Statistics 252 Winter 2007 Midterm #1 – Solutions

1. (a) Since $B(\hat{\theta}) = E(\hat{\theta}) - \theta$, we must first compute $E(\hat{\theta})$. To determine $E(\hat{\theta})$, we need to find the density function of $\hat{\theta}$, which requires us first to find the distribution function of $\hat{\theta}$. Since $Y_1, \ldots, Y_n$ are i.i.d., we find

$$P(\hat{\theta} > x) = P(\min\{Y_1, \ldots, Y_n\} > x) = P(Y_1 > x, \ldots, Y_n > x) = [P(Y_1 > x)]^n.$$

We know the density of Y_1, and so we compute

$$P(Y_1 > x) = \int_x^\infty f_Y(y)\,dy = \int_x^\infty 2\theta^2 y^{-3}\,dy = \theta^2 x^{-2}.$$

Therefore, we find

$$P(\hat{\theta} > x) = [P(Y_1 > x)]^n = \theta^{2n} x^{-2n} \quad \text{for } x \geq \theta$$

and so the distribution function for $\hat{\theta}$ is

$$F_{\hat{\theta}}(x) = P(\hat{\theta} \leq x) = 1 - P(\hat{\theta} > x) = 1 - \theta^{2n} x^{-2n}$$

for $x \geq \theta$. Finally, we differentiate to conclude that the density function for $\hat{\theta}$ is

$$f_{\hat{\theta}}(x) = F'_{\hat{\theta}}(x) = 2n\theta^{2n} x^{-2n-1} \quad \text{for } x \geq \theta.$$

Now we can determine $E(\hat{\theta})$ via

$$E(\hat{\theta}) = \int_{-\infty}^\infty x \cdot f_{\hat{\theta}}(x)\,dx = 2n\theta^{2n} \int_\theta^\infty x^{-2n}\,dx = 2n\theta^{2n} \cdot \frac{\theta^{-2n+1}}{2n-1} = \frac{2n}{2n-1}\,\theta. \qquad (*)$$

Hence, the bias of $\hat{\theta}$ is given by

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta = \frac{2n}{2n-1}\,\theta - \theta = \frac{\theta}{2n-1}.$$

(b) From $(*)$, we conclude that

$$c = \frac{2n-1}{2n}.$$

(c) As for the mean-square error, we have from a result in class that $\mathrm{MSE}(\hat{\theta}_1) = \mathrm{Var}(\hat{\theta}_1) + [B(\hat{\theta}_1)]^2$. Since $\hat{\theta}_1$ is unbiased, $\mathrm{MSE}(\hat{\theta}_1) = \mathrm{Var}(\hat{\theta}_1)$. Since

$$\mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}\left(\frac{2n-1}{2n}\hat{\theta}\right) = \left(\frac{2n-1}{2n}\right)^2 \mathrm{Var}(\hat{\theta})$$

we must compute $\mathrm{Var}(\hat{\theta})$. Therefore,

$$E(\hat{\theta}^2) = \int_{-\infty}^\infty x^2 \cdot f_{\hat{\theta}}(x)\,dx = 2n\theta^{2n} \int_\theta^\infty x^{-2n+1}\,dx = 2n\theta^{2n} \cdot \frac{\theta^{-2n+2}}{2n-2} = \frac{2n}{2n-2}\,\theta^2$$

and so

$$\mathrm{Var}(\hat{\theta}) = E(\hat{\theta}^2) - [E(\hat{\theta})]^2 = \frac{2n}{2n-2}\,\theta^2 - \frac{(2n)^2}{(2n-1)^2}\theta^2.$$

We then conclude that

$$\mathrm{MSE}(\hat{\theta}_1) = \left(\frac{2n-1}{2n}\right)^2 \left[\frac{2n}{2n-2} - \frac{(2n)^2}{(2n-1)^2}\right] \theta^2 = \frac{\theta^2}{4n(n-1)}.$$

2. If $X \sim$Uniform$(0,\theta)$, then

$$f_X(x) = \frac{1}{\theta} \quad \text{for } 0 \le x \le \theta$$

and so

$$F_X(x) = \frac{x}{\theta} \quad \text{for } 0 \le x \le \theta.$$

Since X_1, X_2, X_3, X_4 are i.i.d. Uniform$(0,\theta)$, we conclude

$$P(Y \le y) = P(\max\{Y_1,\ldots,Y_n\} \le y) = P(X_1 \le y,\ldots,X_4 \le y) = [P(X_1 \le y)]^4 = \frac{y^4}{\theta^4}.$$

Thus, the density of Y is

$$f_Y(y) = \frac{4y^3}{\theta^4} \quad \text{for } 0 \le y \le \theta.$$

3. (a) The moment generating function of X_1 is

$$m_{X_1}(t) := E\left(e^{tX_1}\right) = E\left(\exp\{t(Y_1+\cdots+Y_n)\}\right) = E\left(\exp\{tY_1\}\right)\cdots E\left(\exp\{tY\}\right)$$

since the Y_i are i.i.d. We know that

$$m_{Y_1}(t) = E\left(\exp\{tY_1\}\right) = \frac{1}{1-\theta t}$$

since $Y_1 \sim$ Exp(θ), and so we conclude that

$$m_{X_1}(t) = \left[\frac{1}{1-\theta t}\right]^n$$

which is the moment generating function of a Gamma(n,θ) random variable.

(b) Since $Y_1,\ldots,Y_n$ are i.i.d. Exp(θ) random variables, we find

$$P(X_2 > x) = P(\min\{Y_1,\ldots,Y_n\} > x) = P(Y_1 > x,\ldots,Y_n > x) = [P(Y_1 > x)]^n = \left[e^{-x/\theta}\right]^n$$

and so

$$P(X_2 \le x) = 1 - e^{-nx/\theta}$$

which we recognize as the distribution function of an Exp$\left(\frac{\theta}{n}\right)$ random variable.

(c) We find

$$E(\hat{\theta}_1) = \frac{1}{n}E(X_1) = \frac{1}{n}\cdot n\theta = \theta$$

where we have used the fact that the mean of a Gamma(n,θ) random variable is $n\theta$.

(d) We find

$$E(\hat{\theta}_2) = nE(X_2) = n\cdot\frac{\theta}{n} = \theta$$

where we have used the fact that the mean of an Exp(θ/n) random variable is θ/n.

(e) We find

$$\operatorname{Var}(\hat{\theta}_1) = \frac{1}{n^2}\operatorname{Var}(X_1) = \frac{1}{n^2}\cdot n\theta^2 = \frac{\theta^2}{n}$$

where we have used the fact that the variance of a Gamma(n,θ) random variable is $n\theta^2$. We also find

$$\operatorname{Var}(\hat{\theta}_2) = n^2\operatorname{Var}(X_2) = n^2\cdot\frac{\theta^2}{n^2} = \theta^2$$

where we have used the fact that the variance of an Exp(θ/n) random variable is θ^2/n^2. We prefer the unbiased estimator with the smallest variance which, in this case, is $\hat{\theta}_1$.

4. (a) As shown in class,

$$\overline{X} \sim \mathcal{N}\left(\mu_1, \frac{\sigma_1^2}{n}\right) \quad \text{and} \quad \overline{Y} \sim \mathcal{N}\left(\mu_2, \frac{\sigma_2^2}{m}\right).$$

We also showed that the sum of independent normals is again normal, and so

$$\overline{X} - \overline{Y} \sim \mathcal{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right).$$

(b) If

$$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

then $Z \sim \mathcal{N}(0,1)$, and so

$$\begin{aligned} 1-\alpha &= P(-z_{\alpha/2} \le Z \le z_{\alpha/2}) \\ &= P\left(\overline{X} - \overline{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \le \mu_1 - \mu_2 \le \overline{X} - \overline{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right) \end{aligned}$$

Therefore, the required $(1-\alpha)$ confidence interval for $(\mu_1 - \mu_2)$ is

$$\left[\overline{X} - \overline{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \overline{X} - \overline{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right].$$

5. (a) Suppose that $\hat{\theta}$ is an estimator. It is important to determine the sampling distribution of $\hat{\theta}$ in order to analyze it. In particular, we are interested in the bias and the mean-square error of our estimators, and in order to compute these numbers it is necessary to know the distribution of $\hat{\theta}$. (Given the distribution, we can determine the density, and hence compute various moments.) It is also important to know the sampling distribution in order to construct exact confidence intervals. For instance, if a transformation of the estimator yields a distribution which is parameter-free, then this can be used in the construction of confidence intervals via the pivotal method.

(b) Given random variables $Y_1, \ldots, Y_n$, a statistic is simply a single-valued function $g(Y_1, \ldots, Y_n)$ of those random variables. The dual nature of the term arises as follows. Consider an experiment. Before the experiment is performed, the outcome is unknown, and after the experiment is performed, the outcome is, of course, known. Therefore, $g(Y_1, \ldots, Y_n)$ is unknown in advance and is *a priori* a random variable. Once the experiment is performed and the data $y_1, \ldots, y_n$ are known, $g(y_1, \ldots, y_n)$ is simply a number summarizing that data.