Solutions to Exercise Set 8.

18.2. (a) Since the likelihood, $L_n(\theta_1,\theta_2) = \theta_1^K(1-\theta_1)^{n-K}\theta_2^{-n}\exp\{-S/\theta_2\}$, depends on the observations only through K and S, (K,S) is sufficient for (θ_1,θ_2).

(b) $\hat\theta_1 = K/n$ and $\hat\theta_2 = S/n$.

(c) $\mathcal{I}(\theta_1,\theta_2) = \begin{pmatrix} \frac{1}{\theta_1(1-\theta_1)} & 0 \\ 0 & 1/\theta_2^2 \end{pmatrix}$.

(d) $\sqrt{n}(\hat\theta_1-\theta_1,\hat\theta_2-\theta_2) \xrightarrow{\mathcal{L}} \mathcal{N}((0,0),\begin{pmatrix} \theta_1(1-\theta_1) & 0 \\ 0 & \theta_2^2 \end{pmatrix})$.

18.6. (a) By the Central Limit Theorem, $\sqrt{n}((\overline{X}_n,\overline{Y}_n)-(\mu_x,\mu_y)) \xrightarrow{\mathcal{L}} \mathcal{N}((0,0),\Sigma)$. Then we apply Cramér's Theorem with $g(x,y)=x/y$ and $\dot g(x,y) = (1/y,-x/y^2)$, so that $\dot g(\mu_x,\mu_y) = (1/\mu_y)(1,-\theta)$. We find

$$\begin{aligned}\sqrt{n}(\theta_n^*-\theta) &\xrightarrow{\mathcal{L}} \mathcal{N}(0,\frac{1}{\mu_y^2}(1,-\theta)\begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix}\begin{pmatrix} 1 \\ -\theta \end{pmatrix}) \\ &= \mathcal{N}(0,\frac{1}{\mu_y^2}(\sigma_x^2 - 2\theta\sigma_{xy} + \theta^2\sigma_y^2)).\end{aligned}$$

(b) For the exponential distribution of Y, $\mu_y = 1$ and $\sigma_y^2 = 1$. Then $\mathrm{E}(X) = \mathrm{E}(\mathrm{E}(X|Y)) = \mathrm{E}(\theta Y) = \theta$. Similarly, $\mathrm{E}(X^2) = \mathrm{E}(\mathrm{E}(X^2|Y)) = \mathrm{E}(1+\theta^2Y^2) = 1+2\theta^2$, $\sigma_x^2 = 1+\theta^2$, $\mathrm{E}(XY) = \mathrm{E}(\mathrm{E}(XY|Y)) = \mathrm{E}(Y\mathrm{E}(X|Y)) = \mathrm{E}(\theta Y^2) = 2\theta$, and $\sigma_{xy} = \theta$. So in this case, $\sqrt{n}(\theta_n^*-\theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$.

(c) In case (b), the joint density of (X,Y) is $f(x,y|\theta) = (2\pi)^{-1/2}e^{-y}e^{-(x-\theta y)^2/2}$ for $y>0$. We find

$$(\partial/\partial\theta)\log f(x,y|\theta) = (x-\theta y)y.$$

From this we can see that the maximum likelihood estimate of θ is $\hat\theta_n = (\sum X_iY_i/\sum Y_i^2)$. From $(\partial/\partial\theta)^2 f(x,y|\theta) = -y^2$, we see that Fisher information is $\mathcal{I}(\theta) = \mathrm{E}(Y^2) = 2$. Therefore, $\sqrt{n}(\hat\theta_n-\theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1/2)$. The asymptotic efficiency of θ_n^* relative to $\hat\theta_n$ is only 50%.

19.3. (a) The density is $f(x|\mu,\sigma) = \sigma^{-1}\exp\{-e^{-(x-\mu)/\sigma} - (x-\mu)/\sigma\}$. Let $Y = (X-\mu)/\sigma$. Since $\theta = (\mu,\sigma)$ is a location-scale parameter, the distribution of Y does not depend on θ. (It is $G_3(y)$.) We have $\partial Y/\partial\mu = -1/\sigma$ and $\partial Y/\partial\sigma = -Y/\sigma$. Using $\ell = -\log\sigma - e^{-y} - y$, we find

$$\frac{\partial\ell}{\partial\mu} = -\frac{e^y}{\sigma} + \frac{1}{\sigma} \qquad \text{and} \qquad \frac{\partial\ell}{\partial\sigma} = -\frac{1}{\sigma} - \frac{ye^{-y}}{\sigma} + \frac{y}{\sigma}$$

so that $\mathrm{E}e^{-Y} = 1$ and $\mathrm{E}Y = \mathrm{E}Ye^{-Y} + 1$. We also have

$$\frac{\partial^2 \ell}{\partial \mu^2} = -\frac{e^{-y}}{\sigma^2}, \quad \frac{\partial^2 \ell}{\partial \mu \partial \sigma} = -\frac{ye^{-y}}{\sigma^2} + \frac{e^{-y}}{\sigma^2} - \frac{1}{\sigma^2} \quad \text{and} \quad \frac{\partial^2 \ell}{\partial \sigma^2} = \frac{1}{\sigma^2} - \frac{y^2 e^{-y}}{\sigma^2} + \frac{2ye^{-y}}{\sigma^2} - \frac{2y}{\sigma^2}.$$

Taking expectations and simplifying, we find

$$\mathcal{I}(\mu, \sigma) = \frac{1}{\sigma^2} \begin{pmatrix} 1 & \mathrm{E}Ye^{-Y} \\ \mathrm{E}Ye^{-Y} & \mathrm{E}Y^2 e^{-Y} + 1 \end{pmatrix} = \frac{1}{\sigma^2} \begin{pmatrix} 1 & -.42278 \\ -.42278 & 1.82367 \end{pmatrix}.$$

If $X = -\log(Y)$, then X has an exponential distribution with mean 1 $\mathrm{E}Ye^{-Y} = \mathrm{E}Y - 1 = \gamma - 1$ and and Maple tells us that $\mathrm{E}Y^2 e^{-Y} = \mathrm{E}(\log(X)^2 X) = (\pi^2/6) - 2\gamma + \gamma^2$, from which the above numbers follow.

(b) If $\hat{\theta}_n$ is an unbiased estimate of $g(\mu, \sigma) = \mu/\sigma$, then with $\dot{g}(\mu, \sigma) = (1/\sigma, -\mu/\sigma^2)$, we have

$$\mathrm{Var}_\theta(\hat{\theta}_n) \geq \frac{1}{n} \dot{g} \mathcal{I}^{-1} \dot{g}^T = \frac{1}{n}[1.1087 - .5140\frac{\mu}{\sigma} + .6079\frac{\mu^2}{\sigma^2}].$$

(The minimum value occurs at $\mu/\sigma = .4228$.)