8.3 The LRT statistic is

$$\lambda(y) = \frac{\sup_{\theta \le \theta_0} L(\theta|y_1, \ldots, y_m)}{\sup_{\Theta} L(\theta|y_1, \ldots, y_m)}.$$

Let $y = \sum_{i=1}^{m} y_i$, and note that the MLE in the numerator is $\min\{y/m, \theta_0\}$ (see Exercise 7.12) while the denominator has y/m as the MLE (see Example 7.2.7). Thus

$$\lambda(y) = \begin{cases} 1 & \text{if } y/m \le \theta_0 \\ \frac{(\theta_0)^y (1-\theta_0)^{m-y}}{(y/m)^y (1-y/m)^{m-y}} & \text{if } y/m > \theta_0, \end{cases}$$

and we reject H_0 if

$$\frac{(\theta_0)^y (1-\theta_0)^{m-y}}{(y/m)^y (1-y/m)^{m-y}} < c.$$

To show that this is equivalent to rejecting if $y > b$, we could show $\lambda(y)$ is decreasing in y so that $\lambda(y) < c$ occurs for $y > b > m\theta_0$. It is easier to work with $\log \lambda(y)$, and we have

$$\log \lambda(y) = y \log \theta_0 + (m-y) \log(1-\theta_0) - y \log\left(\frac{y}{m}\right) - (m-y) \log\left(\frac{m-y}{m}\right),$$

and

$$\begin{aligned} \frac{d}{dy} \log \lambda(y) &= \log \theta_0 - \log(1-\theta_0) - \log\left(\frac{y}{m}\right) - y\frac{1}{y} + \log\left(\frac{m-y}{m}\right) + (m-y)\frac{1}{m-y} \\ &= \log\left(\frac{\theta_0}{y/m} \frac{\left(\frac{m-y}{m}\right)}{1-\theta_0}\right). \end{aligned}$$

For $y/m > \theta_0$, $1 - y/m = (m-y)/m < 1 - \theta_0$, so each fraction above is less than 1, and the log is less than 0. Thus $\frac{d}{dy} \log \lambda < 0$ which shows that λ is decreasing in y and $\lambda(y) < c$ if and only if $y > b$.

8.5 a. The log-likelihood is

$$\log L(\theta,\nu|\mathbf{x}) = n\log\theta + n\theta\log\nu - (\theta+1)\log\left(\prod_i x_i\right), \quad \nu \le x_{(1)},$$

where $x_{(1)} = \min_i x_i$. For any value of θ, this is an increasing function of ν for $\nu \le x_{(1)}$. So both the restricted and unrestricted MLEs of ν are $\hat{\nu} = x_{(1)}$. To find the MLE of θ, set

$$\frac{\partial}{\partial\theta}\log L(\theta, x_{(1)}|\mathbf{x}) = \frac{n}{\theta} + n\log x_{(1)} - \log\left(\prod_i x_i\right) = 0,$$

and solve for θ yielding

$$\hat{\theta} = \frac{n}{\log(\prod_i x_i/x_{(1)}^n)} = \frac{n}{T}.$$

$(\partial^2/\partial\theta^2)\log L(\theta, x_{(1)}|\mathbf{x}) = -n/\theta^2 < 0$, for all θ. So $\hat{\theta}$ is a maximum.

b. Under H_0, the MLE of θ is $\hat{\theta}_0 = 1$, and the MLE of ν is still $\hat{\nu} = x_{(1)}$. So the likelihood ratio statistic is

$$\lambda(\mathbf{x}) = \frac{x_{(1)}^n/(\prod_i x_i)^2}{(n/T)^n x_{(1)}^{n^2/T}\Big/(\prod_i x_i)^{n/T+1}} = \left(\frac{T}{n}\right)^n \frac{e^{-T}}{(e^{-T})^{n/T}} = \left(\frac{T}{n}\right)^n e^{-T+n}.$$

$(\partial/\partial T)\log\lambda(\mathbf{x}) = (n/T) - 1$. Hence, $\lambda(\mathbf{x})$ is increasing if $T \le n$ and decreasing if $T \ge n$. Thus, $T \le c$ is equivalent to $T \le c_1$ or $T \ge c_2$, for appropriately chosen constants c_1 and c_2.

c Under the null hypothesis $H_0 : \theta = 1$, the Pareto densities form a scale family

$$\begin{aligned} f(x\,|\,1,\nu) &= \frac{v}{x^2} I_{[\nu,\infty)}(x) = \frac{1}{\nu}\frac{1}{(x/\nu)^2} I(x/\nu \ge 1) \\ &= \frac{1}{\nu}\psi\left(\frac{x}{\nu}\right) \quad \text{where} \quad \psi(x) = \frac{1}{x^2} I(x \ge 1)\,. \end{aligned}$$

The statistic T is scale invariant:

$$\begin{aligned} T(c\boldsymbol{x}) &= \log\left[\frac{\prod_{i=1}^n cx_i}{(\min_i cx_i)^n}\right] = \log\left[\frac{c^n\prod_{i=1}^n x_i}{(c\ \min_i x_i)^n}\right] = \log\left[\frac{c^n\prod_{i=1}^n x_i}{c^n(\min_i x_i)^n}\right] \\ &= \log\left[\frac{\prod_{i=1}^n x_i}{(\min_i x_i)^n}\right] = T(\boldsymbol{x})\,. \end{aligned}$$

Thus, when $\theta = 1$, the distribution of T does not depend on the value of the scale parameter ν (i.e., it is ancillary), so we may assume $\nu = 1$. In this case $X_1, \ldots, X_n$ are iid with the density $\psi(x) = 1/x^2$ for $x \ge 1$.

We now show that if a random variable X has density ψ, then $Y = \log X$ is exp(1), an exponential random variable with mean 1. The function $g(x) = \log(x)$ is differentiable and monotonic so that we may apply the general result from Chapter 2:

$$f_Y(y) = f_X(g^{-1}(x))\left|\frac{d}{dy}g^{-1}(y)\right| \quad \text{for} \quad y \in \mathcal{Y}\,.$$

In our case, $f_X = \psi$ and $g(x) = \log(x)$ with inverse $g^{-1}(y) = e^y$. The support of X is $\mathcal{X} = [1,\infty)$, so that the support of $Y = \log(X)$ is $\mathcal{Y} = [0,\infty)$. Plugging this into the above leads to

$$\begin{aligned} f_Y(y) &= \psi(e^y)\left|\frac{d}{dy}e^y\right| \quad \text{for} \quad y \in [0,\infty) \\ &= \frac{1}{(e^y)^2}\,e^y = e^{-y} \text{ for } y \ge 0 \end{aligned}$$

which is the exp(1) pdf.

Let $Y_i = \log X_i$ for $i = 1, \ldots, n$. The statistic T may be written as

$$
\begin{aligned}
T &= \log\left[\frac{\prod_{i=1}^{n} X_i}{(\min_i X_i)^n}\right] = \log\left[\prod_{i=1}^{n} X_i\right] - \log\left[(\min_i X_i)^n\right] = \sum_{i=1}^{n}\log X_i - n\,\log(\min_i X_i) \\
&= \sum_{i=1}^{n}\log X_i - n\,\min_i(\log X_i) = \sum_{i=1}^{n} Y_i - n\,\min_i Y_i = \sum_{i=1}^{n}(Y_i - Y_{(1)}) \quad \text{where } Y_{(1)} = \min Y_i.
\end{aligned}
$$

Now use results about exponential random variables from last semester (which were derived using the memoryless property). Think of $Y_1, \ldots, Y_n$ as the lifetimes of n light bulbs which are all turned on at time zero. At the time $Y_{(1)}$ at which the first bulb burns out, the remaining bulbs (because of the memoryless property) are "good as new", that is, the remaining lifetime $Y_i - Y_{(1)}$ is exp(1). Thus, in the summation $T = \sum_{i=1}^{n}(Y_i - Y_{(1)})$, one of the n terms is zero (corresponding to the first bulb to burn out for which $Y_i = Y_{(1)}$), and the other $n-1$ terms are iid exp(1) rv's. Thus T has the same distribution as the sum of $n-1$ exponential rv's. But $\exp(1) \equiv \text{Gamma}(\alpha = 1, \beta = 1)$, so that by the closure property for the summation of Gamma rv's, $T \sim \text{Gamma}(\alpha = n-1, \beta = 1)$.

The χ^2 distributions are a special case of the Gamma distributions, and our result may be restated in terms of χ^2 distributions as follows. Let Z_i denote iid exp(1) rv's. We argued above that $T \overset{d}{=} \sum_{i=1}^{n-1} Z_i$. Now note that $2Z_i \sim \exp(2) \equiv \text{Gamma}(\alpha = 1, \beta = 2) \equiv \chi^2_2$. Thus, by the closure property for χ^2 rv's,

$$
2T \overset{d}{=} \sum_{i=1}^{n-1} 2Z_i \sim \chi^2_{2(n-1)}\,.
$$

8.6 a.

$$\lambda(\mathbf{x},\mathbf{y}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{x},\mathbf{y})}{\sup_{\Theta} L(\theta|\mathbf{x},\mathbf{y})} = \frac{\sup_\theta \prod_{i=1}^n \frac{1}{\theta}e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\theta}e^{-y_j/\theta}}{\sup_{\theta,\mu} \prod_{i=1}^n \frac{1}{\theta}e^{-x_i/\theta} \prod_{j=1}^m \frac{1}{\mu}e^{-y_j/\mu}}$$

$$= \frac{\sup_\theta \frac{1}{\theta^{m+n}} \exp\left\{-\left(\sum_{i=1}^n x_i + \sum_{j=1}^m y_j\right)/\theta\right\}}{\sup_{\theta,\mu} \frac{1}{\theta^n} \exp\{-\sum_{i=1}^n x_i/\theta\} \frac{1}{\mu^m} \exp\left\{-\sum_{j=1}^m y_j/\mu\right\}}.$$

Differentiation will show that in the numerator $\hat{\theta}_0 = (\sum_i x_i + \sum_j y_j)/(n+m)$, while in the denominator $\hat{\theta} = \bar{x}$ and $\hat{\mu} = \bar{y}$. Therefore,

$$\lambda(\mathbf{x},\mathbf{y}) = \frac{\left(\frac{n+m}{\sum_i x_i + \sum_j y_j}\right)^{n+m} \exp\left\{-\left(\frac{n+m}{\sum_i x_i + \sum_j y_j}\right)\left(\sum_i x_i + \sum_j y_j\right)\right\}}{\left(\frac{n}{\sum_i x_i}\right)^n \exp\left\{-\left(\frac{n}{\sum_i x_i}\right)\sum_i x_i\right\} \left(\frac{m}{\sum_j y_j}\right)^m \exp\left\{-\left(\frac{m}{\sum_j y_j}\right)\sum_j y_j\right\}}$$

$$= \frac{(n+m)^{n+m}}{n^n m^m} \frac{\left(\sum_i x_i\right)^n \left(\sum_j y_j\right)^m}{\left(\sum_i x_i + \sum_j y_j\right)^{n+m}}.$$

b.

$$\lambda = \frac{(n+m)^{n+m}}{n^n m^m} \left(\frac{\sum_i x_i}{\sum_i x_i + \sum_j y_j}\right)^n \left(\frac{\sum_j y_j}{\sum_i x_i + \sum_j y_j}\right)^m = \frac{(n+m)^{n+m}}{n^n m^m} T^n (1-T)^m.$$

Therefore λ is a function of T. λ is a unimodal function of T which is maximized when $T = \frac{n}{m+n}$. Rejection for $\lambda \le c$ is equivalent to rejection for $T \le a$ or $T \ge b$, where a and b are constants that satisfy $a^n(1-a)^m = b^n(1-b)^m$.

c. When H_0 is true, $\sum_i X_i \sim \text{gamma}(n,\theta)$ and $\sum_j Y_j \sim \text{gamma}(m,\theta)$ and they are independent. So by an extension of Exercise 4.19b, $T \sim \text{beta}(n,m)$.

8.15 From the Neyman-Pearson lemma the UMP test rejects H_0 if

$$\frac{f(x \mid \sigma_1)}{f(x \mid \sigma_0)} = \frac{(2\pi\sigma_1^2)^{-n/2} e^{-\Sigma_i x_i^2/(2\sigma_1^2)}}{(2\pi\sigma_0^2)^{-n/2} e^{-\Sigma_i x_i^2/(2\sigma_0^2)}} = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left\{\frac{1}{2}\sum_i x_i^2 \left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)\right\} > k$$

for some $k \geq 0$. After some algebra, this is equivalent to rejecting if

$$\sum_i x_i^2 > \frac{2\log\left(k\left(\sigma_1/\sigma_0\right)^n\right)}{\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)} = c \quad \left(\text{because } \frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2} > 0\right).$$

This is the UMP test of size α, where $\alpha = P_{\sigma_0}(\sum_i X_i^2 > c)$. To determine c to obtain a specified α, use the fact that $\sum_i X_i^2/\sigma_0^2 \sim \chi_n^2$. Thus

$$\alpha = P_{\sigma_0}\left(\sum_i X_i^2/\sigma_0^2 > c/\sigma_0^2\right) = P\left(\chi_n^2 > c/\sigma_0^2\right),$$

so we must have $c/\sigma_0^2 = \chi^2_{n,\alpha}$, which means $c = \sigma_0^2 \chi^2_{n,\alpha}$.

8.20 By the Neyman-Pearson Lemma, the UMP test rejects for large values of $f(x|H_1)/f(x|H_0)$. Computing this ratio we obtain

x	1	2	3	4	5	6	7
$\frac{f(x\|H_1)}{f(x\|H_0)}$	6	5	4	3	2	1	.84

The ratio is decreasing in x. So rejecting for large values of $f(x|H_1)/f(x|H_0)$ corresponds to rejecting for small values of x. To get a size α test, we need to choose c so that $P(X \leq c|H_0) = \alpha$. The value $c = 4$ gives the UMP size $\alpha = .04$ test. The Type II error probability is $P(X = 5, 6, 7|H_1) = .82$.