2007 FYE Solutions

January 18, 2008

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

1.

$$\begin{aligned}
E(L(\mu,d)) &= \frac{1}{2}\left[E(L(\mu=1,d)) + E(L(\mu=0,d))\right] \\
&= \frac{1}{2}\left[5 \times Pr(\bar{y} \leq b) + 1 \times Pr(b < \bar{y} \leq c) + 0 \times Pr(c < \bar{y})\right] \\
&\quad + \frac{1}{2}\left[0 \times Pr(b \leq \bar{y}) + 1 \times Pr(b \leq \bar{y} \leq c) + 5 \times Pr(c < \bar{y})\right] \\
&= \frac{1}{2}\left[4\Phi((b-1)\sqrt{n}) + \Phi((c-1)\sqrt{n}) - 4\Phi(c\sqrt{n}) - \Phi(b\sqrt{n}) + 5\right]
\end{aligned}$$

Differentiating $E(L(\mu,d))$ with respect c, setting equal to 0, and solving for c yields,

$$\begin{aligned}
0 = \frac{dE(L(\mu,d))}{dc} &= \frac{\sqrt{n}}{2}\phi((c-1)\sqrt{n}) - \frac{4\sqrt{n}}{2}\phi(c\sqrt{n}) \\
\Rightarrow 4\phi(c\sqrt{n}) &= \phi((c-1)\sqrt{n}) \\
\Rightarrow 4\exp\left\{-\frac{nc^2}{2}\right\} &= \exp\left\{-\frac{n(c-1)^2}{2}\right\} \\
\Rightarrow \ln(4) - \frac{nc^2}{2} &= -\frac{n(c-1)^2}{2} \\
\Rightarrow \ln(4) &= cn - \frac{n}{2} \\
\Rightarrow c &= \frac{\ln(4)}{n} + \frac{1}{2}
\end{aligned}$$

In similar fashion, the value for b can be found to be,

$$b = \frac{1}{2} - \frac{\ln(4)}{n}$$

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

2. (a) Using a Taylor expansion of $g(\bar{X}_n)$ about λ,

$$g(\bar{X}_n) = g(\lambda) + g'(\lambda)(\bar{X}_n - \lambda) + \frac{g''(\lambda)}{2}(\bar{X}_n - \lambda)^2 + (\text{Higher Order Terms})$$

Because we want to be correct to order $1/n$, we can ignore the higher order terms. Now take expectations of both sides,

$$\begin{aligned} E(g(\bar{X}_n)) &= E(g(\lambda)) + g'(\lambda)E(\bar{X}_n - \lambda) + \frac{g''(\lambda)}{2}E((\bar{X}_n - \lambda)^2) \\ &= g(\lambda) + 0 + \frac{g''(\lambda)}{2}Var(\bar{X}_n) \\ &= g(\lambda) + \frac{\lambda g''(\lambda)}{2n} \end{aligned}$$

Finally, take the variance of both sides,

$$\begin{aligned} Var(g(\bar{X}_n)) &= Var(g(\lambda)) + g'(\lambda)^2 Var(\bar{X}_n - \lambda) + \left[\frac{g''(\lambda)}{2}\right]^2 Var((\bar{X}_n - \lambda)^2) \\ &= 0 + g'(\lambda)^2\frac{\lambda}{2} + (\text{Higher Order Terms}) \\ &= g'(\lambda)^2\frac{\lambda}{2} \end{aligned}$$

The higher order terms are ignored because they will tend to 0 for large n by the central limit theorem.

(b) Because the support of X_i is $\{0, 1, 2, \ldots\}$, we can just take $g(\bar{X}_n) = Y_n = \bar{X}_n^p$. Now using the results from (a),

$$\begin{aligned} Var(g(\bar{X}_n)) &= g'(\lambda)^2\frac{\lambda}{2} \\ &= p^2\lambda^{2p-2}\frac{\lambda}{2} \\ &= \frac{p^2\lambda^{2p-1}}{n} \end{aligned}$$

which is free of λ if and only if $p = 1/2$. Plugging $p = 1/2$ into the answers from (a) yields,

$$\begin{aligned} Var(g(\bar{X}_n)) &= \frac{1}{4n} \\ E(g(\bar{X}_n)) &= \lambda^{1/2} - \frac{\lambda^{-1/2}}{8n} \end{aligned}$$

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

3. (a) Under the given assumptions and definitions, the implied distribution for $\mathbf{y} = (y_1, \ldots, y_n)'$ is $N(\mathbf{X}\beta, \sigma^2\mathbf{I})$. Additionally, the least squares estimate of β is $\hat{\beta} = (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'y}$. Under the theorems of the multivariate normal distribution, because $\hat{\beta}$ is a linear combination of y, the distribution of $\hat{\beta}$ is normal with expectation,

$$\begin{aligned} E(\hat{\beta}) &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}E(\mathbf{y}) \\ &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'X}\beta \\ &= \beta \\ Var(\hat{\beta}) &= (\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'}Var(\mathbf{Y})\mathbf{X}(\mathbf{X'X})^{-\mathbf{1}} \\ &= \sigma^2(\mathbf{X'X})^{-\mathbf{1}}\mathbf{X'IX}(\mathbf{X'X})^{-\mathbf{1}} \\ &= \sigma^2(\mathbf{X'X})^{-\mathbf{1}} \end{aligned}$$

(b) TRUE. Let $\beta_0 = \alpha_0 - (\bar{\mathbf{y}} + \alpha_1\bar{\mathbf{x}}_1 + \alpha_2\bar{\mathbf{x}}_2)$, $\beta_1 = \alpha_1$, and $\beta_2 = \alpha_2$. Then,

$$\begin{aligned} E(\mathbf{y}) &= \beta_0 + \beta_1\mathbf{x}_1 + \beta_2\mathbf{x}_2 \\ &= \alpha_0 - (\bar{\mathbf{y}} + \alpha_1\bar{\mathbf{x}}_1 + \alpha_2\bar{\mathbf{x}}_2) + \alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 \\ \Rightarrow E(\mathbf{y} - \bar{\mathbf{y}}) &= \alpha_0 + \alpha_1(\mathbf{x} - \bar{\mathbf{x}}_2) + \alpha_2(\mathbf{x} - \bar{\mathbf{x}}_2) \\ \Rightarrow E(\mathbf{y}_c) &= \alpha_0 + \alpha_1\mathbf{x}_{c1} + \alpha_2\mathbf{x}_{c2} \end{aligned}$$

(c) FALSE. Notice that when you expand $E(y_i^2) = (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})^2$ you get terms such as $x_{i1}x_{i2}$ which are not included in the form given in part (c).

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

4. (a)

$$\int_0^t \lambda e^{-\lambda x} dx \quad = \quad 1 - e^{-\lambda t}$$

(b) Because each T_i is independent with constant "success" probability given by (a), we know $X_t \sim Bin(N, 1 - exp\{-\lambda t\})$.

(c) Recall the c.d.f. of the minimum of *independent* random variables is given by,

$$F_{W_1}(t) = 1 - [1 - F_{W_i}(t)]^n$$

where $F_{W_i}(t)$ is the c.d.f of W_i (if you don't know this then it can be derived quite easily). Therefore, the probability distribution for W_1 is given by

$$\begin{aligned} F_{W_1}(t) &= 1 - [1 - (1 - \exp\{-\lambda t\})]^n \\ &= 1 - \exp\{-\lambda n t\} \\ &\sim Ex(n\lambda) \end{aligned}$$

(d) Recall the c.d.f. of a maximum of *independent* random variables is given by,

$$\begin{aligned} F_{W_n}(t) &= [F_{W_n}(t)]^n \\ &= [1 - \exp\{\lambda t\}]^n. \end{aligned}$$

Now, differentiating yeilds the density function,

$$f_{W_n}(t) \quad = \quad \lambda n[1 - \exp\{\lambda t\}]^{n-1} \exp -\lambda t$$

(e) Consider the moment generating function of Y_i,

$$\begin{aligned} E(e^{tY_i}) &= E(e^{t(T_i+U_i)}) \\ &= E(e^{tT_i})E(e^{tU_i}) \\ &= \left(\frac{\lambda}{\lambda - t}\right)^2 \end{aligned}$$

which is the moment generating function of a $Ga(2, \lambda)$ random variable.

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

5. (a)

$$\begin{aligned} E(Y_n) &= E\left(\prod_{i=1}^{n} X_i\right) \\ &= \prod_{i=1}^{n} E(X_i) \\ &= \left(\frac{\theta}{2}\right)^N \end{aligned}$$

(b)

$$\lim_{n\to\infty} Y_n = \lim_{n\to\infty} \exp\left\{ n \left[\frac{1}{n} \sum_{i=1}^{n} \ln(X_i) \right] \right\}$$

By the strong law of large numbers, the quantity $(1/n)\sum_{i=1}^{n} \ln(X_i)$ will converge almost surely to its expectation. Recalling your calculus,

$$\begin{aligned} E(\ln(X_i)) &= \int_0^\theta \ln(X_i)\frac{1}{\theta} dX_i \\ &= X_i \ln(X_i) - X_i|_0^\theta \\ &= \theta\ln(\theta) - \theta. \end{aligned}$$

Therefore,

$$\lim_{n\to\infty} Y_n = \lim_{n\to\infty} \exp\{n[\theta\ln(\theta) - \theta]\}$$

which only exists if $\theta\ln(\theta) - \theta \leq 0$. Therefore, the limit will only exist if $\theta \leq e$.

(c) If $\theta = 1$ then the variable $\xi = \theta$ will dominate the sequence $\{|Y_N|\}$ for all n. Therefore, the Lebesgue's dominated convergence theorem will hold and,

$$E(Y_n) \to 0$$

Notice when $\theta = 2.5$,

$$\begin{aligned} \lim_{n\to\infty} E(Y_n) &= \lim_{n\to\infty} \left(\frac{2.5}{2}\right)^n \\ &= \infty. \end{aligned}$$

Therefore, the Lebesgues dominated convergence theorem will not hold for $\theta = 2.5$ because the expectation does not exist.

These solutions were prepared by students studying to take the 2008 FYE. These solutions have *not* been reviewed by faculty members. Therefore, the solutions presented may contain errors. Use at your own risk.

(a) According to the Neyman-Pearson lemma, the most powerful test will be the test based on the likelihood ratio. Therefore, the likelihood ratio is the test statistic and the p-value is calculated as the probability of observing a test statistic as or more extreme than the observed test statistic under the null hypothesis.

	$x=1$	$x=2$	$x=3$	$x=4$	$x=5$
$\frac{f_0(x)}{f_1(x)}$	5	1/5	2	1/4	1
p-value	1	.05	.50	.15	.30

(b) The power of a test is the probability of rejecting a false null hypothesis for a given rejection region. Notice if the rejection region $\{x = 2\}$ is used then the size of the test is $\alpha = .05$. However using the rejection region $\{x = 2, x = 4\}$ will yield a test of size $\alpha = .05 + .10 = .15$ which is the desired size. The power is then $Pr_{f_1(x)}(x = 2 \text{ or } x = 4) = .25 + .40 = 0.65$.

(c) Just use Bayes theorem to get,

$$\pi(\theta = 0|X = x) \quad = \quad \frac{f_0(x)}{f_0(x) + f_1(x)}$$

which will give the posterior probabilities,

	$x=1$	$x=2$	$x=3$	$x=4$	$x=5$
Posterior Probability	5/6	1/6	2/3	1/5	1/2