Errors Incurred When Using Iwao's Sequential Decision Rule in Insect Sampling

JAN P. NYROP AND GARY A. SIMMONS

IPM Program, New York State Agricultural Experiment Station, Geneva, New York 14456;
Department of Entomology, Michigan State University, East Lansing, Michigan 48824

ABSTRACT Sequential decision rules based upon defining a relationship between the mean and variance of a population and using this relationship in a confidence interval about a sum of counts on sample units have been developed and used for sampling insect pests. The actual error rates that arise from the use of these rules are much larger than the nominal ones. This occurs for three reasons: The confidence interval is incorrectly used as a sequential decision rule; it is assumed that random variables are normally distributed; and variances are not constant. If these sequential decision rules are to be used, we suggest that actual error rates be determined through simulation.

WHEN SAMPLING to assess the density or intensity (or any index of either) of a pest population it is often not necessary to estimate a mean precisely; rather, it is often desirable to classify it as above or below a specified level. In such cases, sequential hypothesis testing provides an attractive alternative to sampling schemes that employ a fixed number of observations per sample. Fixed-sample-size methods often provide inadequate precision at low population densities and excessive sample sizes at high densities. With sequential procedures, sample size is dependent on the outcome of each successive observation as it relates to the sum or average of previous observations in a given set of samples. When properly constructed, these procedures, on the average, are as reliable as fixed-sample-size procedures and require fewer observations. For this reason, they are attractive sampling schemes when cost and time efficiency are important.

Most applications of procedures for sequential hypothesis testing in insect sampling have employed Wald's Sequential Probability Ratio test (SPRT) (Wald 1947). A sequential procedure proposed by Iwao (1975) is purported to overcome some of the problems inherent to the SPRT and has recently been applied to insect sampling (Coggin and Dively 1982, Shaw et al. 1983). The basis for each of these hypothesis-testing procedures is quite different. Use of the SPRT necessitates that the distribution of a sampled population can be described by a mathematical model (e.g., Poisson, negative binomial) with all parameters other than the mean fixed regardless of population density. However, population distributions often change with changes in the density, quality, and age of a population, as well as spatial and temporal changes in the environment. Iwao's procedure depends on defining a relationship between the mean (μ) and variance (σ^2) of a population:

$$\sigma^2 = f(\mu) \tag{1}$$

Since such a relationship can often be found when distributional models change in relation to the mean (e.g., variable k of a negative binomial), the procedure is appealing.

The properties, assumptions, and errors associated with SPRT's have been studied and are well documented (Wald 1947, Davies 1956, Onsager 1976, Fowler 1978). Comparable knowledge of Iwao's procedures does not exist. Hence, the purpose of this paper is to establish the statistical principles upon which Iwao's procedure rests and to indicate errors that arise through its use. It will be shown that these errors are actually much larger than the nominal errors established during construction of the procedure. The ideas presented are illustrated by simulating samples from populations described by mean/variance relationships.

Basis for Iwao's Sequential Sampling Procedure

In statistical inference there are three types of problems, each with an element of uncertain probability associated with it: hypothesis-testing problems, multiple-decision problems, and estimation problems. Sequential sampling is an approach to hypothesis testing. The problem is to decide whether a population parameter θ belongs to either set ξ_0 or set ξ_1. More formally, two hypotheses are constructed, the working and alternate hypotheses:

$$H_0 : \theta \epsilon \xi_0$$
$$H_1 : \theta \epsilon \xi_1$$

A sample is drawn from the population and one of three possible actions is taken: the working hypothesis is accepted, the working hypothesis is rejected and it is assumed that the alternate hypothesis is correct, or another sample is taken. To arrive at an action, a decision function $d(x_1, x_2, \ldots, x_n)$ is constructed which separates an n-dimensional sample $(x_1, x_2, \ldots, x_n)$ into three parts: A_0, A_1, and

A_2. If a random sample x produces a point in A_0, then H_0 is accepted. If a random sample x produces a point in A_1, then H_0 is rejected and H_1 is assumed to be correct. If a random sample x produces a point in A_2, another sample is taken.

There are two types of errors associated with such a decision function. A Type I error corresponds to the probability (α) that a random sample x is an element of A_1 given that θ belongs to set ξ_0. A Type II error corresponds to the probability (β) that a random sample x is an element of A_0 given that θ belongs to the set ξ_1. For a given sample size, the value of α can be decreased only by shifting sample points from A_1 to A_0 and this will increase β. Hence, there is no partition of the sample space that will simultaneously minimize α and β. This may not, however, impose a handicap on a decision function since in many cases it is not necessary to minimize both α and β. This is because it may be more important not to make one of the errors in the context of the decision to be made.

Iwao's decision function depends first on describing a mean/variance relationship as in equation 1. The working and alternate hypotheses are then formulated as follows:

$$H_0 : \mu < \gamma$$
$$H_1 : \mu > \gamma$$

A decision function is constructed by computing a confidence interval about γ in terms of the total number of organisms found in n observations. These α-level confidence bounds are:

$$\gamma n \pm z_{\alpha/2} \sqrt{nf(\gamma)} \qquad (2)$$

The term $\sqrt{nf(\gamma)}$ is the standard deviation of the variable γn. The variable $z_{\alpha/2}$ is a standard normal deviate such that $P(|z| > z_{\alpha/2}) = \alpha$. Here, $P(\ )$ reads "the probability that." An interpretation of the confidence limits described in equation 2 is that the sum of animals in n observations (T_n) will lie within the confidence limits $100(1 - \alpha)$ percent of the time if the population mean is γ.

In Iwao's procedure, the space A_0 is upperbounded by the lower limit of the confidence interval in equation 2. The space A_1 is lower-bounded by the upper limit of the confidence interval in equation 2. Sampling is discontinued whenever T_n falls within A_0 or A_1 and a decision is made either to accept or reject the working hypothesis. If T_n falls within A_2, an additional sample is taken. If all sample points fall in A_2, sampling is discontinued when n reaches some maximum ($n$max). Iwao does not provide guidelines for decision making at this point; however, two alternatives are available. First, a mean based on $n$max samples may be computed and compared to γ. Because the variance is assumed known, a maximum sample size can be chosen for a desired level of precision. Second, no decision may be made and sampling may be initiated again after a short period of time. In this paper we adopt the first rule.

Table 1. Frequency of sequentially classifying the mean of n samples ($\bar{x}_n = \sum_{i=1}^{n} x_i/n$) as >5 when $x_i \sim N(4.7, 10)$ based on the following rule: $\mu > 5$ if $\bar{x}_n > 5 + 1.96\sqrt{10/n}$

$n$max	Frequency of error
1	0.021
3	0.039
5	0.049
7	0.053
9	0.065
11	0.064
13	0.067
15	0.077
17	0.061
19	0.085
21	0.086
23	0.076
25	0.086

This rule is applied for each sample up to a maximum of $n$max samples. Frequencies are based on 2,000 simulations for each $n$max.

For a composite hypothesis, the error probability α is defined as $P(x\epsilon A_1|\theta) \leq \alpha$ for all $\theta\epsilon\xi_0$. For Iwao's procedure, this probability is maximized when $\mu = \gamma$. To evaluate the decision rule, a more useful criterion is the operating characteristic (OC) defined as $P(\text{accept } H_0|H_i \text{ is true})$ where H_i is any simple hypothesis; $H_i : \mu = \gamma_i$. Note that the OC for $\mu = \gamma = 1 - \alpha$. A further useful criterion is the power defined as $1 - OC$. With Iwao's procedure, the OC and power are symmetric about γ. When H_0 is true, the power describes the probability of making an error, and when H_1 is true, the OC describes the error probabilities.

The OC is composed of two parts: one part associated with sequential decision making (OC_{seq}) and one part associated with computing a sample mean from $n$max observations and basing a decision on the sample mean (OC_{nmax}). In terms of relative frequencies, OC_{seq} and OC_{nmax} are computed from the number of correct and incorrect sequential and terminal decisions divided by the total number of decisions made. If $\mu < \gamma$, then OC_{seq} is the number of correct sequential decisions divided by the total number of decisions. When $\mu > \gamma$, OC_{seq} is the number of incorrect sequential decisions divided by the total number of decisions. OC_{nmax} is similarly defined.

Assumptions and Errors in Iwao's Procedures

Suppose α is specified as α^0 to construct confidence bounds in equation 2. It might then be assumed that the probability of making a sequential error (provided by the respective parts of the OC and power curves) will be $\leq \alpha^0$. However, this will never be true: Sequential errors will actually be much larger than the specified level over some range of the population mean. This deviation re-
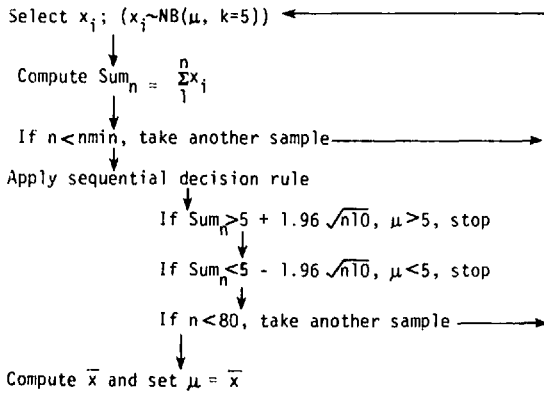
Select x_i; $(x_i \sim NB(\mu, k=5))$

Compute $Sum_n = \sum_1^n x_i$

If $n < nmin$, take another sample

Apply sequential decision rule

If $Sum_n > 5 + 1.96\sqrt{n10}$, $\mu > 5$, stop

If $Sum_n < 5 - 1.96\sqrt{n10}$, $\mu < 5$, stop

If $n < 80$, take another sample

Compute $\bar{x}$ and set $\mu = \bar{x}$

Fig. 1. Iwao's sequential decision rule applied to a population of negative binomially (NB) distributed variates where $nmax = 80$ and the maximum nominal error rate is 0.05.

sults from an incorrect application of the confidence bounds in equation 2 as a sequential decision rule, from the naive assumption that T_n is normally distributed, and because the mean/variance relationship is usually not specified such that the variance is constant with respect to the mean, i.e., $\sigma^2 = C$. In this section, the details of these errors are discussed.

The confidence bounds in equation 2 may be rewritten as $\mu_0 \pm z_{\alpha/2}\sqrt{f(\mu_0)/n}$ by dividing by n. Now suppose samples are to be drawn from a population which is normally distributed with mean μ and variance σ^2. We wish to determine whether μ is greater than or less than μ_0 when $\sigma_0^2 = \sigma^2$. Note that a constant variance has been assumed. Because of the symmetry of the normal distribution, errors for the case $\mu < \mu_0$ and errors for the case $\mu > \mu_0$ will be similar. It therefore suffices to discuss only one of these situations.

If $\mu < \mu_0$, the first probability of making an incorrect sequential classification with the first observation (x_1) is $P(x_1 > C_1)$, where $C_1 = \mu_0 + z_{\alpha/2}\sigma$. When α is chosen to be 0.05, this conditional error (conditional on the fact that $\mu < \mu_0$) is less than 0.025. For two samples, (x_1, x_2) the probability of an incorrect classification is $P(x_1 > C_1) + P[(x_1 + x_2)/2 > C_2 | C_0 < x_1 < C_1]$ where $C_0 = \mu_0 - z_{\alpha/2}\sigma$ and $C_2 = \mu_0 + z_{\alpha/2}\sigma/\sqrt{2}$. The second probability is conditional on $C_0 < x_1 < C_1$. The first and second events are mutually exclusive and the probability of their union is the sum of their probabilities. Note that the second term in this sum must be greater than 0. Hence, the probability of a sequential error when the maximum number of samples to be taken is two is greater than when only one sample is taken before terminating the sampling process. The probability of a sequential error therefore increases in an asymptotic manner as the maximum number of samples which can be taken increases.

Changes in the sequential error as a function of

the maximum sample size are dramatic. As an example, let $\mu = 4.7$, $\mu_0 = 5.0$, $\sigma^2 = 10$. For $nmax = 1$, $P(\text{accept } H_1 | \mu = 4.7) = P(Z_1 > \overset{*}{C}_1) = 0.0202$ where $Z_1 = (x_1 - \mu)/\sigma$ and $\overset{*}{C}_1 = (\mu_0 + 1.96\sigma - \mu)/\sigma$. For $nmax = 2$, $P(\text{accept } H_1 | \mu = 4.7) = P(Z_1 > C_1) + P[(Z_1 + Z_2)/2 > C_2 | C_0 < x_1 < C_1]$ where $Z_2 = (x_2 - \mu)/\sigma$, $\overset{*}{C}_0 = (\mu_0 - 1.96\sigma - \mu)/\sigma$, and $\overset{*}{C}_2 = (\mu_0 + 1.96\sigma/\sqrt{2} - \mu)/\sigma$. The following manipulations may be made on the second term:

$$P[(Z_1 + Z_2)/2 > \overset{*}{C}_2 | \overset{*}{C}_0 < Z_1 < \overset{*}{C}_1]$$

$$= P(Z_2 > 2\overset{*}{C}_2 - Z_1, \overset{*}{C}_0 < Z_1 < \overset{*}{C}_1)/$$
$$P(\overset{*}{C}_0 < Z_1 < \overset{*}{C}_1) \qquad (3a)$$

$$= \int_{-\infty}^{z_1 - 2\overset{*}{C}_2} \int_{\overset{*}{C}_0}^{\overset{*}{C}_1} \psi(Z_1)\psi(Z_2) \, dZ_1 \, dZ_2/$$

$$\int_{\overset{*}{C}_0}^{\overset{*}{C}_1} \psi(Z_1) \, dZ_1 \qquad (3b)$$

where ψ is the standard normal density function,

$$= \int_{\overset{*}{C}_0}^{\overset{*}{C}_1} \psi(Z_1)\phi(Z_1 - 2\overset{*}{C}_2) \, dZ_1/$$

$$\int_{\overset{*}{C}_0}^{\overset{*}{C}_1} \psi(Z_1) \, dZ_1 \qquad (3c)$$

where ϕ is the standard normal cumulative probability function. The integral in equation 3c must be evaluated numerically and has the solution 0.0128. Therefore, the probability of a sequential error is $0.0128 + 0.0202 = 0.0330$. This is much greater than the assumed maximum nominal error of 0.025.

Further extensions of the above analysis are impractical. However, a general picture of the relation between the maximum sample size and sequential error can be obtained through computer simulation. For the population described above, the outcome of 2,000 simulated applications of the decision rule per $nmax$ are listed in Table 1. With a maximum nominal sequential error rate of 0.025, an actual error rate of 0.086 was obtained when $nmax = 25$. This is almost 3.5-fold the assumed maximum sequential error rate. Note that the error rates are distributed about an asymptotic limit which is ca. 0.08.

The important fact is that the application of a confidence bound as a sequential decision rule will not yield an error rate equal to the confidence level chosen when constructing the confidence interval. This is because the actual error rate is a sum of conditional probabilities which must be greater than the assumed nominal error rate. In order for the overall error rate of the sequential portion of the test to be equal to some predetermined level, a new $z_\alpha = z_{\alpha,n}$ would have to be determined for each n. This has in fact been done
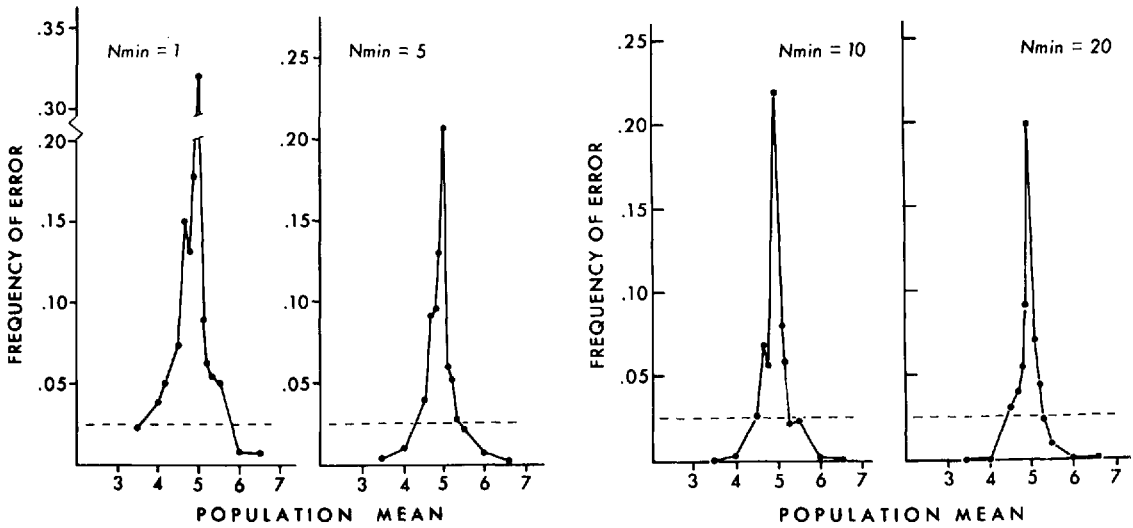
Fig. 2. Frequency of error for the sequential portion of Iwao's decision rule when the minimum number of samples taken before applying the decision rule ranged from 1 to 20. The x_i were distributed as negative binomial variates with $k = 5$ and the null hypothesis was $\mu \leq 5.0$. The dashed line is the conditional maximum nominal error rate except when the population mean $= 5$; then the nominal error rate is 0.05. Error rates are conditional on the sample population mean being either $<$ or > 5.

for a number of sequential t tests (*see* Rushton 1950, Fowler and O'Regan 1974).

A further extension of this concept can be made with respect to the total probability of error with Iwao's procedure. The total probability of error is equal to the probability of a sequential error plus the probability of a terminal error, since these events are again mutually exclusive. Therefore, if a sequential test was constructed to yield an actual error rate of 0.05 and $n$max was selected so that a terminal decision also had a maximum error rate of 0.05 for a particular μ, the total error rate for the test would be 0.10 for this mean. Both error rates would change as the mean changed.

In addition to a misapplication of a confidence interval as a sequential test, violation of the normality assumptions also contributes to increased error rates. A fundamental theorem in statistics, known as the Central Limit Theorem, states that if x_i are independent, identically distributed random variables with mean μ and variance σ^2, then

for $S_n = \sum_{i=1}^{n} x_i$

$$\lim_{n \to \infty} P[(S_n - n\mu)/\sigma\sqrt{n} < Z)] = \phi(Z) \quad (4)$$

where $\phi(Z)$ is the cumulative distribution function for the standard normal variate Z. The term $\lim_{n \to \infty}$ and what follows reads "the limit of $P(\ )$ is $\phi(Z)$ as n converges to infinity." Equation 4 may be rewritten as:

$$\lim_{n \to \infty} P(S_n - Z_{\alpha/2}\sigma\sqrt{n} < n\mu$$
$$< S_n + Z_{\alpha/2}\sigma\sqrt{n}) = 1 - \alpha \quad (5)$$

This is identical to equation 2 with the added requirement that if x is not normally distributed, then the probability converges to $1 - \alpha$ as n gets large. If x is normally distributed, the limit is not necessary. In practice, n equal to 20 is often sufficiently large for the statement to be useful.

The important point is that when x is not normal, a decision based on a normal model should not be made unless n is large enough to invoke the Central Limit Theorem. Hence, some number of samples should be taken before applying the decision rule. Samples of insects most often have a contagious distribution. As a result, considerable deviations in $(1 - \alpha)$ may result unless the criteria for the Central Limit Theorem are met.

The probability statement in equation 5 states that for n large or x normally distributed, the probability of S_n falling within the bounds $n\mu \pm Z_{\alpha/2}\sigma/\sqrt{n}$ is $1 - \alpha$. Note that $Z_{\alpha/2,n}$ must be replaced with $\overset{*}{Z}_{\alpha/2,n}$ if a sequential test is to be used. This is not necessary to illustrate the effect of changes in the population variance. Suppose samples are taken from a population with a mean equal to $\mu + \delta$. One might then expect that the probability of a Type I error would be less than $\alpha/2$. However, this will only be the case provided the variance is a constant.

As an example, let x be normally distributed with $\mu = 5$ and $\sigma = 2$. For a sample size of 10, the 95% confidence limits about α are $5 \pm 1.96 \, 2/\sqrt{10}$. Suppose $\mu + \delta = 6$ and $\sigma = 4$ describe the population from which samples are to be drawn. The probability of classifying the population mean as less than 5 may be found by solving for Z in $3.76 = 6 - (Z \cdot 4/\sqrt{10})$. The value of Z is 1.771 which
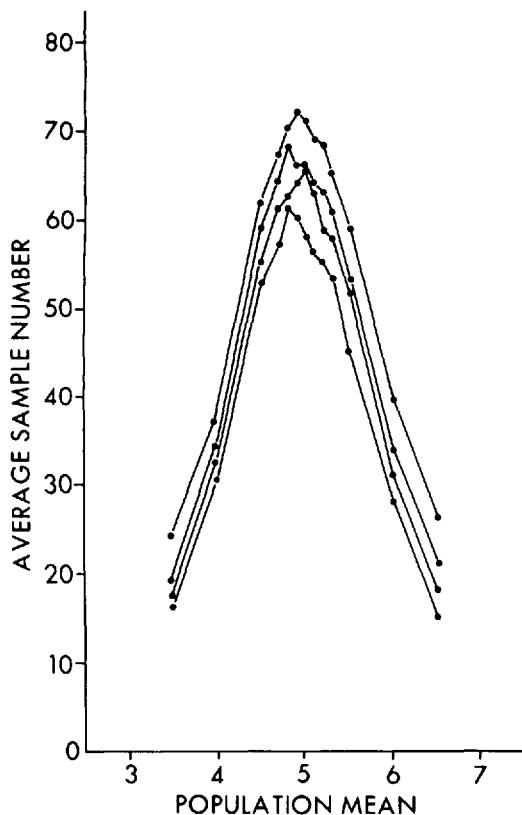
Fig. 3. Average sample number curves for Iwao's decision rule when the minimum number of samples taken before applying the sequential part of the rule was 1, 5, 10, and 20. The upper line corresponds to 20 and the lines decrease in a natural progression. The x_i were distributed as negative binomial variates with $k = 5$ and the null hypothesis was $\mu \leq 5.0$. Sampling was terminated after 80 samples were taken.

corresponds to an α of 0.0383. If the variance had been a constant, then Z would equal 3.541 and α would be 0.0002. The magnitude of these errors will depend on the rate at which the variance changes with the mean and may be confounded by the distributional characteristics of the variates sampled.

An Empirical Example

To illustrate the ideas presented, application of Iwao's sequential procedure was simulated by sampling from a population described by a negative binomial distribution with constant k. The linear relationship between mean crowding ($\overset{*}{\mu}$) and the mean (μ); $\overset{*}{\mu} = a + b\mu$ (Iwao 1968) was used to describe the mean-to-variance relationship: $\sigma^2 = (a + 1)\mu + (b - 1)\mu^2$. In this case, a was set to 0 and b to 1.2. The variable k was specified as $\mu^2/(\sigma^2 - \mu) = 1/(b - 1) = 5$. The working hypothesis was that $\mu < 5$. Hence, we wished to determine

whether the mean of the sample population was greater than or less than 5.

In the simulation, the decision rule was replicated 500 times for each sample population mean. These means ranged from 3.5 to 6.5. In addition, all means were sampled using a variable minimum number of samples ($n$min) prior to initiating the sequential decision process. As $n$min increased, the assumptions of normality were more closely met. Simulations were replicated for minimum sample sizes of 1, 5, 10, and 20. The variable $n$max, which is the maximum number of samples taken before making a terminal decision, was arbitrarily set at 80. The entire decision rule is outlined in Fig. 1. To summarize the simulation results, the following statistics were computed: the frequency of sequential and terminal errors, the proportion of terminal decisions, and the average sample size.

Frequencies of sequential decision errors are presented in Fig. 2. When the population mean is less than 5 or greater than 5 the assumed nominal error rate is less than 0.025. When the sample population mean equals 5, the assumed sequential error rate is 0.05. Obviously, significant deviations from these rates occurred. Small improvements were made in the error rates as $n$min increased from 1 to 20. Of course, the average sample size also increased (Fig. 3). Much of the remaining difference between the nominal error rates and the actual error rates is attributable to the incorrect use of a confidence interval as a sequential test and a non-constant variance. Note that the sequential error probabilities are not symmetrical about 5. When μ is some positive deviation from 5, these error probabilities are less than those for the same negative deviation from 5. This is because the sample population is positively skewed.

The frequency of terminal decisions and the frequency of errors with these decisions are portrayed in Fig. 4. Here, the frequency of errors are adjusted for the number of terminal decisions made to portray them better. As $n$min increased, the percentage of decisions that were terminal decisions increased slightly. Error rates for the terminal decisions are not affected by $n$min since they result from a fixed sample size decision rule. The variation in these curves is due to variation in the random number generators of the simulation. These error rates could be reduced by increasing $n$max.

Discussion

As expected from the theoretical analysis of errors that arise when using Iwao's sequential procedure, the simulation clearly shows that the sequential portion of this decision rule does not operate in the manner proposed by Iwao nor in the manner assumed by those who have used the rule. It is apparent that the procedure cannot be accepted as originally formulated since the actual error rates deviate considerably from the assumed
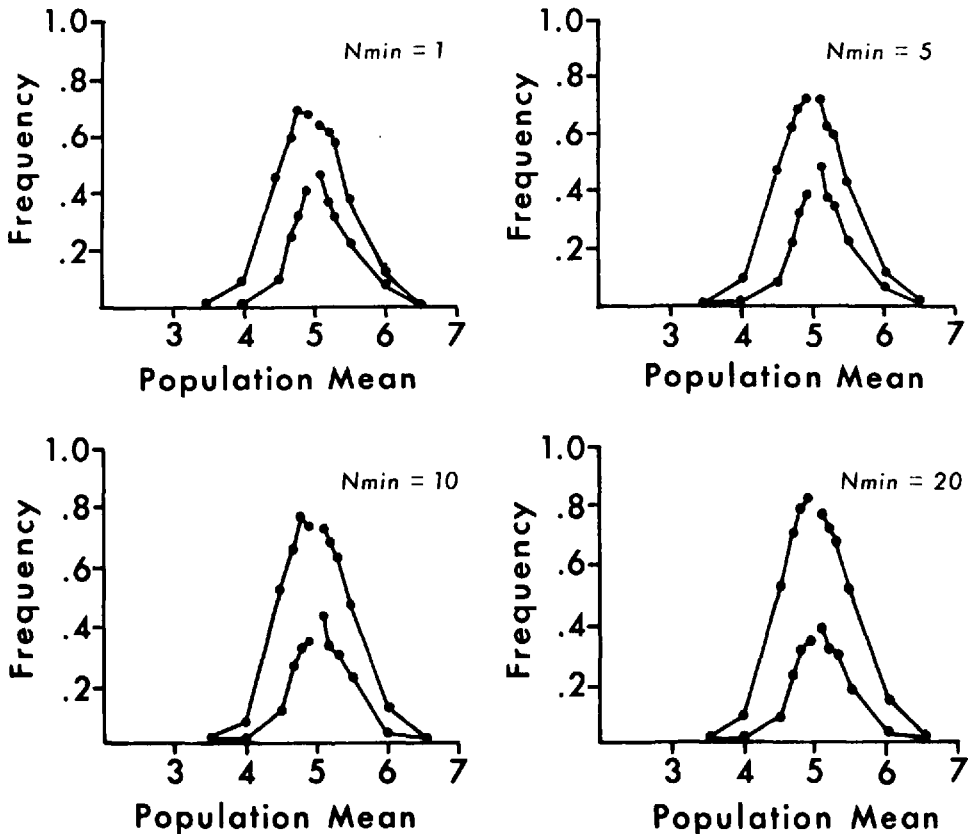
Fig. 4. Frequency of terminal decisions (upper lines) and frequency of terminal decision errors (lower line) for Iwao's decision rule when the minimum number ($n$min) of samples taken before applying the decision rule ranges from 1 to 20. Terminal decision errors were adjusted for the number of terminal decisions made. The x_i were distributed as negative binomial variates with $k = 5$, the null hypothesis was $\mu \leq 5.0$ and terminal decisions were made if the number of samples taken exceeded 80. Values are meaningless for $\mu = 5$ and are not presented.

ones. This does not mean the error rates are unacceptable. It does indicate however, that they must be examined for each instance in which the procedure is to be used and evaluated with respect to the consequences of making an incorrect decision. Such an analysis must rely on simulation. The use of the procedure should be dictated by the costs incurred when making a mistake, i.e., classifying a pest population as less than an economic threshold when it actually exeeds it, the costs incurred through sampling, and the performance of this sequential decision rule compared with other sequential rules. Even in instances when the assumptions of these other rules are violated, they may perform better in terms of error and sample size than Iwao's procedure.

Acknowledgment

References Cited

Coggin, D. L., and G. P. Dively. 1982. Sequential sampling plan for the armyworm in Maryland small grains. Environ. Entomol. 11: 169–172.

Davies, O. L. 1956. Design and analysis of industrial experiments. Hafner, New York.

Fowler, G. W. 1978. Errors in sampling plans based on Wald's sequential probability ratio test. U.S. Dep. Agric. For. Serv. Gen. Tech. Rep. NC-46: 1–13.

Fowler, G. W., and W. G. O'Regan. 1974. One sided truncated sequential t-test: application to natural resource sampling. U.S. Dep. Agric. For. Serv. Res. Pap. PSW-100.

Iwao, S. 1968. A new regression method for analyzing the aggregation pattern of animal populations. Res. Popul. Ecol. 10: 1–20.

Iwao, S. 1975. A new method of sequential sampling

to classify populations relative to a critical density. Ibid. 16: 281–288.

Onsager, J. A. 1976. The rationale of sequential sampling, with emphasis on its use in pest management. U.S. Dep. Agric. Tech. Bull. No. 1526.

Rushton, S. 1950. On a sequential t-test. Biometrica 37: 326–333.

Shaw, P. B., H. Kido, D. L. Flaherty, W. W. Barnett, and H. L. Andris. 1983. Spatial distributions of infestations of Platynota stultana (Lepidoptera: Tortricidae) in California vineyards and a plan for sequential sampling. Environ. Entomol. 12: 60–65.

Wald, A. 1947. Sequential analysis. John Wiley and Sons, New York.