ECON550: Problem Set 10

Nicholas Wu

Fall 2020

Problem 1

4.5.1 From the book, the power function is

$$\Phi\left(-z_{1-\alpha} - \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma}\right)$$

Note that the expression inside the Φ function has a positive coefficient on μ, and hence is increasing. Since Φ is an increasing transformation, this is an increasing transformation of an increasing function, and hence the power function is increasing in μ. Therefore, the the sup of the power function just occurs at μ_0, and hence the size is

$$\Phi(-z_{1-\alpha}) = \alpha$$

4.6.4 The critical region for the one-sided test is

$$C_1 = \left\{\frac{\sqrt{n}(\bar{x} - \mu)}{s} \geq t_{n-1,1-\alpha}\right\}$$

Note that by Neyman-Pearson, this is equivalent to a likelihood ratio test $\mathcal{L}(\theta_0, x)/\mathcal{L}(\theta_1, x) \leq k_\alpha$.

For the two-sided test:

$$C_2 = \left\{\left|\frac{\sqrt{n}(\bar{x} - \mu)}{s}\right| \geq t_{n-1,1-\alpha/2}\right\}$$

Let γ_1 denote the power of the one-sided test, γ_2 the power of the two-sided test. Then we have

$$\gamma_1 - \gamma_2 = \int_{C_1} \mathcal{L}(\theta_1, x)\ dx - \int_{C_2} \mathcal{L}(\theta_1, x)\ dx$$

$$= \int_{C_1 \cap C_2^c} \mathcal{L}(\theta_1, x)\ dx - \int_{C_2 \cap C_1^c} \mathcal{L}(\theta_1, x)\ dx$$

Applying $\mathcal{L}(\theta_0, x)/\mathcal{L}(\theta_1, x) \leq k_\alpha$, we get

$$\gamma_1 - \gamma_2 > \frac{1}{k_\alpha}\left(\int_{C_1 \cap C_2^c} \mathcal{L}(\theta_0, x)\ dx - \int_{C_2 \cap C_1^c} \mathcal{L}(\theta_0, x)\ dx\right) = \frac{1}{k_\alpha}\left(\int_{C_1} \mathcal{L}(\theta_0, x)\ dx - \int_{C_2} \mathcal{L}(\theta_0, x)\ dx\right) = \frac{1}{k_\alpha}(\alpha - \alpha) = 0$$

(strict inequality because the measure on the desired regions are positive). So for $\mu > \mu_0$, the one-sided test has larger power.

4.6.8 Let p' denote the proportion of drivers who wear seatbelts after the ad campaign.

(a) The null hypothesis: $p' = p = 0.14$. Alternative hypothesis: $p' > 0.14$.

(b) Let $\hat{p}$ be the proportion of the drivers sampled wearing a seat belt. Assuming uniformly random sampling, by the CLT, $\sqrt{n}(\hat{p} - p')/\sqrt{p'(1-p')} \to N(0,1)$. So we take the test statistic:

$$T = \frac{\sqrt{n}(\hat{p} - 0.14)}{\sqrt{0.14(0.86)}} \approx 2.539$$

The critical region for $\alpha = 0.01$ of this statistic is $\{T : T \geq z_{0.99} = 2.32635\}$

(c) We reject at the $\alpha = 0.01$ level, and the p-value is approximately 0.005558. It is likely the campaign was successful.

4.6.5

(a) The test statistic is $\sqrt{16}(10.4 - 10.1)/(0.4) = 3 > 1.753$. So we reject at the 5% significance level.

(b) The p-value is approximately 0.0045.

Problem 2

Let the observed statistic value be T.

(a) For the two-tailed test, we find p such that $z_{1-p/2} = |T| = 2.6$. Using a computational aid, $p = 0.009322$.

(b) This time, we take p such that $z_{1-p/2} = |T| = 1.96$. $p = 0.05$

(c) If we only use the one-tailed test, we need to find p such that $z_{1-p} = T$. For $T = -2.6$, $p = 0.995$. For $T = 1.96$, $p = 0.025$.

Problem 3

(a) Note that g is one-to-one and onto from $(-1,1)$ and $\mathbb{R}$, so $\rho = \rho_0$ is an equivalent statement to $g(\rho) = g(\rho_0)$. Then from the previous problem set, we know

$$\sqrt{n}(g(\hat{\rho}) - g(\rho_0)) \to N(0,1)$$

So our test statistic is $T = \sqrt{n}(g(\hat{\rho}) - g(\rho_0))$ and we reject if $|T| > z_{1-\alpha/2}$.

(b) We know then that

$$\left[g(\hat{\rho}) - \frac{z_{1-\alpha/2}}{\sqrt{n}}, g(\hat{\rho}) + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right]$$

is a $1-\alpha$ confidence interval for $g(\rho)$. Using the bijective property of g, we can recover a $1-\alpha$ confidence interval for ρ:

$$\left[g^{-1}\left(g(\hat{\rho})-\frac{z_{1-\alpha/2}}{\sqrt{n}}\right), g^{-1}\left(g(\hat{\rho})+\frac{z_{1-\alpha/2}}{\sqrt{n}}\right)\right]$$

Problem 4

8.1.2 The UMP test is the likelihood ratio test by Neyman-Pearson. The likelihood ratio condition for the critical region is given by

$$\frac{\frac{1}{4}e^{-(x_1+x_2)/2}}{\frac{1}{16}e^{-(x_1+x_2)/4}} \leq k$$

$$e^{-(x_1+x_2)/4} \leq \frac{k}{4}$$

$$-(x_1+x_2)/4 \leq \ln\left(\frac{k}{4}\right)$$

$$x_1+x_2 \geq -4\ln\left(\frac{k}{4}\right)$$

So the UMP test just uses x_1+x_2.

8.1.5 By Neyman-Pearson, the UMP is the likelihood ratio test with critical region:

$$\left\{\{x_i\}, \frac{1}{2^n\prod_i x_i} \leq k_n\right\}$$

Equivalently, the condition is

$$\prod_i x_i \geq \frac{1}{2^n k_n}$$

as desired.

8.1.7 We again invoke Neyman-Pearson and consider the likelihood ratio test with critical region given by the condition:

$$\frac{\exp\left(-\frac{1}{200}\sum_i(x_i-75)^2\right)}{\exp\left(-\frac{1}{200}\sum_i(x_i-78)^2\right)} \leq k_n$$

$$\exp\left(-\frac{1}{200}\sum_i(x_i-75)^2+\frac{1}{200}\sum_i(x_i-78)^2\right) \leq k_n$$

$$\sum_i(x_i-78)^2-(x_i-75)^2 \leq 200\ln k_n$$

$$\sum_i(2x_i-153)(-3) \leq 200\ln k_n$$

$$\sum_i x_i \geq \frac{153n}{2}-\frac{100}{3}\ln k_n$$

$$\bar{x} \geq \frac{153}{2} - \frac{100}{3n} \ln k_n$$

as desired.

8.2.3

$$\begin{aligned}
\gamma(\theta) &= P_\theta(\bar{X}_n \geq 3/5) \\
&= P_\theta(\sum X_i \geq 3n/5) \\
&= P_\theta(N(n\theta, 4n) \geq 3n/5) \\
&= P_\theta\left(N(0,1) \geq \frac{3\sqrt{n}}{10} - \frac{\theta\sqrt{n}}{2}\right) \\
&= 1 - \Phi\left(\frac{3\sqrt{n}}{10} - \frac{\theta\sqrt{n}}{2}\right) \\
&= 1 - \Phi\left(\frac{3 - 5\theta}{2}\right)
\end{aligned}$$

8.2.4 Since x, y are independent, $\bar{x} - \bar{y} \sim N(\theta, 625/n)$. Take the test cutoff value as k. The power function is then

$$\gamma(\theta) = 1 - \Phi\left(\frac{k - \theta}{25/\sqrt{n}}\right)$$

$$\gamma(0) = 1 - \Phi\left(\frac{k\sqrt{n}}{25}\right) = 0.05$$

$$\gamma(10) = 1 - \Phi\left(\frac{k - 10}{25/\sqrt{n}}\right) = 0.9$$

Solving, $n \approx 54$ and $k = 5.60$.

8.2.7 $\bar{X}_n$ is distributed as $N(\theta, 4)$. Then $(\bar{X}_n - 75)/2 \geq 1.28155$ is the desired critical region.

8.3.5 The Neyman-Pearson test critical region is

$$C_1 = \left\{(x_1, ...x_n) \ : \ \frac{\mathcal{L}(\theta_0; x_1, ...x_n)}{\mathcal{L}(\theta_1; x_1, ...x_n)} \leq k_{\alpha,1}\right\}$$

The likelihood ratio principle critical region is

$$C_2 = \left\{(x_1, ...x_n) \ : \ \frac{\mathcal{L}(\theta_0; x_1, ...x_n)}{\max\left(\mathcal{L}(\theta_0; x_1, ...x_n), \mathcal{L}(\theta_1; x_1, ...x_n)\right)} \leq k_{\alpha,2}\right\}$$

However, $k_{\alpha,2} < 1$, which implies that

$$\mathcal{L}(\theta_0; x_1, ...x_n) < \max\left(\mathcal{L}(\theta_0; x_1, ...x_n), \mathcal{L}(\theta_1; x_1, ...x_n)\right)$$

which implies that $\mathcal{L}(\theta_1; x_1, ...x_n) > \mathcal{L}(\theta_0; x_1, ...x_n)$, so

$$\max\left(\mathcal{L}(\theta_0; x_1, ...x_n), \mathcal{L}(\theta_1; x_1, ...x_n)\right) = \mathcal{L}(\theta_1; x_1, ...x_n)$$

Hence

$$C_2 = \left\{(x_1, ...x_n) \ : \ \frac{\mathcal{L}(\theta_0; x_1, ...x_n)}{\mathcal{L}(\theta_1; x_1, ...x_n)} \leq k_{\alpha,2}\right\}$$

Taking $k_{\alpha,2} = k_{\alpha_1}$ to match size, we have $C_1 = C_2$.

8.3.7 Since θ_1 is unspecified, the likelihood maximizing estimate for θ_1 is $\bar{X}_n$. The likelihood maximizing estimate for θ_2 is $n^{-1}\sum(X_i - \bar{X}_n)^2$. So the likelihood decision rule is given by

$$\frac{\mathcal{L}(\bar{X}_n, \theta_2; x_1...x_n)}{\mathcal{L}(\bar{X}_n, n^{-1}\sum(X_i - \bar{X}_n)^2; x_1, ...x_n)} \leq k_\alpha$$

$$\frac{\theta_2^{-n/2}\exp\left(-\frac{\sum(X_i - \bar{X}_n)^2}{2\theta_2}\right)}{\left(n^{-1}\sum(X_i - \bar{X}_n)^2\right)^{-n/2}\exp\left(-\frac{\sum(X_i - \bar{X}_n)^2}{2n^{-1}\sum(X_i - \bar{X}_n)^2}\right)} \leq k_\alpha$$

Denote $\hat{\theta}_2 = n^{-1}\sum(X_i - \bar{X}_n)^2$. Then the condition becomes

$$\left(\frac{\theta_2}{\hat{\theta}_2}\right)^{-n/2}\exp\left(\frac{n}{2}\left(1 - \frac{\hat{\theta}_2}{\theta_2}\right)\right) \leq k_\alpha$$

$$-\frac{n}{2}\ln\left(\frac{\theta_2}{\hat{\theta}_2}\right) + \frac{n}{2}\left(1 - \frac{\hat{\theta}_2}{\theta_2}\right) \leq \ln k_\alpha$$

$$\ln\left(\hat{\theta}_2\right) - \frac{\hat{\theta}_2}{\theta_2} \leq \frac{2}{n}\ln k_\alpha + \ln\theta_2 - 1$$

At a fixed k_α, the LHS is concave in $\hat{\theta}_2$, with a maximum at θ_2, so there will be some $k_1 < k_2$, such that for $\hat{\theta}_2 < k_1$ the condition fails and we reject, or $\hat{\theta}_2 > k_2$ the condition also fails and we reject. Taking $c_1 = nk_1$, $c_2 = nk_2$ gives us the desired test condition.