VARIANCE ESTIMATION FOR A TIME SERIES WITH LINEAR TREND

Stephen M. Woodruff
BUREAU OF LABOR STATISTICS

Introduction

This is a continuation of the paper entitled, "A Comparison of Some Estimators in Sampling for a Time Series with Linear Trend" (1), in which a best linear estimator was chosen. The best way to estimate the variance of this estimate is the subject of this paper.

The problem of variance estimation for a time series with linear trend is studied using model based procedures. The work done by Royall and Cumberland ((2), (3)) is refined and adapted to the special problems of the Current Employment Survey (790 Survey) at the Bureau of Labor Statistics. A variety of variance estimation techniques are examined; including a variation of jackknife estimation and shrinkage toward a constant.

The variance estimator that performed the best is a refinement of the estimator that was suggested by Richard Royall (2). This refinement uses a variation of jackknife estimation to estimate a conditional variance. The population totals, for which a variance estimator is needed, are changing with time. The units which make up the population follow a linear model. Let the characteristic of interest for unit i at time k be denoted $Y_k(i)$. The expected value for $Y_k(i)$ given $Y_{k-1}(i) = y_{k-1}(i)$ is proportional to $y_{k-1}(i)$ where lower case $y_k(i)$ is used to denoted the realization of $Y_k(i)$. The model also states that the units in the population are uncorrelated and the conditional variance for unit i at time k is proportional to the characteristic of interest at time k-1. This model is given algebraically as follows:

$$E(Y_k(i)/k-1) = \beta_k y_{k-1}(i) \qquad \text{for } k=2,3,\ldots. \tag{1.1}$$

$$\text{Cov}\,(Y_k(i), Y_k(j) \,/\, k-1) = \delta_{ij}\, \sigma_k^2\, y_{k-1}(i)$$

where $\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

σ^2 is a constant,

and $E(\,\cdot\,/\,k-1)$ denotes conditional expectation given outcomes up to and including time k-1.

If S is used to denote the entire population then the realized population total at time k is:

$$y_k(S) = \sum_{i \in S} y_k(i)$$

If s_k is used to denote the sample at time k from S then define

$$y_k(s_k) = \sum_{i \in s_k} y_k(i) \text{ for } k=2, 3, \ldots .$$

$$y_1(s_1) = y_1(S) \text{ since } s_1 = S .$$

(Time k=1 is called the benchmark month.)

Upper case will be used to denote random variables and lower case will be used to denote their realized values.

The problem of estimating $y_k(S)$ was studied in (1) and the solution, $t_k(S)$, which is suggested there is:

$$t_1(S) = y_1(S) \text{ since } s_1 = S \tag{1.2}$$

$$\text{and } t_k(S) = y_k(s_k) + \hat{\beta}_k t_{k-1}(r_k) \quad k=2,3,\ldots .$$

where $r_k = S - s_k$ (the complement of s_k in S) and $\hat{\beta}_k$ is an estimate of β_k given by:

$$\hat{\beta}_k = y_k(s_k s_{k-1}) / y_{k-1}(s_k s_{k-1})$$

where $s_k s_{k-1}$ is used to denote $s_k \cap s_{k-1}$, the intersection of s_k and s_{k-1}.

$$\text{Thus } t_k(i) = \begin{cases} y_k(i) & \text{if } i \in s_k \\ \hat{\beta}_k t_{k-1}(i) & \text{if } i \in r_k = S - s_k \end{cases}$$

is used to estimate the outcome in the i^{th} unit at time k.

A common β_k, which is independent of i in model (1.1) provides the link between sample and non sample units. The sample units are used to estimate β_k with $\hat{\beta}_k$. Then $\hat{\beta}_k$ is used to estimate the nonsample $y_k(i)$ from $t_{k-1}(i)$.

If the model (1.1) holds then $t_k(S)$ is unbiased in the sense that

$$E(T_k(S) - Y_k(S)) = 0.$$

where

$$T_k(i) = \begin{cases} Y_k(i) & \text{if } i \in s_k \\ \dfrac{Y_k(s_k s_{k-1})}{Y_{k-1}(s_k s_{k-1})}\, t_{k-1}(i) & \text{if } i \in r_k \quad k = 2,3, \text{---} \end{cases}$$

and $T_1(i) = Y_1(i)$ for all $i \in S$.

The error variance of $T_k(S)$ is defined as the variance of $(T_k(S) - Y_k(S))$ which I'll denote:

$$V(T_k(S) - Y_k(S)) = E\,(T_k(S) - Y_k(S))^2$$

Note that all expectations are taken with respect to the model (1.1) and not with respect to the sampling distribution by which s_k was chosen.

If we condition on the outcomes up to time k-1 then the model (1.1) allows the error variance to be written as:

$$V(T_k(S) - Y_k(S)) = \beta_k^2 V(T_{k-1}(S) - Y_{k-1}(S)) + E(V(T_k(S) - Y_k(S)/k-1)) \text{ for } k > 1 \quad (1.4)$$
$$\text{and} \quad = 0 \quad \text{for } k=1$$

The problem of estimating the error variance of $t_k(S)$ then reduces to estimating β_k^2 and the expected value of the conditional error variance, $V(T_k(S) - Y_k(S)/k-1)$ for each k. What follows is a study of these two estimation problems.

This study contains both theory and empirical work. A variety of variance estimators are tried. One estimator which is based on a variation of the jackknife variance estimator appears to be best. It is given explicitly as V_{k5} in (2.9).

Description of Variance Estimators

1) The iterative formula (1.4) can be written as the following difference equation.

$$V_k = \beta_k^2 V_{k-1} + E(V(T_k(S) - Y_k(S)/k-1)) \qquad k=2,3,.. \qquad (2.1)$$

where $V_k = V(T_k(S) - Y_k(S))$.

(note that $V_1 = 0$)

Now note that $V(T_k(S) - Y_k(S)/k-1) =$

$$= V(T_k(r_k) - Y_k(r_k)/k-1)$$
$$= V((Y_k(s_k s_{k-1})/Y_{k-1}(s_k s_{k-1})) \cdot T_{k-1}(r_k) - Y_k(r_k)/k-1)$$
$$= V((Y_k(s_k s_{k-1})/y_{k-1}(s_k s_{k-1})) \cdot t_{k-1}(r_k) - Y_k(r_k)/k-1)$$
$$= (t_{k-1}(r_k)/y_{k-1}(s_k s_{k-1}))^2 \; V(Y_k(s_k s_{k-1})/k-1) + V(Y_k(r_k)/k-1)$$

First note that the second part of (1.1) which states that $V(Y_k(i)/k-1) = \sigma^2 y_{k-1}(i)$ implies:

$$V(Y_k(r_k)/k-1) = (y_{k-1}(r_k)/y_{k-1}(s_k s_{k-1})) \times V(Y_k(s_k s_{k-1})/k-1).$$

Next note that $y_{k-1}(r_k)$ is unobservable but $t_{k-1}(r_k)$ is its estimator. Therefore

$$V(Y_k(r_k)/k-1) \doteq (t_{k-1}(r_k)/y_{k-1}(s_k s_{k-1})) \times V(Y_k(s_k s_{k-1})/k-1)$$

(2.1) can be rewritten approximately as.

$$V_k \doteq \beta_k^2 V_{k-1} + E((A_k+1) \cdot A_k \cdot V(Y_k(s_k s_{k-1})/k-1))$$
$$(V_1 = 0) \qquad k=2,3,.... \qquad (2.2)$$

where $A_k = (t_{k-1}(r_k)/y_{k-1}(s_k s_{k-1}))$.

If $E((A_k+1) \cdot A_k \cdot V(Y_k(s_k s_{k-1})/k-1))$ is estimated by $(A_k+1) \cdot A_k \cdot \hat{V}(Y_k(s_k s_{k-1})/k-1)$ where $\hat{V}(Y_k(s_k s_{k-1})/k-1)$ is an estimator of $V(Y_k(s_k s_{k-1})/k-1)$ then (2.2) implies $\hat{V}_k$ is approximately:

$$\hat{V}_k \doteq \hat{\beta}_k^2 \hat{V}_{k-1} + (A_k+1)A_k \hat{V}(Y_k(s_k s_{k-1})/k-1) \quad (2.3)$$

where $\hat{\beta}_k$ is an estimator of β_k.

The empirical results which were presented in (1) and the equation (2.2) indicate a linear relation between time, k, and $V(T_k(S)-Y_k(S))$ when the sets s_k remain relatively stable. That is, $y_{k-1}(s_k s_{k-1})$ varies only slightly with k for k > 1. In this case an approximation to (2.2) is:

$$V_1 = 0$$
$$V_k = \beta_k^2 V_{k-1} + G \qquad (2.4)$$

where G is constant.

In addition, if β_k is close to one then (2.4) is approximately:

$$V_k = V_{k-1} + G$$

The solution to this difference equation is:

$$V_k = (k-1)G \qquad (2.5)$$

Thus V_k is roughly proportional to k-1, the distance from the benchmark month.

The assumptions that support the expression for V_k given in (2.5) are met in many strata of the 790 survey (4). The computer simulations mentioned in (1) also give empirical support to (2.5).

An implication of this result is that any estimator of V_k should also be proportional to the distance from the benchmark month.

$$\hat{V}_k \doteq (k-1)\hat{G} \qquad (2.6)$$

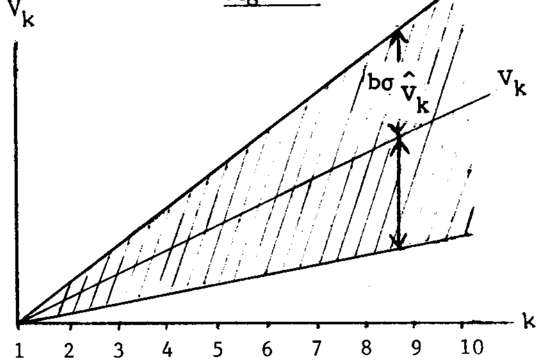
This implies that the variance of the variance estimator is proportional to the square of the distance from the benchmark month. This means that variance estimators must become rather unstable as distance from the benchmark month increases and that the relative error of the variance estimator remains approximately constant as time passes.

Tchebycheff's inequality implies that for any positive real number b

$$P(|\hat{V}_k - V_k| > b\sigma_{\hat{V}_k}) < \frac{1}{b^2} .$$

Thus if V_k is graphed as a function of k and a confidence region for $\hat{V}_k$ is graphed about V_k this confidence region will be fan shaped as shown in figure 1.

Figure 1.



The estimation problem given by (2.3) can be solved a variety of ways depending upon how $\hat{\beta}_k$ and $\hat{V}(Y_k(s_k s_{k-1})/k-1))$ are chosen.

Thus for each distinct pair $(\hat{\beta}_K, \hat{V}(Y_k(s_k s_{k-1})/\rho-1))$, (2.3) gives a different estimator of variance and variety of different $\hat{\beta}$s and $\hat{V}$s were tried.

2) $\hat{\beta}_k$ is estimated in 2 ways. The first method is ratio of y values one month apart for matched units.

$$\hat{\beta}_{ka} = y_k(s_k s_{k-1})/y_{k-1}(s_k s_{k-1})$$

The second method is a shrinkage estimator that makes use of prior knowledge about β_k. In many strata of the 790 Survey β_k is usually near one. This suggests an estimator, $\hat{\beta}_{kb}$, defined as follows:

$$\hat{\beta}_{kb} = (1-\hat{\beta}_{ka})^3 / ((1-\hat{\beta}_{ka})^2 + d_k^2) + 1$$

where d_k^2 is an estimator of the variance of $\hat{\beta}_{ka}$. This method is discussed in (5). When model (1.1) obtains then

$$V(\hat{\beta}_{ka}) = E(\sigma^2/y_{k-1}(s_k s_{k-1}))$$

and $\hat{V}(Y_k(s_k s_{k-1})/k-1) \doteq \sigma^2 y_{k-1}(s_k s_{k-1})$

These last two equations suggest that d_k^2 may be choosen to be

$$d_k^2 = \hat{V}(Y_k(s_k s_{k-1})/k-1)/y_{k-1}^2(s_k s_{k-1}) \qquad (2.7)$$

Next the $V(Y_k(s_k s_{k-1})/k-1)$ needs to be estimated. This term is the conditional variance of a sum of uncorrelated random variables. Thus it may be written as:

$$\sum_{i \in s_k s_{k-1}} V(Y_k(i)/k-1) \qquad .(2.8)$$

An approximately unbiased estimator for each term in this sum is

$$(y_k(i) - \hat{\beta}_{ka} y_{k-1}(i))^2 \ .$$

If a constant K_i is choosen so that $E(K_i(y_k(i) - \hat{\beta}_{ka} y_{k-1}(i))^2 / k-1) = V(Y_k(i)/k-1)$ then (2.8) can be estimated with

$$\hat{V}_{ka} = \sum_{i \in s_k s_{k-1}} K_i(y_k(i) - \hat{\beta}_{ka} y_{k-1}(i))^2$$

where $K_i = \left[1 - y_{k-1}(i)/y_{k-1}(s_k s_{k-1}) \right]^{-1}$

Another method is to choose one K such that when (1.1) holds then

$$E(K \sum_{i \in s_k s_{k-1}} (y_k(i) - \hat{\beta}_{ka} y_{k-1}(i))^2 / k-1)$$
$$= V(Y_k(s_k s_{k-1})/k-1) \ .$$

Then $K = \left[1 - (\sum_{s_k s_{k-1}} (y_{k-1}^2(i)) / y_{k-1}^2(s_k s_{k-1}))\right]^{-1}$

Call this estimator $\hat{V}_{kb}$.

These two estimators for $V(Y_k(s_k s_{k-1})/k-1)$ can be combined with the information contained in (2.4) to develop regression and composite estimators for V_k. If we let V_{k1} represent any one of the four possible estimators for V_k which can be developed from, $\hat{\beta}_{ka}$ or $\hat{\beta}_{kb}$, and, $\hat{V}_{ka}$ or $\hat{V}_{kb}$, using (2.3) then define a new estimator for V_k as

$V_{k2} = B_k \cdot k + C_k$ where B_k and C_k are chosen so that:

$$\sum_{i=1}^{k} (B_k i + C_k - V_{i1})^2 \text{ is minimized.}$$

V_{k2} is a regression estimator for V_k. For k sufficiently large V_{k1} and V_{k2} are nearly independent. This suggests a variance estimator which is a composite of both V_{k1} and V_{k2}. The weights for this composite estimator are a function of the variances of V_{k1} and V_{k2}.

If the model (1.1) holds then the variance of V_{k1} can be estimated with

$$L = (1/k) \sum_{i=1}^{k} (B_k i + C_k - V_{i1})^2 \ .$$

Some additional calculation gives an approximation to the variance of V_{k2} as:

$$(2(2k+1)/(k(k-1)))\ L$$

Thus a third estimator for V_k is given by:

$$V_{k3} = \frac{2(2k+1)}{k(k-1)+2(2k+1)} V_{k1} + \frac{k(k-1)}{k(k-1)+2(2k+1)} V_{k2}$$

where $V_{k2} = \sum_{i=1}^{k} f_i\ V_{i1}$

and $f_i = (12/(k^2 - 1))(i - ((k+1)/2))$

$+ (12/k(k-1))(((2k+1)/6) - i/2)$

A fourth estimator which is crude but did well in the computer simulations is

$$V_{k4} = (.5)V_{k1} + (.5)(k-1)V_{2,1} .$$

V_{k4} is an equally weighted composite of V_{k1} and the extrapolation of V_{21} to time k. This extrapolation is suggested by (2.6).

The next estimator for V_k is based on a jackknife estimator for

$V(Y_k(s_k s_{k-1})/k-1)$.

The basic idea of the jackknife estimator is the deletion of one term or unit and the computation of pseudo estimators with this deleted sample. These pseudo estimates are then used to estimate the variance.

If $V(Y_k(S_k S_{k-1})/k-1)$ is estimated from each of the samples from which one element has been deleted and these estimates are averaged then another estimate of variance can be obtained. Computer simulations showed that this procedure works quite well.

As an example let $z_1\ z_2\ z_3 \cdots z_n$ be independent identically distributed random variables with mean μ and variance σ^2.

Let $C_i = \frac{1}{n-2} \sum_{j \neq i} (z_j - x_i)^2$

where $x_i = \frac{1}{n-1} \sum_{k \neq i} z_k$

Then $\frac{1}{n} \sum_{i=1}^{n} C_i$ is an excellent estimator of σ^2.

In fact, $\frac{1}{n} \sum_{i=1}^{n} C_i$

$= \frac{1}{n-1} \sum_{i=1}^{n} (z_i - \bar{z})^2$ in this simple case.

This same method is applicable to estimating $V(Y_k(s_k s_{k-1})/k-1)$. For example define.

$$v_{kbj} = K_j \sum_{i \in s_k s_{k-1} - j} (y_k(i) - \hat{\beta}_{kaj} y_{k-1}(i))^2$$

where $\hat{\beta}_{kaj} = y_k(s_k s_{k-1} - j)/y_{k-1}(s_k s_{k-1} - j)$

$(s_k s_{k-1} - j)$ is $s_k \cap s_{k-1}$ with its j^{th} element deleted and

$$Kj = \left[1 - \left(\sum_{s_k s_{k-1} - j} (y_{k-1}^2(i)) \right) / y^2(s_k s_{k-1} - j)\right]^{-1}.$$

Then $V_{k5} = A_k(A_k+1)(1/n(s_k s_{k-1}))\times$

$$\sum_{j \in s_k s_{k-1}} v_{kbj} + \beta_{ka}^2 V_{k-1,5} \qquad (2.9)$$

where $n(s_k s_{k-1})$ = no of units in $s_k s_{k-1}$.

Thus V_{k5} is a version of (2.3) where $\hat{V}(Y_k(s_k s_{k-1})/k-1)$ is estimated using a modified jack knife.

The estimators V_{k1} through V_{k5} were studied via computer simulation. The description and results of this simulation follow in the next section.

Simulation Results

These variance estimators were tested on a universe of 300 units. Each unit had data for 20 months which was generated to mimic an SIC in the National Survey of Current Employment (790 Survey). A sample of 52 units was selected according to an optimum stratified random sampling plan. For details see (1).

The data for the first month of the 20 month period was generated from a lognormal distribution. The parameters of this distribution were maximum likelihood estimates obtained from actual 790 data in one SIC. The density function for the lognormal that was used to generate the employment data for month one was:

$$f(x) = (1/((ax - b)\sqrt{2\pi}))EXP(-(1/2a^2))$$

where $a = 1.392$, $b = .5$ and $c = 3.158$.

Three hundred random numbers were generated from this density to provide data for the first month. Then the $y_k(i)$ for $k = 2, 3, \cdots 20$ and $i = 1, 2, 3 \cdots 300$, were generated from the model:

$$y_k(i) = \beta_k y_{k-1}(i) + N(0, (.3)y_k(i)) \qquad (3.1)$$

where $N(a, b)$ is the normal distribution with mean a and variance b.

Non response was modeled into this simulation. Although the same sample of 52 units was kept throughout the 20 month time period, nonresponse caused the usable sample to vary from month to month. The sample of respondents, s_k, was a simple random subsample from the 52 original sample units. The size of this subsample was fixed at $n(s_k) = 35$.

The β_k are listed below.

k	β_k	k	β_k
1	.997	11	.995
2	1.004	12	.992
3	1.009	13	.989
4	1.012	14	.987
5	1.012	15	.986
6	1.012	16	.987
7	1.010	17	.990
8	1.007	18	.995
9	1.003	19	1.002
10	.999	20	1.012

One replication of this simulation consists of constructing the universe of 20 months of data for the 300 units. Then sampling and subsampling to get the s_k for k = 1, 2, ---- 20. Note that s_1 is all 300 units since month 1 is the benchmark month. Then the basic estimator given by (1.2) was computed and the various estimators of its variance were computed for each month.

Averages of the variance estimators were then computed over a number of replications (15 to 50 depending on case). These estimated expected values of the variance estimators were then compared to the target values.

The target values were obtained from 250 replications of the simulation where the error variance was estimated by comparing actual population totals with estimators of these population totals for each of the 20 months. Therefore the target value for month k was

$$(1/250) \sum_{i=1}^{250} z_i \quad \text{where } z_i = (t_k(S)-Y_k(S))^2$$

from the i^{th} replication.

This is an unbiased estimate of the error variance $V(T_k(S)-Y_k(S)) = E(T_k(S)-Y_k(S))^2$.

These 250 replications were necessary to insure that the relative error of the target values was kept below 10%.

In addition to the variance estimators V_{k1} through V_{k5}, other artificial variance estimators were simulated for comparison purposes. For example, (2.3) was applied with the exact values of β_k and $V(Y_k(s_k s_{k-1})/k-1)$. This gives some indication of what is lost in accuracy when these quantities are estimated. This also indicates how much error is produced by the assumptions leading to (2.3).

Results of the Simulations

The two tables in this section briefly summarize the results of the computer simulations which were described in part III. Estimates of the expected value of the variance estimators are given for time periods k = 5, 10, 15 and 20. The right hand column gives the number of replicates which were used for each estimator. Each row of these tables corresponds to the estimator indicated in the left hand column. V_{k1} through V_{k5} are defined in section II.

The ordered pair notation is used for the artificial estimators. In this case (x,z) refers to the estimator derived from (2.3) with a $\hat{\beta}_k$ of x and a $\hat{V}$ of z.

In both table 1 and table 2, V_{k1} through V_{k4} are based on (2.3) with $\hat{\beta}_{ka}$ and $\hat{V}_{kb}$. V is used to denote $V(Y_k(s_k s_{k-1})/k-1)$ and for the last three rows of tables 1 and 2, $\hat{V}_{ka}$ and $\hat{\beta}_{ka}$ are used in $\hat{V}$.

The population of 300 units that was used in the simulation for table 1 contains many small units. In order to avoid negative values of $y_k(i)$ a reflecting barrier was set up at zero. If for some k and i a negative $y_k(i)$ was generated by (3.1) then $y_k(i)$ was set equal to one and $y_{k+1}(i)$ was generated using this new $y_k(i)$ of one. It is immediately noted that this relatively minor pertubation of the model created a postive bias in all variance estimators.

The reason for this upward bias is that the reflecting barrier tends to check the natural dispersal of the population as k increases. The variance estimator which is strongly based on the model does not catch this because the sample contains mostly large units. These units would be the ones least effected by the reflecting barrier.

Table 2 summarizes the simulation results when the reflecting barrier is removed. This was made possible by adding 100 to $y_1(i)$ for each i=1, 2,... 300. When

TABLE 1. (in 1000's)

ESTIMATOR	k= 5	10	15	20	REPS
V_{k1}	18.2	40.6	56.3	78.4	20
V_{k2}	17.8	41.0	61.1	80.7	20
V_{k3}	18.0	40.9	60.0	80.3	20
V_{k4}	17.3	38.7	56.7	80.0	20
V_{k5}	15.4	38.1	54.7	73.9	50
(β_k, V)	16.5	38.1	55.1	74.4	15
$(\beta_{k\beta}, V)$	16.3	37.6	56.4	77.7	15
$(\beta_k, \hat{V})$	18.7	41.1	58.1	78.2	15
$(\beta_{k\beta}, \hat{V})$	18.5	40.5	59.4	81.3	15
$(\beta_{ka}, \hat{V})$	18.3	40.0	59.6	81.2	15
TARGET	12.9	24.9	36.3	48.0	250
$\hat{\sigma}$ target	1.2	2.0	2.8	4.2	250

		k=	5	10	15	20	REPS
TABLE 2. (in 1000's)		V_{k1}	374.5	845.3	1175.3	1610.0	20
		V_{k2}	368.0	852.5	1268.8	1672.6	20
	ESTIMATOR	V_{k3}	371.4	850.2	1247.5	1661.5	20
		V_{k4}	348.3	785.0	1151.3	1570.0	20
		V_{k5}	329.8	815.2	1171.0	1575.8	50
		(β_k, V)	355.3	823.5	1174.8	1575.3	15
		$(\beta_{k\beta}, V)$	350.6	808.0	1184.8	1607.5	15
		(β_k, V_{k1})	376.7	858.6	1209.7	1621.7	15
		$(\beta_{k\beta}, V_{k1})$	371.9	841.7	1219.1	1653.9	15
		(β_{ka}, V_{k1})	369.9	835.3	1237.0	1672.3	15
		TARGET	293.9	711.4	1134.5	1371.8	250
		$\hat{\sigma}$ target	26.1	65.8	116.9	136.2	250

this is done, none of the $y_k(i)$ which are generated by (3.1) are negative.

Table 2 still indicates a positive bias in the variance estimators but not nearly so much as is present in table 1.

Note that V_{k1} is not noticably improved by using the regression and composite estimation schemes. Thus V_{2k}, V_{3k} and V_{4k} do not seem to be noticably better than V_{1k}.

This is probably because V_{1k} already makes full use of the model (1.1) and thus the additional smoothing that is attempted using regression and composite estimation is redundent.

In the lower half of both tables 1 and 2 one can compare the effect of replacing estimates of β_k and $V(Y_k(s_k s_{k-1})/k-1)$ with their exact values. It appears that the exact knowledge of these 2 parameters provides relatively minor improvement in the variance estimator given by (2.3). This means that the bias in these estimators is due to the approximations which were used in the derivation of (2.3) and the use of a reflecting barrier at zero in the simulation. A good estimate of $V(Y_k(s_k s_{k-1})/k-1)$ seems to be more important than a good estimate of β_k. It should be noted that V_{k5} (the jackknife estimator) does as well as the version which uses the exact values of β_k and $V(Y_k(S_k S_{k-1})/k-1)$. Thus it is unlikely that V_{k5} can be significantly improved.

The bias of these estimators also seems to be consistently positive. Therefore inferences from confidence intervals should be conservative.

Table 1 shows that these variance estimators are very sensitive to the type of model failure that is produced by the reflecting barrier at zero. The entries in table 1 are generally about 40% higher than the target values.

Conclusions

The best variance estimator, based on the modest amount of simulation testing done for this paper, is V_{k5}, (2.9). It makes use of the modified jackknife estimator to get a good estimate for $V(Y_k(s_k s_{k-1})/k-1)$. V_{k5} was tested using $\hat{\beta}_{ka}$. V_{k5} would usually be improved slightly using $\hat{\beta}_{kb}$.

The second important finding was the use of shrinkage to estimate the β_k. Although this nonlinear shrinkage estimator, $\hat{\beta}_{kb}$, helps only slightly in variance estimation, it does provide considerable gains when it is used in $T_k(S)$. Thus the new estimator of choice is $T_k(S)$ using $\hat{\beta}_{kb}$ in place of $\hat{\beta}_{ka}$. The modified jackknife can be applied equally well to this new basic estimator.

The modified jackknife estimator provides another alternative for variance estimation in complicated situations. It can certainly be applied to a much wider class of problems than considered here.

My thanks are due to Alfreda Reeves and Tawanna Neal, who typed and retyped this paper.

References:

(1) Woodruff, Stephen M., "A Comparison of some Estimators in Sampling for a Time Series with Linear Trend," 1982 ASA Proceedings, Survey Research Methods.

(2) Royall, Richard M., "Study of Role of Probability Models in 790 Survey Design and Estimation", BLS contract 80-98, 1981.

(3) Royall, Richard M., Cumberland, William, G., "Variance Estimation in Finite Population Sampling", JASA 1978 Vol. 73, #362.

(4) West, Sandra A. (1981) "Linear Models for Monthly all Employment Data", Bureau of Labor Statistics.

(5) Bibby, John; Toutenburg, Helge, "Prediction and Improved Estimation in Linear Models", Wiley 1979.