

האוניברסיטה העברית בירושלים
המחלקה לסטטיסטיקה

מודלים סטטיסטיים ויישומיהם (52518)
פרופ' דוד צוקר
תשע"ב סמסטר א', מועד ב'

תאריך: כ"ד באדר תשע"ב, 18.3.12

משך הבחינה: שעתיים וחצי

חומר עזר: מחשב CIS ודף רשימות (שני דפים בגודל A4, שני צדדים)

חל איסור מוחלט להעתיק. תלמיד שיתפס יורחക לשנה מלימודיו.

בצלחה!!

שאלה 1 (25 נקודות)

להלן נתונים מסקר בנושא מידת התלונות מפורטים. המשתנים הינם התלונות (נמוכה/בינונית/גבוהה), גיל (זכר/UMB), ומין (זכר/נקבה). חשבו סטטיסטי מתאים לבדוק את ההשערה שהتلונות בלתי תלויות במין וגיל. יש להציג את הערך המספרי של הסטטיסטי והתפלגותו תחת השערת האפס. ניתן להסתמך על תוצאות שהוכחו בשיעורים או תרגילים.

הנתונים:

גיל

התלונות	מין	UMB	זכר	גיל
נמוכה	זכר	20	5	3
	נקבה	3	4	15
בינונית	זכר	20	20	15
	נקבה	15	16	10
גבוהה	זכר	30	20	12
	נקבה	20	15	12

שאלה 2 (25 נקודות)

בנוסף נמצא ניתוח רגרסיה לוגיסטיבית ב- SAS (תוכנית ופלט) על נתונים מחקר בנושא משקל לידי נמוֹך.

הערה: $\text{predval} = p(\mathbf{X}_i, \hat{\beta})$

[10] א. חשבו ר"ס ברמה 95% להסתברות שאשה שמשקלה 62 ק"ג בועסת האחרונה לפני ההריון ועננה במהלך ההריון תלד תינוק עם משקל לידי נמוֹך.

[15] ב. בעזרת התוצאות בפלט והשיטות שלמדנו בקורס, בדקו את טיב ההתאמה של המודל לנתונים.

שאלה 3 (25 נקודות)

נתיחס לניתוח שונות דו-כיווני עם $2 = J = I$. נניח שתאי אופציות הבאות לבחירת המשקלות:

א. $\pi_1 = 0.65, \pi_2 = 0.35, \tau_1 = 0.3, \tau_2 = 0.7$

ב. $\pi_1 = 0.5, \pi_2 = 0.5, \tau_1 = 0.55, \tau_2 = 0.45$

תנו דוגמא של μ -ים כך שלפי משקלות אי מתקיים $\alpha_2 < \alpha_1$ ולפי משקלות ב' מתקיים $\alpha_2 > \alpha_1$.

שאלה 4 (25 נקודות)

נתיחס לנתונים דלהלן. נניח כי הנתונים מתנהגים לפי המודל

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

חשבו סטטיסטי F מותאים לבדוק את השערת האפס הבאה:

$$H_0: \mu_4 - \mu_3 = \mu_3 - \mu_2 = \mu_2 - \mu_1$$

יש להגיע לתשובה מספרית. בנוסף, ציינו את דרגות החופש של הסטטיסטי F.

: 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

נספח

```
*****
Data on low birth weights. The variables we use are:
LOW      low birth weight (0 if weight >= 2500 g, 1 if weight < 2500 g).
LWT      weight in pounds of the mother at last menstrual period
SMOKE    smoking status during pregnancy ( = 1 if yes, = 0 if no)

*****
options ls=132 nocenter nodate;

data indat;
infile 'c:\zucker\binary\examples\birthdat.txt';
input id low age lwt race smoke pt1 ht iu ftv bwt;
weight = 0.45359237*lwt; * convert weight to kilograms;

proc logistic descending;
model low = weight smoke / covb;
output out=odat p=predval;
run;

* This block of code breaks up the dataset into 5 subgroups based on weight;
proc rank; var weight; ranks w_rank;
data newdat;
set;
grp = 1 + (weight>50) + (weight>55) + (weight>60) + (weight>70);

proc means;
class grp;
var weight predval low;
run;
```

The LOGISTIC Procedure

Model Information

Data Set	WORK.INDAT
Response Variable	low
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	188
Number of Observations Used	188

[SNIP]

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.6302	0.7947	0.6289	0.4278
weight	1	-0.0295	0.0134	4.8492	0.0277
smoke	1	0.7017	0.3256	4.6431	0.0312

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits
weight	0.971	0.946 0.997
smoke	2.017	1.065 3.819

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.8	Somers' D	0.289
Percent Discordant	35.0	Gamma	0.292
Percent Tied	1.2	Tau-a	0.125
Pairs	7611	c	0.644

Estimated Covariance Matrix

Parameter	Intercept	weight	smoke
Intercept	0.631488	-0.01025	-0.05424
weight	-0.01025	0.00018	0.000126
smoke	-0.05424	0.000126	0.106037

The MEANS Procedure

grp	Obs	Variable	N	Mean	Std Dev	Minimum	Maximum
1	53	weight	53	45.7529209	3.5727393	36.2873896	49.8951607
		predval	53	0.4070434	0.0929614	0.3009761	0.5648060
		low	53	0.4716981	0.5039755	0	1.0000000
2	42	weight	42	53.3403027	1.3883748	50.8023454	54.8846768
		predval	42	0.3369889	0.0769758	0.2709246	0.4581561
		low	42	0.2380952	0.4310805	0	1.0000000
3	31	weight	31	58.0159273	1.4737906	55.3382691	59.8741928
		predval	31	0.3026276	0.0748983	0.2428313	0.4251515
		low	31	0.3548387	0.4863735	0	1.0000000
4	29	weight	29	64.4257577	3.2471882	60.3277852	69.8532250
		predval	29	0.2736474	0.0746484	0.1928208	0.3896119
		low	29	0.2413793	0.4354942	0	1.0000000
5	33	weight	33	83.1036203	11.6539687	70.3068174	113.3980925
		predval	33	0.1804852	0.0609923	0.0693540	0.3222373
		low	33	0.1818182	0.3916747	0	1.0000000