

Confidence Intervals from Single Observations in Forest Research

Harry T. Valentine, George M. Furnival, and Timothy G. Gregoire

ABSTRACT. A procedure for constructing confidence intervals and testing hypotheses from a single trial or observation is reviewed. The procedure requires a prior, fixed estimate or guess of the outcome of an experiment or sampling. Two examples of applications are described: a confidence interval is constructed for the expected outcome of a systematic sampling of a forested tract, and a hypothesis is tested in connection with a watershed experiment. Potential misuses of the procedure also are discussed. *FOR. SCI.* 37(1):370-373.

ADDITIONAL KEY WORDS. Hypothesis, significance test, systematic sampling, watershed experiment.

SCIENTIFIC EXPERIMENTS OR SAMPLINGS of large scope or expense sometimes consist of a single unreplicated trial or observation. NASA's experiment to determine whether microbial life exists on Mars was a notable example. On Earth, unreplicated trials are commonplace in industrial research and development. Typically, a single prototype is constructed or a single batch is run and the outcome evaluated. If the results in some sense meet expectations, a new model or a new process is one step closer to production. If not, theory is re-examined, plans are revised, and a different approach is attempted. Single unreplicated trials also have been reported in the annals of forest research. The classic example is the watershed experiment, where one watershed receives a treatment and the outcome is compared to that of a nearby untreated watershed (e.g., Bormann and Likens 1979, Hornbeck et al. 1986). Other common examples include unreplicated paired-tree experiments and unreplicated systematic samplings of forest inventories. Pinney (1989) has advocated single-tree investigations in agroforestry research as alternatives to more costly experiments.

Statisticians and journal editors, among others, often give short shrift to single-observation experiments or samplings. Many would argue that a single observation provides an inadequate basis for inference. Except for the simplest one-parameter models, inferential procedures did not exist until when Abbott and Rosenblatt (1963) proved that confidence intervals for the midpoint of a distribution could be constructed from a single observation. Recently, Furnival et al. (1989) reported an improved procedure that furnishes shorter intervals. The procedure requires an estimate or educated guess of the outcome of the unreplicated trial or observation that is both independent and fixed in advance of the actual experiment or sampling. The widths of the intervals depend upon how much is known about the distribution of the population of observations. The widest intervals eventuate when nothing is known, and we assume that the distribution is continuous, symmetric, and unimodal. Relatively narrower confidence intervals are calculable for a known continuous distribution, such as the normal, which can be standardized by the transformation:

$$z = \frac{y - \mu}{\sigma} \quad (1)$$

where z , y , μ , and σ , respectively, are the standardized variate, variate, mean, and standard deviation.

In this note, we demonstrate how to construct confidence intervals and test hypotheses with a single observation from either an unknown or normal distribution, using examples from forestry research. We then discuss some of the problems that arise in application.

PROCEDURE

CONFIDENCE INTERVALS

Let $\hat{\mu}$ denote an estimate of the outcome of an unreplicated trial or observation that is fixed in advance of the actual experiment or sampling, and let y denote the actual outcome. Our

objective is a confidence interval for the mean (μ) of all such trials or observations. Under the assumption that y derives from a continuous, symmetric, unimodal distribution with mean μ , Furnival et al. (1989) proved that the interval:

$$\frac{y + \bar{\mu}}{2} \pm k|y - \bar{\mu}| \quad (2)$$

will contain μ with probability of at least $1 - \alpha$ where

$$k = \frac{1 - \alpha + \sqrt{1 - 2\alpha}}{2\alpha} \quad 0 < \alpha \leq 0.5 \quad (3)$$

For $k = \alpha = 0.5$, the interval also will hold for a unique median for any distribution, whether continuous, symmetric, and unimodal, or not.

If y is known or assumed to derive from a normal distribution, then the appropriate value of k is smaller than that furnished by (3), and the resultant confidence interval is narrower. There is no neat formula that furnishes k for a given α when y is normally distributed, but k must satisfy (Furnival et al. 1989):

$$\alpha = \Phi(rp) - \Phi(p) \quad (4)$$

where

$$r = \left(\frac{k + 0.5}{k - 0.5} \right)^2 \quad k > 0.5 \quad (5)$$

$$p = \sqrt{\frac{2 \cdot \ln(r)}{r^2 - 1}} \quad (6)$$

and $\Phi(\cdot)$ is the standard normal distribution function. Values of k for selected values of α are given in Table 1 for both normal and unknown symmetric, unimodal distributions.

TESTS OF HYPOTHESES

The confidence interval, (2), can be used in the conventional two-tailed test of a hypothesis $H_0: \mu = \mu_0$ against the alternative $H_1: \mu \neq \mu_0$. H_0 is accepted if the confidence interval includes μ_0 , but not otherwise. A one-tailed test is appropriate for a null hypothesis such as $H_0: \mu < \mu_0$ against the alternative $H_1: \mu \geq \mu_0$. H_0 is accepted if:

$$\mu_0 > \frac{y + \bar{\mu}}{2} + k|y - \bar{\mu}|. \quad (7)$$

TABLE 1.

Values of k for selected levels of α when a single experimental or sample observation derives from either a normal or unknown, symmetric, unimodal distribution.

	α						
Distribution	0.50	1/3	0.25	0.20	0.10	0.05	0.01
Normal	0.50	1.26	1.80	2.31	4.79	9.66	48.39
Unknown	0.50	1.87	2.91	3.94	8.97	18.99	99.00

A null hypothesis $H_0: \mu > \mu_0$ is accepted if:

$$\mu_0 < \frac{y + \bar{\mu}}{2} - k|y - \bar{\mu}|. \quad (8)$$

The significance levels for a given k are the same for both one- and two-tailed tests.

EXAMPLES

SYSTEMATIC SAMPLING

Watershed 6, a 13-ha tract on the Hubbard Brook Experimental Forest, West Thornton, NH, is divided into 208 permanent, monumented 25 m $\times$ 25 m plots. In 1982, a 100% inventory of the 67-year-old tree crop on these plots was conducted, providing us with the means of simulating a systematic sampling of the volume of trees on the watershed. As a prelude to the simulated sampling, we obtained a predicted volume of the tree crop (55.8 m³/ha) from yield table 8 of Solomon and Leak (1986), which we used as our prior estimate $\bar{\mu}$. For our systematic sampling, we divided the tract into a stack of 5 blocks running from the bottom to the top of the watershed. The sample of each block consisted of the middle strip (row) of plots within the block. Thus, our sampling simulated a conventional forestry strip cruise, where the strips consisted of rows of plots (32 plots in all) spaced equidistantly and perpendicular to the elevational gradient of the watershed. The resultant estimate of the volume of the tree crop volume was $y = 54.0$ m³/ha.

We assume that y derives from a normal distribution. To construct a 75% confidence interval for the true volume of the tract, we use $k = 1.80$ from Table 1 corresponding to $\alpha = 0.25$. Application of formula (2) obtains:

$$\frac{54.0 + 55.8}{2} \pm 1.80 \cdot |55.8 - 54.0|$$

or 54.9 ± 3.3 m³/ha. This interval includes the true volume, 57.3 m³/ha. Alternative analyses for unreplicated systematic samplings can be found in the forestry literature and in standard sampling texts (see, e.g., Matérn 1960; Cochran 1963).

WATERSHED EXPERIMENT

An experimental strip cutting was performed on Watershed 4 of the Hubbard Brook Experimental Forest to observe changes in water yield and quality (Hornbeck et al. 1986). One-third of the tree crop was effectively clearcut in 25-m-wide strips spaced equidistantly and perpendicular to the elevational gradient of the watershed. Stream flow in the year following the strip cutting was expected to increase by 90 mm based on published reports of other clearcutting experiments at different locations. However, the experiment netted an increase of only 22 mm.

Now, suppose that the experimenters had fixed $\bar{\mu} = 90$ mm as their prior estimate of the outcome of the experiment, and had wished to test the null hypothesis that water yield is unchanged for a span of a year following a strip cutting ($H_0: \mu = 0$ mm) against the alternative hypothesis that water yield is increased ($H_1: \mu > 0$ mm). The outcome $y = 22$ mm derives from an unknown distribution, so to construct a 75% confidence interval for μ , we take $k = 2.91$ from Table 1. Applying equation (2) we obtain the very wide interval 56.0 ± 197.9 mm that includes 0 mm. Therefore, we find no convincing evidence that an increase in water yield, as found in this single experiment, should be expected in general following a 33% strip cutting of a forested watershed.

DISCUSSION

Our confidence interval, (2), is centered midway between the prior estimate of the outcome, $\bar{\mu}$, and the actual outcome, y , of an experiment or sampling. The width of the interval for a given α is proportional to the absolute difference between $\bar{\mu}$ and y so a prior estimate that is a wild guess could yield a very wide interval. The confidence interval also is affected by the choice of α . The narrowest interval results where $\alpha = 0.5$; y is either the left or

right bound of the interval, and $\bar{\mu}$ is the other bound. However, a probability of 0.5 is too low to be of any real significance. On the other hand, small values of α (e.g., 0.05, 0.01) produce intervals that tend to be too wide to be of any practical value (i.e., the power of a test is too weak). Therefore, when one has only a single observation to work with, an α of 0.25 or 0.2 is perhaps the best compromise. Previously, Gabriel (1964) advocated larger α values than the conventional 0.05 for the practical problems faced by experimenters.

As was noted, the prior estimate, $\bar{\mu}$, should be both fixed (e.g., chiseled in stone) and independent of the experiment or sampling. Changing $\bar{\mu}$ after the outcome would render the confidence interval meaningless. An investigator also must be careful to avoid an inadvertent transgression of the requirement of independence. For example, foresters often remeasure the same plots over time to track standing crop in a forest or forested region. Suppose that an estimate of volume/unit area based on measurements of one or more plots at time 1 is projected to time 2 with a yield model. At time 2, to save money, it is decided that a sample of one plot will be measured. If this plot is selected at random from the forest, then the projection is an independent, prior estimate of the volume/unit area of the forest. However, because of within-plot serial correlation, the projection would not be an independent prior estimate if the plot measured at time 2 were restricted to one of those measured at time 1. By contrast, a restricted remeasurement of one of the original plots would be permissible if growth, rather than volume, were the parameter of interest. But, of course, restricted remeasurement of growth in a plot from time 2 to time 3 should be avoided if the prior estimate of growth is a projection based wholly or partly on the growth from time 1 to time 2 in that same plot.

In conclusion, we generally do not encourage experiments or samplings consisting of a single trial or observation. However, we have decided to advance our procedure because we recognize that sometimes replication is impractical or impossible, and the information to be gleaned from a single trial or observation could be important, if not critical. Indeed, mankind, wittingly or unwittingly, is currently conducting any number of unreplicated experiments on a global scale.

LITERATURE CITED

- ABBOTT, J.H., and J.I. ROSENBLATT. 1963. Two stage estimation with one observation on the first stage. *Ann. Inst. Stat. Math. (Tokyo)*. 4:229-235.
- BORMANN, F.H., and G.E. LIKENS. 1979. *Pattern and process in a forested ecosystem*. Springer-Verlag, New York. 253 p.
- COCHRAN, W.G. 1963. *Sampling techniques*. Wiley, New York. 413 p.
- FURNIVAL, G.M., T.G. GREGOIRE, and H.T. VALENTINE. 1989. Confidence intervals and significance tests for a single trial. *Comm. Stat. Theory Meth.* 18(10):3749-3761.
- GABRIEL, K.R. 1964. A procedure for testing the homogeneity of all sets of means in analysis of variance. *Biometrics* 20:459-477.
- HORNBECK, J.W., ET AL. 1986. Clearcutting northern hardwoods: effects on hydrologic and nutrient ion budgets. *For. Sci.* 32:667-686.
- MATERN, B. 1960. Spatial variation. *Medd. fr. Statens Skogsforskningsinstitut* 49(5):1-144.
- PINNEY, A. 1989. Studying the single tree. *Agroforestry today* 1(3):4-6.
- SOLOMON, D.S., and W.B. LEAK. 1986. Simulated yields for managed northern stands. *USDA For. Serv. Res. Pap. NE-578*. 24 p.

Copyright 1991 by the Society of American Foresters
Manuscript received January 22, 1990.

AUTHORS AND ACKNOWLEDGMENTS

Harry T. Valentine is Research Forester, USDA Forest Service, Northeastern Forest Experiment Station, P.O. Box 640, Durham, NH 03824-0640; George M. Furnival is J.P. Weyerhaeuser, Jr., Professor of Forest Management, Yale School of Forestry and Environmental Studies, New Haven, CT 06511; and Timothy G. Gregoire is Assistant Professor of Forest Biometrics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.