

מודלים סטטיסטיים וישומיהם 52518

תאריך: 13.4.2008

המורה: פרופ' ש. אומן

משך הבחינה: שתיים וחצי, כולל הארכה של חצי שעה

חומר עזר: אך ורק מחשב כיס ודפי הנוסחאות שהולקו עם טופס הבחינה.

בהצלחה!

שאלה 1 (35 נקודות)

חקלאי מעוניין לבדוק האם הבדלים באיכות הקרקע משפיעים על היבול במטע מסוים של עצי אפרסק. לשם כך הוא בוחר, באופן מקרי, המישה עצים מהפרדס. מכל עץ הוא קוטף איבעה פירות בשלים ושוקל כל אחד. התוצאות (בגרמים) רשומות למטה.

עץ	1	2	3	4	5
1	150	210	150	120	100
2	200	120	170	140	120
3	100	160	130	130	90
4	140	130	140	160	140
ממוצע	147.50	155.00	147.50	137.50	112.50

- בדקו ברמת מובהקות 5%, את ההשערה שאיכות הקרקע אינה משפיעה על משקל היבול.
- מצאו חסם סמך עליון (95%) ליהס של השונות במשקל הפירות בין העצים, לשונותו בתוך העצים.

שאלה 2 (30 נקודות)

להלן נמצאות תוצאות דגימה של ארבעה משתנים איכותיים, a,b,c,d, כל אחד בעל שלוש רמות.

- א. הסבר את המשמעות של הערך 11.41 אשר מסומן בעיגול. כלומר, הגדר במדויק את H_0 הנבדקת ע"י סטטיסטי זה, הגדר את H_1 , תן את דגושה של הסטטיסטי, הסבר איך מחושבות דרגות החופש ודא שהן אכן שוות ל-16. מהי המסקנה האופראטיבית מה- p-value אשר התקבל?
- ב. אותן השאלות שב-א' לעיל, אבל עכשיו בהתייחס לערך 1.54 אשר מסומן בעיגול.

***** THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.
SIMULTANEOUS TEST THAT ALL K+1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.	ITER
0-MEAN	80	1427.65	.00000	1534.85	.00000	
1	72	348.10	.00000	352.28	.00000	2
2	48	66.72	.03812	66.90	.03691	4
3	16	11.41	.78356	11.35	.78735	4
4	0	0.	1.	0.	1.	

***** ASSOCIATION OPTION FOR ALL TERMS OF ORDER LESS THAN OR EQUAL TO 3

EFFECT	PARTIAL ASSOCIATION			ITER	MARGINAL ASSOCIATION		
	D.F.	CHISQ	PROB		D.F.	CHISQ	PROB
a.	2	314.13	.0000				
b.	2	646.86	.0000				
c.	2	118.44	.0000				
d.	2	.12	.9400				
ab	4	160.00	.0000	3	4	160.17	.0000
ac	4	70.71	.0000	3	4	69.23	.0000
ad	4	23.37	.0001	3	4	24.22	.0001
bc	4	22.90	.0001	3	4	21.62	.0002
bd	4	1.54	.8202	3	4	2.59	.6292
cd	4	3.60	.4625	3	4	3.00	.5579
abc	8	28.45	.0004	4	8	26.90	.0007
abd	8	7.53	.4807	4	8	7.97	.4361
acd	8	10.53	.2299	4	8	8.22	.4118
bcd	8	11.85	.1578	3	8	8.74	.3650

הותאם מודל של רגרסיה לוגיסטית לנתונים על שרידות לאחר ניתוח. המודל אשר הותאם הוא

$$\text{logit}[P(Y=1|A, S)] = \hat{\beta}_0 + \hat{\beta}_1 A + \hat{\beta}_2 S$$

כאשר

$$\left. \begin{array}{l} 1 \text{ אם החולה שרד 6 חודשים לפחות} \\ 0 \text{ אם לא} \end{array} \right\} = Y$$

$$A = \text{הגיל של החולה (בשנים)}$$

$$\left. \begin{array}{l} 1 \text{ אם החולה עישן} \\ 0 \text{ אם לא} \end{array} \right\} = S$$

אנו מעוניינים במנה בין יחס הסיכויים לשרוד, עבור חולה בן 55 אשר מעשן, לבין יחס הסיכויים לשרוד עבור חולה בן 50 אשר אינו מעשן. שימו לב שאנו משווים בין יחסים כאשר יש הבדל הן בגיל, הן בהרגלי עישון. נסמן ב- θ את מנת יחס הסיכויים הזה. להלן רשומות התוצאות אשר התקבלו.

- א. חשב אומדן נקודתי ל- θ
 ב. חשב רווח סמך, ברמה של 95%, ל- θ .

TERM	COEFFICIENT	STANDARD ERROR	COEF/SE	EXP(COEF)	95% C.I. OF EXP(COEF)	
					LOWER-BND	UPPER-BND
A	-.5473	.335	-1.63	.579	.299	1.12
S	-1.753	.695	-2.52	.173	.439E-01	.683
CONSTANT	-.1258	.945	-.133	.882	.137	5.69

COVARIANCE MATRIX OF COEFFICIENTS

	A	S	CONSTANT
A	0.112		
S	0.115	0.483	
CONSTANT	0.205	0.310	0.893