ST2132 Tutorial 11 Solution 2210

1. (a) Substituting x and a by X_i and $np_i(\hat{\theta})$ the expression in Tutorial 10 Question 4, we get

$$G \approx 2\sum_{i=1}^{r}\left\{(X_i - np_i(\hat{\theta})) + \frac{(X_i - np_i(\hat{\theta}))^2}{2np_i(\hat{\theta})}\right\}$$

Since $\sum_{i=1}^{r} X_i = \sum_{i=1}^{r}(np_i(\hat{\theta})) = n$, the first sum is 0. The second sum is X^2.

(b) Substituting x and a by X_i and $\bar{X}$,

$$G \approx 2\sum_{i=1}^{n}\left\{(X_i - \bar{X}) + \frac{(X_i - \bar{X})^2}{2\bar{X}}\right\} = \sum_{i=1}^{n}\frac{(X_i - \bar{X})^2}{\bar{X}}$$

the first sum being 0.

2. The expected counts under H_0 are

$$3839 \times \left(\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}\right) \approx (2159.4, 719.8, 719.8, 239.9)$$

(a) Using R, we get $x^2 \approx 287.7$. $P \approx \Pr(\chi_3^2 \geq 287.7) \approx 0$.

(b) $g \approx 387.5$. $P \approx \Pr(\chi_3^2 \geq 387.5) \approx 0$.

The fit is bad.

3. (a) The loglikelihood for Ω is

$$\ell(\mu) = -\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{2\sigma^2}$$

$$\ell'(\mu) = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sigma^2}, \qquad \ell''(\mu) = -\frac{n}{\sigma^2}$$

Solving $\ell'(\mu) = 0$ gives $\bar{X}$ as the ML estimator of μ, so that

$$\ell_1 = \ell(\bar{X}) = -\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{2\sigma^2}$$

Since Ω_0 has only one model, with $\mu = 0$, ℓ_0 is the loglikelihood of that model, which has value $\ell(0)$:

$$\ell_0 = \ell(0) = -\frac{\sum_{i=1}^{n} X_i^2}{2\sigma^2}$$

Since $\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 + n\bar{X}^2$,

$$G = 2(\ell_1 - \ell_0) = \frac{n\bar{X}^2}{\sigma^2}$$

(b) Under H_0, $\bar{X} \sim \text{Normal}(0, \sigma^2/n)$, so $\sqrt{n}\bar{X}/\sigma \sim \text{Normal}(0,1)$. Therefore, $G \sim \chi_1^2$.

4. (a) The loglikelihood for Ω is from slide 6 of Parameter Estimation II. Since the ML estimators of μ and σ^2 are $\bar{X}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$,

$$\ell_1 = -n \log \hat{\sigma} - \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{2\hat{\sigma}^2} = -\frac{n}{2} \log \hat{\sigma}^2 - \frac{n}{2}$$

The loglikelihood for Ω_0 is

$$\ell(\sigma) = -n \log \sigma - \frac{\sum_{i=1}^n X_i^2}{2\sigma^2}$$

$$\ell'(\sigma) = -\frac{n}{\sigma} + \frac{\sum_{i=1}^n X_i^2}{\sigma^3}, \qquad \ell''(\sigma) = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^n X_i^2}{\sigma^4}$$

So the ML estimator of σ^2 is $\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$, and

$$\ell_0 = \ell(\hat{\mu}_2) = -\frac{n}{2} \log \hat{\mu}_2 - \frac{n}{2}$$

(b)

$$G = 2(\ell_1 - \ell_0) = n \log \left(\frac{\hat{\mu}_2}{\hat{\sigma}^2} \right)$$

Now $\hat{\mu}_2 = \hat{\sigma}^2 + n\bar{X}^2$, which is the same as $\sum_{i=1}^n X_i^2 = \sum_{i=1}^n (X_i - \bar{X})^2 + n\bar{X}^2$ in previous question. Since $\log(1+x) \approx x$ for x near 0,

$$G \approx n \log \left(1 + \frac{n\bar{X}^2}{\hat{\sigma}^2} \right) \approx \frac{n\bar{X}^2}{\hat{\sigma}^2}$$

(c) Since $\hat{\sigma}^2 = (n-1)S^2/n$,

$$G \approx \frac{n}{n-1} \frac{n\bar{X}^2}{S^2}$$

For large n, $n/(n-1) \approx 1$. From Tutorial 9 Question 4, $n\bar{X}^2/S^2 \sim F_{1,n-1}$, which is approximately χ_1^2 from Probability Review II slide 18. So approximately $G \sim \chi_1^2$, under H_0.

5. (a) Adapting from Parameter Estimation II slide 6, the loglikelihood for Ω is

$$\ell(\lambda_1, \lambda_2) = \sum_{j=1}^{n_1} X_{1j} \log \lambda_1 - n_1\lambda_1 + \sum_{j=1}^{n_2} X_{2j} \log \lambda_2 - n_2\lambda_2$$

$$\frac{\partial \ell}{\partial \lambda_i} = \frac{\sum_{j=1}^{n_i} X_{ij}}{\lambda_i} - n_i$$

Setting both derivatives to 0 gives the ML estimators $\hat{\lambda}_i = \bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$. Hence

$$\ell_1 = \ell(\hat{\lambda}_1, \hat{\lambda}_2) = \sum_{i=1}^{2} \left(\sum_{j=1}^{n_i} X_{ij} \log \bar{X}_i - n_i\bar{X}_i \right)$$

The loglikelihood for Ω_0 is

$$\ell(\lambda) = \sum_{j=1}^{n_1} X_{1j} \log \lambda - n_1\lambda + \sum_{j=1}^{n_2} X_{2j} \log \lambda - n_2\lambda = \sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij} \log \lambda - (n_1 + n_2)\lambda$$

Setting

$$\ell'(\lambda) = \frac{\sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij}}{\lambda} - (n_1 + n_2)$$

to 0 gives the ML estimator $\hat{\lambda} = \bar{X} = \frac{1}{n_1+n_2}\sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij}$. Hence

$$\ell_0 = \ell(\hat{\lambda}) = \sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij} \log \bar{X} - (n_1 + n_2)\bar{X}$$

(c) Note that $(n_1 + n_2)\bar{X} = n_1\bar{X}_1 + n_2\bar{X}_2$. Hence

$$G = 2(\ell_1 - \ell_0) = 2\sum_{i=1}^{2}\sum_{j=1}^{n_i} X_{ij} \log\left(\frac{\bar{X}_i}{\bar{X}}\right)$$

$k_1 = 2$, $k_0 = 1$. Under H_0, for large n_1 and n_2, approximately $G \sim \chi_1^2$.