STAT 3202: Practice 02

Autumn 2018, OSU

Exercise 1

Let $X_1, X_2, \dots, X_n$ be iid $N(\theta, 1)$ and consider $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Show that $\bar{X}_n$ is a consistent estimator of θ.

Solution:

First, note that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator as we have seen many times before.

$$\mathrm{E}\left[\bar{X}_n\right] = \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} \cdot n\theta = \theta$$

Also, as we have seen before, the variance of $\bar{X}_n$ is given by

$$\mathrm{Var}\left[\bar{X}_n\right] = \frac{\sigma^2}{n} = \frac{1}{n}$$

Since we have an unbiased estimator, we simply need for the variance to vanish as n goes to infinity. We see that

$$\lim_{n \to \infty} \mathrm{Var}\left[\bar{X}_n\right] = \lim_{n \to \infty} \frac{1}{n} = 0$$

Thus $\bar{X}_n$ is a consistent estimator for θ.

Exercise 2

Suppose that $X_1, X_2, \dots, X_n$ are an iid sample from the distribution

$$f(x; \theta) = \frac{1}{2}(1 + \theta x), \quad -1 < x < 1, -1 < \theta < 1.$$

Show that $3\bar{X}_n$ is a consistent estimator of θ.

Solution:

First, recall that we have previously calculated,

$$\mathrm{E}\left[X_i\right] = \frac{\theta}{3}$$

and

$$\mathrm{Var}\left[X_i\right] = \frac{3-\theta^2}{9}$$

Thus, $3\bar{X}_n$ is an unbiased estimator of θ since

$$\mathrm{E}\left[3\bar{X}_n\right] = 3 \cdot \mathrm{E}\left[X_i\right] = 3 \cdot \frac{\theta}{3} = \theta$$

Then, we calculate the variance of the proposed estimator.

$$\mathrm{Var}\left[3\bar{X}_n\right] = 9 \cdot \frac{\mathrm{Var}[X_i]}{n} = 9 \cdot \frac{3-\theta^2}{9n} = \frac{3-\theta^2}{n}$$

Then, since $3\bar{X}_n$ is an unbiased estimator of θ and

$$\lim_{n\to\infty} \mathrm{Var}\left[3\bar{X}_n\right] = \lim_{n\to\infty} \frac{3-\theta^2}{n} = 0$$

we conclude that $3\bar{X}_n$ is a consistent estimator of θ.

Exercise 3

Let $Y_1, Y_2, \ldots, Y_n$ be a random sample such that

- $\mathrm{E}[Y_i] = \mu$
- $\mathrm{Var}[Y_i] = \sigma^2$.

Suggest a consistent estimator for μ^2.

Solution:

First note that,

$$\mathrm{E}\left[\bar{Y}_n\right] = \mu$$

and

$$\lim_{n\to\infty} \mathrm{Var}\left[\bar{Y}_n\right] = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$$

Thus, $\bar{Y}_n$ is a consistent estimator of μ.

Therefore, $\bar{Y}_n^2$ is a consistent estimator of μ^2.

Exercise 4

Let $X_1, X_2, \ldots, X_n$ be iid $N(\mu_X, \sigma_X^2)$. Also, let $Y_1, Y_2, \ldots, Y_n$ be iid $N(\mu_Y, \sigma_Y^2)$.

Suggest a consistent estimator for $\mu_X - \mu_Y$.

Solution:

First note that $\bar{X}$ is a consistent estimator for μ_X and $\bar{Y}$ is a consistent estimator for μ_Y.

Because of this, we can say that $-\bar{Y}$ is a consistent estimator for $-\mu_Y$. Then finally we can conclude that $\bar{X} - \bar{Y}$ is a consistent estimator for $\mu_X - \mu_Y$.

Or, more directly note that

$$\mathrm{E}\left[\bar{X} - \bar{Y}\right] = \mathrm{E}\left[\bar{X}\right] - \mathrm{E}\left[\bar{Y}\right] = \mu_X - \mu_Y$$

thus $\bar{X} - \bar{Y}$ is an unbiased estimator for $\mu_X - \mu_Y$.

Also note that,

$$\mathrm{Var}\left[\bar{X} - \bar{Y}\right] = \frac{\mathrm{Var}\left[\bar{X}\right]}{n} + \frac{\mathrm{Var}\left[\bar{Y}\right]}{n} = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{n}$$

Then, since

$$\lim_{n \to \infty} \mathrm{Var}\left[\bar{X} - \bar{Y}\right] = \lim_{n \to \infty} \left(\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{n}\right) = 0$$

we can conclude that $\bar{X} - \bar{Y}$ is a consistent estimator for $\mu_X - \mu_Y$.

Exercise 5

Let $X_1, X_2, \ldots, X_n$ be iid $N(\mu_X, \sigma^2)$. Also, let $Y_1, Y_2, \ldots, Y_n$ be iid $N(\mu_Y, \sigma^2)$. Note that both distributions have the same variance.

Show that

$$\frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2 + \sum_{i=1}^n \left(Y_i - \bar{Y}\right)^2}{2n - 2}$$

is a consistent estimator for σ^2.

Hint: Note that

$$\frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$\frac{\sum_{i=1}^n \left(Y_i - \bar{Y}\right)^2}{\sigma^2} \sim \chi^2_{n-1}$$

Also, recall that, if $W \sim \chi^2_k$, then $\mathrm{E}[W] = k$ and $\mathrm{Var}[W] = 2k$.

Solution:

First, using the hint, we can show that the estimator is unbiased.

$$\begin{aligned}
\mathrm{E}\left[\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2+\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{2n-2}\right] &= \mathrm{E}\left[\frac{\sigma^2\cdot\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{\sigma^2}+\sigma^2\cdot\frac{\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{\sigma^2}}{2n-2}\right] \\
&= \frac{\sigma^2}{2n-2}\cdot\mathrm{E}\left[\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{\sigma^2}\right]+\frac{\sigma^2}{2n-2}\cdot\mathrm{E}\left[\frac{\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{\sigma^2}\right] \\
&= \frac{\sigma^2}{2n-2}\cdot(n-1)+\frac{\sigma^2}{2n-2}\cdot(n-1) \\
&= \frac{n\sigma^2-\sigma^2+n\sigma^2-\sigma^2}{2n-2}=\frac{2n-2}{2n-2}\cdot\sigma^2=\sigma^2
\end{aligned}$$

Next, we calculate the variance.

$$\begin{aligned}
\mathrm{Var}\left[\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2+\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{2n-2}\right] &= \mathrm{Var}\left[\frac{\sigma^2\cdot\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{\sigma^2}+\sigma^2\cdot\frac{\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{\sigma^2}}{2n-2}\right] \\
&= \left(\frac{\sigma^2}{2n-2}\right)^2\cdot\mathrm{Var}\left[\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{\sigma^2}\right]+\left(\frac{\sigma^2}{2n-2}\right)^2\cdot\mathrm{Var}\left[\frac{\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{\sigma^2}\right] \\
&= \left(\frac{\sigma^2}{2n-2}\right)^2\cdot 2(n-1)+\left(\frac{\sigma^2}{2n-2}\right)^2\cdot 2(n-1) \\
&= \frac{2\left(\sigma^2\right)^2}{2n-2}
\end{aligned}$$

Then finally we see that

$$\lim_{n\to\infty}\mathrm{Var}\left[\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2+\sum_{i=1}^n\left(Y_i-\bar{Y}\right)^2}{2n-2}\right]=\lim_{n\to\infty}\frac{2\left(\sigma^2\right)^2}{2n-2}=0$$

Thus the proposed estimator is consistent.

Exercise 6

Let $Y_1, Y_2, \ldots, Y_n$ be iid observations from a Poisson distribution with parameter λ. Show that $U=\sum_{i=1}^n Y_i$ is sufficient for λ.

Solution:

First, recall that the pmf for a Poisson random variable is given by

$$f(y \mid \lambda) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \ldots, \lambda > 0$$

Then we have

$$\begin{aligned} f(y_1, y_2, \ldots, y_n \mid \lambda) &= \prod_{i=1}^{n} \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \\ &= \frac{\lambda^{\sum_{i=1}^{n} y_i} e^{-n\lambda}}{\prod_{i=1}^{n} y_i!} \\ &= \lambda^u e^{-n\lambda} \frac{1}{\prod_{i=1}^{n} y_i!} \end{aligned}$$

Here we have

$$U = \sum_{i=1}^{n} Y_i$$

$$g(u, \lambda) = \lambda^u e^{-n\lambda}$$

$$h(y_1, y_2, \ldots, y_n) = \frac{1}{\prod_{i=1}^{n} y_i!}$$

Thus, by the Factorization Theorem,

$$U = \sum_{i=1}^{n} Y_i$$

is a sufficient statistic for λ.

Exercise 7

Let $X_1, X_2, \ldots, X_n$ be iid observations from a distribution with density

$$f(x \mid \theta) = \frac{\theta}{(1+x)^{\theta+1}}, \quad 0 < \theta < \infty, 0 < x < \infty$$

Find a sufficient statistic for θ.

Solution:

First note that

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n \mid \theta) &= \prod_{i=1}^{n} \frac{\theta}{(1+x_i)^{\theta+1}} \\
&= \frac{\theta^n}{\left[\prod_{i=1}^{n}(1+x_i)\right]^{\theta+1}} \\
&= \frac{\theta^n}{u^{\theta+1}}
\end{aligned}
$$

Here we have

$$
\begin{aligned}
U &= \prod_{i=1}^{n}(1+X_i) \\
g(u,\theta) &= \frac{\theta^n}{u^{\theta+1}} \\
h(x_1, x_2, \ldots, x_n) &= 1
\end{aligned}
$$

Thus, by the Factorization Theorem,

$$
U = \prod_{i=1}^{n}(1+X_i)
$$

is a sufficient statistic for θ.

Exercise 8

Let $X_1, X_2, \ldots, X_n$ be iid observations from a normal distribution with a unknown mean, μ, and known variance $\sigma^2 = 9$.

Show that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for μ, then use this statistic to create an estimator that is both unbiased and sufficient for estimating μ.

Solution:

First recall that the pdf for a normal distribution is given by

$$
f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma^2 > 0
$$

In this case, with $\sigma^2 = 9$, we have

$$
f(x \mid \mu) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{18}(x-\mu)^2\right]
$$

Now note that

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n \mid \theta) &= \prod_{i=1}^{n} \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{18}\left(x_i - \mu\right)^2\right] \\
&= \left(\frac{1}{3\sqrt{2\pi}}\right)^n \prod_{i=1}^{n} \exp\left[-\frac{1}{18}\left(x_i^2 - 2\mu x_i + \mu^2\right)\right] \\
&= \left(\frac{1}{3\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{18}\sum_{i=1}^{n} x_i^2\right] \exp\left[\frac{2\mu}{18}\sum_{i=1}^{n} x_i\right] \exp\left[-\frac{1}{18}n\mu^2\right] \\
&= \left(\frac{1}{3\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{18}\sum_{i=1}^{n} x_i^2\right] \exp\left[\frac{2\mu}{18}u\right] \exp\left[-\frac{1}{18}n\mu^2\right]
\end{aligned}
$$

Here we have

$$U = \sum_{i=1}^{n} X_i$$

$$g(u, \mu) = \exp\left[\frac{2\mu}{18}u\right] \exp\left[-\frac{1}{18}n\mu^2\right]$$

$$h(x_1, x_2, \ldots, x_n) = \left(\frac{1}{3\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{18}\sum_{i=1}^{n} x_i^2\right]$$

Thus, by the Factorization Theorem,

$$U = \sum_{i=1}^{n} X_i$$

is a sufficient statistic for μ.

Then, because $f(u) = \frac{1}{n}u$ is a one-to-one function,

$$U^\star = \frac{1}{n}\sum_{i=1}^{n} X_i$$

is also a sufficient statistic for μ, which also happens to be unbiased, which we have seen many times before.

Exercise 9

Let $Y_1, Y_2, \ldots, Y_n$ be iid observations from a distribution with density

$$f(y \mid \beta) = \frac{y}{\beta} \cdot \exp\left(\frac{-y^2}{2\beta}\right), \quad y \geq 0, \beta > 0$$

Find a sufficient statistic for β.

Solution:

First note that

$$\begin{aligned} f(y_1, y_2, \ldots, y_n \mid \theta) &= \prod_{i=1}^{n} \frac{y_i}{\beta} \cdot \exp\left(\frac{-y_i^2}{2\beta}\right) \\ &= \beta^{-n} \left(\prod_{i=1}^{n} y_i\right) \exp\left[-\frac{\sum_{i=1}^{n} y_i^2}{2\beta}\right] \\ &= \beta^{-n} \left(\prod_{i=1}^{n} y_i\right) \exp\left[-\frac{u}{2\beta}\right] \end{aligned}$$

Here we have

$$U = \sum_{i=1}^{n} Y_i^2$$

$$g(u, \beta) = \beta^{-n} \exp\left[-\frac{u}{2\beta}\right]$$

$$h(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} y_i$$

Thus, by the Factorization Theorem,

$$U = \sum_{i=1}^{n} Y_i^2$$

is a sufficient statistic for β.

Exercise 10

Let $X_1, X_2, \ldots, X_n$ be iid observations from a distribution with density

$$f(x \mid \alpha, \beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^{\alpha}}, \quad x \geq 0, \alpha > 0, \beta > 0$$

Let α be a known constant and β be unknown. Find a sufficient statistic for β.

Solution:

First note that

$$
\begin{aligned}
f(x_1, x_2, \ldots, x_n \mid \theta) &= \prod_{i=1}^{n} \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-(x/\beta)^{\alpha}} \\
&= \alpha^n \beta^{-n} \frac{\prod_{i=1}^{n} x_i^{\alpha-1}}{\prod_{i=1}^{n} \beta^{\alpha-1}} \exp\left[-\frac{\sum_{i=1}^{n} x_i^{\alpha}}{\beta^{\alpha}}\right] \\
&= \alpha^n \beta^{-n} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \beta^{n-n\alpha} \exp\left[-\frac{\sum_{i=1}^{n} x_i^{\alpha}}{\beta^{\alpha}}\right] \\
&= \alpha^n \beta^{-n\alpha} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left[-\frac{\sum_{i=1}^{n} x_i^{\alpha}}{\beta^{\alpha}}\right] \\
&= \alpha^n \beta^{-n\alpha} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left[-\frac{u}{\beta^{\alpha}}\right]
\end{aligned}
$$

Here we have

$$
U = \sum_{i=1}^{n} X_i^{\alpha}
$$

$$
g(u, \beta) = \beta^{-n\alpha} \exp\left[-\frac{u}{\beta^{\alpha}}\right]
$$

$$
h(x_1, x_2, \ldots, x_n) = \alpha^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1}
$$

Thus, by the Factorization Theorem,

$$
U = \sum_{i=1}^{n} X_i^{\alpha}
$$

is a sufficient statistic for β.
