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

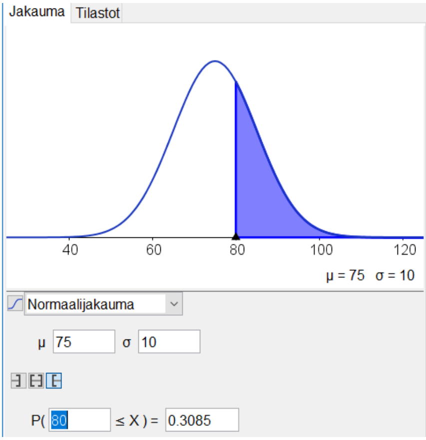
e) Jos perusjoukon keskihajonta pienenee ja otoskoko säilyy samana, perusjoukon keskiarvo tiedetään entistä tarkemmin. Tällöin perusjoukon keskiarvon virhemarginaali pienenee.

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 ≈ 0,31.

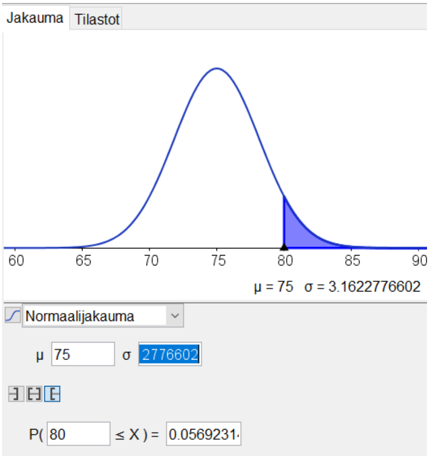
Vastaus: 0,31

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.

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

$$P(237 \text{ asiakasta}) = \binom{238}{237} \cdot 0{,}91^{237} \cdot 0{,}09^{1} = 4{,}203\ldots \cdot 10^{-9}$$
$$P(238 \text{ asiakasta}) = 0{,}91^{238} = 1{,}785\ldots \cdot 10^{-10}$$

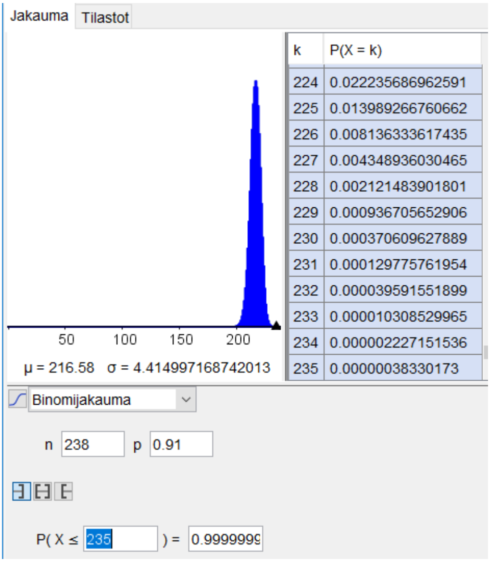
Lasketaan kysytty todennäköisyys vastatapahtuman avulla.

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

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

Tapa 2

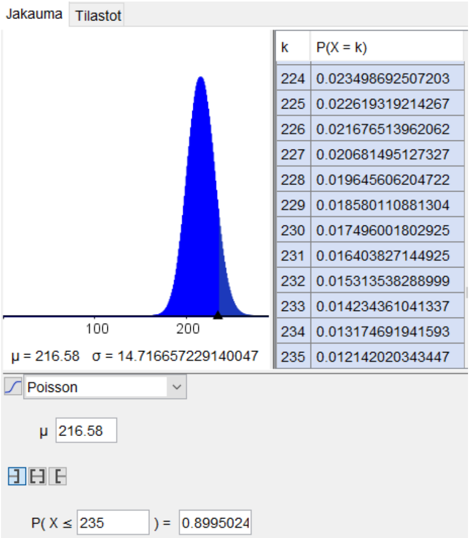
Määritetään kysytty todennäköisyys sopivalla ohjelmalla.



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.

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

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

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

Tehdään χ^2 -yhteensopivuustesti sopivalla ohjelmalla.

Χ²-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

Χ²-testi

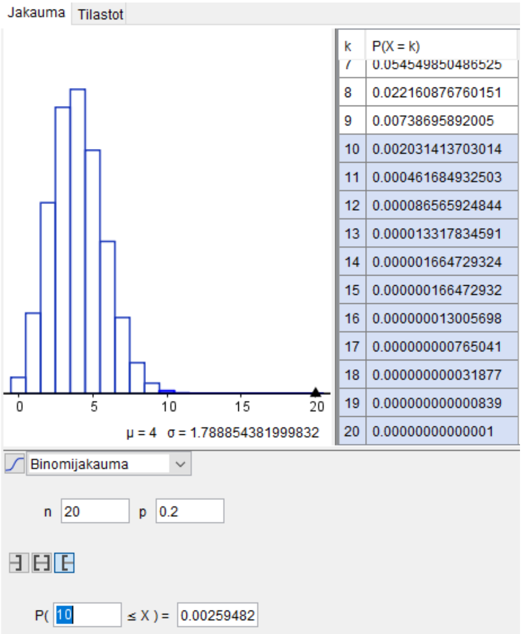
df	6
Χ²	672.2702278590008
P	0

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 %