

Tilastotieteen peruskurssi  
Tentti 04.11.2016

Tentissä saa olla mukana:

- Kirjoitusvälineet
- Ylioppilaskirjoituksiin hyväksyty laskin
- Tenttipaperin yhteydessä jaettavat liitteet

1. Kuvaa lyhyesti mitä vikaa on seuraavissa tilanteissa. **Kuhunkin kohtaan vastauksen maksimipituus on 5 riviä, ylimääräiset rivit jätetään huomiotta.**
  - (a) Yliopistossamme tehdään opintojen arviointiperusteisiin muutoksia, joiden tarkoituksena on lisätä opiskelijoiden tyytyväisyyttä. Nolla-hypoteesia, että "tyytyväisyys lisääntyi", testataan vastahypoteesia "muutosta tyytyväisyydessä ei tapahtunut" vastaan.
  - (b) Tutkimusraportissa lukee, että tulokset ovat tilastollisesti merkitseviä ja P-arvo on 0.95.
  - (c) t-testisuureen arvo on 0.018. Koska tämä on pienempi kuin  $\alpha = 0.05$ , nollahypoteesi hylätään.
2. Viljelykokeessa on 250 koealuetta. Alueilla kasvatetaan vehnää ja muuttuja Satotavoite on määritelty siten, että se saa arvon 1, kun haluttu sadon määrä on saavutettu kyseisellä koealueella. Jos satotavoitetta ei ole saavutettu, niin kyseinen muuttuja saa arvon nolla. Muuttuja Lannoitus saa arvon 1, jos kyseinen koealue on lannoitettu ja arvon 0 jos aluetta ei ole lannoitettu. Tulokset on kirjattu alla olevaan taulukkoon.

		Lannoitus	
		0	1
Satotavoite	0	115	60
	1	35	40

Tarkastele testaamalla, vaikuttaako Lannoitus Satotavoitteeseen. Jos vaikuttaa, niin selitä perustellen miten. Kirjoita testattavat hypoteesit näkyviin. Voit käyttää merkitsevyystasoa  $\alpha = 0.05$ .

3. Eräs suuri pankki halusi tarkastella miten perehdyttämisaian pituus vaikuttaa aikaan, jossa työntekijä saa tehtyä valmiiksi tietyn tyyppisen projektin. Muuttuja koulutus kuvaa perehdyttämiseen käytettyä aikaa tunneissa ja muuttuja projekti kuvaa työntekijän projektin valmiiksi saamiseksi käyttämää aikaa (tunneissa). Analyysi suoritettiin yksinkertaisella regressiolla, jossa vasteena

Käännä

on muuttuja projekti ja selittäjänä on muuttuja koulutus. Alla on ko. analyysistä saatu tuloste:

```
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 27.46608    1.13830  24.129 3.33e-07 ***
koulutus    -0.44470    0.05617  -7.916 0.000216 ***
```

Kirjoita näkyviin muodostetun mallin yhtälö. Tee parametrien tulkinnat. Onko koulutus tilastollisesti merkitsevä tekijä (kirjoita näkyviin myös hypoteesit).

Sovitetun mallin selitysasteesi on saatu 0.9126. Mitä tämä tarkoittaa?

Mikä on arviosi projektin valmiiksi saattamiseen tarvittavasta ajasta, jos työntekijään on perehdytetty 22 tuntia?

4. Astronomi arvelee, että erään tähden etäisyys maasta olisi 16 valovuotta. Hän kuitenkin suorittaa vielä joitakin etäisyysmittauksia. Kuudella mittauksella hän saa seuraavat tulokset (valovuosia):

15.1, 14.8, 14.0, 15.2, 14.7, 14.5

Muodosta näiden mittausten perusteella, astronomin puolesta, 95%:n luottamusväli tähden ja maapallon välisen etäisyyden odotusarvolle. Selitä perustelun voidaanko ajatella, että etäisyys on noin 16 valovuotta.