[MTTTP5] Tilastollisen päättelyn perusteet, syksy 2018
HARJOITUS 5 viikko 48

Ratkaisuja

1.

99%:n luottamusväli odotusarvolle, kun jakauman varianssi on tunnettu, on $\overline{X} \pm z_{0,01/2} \cdot \sigma/\sqrt{n}$. Nyt $z_{0,01/2} = z_{0,005} = 2,58$, koska P(Z > 2,58) = 1 - Φ(2,58) = 1 - 0,9951 = 0,0049 $\approx$ 0,005. Luottamusväli on 1,846± 2,58 · 0,054 / √16 eli 1,846 ± 0,035 eli (1,811, 1,881) .

2.

95%:n luottamusväli odotusarvolle, kun jakauman varianssi on tunnettu, on $\overline{X} \pm z_{0,05/2} \cdot \sigma/\sqrt{n}$. Nyt $z_{0,05/2} = z_{0,025} = 1,96$, koska P(Z > 1,96) = 1 - Φ(1,96) = 1 – 0,975 = 0,025 . Luottamusväli on 336± 1,96 · σ / √9 eli 336 ± 13 / 2 eli (329,5, 342,5) , jolle valmistajan väite 340 kuuluu ja siis uskomme väitteen. Oletuksen (luottamusvälin pituus 13 g) mukaan 1,96 · σ / √9 = 6,5, josta · $\sigma = 9,95$.

3.

100(1-α)%:n luottamusväli odotusarvolle, kun jakauman varianssi on tunnettu, on $\overline{X} \pm z_{\alpha/2}\, \sigma / \sqrt{n}$. Otoskeskiarvo on (21,4 + 18,2) / 2 = 19,8, joten luottamusväli on 19,8 ± $z_{\alpha/2}$· 5 / √25 ja yläraja 21,4 = 19,8 + $z_{\alpha/2}$· 5 / √25 , josta saadaan $z_{\alpha/2}$· 5 / √25 = 1,6. Tästä $z_{\alpha/2}$= 1,6, jolloin α / 2 = 1- 0,9452 = 0,0548 ja α = 0,1096. On siis laskettu 89,04%:n luottamusväli.

4.

100(1-α)%:n luottamusväli odotusarvolle, kun jakauman varianssi on tuntematon, on $\overline{X} \pm t_{\alpha/2;n-1} s/\sqrt{n}$.
Otoskeskiarvo on 72,66, varianssi s^2 =(52795,02 - 10· 72,66^2) / 9 = 0,0293 (kaavasta 1.2), keskihajonta s = √0,0293 = 0,1712.
$t_{0,05/2;10-1}$= 2,262, joten 95%:n luottamusväli on 72,66± 2,262 · 0,1712 / √10 eli 72,66 ± 0,12, joka ei sisällä toivottua tulosta. Päätellään koneen toimivan väärin.

5.

Pala	1	2	3	4	5	6	7	8	9	10
A	63	109	82	156	161	155	47	141	92	149
B	129	105	76	207	253	146	62	160	90	177
Erotus	66	-4	-6	51	92	-9	15	19	-2	28

Koska riippuvat otokset, tarkastellaan erotusta. 100(1-α)%:n luottamusväli odotusarvolle, kun jakauman varianssi on tuntematon, on $\overline{X} \pm t_{\alpha/2;n-1} s/\sqrt{n}$. Erotuksista laskettu keskiarvo on 25, s^2 = 10678 / 9 (kaavasta 1.2) ja keskihajonta s = 34,44, $t_{0,05/2;10-1}$ = 2,262. 95%:n luottamusväli on

$25 \pm 2{,}262 \cdot 34{,}44 / \sqrt{10}$ eli $25 \pm 24{,}64$, joka ei sisällä nollaa. Näin voidaan ajatella erotuksen olevan peräisin jakaumasta, jonka odotusarvo ei ole nolla. Siis menetelmillä on eroja.

6. SPSS-ohje: Tallenna ensin havaintomatriisi, johon kaksi muuttujaa ja 10 tilastoyksikköä. Tee riippuvien otosten t-testi (joka laskee myös vastaavan luottamusväli) odotusarvojen erotukselle Analyze-> Compare Means-> Paired Samples t-test … Voit myös laskea erotukset ja sen jälkeen Analyze-> Compare Means-> One Sample t-test (muuttujaksi erotus).
Ks. SPSS-tuloste
http://www.sis.uta.fi/tilasto/mtttp3/kevat2015/sailymisajat.pdf

7. SPSS-ohje: Valitse tarkasteluun vain keskusta-alue (Data->Select Cases…Kaupunginosa = 'Keskusta'). Luottamusväli odotusarvolle Analyze-> Compare Means-> One Sample t-test
Ks. SPSS-tuloste
http://www.sis.uta.fi/tilasto/mtttp3/kevat2015/H2_T4.pdf

8. Tarkastellaan tyttöjen prosenttiosuutta. Prosenttiosuuden 95%:n luottamusväli on $p \pm 1{,}96 \sqrt{\frac{p(100 - p)}{n}}$. Nyt p = 42,5, n = 200 .
Luottamusväli 42,5 ± 6,85 eli (35,65, 49,35), jolle 50 ei kuulu. Päätellään, että tyttöjä ja poikia ei synny yhtä paljon.

SPSS-ohje: Prosenttiosuuden luottamusväli. Muodosta Sex-muuttujan frekvenssijakauma Analyze-> Descriptive statistics -> Frequencies…, ks. http://cs.uef.fi/statistics/newspss/index.php/fi/2b ja http://www.fsd.uta.fi/menetelmaopetus/frekvenssi/harjoitus2.html .
Laske käsin luottamusväli (kaava 4.3).

9. 95 %:n luottamusväli odotusarvojen erotukselle kaavalla (4.5).
Taulukkoarvo $t_{0,025;72} \approx 2{,}000$.
$s^2 = ((38-1) \times 5{,}9121^2 + (36-1) \times 3{,}2049^2) / (38+36-2) = 22{,}95$, $s = 4{,}79$,
a) $4{,}79\sqrt{\frac{1}{38} + \frac{1}{36}} = 1{,}114$
b) -2,6974 – 2,000x1,114 = -4,9254
c) -2,6974 + 2,000x1,114= -0,4694.

On eroja, koska nolla ei kuulu luottamusvälille.

10. SPSS-ohje: Muodosta aluksi neliöhinta, Transform -> Compute (Nelioh = velatonhinta/neliöt).
Keskiarvot Analyze -> Compare Means -> Means-> … (Dependent = Nelioh, Independent = Sijainti[postinumero])

Laatikko-jana-kuvio Graphs -> Boxplot -> Simple -> (Variable Nelioh, Category Sijainti).

Luottamusväli odotusarvojen erotukselle (ja riippumattomien otosten t-testi) Analyze-> Compare Means-> Independent Samples t-test … (Dependent = Nelioh, Independent = Sijainti[postinumero], Group 1:33540, Group 2:33720), ks. myös
http://cs.uef.fi/statistics/newspss/index.php/fi/7 ja
http://www.fsd.uta.fi/menetelmaopetus/hypoteesi/harjoitus1.html

Ks. SPSS-tulostus:
http://www.sis.uta.fi/tilasto/tiltp2/syksy2010/p2h88.pdf , luottamusväli ei sisällä nollaa, joten keskimääräisissä hinnoissa on eroja. Kalevassa keskimäärin 576-833 euroa korkeampi.

11. SPSS-ohje: Tarkastelu erikseen kaupunginosittain Data -> Split File ->Compare groups -> ryhmittelymuuttujana Sijainti[postinumero].

Analyze-> Compare Means-> Independent Samples t-test … (Dependent = Nelioh, Independent = kunto], Group 1:3, Group 2:4, Option -> 99%:n luottamusväli

Ks. SPSS-tulostus:
http://www.sis.uta.fi/tilasto/tiltp2/syksy2010/p2h89.pdf . Kunto vaikuttaa keskimääräiseen neliöhintaan keskustassa ja Härmälässä mutta ei Kalevassa eikä Hervannassa.