STAT 611 Homework 8 Solutions

1. If $\alpha < p(\mathbf{x})$,

$$\sup_{\theta\in\Theta_0} P(W(\mathbf{X}) \geq c_\alpha) = \alpha < p(\mathbf{x}) = \sup_{\theta\in\Theta_0} P(W(\mathbf{X}) \geq W(\mathbf{x}))$$

Thus $W(\mathbf{x}) < c_\alpha$ and we could not reject H_0 at level α having observed $\mathbf{x}$. On the other hand, if $\alpha \geq p(\mathbf{x})$,

$$\sup_{\theta\in\Theta_0} P(W(\mathbf{X}) \geq c_\alpha) = \alpha \geq p(\mathbf{x}) = \sup_{\theta\in\Theta_0} P(W(\mathbf{X}) \geq W(\mathbf{x}))$$

Either $W(\mathbf{x}) \geq c_\alpha$ in which case we could reject H_0 at level α having observed $\mathbf{x}$ or $W(\mathbf{x}) < c_\alpha$. But, in the latter case we could use $c'_\alpha < W(\mathbf{x})$ and have $\{\mathbf{x}' : W(\mathbf{x}' \geq c'_\alpha)\}$ define a size α rejection region. Then we could reject H_0 at level α having observed $\mathbf{x}$.

2. (a) For $Y = -(\log X)^{-1}$, the pdf of Y is $f_Y(y) = \frac{\theta}{y^2}e^{-\theta/y}$ for $0 < y < \infty$ and

$$P(Y/2 \leq \theta \leq Y) = \int_\theta^{2\theta} \frac{\theta}{y^2}e^{-\theta/y}S = 0.239$$

(b) Since $f_X(x) = \theta x^{\theta-1}$ for $0 < x < 1$, $T = X^\theta$ is a good guess at a pivot, and it is since $f_T(t) = 1$, $0 < t < 1$. Thus a pivotal interval is formed from $P(a < X^\theta < b) = b - a$ and is

$$\left\{\theta : \frac{\log b}{\log x} \leq \theta \leq \frac{\log a}{\log x}\right\}$$

Since $X^\theta \sim Uniform(0,1)$, the interval will have confidence 0.239 as long as $b - a = 0.239$.

(c) The interval in part (a) is a special case of the one in part (b). To find the best interval, we minimize $\log b - \log a$ subject to $b - a = 1 - \alpha$ or $b = 1 - \alpha + a$. Thus, we want to minimize $\log(1 - \alpha + a) - \log a = \log(1 + (1-\alpha)/a)$, which is minimized by taking a as big as possible. Thus, take $b = 1$ and $a = \alpha$, and the best $1 - \alpha$ pivotal interval is

$$\left\{\theta : 0 \leq \theta \leq \frac{\log \alpha}{\log x}\right\}$$

Thus the interval in part (a) is nonoptimal. A shorter interval with confidence coefficient 0.239 is $\{\theta : 0 \leq \theta \leq] log(1 - 0.239)/\log x\}$.

3. (a) Consider the statistic $p_n(\bar{X}) = P(N(\mu, 1) \geq \sqrt{n}\bar{X}|H_0) = 1 - \Phi(\sqrt{n}\bar{X})$. We have that

$$\begin{aligned} P[p_(\bar{X}) \leq \alpha|H_0] &= P[1 - \Phi(\sqrt{n}\bar{X}) \leq \alpha] \\ &= P[\Phi(\sqrt{n}\bar{X}) \geq 1 - \alpha] \\ &= 1 - (1 - \alpha) = \alpha \end{aligned}$$

since Φ is a continuous CDF and hence distributed as Uniform$(0, 1)$. This proves that p_n is a valid p-value and that the Type I error of the test that rejects H_0 if $p_n(\bar{X}) \leq \alpha$ is equal to α. The type II error of the test is

$$P[p_n(\bar{X}) > \alpha|H_1] = P[\sqrt{n}(\bar{X} - \mu_1) < \Phi^{-1}(1 - \alpha) - \sqrt{n}\mu_1] = \Phi(\Phi^{-1}(1 - \alpha) - \sqrt{n}\mu_1)$$

using the fact that under H_1, $\sqrt{n}(\bar{X} - \mu) \sim N(0, 1)$.

(b) Define $\hat{p}_{10l} = p_{10l}(X_1, X_{10l})$. The type I error is given by

$$
\begin{aligned}
&P(\cup_{l=1}^{L}\{\hat{p}_{10k} > \alpha, k = 1, \cdots, l-1 \; and \; \hat{p}_{10l} \leq \alpha\}) \\
&= \alpha + (1-\alpha) \sum_{l=2}^{L} P(\{\hat{p}_{10k} > \alpha, k = 1, \cdots, l-1 \; and \; \hat{p}_{10l} \leq \alpha\} | \hat{p}_{10} > \alpha)
\end{aligned}
$$

In this case, computing the type I error is more challenging because the statistics $\hat{p}_{10l}$ are not independent. The same is true for the type-II error. Let $p_{10}^{(l)} = p_{10}(\bar{X}^l), l = 0, \cdots, L-1$. The type-I error is given by,

$$
P(p_{10}^{(l)} \leq \alpha \; for \; some \; l | H_0) = 1 - P(p_{10}^{(l)} > \alpha \; for \; all \; l | H_0) = 1 - (1-\alpha)^L
$$

where we use that $bar X^{(l)}, l = 0, \cdots, L-1$ are independent. Similarly, the type-II error is equal to $(\Phi(\Phi^{-1}(1-\alpha) - \sqrt{n}\mu_1))^L$. The expected number of samples collected is given by

$$
10 \sum_{l=0}^{L-1} (l+1) P(p_{10} \leq \alpha) [P(p_{10} > \alpha)]^l + 10L[P(p_{10} > \alpha)]^L
$$