STA 250/MTH 342 Intro to Mathematical Statistics

Assignment 4, Model Solutions

In addition to the final answer and some common steps, the bolded or boxed parts indicate that they are important but are easy to be ignored.

Solution 1: Exercise 8.1.8 (page 468)

For a random sample that is to be taken from the Bernoulli distribution with unknown parameter p,

$$Var(X) = p(1-p), \text{ (Definition 5.2.1, page 276)}$$

$$\begin{aligned}
E_p(|\bar{X}_n - p|^2) &= Var(\bar{X}_n), \text{ definition of the variance} \\
&= Var(\frac{\sum_{i=1}^{n} X_i}{n}) \\
&= \frac{\sum_{i=1}^{n} Var(X_i)}{n^2} \\
&= \frac{nVar(X_i)}{n^2} \\
&= \frac{Var(X_i)}{n} \\
&= \boxed{\frac{p(1-p)}{n}}
\end{aligned} \tag{1}$$

This variance will be a maximum when p=1/2, at which points is 1/(4n). Therefore, this variance will be not greater than 0.01 for all values of $p(0 \leq p \leq 1)$ if and only if $1/(4n) \leq 0.01$ or, equivalently, if and only if $n \geq 25$.

Solution 2: Exercise 8.7.4 (page 512)

If X has the geometric distribution with parameter p,

$$\boxed{E(X) = \frac{1-p}{p} = \frac{1}{p} - 1}, \text{ (Equation 5.5.7, page 299)}$$

To find a statistic $\delta(X)$ that is a an unbiased estimator of $\frac{1}{p}$, that is,

$$E(\delta(X)) = \frac{1}{p}$$

Therefore, $\delta(X) = X + 1$ is an unbiased estimator of $\frac{1}{p}$, that is,

$$\boxed{E(\delta(X)) = E(X+1) = E(X) + 1 = \frac{1}{p} - 1 + 1 = \frac{1}{p}}$$

Solution 3: Exercise 8.7.8 (page 513)

To find the unbiased estimator of p,$\delta(X)$, that is, we need look for a statistic $\delta(X)$ that satisfies:

$$\boxed{E(\delta(X)) = p}$$

Then,

$$\boxed{E[\delta(X)] = \sum_{x=o}^{\infty} \delta(X) p(1-p)^x}$$

Therefore, Since this relationship must be satisfied for all values of $p(1-p)$, it follows that the coefficient $\delta(0)$ in the power series must be equal to 1, and the coefficient $\delta(x)$ of $(1-p)^x$ must be equal to 0 for $x = 1, 2, \cdots$, that is,

$$E[\delta(X)] = 1 \times p(1-p)^0 + 0 \times p(1-p)^1 + 0 \times p(1-p)^2 + \cdots = p$$

Therefore,

$$\delta(X) = \begin{cases} 1 & X = 0 \\ 0 & X = 1, 2, \cdots \end{cases}$$

Solution 4: Exercise 8.7.10 (page 513)

Let X denote the number of failures that are obtained before k sucesses have been obtained. Then X has the negative binomial distribution with parameters k and p and $N = X + k$. **Your homework has to clarify the notation of x,k, or other possible notations in this negative binomial distribution you use.**

$$\boxed{\Pr(X = x) = \binom{k+x-1}{x} p^k (1-p)^x, \quad x = 0, 1, \cdots.}$$

Therefore

$$E(\frac{k-1}{N-1}) = E(\frac{k-1}{X+k-1}) = \boxed{\sum_{x=0}^{\infty} \frac{k-1}{x+k-1} \binom{k+x-1}{x} p^k (1-p)^x}$$

$$= \sum_{x=0}^{\infty} \frac{k-1}{x+k-1} \times \frac{(k+x-1)!}{x!(k-1)!} p^k (1-p)^x = \sum_{x=0}^{\infty} \frac{(k+x-2)!}{x!(k-2)!} p^k (1-p)^x$$

$$= \boxed{p \sum_{x=0}^{\infty} \binom{(k-1)+x-1}{x} p^{k-1} (1-p)^x}.$$

Note that **the last sum is just the sum of the probabilities of a negative binomial distribution with parameters $k-1$ and p, which is hence 1**. This completes the proof.

Remark. For those who might be interested to know why the probabilities of a negative binomial distribution sum to one, here comes a short proof.

$$\sum_{x=0}^{\infty}\binom{k+x-1}{x}p^k(1-p)^x = \sum_{x=0}^{\infty}\frac{p^k(1-p)^x}{\Gamma(x+1)\Gamma(k)}\Gamma(k+x) = \sum_{x=0}^{\infty}\frac{p^k(1-p)^x}{\Gamma(x+1)\Gamma(k)}\int_0^{\infty}\mathrm{e}^{-u}u^{k+x-1}\,\mathrm{d}u$$
$$= \frac{p^k}{\Gamma(k)}\int_0^{\infty}\mathrm{e}^{-u}u^{k-1}\left(\sum_{x=0}^{\infty}\frac{(1-p)^x u^x}{x!}\right)\mathrm{d}u = \frac{p^k}{\Gamma(k)}\int_0^{\infty}\mathrm{e}^{-u}u^{k-1}\mathrm{e}^{u(1-p)}\,\mathrm{d}u$$
$$\overset{y=up}{=\!=} \frac{p^k}{\Gamma(k)}\int_0^{\infty}\mathrm{e}^{-y}\left(\frac{y}{p}\right)^{k-1}\frac{1}{p}\,\mathrm{d}y = \frac{1}{\Gamma(k)}\int_0^{\infty}\mathrm{e}^{-y}y^{k-1}\,\mathrm{d}y = 1.$$

Solution 5: Exercise 8.7.14 (page 513)

For a fixed θ, the cumulative distribution function of Y_n is $F_n(x) = 0$ for $x \in (-\infty, 0)$, $F_n(x) = 1$ for $x \in (\theta, \infty)$ and

$$\boxed{F_n(x) = \Pr(Y_n \le x) = \Pr(X_1, \cdots, X_n \le x) = \prod_{i=1}^{n}\Pr(X_i \le x) = \frac{x^n}{\theta^n}}$$

So the pdf of Y_n is supported on $[0, \theta]$ (i.e. it is zero outside the interval) and

$$f_n(x) = F_n'(x) = n\frac{x^{n-1}}{\theta^n}, \quad \text{for all } x \in [0, \theta].$$

Based on this we have

$$E[Y_n] = \int_0^{\theta} tn\frac{t^{n-1}}{\theta^n}\,\mathrm{d}t = \frac{n}{n+1}\theta,$$
$$E[Y_n^2] = \int_0^{\theta} t^2 n\frac{t^{n-1}}{\theta^n}\,\mathrm{d}t = \frac{n}{n+2}\theta^2.$$

So $E\left[\frac{n+1}{n}Y_n\right] = \frac{n+1}{n}E[Y_n] = \theta$. This completes the proof. We now compute the MSE's. First,

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

So

$$\mathsf{MSE}_{(n+1)Y_n/n}(\theta) = \mathrm{Var}\left(\frac{n+1}{n}Y_n\middle|\,\theta\right) + B_{(n+1)Y_n/n}(\theta)^2 = \left(\frac{n+1}{n}\right)^2\frac{n\theta^2}{(n+2)(n+1)^2} + 0 = \frac{\theta^2}{n(n+2)},$$
$$\mathsf{MSE}_{Y_n}(\theta) = \mathrm{Var}(Y_n|\theta) + B_{Y_n}(\theta)^2 = \frac{n\theta^2}{(n+2)(n+1)^2} + \left(\frac{n\theta}{n+1} - \theta\right)^2 = \frac{2\theta^2}{(n+1)(n+2)}.$$

Remark. For those who are interested to know more about the order statistic, here comes a general theorem.

Theorem Suppose that $Y_1, Y_2, \cdots, Y_n$ is a random sample of continuous random variables, each having pdf $f_Y(y)$ and cdf $F_Y(y)$. Then

a The pdf of the largest order statistic is

$$f_{Y_{max}}(y) = f_{Y'_n}(y) = n[F_Y(y)]^{n-1} f_Y(y)$$

b The pdf of the smallest order statistic is

$$f_{Y_{min}}(y) = f_{Y'_1}(y) = n[1 - F_Y(y)]^{n-1} f_Y(y)$$

Solution 6: Exercise 8.7.15 (page 513) (a). We have

$$\begin{aligned}
&\theta^3 + \theta^2(1-\theta) + 2\theta(1-\theta) + \theta(1-\theta)^2 + (1-\theta)^3 \\
=&\theta^3 + (1-\theta)[\theta^2 + 2\theta + \theta - \theta^2] + (1-\theta)^3 \\
=&\theta^3 + 3\theta(1-\theta) + (1-\theta)^3 \\
=&\theta^3 + 3\theta - 3\theta^2 + 1 - 3\theta + 3\theta^2 - \theta^3 = 1.
\end{aligned}$$

(b). We have

$$\begin{aligned}
E_\theta \delta_c(X) =&\theta^3 + (2-2c)\theta^2(1-\theta) + 2c\theta(1-\theta) + (1-2c)\theta(1-\theta)^2 \\
=&\theta^3 + (2\theta^2 - 2c\theta^2 + 2c\theta + \theta - \theta^2 - 2c\theta + 2c\theta^2)(1-\theta) \\
=&\theta^3 + (\theta^2 + \theta)(1-\theta) \\
=&\theta^2 + \theta(1-\theta^2) = \theta^3 + \theta - \theta^3 = \theta,
\end{aligned}$$

which is an unbiased estimator of θ.

(c). We have

$$\begin{aligned}
\text{Var}_\theta \delta_c(X) =&E_\theta \left[(\delta_c(X) - E_\theta \delta_c(X))^2\right] \\
=&E_\theta[(\delta_c(X) - \theta)^2] \\
=&(1-\theta)^2\theta^3 + (2-2c-\theta)^2\theta^2(1-\theta) + 2(c-\theta)^2\theta(1-\theta) + \\
&(1-2c-\theta)^2\theta(1-\theta)^2 + \theta^2(1-\theta)^3 \\
=&\theta(1-\theta)[Ac^2 + Bc + C],
\end{aligned}$$

where

$$\begin{aligned}
A &= 6, \\
B &= -4(\theta+1),
\end{aligned}$$

and C is free of c. Therefore

$$\text{Var}_{\theta_0} \delta_c(X) = 6\theta_0(1-\theta_0)[(c - \frac{1+\theta_0}{3})^2 + C'],$$

where again C' is free of c. We see that $\text{Var}_\theta \delta_c(X)$ achieves its minimum if and only if $c = c_0 := \frac{\theta+1}{3}$.

~~END~~