STAT610 - HWK Solution 11

4.4.6 The length of the confidence interval is given by

$$l(\alpha_1,\alpha_2) = \left(\bar{X} + z(1-\alpha_2)\frac{\sigma}{\sqrt{n}}\right) - \left(\bar{X} - z(1-\alpha_1)\frac{\sigma}{\sqrt{n}}\right) = \frac{\sigma}{\sqrt{n}}(z_{1-\alpha_1} + z_{1-\alpha_2})$$

- If $\alpha_1 + \alpha_2 < \alpha$, let $\alpha_2' = \alpha - \alpha_1 > \alpha_2$, then $z_{1-\alpha_2'} < z_{1-\alpha_2}$ which means the interval using α_1 and α_2' is shorter. So the minimum length can be obtained only when $\alpha_1 + \alpha_2 = \alpha$.
- If $\alpha_1 < \alpha_2$, let $\alpha_1' = \alpha_2' = \frac{1}{2}(\alpha_1 + \alpha_2)$, then

$$z_{1-\alpha_1} > z_{1-\alpha_1'} = z_{1-\alpha_2'} > z_{1-\alpha_2}$$

By mean-value theorem

$$\frac{1}{2}(\alpha_2 - \alpha_1) = \int_{z_{1-\alpha_1'}}^{z_{1-\alpha_1}} \phi(t)dt = \phi(t_1^*)[z_{1-\alpha_1} - z_{1-\alpha_1'}]$$

$$\frac{1}{2}(\alpha_2 - \alpha_1) = \int_{z_{1-\alpha_2}}^{z_{1-\alpha_2'}} \phi(t)dt = \phi(t_2^*)[z_{1-\alpha_2'} - z_{1-\alpha_2}]$$

for some $z_{1-\alpha_2} < t_2^* < z_{1-\alpha_2'} = z_{1-\alpha_1'} < t_1^* < z_{1-\alpha_1}$. Without loss of generality, assume $\alpha < 1/2$, then $0 < t_2^* < t_1^*$.

$$\Rightarrow \phi(t_1^*) < \phi(t_2^*) \Rightarrow z_{1-\alpha_1} - z_{1-\alpha_1'} > z_{1-\alpha_2'} - z_{1-\alpha_2}$$
$$\Rightarrow z_{1-\alpha_1'} + z_{1-\alpha_2'} < z_{1-\alpha_1} + z_{1-\alpha_2}$$

Therefore, $\alpha_1 = \alpha_2 = \alpha/2$ gives the shortest interval.

4.5.1 (a) From B.3.4, B.3.5,

$$2\theta\sum_{i=1}^{n_1} X_i \sim \chi^2_{2n_1}, 2\lambda\sum_{i=1}^{n_2} Y_i \sim \chi^2_{2n_2} \Rightarrow \frac{2\theta\sum_{i=1}^{n_1} X_i/(2n_1)}{2\lambda\sum_{i=1}^{n_2} Y_i/(2n_2)} = \frac{\theta}{\lambda}\frac{\bar{X}}{\bar{Y}} = \Delta\frac{\bar{X}}{\bar{Y}} \sim F_{2n_1,2n_2}$$

$$P\left(\frac{\bar{Y}}{\bar{X}}f_{\alpha/2} \leq \Delta \leq \frac{\bar{Y}}{\bar{X}}f_{1-\alpha/2}\right) = P\left(f_{\alpha/2} \leq \Delta\frac{\bar{X}}{\bar{Y}} \leq f_{1-\alpha/2}\right) = 1-\alpha$$

(b) $H : \Delta = 1$ vs. $K : \Delta \neq 1$

$$1 - P_{\Delta=1}\left(f_{\alpha/2} \leq \frac{\bar{X}}{\bar{Y}} \leq f_{1-\alpha/2}\right) = 1-(1-\alpha) = \alpha$$

(c) $n_1 = 20, n_2 = 7, \bar{x} = 33.95, \bar{y} = 15.857, f_{0.05} = 0.5134, f_{0.95} = 2.2664$, so the 90% CI is [0.24,1.06]. Since $1 \in [0.24, 1.06]$, H is not rejected at level $\alpha = 0.1$.

4.6.1 $2\sum X_i/\theta \sim \chi^2_{2np}$ is a pivot quantity, so a level (1-α) UCB for θ is given by

$$\overline{\theta}^* = \frac{2\sum X_i}{\chi^2_{2np,\alpha}}$$

On the other hand, for $\theta_1 < \theta_2$, $f(x;\theta_2)/f(x;\theta_1) = (\theta_1/\theta_2)^{np}\exp\left((1/\theta_1 - 1/\theta_2)\sum X_i\right)$ is an increasing function of $T = \sum X_i$. Thus, it's a MLR family in T. Then the level α UMP test for $H: \theta = \theta_0$ versus $K: \theta < \theta_0$ is given by

$$\text{Reject } H \text{ when } \sum X_i < \frac{\theta_0}{2}\chi^2_{2np,\alpha} \Leftrightarrow \theta_0 > \frac{2\sum X_i}{\chi^2_{2np,\alpha}} = \overline{\theta}^*$$

By Theorem 4.6.1, $\overline{\theta}^*$ is UMA at level (1-α).

4.7.1 (a) From B.3, $\theta \sim \beta(r,s) \Rightarrow \lambda = \frac{s\theta}{r(1-\theta)} \sim F_{2r,2s}$

(b) $X|\theta \sim B(n,\theta)$, $\quad \theta \sim \beta(r,s) \Rightarrow \theta|X \sim \beta(r+X, n-X+s)$.

$$\Rightarrow \frac{n-X+s}{r+X}\frac{\theta}{1-\theta}\bigg| X = \frac{n-X+s}{r+x}\frac{r}{s}\lambda\bigg| X \sim F_{2(r+X),2(n-X+s)}$$

Let f_α be αth quantile of $F_{2(r+X),2(n-X+s)}$.

$$1-\alpha = P\left(\frac{r(n-X+s)}{s(r+X)}\lambda \geq f_\alpha \bigg| X\right) = P\left(\lambda \geq \frac{s(r+X)}{r(n-X+s)}f_\alpha \bigg| X\right)$$

Therefore, $\underline{\lambda} = \frac{s(r+X)}{r(n-X+s)}f_\alpha$ is a LCB for λ. Similarly, $\bar{\lambda} = \frac{s(r+X)}{s(n-X+s)}f_{1-\alpha}$ is a UCB for λ. Since

$$1-\alpha = P(\lambda \geq \underline{\lambda}) = P\left(\theta \geq \frac{\underline{\lambda}r}{s+\underline{\lambda}r}\right)$$

Therefore, $\frac{\underline{\lambda}r}{s+\underline{\lambda}r}$ is a LCB for θ. Similarly, $\frac{\bar{\lambda}r}{s+\bar{\lambda}r}$ is a UCB for θ.

4.8.2 (a) $(\bar{X}_n - X_{n+1})/\sqrt{(1+\frac{1}{n})\sigma_0^2} \sim N(0,1)$, so

$$P\left(X_{n+1} \geq \bar{X} - z_{1-\alpha}\sigma_0\sqrt{1+\frac{1}{n}}\right) = 1-\alpha$$

$$P\left(X_{n+1} \leq \bar{X} - z_{\alpha}\sigma_0\sqrt{1+\frac{1}{n}}\right) = 1-\alpha$$

Therefore $\bar{X} - z_{1-\alpha}\sigma_0\sqrt{1+\frac{1}{n}}$ and $\bar{X} - z_\alpha\sigma_0\sqrt{1+\frac{1}{n}}$ are level (1-α) lower and upper prediction bounds for X_{n+1}.

(b) $\bar{X}_n - X_{n+1} \sim N(0, (1+1/n)\,\sigma^2)$, $\frac{n-1}{\sigma^2}S^2 = \frac{\sum(X_i - \bar{X}_n)^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow \frac{\bar{X}_n - X_{n+1}}{S\sqrt{1+\frac{1}{n}}} \sim t_{n-1}$.

$$P\left(X_{n+1} \geq \bar{X}_n - t_{1-\alpha}S\sqrt{1+\frac{1}{n}}\right) = 1-\alpha$$

$$P\left(X_{n+1} \leq \bar{X}_n - t_{\alpha}S\sqrt{1+\frac{1}{n}}\right) = 1-\alpha$$

Therefore $\bar{X}_n - t_{1-\alpha}S\sqrt{1+\frac{1}{n}}$ and $\bar{X}_n - t_\alpha S\sqrt{1+\frac{1}{n}}$ are level (1-α) lower and upper prediction bounds for X_{n+1}.

(c) Let $X_{(1)}, \cdots, X_{(n)}$ denote the order statistics. $U_i = F(X_i) \sim U(0,1), i = 1, \cdots, n$.

$$\begin{aligned}
P(X_{(j)} \le X_{n+1}) &= P[F(X_{(j)}) \le F(X_{n+1})] = P[U_{(j)} \le U_{n+1}] = E[I(U_{(j)} \le U_{n+1})] \\
&= E[E\{I(U_{(j)} \le U_{n+1}) | U_{(j)}\}] = E[1 - U_{(j)}] \\
&= 1 - E[U_{(j)}] = 1 - \frac{j}{n+1} \qquad \text{(By Problem B.2.9)} \\
&\ge 1 - \alpha
\end{aligned}$$

For any $j \le \alpha(n+1), j \in \mathbb{N}$, $P(X_{(j)} \le X_{n+1}) \ge 1 - \alpha$. Therefore $X_{([\alpha(n+1)])}$ is a distribution free lower prediction bound for X_{n+1}. Similarly, $X_{([(1-\alpha)(n+1)]+1)}$ is a distribution free upper prediction bound for X_{n+1}.

4.9.9 (a) Since $X_i | X_1, \cdots, X_{i-1} \sim N(\theta X_{i-1}, \sigma^2)$,

$$f(x_1, \cdots, x_n) = f(x_1) f(x_2 | X_1) \cdots f(x_n | x_1, \cdots, x_{n-1}) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left(-\frac{\sum (x_i - \theta x_{i-1})^2}{2\sigma^2} \right)$$

(b) $H : \theta = 0$ vs $K : \theta \ne 0$. Assume σ^2 is known, the MLE of $\theta \in \mathbb{R}$ is $\hat{\theta} = \dfrac{\sum x_i x_{i-1}}{\sum x_{i-1}^2}$.

$$\begin{aligned}
\lambda &= \frac{\sup_\theta p(x;\theta)}{\sup_{\theta=0} p(x;\theta)} = \frac{p(x;\hat{\theta})}{p(x;0)} \\
&= \exp\left(-\frac{1}{2\sigma^2} \left\{ -2\hat{\theta} \sum x_i x_{i-1} + \hat{\theta}^2 \sum x_{i-1}^2 \right\} \right) = \exp\left(\frac{1}{2\sigma^2} \frac{(\sum x_i x_{i-1})^2}{\sum x_{i-1}^2} \right)
\end{aligned}$$

Therefore $\lambda > c \Leftrightarrow \dfrac{(\sum x_i x_{i-1})^2}{\sum x_{i-1}^2} > c$.