AALTO-YLIOPISTO, PERUSTIETEIDEN KORKEAKOULU
Matematiikan ja systeemianalyysin laitos

Mat-2.2103 Koesuunnittelu ja tilastolliset mallit

Seppälä

Tentti, 11.5.2015

Kirjoita selvästi jokaiseen koepaperiin:

- kurssin koodi ja nimi
- opiskelijanumero, TEKSTATEN sukunimi, etunimet
- koulutusohjelma, vuosikurssi
- allekirjoitus

Sallitut apuvälineet: YO-lautakunnan hyväksymä laskin, Mellinin kaava- ja taulukkokokoelma.

---

1. Vastaa seuraaviin väittämiin Kyllä tai Ei (perusteluja ei vaadita). Oikeasta vastauksesta 1p, väärästä vastauksesta -1p ja tyhjästä 0p. Jos tehtävän kokonaispisteet jäävät alle nollan, niin tehtävää ei huomioida arvostelussa.

   (a) Jos vasteeseen vaikuttaa kaksi tekijää A ja B, joilla molemmilla on kolme tasoa, niin pää- ja yhdysvaikutusten testaamiseen riittää yhdeksän havaintoa.

   (b) Tutkitaan faktorien A, B, C ja D vaikutusta vasteeseen. Suunnitellaan 2^{4-1} osafaktorikoe, jonka määrittelevä relaatio on $I = ABCD$. Käsittelykombinaatio b on mukana tässä koesuunnitelmassa.

   (c) Edellisen kohdan koeasetelman resoluutio on 4.

   (d) Testin p-arvo voi olla lähellä nollaa, vaikka testisuureen arvo olisi todella suuri.

   (e) Jos yhden selittäjän lineaarisessa regressiomallissa vasteen ja selittäjän välinen korrelaatio on positiivinen, niin ainakin toinen regressiokertoimista on positiivinen.

   (f) Nollahypoteesi saatetaan joissain tapauksissa hylätä, vaikka se olisi totta.

2. Meijerissä verrataan kolmea pesuliuosta (A,B,C). Neljänä päivänä valitaan satunnaisesti kolme maitosäiliötä, joista kukin pestään yhdellä liuoksella. Pesun jälkeen säiliöistä mitataan bakteerien määrä. Tulokset:

	Päivä			
	1	2	3	4
liuos A	6	8	7	6
liuos B	4	11	5	2
liuos C	10	14	10	8

   Testaa merkitsevyystasolla 0.05, onko liuosten tehokkuudessa eroja.

   Aputulos: Kaikkien havaintojen neliöiden summa = 811, kaikkien havaintojen summan neliö = 8281.

Käännä

3. Tietyn lisäaineen pitoisuus mustaviinimarjahyytelössä ei saa ylittää arvoa 35 mg/kg. Suuresta hyytelöerästä otetaan 25 näytettä, joiden lisäainepitoisuudet mitataan. Oletetaan mittaustulosten noudattavan normaalijakaumaa, jonka odotusarvo on todellinen pitoisuus θ ja keskihajonta 4 mg/kg (tunnetaan historiadatan perusteella). Testataan siis hypoteeseja

$$\begin{aligned} H_0 &: \quad \theta = 35 \\ H_1 &: \quad \theta > 35 \end{aligned}$$

laskemalla otoksen aritmeettinen keskiarvo $\overline{x}$.

a) Millä $\overline{x}$:n arvoilla H_0 hylätään, kun käytetään merkitsevyystasoa 0.1?

b) Mikä on testin voimakkuus, kun todellinen keskimääräinen lisäainepitoisuus on 38 mg/kg? (Siis millä todennäköisyydellä H_0 hylätään, kun $\theta = 38$?)

4. Kokeessa tutkitiin ruokavalion vaikutusta koehenkilöiden keskimääräiseen painonnousuun neljässä viikossa. Kussakin ruokavalioryhmässä oli 10 koehenkilöä. Kokeen tulokset on esitelty alla ovessa taulukossa siten, että $\bar{y}_i$ (kg) on keskimääräinen painonnousu ja s_i^2 (kg^2) painonnousun otosvarianssi ruokavalioryhmässä i.

	$\bar{y}_i$	s_i^2	n_i
1. Kaikkiruokaiset	2.03	0.21	10
2. Ei punaista lihaa	1.58	0.26	10
3. Ei kalaa	1.93	0.15	10
4. Kasvissyöjät	1.26	0.12	10

a) Olkoon $C_1 = \frac{1}{3}(\mu_1+\mu_2+\mu_3)-\mu_4$. Osoita, että C_1 on kontrasti ja testaa nollahypoteesia $H_0: \ C_1 = 0$ merkitsevyystasolla 0.05. Vinkki: $MSE = \sum_i (n_i-1)s_i^2/(N-k)$. Minkä tulkinnan annat nollahypoteesille?

b) Kontrastit $C = \sum c_i\mu_i$ ja $\Gamma = \sum d_i\mu_i$ ovat ortogonaalisia, jos $\sum_i (c_i d_i)/n_i = 0$. Osoita, että C_1 on ortogonaalinen kontrastin $C_2 = \mu_1 - \frac{1}{2}(\mu_2+\mu_3)$ kanssa. Minkä tulkinnan annat nollahypoteesille $H_0: \ C_2 = 0$?

c) Määritä kontrasti C_3, joka on ortogonaalinen kontrastien C_1 ja C_2 kanssa. Minkä tulkinnan annat nollahypoteesille $H_0: \ C_3 = 0$?

5. Tutkitaan puuvillan osuuden vaikutusta synteettisen kuidun vetolujuuteen siten, että osuus on 15%, 25% tai 30%. Tutkimuksessa mitataan viiden kustakin kuitutyypistä valmistetun kangaspalan vetolujuus. Osa mittauksista on kuitenkin selvästi virheellisiä, ja ne on päätetty jättää pois analyysista. Käyttökelpoiset koetulokset on esitetty alla olevassa taulukossa.

Puuvillan osuus		
15%	25%	35%
2	9	2
2	13	10
10	14	5
6		6
4		

Voidaan olettaa, että eri kuitutyyppeihin liittyvien havaintojen varianssit ovat yhtä suuret. Testaa 5 % merkitsevyystasolla, onko puuvillan osuudella merkitystä kuidun vetolujuuteen.

Aputulos: Kaikkien havaintojen neliöiden summa = 771, kaikkien havaintojen summan neliö = 6889.

Kaavoja

Yksisuuntainen varianssianalyysi

Neliösumma	Vapausasteet
$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{1}{N}T^2$	$N-1$
$SSG = \sum_{i=1}^{k} \frac{1}{n_i}T_i^2 - \frac{1}{N}T^2$	$k-1$
SSE	$N-k$

Kontrastien testaus

Hypoteesit:

$$H_0 : \Gamma = \sum_{i=1}^{k} c_i\mu_i = 0, \qquad H_1 : \Gamma = \sum_{i=1}^{k} c_i\mu_i \neq 0$$

t-testisuure:

$$t = \frac{\sum_{i=1}^{k} c_i\bar{y}_i}{\sqrt{MSE\sum_{i=1}^{k}\frac{c_i^2}{n_i}}}$$

Jos H_0 pätee, niin $t \sim t(N-k)$.

Kaksisuuntainen varianssianalyysi

Neliösumma	Vapausasteet
$SST = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} y_{kij}^2 - \frac{1}{IJK}T_{\cdot\cdot\cdot}^2$	$IJK-1$
$SSA = \frac{1}{JK}\sum_{i=1}^{I} T_{\cdot i\cdot}^2 - \frac{1}{IJK}T_{\cdot\cdot\cdot}^2$	$I-1$
$SSB = \frac{1}{IK}\sum_{j=1}^{J} T_{\cdot\cdot j}^2 - \frac{1}{IJK}T_{\cdot\cdot\cdot}^2$	$J-1$
$SS = \frac{1}{K}\sum_{i=1}^{I}\sum_{j=1}^{J} T_{\cdot ij}^2 - \frac{1}{IJK}T_{\cdot\cdot\cdot}^2$	
$SSAB$	$(I-1)(J-1)$
SSE	$IJ(K-1)$

$$SS = SSA + SSB + SSAB$$

Latinalaisten neliöiden koeasetelma

Neliösumma	Vapausasteet
$SST = \sum_{i=1}^{P}\sum_{j=1}^{P}\sum_{k=1}^{P} y_{ijk}^2 - \frac{1}{P^2}T_{...}^2$	$P^2 - 1$
$SSA = \frac{1}{P}\sum_{k=1}^{P} T_{..k}^2 - \frac{1}{P^2}T_{...}^2$	$P - 1$
$SSR = \frac{1}{P}\sum_{i=1}^{P} T_{i..}^2 - \frac{1}{P^2}T_{...}^2$	$P - 1$
$SSC = \frac{1}{P}\sum_{j=1}^{P} T_{.j.}^2 - \frac{1}{P^2}T_{...}^2$	$P - 1$
SSE	$(P-2)(P-1)$

Satunnaistettu täydellinen lohkoasetelma

Neliösumma	Vapausasteet
$SST = \sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}^2 - \frac{1}{IJ}T_{..}^2$	$IJ - 1$
$SSA = \frac{1}{J}\sum_{i=1}^{I} T_{i.}^2 - \frac{1}{IJ}T_{..}^2$	$I - 1$
$SSB = \frac{1}{I}\sum_{j=1}^{J} T_{.j}^2 - \frac{1}{IJ}T_{..}^2$	$J - 1$
SSE	$(I-1)(J-1)$

2^2-faktorikoe

Neliösumma	Vapausasteet
$SST = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} y_{kij}^2 - 4n\bar{y}_{...}^2$	$4n - 1$
$SSA = \frac{1}{4n}(ab + a - b - (1))^2$	1
$SSB = \frac{1}{4n}(ab - a + b - (1))^2$	1
$SSAB = \frac{1}{4n}(ab - a - b + (1))^2$	1
SSE	$4(n-1)$