131B HW#6 solution

8.7 Unbiased Estimators

12. (a) Let X denote the value of the characteristic for a person chosen at random from the total population, and let A_i denote the event that the person belongs to stratum $i (i = 1, \ldots, k)$. Then

$$\mu = E(X) = \sum_{i=1}^{k} E(X|A_i)P(A_i) = \sum_{i=1}^{k} \mu_i p_i.$$

Also,

$$E(\mu) = \sum_{i=1}^{k} p_i E(\bar{X}_i) = \sum_{i=1}^{k} p_i \mu_i = \mu.$$

(b) Since the samples are taken independently of each other, the variables $\bar{X}_1, \ldots, \bar{X}_k$ are independent. Therefore,

$$\text{var}(\hat{\mu}) = \sum_{i=1}^{k} p_i^2 \text{var}(\bar{X}_i) = \sum_{i=1}^{k} \frac{p_i^2 \sigma_i^2}{n_i}.$$

Hence, the values of $n_1, \ldots, n_k$ must be chosen to minimize $v = \sum_{i=1}^{k} \frac{(p_i \sigma_i)^2}{n_i}$, subject to the constraint that $\sum_{i=1}^{k} n_i = n$. If we let $n_k = n - \sum_{i=1}^{k-1} n_i$, then

$$\frac{\partial v}{\partial n_i} = \frac{-(p_i \sigma_i)^2}{n_i^2} + \frac{(p_k \sigma_k)^2}{n_k^2} \quad \text{for } i = 1, \ldots, k-1.$$

When each of these partial derivatives is set equal to 0, it is found that $n_i/(p_i \sigma_i)$ has the same value for $i = 1, \ldots, k$. Therefore, $n_i = c p_i \sigma_i$ for some constant c. It follows that $n = \sum_{j=1}^{k} h_j = c \sum_{j=1}^{k} p_j \sigma_j$. Hence, $c = n / \sum_{j=1}^{k} p_j \sigma_j$ and, in turn,

$$n_i = \frac{n p_i \sigma_i}{\sum_{j=1}^{k} p_j \sigma_j}.$$

This analysis ignores the fact that the values of $n_1, \ldots, n_k$ must be integers. The integers $n_1, \ldots, n_k$ for which v is a minimum would presumably be near the minimizing values of $n_1, \ldots, n_k$ which have just been found.

14. For $0 < y < \theta$, the c.d.f. of Y_n is

$$F(y|\theta) = P(Y \leq y|\theta) = P(X_1 \leq y, \ldots, X_n \leq y|\theta) = (\frac{y}{\theta})^n.$$

Therefore, for $0 < y < \theta$, the p.d.f. of Y_n is

$$f(y|\theta) = \frac{d}{dy}F(y|\theta) = \frac{ny^{n-1}}{\theta^n}.$$

It now follows that

$$E_\theta(Y_n) = \int_0^\theta y\frac{ny^{n-1}}{\theta^n}dy = \frac{n}{n+1}\theta.$$

Hence, $E_\theta((n+1)Y_n/\theta) = \theta$, which means that $(n+1)Y_n/n$ is an unbiased estimator of θ.

8.8 Fisher Information

4. The p.d.f. of the normal distribution is

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

Then

$$\begin{aligned}
\log f(x|\sigma) &= -\log\sigma - \frac{x^2}{2\sigma^2} - \log(\sqrt{2\pi}), \\
\frac{d\log f(x|\sigma)}{d\sigma} &= -\frac{1}{\sigma} + \frac{x^2}{\sigma^3}, \\
\frac{d^2\log f(x|\sigma)}{d\sigma^2} &= \frac{1}{\sigma^2} - \frac{3x^2}{\sigma^4}.
\end{aligned}$$

Therefore,

$$I(\sigma) = -E_\sigma\left[\frac{d^2\log f(X|\sigma)}{d\sigma^2}\right] = -\frac{1}{\sigma^2} + \frac{3E(X^2)}{\sigma^4} = -\frac{1}{\sigma^2} + \frac{3\sigma^2}{\sigma^4} = \frac{2}{\sigma^2}.$$

7. The p.f. of the Bernoulli distribution is

$$f(x|p) = p^x(1-p)^{1-x}$$

Then

$$\begin{aligned}
\log f(x|p) &= x \log p + (1-x)\log(1-p), \\
\frac{d \log f(x|p)}{dp} &= \frac{x}{p} - \frac{1-x}{1-p}, \\
\frac{d^2 \log f(x|p)}{dp^2} &= -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}.
\end{aligned}$$

Therefore,

$$I(p) = -E_p\left[\frac{d^2 \log f(X|p)}{dp^2}\right] = \frac{EX}{p^2} + \frac{1-EX}{(1-p)^2} = \frac{p}{p^2} + \frac{1-p}{(1-p)^2} = \frac{1}{p(1-p)}.$$

Clearly, $E_p(\bar{X}_n) = p$ and $\mathrm{var}_p(\bar{X}_n) = p(1-p)/n$. It follows that $\mathrm{var}(\bar{X}_n)$ equals the lower bound $1/[nI(p)]$ in the information inequality. Thus, $\bar{X}_n$ is an efficient estimator of p.

10. Let $m(\sigma) = \log \sigma$, then $m'(\sigma) = 1/\sigma$ and $[m'(\sigma)]^2 = 1/\sigma^2$. Also, it was shown in Exercise 4 that $I(\sigma) = 2/\sigma^2$. Therefore, if T is an unbiased estimator of $\log \sigma$, it follows from the Eq. (8.8.14) that

$$\mathrm{var}(T) \geq \frac{1}{\sigma^2} \cdot \frac{\sigma^2}{2n} = \frac{1}{2n}.$$

18. The derivative of the log-likelihood with respect to p is

$$\frac{d \log f(x|p)}{dp} = \frac{d}{dp}\left[\log \binom{r+x-1}{x} + r\log(p) + x\log(1-p)\right] = \frac{r}{p} - \frac{x}{1-p}.$$

The mean of $\frac{d \log f(x|p)}{dp}$ is clearly 0, so

$$I(p) = \mathrm{var}_p\left[\frac{d \log f(X|p)}{dp}\right] = \frac{\mathrm{var}_p(X)}{(1-p)^2} = \frac{r(1-p)/p^2}{(1-p)^2} = \frac{r}{p^2(1-p)}.$$

8.9 Supplementary Exercises

14. (a) Since Y has a Poisson distribution with mean $n\theta$, it follows that

$$\begin{aligned}
E(\exp(-cY)) &= \sum_{y=0}^{\infty} \frac{\exp(-cy)\exp(-n\theta)(n\theta)^y}{y!} = \exp(-n\theta)\sum_{y=0}^{\infty}\frac{(n\theta\exp(-c))^y}{y!} \\
&= \exp(-n\theta)\exp[n\theta\exp(-c)] = \exp(n\theta[\exp(-c)-1]).
\end{aligned}$$

Since this expectation must be $\exp(-\theta)$, it follows that $n(\exp(-c)-1) = -1$ and $c = \log[n/(n-1)]$.

(b) To obtain the Fisher information, consider the p.f. of the Poisson distribution

$$f(x|\theta) = \frac{\theta^x}{x!}\exp(-\theta)$$

Then

$$\begin{aligned}
\log f(x|\theta) &= x\log\theta - \theta - \log(x!),\\
\frac{d\log f(x|\theta)}{d\theta} &= \frac{x}{\theta} - 1,\\
\frac{d^2\log f(x|\theta)}{d\theta^2} &= -\frac{x}{\theta^2}.
\end{aligned}$$

Therefore,

$$I(\theta) = -E_\theta\left[\frac{d^2\log f(X|\theta)}{d\theta^2}\right] = \frac{E(X)}{\theta^2} = \frac{\theta}{\theta^2} = \frac{1}{\theta}.$$

Since $m(\theta) = \exp(-\theta)$, $[m'(\theta)]^2 = \exp(-2\theta)$. Hence, from Eq. (8.8.14),

$$\mathrm{var}(\exp(-cY)) \geq \frac{\theta\exp(-2\theta)}{n}.$$

18. The p.d.f. of the exponential distribution is $f(x|\beta) = \beta\exp(-\beta x)$, and the joint p.d.f. is $f_n(\mathbf{x}|\beta) = \beta^n\exp\left(-\beta\sum_{i=1}^n x_i\right)$. Then

$$\log f_n(\mathbf{x}|\beta) = n\log\beta - \beta\sum_{i=1}^n x_i$$

$$\frac{d\log f_n(\mathbf{x}|\beta)}{d\beta} = \frac{n}{\beta} - \sum_{i=1}^n x_i$$

That is,

$$\sum_{i=1}^n x_i = -\frac{d\log f_n(\mathbf{x}|\beta)}{d\beta} + \frac{n}{\beta}.$$

Therefore, $T = \sum_{i=1}^n X_i$ is an efficient estimator. It is known that $E(X_i) = 1/\beta$ and $\mathrm{var}(X_i) = 1/\beta^2$. Hence, $E(T) = n/\beta$ and $\mathrm{var}(T) = n/\beta^2$.

Obviously, $\bar{X}_n = T/n$ is an efficient estimator of $1/\beta$, as $E(\bar{X}_n) = 1/\beta$ and $\mathrm{var}(\bar{X}_n) = 1/(n\beta^2)$, and it can be verified that

$$\mathrm{var}(\bar{X}_n) = [m'(\beta)]^2/[nI(\beta)],$$

where $m(\beta) = 1/\beta$ and $I(\beta) = 1/\beta^2$.