The Great Lakes Entomologist

Volume 15
Number 2 - Summer 1982 *Number 2 - Summer 1982*

Article 5

June 1982

Accuracy and Precision of Insect Density and Impact Estimates

Gary W. Fowler
University of Michigan

John A. Witter
University of Michigan

ACCURACY AND PRECISION OF INSECT DENSITY AND IMPACT ESTIMATES

Gary W. Fowler and John A. Witter[1]

ABSTRACT

In estimating insect density and impact, entomologists are understandably interested in accuracy of estimation, but they almost always are dealing with precision because of bias due to an invalid estimator, probability sampling, or nonsampling errors. Definitions related to statistical estimation are reviewed and the concepts of accuracy and precision examined. Interval estimation and optimum sample size determination related to accuracy and precision, using the concept of allowable error, are examined. Criteria for selecting the best estimator in terms of accuracy and precision are presented, and the distortion of probability statements due to bias is discussed. Accuracy and precision are compared and contrasted using two examples: (1) estimating insect density and (2) estimating insect impact. Adjusted and more accurate estimators can be obtained if the bias of an estimator can be estimated from a preliminary sample.

In using statistical procedures to estimate insect densities and impact, entomologists must decide whether to emphasize the accuracy or the precision of their estimates when referring to the error of estimation. The terms accuracy and precision are not synonymous in most situations, and there is confusion among some entomologists as to whether they are working with accuracy or precision. Interpretation of point and interval estimates and the determination of optimum sample size depend on which concept is being used.

The objectives of this paper are to (1) review some definitions related to estimation, (2) compare and contrast accuracy and precision and present the case for using the term precision instead of accuracy, (3) examine interval estimation and the determination of optimum sample size, (4) compare and contrast unbiased and biased estimators, and (5) examine two case studies: (A) estimating insect density and (B) estimating insect impact.

Most entomologists are familiar with the definitions given in Table 1 (p. 114), which is included for those who might find a quick review useful.

ACCURACY AND PRECISION

The precision of an estimator refers to the magnitude of the deviations of sample estimates ($\bar{x}$) from their mean or expected value $E(\bar{x})$ and is usually measured by the variance of the estimator $Var(\bar{x})$ or its standard error $\sqrt{Var(\bar{x})}$. The actual sampling error, or precision, of an estimate, $\bar{x} - E(\bar{x})$, is almost always unknown. The sample standard error $s_{\bar{x}}$ is an estimate of the average precision of the estimator $\bar{x}$.

The accuracy of an estimator refers to the magnitude of the deviations of the sample estimates ($\bar{x}$) from the population parameter μ being estimated and is usually measured by the mean square error of the estimator $MSE(\bar{x})$ or its standard error $\sqrt{MSE(\bar{x})}$ (Cochran 1977, Lindgren 1962, Raj 1968).

[1]Forest Resources Program, School of Natural Resources, The University of Michigan, Ann Arbor, MI 48109.

The true sampling error, or accuracy, of an estimate, $\bar{x} - \mu$, is almost always unknown. If an estimator is unbiased in probability sampling, valid for the situation, and nonsampling errors are not present, $E(\bar{x}) = \mu$, $MSE(\bar{x}) = Var(\bar{x})$, and accuracy and precision are identical. In this case, $s_{\bar{x}}$ is an estimate of the average accuracy of the estimator $\bar{x}$. However, if an estimator is biased in probability sampling, invalid for the situation, or nonsampling errors are present, an estimate of the accuracy of the estimator $\bar{x}$ cannot be obtained because the magnitude of the bias is usually unknown. All that can be calculated is $s_{\bar{x}}$ which is only a measure of precision.

Entomologists estimating insect density and impact should emphasize the precision and not the accuracy of their estimates. We agree that it is accuracy that entomologists are usually interested in, but it is only precision that they will be able to measure in most cases due to one or more of three types or sources of bias: (1) bias in probability sampling, (2) bias due to use of an estimator invalid for the situation, and (3) bias caused by nonsampling errors.

Biased estimates will be obtained from estimators that are biased in probability sampling (i.e., the mean or expected value of the estimator is not equal to the population parameter being estimated). Regression and ratio estimators are biased estimators, and the biases associated with these estimators can be minimized by taking large samples.

Estimators can yield biased estimates when the estimator is invalid for the situation. Examples of such estimators include (A) the use of simple random sampling estimators, based on equal probabilities of selection, when sampling with probabilities proportional to the size of the sampling unit (p.p.s.) should be used; (B) the use of equal selection probabilities at one or more stages of a multi-stage sampling estimator when unequal selection probabilities should be used, and (C) the use of simple random sampling estimators when data have been collected using systematic sampling. The biases associated with such estimators can be alleviated in many cases by using the correct probabilities of selection.

Estimators that are unbiased in probability sampling and valid for the situation can, and almost always do, produce biased estimates due to one or more nonsampling errors. Nonsampling errors include all errors arising in the course of collecting and processing the data. Some of the more common nonsampling errors associated with insect density and impact estimates are listed in Table 2 (p. 116). Nonsampling errors can be minimized by (1) being aware of such errors, (2) careful organization of the sampling process, (3) hiring the best personnel for the job, (4) thoroughly training personnel, and (5) frequently checking all phases of the sampling process with error checking routines (e.g., using a facsimile tally sheet of the previous year's data as a check in the field when collecting next year's data). Even though such errors can be minimized, they cannot be eliminated entirely.

For the above reasons, entomologists should emphasize the precision of their estimates. Accuracy can only be referred to when estimates are unbiased or the bias associated with biased estimates is known. This is hardly ever possible in real-world situations.

INTERVAL ESTIMATION AND OPTIMUM SAMPLE SIZE

It is almost never known how close a point estimate $\bar{x}$ is to the population mean μ. The goodness or reliability of an estimate is usually determined by a normal-based confidence interval or a distribution-free error bound. Such interval estimates are obtained from an interval estimator.

An interval estimator is a rule for constructing an interval estimate from sample data so that a confidence statement can be made about the population parameter being estimated. The upper and lower bounds of the interval estimate are called confidence or error bound limits. The confidence coefficient, usually expressed in percent, indicates how many times in the long run such intervals, based on the same sample size and identical sampling procedures, will include the population parameter being estimated.

Interval Estimation. Most entomologists use two types of $(1 - \alpha)\%$ confidence intervals for μ based on a simple random sample from a population that has a normal distribution:

(1) σ^2 known: $\bar{x} \pm Z_{\alpha/2}\sigma_{\bar{x}}$
(2) σ^2 unknown: $\bar{x} \pm t_{\alpha/2,\, n-1}s_{\bar{x}}$

where σ^2 is the population variance, $Z_{\alpha/2}$ and $t_{\alpha,\ n-1}$ are the upper critical values of the standard normal distribution and Student's t-distribution with n − 1 degrees of freedom, respectively, and α is the level of significance.

Some entomologists may not be familiar with the 75% error bound for μ based on a simple random sample with no assumption made about the population distribution:

(3) $\bar{x} \pm 2s_{\bar{x}}$.

For the (1 − α)% confidence intervals, one states with (1 − α)% confidence that the true population mean is within the calculated interval if the underlying assumptions are true. For the 75% error bound based on Tchebysheff's Theorem (Fowler and Hauke 1979, Hogg and Craig 1965, Scheaffer et al. 1979), one states with at least 75% confidence that the true population mean is within the calculated interval regardless of the underlying distribution of X. All three procedures assume a finite mean and variance.

The 75% error bound is based on the assumption that the population variance σ^2 is known. Since equation (3) is based on $s_{\bar{x}}$, the error bound coefficient is approximate. This approximation is very good if the sample size is greater than 10. Seventy-five percent is a lower bound on the confidence coefficient. Most distributions would yield an actual confidence coefficient between 80 and 90%. If the population were normally distributed, the 75% error bound has an actual confidence coefficient equal to approximately 95% because $Z_{0.025}$ = 1.96 is very close to 2.

We strongly prefer the 75% error bound approach when the population distribution is unknown or not normal and the sample size is relatively small. It puts a lower bound on the confidence coefficient regardless of the population distribution while at the same time says that the confidence coefficient is approximately 95% if the distribution is normal. Also, the 75% error bound is simpler to calculate than either of the (1 − α)% confidence intervals. The two 95% confidence intervals and the 75% error bound yield essentially the same results in terms of the width of the interval estimate and confidence coefficient for larger sample sizes, due to the Central Limit Theorem (Hogg and Craig 1965).

All three interval estimators yield unbiased interval estimates if based on observations from a simple random sample and none of the three types of bias are present. The confidence coefficients assume unbiased interval estimates when inferences are to be made about the population mean μ. The actual confidence coefficients related to estimation of μ are unknown if interval estimates are biased. Thus, we refer to the accuracy of $\bar{x}$ in estimating μ if interval estimates are unbiased; however, if interval estimates are biased, we refer to the accuracy of $\bar{x}$ in estimating $E(\bar{x}) = \mu + B$.

The goodness (accuracy or precision) of the estimate $\bar{x}$ of μ can be measured by the half-width, d, of the confidence interval ($Z_{\alpha/2}\sigma_{\bar{x}}$ or $t_{\alpha/2,n-1}s_{\bar{x}}$) or the 75% error bound ($2s_{\bar{x}}$); d is called the allowable error AE. The allowable error percent AE% = (d/$\bar{x}$) · 100 is the most common measure of the goodness of $\bar{x}$. In entomology sampling problems, AE% is usually set between 5 and 40. Once a sample has been taken and an interval estimate calculated, the AE% associated with that sample can be determined.

Optimum Sample Size. An important problem in estimation is to determine prior to sampling the sample size n necessary to yield a confidence interval or error bound half-width no larger than some specified size d. The formula to determine this optimum sample size for the 75% error bound, disregarding the finite population correction, is obtained by solving equation (3) for n to yield

(4) $n = (2s/d)^2$.

Similar formulas can be obtained from equations (1) and (2) for the two normal based confidence intervals.

To determine n, d = AE is set at some desired value (e.g., by administrative fiat) or obtained by setting the AE%, taking a preliminary sample (or using a previous sample from the same or a similar population), and calculating d = AE% · $\bar{x}$. The sample standard deviation s must also be determined from a preliminary sample or a previous sample from the same or a similar population.

Equation (4) can be modified to use cv, the sample coefficient of variation, and AE% to yield

$$(5)\ n = \left(\frac{2s}{d}\right)^2 = \left(\frac{2s/\bar{x}(100)}{d/\bar{x}(100)}\right)^2 = \left(\frac{2\ cv}{AE\%}\right)^2.$$

Equations (4) and (5) yield identical results.

For small sample sizes, the two normal-based procedures with $\alpha = 0.05$ and the 75% error bound procedure yield considerably different optimum sample sizes. For larger sample sizes, the three procedures yield approximately the same results. We strongly prefer the 75% error bound procedure as discussed earlier.

Some entomologists use $d = AE = \sigma_{\bar{x}}$ or $s_{\bar{x}}$ as a measure of the goodness of $\bar{x}$. In this case, $AE\% = s_{\bar{x}}/\bar{x}(100) = cvm$, the coefficient of variation of the mean. This value of d is associated with an approximate 68% confidence interval when X is normally distributed and σ^2 is known. Since the allowable error associated with the 75% error bound ($d = 2s_{\bar{x}}$) is twice that associated with $d = s_{\bar{x}}$, the sampler should make sure which form of AE and AE% is appropriate by determining what confidence level is correct for the situation. If $AE = s_{\bar{x}}$ or $AE = 2s_{\bar{x}}$, one is approximately 68% or 95% confident that the interval calculated will include μ, respectively, assuming X is normally distributed and that $\bar{x}$ is an unbiased estimate.

Goodness usually refers to precision since insect density and impact estimates are almost always biased.

COMPARISON OF ESTIMATORS

At first glance, entomologists might choose an unbiased estimator over a biased one since biased estimators have a bad reputation because bias introduces complications and may distort inferences about population parameters. Some samplers consider only the bias due to probability sampling or invalid estimators. This type of bias is usually relatively small compared to nonsampling biases (Table 2, p. 116).

Given an estimator $\bar{x}$, bias is defined as $B = E(\bar{x}) - \mu$. If $E(\bar{x}) = \mu$, $B = 0$, $\bar{x}$ is an unbiased estimator yielding unbiased estimates, and the sampling distribution $f(\bar{x})$ is centered around μ (Fig. 1). When $\bar{x}$ is biased, $E(\bar{x}) = \mu + B$, its sampling distribution is not centered around μ but $E(\bar{x})$ (Fig. 2).

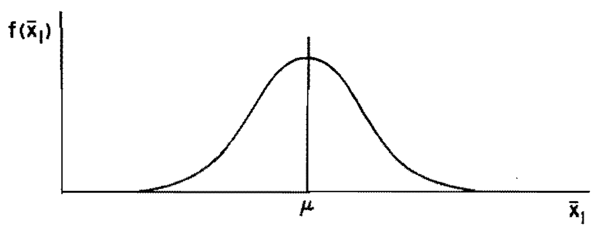

Fig. 1. Hypothetical normal distribution for the unbiased estimator $\bar{x}_1$ where $E(\bar{x}_1) = \mu$.

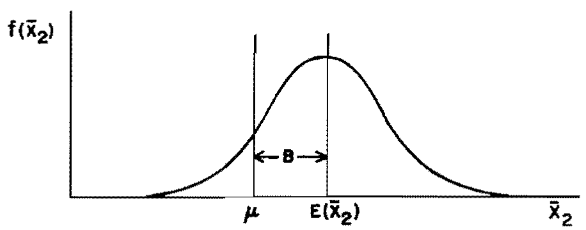

Fig. 2. Hypothetical normal distribution for the biased estimator $\bar{x}_2$ where $E(\bar{x}_2) = \mu + B$ and B is positive.

Even though biased estimators have a bad reputation and there is a misconception that unbiased estimators are always more accurate than biased ones, some biased estimators may be more accurate than unbiased ones. For example (Fig. 3), the biased estimator $\bar{x}_2$ may or may not be more accurate than the unbiased estimator $\bar{x}_1$ since the variance of $\bar{x}_2$ is considerably smaller than the variance of $\bar{x}_1$. For a given estimator, accuracy is a function of the magnitude of both the variance and the bias. The probability that the sample estimates will fall in the interval (a,b) is much larger in the case of $\bar{x}_2$ than $\bar{x}_1$.

In choosing between the two estimators when at least one of them is biased, the sampler must develop a criterion to select the "best" estimator and be aware of the effects of both bias and variance on probability statements.

Selection Criteria. The precision of an estimator $\bar{x}$ is measured by the variance of the estimator $\sigma_{\bar{X}}^2 = \sigma^2/n$ where σ^2 is the variance of the random variable X; $\sigma_{\bar{X}}^2$ measures deviations from $E(\bar{x})$, not from μ. If $\bar{x}$ is unbiased, then $E(\bar{x}) = \mu$ and the accuracy of $\bar{x}$ is measured by $\sigma_{\bar{x}}^2$. If precision is the criterion for selecting the best estimator, then the estimator that has the smallest variance is the "best" or most precise one.

The accuracy of an estimator is measured by the mean square error (MSE) of the estimator where $MSE(\bar{x}) = \sigma_{\bar{x}}^2 + B^2$. If an estimator is unbiased, $MSE(\bar{x}) = \sigma_{\bar{X}}^2$ and its accuracy and precision are identical. $MSE(\bar{x})$ measures deviations of $\bar{x}$ from μ and appears to be a more desirable criterion for choosing the "best" or most accurate estimator.

The relative statistical efficiency (Arvanitis and Fowler 1979, Fowler 1979, Fowler and Simmons 1980, Lindgren 1962) of estimator $\bar{x}_1$ compared to estimator $\bar{x}_2$ is $e(\bar{x}_1,\bar{x}_2) = MSE(\bar{x}_2)/MSE(\bar{x}_1)$. If $e(\bar{x}_1,\bar{x}_2) < 1$, $\bar{x}_2$ is more efficient (accurate) than $\bar{x}_1$ with the reverse being true when $e(\bar{x}_1,\bar{x}_2) > 1$. If both $\bar{x}_1$ and $\bar{x}_2$ are unbiased,

$$e(\bar{x}_1,\bar{x}_2) = \sigma_{\bar{X}_2}^2/\sigma_{\bar{X}_1}^2.$$

If $\bar{x}_1$ is unbiased and $\bar{x}_2$ is biased,

$$e(\bar{x}_1,\bar{x}_2) = \frac{\sigma_{\bar{X}_2}^2 + B^2}{\sigma_{\bar{X}_1}^2} = \frac{\sigma_2^2/n + B^2}{\sigma_1^2/n}.$$

If $\sigma_{\bar{X}_2}^2 < \sigma_{\bar{X}_1}^2$, $\bar{x}_2$ is more precise than $\bar{x}_1$. In addition, $\bar{x}_2$ is more accurate than $\bar{x}_1$ for sample sizes less that n^*, where n^* is determined by setting $MSE(\bar{x}_1) = MSE(\bar{x}_2)$ and solving for $n = n^*$ (i.e., n^* is that value of n where $MSE(\bar{x}_1)$ and $MSE(\bar{x}_2)$ are equal). The estimator $\bar{x}_1$ is more accurate than $\bar{x}_2$ for sample sizes greater than n^*.

Entomologists are usually interested in accuracy. If they want only to make a point estimate of μ, the estimator with the smallest MSE is preferable. However, most estimates will be biased because of nonsampling errors with the bias unknown. Thus, entomologists will be able to estimate only the variance of the estimator and choose the estimator that has the smallest variance (i.e., highest precision). If the bias is known or can be estimated, the MSE should be calculated so that the estimator with the smallest MSE (i.e., highest accuracy) can be chosen. If an estimate of the bias is available from a preliminary sample, adjusted and more accurate estimates can be obtained.

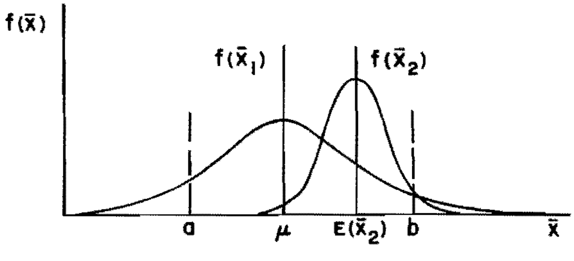

Fig. 3. Hypothetical normal distributions for the unbiased ($\bar{x}_1$) and biased ($\bar{x}_2$) estimators where $E(\bar{x}_1) = \mu$ and $E(\bar{x}_2) = \mu + B$ and B is positive.

It is not entirely correct that two estimators that have the same MSE are equally accurate (Cochran 1977). The frequency distribution of errors ($\bar{x} - \mu$) will not be the same for two estimators with different size biases. However, if $|B|/\sigma_{\bar{x}} < 0.5$, the two frequency distributions are almost identical with respect to absolute errors $|\bar{x} - \mu|$.

Probability Statements. The estimator with the smallest mean square error may be the best one if entomologists only want to make a point estimate of a population parameter. However, if probability statements about population parameters are also needed to construct confidence intervals or test hypotheses, it should be realized that bias distorts such probability statements (Cochran 1977, Hansen et al. 1960, Raj 1968).

For an unbiased estimator $\bar{x}_1$ with $E(\bar{x}_1) = \mu$, assuming normality, variance $\sigma_{\bar{x}_1}^2$ known, and level of significance $\alpha = 0.05$,

(6) $P(\bar{x}_1 - 1.96\sigma_{\bar{x}_1} \leq \mu \leq \bar{x}_1 + 1.96\sigma_{\bar{x}_1}) = 0.95$

from which the 95% C.I. for μ can be computed as $(\bar{x}_1 \pm 1.96\sigma_{\bar{x}_1})$. In terms of the accuracy of $\bar{x}_1$, one can say that $\bar{x}_1$ is in error by $1.96\sigma_{\bar{x}_1}$ or more with probability 0.05 (see $f(\bar{x}_1)$ in Fig. 4).

However, for a biased estimator $\bar{x}_2$ with $E(\bar{x}_2) = \mu + B$, variance $\sigma_{\bar{x}_2}^2$ known, and bias B unknown,

(7) $P(\bar{x}_2 - 1.96\sigma_{\bar{x}_2} \leq \mu \leq \bar{x}_2 + 1.96\sigma_{\bar{x}_2}) \leq 0.95$

which yields an actual α larger than 0.05 and a less than 95% C.I. for μ that can be computed as $(\bar{x}_2 \pm 1.96\sigma_{\bar{x}_2})$. Since $E(\bar{x}_2) = \mu + B$, the effect of bias on the probability statement depends on $B/\sigma_{\bar{x}_2}$. The larger $|B|/\sigma_{\bar{x}_2}$, the larger the actual α will be compared to the nominal α and the smaller the actual confidence coefficient will be compared to the nominal confidence coefficient. If $B = 0$, equation (7) reduces to equation (6).

In terms of the accuracy of $\bar{x}_2$, we can say that $\bar{x}_2$ is in error by $1.96\sigma_{\bar{x}_2} - B$ or more with probability greater than 0.05 (see $f(\bar{x}_2)$ in Fig. 4), depending on the size of $|B|/\sigma_{\bar{x}_2}$. For $\alpha = 0.05$, the actual values of α are 0.0509, 0.0546, 0.0604, 0.0790, and 0.1700 for $|B|/\sigma_{\bar{x}2} = 0.10$, 0.20, 0.30, 0.50, and 1.00, respectively. If the bias is no larger than 10% of $\sigma_{\bar{x}_2}$, the effect of bias on probability statements is negligible. Even with the bias as large as 30% of $\sigma_{\bar{x}_2}$, the effect is quite modest. Biases associated with insect density and impact estimates are often considerably larger than 30% of $\sigma_{\bar{x}_2}$ (see next section).

If the bias is known, the 95% C.I. for μ is $(\bar{x}_2 - B) \pm 1.96\sigma_{\bar{x}_2}$ as the effect of the bias on probability statements has been eliminated. However, bias is usually not known in entomology estimation studies. If an estimate of the bias ($\hat{B}$) can be obtained, $\hat{B}$ can be substituted for B, reducing the effect of bias (Fowler and Simmons 1980).

The effect of bias increases as sample size increases since B is constant and $\sigma_{\bar{x}_2}^2$ decreases with sample size. To find the largest sample size n′ such that the effect of bias is modest (actual $\alpha = 0.0604$ when nominal $\alpha = 0.05$), set $|B|/\sigma_{\bar{x}2} = 0.3$ and solve for n—this value of n is n′.

It should be pointed out that some biased estimators may be more cost-effective than unbiased ones.

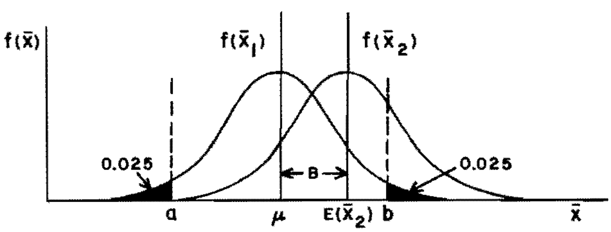

Fig. 4. Hypothetical normal distributions for the unbiased ($\bar{x}_1$) and biased ($\bar{x}_2$) estimators with $E(\bar{x}_1) = \mu$ and $E(\bar{x}_2) = \mu + B$ where B is positive. The $p(a \leq \bar{x}_1 \leq b) = 0.95$ and $p(a \leq \bar{x}_2 \leq b) < 0.95$ for $\alpha = 0.05$ where $a = -1.96\sigma_{\bar{x}_1}$ or $-1.96\sigma_{\bar{x}_2} - B$ and $b = 1.96\sigma_{\bar{x}_1}$ or $1.96\sigma_{\bar{x}_2} - B$ for $\bar{x}_1$ and $\bar{x}_2$, respectively; $\sigma_{\bar{x}_1}$ and $\sigma_{\bar{x}_2}$ are equal.

DOI: 10.22543/0090-0222.1420

EXAMPLE 1: ESTIMATING INSECT DENSITY

Twenty branch tips each 70 cm long were clipped from the midcrown of several balsam fir, *Abies balsamea* L., trees in the Ottawa National Forest in Michigan's Upper Peninsula during August 1979. A counter determined the number of spruce budworm, *Choristoneura fumiferana* (Clemens), egg masses on each branch, and a checker determined the number of egg masses missed by the counter on each branch (Fowler and Simmons 1980). The number of egg masses found by the counter will be called the observed number of egg masses while the number of egg masses found by the observer plus the additional number of egg masses found by the checker will be called the actual number of egg masses. The difference between the actual and observed number of egg masses is the observer error.

The set of 20 branches represents an artificial population of foliage surface area and in no way is representative of the real population from which the branches were selected. This population will be used to demonstrate the difference between accuracy and precision in estimating egg mass density.

The population N = 20 branches yielded a total foliage surface area of 45,030 cm^2, a total actual number of egg masses of 619, and an egg mass density of $\mu_A = 13.75$ egg masses per 1000 cm^2 of foliage surface area. The total observed number of egg masses was 524, which yielded an egg mass density of $\mu_O = 11.64$. The bias due to the nonsampling error of counter missed egg masses was $B = \mu_O - \mu_A = -2.11$, meaning that the nonsampling error for this population was -15.3% of μ_A. The means μ_A and μ_O were determined using probabilities proportional to branch surface area (i.e., the weight of each branch in determining the population mean is equal to its surface area).

The number of egg masses observed, actual number of egg masses present, foliage surface area, observed egg mass density, and actual egg mass density for the 20 branches are shown in Table 3. Assume that the egg mass density for the population of 20 branches is to be

Table 3. Number of egg masses observed, actual number of egg masses present, branch foliated surface area, and observed and actual egg mass density for the population of 20 branches.

Branch Number	Egg Masses Observed	Egg Masses Present	Surface Area (cm^2)	Egg Mass Density	
				Observed	Actual
1	41	50	2880	14.24	17.36
2	58	72	4740	12.24	15.19
3	13	16	1920	6.77	8.33
4	41	52	2640	15.53	19.70
5	6	10	960	17.59	20.37
6	28	32	1680	16.67	19.05
7	38	44	2160	6.25	10.42
8	25	30	1470	17.01	20.41
9	14	16	1320	10.61	12.12
10	35	39	3000	11.67	13.00
11	24	27	2880	10.53	11.84
12	12	13	1680	7.14	7.74
13	2	2	1680	1.19	1.19
14	30	32	2400	12.50	13.33
15	60	77	3240	18.52	23.77
16	28	33	4620	6.06	7.14
17	12	13	1440	8.33	9.03
18	23	24	1920	11.98	12.50
19	23	25	1800	12.78	13.89
20	11	12	1200	9.17	10.00

estimated by selecting a simple random sample of n branches (i.e., sampling probabilities are equal). This is an invalid or biased sampling procedure since each branch has a different surface area and these surface areas were used to determine the population means μ_A and μ_O. A valid or unbiased sampling procedure would be to select branches with probabilities determined by surface area (i.e., probabilities proportional to size). However, such a procedure is not feasible since the size distribution of branches is not known in advance.

The per branch population mean $\bar{X}_A$ and variance σ_A^2 of the 20 branches for the actual egg mass density are 13.32. and 29.3231, respectively. The bias due to the procedure of random sampling with equal probabilities of selection is $B_1 = \bar{X}_A - \mu_A = -0.43$. The per branch population mean $\bar{X}_O$ and variance σ_O^2 of the 20 branches for the observed egg mass density are 11.34 and 19.6089, respectively. The bias $B_2 = \bar{X}_O - \mu_A = -2.41$, where 12.4% $(\bar{X}_O - \mu_O = -0.30)$ and 87.6% $(\mu_O - \mu_A = -2.11)$ of B_2 are due to the invalid sampling procedure and counter missed egg masses, respectively. The part of B_2 due to counter missed egg masses is approximately 7 times larger than that due to the invalid sampling procedure.

The estimator $\bar{x}_A = \sum_{i=1}^{n} x_{Ai}/n$ based on the actual egg mass density has a variance $\sigma_{\bar{x}_A}^2 = \sigma_A^2/n$, assuming sampling with replacement, while the estimator $\bar{x}_O = \sum_{i=1}^{n} x_{Oi}/n$ based on the observed egg mass density has a variance of $\sigma_{\bar{x}_O}^2 = \sigma_O^2/n$.

In terms of precision, the relative efficiency of $\bar{x}_O$ to $\bar{x}_A$ is

$$e(\bar{x}_O,\bar{x}_A) = \frac{\sigma_{\bar{x}_A}^2}{\sigma_{\bar{x}_O}^2} = \frac{\sigma_A{}^2/n}{\sigma_O{}^2/n} = \frac{29.3231}{19.6089} = 1.50$$

for any sample size n, showing that $\bar{x}_O$ is considerably more precise than $\bar{x}_A$.

In terms of accuracy, the relative efficiency of $\bar{x}_O$ to $\bar{x}_A$ is

$$e(\bar{x}_O,\bar{x}_A) = \frac{MSE(\bar{x}_A)}{MSE(\bar{x}_O)} = \frac{\sigma_{\bar{x}_A}^2 + B_1{}^2}{\sigma_{\bar{x}_O}^2 + B_2{}^2} = \frac{29.3231/n + 0.1849}{19.6089/n + 5.8081},$$

showing relative efficiency is a function of sample size n. The estimator $\bar{x}_A$ is more accurate than $\bar{x}_O$ for sample sizes greater than n = 1 and becomes more and more accurate as n increases. For example, $e(\bar{x}_O,\bar{x}_A) = 0.95$, 0.62, 0.40, and 0.24 for n = 2, 5, 10, and 20, respectively.

The precision of an estimator as a percentage is commonly calculated as $(\sigma_{\bar{x}}/\mu)100$, which is the coefficient of variation of the mean. This is the allowable error percent related to a 68.27% confidence interval under normality. For the 75% error bound, the precision of an estimator or the allowable error percent is calculated as $(2\sigma_{\bar{x}}/\mu)100$, which is related to a 95.45% confidence interval under normality and twice as large as precision related to the 68.27% confidence interval. If the estimator is unbiased, one refers to accuracy as well as precision.

The accuracy of an estimator as a percentage is calculated as $(\sqrt{MSE(\bar{x})}/\mu)100$, which reduces to $(\sigma_{\bar{x}}/\mu)100$ when $\bar{x}$ is unbiased.

The precision and accuracy of $\bar{x}_A$ and $\bar{x}_O$ for n = 2, 5, 10, and 20 are shown in Table 4. Precision and accuracy for both $\bar{x}_A$ and $\bar{x}_O$ increase as n increases with $\bar{x}_O$ always more precise than $\bar{x}_A$ because $\bar{x}_O$ has the smaller variance. For n constant, both $\bar{x}_A$ and $\bar{x}_O$ have higher precision than accuracy with $\bar{x}_A$ being more accurate than $\bar{x}_O$. As n increases $\bar{x}_A$ becomes increasingly more accurate than $\bar{x}_O$.

Since both $\bar{x}_A$ and $\bar{x}_O$ are biased and the bias is usually unknown, when we calculate sample estimates we are almost always working with precision when what we want is accuracy. For n = 10, $\bar{x}_A$ has a precision $((\sigma_{\bar{x}}/\mu)100)$ of 12.5%. We would be somewhat underestimating the actual accuracy of 12.8% if we erroneously called it 12.5%. A sample size somewhat larger than 10, on the average, would be needed to yield an accuracy of 12.5%. For n = 10, $\bar{x}_O$ has a precision of 10.2%. If we erroneously called this accuracy, we would be considerably underestimating the actual accuracy of 20.3%. No sample size would yield an accuracy of 10.2% because $(B_2/\mu_A)100 = 17.5\%$, which is as accurate as $\bar{x}_O$ can be. Sometimes biased estimators cannot yield the desired accuracy.

DOI: 10.22543/0090-0222.1420

Table 4. Precision and accuracy of $\bar{x}_A$ and $\bar{x}_O$ for n = 2, 5, 10, and 20.

	$(\sigma_{\bar{x}}/\mu)100$		$(2\sigma_{\bar{x}}/\mu)100$		$(\sqrt{MSE}/\mu)100$	
n	$\bar{x}_A$	$\bar{x}_O$	$\bar{x}_A$	$\bar{x}_O$	$\bar{x}_A$	$\bar{x}_O$
2	27.8	22.8	55.9	45.6	28.0	28.7
5	17.6	14.4	35.2	28.8	17.9	22.7
10	12.5	10.2	25.0	20.4	12.8	20.3
20	8.8	7.2	17.6	14.4	9.3	18.9

As discussed earlier, bias distorts probability statements. For $\bar{x}_A$, $|B|/\sigma_{\bar{x}_A}$ = 0.03, 0.18, 0.25, and 0.36 for n = 2, 5, 10, and 20, respectively. For $\bar{x}_O$, $|B|/\sigma_{\bar{x}_O}$ = 0.78, 1.22, 1.72, and 2.43 for n = 2, 5, 10, and 20, respectively. Probability statements are seriously distorted for $\bar{x}_O$ for all values of n, while probability statements are only moderately distorted even at n = 20 for $\bar{x}_A$. For both $\bar{x}_A$ and $\bar{x}_O$, the distortion increases as sample size increases.

The bias due to the invalid sampling rule associated with $\bar{x}_A$ is considerably smaller than the bias due to the invalid sampling rule and counter missed egg masses associated with $\bar{x}_O$. Even though $\bar{x}_O$ is always more precise than $\bar{x}_A$, $\bar{x}_A$ is more accurate than $\bar{x}_O$ for sample sizes greater than 1, especially for larger sample sizes. The superior estimator is clearly $\bar{x}_A$ as it has a relatively small bias.

In comparing two biased estimators or a biased estimator with an unbiased one, it should be noted that bias affects both the precision and accuracy of an estimator. For the above example, the estimator with both bias due to use of an invalid estimator and counter missed egg masses ($\bar{x}_O$) had a considerably smaller variance and larger bias than the estimator with just bias due to use of an invalid estimator ($\bar{x}_A$). The effect of bias on the variance should not be overlooked.

More accurate estimators can be obtained if the bias can be estimated from a preliminary sample and this estimate used to develop an adjusted estimator (Fowler and Simmons 1980).

EXAMPLE 2: ESTIMATING INSECT IMPACT

Impact is often defined as the cumulative net effects of a given pest on the realized value of a tree species, forest type, or management unit with respect to different resource uses and management objectives.

The spruce budworm is the most important pest in North American spruce-fir forests. The current budworm outbreak covers over 60 million ha of spruce-fir forests with losses estimated at 283 million m^3 in 1978 (Witter 1981). In the Lake States, the current outbreak began in the 1960's and mortality of balsam fir was first reported in the eastern part of Michigan's Upper Peninsula in 1971 (Hastings and Mosher 1976). Currently, damage in individual stands varies from light defoliation to nearly complete stand mortality. Little information was available on the impact of the spruce budworm in Michigan. Therefore, the Michigan Impact Plot System was established in 1978 and 1979 to obtain a database for quantifying the impact of the spruce budworm on forest growth and productivity in the Ottawa and Hiawatha National Forests.

Mog and Witter (1979), Witter and Mog (1981), and Mog et al. (1982) described the sampling units being used in the Michigan Impact Plot System as: (1) primary sampling unit (PSU), a forest compartment, (2) secondary sampling unit (SSU), a spruce-fir stand, and (3) tertiary sampling unit (TSU), circular plots of various radii. The PSU's were weighted according to their acreages of spruce-fir and then selected randomly from each national forest. The SSU's were weighted according to their acreages of spruce-fir, and two SSU's were selected randomly from each PSU. Each SSU was divided into approximately two equal parts with a composite ground sampling unit located in each part. The composite ground sampling unit consisted of three concentric circular plots of various areas (0.02, 0.04, and 0.08 ha) established around a common plot center. A TSU was one of these plots, depending upon the parameter being measured and evaluated.

In order to determine the number of TSU's to sample in a stand, Karpinski and Witter (in press) investigated the effects of plot size (0.02, 0.04, and 0.08 ha) and number of plots (2–5 plots) on the precision of estimates of several impact parameters. In general, AE% decreased more rapidly as the number of plots increased compared to increasing the plot size. The time needed to evaluate three 0.04 ha plots was approximately the same as that needed to evaluate two 0.08 ha plots. Results indicated that three 0.04 ha plots yielded slightly more precise estimates than two 0.08 ha plots.

Mog et al. (1982) presented the spruce budworm impact data at the national forest and forest district levels for the years 1978–1980 for the following parameters: (1) percent tree mortality, (2) total dead volume, (3) dead volume per ha, (4) live volume per ha, (5) defoliation ranking, (6) frequency of top-kill, and (7) incidence of spruce budworm feeding on saplings and regeneration. Estimates of all parameters are based on three-stage cluster sampling techniques, in accordance with the prescribed design.

Table 5 shows estimates of live volume (m^3/ha), with standard errors of these estimates in parentheses, of balsam fir on the Ottawa and Hiawatha National Forests and the Iron River Ranger District of the Ottawa National Forest. Live volume per ha was estimated from 0.08 ha plots. The allowable error percent (AE%), or coefficient of variation of the mean, was 8.4% (1978), 8.7% (1979), and 9.3% (1980) for the Ottawa National Forest, 23.6% (1978), 21.3% (1979), and 24.2% (1980) for the Hiawatha National Forest, and 14.7% (1978), 11.8% (1979), and 11.4% (1980) for the Iron River District of the Ottawa National Forest.

The set of stands chosen for the Michigan Impact Plot System could have been more homogeneous by (1) choosing only those stands with a high percentage of fir throughout all districts, (2) sampling in just one or two districts rather than all districts, or (3) traveling the main roads and selecting stands within easy walking distance. Such a set of homogeneous stands probably would yield more precise estimates than the set actually chosen, but the bias associated with such estimates in estimating impact parameters for the entire population of spruce-fir could be very large. The accuracy of such estimates could be very low and the distortion of probability statements due to bias severe. The estimates based on the Michigan Impact Plot System are probably not as precise as those based on a more homogeneous set of stands, but they are more representative and in all probability are considerably more accurate with distinctly smaller biases.

Entomologists using the above results for management planning might use the AE% as a measure of the accuracy of the estimates for the population of spruce-fir on a given national forest or ranger district. We strongly believe that the above AE%'s can only be used as measures of precision because the estimates are biased causing the true unknown accuracy percentages to be higher than the above AE%'s.

Table 5. Estimates of live volume (m^3/ha) of balsam fir for the Ottawa and Hiawatha National Forests and the Iron River District of the Ottawa National Forest.

YEAR	OTTAWA N.F.	HIAWATHA N.F.	IRON RIVER DISTRICT
1978	41.8 (3.5)[a]	53.4 (12.6)	44.1 (6.5)
1979	44.8 (3.9)	47.4 (10.1)	45.8 (5.4)
1980	43.1 (4.0)	43.4 (10.5)	44.8 (5.1)

[a]Values in parentheses are standard errors of the estimates.

Bias was caused by the approximate nature of the sampling procedure (Mog 1981). Not all spruce-fir stands in a given national forest or ranger district were included in the sampling universe. The unequal sampling probabilities associated with the first two stages of the three-stage sampling process were only approximate. The selection of the plots within a stand was only approximately random. Bias was also caused by many of the nonsampling errors listed in Table 2, even though it was minimized using procedures discussed earlier in the Accuracy and Precision Section.

Even though the bias caused by nonsampling errors can be minimized, it will almost always be larger than the bias due to probability sampling and the use of invalid estimators for most impact estimates. In any case, bias will always be present and the variance of the estimator refers to precision and not accuracy.

COMMENTS

Even though entomologists are understandably interested in accuracy, they are almost always dealing with precision. Insect density and impact estimates are biased due to (1) an invalid estimator, (2) probability sampling, or (3) nonsampling errors. Bias due to nonsampling errors, such as incomplete coverage of the target population and sampling only 70-cm branch tips from the midcrown of balsam fir trees, is almost alway larger than the other two types of bias. Even though nonsampling biases can be minimized by careful planning and execution of the sampling process, such biases are always present.

It has been argued that if these biases occurred according to some random process, then positive and negative biases would tend to cancel out over a sufficiently large sample. However, many of these sources of bias are systematic in nature. The cumulative effect of the various biases on an estimate is not always negligible. It is our belief that in most density and impact studies the various biases do not cancel out and the net effect is a large bias.

Bias affects both the precision and accuracy of estimates in that bias is a component of both the variance and mean square error.

Erroneously using the AE% as the accuracy percentage underestimates the true accuracy of the estimate. Confusing accuracy with precision could lead to estimates less accurate than desired, taking fewer observations in a sample than needed to obtain some desired accuracy, and possibly making the wrong decision.

Even though precision is what they deal with, entomologists should think in terms of accuracy. The allocation of resources between reducing sampling error and minimizing bias should be made such that the difference between accuracy and precision is minimized in a cost-effective approach. If estimates of the bias are available, the approximate mean square error should be used. Final sample sizes should be larger than those determined by optimum sample size procedures, if possible, to decrease the unknown accuracy percentage.

In choosing an estimator, it should be remembered that some biased estimators can be more accurate than unbiased estimators, bias distorts probability statements, some biased estimators may be more cost-effective than unbiased ones, and biased estimates may cause incorrect management decisions.

If some of the biases associated with an estimator can be feasibly estimated from a preliminary sample, more accurate adjusted estimators can be obtained. Thus, bias can be reduced and the difference between accuracy and precision decreased. In any case, most estimates will be biased, more or less, and the entomologist should be aware of the resulting difference between accuracy and precision.

Many insect density and impact estimates have high precision (small variances) but low accuracy (large mean square errors) because of bias. Thus, the accuracy of estimates with different biases cannot be compared by examining their variances. Inasmuch as it is cost-effective, bias should be minimized as much as is feasibly possible. However, whatever the case, entomologists deal with precision and should attempt to minimize the difference between precision and accuracy.

Table 1. Definitions related to statistical estimation.[a]

POPULATION—the aggregate of items, elements, or units of interest in a well-defined group. The target population (the population about which information is desired) should coincide with the sampled population (the population to be sampled).

SAMPLE—a subset of the population obtained using some selection procedure.

SAMPLING UNIT—the population is divided into a finite number of distinct and identifiable units called sampling units. Each sampling unit consists of one or more elements from the population.

SAMPLING FRAME—a list of all of the sampling units in the population. The sampling frame provides the basis for the selection and identification of units in the sample.

POPULATION PARAMETER—a quantitative characteristic describing the population such as the population mean μ.

SAMPLE STATISTIC—a quantitative characteristic describing the sample which is used to estimate an unknown population parameter. For example, the sample mean $\bar{x}$ is used to estimate the population mean μ.

VARIABLE—a characteristic of interest, represented by a symbol like X, that can take on any value of a property or attribute observed on each element of a population. X usually varies from element to element.

RANDOM VARIABLE—a variable, like X, whose value is determined by a random selection procedure. Each sampling unit in the population has a certain probability of being selected. X takes on a particular value, depending on which sampling unit is selected.

ESTIMATOR—a rule or formula that shows how to calculate a sample statistic for the elements chosen in a given sample. For example.

$$\bar{x} = \sum_{i=1}^{n} x_i/n$$

is an estimator for μ.

ESTIMATE—the sample statistic or number calculated from the n sample observations (values) taken from a population using a particular estimator. For example, the $n = 4$ observations $\bar{x}_1 = 2$, $\bar{x}_2 = 4$, $\bar{x}_3 = 3$, and $\bar{x}_4 = 1$ taken in a simple random sample yield the estimate $\bar{x} = (2+4+3+1)/4 = 2.5$.

EXPECTED VALUE—the expected value of an estimate is the mean of the estimates for all possible samples that can be taken from the population. For example, the expected value of the sample mean is represented by $E(\bar{x})$.

UNBIASED ESTIMATOR—an estimator is unbiased if the mean or expected value of all possible estimates is equal to the population parameter being estimated. For example, $E(\bar{x}) = \mu$. Estimates based on an unbiased estimator are called unbiased estimates.

BIAS—the difference between the expected value of the estimate and the true population parameter being estimated. The bias in estimating the population mean μ is $B = E(\bar{x}) - \mu$. Bias is caused by (1) using an estimator that is not valid for a particular situation, (2) using an estimator that is biased in probability sampling, or (3) non-sampling errors (errors arising in the course of collecting and processing the data). If $\bar{x}$ is an unbiased estimator, $B = O$.

BIASED ESTIMATOR—an estimator is biased if the mean of all possible estimates is not equal to the population parameter being estimated (e.g., $E(\bar{x}) = \mu + B$). Estimates based on biased estimators are called biased estimates.

VARIANCE OF A RANDOM VARIABLE X—a measure of the dispersion or spread of the distribution of X from the mean or expected value of X. More exactly, the variance of X is the average squared deviation from E(X) of all the values of X, namely, $\sigma^2 = Var(X) = E[X - E(X)]^2$. If $E(X) = \mu$, $\sigma^2 = E(X-\mu)^2$. For simple random sampling,

$$s^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2/(n-1)$$

is an unbiased estimator of σ^2.

Table 1. (Continued)

VARIANCE OF AN ESTIMATOR—a measure of the dispersion or spread of the distribution of all possible estimates from their mean or expected value. In estimation of the mean, the variance of the estimator $\bar{x}$ is the average squared deviation from $E(\bar{x})$ of all estimates $\bar{x}$, namely,

$$\sigma_{\bar{x}}^2 = \text{Var}(\bar{x}) = E[\bar{x} - E(\bar{x})]^2.$$

If the estimator $\bar{x}$ yields unbiased estimates of μ, $E(\bar{x}) = \mu$ and $\sigma_{\bar{x}}^2 = E(\bar{x} - \mu)^2$. For simple random sampling,

$$s_{\bar{x}}^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2/[n(n-1)],$$

the sample variance of the mean, is an unbiased estimator of $\sigma_{\bar{x}}^2$.

MEAN SQUARE ERROR (MSE) OF AN ESTIMATOR—a measure of the dispersion or spread of the distribution of all possible estimates from the population parameter being estimated. In estimation of the mean, the MSE is the average squared deviation from μ of all estimates $\bar{x}$ where $MSE(\bar{x}) = E(\bar{x} - \mu)^2 = E[\bar{x} - E(\bar{x})]^2 + [E(\bar{x}) - \mu]^2 = \text{Var}(\bar{x}) + B^2$. Thus, the MSE is equal to the variance of the estimator plus the square of the bias. If the estimator is unbiased, MSE is equal to the variance of the estimator. The MSE of a biased estimator cannot usually be estimated from a sample because the bias usually is unknown and cannot be estimated.

RELATIVE EFFICIENCY-the ratio of the MSE's or variances of two estimators, depending on how efficiency is defined. For estimation of the mean, the relative efficiency of estimator $\bar{x}_1$ compared to $\bar{x}_2$ is

$$e(\bar{x}_1,\bar{x}_2) = MSE(\bar{x}_2)/MSE(\bar{x}_1) \text{ or } e(\bar{x}_1,\bar{x}_2) = \text{Var}(\bar{x}_1)/\text{Var}(\bar{x}_2).$$

COEFFICIENT OF VARIATION (CV)—a relative measure of the variability of the random variable X. It is the ratio of the standard deviation of X to the mean or expected value of X times 100 to yield percentage values. The standard deviation is the square root of the variance. The population $CV = (\sqrt{VAR(X)}/E(X))100$ is estimated by the sample $CV = (s/\bar{x})100$.

COEFFICIENT OF VARIATION OF THE MEAN (CVM)—a relative measure of the variability of the sample mean $\bar{x}$. It is the ratio of the standard error of $\bar{x}$ to the mean or expected value of $\bar{x}$ times 100. The standard error of the mean is the square root of the variance of the mean. The population $CVM = (\sqrt{VAR(\bar{x})}/E(\bar{x}))100$ is estimated by the sample $cvm = (s_{\bar{x}}/\bar{x})\ 100$.

SAMPLING ERROR—the difference between the sample estimate, say, $\bar{x}$, and the mean of all possible estimates $(E(\bar{x}))$ based on the same sample size and the same selection procedure. Sampling error is almost always unknown, and its size depends on the particular sample chosen, sample size, population variance, and particular sampling method used. If the estimator is unbiased, we have the true sampling error which is the difference between the estimate and the population parameter being estimated due to chance alone (sampling error $= \bar{x} - \mu$), where $E(\bar{x}) = \mu$. Sampling error is not an error in the true sense of the word as it is due to the sample being based on only a portion of the population elements. If the estimator is biased, we have the actual sampling error which is the difference between the estimate and the mean of all possible estimates (sampling error $= \bar{x} - E(\bar{x})$), where $E(\bar{x}) = \mu + B$.

NONSAMPLING ERRORS—those errors that arise in the course of collecting and processing the data to include observational or response errors, incomplete coverage errors, compiling errors, and any number of other errors that can and usually do occur during the sampling process. These errors may be equally as or more important than sampling errors. It is desirable and possible to minimize such errors.

ERROR OF ESTIMATION—the deviation of the estimate from the population parameter being estimated. The true sampling error is equal to the error of estimation if the estimator is unbiased. The actual sampling error plus the bias, which is the true sampling error, is

Table 1. (Continued)

equal to the error of estimation if the estimator is biased. In estimation of the mean, the average size of the error of estimation ($\bar{x} - \mu$) is given by the variance and the mean square error of the estimator for unbiased and biased estimators, respectively. The most common measure used is the standard error, which is the square root of the variance or the mean square error.

ACCURACY—the accuracy of a sample estimate is the difference between the sample estimate and the true population parameter being estimated.

PRECISION—the precision (reliability) of a sample estimate is the difference between the sample estimate and the mean of the estimates of all possible samples that can be taken from the population.

[a]More detailed discussions of the above terms and concepts can be found in Cochran (1977, p. 12–16,359–396), Hansen et al. (1960, p. 16–26,34–39,56–109), Lindgren (1962, p. 266–291), Raj (1968, p. 26–30,165–186), and Sukhatme and Sukhatme (1970, p. 10–17,28,444–484).

Table 2. Some nonsampling errors associated with insect density and impact estimates.

1. Modifications of the theoretical sampling procedure to include (a) substitution of one observation for another, (b) incomplete samples or nonresponse, (c) keeping plots away from the forest edge, and (d) not using random selection procedures (e.g., sample selected with unknown probabilities, representative sample, and convenient sample).
2. Incomplete coverage of population of interest; the sampling population is only a subset of the target population (e.g., sampling spruce-fir trees near roads rather than throughout a forest region and sampling 70 cm branch tips at midcrown rather than whole branches throughout a tree).
3. Decision on sample design based on preliminary sample results.
4. Use of sample estimates from other populations.
5. Selecting units nearest to randomly selected points.
6. Observation (response or measurement) errors, due to human mistakes or inaccurate instruments, to include (a) insect-counting errors, (b) plot location and layout errors, (c) missing trees in sample plots, (d) inaccurate counts of saplings, (e) inaccurate tree diameter and height measurement, (f) calling live trees dead and vice-versa, (g) errors in determining crown position, (h) errors in estimating tree defoliation, and (i) errors in tree species identification.
7. Volume table errors.
8. Coding, notekeeping, editing, and keypunching errors.
9. Analysis and interpretation errors.

ACKNOWLEDGMENTS

Thanks are due to L. G. Arvanitis, J. Carow, R. P. Ford, A. L. Jensen, C. Karpinski, Jr., A. M. Lynch, B. A. Montgomery, J. P. Nyrop, R. L. Patterson, P. A. Rush, G. A. Simmons, K. D. Smith, L. Waisanen, and W. Waters for their timely and helpful comments. This work was supported in part by the Canada-United States Spruce Budworms Program, the McIntire-Stennis Cooperative Forestry Research Program, and the School of Natural Resources, The University of Michigan.

https://scholar.valpo.edu/tgle/vol15/iss2/5
DOI: 10.22543/0090-0222.1420

LITERATURE CITED

Arvanitis, L. G. and G. W. Fowler. 1979. Some aspects of biased sampling estimators. Proc., For. Res. Inventories Workshop. Colorado State Univ. Ft. Collins. Vol. I:298–309.

Cochran, W. G. 1977. Sampling techniques. 3rd Ed. John Wiley & Sons, Inc. New York.

Fowler, G. W. 1979. Sampling of natural resource populations: comparison of estimators. Res. Inv. Notes BLM 25:1–7.

Fowler, G. W. and D. Hauke. 1979. A distribution-free method for interval estimation and sample size determination. Res. Inv. Notes. BLM 19:1–8.

Fowler, G. W. and G. A. Simmons. 1980. More accurate estimators for spruce budworm egg mass density (Lepidoptera: Tortricidae). Great Lakes Entomol. 13:151–157.

Hansen, H., W. M. Hurwitz, and W. G. Madow. 1960. Sample survey methods and theory. Vol. I. Methods and application. John Wiley & Sons, Inc. New York.

Hastings, A. R. and D. C. Mosher. 1976. An aerial survey of defoliation and mortality caused by the spruce budworm in the Upper Peninsula of Michigan—1976, with management considerations. USDA For. Serv. Eval. Rep. 5-25-76.

Hogg, R. V. and A. T. Craig. 1965. Introduction to mathematical statistics. The MacMillan Co. New York.

Lindgren, B. W. 1962. Statistical theory. MacMillan Co. New York.

Karpinski, C., Jr. and J. A. Witter. Plot size and number of plots needed to assess spruce budworm impact. Canadian J. For. Res. (in press).

Mog, T. P. 1981. Impact of *Choristoneura fumiferana* (Clemens) in Michigan's Upper Peninsula. Ph.D. dissert. Univ. Michigan, Ann Arbor.

Mog, T. P., A. M. Lynch, and J. A. Witter. 1982. Impact of the spruce budworm on the Ottawa and Hiawatha National Forests, 1978–1980. Great Lakes Entomol. 15:1–24.

Mog, T. P. and J. A. Witter. 1979. Field techniques for assessing the impact of the spruce budworm (Lepidoptera: Tortricidae). Great Lakes Entomol. 12:213–218.

Raj, D. 1968. Sampling theory. McGraw-Hill Book Co. New York.

Scheaffer, R. L., W. Mendenhall, and L. Ott. 1979. Elementary survey sampling. Duxbury Press. North Scituate, Mass.

Sukhatme, P. V. and B. V. Sukhatme. 1970. Sampling theory of surveys with applications. Iowa State Univ. Press, Ames.

Witter, J. A. 1981. Techniques for asessing the impact of the spruce budworm in Michigan. p. 529–540, *in* Proc., XVII IUFRO World Cong. Div. 2.

Witter, J. A. and T. P. Mog. 1981. An integrated approach for assessing spruce budworm damage and developing a hazard rating system and stand models for spruce-fir stands in Michigan's Upper Peninsula, , p. 45–51 *in* Hedden, R. L., S. J. Barras, and J. E. Coster (Tech. Co-ord.). Hazard-rating systems in forest insect pest management: symposium proceedings (Athens, GA., July 1980). USDA For. Serv. Tech. Rep. WO-27.