respectively. We then have the following probabilities:

$$Pr\{A_1\} = Pr\{A_2\} = Pr\{A_3\} = \tfrac{1}{2}$$

$$Pr\{A_1A_2\} = Pr\{A_2A_3\} = Pr\{A_1A_3\} = \tfrac{1}{4}$$

$$= Pr\{A_1\}Pr\{A_2\} = Pr\{A_2\}Pr\{A_3\} = Pr\{A_1\}Pr\{A_3\}.$$

But

$$Pr\{A_1A_2A_3\} = \tfrac{1}{4} \neq Pr\{A_1\}Pr\{A_2\}Pr\{A_3\}.$$

Therefore, the three events A_1, A_2, A_3 are not mutually independent, but any two of them are.

Example 2. The same result is obtained as in Example 1 if one follows this procedure: First, write each combination of $n - 1$ symbols on different balls, one combination per ball; then write each combination of $n - 3$ symbols on different balls, one combination per ball; etc. If n is odd we end up writing each combination of two symbols on different balls, one combination per ball, and then we take one more ball and leave it blank; if n is even, we stop after writing each single symbol on different balls, one symbol per ball.

REFERENCES

1. Feller, W. *An Introduction to Probability Theory and Its Applications*, Vol. I, Second Edition, John Wiley, New York, 1957.
2. Tucker, H. *A Graduate Course in Probability*, Academic Press, New York, 1967.

Completeness and Unbiased Estimation

STEPHEN M. STIGLER
University of Wisconsin, Madison

Since Lehmann and Scheffé [1] introduced it in 1950, the concept of a *complete* family of probability distributions has been widely applied in estimation problems and has become a standard topic in mathematical statistics courses. Completeness is well-suited for treatment in such courses: the definition is easily stated and the concept can be immediately applied (through the Rao-Blackwell and Lehmann-Scheffé theorems) to find best unbiased estimators in a wide variety of settings. However, while a student encountering completeness for the first time is very likely to appreciate its usefulness, he is just as likely to be puzzled by its name, and wonder what connection (if any) there is between the statistical use of the term "complete" and the dictionary definition: lacking none of the parts, whole, entire. Inspection of a number of the current texts ([2]–[9]) which define "complete" reveals that none of them comments on the reason for the choice of terminology. (One might say they avoid this aspect of the subject completely). The purpose of this note is to present a simple example which illuminates the motivation behind the use of the word "complete," and which in addition highlights one facet of the definition of "unbiased" often overlooked in introductory courses.

We begin with the usual definition:

Definition:

Let $\mathcal{P} = \{P_\theta : \theta \in \mathrm{A}\}$ be a family of probability distributions of a random variable (or statistic) X, indexed by the parameter set A. We say $\mathcal{P}$ is *complete* if for any function ϕ satisfying

$$E_\theta[\phi(X)] = 0 \quad \text{for all} \quad \theta \in \mathrm{A}, \tag{1}$$

it must also be true that $\phi(x) = 0$ for all x (except possibly for a set of x having probability zero for all $\theta \in \mathrm{A}$). (E_θ denotes expectation with respect to the distribution P_θ).

A common verbalization of this definition is that $\mathcal{P}$ is complete if there exist no unbiased estimators of zero except for the trivial estimator $\phi(x) \equiv 0$.

The intuitive motivation for the choice of the word "complete" can be expressed as follows: Requiring that the function ϕ satisfy condition (1) above represents a restriction on ϕ. The larger the family $\mathcal{P}$ (or the larger A), the greater the restriction on ϕ. When the family $\mathcal{P}$ is augmented to the point that the condition (1) rules out all ϕ except the trivial $\phi(x) \equiv 0$, $\mathcal{P}$ is said to be complete.

As an example of what is meant, consider the family of distributions $\mathcal{P} = \{P_N : N \geqq 1\}$, where we define

$$P_N\{X = k\} = \begin{cases} 1/N & \text{for} \quad k = 1, 2, \ldots, N \\ 0 & \qquad \text{otherwise.} \end{cases}$$

It is well-known (and easily shown-for example, by induction) that in order that $E_N[\phi(X)] = 0$ for all $N \geqq 1$, we must have $\phi(k) = 0$ for $k = 1, 2, \ldots$. In other words, $\mathcal{P}$ is complete. However, $\mathcal{P}$ is just barely complete. For if even one of the infinity of probability distributions in $\mathcal{P}$ is removed, the resulting family is incomplete. That is, for any integer $n \geq 1$, the family $\mathcal{P} - \{P_n\}$ is *not* complete.

To see that this is so, consider the function ϕ_0 given by

$$\phi_0(k) = \begin{cases} 0 & \text{for} \quad k = 1, 2, \ldots, n-1, \\ & \qquad\quad n+2, n+3, \ldots \\ a & \text{for} \quad k = n \\ -a & \text{for} \quad k = n+1, \end{cases} \tag{2}$$

where a is any nonzero constant. It is easily seen that

$$E_N[\phi_0(X)] = \begin{cases} 0 & \text{for } N \neq n \\ a/N & \text{for } N = n, \end{cases}$$

and thus ϕ_0 is an unbiased estimator of zero *for the family* $\mathcal{P} - \{P_n\}$. (The *only* nontrivial unbiased estimators of zero are given by (2) for various choices of a). This serves as a simple example that completeness is a property of the family of distributions rather than the random variable or the parametric form, that the statistical definition of "complete" *is* related to every-day usage, and that removing even one point from a parameter set may alter the completeness of the family.

As a sidelight to this example, the Lehmann-Scheffé theorem tells us that since $\mathcal{P}$ is complete, the unbiased estimator $\phi_1(X) = 2X - 1$ is the minimum variance unbiased estimator of N for the family $\mathcal{P}$. But since $\mathcal{P} - \{P_n\}$ is not complete, the theorem cannot be applied to this family, and in fact ϕ_1 is *not* the minimum variance unbiased estimator of N for the family $\mathcal{P} - \{P_n\}$. The minimum variance unbiased estimator for N for the family $\mathcal{P} - \{P_n\}$ is ϕ_2 given by

$$\phi_2(k) = \begin{cases} 2k - 1 & \text{for } k \neq n, k \neq n + 1 \\ 2n & \text{for } k = n, n + 1. \end{cases}$$

That this is so can either be shown directly or by using a generalization of the Lehmann-Scheffé theorems (see Rao [10], p. 257) which states that a necessary and sufficient condition that ϕ_2 have minimum variance is that it be uncorrelated with all unbiased estimators of zero (i.e., those given by (2)). A simple calculation shows that

$$\text{var}_N[\phi_2(X)] = \begin{cases} \text{var}_N[\phi_1(X)] & \text{for } N < n \\ \text{var}_N[\phi_1(X)] - 2/N & \text{for } N > n. \end{cases}$$

It is interesting to note that while ϕ_2 has minimum variance among all unbiased estimators of N for the family $\mathcal{P} - \{P_n\}$, it is not unbiased for the family $\mathcal{P}$, since $E_n[\phi_2(X)] = n + 1/n$. This points up the fact that unbiasedness, like completeness, is neither a property of the statistic nor the parametric form, but rather a property of the family of distributions of the statistic.

REFERENCES

[1] Lehmann, E. L., and Scheffé, H., "Completeness, similar regions, and unbiased estimation," *Sankhyā* **10** (1950), pp. 305–340, and **15** pp. 219–236.
[2] Bickel, P. J., *Mathematical Statistics* (Preliminary Edition). San Francisco: Holden-Day, 1971.
[3] Hogg, R. V., and Craig, A. T., *Introduction to Mathematical Statistics* (Third Edition). London: Macmillan, 1970.
[4] Lehmann, E. L., *Testing Statistical Hypotheses*. New York: Wiley, 1959.
[5] Mood, A. M., and Graybill, F. A., *Introduction to the Theory of Statistics* (Second Edition). New York: McGraw-Hill, 1963.
[6] Roussas, G. G., *A First Course in Mathematical Statistics*. Reading: Addison-Wesley, 1972.
[7] Savage, I. R., *Statistics: Uncertainty and Behavior*. Boston: Houghton Mifflin, 1968.
[8] Zacks, S., *The Theory of Statistical Inference*. New York: Wiley, 1971.
[9] Ferguson, T. S., *Mathematical Statistics: A Decision Theoretic Approach*. New York: Academic Press, 1967.
[10] Rao, C. R., *Linear Statistical Inference and its Applications*. New York: Wiley, 1965.

A Property of Certain Distributions

NATHAN MANTEL
Biometry Branch
National Cancer Institute

Following a letter by Gross [1969] describing a class of discrete distributions having no moments or only limited moments, Lachenbruch and Brogan [1971] gave a parallel class of continuous distributions, related to ratios of exponential variables, having similar properties.

Except for scale factors, the distributions given by Lachenbruch and Brogan can be recognized as those of Snedecor's F. The limited existence of moments for F is known [see Wilks, 1962, p. 187] so that the r'th moment of $F(k_1, k_2)$, where k_1 and k_2 designate numerator and denominator degrees of freedom (d.f.) respectively, exists only for $-k_1 < 2r < k_2$. For Lachenbruch and Brogan's $\theta_2 = k_2/2 = 1$ no moments exist, while moments only up to $\theta_2 - 1$ exist for higher values of θ_2. The standard example of the Cauchy distribution as one having no moments follows from the results for the F distribution if we take the liberty of interpreting the first moment of the Cauchy as the half-moment of $F(1, 1)$. (See also Mantel [1969], who effectively gives some of Lachenbruch and Brogan's distributions in c.d.f. form.)

Other interesting properties than those of existence or non-existence of moments attend distributions related to the normal, chi squared, and Snedecor's F. The normal distribution, $N(0, \sigma^2)$, has the property that the average of n independent observations is distributed like $1/(n)^{1/2}$ times a single observation while for the Cauchy (which is the ratio of 2 independent normal $N(0, 1)$ observations) the average is distributed like a single observation; the corresponding totals for n observations are distributed like $(n)^{1/2}$ and n times a single observation respectively.

An exercise brought to my attention by Professor D. F. Kerridge has indicated the existence of a distribution for which an average of n independent observa-