

## האוניברסיטה העברית בירושלים המחלקה לסטטיסטיקה

מודלים סטטיסטיים ויישומיהם (52518)

פרופ' דוד צוקר

תשע"ז סמסטר א', מועד א'

תאריך: י"ב בשבט תשע"ז, 8.2.17

משך הבחינה: שעתיים וחצי

חומר עזר: מחשב כיס ודפי רשימות (שני דפים בגודל A4, שני צדדים)

חל איסור מוחלט להעתיק. תלמיד שייתפס יורחק לשנה מלימודיו.

בהצלחה!!

שאלה 1 (20 נקודות)

נתייחס לנתונים דלהלן. נניח כי הנתונים מתנהגים לפי המודל

$$Y_{ij} = \mu_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \text{ ב"ת}$$

חשבו סטטיסטי  $F$  לבדוק את ההשערה  $H_0: \mu_1 = \mu_2 \text{ and } \mu_3 = \mu_4$ . יש להגיע לתשובה מספרית.

ציינו את דרגות החופש של הסטטיסטי.

הנתונים:

GROUP		N	mean	SD
1		40	4.04	8.79
2		110	3.67	7.44
3		80	0.46	7.83
4		60	0.83	8.12

## שאלה 2 (20 נקודות)

נתייחס למודל של ניתוח שונות דו-כיווני ללא אינטראקציה :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

$$\varepsilon_{ijk} \sim N(0, \sigma^2) \text{ ב"ת, } \sigma^2 \text{ לא ידוע}$$

$$i = 1, \dots, 3; j = 1, \dots, 2; k = 1, \dots, n_{ij}$$

$$\sum_{i=1}^3 \alpha_i = 0, \sum_{j=1}^2 \beta_j = 0$$

נניח הנתונים הרשומים להלן מתנהגים לפי המודל הנ"ל.

קחו כידוע שה- MSE שווה ל- 147.125.

[10] א. חשבו ר"ס ברמה 95% עבור  $\alpha_1 - \alpha_2$  כקונטרסט יחיד.

[10] ב. חשבו ר"ס בו-זמניים ברמה 95% עבור  $\alpha_1 - \alpha_2$  ו-  $\alpha_1 - \alpha_3$ .

יש להשתמש בלוח של ערכים קריטיים הנמצא בנספח.

הנתונים :

$i$	$j$	התצפיות
1	1	92,80
1	2	65,75
2	1	86,91
2	2	56,69
3	1	56,34
3	2	57,54

שאלה 3 (20 נקודות)

להלן נתונים ממחקר עם משתנה מוסבר בינארי  $Y$  (0/1) ומשתנה מסביר (יחיד)  $X$  עם שלושה ערכים אפשריים: 0,1,2. גודל המדגם כולו הינו 300. נעבוד עם המודל  $\Pr(Y=1|X=x) = G(\beta_0 + \beta_1 x)$  כאשר  $G(u) = e^u / (1 + e^u)$ . רשמו את משוואות הנראות עבור  $\beta_0, \beta_1$ . החל מהערכים ההתחלתיים  $\hat{\beta}_0^{(0)} = 1.3, \hat{\beta}_1^{(0)} = 0.8$ , בצעו צעד אחד של שיטת ניוטון-רפסון לקראת פתרון המשוואות.

הנתונים:

		מס' תצפיות עם $Y=0$	מס' תצפיות עם $Y=1$
$X=0$		70	30
$X=1$		40	60
$X=2$		15	85

שאלה 4 (20 נקודות)

להלן פלט מהרצת מודל ניתוח שונות חד-כיווני עם אפקט מקרי על נתונים ממחקר עם מספר בתי ספר ומספר תלמידים בכל בית ספר. נזכור כי המודל הינו

$$Y_{ij} = \mu + a_i + \epsilon_{ij}, \quad a_i \sim N(0, \sigma_a^2), \quad \epsilon_{ij} \sim N(0, \sigma_\epsilon^2), \quad \text{all random variables independent}$$

חשבו ר"ס עבור המתאם (קורלציה) בין שני תלמידים באותו בית ספר.

פלט

$$\hat{\mu} = 2.54, \quad \text{Var}(\hat{\mu})^{1/2} = 0.30$$

$$\hat{\sigma}_a^2 = 1.69, \quad \hat{\sigma}_\epsilon^2 = 1.24$$

Estimated Asymptotic Covariance Matrix of  $(\hat{\sigma}_a^2, \hat{\sigma}_\epsilon^2)$ :

$$\begin{pmatrix} 0.32 & -0.00026 \\ -0.00026 & 0.0066 \end{pmatrix}$$

שאלה 5 (20 נקודות)

נערך מחקר עם 4 משתנים קטגוריאליים A,B,C,D. נגדיר

$$\psi_{kl}^{(i_1, i_2, j_1, j_2)} = \frac{\Pr(B = j_2 | A = i_2, C = k, D = l) / \Pr(B = j_1 | A = i_2, C = k, D = l)}{\Pr(B = j_2 | A = i_1, C = k, D = l) / \Pr(B = j_1 | A = i_1, C = k, D = l)}$$

נניח כי מתקיים המודל הלוג-ליניארי  $(AB, AC, AD, BCD)$ . הראו כי  $\psi_{kl}^{(i_1, i_2, j_1, j_2)}$  לא תלוי ב-  $k$  או  $l$ .

## נספח א'

לוח של ערכים של  $c$  כך ש-  $\Pr(|T_1| \leq c, |T_2| \leq c) = 0.95$

באשר  $(T_1, T_2)$  מפולג רב- $t$  עם 8 ד"ח ומתאם  $\rho$

$\rho$	$c$
0.00	2.7182
0.05	2.7177
0.10	2.7165
0.15	2.7144
0.20	2.7115
0.25	2.7077
0.30	2.7029
0.35	2.6972
0.40	2.6903
0.45	2.6824
0.50	2.6731
0.55	2.6625
0.60	2.6501
0.65	2.6359
0.70	2.6193
0.75	2.5999
0.80	2.5766
0.85	2.5840
0.90	2.5111
0.95	2.4581
1.00	2.3064