Statistics 252 "Practice Exam" (Solutions) – Winter 2006

1. (a) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, since $\overline{Y} \sim \mathcal{N}(\mu, 9/n)$,

$$0.05 = P_{H_0}(\text{reject } H_0) = P(\overline{Y} > c|\mu = 0) = P\left(\frac{\overline{Y} - 0}{3/\sqrt{n}} > \frac{c - 0}{3/\sqrt{n}}\right) = P(Z > c\sqrt{n}\,/\,3),$$

where $Z \sim \mathcal{N}(0, 1)$. From Table 4, we find that $P(Z > 1.65) = 0.05$. We, therefore, must have

$$\frac{c\sqrt{n}}{3} = 1.96 \quad \text{or} \quad c = \frac{4.95}{\sqrt{n}}.$$

1. (b) By definition, the power of an hypothesis test is the probability under H_A that H_0 is rejected. Hence, when $\mu = 1$, $n = 36$, we find $c = 0.825$, so that

$$\text{power} = P(\overline{Y} > 0.825|\mu = 1) = P\left(\frac{\overline{Y} - 1}{3/\sqrt{36}} > \frac{0.825 - 1}{3/\sqrt{36}}\right) = P(Z > -0.36) = 1 - 0.3594 = 0.6406$$

where $Z \sim \mathcal{N}(0, 1)$. (The last step follows from Table 4.)

1. (c) As in **(a)** and **(b)**,

$$\text{power} = P\left(\overline{Y} > \frac{4.95}{\sqrt{n}}|\mu = 1\right) = P\left(Z > \frac{4.95/\sqrt{n} - 1}{3/\sqrt{n}}\right) = P(Z > 1.65 - \sqrt{n}/3).$$

Hence, as n increases ($\to \infty$), $1.65 - \sqrt{n}/3$ decreases monotonically ($\to -\infty$), so that the the power increases monotonically ($\to 1$). In particular, if $m > n$, then

$$P(Z > 1.65 - \sqrt{n}/3\,) < P(Z > 1.65 - \sqrt{m}/3\,).$$

This indeed makes sense intuitively. As the sample size increases, it becomes easier to detect that $\mu = 1$ is false.

2. (a) By definition, the significance level α is the probability of a Type I error; that is, the probability under H_0 that H_0 is rejected. Hence, since $\overline{Y} \sim \mathcal{N}(\mu, 4/n)$,

$$\begin{aligned}\alpha = P_{H_0}(\text{reject } H_0) = P(\overline{Y} > 3.92/\sqrt{n}|\mu = 0) &= P\left(\frac{\overline{Y} - 0}{2/\sqrt{n}} > \frac{3.92/\sqrt{n} - 0}{2/\sqrt{n}}\right)\\ &= P(Z > 1.96) = 0.025,\end{aligned}$$

where $Z \sim \mathcal{N}(0, 1)$. (The last step follows from Table 4.) Hence, we see that the hypothesis test does, in fact, have significance level $\alpha = 0.025$.

2. (b) By definition, the power of an hypothesis test is the probability under H_A that H_0 is rejected. Hence, when $\mu = 0.5$, we find

$$\begin{aligned}\text{power} = P_{H_A}(\text{reject } H_0) = P(\overline{Y} > 3.92/\sqrt{n}|\mu = 0.5) &= P\left(\frac{\overline{Y} - 0.5}{2/\sqrt{n}} > \frac{3.92/\sqrt{n} - 0.5}{2/\sqrt{n}}\right)\\ &= P(Z > 1.96 - 0.25\sqrt{n}\,)\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$. If we desire the test to have power 0.9, then using Table 4, we find $P(Z > -1.28) = 0.90$. Thus, we require that n satisfy

$$1.96 - 0.25\sqrt{n} = -1.28 \quad \text{or} \quad n \approx 168.$$

(In fact, we can take $n \geq 168$ to guarantee that the test will have power (at least) 0.9 when $\mu = 0.5$.

3. Draw a picture! From the scenario presented, we know that John rejects H_0 iff $p \leq 0.01$, and that George rejects H_0 iff $p \leq 0.05$. Since Ringo's p-value is smaller than 0.03, we can conclude immediately that George will reject the null hypothesis. However, John cannot make a decision. We are only told that Ringo's p-value is smaller than 0.03. We do not know, therefore, how it compares to John's desired significance level of $\alpha = 0.01$. (It could be the case that $0.01 < p < 0.03$ or it could be the case that $p < 0.01 < 0.03$. These yield different conclusions for John.)

4. Consider an hypothesis test of $H_0 : \theta = \theta_0$ against H_A where H_A could be any one of $H_A : \theta \neq \theta_0$, $H_A : \theta > \theta_0$, or $H_A : \theta < \theta_0$. The significance level α is simply the probability of a Type I error. A Type I error occurs if H_0 is rejected when, in fact, H_0 is true. Thus,

$$\alpha = P(\text{Type I error}) = P_{H_0}(\text{reject } H_0).$$

5. In this problem, we find that $\alpha = P(\overline{Y} < c|\mu = 0)$ and $\beta = P(\overline{Y} > c|\mu = -1/2)$. Since $\overline{Y} \sim \mathcal{N}(\mu, \sigma^2/n) = \mathcal{N}(\mu, 0.25)$, we conclude that

$$\alpha = P(\overline{Y} < c|\mu = 0) = P\left(\frac{\overline{Y} - 0}{\sqrt{0.25}} < \frac{c - 0}{\sqrt{0.25}}\right) = P(Z < 2c)$$

and

$$\beta = P(\overline{Y} > c|\mu = -1/2) = P\left(\frac{\overline{Y} + 1/2}{\sqrt{0.25}} > \frac{c + 1/2}{\sqrt{0.25}}\right) = P(Z > 2c + 1)$$

where $Z \sim \mathcal{N}(0,1)$. In order for $\alpha = \beta$, we require that $P(Z < 2c) = P(Z > 2c + 1)$. Since the standard normal distribution is symmetric about 0, we see that we must have $-2c = 2c + 1$ or $c = -1/4$. (DRAW A PICTURE TO SEE WHERE THE MINUS SIGN COMES FROM!) Consulting Table 4, we find that with $c = -1/4$, the significance level of the this test is

$$\alpha = P(Z < -1/2) = 0.3085.$$

6. (a) The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(y_i|\theta) = (\theta - 1)^n \left(\prod_{i=1}^{n} y_i\right)^{-\theta}$$

so that the log-likelihood function is

$$\ell(\theta) = n\log(\theta - 1) - \theta \sum_{i=1}^{n} \log(y_i).$$

Hence $\ell'(\theta) = 0$ implies

$$0 = \frac{n}{\theta - 1} - \sum_{i=1}^{n} \log(y_i).$$

Since

$$\ell''(\theta) = -\frac{n}{(\theta-1)^2} < 0,$$

we conclude that

$$\hat{\theta}_{\text{MLE}} = 1 + \frac{n}{\sum_{i=1}^{n} \log(Y_i)}.$$

6. (b) If we let $u = \prod_{i=1}^{n} y_i$, then we can write

$$L(\theta) = g(u, \theta) \cdot h(y_1, \ldots, y_n)$$

where

$$h(y_1, \ldots, y_n) = 1 \quad \text{and} \quad g(u, \theta) = (\theta - 1)^n \left(\prod_{i=1}^{n} y_i \right)^{-\theta}$$

so by the Factorization Theorem we conclude that

$$\prod_{i=1}^{n} Y_i$$

is sufficient for θ. Recall that any one-to-one function of a sufficient statistic is also sufficient. Therefore, if we let

$$T(U) = 1 + \frac{n}{\log U},$$

then since T is one-to-one, we find that

$$T\left(\prod_{i=1}^{n} Y_i\right) = 1 + \frac{n}{\sum_{i=1}^{n} \log(Y_i)} = \hat{\theta}_{\text{MLE}}$$

is sufficient for θ.

6. (c) Since

$$\log f(y|\theta) = \log(\theta - 1) - \theta \log(y),$$

we find

$$\frac{\partial^2}{\partial \theta^2} \log f(y|\theta) = \frac{-1}{(\theta-1)^2}.$$

Thus,

$$I(\theta) = -\mathbb{E}\left(\frac{\partial^2}{\partial \theta^2} \log f(Y|\theta) \right) = \frac{1}{(\theta-1)^2}.$$

6. (d) The rejection region of a significance level 0.10 test of $H_0 : \theta = 8$ vs. $H_A : \theta \neq 8$ based on the Fisher information and the MLE is

$$\left\{ \sqrt{nI(\hat{\theta}_{\text{MLE}})} \left| \hat{\theta}_{\text{MLE}} - \theta_0 \right| \geq z_{0.05} \right\}$$

or

$$\left\{ \left| \sqrt{n} - \frac{7 \sum_{i=1}^{n} \log(Y_i)}{\sqrt{n}} \right| \geq 1.65 \right\}.$$

Since $n = 25$ and $\sum \log y_i = 5$, we find

$$\left| \sqrt{n} - \frac{7 \sum_{i=1}^{n} \log(y_i)}{\sqrt{n}} \right| = \left| 5 - \frac{7 \cdot 5}{5} \right| = 2 > 1.65,$$

and so at the $\alpha = 0.10$ level, we reject the hypothesis $H_0 : \theta = 8$ in favour of $H_A : \theta \neq 8$.

Alternative Solution: An approximate 90% confidence interval for θ is given by

$$\hat{\theta}_{\text{MLE}} \pm 1.65 \, \frac{1}{\sqrt{n \, I(\hat{\theta}_{\text{MLE}})}}.$$

Since $n = 25$ and $\sum \log y_i = 5$, we conclude that

$$\hat{\theta}_{\text{MLE}} = 1 + \frac{25}{5} = 6$$

and

$$I(\hat{\theta}_{\text{MLE}}) = \frac{1}{(\hat{\theta}_{\text{MLE}} - 1)^2} = \frac{1}{25}.$$

Hence, an approximate 90% confidence interval for θ is

$$6 \pm 1.65.$$

We see that since 8 does NOT lie in the confidence interval, the hypothesis test–confidence interval duality allows us to reject the hypothesis $H_0 : \theta = 8$ in favour of $H_A : \theta \neq 8$ at the $\alpha = 0.10$ level.