An i.i.d. sample $X_1, \ldots, X_n$ is taken from a discrete distribution F with probability function $f_\theta(x) = (\theta - 1)^{x-1}\theta^{-x}$ for $x = 1, 2, 3, \ldots$ where $\theta \in \Omega = (1, \infty)$ is an unknown parameter. For this distribution, $\mathsf{E}(X) = \theta$ and $\mathsf{Var}(X) = \theta(\theta - 1)$.

1. (6 points) For sample $s = (x_1, \ldots, x_n)$, show that $\bar{x}$ is a minimal sufficient statistic.

   Solution: The likelihood can be written as a function of $\bar{x}$ as follows.

$$\begin{aligned} L(\theta \mid s) &= \prod_{i=1}^{n} (\theta - 1)^{x_i - 1}\theta^{-x_i} \\ &= (\theta - 1)^{n(\bar{x}-1)}\theta^{-n\bar{x}} \end{aligned}$$

   Thus, by the factorization theorem, $\bar{x}$ is a sufficient statistic. To show that $\bar{x}$ is a minimal sufficient statistic, we consider the likelihood ratio for two separate samples, $s_1 = (x_1, \ldots, x_n)$ and $s_2 = (y_1, \ldots, y_n)$.

$$\begin{aligned} \frac{L(\theta \mid s_1)}{L(\theta \mid s_2)} &= \frac{(\theta - 1)^{n(\bar{x}-1)}\theta^{-n\bar{x}}}{(\theta - 1)^{n(\bar{y}-1)}\theta^{-n\bar{y}}} \\ &= \left(\frac{\theta - 1}{\theta}\right)^{n(\bar{x}-\bar{y})} \end{aligned}$$

   If this ratio does not depend on θ, then the exponent must be zero, so $n(\bar{x} - \bar{y}) = 0$ which implies that $\bar{x} = \bar{y}$. Thus, $\bar{x}$ is a minimal sufficient statistic.

   Note that putting the likelihood ratio in the form $g(\theta)^a$ is necessary. Most everyone expressed the likelihood ratio in the form $g(\theta)^a h(\theta)^b$ and concluded that $a = b = 0$, but this is not necessarily true. For example, if $g(\theta) = \theta^2$, $h(\theta) = \theta$, $a = 1$, and $b = -2$, then $g(\theta)^a h(\theta)^b = 1$ does not depend on θ, but $a \neq 0$ and $b \neq 0$.

2. (6 points) Find the maximum likelihood estimate $\hat{\theta}$ of θ in terms of $\bar{x}$.

   Solution: We begin the solution using the standard method of taking derivatives of the log-likelihood, setting this expression equal to zero, and solving for θ, but note that there is a subtlety when the sample mean is one. The sample are all positive integers, but there is a positive probability that all x_i are equal to one, and this case deserves special attention.

   If $\bar{x} > 1$, the following holds.

$$\begin{aligned} L(\theta \mid s) &= (\theta - 1)^{n(\bar{x}-1)}\theta^{-n\bar{x}} \\ \log L(\theta \mid s) = \ell(\theta \mid s) &= n(\bar{x} - 1)\log(\theta - 1) - n\bar{x}\log\theta \\ \ell'(\theta \mid s) &= \frac{n(\bar{x} - 1)}{\theta - 1} - \frac{n\bar{x}}{\theta} = 0 \\ \frac{n(\bar{x} - 1)}{\theta - 1} &= \frac{n\bar{x}}{\theta} \\ n(\bar{x} - 1)\theta &= n\bar{x}(\theta - 1) \\ n\bar{x}\theta - n\theta &= n\bar{x}\theta - n\bar{x} \\ \hat{\theta} &= \bar{x} \end{aligned}$$

   Thus, $\hat{\theta} = \bar{x}$ is the only possible maximum. We verify that $\hat{\theta}$ is a maximum by checking the sign of the second derivative when $\theta = \bar{x}$.

$$\begin{aligned} \ell''(\theta \mid s)\Big|_{\theta=\bar{x}} &= \frac{-n(\bar{x}-1)}{(\theta-1)^2} + \frac{n\bar{x}}{\theta^2}\Big|_{\theta=\bar{x}} \\ &= \frac{-n}{\bar{x}-1} + \frac{n}{\bar{x}} \\ &= \frac{-n}{\bar{x}(\bar{x}-1)} < 0 \end{aligned}$$

However, when $\bar{x} = 1$, the likelihood takes a different form.

$$L(\theta \mid s, \bar{x} = 1) = \theta^{-n}$$

This function is decreasing for all $\theta > 1$. The maximum in the interval $[1, \infty)$ occurs at the endpoint 1, which is not in the parameter space $\Omega = (1, \infty)$. Technically, this is a situation where the maximum likelihood estimate does not exist. If the parameter space is expanded so that $\Omega = [1, \infty)$, then $\hat{\theta} = \bar{x}$ is the maximum likelihood estimate, but the justification that this is the maximum is different depending on whether or not $\bar{x}$ is equal to 1.

3. (4 points) Find expressions for the bias, variance, and mean square error of $\hat{\theta}$.

Solution: By definition, $\mathsf{Bias}(\hat{\theta}) = \mathsf{E}(\hat{\theta} - \theta)$. Thus,

$$\mathsf{Bias}(\hat{\theta}) = \mathsf{E}\bar{X} - \theta = \left(\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}X_i\right) - \theta = \theta - \theta = 0$$

We are given that $\mathsf{Var}(X) = \theta(\theta - 1)$.

$$\mathsf{Var}(\hat{\theta}) = \mathsf{Var}\bar{X} = \left(\frac{1}{n^2}\sum_{i=1}^{n} \mathsf{Var}X_i\right) = \frac{\theta(\theta-1)}{n}$$

The mean square error is the square of the bias plus the variance, so

$$\mathsf{MSE}(\hat{\theta}) = \frac{\theta(\theta-1)}{n}$$

4. (2 points) Evaluate $\hat{\theta}$ for the sample $s = (4, 8, 10, 9, 1)$ and find numerical estimates of the bias, variance, and mean square error.

Solution: The sample mean is 6.4, so $\hat{\theta} = 6.4$ for this sample. Evaluating the bias, variance, and MLE by plugging in this value gives $\widehat{\mathsf{Bias}} = 0$, $\widehat{\mathsf{Var}(\hat{\theta})} = 6.912$, and $\widehat{\mathsf{MSE}(\hat{\theta})} = 6.912$.

5. (2 points) An alternative parameterization uses $\psi = 1/\theta$. Find the maximum likelihood estimate of ψ.

Solution: $\psi = 1/\theta$ is a one-to-one function of θ. Thus, $\hat{\psi} = 1/\hat{\theta} = 1/6.912 = 0.1447$.