Solutions Assignment 2

11

3+8

7.9.2:
Instead of the last sentence ("Explain why …"), do the following:

(a) Give a one-sentence argument that shows that there is another estimator that makes delta_1 (as given in problem) inadmissible.

Because $T(X_1, \dots, X_n) = X_{(n)}$ is a sufficient statistic for θ (see lecture notes or Ex. 7.7.5 in textbook) and δ_1 is not a function of T, there is another estimator by the Rao-Blackwell theorem that makes δ_1 inadmissible.

(b) Find such an estimator and compare its MSE to the MSE of delta_1 through an explicit calculation.

Because δ_1 is unbiased,

$$MSE(\delta_1;\theta) = Var(\delta_1(X_1,\dots,X_n)) = Var(2\overline{X}_n) = 4Var(\overline{X}_n) = \frac{4}{n}Var(X_1)$$
$$= \frac{4}{n} \times \frac{\theta^2}{12} = \frac{\theta^2}{3n}$$

Note: Some students may propose $T'(X_n) := (1+\frac{1}{n})X_{(n)}$, which has an even better MSE: $MSE(T',\theta) = \frac{1}{n(n+2)}$

Consider $T(X_1,\dots,X_n) := X_{(n)}$. The cdf of $X_{(n)}$ has been obtained in Ex 6.3.15, from which we get the pdf of $X_{(n)}$ as

$$f_n(x) = F_n'(x) = \frac{n}{\theta^n} x^{n-1} \mathbb{1}_{[0,\theta]}(x).$$ Therefore,

$$E[X_{(n)}] = \int_0^\theta x f_n(x)\,dx = \int_0^\theta \frac{n}{\theta^n} x^n\,dx = \frac{n}{(n+1)\theta^n} x^{n+1}\Big|_0^\theta = \frac{n}{n+1}\theta,$$

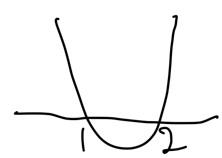
$$E[(X_{(n)})^2] = \int_0^\theta x^2 f_n(x)\,dx = \int_0^\theta \frac{n}{\theta^n} x^{n+1}\,dx = \frac{n}{(n+2)\theta^n}\theta^{n+2} = \frac{n}{n+2}\theta^2$$

$$\Rightarrow MSE(T;\theta) = Var(T) + (Bias(T;\theta))^2 = E[X_{(n)}^2] - E[X_{(n)}]^2 + (E[X_{(n)}]-\theta)^2$$
$$= \left(\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right)\theta^2 + \frac{1}{(n+1)^2}\theta^2 = \frac{2}{(n+2)(n+1)}\theta^2$$

NOT GRADED

We can now verify that $MSE(T;\theta) \le MSE(\delta_1;\theta)$:

$$\frac{2}{(n+2)(n+1)} \le \frac{1}{3n} \iff 6n \le n^2+3n+2 \iff n^2-3n+2 \ge 0$$



The function n^2-3n+2 has zeros $n=1$ and $n=2$ and is strictly positive outside $[1,2]$. Therefore $MSE(T;\theta) \le MSE(\delta_1;\theta)$ for all n, with strict inequality for $n \ge 3$.

NOT GRADED

7.9.6 (see also lecture notes)

It was shown in Exercise 6 of Sec. 7.7 that $\prod_{i=1}^{n} X_i$ is a sufficient statistic in this problem. Since the value of $\overline{X}_n$ cannot be determined from the value of the sufficient statistic alone, $\overline{X}_n$ is inadmissible. by the Rao-Blackwell theorem.

4

7.7.4

In Exercises 1–11, let t denote the value of the statistic T when the observed values of $X_1, \ldots, X_n$ are $x_1, \ldots, x_n$. In each exercise, we shall show that T is a sufficient statistic by showing that the joint p.f. or joint p.d.f. can be factored as in Eq. (7.7.1).

The joint p.d.f. is

$$f_n(\boldsymbol{x} \mid \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{t}{2\sigma^2} \right\}.$$

More details

$$f_n(x;\sigma^2) = \underbrace{1}_{u(x)} \times \underbrace{\frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{2\sigma^2}}}_{V(r(x),\,\sigma^2)} \quad = r(x)$$

By the factorization theorem, $r(X_n) = \sum_{i=1}^{n}(X_i-\mu)^2$ is a sufficient statistic.

Students must (somehow) indicate u, v, and r to get the points.

Both blue and green versions are acceptable.
But: For the green version, if the underlined parts are missing/wrong: no points!

5

7.7.12

The likelihood function is

$$\frac{\alpha^n x_0^{\alpha n}}{\prod_{i=1}^n x_i]^{\alpha+1}}, \tag{S.7.9}$$

for all $x_i \geq x_0$.

More details:

$$f_n(\underline{x}; \alpha, x_0) = \prod_{i=1}^n \frac{\alpha x_0^\alpha}{x_i^{\alpha+1}} \mathbb{1}_{[x_0,\infty)}(x_i) = \frac{\alpha^n x_0^{\alpha n}}{\prod_{i=1}^n x_i^{\alpha+1}} \prod_{i=1}^n \mathbb{1}_{(0,x_i]}(x_0) = \frac{\alpha^n x_0^{\alpha n}}{\prod_{i=1}^n x_i^{\alpha+1}} \mathbb{1}_{(0,x_{(1)}]}(x_0)$$

(a) If x_0 is known, α is the parameter, and (S.7.9) has the form $u(\boldsymbol{x})v[r(\boldsymbol{x}),\alpha]$, with $u(\boldsymbol{x}) = 1$ if all $x_i \geq x_0$ and 0 if not, $r(\boldsymbol{x}) = \prod_{i=1}^n x_i$, and $v[t,\alpha] = \alpha^n x_0^{\alpha n}/t^{\alpha+1}$. So $\prod_{i=1}^n X_i$ is a sufficient statistic.

(b) If α is known, x_0 is the parameter, and (S.7.9) has the form $u(\boldsymbol{x})v[r(\boldsymbol{x}),x_0]$, with $u(\boldsymbol{x}) = \alpha^n/[\prod_{i=1}^n x_i]^{\alpha+1}$, $r(\boldsymbol{x}) = \min\{x_1,\dots,x_n\}$, and $v[t,x_0] = 1$ if $t \geq x_0$ and 0 if not. Hence $\min\{X_1,\dots,X_n\}$ is a sufficient statistic.

More details on b:

$$f_n(\underline{x}; x_0) = \underbrace{\frac{\alpha}{\prod_{i=1}^n x_i^{\alpha+1}}}_{=u(x)} \times \underbrace{x_0^{\alpha n} \mathbb{1}_{(0, \overset{r(x)}{x_{(1)}}]}(x_0)}_{=V(r(x),\, x_0)}$$

$$\Rightarrow r(\underline{X}) = X_{(1)} \text{ is a sufficient statistics.}$$

NOT GRADED

7.8.4

In Exercises 1–4, let t_1 and t_2 denote the values of T_1 and T_2 when the observed values of $X_1,\dots,X_n$ are $x_1,\dots,x_n$. In each exercise, we shall show that T_1 and T_2 are jointly sufficient statistics by showing that the joint p.d.f. of $X_1,\dots,X_n$ can be factored as in Eq. (7.8.1).

Again let the function h be as defined in Example 7.8.4. Then the joint p.d.f. can be written as follows:

$$f_n(\boldsymbol{x} \mid \theta) = \frac{h(\theta,t_1)h(t_2,\theta+3)}{3^n}.$$

$$\theta \leq x_i \leq \theta+3 \Leftrightarrow x_i - 3 \leq \theta \leq x_i$$

$$f_n(\underline{x};\theta) = \prod_{i=1}^n f(x_i;\theta) = \prod_{i=1}^n \frac{1}{3} \mathbb{1}_{[\theta,\theta+3]}(x_i) = 3^{-n} \prod_{i=1}^n \mathbb{1}_{[x_i-3,\, x_i]}(\theta)$$

$\theta \leq x_i$ for all i
$\Leftrightarrow \theta \leq x_{(1)}$
$\theta \geq x_i - 3$ for all i
$\Leftrightarrow \theta \geq x_{(n)} - 3$

$$= \underbrace{3^{-n}}_{u(x)} \times \underbrace{\mathbb{1}_{[x_{(n)}-3;\, x_{(1)}]}(\theta)}_{V((x_{(1)},\, x_{(n)});\,\theta)}$$

$r(x) = (x_{(1)}, x_{(n)})$

$$\Rightarrow r(\underline{X}) := (X_{(1)}, X_{(n)}) \text{ is a sufficient Statistic for } \theta.$$

NOT GRADED

8.1.2

It is known that $\overline{X}_n$ has the normal distribution with mean θ and variance $4/n$. Therefore,

$$E_\theta(|\overline{X}_n - \theta|^2) = \mathrm{Var}_\theta(\overline{X}_n) = 4/n,$$

and $4/n \le 0.1$ if and only if $n \ge 40$.

$$= P\left(\left| \frac{\overline{X}_n - \theta}{\sqrt{4/n}} \right| \le \frac{0.1}{\sqrt{4/n}} \right)$$

$$= \Phi(0.05\sqrt{n}) - \underbrace{\Phi(-0.05\sqrt{n})}_{= 1 - \Phi(0.05\sqrt{n})}$$

NOT GRADED

8.1.4 $\sim N(0,1)$

If Z ~~is defined as in the solution of Exercise 3,~~ then

$$\Pr(|\overline{X}_n - \theta| \le 0.1) = \Pr(|Z| \le 0.05\sqrt{n}) = 2\Phi(0.05\sqrt{n}) - 1.$$

Therefore, this value will be at least 0.95 if and only if $\Phi(0.05\sqrt{n}) \ge 0.975$. It is found from p. 860 ~~a table of values of Φ~~ that we must have $0.05\sqrt{n} \ge 1.96$. Therefore, we must have $n \ge 1536.64$ or, since n must be an integer, $n \ge 1537$.

2

8.2.10 (also find how many degrees of freedom the chi^2 distribution has)

Each of the variables $X_1 + X_2 + X_3$ and $X_4 + X_5 + X_6$ will have the normal distribution with mean 0 and variance 3. Therefore, if each of them is divided by $\sqrt{3}$, each will have a standard normal distribution. Therefore, the square of each will have the χ^2 distribution with one degree of freedom and the sum of these two squares will have the χ^2 distribution with two degrees of freedom. In other words, $Y/3$ will have the χ^2 distribution with two degrees of freedom.

(only grade results for this exercise)

7

8.4.2

Since $\hat{\mu} = \overline{X}_n$ and $\hat{\sigma}^2 = S_n^2/n$, it follows ~~from the definition of U in Eq. (8.4.4)~~ that

(n-1)

$$\Pr(\hat{\mu} > \mu + k\hat{\sigma}) = \Pr\left(\frac{\overline{X}_n - \mu}{\hat{\sigma}} > k\right) = \Pr[U > k(n-1)^{1/2}].$$

where $U = \frac{\overline{X}_n - \mu}{\sqrt{S_n^2/n}}$

Since U has the t distribution with $n-1$ degrees of freedom and $n = 17$, we must choose k such that $\Pr(U > 4k) = 0.95$. It is found from a table of the t distribution with 16 degrees of freedom that $\Pr(U < 1.746) = 0.95$. Hence, by symmetry, $\Pr(U > -1.746) = 0.95$. It now follows that $4k = -1.746$ and $k = -0.4365$.

NOT GRADED

8.5.4

Since $\sqrt{n}(\overline{X}_n - \mu)/\sigma$ has a standard normal distribution, $\Pr\left[-1.96 < \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} < 1.96\right] = 0.95$. This relation can be rewritten in the form

$$\Pr\left(\overline{X}_n - \frac{1.96\sigma}{\sqrt{n}} < \mu < \overline{X}_n + \frac{1.96\sigma}{\sqrt{n}}\right) = 0.95.$$

Therefore, the interval with endpoints $\overline{X}_n - 1.96\sigma/\sqrt{n}$ and $\overline{X}_n + 1.96\sigma/\sqrt{n}$ will be a confidence interval for μ with confidence coefficient 0.95. The length of this interval will be $3.92\sigma/\sqrt{n}$. It now follows that $3.92\sigma/\sqrt{n} < 0.01\sigma$ if and only if $\sqrt{n} > 392$. This means that $n > 153664$ of $n = 153665$ or more.

6

8.5.6

The exponential distribution with mean μ is the same as the gamma distribution with $\alpha = 1$ and $\beta = 1/\mu$. ✓ (the reproductive property) Therefore, by Theorem 5.7.7, $\sum_{i=1}^{n} X_i$ will have the gamma distribution with parameters $\alpha = n$ and $\beta = 1/\mu$. In turn, it follows from Exercise 1 of Sec. 5.7 that (scaling property) $2\sum_{i=1}^{n} X_i/\mu$ has the gamma distribution with parameters $\alpha = n$ and $\beta = \frac{1}{2}$. which is by a theorem in class just ~~It follows from Definition 8.2.1 that $2\sum_{i=1}^{n} X_i/\mu$ has~~ the χ^2 distribution with $2n$ degrees of freedom. [~~Constants c_1 and c_2 which satisfy the relation given in the hint for this exercise will then each be 1/2 times some quantile of the χ^2 distribution with $2n$ degrees of freedom. There are an infinite number of pairs of values of such quantiles, one corresponding to each pair of numbers $q_1 \geq 0$ and $q_2 \geq 0$ such that $q_2 - q_1 = \gamma$. For example, with $q_1 = (1-\gamma)/2$ and $q_2 = (1+\gamma)/2$ we can let c_i be 1/2 times the q_i quantile of the χ^2 distribution with $2n$ degrees of freedom for $i = 1, 2$. It now follows that~~

$$\Pr\left(\frac{1}{c_2}\sum_{i=1}^{n} X_i < \mu < \frac{1}{c_1}\sum_{i=1}^{n} X_i\right) = \gamma.$$

~~Therefore, the interval with endpoints equal to the observed values of $\sum_{i=1}^{n} X_i/c_2$ and $\sum_{i=1}^{n} X_i/c_1$ will be a confidence interval for μ with confidence coefficient γ.~~]

(Note that γ here is $1-\alpha$ in class)

In other words, $\Psi(X_1,\dots,X_n;\mu) = \frac{2}{\mu}\sum_{i=1}^{n} X_i$ is a pivot for μ. ✓

Choose $\beta_1, \beta_2 \in [0, 1-\gamma]$ such that $\beta_1 + \beta_2 = 1-\gamma$. ✓ Because $\Psi \sim \chi^2_{2n}$, ✓ it follows that

(either specific or general β_1, β_2 that add up to $1-\gamma$)

as seen in class

$$\mathbb{P}\left(q(\chi^2_{2n};\beta_1) \leq \Psi(X_1,\dots,X_n;\mu) \leq q(\chi^2_{2n};1-\beta_2)\right) = \gamma$$

$$\Longleftrightarrow \mathbb{P}\left(q(\chi^2_{2n};\beta_1) \leq \frac{2}{\mu}\sum_{i=1}^{n} X_i \leq q(\chi^2_{2n};1-\beta_2)\right) = \gamma \quad ✓$$

$$\Longleftrightarrow \frac{2}{q(\chi^2_{2n};1-\beta_2)}\sum_{i=1}^{n} X_i \leq \mu \leq \frac{2}{q(\chi^2_{2n};\beta_1)}\sum X_i$$

$\Longrightarrow$ A γ-CI for μ is given by

$$\left(\overset{c_1}{\frac{2}{q(\chi^2_{2n};1-\beta_2)}}\sum X_i,\ \overset{c_2}{\frac{2}{q(\chi^2_{2n};\beta_1)}}\sum X_i\right) \quad ✓$$

(or a specific interval like the one below)

where $\beta_1, \beta_2 \in [0, 1-\gamma]$ can be chosen arbitrarily as long as $\beta_1 + \beta_2 = 1-\gamma$.

For example, if $\beta_1 = \beta_2 = \frac{1-\gamma}{2}$, we obtain

$$\left(\underset{c_1}{\frac{2}{q(\chi^2_{2n};\frac{1+\gamma}{2})}}\sum X_i,\ \underset{c_2}{\frac{2}{q(\chi^2_{2n};1-\frac{\gamma}{2})}}\sum X_i\right)$$

NOT GRADED

8.7.4 back of textbook

If X has the geometric distribution with parameter p, then it follows from ~~Eq. (5.5.7)~~ that $E(X) = (1-p)/p = 1/p - 1$. Therefore, $E(X+1) = 1/p$, which implies that $X+1$ is an unbiased estimator of $1/p$.

NOT GRADED

8.9.6 (i.e., $\hat{\theta}_n = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbb{1}_A(X_i) = \overline{Y}_n$ where $Y_i = \mathbb{1}_A(X_i)$)

Let $\hat{\theta}_n$ be the proportion of the n observations that lie in the set A. Since each observation has probability θ of lying in A, the observations can be thought of as forming n Bernoulli trials, each with probability θ of success. Hence, $E(\hat{\theta}_n) = \theta$ and $\mathrm{Var}(\hat{\theta}_n) = \theta(1-\theta)/n$.

NOT GRADED

Ex I: Let X_1, …, X_n be iid Bernoulli(p). Show that there is no unbiased estimator of theta=log(p).

If $T=\varphi(X_1,\dots,X_n)$ were an unbiased estimator, we would have

$$E[T] = \sum_{x_i\in\{0,1\}} \varphi(x_1,\dots,x_n)\,\mathbb{P}(X_1=x_1,\dots,X_n=x_n) = \sum_{x_i\in\{0,1\}} \varphi(x_1,\dots,x_n)\, p^{\sum x_i}(1-p)^{n-\sum x_i}$$

$$\overset{!}{=} \log p$$

However, as $p\to 0$,

$$\left|\sum_{x_i\in\{0,1\}} \varphi(x_1,\dots,x_n)\, p^{\sum x_i}(1-p)^{n-\sum x_i}\right| \le \sum_{x_i\in\{0,1\}} |\varphi(x_1,\dots,x_n)|$$

is bounded, but

$$|\log p| \to \infty,$$

a contradiction.

6

Ex II: For a sample of 9 iid normal observations with sample standard deviation s_n = 3.2, construct a 99%-confidence interval of the form (0,b) for the variance σ^2.

Because

(pivot!) $$\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2_{n-1},$$

we know that

$$P\left(\frac{(n-1)S_n^2}{\sigma^2} \; > q(\chi^2_{n-1}, 0.01) \; \cancel{< \chi^2_{0.01,n-1}}\right) = 0.99,$$

$q(\chi^2_{n-1}, 0.01) = q(\chi^2_8, 0.01) = 1.647$

where $\cancel{\chi^2_{0.01,n-1} = \chi^2_{0.01,8} = 20.090}$. The inside of the probability statement above is equivalent to

$$\frac{(n-1)S_n^2}{\sigma^2} > 1.647 \; \cancel{< 20.09} \iff \sigma^2 < \frac{(n-1)S_n^2}{1.647\;\cancel{20.09}} = \frac{8\times 3.2^2}{1.647\;\cancel{20.09}} = 49.74\;\cancel{4.08}$$

Thus, an upper ~~a lower~~ 99%-confidence bound for σ^2 is given by

$\cancel{(4.08, \infty)}$.

(0, 49.74)

Sum: 41