Rice STAT 533/ GS01 1283 Homework 2 Solutions
February 6, 2020
Request: Please email the instructor if you find any mistakes in this document.

Optional Exercise: The optional exercise was to proof the claim made in Equation 6.61 on p428 or provide a counter-example. In older printings of the book, which state "for some $c_0 > 0$", the claim is false. For newer printings of the book, which state "for some $c_0 \geq 1$", the claim is true.

$c_0 \geq 1$ proof:

$$\{x : \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} > c_0\} = \{x : \frac{f_{\theta_0}(x)}{f_{\theta_1}(x)} < \frac{1}{c_0}\}$$
$$= \{x : \frac{f_{\theta_0}(x)}{\max(f_{\theta_1}(x), f_{\theta_0}(x)}) < \underbrace{\frac{1}{c_0}}_{\equiv c}\}.$$

The second equality requires $c_0 \geq 1$ and hence $1/c_0 \leq 1$.

$c_0 > 0$ counterexample: Let $f_0 = f_{\theta_0}$ be $N(0,1)$ and $f_1 = f_{\theta_1}$ be N(1,1) and X a single observation from one of these distributions. Consider rejecting when x ¿ 0. This is a UMP size 1/2 test with $c_0 = f_1(0)/f_0(0)$. $L(x) = f_0(x)/\max(f_0(x), f_1(x)) = 1$ for all x ¡ 1/2 (where L is the likelihood ratio). This is because $f_0(x) > f_1(x)$ on $x < 1/2$. If you reject for all x such that $L(x) < c <= 1$, then you will not reject on $(-\infty, 1/2)$. So there will be no way to choose a c which rejects on $(0, 1/2)$, as is done with the UMP test.

Section 6.6 Exercise 91:

There is a typo in this problem. The statistic W should be

$$W = \frac{\sum_{i=1}^n (X_{i1} - \bar{X}_i)(X_{i2} - \bar{X}_2)}{\sqrt{\sum_{i=1}^n (X_{i1} - \bar{X}_1)^2 \sum_{i=1}^n (X_{i2} - \bar{X}_2)^2}}$$

This is the sample correlation coefficient, so sensibly we will reject H_0 when $|W|$ is large. To derive W from a LRT perspective, one needs to compute restricted (parameters optimized over null) and unrestricted MLE. The unrestricted MLE are the usual estimators $\widehat{\mu} = \bar{X} = (\bar{X}_1, \bar{X}_2)^T$ and sample covariance $\widehat{\Sigma} = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^T$. For the restricted MLE, the mean is the same and the covariance matrix $\widetilde{\Sigma}$ is diagonal with j,j element $\frac{1}{n}\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2$. Then one can show that for any $c \in [0,1]$ there exists c' such that

$$\{x : \frac{l(\bar{X}, \widetilde{\Sigma})}{l(\bar{X}, \widehat{\Sigma})} < c\} = \{x : W > c'\}.$$

Deriving the distribution of W is difficult, but from the internet under the null hypothesis of $\rho = 0$

$$f_W(w) = \frac{(1 - w^2)^{(n-4)/2}}{B(1/2, (n-2)/2)}$$

where B is the Beta function. This distribution is symmetric about 0, so to construct a size α test we reject when $|W| > w_{\alpha/2}$ where $w_{\alpha/2}$ satisfies $F_W(w_{\alpha/2}) = \alpha/2$.

Section 6.6 Exercise 94a:

The unrestricted MLEs are $\widehat{p}_1 = \frac{n_1}{X_1}$ and $\widehat{p}_2 = \frac{n_2}{X_2}$ and the restricted MLE under the null is $\widetilde{p} = \widetilde{p}_1 = \widetilde{p}_2 = \frac{n_1+n_2}{x_1+x_2} \equiv \frac{n}{x}$ where $n \equiv n_1 + n_2$ and $x = x_1 + x_2$. The LR test statistic is

$$\lambda((x_1, x_2)) = \frac{l(\widetilde{p})}{l(\widehat{p})} = \frac{\overbrace{(\frac{n}{x})^n(1-\frac{n}{x})^{x-n}}^{\equiv g(x)}}{\underbrace{(\frac{n_1}{x_1})^{n_1}(1-\frac{n_1}{x_1})^{x_1-n_1}(\frac{n_2}{x_2})^{n_2}(1-\frac{n_2}{x_2})^{x_2-n_2}}_{h(x_1,x_2)}}$$

A LRT rejects when λ is small which is equivalent to rejecting for $\{(x_1, x_2) : g(x) < ch(x_1, x_2)\}$ for some c. Note that since x is a sufficient statistic under the null, the distribution of $h(x_1, x_2)$ conditioned on x will not depend on p. Thus we can determine c_α such that $P(g(x) < c_\alpha h(x_1, x_2)|x) \leq \alpha$.

Section 6.6 Exercise 97:

Verify 6.65: Taylor expand $R(\widetilde{\theta})$ around $R(\theta)$. We have

$$R(\widetilde{\theta}) = R(\theta) + \frac{\partial R}{\partial \theta}(\theta)^T(\widetilde{\theta} - \theta) + o_P(||\widetilde{\theta} - \theta||)$$

Noting that under H_0, $\widetilde{\theta} \to \theta$ and $R(\widetilde{\theta}) = R(\theta) = 0$ we have

$$0 = C(\theta)^T(\widetilde{\theta} - \theta) + o_P(||\widetilde{\theta} - \theta||)$$

Since $||\sqrt{n}(\widetilde{\theta} - \theta)|| \to_d ||N(0, \Sigma)||$ we have

$$C(\theta)^T(\widetilde{\theta} - \theta) = o_P(n^{-1/2})$$

Verify 6.66:

Let $s_i(\theta) = \frac{\partial}{\partial\theta} \log f(x_i|\theta)$. Then

$$s_i(\widetilde{\theta}) = s_i(\theta) - I(\theta)(\widetilde{\theta} - \theta) + o_P(n^{-1/2})$$

Summing across i we have

$$s_n(\widetilde{\theta}) = s_n(\theta) - I_n(\theta)(\widetilde{\theta} - \theta) + o_P(n^{1/2})$$

From the Lagrange multiplier restriction we have

$$s_n(\theta) - I_n(\theta)(\widetilde{\theta} - \theta) + C(\widetilde{\theta})\lambda_n = o_p(n^{1/2})$$

Section 6.6 Exercise 100:

Parameterize the multinomial distribution for k categories with $p' \in \Theta' = \{p' : \sum_{j=1}^{k-1} p'_j < 1, p_j > 0\}$ and $p'_k \equiv 1 - \sum_{j=1}^{k-1} p_j$. The resulting pdf (wrt counting measure) is

$$f_p(x) = \frac{n!}{x_1! \ldots x_k!} \left(\prod_{j=1}^{k-1} p_j^{x_j}\right)(1 - \sum_{j=1}^{k=1} p_j)^{x_k}$$

See Example 2.7 on page 98 for a discussion of the Multinomial distribution and this parameterization. Here Θ' contains an open set and satisfies the regularity conditions required by Theorem 6.5 (asymptotic distribution of Wald and Rao tests).

Let $\xi_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ be a length k vector with a 1 in the jth element if the result of trial i was category j. We have $X = \sum_{i=1}^{n} \xi_i$. We test the hypotheses

$$\begin{aligned} H_0 &: p' = p'_0 \\ H_1 &: p' \neq p'_0 \end{aligned}$$

This null hypothesis can be restated as $R(p') = p' - p'_0$. Therefore $C(p') = \frac{\partial}{\partial p'} R(p') = I$.

<u>Wald Test:</u>

The Wald statistic is

$$\begin{aligned} W_n &= R(\widehat{p}')^T (C(\widehat{p}')^T I_n^{-1}(\widehat{p}') C(\widehat{p}'))^{-1} R(\widehat{p}') \\ &= (\widehat{p}' - p'_0)^T I_n(\widehat{p}')(\widehat{p}' - p'_0) \end{aligned}$$

Computing partial derivatives of the log likelihood one obtains

$$\mathbb{E}[\frac{-\partial \log f}{\partial p'_l \partial p'_j}] == \begin{cases} \frac{1}{1-\sum_{j=1}^{k-1} p'_j} & l \neq j \\ \frac{1}{p'_j} + \frac{1}{1-\sum_{j=1}^{k-1} p'_j} & l = j \end{cases}$$

Thus $I_n(\widehat{p}') = nI(\widehat{p}') = nA + nB$ where A is diagonal with $A_{jj} = 1/\widehat{p}'_j$ and all elements of B are $(1 - \sum_{j=1}^{k-1} \widehat{p}'_j)^{-1} = \widehat{p}'^{-1}_k$.

After manipulating matrices we have

$$\begin{aligned} W_n &= (\widehat{p}' - p'_0)^T (nA + nB)(\widehat{p}' - p'_0) \\ &= \sum_{j=1}^{k} \frac{(X_j - np'_{0j})^2}{X_j} \\ &\equiv \widetilde{\chi}^2 \end{aligned}$$

<u>Rao–Score Test:</u>

Need to compute

$$R_n = s_n(p'_0)^T I_n^{-1}(p_0) s_n(p'_0)$$

The score is

$$s_{n,j}(p'_0) = \frac{\partial}{\partial p'_j} \log l(p'_0) = \frac{X_j}{p_{0j}} - \frac{X_k}{p_{0k}}$$

Note that we can compute the inverse of I_n (originally computed for the Wald), using the Sherman Woodbury Morrison formula since A is diagonal and B is a rank one update. Therefore

$$I_n^{-1}(p'_0) = n^{-1} A^{-1} + n^{-1} p'_0 p'^T_0$$

where A^{-1} is diagonal with $A^{-1}_{jj} = p_{0j}$.

After some matrix algebra with one has

$$R_n = \sum_{j=1}^{k} \frac{(X_j - p'_{0j} n)^2}{np'_{0j}}.$$