Problem 1

The likelihood function is:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n}\Big[\frac{\theta}{2\sqrt{y_i}}e^{-\theta\sqrt{y_i}}\Big] \\ &= \frac{\theta^n}{2^n}e^{-\theta\sum_{i=1}^{n}\sqrt{y_i}}\prod_{i=1}^{n}\Big(\frac{1}{\sqrt{y_i}}\Big) \end{aligned}$$

And

$$\begin{aligned} l(\theta) &= \ln L(\theta) \\ &= n\ln\theta - n\ln 2 - \frac{1}{2}\sum_{i=1}^{n}\ln y_i - \theta\sum_{i=1}^{n}\sqrt{y_i} \end{aligned}$$

Make $dl(\theta)/d\theta = 0$, then we have:

$$\frac{n}{\theta} - \sum_{i=1}^{n}\sqrt{y_i} = 0$$

Thus, we can get MLE $\hat{\theta}$:

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n}\sqrt{y_i}}$$

And by the sample given:

$$\hat{\theta} = \frac{4}{\sqrt{6.2}+\sqrt{7.0}+\sqrt{2.5}+\sqrt{4.2}} = 0.4563$$

Now we are computing $P(2 \le \hat{\theta} \le 3)$, since this is a continuous random variable:

$$P(2 \le \hat{\theta} \le 3) = P(\hat{\theta} \le 3) - P(\hat{\theta} \le 2)$$

Let's just consider $P(\hat{\theta} \le 3)$ first, we have:

$$\begin{aligned} P(\hat{\theta} \le 3) &= P\Big(\frac{n}{\sum_{i=1}^{n}\sqrt{y_i}} \le 3\Big) \\ &= P\Big(\frac{\sum_{i=1}^{n}\sqrt{y_i}}{n} \ge \frac{1}{3}\Big) \end{aligned}$$

By Central Limit Theorem, as $n \to \infty$, we have:

$$P\Big(\frac{\frac{\sum_{i=1}^{n}\sqrt{y_i}}{n} - \mu}{\sigma} \le x\Big) \to \Phi(x)$$

Where $\Phi(\cdot)$ is the distribution function of stand normal random variable, and μ is the mean of $\frac{\sum_i\sqrt{y_i}}{n}$, σ^2 is the variance of $\frac{\sum_i\sqrt{y_i}}{n}$. First, we need to find what's the distribution of $\sqrt{y_i}$, as y_i's are i.i.d., we only need to find $\sqrt{y_1}$'s. First, we can easily to get y_1's CDF:

$$F_Y(y) = 1 - e^{\theta\sqrt{y}}$$

And

$$\begin{aligned} F_{\sqrt{Y}}(y) &= P(\sqrt{Y} \leq y) \\ &= P(Y \leq y^2) \\ &= F_Y(y^2) \\ &= 1 - e^{\theta y} \end{aligned}$$

Thus $\sqrt{y_1}$ actually follows exponential distribution, which has mean $\frac{1}{\theta}$ and variance $\frac{1}{\theta^2}$, moreover:

$$\mu = E\Big(\frac{\sum_i \sqrt{y_i}}{n}\Big) = \frac{1}{\theta}$$

and variance σ^2:

$$\sigma^2 = Var\Big(\frac{\sum_i \sqrt{y_i}}{n}\Big) = \frac{1}{n^2}\sum_{i=1}^{n} Var(\sqrt{y_i}) == \frac{1}{n\theta^2}$$

Therefore, we have:

$$\begin{aligned} P\Big(\frac{\sum_{i=1}^n \sqrt{y_i}}{n} \geq \frac{1}{3}\Big) &= 1 - P\Big(\frac{\sum_{i=1}^n \sqrt{y_i}}{n} \leq \frac{1}{3}\Big) \\ &= 1 - P\Big(\frac{\frac{\sum_{i=1}^n \sqrt{y_i}}{n} - \frac{1}{\theta}}{\frac{1}{\sqrt{n}\theta}} \leq \frac{\frac{1}{3} - \frac{1}{\theta}}{\frac{1}{\sqrt{n}\theta}}\Big) \\ &\approx 1 - \Phi\Big(\sqrt{n}(\frac{\theta}{3} - 1)\Big) \end{aligned}$$

In the same way, we have:

$$P(\hat{\theta} \leq 2) \approx 1 - \Phi\Big(\sqrt{n}(\frac{\theta}{2} - 1)\Big)$$

Finally,

$$P(2 \leq \hat{\theta} \leq 3) = \Phi\Big(\sqrt{n}(\frac{\theta}{2} - 1)\Big) - \Phi\Big(\sqrt{n}(\frac{\theta}{3} - 1)\Big))$$

`Problem 2`

(a) In this problem, we can't find a θ to make $\frac{dL(\theta)}{d\theta} = 0$, so we are going to find the one make $L(\theta)$ largest.

$$L(\theta) = \frac{1}{\theta^n}, 0 \leq y_1, \ldots, y_n \leq \theta$$

Thus, the smaller θ, the larger $L(\theta)$. Notice θ need to be bigger than every y_i, then the smallest θ we can get is actually $\max_i y_i$.

$$\hat{\theta} = \max_{i=1,\ldots,n} \{y_1, y_2, \ldots, y_n\} = 14.2$$

Review section 3.10, $\hat{\theta}$ is actually the order statistic Y_n', its distribution function is:

$$F_{\hat{\theta}}(y) = [F_Y(y)]^n = \begin{cases} 0, & y < 0 \\ \frac{y^n}{\theta^n}, & 0 \leq y \leq \theta \\ 1, & y > 1 \end{cases}$$

(b) Same as part(a), the derivative can't be 0, we just need to find the $\hat{\theta}_1$ and $\hat{\theta}_2$, to make $L(\theta_1, \theta_2)$ largest, then they are MLE.

$$L(\theta_1, \theta_2) = \frac{1}{(\theta_2 - \theta_1)^n}, \theta_1 \leq y_1, \ldots, y_n \leq \theta_2$$

Thus, the problem become to find largest θ_1 and smallest θ_2. Similar as above, we have:

$$\hat{\theta}_1 = Y_1' = \min_{i=1,\ldots,n} \{y_1, y_2, \ldots, y_n\} = 1.8$$

$$\hat{\theta}_2 = Y_n' = \max_{i=1,\ldots,n} \{y_1, y_2, \ldots, y_n\} = 14.2$$

Problem 3

The maximum likelihood function is:

$$L(\theta) = \frac{2^n \prod_{i=1}^n y_i}{\theta^{2n}}$$

Thus, as the same discussion in 5.2.10, $\hat{\theta} = Y_n' = \max_{i=1,\ldots,n}\{y_1, \ldots, y_n\}$. By theorem in section 3.10, we have PDF of $\hat{\theta}$:

$$\begin{aligned} f_{\hat{\theta}}(y) &= n[F_Y(y)]^{n-1} f_Y(y) \\ &= \frac{2n}{\theta^{2n}} y^{2n-1} \end{aligned}$$

Then the expectation is:

$$E(\hat{\theta}) = \int_0^{\theta} y \cdot \frac{2n}{\theta^{2n}} y^{2n-1} dy = \frac{2n}{2n+1}\theta$$

`Problem 4`

(a) (The additional question cancelled, but if some one do that, he can get extra credits)
The likelihood function is:

$$\begin{aligned} l(\alpha) &= \ln L(\alpha) \\ &= n \ln \alpha + n \ln \beta - \alpha \sum_{i=1}^{n} y_i^{\beta} + (\beta - 1) \sum_{i=1}^{n} \ln y_i \end{aligned}$$

To make the derivative of $l(\theta)$ be 0, we can have:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} y_i^{\beta}}$$

To compute $E(\hat{\alpha})$, first, we claim Y_i^{β} follows exponential distribution $exp(\alpha)$ (Proof skipped, it can be showed by CDF). Then the $\sum_{i=1}^{n} y_i^{\beta}$ follows Erlang Distribution, which has PDF:

$$f_{\sum_{i=1}^{n} y_i^{\beta}}(y) = \frac{\alpha^n y^{n-1} e^{-\alpha y}}{(n-1)!}, y \geq 0$$

Thus the expectation is:

$$\begin{aligned} E(\hat{\alpha}) &= E_{\sum_{i=1}^{n} y_i^{\beta}}\left(\frac{n}{\sum_{i=1}^{n} y_i^{\beta}}\right) \\ &= \int_0^{\infty} \frac{n}{y} \cdot \frac{\alpha^n y^{n-1} e^{-\alpha y}}{(n-1)!} dy \\ &= \frac{n}{n-1}\alpha \int_0^{\infty} \frac{\alpha^{n-1} y^{n-2} e^{-\alpha y}}{(n-2)!} dy \\ &= \frac{n}{n-1}\alpha \end{aligned}$$

(b) We have already gotten the likelihood function in part (a)

$$l(\alpha, \beta) = n \ln \alpha + n \ln \beta - \alpha \sum_{i=1}^{n} y_i^{\beta} + (\beta - 1) \sum_{i=1}^{n} \ln y_i$$

Make $\partial l(\alpha,\beta)/\partial\beta = 0$ and $\partial l(\alpha,\beta)/\partial\alpha = 0$, we get:

$$\begin{cases} \frac{n}{\beta} + \sum_{i=1}^n \ln y_i - \alpha\beta \sum_{i=1}^n y_i^\beta \ln y_i = 0 \\ \frac{n}{\alpha} - \sum_{i=1}^n y_i^\beta = 0 \end{cases}$$

`Problem 5`

Let $X_1 = k_1, \cdots, X_n = k_n$. Then given parameters p and n, the probability that $X_i = p_i$ is $\binom{n}{k_i} p^{k_i}(1-p)^{n-k_i}$. So we have

$$L(p) = \binom{n}{k_1} p^{k_1}(1-p)^{n-k_1} \cdots \binom{n}{k_n} p^{k_n}(1-p)^{n-k_n}.$$

We take natural log to make things easier to work with:

$$\ln(L(p)) = \ln\left(\binom{n}{k_1}\cdots\binom{n}{k_2}\right) + k_1 \ln p + \cdots + k_n \ln p + (n-k_1)\ln(1-p) + \cdots (n-k_n)\ln(1-p)$$

To maximize we set the derivative to 0.

$$\frac{d\ln L(p)}{dp} = \frac{k_1}{p} + \cdots + \frac{k_n}{p} + \frac{n-k_1}{1-p} + \cdots \frac{n-k_n}{1-p} = 0$$

Simplifying (common denominator) we have

$$(k_1 + \cdots + k_n)(1-p) - n^2 p + (k_1 + \cdots + k_n)p = \sum_{i=1}^n k_i - n^2 p = 0.$$

So we have that $\sum k_i = n^2 p$, giving that the maximizer is $p = \frac{k_1+\cdots+k_n}{n^2}$. So the maximum likelihood estimator is $\frac{X_1+\cdots+X_n}{n^2}$.

`Problem 6`

First, given $Y_1, \cdots, Y_{16}$ as data points, we know from Example 5.2.4 that the maximum likelihood estimator for μ will be $\hat{\mu} = \frac{1}{16}\sum Y_i = \overline{Y}$. We also know, by Theorem 4.3.3, that

$$\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{Y} - 20}{10/\sqrt{16}}$$

has a standard normal distribution. Thus the probability that $\overline{Y} = \hat{\mu}$ will be between 19 and 21 is $P(\frac{19-20}{2.5} \leq Z \leq \frac{21-20}{2.5}) \approx .31$.

If n is increased, then $\sigma/\sqrt{n}$ decreases. This means that the probability that the estimator is in the given range is the same as the probability that Z is in a larger range than before, and thus the probability increases.

Problem 7

Look at Theorem 3.10.1 for the pdf of the smallest order statistic. We'll find $E(Y_{\min})$ and then transform so that the result will have expected value θ.

$E(Y_{\min}) = \int_0^\theta yn(1-\frac{y}{\theta})^{n-1}\frac{1}{\theta}dy = \frac{n}{\theta}\int_0^1 \theta x(1-x)^{n-1}\theta dx =$ $n\theta\int_0^1 x(1-x)^{n-1}dx$ by u-substitution. That integral is easily done by integration by parts (let your u be x, and your dv be the $(1-x)^{n-1}$). The result is that $E(Y_{\min}) = \frac{1}{n+1}\theta$. Hence, $(n+1)Y_{\min}$ is an unbiased estimator for θ.

Problem 8

Again, recalling Theorem 3.10.1, we calculate E$(nY_{\min}) = n^2\int_0^\infty ye^{-\frac{(n-1)y}{\theta}}\frac{1}{\theta}e^{-\frac{y}{\theta}}dy$ $= n^2\frac{1}{\theta}\int_0^\infty ye^{-\frac{ny}{\theta}}dy = \theta\int_0^\infty xe^{-x}dx$ after u-substitution with $x = \frac{ny}{\theta}$. The integral we're now left with is 1 (use integration by parts with u = x), so we just end up with θ, making our estimator unbiased.

What about $\frac{1}{n}\sum_0^n Y_i$? Well, $E(\frac{1}{n}\sum_0^n Y_i) = \frac{1}{n}\sum_0^n E(Y) = E(Y)$, so we just need to show the mean of Y is θ. E(Y) $= \frac{1}{\theta}\int_0^\infty ye^{-\frac{y}{\theta}}dy = \theta\int_0^\infty xe^{-x}dx$ by u-substitution, and as before, this equals θ. Hence our estimator is unbiased.

Problem 9

We have var$(\hat{\theta}_1) = \frac{36}{25}var(Y_{\max})$ and var$(\hat{\theta}_2) = 36$var$(Y_{\min})$. As the problem hints at, var$(Y_{\max})$ and var$(Y_{\min})$ are equal. To reason why is that there's a symmetry between the two of the form $x \to \theta - x$. In other words, the min and the max behave the same, just on opposite sides of the interval $[0, \theta]$.

So, given that, it's clear that $36v$ is worse than $\frac{36}{25}v$ (ie, lower variance is better), meaning $\hat{\theta}_1$ is the estimator we should prefer.

Does this make intuitive sense? At first, perhaps not, since I just said that there's a symmetry between the min and the max. However, that symmetry is broken when we convert them into an estimate for θ. The reason is that on average, the min is $\frac{1}{n+1}$ of the way to θ, and the max is $\frac{n}{n+1}$ of the way there. So we guess that θ is n+1 times the min and $\frac{n+1}{n}$ times the max. But that makes the variations in the min get magnified more than they do for the max.

Problem 10

The relative efficiency of $\hat{\lambda}_1$ to $\hat{\lambda}_2$ is the ratio $\frac{var(\hat{\lambda}_2)}{var(\hat{\lambda}_1)}$. $var(\hat{\lambda}_1)$ is simply the variance of a Poisson distribution, call it v. $var(\hat{\lambda}_2)$ is $\frac{1}{n^2}\sum_1^n var(X_i) = \frac{1}{n}var(X) = \frac{1}{n}v$. So the relative efficiency is simply $\frac{1}{n}$.