

**Computation of Confidence Limits for
Linear Functions of the Normal Mean and
Variance**

B.F. Lyon
C.E. Land

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831; prices available from 615-576-8401, FTS 626-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**COMPUTATION OF CONFIDENCE LIMITS FOR LINEAR FUNCTIONS OF THE NORMAL
MEAN AND VARIANCE**

B. F. Lyon

Toxicology and Risk Analysis Section
Life Sciences Division
Oak Ridge National Laboratory

C.E. Land

Radiation Epidemiology Branch
National Cancer Institute
Bethesda, MD 20892-7362

Date Issued—September 1999

Prepared by
Toxicology and Risk Analysis Section
Life Sciences Division
Oak Ridge National Laboratory

OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831-6285
managed by
Lockheed Martin Energy Research Corp.,
for the U.S. Department of Energy under contract no. DE-AC05-96OR22464.

PREFACE

A program is described that calculates exact and optimal (uniformly most accurate unbiased) confidence limits for linear functions of the normal mean and variance. The program can therefore also be used to calculate confidence limits for monotone transformations of such functions (e.g., lognormal means). The accuracy of the program has been thoroughly evaluated in terms of coverage probabilities for a wide range of parameter values.

CONTENTS

1. INTRODUCTION	1
2. BACKGROUND	1
3. METHOD	3
4. RESULTS AND APPLICATIONS	10
5. SUMMARY	14
APPENDIX	14
BIBLIOGRAPHY	15

TABLES

1. Example confidence limits and execution times (in seconds) on a Pentium 75MH Processor, Tolerance=1.0E-08.. 17

FIGURES

1. Calculation of lower, one-sided 0.95 confidence limit and lower, one-sided 0.05 limit (which is also the upper one-sided limit of level 0.95) 16

THIS PAGE INTENTIONALLY LEFT BLANK

1. INTRODUCTION

If a known transformation of a random variable X is normally distributed with mean μ and variance σ^2 , then the mean, variance, and any other distributional property of X can be expressed in terms of μ and σ^2 . For example, if X is lognormally distributed, i.e., $X \sim \Lambda(\mu, \sigma^2)$ or (equivalently) $Y = \log(X) \sim N(\mu, \sigma^2)$, then the expected value, variance, median, and mode of X are, respectively, $E(X) = \exp(\mu + \sigma^2/2)$, $var(X) = \exp(2\mu + \sigma^2)(\exp(\sigma^2) - 1)$, $med(X) = \exp(\mu)$, and $mode(X) = \exp(\mu - \sigma^2)$. Exact and optimal (uniformly most accurate unbiased) confidence limit procedures have been developed for linear functions of μ and σ^2 (Land, 1971, 1973) and, therefore, because confidence limits for a parameter are invariant under smooth, monotone transformations of that parameter, for the mean and mode of a lognormal distribution. In fact, the lognormal distribution is the only one whose mean can be expressed as a function of a non-trivial linear combination of μ and σ^2 (Land, 1971), but other functions, including those arising in connection with other normalizing transformations, can be approximated locally by linear functions for which exact limits can be constructed that define approximate limits for the original parametric functions of interest (Land, 1974, 1988).

Tables have been published to facilitate the calculation of confidence limits for arbitrary linear functions of μ and σ^2 (Land, 1975), but their use is often tedious, requiring repeated interpolation and calculation. An unpublished Fortran program to compute confidence limits directly from sample estimates of μ and σ^2 has been available from the second author, and has been used by a number of investigators to analyze lognormal data sets.

The present paper introduces a more efficient computational algorithm, and documents the program for prospective users. An option has been added which makes it easy for the user to generate tables of confidence limits. Finally, the accuracy of the program has been thoroughly evaluated in terms of coverage probabilities for a wide range of parameter values.

2. BACKGROUND

Suppose we have statistically independent estimates $y \sim N(\mu, \sigma^2/\gamma^2)$ and s^2 distributed as σ^2/ν times chi-square with ν degrees of freedom ($s^2 \sim (\sigma^2/\nu)\chi_\nu^2$), where γ and

ν are known. From Land (1971, 1988) the critical value $t_\alpha(m)$ for testing, at level α , the null hypothesis $\mu + \lambda\sigma^2 = m$ against the one-sided alternative, $\mu + \lambda\sigma^2 < m$, for known λ , is found by solving for $t_\alpha(m)$ the integral equation,

$$\int_{-\infty}^{t_\alpha(m)} f_m(t) dt = \alpha \int_{-\infty}^{\infty} f_m(t) dt$$

where

$$f_m(t) = (\nu + t^2)^{-(\nu+1)/2} \exp(w_m t / (\nu + t^2)^{1/2})$$

and

$$w_m = -\lambda \gamma^2 (\nu s^2 / \gamma^2 + (y - m)^2)^{1/2}$$

The corresponding confidence interval of level $1 - \alpha$ for $\mu + \lambda\sigma^2$ is the set of values m such that $T(m) = \gamma(y - m)/s \geq t_\alpha(m)$. The upper confidence limit is the largest such value m , and is found by solving the equation, $T(m) = t_\alpha(m)$, for m . A one-sided lower confidence limit of level $1 - \alpha$ is the same as an upper limit of level α .

The two-sided case is somewhat more complicated. The null hypothesis $\mu + \lambda\sigma^2 = m$ is rejected in favor of the alternative $\mu + \lambda\sigma^2 \neq m$ if $T(m) < t_{1,\alpha}(m)$ or $T(m) > t_{2,\alpha}(m)$, where the critical values $t_1 = t_{1,\alpha}(m)$ and $t_2 = t_{2,\alpha}(m)$ are defined by the two integral equations

$$\int_{t_1}^{t_2} f_m(t) dt = (1 - \alpha) \int_{-\infty}^{\infty} f_m(t) dt \tag{2.1}$$

$$\int_{t_1}^{t_2} \frac{t}{\sqrt{\nu + t^2}} f_m(t) dt = (1 - \alpha) \int_{-\infty}^{\infty} \frac{t}{\sqrt{\nu + t^2}} f_m(t) dt$$

The level $1 - \alpha$ confidence interval for $\mu + \lambda\sigma^2$, therefore, is the set of numbers m such that $t_{1,\alpha}(m) \leq T(m) \leq t_{2,\alpha}(m)$. The upper limit is found by solving for m the equation, $T(m) = t_{1,\alpha}(m)$, and the lower limit by solving $T(m) = t_{2,\alpha}(m)$.

From (2.1), it follows that, for fixed s^2 and ν , any level $1 - \alpha$ confidence limit for $\mu + \lambda\sigma^2$, for arbitrary y , s^2 , γ , and $\lambda > 0$, is equal to $y + (\nu + 1)/(2\lambda\gamma^2) m^*$, where m^* is the corresponding confidence limit for $\mu + (1/2)\sigma^2$ given a simple, single-sample model with sample mean $y^* = 0$, $\gamma^* = (\nu + 1)^{1/2}$, and variance $s^{*2} = 2\lambda\gamma s^2/(\nu + 1)^{1/2}$ (Land, 1973). Moreover, given $y = 0$, confidence limits are symmetric with respect to changes in sign of λ , in the sense that a level $1 - \alpha$ upper limit for $\mu - \lambda\sigma^2$ is identical to the corresponding level $1 - \alpha$ lower limit for $\mu + \lambda\sigma^2$, with a change of sign. Thus, it is sufficient to develop computational procedures for the single-sample case with $y = 0$, $\gamma = (\nu + 1)^{1/2}$, and $\lambda = 1/2$. For that case and sampling model, $T(m) = -(\nu + 1)^{1/2} m/s$, and

$$w_m = - \left(\frac{\nu + 1}{2} \right) (\nu s^2/(\nu + 1) + m^2)^{1/2}.$$

3. METHOD

For both the one-sided and two-sided limits, the basic method of approximation is the secant method. The application to the one-sided case is straightforward. Here, we consider only the case of a one-sided, lower confidence limit of level $1 - \alpha$ since it is also a one-sided upper limit of level α . A solution (there is only one) is found for $G(T(m)) = 1 - \alpha$, where

$$G(T(m)) = \frac{\int_{-\infty}^{T(m)} f_m(t) dt}{\int_{-\infty}^{\infty} f_m(t) dt} \quad (3.1)$$

The two-sided case is more complex, and requires some additional notation. The upper limit of level $1 - \alpha$ is that number m such that $T(m) = t_{1,\alpha}(m)$; i.e., it is necessary to find the unique pair of numbers m and t_2 such that both $G_1(T(m), t_2) = 1 - \alpha$ and $G_2(T(m), t_2) = 1 - \alpha$, where

$$G_1(T(m), t_2) = \frac{\int_{T(m)}^{t_2} f_m(t) dt}{\int_{-\infty}^{\infty} f_m(t) dt} \quad (3.2)$$

$$G_2(T(m), t_2) = \frac{\int_{T(m)}^{t_2} t(\nu + t^2)^{-1/2} f_m(t) dt}{\int_{-\infty}^{\infty} t(\nu + t^2)^{-1/2} f_m(t) dt}$$

Similarly, the lower limit is that number m such that $T(m) = t_{2,\alpha}(m)$, and is obtained by finding the unique pair of numbers t_1 and m such that both $G_1(t_1, T(m)) = 1 - \alpha$ and $G_2(t_1, T(m)) = 1 - \alpha$. Note that the upper and lower limits are obtained separately.

For the one-sided limit, the secant method alone is used to find the zero of the function $G(T(m)) - \alpha$. For the upper two-sided limit, the secant method is applied to find the zero of the function $G_2(T(m), t_2) - (1 - \alpha)$, subject to the constraint that $G_1(T(m), t_2) = 1 - \alpha$; for a given m and α , this t_2 , which depends on m and α , is found by numerically inverting the integral defining G_1 . The lower two-sided limit is found in the same manner, after switching the order of the arguments of G_1 and G_2 .

For $\nu = 2$ all integrals are evaluated directly, while in other cases they are approximated using the adaptive quadrature method described in Burden and Faires (1989). With this method, subintervals are determined so that the integral is approximated with the desired accuracy using Simpson's rule on each subinterval. This method is generally faster than simpler integration methods to achieve that same accuracy because the ultimate subdivision that is used need not be uniformly spaced over the entire interval of integration; the subintervals can be selected based on the desired accuracy and the variability of the function to be integrated (for a more complete description, see Burden and Faires 1989).

Although the method is in general straightforward, several numerical considerations warrant discussion.

Conversion to integral over $(-1, 1)$

By the change of variable $x(t) = t/\sqrt{\nu + t^2}$, and for $-\infty < A < B < \infty$, simple substitutions show that

$$\int_A^B f_m(t) dt = \int_{A'}^{B'} g_m(x) dx$$

where $A' = x(A) = A/\sqrt{\nu + A^2}$, $B' = x(B)$, and

$$g_m(x) = \nu^{-\nu/2} (1 - x^2)^{\nu/2-1} \exp(w_m x)$$

for $-1 < x < 1$. Also,

$$\int_{-\infty}^{\infty} f_m(t) dt = \sqrt{\pi} \nu^{-\nu/2} \left(\frac{w_m}{2}\right)^{-(\nu-1)/2} \Gamma\left(\frac{\nu}{2}\right) I_{(\nu-1)/2}(w_m)$$

where $I_\nu(u)$ is the modified Bessel of the third kind.

Similarly

$$\int_A^B t(\nu + t^2)^{-1/2} f_m(t) dt = \int_{A'}^{B'} x g_m(x) dx$$

and

$$\int_{-\infty}^{\infty} t(\nu + t^2)^{-1/2} f_m(t) dt = -\sqrt{\pi} \nu^{-\nu/2} \left(\frac{w_m}{2}\right)^{-(\nu-1)/2} \Gamma\left(\frac{\nu}{2}\right) I_{(\nu+1)/2}(w_m)$$

These formulas are used for the integrals in the functions G , G_1 and G_2 . When possible, the Bessel function is evaluated using the algorithm written by Cody and Stolz (Cody and Stolz, 1989).

With the exception of the cases when the Bessel function routine is used, the integrals are evaluated using the adaptive quadrature method described above.

Bessel Function Considerations

For numerical reasons, the product $e^{-u} I_\nu(u)$ is evaluated instead of $I_\nu(u)$ alone. Nevertheless, numerical problems still arise and the integral must be evaluated directly in these cases.

The first such case occurs for large values ($\gg 1000$) of the argument $u = w_m$, even when the product $e^{-u} I_\nu(u)$ is of moderate size. In the program, the integral is evaluated directly when the argument is larger than 800; a different cutoff value can be selected by modifying the associated control file (“Conf95.opt”).

A second case occurs when the argument u is too small relative to the order, $(\nu - 1)/2$. This is because $I_\nu(u)$ vanishes (quite quickly) as the ratio u/ν goes to zero. For this reason, if the ratio of the argument and the order is less than 0.26 (again, not necessarily an optimal value but sufficient) then the Bessel function is not used and the integral is evaluated directly.

Normalization of Integrand

In general, $g_m(x)$ can vary widely for different values of ν , s and ξ . To ameliorate this, $g_m(x)$ is normalized by its maximum value on the interval $[-1, 1]$, which occurs at

the point

$$x_{max} = \frac{\nu/2 - 1 - \sqrt{(\nu/2 - 1)^2 + w_m^2}}{w_m}.$$

For extreme values of s (large and small) and large ν , exponential overflow prevents the evaluation of the $g_m(x)$ to perform the normalization itself. This is circumvented by scaling $f_m(t)$ by the exponential of $w(s)(\nu+1)$, where $w(s)$ is the linear function that maps 0.1 to -1 and 7 to 0. The same normalization process is used successfully for $x g_m(x)$ when calculating the two-tailed limits.

Inversion of Integral

After the scale change from t to x , and defining $X(m) = T(m)/\sqrt{\nu + T(m)^2}$, finding $x_1 = t_1/\sqrt{\nu + t_1^2}$ for fixed m is equivalent to finding x_1 such that the integral from $X(m)$ to x_1 of $g_m(x)$ is equal to some constant, C . Inversion of the integral is performed by calling the adaptive quadrature integration procedure as if to approximate the integral over the entire interval $[X(m), 1]$. The order of the calculation is such that the current subinterval over which the integral is being approximated is always the left-most interval not yet considered. In this way, the integral is built up from the left. The procedure is stopped when the current value of the integral is sufficiently close to the desired value, at which time the current upper limit of the right-most subinterval approximated is the desired value x_1 . If the approximation exceeds the desired value by more than the desired accuracy, then a finer subdivision is used. Determination of $x_2 = t_2/\sqrt{\nu + t_2^2}$ is similar, except that the integrals over the right-most subintervals are approximated first.

Limits of Integration

The functions integrated over the region $[-1, 1]$ have a single maximum, and generally vanish rapidly as the distance between x and the maximum increases, especially for large s and ν . Depending on the particular interval of integration, this fact can cause the adaptive quadrature routine to terminate prematurely and return a value of 0. The difficulty occurs because the routine terminates when the difference between the approximate integral (using a 3-point Simpson's rule) over an interval $[a, b]$, and the sum of the approximate integrals over $[a, (a+b)/2]$ and $[(a+b)/2, b]$, is sufficiently small. If the points a and b lie too far from the maximum, then the approximation to all integrals may be so small that the termination

condition is met. If the interval does not contain the maximum, then this is not a problem in practice. However, additional steps are taken to avert this from happening when the interval does contain the maximum.

For the one-sided limits, the evaluation of integrals in the definition of G (3.1) is equivalent to evaluating

$$\frac{\int_{-1}^b F(x) dx}{\int_{-1}^1 F(x) dx}$$

where $F(x) = g_m(x)$ and $b \in [-1, 1]$. The denominator is evaluated either using the Bessel function routine or by the adaptive quadrature routine, as discussed above. If the adaptive quadrature routine is used, then to avoid premature (and inaccurate) termination of the routine, the integral is evaluated as

$$\int_{-1}^1 F(x) dx = \int_{-1}^{(x_{max}-1)/2} F(x) dx + \int_{(x_{max}-1)/2}^{x_{max}} F(x) dx + \int_{x_{max}}^{(x_{max}+1)/2} F(x) dx + \int_{(x_{max}+1)/2}^1 F(x) dx$$

and the adaptive quadrature routine is called on each subinterval. This successfully prevents premature termination of the adaptive quadrature routine for the cases of interest.

If $b < x_{max}$ then the adaptive quadrature routine is applied directly to the numerator. However, if $b > x_{max}$ then the ratio of integrals is evaluated as

$$1 - \frac{\int_{b}^1 F(x) dx}{\int_{-1}^1 F(x) dx}$$

This is done to avoid redundant calculations in the numerator and denominator.

For the two-sided limits, evaluation of $G_1(a, b)$ and $G_2(a, b)$ (equation 3.2) involves approximating expressions of the form

$$\frac{\int_{a'}^{b'} F(x) dx}{\int_{-1}^1 F(x) dx}$$

where $F(x) = g_m(x)$ or $x g_m(x)$ and $a' = a/\sqrt{\nu + a^2}$, $b' = b/\sqrt{\nu + b^2}$. For $|\alpha - 0.5| > 0.25$, it is expected that it will generally be the case that $a' < x_{max} < b'$. To avoid redundant calculations as discussed above, the ratio of integrals is evaluated as

$$1 - \frac{\int_{-1}^{a'} F(x) dx + \int_{b'}^1 F(x) dx}{\int_{-1}^1 F(x) dx}$$

The adaptive quadrature routine is applied directly to the integrals in the numerator, while the denominator is evaluated as

$$\int_{-1}^1 F(x) dx = \int_{-1}^{x_{max}} F(x) dx + \int_{x_{max}}^{b'} F(x) dx + \int_{-1}^{a'} F(x) dx + \int_{b'}^1 F(x) dx$$

This approach is successful for all cases of interest, even if $|\alpha - 0.5| \leq 0.25$.

Initial Guesses for the Secant Method for One-Sided Limits

Two initial guesses are required for the secant method.

For the one-sided limits, the initial guesses are selected so that in general the sequence of approximations generated by the secant method is always on one side of the root. By satisfying this property, the sequence of approximations does not oscillate about the root, which can result in divergence of the secant method, nor are additional (time-consuming) measures required to address intermediate situations that may lead to possible divergence. Figure 1 illustrates the general behavior of the function $G(T(m))$, which is a decreasing function of m . In the graph, $G(T(m)) = 0.95$ gives the lower one-sided level 0.95 confidence limit and $G(T(m)) = 0.05$ gives the upper one-sided limit 0.95 confidence limit. The two initial guesses are chosen towards the middle of the graph, and then the secant method generates approximations that gradually move out to the root. This is done using an approximate confidence interval method suggested by David Cox (Land, 1972), which uses the fact that $y + s^2/2$ is asymptotically normal with mean $\mu + \sigma^2/2$ and standard deviation $\beta(\sigma, \nu) = \sqrt{\sigma^2/(\nu + 1) + \sigma^4/(2(\nu + 2))}$. Thus, an approximate upper confidence limit of level α for $\mu + \sigma^2/2$, when $y = 0$, is $s^2/2 + \Phi(\alpha)\beta(s, \nu)$. Usable initial guesses, closer to the middle of the graph than the solution, are of the form $s^2/2 + C \Phi(\alpha)\beta(s, \nu)$. We found that

pairs of initial guesses corresponding to $C = 0$ and $C = 0.15$ worked well for the tested range of values ν , α , and s , providing enough separation between the guesses to ensure convergence. The percentiles $\Phi(\alpha)$ are calculated with accuracy 10^{-4} using Algorithm 66 available from STATLIB at Carnegie Mellon University, which computes the cumulative distribution of the standard normal distribution, and a simple root finding procedure.

Other Modifications of Secant Method

For the one-sided limits, the secant method as described above is used to approximate the root. However, for the two-sided limits additional steps are necessary.

First, guesses for the two-tailed limits of level α are estimated based on the assumption that the one-sided limits of level $1 - (1 - \alpha)/2$ should be “close,” although some fine-tuning is necessary. For the lower two-tailed limit, the first guess is the one-sided lower confidence limit (calculated using the algorithm described here) of level $(1 - \alpha)/2$. The second guess is the first guess divided by 1.1. For the upper two-tailed limit, it was found that the upper one-sided limit of level $1 - (1 - \alpha)/2$ led to numerical problems, and so the first guess is set at the slightly larger value corresponding to level $1 - (1 - \alpha)/1.1$, and the second guess is 1.05 times the first. The initial guesses are then modified as necessary to ensure that the root lies between the two guesses. Then, at each step the approximation generated by the secant method is checked to see if it lies between the two previous approximations. If it does not, then the midpoint of the two previous approximations is used as the next approximation. Finally, the approximations are retained so that the root always lies between them. These extra steps increase the execution time, but avoid numerical instabilities that are associated with use of the secant method alone.

In the course of performing verification of the algorithms (section 4), it was necessary to use values of s well below 0.1, which is the lower bound of previously published values. In these cases, calculation of the upper two-sided limits requires finding the zero of a function that has a large derivative at the root. The approximations generated by the secant method, while still in the interval containing the root, converge quite slowly, and usually the maximum number of iterations set by the user is exceeded. For this reason, when the maximum number of iterations is exceeded when calculating the two-sided limits, the bisection method alone is used until the difference between approximations is below a

value set by the user.

4. RESULTS AND APPLICATIONS

Algorithm Verification

Although the results of the program have been checked with published tables (Land 1971) as well as the earlier program, an additional independent method was also used to verify the accuracy of the predicted values. This method, described below, is considered to be a more reliable means of assessing the accuracy of the program.

If a sample is to be drawn from a normal distribution with mean and variance μ and σ^2 , respectively, then the set of all possible values y , s^2 is equivalent to the set of all values

$$y_p = \mu + (\sigma/\sqrt{n}) \Phi(p),$$

$$s_q^2 = \sigma^2 \chi_{n-1}^2(q)/(n-1)$$

for $p, q \in [0, 1]$, where $\Phi(p)$ is the p th percentile of the standard normal distribution and $\chi_{n-1}^2(q)$ is the q th percentile of the chi-squared distribution with $n-1$ degrees of freedom. Let $C(p, q; \mu, \sigma, \alpha)$ denote the upper confidence limit calculated using the algorithm described here of level α on $\mu + \frac{1}{2}\sigma^2$ using the sample estimates y_p and s_q . For a given level of confidence α , the set of pairs (p, q) for which the calculated upper confidence limit is below the true value $\mu + \frac{1}{2}\sigma^2$, i.e., for which

$$C(p, q; \mu, \sigma, \alpha) \leq \mu + \frac{1}{2}\sigma^2,$$

should be a region with area $1-\alpha$. Similarly, if $C_L(p, q; \mu, \sigma, \alpha)$ and $C_U(p, q; \mu, \sigma, \alpha)$ denote the lower and upper two-sided confidence limits for $\mu + \frac{1}{2}\sigma^2$, then the set of pairs (p, q) for which

$$C_U(p, q; \mu, \sigma, \alpha) \leq \mu + \frac{1}{2}\sigma^2$$

and

$$C_L(p, q; \mu, \sigma, \alpha) \geq \mu + \frac{1}{2}\sigma^2$$

should have area $1-\alpha$. Since $C(p, q; \mu, \sigma, \alpha)$, $C_L(p, q; \mu, \sigma, \alpha)$, and $C_U(p, q; \mu, \sigma, \alpha)$ are nondecreasing functions of p (i.e., the confidence limits are nondecreasing functions of the

sample mean y), the boundaries of the “failure” regions can be determined, and the areas thereby estimated.

For a given q , let $p(q)$, $p_L(q)$, and $p_U(q)$ denote the “critical” values of p such that

$$\begin{aligned} C(p(q), q; \mu, \sigma, \alpha) &= \mu + \frac{1}{2}\sigma^2, \\ C_L(p_L(q), q; \mu, \sigma, \alpha) &= \mu + \frac{1}{2}\sigma^2, \\ &\text{and} \\ C_U(p(q), q; \mu, \sigma, \alpha) &= \mu + \frac{1}{2}\sigma^2, \end{aligned}$$

respectively. Then we should have

$$\begin{aligned} \int_0^1 p(q) dq &= 1 - \alpha \\ \int_0^1 p_L(q) dq + \left(1 - \int_0^1 p_U(q) dq\right) &= 1 - \alpha \end{aligned}$$

This approach avoids the problem of sampling variation that accompanies Monte Carlo methods of validation. Further, since the confidence limits are additive in the sample mean y , one can show (see Appendix for details) that

$$p(q) = \Phi^{-1} \left(\frac{(\mu + \frac{1}{2}\sigma^2 - C(0.5, q; \mu, \sigma, \alpha))}{\sigma/\sqrt{n}} \right) \quad (4.1)$$

where Φ^{-1} is the inverse of the standard normal distribution function. A similar relationship holds between $p_L(q)$ and $C_L(p, q; \mu, \sigma, \alpha)$ and between $p_U(q)$ and $C_U(p, q; \mu, \sigma, \alpha)$ in the two-sided case.

The algorithm was checked in this fashion with $\mu = 0$ for the one-sided limits over the following range: $\nu = 2, 3, 5, 10, 100, 1000$, $\sigma = 0.1, 1, 10$ and $\alpha = 0.001, 0.005, 0.01, 0.1, 0.9, 0.95, 0.99, 0.995, 0.999$, with the integrals approximated with tolerance 10^{-6} , and the confidence limits themselves approximated with tolerance 10^{-8} . Similarly, the two-sided limits were checked for $\nu = 2, 3, 5, 10, 100, 1000$, $\sigma = 0.1, 1, 10$ and $\alpha = 0.9, 0.95, 0.99, 0.995$, and 0.999 . In all cases, the absolute errors between the integrals and $1 - \alpha$ were at most 10^{-6} . It should be noted that with this method there are many cases where the sample s is much less than 0.1 or greater than 10 (the previous limits of published values),

depending on the value of q and σ . In particular, for a given ν , s , and the tolerance ε used here of 10^{-6} , s ranges over the interval $(\sigma\sqrt{u_\varepsilon(\nu)/\nu}, \sigma\sqrt{u_{1-\varepsilon}(\nu)/\nu})$. For $\nu = 2$, range of values for s is $(0.0001, 37.6)$, while for $\nu = 1000$ it is the interval $(0.09, 11.1)$. In all cases, the chi-squared distribution was calculated using routines from the CDFLIB library of Fortran routines (Brown and Lovato, 1993).

Although execution times vary depending on the relative error desired and particular values of ν , s and α , some sample output and timings (in seconds) required for the method to converge with tolerance 10^{-8} are provided in Table 1 (these calculations were performed on a personal computer with a Pentium 75MH processor). The two-tailed limits take about 10 times longer to compute than the one-tailed limits. This is because two one-tailed limits are computed as initial guesses for the secant method, and because an integral must be inverted during each iteration.

Examples

The program provides exact and optimal (uniformly most accurate unbiased) confidence limits for linear functions of the normal mean μ and variance σ^2 and, therefore, of monotone transformations of such functions such as the lognormal mean and other lognormal moments about zero, $EX^k = \exp(k\mu + k^2\sigma^2/2)$, and the lognormal mode, $MX = \exp(\mu - \sigma^2)$. Suppose that we have a simple random sample of size 20 on a normally distributed random variable, which is the logarithm of the variable of interest, X . Suppose that the sample mean and standard deviation in the logarithmic scale are $y = 2.0$, and $s = 0.5$. The exact and optimal 95% upper confidence limit for $EX = \exp(\mu + \sigma^2/2)$ is the exponential of the value obtained from the program with input values y , s , $\lambda = 0.5$ (the default value), $\nu = 19$, and $\gamma^2 = 20$ (the default value given $\nu = 19$). That is, the 95% upper confidence limit for EX is $\exp(2.334) = 10.53$. The corresponding limit for $MX = \exp(\mu - \sigma^2)$ is $\exp(1.950) = 7.03$.

Other, intrinsically nonlinear, parametric functions of interest include the lognormal central moments,

$$m_2 = \exp(2\mu + \sigma^2) (\exp(\sigma^2) - 1)$$

$$m_3 = \exp(3\mu + 1.5\sigma^2) (\exp(\sigma^2) - 1)^2 (\exp(\sigma^2) + 2),$$

etc., and the means of variates that can be transformed to normality other than by the logarithmic transformation; e.g., by the square root, cube root, arcsine(square root), and hyperbolic arcsine (square root) transformations; i.e.,

$$\begin{aligned} E(Y^2) &= \mu^2 + \sigma^2, \\ E(Y^3) &= \mu^3 + 3\mu\sigma^2, \\ E(\sin^2 Y) &= \frac{1}{2} (1 - \cos(2\mu)\exp(-2\sigma^2)), \\ E(\sinh^2 Y) &= \frac{1}{2} (\cosh(2\mu)\exp(\sigma^2) - 1), \end{aligned}$$

respectively, where Y is normally distributed with mean μ and variance σ^2 .

Exact solutions for most intrinsically nonlinear functions of μ and σ^2 are not known, but the functions can be approximated locally by linear functions, yielding approximate confidence limits. For example, given estimates y for μ and s^2 for σ^2 , the function $\mu^2 + \sigma^2$ can be approximated locally by the linear truncation of its Taylor series expansion about the point $\mu = y$, $\sigma^2 = s^2$: $y^2 + s^2 + 2y(\mu - y) + (\sigma^2 - s^2) = -y^2 + 2y(\mu + \lambda\sigma^2)$, where $\lambda = 1/(2y)$. Suppose once more that $y = 2.0$ and $\sigma = 0.5$ represent the sample mean and standard deviation of a normal random sample of size 20. Then $\nu = 19$ and $\gamma^2 = 20$, and $\lambda = 0.25$. Using the program, two-sided confidence limits at level 0.95 for $\mu + .25\sigma^2$ are 1.835 and 2.312, which correspond to 3.340 and 5.248, respectively, for the quantity $-y^2 + 2y(\mu + \lambda\sigma^2)$. Thus, 3.340 and 5.248 are approximate 95% limits for $\mu^2 + \sigma^2$. A refinement of the method, discussed in Land (1988), involves a second iteration in which the function is expanded about the maximum likelihood estimates of μ and σ^2 constrained to the approximate confidence limit in the first iteration (e.g., the maximum likelihood point on the curve $\mu^2 + \sigma^2 = 5.248$).

The linearization method was evaluated by Monte Carlo simulation (Land 1974) for four of the mean value functions given above; that is, for $E(Y^2) = 10$ for σ^2 between 0.25 and 0.40, $E(Y^3) = 10$ and $E(\sin^2 Y) = 10$ for σ^2 between 0.01 and 0.20, and $E(\sinh^2 Y) = 10$ for σ^2 between 0.01 and 1.0 (Land 1974). For these parametric values and ν between 10 and 1000, coverage probabilities were estimated at between 0.892 and 0.909 for the one-sided lower limits and between 0.892 and 0.913 for upper limits at nominal confidence level 0.90.

5. SUMMARY

The program described here provides exact and optimal (uniformly most accurate unbiased) confidence limits for linear functions of the normal mean μ and variance σ^2 . The program (including source code) is available without charge from the authors.

APPENDIX

In this short appendix a proof is provided for the identity shown in equation (4.1). Only the one-sided interval case is discussed, but the same proof holds for the lower and upper limits in the two-sided case. In Section 2 it is noted that, if y is the sample mean, any level $1 - \alpha$ confidence limit for $\mu + \lambda s^2$ is equal to $y + f$ where f does not depend on y (Land, 1973). Let C denote the confidence limit using the sample mean y_1 . Then the calculated confidence limit using the sample mean y_2 will obviously be equal to $C + y_2 - y_1$. If we write $y_i = \mu + \Phi(p_i)\sigma/\sqrt{n}$, where $\Phi(p_i)$ is the p_i th normal quantile, then, using the notation defined above, we have that

$$\begin{aligned} C(p_2, q; \mu, \sigma, \alpha) &= C(p_1, q; \mu, \sigma, \alpha) + y_2 - y_1 \\ &= C(p_1, q; \mu, \sigma, \alpha) + \sigma/\sqrt{n} (\Phi(p_2) - \Phi(p_1)) \end{aligned}$$

In particular, letting $p_1 = 0.5$ (in which case $\Phi(p_1) = 0$),

$$C(p_2, q; \mu, \sigma, \alpha) = C(0.5, q; \mu, \sigma, \alpha) + \Phi(p_2)\sigma/\sqrt{n}$$

This identity holds for arbitrary p_2 . By definition, the “critical” $p(q)$ satisfies

$$C(p(q), q; \mu, \sigma, \alpha) = \mu + \frac{1}{2}\sigma^2$$

Using this identity in the previous equation shows that

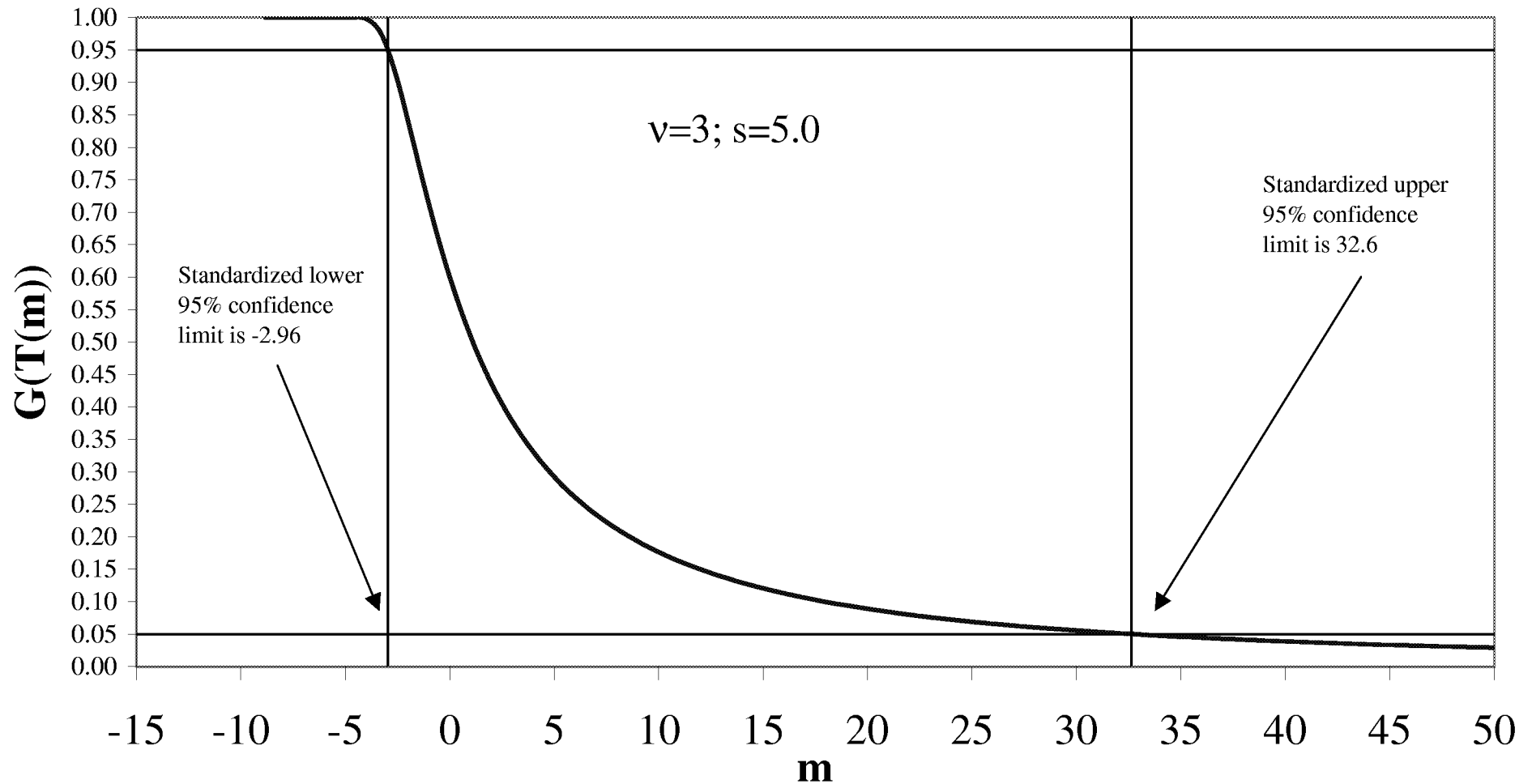
$$\mu + \frac{1}{2}\sigma^2 = C(0.5, q; \mu, \sigma, \alpha) + \Phi(p(q))\sigma/\sqrt{n}$$

and solving for $p(q)$ yields the identity in equation (4.1).

BIBLIOGRAPHY

- Brown, B.W. and J. Lavoto (1993). CDFLIB Library of Fortran Routines for Cumulative Distribution Functions, Inverses, and Other Parameters. Available from StatLib, a system for distributing statistical software, datasets, and information by electronic mail, FTP and WWW (lib.stat.cmu.edu).
- Burden, R.L. and J.D. Faires (1989). *Numerical Analysis*, Fourth Edition, PWS-Kent Publishing Company, Boston.
- Cody, W. J. and L. Stoltz (1989), Fortran routine to calculate Bessel functions $I_{N+\alpha}(x)$ for non-negative argument x , and non-negative order $N+\alpha$. Available from the Netlib collection of mathematical software, papers, and databases (www.netlib.org).
- Land, C.E. (1971). "Confidence Intervals for Linear Functions of the Normal Mean and Variance," *Annals of Mathematical Statistics*, 43, 1187-1205.
- (1973). "Standard Confidence Limits for Linear Functions of the Normal Mean and Variance," *Journal of the American Statistical Association*, 68, 960-963.
- (1974). "Confidence interval estimation for means after data transformations to normality," *Journal of the American Statistical Association*, 69, 795-802.
- (1975). "Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance," *Selected Tables in Mathematical Statistics*, Volume III, 385-419.
- (1988). "Hypothesis Tests and Interval Estimates," in *Lognormal Distributions, Theory and Applications*, E.L. Crow and K. Shimizu, eds. Marcel Dekker, 1988.

Figure 1. Calculation of lower, one-sided 0.95 confidence limit and lower, one-sided 0.05 limit (which is also the upper one-sided limit of level 0.95).



DISTRIBUTION

- 1-50. B.F. Lyon
- 51. Central Research Library
- 52. Document Reference Center
- 53. ORNL Patent Section
- 54-55. ORNL Laboratory Records
- 56. ORNL Laboratory Records, RC

- 57. Office of Assistant Manager, Energy Research and Development, DOE-ORO, P.O. Box 2008, Oak Ridge, TN 37831-6269
- 58-59. Office of Science and Technical Information, P.O. Box 62, Oak Ridge, TN 37831