EXERCISE 1 - SOLUTION:

(a) Build the log likelihood

$$\begin{aligned} l(\theta|x_1,\ldots,x_n) &= \log(p_\theta(x_1,\ldots,x_n)) \\ &= \log(\theta^n \cdot (1-\theta)^{\sum_{i=1}^n x_i}) \\ &= n\log(\theta) + \sum_{i=1}^n x_i \cdot \log(1-\theta) \end{aligned}$$

Take the derivatives w.r.t. θ:

$$\begin{aligned} \frac{d}{d\theta}\, l(\theta|x_1,\ldots,x_n) &= \frac{n}{\theta} + \sum_{i=1}^n x_i \cdot \frac{1}{1-\theta} \cdot (-1) \\ \frac{d^2}{d\theta^2}\, l(\theta|x_1,\ldots,x_n) &= \frac{-n}{\theta^2} - \sum_{i=1}^n x_i \cdot \frac{1}{(1-\theta)^2} \end{aligned}$$

Set the first derivative equal to zero:

$$\frac{n}{\theta} + \sum_{i=1}^n x_i \cdot \frac{1}{1-\theta} \cdot (-1) = 0 \Leftrightarrow \theta = \frac{n}{n + \sum_{i=1}^n x_i}$$

This critical point is a maximum, as $\frac{d^2}{d\theta^2}\, l(\theta|y_1,\ldots,y_n) < 0$ for all θ.
So $\hat{\theta}_{ML} = \frac{n}{n+\sum_{i=1}^n X_i}$.

(b) Set $\bar{X} = E[X_1] = \frac{1-\theta}{\theta}$ and solve for θ. This yields: $\hat{\theta}_{MOM} = \frac{1}{\bar{X}+1}$.

Comparison with MLE:

$$\hat{\theta}_{MOM} = \frac{1}{\bar{X}+1} = \frac{n}{n} \cdot \frac{1}{\bar{X}+1} = \frac{n}{n+\sum_{i=1}^n X_i} = \hat{\theta}_{ML}$$

(c) Expected Fisher information:

$$\begin{aligned} I(\theta) = E_\theta[(\frac{d}{d\theta}\, l(\theta|X_1))^2] &= E_\theta[(\frac{1}{\theta} - X_1 \cdot \frac{1}{1-\theta})^2] \\ &= E_\theta[\frac{1}{\theta^2} - 2 \cdot \frac{1}{\theta} \cdot X_1 \cdot \frac{1}{1-\theta} + X_1^2 \cdot \frac{1}{(1-\theta)^2}] \\ &= \frac{1}{\theta^2} - 2 \cdot \frac{1}{\theta} \cdot E_\theta[X_1] \cdot \frac{1}{1-\theta} + E_\theta[X_1^2] \cdot \frac{1}{(1-\theta)^2} \\ &= \frac{1}{\theta^2} - 2 \cdot \frac{1}{\theta} \cdot \frac{1-\theta}{\theta} \cdot \frac{1}{1-\theta} + E_\theta[X_1^2] \cdot \frac{1}{(1-\theta)^2} \\ &= \frac{1}{\theta^2} - 2 \cdot \frac{1}{\theta^2} + (\frac{1-\theta}{\theta^2} + \frac{(1-\theta)^2}{\theta^2}) \cdot \frac{1}{(1-\theta)^2} \\ &= \frac{-1}{\theta^2} + \frac{1}{(1-\theta)\theta^2} + \frac{1}{\theta^2} \\ &= \frac{1}{(1-\theta)\theta^2} \end{aligned}$$

In between, we have used: $E[X_1^2] = Var(X_1) + E[X_1]^2 = \frac{1-\theta}{\theta^2} + \frac{(1-\theta)^2}{\theta^2}$.

(d) The LRT is given by:

$$\lambda(X_1) = \frac{0.4 \cdot (1-0.4)^{X_1}}{0.2 \cdot (1-0.2)^{X_1}} = 2 \cdot (\frac{3}{4})^{X_1}$$

$\lambda(.)$ is monotone (decreasing) in X_1. Thus, reject the null hypothesis if X_1 is 'too' large. To this end, determine the smallest $n \in \mathbb{N}$ for which:

$$P_{\theta=0.4}(X_1 > n) \leq 0.05 \Leftrightarrow 1 - F_{\theta=0.4}(n) \leq 0.05$$

Under H_0, X_1 has a geometric distribution with $\theta = 0.4$.
Approach: For $x \in \mathbb{R}$:

$$1 - (1 - (1-0.4)^{x+1}) = 0.05 \Leftrightarrow (x+1)\log(0.6) = \log(0.05) \Leftrightarrow x = \log(0.05)/\log(0.6) - 1$$

So $x \approx 4.86$ and thus $n = 5$. Reject H_0 if $X_1 > 5 \Leftrightarrow \lambda(X_1) > 2 \cdot (\frac{3}{4})^5$

Power at $\theta = 0.2$: $P_{\theta=0.2}(X_1 > 5) = 1 - F_{\theta=0.2}(5) = (1-0.2)^{5+1} \approx 0.26$.

P-value for $X_1 = 9$: Under H_0, a value greater or equal to 9 has probability:
$P_{\theta=0.4}(X_1 > 8) = 1 - F_{\theta=0.4}(8) = (1-0.4)^{8+1} \approx 0.01$. So, p-value: ≈ 0.01.

EXERCISE 2 - SOLUTION:
(a) Build the log likelihood and take its derivative:

$$\begin{aligned}
l(\sigma^2|x_1,\ldots,x_n) &= \log\left(p_{\sigma^2}(x_1,\ldots,x_n)\right) \\
&= \log\left((2\pi)^{n/2}\cdot(\frac{1}{\sigma^2})^{n/2}\cdot e^{-\frac{1}{2}\cdot\sum\limits_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}}\right) \\
&= \frac{n}{2}\cdot\log(2\pi)-\frac{n}{2}\log(\sigma^2)-0.5\cdot\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2}
\end{aligned}$$

Take the derivatives w.r.t. σ^2:

$$\frac{d}{d\theta}\ l(\sigma^2|x_1,\ldots,x_n) = -\frac{n}{2}\cdot\frac{1}{\sigma^2}+0.5\cdot\sum_{i=1}^{n}(x_i-\mu)^2\cdot\frac{1}{\sigma^4}$$

Set the derivative equal to zero:

$$\begin{aligned}
& -\frac{n}{2}\cdot\frac{1}{\sigma^2}+0.5\cdot\sum_{i=1}^{n}(x_i-\mu)^2\cdot\frac{1}{\sigma^4}=0 \\
\Leftrightarrow\ & -\frac{n}{2}\cdot\sigma^2+0.5\cdot\sum_{i=1}^{n}(x_i-\mu)^2=0 \\
\Leftrightarrow\ & -\frac{n}{2}\cdot\sigma^2=-0.5\cdot\sum_{i=1}^{n}(x_i-\mu)^2 \\
\Leftrightarrow\ & \sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}$$

So: $\widehat{\sigma^2_{ML}} = \frac{1}{n}\sum\limits_{i=1}^{n}(x_i-\mu)^2$.

(b) Mean Squared Error:

$$\begin{aligned}
E(\widehat{\sigma^2_{ML}}) &= E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right] \\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}E\left[X_i^2-2X_i\cdot\mu+\mu^2\right] \\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}(E[X_i^2]-2E[X_i]\cdot\mu+\mu^2) \\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}(\sigma^2+\mu^2-2\mu^2+\mu^2) \\
&= \frac{1}{n}\cdot\sum_{i=1}^{n}\sigma^2=\sigma^2
\end{aligned}$$

In between, we have used $E[X_i]=\mu$ and $E[X_i^2]=Var(X_i)+E[X_i]^2=\sigma^2+\mu^2$. Thus, the ML estimator is unbiased. Therefore:

$$\begin{aligned}
MSE(\widehat{\sigma^2_{ML}}) &= Var\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right) \\
&= \frac{1}{n^2}\cdot Var\left(\sum_{i=1}^{n}(X_i-\mu)^2\right) \\
&= \frac{1}{n^2}\cdot Var\left(\frac{\sigma^2}{\sigma^2}\cdot\sum_{i=1}^{n}(X_i-\mu)^2\right) \\
&= \frac{\sigma^4}{n^2}\cdot Var\left(\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\right) \\
&= \frac{\sigma^4}{n^2}\cdot 2n \\
&= \frac{2\sigma^4}{n}
\end{aligned}$$

The ML estimator is asymptotically consistent, as it is (asymptotically) unbiased and its variance converges to 0 for $n \to \infty$.

(c) Confidence interval: According to the hint:

$$P(q_{0.05} < \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \leq q_{0.95}) = 0.9$$

where $q_{0.05}$ and $q_{0.95}$ are the 0.05 and 0.95 quantiles of the χ^2 distribution with n degrees of freedom.

$$\begin{aligned}
& P(q_{0.05} < \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \leq q_{0.95}) = 0.9 \\
\Leftrightarrow\ & P(\frac{1}{q_{0.05}} > \frac{\sigma^2}{\sum\limits_{i=1}^{n}(X_i-\mu)^2} \geq \frac{1}{q_{0.95}}) = 0.9 \\
\Leftrightarrow\ & P(\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{q_{0.05}} > \sigma^2 \geq \frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{q_{0.95}}) = 0.9
\end{aligned}$$

Thus, $\left[\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_{0.95}}, \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{q_{0.05}}\right]$ is a 90% confidence interval for σ^2.

EXERCISE 3 - SOLUTION:
Let X and x denote the random sample and the observed data.
(a) Build the log likelihood:

$$\begin{aligned} l(\lambda|X) &= \log\left(p_\lambda(X)\right) \\ &= \log\left(\prod_{i=1}^{n} \lambda \cdot x_i \cdot e^{-\lambda \cdot x_i \cdot y_i}\right) \\ &= n \cdot \log(\lambda) + \sum_{i=1}^{n} \log(x_i) - \lambda \sum_{i=1}^{n} x_i y_i \end{aligned}$$

Set the derivative equal to zero:

$$\frac{d}{d\lambda}\, l(\lambda|X) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i y_i = 0$$

Critical point: $\lambda = \frac{n}{\sum\limits_{i=1}^{n} x_i y_i}$

Second derivative:

$$\frac{d^2}{d\lambda^2}\, l(\lambda|X) = \frac{-n}{\lambda^2} < 0 \quad \text{for all } \lambda$$

So $\hat{\lambda}_{ML} = \frac{n}{\sum\limits_{i=1}^{n} x_i y_i}$.

(b) Consider the joint density:

$$\begin{aligned} p_\lambda(X) &= \prod_{i=1}^{n} \lambda \cdot x_i \cdot e^{-\lambda \cdot x_i \cdot y_i} \\ &= \lambda^n \cdot \prod_{i=1}^{n} x_i \cdot e^{-\lambda \cdot \sum_{i=1}^{n} x_i \cdot y_i} \\ &= \lambda^n \cdot \prod_{i=1}^{n} x_i \cdot e^{-\lambda \cdot \frac{n}{\hat{\lambda}_{ML}}} \\ &= \prod_{i=1}^{n} x_i \cdot \left(\lambda^n \cdot e^{\frac{-\lambda \cdot n}{\hat{\lambda}_{ML}}}\right) \\ &= h(x) \cdot g(\lambda, \hat{\lambda}_{ML}) \end{aligned}$$

where $h(x) = \prod_{i=1}^{n} x_i$ and $g(\lambda, \hat{\lambda}_{ML}) = \lambda^n \cdot e^{-\lambda \cdot \sum_{i=1}^{n} \frac{n}{\hat{\lambda}_{ML}}}$.

According to the factorization theorem, $\hat{\lambda}_{ML}$ is a sufficient statistic.

(c) Asymptotic distribution of MLE:

$\sqrt{n} \cdot (\hat{\lambda}_{ML} - \lambda) \sim \mathcal{N}(0, I(\lambda)^{-1}) \Leftrightarrow \hat{\lambda}_{ML} \sim \mathcal{N}(\lambda, \frac{1}{n \cdot I(\lambda)})$.

Compute the expected Fisher information. Using the second derivative and $n = 1$:

$$I(\lambda) = (-1) \cdot E_\lambda[\frac{d^2}{d\lambda^2}\, l(\lambda|X_1)] = (-1) \cdot E_\lambda[\frac{-1}{\lambda^2}] = \frac{1}{\lambda^2}$$

Asymptotically, we thus have:

$$\sqrt{n} \cdot (\hat{\lambda}_{ML} - \lambda) \sim \mathcal{N}(0, \lambda^2) \Leftrightarrow \hat{\lambda}_{ML} \sim \mathcal{N}(\lambda, \frac{\lambda^2}{n})$$

It follows from the asymptotic distribution: $\sqrt{I(\lambda)} \cdot \sqrt{n} \cdot (\hat{\lambda}_{ML} - \lambda) \sim \mathcal{N}(0, 1)$. Thus

$$\begin{aligned}
& P(q_{0.025} < \sqrt{I(\lambda)} \cdot \sqrt{n} \cdot (\hat{\lambda}_{ML} - \lambda) \leq q_{0.975}) = 0.95 \\
& P(-2 < \sqrt{I(\lambda)} \cdot \sqrt{n} \cdot (\hat{\lambda}_{ML} - \lambda) \leq 2) = 0.95 \\
\Leftrightarrow \quad & P(\frac{-2}{\sqrt{I(\lambda)} \cdot \sqrt{n}} < \hat{\lambda}_{ML} - \lambda \leq \frac{2}{\sqrt{I(\lambda)} \cdot \sqrt{n}}) = 0.95 \\
\Leftrightarrow \quad & P(\frac{2}{\sqrt{I(\lambda)} \cdot \sqrt{n}} > \lambda - \hat{\lambda}_{ML} \geq \frac{-2}{\sqrt{I(\lambda)} \cdot \sqrt{n}}) = 0.95 \\
\Leftrightarrow \quad & P(\hat{\lambda}_{ML} + \frac{2}{\sqrt{I(\lambda)} \cdot \sqrt{n}} > \lambda \geq \hat{\lambda}_{ML} + \frac{-2}{\sqrt{I(\lambda)} \cdot \sqrt{n}}) = 0.95
\end{aligned}$$

So the 95% confidence interval for λ is: $\left[\hat{\lambda}_{ML} - \frac{2}{\sqrt{I(\lambda)\cdot\sqrt{n}}}, \hat{\lambda}_{ML} + \frac{2}{\sqrt{I(\lambda)\cdot\sqrt{n}}}\right]$

Note: $I(\theta)$ has to be replaced by $I(\hat{\theta}_{ML})$, what is asymptotically correct.

This yields:

$$\left[\hat{\lambda}_{ML} - \frac{2}{\sqrt{I(\hat{\lambda}_{ML})} \cdot \sqrt{n}}, \hat{\lambda}_{ML} + \frac{2}{\sqrt{I(\hat{\lambda}_{ML})} \cdot \sqrt{n}}\right]$$

where $\sqrt{I(\hat{\lambda}_{ML})} = \frac{1}{\hat{\lambda}_{ML}}$.

Plugging in the MLE from (a) yields:

$$\left[\frac{n}{\sum\limits_{i=1}^{n} x_i y_i} - \frac{2\sqrt{n}}{\sum\limits_{i=1}^{n} x_i y_i} \ , \ \frac{n}{\sum\limits_{i=1}^{n} x_i y_i} + \frac{2\sqrt{n}}{\sum\limits_{i=1}^{n} x_i y_i}\right]$$

EXERCISE 4 - SOLUTION:

(a) Build the log likelihood:

$$\begin{aligned} l(\theta|x) &= \log\left(p_\theta(x)\right) \\ &= \log\left(\prod_{i=1}^{n} \theta^{x_i} \cdot \frac{1}{x_i!} \cdot e^{-\theta}\right) \\ &= \sum_{i=1}^{n} x_i \cdot \log(\theta) - \sum_{i=1}^{n} \log(x_i!) - n\theta \end{aligned}$$

Build the derivative and set it equal to zero:

$$\frac{d}{d\theta}\, l(\theta|x) = \frac{\sum_{i=1}^{n} x_i}{\theta} - n = 0$$

Critical point: $\theta = \bar{x}$ Second derivative:

$$\frac{d^2}{d\theta^2}\, l(\theta|x) = -\frac{\sum_{i=1}^{n} x_i}{\theta^2} < 0 \quad \forall\theta$$

So $\hat{\theta}_{ML} = \bar{X}$

(b) Cramer-Rao bound. Use the 2nd derivative:

$$E_\theta[\frac{d^2}{d\theta^2}\, l(\theta|x)] = E_\theta[-\frac{\sum\limits_{i=1}^{n} X_i}{\theta^2}] = -\frac{1}{\theta^2}\sum_{i=1}^{n} E_\theta[X_i] = -\frac{n}{\theta}$$

Cramer-Rao bound is equal to $\left(-E_\theta[\frac{d^2}{d\theta^2}\, l(\theta|x)]\right)^{-1} = \frac{\theta}{n}$.

(c) The MLE is unbiased, as:

$$E_\theta[\bar{X}] = \frac{1}{n}\sum_{i=1}^{n} E_\theta[X_i] = \theta$$

The variance of the MLE is:

$$Var_\theta(\bar{X}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \theta = \frac{\theta}{n}$$

Hence, the MLE attains the Carmer-Rao bound.

(d) Build LRT and use Neyman Pearson to test $H_0 : \theta = 7$ vs. $H_1 : \theta = \theta_1$, where $\theta_1 > 7$.

$$\lambda(X) = \left(\frac{7}{\theta_1}\right)^{\sum_{i=1}^{4} X_i} \cdot e^{-7+\theta_1}$$

$\lambda(X)$ is monotone (increasing) in $\sum\limits_{i=1}^{4} X_i$, as $\theta_1 < 7$. Reject the null hypothesis if $\sum\limits_{i=1}^{4} X_i$ is 'too small'. For the upper bound ($\theta_0 = 7$) of Θ_0, $\sum\limits_{i=1}^{4} X_i$ has a Poisson distribution with parameter $n \cdot \theta_0 = 28$. Table 1 shows that $x = 20$ is the smallest integer with:

$$P_{\theta_0}(\sum_{i=1}^{4} X_i \leq 20) \geq 0.05$$

Therefore

$$P_{\theta_0}(\sum_{i=1}^{4} X_i \leq 19) < 0.05$$

and consequently values lower or equal to 19 have a probility less than $\alpha = 0.05$ under H_0. The rejection region of the UMP is thus: $RR = \{0, 1, 2, \dots, 19\}$. For the observed sample with $\sum_{i=1}^{n} x_i = 16$, the UMP test rejects the null hypothesis.

EXERCISE 5 - SOLUTION:

(a) Compute:

$$E[\frac{d}{d\theta} l_{X_1}(\theta)] = \int (\frac{d}{d\theta} \log f_\theta(x)) f_\theta(x)\ dx = \int \frac{\frac{d}{d\theta} f_\theta(x)}{f_\theta(x)} f_\theta(x)\ dx = \frac{d}{d\theta} \int f_\theta(x)\ dx = \frac{d}{d\theta}\ 1 = 0$$

$$Var[\frac{d}{d\theta} l_{X_1}(\theta)] = E[(\frac{d}{d\theta} l_{X_1}(\theta))^2] - E[\frac{d}{d\theta} l_{X_1}(\theta)]^2 = E[(\frac{d}{d\theta} l_{X_1}(\theta))^2] - 0^2 = I(\theta) \quad \text{by definition}$$

(b) Central Limit Theorem (CLS): For $Y_1, \dots, Y_n$ i.i.d. sample, asymptotically:

$$\sqrt{n} \cdot \frac{\bar{Y} - E[Y_1]}{\sqrt{Var(Y_1)}} \sim \mathcal{N}(0, 1)$$

Set: $Y_i = \frac{d}{d\theta} l_{X_1}(\theta)$, then it follows from (a): $E[Y_1] = 0$ and $Var(Y_1) = I(\theta)$. Moreover:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i = \frac{1}{n} \sum_{i=1}^{n} \frac{d}{d\theta} l_{X_i}(\theta) = \frac{1}{n} \cdot \frac{d}{d\theta} \sum_{i=1}^{n} l_{X_i}(\theta) = \frac{1}{n} \cdot \frac{d}{d\theta} l_X(\theta)$$

Therefore according to the CLS, asymptotically:

$$\sqrt{n} \cdot \frac{\bar{Y} - E[Y_1]}{\sqrt{Var(Y_1)}} = \sqrt{n} \cdot \frac{\frac{1}{n} \cdot \frac{d}{d\theta} l_X(\theta) - 0}{\sqrt{I(\theta)}} = \frac{\frac{d}{d\theta} l_X(\theta)}{\sqrt{n} \cdot \sqrt{I(\theta)}} \sim \mathcal{N}(0, 1)$$