166\ \text{(cm)}$$

Moodi on sen luokan luokkakeskus, jonka frekvenssi on suurin. Suurin frekvenssi (8) on luokassa 170–174 cm, joten moodi

$$\text{Mo}=\frac{169,5+174,5}{2}=172\,\text{(cm)}$$

Mediaanin määrittämiseksi muodostetaan summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Pituus (cm)	***f***	***sf***	***sf*%**
155–159	7	7	$\frac{7}{30} = 0{,}233... \approx 23\%$
160–164	6	13	$\frac{13}{30} = 0{,}433... \approx 43\%$
165–169	6	19	0,633…≈ 63 %
170–174	8	27	0,9 = 90 %
175–179	1	28	0,933…≈ 93 %
180–184	2	30	1 = 100 %

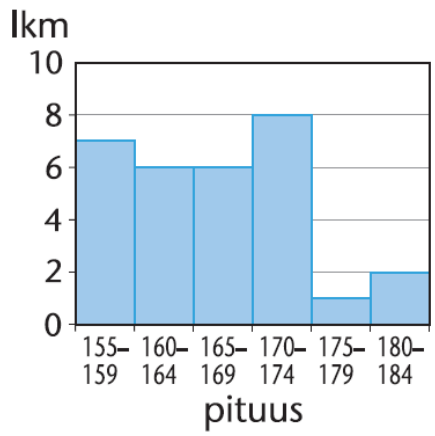
Mediaani on sen luokan luokkakeskus, jonka suhteellinen summafrekvenssi on ensimmäisenä vähintään 50 %. Tällainen luokka on 165–169 cm, joten

$$\mathrm{Md} = \frac{164{,}5 + 169{,}5}{2} = 167\ (\mathrm{cm})$$

Keskihajonta voidaan laskea laskimen tilastotoiminnolla

$$s = \sqrt{\frac{7(157 - 166{,}33...)^2 + ... + 2(182 - 166{,}33...)^2}{30 - 1}}$$
$$= 7{,}396...$$
$$\approx 7{,}40\ (\mathrm{cm})$$

Piirretään frekvenssijakauman kuvaaja, histogrammi. Koska muuttuja (pituus) on jatkuva, pylväät piirretään yhteen.

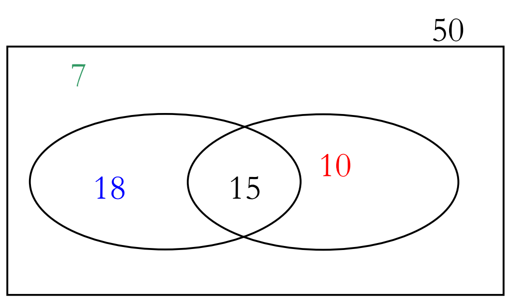


Vastaus: $\overline{x} \approx 166\,\text{cm}$, $s \approx 7{,}40$ cm, Md = 167 cm, Mo = 172 cm

2. Havainnollistetaan tilannetta Venn-diagrammilla.

Sähköpostia luetaan

- puhelimella 33
- puhelimella ja iPadilla 15
- muulla laitteella 7
- pelkästään iPadilla 50 – 7 – 33 = 10



a) $P(\text{pelkästään iPad}) = \dfrac{10}{50} = 0{,}2$

b) $P(\text{puhelin tai iPad}) = \dfrac{43}{50} = 0{,}86$

c) $P(\text{pelkästään puhelin}) = \dfrac{18}{50} = 0{,}36$

Vastaus: a) 0,2 b) 0,86 c) 0,36

3. **a)** $P(\text{tarttuu oikeaan kirjaan}) = \dfrac{1}{5} = 0{,}2$

b) Kurssilla on 26 oppilasta. $P(\text{kaikki tarttuvat oikeaan kirjaan}) = \left(\dfrac{1}{5}\right)^{26} \approx 6{,}7 \cdot 10^{-19}$

c) $P(\text{ässä tai hertta kymppi}) = \dfrac{4+1}{52} = \dfrac{5}{52} \approx 0{,}096$

Vastaus: a) $\dfrac{1}{5}$ b) $6{,}7 \cdot 10^{-19}$ c) $\dfrac{5}{52}$

4. *P*(käyttää laseja) = 0,6
P(ei käytä laseja) = 0,4
Valitaan satunnaisesti 8 henkilöä.

a) $P(\text{joukosta 5 käyttää laseja, 3 ei käytä}) = \dbinom{8}{5} \cdot 0{,}6^5 \cdot 0{,}4^3 = 0{,}2786... \approx 0{,}28$

b) *P*(ainakin kaksi käyttää laseja)
= 1 - *P*(korkeintaan yksi käyttää laseja)
= 1 – [*P*(0 käyttää) + *P*(1 käyttää)]

$$= 1 - \left[0{,}4^8 + \dbinom{8}{1} \cdot 0{,}6 \cdot 0{,}4^7 \right]$$

$= 0{,}9914... \approx 0{,}99$

Vastaus: a) 0,28 b) 0,99

5. Nopeuden jakauma on normaali, $\mu = 77\,\text{km/h}$, $\sigma = 7{,}0\,\text{km/h}$.

Normitetaan nopeudet $80\dfrac{\text{km}}{\text{h}}$ ja $95\dfrac{\text{km}}{\text{h}}$.

$$z_{80} = \frac{80-77}{7} = \frac{3}{7} = 0{,}4285... \approx 0{,}43$$
$$z_{95} = \frac{95-77}{7} = \frac{18}{7} = 2{,}5714... \approx 2{,}57$$

P(nopeus välillä 80 km/h - 95 km/h)

$= \Phi(2{,}57) - \Phi(0{,}43)$

$= 0{,}9949 - 0{,}6664$

$= 0{,}3285$

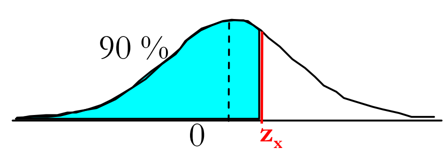
$\approx 33\,\%$

Vastaus: 33 %

6. Tilavuus on normaalisti jakautunut, keskiarvo $\mu = 200$ ja keskihajonta $\sigma = 4$.

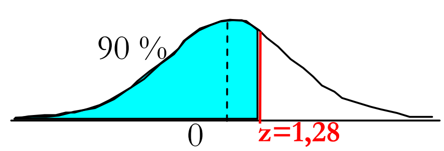
Jos annoksen ylivalumisen todennäköisyys on alle 10 %, kahvimukin on oltava niin suuri, että 90 % annoksista on pienempiä kuin tämä tilavuus. Olkoon tämä tilavuus x (cm^3). Sitä vastaa normitettu arvo

$$z_x = \frac{x-200}{4}$$



Toisaalta tiedetään, että tätä normitettua arvoa vastaava prosenttiluku on 90 %. Taulukosta nähdään, että $\Phi(1{,}28) = 0{,}8997$ (89,97 %). Siis

$$z_x \approx 1{,}28$$



Saadaan siis yhtälö

$$
\begin{aligned}
\frac{x-200}{4} &= 1{,}28 \quad \Big|\cdot 4 \\
x-200 &= 5{,}12 \\
x &= 205{,}12 \\
x &\approx 205
\end{aligned}
$$

Mukin on oltava 205 cm^3.

Vastaus: 205 cm^3

Ekstrat

1. **a)** Otos on valittava mahdollisimman kattavasti. Yksi tapa otoksen keräämiseen on se, että valitaan satunnaisesti muutama koulu koko Suomesta, joista aineisto kerätään satunnaisotannalla.

b) Näyte ei edusta koko Suomen lukiolaisia. Tällainen aineisto saadaan esimerkiksi, kun kysytään tietoja vain oman koulun opiskelijoilta.

c) Kysytään kaikilta lukiolaisilta.

2. **a)** Matkapuhelintehtaan laadunvalvonta on otantatutkimus, koska kaikkia matkapuhelimia on yleensä mahdotonta testata.

b) Ravintolan tekemä asiakastyytyväisyyskysely voi olla joko otanta- tai kokonaistutkimus. Kokonaistutkimus on kyseessä silloin, kun kaikilta ravintolan asiakkailta voidaan saada jollain tavalla vastaus. Tällöin tutkimuksen kohteena olevan joukon täytyy olla suhteellisen pieni (esimerkiksi tiettynä päivänä lounasaikaan käyvät asiakkaat). Yleensä asiakastyytyväisyys-tutkimukset ovat kuitenkin otantatutkimuksia, koska kaikilta asiakkailta on aika vaikeaa saada tietoja.

c) Kansanedustajien mielipidekysely voidaan toteuttaa kokonaistutkimuksena. Tällöin kaikilta kansanedustajilta on saatava vastaukset tutkittaviin asioihin. Varsinkin tietokantoja hyväksikäyttäen on helppo tehdä kansanedustajista kokonaistutkimuksia. Myös otantatutkimus on kansanedustajien mielipidekyselyissä mahdollinen.

d) Eduskuntavaalien äänestystuloksen ennustus on aina otantatutkimus. Ennustuksen tekemiseen ei koskaan käytetä kaikkien äänestäjien tietoja. Kun äänestyksen tulos on tiedossa, kyseessä ei ole enää ennustus.

3. **a)** Kyseessä on **kokonaistutkimus**, mikäli kaikilta aamulennon asiakkailta saadaan vastaus kerättyä. Muuten kyseessä on otantatutkimus.

b) Kyseessä on **otantatutkimus**, koska kaikilta lomamatkalaisilta ei mitenkään saada haluttuja tietoja.

c) Kyseessä on **otantatutkimus**, koska kaikilta lentokentälle saapuvilta matkustajilta ei pystytä haluttua tietoa keräämään.

4. Tehtävässä on useita oikeita vastauksia. olennaista on, että perustelut tietyn keräämistavan valintaan ovat järkeviä ja tilanteeseen sopivia.

a) kysely: tehdään kyselylomake, joka lähetetään otokseen valituille lukiolaisille

haastattelu: haastatellaan itse otokseen valittuja lukiolaisia

tietokanta: käytetään valmiiksi kerättyjä tietoja lukiolaisista. Tämä mahdollista vain, jos tutkijalla on pääsy tarvittavaan tietokantaan.

b) kysely: tehdään kyselylomake, joka lähetetään otokseen valituille lukiolaisille

haastattelu: haastatellaan itse otokseen valittuja lukiolaisia

havainnointi: havainnoidaan esimerkiksi juoma-automaatin läheisyydessä. Tämän keinon käyttäminen riippuu siitä, mikä on tutkimuksen kohteena oleva joukko. Koko Suomen lukiolaisia tutkittaessa havainnointi ei ole hyvä tapa tiedon keräämiseen.

systemaattiset koejärjestelyt: annetaan mahdollisuus maistamalla valita tietyistä juomista. Tämä tapa ei myöskään ole hyvä kaikkien tutkimusten kohdalla.

c) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille lukiolaisille

haastattelu: haastatellaan itse otokseen valittuja lukiolaisia

d) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille. Tämän tavan ongelmana on vastaajien rehellisyys: kaikki eivät välttämättä vastaa kyselylomakkeella rehellisesti.

haastattelu: kysytään itse otokseen valituista. Tämän tavan ongelmana on vastaajien rehellisyys: kaikki eivät välttämättä vastaa kysymyksiin rehellisesti. Haastattelussa on ehkä kuitenkin hankalampi valehdella kuin kyselylomakkeessa.

havainnointi: tarkkaillaan itse ruokalan jonoja ja etuilijoita. Tämä tapa antaa oikeimman mahdollisen kuvan, jos tutkittava populaatio on sopivan pieni tai otokseen kuuluvat tutkimuksen kohteet helposti havainnoitavissa.

5. Tehtävässä on useita oikeita vastauksia. olennaista on, että perustelut tietyn keräämistavan valintaan ovat järkeviä ja tilanteeseen sopivia.

a) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille asiakkaille.

haastattelu: haastatellaan itse otokseen valittuja asiakkaita

b) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille työntekijöille

haastattelu: haastatellaan itse otokseen valittuja työntekijöitä

tietokanta: etsitään halutut tiedot valmiista tietokannasta. Tämä tapa on nopein ja luotettavin, jos vain on pääsy tarvittavaan tietokantaan.

c) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille

haastattelu: haastatellaan itse otokseen valittuja

systemaattiset koejärjestelyt: pyydetään otokseen valittuja maistamaan ja kertomaan saman tien valintansa. Tämä on paras tapa tiedon keräämiseen, jos maistelukokeen järjestäminen vain on mahdollista.

d) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille asiakkaille

haastattelu: haastatellaan itse otokseen valittuja asiakkaita

havainnointi: havainnoidaan itse asiakkaiden kassavalintoja

6. Tehtävässä on useita oikeita vastauksia. olennaista on, että perustelut tietyn keräämistavan valintaan ovat järkeviä ja tilanteeseen sopivia.

a) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille opiskelijoille.

haastattelu: haastatellaan itse otokseen valittuja opiskelijoita

havainnointi: havainnoidaan itse esimerkiksi luokkien ulkopuolella

tietokanta: käytetään hyväksi koulun tietokantaa (kurssipäiväkirjoja) opiskelijoiden myöhästymisistä

b) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille opiskelijoille ja opettajille.

haastattelu: haastatellaan itse otokseen valittuja opiskelijoita ja opettajia

c) kysely: tehdään kyselylomake, joka lähetetään/annetaan otokseen valituille opiskelijoille. Tällaisen kyselylomakkeen laatiminen ja siihen vastaaminen voi kuitenkin olla aika vaikeaa.

haastattelu: haastatellaan itse otokseen valittuja opiskelijoita. Kysymysten laatiminen ja niihin vastaaminen voi kuitenkin olla vaikeaa.

havainnointi: tarkkaillaan itse, miten opiskelija valitsee paikkansa tyhjässä luokassa. Tämän ongelma on kuitenkin se, ettei luokka ole kovinkaan usein tyhjä, kun yksittäinen opiskelija sinne menee.

systemaattiset koejärjestelyt: järjestetään olosuhteet, joissa opiskelija pääsee aina tyhjään luokkaan valitsemaan haluamansa paikan. Jos koejärjestelyt pystytään järjestämään, tämä tapa on luotettavin tiedon keräämisessä.

7. **a)** Jatkuva muuttuja, koska huippunopeus voi saada positiivisia reaalilukuarvoja.

b) Diskreetti muuttuja, koska vaihteiden määrä on positiivinen kokonaisluku.

c) Jatkuva muuttuja, koska bensiinin kulutus voi saada positiivisia reaalilukuarvoja.

d) Diskreetti muuttuja, koska mittarissa näkyvä suurin nopeus on positiivinen kokonaisluku.

8. **a) kvantitatiivisia**: pinta-ala, väkiluku, verojen osuus BKT:sta
kvalitatiivisia: valtiomuoto

b) jatkuvia: pinta-ala, verojen osuus BKT:sta
diskreettejä: väkiluku, valtiomuoto

c) laatueroasteikko: valtiomuoto
suhdelukuasteikko: pinta-ala, väkiluku, verojen osuus BKT:sta

9. **a)** Muuttuja on mitattu **suhdelukuasteikolla**, sillä muuttujalla on absoluuttinen nollakohta. Suhteen laskeminen on myös järkevä laskutoimitus.

b) Muuttuja on mitattu **järjestysasteikolla**, koska arvot 1–4 voidaan laittaa paremmuusjärjestykseen, mutta laskutoimitukset eivät ole mielekkäitä. Ei esimerkiksi voida sanoa, että muuttujan arvo 4 on kaksi kertaa parempi kuin arvo 2.

c) Muuttuja on mitattu **laatueroasteikolla**, koska muuttujan arvot voidaan luokitella, mutta esimerkiksi arvojen paremmuusjärjestykseen laittaminen on mahdotonta.

d) Muuttuja on mitattu **suhdelukuasteikolla**, koska muuttujalla on absoluuttinen nollakohta ja arvojen välisten suhteiden laskeminen on järkevää.

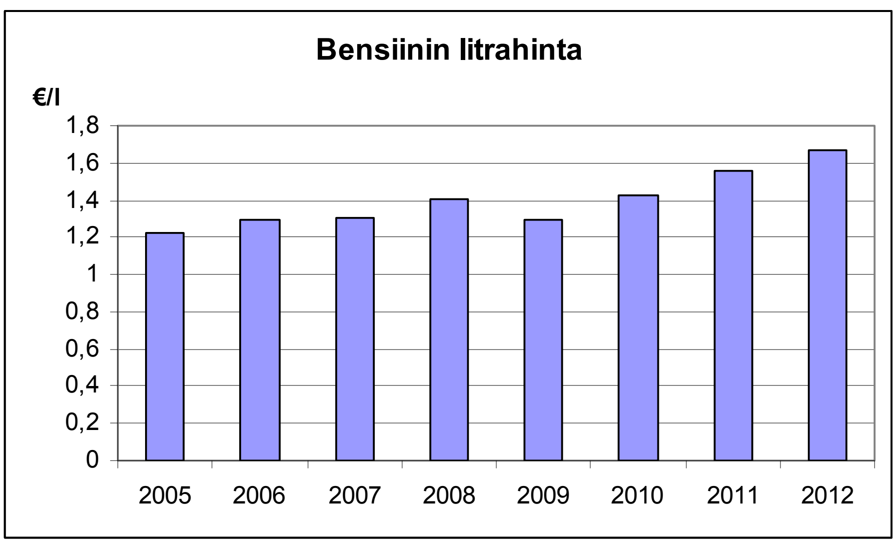
10. **Auton merkki** ja **väri** ovat laatueroasteikollisia muuttujia, koska niiden avulla voidaan vain luokitella.

Kuntoarvio on järjestysasteikollinen muuttuja, koska sen avulla arvot voidaan laittaa järjestykseen, mutta laskutoimitukset eivät ole mielekkäitä.

Rekisteröintivuosi on välimatka-asteikollinen muuttuja, koska vuosiluvun absoluuttinen nollakohta voidaan sopia ajanlaskutavasta riippuen. Monissa yhteyksissä vuosilukua kuitenkin käsitellään suhdelukuasteikollisena muuttujana, jos sekaantumisen vaaraa eri ajanlaskutapojen välillä ei ole.

Ajetut kilometrit, bensatankin tilavuus ja hinta-arvio ovat suhdelukuasteikollisia muuttujia. Kaikilla näillä on absoluuttinen nollakohta.

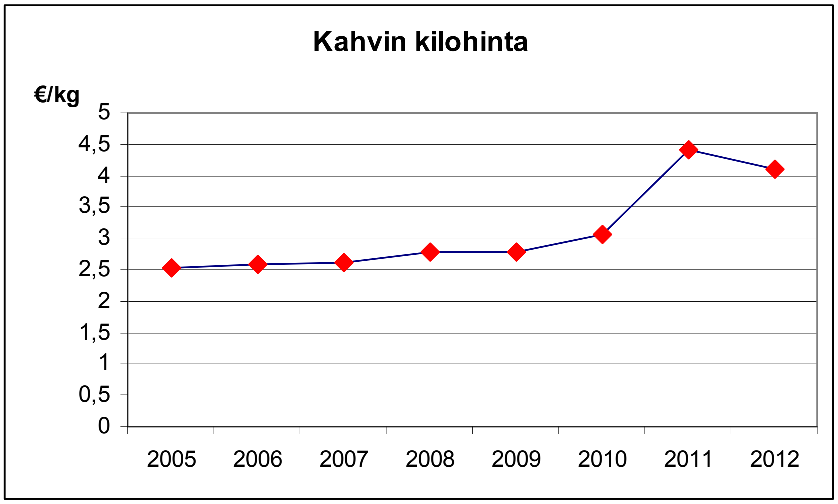
11. Aikasarjan kuvaaja: pylväsdiagrammi. Pylväät ovat aikajärjestyksessä. Pylvään pituus määräytyy muuttujan arvon mukaan.



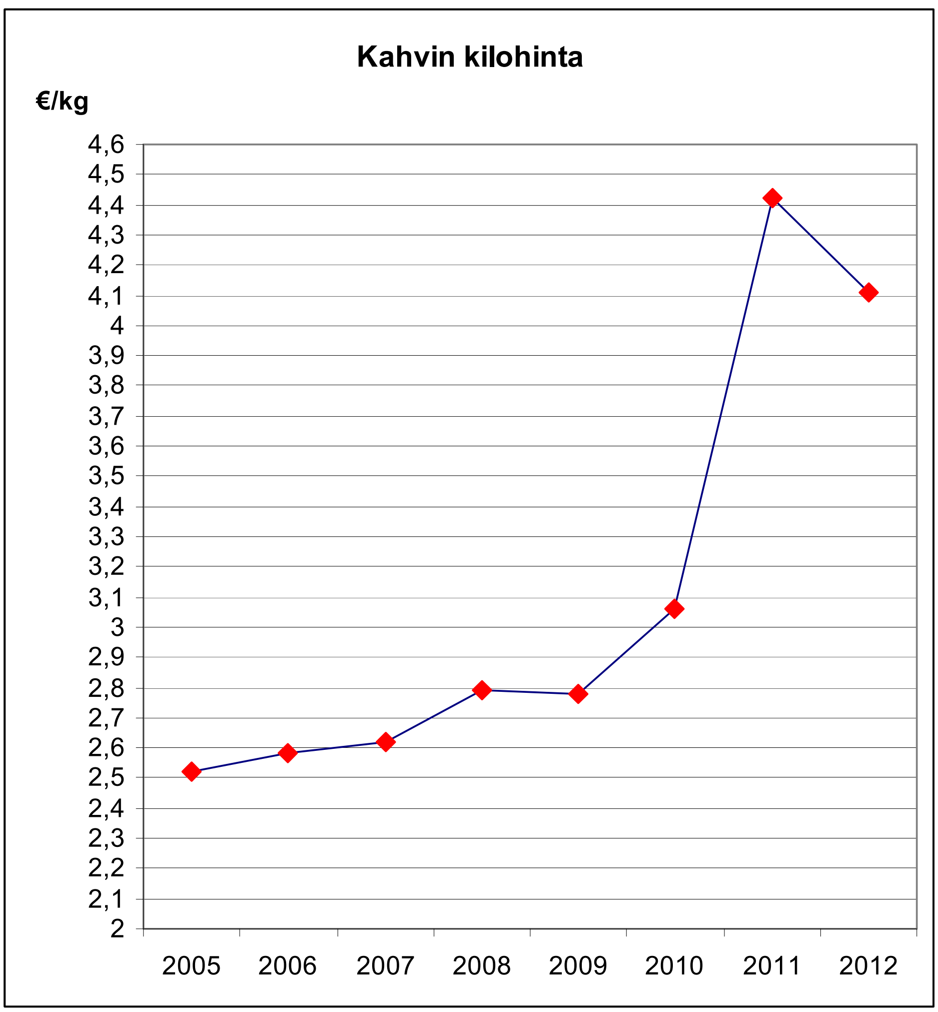
12. –

13. Piirretään trendikäyrät. Muuttujan arvot merkitään pisteinä aikajärjestyksessä. Pisteet yhdistetään viivalla.

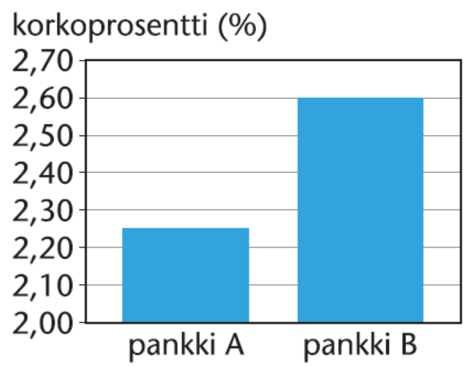
a)



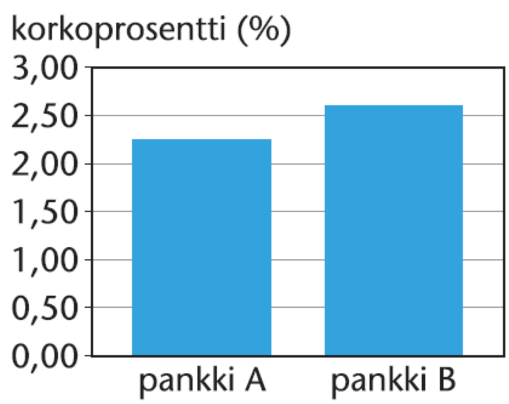
b)



14. **a)** Kun tilanne esitetään alemman koron tarjonneen pankin näkökulmasta, halutaan alemman koron näyttävän paljon edullisemmalta vaihtoehdolta kuin korkeamman koron. Korkojen eroa halutaan siis suurentaa.



b) Kalliimman koron tarjonneen pankin kannalta korkojen välisen eron tulisi näyttää mahdollisimman pieneltä.



15. Tilastoilla valehteluun on useita eri tapoja tilanteesta riippuen. Kirjassa olevien tapojen lisäksi tilastoilla voi valehdella esimerkiksi

- valitsemalla otos (näyte) sopivasti niin, että saadaan haluttuja tuloksia.
- jättämällä asteikoista numeroinnin kokonaan pois, jolloin muutosten todellisen suuruuden tulkitseminen on mahdotonta.
- käyttämällä pylväissä sopivia värejä. Eri värien käyttö esimerkiksi pylväissä saattaa auttaa halutun mielikuvan syntymiseen.
- taustakuvia käyttämällä. Esimerkiksi pylväsdiagrammien taakse lisätyt taustakuvat haittaavat diagrammin luettavuutta.

- jättämällä taustaviivat kokonaan pois. Kun lukemista helpottavat taustaviivat poistetaan, tilaston lukeminen vaikeutuu ja eroja on hankalampi hahmottaa.
- lisäämällä kuvioon ylimääräisiä kuvia. Ylimääräisillä kuvilla voidaan kiinnittää huomio epäolennaisiin asioihin.
- lisäämällä tarpeettomia selityksiä. Valmiilla kuviin lisätyillä selityksillä voidaan lukija saada näkemään tilastoista juuri se, mitä lukijan halutaankin näkevän.

16. Muodostetaan taulukon perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Poissaolo-päivät	***f***	***sf***	***sf*%**
0	8	8	$\frac{8}{56} = 0{,}142... \approx 14\,\%$
1	6	14	0,25 = 25 %
2	5	19	0,339…≈ 34 %
3	7	26	0,464… ≈ 46 %
4	6	32	0,571… ≈ 57 %
5	10	42	0,75 = 75 %
6	6	48	0,857… ≈ 86 %
7	4	52	0,928… ≈ 93 %
9	3	55	0,982… ≈ 98 %
11	1	56	1,00 = 100 %

a) Alakvartiilin kohdalla suhteellinen summafrekvenssi on vähintään 25 %. Alakvartiili on siis **1**.

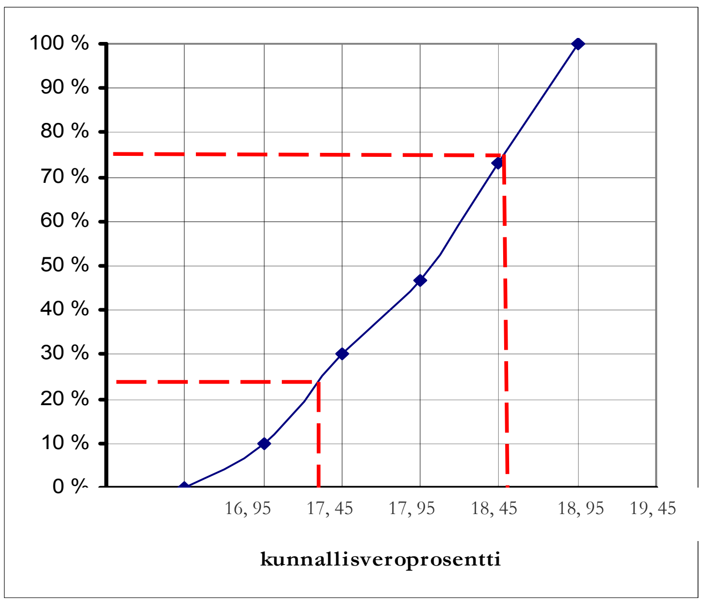
b) Mediaanin kohdalla suhteellinen summafrekvenssi on vähintään 50 %. Mediaani on siis **4**.

c) Yläkvartiilin kohdalla suhteellinen summafrekvenssi on vähintään 75 %. Yläkvartiili on siis **5**.

17. Muodostetaan taulukon perusteella summafrekvenssijakauma ja suhteellinen summafrekvenssijakauma.

Kunnallisvero-prosentti	***f***	***sf***	***sf*%**
17,0–17,4	3	3	10 %
17,5–17,9	6	9	30 %
18,0–18,4	5	14	47 %
18,5–18,9	8	22	73 %
19,0–19,4	8	30	100 %

Piirretään suhteellisen summafrekvenssijakauman kuvaaja.



a) Alakvartiili kuvan perusteella on noin 17,8 ja yläkvartiili noin 19,0.

b) Edellisen kohdan likiarvojen perusteella kvartiiliväli on [17,8; 19,0]. Kvartiilivälin pituus on

19,0 – 17,8 = 1,2

c) Vaihteluväli on[17,0; 19,4], joten vaihteluvälin pituus on

$$19{,}4 - 17{,}0 = 2{,}4$$

Kvartiiliväli on vaihteluvälistä prosentteina

$$\frac{1{,}2}{2{,}4} = 0{,}5 = 50\,\%$$

18. Muodostetaan kuvaajan perusteella tarvittavat frekvenssijakaumat.

Arvosana	***f***	***sf***	***sf*%**
4	2	2	6 %
5	5	7	19 %
6	3	10	28 %
7	8	18	50 %
8	12	30	83 %
9	4	34	94 %
10	2	36	100 %

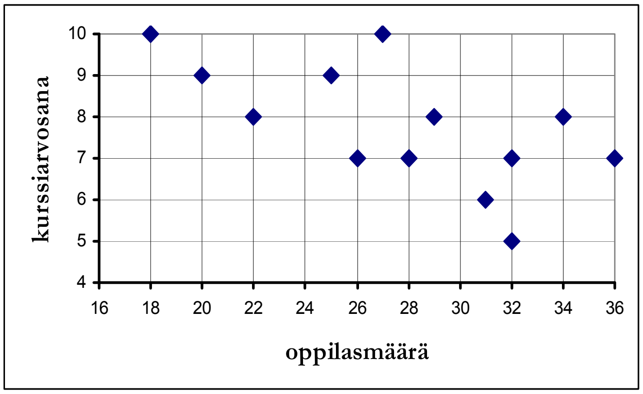
Alakvartiilin kohdalla suhteellinen summafrekvenssi on vähintään 25 %, joten **alakvartiili on 6**.

Yläkvartiilin kohdalla suhteellinen summafrekvenssi on vähintään 75 %, joten **yläkvartiili on 8**.

Tässä aineistossa tämä tarkoittaa, että noin neljännes (25 %) opiskelijoista saa arvosanan, joka on korkeintaan 6 (eli 4, 5 tai 6) ja noin neljännes saa arvosanan joka on vähintään 8 (eli 8, 9 tai 10).

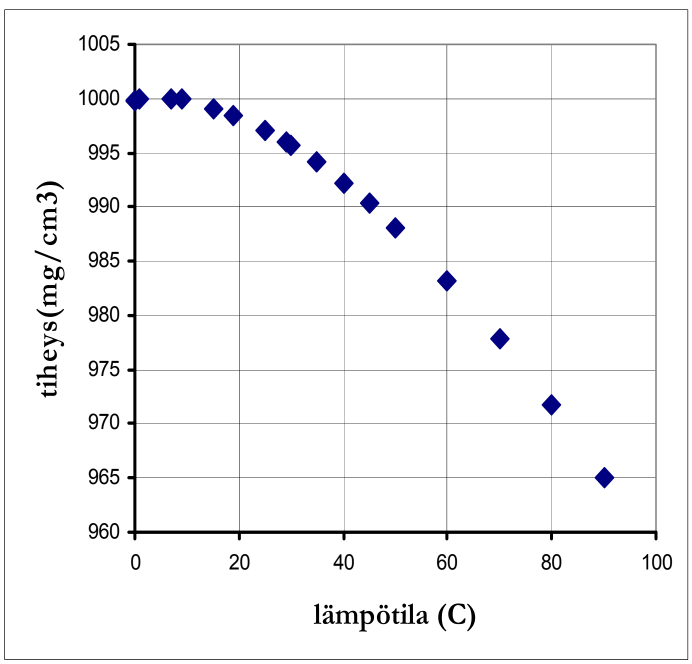
Puolet opiskelijoiden arvosanoista on ala- ja yläkvartiilin välissä eli noin puolet opiskelijoista saa tässä tapauksessa arvosanakseen 7.

19. Piirretään yhteisjakauman kuvaaja. Koska tarkastellaan kurssiarvosanan riippuvuutta oppilasmäärästä, kurssiarvosana on pystyakselilla ja oppilasmäärä vaaka-akselilla.



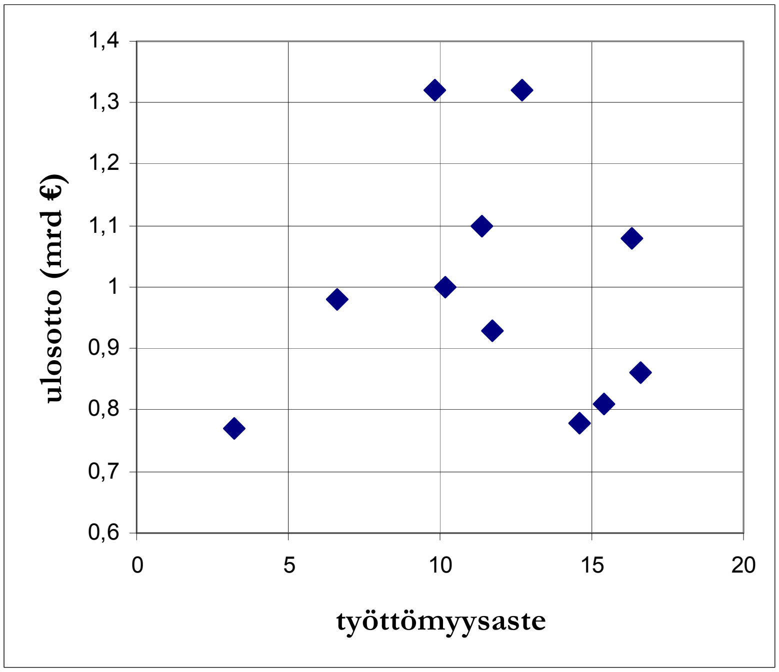
Oppilasmäärän ja kurssiarvosanan välillä on riippuvuutta. Koska pistejoukkoa kuvaa parhaiten suora, riippuvuus on **lineaarista**.

20. Piirretään yhteisjakauman kuvaaja.



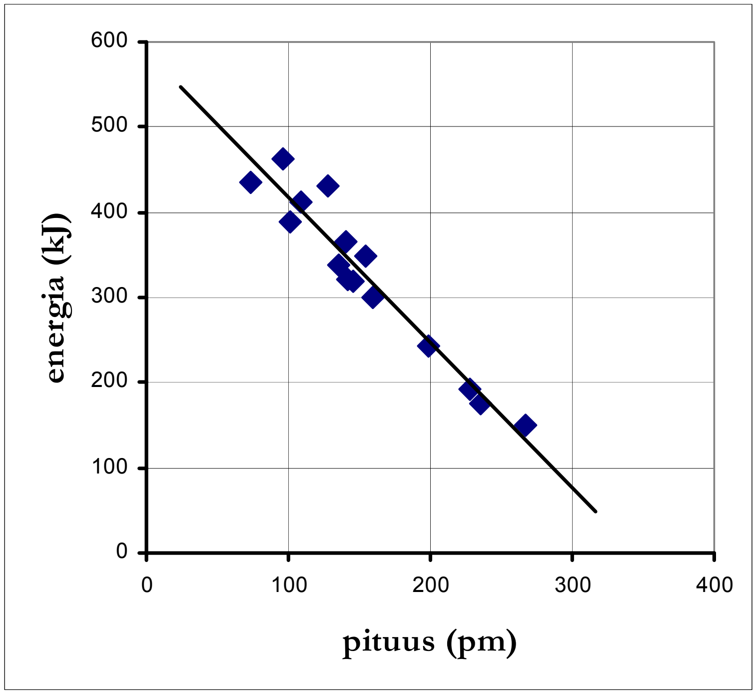
Tilannetta kuvaa parhaiten **polynominen malli**.

21. Piirretään yhteisjakauman kuvaaja.



Pisteet hajaantuvat, eikä pisteparveen voi sovittaa mitään käyrää. Työttömyysasteen ja ulosottoon lähetettyjen verorästien määrän välillä **ei ole riippuvuutta**.

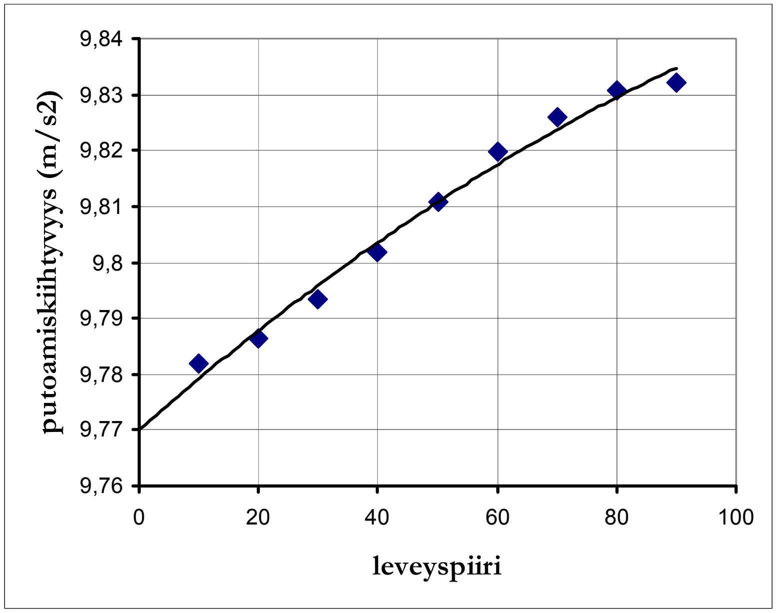
22. Piirretään yhteisjakauman kuvaaja ja sovitetaan siihen regressiokäyrä.



a) Kun sidoksen pituus on 300 pm, sen energia on noin 80 kJ.

b) Kun sidoksen pituus on 50 pm, sen energia on noin 500 kJ.

23. Piirretään yhteisjakauman kuvaaja.



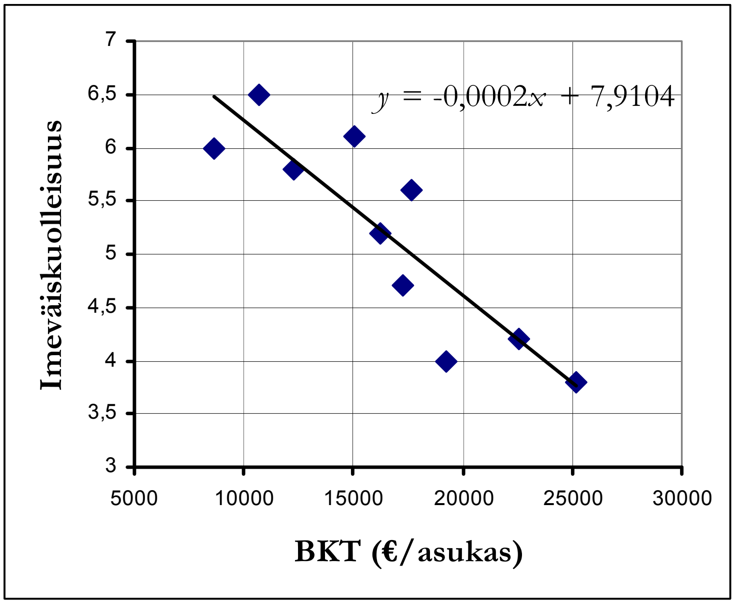
Päiväntasaajalle leveyspiirin asteluku on 0º. Kuvaajan perusteella putoamiskiihtyvyys päiväntasaajalla on 9,770 m/s^2.

24. **a)** Koska tarkoitus on tutkia bruttokansantuotteen vaikutusta, bruttokansantuotetta kuvataan x-akselilla ja imeväiskuolleisuutta y-akselilla.

Syötetään pisteet (x, y) laskimeen tai taulukkolaskentaohjelmaan.
Laskin antaa regressiosuoran yhtälön

$y = A + Bx$ kertoimet

Taulukkolaskentaohjelmassa pistejoukon kuvaajaan voidaan sijoittaa regressiokäyrä, jolloin suoranyhtälö saadaan näkyviin.



Regressiosuoran yhtälö on:

$$y = -0{,}0002x + 7{,}9104$$

b) Korrelaatiokerroin (joko laskimen tilastotoiminnoilla tai taulukkolaskentaohjelmalla) on

$$r \approx 0{,}88$$

c) Lasketaan regressiosuoran yhtälön avulla imeväiskuolleisuus (y), kun BKT (x) on 26 863 € asukasta kohti.

$$\begin{aligned} y &= -0{,}0002 \cdot 26863 + 7{,}9104 \\ &= 2{,}5378 \\ &\approx 2{,}5 \end{aligned}$$

Imeväiskuolleisuus on tällöin noin 2,5 lasta 1000 syntynyttä lasta kohti.

25. **a)** Riippuvuutta on havaittavissa kaikissa muissa kuvissa paitsi kuvassa *B*. Siinä pisteet hajaantuvat eivätkä noudata mitään käyrää.

b) Kuvassa *A* ja *D* riippuvuus on lineaarista. Näissä korrelaatiokertoimen itseisarvo on suurin. Koska kuvan *A* pistejoukkoa kuvaa parhaiten laskeva suora, sen korrelaatiokerroin on negatiivinen.

Kuva *A*: $r \approx -0{,}87$

Kuvassa *B* pistejoukko hajaantuu. Korrelaatiokerroin on tällöin hyvin pieni. Koska kuvassa *C* lineaarinen riippuvuus on vielä pienempää kuin tässä kuvassa, tämän korrelaatiokerroin on annetuista vaihtoehdoista – 0,15.

Kuva *B*: $r \approx -0{,}15$

Kuvassa *C* lineaarinen riippuvuus ei ole kovinkaan voimakasta. Pistejoukkoon sopii parhaiten toisen asteen käyrä. Koska korrelaatiokertoimella mitataan lineaarisen riippuvuuden voimakkuutta, korrelaatiokerroin on tässä tapauksessa hyvin lähellä nollaa.

Kuva *C*: $r \approx 0$

Kuvassa *D* lineaarinen riippuvuus on hyvin voimakasta. Koska pistejoukkoa kuvaa parhaiten nouseva suora, korrelaatiokerroin on positiivinen.

Kuva *D*: $r \approx 0{,}97$

c) Kuva *A*: laskeva suora
Kuva *B*: ei mikään
Kuva *C*: toisen asteen käyrä
Kuva *D*: nouseva suora

26. **a)** Korrelaatiokerroin on positiivinen silloin, kun pistejoukkoa kuvaa nouseva suora. Kuvan *B* pistejoukkoa kuvaa nouseva suora.

b) Korrelaatiokerroin on negatiivinen, kun pistejoukkoa kuvaa laskeva suora. Kuvan *A* pistejoukkoa kuvaa laskeva suora.

c) Regressiokäyrä on kuvassa *B* nouseva suora.

d) Regressiokäyrä on kuvassa *A* laskeva suora.

e) Korrelaatio on negatiivinen kuvassa *A*.

f) Korrelaatiokerroin on noin 0,93 kuvassa *B*.

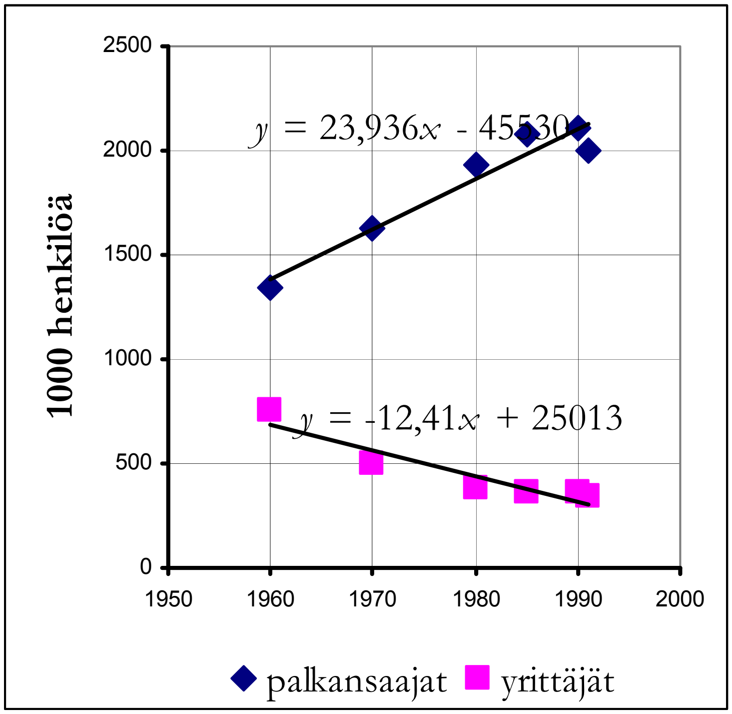
Vastaus: Kuva *A*: b, d, e
Kuva *B*: a, c, f

27. **a)** Pistejoukkoa kuvaa nouseva suora. Korrelaatiokerroin on täten positiivinen. Koska pisteet asettuvat hyvin lähelle suoraa, korrelaatiokertoimen arvo on suuri, joten se kuuluu luokkaan

$0{,}5 < r < 1$

b) Voimakas positiivinen korrelaatio tarkoittaa, että mitä suurempi on ilman lämpötila, sitä enemmän jäätelöä myydään.

28. Havainnollistetaan palkansaajien ja yrittäjien määrien kehitystä samassa diagrammissa.



Lasketaan regressiosuoran yhtälön kertoimet joko taulukkokirjan kaavoilla, laskimen tilastotoiminnoilla tai taulukkolaskentaohjelmalla.

Regressiosuorat ovat
palkansaajat: $y = 23{,}936x - 45\,530$

yrittäjät: $y = -12{,}410x + 25\,013$

Lasketaan regressiosuorien yhtälöiden avulla ennusteet vuoden 1994 palkansaajien ja yrittäjien määrille.

Palkansaajat:

$$\begin{aligned} y &= 23{,}936 \cdot 1994 - 45530 \\ &= 2198{,}384 \\ &\approx 2198 \qquad \text{(tuhatta henkilöä)} \end{aligned}$$

Palkansaajia vuonna 1994 on 2 198 000.

Yrittäjät:

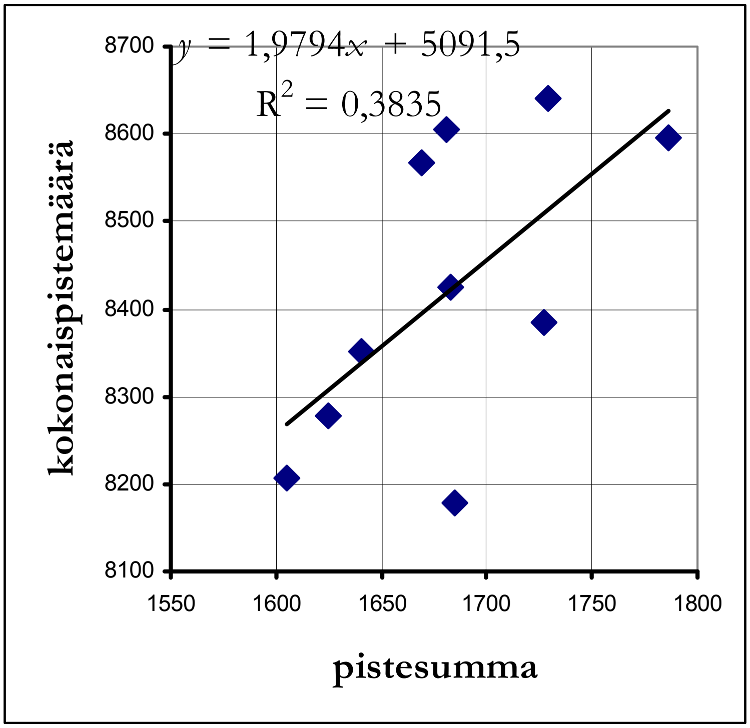
$$\begin{aligned} y &= -12{,}410 \cdot 1994 + 25013 \\ &= 267{,}46 \\ &\approx 267 \qquad \text{(tuhatta henkilöä)} \end{aligned}$$

Yrittäjiä vuonna 1994 on 267 000.

29. Lasketaan kilpailijoiden 100 metrin juoksun ja kuulantyönnön pistemäärät yhteen.

pistesumma ***x***	**kokonaispisteet** ***y***
1729	8641
1681	8606
1786	8595
1669	8567
1683	8425
1727	8385
1641	8351
1625	8277
1605	8206
1685	8178

Merkitään pisteet (x, y) koordinaatistoon. Näin saatu kuvio on korrelaatiodiagrammi.



Lasketaan regressiosuoran kertoimet ja korrelaatiokerroin taulukkokirjan kaavoilla, laskimen tilastotoiminnoilla tai taulukkolaskentaohjelmalla.

Regressiosuora $y = A + Bx$ kertoimet ovat $A \approx 5091,5$ ja $B \approx 1,9794$.

Korrelaatiokerroin $r \approx 0,6193$.

Hämäläisen pistesumma on $858 + 732 = 1590$.

Kokonaispistemäärä on tällöin

$$\begin{aligned} y &= 5091,5 + 1,9794 \cdot 1590 \\ &= 8238,746 \\ &\approx 8239 \quad \text{(pistettä)} \end{aligned}$$