

[MTTTP5] Tilastollisen päättelyn perusteet, syksy 2018

HARJOITUS 7 viikko 50

RYHMÄT:

ma	10.15–11.45	Is. C6	Leppälä
ma	12.15–13.45	Is. C6	Männikkö
ti	08.30–10.00	Is. C6	Männikkö
ti	10.15–11.45	Is. C6	Männikkö
ti	12.15–13.45	Is. C6	Leppälä

Nämä ovat kurssin viimeiset harjoitukset. Sopivan ryhmän voi valita vapaasti. Tentti to 13.12.2018 klo 9.15–12.00 Is. A1. Tähän tenttiin voi osallistua, jos on tehnyt kurssin aikana vähintään 30 % laskuharjoituksista.

Aihe: hypoteesien testaus

1. Kymmenen vuotta sitten eräessä yliopistossa tehdyn tutkimuksen mukaan 75 % yliopiston opiskelijoista uskoi löytävänsä koulutustaan vastaavaa työtä valmistuttuaan. Haluttiin tutkia, oliko tämä prosenttiosuus muuttunut. Kysyttiin mielipidettä satunnaisesti valituilta opiskelunsa aloittaneilta. Kyselyyn vastasi 300 opiskelijaa, joista 240 uskoi työllistyvänsä koulutustaan vastaaviin tehtäviin valmistumisen jälkeen. Onko kymmenessä vuodessa tapahtunut muutosta? Määritä tilanteeseen liittyvät hypoteesit ja suorita testaus. Laske myös pienin riskitaso, jolla H_0 voidaan hylätä.

2. Kasvitieteilijä olettaa, että erään kasvin kukista $\frac{1}{4}$ on valkoisia ja loput punaisia. Hän kasvatti 200 kasvia, joista 30 oli valkokukkaista. Onko kasvitieteilijän syytä muuttaa käsitystään? Määritä tilanteeseen liittyvät hypoteesit ja suorita testaus. Laske myös pienin riskitaso, jolla H_0 voidaan hylätä.

3. Eräessä tehdas valmistaa tiettyjä komponentteja koneilla A ja B. Tarkastellaan komponenttien pituuksia. Tiedetään, että molemmat koneet tuottavat keskimäärin samanmittaisia komponentteja. Oletetaan, että molempien koneiden tuottamien komponenttien pituus vaihtelee normaalijakauman mukaisesti keskihajontana 0,005 cm. Molempia koneita joudutaan korjaamaan. Korjauksen jälkeen tehdään laadunvalvontaa ja valitaan satunnaisesti 100 komponenttia molempien koneiden tuotannosta. Valittujen komponenttien pituuksien keskiarvojen erotuksen itseisarvoksi saadaan 0,002 cm. Tuottavatko molemmat koneet korjauksenkin jälkeen keskimäärin samanmittaisia komponentteja? Aseta tilanteeseen sopiva H_0 ja H_1 . Suorita testaus 1 %:n riskitasolla. Laske pienin riskitaso, jolla H_0 voidaan hylätä.

4. Tutkitaan kerrostalohuoneistojen myyntihintoja Tampereen Hervannassa, aineistona Tre_myydyt_asunnot_2010.sav sivulta <https://coursepages.uta.fi/mtt1/esimerkkiaineistoja/>. Halutaan selvittää, vaikuttaako huoneiston kunto keskineliöhintaan. Ohessa on osa analysointituloksista. Aseta tilanteeseen liittyvät hypoteesit. Täydennä analysointituloksiin kohdat a) – c). Suorita testaus. Mitkä ovat testin käyttöön liittyvät oletukset? Mitä Levenen testillä testataan? Mikä päättely tässä sen perusteella tehdään?

Group Statistics

Huoneiston kunto		N	Mean	Std. Deviation
Neliöhinta	hyvä	91	1815,6627	428,81497
	tydyttävä	37	1636,1252	393,10922

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
Neliöhinta	Equal variances assumed	1,272	,262	a)	b)	c) Arvioi	179,53750	81,68055
	Equal variances not assumed			2,281	72,475			

5. Tutkit sairauden vuoksi työstä poissaoloja. Haluat vertailla päivätyötä ja yötyötä tekeviä. Sinulla on *ennakkoarvio, että yötyöläiset sairastavat enemmän*. Teet satunnaisotokset molemmista työntekijäryhmistä ja lasket heidän sairauspoissaolopäivien lukumäärät kuluneen vuoden ajalta saaden tulokset:

Yötyö	15	10	10	7	7	4	9	6	10	12
Päivätyö	8	9	2	0	0	9	9	7	3	3

Muodosta havaintomatriisi. Aseta tilanteeseen sopiva H_0 ja H_1 . Testaa väitettäsi 1 %:n riskitasolla. Arvioi pienintä riskitasoa, jolla H_0 voidaan hylätä.

6. Ohessa on osa tehtävän 5 analyysin SPSS-tulosteesta. Täytä kohdat a) – h) tehtävässä 5 saamiesi laskutulosten perusteella.

Group Statistics

Työvuoro		N	Mean	Std. Deviation	Std. Error Mean
Poissaolopäivät	Yötyö	10	a)	c)	1,000
	Päivätyö	10	b)	d)	1,193

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
Poissaolopäivät	Equal variances assumed	2,009	,173	h)	g)	,019	e)	f)
	Equal variances not assumed			2,570	17,469		,020	4,000

7. Tutkitaan viljelymenetelmien A ja B vaikutusta satoon. Jaetaan pelto 16 lohkoksi ja arvotaan menetelmät näihin lohkoihin siten, että molempia menetelmiä käytetään yhtä monta kertaa. Tämän jälkeen mitataan jokaisesta lohkoksi sato (tulokset ohessa). Aseta tilanteeseen liittyvät hypoteesit ja määritä testisuureen arvo. Mitä jakaumaa testisuure noudattaa nollahypoteesisi ollessa tosi? Suorita testaus 1 %:n riskitasolla. Määritä pienin riskitaso, jolla nollahypoteesisi voidaan hylätä. Käytä SPSS -ohjelmaa (tai vastaavaa).

A 82	A 68	B 131	B 113
B 121	A 109	A 95	A 102
B 100	A 106	B 116	B 89
B 99	B 119	A 81	A 80

8. Tarkastele Tampereella kerrostaloasuntojen vuokria käyttäen sivulla <https://coursepages.uta.fi/mtt1/esimerkkiaineistoja/> olevaa aineistoa Tre_vuokra-asunnot_2011.sav. Tutki eroavatko keskustan ja Hervannan keskineliövuokrat toisistaan.

Tee analyysin tuloksiin liittyvät tehtävät Moodlessa kohdassa Harjoitus 7, tehtävä 8 (<https://learning2.uta.fi/mod/quiz/view.php?id=640232>). Tee nämä Moodle-tehtävät viimeistään ke 12.12. klo 9. Lue myös tehtävistä automaattisesti tuleva palaute.