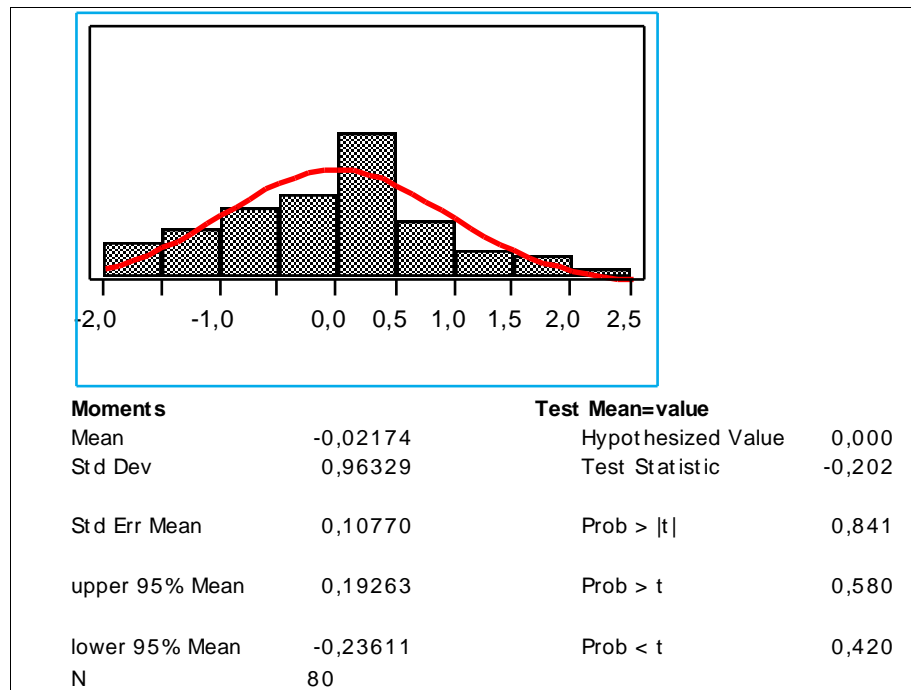


MTTTP5

Lisätehtäviä ratkaisuihin luentomonisteen lukuun 6 liittyen

1. Valitaan 25 alkion satunnaisotos jakaumasta $N(\mu, 25)$. Olkoon $H_0 : \mu = 12$. Hylätään H_0 , jos otoskeskiarvo on suurempi kuin 13,96. Mikä on testissä käytetty α :n arvo?
2. Perunalastujen valmistaja ilmoittaa perunalastupussin keskipainoksi 340 g. Tutkitaan väitettä ja tehdään 16 pussin satunnaisotos ja saadaan keskipainoksi 336 g ja keskihajonnaksi 11 g. Testaa kauppiaan väitettä 5% riskitasolla. Arvioi todennäköisyyttä saada otoksesta saatua testisuureen arvoa harvinaisempi arvo eli arvioi pienintä riskitasoa, jolla valmistajan väite voidaan hylätä.
3. Tarkastellaan oheisia analysointituloksia. Mille populaation parametrille löytyy luottamusväli? Mitä kaavaa käyttäen se on laskettu? Mitä hypoteesia on testattu? Kirjoita testisuureen kaava. Määritä pienin riskitaso, jolla H_0 voidaan hylätä.



4. Tutkitaan onko miesten ja naisten kolesteroliarvot (tietyllä tavalla rajatussa populaatiossa) keskimäärin samoja. Ohessa analysointituloksia. Mitä graafista esitystä olisi ollut mahdollista käyttää? Kirjoita käytetyn testisuureen kaava. Sijoita testisuureen kaavaan aineistosta lasketut luvut. Mitä t-jakauman taulukkoarvoa on käytetty laskettaessa vastaavaa luottamusväliä? Tulkitse tulos.

Group Statistics

sukupuoli		N	Mean	Std. Deviation	Std. Error Mean
kolesteroli	mies	78	7,3126	1,5672	,1775
	nainen	19	6,5926	1,6063	,3685

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
kolesteroli	Equal variances assumed	,059	,808	1,787	95	,077	,7199	,4029	-7,99E-02	1,5197
	Equal variances not assumed			1,760	26,977	,090	,7199	,4090	-,1193	1,5592

5. Tutkittiin opetusmenetelmän vaikutusta oppimiseen. Kurssilla luennot esitettiin toiselle ryhmälle televisioituna ja toiselle ryhmälle tavalliseen tapaan. Osallistujille tehtiin testi sekä ennen että jälkeen kurssin ja laskettiin testipistemäärien erotukset (selitettävä muuttuja). Ohessa saatuja analysointituloksia. Täydennä kohdat a) - c) sekä suorita testaus ja tee johtopäätelmät. Määritä pienin riskitaso, jolla nollahypoteesi voidaan hylätä. Mitkä ovat testin käyttöön liittyvät oletukset?

Group Statistics

OPETUSTAPA	N	Mean	Std. Deviation	Std. Error Mean
Tavallinen	a)	13,748	13,691	2,542
TV	49	17,090	10,883	1,555

Independent Samples Test

Levene's Test for Equality of Variances

F	Sig.
3,637	,060

t-test for Equality of Means

	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	
Equal variances assumed	c)	76	,238	b)	2,810	
Equal variances not assumed		-1,121	48,87	,268	****	2,980

6. On generoitu 100 alkion otos Tas(0,1):sta. Ohessa generointiin liittyviä tuloksia.

Moments	
Mean	0,5155
Std Dev	0,2788
Std Err Mean	0,0279
upper 95% Mean	0,5708
lower 95% Mean	0,4602
n	100

Testaa otoksen perusteella sitä, onko otos peräisin jakaumasta, jonka odotusarvo on kyseisen tasajakauman odotusarvo. Arvioi p-arvoa. Miten luottamusväli on laskettu?

Ratkaisuja

1. Kun H_0 on tosi, niin

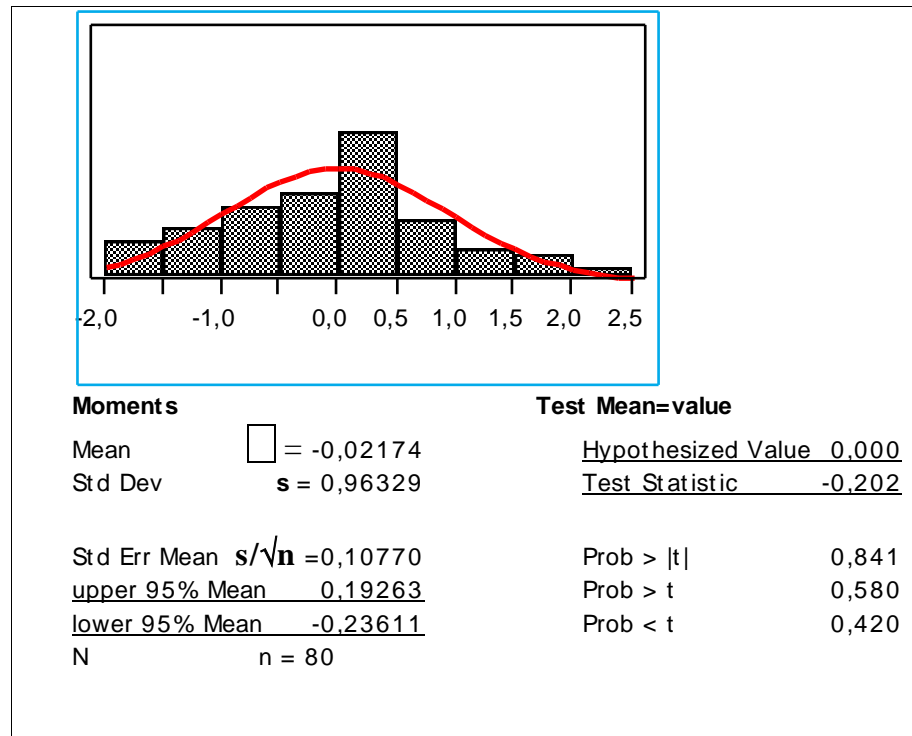
$$P(\bar{X} > 13.96) = 1 - P(\bar{X} \leq 13.96) = 1 - \Phi\left(\frac{13.96 - 12}{5/\sqrt{25}}\right) = 1 - \Phi(1.96) = 1 - 0.975 = 0.025$$

joten $\alpha = 0.025$.

2. $H_0 : \mu = 340$, $H_1 : \mu < 340$. Kun H_0 on tosi, niin $t = \frac{\bar{X} - 340}{s/\sqrt{n}} \sim t(n-1)$.

$t_{\text{hav}} = \frac{336 - 340}{11/\sqrt{16}} = -1.45$. $t_{0.05;15} = 1.753$, $t_{0.1;15} = 1.341$. Koska $t_{\text{hav}} > -1.753$, niin H_0 hyväksytään 5%:n riskitasolla tarkasteltuna. Koska $-1.341 < t_{\text{hav}} < -1.753$, niin p pienin riskitaso, jolla H_0 voidaan hylätä (yksisuuntaisessa testissä) on $0.05 < p < 0.1$

3.



100(1- α)%:n luottamusväli odotusarvolle, kun jakauman varianssi on tuntematon, $\bar{X} \pm t_{\alpha/2; n-1} s/\sqrt{n}$, missä $\alpha = 0.05$. Tässä $H_0: \mu = 0$ ja testisuureen

kaava $t = \frac{\bar{X} - 0}{s/\sqrt{n}}$, josta $t_{hav} = -0.202$ sekä kaksisuuntaisessa testissä $p = 0.841$.

4.

Selitettävä muuttuja on kvantitatiivinen *kolesteroliarvo* ja selittävä luokitteluasteikollinen *sukupuoli*. Graafisesti näitä voi tarkastella pisteparvena tai laatikko-jana -kuviona.

Hypoteesit $H_0: \mu_M = \mu_N$, $H_1: \mu_M \neq \mu_N$. Käytetään riippumattomien otosten t -

testiä, jolloin $t = \frac{\bar{Y}_M - \bar{Y}_N}{S \sqrt{\frac{1}{n_M} + \frac{1}{n_N}}}$, missä $S^2 = \frac{(n_M - 1)S_M^2 + (n_N - 1)S_N^2}{n_M + n_N - 2} = \frac{SS_M + SS_N}{n_M + n_N - 2}$

$$S^2 = \frac{(78 - 1)1.5672^2 + (19 - 1)1.6063^2}{78 + 19 - 2} = 2.4796, \text{ joten}$$

$$t_{\text{hav}} = \frac{7.3126 - 6.5926}{1.5748 \sqrt{\frac{1}{78} + \frac{1}{19}}} = 1.787 \text{ tai suoraan}$$

(tuloksista saadaan Std. Error Difference = 0.4029 = $1.5748 \sqrt{\frac{1}{78} + \frac{1}{19}}$)

$$t_{\text{hav}} = \frac{7.3126 - 6.5926}{0.4029} = 1.787. \text{ Laskettu } 95\% \text{:n luottamusväli, joten kaavassa}$$

käytetty $t_{0.05/2;78+19-2} \approx 1.98$. Taulukkoarvo voidaan määrittää myös

seuraavasti: luottamusvälin yläraja $1.5197 = 0.7199 + t_{0.05/2;95} \cdot 0.4029$, josta

$$t_{0.05/2;95} = 0.7998 / 0.4029 = 1.9851.$$

Koska (kaksisuuntaisessa testissä) pienin riskitaso, jolla H_0 voidaan hylätä on 0.077 (Sig. (2-tailed)), hyväksytään H_0 ja tehdään johtopäätelmä, että kolesteroliarvot ovat miehillä ja naisilla keskimäärin samoja. Saman johtopäätelmän voi tehdä luottamusvälin avulla. Koska nolla kuuluu odotusarvojen erotuksen luottamusvälille, voidaan erotusta pitää nollassa.

5. Group Statistics

OPETUSTAPA	N	Mean	Std. Deviation	Std. Error Mean
Tavallinen	a) 76-49 +2=29	13,748	13,691	2,542
TV	49	17,090	10,883	1,555

Independent Samples Test

Levene's Test for Equality of Variances

F Sig. **Tässä $H_0 : \sigma_{\text{tav}}^2 = \sigma_{\text{TV}}^2$**

3,637 ,060 **Koska $p = 0.06$, voidaan H_0 hyväksyä ja olettaa populaatioiden varianssin yhtä suuriksi.**

t-test for Equality of Means

	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Equal variances assumed	c) -3.342	76	,238	b) 13.748-17.090	2,810
	= -3.342/2,810			= -3.342	
	= -1.189				
Equal variances not assumed	-1,121	48,87	,268	*****	2,980

Kuten tehtävässä 4 (ks. kaavat sieltä) käytetään riippumattomien otosten t -testiä. Hypoteesit $H_0 : \mu_{\text{tav}} = \mu_{\text{TV}}$, $H_1 : \mu_{\text{tav}} \neq \mu_{\text{TV}}$. Pienin riskitaso, jolla H_0 voidaan kaksisuuntaisessa testissä hylätä on $0.238 > 0.05$ (esim.), joten hyväksytään H_0 . Voidaan tietysti myös katso taulukkoarvo $t_{0.05/2;76} \approx 2$. Koska $|-1.189| < 2$, H_0 hyväksytään 5%:n riskitasolla tarkastellen.

Riippumattomien otosten t-testissä oletetaan, että meillä on riippumattomat satunnaisotokset kahdesta normaalijakaumasta, joiden varianssit tuntemattomia, mutta yhtä suuria.

6.

Kyseisen tasajakauman odotusarvo on $(0+1)/2 = 0.5$. $H_0: \mu = 0.5$, $H_1: \mu \neq 0.5$. Jos oletetaan varianssi tuntemattomaksi niin, kun H_0 on tosi $t = \frac{\bar{X} - 0.5}{s/\sqrt{n}} \sim t(n-1)$.

Nyt $t_{\text{hav}} = \frac{0.5155 - 0.5}{0.2788/\sqrt{100}} = 0.5560$, $t_{0.025;99} \approx 1.98$, joten H_0 hyväksytään 5%:n riskitasolla tarkastellen. Koska $t_{0.1;99} \approx 1.289$, niin p-arvo > 0.2 .

Voidaan olettaa myös, että tunnetaan populaation (tasajakauman) varianssi, joka on $(1-0)^2/12 \approx 0.08333$, joten $\sigma = 0.2887$. Tällöin testisuure

$Z = \frac{\bar{X} - 0.5}{0.2887/\sqrt{n}} \sim N(0,1)$, kun tosi H_0 ja $z_{\text{hav}} = 0.537$. Nyt $z_{0.025} \approx 1.96$ ja H_0

hyväksytään 5%:n riskitasolla tarkastellen. Pienin riskitaso, jolla H_0 voidaan hylätä, on $P(|Z| > 0.537) = 2P(Z > 0.537) = 2(1 - P(Z < 0.537)) = 2(1 - \Phi(0.537)) = 2(1 - 0.7054) = 0.5892$.

100(1- α)%:n luottamusväli odotusarvolle, kun jakauman varianssi on tuntematon, on

$\bar{X} \pm t_{\alpha/2; n-1} s/\sqrt{n}$. Tässä käytetty tätä kaavaa (eli on oletettu varianssi

tuntemattomaksi). Luottamusvälin yläraja saadaan $0.5155 + 1.98 \cdot 0.2788/\sqrt{10} = 0.5707$. Jos olettaisi varianssin tunnetuksi yläraja olisi $0.5155 + 1.96 \cdot 0.2887/\sqrt{10} = 0.5721$.