Solutions to Exercise Set 9.

18.4. (a) The observations are i.i.d. with density $f(x|\theta) = \begin{cases} (1/\theta)e^{-x/\theta} & \text{if } 0 < x < T \\ e^{-T/\theta} & \text{if } x = T \end{cases}$ with respect to the measure $\nu(x)$ that equals Lebesgue measure on $(0, T)$ and gives mass one to the point $x = T$. If F_n denotes the number of failures $F_n = \sum_1^n \mathrm{I}(X_i < T)$, and $S_n = \sum_1^n X_i$ denotes the "total time on test", then the likelihood function is $L_n(\theta) = \prod_1^n f(x_i|\theta) = (1/\theta^{F_n})e^{-S_n/\theta}$. The first derivative of $\log L_n(\theta)$ is $-F_n/\theta + S_n/\theta^2$. Setting this to zero and solving yields the MLE, $\hat{\theta}_n = S_n/F_n$. Note that this is interpreted as $+\infty$ if $F_n = 0$.

(b) We find Fisher information as the negative of the expectation of the second derivative of the density.

$$\mathcal{I}(\theta) = -\mathrm{E}(\frac{F_1}{\theta^2} - \frac{2X_1}{\theta^3}) = \frac{2\mathrm{E}X_1}{\theta^3} - \frac{\mathrm{E}(F_1)}{\theta^2}.$$

Note that since $0 = \mathrm{E}(\frac{\partial}{\partial\theta} L_1(\theta)) = -\mathrm{E}(F_1)/\theta + \mathrm{E}(X_1)/\theta^2$, we have $\mathrm{E}(X_1) = \theta\mathrm{E}(F_1)$. And since $\mathrm{E}(F_1) = \mathrm{P}(\text{failure}) = 1 - e^{T/\theta}$, we have $\mathrm{E}(X_1) = \theta(1 - e^{-T/\theta})$. Hence, $\mathcal{I}(\theta) = (1 - e^{-T/\theta})/\theta^2$, and $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \theta^2/(1 - e^{-T/\theta}))$ (even though $\mathrm{E}(\hat{\theta}_n) = +\infty$).

19.5. (a) The log likelihood function is $\ell_n(\alpha, \beta) = \log L_n(\alpha, \beta) = -n \log(\sqrt{2\pi}\sigma) - (1/2\sigma^2) \sum_1^n (y_i - \alpha - \beta x_i)^2$. The partial derivatives with respect to α and β are

$$\frac{\partial \ell_n}{\partial \alpha} = \frac{1}{\sigma^2} \sum (y_i - \alpha - \beta x_i), \qquad \frac{\partial \ell_n}{\partial \beta} = \frac{1}{\sigma^2} \sum (y_i - \alpha - \beta x_i) x_i.$$

and the second partial derivatives are

$$\frac{\partial^2 \ell_n}{\partial \alpha^2} = -\frac{n}{\sigma^2}\,, \qquad \frac{\partial^2 \ell_n}{\partial \alpha \partial \beta} = -\frac{1}{\sigma^2} \sum x_i\,, \qquad \frac{\partial^2 \ell_n}{\partial \beta^2} = -\frac{1}{\sigma^2} \sum x_i^2$$

So Fisher Information is

$$\mathcal{I}_n(\alpha, \beta) = \frac{n}{\sigma^2} \begin{pmatrix} 1 & \overline{x}_n \\ \overline{x}_n & \frac{1}{n} \sum x_i^2 \end{pmatrix}$$

(b) The inverse of Fisher information is

$$\mathcal{I}_n^{-1}(\alpha, \beta) = \frac{\sigma^2}{n s_x^2} \begin{pmatrix} \frac{1}{n} \sum x_i^2 & -\overline{x}_n \\ -\overline{x}_n & 1 \end{pmatrix},$$

so for any unbiased estimate $\hat{\beta}_n$ of β we have $\mathrm{Var}(\hat{\beta}_n) \geq \sigma^2/(n s_x^2)$. The MLE (and least squares estimate) of β is $\hat{\beta}_n = s_{xy}/s_x^2$. It may be written in the form

$$\hat{\beta}_n = \frac{\sum Y_i (x_i - \overline{x}_n)}{n s_x^2} = \beta + \frac{\sum \epsilon_i (x_i - \overline{x}_n)}{n s_x^2}$$

From this we can see that $\hat{\beta}_n$ is unbiased and has variance $\sigma^2/(ns_x^2)$, attaining the CR lower bound.

(c) If α is known, the CR bound is given by the reciprocal of the lower right corner of $\mathcal{I}(\theta)$, namely, $\mathrm{Var}(\hat{\beta}_n) \geq \sigma^2/\sum x_i^2$. In this case, the MLE (or least squares estimate) is

$$\hat{\beta}_n = \frac{\sum(Y_i - \alpha)}{\sum x_i^2} = \beta + \frac{\sum \epsilon_i x_i}{\sum x_i^2}$$

Thus $\hat{\beta}_n$ is unbiased and has variance $\sigma^2/\sum x_i^2$, attaining the lower bound.

20.3. (a) Since $F(x|\theta) = x^\theta$ on (0,1), the median, μ, satisfies $\mu^\theta = 1/2$. So, $\mu(\theta) = 2^{-1/\theta}$.

(b) Since $f(\mu(\theta)|\theta) = \theta 2^{-(\theta-1)/\theta}$, we have from the results of Chapter 13, $\sqrt{n}(\hat{\mu}_n - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2^{-2/\theta}/\theta^2)$.

(c) Fisher information is $1/\theta^2$. To find the Cramér-Rao bound for an unbiased estimate of $\mu = 2^{-1/\theta}$, we have $g(\theta) = 2^{-1/\theta}$, and $\dot{g}(\theta) = (2^{-1/\theta}\log 2)/\theta^2$, so an unbiased estimate of μ has variance at least $\dot{g}(\theta)^2/n\mathcal{I}(\theta) = 2^{-2/\theta}(\log 2)^2/n\theta^2$.

(d) The asymptotic efficiency is $[2^{-2/\theta}(\log 2)^2/\theta^2]/[2^{-2/\theta}/\theta^2] = (\log 2)^2 = .48\cdots$.