Инжењерска статистика
Вежба 7

Постављање модела - двофакторска

$$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + e_{ijk}, i = 1, a, j = 1, b, k = 1, n$$

Системи хипотеза

$$H_0\left(\tau_i = 0, i = 1...a\right) \qquad H_1\left(\exists \tau_i \neq 0, i = 1...a\right)$$

$$H_0\left(\beta_j = 0, j = 1...b\right) \qquad H_1\left(\exists \beta_j \neq 0, j = 1...b\right)$$

$$H_0\left((\tau\beta)_{ij} = 0, i = 1...a, j = 1...b\right) \qquad H_1\left(\exists (\tau\beta)_{ij} \neq 0, i = 1...a, j = 1...b\right)$$

Поступак:

1. Суме квадрата
2. Број степени слободе
3. Оцене варијанси
4. Прорачунска вредност F теста и однос оцена варијанси
5. Формирање ANOVA табеле

Суме квадрата

$$T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk} \qquad N = a \cdot b \cdot n$$

$$SK_T = SK_A + SK_B + SK_{AB} + SK_e$$

$$SK_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{T^2}{N}$$

$$SK_A = \frac{1}{bn}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N}$$

$$SK_B = \frac{1}{an}\sum_{j=1}^{b}\left(\sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N}$$

$$SK_{AB} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N} - SK_A - SK_B$$

$$SK_e = SK_T - SK_A - SK_B - SK_{AB}$$

Број степени слободе

$$f_T = N - 1$$

$$f_A = a - 1$$

$$f_B = b - 1$$

$$f_{AB} = f_A \cdot f_B$$

$$f_e = f_T - f_A - f_B - f_{AB}$$

Оцена варијансе

$$S_A^2 = \frac{SK_A}{f_A}$$

$$S_B^2 = \frac{SK_B}{f_B}$$

$$S_{AB}^2 = \frac{SK_{AB}}{f_{AB}}$$

$$S_e^2 = \frac{SK_e}{f_e}$$

$$\Rightarrow$$

Прорачунска вредност *F*-теста

$$F_{0A} = \frac{S_A^2}{S_e^2}$$

$$F_{0B} = \frac{S_B^2}{S_e^2}$$

$$F_{0AB} = \frac{S_{AB}^2}{S_e^2}$$

Таблична вредност F-теста

$$F_{\alpha,f_A,f_e}, F_{\alpha,f_B,f_e}, F_{\alpha,f_{AB},f_e}$$

Одговарајући фактор је утицајан ако је: $F_{\alpha,f_A,f_e} \vee F_{\alpha,f_B,f_e} \vee F_{\alpha,f_{AB},f_e} < F_0$

Табела анализе варијансе

Извор варијације	Сума квадрата	Број степени слободе	Оцена варијансе	F
Фактор А	SK_A	f_A	S_A^2	F_{0A}
Фактор Б	$SK_Б$	$f_Б$	$S_Б^2$	$F_{0Б}$
Интеракција АБ	$SK_{АБ}$	$f_{АБ}$	$S_{АБ}^2$	$F_{0АБ}$
Грешка	SK_e	f_e	S_e^2	
Укупно	SK_T	f_T		

Пример

Сматра се да на број неисправних производа који се производе у одређеном производном процесу, утичу материјал и смена у којој радници раде. Спроведен је експеримент са ова два фактора. Мерен је број неисправних комада у случајно изабраним серијама величине 1000 комада. Испитивана су 3 могућа материјала (фактор А), а рад је вршен у 3 смене. Резултати су приказани у табели:

А1			А2			А3		
Б1	Б2	Б3	Б1	Б2	Б3	Б1	Б2	Б3
18	34	45	16	15	45	22	76	44
2	14	66	2	16	48	33	43	47
3	34	42	3	34	52	23	34	25

Испитивани фактори: *А* – врста материјала *Б* - смена

Број нивоа фактора: *a*=3 *b*=3

Број мерења у оквиру нивоа (репликација): *n*=3

Тип експеримента: балансиран

$$N = a \cdot b \cdot n = 3 \cdot 3 \cdot 3 = 27$$

$$y_{11} = 18 + 2 + 3 = 23 \qquad y_{12} = 82 \qquad y_{13} = 153$$

$$y_{21} = 16 + 2 + 3 = 21 \qquad y_{22} = 65 \qquad y_{23} = 145$$

$$y_{31} = 22 + 33 + 23 = 78 \qquad y_{32} = 153 \qquad y_{33} = 116$$

$$y_{1j} = 23 + 82 + 153 = 258$$

$$y_{2j} = 21 + 65 + 145 = 231$$

$$y_{3j} = 78 + 153 + 116 = 347$$

$$y_{i1} = 122 \qquad y_{i2} = 300 \qquad y_{i3} = 414$$

$$T = 258 + 231 + 347 = 836$$

$$SK_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{T^2}{N} = 9696{,}963$$

$$SK_A = \frac{1}{bn}\sum_{i=1}^{a}\left(\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N} = 818{,}74$$

$$SK_B = \frac{1}{an}\sum_{j=1}^{b}\left(\sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N} = 4812{,}74$$

$$SK_{AB} = \frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\sum_{k=1}^{n} y_{ijk}\right)^2 - \frac{T^2}{N} - SK_A - SK_B =$$

$$= 1584{,}189$$

$$SK_e = SK_T - SK_A - SK_B - SK_{AB} =$$

$$= 2481{,}3341$$

$$f_T = N - 1 = 26 \qquad f_{AB} = (a-1)(b-1) = 4$$

$$f_A = a - 1 = 2 \qquad f_e = f_T - f_A - f_B - f_{AB} = 18$$

$$f_B = b - 1 = 2$$

Оцене варијанси

$$S_A^2 = \frac{SK_A}{f_A} = \frac{818,74}{2} = 414,37 \qquad S_B^2 = \frac{SK_B}{f_B} = \frac{4812,74}{2} = 2406,37$$

$$S_{AB}^2 = \frac{SK_{AB}}{f_{AB}} = \frac{1584,189}{4} = 396,04 \qquad S_e^2 = \frac{SK_e}{f_e} = \frac{2481,334}{18} = 137,85$$

Табела анализе варијансе

Извор варијације	Сума квадрата	Број степени слободе	Оцена варијансе	F
Фактор А	SK_A	f_A	S_A^2	5.939
Фактор Б	$SK_Б$	$f_Б$	$S_Б^2$	17.45
Интеракција АБ	$SK_{АБ}$	$f_{АБ}$	$S_{АБ}^2$	11.50
Грешка	SK_e	f_e	S_e^2	
Укупно	SK_T	f_T		

$$F_{\alpha,f_A,f_e} = F_{0.05,2,18} = 3.55$$

$$F_{\alpha,f_A,f_e} = F_{0.05,2,18} = 3.55$$

$$F_{\alpha,f_A,f_e} = F_{0.05,4,18} = 2.93$$

Непараметарски тестови

Непараметарски тестови се користе када:

- **се доказује да ли узорак одговара некој расподели**
  - **Hi квадрат тест**
  - **Тест Kolmogorova**
- **поређење две расподеле без обзира на њихов облик**
  - **Тест Kolmogorov-Smirnova**
  - **U*-тест Mann-Withney-а**
  - **Тест медијане**
  - **Тест разлика медијана**

Тест Kolmogorova-бројеви

Тестирати хипотезу да је узорак од 20 елемената узет из популације са нормалном расподелом. Вредности појединих мерења приказане су у табели:

0.464	0.137	2.455	-0.323	-0.068
0.906	-0.513	-0.525	0.595	0.881
-0.482	1.678	-0.057	-1.229	-0.486
-1.787	-0.261	1.237	1.046	-0.508

Процедура тестирања хипотеза

1. Идентификација расподеле по којој се подаци распоређују ако претпоставка расподеле није дата конструише се хистограм. Задатком је дата претпоставка о нормалној расподели.

2. Постављање система хипотеза

$$H_0\left(F_n(x) = F_0(x)\right)$$

$$H_1\left(F_n(x) \neq F_0(x)\right)$$

3. Праг значајности: усваја се $\alpha = 0.05$

4. Статистика одлучивања: $D_n = \max_x \left|F_n(x) - F_0(x)\right|$

5. Критична област одбацивања H_0 $\lambda_0 > \lambda_\alpha$

6. Прорачун

6.1. Сортирање података у растућем поретку

6.2. Одређивање емпиријских вредности на основу таблица

$$\forall x_i \in (-4,+4) \Rightarrow \quad F_{0i} = \Phi(x_i) \sim \mathcal{N}(0,1^2)$$

нпр. $x_1 = -1.787 \Rightarrow \Phi(-1.79) = 0.037$

6.3. Одређивање теоријских вредности на основу расподеле

$$F_n(x) = \begin{cases} 0, & x < x_1 \\ \dfrac{k}{n}, & x_k < x < x_{k+1} \\ 1, & x_n < x \end{cases}$$

нпр. $F_n(x_1) = \dfrac{1}{n} = \dfrac{1}{20} = 0.05$

6.4. Одређивање прорачунске вредности статистике одлучивања

$$D_i = \left|F_n(x_i) - F_0(x_i)\right|$$

нпр. $D_1 = |0.05 - 0.037| = 0.013$

6.5. Одређивање D_n

$$D_n = \max_x \left| F_n(x) - F_0(x) \right|$$

$$D_n = 0.150$$

$$\lambda_0 = D_n \sqrt{n} = 0.67$$

$$\lambda_\alpha = \lambda_{0.05} = 1.355$$

7. Одлука

$$\lambda_\alpha > \lambda_0 \Leftrightarrow 1.355 > 0.67$$

M. Нулта хипотеза се не може одбацити.

I. Подаци се распоређују по нормалној расподели.

	x_i	$F_0(x)$	$F_n(x)$	$\lvert F_n(x) - F_0(x) \rvert$
1	-1.787	0.037	0.05	0.013
2	-1.229	0.110	0.10	0.01
3	-0.525	0.300	0.15	0.150
4	-0.513	0.304	0.20	0.104
5	-0.508	0.306	0.25	0.056
6	-0.486	0.315	0.30	0.015
7	-0.482	0.315	0.35	0.035
8	-0.323	0.373	0.40	0.027
9	-0.261	0.397	0.45	0.053
10	-0.068	0.473	0.50	0.027
11	-0.057	0.477	0.55	0.073
12	0.137	0.554	0.60	0.046
13	0.464	0.679	0.65	0.029
14	0.595	0.724	0.70	0.024
15	0.881	0.811	0.75	0.061
16	0.906	0.818	0.80	0.018
17	1.046	0.852	0.85	0.002
18	1.237	0.892	0.90	0.008
19	1.678	0.953	0.95	0.003
20	2.455	0.993	1	0.007

Домаћи задатак 5

- Урадити одговарајући задатак из фајла окаченог на сајту предмета