

Tilastotieteen peruskurssi  
Tentti 23.03.2018

Tentissä saa olla mukana:

- Kirjoitusvälineet
- Laskin
- Tenttipaperin yhteydessä jaettavat liitteet

1. Viljelykokeessa on 250 koealuetta. Alueilla kasvatetaan vehnää. 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. Käyttää merkitsevyytensä  $\alpha = 0.05$ .

2. 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 käyttäen regressioanalyysia, jossa vasteena 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 ***
```

**Käännä**

Kirjoita näkyviin muodostetun mallin yhtälö (mallin systemaattinen osa). Tulkitse tulosteesta koulutuksen kerroin ( $\beta_1$ ). Tulkitse tulosteesta muuttujaan koulutus liittyvä testi ts. 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?

3. 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

Mittausten populaatiovarianssi  $\sigma^2$  on 0.2. Toisin sanoen mittaustuloksena on lukeman, joka on muotoa "todellinen etäisyys" + "satunnainen virhe". Tämä satunnainen virhe on normaalijakautunut odotusarvolla 0 ja varianssilla  $\sigma^2 = 0.2$ .

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

4. Tarkastellaan kuvitteellista tutkimusta, johon osallistui 5 miestä. Tutkimuk-  
sessa nämä miehet alkoivat tehdä ampumahiihtoharjoittelua. Miesten painoin-  
deksi mitattiin ennen harjoittelun alkamista ja 6:n kuukauden harjoittelun jäl-  
keen. Mittaustulokset ovat alla:

	BMI				
	1	2	3	4	5
6 kuukautta	21.61	22.64	21.35	24.44	22.83
0 kuukautta	26.24	29.94	22.77	25.54	26.10

Testaa, käyttäen kurssilla esitettyä menetelmää, vaikuttaako vaikuttaako am-  
pumahiihtoharjoittelu painoindeksiin. Kirjoita näkyviin testaamasi hypotee-  
sit. Käytä merkitsevyystasoa  $\alpha = 0.05$ .