Assignment 7 Solutions – M. Lacey, 03/18/02

Chapter 9, Problem 17
We compute the likelihood ratio test statistic $-2\log\Lambda = 2\sum_i O_i \log\left(\frac{O_i}{E_i}\right)$ using the observed and expected values provided in the table on p.314:

$$2\left(56\log\left(\frac{56}{34.9}\right) + 104\log\left(\frac{104}{85.1}\right) + \cdots + 20\log\left(\frac{20}{5}\right)\right) = 54.6.$$

Since $P(\chi^2_{(6)} > 54.6) \approx 0$, the likelihood ratio test strongly rejects the Poisson model, agreeing with the results from Pearson's chi-square test in Example B.

Chapter 9, Problem 25
To determine whether there is a "holiday effect" related to Passover for Jewish individuals, the appropriate test statistic will compare the observed values 922 and 997 to the values expected under the assumption of no effect, 959.5 for each of the two weeks. Using the likelihood ratio test:

$$2\left(922\log\left(\frac{922}{959.5}\right) + 997\log\left(\frac{997}{959.5}\right)\right) = 2.93.$$

$P(\chi^2_{(1)} > 2.93) = 0.087$, so we may reject the null hypothesis at the 0.1 level (but not at the 0.05 level). For the group of Chinese and Japanese males, the test statistic is

$$2\left(418\log\left(\frac{418}{426}\right) + 434\log\left(\frac{434}{426}\right)\right) = 0.30,$$

and since $P(\chi^2_{(1)} > 0.30) = 0.58$ there's no reason to reject the null hypothesis for this group. Based on these results, there appears to be some evidence that the holiday effect applies.

Chapter 9, Problem 30
To show that $\sum_{i=1}^{2}\frac{(X_i-np_i)^2}{np_i} = \frac{(X_1-np_1)^2}{np_1(1-p_1)}$, simply recognize that for a two-celled multinomial distribution $p_2 = 1-p_1$ and $X_2 = n - X_1$. Then

$$\begin{aligned}\sum_{i=1}^{2}\frac{(X_i-np_i)^2}{np_i} &= \frac{(X_1-np_1)^2}{np_1} + \frac{((n-X_1)-n(1-p_1))^2}{n(1-p_1)} \\ &= \frac{(X_1-np_1)^2(1-p_1)}{np_1(1-p_1)} + \frac{(np_1-X_1)^2(p_1)}{np_1(1-p_1)} = \frac{(X_1-np_1)^2}{np_1(1-p_1)}.\end{aligned}$$

Chapter 9, Problem 33
(a) The likelihood test statistic for the observed coin tosses is given by

$$2\left(9207\log\left(\frac{9207}{8975}\right) + 8743\log\left(\frac{8743}{8975}\right)\right) = 12.0.$$

This is a highly significant discrepancy, with $P(\chi^2_{(1)} > 12) = 0.0005$.

(b) The null hypothesis is that the 3590 sets of five tosses X follow a binomial distribution with $n = 5$ and $p = \frac{1}{2}$. Under this distribution, the expected counts E_i are calculated as follows:

$$\begin{aligned}
E_0 = &\ (3590)(P(X=0)) &= 3590 * (1/2)^5 &= 112.2\\
E_1 = &\ (3590)(P(X=1)) &= 3590 * 5 * (1/2)^5 &= 560.9\\
E_2 = &\ (3590)(P(X=2)) &= 3590 * 10 * (1/2)^5 &= 1121.9\\
E_3 = &\ (3590)(P(X=3)) &= 3590 * 10 * (1/2)^5 &= 1121.9\\
E_4 = &\ (3590)(P(X=4)) &= 3590 * 5 * (1/2)^5 &= 560.9\\
E_5 = &\ (3590)(P(X=2)) &= 3590 * (1/2)^5 &= 112.2
\end{aligned}$$

Based on these expected counts, the likelihood test statistic is given by

$$2\left(100\log\left(\frac{100}{112.2}\right) + 524\left(\log\frac{524}{560.9}\right) + 1080\left(\log\frac{1080}{1121.9}\right) + \cdots + 105\left(\log\frac{105}{112.2}\right)\right) = 20.9,$$

and $P(\chi^2_{(5)} > 20.9) = 0.00085$.

(c)To determine whether or not the coin tosses are consistent with another distribution with a fixed probability of heads p, we can estimate p using the data and fit the values for the corresponding binomial distribution. The estimate $\hat{p} = \frac{9207}{17950} = 0.513$, and the expected values according to this distribution are

$$\begin{aligned}
E_0 = &\ (3590)(P(X=0)) &= 3590 * (.487)^5 &= 98.3\\
E_1 = &\ (3590)(P(X=1)) &= 3590 * 5 * (.513)(.487)^4 &= 518\\
E_2 = &\ (3590)(P(X=2)) &= 3590 * 10 * (.513)^2(.487)^3 &= 1091.3\\
E_3 = &\ (3590)(P(X=3)) &= 3590 * 10 * (.513)^3(.487)^2 &= 1149.5\\
E_4 = &\ (3590)(P(X=4)) &= 3590 * 5 * (.513)^4(.498) &= 605.4\\
E_5 = &\ (3590)(P(X=2)) &= 3590 * (.513)^5 &= 127.6,
\end{aligned}$$

and the corresponding likelihood test statistic is

$$2\left(100\log\left(\frac{100}{98.3}\right) + 524\left(\log\frac{524}{518}\right) + 1080\left(\log\frac{1080}{1091.3}\right) + \cdots + 105\left(\log\frac{105}{127.6}\right)\right) = 8.72,$$

and $P(\chi^2_{(4)} > 8.72) = 0.07$. While we can't reject the null hypothesis at the 0.05 level, the model doesn't appear to be a very good fit to the data.