Notice that $L(\theta)$ is a function only of θ and $\overline{y}$ and that if

$$g(\overline{y}, \theta) = \frac{e^{-n\overline{y}/\theta}}{\theta^n} \quad \text{and} \quad h(y_1, y_2, \ldots, y_n) = 1,$$

then

$$L(y_1, y_2, \ldots, y_n \mid \theta) = g(\overline{y}, \theta) \times h(y_1, y_2, \ldots, y_n).$$

Hence, Theorem 9.4 implies that $\overline{Y}$ is a sufficient statistic for the parameter θ. ∎

Theorem 9.4 can be used to show that there are many possible sufficient statistics for any one population parameter. First of all, according to Definition 9.3 or the factorization criterion (Theorem 9.4), the random sample itself is a sufficient statistic. Second, if $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a distribution with a density function with parameter θ, then the set of order statistics $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$, which is a function of $Y_1, Y_2, \ldots, Y_n$, is sufficient for θ. In Example 9.5, we decided that $\overline{Y}$ is a sufficient statistic for the estimation of θ. Theorem 9.4 could also have been used to show that $\sum_{i=1}^{n} Y_i$ is another sufficient statistic. Indeed, for the exponential distribution described in Example 9.5, any statistic that is a one–to–one function of $\overline{Y}$ is a sufficient statistic.

In our initial example of this section, involving the number of successes in n trials, $Y = \sum_{i=1}^{n} X_i$ reduces the data $X_1, X_2, \ldots, X_n$ to a single value that remains sufficient for p. Generally, we would like to find a sufficient statistic that reduces the data in the sample as much as possible. Although many statistics are sufficient for the parameter θ associated with a specific distribution, application of the factorization criterion typically leads to a statistic that provides the "best" summary of the information in the data. In Example 9.5, this statistic is $\overline{Y}$ (or some one-to-one function of it). In the next section, we show how these sufficient statistics can be used to develop unbiased estimators with minimum variance.

Exercises

9.37 Let $X_1, X_2, \ldots, X_n$ denote n independent and identically distributed *Bernoulli* random variables such that

$$P(X_i = 1) = p \quad \text{and} \quad P(X_i = 0) = 1 - p,$$

for each $i = 1, 2, \ldots, n$. Show that $\sum_{i=1}^{n} X_i$ is sufficient for p by using the factorization criterion given in Theorem 9.4.

9.38 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a normal distribution with mean μ and variance σ^2.

a If μ is unknown and σ^2 is known, show that $\overline{Y}$ is sufficient for μ.

b If μ is known and σ^2 is unknown, show that $\sum_{i=1}^{n}(Y_i - \mu)^2$ is sufficient for σ^2.

c If μ and σ^2 are both unknown, show that $\sum_{i=1}^{n} Y_i$ and $\sum_{i=1}^{n} Y_i^2$ are jointly sufficient for μ and σ^2. [Thus, it follows that $\overline{Y}$ and $\sum_{i=1}^{n}(Y_i - \overline{Y})^2$ or $\overline{Y}$ and S^2 are also jointly sufficient for μ and σ^2.]

9.39 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Poisson distribution with parameter λ. Show by conditioning that $\sum_{i=1}^{n} Y_i$ is sufficient for λ.

9.40 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Rayleigh distribution with parameter θ. (Refer to Exercise 9.34.) Show that $\sum_{i=1}^{n} Y_i^2$ is sufficient for θ.

9.41 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Weibull distribution with known m and unknown α. (Refer to Exercise 6.26.) Show that $\sum_{i=1}^{n} Y_i^m$ is sufficient for α.

9.42 If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a geometric distribution with parameter p, show that $\overline{Y}$ is sufficient for p.

9.43 Let $Y_1, Y_2, \ldots, Y_n$ denote independent and identically distributed random variables from a power family distribution with parameters α and θ. Then, by the result in Exercise 6.17, if $\alpha, \theta > 0$,

$$f(y \mid \alpha, \theta) = \begin{cases} \alpha y^{\alpha-1}/\theta^{\alpha}, & 0 \le y \le \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

If θ is known, show that $\prod_{i=1}^{n} Y_i$ is sufficient for α.

9.44 Let $Y_1, Y_2, \ldots, Y_n$ denote independent and identically distributed random variables from a Pareto distribution with parameters α and β. Then, by the result in Exercise 6.18, if $\alpha,\ \beta > 0$,

$$f(y \mid \alpha, \beta) = \begin{cases} \alpha\beta^{\alpha} y^{-(\alpha+1)}, & y \ge \beta, \\ 0, & \text{elsewhere.} \end{cases}$$

If β is known, show that $\prod_{i=1}^{n} Y_i$ is sufficient for α.

9.45 Suppose that $Y_1, Y_2, \ldots, Y_n$ is a random sample from a probability density function in the (one-parameter) exponential family so that

$$f(y \mid \theta) = \begin{cases} a(\theta)b(y)e^{-[c(\theta)d(y)]}, & a \le y \le b, \\ 0, & \text{elsewhere,} \end{cases}$$

where a and b do not depend on θ. Show that $\sum_{i=1}^{n} d(Y_i)$ is sufficient for θ.

9.46 If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from an exponential distribution with mean β, show that $f(y \mid \beta)$ is in the exponential family and that $\overline{Y}$ is sufficient for β.

9.47 Refer to Exercise 9.43. If θ is known, show that the power family of distributions is in the exponential family. What is a sufficient statistic for α? Does this contradict your answer to Exercise 9.43?

9.48 Refer to Exercise 9.44. If β is known, show that the Pareto distribution is in the exponential family. What is a sufficient statistic for α? Argue that there is no contradiction between your answer to this exercise and the answer you found in Exercise 9.44.

***9.49** Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the uniform distribution over the interval $(0, \theta)$. Show that $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

***9.50** Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the uniform distribution over the interval (θ_1, θ_2). Show that $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ and $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ are jointly sufficient for θ_1 and θ_2.

***9.51** Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the probability density function

$$f(y \mid \theta) = \begin{cases} e^{-(y-\theta)}, & y \ge \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

Show that $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

***9.29** Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample of size n from a power family distribution (see Exercise 6.17). Then the methods of Section 6.7 imply that $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ has the distribution function given by

$$F_{(n)}(y) = \begin{cases} 0, & y < 0, \\ (y/\theta)^{\alpha n}, & 0 \le y \le \theta, \\ 1, & y > \theta. \end{cases}$$

Use the method described in Exercise 9.26 to show that $Y_{(n)}$ is a consistent estimator of θ.

9.30 Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables, each with probability density function

$$f(y) = \begin{cases} 3y^2, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Show that $\overline{Y}$ converges in probability to some constant and find the constant.

9.31 If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a gamma distribution with parameters α and β, show that $\overline{Y}$ converges in probability to some constant and find the constant.

9.32 Let $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the probability density function

$$f(y) = \begin{cases} \dfrac{2}{y^2}, & y \ge 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Does the law of large numbers apply to $\overline{Y}$ in this case? Why or why not?

9.33 An experimenter wishes to compare the numbers of bacteria of types A and B in samples of water. A total of n independent water samples are taken, and counts are made for each sample. Let X_i denote the number of type A bacteria and Y_i denote the number of type B bacteria for sample i. Assume that the two bacteria types are sparsely distributed within a water sample so that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ can be considered independent random samples from Poisson distributions with means λ_1 and λ_2, respectively. Suggest an estimator of $\lambda_1/(\lambda_1+\lambda_2)$. What properties does your estimator have?

9.34 The Rayleigh density function is given by

$$f(y) = \begin{cases} \left(\dfrac{2y}{\theta}\right)e^{-y^2/\theta}, & y > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

In Exercise 6.34(a), you established that Y^2 has an exponential distribution with mean θ. If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from a Rayleigh distribution, show that $W_n = \frac{1}{n}\sum_{i=1}^{n} Y_i^2$ is a consistent estimator for θ.

9.35 Let $Y_1, Y_2, \ldots$ be a sequence of random variables with $E(Y_i) = \mu$ and $V(Y_i) = \sigma_i^2$. Notice that the σ_i^2's are not all equal.

a What is $E(\overline{Y}_n)$?

b What is $V(\overline{Y}_n)$?

c Under what condition (on the σ_i^2's) can Theorem 9.1 be applied to show that $\overline{Y}_n$ is a consistent estimator for μ?

9.36 Suppose that Y has a binomial distribution based on n trials and success probability p. Then $\hat{p}_n = Y/n$ is an unbiased estimator of p. Use Theorem 9.3 to prove that the distribution of speaking, the factorization criterion presented in Section 9.4 can be applied to find sufficient statistics that best summarize the information contained in sample data about parameters of interest. For the distributions that we consider in this text, an MVUE for a target parameter θ can be found as follows. First, determine the best sufficient statistic, U. Then, find a function of U, $h(U)$, such that $E[h(U)] = \theta$.

This method often works well. However, sometimes a best sufficient statistic is a fairly complicated function of the observable random variables in the sample. In cases like these, it may be difficult to find a function of the sufficient statistic that is an unbiased estimator for the target parameter. For this reason, two additional methods of finding estimators—the method of moments and the method of maximum likelihood—are presented in the next two sections. A third important method for estimation, the method of least squares, is the topic of Chapter 11.

Exercises

9.56 Refer to Exercise 9.38(b). Find an MVUE of σ^2.

9.57 Refer to Exercise 9.18. Is the estimator of σ^2 given there an MVUE of σ^2?

9.58 Refer to Exercise 9.40. Use $\sum_{i=1}^{n} Y_i^2$ to find an MVUE of θ.

9.59 The number of breakdowns Y per day for a certain machine is a Poisson random variable with mean λ. The daily cost of repairing these breakdowns is given by $C = 3Y^2$. If $Y_1, Y_2, \dots, Y_n$ denote the observed number of breakdowns for n independently selected days, find an MVUE for $E(C)$.

9.60 Let $Y_1, Y_2, \dots, Y_n$ denote a random sample from the probability density function

$$f(y \mid \theta) = \begin{cases} \theta y^{\theta-1}, & 0 < y < 1,\ \theta > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

a Show that this density function is in the (one-parameter) exponential family and that $\sum_{i=1}^{n} -\ln(Y_i)$ is sufficient for θ. (See Exercise 9.45.)

b If $W_i = -\ln(Y_i)$, show that W_i has an exponential distribution with mean $1/\theta$.

c Use methods similar to those in Example 9.10 to show that $2\theta \sum_{i=1}^{n} W_i$ has a χ^2 distribution with $2n$ df.

d Show that

$$E\left(\frac{1}{2\theta \sum_{i=1}^{n} W_i}\right) = \frac{1}{2(n-1)}.$$

[*Hint:* Recall Exercise 4.112.]

e What is the MVUE for θ?

9.61 Refer to Exercise 9.49. Use $Y_{(n)}$ to find an MVUE of θ. (See Example 9.1.)

9.62 Refer to Exercise 9.51. Find a function of $Y_{(1)}$ that is an MVUE for θ.

9.63 Let $Y_1, Y_2, \dots, Y_n$ be a random sample from a population with density function

$$f(y \mid \theta) = \begin{cases} \dfrac{3y^2}{\theta^3}, & 0 \le y \le \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

In Exercise 9.52 you showed that $Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

a Show that $Y_{(n)}$ has probability density function

$$f_{(n)}(y \mid \theta) = \begin{cases} \dfrac{3ny^{3n-1}}{\theta^{3n}}, & 0 \le y \le \theta, \\ 0, & \text{elsewhere.} \end{cases}$$

b Find the MVUE of θ.

9.64 Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a normal distribution with mean μ and variance 1.

a Show that the MVUE of μ^2 is $\widehat{\mu^2} = \overline{Y}^2 - 1/n$.

b Derive the variance of $\widehat{\mu^2}$.

***9.65** In this exercise, we illustrate the direct use of the Rao–Blackwell theorem. Let $Y_1, Y_2, \ldots, Y_n$ be independent Bernoulli random variables with

$$p(y_i \mid p) = p^{y_i}(1-p)^{1-y_i}, \qquad y_i = 0, 1.$$

That is, $P(Y_i = 1) = p$ and $P(Y_i = 0) = 1 - p$. Find the MVUE of $p(1-p)$, which is a term in the variance of Y_i or $W = \sum_{i=1}^{n} Y_i$, by the following steps.

a Let

$$T = \begin{cases} 1, & \text{if } Y_1 = 1 \text{ and } Y_2 = 0, \\ 0, & \text{otherwise.} \end{cases}$$

Show that $E(T) = p(1-p)$.

b Show that

$$P(T = 1 \mid W = w) = \frac{w(n-w)}{n(n-1)}.$$

c Show that

$$E(T \mid W) = \frac{n}{n-1}\left[\frac{W}{n}\left(1 - \frac{W}{n}\right)\right] = \frac{n}{n-1}\overline{Y}(1-\overline{Y})$$

and hence that $n\overline{Y}(1-\overline{Y})/(n-1)$ is the MVUE of $p(1-p)$.

***9.66** The likelihood function $L(y_1, y_2, \ldots, y_n \mid \theta)$ takes on different values depending on the arguments $(y_1, y_2, \ldots, y_n)$. A method for deriving a *minimal* sufficient statistic developed by Lehmann and Scheffé uses the ratio of the likelihoods evaluated at two points, $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$:

$$\frac{L(x_1, x_2, \ldots, x_n \mid \theta)}{L(y_1, y_2, \ldots, y_n \mid \theta)}.$$

Many times it is possible to find a function $g(x_1, x_2, \ldots, x_n)$ such that this ratio is free of the unknown parameter θ if and only if $g(x_1, x_2, \ldots, x_n) = g(y_1, y_2, \ldots, y_n)$. If such a function g can be found, then $g(Y_1, Y_2, \ldots, Y_n)$ is a minimal sufficient statistic for θ.

a Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a Bernoulli distribution (see Example 9.6 and Exercise 9.65) with p unknown.

i Show that

$$\frac{L(x_1, x_2, \ldots, x_n \mid p)}{L(y_1, y_2, \ldots, y_n \mid p)} = \left(\frac{p}{1-p}\right)^{\Sigma x_i - \Sigma y_i}.$$