

האוניברסיטה העברית בירושלים

המחלקה לסטטיסטיקה

מודלים סטטיסטיים ויישומיהם (52518)

פרופ' דוד צוקר

תשע"ו סמסטר א', מועד ב'

תאריך: ה' בניסן תשע"ו, 13.4.16

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 (25 נקודות)

נתייחס לנתונים דלהלן. נניח כי הנתונים מתנהגים לפי המודל

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

חשבו סטטיסטי F לבדוק את השערת האפס הבאה:

$$H_0: \mu_3 - \mu_1 = 2(\mu_2 - \mu_1) \text{ and } \mu_4 - \mu_1 = 4(\mu_2 - \mu_1)$$

יש להגיע לתשובה מספרית.

בנוסף, ציינו (בצורה מלאה) את התפלגותו של הסטטיסטי תחת H_0 .

רמז: תחת H_0 ניתן לבטא כל אחד מה- μ_i במונחים של שני פרמטרים $\beta_1^{(0)}, \beta_2^{(0)}$. למשל, אפשר לקחת

$\beta_1^{(0)} = \mu_1$. אז, ניתן לרשום את המודל בצורה $Y^{(0)} = X^{(0)}\beta^{(0)} + \epsilon$, עבור מטריצה $X^{(0)}$ מתאימה.

עליכם להגדיר בצורה נכונה את $X^{(0)}$ ו- $\beta^{(0)}$.

הנתונים:

GROUP	N	mean	SD
1	49	4.04	8.79
2	48	-0.46	7.83
3	48	-3.67	7.44
4	53	-8.89	9.86

שאלה 2 (10 נקודות)

להלן תצפיות ממדגם מקרי פשוט:

1.81, -0.55, 2.50, 0.92, 1.19, -1.12, -1.03, 0.44

חשבו ושרטטו את פונקצית ההתפלגות האמפירית \hat{F}_n .

שאלה 3 (25 נקודות)

נתייחס לנתונים הבאים (המספרים הינם תצפיות Y_{ijk}).

	Age ≤ 45	Age > 45
0 mg	12,13	11,12
50 mg	9,10	16,18
100 mg	15,16	17,16
150 mg	7,18	12,18

להלן לוח הממוצעים \bar{Y}_{ij} :

	Age ≤ 45	Age > 45
0 mg	12.5	11.5
50 mg	9.5	17
100 mg	15.5	16.5
150 mg	12.5	15

בנוסף,

$$\sum_{i=1}^4 \sum_{j=1}^2 \sum_{k=1}^2 Y_{ijk}^2 = 3206$$

ננתח את הנתונים האלה באמצעות ניתוח שונות דו-כיווני ללא אינטראקציה.

א. נסחו המודל בצורה $Y = X\eta + \epsilon$, עם ההגדרות המתאימות של X ו- η . וחשבו המטריצה $(X^T X)^{-1}$. השתמשו במשקולות $\pi_i = 1/I \quad \forall i$ ו- $\tau_j = 1/J \quad \forall j$. שימו לב, $(X^T X)^{-1}$ מטריצה סימטרית. [6]

תזכורת: נוסחת הפיכת מטריצה.

$$(A^{-1})_{ji} = (-1)^{i+j} \det(A_{ij}) / \det(A)$$

כאשר, A_{ij} הינה המטריצה המתקבלת ממחיקת השורה i והעמודה j במטריצה A .

ב. מצאו את האומדים של α_i , עבור $i = 1, \dots, 4$. [3]

ג. חשבו רווח סמך בו-זמני (ברמת ביטחון של 95%) לפי שיטת שאפה, עבור הקונטראסט הבא.

$$\psi(\alpha) = \frac{\alpha_3 + \alpha_4}{2} - \alpha_1$$

(הניחו כי הקונטראסט אינו נקבע מראש). [16]

שאלה 4 (25 נקודות)

במסגרת של מודלים לוג-ליניאריים, נגדיר

$$\Psi_{ii'jj'}^k = \left(\frac{\Pr(B = j' | A = i', C = k)}{\Pr(B = j | A = i', C = k)} \right) / \left(\frac{\Pr(B = j' | A = i, C = k)}{\Pr(B = j | A = i, C = k)} \right)$$

הוכיחו כי המודל (AB, AC, BC) שקול למצב בו $\Psi_{ii'jj'}^k$ לא תלוי ב- k , ז"א מנת יחס הסיכויים של כל שני משתנים אינה תלויה בערכו של המשתנה השלישי.

שאלה 5 (15 נקודות)

נתייחס לחומר בנספח 1 עבור הרצה של רגרסיה לוגיסטית.

נסמן ב- x_1 את המצב $\text{term} = 24, \text{amount} = 10000, \text{age} = 30, \text{land} = 0$.

נסמן ב- x_2 את המצב $\text{term} = 36, \text{amount} = 15000, \text{age} = 40, \text{land} = 1$.

בנוסף, עבור כל x אפשרי נסמן $p(x) = \Pr(Y = 1 | X = x)$.

א. חשבו רווח סמך (ברמת ביטחון של 95%) עבור $p(x_1)$.

ב. חשבו רווח סמך (ברמת ביטחון של 95%) עבור $\frac{p(x_2)}{1-p(x_2)} / \frac{p(x_1)}{1-p(x_1)}$.

נספח 1

Data on 1000 loan customers of a German bank.

The variables are as follows:

gc: 0/1 indicator of whether the person is a good risk for the requested loan (1) or not 0)

term: duration of requested loan

amount: amount requested (in German Marks)

age: age of the customer (years)

land: 0/1 indicator of whether the customer owns land (1) or does not (0)

We apply a log transformation to term, amount, and age before running the regression.

```
lterm = log(term)
```

```
amtlog = log(amount)
```

```
l.age = log(age)
```

Results of Logistic Regression Analysis

	Coef	S.E.	Wald	Z	Pr(> Z)
Intercept	-1.1179	1.1539	-0.97	0.3327	
lterm	-0.9297	0.1714	-5.43	<0.0001	
amtlog	0.1870	0.1230	1.52	0.1284	
l.age	0.9560	0.2565	3.73	0.0002	
land	-0.7341	0.1980	-3.71	0.0002	

Covariance Matrix of Estimated Coefficients

	Intercept	lterm	amtlog	l.age	land
Intercept	1.33159839	0.0137349459	-0.0752634941	-0.2258012175	0.0688971954
lterm	0.01373495	0.0293662476	-0.0134884945	0.0013452636	-0.0009395372
amtlog	-0.07526349	-0.0134884945	0.0151238319	-0.0007412548	-0.0036072797
l.age	-0.22580122	0.0013452636	-0.0007412548	0.0658113459	-0.0128218795
land	0.06889720	-0.0009395372	-0.0036072797	-0.0128218795	0.0392093907

נספח 2

לוח של התפלגות F

ערכים קריטיים עבור $\alpha = 0.05$ בהתפלגות F עם d_1 ד"ח במונה ו- d_2 ד"ח במכנה

$d_2 \backslash d_1$	1	2	3	4	5	6	7	8	9	10
1	161.45	18.51	10.13	7.71	6.61	5.99	5.59	5.32	5.12	4.96
2	199.50	19.00	9.55	6.94	5.79	5.14	4.74	4.46	4.26	4.10
3	215.71	19.16	9.28	6.59	5.41	4.76	4.35	4.07	3.86	3.71
4	224.58	19.25	9.12	6.39	5.19	4.53	4.12	3.84	3.63	3.48
5	230.16	19.30	9.01	6.26	5.05	4.39	3.97	3.69	3.48	3.33
6	233.99	19.33	8.94	6.16	4.95	4.28	3.87	3.58	3.37	3.22
7	236.77	19.35	8.89	6.09	4.88	4.21	3.79	3.50	3.29	3.14
8	238.88	19.37	8.85	6.04	4.82	4.15	3.73	3.44	3.23	3.07
9	240.54	19.38	8.81	6.00	4.77	4.10	3.68	3.39	3.18	3.02
10	241.88	19.40	8.79	5.96	4.74	4.06	3.64	3.35	3.14	2.98
11	242.98	19.40	8.76	5.94	4.70	4.03	3.60	3.31	3.10	2.94
12	243.91	19.41	8.74	5.91	4.68	4.00	3.57	3.28	3.07	2.91
13	244.69	19.42	8.73	5.89	4.66	3.98	3.55	3.26	3.05	2.89
14	245.36	19.42	8.71	5.87	4.64	3.96	3.53	3.24	3.03	2.86
15	245.95	19.43	8.70	5.86	4.62	3.94	3.51	3.22	3.01	2.85
16	246.46	19.43	8.69	5.84	4.60	3.92	3.49	3.20	2.99	2.83
17	246.92	19.44	8.68	5.83	4.59	3.91	3.48	3.19	2.97	2.81
18	247.32	19.44	8.67	5.82	4.58	3.90	3.47	3.17	2.96	2.80
19	247.69	19.44	8.67	5.81	4.57	3.88	3.46	3.16	2.95	2.79
20	248.01	19.45	8.66	5.80	4.56	3.87	3.44	3.15	2.94	2.77
21	248.31	19.45	8.65	5.79	4.55	3.86	3.43	3.14	2.93	2.76
22	248.58	19.45	8.65	5.79	4.54	3.86	3.43	3.13	2.92	2.75
23	248.83	19.45	8.64	5.78	4.53	3.85	3.42	3.12	2.91	2.75
24	249.05	19.45	8.64	5.77	4.53	3.84	3.41	3.12	2.90	2.74
25	249.26	19.46	8.63	5.77	4.52	3.83	3.40	3.11	2.89	2.73
26	249.45	19.46	8.63	5.76	4.52	3.83	3.40	3.10	2.89	2.72
27	249.63	19.46	8.63	5.76	4.51	3.82	3.39	3.10	2.88	2.72
28	249.80	19.46	8.62	5.75	4.50	3.82	3.39	3.09	2.87	2.71
29	249.95	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.87	2.70
30	250.10	19.46	8.62	5.75	4.50	3.81	3.38	3.08	2.86	2.70
31	250.23	19.46	8.61	5.74	4.49	3.80	3.37	3.07	2.86	2.69
32	250.36	19.46	8.61	5.74	4.49	3.80	3.37	3.07	2.85	2.69
33	250.48	19.47	8.61	5.74	4.48	3.80	3.36	3.07	2.85	2.69
34	250.59	19.47	8.61	5.73	4.48	3.79	3.36	3.06	2.85	2.68
35	250.69	19.47	8.60	5.73	4.48	3.79	3.36	3.06	2.84	2.68
36	250.79	19.47	8.60	5.73	4.47	3.79	3.35	3.06	2.84	2.67
37	250.89	19.47	8.60	5.72	4.47	3.78	3.35	3.05	2.84	2.67
38	250.98	19.47	8.60	5.72	4.47	3.78	3.35	3.05	2.83	2.67
39	251.06	19.47	8.60	5.72	4.47	3.78	3.34	3.05	2.83	2.66
40	251.14	19.47	8.59	5.72	4.46	3.77	3.34	3.04	2.83	2.66