Solutions to exercises - Week 42

Complete sufficient statistics and best unbiased estimators:

- Exercises 7.47, 7.52, 7.59 and 7.60
- Additional exercise

Exercise 7.47

A circle has radius r and area $A = \pi r^2$

We make measurements $X_1, \ldots, X_n$ of the radius

We assume that $X_i = r + \varepsilon_i$, where the ε_i's are iid and $n(0,\sigma^2)$-distributed

Thus the X_i's are iid and $n(r,\sigma^2)$-distributed

The $n(r,\sigma^2)$ -distribution may be written as an exponential family:

$$f(x\,|\,r,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{r^2}{2\sigma^2}\right)\exp\left(\frac{r}{\sigma^2}x - \frac{1}{2\sigma^2}x^2\right)$$

Hence

$$\left(\sum_{i=1}^{n} X_i \,,\, \sum_{j=1}^{n} X_i^2\right)$$

is a complete sufficient statistics

We have $\bar{X} \sim n(r,\sigma^2/n)$

Note that

$$A = \pi r^2 = \pi(\mathrm{E}\bar{X})^2 = \pi\left(\mathrm{E}\bar{X}^2 - \mathrm{Var}\bar{X}\right)$$

Thus an unbiased estimator of A is

$$\hat{A} = \pi\left(\bar{X}^2 - S^2/n\right)$$

The estimator is based on a complete sufficient statistic, and hence it is the best unbiased estimator

Exercise 7.52

Let $X_1, X_2, \ldots, X_n$ be iid Poisson(λ)

The Poisson pmf is given by

$$f(x\,|\,\lambda) = \begin{cases} \dfrac{\lambda^x}{x!}e^{-\lambda} & \text{for } x = 0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

This may be written as an exponential family:

$$f(x\,|\,\lambda) = \underbrace{I_{\{0,1,2,\ldots\}}(x)\,\frac{1}{x!}}_{h(x)}\;\underbrace{e^{-\lambda}}_{c(\lambda)}\;\exp\{\underbrace{\log(\lambda)}_{w_1(\lambda)}\;\underbrace{x}_{t_1(x)}\}$$

Hence $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic

a) $\bar{X} = (1/n)\sum_{i=1}^{n} X_i$ is an unbiased estimator for λ that is based on a complete and sufficient statistic

Hence $\bar{X}$ is the best unbiased estimator for λ

b) $S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 / (n-1)$ is an unbiased estimator for the population variance, which for the Poisson distribution equals λ

Now $\bar{X}$ is a complete sufficient statistic (any one-to-one function of a complete sufficient statistic is itself a complete sufficient statistic)

We have that $\mathrm{E}(S^2 \mid \bar{X})$ is an unbiased estimator for λ that is a function of $\bar{X}$

By the uniqueness this gives $\mathrm{E}(S^2 \mid \bar{X}) = \bar{X}$

Now we have

$$\mathrm{Var}S^2 = \mathrm{Var}\left[\mathrm{E}(S^2 \mid \bar{X})\right] + \mathrm{E}\left[\mathrm{Var}(S^2 \mid \bar{X})\right]$$

$$= \mathrm{Var}\bar{X} + \mathrm{E}\left[\mathrm{Var}(S^2 \mid \bar{X})\right] > \mathrm{Var}\bar{X}$$

c) A general theorem is as follows:

Let $T = T(\mathbf{X})$ be a complete sufficient statistic, and let $T' = T'(\mathbf{X})$ be another statistic such that $\mathrm{E}T' = \mathrm{E}T$. Then $\mathrm{E}(T' \mid T) = T$ and $\mathrm{Var}T' > \mathrm{Var}T$

Exercise 7.59

Let $X_1, X_2, \ldots, X_n$ be iid $n(\mu, \sigma^2)$ random variables, where both parameters are unknown

$\left(\sum_{i=1}^{n} X_i \,,\, \sum_{j=1}^{n} X_i^2\right)$ is a complete sufficient statistic

Then $(\bar{X}, S^2)$ is also a complete sufficient statistic

Now $T = (n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$

From exercise 3.17 (with $\alpha = (n-1)/2$, $\beta = 2$) we have that

$$\mathrm{E}T^{p/2} = 2^{p/2} \frac{\Gamma((n-1)/2 + p/2)}{\Gamma((n-1)/2)} = 2^{p/2} \frac{\Gamma((n+p-1)/2)}{\Gamma((n-1)/2)}$$

Thus we have

$$\mathrm{E}\left(\frac{(n-1)S^2}{\sigma^2}\right)^{p/2} = 2^{p/2} \frac{\Gamma((n+p-1)/2)}{\Gamma((n-1)/2)}$$

It follows that

$$\mathrm{E}S^p = \mathrm{E}\left(\frac{\sigma^2}{n-1} \frac{(n-1)S^2}{\sigma^2}\right)^{p/2} = \frac{\sigma^p}{(n-1)^{p/2}} 2^{p/2} \frac{\Gamma((n+p-1)/2)}{\Gamma((n-1)/2)}$$

Thus

$$\frac{(n-1)^{p/2}}{2^{p/2}} \frac{\Gamma((n-1)/2)}{\Gamma((n+p-1)/2)} S^p$$

is an unbiased estimator of σ^p

The estimator is based on a complete sufficient statistic, and hence it is UMVUE

Exercise 7.60

$X_1, X_2, \ldots, X_n$ iid $\text{gamma}(\alpha, \beta)$ distributed, α known

The gamma pdf is given by

$$f(x \mid \beta) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

This may be written as an exponential family:

$$f(x \mid \beta) = \underbrace{I_{\{x>0\}}(x) \frac{1}{\Gamma(\alpha)} x^{\alpha-1}}_{h(x)} \underbrace{\frac{1}{\beta^\alpha}}_{c(\beta)} \exp\left\{ \underbrace{-\frac{1}{\beta}}_{w_1(\beta)} \underbrace{x}_{t_1(x)} \right\}$$

Thus $\sum_{j=1}^{n} X_i$ is a complete sufficient statistics

Now $\sum_{j=1}^{n} X_i \sim \text{gamma}(n\alpha, \beta)$

By the result in exercise 3.17 we obtain

$$\mathrm{E}\left[\left(\sum_{j=1}^{n} X_i \right)^{-1} \right] = \beta^{-1} \frac{\Gamma(n\alpha - 1)}{\Gamma(n\alpha)} = \beta^{-1} \frac{\Gamma(n\alpha - 1)}{(n\alpha - 1)\Gamma(n\alpha - 1)} = \frac{1}{\beta(n\alpha - 1)}$$

Thus $(n\alpha - 1) / \sum_{i=1}^{n} X_i$ is an unbiased estimator for $1/\beta$

The estimator is based on a complete sufficient statistic, and hence it is UMVUE

Additional exercise

Assume that $X_1, X_2, \ldots, X_n$ are iid and Poisson distributed with mean λ

a) Find a sufficient and complete statistic for λ

Solution:

By exercise 7.52 we have that $T = \sum_{j=1}^{n} X_i$ is a complete sufficient statistics

b) Find an unbiased estimator for $\tau(\lambda) = e^{-\lambda}$ based on X_1

Solution:

Let $W = I\{X_1 = 0\}$

Then $\mathrm{E}W = P(X_1 = 0) = e^{-\lambda} = \tau(\lambda)$

c) Find the best unbiased estimator for $\tau(\lambda) = e^{-\lambda}$

Solution:

The best unbiased estimator is given as

$$\phi(T) = \mathrm{E}(W \mid T) = P(X_1 = 0 \mid \sum_{i=1}^{n} X_i)$$

Now we have that

$$P\left(X_1 = 0 \mid \sum_{i=1}^{n} X_i = t \right) = \frac{P\left(X_1 = 0 \ , \ \sum_{i=1}^{n} X_i = t \right)}{P\left(\sum_{i=1}^{n} X_i = t \right)}$$

$$= \frac{P\left(X_1 = 0 \ , \ \sum_{i=2}^{n} X_i = t\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)} = \frac{P(X_1 = 0) \cdot P\left(\sum_{i=2}^{n} X_i = t\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)}$$

$$= \frac{e^{-\lambda} \cdot \frac{[(n-1)\lambda]^t}{t!} e^{-(n-1)\lambda}}{\frac{(n\lambda)^t}{t!} e^{-n\lambda}} = \left(1 - \frac{1}{n}\right)^t$$

Thus the best unbiased estimator is given as

$$\hat{\tau} = \left(1 - \frac{1}{n}\right)^{\sum_{i=1}^{n} X_i} = \left[\left(1 - \frac{1}{n}\right)^n\right]^{\bar{X}}$$