Using the factorization criterion, we can show $\sum_{i=1}^{n} Y_i$ and the product $\prod_{i=1}^{n} Y_i$ to be sufficient statistics for the gamma density function. Because the method-of-moments estimators $\hat{\alpha}$ and $\hat{\beta}$ are not functions of these sufficient statistics, we can find more efficient estimators for the parameters α and β. However, it is considerably more difficult to apply other methods to find estimators for these parameters.

To summarize, the method of moments finds estimators of unknown parameters by equating corresponding sample and population moments. The method is easy to employ and provides consistent estimators. However, the estimators derived by this method are often not functions of sufficient statistics. As a result, method-of-moments estimators are sometimes not very efficient. In many cases, the method-of-moments estimators are biased. The primary virtues of this method are its ease of use and that it sometimes yields estimators with reasonable properties.

Exercises

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

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

Find an estimator for θ by the method of moments. Show that the estimator is consistent. Is the estimator a function of the sufficient statistic $-\sum_{i=1}^{n} \ln(Y_i)$ that we can obtain from the factorization criterion? What implications does this have?

9.70 Suppose that $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from a Poisson distribution with mean λ. Find the method-of-moments estimator of λ.

9.71 If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the normal distribution with known mean $\mu = 0$ and unknown variance σ^2, find the method-of-moments estimator of σ^2.

9.72 If $Y_1, Y_2, \ldots, Y_n$ denote a random sample from the normal distribution with mean μ and variance σ^2, find the method-of-moments estimators of μ and σ^2.

9.73 An urn contains θ black balls and $N - \theta$ white balls. A sample of n balls is to be selected without replacement. Let Y denote the number of black balls in the sample. Show that $(N/n)Y$ is the method-of-moments estimator of θ.

9.74 Let $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from the probability density function given by

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

a Find an estimator for θ by using the method of moments.

b Is this estimator a sufficient statistic for θ?

9.75 Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from the probability density function given by

$$f(y \mid \theta) = \begin{cases} \dfrac{\Gamma(2\theta)}{[\Gamma(\theta)]^2}(y^{\theta-1})(1-y)^{\theta-1}, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the method-of-moments estimator for θ.

9.76 Let $X_1, X_2, X_3, \ldots$ be independent Bernoulli random variables such that $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$ for each $i = 1, 2, 3, \ldots$. Let the random variable Y denote the number of trials necessary to obtain the first success—that is, the value of i for which $X_i = 1$ first occurs. Then Y has a geometric distribution with $P(Y = y) = (1 - p)^{y-1}p$, for $y = 1, 2, 3, \ldots$. Find the method-of-moments estimator of p based on this single observation Y.

9.77 Let $Y_1, Y_2, \ldots, Y_n$ denote independent and identically distributed uniform random variables on the interval $(0, 3\theta)$. Derive the method-of-moments estimator for θ.

9.78 Let $Y_1, Y_2, \ldots, Y_n$ denote independent and identically distributed random variables from a power family distribution with parameters α and $\theta = 3$. Then, as in Exercise 9.43, if $\alpha > 0$,

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

Show that $E(Y_1) = 3\alpha/(\alpha + 1)$ and derive the method-of-moments estimator for α.

***9.79** Let $Y_1, Y_2, \ldots, Y_n$ denote independent and identically distributed random variables from a Pareto distribution with parameters α and β, where β is known. Then, if $\alpha > 0$,

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

Show that $E(Y_i) = \alpha\beta/(\alpha - 1)$ if $\alpha > 1$ and $E(Y_i)$ is undefined if $0 < \alpha < 1$. Thus, the method-of-moments estimator for α is undefined.

9.7 The Method of Maximum Likelihood

In Section 9.5, we presented a method for deriving an MVUE for a target parameter: using the factorization criterion together with the Rao–Blackwell theorem. The method requires that we find some function of a minimal sufficient statistic that is an unbiased estimator for the target parameter. Although we have a method for finding a sufficient statistic, the determination of the function of the minimal sufficient statistic that gives us an unbiased estimator can be largely a matter of hit or miss. Section 9.6 contained a discussion of the method of moments. The method of moments is intuitive and easy to apply but does not usually lead to the best estimators. In this section, we present the method of maximum likelihood that often leads to MVUEs.

We use an example to illustrate the logic upon which the method of maximum likelihood is based. Suppose that we are confronted with a box that contains three balls. We know that each of the balls may be red or white, but we do not know the total number of either color. However, we are allowed to randomly sample two of the balls without replacement. If our random sample yields two red balls, what would be a good estimate of the total number of red balls in the box? Obviously, the number of red balls in the box must be two or three (if there were zero or one red ball in the box, it would be impossible to obtain two red balls when sampling without replacement). If there are two red balls and one white ball in the box, the probability of randomly selecting two red balls is

$$\frac{\binom{2}{2}\binom{1}{0}}{\binom{3}{2}} = \frac{1}{3}.$$

***9.93** Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a population with density function

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

In Exercise 9.53, you showed that $Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ is sufficient for θ.

a Find the MLE for θ. [*Hint:* See Example 9.16.]

b Find a function of the MLE in part (a) that is a pivotal quantity.

c Use the pivotal quantity from part (b) to find a $100(1-\alpha)\%$ confidence interval for θ.

***9.94** Suppose that $\hat{\theta}$ is the MLE for a parameter θ. Let $t(\theta)$ be a function of θ that possesses a unique inverse [that is, if $\beta = t(\theta)$, then $\theta = t^{-1}(\beta)$]. Show that $t(\hat{\theta})$ is the MLE of $t(\theta)$.

***9.95** A random sample of n items is selected from the large number of items produced by a certain production line in one day. Find the MLE of the ratio R, the proportion of defective items divided by the proportion of good items.

9.96 Consider a random sample of size n from a normal population with mean μ and variance σ^2, both unknown. Derive the MLE of σ.

9.97 The geometric probability mass function is given by

$$p(y \mid p) = p(1-p)^{y-1}, \qquad y = 1, 2, 3, \ldots.$$

A random sample of size n is taken from a population with a geometric distribution.

a Find the method-of-moments estimator for p.

b Find the MLE for p.

9.8 Some Large-Sample Properties of Maximum-Likelihood Estimators (Optional)

Maximum-likelihood estimators also have interesting large-sample properties. Suppose that $t(\theta)$ is a differentiable function of θ. In Section 9.7, we argued by the invariance property that if $\hat{\theta}$ is the MLE of θ, then the MLE of $t(\theta)$ is given by $t(\hat{\theta})$. Under some conditions of regularity that hold for the distributions that we will consider, $t(\hat{\theta})$ is a *consistent* estimator for $t(\theta)$. In addition, for large sample sizes,

$$Z = \frac{t(\hat{\theta}) - t(\theta)}{\sqrt{\left[\dfrac{\partial t(\theta)}{\partial \theta}\right]^2 \Big/ nE\left[-\dfrac{\partial^2 \ln f(Y \mid \theta)}{\partial \theta^2}\right]}}$$

has approximately a standard normal distribution. In this expression, the quantity $f(Y \mid \theta)$ in the denominator is the density function corresponding to the continuous distribution of interest, evaluated at the random value Y. In the discrete case, the analogous result holds with the probability function evaluated at the random value Y, $p(Y \mid \theta)$ substituted for the density $f(Y \mid \theta)$. If we desire a confidence interval for $t(\theta)$, we can use quantity Z as a pivotal quantity. If we proceed as in Section 8.6, we obtain balance between the risks of type I and type II errors, but both α and β remain disconcertingly large. *How can we reduce both α and β?* The answer is intuitively clear: Shed more light on the true nature of the population by increasing the sample size. For almost all statistical tests, if α is fixed at some acceptably small value, β decreases as the sample size increases. ■

In this section, we have defined the essential elements of any statistical test. We have seen that two possible types of error can be made when testing hypotheses: type I and type II errors. The probabilities of these errors serve as criteria for evaluating a testing procedure. In the next few sections, we will use the sampling distributions derived in Chapter 7 to develop methods for testing hypotheses about parameters of frequent practical interest.

Exercises

10.1 Define α and β for a statistical test of hypotheses.

10.2 An experimenter has prepared a drug dosage level that she claims will induce sleep for 80% of people suffering from insomnia. After examining the dosage, we feel that her claims regarding the effectiveness of the dosage are inflated. In an attempt to disprove her claim, we administer her prescribed dosage to 20 insomniacs and we observe Y, the number for whom the drug dose induces sleep. We wish to test the hypothesis $H_0: p = .8$ versus the alternative, $H_a: p < .8$. Assume that the rejection region $\{y \le 12\}$ is used.

- **a** In terms of this problem, what is a type I error?
- **b** Find α.
- **c** In terms of this problem, what is a type II error?
- **d** Find β when $p = .6$.
- **e** Find β when $p = .4$.

10.3 Refer to Exercise 10.2.

- **a** Find the rejection region of the form $\{y \le c\}$ so that $\alpha \approx .01$.
- **b** For the rejection region in part (a), find β when $p = .6$.
- **c** For the rejection region in part (a), find β when $p = .4$.

10.4 Suppose that we wish to test the null hypothesis H_0 that the proportion p of ledger sheets with errors is equal to .05 versus the alternative H_a, that the proportion is larger than .05, by using the following scheme. Two ledger sheets are selected at random. If both are error free, we reject H_0. If one or more contains an error, we look at a third sheet. If the third sheet is error free, we reject H_0. In all other cases, we accept H_0.

- **a** In terms of this problem, what is a type I error?
- **b** What is the value of α associated with this test?
- **c** In terms of this problem, what is a type II error?
- **d** Calculate $\beta = P(\text{type II error})$ as a function of p.