Skewness-Adjusted Confidence Intervals in Stratified Biological Surveys

David FLETCHER and Raymond WEBSTER

The purpose of this paper is to make available to applied statisticians and biologists a simple formula to calculate a skewness-adjusted confidence interval for a population mean from a stratified random sample. The resulting interval is more reliable, in terms of coverage error, than the simple "plus and minus two standard errors" approach. The theory underlying this new interval was outlined by Hall. Our aim here is to promote the application of this theory to stratified biological surveys. In doing this, we compare the formula given with use of the bootstrap, a more computer intensive procedure. The results of a large-scale cockle survey are used to generate an example of a skewed population; a simulation study using this population indicates the extent to which these intervals can be more reliable. Further improvement in the coverage error can be obtained by using the bootstrap in combination with this formula. The extent to which these methods differ is illustrated for a recent trawl survey. Finally, the use of a pair of confidence curves is suggested as a means of overcoming some of the limitations inherent in using confidence intervals for skewed populations.

Key Words: Biomass estimation; Bootstrap; Confidence interval; Error rate; Lognormal distribution; Skewness.

1. INTRODUCTION

In many biological surveys, a stratified sample is used to estimate animal or plant biomass in a given area. The stratified sample mean is used to provide an estimate of the mean biomass density, which is multiplied by the area to give an estimate of total biomass. Existing methods for calculating a confidence interval around this estimate implicitly assume that the stratified sample mean has a near-normal sampling distribution. In biological surveys this may be very far from the truth: clumping in the spatial distribution of animal and plant species can lead to a high degree of skewness in the population, which can then carry over to the distribution of the stratified sample mean. For a positively skewed measurement, the usual confidence intervals will tend to underestimate both confidence limits.

A general alternative in such situations is to employ a bootstrap procedure (Hall 1992). In this context, for example, we can do this by resampling the studentized version

David Fletcher is a Lecturer, and Raymond Webster is a Tutor at the Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin, New Zealand.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 1, Number 1, Pages 120–130

of the stratified sample mean. An alternative approach is suggested here, as it achieves the same degree of improvement without the need for resampling. It is based on a general method for skewness adjustment suggested by Hall (1992).

2. Stratified Sampling

The notation we adopt is as follows. A population of N units is divided into L strata, and in stratum h we have:

N_h	total number of stratum units
n_h	number of stratum units in the sample
μ_h, σ_h^2	stratum mean and variance
$W_h = N_h/N$	stratum weight
$\bar{x}, s_h^2$	sample stratum mean and variance

It is assumed that the sampling fraction in each stratum is negligible (i.e., that $n_h \ll N_h$ for all h). This will be true in many biological surveys and has the advantage of simplifying the technical discussion. We also focus discussion on producing a confidence interval for the population mean μ, as analogous results automatically follow for the population total.

The stratified sample mean and its standard error are given by

$$\bar{x}_{st} = \sum_h W_h \bar{x}_h \quad \text{and} \quad \text{S.E.}(\bar{x}_{st}) = \sqrt{\sum_h W_h^2 \sigma_h^2 / n_h}.$$

The studentized version of this is written as $T_{st} = (\bar{x}_{st} - \mu)/\text{se}(\bar{x}_{st})$, where

$$\text{se}(\bar{x}_{st}) = \sqrt{\sum_h W_h^2 s_h^2 / n_h}$$

is the estimated standard error of $\bar{x}_{st}$.

If the population is normally distributed within each stratum, T_{st} has approximately a t-distribution with effective degrees of freedom given by

$$n_e = \left(\sum_h g_h s_h^2 \right)^2 \Big/ \sum_h \frac{g_h^2 s_h^4}{(n_h - 1)},$$

where $g_h = N_h(N_h - n_h)/n_h$ (see Cochran 1977, p. 96). An approximate $(100 - \alpha)\%$ confidence interval for μ is then given by

$$\bar{x}_{st} \pm t(n_e, \alpha)\text{se}(\bar{x}_{st}), \tag{2.1}$$

where $t(n_e, \alpha)$ is the upper $(\alpha/2)\%$ point for the t-distribution with $[n_e]$ degrees of freedom.

When the population is highly skewed within some or all of the strata, the distribution of T_{st} will also be skewed. If L is large, or all the n_h are large, a central limit argument implies that the level of skewness in T_{st} will be small enough for $\bar{x}_{st} \pm 2\text{se}(\bar{x}_{st})$ to provide reasonable 95% confidence limits. An obvious problem with this simple approach is in deciding when these limits are reasonable. We now consider two alternative methods that are generally more reliable.

3. ALTERNATIVE METHODS

A common method of coping with skewness is to use standard methods of analysis on a log-scale and then back-transform the resulting confidence limits [see Crow and Shimizu (1988) for a thorough survey]. This approach could be used within each stratum, but does not lend itself so easily to the case of the whole population. Furthermore, neither of the methods to be discussed later make an assumption of lognormality. We do not therefore consider any log-transformation approach.

In considering the reliability of any method, we can make use of an expression for the cdf of T_{st} provided by Abramovitch and Singh (1985). They showed that

$$P(T_{st} \leq x) = \Phi(x) + n^{-1/2}\gamma\left(\frac{1}{3}x^2 + \frac{1}{6}\right)\phi(x) + o(n^{-1/2}), \tag{3.1}$$

where $n = \sum_h n_h$ and $\phi(x)$, $\Phi(x)$ are the N(0,1) pdf and cdf, respectively. The term γ reflects the amount of skewness in the distribution of T_{st}, and is given by

$$\gamma = n^{1/2}\frac{\sum_h W_h^3 \mu_{3h}/n_h^2}{\text{S.E.}(\bar{x}_{st})^3},$$

where μ_{3h} is the third central moment in stratum h.

Thus it can be seen that using $\pm 2\text{se}(\bar{x}_{st})$ amounts to employing only the first term in (3.1). The alternatives to be suggested here improve upon this by making some allowance for the second term.

A natural first approach would be to use the bootstrap to estimate the true cdf of T_{st} (Hall 1992). This involves generating B bootstrap samples; for each such sample, the observations within each stratum are independently resampled, and the resulting mean and variance noted. This leads to the calculation of

$$T_{st(j)} = \frac{\bar{x}_{st(j)} - \bar{x}_{st}}{\text{se}(\bar{x}_{st(j)})},$$

where $\bar{x}_{st(j)}$ and $\text{se}(\bar{x}_{st(j)})$ are the stratified sample mean and its estimated standard error calculated from the jth bootstrap sample ($j = 1, \ldots, B$). A $(100 - \alpha)\%$ confidence interval for μ is then given by

$$\left[\bar{x}_{st} - t_U \text{se}(\bar{x}_{st}), \bar{x}_{st} - t_L \text{se}(\bar{x}_{st})\right],$$

where t_L and t_U are the observed lower and upper $\alpha/2$ percentiles of the distribution of the $T_{st(j)}$ ($j = 1, \ldots, B$). As is generally the case with using this studentized version of the bootstrap, the resulting estimate of the cdf of T_{st} allows for the second term in (3.1), and so improves upon the N(0,1) approximation (Hall 1992, p. 84).

An alternative to the bootstrap is provided by a general method for skewness-correction, also described in Hall (1992). The idea is to find an invertible transformation $F_{st} = f(T)$ such that the cdf of F_{st} is closer to $\Phi(.)$ than is the cdf of T_{st}. A $(100 - \alpha)\%$ confidence interval for μ is then given by

$$\left[\bar{x}_{st} - f^{-1}(z_\alpha)\text{se}(\bar{x}_{st}), \bar{x}_{st} - f^{-1}(-z_\alpha)\text{se}(\bar{x}_{st})\right],$$

where z_α is the upper $(\alpha/2)\%$ point of the N(0,1) distribution.

Hall (1992, p.123) suggested using a cubic transformation whose coefficients depend on the form of (3.1). Direct application of this method leads to the use of

$$F_{st} = f(T_{st}) = \frac{1}{6}n^{-1/2}\hat{\gamma} + T_{st} + \frac{1}{3}n^{-1/2}\hat{\gamma}T_{st}^2 + \frac{1}{3}\left(\frac{1}{3}n^{-1/2}\hat{\gamma}\right)^2 T_{st}^3 ,$$

where

$$\hat{\gamma} = n^{1/2}\frac{\sum_h W_h^3 m_{3h}/n_h^2}{\text{se}(\bar{x}_{st})^3} \quad \text{and} \quad m_{3h} = \sum_i (x_{hi} - \bar{x}_h)^3/n_h .$$

The inverse of this transformation, used to calculate the confidence interval, is given by

$$f^{-1}(z) = \left(\frac{1}{3}n^{-1/2}\hat{\gamma}\right)^{-1}\left\{\left[1 + n^{-1/2}\hat{\gamma}\left(z - \frac{1}{6}n^{-1/2}\hat{\gamma}\right)\right]^{1/3} - 1\right\}.$$

By virtue of (3.1) it can be shown that

$$P(F_{st} \le x) = \Phi(x) + o(n^{-1/2}), \tag{3.2}$$

and so the cdf of F_{st} is closer to $\Phi(.)$, as required.

In fact, this method can be improved upon by using bootstrap resampling to estimate the cdf of F_{st}, thereby allowing for some of the nonnormality corresponding to the second term in (3.2) (see Hall 1992, p.124). This involves the calculation of

$$F_{st(j)} = \frac{1}{6}n^{-1/2}\hat{\gamma}_{(j)} + T_{st(j)} + \frac{1}{3}n^{-1/2}\hat{\gamma}_{(j)}T_{st(j)}^2 + \frac{1}{3}\left(\frac{1}{3}n^{-1/2}\hat{\gamma}_{(j)}\right)^2 T_{st(j)}^3,$$

where $\hat{\gamma}_{(j)}$ is the estimate of γ obtained from the jth bootstrap sample ($j = 1, \ldots, B$). A $(100 - \alpha)\%$ confidence interval for μ is then given by

$$\left[\bar{x}_{st} - f^{-1}(y_U)\text{se}(\bar{x}_{st}), \bar{x}_{st} - f^{-1}(y_L)\text{se}(\bar{x}_{st})\right] ,$$

where y_L and y_U are the observed lower and upper $\alpha/2$ percentiles of the distribution of the $F_{st(j)} (j = 1, \ldots, B)$.

4. SIMULATION STUDY

To assess the actual coverage rates that we might expect for a biological population, a simulation study was carried out. The full set of results are contained in Webster (1993), with the most practically relevant features summarized below. We note here that neither the percentile bootstrap nor the bias-corrected percentile bootstrap (Efron 1982) were as reliable as the studentized version described in Section 3.

The four methods under consideration will be referred to as follows:

- NT is the method that uses $\Phi(.)$ as an estimate of the cdf of T_{st}
- BT is the method that uses the bootstrap to estimate the cdf of T_{st}
- NF is the method that uses $\Phi(.)$ as an estimate of the cdf of F_{st}
- BF is the method that uses the bootstrap to estimate the cdf of F_{st}

Table 1. Distributions Used to Generate Observations in the Simulation Study

		Lognormal parameters	
Density	δ	μ	σ^2
Low	0.44	4.12	2.31
High	0.34	5.97	1.56

NOTE: δ = probability of a zero; μ and σ^2 are the mean and variance on the log-scale.

The simulation study was based on the results of a cockle survey carried out near Dunedin, New Zealand, in 1991 and 1992 (Stewart, Keogh, Fletcher, and Mladenov 1992). The stratum areas in the survey were large enough to make the sampling fractions negligible. The observations were generated using a delta distribution: each observation had a probability of being zero, with a lognormal distribution being used to generate the nonzero values (Pennington 1983). For every type of survey structure considered, half the strata were assumed to be high density and half low density. The parameters of the delta distributions chosen are shown in Table 1.

In each case, the stratum weights were set to be in the ratio 9:1 (low:high density strata). This choice was made partly to reflect the pattern found in the actual survey, and partly because it meant that equal sample sizes in the different strata corresponded approximately to Neyman allocation.

To cover a range of survey structures, three basic types of survey were considered (see Table 2). For each total sample size, survey type I consisted of as many strata as possible with just three units sampled in each (the minimum needed to allow skewness to be estimated). Survey type III involved only two strata, a larger sample size being used within each of these. Survey type II was chosen to be intermediate between the two extremes. The total sample sizes were chosen to cover a range of values for which it was thought the central limit argument discussed in Section 2 might not be reasonable.

In what follows we will be considering the coverage errors associated with different confidence intervals, and it is important to distinguish between two-sided and one-sided error rates. The former are concerned with overall coverage, while the latter focus on the error associated with each end of the interval. We shall be concerned here with one-sided error rates, as it will be important to assess the reliability of both the upper and lower limits.

Table 2. Survey Structures Used in the Simulations

Total sample	*I*		*II*		*III*	
size (n)	L	n_h	L	n_h	L	n_h
30	10	3	6	5	2	15
48	16	3	8	6	2	24
90	30	3	10	9	2	45
144	48	3	12	12	2	72

Table 3. Lower Limit Error Rates (%) for Four Different Methods of Calculating a 95% Confidence Interval, Where n is the Overall Sample Size and the Survey Structures I, II, and III are as Shown in Table 2

	n=30			n=48			n=90			n=144		
Method	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
NT	0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.1	0.0	0.2
BT	0.1	0.1	0.6	0.0	0.2	1.0	0.1	0.8	0.7	0.2	0.5	1.7
NF	0.2	0.7	0.9	0.1	0.7	1.0	0.2	1.1	0.8	0.6	0.8	2.2
BF	0.4	1.6	2.0	0.1	1.6	1.5	0.3	1.9	2.1	0.6	2.0	2.8

Error rates were calculated for the lower and upper 95% confidence limits. These were the percentage of simulations for which the true population mean was below the lower limit and above the upper limit, respectively. One thousand simulations were used each time so that the error rates could be estimated (with 95% confidence) to within 1 percentage point, assuming the true error rates to be around 2.5%.

For those methods using the bootstrap, 500 resamples were selected each time, ensuring (with approximately 85% confidence) that the upper and lower 2.5% points of the resulting bootstrap distribution would correspond to the upper and lower $(2.5 \pm 1)\%$ points of the bootstrap distribution that would have been obtained using an infinite number of resamples.

To gain some understanding of any differences in error rates, it is useful to compare mean interval widths (Rao and Wu 1988). This was done here by calculating standardized confidence interval half-widths, defined as

$$W_L = \frac{\bar{w}_L}{2\text{se}(\bar{x}_{st})} \quad \text{and} \quad W_U = \frac{\bar{w}_U}{2\text{se}(\bar{x}_{st})},$$

where $\bar{w}_L$ is the mean (over all simulations) of the width of the lower part of the confidence interval, and $\bar{w}_U$ the corresponding mean for the upper part of the interval. The denominator is the half-width (lower or upper) of an interval based on the $\pm 2\text{se}(\bar{x}_{st})$ rule. An independent estimate of the standard error of $\bar{x}_{st}$ was used to calculate this, namely, the standard deviation of an additional set of 3000 stratified sample means.

Table 4. Upper Limit Error Rates (%) for Four Different Methods of Calculating a 95% Confidence Interval, where n is the Overall Sample Size and the Survey Structures I, II, and III are as Shown in Table 2

	n=30			n=48			n=90			n=144		
Method	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
NT	17.5	20.3	22.9	14.6	17.0	20.0	13.0	16.0	16.9	14.1	13.0	14.8
BT	12.4	11.2	11.9	10.2	10.2	10.4	9.4	9.7	9.1	10.7	7.7	8.4
NF	19.4	15.6	12.9	16.9	12.8	10.0	13.2	10.6	9.0	14.2	8.1	8.4
BF	10.9	9.6	10.6	8.7	8.4	8.7	7.8	8.3	7.9	9.6	6.7	7.6

Table 5. Lower Standardized Interval Half-Widths for Four Different Methods of Calculating a 95% Confidence Interval, Where n is the Overall Sample Size and the Survey Structures I, II, and III are as Shown in Table 2

	n=30			*n=48*			*n=90*			*n=144*		
Method	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
NT	1.06	0.80	0.64	0.98	0.82	0.68	1.20	0.86	0.80	1.16	0.82	0.76
BT	2.36	0.56	0.46	1.60	0.58	0.48	3.02	0.62	0.58	1.92	0.60	0.58
NF	0.54	0.48	0.42	0.54	0.54	0.46	0.68	0.60	0.56	0.70	0.58	0.56
BF	1.38	0.42	0.40	1.12	0.46	0.44	1.90	0.54	0.52	1.26	0.54	0.54

A MATLAB program for calculating the four types of confidence intervals described here is given in the Appendix.

5. SIMULATION RESULTS

The error rates are shown in Tables 3 and 4, from which we note the following:

1. In virtually all cases, the lower error rate is too low and the upper one too high. This corresponds to both limits being too low in general.
2. The upper error rates are clearly worse than the lower ones. This is to be expected, as high positive skewness will cause the upper limit to be poorly defined.
3. The lower error rate appears to improve as the sample size within strata increases. This also appears to be true for the upper error rate, with the exception of the NT method.
4. NT is uniformly worse than the other three methods, particularly for the upper limit. This is not surprising.
5. The BT and NF methods perform about equally well in general, as would be expected, although the latter is not as good on the upper limit when the stratum sample size is small. This is also to be expected, as the NF method involves estimating the skewness within each stratum.
6. BF is consistently the best method, again as expected. It appears to provide a significant improvement over the NF method, particularly for the upper limit when the stratum sample size is small.

The standardized interval half-widths are shown in Tables 5 and 6. These indicate that

1. BT, NF, and BF produce asymmetric intervals (for each of these methods, both the lower and upper limits are generally higher than for the NT method), and
2. the BT method generally results in wider intervals than the NF method even when the error rates are about the same.

Table 6. Upper Standardized Interval Half-Widths for Four Different Methods of Calculating a 95% Confidence Interval, Where *n* is the Overall Sample Size and the Survey Structures I, II, and III are as Shown in Table 2

	n=30			*n*=48			*n*=90			*n*=144		
Method	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*	*I*	*II*	*III*
NT	1.06	0.80	0.64	0.98	0.82	0.68	1.20	0.86	0.80	1.16	0.82	0.76
BT	3.48	3.00	2.94	2.30	3.08	2.96	3.26	3.02	2.96	2.20	1.96	1.88
NF	0.86	1.74	1.86	0.82	1.90	1.98	1.04	2.14	2.30	1.04	1.98	2.06
BF	5.28	3.54	2.78	4.02	3.48	2.88	4.84	3.42	3.04	3.94	2.84	2.58

6. DISCUSSION

The main purpose of this paper has been to make available to applied statisticians and biologists a simple method of skewness correction for confidence intervals from stratified surveys. This formula has been derived from the work of Abramovitch and Singh (1985) and Hall (1992), references that may not be readily accessible to our intended audience. On its own, the formula can improve coverage error to the same extent that is achieved by using the bootstrap procedure, unless the sample size within each stratum is very small. Furthermore, by using it in conjunction with the bootstrap, further improvement in the coverage error is possible.

The simulation results in Section 5 indicate that even these improvements cannot ensure that the actual error rates are as close to their nominal levels as we would like. This is especially true for the upper confidence limit. This is not surprising, as determination of this upper limit is bound to be extremely difficult when there is a high level of skewness in the population. It can be argued that in such cases a confidence interval is in itself an inappropriate way of expressing the uncertainty associated with an estimate. We suggest that confidence curves may help here. These are found by simply applying the skewness-correction method for a range of values of α.

We illustrate this idea using some trawl survey data taken from Clark and Tracey (1991), as the usefulness of these curves is particularly marked for these data. A total of 72 trawls were carried out in 14 strata, with the number of trawls per stratum varying from 3 to 15. The four different 95% confidence intervals for total biomass are shown in Table 7. As would be expected from the results in Section 5, the differences between the upper limits are quite marked, with the BF method providing the most reliable limits, especially

Table 7. 95% Confidence Intervals for Total Biomass (Metric Tons) From a Trawl Survey Involving 72 Trawls Taken in 14 Strata

Method	*Lower*	*Upper*
NT	5,539	107,925
BT	17,697	317,649
NF	19,376	171,832
BF	22,922	360,087

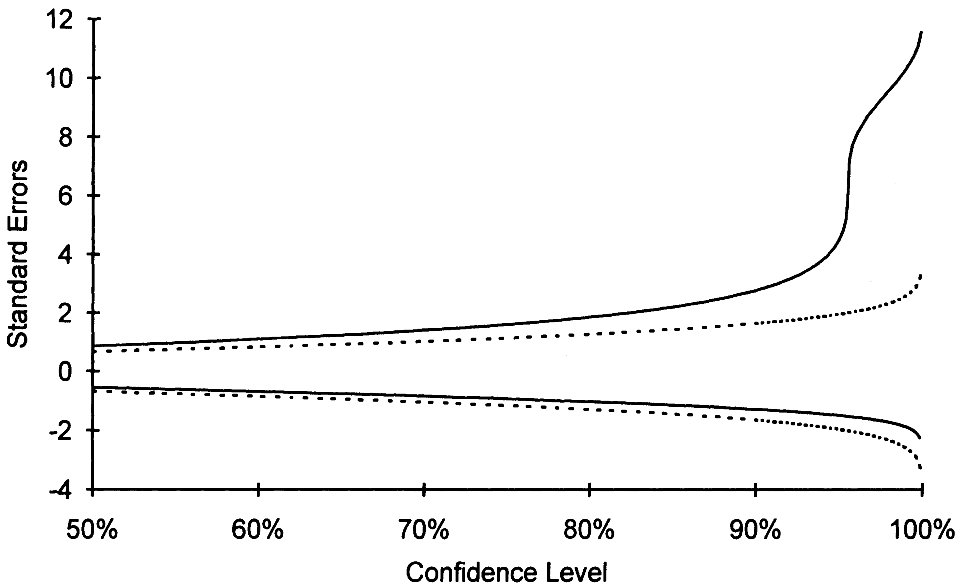

Figure 1. Confidence Curves for Trawl Survey Data Using the NT Method (dotted lines) and the NF Method (solid lines).

at the upper end (further simulations, based on these data, support this conclusion). The confidence curves for the NF method ($\hat{\gamma} = 4.1$) are shown in Figure 1, together with those for the NT method as a comparison. Note that the vertical axis is scaled in terms of the standard error of the stratified sample mean.

This plot indicates that the 95% upper limit produced by the skewness-corrected method is unstable in the sense that this limit can vary substantially if we go from a confidence level of 90% to one of 95% or higher. Such curves may prove valuable in assessing the uncertainty associated with these limits, particularly in those cases where the standard error is large relative to the estimate.

APPENDIX: MATLAB PROGRAM

A.1 Main Program

This program calculates a confidence interval for the population mean using each of the four methods discussed in this paper. Functions used by this program are given in part 2 of the Appendix.

The workspace '`Datafile`' contains the stratum weights (`W`), sample sizes (`n`), and data (`Data`). `W` and `n` are row vectors; `Data` is a column vector with the data in the order defined by `W` and `n`.

```
clear; load Datafile
```

`B` is the number of times to resample, `C` is the desired confidence level, and `z` the upper $100(1 - C)/2$ percentile for N(0,1).

```
B=500; C=0.95; z=1.96;
```

m, v, and m3 contain the stratum means, variances and 3rd central moments. est is the estimated population mean, se the estimated standard error and sk= $\hat{\gamma}/\sqrt{n}$.

```
ic=0;
for i=1:length(n)
ii=ic+[1:n(i)]; ic=ic+n(i);
[m(i),v(i),m3(i)]=statistics(Data(ii));
end
[est,se,sk]=estimates(n,W,m,v,m3);
```

Resampling of the data for the BT and BF intervals: mb, vb, m3b, estb, seb, and skb are resample values of m, v, m3, est, se, and sk. Tb and Fb are vectors containing the B values of T_{st} and F_{st}.

```
rand('uniform')
for j=1:B
ic=0;
for i=1:length(n)
ii=ic+round(n(i)*rand(1,n(i))+0.5); ic=ic+n(i);
[mb(i),vb(i),m3b(i)]=statistics(Data(ii));
end
[estb,seb,skb]=estimates(n,W,mb,vb,m3b);
Tb=[Tb; (estb-est)/seb]; Fb=[Fb; f((estb-est)/seb,skb)];
end
```

Tb and Fb are sorted in order to identify the required percentiles, L and U. NT, NF, BT, and BF contain the required intervals.

```
Tb=sort(Tb); Fb=sort(Fb); L=round(0.5*(1-C)*(B+1));
U=round(0.5*(1+C)*(B+1));
NT=[est-z*se est-(-z)*se];
NF=[est-finv(z,sk)*se est-finv(-z,sk)*se];
BT=[est-Tb(U)*se est-Tb(L)*se];
BF=[est-finv(Fb(U),sk)*se est-finv(Fb(L),sk)*se];
```

A.2 FUNCTIONS USED BY THE MAIN PROGRAM

Each of the functions shown below is stored as a separate file.

Calculates the cube root of x, allowing for negative values of x:

```
function y=cuberoot(x)
y=sign(x)*(abs(x))^(1/3);
```

Calculates the stratified sample mean, the estimated standard error and $\hat{\gamma}/\sqrt{n}$:

```
function [est,se,sk]=estimates(n,W,m,v,m3)
est=sum(W.*m); se=sqrt(sum((W.^2).*v./n));
sk=sum((W.^3).*m3./(n.^2))/(se^3);
```

Calculates F_{st}:

```
function y=f(T,sk)
y=sk/6+T+sk.*T.^2/3+(sk/3).^2.*T.^3/3;
```

Calculates the inverse of F_{st}:

```
function y=finv(z,sk)
y=(cuberoot(1+sk*(z-sk/6))-1)/(sk/3);
```

Calculates the mean, variance and 3rd central moment of the data in d:

```
function [m,v,m3]=statistics(d)
m=mean(d); v=std(d)^2; m3=mean((d-m).^3);
```

ACKNOWLEDGMENTS

We would like to thank Jake Keogh for discussing this work from a biologist's perspective, and are grateful to a referee for comments concerning presentation.

[Received November 1994. Revised July 1995.]

REFERENCES

Abramovitch, L., and Singh, K. (1985), "Edgeworth Corrected Pivotal Statistics and the Bootstrap," *The Annals of Statistics,* 13, 116–132.

Clark, M. R., and Tracey, D. M. (1991), *Report of an Orange Roughy Trawl Survey in Southern New Zealand Waters, June-July 1991,* Wellington, New Zealand: MAF Fisheries.

Cochran, W. G. (1977), *Sampling Techniques* (3rd ed.), New York: John Wiley.

Crow, E. L., and Shimizu, K. (ed.) (1988), *Lognormal Distributions. Theory and Applications,* New York: Marcel Dekker.

Efron, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans,* Philadelphia: SIAM.

Hall, P. (1992), *The Bootstrap and Edgeworth Expansion,* New York: Springer–Verlag.

MATLAB ver. 4.1 (1993), The MathWorks, Inc., Natick, MA.

Pennington, M. (1983), "Efficient Estimators of Abundance for Fish and Plankton Surveys," *Biometrics,* 36, 707–719.

Rao, J. N. K., and Wu, C. F. J. (1988), "Resampling Inference With Complex Survey Data," *Journal of the American Statistical Association,* 83, 231–241.

Stewart, B., Keogh, J., Fletcher, D., and Mladenov, P. (1992), *Biomass Survey of the New Zealand Littleneck Clam in Papanui and Waitati Inlets, Otago, 1991/1992,* Marine Science and Aquaculture Research Centre, University of Otago, Dunedin, New Zealand.

Webster, R. (1993), "Confidence Interval Methods for the Stratified Sample Mean," unpublished M.S. thesis, University of Otago, Dunedin, New Zealand.