KERTAUS

K1.

a) Muuttuja oli pussin massa.

Vastaus: pussin massa

b) Otos oli 100 liukuhihnalta otettua juureslastupussia.

Vastaus: 100 liukuhihnalta otettua juureslastupussia

c) Perusjoukko oli kaikkien koneen valmistamien juureslastupussien joukko.

Vastaus: kaikkien koneen valmistamien juureslastupussien joukko

d) Koska pusseja otettiin hihnalta säännöllisin välein, otantamenetelmä oli systemaattinen satunnaisotanta.

Vastaus: systemaattinen satunnaisotanta

K2.

a) Maalien määrä on luku, joten maalien määrä on kvantitatiivinen muuttuja.
Maalien määrä saa vain positiivisia kokonaislukuarvoja, joten se on diskreetti muuttuja.
Maalien määrien suhde on mielekästä laskea, joten maalien määrää mitataan suhdeasteikolla.

Vastaus: kvantitatiivinen, diskreetti, suhdeasteikolla

b) Ruokavaliota ei ole mielekästä ilmoittaa lukuina, joten ruokavalio on kvalitatiivinen muuttuja ja sitä mitataan luokitteluasteikolla.
Ruokavaliolle on vain kolme vaihtoehtoa, eikä niiden väleissä ole mitään, joten ruokavalio on diskreetti muuttuja.

Vastaus: kvalitatiivinen, diskreetti, luokitteluasteikolla

c) Maitotölkin tilavuus ilmoitetaan lukuna, joten se on kvantitatiivinen muuttuja.
Tilavuus voi saada minkä tahansa lukuarvon kahden tilavuuden välillä, joten se on jatkuva muuttuja.
Tölkkien tilavuuksien suhde on mielekästä laskea, joten tölkin tilavuutta mitataan suhdeasteikolla.

Vastaus: kvantitatiivinen, jatkuva, suhdeasteikolla

d) Ruoan maulle on vain kolme vaihtoehtoa, eikä niiden väleissä ole mitään, joten ruokavalio on diskreetti muuttuja.
Maut voidaan järjestää paremmuusjärjestykseen, mutta niiden välimatkaa ei ole mielekästä laskea, joten ruoan makua mitataan järjestysasteikolla ja ruoan maku on kvalitatiivinen muuttuja.

Vastaus: kvalitatiivinen, diskreetti, järjestysasteikolla

e) Kellonaika ilmoitetaan kahden luvun avulla, joten se on kvantitatiivinen muuttuja.
Kellonaika ei voi saada mitään arvoja esimerkiksi kellonaikojen 23:58 ja 23:59 välillä, joten se on diskreetti muuttuja.
Kellonaikojen välimatka voidaan laskea, mutta niiden suhdetta ei ole mielekästä laskea, joten kellonaikaa mitataan välimatka-asteikolla.

Vastaus: kvantitatiivinen, diskreetti, välimatka-asteikolla

K3. Jos otantamenetelmäksi valitaan satunnaisotanta, on hankittava ensin lista kaikista Suomen palvelutaloissa asuvista henkilöistä, ja tämän jälkeen arvottava heitä listalta haluttu määrä. Otos on siis periaatteessa yksinkertainen toteuttaa, mutta listaa kaikista palvelutaloissa asuvista henkilöistä on kuitenkin vaikea saada.

Jos otantamenetelmäksi valitaan ryväsotanta, on hankittava lista kaikista Suomen palvelutaloista, arvottava niistä muutama palvelutalo, ja sen jälkeen arvottava näiden palvelutalojen asukkaista haluttu määrä satunnaisotannalla. Jos lista kaikista Suomen palvelutaloista onnistutaan hankkimaan ja arvotuista palvelutaloista saadaan tutkimuslupa, niin tutkimus voidaan toteuttaa.

Tutkimus lienee helpompaa toteuttaa ryväsotannalla kuin satunnaisotannalla.

Vastaus: –

K4. Korttipakassa on neljä maata, jotka ovat kaikki yhtä todennäköisiä, joten yksittäinen arvaus on oikein todennäköisyydellä $p = \frac{1}{4}$. Kun arvausten määrä on $n = 6$, oikeiden vastausten lukumäärän odotusarvo on $\mu = np = 6 \cdot \frac{1}{4} = \frac{3}{2} = 1{,}5$ ja keskihajonta

$$s = \sqrt{np(1-p)} = \sqrt{6 \cdot \frac{1}{4} \cdot \left(1 - \frac{1}{4}\right)} = 1{,}060... \approx 1{,}06.$$

Vastaus: $\mu = 1{,}5$; $s = 1{,}06$

K5. **a)** Arpakuutiossa on kuusi silmälukua, joista tapahtumalle ”silmäluku on 2 tai 3” suotuisia on kaksi. Tapahtuman todennäköisyys saadaan jakamalla suotuisten alkeistapausten määrä kaikkien alkeistapausten määrällä, joten tapahtuman ”silmäluku on 2 tai 3” todennäköisyys on

$$\frac{2}{6}^{(2} = \frac{1}{3}.$$

Kun heitetään arpakuutiota 7 kertaa, binomitodennäköisyyden avulla saadaan

P(saadaan täsmälleen kahdesti silmäluku 2 tai 3)

$$\begin{aligned}&= \binom{7}{2} \cdot \left(\frac{1}{3}\right)^2 \cdot \left(1 - \frac{1}{3}\right)^{7-2} \\ &= \binom{7}{2} \cdot \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^5 \\ &= 0,307... \\ &\approx 0,31.\end{aligned}$$

Vastaus: 0,31

b) Binomitodennäköisyyden avulla saadaan

P(saadaan täsmälleen kerran silmäluku 2 tai 3)

$$\begin{aligned}&= \binom{7}{1} \cdot \left(\frac{1}{3}\right)^1 \cdot \left(1 - \frac{1}{3}\right)^{7-1} \\ &= \binom{7}{1} \cdot \left(\frac{1}{3}\right)^1 \cdot \left(\frac{2}{3}\right)^6 \\ &= 0,204... \\ &\approx 0,20.\end{aligned}$$

Vastaus: 0,20

c) Silmäluku 2 tai 3 saadaan vähintään 3 kertaa, jos se saadaan 3, 4, 5, 6 tai 7 kertaa. Lasketaan kysytty todennäköisyys vastatapahtuman avulla. Tapahtuman ”silmäluku 2 tai 3 saadaan vähintään kolme kertaa” vastatapahtuma on ”silmäluku 2 tai 3 saadaan 0, 1 tai 2 kertaa”.

Kohdissa a ja b on laskettu todennäköisyydet sille, että silmäluku 2 tai 3 saadaan täsmälleen kaksi kertaa tai täsmälleen kerran, joten lasketaan vielä todennäköisyys sille, että silmälukuja 2 ja 3 ei saada kertaakaan.

P(silmälukuja 2 ja 3 ei saada kertaakaan)

$$= \binom{7}{0} \cdot \left(\frac{1}{3}\right)^0 \cdot \left(1 - \frac{1}{3}\right)^{7-0}$$

$$= 1 \cdot 1 \cdot \left(\frac{2}{3}\right)^7$$

$$= 0{,}0585...$$

P(saadaan vähintään kolme kertaa silmäluku 2 tai 3)
= 1 − P(saadaan korkeintaan kaksi kertaa silmäluku 2 tai 3)
= 1 − [P(silmäluku 2 tai 3 saadaan 0, 1 tai 2 kertaa)]
= 1 − (0,307… + 0,204… + 0,0585…)
= 0,429…
≈ 0,43

Vastaus: 0,43

K6. Sähköpostiin saapuva viesti on roskapostia todennäköisyydellä $\frac{1}{4}$.

Oletetaan, että osoitteet, josta viestit tulevat, eivät riipu toisistaan, jolloin tilanne voidaan tulkita toistokokeeksi.

Kahdeksasta viestistä suurin osa on roskapostiviestejä, jos roskapostiviestien määrä on 5, 6, 7 tai 8.
Määritetään kysytty todennäköisyys binomijakauman avulla, kun toistojen määrä on $n = 8$, yksittäisen viestin todennäköisyys olla roskaposti on $p = 0{,}25$ ja roskapostiviestien määrä on $k = 5$, 6, 7 tai 8.

$$\begin{aligned} P(5 \text{ roskapostiviestiä}) &= \binom{8}{5} \cdot 0{,}25^5 \cdot (1-0{,}25)^{8-5} \\ &= \binom{8}{5} \cdot 0{,}25^5 \cdot 0{,}75^3 = 0{,}0230... \end{aligned}$$

$$P(6 \text{ roskapostiviestiä}) = \binom{8}{6} \cdot 0{,}25^6 \cdot 0{,}75^2 = 0{,}00384...$$

$$P(7 \text{ roskapostiviestiä}) = \binom{8}{7} \cdot 0{,}25^7 \cdot 0{,}75^1 = 0{,}00421...$$

$$P(8 \text{ roskapostiviestiä}) = \binom{8}{8} \cdot 0{,}25^8 \cdot 0{,}75^0 = 0{,}0000152...$$

Todennäköisyys sille, että roskapostiviestejä on 5, 6, 7 tai 8, on
0,0230… + 0,00384… + 0,00421… + 0,0000152… = 0,0272… ≈ 0,027.

Todennäköisyys sille, että suurin osa kahdeksasta viestistä on roskapostiviestejä on 0,027 = 2,7 %

Vastaus: 2,7 %

K7. **a)** Todennäköisyys sille, että yksittäisellä heitolla saadaan silmäluku 4, on $p=\frac{1}{6}$ ja heittojen määrä $n = 5$. Määritetään tapahtuman ”silmäluku on 4” tapahtumakertojen määrien X todennäköisyydet binomitodennäköisyyksinä.

$$P(X=0)=\binom{5}{0}\cdot\left(\frac{1}{6}\right)^{0}\cdot\left(1-\frac{1}{6}\right)^{5-0}=0,401...$$

$$P(X=1)=\binom{5}{1}\cdot\left(\frac{1}{6}\right)^{1}\cdot\left(1-\frac{1}{6}\right)^{5-1}=0,401...$$

$$P(X=2)=\binom{5}{2}\cdot\left(\frac{1}{6}\right)^{2}\cdot\left(1-\frac{1}{6}\right)^{5-2}=0,160...$$

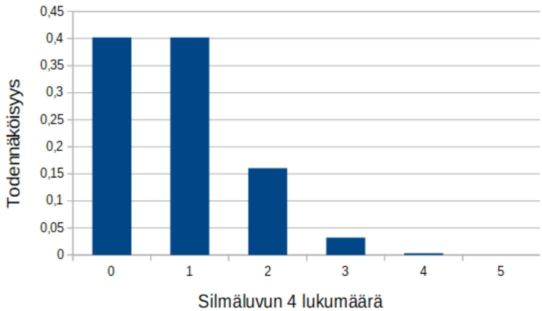
$$P(X=3)=\binom{5}{3}\cdot\left(\frac{1}{6}\right)^{3}\cdot\left(1-\frac{1}{6}\right)^{5-3}=0,0321...$$

$$P(X=4)=\binom{5}{4}\cdot\left(\frac{1}{6}\right)^{4}\cdot\left(1-\frac{1}{6}\right)^{5-4}=0,00321...$$

$$P(X=5)=\binom{5}{5}\cdot\left(\frac{1}{6}\right)^{5}\cdot\left(1-\frac{1}{6}\right)^{5-5}=0,000128...$$

Piirretään pylväskaavio sopivalla ohjelmalla.

Vastaus:



b) Yksittäisellä heitolla tulee klaava todennäköisyydellä $p = \frac{1}{2}$ ja heittojen lukumäärä $n = 5$. Määritetään tapahtuman ”tulee klaava” tapahtumakertojen määrien X todennäköisyydet binomitodennäköisyyksinä.

$$P(X=0) = \binom{5}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(1-\frac{1}{2}\right)^{5-0} = \left(\frac{1}{2}\right)^5 = 0{,}03125$$

$$P(X=1) = \binom{5}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(1-\frac{1}{2}\right)^{5-1} = 5 \cdot \left(\frac{1}{2}\right)^5 = 0{,}15625$$

$$P(X=2) = \binom{5}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(1-\frac{1}{2}\right)^{5-2} = 10 \cdot \left(\frac{1}{2}\right)^5 = 0{,}3125$$

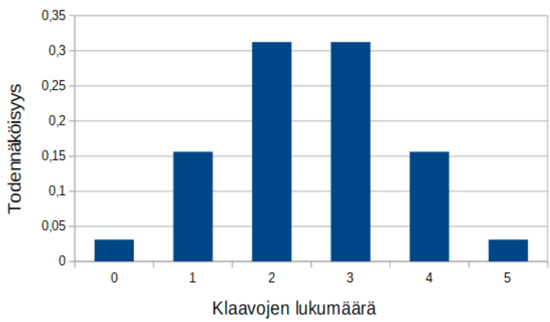
$$P(X=3) = \binom{5}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(1-\frac{1}{2}\right)^{5-3} = 10 \cdot \left(\frac{1}{2}\right)^5 = 0{,}3125$$

$$P(X=4) = \binom{5}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(1-\frac{1}{2}\right)^{5-4} = 5 \cdot \left(\frac{1}{2}\right)^5 = 0{,}15625$$

$$P(X=5) = \binom{5}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(1-\frac{1}{2}\right)^{5-5} = \left(\frac{1}{2}\right)^5 = 0{,}03125$$

Piirretään pylväskaavio sopivalla ohjelmalla.

Vastaus:



c) Silmäluvun 4 lukumäärän odotusarvo on

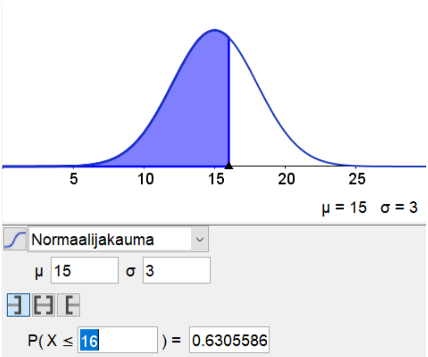
$\mu = np = 5 \cdot \frac{1}{6} = 0,833... \approx 0,83$ ja klaavojen lukumäärän odotusarvo

$\mu = np = 5 \cdot \frac{1}{2} = 2,5.$

Vastaus: a-kohdassa $\mu = 0,83$ ja b-kohdassa $\mu = 2,5$

K8. Satunnaismuuttuja X noudattaa normaalijakaumaa, jonka keskiarvo on 15 ja keskihajonta 3. Määritetään kysytyt todennäköisyydet sopivalla ohjelmalla.

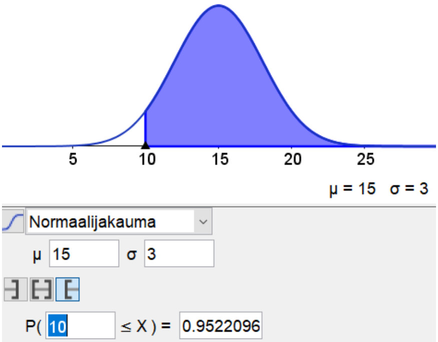
a)



Kysytty todennäköisyys on 0,630… ≈ 0,63.

Vastaus: 0,63

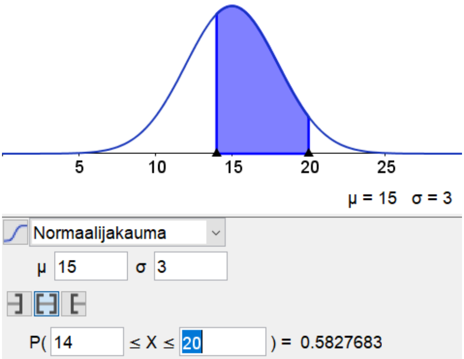
b)



Kysytty todennäköisyys on 0,952… ≈ 0,95.

Vastaus: 0,95

c)

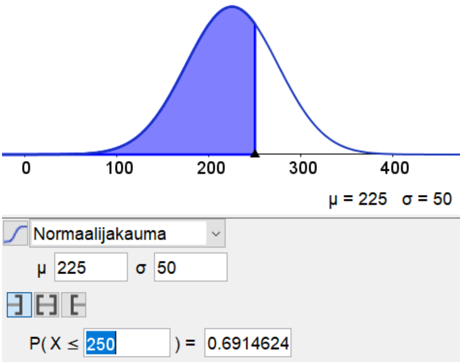


Kysytty todennäköisyys on 0,582… ≈ 0,58.

Vastaus: 0,58

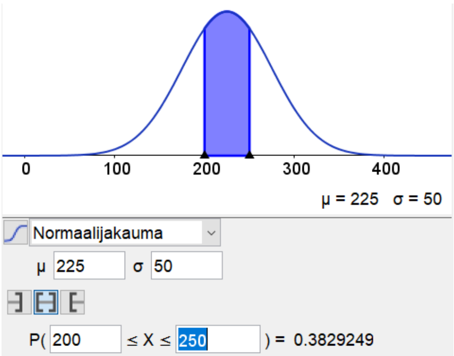
K9. Partahylkeen massa noudattaa normaalijakaumaa, jonka keskiarvo on 225 kg ja keskihajonta 50 kg. Tietyn painorajan ylittävien tai alittavien partahylkeiden osuus on yhtä suuri kuin todennäköisyys, jolla yksittäisen hylkeen massa ylittää tai alittaa kyseisen rajan. Määritetään kysytyt osuudet todennäköisyyksinä sopivan ohjelman avulla.

a)



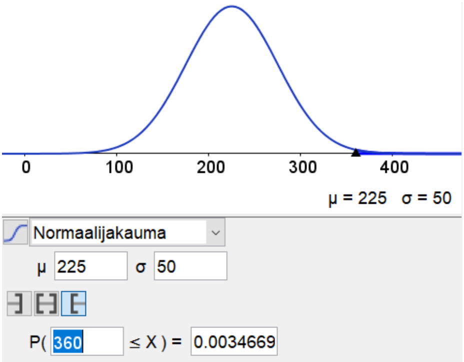
Vastaus: 69 %

b)



Vastaus: 38 %

c)



Vastaus: 0,35 %

K10. **a)** Banaanien massojen keskiarvo on $\bar{x} = 120$ g ja keskihajonta s = 15 g, joten banaanin massaa x = 140 g vastaava normitettu arvo on

$$z = \frac{x - \bar{x}}{s} = \frac{140 - 120}{15} = 1{,}333... \approx 1{,}33.$$

Vastaus: z = 1,33

b) Koska banaanin massaa 140 g vastaava normitettu arvo on 1,33, niin kysytty todennäköisyys on normaalijakauman kertymäfunktion arvo kohdassa 1,33. Taulukosta nähdään, että kertymäfunktion arvo on 0,9082, joten kysytty todennäköisyys on 0,9082 ≈ 0,91.

Vastaus: 0,91

K11. Satunnaismuuttuja Z noudattaa normitettua normaalijakaumaa eli normaalijakaumaa, jonka keskiarvo on 0 ja keskihajonta 1. Määritetään kysytyt todennäköisyydet taulukon avulla.

a) Taulukosta nähdään, että $P(Z < 1) = 0{,}8413 \approx 0{,}84$.

Vastaus: 0,84

b) Taulukossa ei ole negatiivisia muuttujan arvoja vastaavia kertymäfunktion arvoja, joten käytetään symmetriaa.

$P(Z > -2) = P(Z \leq 2)$

Taulukosta nähdään, että $P(Z \leq 2) = 0{,}9772$, joten kysytty todennäköisyys on $0{,}9772 \approx 0{,}98$.

Vastaus: 0,98

c) Kysytty todennäköisyys on

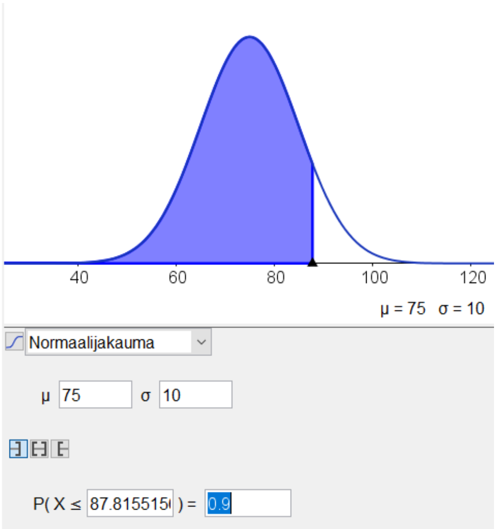
$$
\begin{aligned}
&P(-2 < Z < 1) \\
&= P(Z < 1) - P(Z < -2) \\
&= P(Z < 1) - [1 - P(Z > -2)] \\
&= P(Z < 1) - 1 + P(Z > -2).
\end{aligned}
$$

Kohtien a ja b perusteella $P(Z < 1) = 0{,}8413$ ja $P(Z > -2) = 0{,}9772$, joten

$$P(Z < 1) - 1 + P(Z > -2) = 0{,}8413 - 1 + 0{,}9772 = 0{,}8185 \approx 0{,}82$$

Vastaus: 0,82

K12. Heinän pituuden keskiarvo on 75 cm ja keskihajonta 10 cm. Etsitään sopivan ohjelman avulla pituus, jonka alapuolella on 90 % heinistä.



90 % heinistä on alle 87,815… cm ≈ 88 cm pitkiä.

Vastaus: 88 cm

K13. Partavesipullojen tilavuuksien keskihajonta on 3,8 ml, ja 95 %:n pulloista on oltava tilavuudeltaan yli 50,0 ml. Tällöin tilavuudeltaan korkeintaan 50,0 ml on 100 % − 95 % = 5 % pulloista. Merkitään keskiarvoa kirjaimella x ja ratkaistaan se sopivan ohjelman avulla.

Normaalijakauma(x,3.8,50)=0.05

RatkaiseNumeerisesti: $\{x = 56.25044378241\}$

Tilavuuden keskiarvoksi on valittava 56,250… ml ≈ 56,3 ml.

Vastaus: $\bar{x} = 56{,}3$ ml

K14. Luottamusvälin päätepisteet saadaan, kun otoksessa saadusta suhteellisesta osuudesta vähennetään virhemarginaali ja siihen lisätään virhemarginaali.

Luottamusvälin päätepisteet ovat siis 32,6 % − 3,6 % = 29,0 % ja 32,6 % + 3,6 % = 36,2 %.

Vastaus: [29,0 %; 36,2 %]

K15. Määritetään suhteellisen osuuden luottamusväli, kun suhteellinen osuus otoksessa on p = 0,93, otoskoko n = 303 ja 99,9 %:n luottamustason kriittinen arvo 3,29.

$$\left[0{,}93-3{,}29\cdot\sqrt{\frac{0{,}93(1-0{,}93)}{303}};0{,}93+3{,}29\cdot\sqrt{\frac{0{,}93(1-0{,}93)}{303}}\right]$$
$$=[0{,}88177...;0{,}97822...]$$

Suhteellisen osuuden luottamusvälin alaraja on 88,177… % ja yläraja 97,822… %

Luottamusvälin päätepisteet on aina pyöristettävä poispäin keskiarvosta, joten luottamusväli on [88,1 %; 97,9 %].

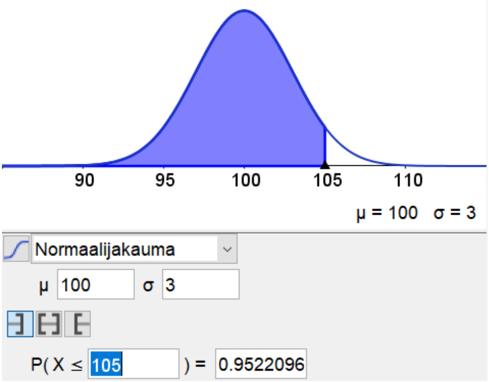
Vastaus: [88,1 %; 97,9 %]

K16. **a)** Lasketaan keskiarvon keskivirhe 25 henkilön otoksessa.

$$s_{\bar{x}} = \frac{15}{\sqrt{25}} = 3$$

Vastaus: $s_{\bar{x}} = 3$

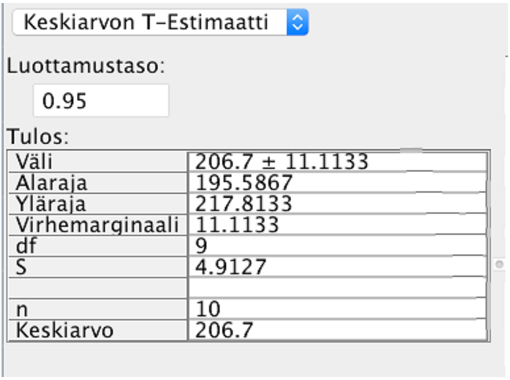
b) Kohdan a perusteella 25 henkilön otoksen keskiarvo noudattaa normaalijakaumaa, jonka keskiarvo on 100 ja keskihajonta 3. Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



25 henkilön otoksessa älykkyysosamäärien keskiarvo on alle 105 todennäköisyydellä 0,952… ≈ 0,95.

Vastaus: 0,95

K17. Kopioidaan taulukko sopivaan ohjelmaan. Lasketaan ohjelman avulla appelsiinien massojen keskiarvon 95 %:n luottamusväli.



Väli	206.7 ± 11.1133
Alaraja	195.5867
Yläraja	217.8133
Virhemarginaali	11.1133
df	9
S	4.9127
n	10
Keskiarvo	206.7

95 %:n luottamusväli on [195,586… g; 217,813… g].

Luottamusvälin päätepisteet on aina pyöristettävä poispäin keskiarvosta, joten luottamusväli on [195 g, 218 g].

Vastaus: [195 g, 218 g]

K18. Ruuvien pituuden keskiarvo otoksessa on luottamusvälin ala- ja ylärajan keskiarvo, joten $\bar{x} = \dfrac{12{,}9\text{ mm} + 13{,}1\text{ mm}}{2} = 13{,}0\text{ mm}.$

Luottamusväli on $\left[\bar{x} - \text{kriittinen arvo} \cdot \dfrac{s}{\sqrt{n}}, \bar{x} + \text{kriittinen arvo} \cdot \dfrac{s}{\sqrt{n}}\right].$

Muodostetaan yhtälö sijoittamalla välin ylärajan lausekkeeseen keskiarvo $\bar{x} = 13{,}0$ mm, 99 %:n kriittinen arvo 2,58 ja otoskoko n = 100.

Ratkaistaan saadusta yhtälöstä keskihajonta s.

$$
\begin{aligned}
13 + 2{,}58 \cdot \frac{s}{\sqrt{100}} &= 13{,}1 \\
13 + 2{,}58 \cdot \frac{s}{10} &= 13{,}1 \\
2{,}58 \cdot \frac{s}{10} &= 0{,}1 && \| \cdot 10 \\
2{,}58s &= 1 && \| :2{,}58 \\
s &= 0{,}387...
\end{aligned}
$$

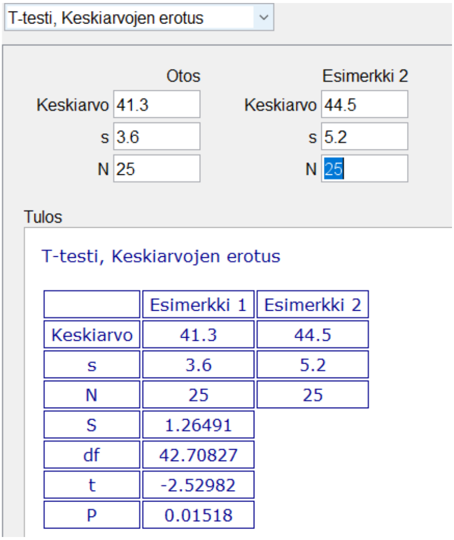
Otoksen keskihajonta oli s = 0,387… mm ≈ 0,39 mm.

Vastaus: s = 0,39 mm

K19. Nollahypoteesi hylätään, jos p-arvo on riskitasoa suurempi. Kun p-arvo on 0,0078 = 0,78 %, nollahypoteesi hylätään 1 %:n riskitasolla, mutta ei 0,1 %:n riskitasolla.

Vastaus: hylätään 1 %:n riskitasolla, ei hylätä 0,1 %:n riskitasolla

K20. Otosten keskiarvot ovat $\overline{x}_1 = 41,3$ ja $\overline{x}_2 = 44,5$, keskihajonnat $s_1 = 3,6$ ja $s_2 = 5,2$ ja otoskoot $n_1 = n_2 = 25$. Määritetään keskiarvojen eron t-testin p-arvo sopivalla ohjelmalla.

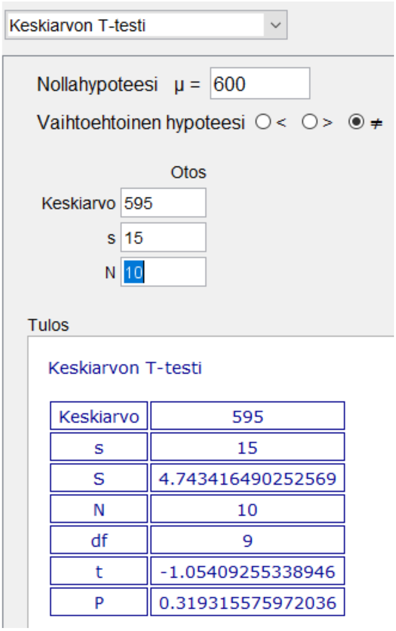


	Esimerkki 1	Esimerkki 2
Keskiarvo	41.3	44.5
s	3.6	5.2
N	25	25
S	1.26491	
df	42.70827	
t	-2.52982	
P	0.01518	

Ohjelman antama p-arvo on 0,01518. p-arvo on pienempi kuin valittu riskitaso 5 % = 0,05, joten perusjoukkojen A ja B keskiarvoilla voidaan sanoa olevan eroa.

Vastaus: Voidaan sanoa.

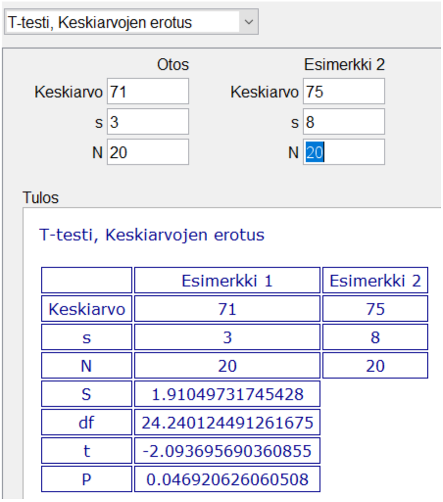
K21. Tehdään nollahypoteesi H_0: ”Ilmoitettu keskiarvo pitää paikkansa”. Tehdään keskiarvon *t*-testi sopivalla ohjelmalla.



Ohjelma antaa *p*-arvoksi 0,31931…, joten nollahypoteesin hylkäämisessä otettaisiin 31,931… %:n riski tehdä väärä johtopäätös. Riski on niin suuri, ettei voida sanoa ilmoitetun keskiarvon olevan väärä.

Vastaus: Ei voida.

K22. Otosten keskiarvot ovat prosentteina $\overline{x}_1 = 71$ ja $\overline{x}_2 = 75$ keskihajonnat $s_1 = 3$ ja $s_2 = 8$ ja otoskoot $n_1 = n_2 = 20$. Määritetään keskiarvojen eron t-testin p-arvo sopivalla ohjelmalla.



	Esimerkki 1	Esimerkki 2
Keskiarvo	71	75
s	3	8
N	20	20
S	1.91049731745428	
df	24.240124491261675	
t	-2.093695690360855	
P	0.046920626060508	

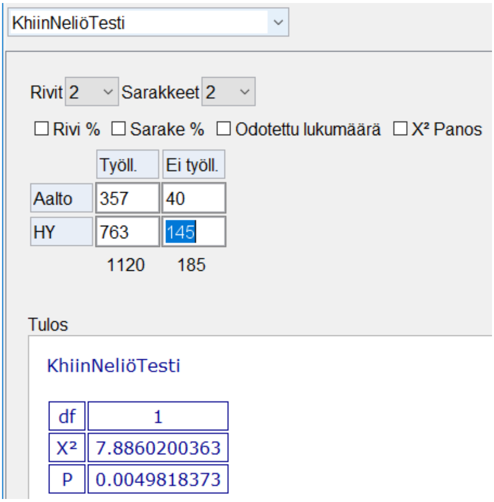
Ohjelman antama p-arvo on 0,04692…= 4,692… %, joten jos sanotaan rautamalmin laaduissa olevan eroa, otetaan 4,692… %:n riski. Riski on suurempi kuin 1 %, joten ei voida 1 %:n riskitasolla sanoa, että kaivosten rautamalmin laaduissa on eroa.

Vastaus: Ei voida.

K23. Tehdään nollahypoteesi H_0: ”Aalto-yliopistosta valmistuneista työllistyy yhtä suuri osuus kuin Helsingin yliopistosta.”

Aalto-yliopiston 397 vastaajista työllistyi 90 % eli
$0{,}9 \cdot 397 = 357{,}3 \approx 357$. Loput $397 - 357 = 40$ ei työllistynyt.
Helsingin yliopiston vastaavat luvut olivat $0{,}84 \cdot 908 = 762{,}72 \approx 763$ ja $908 - 763 = 145$.

Tehdään χ^2 -riippumattomuustesti sopivalla ohjelmalla.



Ohjelman antama p-arvo on 0,00498…, joten nollahypoteesin hylkäämisessä otetaan 0,498… %:n riski tehdä väärä johtopäätös. Riski on niin pieni, että nollahypoteesi voidaan hylätä. Voidaan siis sanoa, että Aalto-yliopistosta valmistuneet työllistyvät Helsingin yliopistosta valmistuneita paremmin.

Vastaus: Voidaan sanoa.

KOKOAVIA TEHTÄVIÄ

ILMAN TEKNISIÄ APUVÄLINEITÄ

1. **a)** Lasketaan kesäfestivaalin A kävijämäärää 118 000 vastaava normitettu arvo, kun keskiarvo on $\overline{x} = 86\,000$ ja keskihajonta $s = 16\,000$.

$$\frac{118\,000 - 86\,000}{16\,000} = \frac{32\,000}{16\,000} = 2$$

Festivaalin A kävijämäärää vastaava normitettu arvo oli 2.

Vastaus: 2

b) Lasketaan festivaalin B kävijämäärää 74 000 vastaava normitettu arvo, kun keskiarvo on $\overline{x} = 53\,000$ ja keskihajonta $s = 7000$.

$$\frac{74\,000 - 53\,000}{7000} = \frac{21\,000}{7000} = 3$$

Festivaalin B vastaava normitettu arvo oli 3, joka oli suurempi kuin festivaalin A vastaava normitettu arvo. Kävijöitä oli siis aiempaan suhteutettuna enemmän festivaalilla B.

Vastaus: festivaalilla B

2. Käyrän alle jäävä tummennetun alueen pinta-ala kuvaa todennäköisyyttä.

a) Kuvan perustella todennäköisyys sille, että satunnaisesti valittu 17-vuotias japanilaismies on alle 170 cm pitkä, on 0,44.

Vastaus: 0,44

b) Komplementtisäännön ja a-kohdan perusteella todennäköisyys sille, että satunnaisesti valittu 17-vuotias mies on yli 170 cm pitkä, on $1 - 0{,}44 = 0{,}56$.

Vastaus: 0,56

c) Kuvien perusteella todennäköisyys sille, että satunnaisesti valittu japanilaismies on 170-180 cm pitkä, on

$$\begin{aligned}&P(170-180 \text{ cm pitkä})\\&= P(180 \text{ cm pitkä}) - P(170 \text{ cm pitkä})\\&= 0{,}95 - 0{,}44\\&= 0{,}51.\end{aligned}$$

Vastaus: 0,51

3. Inkeri joutuu pysähtymään yksittäisiin liikennevaloihin
100 % −70 % = 30 %:n todennäköisyydellä ja 70 %:n todennäköisyydellä ei joudu pysähtymään niihin.

Binomitodennäköisyytenä voidaan siis laskea

$$P(\text{Inkeri pysähtyy täsmälleen kahdesti}) = \binom{5}{2} \cdot 0{,}3^2 \cdot 0{,}7^3 \text{ ja}$$

$$P(\text{Inkeri pysähtyy täsmälleen kolmesti}) = \binom{5}{3} \cdot 0{,}3^2 \cdot 0{,}7^3.$$

Tapahtuma A liittyy siis todennäköisyyteen IV ja tapahtuma B todennäköisyyteen II.

Koska Saku joutuu pysähtymään punaisiin valoihin keskimäärin 0,7 kertaa, niin Poisson-jakauman avulla saadaan pysähtymisten määrien todennäköisyyksiksi

$$P(\text{Saku pysähtyy täsmälleen kahdesti}) = \frac{0{,}7^2}{2!} e^{-0{,}7} \text{ ja}$$

$$P(\text{Saku pysähtyy täsmälleen kolmesti}) = \frac{0{,}7^3}{3!} e^{-0{,}7}.$$

Tapahtuma C liittyy siis todennäköisyyteen III ja tapahtuma D todennäköisyyteen I.

Vastaus: A: IV, B: II, C: III ja D: I

4. **a)** Koska $\frac{14+18}{2}=16,$ niin otoksessa massojen keskiarvo oli 16 kg.
Väite on siis tosi.

Vastaus: tosi

b) Perusjoukon massojen keskiarvon 99 %:n luottamusväli on [14 kg, 18kg], mutta massojen keskiarvo voi 1 %:n todennäköisyydellä olla tämän välin ulkopuolella, siis myös yli 18 kg. Väite on siis tosi.

Vastaus: tosi

c) Massojen keskiarvon 99 %:n luottamusväli on [14 kg, 18 kg], eli massojen keskiarvo on 99 %:n varmuudella vähintään 14 kg mutta alle 18 kg. 99 %:n varmuutta siitä, että massojen keskiarvo olisi 16 kg, ei kuitenkaan ole, joten väite on epätosi.

Vastaus: epätosi

d) Massojen keskiarvon 99 %:n luottamusväli on [14 kg, 18 kg], eli massojen keskiarvo on 1 %:n todennäköisyydellä tämän välin ulkopuolella. Massat noudattavat normaalijakaumaa, joka on symmetrinen keskiarvon suhteen, joten massojen keskiarvo on 0,5 %:n todennäköisyydellä alle 14 kg ja 0,5 %:n todennäköisyydellä yli 18 kg. Väite on siis tosi.

Vastaus: tosi

e) Massojen keskiarvon 99 %:n luottamusväli on [14 kg, 18 kg]. Jos halutaan tietää tätä suuremmalla varmuudella väli, jolla keskiarvo on, on väliä pidennettävä. Väli [15 kg, 17 kg] on kuitenkin lyhempi kuin väli [14 kg, 18 kg], joten se ei voi olla massojen 99,9 %:n luottamusväli. Väite on siis epätosi.

Vastaus: epätosi

f) Otoksessa massojen keskiarvo oli a-kohdan perusteella 16 kg. Tilastollisessa testauksessa testataan, millä todennäköisyydellä saadaan otoksessa saatu tulos, jos nollahypoteesi on voimassa. Jos otoksessa saatu keskiarvo on sama kuin nollahypoteesin mukainen keskiarvo, ei testausta tarvita, koska poikkeamaa ei ole. Otoksen perusteella ei ole syytä hylätä väitettä millään riskitasolla.

Vastaus: epätosi

5. Lasketaan keskiarvon keskivirhe, kun otoskoko on $n = 100$. Koska otoskoko on suuri, perusjoukon keskihajontana voidaan käyttää otoksen keskihajontaa $s = 50$.

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{50}{\sqrt{100}} = \frac{50}{10} = 5$$

Jos sadan juuston erän massa on yli 51 kg, niin juuston keskimääräinen massa on erässä yli $\frac{51\,\text{kg}}{100} = 0{,}51\text{ kg} = 510\text{ g}.$ Lasketaan tätä vastaava normitettu arvo.

$$z = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{510 - 500}{5} = \frac{10}{5} = 2$$

Normaalijakauman kertymäfunktion arvojen taulukosta luetaan, että $P(Z < 2) = 0{,}9772$.
Tällöin $P(Z > 2) = 1 - P(Z < 2) = 1 - 0{,}9772 = 0{,}0228$.
Sadan juuston erän massa on yli 51 kg todennäköisyydellä $0{,}0228 \approx 0{,}023$.

Vastaus: 0,023

6. Todennäköisyysjakaumia ovat esimerkiksi binomijakauma, Poisson-jakauma ja normaalijakauma.

Binomijakaumaa käytetään, kun halutaan tietää millä todennäköisyydellä jokin tapahtuma tapahtuu toistokokeessa tietyn määrän kertoja, ja tapahtuman todennäköisyys yksittäisessä toistossa tiedetään.

Poisson-jakaumaa käytetään, kun ei tiedetä tapahtuman todennäköisyyttä yksittäisessä toistossa eikä välttämättä edes toistojen määrää, mutta tiedetään tapahtuman tapahtumakertojen määrän odotusarvo.

Normaalijakaumaa käytetään, kun tiedetään suureen noudattavan normaalijakaumaa ja halutaan laskea, millä todennäköisyydellä suureen arvo on jollain tietyllä välillä.

Vastaus: –

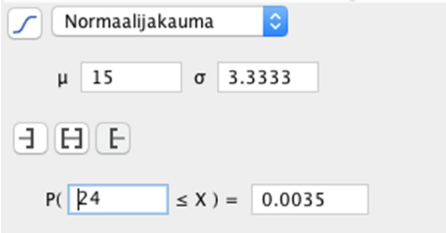
7. χ^2 -yhteensopivuustestillä testataan, noudattaako aineisto jotakin valittua jakaumaa, kuten tasaista jakaumaa. Voidaan esimerkiksi testata, ovatko kolme puhelinmallia yhtä suosittuja, jos otoksen perusteella saadaan tulos, että puhelinmalli A on 15 henkilöllä, puhelinmalli B 19 henkilöllä ja puhelinmalli C 21 henkilöllä. Tällöin havaintoyksiköitä ovat henkilöiden puhelimet ja tilastomuuttujana puhelinmalli.

Vastaus: Testillä testataan, noudattaako aineisto jotakin valittua jakaumaa, kuten tasaista jakaumaa.

APUVÄLINEET SALLITTU

8. Keskihajonta on $s = 3\text{ h } 20\text{ min} = 3\frac{1}{3}\text{ h}$ ja keskiarvo on $\overline{x} = 15\text{ h}$.

Määritetään sopivalla ohjelmalla todennäköisyys, että kuvat ovat valmiita vasta 24 tunnin jälkeen.



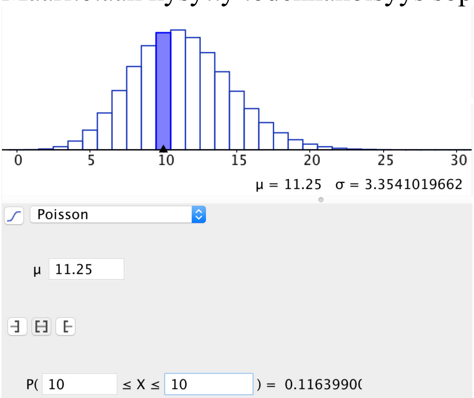
Tilauksista noin 0,35 % joudutaan antamaan ilmaiseksi.

Vastaus: 0,35 %

9. Sovelletaan Poisson-jakaumaa.

a) Tunnissa tulee keskimäärin 45 asiakasta, joten 15 minuutissa tulee keskimäärin $\frac{45}{4} = 11{,}25$ asiakasta.

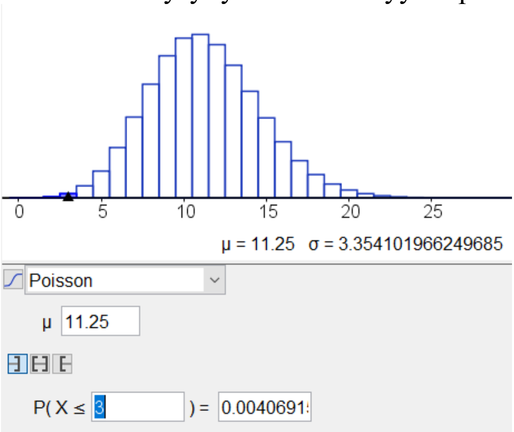
Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



Todennäköisyys sille, että 15 minuutissa tulee täsmälleen 10 asiakasta, on $0{,}116\ldots \approx 0{,}12$.

Vastaus: 0,12

b) Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



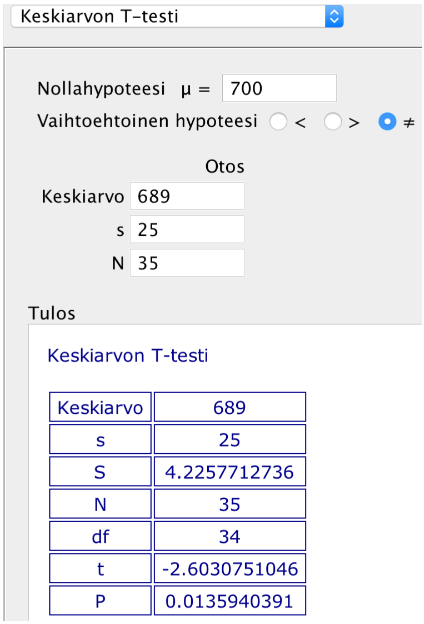
Todennäköisyys sille, että 15 minuutissa tulee korkeintaan 3 asiakasta, on 0,00406… ≈ 0,0041.

Vastaus: 0,0041

10. Tehdään *t*-testi sopivalla ohjelmalla.

Ohjelma antaa *p*-arvoksi 0,01359… = 1,359… % ≈ 1,4 %.
Otetaan siis 1,4 %:n riski, jos sanotaan keskiarvon poikkeavan 700 ml:sta.
Keskitilavuuden voidaan sanoa poikkeavan 700 ml:sta 1,4 %:n riskitasolla.

Vastaus: 1,4 %

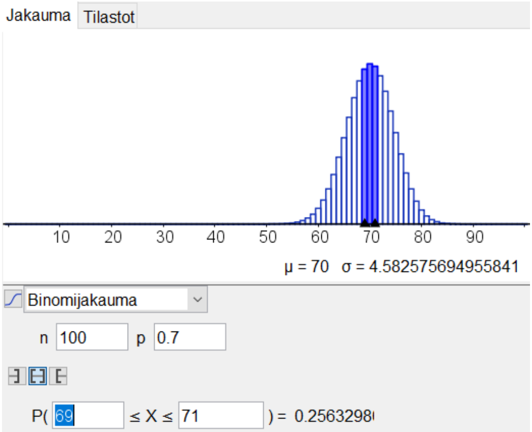


11. **a)** Tilastollinen todennäköisyys sille, että näkkileipä putoaa yhdellä kerralla voipuoli alaspäin, on $\frac{14}{20} = 0,7.$

Määritetään sopivalla ohjelmalla todennäköisyys sille, että sadan heiton sarjassa näkkileipä putoaa alaspäin 69–71 kertaa.

Näkkileipä putoaa voipuoli alaspäin 69–71 kertaa todennäköisyydellä 0,256… ≈ 0,26.

Vastaus: 0,26

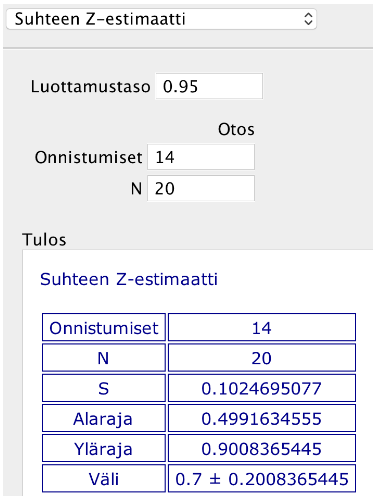


b) Määritetään sopivalla ohjelmalla suhteellisen osuuden 95 %:n luottamusväli, kun osuus otoksessa oli $p = 70\ \% = 0,7$, otoskoko $n = 20$ ja onnistumisten lukumäärä $k = 0,7 \cdot 20 = 14$.

Luottamusväli on [0,4991…; 0,9008…] = [49,91… %, 90,08… %].

Koska luottamusvälin päätepisteet on pyöristettävä keskiarvosta poispäin, niin 95 %:n luottamusväli on [49 %, 91 %].

Vastaus: [49 %, 91 %]



Onnistumiset	14
N	20
S	0.1024695077
Alaraja	0.4991634555
Yläraja	0.9008365445
Väli	0.7 ± 0.2008365445

12. Määritetään virhemarginaali sopivalla ohjelmalla, kun otoksen keskiarvo on $\overline{x} = 65 \text{ min} = 1 \text{ h } 65 \text{ min}$, luottamustaso p = 95 %, keskihajonta s = 25 min ja otoskoko n = 100.

Virhemarginaali on 4,899… min ≈ 5 min.

Vastaus: 5 min

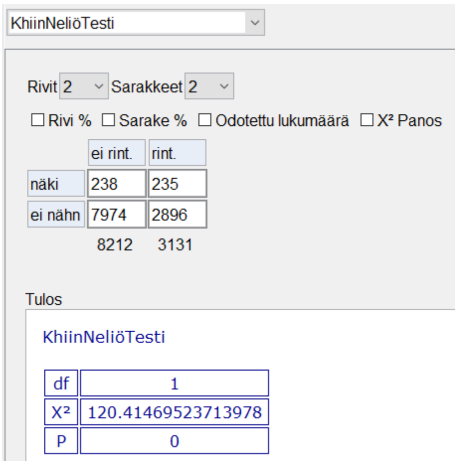
13. Nollahypoteesi on H_0: ”Rintamalla palvelleet näkevät saman verran painajaisia kuin muut sodassa olleet”.
Lasketaan painajaisia usein nähneiden miesten määrät kussakin ryhmässä.

	Ei ollut rintamalla (n = 8212)	Oli rintamalla, ei haavoittunut (n = 1848)	Oli rintamalla, haavoittui muttei pysyvästi (n = 881)	Oli rintamalla, invalidisoitui (n = 402)
Näki usein painajaisia	$0{,}029 \cdot 8212 = 238{,}15 \approx 238$	$0{,}07 \cdot 1848 = 129{,}36 \approx 129$	$0{,}07 \cdot 881 = 61{,}67 \approx 62$	$0{,}109 \cdot 402 = 43{,}82 \approx 44$

Lasketaan myös niiden miesten määrät, jotka eivät nähneet usein painajaisia.

	Ei ollut rintamalla (n = 8212)	Oli rintamalla, ei haavoittunut (n = 1848)	Oli rintamalla, haavoittui muttei pysyvästi (n = 881)	Oli rintamalla, invalidisoitui (n = 402)
Näki usein painajaisia	238	129	62	44
Ei nähnyt usein painajaisia	8212 − 238 = 7974	1848 – 129 = 1719	881 – 62 = 819	402 – 44 = 358

Rintamalla olleita usein painajaisia näkeviä oli yhteensä
129 + 62 + 44 = 235.
Rintamalla olleita, jotka eivät nähneet usein painajaisia, oli yhteensä
1719 + 819 + 358 = 2896.
Tehdään χ^2 -yhteensopivuustesti sopivalla ohjelmalla.



Ohjelman antama p-arvo on 0, joten nollahypoteesin hylkäämisessä otetaan niin pieni riski, ettei ohjelma voi sitä edes ilmoittaa. Nollahypoteesi hylätään siis millä tahansa riskitasolla. Rintamalla palvelleet näkevät enemmän painajaisia kuin muut sodassa olleet.

Vastaus: Väite ei pidä paikkaansa millään riskitasolla.

14. Osan sisältämän metallin määrän keskiarvo on $\overline{x} = 22$ g. Kun muuttuja X on yksittäisen osan sisältämän metallin määrä ja täsmälleen 0,1 % osista sisältää korkeintaan 20 g metallia, niin
$P(X < 20) = 0{,}001$.

Merkitään keskihajontaa kirjaimella s. Muodostetaan yhtälö ja ratkaistaan se sopivalla ohjelmalla.

Normaalijakauma(22,s,20)=0.001

RatkaiseNumeerisesti: $\{s = \mathbf{0.6472005344}\}$

Keskihajonta on pyöristettävä alaspäin, joten se saa olla korkeintaan 0,64.

Vastaus: $s = 0{,}64$

HARJOITUSKOE

H1. **a)** Esimerkiksi pituus on jatkuva ja kvantitatiivinen. Kahden pituuden suhde on mielekästä laskea, joten pituutta mitataan suhdeasteikolla.

Vastaus: esim. pituus

b) Esimerkiksi suosikkikahvi on diskreetti ja kvalitatiivinen muuttuja. Lattea ja cappuccinoa ei ole mielekästä asettaa suuruusjärjestykseen, joten suosikkikahvia mitataan luokitteluasteikolla.

Vastaus: esim. suosikkikahvi

c) Esimerkiksi vuosiluku on diskreetti ja kvantitatiivinen muuttuja. Vuosiluvulla ei ole absoluuttista nollapistettä, joten suhdeasteikkoa ei ole mielekästä käyttää. Vuosilukujen välimatkojen laskeminen on kuitenkin mielekästä, joten vuosilukua mitataan välimatka-asteikolla.

Vastaus: esim. vuosiluku

H2. **a)** Lasketaan normitetut arvot vähentämällä pituudesta keskiarvo ja jakamalla tulos keskihajonnalla.

$$\frac{175\text{ cm}-161\text{ cm}}{5\text{ cm}}=\frac{14\text{ cm}}{5\text{ cm}}=2\frac{4}{5}=2{,}8$$

$$\frac{175\text{ cm}-181\text{ cm}}{6\text{ cm}}=\frac{-6\text{ cm}}{6\text{ cm}}=-1$$

Vastaus: 2,8 ja −1

b) Kohdan a perusteella 175 cm pitkän kivikauden miehen pituus olisi poikennut keskiarvosta 2,8 keskihajonnan verran ylöspäin. Tämä vastaisi nykyajan suomalaismiestä, jonka pituus on
$181\text{ cm} + 2{,}8 \cdot 6\text{ cm} = 197{,}8\text{ cm} \approx 198\text{ cm}$

Vastaus: 198 cm

H3.

a) Jos otoskoko kasvaa, saadaan yhä varmempaa tietoa keskiarvosta, jolloin keskiarvon virhemarginaali pienenee.
Tämän voi päätellä myös seuraavasti: otoskoko on keskiarvon keskivirheen kaavassa jakajana, ja jakajan kasvaessa osamäärä pienenee. Otoskoon kasvaessa keskiarvon keskivirhe siis pienenee. Koska keskiarvon virhemarginaali on suoraan verrannollinen keskiarvon keskivirheeseen, myös keskiarvon keskivirhe tällöin pienenee.

Vastaus: pienenee

b) Jos luottamustasoa kasvatetaan, on tiedettävä entistä suuremmalla todennäköisyydellä, millä välillä keskiarvo tai suhteellinen osuus on. Tällöin luottamusvälin pituus kasvaa.

Vastaus: kasvaa

c) Luottamusvälin pituus on puolet virhemarginaalista, joten virhemarginaalin pienentyessä luottamusvälin pituus pienenee.

Vastaus: pienenee

d) p-arvo on todennäköisyys sille, että nollahypoteesin hylkääminen on väärä johtopäätös, joten p-arvon kasvaessa riski tehdä väärä johtopäätös nollahypoteesia hylättäessä kasvaa.

Vastaus: kasvaa

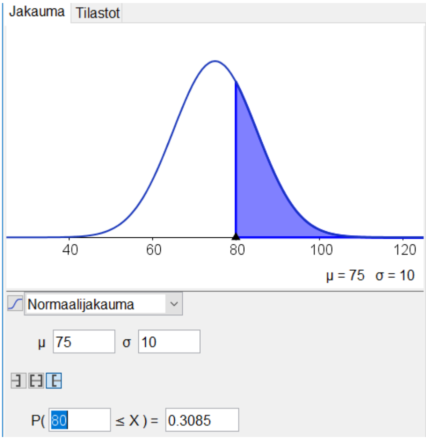
e) Jos perusjoukon keskihajonta pienenee ja otoskoko säilyy samana, perusjoukon keskiarvo tiedetään entistä tarkemmin. Tällöin perusjoukon keskiarvon virhemarginaali pienenee.

Vastaus: pienenee

f) Keskiarvon keskivirhe tarkoittaa otoksen keskiarvon keskihajontaa, joten jos keskiarvon keskivirhe kasvaa, niin otoksen keskiarvon keskihajonta kasvaa.

Vastaus: kasvaa

H4. **a)** Sekunnissa myytyjen hampurilaisten määrä noudattaa normaalijakaumaa, jonka odotusarvo on $\mu = 75$ ja keskihajonta $s = 10$. Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



Ohjelma antaa todennäköisyydeksi 0,3085, joten ketju myy satunnaisesti valitun sekunnin aikana yli 80 hampurilaista todennäköisyydellä $0{,}3085 \approx 0{,}31$.

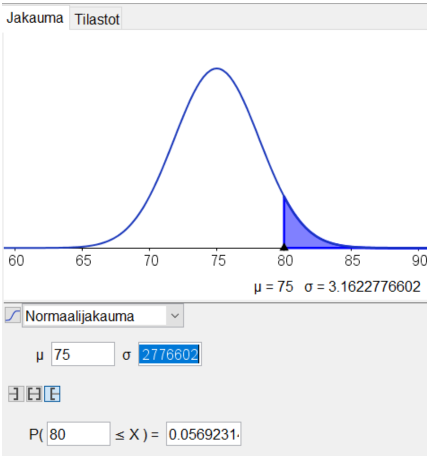
Vastaus: 0,31

b) Ajatellaan, että otetaan 10 sekunnin otos. 10 sekunnin aikana myydään yli 800 hampurilaista, jos otoksen keskiarvo on yli $\frac{800}{10} = 80$ hampurilaista sekunnissa.

Otoksen keskiarvo noudattaa normaalijakaumaa, jonka odotusarvo on $\mu = 75$ ja keskihajonta keskiarvon keskivirhe

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{10}{\sqrt{10}} = 3{,}162...$$

Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



Ohjelma antaa todennäköisyydeksi 0,0569…, joten ketju myy satunnaisesti valitun 10 sekunnin aikana yli 800 hampurilaista todennäköisyydellä 0,0569 ≈ 0,057.

Vastaus: 0,057

H5. **a)** Paikalle saapuvien asiakkaiden lukumäärän odotusarvo on $\mu = np = 0{,}91 \cdot 238 = 216{,}58 \approx 217$.

Vastaus: 217 asiakasta

b) **Tapa 1**

Kaikki saapuvat asiakkaat voidaan päästää sisään, jos heitä on 0, 1, 2, …, 234 tai 235. Kaikkia saapuvia asiakkaita ei voida päästää sisään, jos heitä on 236, 237 tai 238. Lasketaan näiden kolmen asiakasmäärän todennäköisyydet.

$$
\begin{aligned}
P(236 \text{ asiakasta}) &= \binom{238}{236} \cdot 0{,}91^{236} \cdot (1-0{,}91)^{238-236} \\
&= \binom{238}{236} \cdot 0{,}91^{236} \cdot 0{,}09^{2} = 4{,}926... \cdot 10^{-8}
\end{aligned}
$$

$$P(237 \text{ asiakasta}) = \binom{238}{237} \cdot 0{,}91^{237} \cdot 0{,}09^{1} = 4{,}203... \cdot 10^{-9}$$

$$P(238 \text{ asiakasta}) = 0{,}91^{238} = 1{,}785... \cdot 10^{-10}$$

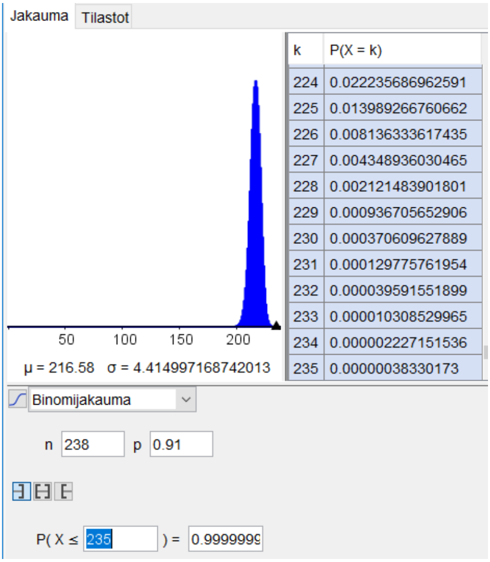
Lasketaan kysytty todennäköisyys vastatapahtuman avulla.

$$
\begin{aligned}
&P(\text{kaikki voidaan päästää sisään}) \\
&= 1 - P(\text{kaikkia ei voida päästää sisään}) \\
&= 1 - [P(236 \text{ asiakasta}) + P(237 \text{ asiakasta}) + ...P(240 \text{ asiakasta})] \\
&= 1 - \left(4{,}926... \cdot 10^{-8} + 4{,}203... \cdot 10^{-9} + 1{,}785.. \cdot 10^{-10}\right) \\
&= 0{,}999999946...
\end{aligned}
$$

Kaikki saapuvat asiakkaat voidaan päästää sisään todennäköisyydellä 0,999999946… ≈ 0,9999999

Tapa 2

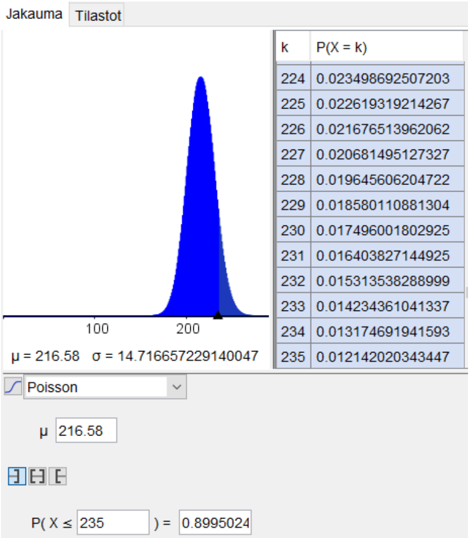
Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



Ohjelma antaa todennäköisyydeksi 0,999999946…, joten kaikki saapuvat asiakkaat voidaan päästää sisään todennäköisyydellä 0,999999946… ≈ 0,9999999.

Vastaus: 0,9999999

c) Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



Kaikki saapuvat asiakkaat voidaan päästä sisään todennäköisyydellä 0,899… ≈ 0,90

Vastaus: 0,90

H6. Tehdään nollahypoteesi H_0: ”Autojen liikennemäärät ovat teorian mukaisia.”

Kopioidaan taulukko sopivaan ohjelmaan ja lasketaan kulkeneiden ajoneuvojen kokonaismäärä.

	A	B
1	Silta	Ajoneuvojen määrä
2	Hakaniemen silta	3367
3	Hämeentien silta	2257
4	Kulosaaren silta	4868
5	Länsiväylä	6643
6	Meilahden silta	655
7	Munkkiniemen silta	3202
8	Pitkäsilta	2268
9	Yhteensä	=SUMMA(B2:B8)

Ajoneuvojen kokonaismääräksi saadaan 23 260.

Merkitään teorian mukaista Meilahden sillalla kulkevien ajoneuvojen määrää kirjaimella x. Tällöin muiden siltojen ajoneuvojen määrät ovat seuraavat:

Länsiväylä $10x$
Kulosaaren silta $10x$
Pitkäsilta $5x$
Hakaniemen silta $5x$
Munkkiniemen silta $5x$
Hämeentien silta $5x$

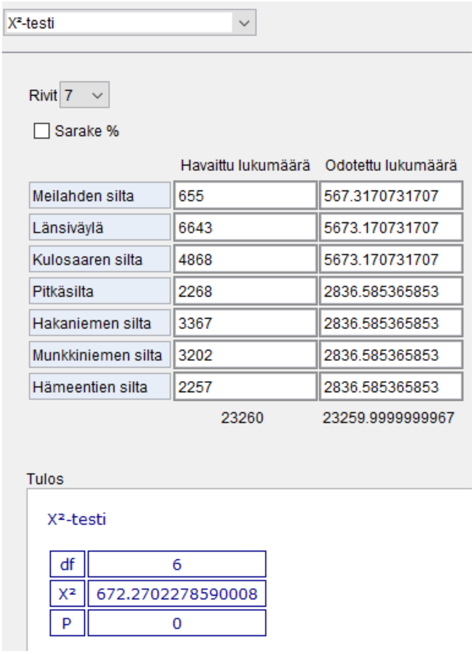
Muodostetaan yhtälö ja ratkaistaan siitä teorian mukainen Meilahden sillalla kulkevien ajoneuvojen määrä x.

$$\begin{aligned} x + 2 \cdot 10x + 4 \cdot 5x &= 23\,260 \\ 41x &= 23\,260 \qquad \|:41 \\ x &= 567{,}317\ldots \end{aligned}$$

Muilla silloilla kulkevien ajoneuvojen määrät ovat tällöin seuraavat:
Länsiväylä ja Kulosaaren silta $10 \cdot 567{,}317\ldots = 5673{,}170\ldots$

Pitkäsilta, Hakaniemen silta, Munkkiniemen silta ja Hämeentien silta
$5 \cdot 567{,}317\ldots = 2836{,}585\ldots$

Tehdään χ^2 -yhteensopivuustesti sopivalla ohjelmalla.



X²-testi

Rivit 7

Sarake %

	Havaittu lukumäärä	Odotettu lukumäärä
Meilahden silta	655	567.3170731707
Länsiväylä	6643	5673.170731707
Kulosaaren silta	4868	5673.170731707
Pitkäsilta	2268	2836.585365853
Hakaniemen silta	3367	2836.585365853
Munkkiniemen silta	3202	2836.585365853
Hämeentien silta	2257	2836.585365853
	23260	23259.9999999967

Tulos

X²-testi

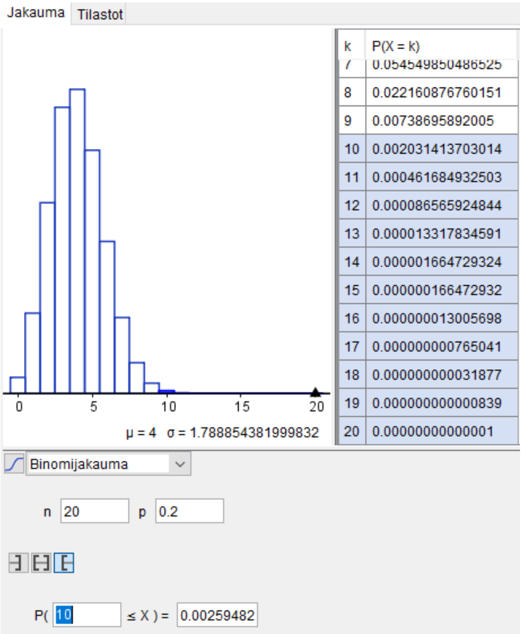
df	6
X²	672.2702278590008
P	0

Ohjelman antama p-arvo on 0, joten nollahypoteesi hylkäämisessä otetaan niin pieni riski, ettei ohjelma voi sitä edes ilmoittaa. Teoria siis ei ole uskottava.

Vastaus: Teoria ei ole uskottava.

H7. Opiskelija tietää vastauksen 10 kysymykseen ja arvaa loput 20 vastausta. Hän saa yksittäisen arvaamansa vastauksen oikein todennäköisyydellä $\frac{1}{5}=0,2.$ Jotta opiskelija saisi 30 vastauksesta oikein vähintään 20, hänen on onnistuttava arvaamaan 20 vastauksesta vähintään 10 oikein.

Lasketaan binomijakauman avulla sopivalla ohjelmalla todennäköisyys sille, että opiskelija saa 20 vastauksesta oikein vähintään 10.



Ohjelma antaa todennäköisyydeksi 0,00259… Siis opiskelija saa vähintään 20 vastausta oikein todennäköisyydellä 0,00259…

Kun 104 000 lukiolaista osallistuu kilpailuun, vähintään 20 vastausta oikein saaneiden opiskelijoiden odotusarvo on
104 000 · 0,00259… = 269,862…

Palkintojen hinnan odotusarvoksi tulee
269,862… · 100 € = 26 986,204… € ≈ 27 000 €.

Vastaus: 27 000 €

H8. Suhteellisen osuuden luottamusväli on

$$\left[p - \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}}, p + \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}}\right].$$

Luottamusvälin pituus on ylä- ja alarajan erotus eli

$$p + \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}} - \left(p - \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}}\right)$$
$$= p + \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}} - p + \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}}$$
$$= 2 \cdot \text{kriittinen arvo} \cdot \sqrt{\frac{p(1-p)}{n}}.$$

Luottamusvälin pituus on siis suoraan verrannollinen lausekkeen $s = \sqrt{\frac{p(1-p)}{n}}$, jossa p on otoksesta laskettu suhteellinen osuus ja n otoskoko, arvoon.

Merkitään alkuperäistä otoskokoa kirjaimella n ja uutta otoskokoa kirjaimella m. Tällöin uusi keskihajonta on $\sqrt{\frac{p(1-p)}{m}}$, ja toisaalta uusi keskihajonta on puolet vanhasta keskihajonnasta. Muodostetaan yhtälö ja ratkaistaan siitä uusi otoskoko m.

$$\sqrt{\frac{p(1-p)}{m}} = \frac{1}{2}\sqrt{\frac{p(1-p)}{n}}$$

Ratkaistaan yhtälöstä m sopivalla ohjelmalla.

```
Ratkaise(sqrt(p(1-p)/m)=(1/2)sqrt(p(1-p)/n),m)
→ {m = 4 n}
```

Uuden otoskoon on oltava vanhaan nähden nelinkertainen, eli otoskokoa on kasvatettava 400 % − 100 % = 300 %.

Vastaus: 300 %