Exercise 2.3
Show (1) is equal to (2)

$$\sum_j \sum_k (Y_{jk} - a_j - b_j x_{jk})^2 = \sum_j \sum_k (Y_{jk} - \mu_{jk} + \mu_{jk} - a_j - b_j x_{jk})^2 \tag{1}$$

$$\sum_j \sum_k (Y_{jk} - [\alpha_j + \beta_j x_{jk}])^2 - K \sum_j (\bar{Y}_j - \alpha_j - \beta_j \bar{x}_j)^2 - \sum_j (b_j - \beta_j)^2 (\sum_k x_{jk}^2 - K\bar{x}_j^2) \tag{2}$$

Note $\mu_{jk} = \alpha_j + \beta_j x_{jk}$. Let $A = (Y_{jk} - \mu_{jk})$ and $B = (\mu_{jk} - a_j - b_j x_{jk})$. Then we have that (1) is equal to $\sum_j \sum_k A^2 + 2AB + B^2$. The A^2 term is as desired in the resulting equation, thus what remains to show is the $2AB + B^2$ is equivalent to the final two terms in (2).

$$2AB = \sum_j \sum_k 2Y_{jk}\mu_{jk} - 2Y_{jk}a_j - 2Y_{jk}b_j x_{jk} - 2\mu_{jk}^2 + 2\mu_{jk}a_j + 2\mu_{jk}b_j x_{jk} \tag{3}$$

$$B^2 = \sum_j \sum_k \mu_{jk}^2 + a_j^2 + (b_j x_{jk})^2 - 2\mu_{jk}a_j - 2\mu_{jk}b_j x_{jk} + 2a_j b_j x_{jk} \tag{4}$$

Using the normal equations (pg. 26, third edition) we get (5) and (6). We will also use the identity $\bar{Y}_j = a_j + b_j \bar{x}_j$ and the fact that $\sum_k Y_{jk} = K\bar{Y}_j$.

$$\sum_j \sum_k 2Y_{jk}b_j x_{jk} = 2\sum_j [b_j^2 (\sum_k x_{jk}^2 - K\bar{x}_j^2) + K\bar{Y}_j b_j \bar{x}_j] \tag{5}$$

$$\sum_j \sum_k 2Y_{jk}\beta_j x_{jk} = 2\sum_j [b_j \beta_j (\sum_k x_{jk}^2 - K\bar{x}_j^2) + K\bar{Y}_j \beta_j \bar{x}_j] \tag{6}$$

When we add (3) and (4) we get

$$\sum_j \sum_k 2Y_{jk}\alpha_j + 2Y_{jk}\beta_j x_{jk} - 2Y_{jk}a_j - 2Y_{jk}b_j x_{jk} - \mu_{jk}^2 + a_j^2 + (b_j x_{jk})^2 + 2a_j b_j x_{jk}, \tag{7}$$

which is equal to (8) plus (9).

$$\sum_j [2K\bar{Y}_j\alpha_j + 2K\bar{Y}_j\beta_j\bar{x}_j - 2K\bar{Y}_j a_j - 2K\bar{Y}_j b_j \bar{x}_j \tag{8}$$
$$-K\alpha_j^2 - 2K\alpha_j\beta_j\bar{x}_j - K\beta_j^2\bar{x}_j^2 + Ka_j^2 + 2Ka_j b_j \bar{x}_j + Kb_j^2\bar{x}_j^2]$$

$$\sum_j [2b_j\beta_j(\sum_k x_{jk}^2 - K\bar{x}_j^2) - 2b_j^2(\sum_k x_{jk}^2 - K\bar{x}_j^2) - \beta_j^2 \sum_k x_{jk}^2 + b_j^2 \sum_k x_{jk}^2 + K\beta_j^2\bar{x}_j^2 - Kb_j^2\bar{x}_j^2] \tag{9}$$

With a little algebra, (8) easily morphs into the second term in (2). Now, we simplify (9) in (10), which gives us the third term in (2). Thus we have shown (1) is equal to (2).

$$2b_j\beta_j(\sum_k x_{jk}^2 - K\bar{x}_j^2) - 2b_j^2(\sum_k x_{jk}^2 - K\bar{x}_j^2) + b_j^2(\sum_k x_{jk}^2 - K\bar{x}_j^2) - \beta_j^2(\sum_k x_{jk}^2 - K\bar{x}_j^2)$$
$$= -\sum_j (b_j^2 - 2b_j\beta_j + \beta_j^2)(\sum_k x_{jk}^2 - K\bar{x}_j^2) = -\sum_j (b_j - \beta_j)^2 (\sum_k x_{jk}^2 - K\bar{x}_j^2) \tag{10}$$

Exercise 3.11
Suppose we have a Pareto distribution with $f(y;\theta) = \theta y^{-\theta-1}$, with $\theta > 0$ and $y \geq 1$. We are asked to find the score statistic U, show $\mathrm{E}(U) = 0$ and thus find $\mathcal{I}(\theta) = \mathrm{Var}(U)$

$$f(y;\theta) = \theta y^{-\theta-1} = \exp(\log\theta - \theta\log y - \log y)$$
$$\Leftrightarrow l(\theta) = \log\theta - \theta\log y - \log y$$
$$\Leftrightarrow U = \frac{d}{d\theta}l(\theta) = \frac{1}{\theta} - \log y \tag{11}$$

From the first line in (11), we note that $a(y) = \log y$, $b(\theta) = -\theta$, $c(\theta) = \log\theta$ and $d(y) = -\log y$. Since $\mathrm{E}(a(Y)) = -c'(\theta)/b'(\theta)$, $\mathrm{E}(\log Y) = -(1/\theta)/-1 = 1/\theta$, we have that $\mathrm{E}(U) = 1/\theta - 1/\theta = 0$. Since we have that the expectation of the score statistic is zero, the fisher information is equal to the variance of the score. By results of the exponential family, we know that

$$\mathrm{Var}(a(Y)) = \frac{b''(\theta)c'(\theta) - b'(\theta)c''(\theta)}{[b'(\theta)]^3}, \tag{12}$$

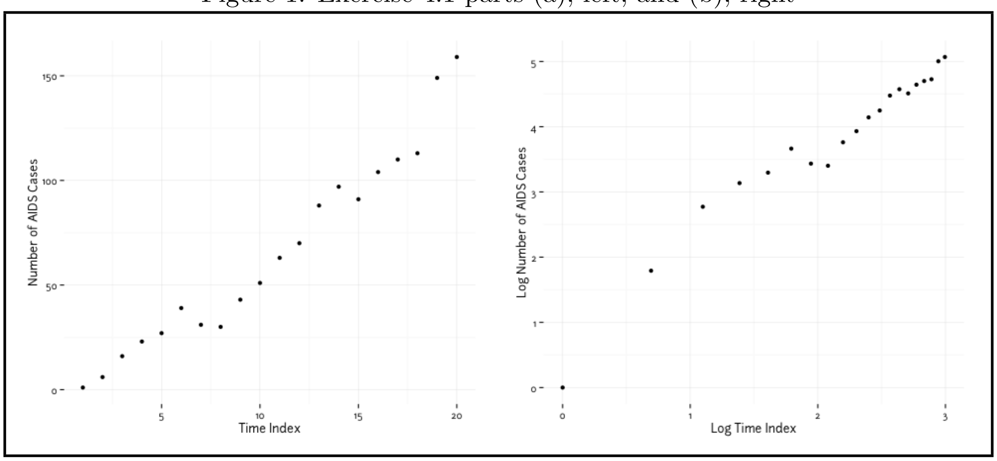
we get $\mathrm{Var}(U) = \mathrm{Var}(\log Y) = [0 - (-1)(-1/\theta^2)]/(-1) = 1/\theta^2$. Notice that the expectation and variance of $\log Y$ are $1/\theta$ and $1/\theta^2$ respectively, the expectation and variance of an exponential random variable with rate θ. Using the method of CDF's, and knowing that $F(y;\theta) = 1 - (1/y)^\theta$ we see that

$$P(\log Y < x) = P(Y < e^x) = 1 - (1/e^x)^\theta = 1 - e^{-\theta x}, \tag{13}$$

which is the CDF of an exponential distribution with rate θ.

Exercise 4.1

Figure 1: Exercise 4.1 parts (a), left, and (b), right



(c) Our model is $g(\mu_i) = \log\lambda_i = \beta_1 + \beta_2 x_i$, with $Y_1, ..., Y_n$ independent Poisson random variables with parameters λ_i. In our model $\eta_i = \log\lambda_i$, so $\mu_i = \lambda_i = e^{\eta_i}$. Since

$$w_{ii} = \frac{1}{\mathrm{Var}(Y_i)}\left(\frac{d\mu_i}{d\eta_i}\right)^2 = \frac{1}{e^{\eta_i}}(e^{\eta_i})^2 = e^{\eta_i} = \exp(\beta_1 + \beta_2 x_i) \tag{14}$$
$$z_i = b_1 + b_2 x_i + (y_i - \lambda_i)(1/\lambda_i) = b_1 + b_2 x_i + \frac{y_i}{\exp(b_1 + b_2 x_i)} - 1 \tag{15}$$

we can use the formula $\mathbf{b}^{(m)} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{z}$, with $\mathbf{W}$ the matrix with diagonal elements $w_{ii} = \exp(b_1 + b_2 x_i)$ evaluated at $\mathbf{b}^{(m-1)}$. Setting $\mathbf{b}^{(0)} = [0\ 2]^T$, converges to the maximum likelihood estimates $\mathbf{b}^{(5)} = [0.995998\ 1.326610]^T$.