Solution for weekly review exercises #12

Prepared by Chen-Yen Lin

Apr. 13, 2011

(a) i. Denote \( T = -\sum_{i=1}^{n} \log X_i \). First derive the MLE of \( \theta \), \( \hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log x_i} = \frac{n}{T} \). Then the likelihood ratio test rejects the null if and only if

\[
\frac{\theta_0^n (\prod_i x_i)^{\theta_0 - 1}}{\hat{\theta}^n (\prod_i x_i)^{\hat{\theta} - 1}} < c
\]

\[
(\theta_0 - 1)(-T) - n \log \left( \frac{n}{T} \right) - \left( \frac{n}{T} - 1 \right)(-T) < c'
\]

\[-\theta_0 T + n \log \left( \frac{T}{\theta_0} \right) < c''
\]

\[
\left\{ \frac{T}{\theta_0} < k_1 \text{ or } \frac{T}{\theta_0} > k_2, \ k_1 < k_2 \right\}
\]

ii. From i. the acceptance region has the form

\[ A(\theta_0) = \left\{ x : k_1 < \frac{T(x)}{\theta_0} < k_2 \right\} \]

Inverting the acceptance region gives the \( 1 - \alpha \) confidence set

\[ C(x) = \left\{ \theta : k_1 < \frac{T}{\theta} < k_2 \right\} \]

where these two constants \( k_1, k_2 \) can be found by solving the system

\[-\frac{T^2}{k_1} + n \log k_1 = -\frac{T^2}{k_2} + n \log k_2
\]

\[ P \left( k_1 < \frac{T}{\theta} < k_2 \right) = 1 - \alpha
\]

iii. An equal-tailed C.I. can be found by setting

\[ P \left[ \chi^2_{\alpha/2}(2n) < 2\theta T < \chi^2_{1-\alpha/2}(2n) \right] = 1 - \alpha
\]

Thus, the C.I. for \( \theta \) is given by

\[ \left( \frac{\chi^2_{\alpha/2}(2n)}{2T}, \frac{\chi^2_{1-\alpha/2}(2n)}{2T} \right) \]
(b)  
i. Considering that \( \frac{X_i}{1-\theta} \) follows a U(0, 1), we have \( \frac{X_{(1)}}{1-\theta} \sim \text{Beta}(1, n) \). Since \( \frac{X_{(1)}}{1-\theta} \) involves random variable and unknown parameter and its distribution does not depend on \( \theta \), it is a pivotal quantity.

ii. Start from the fact that \( \frac{X_{(1)}}{1-\theta} \) is a Beta(1, n), for any pair of constants \((k_1, k_2)\) such that
\[
P\left(k_1 < \frac{X_{(1)}}{1-\theta} < k_2\right) = 1 - \alpha
\]
we can derive a C.I. for \( \theta \). That is
\[
\left(1 - \frac{X_{(1)}}{k_1}, 1 - \frac{X_{(1)}}{k_2}\right)
\]
For simplicity, we consider an equal-tailed interval. That is, \( k_1 \) and \( k_2 \) are \( \alpha/2 \) and \( 1 - \alpha/2 \) percentiles of Beta(1, n) distribution, respectively. Moreover, they can be written more explicitly since the \( \tau \)-percentile of Beta(1, n) has the form \( 1 - \left[1 - \tau\right]^{1/n} \). So, the \( 1 - \alpha \) equal-tailed C.I. for \( \theta \) is given by
\[
\left(1 - \frac{X_{(1)}}{1 - (1 - \alpha/2)^{1/n}}, 1 - \frac{X_{(1)}}{1 - (\alpha/2)^{1/n}}\right)
\]

iii. Derive the CDF of \( X_{(1)} \)
\[
P(X_{(1)} < x) = 1 - P(X_{(1)} > x) = 1 - [P(X_1 > x)]^n = 1 - \left[\frac{1-x}{1-\theta}\right]^n
\]
By setting two constants \( c_1 \) and \( c_2 \) such that
\[
P\left(c_1 < 1 - \left[\frac{1-X_{(1)}}{1-\theta}\right]^n < c_2\right) = 1 - \alpha
\]
We can derive another C.I. for \( \theta \). That is
\[
\left(1 - \frac{1-X_{(1)}}{(1-c_1)^{1/n}}, 1 - \frac{1-X_{(1)}}{(1-c_2)^{1/n}}\right)
\]
Again, for simplicity, we consider an equal-tailed C.I. for \( \theta \). That is, set \( c_1 = \alpha/2 \) and \( c_2 = 1 - \alpha/2 \), then the C.I. for \( \theta \) is given by
\[
\left(1 - \frac{1-X_{(1)}}{(\alpha/2)^{1/n}}, 1 - \frac{1-X_{(1)}}{(1-\alpha/2)^{1/n}}\right)
\]

(c)  
i. 
\[
\frac{f(x|\theta)}{f(y|\theta)} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1-\theta}{\theta}} = \left(\prod_{i=1}^{n} y_i\right)^{\frac{1-\theta}{\theta}} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1-\theta}{\theta}}
\]
The ratio is free from \( \theta \) if and only if \( \prod_{i=1}^{n} x_i = \prod_{i=1}^{n} y_i \). Thus, \( \prod_{i=1}^{n} x_i \) is minimal sufficient for \( \theta \). So is \(-2\log(\prod_{i=1}^{n} x_i)\)
ii. \( Y_1 = \phi(X_1) = -2 \log X_1, \ X_1 = \phi^{-1}(Y_1) = e^{-\frac{Y_1}{2}} \) and the Jacobian term is \( \frac{1}{2} e^{-\frac{Y_1}{2}} \).

So, the density function of \( Y_1 \) is given by

\[
 f_{Y_1}(y) = \frac{1}{\theta} \left( e^{-\frac{y}{2}} \right)^\frac{1}{\theta} \frac{1}{2} e^{-\frac{y}{2}} = \frac{1}{2\theta} e^{-\frac{y}{2\theta}}
\]

That is, \( Y_1 \sim \exp(2\theta) \).

iii. Since \( T = -2 \sum_{i=1}^{n} \log X_i = \sum_{i=1}^{n} Y_i \sim \text{Gamma}(n, 2\theta) \), \( \frac{T}{\theta} \sim \chi^2(2n) \). We can construct a confidence interval for \( \theta \) using \( \chi^2(2n) \) percentiles. i.e.

\[
 P \left( \chi^2_{\alpha/2}(2n) < \frac{T}{\theta} < \chi^2_{1-\alpha/2}(2n) \right),
\]

As a result, an equal-tailed C.I. for \( \theta \) is given by

\[
 \left( \frac{T}{\chi^2_{1-\alpha/2}(2n)}, \frac{T}{\chi^2_{\alpha/2}(2n)} \right)
\]

iv. The expected length of the interval in iii. can be found as

\[
 E \left[ \frac{T}{\chi^2_{\alpha/2}(2n)} - \frac{T}{\chi^2_{1-\alpha/2}(2n)} \right] = \left[ \frac{1}{\chi^2_{\alpha/2}(2n)} - \frac{1}{\chi^2_{1-\alpha/2}(2n)} \right] ET = \left[ \frac{1}{\chi^2_{\alpha/2}(2n)} - \frac{1}{\chi^2_{1-\alpha/2}(2n)} \right] 2n\theta
\]

As \( n \to \infty \), we need to study the behavior of \( \chi^2 \) percentile. By “CLT”, we can approximate the distribution of \( \chi^2(2n) \) as follows

\[
 P[\chi^2(2n) < \chi^2_{\alpha}(2n)] = \alpha
\]

\[
 P \left[ \frac{\sum_{i=1}^{2n} Z_i - 2n}{\sqrt{4n}} < \frac{\chi^2_{\alpha}(2n) - 2n}{\sqrt{4n}} \right] = \alpha
\]

where \( Z_i \overset{iid}{\sim} \chi^2(1) \). Thus, we can approximate the \( \alpha \)-percentile of \( \chi^2(2n) \) as

\[
 \chi^2_{\alpha}(2n) \approx 2n + \sqrt{4n}Z_{\alpha}
\]

Then the expected length will have limit

\[
 \lim_{n \to \infty} \left[ \frac{1}{\chi^2_{\alpha/2}(2n)} - \frac{1}{\chi^2_{1-\alpha/2}(2n)} \right] 2n\theta = \lim_{n \to \infty} \left[ \frac{1}{2n + \sqrt{4n}Z_{\alpha/2}} - \frac{1}{2n + \sqrt{4n}Z_{1-\alpha/2}} \right] 2n\theta = 0
\]