MTTTP5, luento 29.11.2018

Kertausta

$$\boxed{H_0 : \mu = \mu_0}$$

Olk. $X_1, X_2, \ldots, X_n$ on satunnaisotos N(μ, σ^2):sta, missä σ^2 tunnettu. Jos H_0 on tosi, niin

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1). \qquad \text{Kaava 5.1}$$

Esim. Riisipussien pakkauskoneen pitäisi tuottaa keskimäärin kilon pusseja. Painon oletetaan vaihtelevan normaalijakauman mukaisesti keskihajonnan ollessa 2,1 g. Tutkitaan koneen toimivuutta. Punnitaan satunnaisesti valitut 20 pussia, joiden keskipainoksi saadaan 1001 g. Toimiiko kone oikein?

$H_0 : \mu = 1000$

$H_1 : \mu \neq 1000$

Jos H_0 on tosi, niin

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1).$$

Saadaan

$$z_{hav.} = \frac{1001-1000}{2{,}1/\sqrt{20}} = 2{,}13$$

Pienin riskitaso, jolla H_0 voidaan hylätä, on $2P(Z>2{,}13) = 2(1-\Phi(2{,}13)) = 2(1 - 0{,}9834) = 0{,}0338$. Jos valitaan riskitaso, joka on tätä suurempi, niin H_0 hylätään ja H_1 hyväksytään (päätellään: kone ei toimi oikein). Jos valitaan esim. 1 %:n riskitaso, H_0 hyväksytään (päätellään: kone toimii oikein).

Päättely taulukkoarvojen perusteella:

- Jos valitaan 5 %:n riskitaso, niin harvinaisten arvojen raja on $z_{0,05/2} = 1{,}96$. H_0 hylätään, H_1 hyväksytään.
- Jos valitaan 1 %:n riskitaso, niin harvinaisten arvojen raja on $z_{0,01/2} = 2{,}57$. H_0 hyväksytään.

6.1.1 Yhden populaation odotusarvoa koskeva päättely (jatkoa)

$\boxed{H_0 : \mu = \mu_0}$

Olk. $X_1, X_2, \ldots, X_n$ on satunnaisotos N(μ, σ^2):sta, missä σ^2 tuntematon.

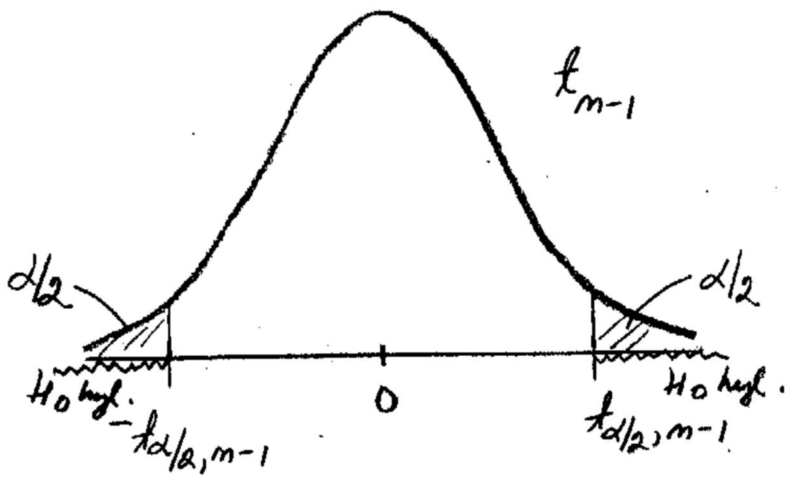
Jos H_0 on tosi, niin

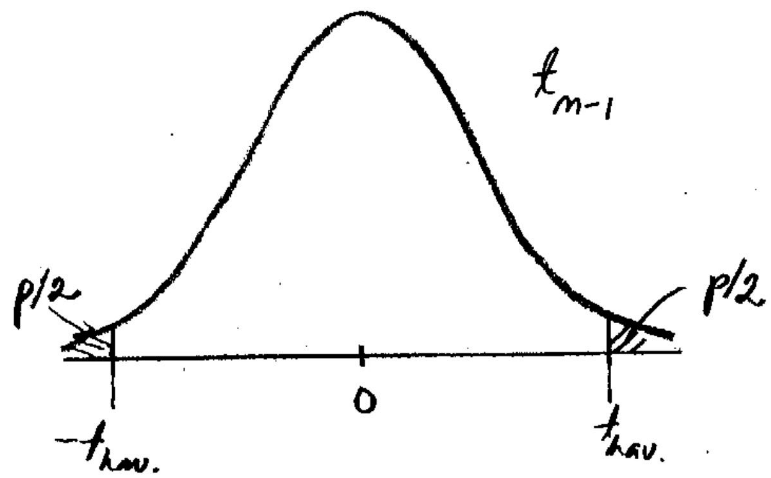
$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1} \,. \qquad \text{Kaava 5.2}$$

- Jos $H_1 : \mu \neq \mu_0$, niin H_0 hylätään riskitasolla α, jos otoksesta laskettu $|t_{hav.}| > t_{\alpha/2,\, n-1}$.

  Pienin riskitaso p, jolla H_0 voidaan hylätä, on

  $2P(t_{n-1} > |t_{hav.}|)$.

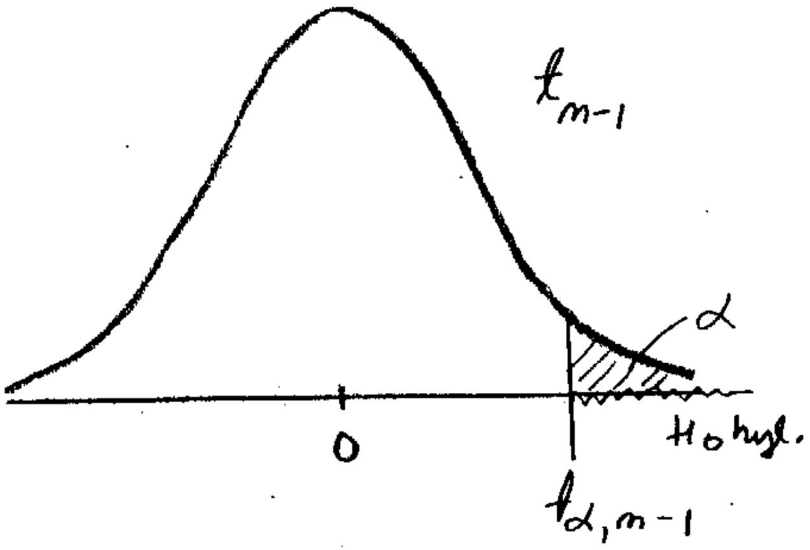


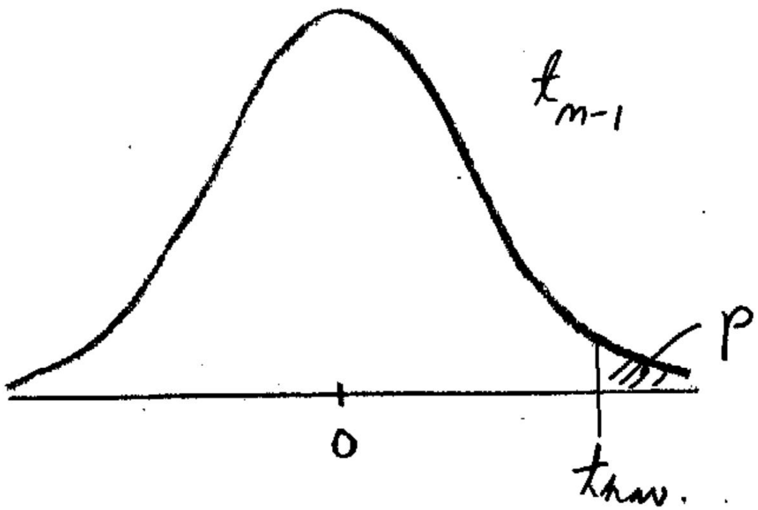


- Jos $H_1 : \mu > \mu_0$, niin H_0 hylätään riskitasolla α, jos otoksesta laskettu $t_{hav.} > t_{\alpha,n-1}$

  Pienin riskitaso p, jolla H_0 voidaan hylätä, on

  $P(t_{n-1} > t_{hav})$.

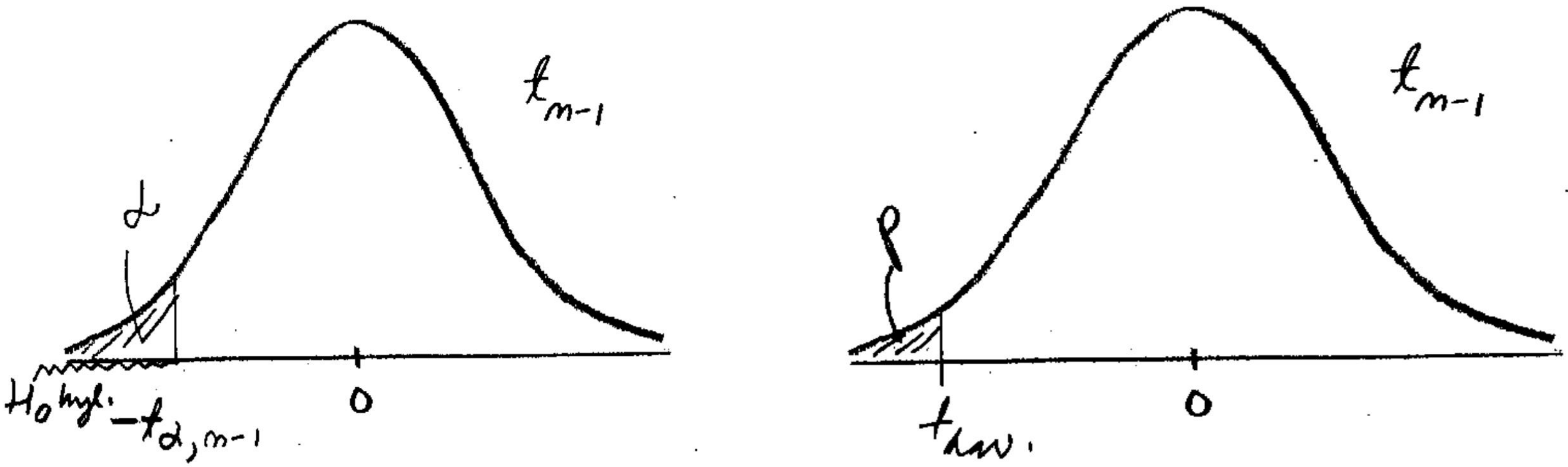




- Jos $H_1 : \mu < \mu_0$, niin H_0 hylätään riskitasolla α, jos otoksesta laskettu $t_{hav.} < -t_{\alpha,n-1}$

  Pienin riskitaso p, jolla H_0 voidaan hylätä, on

  $P(t_{n-1} < t_{hav.})$.



Esim. 6.1.3 Lepakot paikallistavat hyönteisiä lähettämällä korkeataajuista ääntä. Kaiun kuulemiseen kuluvan ajan perusteella ne pystyvät paikallistamaan hyönteiset. Tutkijat arvelivat, että keskimääräinen tunnistusmatka voisi olla yli 35 cm. He keräsivät aineiston mitaten etäisyydet (cm), joista lepakot löysivät hyönteisiä. Mitatut etäisyydet olivat 62, 52, 68, 23, 34, 45, 27, 42, 83, 56, 40. Voidaanko saatujen tulosten perusteella pitää tutkijoiden arvioita oikeana?

$H_0 : \mu = 35$

$H_1 : \mu > 35$

Jos H_0 on tosi, niin

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}\ .$$

Saadaan

$\bar{x}$ = 48,36, s = 18,085, $t_{hav.} = \frac{48{,}36 - 35}{18{,}085/\sqrt{11}} =$ 2,450

Koska $t_{hav.} = 2,450 < t_{0,01;10} = 2,764$, H_0 hyväksytään 1 % riskitasolla.

Koska $t_{hav.} = 2,450 > t_{0,025;10} = 2,228$, H_0 hylätään 2,5 % riskitasolla

Siis $0,01 < p < 0,025$ (SPSS: 0,034/2 = 0,017)

SPSS-ohjeet:

Analyze -> Compare Means -> One-Sample T Test -> Test Variable Matka, Test Value 35

SPSS-tulostus

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Matka	11	48,36	18,085	5,453

One-Sample Test

	Test Value = 35			
	t	df	Sig. (2-tailed)	Mean Difference
Matka	2,451	10	,034	13,364

Esim. Valmistaja ilmoittaa kynttilöittensä keskimääräiseksi palamisajaksi 9,5 tuntia. Tutkit asiaa ja teet valmistajan kynttilöistä 6 kynttilän satunnaisotoksen. Mittaat näiden palamisajat ja saat keskiarvoksi 8,8 tuntia ja keskihajonnaksi 0,48 tuntia. Voitko uskoa valmistajan väitteen?

$$H_0 : \mu = 9{,}5$$

$$H_1 : \mu < 9{,}5$$

Jos H_0 on tosi, niin $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$. Nyt $\mu_0 = 9{,}5$, $n = 6$,

$\bar{x} = 8{,}8$, $s = 0{,}48$, $t_{hav.} = \frac{8{,}8-9{,}5}{0{,}48/\sqrt{6}} = -3{,}57$.

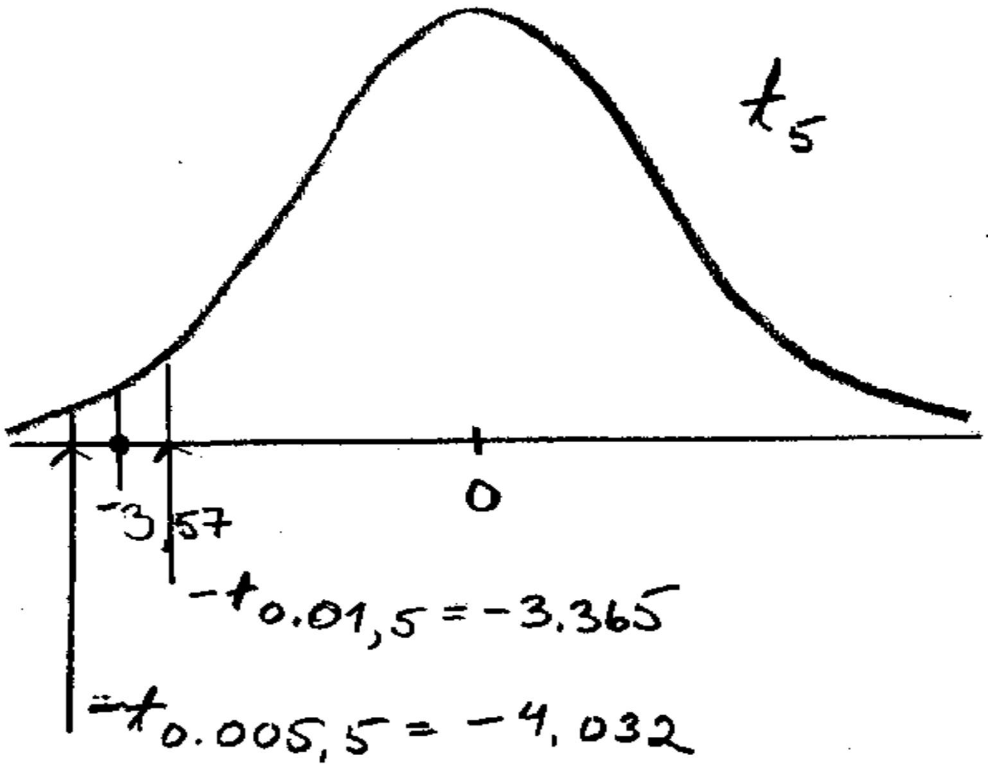
H_0 hyväksytään 0,5 % riskitasolla, koska

$t_{hav.} = -3{,}57 > -t_{0{,}005{,}5} = -4{,}032$

$t_{0{,}01{,}5} = 3{,}365$ H_0 hylätään 1 %:n riskitasolla, koska

$t_{hav.} = -3{,}57 < -t_{0{,}01{,}5} = -3{,}365$.

Siis 0,005 < p < 0,01



Esim. Tutkittiin pH-mittarin toimivuutta. Mitattiin 14 neutraalin (pH = 7) liuoksen pH-arvot, joiksi saatiin 7,01, 7,04, 6,97, 7,00, 6,99, 6,97, 7,04, 7,04, 7,01, 7,00, 6,99, 7,04, 7,07, 6,97. Toimiiko mittari oikein?

$H_0 : \mu = 7$

$H_1 : \mu \neq 7$

Jos H_0 on tosi, niin $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}} \sim t_{n-1}$. Nyt $\mu_0 = 7$, $n = 14$, $\bar{x} = 7{,}01$, $s = 0{,}03162$, $t_{hav.} = \frac{7{,}01-7}{0{,}03162/\sqrt{14}} = 1{,}183$.

Jos valitaan $\alpha = 0,05$, niin $t_{0,05/2;13} = t_{0,025;13} = 2,160$.

H_0 hyväksytään 5 %:n riskitasolla, koska

$-2,160 < t_{hav.} < 2,160$.

Pienin riskitaso, jolla H_0 voidaan hylätä, on > 0,05.

Laskurin http://vassarstats.net/ tulos
http://www.sis.uta.fi/tilasto/mtttp5/syksy2015/pH_mittari.pdf

SPSS-tulostus

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
pH	14	7,0100	,03162	,00845

One-Sample Test

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
pH	1,183	13	,258	,01000	-,0083	,0283

Esim. Eräs yritys tarjoaa valmennuskurssia yliopistoon pyrkijöille. Yritys haluaa tutkia kurssinsa tehokkuutta. Tutkitaan pareja, joilla on samanlaiset lähtötiedot. Toinen osallistuu valmennuskurssille, toinen ei. Saadaan aineisto, jossa pyrkijöiden valintakoepisteet. Onko kurssi tehokas?

	Osallistui	Ei osallistunut	Erotus d
Pari 1	82	75	7
Pari 2	73	71	2
Pari 3	59	52	7
Pari 4	48	46	2
Pari 5	69	70	-1
Pari 6	93	83	10

Tutkitaan, onko erotus peräisin jakaumasta, jonka odotusarvo nolla. Siis

$H_0 : \mu_D = 0$

$H_1 : \mu_D > 0$

Jos H_0 on tosi, niin $t = \frac{\bar{D}}{s/\sqrt{n}} \sim t_{n-1}$. Nyt $\sum d_i = 27$,

$\sum d_i^2 = 207$, $\bar{d} = 4{,}5$ $SS_d = 207 - 27^2/6 = 85{,}5$, s = 4,135, $t_{hav.} = \frac{4{,}5}{4{,}135/\sqrt{6}} = 2{,}666$.

$t_{0,025;5} = 2{,}571$, $t_{0,01;5} = 3{,}365$, joten $0{,}01 < p < 0{,}025$

Ks. SPSS-tuloste

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
D	6	4,50 $= \bar{d}$	4,135 $= s$	1,688 $= s/\sqrt{n}$

Taulukosta

$t_{0,025,5} = 2,571 < 2,666$

H_0 hyl.

$t_{0,01,5} = 3,365 > 2,666$

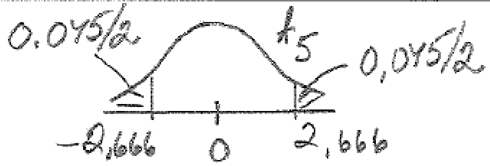
H_0 hyv.

One-Sample Test

	Test Value = 0 $H_0: \mu = 0$					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
D	2,666	5	,045	4,500	,16	8,84

$0,01 < p\text{-arvo} < 0,025$

$t = \frac{4,5-0}{4,135/\sqrt{6}}$



1-suuntaisessa testissä
p-arvo = 0.045/2 = 0.0225