

TILM3555 Tilastollisen päättelyn peruskurssi

Tentti 27.3.2017

Tentissä saa olla mukana kirjoitusvälineet, kaavakokoelma sekä laskin.

Graafisen ja symbolisen laskimen käyttö on kielletty.

Kaavakokoelma jaetaan tentissä.

Merkitse jokaiseen konseptiin nimi, opiskelijanumero, päivämäärä sekä kurssin koodi (TILM3555).

Kirjoita ylös käyttämäsi kaavat ja tekemäsi oletukset. Erityisesti muista ilmoittaa mikä on hypoteesiparisi, jos teet tilastollista testausta.

1. (max 2p.) Eräässä populaatiossa esiintyy erästä ominaisuutta osuudella p ihmisistä. Populaatiosta on kerätty n kpl riippumattomia havaintoja. Mikä on suhteellista osuutta estimoivan suhteellisen frekvenssin tarkka keskivirhe?
2. (max 2p.) Mikä on hypoteesin testauksessa tyypin I virheen määritelmä?
3. (max 2p.) Mikä on hypoteesin testauksen p -arvon määritelmä?
4. (max 2p.) Eräästä tilastollisesta muuttujasta on kerätty havaintoaineisto, josta ollaan laskettu vinouskertoimen arvoksi $g_1 = 1.9$. Esitä kaksi mahdollista muunnosta, jolla vinouskertoimen arvo voitaisiin saada pienemmäksi.
5. Kahden eri pintamateriaalin kulumista testattiin. Materiaalin 1 testikuluman otoskeskiarvoksi $n_1 = 11$ testissä saatiin $\bar{x}_1 = 80$ (tietyissä yksiköissä) otoskeskihajonnan ollessa $s_1 = 2$. Materiaalin 2 testikuluman otoskeskiarvoksi $n_2 = 16$ testissä saatiin $\bar{x}_2 = 81$ ja otoskeskihajonnaksi $s_2 = 1$. Jakaumat voidaan olettaa normaaleiksi samoin varianssein.
 - (a) (max 2p.) Hae otosvariaatiokerroin materiaalin 1 testikulumille.
 - (b) (max 2p.) Hae 95%-luottamusväli pintamateriaalien odotettujen kulumisten erotukselle $\mu_1 - \mu_2$.
6. (max 2p.) Tutkittiin suomalaisten lukiolaispoikien pituutta. Kerättiin havaintoaineisto

$$x_i: 164 \ 167 \ 170 \ 172 \ 175 \ 178 \ 180 \ 183 \ 186$$

$\sum_{i=1}^9 x_i = 1575$, $\sum_{i=1}^9 x_i^2 = 276063$. Käyttäen jotain epäparametrista testiä, voidaanko tasolla 0.05 sanoa tämän aineiston perusteella, että suomalaisten lukiolaispoikien pituuden mediaani poikkeaa suomalaisten 30-40 vuotiaiden mediaanista 180cm?

7. (max 2p.) Eräästä 80-luvulta tutusta televisio-ohjelmasta ollaan tekemässä uutta nykyaikaistettua versiota, jossa pääosassa on nokkela katujen kasvatti, jonka paras kaveri on taidokas hakkeri. Televisioyhtiö järjestää mielipidemittauksen, jossa 100 katsojalta kysyttiin pitääkö hän uudesta vai vanhasta versiosta enemmän. Mielipidemittaukseen osallistujista 60 piti uutta versiota huonompana verrattuna vanhaan versioon. Voidaanko tämän otoksen perusteella sanoa tasolla 0.05, että vanha versio on uutta suositumpi koko populaatiossa? Ilmoita testin hypoteesipari sekä p -arvo.

8. Eräässä yliopistossa pidetään kahta eri oppiaineille tarkoitettua matematiikan peruskurssia (A ja B), ja kummaltakin kurssilta on valittu satunnaisesti 7 opiskelijaa edustamaan kurssiaan tasokokeessa, jossa mitataan opiskelijan oppimia matematiikan tietoja ja taitoja. Seuraavat luvut ovat opiskelijoiden kyseisessä kokeessa saavuttamia pistemääriä

$$\begin{aligned} \text{Kurssi A: } (x_i) & \quad 24 \quad 22 \quad 21 \quad 20 \quad 19 \quad 15 \quad 14 \\ \text{Kurssi B: } (y_i) & \quad 23 \quad 17 \quad 16 \quad 13 \quad 12 \quad 10 \quad 9 \end{aligned}$$

Tarvittaessa voit käyttää hyödyksi tuloksia

$$\begin{aligned} \sum_{i=1}^7 x_i &= 135, & \sum_{i=1}^7 x_i^2 &= 2683 \\ \sum_{i=1}^7 y_i &= 100, & \sum_{i=1}^7 y_i^2 &= 1568 \end{aligned}$$

- (a) (max 2p.) Tutki epäparametrisella tilastollisella merkitsevyydestillä tasolla 0.05 onko kurssin A käyneiden opiskelijoiden pisteet parempi kuin kurssin B käyneiden opiskelijoiden. Muista mainita mikä on hypoteesiparisi.
- (b) (max 2p.) Mitä oletuksia teit populaatiojakaumista (a)-kohdan testiä varten.
9. Kvantitatiiviselle muuttujalle X on otoksesta laskettu seuraavan näköinen frekvenssijakauma:

x_i	-3	-2	-1	0	1	2	3
f_i	2	5	15	10	10	4	4

- (a) (max 2p.) Määrää aineiston moodi
- (b) (max 2p.) Määrää aineiston otoskeskiarvo
10. Tarkastellaan 12 päiväkotilapsen erään testin mukaisesti kerättyä aggressiivisuusaineistoa. Mitä suurempi on testin pisteluku, sitä aggressiivisemmin lapsi käyttäytyi koetilanteessa.

10	11	12	13	14	15
16	22	23	25	35	45

- (a) (max 2p.) Muodosta aineistolle laatikko-janakuvi **käyttäen osuusajattelua**.
- (b) (max 2p.) Näyttääkö aineisto laatikko-janakuvi perusteella vasemmalle vinolta, symmetriseltä vai oikealle vinolta? Aggressiivisuusasteet on mitattu suhteasteikolla. Mitä tämä tarkoittaa?
11. (max 2p.) Eräs tutkimus peruskoulun ensimmäisen luokan oppilaiden sopeutumisesta koulumaailmaan tehtiin siten, että opettajat arvioivat oppilaidensa koulusopeutumista ensimmäisen kouluvuoden lopulla. Saatiin seuraavat tulokset:

	Sopeutunut	
	Hyvin	Huonosti
Tyttö	90	20
Poika	125	45

Testaa tilastollisesti tasolla 0.05, että onko hyvin sopeutuneiden oppilaiden suhteellinen osuus erilainen tyttöjen ja poikien joukossa. Muista mainita mikä on hypoteesiparisi, sekä ilmoita testin p-arvo.