3 (a) Show that the family of joint densities of the $X_i$ is a one-parameter exponential family.

3 (b) Identify a "canonical" parameter $\eta = \eta (x)$ and sufficient statistic $T = T(X_1, \ldots, X_n)$.

4 (c) How do we know the sufficient statistic is complete?

4 (d) State the Lehmann–Scheffé theorem.

3 (e) We know $S^2/(n-1)$ would be unbiased for $\text{var } X_i$ if the $X_i$ were a normal sample. Explain in a sentence why it's also true for our $X_i$, which are Poisson.

7 (f) Find with explanation $E [S^2 | X]$.

5 (g) How do we know $S^2$ is not a sufficient statistic? (Don't just say it doesn't appear in the exponential-family form in part (a)! That's not enough.)

4 (h) Explain whether $\sqrt{\frac{S^2}{(n-1)}}$ is biased up or down for $\sqrt{\text{var }}$.

3 (i) What about $\sqrt{\frac{S^2}{n}}$? Biased up or down?

4 (j) What about $\sqrt{\frac{S^2}{L_n-1}}$, where $L_n^{-1} = \sqrt{\frac{\pi}{\Gamma} (\frac{n}{\pi}) / \Gamma (\frac{n}{\pi})}$? Does this get rid of the bias? Comment.

Now we're going to reparameterize the Poisson family by the "zero-probability $\pi = e^{-\lambda}$, so that $E X_i = -\log \lambda$.

3 (k) Write the parent probability function of $X_i$ in the new form involving $\pi$.

5 (l) Hence find the Cramér–Rao bound for the variance of an unbiased estimator (using the whole sample) of $E X_i = -\log \lambda$.

5 (m) Explain whether this bound is achievable.

4 (n) If so, by what estimator?

7 (o) Assume $n = 2016$: Find the correlation between $\bar{X}$ and $\bar{X}$, where $\bar{X}$ is the sample mean of the first 1,008 observations.
(a) Easy 
(b) Easy 
(c) This is exactly Example 6.23(i)(b) on page 42 of textbook.
The parameter space contains an open set, and the Poisson family
\{ P(\lambda), \lambda > 0 \} is of full rank, by Theorem 6.22. It is complete.
proof using definition would unfortunately not work clearly.
(d) Easy 
(e) The unbiasedness of
\[ \delta^2 = \frac{S^2}{n-1} \]
is distribution free, since we did not use the distribution information to get that
\[ \frac{S^2}{n-1} \]
is unbiased.
(f) 
\[ E \left[ \frac{S^2}{n-1} \right] = \lambda \] and \( \bar{X} \) is \( \lambda \), \( \bar{X} \) is complete sufficient.

\( \bar{X} \) is the unique UMVUE for \( \lambda \).
\( E \left[ \frac{S^2}{n-1} | \bar{X} \right] \) is also the unique UMVUE for \( \lambda \) by Lehmann-Scheffe.

Thus, \( E \left[ \frac{S^2}{n-1} | \bar{X} \right] = \bar{X} \), \( E [ \delta^2 | \bar{X} ] = (n-1) \bar{X} \).

(g) solution 1: \( f(x_1, \ldots, x_n) = \exp \{ \log \lambda \cdot \sum x_i - n \lambda \} h(x) \).

\[ = \exp \{ (s^2 + n\bar{x}^2) \log \lambda - n \lambda \} h(x) \]
We cannot factorize the density into part that depends on \( \lambda \) and \( \bar{x} \) only through \( S^2 \).

Solution 2: It is easy to show that \( \bar{X} \) is minimal sufficient,
if \( S^2 \) is sufficient, then there exists a function \( H \) such that
\[ S^2 = H(\bar{X}) \]. But unfortunately, \( H \) does not exist.

Therefore \( S^2 \) is not sufficient.
(h) \( y = \sqrt{x} \) is a concave function, by Jensen inequality,
\[
E \sqrt{\frac{s^2}{n-1}} \leq \sqrt{E \frac{s^2}{n-1}} = \sqrt{\lambda} = \lambda
\]
So it's biased down.

(i) Since \( \frac{s^2}{n} < \frac{s^2}{n-1} \) for all \( n > 1 \), \( E(\frac{s^2}{n}) < E(\frac{s^2}{n-1}) \leq \sqrt{\lambda} \)
So it's still biased down.

(j) If \( X_i \sim \text{iid } N(\mu, \sigma^2) \) then \( E(\frac{s^2}{\sigma^2}) = (\sigma^2)^{-\frac{1}{2}} L_{n-1, 1} \)
Which means \( \frac{s^2}{L_{n-1, 1}} \) is unbiased for \( \sigma \).
This unbiasedness, or bias correction, is based on the normality of \( X_i \)'s and chi-distribution of \( \frac{L_{n-1, 1}}{\sigma} \), thus it is unbiased for normal iid \( X_i \)'s only.
Therefore this does not get rid of bias for poisson.

(k) \( g_{\pi}(x) = \pi . \frac{(-\log \pi)^x}{x!} \)

(l) \( g_{\pi}(x) = E_{\pi} \left[ \frac{d}{d\pi} \log g_{\pi}(x) \right]^2 = -E_{\pi} \left[ \frac{d^2}{d\pi^2} \log g_{\pi}(x) \right] \)
\[
\frac{d}{d\pi} \log g(x) = \frac{d}{d\pi} \left( \log \pi + x \log (-\log \pi) \right)
= \frac{1}{\pi} - \frac{x}{\log \pi} \cdot (-\frac{1}{\pi}) = \frac{1}{\pi} + \frac{x}{\pi \log \pi}
\]
\[
\frac{d^2}{d\pi^2} \log g(x) = -\frac{1}{\pi^2} - \frac{x}{\pi^2 \log \pi} - \frac{x}{\pi \log \pi} \cdot \frac{1}{\pi} = -\frac{1}{\pi^2} - \frac{x}{\pi^2 \log \pi} - \frac{x}{\pi \log \pi}^2
\]
\( g(x) = E \left[ \frac{1}{\pi^2} + \frac{x}{\pi^2 \log \pi} + \frac{x}{\pi \log \pi} \right] = \frac{1}{\pi^2} + \frac{1}{\pi^2 \log \pi} (-\log \pi) - \frac{\log \pi}{\pi^2 \log \pi}
\]
\( = -\frac{1}{\pi^2 \log \pi} \)
\[ I(\pi) = n \frac{i(\pi)}{\pi^2 \log \pi} \quad \text{and} \quad g(\pi) = \lambda = -\log n \]

Cramer–Rao bound:
\[ \text{Var} \ S(x) \geq \frac{(g'(\pi))^2}{I(\pi)} = \frac{1}{\pi^2} \cdot \frac{n}{\pi^2 \log n} = \frac{\log n}{n} \]

\[(n)\ (o). \]
\[ \text{Var} (\bar{x}) = \frac{\text{Var}(x)}{n} = \frac{\lambda}{n}, \text{it is achieved by } \bar{x}. \]

\[(p)\]
\[ \rho = \frac{\text{Cov} (\bar{x}, \bar{x})}{\sqrt{\text{Var} \bar{x} \cdot \sqrt{\text{Var} \bar{x}}}} \quad \text{Var} \bar{x} = \frac{\text{Var} x}{n} = \frac{\lambda}{2016} \]
\[ \text{Var} \bar{x} = \frac{\lambda}{1008} \]
\[ \text{Cov} (\bar{x}, \bar{x}) = \text{Cov} \left( \frac{1}{2016} \sum_i x_i, \frac{1}{1008} \sum_i x_i \right) \]
\[ = \frac{1}{2016 \times 1008} \quad \text{Cov} \left( \sum_i x_i, \sum_i x_i \right) \]
\[ = \frac{1}{2016 \times 1008} \left[ \text{Cov} \left( \sum_i x_i, \sum_i x_i \right) + \right. \]
\[ \left. \text{Cov} \left( \sum_{1008} x_i, \sum_{1008} x_i \right) \right] \]
\[ = \frac{1}{2016 \times 1008} \quad \text{Var} (\sum_{1008} x_i) \]
\[ = \frac{\sum_i \text{Var} (x_i)}{2016 \times 1008} = \frac{1008}{2016 \times 1008} \quad \text{Var} x = \frac{1}{2016} \lambda \]
\[ \rho = \frac{\frac{1}{2016} \lambda}{\sqrt{\frac{1}{2016 \times 1008} \lambda}} = \sqrt{\frac{1008}{2016}} = \frac{\sqrt{2}}{2} \]