Homework for Chapter 10

蒋翌坤 20307100013

10.8

(a)

Model fit is $\hat{\mu}_{ab} = \frac{n_{ab}+n_{ba}}{2}$, and standard residual is $r_{ab} = \frac{n_{ab}-n_{ba}}{\sqrt{n_{ab}+n_{ba}}}$. The symmetry model fit is shown in the following table. $(\cdot)$ represents fitted value for symmetry model, $[\cdot]$ represents standard residuals.

	Second Purchase				
First Purchase	High Point	Taster's Choice	Sanka	Nescafe	Brim
High Point	93	17	44	7	10
	(93)	(13)	(30.5)	(6.5)	(10)
	[0.00]	[1.57]	[3.46]	[0.28]	[0.00]
Taster's Choice	9	46	11	0	9
	(13)	(46)	(11)	(2)	(6.5)
	[-1.57]	[0.00]	[0.00]	[-2.00]	[1.39]
Sanka	17	11	155	9	12
	(30.5)	(11)	(155)	(9)	(12)
	[-3.46]	[0.00]	[0.00]	[0.00]	[0.00]
Nescafe	6	4	9	15	2
	(6.5)	(2)	(9)	(15)	(2)
	[-0.28]	[2.00]	[0.00]	[0.00]	[0.00]
Brim	10	4	12	2	27
	(10)	(6.5)	(12)	(2)	(27)
	[0.00]	[-1.39]	[0.00]	[0.00]	[0.00]

$X^2 = \sum r_{ab}^2 = 22.47$ with $df = 10$, p-value is 0.01, so the model fits badly because there are large residuals.

(b)

Using software, $G^2 = 12.58$ with $df = 4$, p-value is 0.01, so the model fits badly.

(c)

Using software, $G^2 = 13.8$ with $df = 11$, p-value is 0.25, so the model fits well.

10.16

Using software, we have

Parameter	df	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	0	0.0000	0.0000	.	.
Biometrika	1	0.2690	0.0501	28.84	<.0001
Commun. Stat.	1	3.2180	0.0794	1642.25	<.0001
JASA	1	0.7485	0.0516	210.60	<.0001
JRSS-B	0	0.0000	0.0000	.	.

$G^2 = 8.59$ with $df = 9$, p-value is 0.48, so the model fits well.

From the fitted value, we have *JRSS-B* $\gtrsim$ *Biometrika* $\gtrsim$ *JASA* $\gtrsim$ *Commun. Stat.*

The probability that *Commun. Stat.* article cites *JRSS-B* article is $\frac{e^{3.218}}{1+e^{3.218}} = 0.96$

10.38

Marginal homogeneity means $\pi_{a+} = P(Y_1 = u_a) = P(Y_2 = u_a) = \pi_{+a}$ for $a = 1, \dots, I$.

$\mathbb{E}(\bar{Y}_1) = \mathbb{E}(\sum_a u_a p_{a+}) = \sum_a u_a \pi_{a+}$, $\mathbb{E}(\bar{Y}_2) = \mathbb{E}(\sum_a u_a p_{+a}) = \sum_a u_a \pi_{+a}$

So, $\mathbb{E}(\bar{Y}_1) = \mathbb{E}(\bar{Y}_2)$

$$
\begin{aligned}
\mathrm{Var}(\bar{Y_1} - \bar{Y_2}) &= \mathrm{Var}(\sum_a \sum_b (u_a - u_b) p_{ab}) \\
&= \sum_a \sum_b (u_a - u_b)^2 \mathrm{Var}(p_{ab}) + \sum_{(a,b)\neq(c,d)} (u_a - u_b)(u_c - u_d)\mathrm{Cov}(p_{ab}, p_{cd}) \\
&= \frac{1}{n} \sum_a \sum_b (u_a - u_b)^2 \pi_{ab}(1 - \pi_{ab}) - \frac{1}{n} \sum_{(a,b)\neq(c,d)} (u_a - u_b)(u_c - u_d)\pi_{ab}\pi_{cd} \\
&= \frac{1}{n}\Big[\sum_a \sum_b (u_a - u_b)^2 \pi_{ab} - (\sum_a \sum_b (u_a - u_b)\pi_{ab})^2\Big] \\
&= \frac{1}{n}\Big[\sum_a \sum_b (u_a - u_b)^2 \pi_{ab} - (\sum_a u_a \pi_{a+} - \sum_a u_a \pi_{+a})^2\Big]
\end{aligned}
$$

Substitute $\pi_{ab}, \pi_{a+}, \pi_{+a}$ with p_{ab}, p_{a+}, p_{+a}, we have the estimates for $\mathrm{Var}(\bar{Y_1} - \bar{Y_2})$.

A test of marginal homogeneity can be constructed by using $z = \frac{\bar{Y_1} - \bar{Y_2}}{\sqrt{\mathrm{Var}(\bar{Y_1} - \bar{Y_2})}} \xrightarrow{d} N(0, 1)$