STAT 381 ADDITIONAL EXERCISES

1 Distribution of Order Statistics

DEFINITION Suppose we have n observations $X_1, \ldots, X_n$. Denote by $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$ the ordered set. For any i, $i = 1, \ldots, n$, $X_{(i)}$ is called the *i-th order statistic*. Note that $X_{(1)}$ is the minimum, whereas $X_{(n)}$ denotes the maximum.

PROPOSITION Suppose $X_1, \ldots, X_n$ are iid random variables with a common pdf $f(x)$ and cdf $F(x)$. The pdf of the i-th order statistic has the form

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} \big[F(x)\big]^{i-1} f(x) \big[1 - F(x)\big]^{n-i}.$$

PROOF: If the i-th order statistic is "equal" to x (contributing $f(x)$), then $i-1$ observations necessarily lie below x (contributing $\big[F(x)\big]^{i-1}$) , and the other $n-i$ lie above x (contributing $\big[1 - F(x)\big]^{n-i}$). Finally, the multiplicative factor is the number of ways to choose $i-1$ observations to lie below x, and $n-i$ to exceed x.

EXAMPLE If we let $i = n$ in the above proposition, we obtain the pdf of the maximum of n iid observations,

$$f_{X_{(n)}}(x) = \frac{n!}{(n-1)!(n-n)!} \big[F(x)\big]^{n-1} f(x) \big[1 - F(x)\big]^{n-n} = n\, f(x) \big[F(x)\big]^{n-1}.$$

This is intuitive, since the pdf of $X_{(n)}$ can also be obtained by the following reasoning:

$$F_{X_{(n)}}(x) = \mathbb{P}(X_{(n)} \leqslant x) = \mathbb{P}(X_1 \leqslant x,\, X_2 \leqslant x, \ldots, X_n \leqslant x) = \big[F(x)\big]^n,$$

and, thus, the pdf is $f_{X_{(n)}}(x) = F'_{X_{(n)}}(x) = n\, f(x) \big[F(x)\big]^{n-1}$.

EXERCISE 1 Consider n iid observations with the common pdf $f(x)$ and cdf $F(x)$. Use the formula for the pdf of the i-th order statistic to show that the pdf of the minimum is $f_{X_{(1)}}(x) = n\, f(x) \big[1 - F(x)\big]^{n-1}$. Also, find the pdf by first deriving the expression for the cdf, arguing from the first principles.

EXERCISE 2 Let $X_1, \ldots, X_n$ be iid realizations of a standard uniform random variable. Find the pdf's of: (a) i-th order statistic, $i = 1, \ldots, n$, (b) minimum, and (c) maximum. Specify the name of the distribution and respective parameters.

EXERCISE 3 Let $X_1, \ldots, X_n$ be independent exponential random variables with mean $1/\beta$. Find the densities of: (a) $X_{(i)}$, $i = 1, \ldots, n$, (b) minimum (give the distribution name and specify parameters), and (c) maximum.

2 Maximum Likelihood Estimator

Definition Suppose $X_1, \ldots, X_n$ are iid random variables with a common pmf (discrete case) or pdf (continuous case) $f(x;\theta)$. The *likelihood function* is a function of the unknown parameter θ that is given by

$$L(\theta) = L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i;\theta).$$

Definition An estimator $\widehat{\theta} = \widehat{\theta}(X_1, \ldots, X_n)$ is called the *maximum likelihood estimator* (*MLE*) of θ if it maximizes the likelihood function $L(\theta)$.

Example 1 Let $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$. The likelihood function is

$$L(p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} = p^{\sum_{i=1}^{n} X_i}(1-p)^{n-\sum_{i=1}^{n} X_i}.$$

It is easier to work with the *log-likelihood function*, the natural logarithm of the likelihood function,

$$\ln L(p) = \sum_{i=1}^{n} X_i \ln p + (n - \sum_{i=1}^{n} X_i) \ln(1-p).$$

To maximize the log-likelihood function, we equate to zero the first partial derivative of $\ln L(p; X_1, \ldots, X_n)$ with respect to p, and solve for p. We obtain

$$0 = \frac{d \ln L(p)}{dp} = \frac{\sum_{i=1}^{n} X_i}{p} - \frac{n - \sum_{i=1}^{n} X_i}{1-p}.$$

Thus, $\widehat{p}$, the maximum likelihood estimator of p, satisfies the equation

$$\frac{\sum_{i=1}^{n} X_i}{\widehat{p}} = \frac{n - \sum_{i=1}^{n} X_i}{1-\widehat{p}},$$

from where $\widehat{p} = \sum_{i=1}^{n} X_i/n = \bar{X}$. The MLE $\widehat{p} = \bar{X}$ represents the proportion of successes among n observations, and is an intuitive estimator of p, the probability of success.

Exercise 4 Let $X_1, \ldots, X_n \overset{iid}{\sim} Binomial(N, p)$ where N is known. Show that the MLE of p is $\widehat{p} = \bar{X}/N$.

Exercise 5 Let $X_1, \ldots, X_n \overset{iid}{\sim} Geometric(p)$ with pmf $p(x) = p(1-p)^{x-1},\ \ x = 1, 2, \ldots$. Prove that the MLE of p is $\widehat{p} = 1/\bar{X}$.

Exercise 6 Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Check that the MLE of λ is $\widehat{\lambda} = \bar{X}$.

Exercise 7 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Show that the MLE of θ is $\widehat{\theta} = X_{(n)}$, the nth order statistic (or, simply, the maximum).

Exercise 8 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(a, b)$. Verify that the MLE of a is $\widehat{a} = X_{(1)}$, the first order statistic (i.e., the minimum), and that the MLE of b is $\widehat{b} = X_{(n)}$, the nth order statistic (i.e., the maximum).

Exercise 9 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Exponential* with mean β. Show that the MLE of β is $\widehat{\beta} = \bar{X}$.

Exercise 10 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Exponential* with mean $1/\beta$. Show that the MLE of β is $\widehat{\beta} = 1/\bar{X}$.

Exercise 11 Let $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$. Prove that the MLE of μ is $\widehat{\mu} = \bar{X}$, and the MLE of σ^2 is $\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$.

Exercise 12 Let $X_1, \ldots, X_n \overset{iid}{\sim} Weibull(\alpha)$ where the pdf is defined as

$$f(x; \alpha) = \alpha\, x^{\alpha - 1} \exp\{-x^{\alpha}\}, \quad x > 0, \ \alpha > 0.$$

Show that $\widehat{\alpha}$, the MLE of α, is the solution of the equation

$$\frac{n}{\widehat{\alpha}} + \sum_{i=1}^{n} \ln X_i - \sum_{i=1}^{n} X_i^{\widehat{\alpha}} \ln X_i = 0.$$

This equation has no closed-form solution and has to be solved numerically. Check that if $X_1 = 0.4, X_2 = 0.3$, and $X_3 = 0.6$, the MLE is $\widehat{\alpha} = 1.0067$.

Exercise 13 Let $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$, $0 \leqslant p \leqslant 1/5$. Verify that the MLE of p is $\widehat{p} = \bar{X}$, if $0 \leqslant \bar{X} \leqslant 1/5$, and $1/5$, if $\bar{X} > 1/5$.

Exercise 14 Let $X_1, \ldots, X_n \overset{iid}{\sim} f(x; \theta) = \frac{1}{\beta} e^{-x/\beta}$, $x > 0$, $\beta > 4$. Prove that the MLE of β is $\bar{X}$, if $\bar{X} \geqslant 4$, and 4, if $0 \leqslant \bar{X} < 4$.

Exercise 15 Let $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, 1)$ where $\mu \geqslant 0$. Show that the MLE of μ is $\widehat{\mu} = \bar{X}$ if $\bar{X} \geqslant 0$, and 0, if $\bar{X} < 0$.

Exercise 16 Let $X_1, \ldots, X_n \overset{iid}{\sim} f(x; \theta)$ where the pmf $f(x; \theta)$ is given by the table:

	x		
θ	1	2	4
0	1/4	1/2	1/4
1/3	1/2	0	1/2
1/4	3/5	1/5	1/5

Check that if the observations are $X_1 = 1$, $X_2 = 4$, and $X_3 = 2$, then the MLE of θ is equal to 0.

Theorem 1 (Functional Invariance of MLE) Suppose $X_1, \ldots, X_n \stackrel{iid}{\sim}$ pmf or pdf $f(x;\theta)$. Let g be some continuous function, and let $\delta = g(\theta)$. Denote by $\widehat{\theta}$ the MLE of θ. Then the MLE of δ can be computed as $\widehat{\delta} = g(\widehat{\theta})$.

Exercise 17 Let $X_1, \ldots, X_n \stackrel{iid}{\sim}$ *Bernoulli*(p). Show that the MLE of $\mathbb{V}ar(X_1) = p(1-p)$ is $\bar{X}(1-\bar{X})$.

Exercise 18 Let $X_1, \ldots, X_n \stackrel{iid}{\sim}$ *Geometric*(p). Verify that the MLE of $\mathbb{E}(X_1) = 1/p$ is $\bar{X}$.

Exercise 19 Let $X_1, \ldots, X_n \stackrel{iid}{\sim} X \sim$ *Poisson*(λ). Prove that the MLE of $\mathbb{P}(X_1 = 1) = \lambda \exp\{-\lambda\}$ is $\bar{X}\exp\{-\bar{X}\}$.

Exercise 20 Let $X_1, \ldots, X_n \stackrel{iid}{\sim} Uniform(0, \theta)$. Check that the MLE of $\mathbb{V}ar(X_1) = \theta^2/12$ is $X_{(n)}^2/12$.

Exercise 21 Let $X_1, \ldots, X_n \stackrel{iid}{\sim} p(x, \theta)$ where $p(0, \theta) = \exp\{-\theta\}$ and $p(1, \theta) = 1 - \exp\{-\theta\}$. Prove that the MLE of θ is $\widehat{\theta} = -\ln(1-\bar{X})$.

Exercise 22 Let $X_1, \ldots, X_n \stackrel{iid}{\sim} Normal(\mu, \sigma^2)$. Prove that the MLE of σ is $\widehat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2}$.

3 Method of Moments Estimator

Definition Suppose $X_1, \ldots, X_n$ are iid random variables with a common distribution that depends on k parameters $\theta_1, \ldots, \theta_k$. The *method of moments* (*MM*) estimators of the parameters solve the system of k equations

$$\begin{cases} \mathbb{E}(X_1) = \dfrac{\sum_{i=1}^n X_i}{n} = \bar{X}, \\ \mathbb{E}(X_1^2) = \dfrac{\sum_{i=1}^n X_i^2}{n}, \\ \mathbb{E}(X_1^3) = \dfrac{\sum_{i=1}^n X_i^3}{n}, \\ \quad \ldots \\ \mathbb{E}(X_1^k) = \dfrac{\sum_{i=1}^n X_i^k}{n}. \end{cases}$$

That is, in each equation the theoretical moment is equated to the corresponding empirical moment.

Example 2 Let $X_1, \ldots, X_n \stackrel{iid}{\sim} Normal(\mu, \sigma^2)$. To find the MM estimators of μ and σ^2, we equate the first and second theoretical and empirical moments, respectively:

$$\begin{cases} \mathbb{E}(X_1) = \mu = \dfrac{\sum_{i=1}^n X_i}{n} = \bar{X}, \\ \mathbb{E}(X_1^2) = \sigma^2 + \mu^2 = \dfrac{\sum_{i=1}^n X_i^2}{n}. \end{cases}$$

The solution of this system is $\widehat{\mu} = \bar{X}$, and $\widehat{\sigma}^2 = \frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$. Note that the MM estimators of μ and σ^2 coincide with the corresponding MLEs.

EXERCISE 23 Let $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$. Show that the MM estimator for p is $\widehat{p} = \bar{X}$, the same as the MLE.

EXERCISE 24 Let $X_1, \ldots, X_n \overset{iid}{\sim} Binomial(N, p)$ where N is known. Verify that the MM estimator for p is $\bar{X}/N$, the same as the MLE.

EXERCISE 25 Let $X_1, \ldots, X_n \overset{iid}{\sim} Geometric(p)$. Show that the MM estimator for p is $\widehat{p} = 1/\bar{X}$ and coincides with the MLE.

EXERCISE 26 Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Prove that the MM estimator for λ is $\widehat{\lambda} = \bar{X}$, the same as the MLE.

EXERCISE 27 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Prove that the MM estimator for θ is $\widehat{\theta} = 2\bar{X}$. This estimator is different from the MLE. Check by giving a numeric example that the MM estimator may be smaller than the MLE, and thus, the MM estimator doesn't always make sense.

EXERCISE 28 Let $X \sim Uniform(a, b)$. Show that the MM estimators for a and b have the form

$$\widehat{a} = \bar{X} - \sqrt{3\Big(\frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2\Big)}, \quad \text{and} \quad \widehat{b} = \bar{X} + \sqrt{3\Big(\frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2\Big)}.$$

These estimators are different from the MLE's and don't always make sense.

EXERCISE 29 Let $X \sim$ *Exponential* with mean β. Prove that the MM estimator for β is $\bar{X}$, the same as the MLE.

EXERCISE 30 Let $X \sim$ *Exponential* with mean $1/\beta$. Prove that the MM estimator for β is $1/\bar{X}$, the same as the MLE.

4 Unbiased Estimator

Definition Let $X_1, \ldots, X_n \overset{iid}{\sim}$ pmf or pdf $f(x;\theta)$. Denote by $\widehat{\theta} = \widehat{\theta}(X_1, \ldots, X_n)$ an estimator of θ. The estimator $\widehat{\theta}$ is called *unbiased* if $\mathbb{E}(\widehat{\theta}) = \theta$. An estimator that is not unbiased is called *biased*.

Example 3 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Bernoulli*(p), and consider $\widehat{p} = \bar{X}$, the MLE and MM estimator of p. This estimator is unbiased because $\mathbb{E}(\widehat{p}) = \mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = p$. In fact, for any distribution, an estimator $\bar{X}$ is an unbiased estimator of the mean since $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1)$.

Exercise 31 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Binomial*(N, p) where N is known. Verify that the MLE and MM estimator $\widehat{p} = \bar{X}/N$ is an unbiased estimator of p.

Exercise 32 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Geometric*(p). Show that the MLE and MM estimator $\widehat{p} = 1/\bar{X}$ is a biased estimator of p. Show also that among all estimators of p that are based on X_1 alone, the only unbiased estimator is

$$\widehat{p}(X_1) = \begin{cases} 1, & \text{if } X_1 = 1, \\ 0, & \text{if } X_1 = 2, 3, \ldots. \end{cases}$$

Exercise 33 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Geometric*(p). Show that $\bar{X}$ is an unbiased estimator of the mean $\mathbb{E}(X_1) = 1/p$.

Exercise 34 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Poisson*(λ). Check that the MLE and MM estimator $\widehat{\lambda} = \bar{X}$ is an unbiased estimator of λ.

Exercise 35 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Prove that $X_{(n)}$, the MLE of θ, is biased, whereas $2\bar{X}$, the MM estimator, is unbiased. Show that $\dfrac{n+1}{n} X_{(n)}$ is an unbiased estimator of θ.

Exercise 36 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(a, b)$. Show that $X_{(1)}$, the MLE of a, is biased, and so is $X_{(n)}$, the MLE of b. Derive that $\dfrac{1}{n-1}\big(n\, X_{(1)} - X_{(n)}\big)$ is an unbiased estimator of a, and $\dfrac{1}{n-1}\big(n\, X_{(n)} - X_{(1)}\big)$ is an unbiased estimator of b.

Exercise 37 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Exponential* with mean β. Verify that $\bar{X}$, the MLE and MM estimator of β, is an unbiased estimator of β.

Exercise 38 Let $X_1, \dots, X_n \overset{iid}{\sim}$ *Exponential* with mean $1/\beta$. Verify that $1/\bar{X}$, the MLE and MM estimator of β, is a biased estimator of β. Show also that $\dfrac{n-1}{n\,\bar{X}}$ is an unbiased estimator of β. Hint: Use the fact that $\sum\limits_{i=1}^{n} X_i \sim Gamma(n, \beta)$.

Exercise 39 Let $X_1, \dots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$. Verify that $\bar{X}$, the MLE and MM estimator of μ, is unbiased, whereas $\widehat{\sigma}^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2$, the MLE and MM estimator of σ^2, is biased. Prove that

$$\frac{n}{n-1}\,\widehat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

is an unbiased estimator of σ^2.

5 Consistent Estimator

Definition Let $X_1, \dots, X_n$ be independent with a common density $f(x;\theta)$. An estimator $\widehat{\theta}_n = \widehat{\theta}_n(X_1, \dots, X_n)$ is called a *consistent* estimator of θ, if for any $\varepsilon > 0$, $\mathbb{P}\big(|\widehat{\theta}_n - \theta| \geqslant \varepsilon\big) \to 0$, as $n \to \infty$.

Proposition The Chebyshev inequality states that for any $\varepsilon > 0$,

$$\mathbb{P}\big(|\widehat{\theta}_n - \theta| \geqslant \varepsilon\big) \leqslant \frac{\mathbb{E}\big[(\widehat{\theta}_n - \theta)^2\big]}{\varepsilon^2}.$$

From here, if $\mathbb{E}\big[(\widehat{\theta}_n - \theta)^2\big] \to 0$ as $n \to \infty$, then $\widehat{\theta}_n$ is a consistent estimator of θ.

Definition The quantity $\mathbb{E}\big[(\widehat{\theta}_n - \theta)^2\big]$ is called the *mean square error* and is denoted by MSE. The mean square error can be expressed as the sum of two terms:

$$\begin{aligned} MSE = \mathbb{E}\big[(\widehat{\theta}_n - \theta)^2\big] &= \mathbb{E}\big[(\widehat{\theta}_n - \mathbb{E}(\widehat{\theta}_n) + \mathbb{E}(\widehat{\theta}_n) - \theta)^2\big] \\ &= \mathbb{E}\big[(\widehat{\theta}_n - \mathbb{E}(\widehat{\theta}_n))^2\big] + \overbrace{2\,\mathbb{E}\big[\widehat{\theta}_n - \mathbb{E}(\widehat{\theta}_n)\big]}^{0}\big[\mathbb{E}(\widehat{\theta}_n) - \theta\big] + \big[\mathbb{E}(\widehat{\theta}_n) - \theta\big]^2 \\ &= \mathbb{V}ar\big(\widehat{\theta}_n\big) + \big[\mathbb{E}(\widehat{\theta}_n) - \theta\big]^2. \end{aligned}$$

The quantity $\big[\mathbb{E}(\widehat{\theta}_n) - \theta\big]$ represents the bias of an estimator $\widehat{\theta}_n$. Thus, the formula for the MSE has the form:

$$MSE = \mathbb{V}ar\big(\widehat{\theta}_n\big) + \big[bias(\widehat{\theta}_n, \theta)\big]^2.$$

If $\widehat{\theta}_n$ is an unbiased estimator of θ, then $MSE = \mathbb{V}ar(\widehat{\theta}_n)$, and if this variance tends to zero as n increases, then the estimator of consistent.

A biased estimator for which the bias goes to zero as n goes to infinity, is called *asymptotically unbiased*. Thus, an estimator may be biased, but it is consistent if it is asymptotically unbiased and its variance decreases as the sample size increases.

EXERCISE 40 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Verify that
(a) The unbiased estimator $\dfrac{n+1}{n} X_{(n)}$ is a consistent estimator of θ. Hint: Show that $MSE = \dfrac{\theta^2}{n(n+2)}$.
(b) The MLE $X_{(n)}$ is asymptotically unbiased and is a consistent estimator of θ. Hint: Prove that $bias = -\dfrac{\theta}{n+1}$ and $MSE = \dfrac{2\theta^2}{(n+1)(n+2)}$.
(c) The estimator $\dfrac{n+2}{n+1} X_{(n)}$, which has the smallest MSE among all scalar multiples of $X_{(n)}$ (prove this!), is a consistent estimator of θ. Hint: Show that its $MSE = \dfrac{\theta^2}{(n+1)^2}$.

EXERCISE 41 Consider $X_1, \ldots, X_n$ that come from a $Uniform(0, \theta)$ distribution. Prove that
(a) The MM estimator $2\bar{X}_n$ is unbiased, consistent estimator of θ. Hint: Show that $MSE = \dfrac{\theta^2}{3n}$.
(b) The bias of the estimator $\bar{X}_n$ is independent of n, and thus this estimator is not a consistent estimator of θ.

EXERCISE 42 Let $X_1, \ldots, X_n$ be iid realizations of an exponential random variable with mean $1/\beta$. Check that
(a) The MLE $1/\bar{X}_n$ is asymptotically unbiased and consistent estimator of β. Hint: Prove that its $bias = \dfrac{\beta}{n-1}$ and $MSE = \dfrac{(n+2)\beta^2}{(n-1)(n-2)}$.
(b) The unbiased estimator $\dfrac{n-1}{n\bar{X}_n}$ is consistent. Hint: Prove first that the variance of this estimator is $\dfrac{\beta^2}{n-2}$.

EXERCISE 43 Suppose $X_1, \ldots, X_n$ is a random sample from $Normal(\mu, \sigma^2)$ distribution. Show that
(a) The unbiased estimator $\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ is a consistent estimator of σ^2. Hint: Show

first that the variance is equal to $\dfrac{2\sigma^4}{n-1}$.

(b) The MLE $\dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ is asymptotically unbiased, consistent estimator of σ^2. Hint: Show that its $MSE = \dfrac{2n-1}{n^2}\sigma^4$.

6 Sufficient Statistic, Factorization Theorem

Definition Let $X_1, \ldots, X_n \overset{iid}{\sim}$ pmf or pdf $f(x;\theta)$. A statistic $\widehat{\theta} = \widehat{\theta}(X_1, \ldots, X_n)$ is called a *sufficient* statistic for θ, if the conditional distribution of $X_1, \ldots, X_n$, given $\widehat{\theta}$, does not depend on θ.

It is more practical to find sufficient statistics not using the definition, but rather using the *factorization theorem.*

Theorem 2 (Factorization Theorem) Let $X_1, \ldots, X_n \overset{iid}{\sim}$ pmf or pdf $f(x;\theta)$. Then $\widehat{\theta} = \widehat{\theta}(X_1, \ldots, X_n)$ is a sufficient statistic for θ if and only if there exist two nonnegative functions g and h such that

$$\prod_{i=1}^{n} f(X_i;\theta) = g(X_1, \ldots, X_n)\, h(\widehat{\theta};\theta).$$

This expression is interpreted as saying that the likelihood function for the observations $X_1, \ldots, X_n$ can be written as a product of two functions, one of which depends only on the observations, and the other depends on some statistic that cannot be separated from the parameter. Both functions are multiplicative factors, thus the name *factorization theorem.*

Remark The factorization theorem can be formulated for distributions that depend on several parameters. The vector of estimators $(\widehat{\theta}_1, \ldots, \widehat{\theta}_k)$ is a sufficient statistic for the vector of parameters $(\theta_1, \ldots, \theta_k)$ if and only if there exist two nonnegative functions g and h such that

$$\prod_{i=1}^{n} f(X_i;\theta_1, \ldots, \theta_k) = g(X_1, \ldots, X_n)\, h(\widehat{\theta}_1, \ldots, \widehat{\theta}_k;\theta_1, \ldots, \theta_k).$$

Proposition Any invertible function of a sufficient statistic is itself a sufficient statistic.

Example 4 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Bernoulli*(p). The likelihood function has the form

$$\prod_{i=1}^{n} f(X_i;p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i} = p^{\sum_{i=1}^{n} X_i}\,(1-p)^{n-\sum_{i=1}^{n} X_i}.$$

Now let

$$g(X_1,\ldots,X_n)=1,\ \ \widehat{p}=\sum_{i=1}^n X_i,\ \ \text{and}\ \ h(\widehat{p};p)=p^{\widehat{p}}(1-p)^{n-\widehat{p}}.$$

We see that

$$\prod_{i=1}^n f(X_i;p)=p^{\widehat{p}}(1-p)^{n-\widehat{p}}=g(X_1,\ldots,X_n)\,h(\widehat{p};p).$$

By factorization theorem, $\sum_{i=1}^n X_i$ is a sufficient statistic for p. Since any invertible function is also sufficient, we can conclude that $\bar{X}=\sum_{i=1}^n X_i/n$ is also a sufficient statistic for p.

Exercise 44 Let $X_1,\ldots,X_n \overset{iid}{\sim} Binomial(N,p)$ where N is known. Show that $\bar{X}/N$ is a sufficient statistic for p.

Exercise 45 Let $X_1,\ldots,X_n \overset{iid}{\sim} Geometric(p)$ where N is known. Verify that $\bar{X}$ is a sufficient statistic for p.

Exercise 46 Let $X_1,\ldots,X_n \overset{iid}{\sim} Poisson(\lambda)$. Prove that $\bar{X}$ is a sufficient statistic for λ.

Exercise 47 Let $X_1,\ldots,X_n \overset{iid}{\sim} Uniform(0,\theta)$. Check that $X_{(n)}$ is a sufficient statistic for θ.

Exercise 48 Let $X_1,\ldots,X_n \overset{iid}{\sim} Uniform(a,b)$. Prove that the vector of estimators $(X_{(1)},X_{(n)})$ is a sufficient statistic for the vector of parameters (a,b).

Exercise 49 Let $X_1,\ldots,X_n \overset{iid}{\sim} Exponential(\beta)$. Check that $\bar{X}$ is a sufficient statistic for β.

Exercise 50 Let $X_1,\ldots,X_n \overset{iid}{\sim} Normal(\mu,\sigma^2)$. Verify that the vector of estimators $\left(\bar{X},\dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i-\bar{X})^2\right)$ is sufficient for the vector of parameters (μ,σ^2). Hint: Show first that $(\sum_{i=1}^n X_i,\sum_{i=1}^n X_i^2)$ is sufficient.

7 Uniform Minimum Variance Unbiased Estimator (UMVUE), Rao-Blackwell Theorem

Definition Let $X_1,\ldots,X_n \overset{iid}{\sim}$ pmf or pdf $f(x;\theta)$. An estimator $\widehat{\theta}=\widehat{\theta}(X_1,\ldots,X_n)$ of θ is called a *uniformly minimum variance unbiased estimator* (*UMVUE*), if it is unbiased and its variance is minimal, that is, if $\mathbb{E}(\widehat{\theta}) = \theta$ and $\mathbb{V}ar(\widehat{\theta}) \leqslant \mathbb{V}ar(\tilde{\theta})$ for any unbiased estimator $\tilde{\theta}$.

Rao-Blackwell Theorem If u is a sufficient statistic for θ and $\widehat{\theta}$ is an unbiased estimator of θ, then $\mathbb{E}(\widehat{\theta}|u)$ is the UMVUE for θ.

Example 5 Let $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$. We would like to find the UMVUE for p. We recall that $\bar{X}$ is a sufficient statistic and an unbiased estimator of p. Hence, it is the UMVUE.

Exercise 51 Let $X_1, \ldots, X_n \overset{iid}{\sim} Binomial(N, p)$ where N is fixed. Show that $\bar{X}/N$ is the UMVUE for p.

Exercise 52 Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Verify that $\bar{X}$ the UMVUE for λ.

Exercise 53 Let $X_1, \ldots, X_n \overset{iid}{\sim} Exponential$ with mean β. Check that $\bar{X}$ is the UMVUE for β.

Exercise 54 Let $X \sim Normal(\mu, \sigma^2)$ where σ is known. Verify that $\widehat{\mu} = \bar{X}$ is the UMVUE for μ.

Exercise 55 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(a, b)$. Show that $\dfrac{1}{n-1}\big(n\,X_{(1)} - X_{(n)}\big)$ is the UMVUE for a, and $\dfrac{1}{n-1}\big(n\,X_{(n)} - X_{(1)}\big)$ is the UMVUE for b.

Exercise 56 Let $X_1, \ldots, X_n \overset{iid}{\sim} Exponential$ with mean $1/\beta$. Show that $\dfrac{n-1}{n\,\bar{X}}$ is the UMVUE for β.

Exercise 57 Let $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$. Prove that $\bar{X}$ is the UMVUE for μ, and $\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is the UMVUE for σ^2.

Exercise 58 Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Derive that $\big(1 - 1/n\big)\bar{X} + \bar{X}^2$ is the UMVUE for the second moment $\mathbb{E}(X_1^2) = \lambda(1 + \lambda)$.

Exercise 59 Let $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$. Denote by $\widehat{\sigma}^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Verify that $\bar{X}^2 - \widehat{\sigma}^2/n$ is the UMVUE for μ^2.

Exercise 60 Let $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. Find the UMVUE for $\mathbb{P}(X_1 = 0) = e^{-\lambda}$.

Exercise 61 *rm Let* $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$. *Find the UMVUE for* p^2.

Exercise 62 Let $X_1, \ldots, X_n \overset{iid}{\sim} \frac{1}{\theta}\, e^{-x/\theta}, x > 0, \theta > 0$. Find the UMVUE for $\mathbb{P}(X_1 \leqslant 2) = 1 - e^{-2/\theta}$.

8 Likelihood Ratio Test

Definition Let $X_1, \ldots, X_n$ be iid with pdf $f(x;\theta)$. Suppose we want to test $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$. Define the *likelihood ratio test* as follows. The test statistic is the ratio of the two likelihood functions where the parameter θ assumes the values θ_0 and the MLE $\widehat{\theta}$, respectively, that is, $\Lambda = \dfrac{L(\theta_0)}{L(\widehat{\theta})} = \dfrac{\prod_{i=1}^n f(X_i;\theta_0)}{\prod_{i=1}^n f(X_i;\widehat{\theta})}$. If θ_0 is the true value of θ, then $L(\theta_0)$ is asymptotically the maximum value of $L(\theta)$ (Intuitively, if we sample the entire population, the most likely value of θ is θ_0). Thus, under H_0, Λ should be close to 1, and the decision rule for the test is to reject H_0 if $\Lambda \leqslant c$, where a constant c is such that $\alpha = \mathbb{P}(\Lambda \leqslant c \,|\, H_0$ is true) for a significance level α. The region $\{X_1, \ldots, X_n : \Lambda \leqslant c\}$ is called the *rejection region.*

As a rule, Λ is a very complicated function, and its distribution is very hard to figure out. However, an asymptotic distribution can be used.

Proposition Under H_0, for large n, $-2\ln\Lambda$ has approximately a chi-squared distribution with one degree of freedom.

Definition An *asymptotic likelihood ratio test* with a significance level α has the test statistic $\chi^2 = -2\ln\Lambda$, and rejects H_0 if $\chi^2 \geqslant \chi^2_\alpha(1)$, where $\chi^2_\alpha(1)$ is the $(1-\alpha)$−percentile of a chi-squared distribution with one degree of freedom.

Example. Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$, and suppose we are interested in testing $H_0 : p = p_0$ versus $H_1 : p \neq p_0$. The likelihood ratio is $\Lambda = \dfrac{\prod_{i=1}^n p_0^{X_i}(1-p_0)^{1-X_i}}{\prod_{i=1}^n \bar{X}^{X_i}(1-\bar{X})^{1-X_i}} = \dfrac{p_0^{n\bar{X}}(1-p_0)^{n-n\bar{X}}}{\bar{X}^{n\bar{X}}(1-\bar{X})^{n-n\bar{X}}}$. For large n, we reject H_0 if $\chi^2 = -2\ln\Lambda = -2n\bar{X}\ln\big(\frac{p_0}{\bar{X}}\big) - 2n(1-\bar{X})\ln\big(\frac{1-p_0}{1-\bar{X}}\big)$ exceeds the critical value $\chi^2_\alpha(1)$.

Exercise 63 Let $X_1, \ldots, X_n \overset{iid}{\sim} Binomial(N, p)$ with a known N. Suppose we are testing $H_0 : p = p_0$ against $H_1 : p \neq p_0$. Find the expression for the asymptotic likelihood ratio test statistic. State the decision rule.

Exercise 64 Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Geometric(p)$. Compute the likelihood ratio test statistic for testing $H_0 : p = p_0$ against $H_1 : p \neq p_0$. Assume n is large. Specify the decision rule.

Exercise 65 Assume $X_1, \ldots, X_n \overset{iid}{\sim} Poisson(\lambda)$. We are conducting the likelihood ratio test with $H_0 : \lambda = \lambda_0$ and $H_1 : \lambda \neq \lambda_0$. Find the test statistic for n large. Find the rejection region.

Exercise 66 Consider $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Produce the test statistic for the asymptotic likelihood ratio test with $H_0 : \theta = \theta_0$ and $H_1 : \theta \neq \theta_0$. Specify the decision rule.

Exercise 67 Let $X_1, \ldots, X_n \overset{iid}{\sim}$ *Exponential* with mean β. Write down the asymptotic likelihood ratio test statistic for testing $H_0 : \beta = \beta_0$ versus $H_1 : \beta \neq \beta_0$. Specify the rejection region.

Exercise 68 Consider $X_1, \ldots, X_n \overset{iid}{\sim} Normal(\mu, \sigma^2)$ where σ is given. Find the expression for the asymptotic likelihood ratio test statistic $-2 \ln \Lambda$ and show that it has an exact χ^2-distribution with one degree of freedom. Assume $H_0 : \mu = \mu_0$ and $H_1 : \mu \neq \mu_0$. State the decision rule.

9 Power Function of a Test

Definition The *probability of Type II error* is $\beta = \mathbb{P}(\textit{accept } H_0 \,|\, H_1 \textit{ is true})$. Note that β is a function of θ which range is determined by H_1. Typically, β is computed for a specific value of θ in that range.

Definition A *power* of a statistical test is $\textit{power} = 1 - \beta = \mathbb{P}(\textit{reject } H_0 \,|\, H_1 \textit{ is true})$.

Example Suppose we have a single observation X from a $Binomial(5, p)$ distribution which we use to test $H_0 : p < 1/2$ against $H_1 : p \geqslant 1/2$. For a rejection region $\{X = 5\}$, the power of the test is $\textit{power} = 1 - \beta = \mathbb{P}(X = 5 \,|\, p \geqslant 1/2) = p^5,\ 1/2 \leqslant p \leqslant 1$.

Exercise 69 Take $X \sim Binomial(6, p)$. Suppose we are interested in testing $H_0 : p < 1/3$ against $H_1 : p \geqslant 1/3$. Compute the power of the test is we define the rejection region as: (a) $\{X = 6\}$, (b) $\{X = 5, 6\}$, and (c) $\{X = 4, 5, 6\}$.

Exercise 70 Let $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. Consider the asymptotic likelihood ratio test for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ with a significance level α. The test statistic for this test has been derived in Exercise 66. Present the power of this test as a function of θ.

Exercise 71 Let $X_1, \ldots, X_n$ be a random sample taken from a $Normal(\mu, \sigma^2)$ distribution with some known σ. The testing is done between $H_0 : \mu = \mu_0$ and $H_1 : \mu = \mu_1$ where $\mu_1 > \mu_0$. The rejection region of the test is defined as $\{\bar{X} > k\}$ for some constant k. Suppose that the significance level α is specified. Prove that the power of this test can be written as $\textit{power} = 1 - \Phi\Big(\Phi^{-1}(1 - \alpha) - \dfrac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\Big)$ where Φ denotes the cdf of the standard normal distribution.

SOLUTIONS TO EXERCISES

Exercise 1 In the formula for the pdf of the i-th order statistic we let $i = 1$ to obtain that $f_{X_{(1)}}(x) = \dfrac{n!}{(1-1)!(n-1)!}\big[F(x)\big]^{1-1} f(x)\big[1 - F(x)\big]^{n-1} = n\, f(x)\big[1 - F(x)\big]^{n-1}$. We can also find the pdf of the minimum as follows:

$$1 - F_{X_{(1)}}(x) = \mathbb{P}(X_{(1)} \geqslant x) = \mathbb{P}(X_1 \geqslant x,\, X_2 \geqslant x, \dots, X_n \geqslant x) = \big[1 - F(x)\big]^n,$$

therefore, $F_{X_{(1)}}(x) = 1 - \big[1 - F(x)\big]^n$, and $f_{X_{(n)}}(x) = F'_{X_{(n)}}(x) = n\, f(x)\big[1 - F(x)\big]^{n-1}$.

Exercise 2 We are given that $f(x) = 1$, and $F(x) = x,\ 0 \leqslant x \leqslant 1$. Hence,
(a) the i-th order statistic has the pdf $f_{X_{(i)}}(x) = \dfrac{n!}{(i-1)!(n-i)!} x^{i-1}(1-x)^{n-i}$, that is, $X_{(i)} \sim Beta(i, n-i+1)$.
(b) If we let $i = 1$, we get the pdf of the minimum, $f_{X_{(1)}}(x) = n(1-x)^{n-1}$, that is, $X_{(1)} \sim Beta(1, n)$.
(c) Letting $i = n$, we obtain the pdf of the maximum, $f_{X_{(n)}}(x) = n\, x^{n-1}$, that is, $X_{(n)} \sim Beta(n, 1)$.

Exercise 3 The pdf of X's is $f(x) = \beta\, \exp\{-\beta\, x\}$, and the cdf is $F(x) = 1 - \exp\{-\beta\, x\},\ x > 0,\ \beta > 0$. Therefore, (a) the i-th order statistic has the pdf $f_{X_{(i)}}(x) = \dfrac{n!}{(i-1)!(n-i)!}\big[1 - \exp\{-\beta\, x\}\big]^{i-1} \beta\, \exp\{-\beta\, x\}\big[\exp\{-\beta\, x\}\big]^{n-i} = \dfrac{n!}{(i-1)!(n-i)!}\beta\, \exp\{-\beta\, x\,(n-i+1)\}\big[1 - \exp\{-\beta\, x\}\big]^{i-1}$.
(b) In particular, for $i = 1$, the pdf of the minimum is $f_{X_{(1)}}(x) = n\,\beta\, \exp\{-\beta\, n\, x\}$, that is, $X_{(1)}$ has an exponential distribution with mean $\dfrac{1}{n\beta}$.
(c) The pdf of the maximum is derived by letting $i = n$. We have $f_{X_{(n)}}(x) = n\,\beta\, \exp\{-\beta\, x\}\big[1 - \exp\{-\beta\, x\}\big]^{n-1}$. We can also notice that the cdf of the maximum is $F(x) = \big(1 - \exp\{-\beta\, x\}\big)^n$, which can be obtained by either integrating the density or arguing that all n observations must not exceed x, if the maximum doesn't exceed x.

Exercise 4 The likelihood function has the form

$$L(p) = \prod_{i=1}^{n}\binom{N}{X_i} p^{X_i}(1-p)^{N-X_i}$$

$$= \Big[\prod_{i=1}^{n}\binom{N}{X_i}\Big] p^{\sum_{i=1}^n X_i}(1-p)^{nN - \sum_{i=1}^n X_i}.$$

The log-likelihood function is

$$\ln L(p) = \ln\Big[\prod_{i=1}^{n}\binom{N}{X_i}\Big] + \sum_{i=1}^{n} X_i \,\ln p + \big(nN - \sum_{i=1}^{n} X_i\big)\,\ln(1-p).$$

The MLE $\widehat{p}$ solves the equation

$$0 = \frac{d\ln L(p)}{dp}\Big|_{p=\widehat{p}} = \frac{\sum_{i=1}^{n} X_i}{\widehat{p}} - \frac{nN - \sum_{i=1}^{n} X_i}{1-\widehat{p}}.$$

Hence, the MLE of p is

$$\widehat{p} = \frac{\sum_{i=1}^{n} X_i}{n\,N} = \frac{\bar{X}}{N}.$$

To understand the structure of this estimator, we can rewrite it as

$$\widehat{p} = \frac{\sum_{i=1}^{n}(X_i/N)}{n},$$

which is the average of proportions of successes among N trials.

Exercise 5 The likelihood function has the form

$$L(p) = \prod_{i=1}^{n} p\,(1-p)^{X_i-1} = p^n\,(1-p)^{\sum_{i=1}^{n} X_i - n}.$$

The log-likelihood function is

$$\ln L(p) = n\,\ln p + \big(\sum_{i=1}^{n} X_i - n\big)\,\ln(1-p).$$

The MLE $\widehat{p}$ solves the equation

$$0 = \frac{d\ln L(p)}{dp}\Big|_{p=\widehat{p}} = \frac{n}{\widehat{p}} - \frac{\sum_{i=1}^{n} X_i - n}{1-\widehat{p}},$$

and so,

$$\widehat{p} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}.$$

Since the mean of X_i's is equal to $1/p$, the MLE is an estimator of p derived from estimating the mean by the sample mean $\bar{X}$.

Exercise 6 The likelihood function is

$$L(\lambda) = \prod_{i=1}^{n}\frac{\lambda^{X_i}\,\exp\{-\lambda\}}{X_i!} = \Big[\prod_{i=1}^{n}\frac{1}{X_i!}\Big]\,\lambda^{\sum_{i=1}^{n} X_i}\,\exp\{-n\lambda\},$$

and the log-likelihood function takes the form

$$\ln L(\lambda) = \ln\Big[\prod_{i=1}^{n} \frac{1}{X_i!}\Big] + \sum_{i=1}^{n} X_i \, \ln\lambda - n\lambda.$$

The MLE $\widehat{\lambda}$ is the solution of the equation

$$0 = \frac{d\ln L(\lambda)}{d\lambda}\Big|_{\lambda=\widehat{\lambda}} = \frac{\sum_{i=1}^{n} X_i}{\widehat{\lambda}} - n.$$

Hence,

$$\widehat{\lambda} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

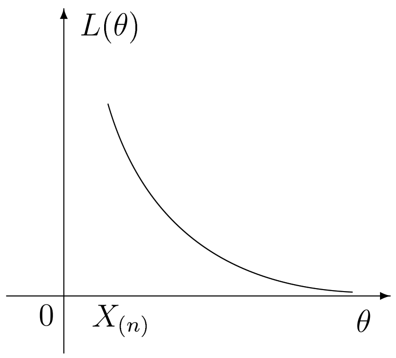
Indeed, it is intuitive to estimate the mean λ by the sample mean$\bar{X}$.

Exercise 7 The likelihood function is derived as

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta}\mathbb{I}\{0 \leqslant X_i \leqslant \theta\} = \frac{1}{\theta^n}\mathbb{I}\{0 \leqslant X_{(n)} \leqslant \theta\}.$$

Here $\mathbb{I}\{A\}$ denotes the indicator function of an event A, that is, it is equal to 1 if A occurs, and 0, otherwise. The last equality is justified by noticing that the events $\{0 \leqslant X_i \leqslant \theta\}$ occur simultaneously for all $i = 1, \dots, n$, if and only if the event $\{0 \leqslant X_{(n)} \leqslant \theta\}$ occurs.

Next, we plot the likelihood function $L(\theta) = L(\theta; X_1, \dots, X_n) = 1/\theta^n, \quad \theta \geqslant X_{(n)}$, against θ to see where it attains the maximum value.



As seen on the graph, the maximum is attained at $X_{(n)}$, thus $\widehat{\theta} = X_{(n)}$ is the MLE of θ. On intuitive level, if $X_1, \dots, X_n$ are observed, and we know that each of them doesn't exceed θ, then our best guess about the value of θ is the maximum of all the observations.

Exercise 8 The likelihood function is

$$L(a, b) = \prod_{i=1}^{n} \frac{1}{b-a}\,\mathbb{I}\{a \leqslant X_i \leqslant b\}$$

$$= \frac{1}{(b-a)^n} \mathbb{I}\{a \leqslant X_{(1)} \leqslant X_{(n)} \leqslant b\}.$$

To maximize this likelihood function, we have to minimize the denominator $(b-a)^n$, or, equivalently, minimize the distance between a and b. Since it must be true that $a \leqslant X_{(1)} \leqslant X_{(n)} \leqslant b$, the distance is minimal when a is equal to $X_{(1)}$ and b is equal to $X_{(n)}$. This leads to conclusion that the MLE of a is $\widehat{a} = X_{(1)}$ and the MLE of b is $\widehat{b} = X_{(n)}$.

Exercise 9 The likelihood function is written as

$$L(\beta) = \prod_{i=1}^{n} \frac{1}{\beta} \exp\{-X_i/\beta\} = \frac{1}{\beta^n} \exp\{-\sum_{i=1}^{n} X_i/\beta\},$$

and the log-likelihood function takes the form

$$\ln L(\beta) = -n \ln \beta - \frac{\sum_{i=1}^{n} X_i}{\beta}.$$

The maximum likelihood estimator of β satisfies the equation

$$0 = \frac{d \ln L(\beta)}{d\beta}\Big|_{\beta=\widehat{\beta}} = -\frac{n}{\widehat{\beta}} + \frac{\sum_{i=1}^{n} X_i}{\widehat{\beta}^2}.$$

From here,

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

We see that it is only reasonable to estimate the mean β by the sample mean $\bar{X}$.

Exercise 10 The likelihood function has the form

$$L(\beta) = \prod_{i=1}^{n} \beta \exp\{-\beta X_i\} = \beta^n \exp\{-\beta \sum_{i=1}^{n} X_i\},$$

and the log-likelihood function is

$$\ln L(\beta) = n \ln \beta - \beta \sum_{i=1}^{n} X_i.$$

Differentiating the log-likelihood function, we get an equation for the MLE $\widehat{\beta}$:

$$0 = \frac{d \ln L(\beta)}{d\beta}\Big|_{\beta=\widehat{\beta}} = \frac{n}{\widehat{\beta}} - \sum_{i=1}^{n} X_i.$$

Thus,

$$\widehat{\beta} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\bar{X}}.$$

Exercise 11 First, we obtain the likelihood function. We write

$$
\begin{aligned}
L(\mu, \sigma^2) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ - \frac{(X_i - \mu)^2}{2\sigma^2} \Big\} \\
&= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\Big\{ - \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{2\sigma^2} \Big\}.
\end{aligned}
$$

Next, we find the log-likelihood function as

$$
\ln L(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{2\sigma^2}.
$$

The maximum likelihood estimators $\widehat{\mu}$ and $\widehat{\sigma}^2$ are solutions of the system of two equations

$$
\begin{cases}
0 = \dfrac{\partial \ln L(\mu, \sigma^2)}{\partial \mu} \Big|_{\substack{\mu=\widehat{\mu}, \\ \sigma^2=\widehat{\sigma}^2}} = \dfrac{\sum_{i=1}^{n} (X_i - \widehat{\mu})}{\widehat{\sigma}^2}, \\
0 = \dfrac{\partial \ln L(\mu, \sigma^2)}{\partial \sigma^2} \Big|_{\substack{\mu=\widehat{\mu}, \\ \sigma^2=\widehat{\sigma}^2}} = -\dfrac{n}{2\widehat{\sigma}^2} + \dfrac{\sum_{i=1}^{n} (X_i - \widehat{\mu})^2}{2\widehat{\sigma}^4},
\end{cases}
$$

so

$$
\widehat{\mu} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}, \quad \text{and} \quad \widehat{\sigma}^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}.
$$

Since μ is the mean of the normal distribution, the estimator is indeed intuitive. The variance is estimated by the average squared distance between each observation and the sample mean, which is a natural measure of spread.

Exercise 12 We derive the likelihood function as follows:

$$
\begin{aligned}
L(\alpha) &= \prod_{i=1}^{n} \alpha X_i^{\alpha-1} \exp\{-X_i^{\alpha}\} \\
&= \alpha^n \Big(\prod_{i=1}^{n} X_i \Big)^{\alpha-1} \exp\{-\sum_{i=1}^{n} X_i^{\alpha}\}.
\end{aligned}
$$

The log-likelihood function is given by

$$
\ln L(\alpha) = n \ln \alpha + (\alpha - 1) \sum_{i=1}^{n} \ln X_i - \sum_{i=1}^{n} X_i^{\alpha}.
$$

Differentiating the log-likelihood function with respect to α and setting the derivative equal to zero, we obtain the equation that the MLE of α solves:

$$
0 = \frac{d \ln L(\alpha)}{d\alpha} \Big|_{\alpha=\widehat{\alpha}} = \frac{n}{\widehat{\alpha}} + \sum_{i=1}^{n} \ln X_i - \sum_{i=1}^{n} X_i^{\widehat{\alpha}} \ln X_i.
$$

There is no explicit solution to this equation, thus is has to be solved numerically. For the observations $X_1 = 0.4, X_2 = 0.3$, and $X_3 = 0.6$, the MLE of α solves

$$\frac{3}{\widehat{\alpha}} + \ln 0.4 + \ln 0.3 + \ln 0.6 - (0.4^{\widehat{\alpha}} \ln 0.4 + 0.3^{\widehat{\alpha}} \ln 0.3 + 0.6^{\widehat{\alpha}} \ln 0.6) = 0.$$

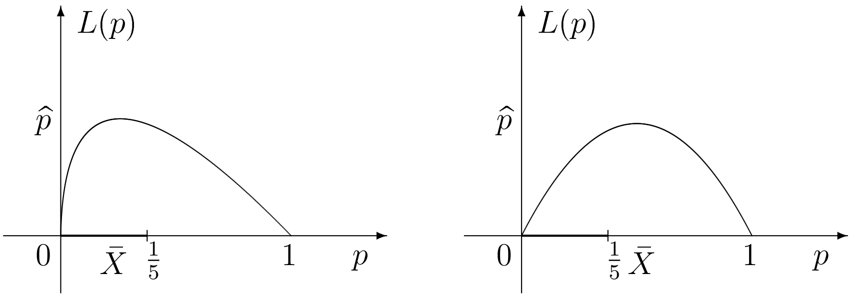
Using Excel, for example, it is easy to verify that $\widehat{\alpha} = 1.51909$.

Exercise 13 In Example 1 we have shown that the maximum of the likelihood function

$$L(p) = p^{\sum_{i=1}^n X_i}(1-p)^{n-\sum_{i=1}^n X_i}$$

is attained when $p = \bar{X}$.

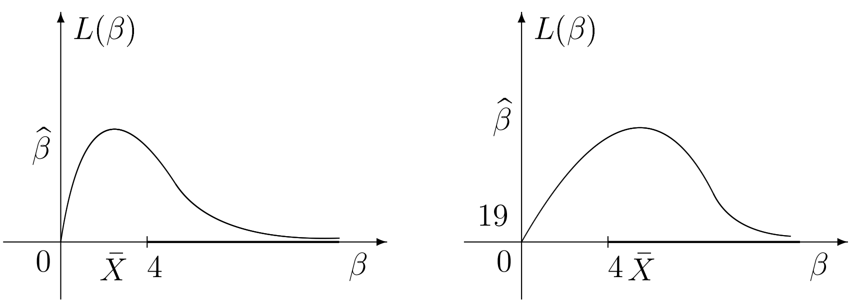
We will plot this likelihood function against values of p when $\bar{X}$ is on either side of $1/5$ to see where the maximums of this function are attained on $[0, 1/5]$.



From the graphs, if $0 \leqslant \bar{X} \leqslant 1/5$, then the maximum of $L(p)$ on the interval $0 \leqslant p \leqslant 1/5$ is attained at $\bar{X}$, whereas when $\bar{X} > 1/5$, then the maximum of the likelihood function on this interval is attained at $1/5$. Thus, the MLE of p is

$$\widehat{p} = \begin{cases} \bar{X}, & \text{if } 0 \leqslant \bar{X} \leqslant 1/5, \\ 1/5, & \text{if } \bar{X} > 1/5. \end{cases}$$

Exercise 14 We know from Exercise 9 that in the general case of $\beta > 0$, the likelihood function $L(\beta) = \dfrac{1}{\beta^n} \exp\{-\sum_{i=1}^n X_i/\beta\}$ attains its maximum at $\widehat{\beta} = \bar{X}$. In this exercise, the values of β are bounded from below by 4. The two graphs below present two possible scenarios: when $0 \leqslant \bar{X} < 4$ and when $\bar{X} \geqslant 4$.



As seen on the graphs, the maximum of the likelihood function is attained on $[4, \infty)$ at $\widehat{\beta} = 4$ if $0 \leqslant \bar{X} < 4$, and at $\widehat{\beta} = \bar{X}$, if $\bar{X} \geqslant 4$.

EXERCISE 15 From Exercise 11, we know that if there are no restrictions on the value of μ, the maximum of the likelihood function

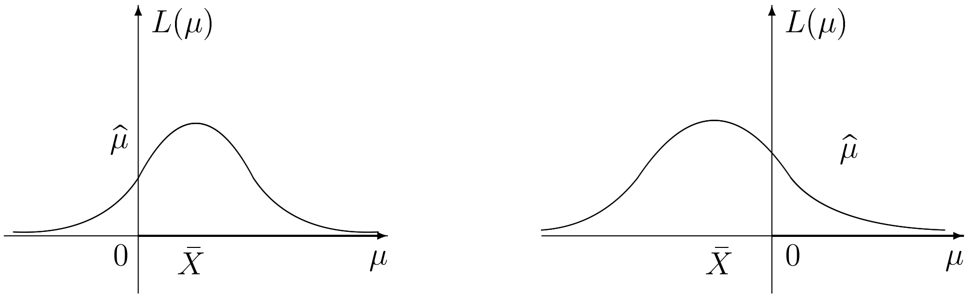
$$L(\mu) = \frac{1}{(2\pi)^{n/2}} \exp\big\{-\sum_{i=1}^{n}(X_i - \mu)^2/2\big\}$$

is attained at $\widehat{\mu} = \bar{X}$. In the present exercise, it is assumed that $\mu \geqslant 0$.

To see how the plot of $L(\mu)$ looks like, we rewrite the likelihood function as

$$\begin{aligned} L(\mu) &= \frac{1}{(2\pi)^{n/2}} \exp\Big\{-\frac{1}{2}\Big(\sum_{i=1}^{n} X_i^2 - 2\mu n\bar{X} + n\mu^2\Big)\Big\} \\ &= \frac{1}{(2\pi)^{n/2}} \exp\Big\{-\frac{1}{2}\sum_{i=1}^{n} X_i^2 + \frac{n}{2}\bar{X}^2\Big\} \exp\Big\{-\frac{n}{2}(\mu - \bar{X})^2\Big\}. \end{aligned}$$

From here we can see that $L(\mu)$ is bell shaped and is centered around $\bar{X}$. Now we plot $L(\mu)$ in the cases $\bar{X} \geqslant 0$ and $\bar{X} < 0$, respectively, to determine at what respective points the maximums are attained on $[0, \infty)$.



As depicted on the graphs, in the case when $\bar{X} \geqslant 0$, the MLE of μ is $\widehat{\mu} = \bar{X}$, while if $\bar{X} < 0$, then $\widehat{\mu} = 0$.

EXERCISE 16 The likelihood function is calculated as

$$L(\theta; X_1, X_2, X_3) = f(1;\theta)f(4;\theta)f(2;\theta) = \begin{cases} (1/4)(1/4)(1/2) = 0.03125, & \text{if } \theta = 0, \\ (1/2)(1/2)(0) = 0, & \text{if } \theta = 1/3, \\ (3/5)(1/5)(1/5) = 0.024, & \text{if } \theta = 1/4. \end{cases}$$

The largest value of the likelihood function is 0.03125 and corresponds to the MLE $\widehat{\theta} = 0$.

Exercise 17 We know from Example 1 that the MLE of p is $\widehat{p} = \bar{X}$, and by Theorem 1, the MLE of $\mathbb{V}ar(X_1) = p(1-p)$ is $\bar{X}(1-\bar{X})$.

Exercise 18 By Exercise 5, the MLE of p is $\widehat{p} = 1/\bar{X}$. Hence, using Theorem 1, we find that the MLE of $\mathbb{E}(X_1) = 1/p$ as $1/(1/\bar{X}) = \bar{X}$.

Exercise 19 By Exercise 6, the MLE of λ is $\widehat{\lambda} = \bar{X}$. Thus, using Theorem 1, we conclude that the MLE of $\mathbb{P}(X_1 = 1) = \lambda \exp\{-\lambda\}$ is $\bar{X}\exp\{-\bar{X}\}$.

Exercise 20 As shown in Exercise 7, the MLE of θ is $\widehat{\theta} = X_{(n)}$. Applying Theorem 1, we get that the MLE of $\mathbb{V}ar(X_1) = \theta^2/12$ is $X_{(n)}^2/12$.

Exercise 21 The likelihood function has the form

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} (1-\exp\{-\theta\})^{X_i} (\exp\{-\theta\})^{1-X_i} \\ &= (1-\exp\{-\theta\})^{\sum_{i=1}^n X_i} (\exp\{-\theta\})^{n-\sum_{i=1}^n X_i}. \end{aligned}$$

This is a Poisson distribution truncated at $x = 1$, or, alternatively, it can be looked at as a Bernoulli distribution with $p = 1 - \exp\{-\theta\}$. The quickest way to find the MLE of θ is to recall from by Example 1 that the MLE of p is $\widehat{p} = \bar{X}$, and now use Theorem 1 to conclude that the MLE of θ solves $\widehat{p} = \bar{X} = 1 - \exp\{-\widehat{\theta}\}$. Thus, $\widehat{\theta} = -\ln(1-\bar{X})$.

Exercise 22 In Exercise 11 we have shown that the MLE of σ^2 is $\widehat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$. We use this result and Theorem 1 to conclude that the MLE of σ is $\widehat{\sigma} = \sqrt{\widehat{\sigma}^2} = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2}$.

Exercise 23 To find the MM estimator of p, we equate the theoretical and empirical first moments. We have

$$\mathbb{E}(X_1) = p = \frac{\sum_{i=1}^n X_i}{n} = \bar{X}.$$

The solution is $\widehat{p} = \bar{X}$, and, thus, the MM estimator coincides with the MLE for p.

Exercise 24 The MM estimator for p solves the equation $\mathbb{E}(X_1) = Np = \bar{X}$, or $\widehat{p} = \bar{X}/N$. It is the same as the MLE.

Exercise 25 The MM estimator for p satisfies

$$\mathbb{E}(X_1) = \frac{1}{p} = \bar{X}.$$

Hence, $\widehat{p} = 1/\bar{X}$, and it coincides with the MLE for p.

Exercise 26 The MM estimator for λ is the solution of the equation

$$\mathbb{E}(X_1) = \lambda = \bar{X},$$

and so, $\widehat{\lambda} = \bar{X}$. It is the same as the MLE.

Exercise 27 To find the MM estimator for θ we write

$$\mathbb{E}(X_1) = \frac{\theta}{2} = \bar{X},$$

thus, $\widehat{\theta} = 2\bar{X}$. This estimator is not the same as $X_{(n)}$, the MLE of θ. Moreover, for some observations, $2\bar{X}$ is smaller than $X_{(n)}$, and hence, the MM estimator doesn't always make sense. For example, if $X_1 = 1, X_2 = 1, X_3 = 2$, and $X_4 = 8$. Then $2\bar{X} = 6$, whereas $X_{(4)} = 8$, so we have an observation that exceeds our MM estimate of θ.

Exercise 28 We find the MM estimators for a and b by solving the system of equations:

$$\begin{cases} \mathbb{E}(X_1) = \dfrac{a+b}{2} = \bar{X}, \\ \mathbb{E}(X_1^2) = \displaystyle\int_a^b \frac{x^2}{b-a}\,dx = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3} = \frac{\sum_{i=1}^n X_i^2}{n}. \end{cases}$$

Hence, $\widehat{a}$ and $\widehat{b}$ satisfy the equations

$$\begin{cases} \widehat{a} + \widehat{b} = 2\bar{X}, \\ \widehat{a}^2 + \widehat{a}\widehat{b} + \widehat{b}^2 = 3\dfrac{\sum_{i=1}^n X_i^2}{n}. \end{cases}$$

Squaring the first equation and subtracting the second, we get

$$\begin{cases} \widehat{a} + \widehat{b} = 2\bar{X}, \\ \widehat{a}\widehat{b} = 4\bar{X}^2 - 3\dfrac{\sum_{i=1}^n X_i^2}{n}. \end{cases}$$

Letting $\widehat{b} = 2\bar{X} - \widehat{a}$ and plugging it into the second equation, we arrive at a quadratic equation. The system becomes

$$\begin{cases} \widehat{a} + \widehat{b} = 2\bar{X}, \\ \widehat{a}^2 - 2\bar{X}\widehat{a} + 4\bar{X}^2 - 3\dfrac{\sum_{i=1}^n X_i^2}{n} = 0. \end{cases}$$

The solution of this system is

$$\widehat{a} = \bar{X} - \sqrt{3\Big(\frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2\Big)}, \text{ and } \widehat{b} = \bar{X} + \sqrt{3\Big(\frac{\sum_{i=1}^n X_i^2}{n} - \bar{X}^2\Big)}.$$

Note that these estimators are not the same as the MLE's $X_{(1)}$ and $X_{(n)}$. In addition, they

may not make sense for some data sets, where the minimum is below $\widehat{a}$ and/or the maximum is above $\widehat{b}$.

Exercise 29 The MM estimator for β is the solution of the equation $\mathbb{E}(X_1) = \beta = \bar{X}$, thus, $\widehat{\beta} = \bar{X}$, and is equal to the MLE.

Exercise 30 The MM estimator $\widehat{\beta}$ satisfies $\bar{X} = \mathbb{E}(X_1) = 1/\widehat{\beta}$. Thus, $\widehat{\beta} = 1/\bar{X}$,the same as the MLE.

Exercise 31 We write $\mathbb{E}(\widehat{p}) = \mathbb{E}(\bar{X}/N) = \mathbb{E}(X_1)/N = Np/N = p$, thus the estimator is unbiased.

Exercise 32 The sum $\sum_{i=1}^n X_i$ of n independent *Geometric*(p) random variables has a $NegativeBinomial(n, p)$ distribution with the pmf

$$P(\sum_{i=1}^n X_i = x) = \binom{x-1}{n-1} p^n (1-p)^{x-n}, \quad x = n, n+1, \dots.$$

So, we write

$$\mathbb{E}(\widehat{p}) = \mathbb{E}\big(\frac{1}{\bar{X}}\big) = \mathbb{E}\big(\frac{n}{\sum_{i=1}^n X_i}\big) = \sum_{x=n}^{\infty} \frac{n}{x}\binom{x-1}{n-1} p^n (1-p)^{x-n} \neq p.$$

Thus, the estimator is biased.

For an estimator $\widehat{p} = \widehat{p}(X_1)$ to be an unbiased estimator of p, it must satisfy the identity $\mathbb{E}(\widehat{p}) = \mathbb{E}(\widehat{p}(X_1)) = \sum_{x=1}^{\infty} \widehat{p}(x)p(1-p)^{x-1} = p$. Since the left-hand side is a polynomial in p, the only solution is

$$\widehat{p}(X_1) = \begin{cases} 1, & \text{if } X_1 = 1, \\ 0, & \text{if } X_1 = 2, 3, \ldots. \end{cases}$$

EXERCISE 33 We have that $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = 1/p$, thus, it is an unbiased estimator.

EXERCISE 34 We write $\mathbb{E}(\widehat{\lambda}) = \mathbb{E}(\bar{X}) = \lambda$, hence, $\widehat{\lambda}$ is an unbiased estimator of λ.

EXERCISE 35 We start by finding the cdf of the largest order statistic:

$$\begin{aligned} F_{X_{(n)}}(x;\theta) &= \mathbb{P}(X_{(n)} \leqslant x) = \mathbb{P}(X_1 \leqslant x, \ldots, X_n \leqslant x) \\ &= \mathbb{P}(X_1 \leqslant x) \cdots \mathbb{P}(X_n \leqslant x), \text{ by independence,} \\ &= \frac{x^n}{\theta^n}, \quad 0 \leqslant x \leqslant \theta. \end{aligned}$$

From here, the density of $X_{(n)}$ is $f_{X_{(n)}}(x;\theta) = F'_{X_{(n)}}(x;\theta) = nx^{n-1}/\theta^n, \;\; 0 \leqslant x \leqslant \theta$. And thus the expected value is derived as

$$\mathbb{E}(X_{(n)}) = \int_0^\theta x\, n \frac{x^{n-1}}{\theta^n}\, dx = \frac{n}{n+1}\,\theta = \big(1 - \frac{1}{n+1}\big)\theta < \theta.$$

We can see that $X_{(n)}$ is a biased estimator of θ, and, in fact, it underestimates θ by $1/(n+1)$th of θ, on average. An unbiased estimator of θ based on the maximum value is $\dfrac{n+1}{n}X_{(n)}$.
In the case of the MM estimator of θ we write $\mathbb{E}(2\bar{X}) = 2\mathbb{E}(X_1) = 2(\theta/2) = \theta$. Thus, it is unbiased.

EXERCISE 36 We derive the cdf of the smallest order statistic. We have

$$\begin{aligned} \mathbb{P}(X_{(1)} \geqslant x) &= \mathbb{P}(X_1 \geqslant x, \ldots, X_n \geqslant x) \\ &= \mathbb{P}(X_1 \geqslant x) \cdots P(X_n \geqslant x), \text{ by independence,} \\ &= \frac{(b-x)^n}{(b-a)^n}, \quad a \leqslant x \leqslant b. \end{aligned}$$

Hence, the cdf of $X_{(1)}$ is

$$F_{X_{(1)}}(x;a,b) = \mathbb{P}(X_{(1)} \leqslant x) = 1 - \mathbb{P}(X_{(1)} \geqslant x) = 1 - \frac{(b-x)^n}{(b-a)^n}, \quad a \leqslant x \leqslant b.$$

The pdf is equal to $f_{X_{(1)}}(x;a,b) = F'_{X_{(1)}}(x;a,b) = n\,(b-x)^{n-1}/(b-a)^n, \quad a \leqslant x \leqslant b$. The expectation is found as

$$\begin{aligned}
\mathbb{E}(X_{(1)}) &= \int_a^b x\,n\,\frac{(b-x)^{n-1}}{(b-a)^n}\,dx = -\int_a^b (b-x-b)\,n\,\frac{(b-x)^{n-1}}{(b-a)^n}\,dx \\
&= -\int_a^b n\,\frac{(b-x)^n}{(b-a)^n}\,dx + b\int_a^b n\,\frac{(b-x)^{n-1}}{(b-a)^n}\,dx \\
&= -\frac{n}{n+1}\,(b-a) + b = a + \frac{1}{n+1}\,(b-a) > a.
\end{aligned}$$

Thus, $X_{(1)}$ is biased, and overestimates the lower endpoint a by $1/(n+1)$th of the length $b-a$ of the interval, on average.

Further, the cdf of the nth order statistic is

$$\begin{aligned}
F_{X_{(n)}}(x;a,b) &= \mathbb{P}(X_{(n)} \leqslant x) = \mathbb{P}(X_1 \leqslant x, \ldots, X_n \leqslant x) \\
&= \mathbb{P}(X_1 \leqslant x)\cdots\mathbb{P}(X_n \leqslant x), \text{ by independence,} \\
&= \frac{(x-a)^n}{(b-a)^n}, \quad a \leqslant x \leqslant b.
\end{aligned}$$

The pdf of $X_{(n)}$ is $f_{X(n)}(x;a,b) = F'_{X_{(n)}}(x;a,b) = n\,(x-a)^{n-1}/(b-a)^n, \quad a \leqslant x \leqslant b$. The mean is computed as

$$\begin{aligned}
\mathbb{E}(X_{(n)}) &= \int_a^b x\,n\,\frac{(x-a)^{n-1}}{(b-a)^n}\,dx = \int_a^b (x-a+a)\,n\,\frac{(x-a)^{n-1}}{(b-a)^n}\,dx \\
&= \int_a^b n\,\frac{(x-a)^n}{(b-a)^n}\,dx + a\int_a^b n\,\frac{(x-a)^{n-1}}{(b-a)^n}\,dx \\
&= \frac{n}{n+1}\,(b-a) + a = b - \frac{1}{n+1}\,(b-a) < b.
\end{aligned}$$

This indicates that $X_{(n)}$ is biased, and, on average, it underestimates the upper endpoint b by $1/(n+1)$th of the length $b-a$ of the interval.

To see what estimators based on $X_{(1)}$ and $X_{(n)}$ are unbiased estimators of a and b, we solve the following system with respect to a and b:

$$\begin{cases}
\mathbb{E}(X_{(1)}) = a + \dfrac{1}{n+1}\,(b-a), \\
\mathbb{E}(X_{(n)}) = b - \dfrac{1}{n+1}\,(b-a).
\end{cases}$$

Adding and subtracting the equations yield

$$\begin{cases}
\mathbb{E}\big(X_{(1)} + X_{(n)}\big) = a + b, \\
\mathbb{E}\big(X_{(n)} - X_{(1)}\big) = \dfrac{n-1}{n+1}\,(b-a),
\end{cases} \qquad \text{or} \qquad
\begin{cases}
\mathbb{E}\big(X_{(1)} + X_{(n)}\big) = a + b, \\
\mathbb{E}\Big[\dfrac{n+1}{n-1}\,\big(X_{(n)} - X_{(1)}\big)\Big] = b - a.
\end{cases}$$

Again adding and subtracting the equations yield

$$\begin{cases} a = \mathbb{E}\Big[\dfrac{1}{2}\big(X_{(1)} + X_{(n)} - \dfrac{n+1}{n-1}\,(X_{(n)} - X_{(1)})\big)\Big] = \mathbb{E}\Big[\dfrac{1}{n-1}\big(n\,X_{(1)} - X_{(n)}\big)\Big], \\ \\ b = \mathbb{E}\Big[\dfrac{1}{2}\big(X_{(1)} + X_{(n)} + \dfrac{n+1}{n-1}\,(X_{(n)} - X_{(1)})\big)\Big] = \mathbb{E}\Big[\dfrac{1}{n-1}\big(n\,X_{(n)} - X_{(1)}\big)\Big]. \end{cases}$$

Thus, $\dfrac{1}{n-1}\big(n\,X_{(1)} - X_{(n)}\big)$ is an unbiased estimator of a, and $\dfrac{1}{n-1}\big(n\,X_{(n)} - X_{(1)}\big)$ is an unbiased estimator of b.

EXERCISE 37 We know that $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = \beta$. Thus, $\bar{X}$ is an unbiased estimator of β.

EXERCISE 38 As the sum of n independent exponential random variables, $\sum_{i=1}^n X_i$ has a Gamma distribution with the pdf $f(x) = \dfrac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}$, $x > 0$. Hence, we expected value of $1/\bar{X}$ can be computed explicitly as follows:

$$\mathbb{E}\big(1/\bar{X}\big) = \mathbb{E}\big(n/\sum_{i=1}^n X_i\big) = \int_0^\infty \frac{n}{x}\,\frac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}\,dx$$
$$= \frac{n(n-2)!}{(n-1)!}\,\beta \int_0^\infty \frac{\beta^{n-1}\, x^{n-2}\, \exp\{-\beta\, x\}}{(n-2)!}\,dx = \frac{n}{n-1}\,\beta.$$

Thus, $1/\bar{X}$ is a biased estimator of β, but $\dfrac{n-1}{n\bar{X}}$ is unbiased.

EXERCISE 39 Since $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = \mu$, $\bar{X}$ is an unbiased estimator of μ. Next, we compute the expected value of $\widehat{\sigma}^2$. We write

$$\mathbb{E}(\widehat{\sigma}^2) = \mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2\Big] = \frac{1}{n}\,\mathbb{E}\big(\sum_{i=1}^n X_i^2 - n\,\bar{X}^2\big)$$
$$= \mathbb{E}(X_1^2) - \mathbb{E}(\bar{X}^2) = \mathbb{V}ar(X_1) + (\mathbb{E}(X_1))^2 - \big[\mathbb{V}ar(\bar{X}) + (\mathbb{E}(\bar{X}))^2\big]$$
$$= \sigma^2 + \mu^2 - \big(\frac{\sigma^2}{n} + \mu^2\big) = \frac{n-1}{n}\,\sigma^2.$$

Hence, $\dfrac{n}{n-1}\widehat{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i - \bar{X})^2$ is an unbiased estimator of σ^2.

EXERCISE 40 (a) The density of $X_{(n)}$ is $f_{X_{(n)}}(x) = \dfrac{nx^{n-1}}{\theta^n}$, $0 < x < \theta$, therefore, the mean squared error of the unbiased estimator $\dfrac{n+1}{n}\,X_{(n)}$ is computed as follows:

$$MSE = \mathbb{V}ar\Big[\frac{n+1}{n}\,X_{(n)}\Big] = \Big(\frac{n+1}{n}\Big)^2\,\mathbb{V}ar\big(X_{(n)}\big)$$

$$= \Big(\frac{n+1}{n}\Big)^2 \Big[\int_0^\theta \frac{n x^{n+1}}{\theta^n}\, dx - \Big(\frac{n\theta}{n+1}\Big)^2\Big] = \Big(\frac{n+1}{n}\Big)^2 \Big[\frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2}\Big]$$

$$= \Big(\frac{n+1}{n}\Big)^2 \frac{n\,\theta^2}{(n+2)(n+1)^2} = \frac{\theta^2}{n(n+2)} \to 0 \text{ as } n \to \infty.$$

Thus, it is a consistent estimator of θ.

(b) The bias of $X_{(n)}$ is equal to $\mathbb{E}(X_{(n)}) - \theta = \dfrac{n}{n+1}\theta - \theta = -\dfrac{\theta}{n+1}$. Since the bias goes to zero as n increases, this estimator is asymptotically unbiased. Its mean square error is

$$MSE = \mathbb{V}ar(X_{(n)}) + \big[bias(X_{(n)},\theta)\big]^2 = \frac{n\,\theta^2}{(n+2)(n+1)^2} + \Big[-\frac{\theta}{n+1}\Big]^2$$

$$= \frac{\theta^2}{(n+1)^2}\Big[\frac{n}{n+2} + 1\Big] = \frac{2\theta^2}{(n+1)(n+2)}.$$

Since $MSE \to 0$ as $n \to \infty$, the estimator is consistent.

(c) First we will show that $\dfrac{n+2}{n+1}X_{(n)}$ has the smallest MSE among all estimators of the form $c\,X_{(n)}$, where $c = c(n)$ is a function of n. We write

$$MSE = \mathbb{V}ar(c\,X_{(n)}) + \big[bias(c\,X_{(n)},\theta)\big]^2 = \frac{c^2\,n\,\theta^2}{(n+2)(n+1)^2} + \Big(\frac{c\,n\,\theta}{n+1} - \theta\Big)^2.$$

We would like to minimize with respect to c the following function

$$\frac{c^2\,n}{(n+2)(n+1)^2} + \Big(\frac{c\,n}{n+1} - 1\Big)^2.$$

Taking derivative with respect to c and setting it equal to zero, we arrive at the identity

$$\frac{2cn}{(n+2)(n+1)^2} + \frac{2n}{n+1}\Big(\frac{cn}{n+1} - 1\Big) = 0,$$

from where $c = \dfrac{n+2}{n+1}$. The MSE of this estimator is

$$MSE = \Big(\frac{n+2}{n+1}\Big)^2 \frac{n\,\theta^2}{(n+2)(n+1)^2} + \theta^2\Big(\frac{(n+2)\,n}{(n+1)^2} - 1\Big)^2$$

$$= \frac{(n+2)\,n\,\theta^2}{(n+1)^4} + \frac{\theta^2}{(n+1)^4} = \frac{\theta^2}{(n+1)^2}.$$

The MSE goes to zero, as n increases, which proves the consistency of the estimator.

EXERCISE 41 (a) The MM estimator $2\bar{X}_n$ is unbiased, and its mean square error is obtained as

$$MSE = \mathbb{V}ar(2\bar{X}_n) = 4\frac{\theta^2}{12n} = \frac{\theta^2}{3n}.$$

The MSE tends to zero when n goes to infinity, implying consistency of the estimator.

(b) The bias of the estimator $\bar{X}_n$ is $bias(\bar{X}_n, \theta) = \mathbb{E}(\bar{X}_n) - \theta = \dfrac{\theta}{2} - \theta = -\dfrac{\theta}{2} \nrightarrow 0$, as $n \to \infty$. It means that this estimator is not asymptotically unbiased, and, consequently, not consistent.

EXERCISE 42 (a) The bias of $1/\bar{X}_n$ is $bias(1/\bar{X}_n, \beta) = \mathbb{E}\big(1/\bar{X}_n\big) - \beta = \displaystyle\int_0^\infty \frac{n}{x} \frac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}\, dx - \beta = \frac{n\beta}{n-1} - \beta = \frac{\beta}{n-1} \to 0$, as $n \to \infty$. Hence, the estimator is asymptotically unbiased. The mean square error is

$$
\begin{aligned}
MSE = \mathbb{V}ar\Big[\frac{1}{\bar{X}_n}\Big] + \big[bias(1/\bar{X}_n, \beta)\big]^2 &= \int_0^\infty \frac{n^2}{x^2} \frac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}\, dx \\
-\Big(\frac{n\beta}{n-1}\Big)^2 + \Big(\frac{\beta}{n-1}\Big)^2 &= \frac{n^2\beta^2}{(n-1)(n-2)} - \frac{n^2\beta^2}{(n-1)^2} + \frac{\beta^2}{(n-1)^2} \\
&= \frac{(n+2)\beta^2}{(n-1)(n-2)} \to 0,
\end{aligned}
$$

as n increases. Therefore, the estimator is consistent.

(b) The mean square error of an unbiased estimator $\dfrac{n-1}{n\bar{X}_n}$ is equal to its variance. We derive the expression for the variance as:

$$
\begin{aligned}
\mathbb{V}ar\Big(\frac{n-1}{n\,\bar{X}}\Big) &= (n-1)^2 \Big[\int_0^\infty \frac{1}{x^2} \frac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}\, dx \\
-\Big(\int_0^\infty \frac{1}{x} \frac{\beta^n\, x^{n-1}\, \exp\{-\beta\, x\}}{(n-1)!}\, dx\Big)^2\Big] &= (n-1)^2 \Big[\frac{\beta^2\,(n-3)!}{(n-1)!} \int_0^\infty \frac{\beta^{n-2}\, x^{n-3}\, \exp\{-\beta\, x\}}{(n-3)!}\, dx \\
-\Big(\frac{\beta\,(n-2)!}{(n-1)!} \int_0^\infty \frac{\beta^{n-1}\, x^{n-2}\, \exp\{-\beta\, x\}}{(n-2)!}\Big)^2\Big] &= (n-1)^2\, \beta^2 \Big[\frac{1}{(n-1)(n-2)} - \frac{1}{(n-1)^2}\Big] \\
&= \beta^2 \Big[\frac{n-1}{n-2} - 1\Big] = \frac{\beta^2}{n-2}.
\end{aligned}
$$

Thus, the mean square error is

$$
MSE = \mathbb{V}ar\Big[\frac{n-1}{n\bar{X}_n}\Big] = \frac{\beta^2}{n-2} \to 0, \text{ as } n \to \infty.
$$

This proves consistency.

EXERCISE 43 (a) The MSE of an unbiased estimator $\dfrac{1}{n-1}\displaystyle\sum_{i=1}^n (X_i - \bar{X}_n)^2$ is equal to its variance, which is $\dfrac{2\sigma^4}{n-1}$. This quantity tends to zero as n increases, implying the consistency of the estimator.

(b) The bias of the MLE $\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2$ is computed as

$$bias\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2,\sigma^2\Big)=\mathbb{E}\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\Big]-\sigma^2$$

$$=\frac{n-1}{n}\,\mathbb{E}\Big[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\Big]-\sigma^2=\frac{n-1}{n}\sigma^2-\sigma^2=-\frac{\sigma^2}{n}\to 0,$$

as n tends to infinity. Thus, this estimator is asymptotically unbiased. Its mean square error is found as

$$\begin{aligned}MSE&=\mathbb{V}ar\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\Big]+\Big[bias\Big(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2,\sigma^2\Big)\Big]^2\\&=\Big(\frac{n-1}{n}\Big)^2\mathbb{V}ar\Big[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\Big]+\Big(-\frac{\sigma^2}{n}\Big)^2\\&=\Big(\frac{n-1}{n}\Big)^2\frac{2\sigma^4}{n-1}+\frac{\sigma^4}{n^2}=\frac{2n-1}{n^2}\sigma^4\to 0\text{ as }n\to\infty,\end{aligned}$$

whence, the estimator is consistent.

Exercise 44 The likelihood function is of the form

$$\begin{aligned}\prod_{i=1}^{n}f(X_i;p)&=\prod_{i=1}^{n}\binom{N}{X_i}p^{X_i}(1-p)^{N-X_i}\\&=\Big[\prod_{i=1}^{n}\binom{N}{X_i}\Big]p^{\sum_{i=1}^{n}X_i}(1-p)^{n\,N-\sum_{i=1}^{n}X_i}.\end{aligned}$$

Now we take

$$g(X_1,\dots,X_n)=\prod_{i=1}^{n}\binom{N}{X_i},\quad \widehat{p}=\sum_{i=1}^{n}X_i,\quad\text{and}\quad h(\widehat{p};p)=p^{\widehat{p}}(1-p)^{n\,N-\widehat{p}}.$$

The likelihood function can be written as the product of g and h, and, therefore, by the factorization theorem, $\sum_{i=1}^{n}X_i$ is sufficient. Since any invertible function of a sufficient statistic is sufficient, $\sum_{i=1}^{n}X_i/(n\,N)=\bar{X}/N$ is also a sufficient statistic for p.

Exercise 45 We write the likelihood function as

$$\prod_{i=1}^{n}f(X_i;p)=\prod_{i=1}^{n}p(1-p)^{X_i-1}=p^n(1-p)^{\sum_{i=1}^{n}X_i-n}.$$

If we suppose that

$$g(X_1,\ldots,X_n) = 1, \quad \widehat{p} = \sum_{i=1}^{n} X_i, \quad \text{and} \quad h(\widehat{p};p) = p^n\,(1-p)^{\widehat{p}-n},$$

then the likelihood function becomes

$$\prod_{i=1}^{n} f(X_i;p) = p^n\,(1-p)^{\widehat{p}-n} = g(X_1,\ldots,X_n)\,h(\widehat{p};p).$$

From here, by factorization theorem, we conclude that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for p, and thus, $\sum_{i=1}^{n} X_i/n = \bar{X}$ is sufficient.

Exercise 46 The likelihood function is of the form

$$\prod_{i=1}^{n} f(X_i;\lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i}}{X_i!}\,\exp\{-\lambda\} = \Big[\prod_{i=1}^{n} X_i!\Big]^{-1} \lambda^{\sum_{i=1}^{n} X_i}\,\exp\{-n\,\lambda\}.$$

We take

$$g(X_1,\ldots,X_n) = \Big[\prod_{i=1}^{n} X_i!\Big]^{-1}, \quad \widehat{\lambda} = \sum_{i=1}^{n} X_i, \quad \text{and} \quad h(\widehat{\lambda};\lambda) = \lambda^{\widehat{\lambda}}\,\exp\{-n\lambda\}.$$

Hence,

$$\prod_{i=1}^{n} f(X_i;\lambda) = g(X_1,\ldots,X_n)\,h(\widehat{\lambda};\lambda),$$

which means that $\sum_{i=1}^{n} X_i$ is sufficient for λ, therefore $\sum_{i=1}^{n} X_i/n = \bar{X}$ is sufficient as well.

Exercise 47 The likelihood function is

$$\prod_{i=1}^{n} f(X_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta}\,\mathbb{I}\{0 \leqslant X_i \leqslant \theta\} = \frac{1}{\theta^n}\,\mathbb{I}\{0 \leqslant X_{(n)} \leqslant \theta\}.$$

We let

$$g(X_1,\ldots,X_n) = 1, \quad \widehat{\theta} = X_{(n)}, \quad \text{and} \quad h(\widehat{\theta},\theta) = \frac{1}{\theta^n}\,\mathbb{I}\{0 \leqslant \widehat{\theta} \leqslant \theta\}.$$

The likelihood function can be factored into g and h. Hence, by the factorization theorem, $X_{(n)}$ is a sufficient statistic for θ.

Exercise 48 We write the likelihood function

$$\prod_{i=1}^{n} f(X_i;a,b) = \prod_{i=1}^{n} \frac{1}{b-a}\,\mathbb{I}\{a \leqslant X_i \leqslant b\} = \frac{1}{(b-a)^n}\,\mathbb{I}\{a \leqslant X_{(1)} \leqslant X_{(n)} \leqslant b\}.$$

If we define $g(X_1,\dots,X_n)=1$, $\widehat{a}=X_{(1)}$, $\widehat{b}=X_{(n)}$, and

$$h(\widehat{a},\widehat{b};a,b)=\frac{1}{(b-a)^n}\,\mathbb{I}\{a\leqslant\widehat{a}\leqslant\widehat{b}\leqslant b\},$$

then the likelihood function is factored into g and h, and, thus, by the factorization theorem, $(X_{(1)},X_{(n)})$ is sufficient for (a,b).

Exercise 49 We have that the likelihood function

$$\prod_{i=1}^{n} f(X_i;\beta)=\prod_{i=1}^{n}\frac{1}{\beta}\,\exp\{-X_i/\beta\}=\frac{1}{\beta^n}\,\exp\{-\sum_{i=1}^{n}X_i/\beta\}.$$

Hence the likelihood function is a product of $g(X_1,\dots,X_n)=1$, and $h(\widehat{\beta};\beta)=\frac{1}{\beta^n}\,\exp\{-\widehat{\beta}/\beta\}$ where $\widehat{\beta}=\sum_{i=1}^{n}X_i$. By the factorization theorem, $\sum_{i=1}^{n}X_i$ is sufficient for β, and therefore $\sum_{i=1}^{n}X_i/n=\bar{X}$ is sufficient.

Exercise 50 The likelihood function has the form

$$\prod_{i=1}^{n} f(X_i;\mu,\sigma^2)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\,\exp\Big\{-\frac{(X_i-\mu)^2}{2\sigma^2}\Big\}$$

$$=\frac{1}{(2\pi\sigma^2)^{n/2}}\,\exp\Big\{-\frac{1}{2\sigma^2}\big(\sum_{i=1}^{n}X_i^2-2\,\mu\sum_{i=1}^{n}X_i+n\,\mu^2\big)\Big\}.$$

Now we let $u=\sum_{i=1}^{n}X_i^2$ and $v=\sum_{i=1}^{n}X_i$, and define $g(X_1,\dots,X_n)=1$, and

$$h(u,v;\mu,\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\,\exp\Big\{-\frac{1}{2\sigma^2}\big(u-2\,\mu\,v+n\,\mu^2\big)\Big\}.$$

We see that the likelihood function factors into g and h, and, thus, by the factorization theorem, (u,v) is sufficient for (μ,σ^2). We now define $\widehat{\mu}=v/n=\bar{X}$ and $\widehat{\sigma}^2=\frac{1}{n-1}(u-v^2/n)=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$. Since $\widehat{\mu}$ and $\widehat{\sigma}^2$ are invertible functions of u and v, we conclude that the vector $(\widehat{\mu},\widehat{\sigma}^2)$ is also a sufficient statistic for (μ,σ^2).

Exercise 51 The estimator $\bar{X}/N$ is sufficient for p, and, thus, it is the UMVUE for p.

Exercise 52 The estimator $\bar{X}$ is sufficient for λ. Therefore, it is the UMVUE for λ.

Exercise 53 The estimator $\bar{X}$ is sufficient for β. Hence, it is the UMVUE for β.

EXERCISE 54 The estimator $\bar{X}$ is sufficient for μ. Thus, it is the UMVUE for μ.

EXERCISE 55 We know that $\dfrac{1}{n-1}\big(n\,X_{(1)} - X_{(n)}\big)$ is an unbiased estimator of a that is based on a sufficient statistic $(X_{(1)}, X_{(n)})$. Likewise, $\dfrac{1}{n-1}\big(n\,X_{(n)} - X_{(1)}\big)$ is an unbiased estimator of b that is based on a sufficient statistic.

EXERCISE 56 The estimator $\dfrac{n-1}{n\,\bar{X}}$ is an unbiased and based on a sufficient statistic for β, $\sum_{i=1}^{n} X_i$, and, therefore, it is the UMVUE for β.

EXERCISE 57 The estimator $\bar{X}$ is sufficient and unbiased for μ. Also, $\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$ is sufficient and unbiased for σ^2. Hence, $\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2$ is the UMVUE for σ^2.

EXERCISE 58 The estimator $\bar{X}$ is a sufficient statistic for λ. We need to find an unbiased estimator of $\lambda(1+\lambda)$ based on $\bar{X}$. We write

$$\mathbb{E}\big[\bar{X}(1+\bar{X})\big] = \mathbb{E}(\bar{X}) + \mathbb{E}(\bar{X}^2) = \mathbb{E}(\bar{X}) + \mathbb{V}ar(\bar{X}) + \big[\mathbb{E}(\bar{X})\big]^2$$
$$= \mathbb{E}(X_1) + \frac{\mathbb{V}ar(X_1)}{n} + \big[\mathbb{E}(X_1)\big]^2 = \lambda + \frac{\lambda}{n} + \lambda^2.$$

It follows that if we subtract $\bar{X}/n$ from $\bar{X}(1+\bar{X})$, we will get an unbiased estimator of $\lambda(1+\lambda)$. Indeed,

$$\mathbb{E}\Big[\bar{X}(1+\bar{X}) - \frac{\bar{X}}{n}\Big] = \lambda + \frac{\lambda}{n} + \lambda^2 - \frac{\lambda}{n} = \lambda(1+\lambda).$$

Thus, by the Lehmann-Scheffé theorem, $\bar{X}(1+\bar{X}) - \dfrac{\bar{X}}{n} = \big(1 - \dfrac{1}{n}\big)\bar{X} + \bar{X}^2$ is the UMVUE for the second moment $\mathbb{E}(X_1^2) = \lambda(1+\lambda)$.

EXERCISE 59 We know that $(\bar{X}, \widehat{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_1 - \bar{X})^2)$ is sufficient for (μ, σ^2). Also, $\mathbb{E}(\bar{X}) = \mu$ and $\mathbb{E}(\widehat{\sigma}^2) = \sigma^2$. Remained to find an unbiased estimator of μ^2. We take $\bar{X}^2$ and compute $\mathbb{E}(\bar{X}^2) = \mathbb{V}ar(\bar{X}) + \big[\mathbb{E}(\bar{X})\big]^2 = \dfrac{\sigma^2}{n} + \mu^2$. Thus, $\bar{X}^2 - \dfrac{\widehat{\sigma}^2}{n}$ is an unbiased estimator of μ^2. Indeed, $\mathbb{E}\Big[\bar{X}^2 - \dfrac{\widehat{\sigma}^2}{n}\Big] = \dfrac{\sigma^2}{n} + \mu^2 - \dfrac{\sigma^2}{n} = \mu^2$. Hence, it is the UMVUE of μ^2.

Exercise 60 Take $\mathbb{I}_{\{X_1=0\}}$. It is an unbiased estimator of $\mathbb{P}(X_1 = 0)$. Computing the conditional expectation, conditioning on a sufficient statistic $\sum_{i=1}^n X_i = x$, we arrive at the UMVUE of $e^{-\lambda}$:

$$\mathbb{E}\Big[\mathbb{I}_{\{X_1=0\}} \mid \sum_{i=1}^n X_i = x\Big] = \mathbb{P}\Big(X_1 = 0 \mid \sum_{i=1}^n X_i = x\Big) = \frac{\mathbb{P}\big(X_1 = 0, \sum_{i=1}^n X_i = x\big)}{\mathbb{P}(\sum_{i=1}^n X_i = x)}$$

$$\overset{ind}{=} \frac{\mathbb{P}(X_1 = 0)\mathbb{P}(\sum_{i=2}^n X_i = x)}{\mathbb{P}(\sum_{i=1}^n X_i = x)} = e^{-\lambda}\,\frac{[(n-1)\lambda]^x\, e^{-(n-1)\lambda}}{x!}\,\frac{x!}{(n\lambda)^x\, e^{-n\lambda}} = \Big(1 - \frac{1}{n}\Big)^{n\bar{X}}.$$

Exercise 61 Take X_1X_2. It is an unbiased estimator of $\mathbb{E}(X_1X_2) = p^2$. The UMVUE of p^2 is

$$\mathbb{E}\Big[X_1X_2 \mid \sum_{i=1}^n X_i = x\Big] = \frac{\mathbb{P}\Big(X_1 = 1, X_2 = 1, \sum_{i=3}^n X_i = x-2\Big)}{\mathbb{P}\big(\sum_{i=1}^n X_i = x\big)}$$

$$= \frac{(p)(p)\binom{n-2}{x-2}p^{x-2}(1-p)^{n-x}}{\binom{n}{x}p^x(1-p)^{n-x}} = \frac{x(x-1)}{n(n-1)} = \frac{n\bar{X}(n\bar{X}-1)}{n(n-1)}.$$

Exercise 62 The UMVUE of $\mathbb{P}(X \leqslant 2) = 1 - e^{-2/\theta}$ is

$$\mathbb{E}\Big[\mathbb{I}_{\{X_1\leqslant 2\}} \mid \sum_{i=1}^n X_i = x\Big] = \int_0^2 \frac{f\Big(X_1 = y, \sum_{i=2}^n X_i = x-y\Big)}{f\Big(\sum_{i=1}^n X_i = x\Big)}\,dy$$

$$= \int_0^2 (1/\theta)e^{-y\theta}\,\frac{(x-y)^{n-2}e^{-(x-y)/\theta}}{\Gamma(n-1)\theta^{n-1}} \Big/ \Big[\frac{x^{n-1}e^{-x/\theta}}{\Gamma(n)\theta^n}\Big]\,dy = (n-1)\int_0^2 (1-y/x)^{n-2}\,dy = 1 - \Big(1 - \frac{2}{n\bar{X}}\Big)^{n-1}.$$

Exercise 63 The MLE of p is $\widehat{p} = \bar{X}/N$. The likelihood ratio for testing $H_0 : p = p_0$ against $H_1 : p \neq p_0$ is computed as follows:

$$\Lambda = \frac{\prod_{i=1}^n \binom{N}{X_i} p_0^{X_i}(1-p_0)^{N-X_i}}{\prod_{i=1}^n \binom{N}{X_i}\big(\frac{\bar{X}}{N}\big)^{X_i}\big(1-\frac{\bar{X}}{N}\big)^{N-X_i}} = \frac{p_0^{n\bar{X}}(1-p_0)^{nN-n\bar{X}}}{\big(\frac{\bar{X}}{N}\big)^{n\bar{X}}\big(1-\frac{\bar{X}}{N}\big)^{nN-n\bar{X}}}.$$

The asymptotic likelihood ratio test statistic is

$$\chi^2 = -2\ln\Lambda = 2n\bar{X}\ln\Big[\frac{\bar{X}(1-p_0)}{(N-\bar{X})p_0}\Big] + 2nN\ln\Big[\frac{N-\bar{X}}{N-p_0}\Big].$$

The decision is to reject the null hypothesis if $\chi^2 \geqslant \chi^2_\alpha(1)$.

Exercise 64 The MLE of p is $\widehat{p} = 1/\bar{X}$. Therefore, the likelihood ratio has the form

$$\Lambda = \frac{\prod_{i=1}^n p_0(1-p_0)^{X_i-1}}{\prod_{i=1}^n \big(\frac{1}{\bar{X}}\big)\big(1-\frac{1}{\bar{X}}\big)^{X_i-1}} = \frac{(\bar{p}_0)^n\,(1-p_0)^{n\bar{X}-n}}{\big(\frac{1}{\bar{X}}\big)^n\,\big(1-\frac{1}{\bar{X}}\big)^{n\bar{X}-n}}.$$

The asymptotic likelihood ratio test statistic is

$$\chi^2 = -2\ln\Lambda = 2n\ln\Big[\frac{1-p_0}{(\bar{X}-1)p_0}\Big] + 2n\bar{X}\ln\Big[\frac{\bar{X}-1}{\bar{X}(1-p_0)}\Big].$$

If this statistic is in excess of the critical value $\chi^2_\alpha(1)$, then H_0 is rejected.

EXERCISE 65 The MLE of λ is $\widehat{\lambda} = \bar{X}$, thus the likelihood ratio is

$$\Lambda = \frac{\prod_{i=1}^n \frac{\lambda_0^{X_i}}{X_i!} \exp\{-\lambda_0\}}{\prod_{i=1}^n \frac{\bar{X}^{X_i}}{X_i!} \exp\{-\bar{X}\}} = \Big(\frac{\lambda_0}{\bar{X}}\Big)^{n\bar{X}} \exp\{-n(\lambda_0 - \bar{X})\},$$

and the asymptotic test statistic is

$$\chi^2 = -2\ln\Lambda = 2n\bar{X}\ln\Big(\frac{\bar{X}}{\lambda_0}\Big) + 2n(\lambda_0 - \bar{X}).$$

The rejection region is of the form $\{X_1, \dots, X_n : \chi^2 \geqslant \chi^2_\alpha(1)\}$.

EXERCISE 66 The MLE of θ is $\widehat{\theta} = X_{(n)}$. The likelihood ratio is written as

$$\Lambda = \frac{\big(1/\theta_0\big)^n \mathbb{I}\{X_{(n)} \leqslant \theta_0\}}{\big(1/X_{(n)}\big)^n} = \Big(\frac{X_{(n)}}{\theta_0}\Big)^n \mathbb{I}\{X_{(n)} \leqslant \theta_0\}.$$

The asymptotic likelihood ratio test statistic is equal to $\chi^2 = -2\ln\Lambda = 2n\ln\theta_0 - 2n\ln X_{(n)}$, and the decision rule is to reject H_0 if either $X_{(n)} > \theta_0$ or $\chi^2 \geqslant \chi^2_\alpha(1)$.

EXERCISE 67 The MLE of β is $\widehat{\beta} = \bar{X}$. Consequently, the likelihood ratio has the expression

$$\Lambda = \frac{\prod_{i=1}^n \big(1/\beta_0\big) \exp\{-X_i/\beta_0\}}{\prod_{i=1}^n \big(1/\bar{X}\big) \exp\{-X_i/\bar{X}\}} = \Big(\frac{\bar{X}}{\beta_0}\Big)^n \exp\Big\{n\Big(1 - \frac{\bar{X}}{\beta_0}\Big)\Big\}.$$

The asymptotic likelihood ratio test statistic is $\chi^2 = -2\ln\Lambda = 2n\ln\beta_0 - 2n\ln\bar{X} + 2n(\bar{X}/\beta_0 - 1)$, and the rejection region is $\{\chi^2 \geqslant \chi^2_\alpha(1)\}$.

EXERCISE 68 The MLE of μ is $\widehat{\mu} = \bar{X}$, and hence the likelihood ratio is of the form

$$\Lambda = \frac{\prod_{i=1}^n (2\pi\sigma^2)^{-1/2} \exp\big\{-\frac{1}{2\sigma^2}(X_i - \mu_0)^2\big\}}{\prod_{i=1}^n (2\pi\sigma^2)^{-1/2} \exp\big\{-\frac{1}{2\sigma^2}(X_i - \bar{X})^2\big\}}$$

$$= \exp\Big\{-\frac{1}{2\sigma^2}\sum_{i=1}^n \big[(X_i - \mu_0)^2 - (X_i - \bar{X})^2\big]\Big\} = \exp\Big\{-\frac{n}{2\sigma^2}(\bar{X} - \mu_0)^2\Big\}.$$

The asymptotic likelihood ratio test statistic is $\chi^2 = -2\ln\Lambda = \dfrac{n}{\sigma^2}(\bar{X} - \mu_0)^2 = \Big(\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\Big)^2 \sim \chi^2(1)$. The decision rule is to reject the null if $\chi^2 \geqslant \chi^2_\alpha(1)$.

EXERCISE 69 For the rejection region $\{X = 6\}$, *power* $= 1 - \beta = p^6,\ 1/3 \leqslant p \leqslant 1$.
(b) For the rejection region $\{X = 5, 6\}$, *power* $= 1-\beta = \binom{6}{5}p^5(1-p)+p^6 = 6p^5(1-p)+p^6 = 6p^5 - 5p^6,\ 1/3 \leqslant p \leqslant 1$.
(c) For the rejection region $\{X = 4, 5, 6\}$, *power* $= 1-\beta = \binom{6}{4}p^4(1-p)^2 + \binom{6}{5}p^5(1-p)+p^6 = 15p^4(1-p)^2 + 6p^5(1-p) + p^6 = 15p^4 - 24p^5 + 10p^6,\ 1/3 \leqslant p \leqslant 1$.

EXERCISE 70 As found in Exercise 66, the test statistic is equal to $\chi^2 = 2n \ln \theta_0 - 2n \ln X_{(n)}$, and the rejection region is $\{X_{(n)} > \theta_0 \text{ or } \chi^2 \geqslant \chi^2_\alpha(1)\}$. Therefore, the power is computed as:

$$\begin{aligned}
power = 1 - \beta &= \mathbb{P}\big(X_{(n)} > \theta_0 \text{ or } \chi^2 \geqslant \chi^2_\alpha(1) \,|\, \theta \neq \theta_0\big) \\
&= \mathbb{P}\big(X_{(n)} > \theta_0 \text{ or } 2n \ln \theta_0 - 2n \ln\, X_{(n)} \geqslant \chi^2_\alpha(1) \,|\, \theta \neq \theta_0\big) \\
&= \mathbb{P}\Big(X_{(n)} > \theta_0 \text{ or } X_{(n)} \leqslant \theta_0\ \exp\Big\{-\frac{\chi^2_\alpha(1)}{2n}\Big\} \,|\, \theta \neq \theta_0\Big) \\
&= 1 - F_{X_{(n)}}(\theta_0) + F_{X_{(n)}}\Big(\theta_0\ \exp\Big\{-\frac{\chi^2_\alpha(1)}{2n}\Big\}\Big)
\end{aligned}$$

$$= 1 - \frac{\theta_0^n}{\theta^n} + \frac{\theta_0^n\ \exp\big\{-\frac{\chi^2_\alpha(1)}{2}\big\}}{\theta^n} = 1 - \Big(1 - \exp\Big\{-\frac{\chi^2_\alpha(1)}{2}\Big\}\Big)\frac{\theta_0^n}{\theta^n},\ \theta \neq \theta_0.$$

EXERCISE 71 The constant k is determined by the significance level α, that is, it solves $\alpha = \mathbb{P}(\bar{X} > k \,|\, \mu = \mu_0) = \mathbb{P}\Big(Z > \dfrac{k - \mu_0}{\sigma/\sqrt{n}}\Big) = 1-\Phi\Big(\dfrac{k - \mu_0}{\sigma/\sqrt{n}}\Big)$. From here, $k = \dfrac{\sigma}{\sqrt{n}}\,\Phi^{-1}(1-\alpha)+\mu_0$. The power of the test is then computed as *power* $= 1 - \beta = \mathbb{P}(\bar{X} > k \,|\, \mu = \mu_1) = 1 - \mathbb{P}\Big(Z \leqslant \dfrac{k - \mu_1}{\sigma/\sqrt{n}}\Big) = 1 - \Phi\Big(\dfrac{k - \mu_1}{\sigma/\sqrt{n}}\Big) = 1 - \Phi\Big(\dfrac{\frac{\sigma}{\sqrt{n}}\,\Phi^{-1}(1-\alpha) + \mu_0 - \mu_1}{\sigma/\sqrt{n}}\Big) = 1 - \Phi\Big(\Phi^{-1}(1-\alpha) - \dfrac{\mu_1 - \mu_0}{\sigma/\sqrt{n}}\Big)$.