18.443 Problem Set 2 Spring 2015
Statistics for Applications
Due Date: 2/20/2015
prior to 3:00pm

Problems from John A. Rice, Third Edition. [*Chapter.Section.Problem*]

1. Problem 8.10.27; instead of 5 components, suppose there are 8 and that the first one fails at 50 days.

   Suppose that certain electronic components have lifetimes that are exponentially distributed

   $$f(t \mid \tau) = (1/\tau)exp(-t/\tau),\ t \geq 0.$$

   Five new components are put on test. The first one fails at 100 days, and no further observations are recorded.

   By the Hint (Example A of Section 3.7), let $T_1, T_2, \ldots, T_n$ be the time until failure of n components (we shall set $n = 8$ below).

   These random variables are i.i.d. with cumulative distribution function

   $$F_T(t) = 1 - exp(-t/\tau)$$

   Let $T_{MIN} = min(T_1, T_2, \ldots, T_n)$ be the shortest time to failure. (If a system operates with components 1-n connected in a series, then T_{MIN} is the time until failure of the sytem.)

   The cumulative distribution function of T_{Min}, $F_{T_{MIN}}(t)$ must satisfy:

   $$\begin{aligned} [1 - F_{T_{MIN}}(t)] &= P(T_{MIN} > t) \\ &= P(T_1 > t, T_2 > t, \ldots, T_n > t) = \prod_{i=1}^{n} P(T_i > t) \\ &= \prod_{j=1}^{n} [1 - F_T(t)] = [1 - F_T(t)]^n \end{aligned}$$

   It follows that the probability density function of T_{MIN} is

   $$\begin{aligned} f_{T_{MIN}}(t) &= -\tfrac{d}{dt}[1 - F_{T_{MIN}}(t)] \\ &= n[1 - F_T(t)]^{n-1} \tfrac{d}{dt}[F_T(t)] \\ &= n[exp(-t/\tau)]^{n-1}(1/\tau)exp(-t/\tau) \\ &= (n/\tau)exp[-t(n/\tau)] \end{aligned}$$

   That is, $T_{MIN} \sim Exponential(rate = n/\tau)$.

   (a). What is the likelihood function of τ? The data consists of $T_{MIN} = 50$ with $n = 8$, so

   $$\begin{aligned} lik(\tau) &= f_{T_{MIN}}(t = 50) = (n/\tau)exp[-t(n/\tau)]|_{t=50,n=8} \\ &= (8/\tau)exp[-50(8/\tau)] \end{aligned}$$

(b). What is the mle of τ?

The mle maximizes $lik(\tau)$ and $\ell(\tau) = \log lik(\tau)$, which is the solution to

$$\begin{aligned} 0 = \ell'(\tau) &= \tfrac{d}{d\tau}\ell(\tau) \\ &= \tfrac{d}{d\tau}[\ln(n/\tau)] + \tfrac{d}{d\tau}[-t(n/\tau)] \\ &= -\tfrac{1}{\tau} - tn(\tfrac{1}{\tau})^2(-1) \\ \Longrightarrow \hat{\tau} &= tn = T_{MIN} \times n = (50 \times 8) = 400. \end{aligned}$$

(c). What is the sampling distribution of the mle?

The mle is $\hat{\tau} = T_{MIN} \times n$ which is n times the minimum of n i.i.d. $Exponential(rate = 1/\tau)$ random variables. The distribution of T_{MIN} is $Exponential(rate = 5/\tau)$ The mle is 5 the minimum of n exponential

$$\begin{aligned} 1 - F_{\hat{\tau}}(u) &= P(\hat{\tau} \geq u) \\ &= P(T_{MIN} \times n \geq u) = P(T_{MIN} \geq u/n) \\ &= [1 - F_T(u/n)]^n \\ &= [exp(-(u/n)/\tau]^n \\ &= [exp(-u/\tau)] \end{aligned}$$

So $F_{\hat{\tau}}(u) = 1 - exp(-u/\tau)$ which is the cdf of an $Exponential(rate = 1/\tau)$ random variable.

(d). What is the standard error of the mle?

The standard error of the mle is the formula for the square root of the variance of the mle, plugging in the mle estimate for the value of the true parameter

$$\begin{aligned} StError(\hat{\tau}) &= \sqrt{Var(\hat{\tau})}|_{\tau=\hat{\tau}} \\ &= \sqrt{\tau^2}|_{\tau=\hat{\tau}} = \hat{\tau} = 400. \end{aligned}$$

The variance formula follows from p. A2 Appendix A for

$Gamma(\alpha = 1, \lambda = 1/\tau)$, which is $Exponential(rate = 1/\tau)$.

2. Problem 8.10.31; answer if George observes three heads and Hilary spins the coin five times.

   George spins a coin three times and observes no heads. He then gives the coin to Hilary. She spins it until the first head occurs, and ends up spinning it four times total. Let θ be the probabiity the coin comes up heads.

   (a). What is the likelihood function?

   Let X_1, X_2, X_3 be the outcomes of George's 3 spins (1=Head, 0=Tail). and Let Y be the number of spins Hilary makes until the first head occurs.

- X_1, X_2, X_3 are i.i.d. (independent and identically distributed) $Bernoulli(\theta)$ random variables
- Y is independent of X_1, X_2, X_3 and has a geometric distribution with parameter $p = \theta$ (see p. A1 of Appendix A)

The likelihood function is the value of the joint pmf function for (X_1, X_2, X_3, Y) as a function of θ for fixed outcome

$$(X_1, X_2, X_3, Y) = (1, 1, 1, 5)$$

Thus:

$$\begin{aligned}
lik(\theta) &= f(x_1, x_2, x_3, y \mid \theta) \\
&= f(x_1 \mid \theta) f(x_2 \mid \theta) f(x_3 \mid \theta) f(y \mid \theta) \\
&= [\prod_{i=1}^{3} \theta^{x_i} (1-\theta)^{1-x_i}] \times [\theta(1-\theta)^{(y-1)}] \\
&= \theta^4 (1-\theta)^4
\end{aligned}$$

(b). What is the MLE of θ?

The MLE solves

$$\begin{aligned}
0 = \ell'(\theta) &= \frac{d}{d\theta} \log[lik(\theta)] \\
&= \frac{d}{d\theta}(4\ln[\theta] + 4\ln[(1-\theta)]) \\
&= \frac{4}{\theta} + \frac{4}{1-\theta} \times (-1) \\
\Longrightarrow \ 4\theta &= 4(1-\theta) \\
\Longrightarrow \hat{\theta} &= 4/8
\end{aligned}$$

Note that the MLE of θ is identical to the MLE for 8 Bernoulli trials which have 4 Heads and 4 Tails.

3. Problem 8.10.53

Let $X_1, \ldots, X_n$ be i.i.d. uniform on $[0, \theta]$

(a). Find the method of moments estimate of θ and its mean and variance.

The first moment of each X_i is

$$\begin{aligned}
\mu_1 &= E[X_i \mid \theta] = \int_0^{\theta} x f(x \mid \theta) dx \\
&= \int_0^{\theta} x \frac{1}{\theta} dx = \frac{1}{\theta}[x^2/2]|_{x=0}^{\theta} \\
&= \frac{1}{\theta}[\theta^2/2] = \theta/2
\end{aligned}$$

The method of moments estimate solves

$$\overline{x} = \frac{1}{n}\sum_1^n x_i = \mu_1 = \theta/2$$

So

$$\hat{\theta}_{MOM} = 2\overline{x}.$$

The mean and variance of $\hat{\theta}_{MOM} = 2\overline{x}$ can be computed from the mean and variance of the i.i.d. X_i:

$$\begin{aligned}
\mu_1 &= E[X_i] = \theta/2 \text{ (shown above)} \\
\sigma^2 &= \mu_2 - (\mu_1)^2 = E[X_i^2] - E[X_i]^2 \\
&= \int_0^\theta x^2 f(x \mid \theta)dx - [\theta/2]^2 \\
&= \frac{1}{\theta}[x^3/3]|_{x=0}^{x=\theta} - [\theta/2]^2 \\
&= [\theta^2/3] - [\theta^2/4] = \theta^2/12
\end{aligned}$$

It follows that

$$\begin{aligned}
E[\hat{\theta}_{MOM}] &= E[2\overline{X}] = 2 \times \frac{1}{n}\sum_1^n E[X_i] = 2 \times \frac{n}{n}\theta = \theta \\
Var[\hat{\theta}_{MOM}] &= Var[2\overline{X}] = 4 \times Var[\overline{X}] \\
&= 4 \times Var[X_i]/n \\
&= 4 \times (\theta^2/12)/n = \theta^2/(3n).
\end{aligned}$$

(b). Find the mle of θ.

The mle $\hat{\theta}_{MLE}$ maximizes the likelihood function

$$\begin{aligned}
lik(\theta) &= f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^n [\frac{1}{\theta} \times \mathbf{1}_{[0,\theta]}(x_i)] \\
&= \frac{1}{\theta^n}\mathbf{1}_{[0,\theta]}(max(x_1, \ldots, x_n)) \\
&= \frac{1}{\theta^n}\mathbf{1}_{[max(x_1,\ldots,x_n),\infty)}(\theta)
\end{aligned}$$

Where $\mathbf{1}_{[a,b]}(x) = \begin{cases} 1, & if \quad x \in [a,b] \\ 0, & if \quad x \notin [a,b] \end{cases}$

The likelihood function is maximized by the minimizing θ. Since $\theta \geq max(x_1, \ldots, x_n)$ the likelihood is maximized with

$$\hat{\theta} = max(x_1, \ldots, x_n)$$

(c). Find the probability density of the mle, and calculate its mean and variance. Compare the variance, the bias, and the mean squared error to those of the method of moments estimate.

The probability density of the mle is the probability density of the maximum member of the sample $X_1, \ldots, X_n$

We compute the cdf (cumulative distribution function) of the mle first.

$$\begin{aligned}
F_{\hat{\theta}_{MLE}}(t) &= P(\hat{\theta}_{MLE} \leq t) \\
&= P(max(X_1, \ldots, X_n) \leq t) = P(X_1 \leq t, X_2 \leq t, \ldots X_n \leq t) \\
&= [P(X_i \leq t)]^n = [\frac{t}{\theta}]^n
\end{aligned}$$

The density of $\hat{\theta}_{MLE}$ is the derivative of the cdf:

$$\begin{aligned}
f_{\hat{\theta}_{MLE}}(t) &= \frac{d}{dt}F_{\hat{\theta}_{MLE}}(t) = n[\frac{t}{\theta}]^{n-1}[\frac{1}{\theta}] \\
&= \frac{nt^{n-1}}{\theta^n}, \ \ 0 < t < \theta
\end{aligned}$$

The mean and variance of $\hat{\theta}_{MLE}$ can be computed directly

$$
\begin{aligned}
E[\hat{\theta}_{MLE}] &= \textstyle\int_0^\theta [t f_{\hat{\theta}_{MLE}}(t)]dt \\
&= \int_0^\theta [t\frac{nt^{n-1}}{\theta^n}]dt \\
&= \frac{n}{\theta^n}[\frac{t^{n+1}}{n+1}]|_{t=0}^{t=\theta} \\
&= \frac{n}{\theta^n}[\frac{\theta^{n+1}}{n+1}] \\
&= \frac{n}{n+1}\theta
\end{aligned}
$$

$$
\begin{aligned}
Var[\hat{\theta}_{MLE}] &= E[\hat{\theta}^2_{MLE}] - (E[\hat{\theta}_{MLE}])^2 \\
&= \textstyle\int_0^\theta [t^2 f_{\hat{\theta}_{MLE}}(t)]dt - (E[\hat{\theta}_{MLE}])^2 \\
&= \int_0^\theta [t^2\frac{nt^{n-1}}{\theta^n}]dt \\
&= \frac{n}{\theta^n}[\frac{t^{n+2}}{n+2}]|_{t=0}^{t=\theta} - (E[\hat{\theta}_{MLE}])^2 \\
&= \frac{n}{\theta^n}[\frac{\theta^{n+2}}{n+2}] - (E[\hat{\theta}_{MLE}])^2 \\
&= \frac{n}{n+2}\theta^2 - (E[\hat{\theta}_{MLE}])^2 \\
&= \frac{n}{n+2}\theta^2 - [\frac{n}{n+1}\theta]^2 \\
&= \theta^2 \times [\tfrac{n}{n+2} - \tfrac{n^2}{(n+1)^2}] \\
&= \theta^2 \times \tfrac{n}{(n+2)(n+1)^2}
\end{aligned}
$$

Now, to compare the variance, the bias, and the mean squared error to those of the method of moments estimate, we apply the formulas:

$$
\begin{aligned}
Bias(\hat{\theta}) &= E[\hat{\theta}] - \theta \\
MSE(\hat{\theta}) &= Var(\hat{\theta}) + [Bias(\hat{\theta})]^2
\end{aligned}
$$

So, we can compare:

$$
\begin{aligned}
Var(\hat{\theta}_{MLE}) &= \theta^2 \times \tfrac{n}{(n+2)(n+1)^2} \\
Bias(\hat{\theta}_{MLE}) &= E[\hat{\theta}_{MLE}] - \theta = -\tfrac{1}{n+1}\theta \\
MSE(\hat{\theta}_{MLE}) &= Var(\hat{\theta}_{MLE}) + [Bias(\hat{\theta}_{MLE})]^2 \\
&= \theta^2 \times [\tfrac{n}{(n+2)(n+1)^2} + (-\tfrac{1}{n+1})^2] \\
&= \theta^2 \times [\tfrac{2n+2}{(n+2)(n+1)^2}]
\end{aligned}
$$

$$
\begin{aligned}
Var(\hat{\theta}_{MOM}) &= \theta^2/3n \\
Bias(\hat{\theta}_{MOM}) &= E[\hat{\theta}_{MOM}] - \theta = 0 \\
MSE(\hat{\theta}_{MOM}) &= Var(\hat{\theta}_{MOM}) + [Bias(\hat{\theta}_{MOM})]^2 \\
&= \theta^2/3n
\end{aligned}
$$

Note that as n grows large, the MLE has variance and MSE which decline to order $O(n^{-2})$ while the MOM estimate declines slower, to order $O(n^{-1})$.

(d). Find a modification of the mle that renders it unbiased.

To adjust $\hat{\theta}_{MLE}$ to make it unbiased, simply multiply it by the factor $(\frac{n+1}{n})$

$$\hat{\theta}^*_{MLE} = (\frac{n+1}{n})\hat{\theta}_{MLE}$$

4. Problem 8.10.57

Solution:

(a). $E[s^2] = E[\frac{\sigma^2}{n-1}\chi^2_{n-1}] = \frac{\sigma^2}{n-1} \times E[\chi^2_{n-1}] = \frac{\sigma^2}{n-1} \times (n-1) = \sigma^2$ and $E[\hat{\sigma}^2] = E[\frac{n-1}{n}s^2] = \frac{n-1}{n}E[S^2] = \frac{n-1}{n}\sigma^2$.

So, s^2 is unbiased.

(b). The MSE of an estimate is the sum of its variance and its squared bias.

First, compute the variances of each estimate:

$$\begin{aligned} Var[s^2] &= Var[\frac{\sigma^2}{n-1}\chi^2_{n-1}] = (\frac{\sigma^2}{n-1})^2 \times Var[\chi^2_{n-1}] \\ &= (\frac{\sigma^2}{n-1})^2 \times 2(n-1) = 2\sigma^4/(n-1) \\ Var[\hat{\sigma}^2] &= Var[(\frac{n-1}{n})s^2] = (\frac{n-1}{n})^2 Var[s^2] \\ &= (\frac{n-1}{n})^2 \times 2\sigma^4/(n-1) \\ &= 2(\frac{n-1}{n^2}) \times \sigma^4 \end{aligned}$$

Then, compute the MSEs of each estimate:

$$\begin{aligned} MSE(s^2) &= Var[s^2] + [Bias(s^2)]^2 = Var[s^2] = 2\sigma^4/(n-1) \\ MSE(\hat{\sigma}^2) &= Var[\hat{\sigma}^2] + [Bias(\hat{\sigma}^2)]^2 \\ &= 2\sigma^4(\frac{n-1}{n^2}) + (-\frac{1}{n})^2\sigma^4 \\ &= 2\sigma^4(\frac{n-1/2}{n^2}) \end{aligned}$$

Simple algegra proves that $MSE(\hat{\sigma}^2) < MSE(s^2)$.

(c). For what values of ρ does $W = \rho\sum_{i=1}^n(x_i - \overline{x})^2$ have minimal MSE?

$$\begin{aligned} MSE(W) &= Var[W] + [Bias(W)]^2 \\ &= \rho^2 Var[\sum_1^n(x_i - \overline{x})^2] + [E(W) - \sigma^2]^2 \\ &= \rho^2[2(n-1)\sigma^4] + [\rho(n-1)\sigma^2 - \sigma^2]^2 \\ &= \sigma^4[\rho^2(2(n-1) + (n-1)^2) - 2\rho(n-1) + 1] \\ &= \sigma^4[\rho^2(n-1)(n+1) - 2\rho(n-1) + 1] \end{aligned}$$

Minimizing with respect to ρ we solve:

$$\begin{aligned} \frac{d}{d\rho} MSE(W) &= 0 \\ \Longrightarrow \rho &= \frac{1}{n+1} \end{aligned}$$

So $\rho = 1/(n+1)$ is the value that minimizes the MSE.