

הסקה על מודל רגרסיה לוגיסטית.

נעזר בחומר משקפים של ד"ר נויה גלאי וד"ר ניצה ברקן

בדיקת השערות על פרמטרים במודל

■ מבחני יחס הנראות (כי-בריבוע):

$$\chi^2 = -2 \left(\ln(L_{reduced}) - \ln(L_{full}) \right)$$

■ מבחני Wald המבוססים על התפלגות נורמלית אסימפטוטית של אומדי הנראות המכסימלית למקדמים.

■ במודל רגרסיה לוגיסטית יכולה להיות חוסר התאמה בין התוצאות של מבחן Wald על המקדמים לבין מבחן יחס הנראות המתאים. מבחן יחס הנראות נחשב לאמין ומדויק יותר.

רווחי סמך לפרמטרים במודל

■ רווח סמך ל- β_j :

מפולג אסימפטוטית

$$\frac{\hat{\beta}_j - \beta_j}{\hat{se}(\hat{\beta}_j)} \sim N(0,1)$$

$$P\left(Z_{\alpha/2} < \frac{\hat{\beta}_j - \beta_j}{\hat{se}(\hat{\beta}_j)} < Z_{1-\alpha/2}\right) = 1 - \alpha$$

$$P\left(-Z_{1-\alpha/2} \hat{se}(\hat{\beta}_j) < \hat{\beta}_j - \beta_j < Z_{1-\alpha/2} \hat{se}(\hat{\beta}_j)\right) = 1 - \alpha$$

$$\beta_j \in \left(\hat{\beta}_j \pm Z_{1-\alpha/2} \hat{se}(\hat{\beta}_j)\right)$$

רווחי סמך ל-OR

■ תזכורת:

$$OR = e^{\beta_1}$$

■ לכן נקבל

$$OR \in \exp\left(\hat{\beta}_1 \pm Z_{1-\alpha/2} \hat{se}(\hat{\beta}_1)\right)$$

■ כאשר

$$\hat{se}(\hat{\beta}_1)$$

הינו אמד לטעות התקן של $\hat{\beta}_j$ מתוך המודל.

מודל רגרסיה לוגיסטית (logistic regression model)

- דוגמא: הקובץ titanic (ד"ר ניצה ברקון)
רשימת המשתנים:
id – מספר תצפית
name – שם הנוסע
pclass – סוג האירוח בספינה (first, secon, third)
age – גיל הנוסע
gender – מין הנוסע 0=זכר, 1=נקבה
survived – משתנה תלוי 1=ניצל, 0=לא ניצל
משתנים מסבירים אינם משתנים בלתי תלויים.

מבחי השערות

■ המשך דוגמא:

- Analysis ← regression ← logistic
- dependent=survived, מימין לשנות את
- ה-response variable ל-1 ← להעביר
- ל-quantitative variables את age ואת gender ←
- להעביר ל-classification variables את pclass,
- מימין לשנות את ה-coding style ל-reference ←
- בלשונית statistics, ב-parameters נסמן
- profile likelihood וגם individual Wald tests, על
- מנת לקבל את רווחי סמך עבור הפרמטרים ← finish

Class Level Information			
Class	Value	Design Variables	
		1	2
pclass	1	1	0
	2	0	1
	3	0	0

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	932.116	623.821
SC	936.637	646.424
-2 Log L	930.116	613.821

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	316.2953	4	<.0001
Score	273.8579	4	<.0001
Wald	186.7844	4	<.0001

מבחני השערה על הפרמטרים

Type III Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
age	1	27.8002	<.0001
pclass	2	73.0003	<.0001
gender	1	164.1793	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.3655	0.2761	24.4663	<.0001
age		1	-0.0429	0.00813	27.8002	<.0001
pclass	1	1	2.6033	0.3047	73.0003	<.0001
pclass	2	1	1.2385	0.2599	22.7112	<.0001
gender		1	2.7979	0.2184	164.1793	<.0001

רווחי סמך לפרמטרים במודל

Profile Likelihood Confidence Interval for Parameters				
Parameter		Estimate	95% Confidence Limits	
Intercept		-1.3655	-1.9170	-0.8331
age		-0.0429	-0.0591	-0.0272
pclass	1	2.6033	2.0214	3.2178
pclass	2	1.2385	0.7358	1.7565
gender		2.7979	2.3807	3.2380

Wald Confidence Interval for Parameters				
Parameter		Estimate	95% Confidence Limits	
Intercept		-1.3655	-1.9066	-0.8244
age		-0.0429	-0.0588	-0.0269
pclass	1	2.6033	2.0061	3.2005
pclass	2	1.2385	0.7292	1.7479
gender		2.7979	2.3699	3.2259

רווחי סמך ל-OR

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
Age	0.958	0.943	0.973
pclass 1 vs 3	13.509	7.435	24.545
pclass 2 vs 3	3.451	2.073	5.743
gender	16.410	10.696	25.175