TRANSACTIONS OF SOCIETY OF ACTUARIES
1981 VOL. 33

ON THE VARIANCE AND MEAN SQUARED ERROR OF DECREMENT ESTIMATORS

STUART A. KLUGMAN

ABSTRACT

Actuarial estimators of decrement probabilities normally use the number of policies or the amount of insurance, rather than the number of lives, as units. In this situation, the binomial distribution is not an appropriate model. This paper defines a general model for arbitrary units and gives formulas for computing the variance and mean squared error of decrement estimators. The cases of random and nonrandom units are discussed, as well as the situation in which the decrement probability depends on the value of the unit. It is observed that, in small studies, the mean squared error is reduced by counting lives, even if the probability of death depends somewhat on the amount of insurance.

I. INTRODUCTION

MOST estimates of decrement probabilities are of the form $\hat{q} = D/G$, where G relates to the number of observations and D bears a similar relationship to the number of successes. If G and D refer to the actual number of observations, and if the observations are independent with success probability q, then $G\hat{q}$ will have a binomial distribution with parameters G and q. It follows that $E(\hat{q}) = q$ and Var $(\hat{q}) = q(1 - q)/G$. If $1 - q \doteqdot 1$ then Var $(\hat{q}) \doteqdot q/G$, the formula recommended in the Society's Part 5 textbook ([2], p. 222).

It is unusual in actuarial studies to use the number of observations (lives) when counting G and D. It is more common to use number of policies (e.g., [7], p. 140) or amount (e.g., [12], pp. 1–55). For a discussion of the merits of each approach see [2], pages 215–18.

In this paper, we will be using the following model. Let X_i be the number of units of exposure for individual i for $i = 1, \ldots, n$. Let $\theta_i = 0$ if individual i is an observed failure and $\theta_i = 1$ if i is a success. Then $G = \Sigma X_i$ and $D = \Sigma \theta_i X_i$. Assume that $\theta_1, \ldots, \theta_n, X_1, \ldots, X_n$ are mutually independent except that θ_i may depend on X_i. Let $q(x) = P(\theta_i = 1 \mid X_i = x)$. This dependence has been clearly established when X_i is the amount of insurance but the nature of $q(x)$ is not well known [11].

This paper ignores some of the problems introduced by the practical aspects of decrement estimation, the most significant being the problem of censoring. Several authors (see, for example, [3], [6], and [13]) have pointed out that the "actuarial" estimator is not consistent and is not based on the likelihood equation of a binomial model. While not reflecting the customary formulas, the model introduced above does have the desired statistical properties. For situations with low withdrawal probabilities, the model used here should provide a good representation.

The problem is to find $E(\hat{q})$ and Var $(\hat{q})$ in this general setting. In Section II the case where $q(x) = q$ for all x will be considered. The solution for general $q(x)$ will be presented in Section III. In Section IV some uses for these results will be given.

II. X_i AND θ_i INDEPENDENT

In this case the desired values are relatively easy to obtain. Two results help considerably. They apply to any pair of random variables X and Y. First, $E(Y) = E[E(Y \mid X)]$, and, second, $\text{Var}\,(Y) = E[\text{Var}\,(Y \mid X)] + \text{Var}\,[E(Y \mid X)]$. To use these results, let $Y = \hat{q}$ and $X = (X_1, \ldots, X_n)$. Then

$$E(Y \mid X) = E(\Sigma\, \theta_i X_i / \Sigma\, X_i \mid X_1, \ldots, X_n) = q\ .$$

Also,

$$\text{Var}\,(Y \mid X) = \text{Var}\,(\Sigma\, \theta_i X_i / \Sigma\, X_i \mid X_1, \ldots, X_n) = q(1 - q)\, \Sigma\, X_i^2 / (\Sigma X_i)^2\ .$$

Finally, $E(\hat{q}) = E(q) = q$, and

$$\begin{aligned}\text{Var}\,(\hat{q}) &= E[q(1 - q)\, \Sigma\, X_i^2/(\Sigma\, X_i)^2] + \text{Var}\,(q) \\ &= q\,(1 - q) E[\Sigma\, X_i^2/(\Sigma\, X_i)^2]\ . \end{aligned} \tag{1}$$

At this point, two situations must be identified. In some cases the values of $X_1, \ldots, X_n$ will be known in advance. In this situation $X_1, \ldots, X_n$ should be treated not as random variables but as fixed quantities. This makes Var $(\hat{q})$ easy to compute, since formula (1) involves the expectation of a constant. In other situations the values of $X_1, \ldots, X_n$ will not be known. This occurs if the calculation of Var $(\hat{q})$ is being done in advance of the data collection, or if the data set is so large as to make the evaluation of $\Sigma\, X_i^2$ and $\Sigma\, X_i$ impractical. In the latter case, formula (1) requires the evaluation of $E[\Sigma\, X_i^2/(\Sigma\, X_i)^2]$, a task that can be difficult even if each X_i has a simple distribution. We can, however, obtain an asymptotic estimate of this expectation. Use the Theorem in the Appendix with $Y_i = X_i^2$, $Z_i = X_i$, and $f(y, z) = y/z^2$. Then the expression

$$n^{1/2}\{n \Sigma X_i^2/(\Sigma X_i)^2 - E(X^2)/[E(X)]^2\}$$

will converge to a normally distributed random variable with mean zero and finite variance. Therefore, a reasonable approximation to n Var ($\hat{q}$) is $q(1 - q)E(X^2)/[E(X)]^2$, where X has the same distribution as each X_i. The two situations described above may be combined if, for the case in which $X_1, \ldots, X_n$ are known, X is defined as assigning probability $1/n$ to each value of X_i.

If $G > n$ (as would be the case when number of policies or amount is used), then

$$\frac{q(1 - q)}{G} \leqslant \frac{q(1 - q)}{n} \leqslant \frac{q(1 - q)E(X^2)}{n[E(X)]^2}\ .$$

This indicates that the value of Var ($\hat{q}$) cannot be less than the value given by a naive formula using exposures, and also that it exceeds the value obtained when number of lives is used. The only way the above relation can produce an equality is if X places all its probability on a single value. Table 1 gives an example of the relative magnitudes of these values. The data are from a sampling of policies issued by the Equitable Insurance

TABLE 1

VARIANCE FORMULAS BASED ON KNOWN AMOUNTS OF INSURANCE

Ages	n	Variance Formula*		
		(1) $10^4/\Sigma X_i$	(2) $10^4/n$	(3) $10^4\Sigma X_i^2/(\Sigma X_i)^2$
48–52	261	.0058	38	79
53–57	355	.0056	28	70
58–62	383	.0061	26	68
63–67	311	.0075	32	95
68–72	300	.0090	33	246
73–77	238	.0118	42	141
78–82	142	.0197	70	518
83–87	76	.0461	132	290
88–92	24	.1087	417	1,577
48–92	2,090	.0011	5	16

*(1) Naive variance based on number of exposures. (2) True variance for estimates based on number of lives. (3) True variance for estimates based on amount. All three formulas omit the factor $q(1 - q)$.

Company of Iowa. Since q is unknown, the term $q(1 - q)$ is left out of each expression.

Exposures are usually calculated by using amount rather than number of policies, for two reasons: first, savings and convenience, since using amounts does not necessitate a check for duplicate policies; second, the belief that q should measure the monetary impact of the loss. As shown in Table 1, however, if q does not depend on the amount of insurance, accuracy is lost by using amount as the unit. It should be noted in addition that using number of policies instead of number of lives also leads to an increase in the variance. As an example, suppose that the number of policies on an individual has a Poisson distribution with parameter λ. Conditioning on the existence of at least one policy yields

$$P(X = x) = \lambda^x e^{-\lambda}/[x!(1 - e^{-\lambda})] \quad \text{for} \quad x = 1, 2, \ldots$$

and

$$E(X^2)/[E(X)]^2 = (\lambda + 1)(1 - e^{-\lambda})/\lambda .$$

The ratio has a maximum value of 1.30, which indicates that the penalty for using number of policies is not severe.

III. θ_i DEPENDENT UPON X_i

The large-amount studies of the Society of Actuaries indicate that θ_i does depend upon X_i. In particular, there is some evidence that $q(x)$ is a decreasing function of x [11]. In this case,

$$E(\hat{q}) = E[E(\Sigma\, \theta_i X_i/\Sigma\, X_i \mid X_1, \ldots, X_n)] = E[\Sigma\, X_i q(X_i)/\Sigma\, X_i] .$$

Arguing as in Section II, we find that an approximation is provided by $q_G = E[Xq(X)]/E(X)$. This is the quantity of interest, since it represents the dollar-weighted probability of death. This will be different from $q_L = E[q(X)]$, which represents the probability of death for the average life. We can compare the merits of $\hat{q}_G = \Sigma\, \theta_i X_i/\Sigma\, X_i$ as opposed to $\hat{q}_L = \Sigma\, \theta_i/n$ by examining their mean squared errors (MSE) with respect to q_G. For $\hat{q}_G$, the approximation is

$$\text{MSE}(\hat{q}_G) \doteq \text{Var}\,(\hat{q}_G) \doteq (\mu_U/\mu_X^2 + \mu_T^2\beta_X/\mu_X^4 - 2\mu_T\mu_U/\mu_X^3)/n$$

(see the Appendix for the derivation and for definitions of the symbols). For $\hat{q}_L$ it is

$$\text{MSE}(\hat{q}_L) = \text{Var}\,(\hat{q}_L) + [E(\hat{q}_L) - q_G]^2 = q_L(1 - q_L)/n + (q_L - q_G)^2 .$$

If $q(x)$ indicates that the probability is significantly affected by the amount, then the increased variability in $\hat{q}_G$ due to the use of amounts may be more than offset by the bias in $\hat{q}_L$.

As an example, consider the Equitable data at ages 48–52. I have selected $q(x) = b + a/x$ as a general form. There is no empirical basis for its selection other than its simplicity. It leads to $q_G = b + a/\mu_X$ and $q_L = b + aE(1/X)$. Furthermore, $\mu_U = b\beta_X + a\mu_X$ and $\mu_T = b\mu_X + a$. For the Equitable data $\mu_X = 6{,}557$, $\beta_X = 88{,}440{,}613$, and $E(1/X) = 0.000344$. This results in $q_G = b + 0.0001525a$, $q_L = b + 0.000344a$, and n Var $(\hat{q}_G) = 0.0001525a + 2.057b + 1.3265a^2/10^9 - 2.057b^2 - 0.000305ab$. Let $q_G = 0.006$, in which case $b = 0.006 - 0.0001525a$. Substituting this in the previous expressions yields

$$\mathrm{MSE}(\hat{q}_G) = (0.012268 - 0.00015926a - 8.9375a^2/10^{13})/n$$

and

$$\mathrm{MSE}(\hat{q}_L) = (0.005964 + 0.0001892a - 3.677a^2/10^8)/n + 3.667a^2/10^8\ .$$

The mean squared error using $\hat{q}_G$ will be smaller for $n \geqslant 0.999976 - 9{,}490/a + 171{,}803/a^2$. A reasonable value for a would be in the neighborhood of 5 (roughly based on results in [11]). This would require $n \geqslant 4{,}975$, a relatively large number of observations. Note that, if $q(x)$ is decreasing, $q_L > q_G$ and the use of number of lives provides for a conservative estimate.

A compromise between the large variance of q_G and the bias of q_L may be obtained by reducing the influence of large-amount policies. For example, X_i could be defined as the face value if the face value is less than c, and as c if the face value is greater than c. This suggestion is endorsed in [2], page 217, and results similar to those presented in this section are given in [10].

IV. PRACTICAL APPLICATIONS

Knowledge of the variance of decrement estimators can prove useful during the planning, graduation, and reporting stages of a decrement study. In the planning stage, the magnitude of the variance reveals the sample size needed to achieve a desired level of accuracy. Several modern graduation methods suggest using the variances as weights. These include Whittaker-Henderson methods [8], Bayesian methods [9], and the fitting of mathematical models [14]. Finally, any well-conceived report should include a statement about the probable error. In all of the above, it would be appropriate, when using a biased estimator, to use the mean squared error in place of the variance.

The method outlined in Section III can be used to decide on the appropriate measurement unit. In particular, for small studies, even if decrement probabilities depend on amounts, a smaller mean squared error may be obtained by using number of lives or number of policies.

V. CONCLUSIONS

The methods presented in this paper enable an investigator to estimate the variance and mean squared error of the estimators when the units of investigation are variable. Two items that appear worthy of investigation are the nature of $q(x)$ and the dependence of $E(X^2)/[E(X)]^2$ on age.

VI. ACKNOWLEDGMENTS

I would like to thank Don Iverson of the Equitable Insurance Company of Iowa for supplying the data. I also thank Bob Hogg and Russ Lenth of the University of Iowa for stimulating my interest in error reduction and for providing the insights that led to this study. In addition, I thank a reviewer for providing some valuable comments and additional references.

APPENDIX

The paper makes use of the following theorem.

THEOREM. *Let (Y_1, Z_1), (Y_2, Z_2), . . . be independent and identically distributed random pairs each with mean (μ_Y, μ_Z) and variance (σ_Y^2, σ_Z^2). Let σ_{YZ} = Cov (Y_i, Z_i), $\bar{Y}_n = \Sigma_{i=1}^n Y_i/n$, $\bar{Z}_n = \Sigma_{i=1}^n Z_i/n$, and let $f(y, z)$ be any real-valued function with first and second derivatives existing in a neighborhood of (μ_Y, μ_Z). Then as $n \longrightarrow \infty$, the distribution of $n^{1/2}[f(\bar{Y}_n, \bar{Z}_n) - f(\mu_Y, \mu_Z)]$ converges to the distribution of a normal random variable with mean zero and variance*

$$\left[\frac{\partial}{\partial y} f(y, z)\bigg|_{(\mu_Y, \mu_Z)}\right]^2 \sigma_Y^2$$

$$+ 2\left[\frac{\partial}{\partial y} f(y, z)\bigg|_{(\mu_Y, \mu_Z)}\right]\left[\frac{\partial}{\partial z} f(y, z)\bigg|_{(\mu_Y, \mu_Z)}\right]\sigma_{YZ}$$

$$+\left[\frac{\partial}{\partial z} f(y, z)\bigg|_{(\mu_Y, \mu_Z)}\right]^2 \sigma_Z^2 .$$

Proof. See Theorems 4.2.3 and 4.2.5 of [1].

To obtain the result in Section III, make the following definitions: $T_i = X_i q(X_i)$, $U_i = X_i^2 q(X_i)$, $V_i = [X_i q(X_i)]^2$, and $\beta_X = E(X_i^2)$. Let μ_T, μ_U, μ_V, and μ_X be the means of T_i, U_i, V_i, and X_i, respectively. Then

$$\begin{aligned}
\mathrm{Var}(\hat{q}_G) &= E[\mathrm{Var}(\Sigma\theta_i X_i/\Sigma X_i|X_1,\ldots,X_n)] + \mathrm{Var}[E(\Sigma\theta_i X_i/\Sigma X_i|X_1,\ldots,X_n)] \\
&= E\{\Sigma X_i^2 q(X_i)[1-q(X_i)]/(\Sigma X_i)^2\} + \mathrm{Var}[\Sigma X_i q(X_i)/\Sigma X_i] \\
&= E[\Sigma(U_i - V_i)/(\Sigma X_i)^2] + \mathrm{Var}(\Sigma T_i/\Sigma X_i) \\
&= E[(\bar{U}_n - \bar{V}_n)/\bar{X}_n^2]/n + \mathrm{Var}(\bar{T}_n/\bar{X}_n) .
\end{aligned} \tag{2}$$

The theorem provides approximations for each term. The first is just $(\mu_U - \mu_V)/(n\mu_X^2)$. For the second, use $f(y, z) = y/z$ and note that $\sigma_Y^2 = \sigma_T^2 = \mu_V - \mu_T^2$, $\sigma_Z^2 = \sigma_X^2 = \beta_X - \mu_X^2$, and $\sigma_{YZ} = \sigma_{TX} = \mu_U - \mu_T\mu_X$. Then

$$\begin{aligned}
n\,\mathrm{Var}\,(\bar{T}_n/\bar{X}_n) \doteq (1/\mu_X)^2(\mu_V - \mu_T^2) &+ 2(1/\mu_X)(-\mu_T/\mu_X^2)(\mu_U - \mu_T\mu_X) \\
&+ (-\mu_T/\mu_X^2)^2(\beta_X - \mu_X^2) .
\end{aligned}$$

Combining these two results yields

$$\mathrm{Var}\,(\hat{q}_G) \doteq (\mu_U/\mu_X^2 + \mu_T^2\beta_X/\mu_X^4 - 2\mu_T\mu_U/\mu_X^3)/n .$$

To complete the result, note that Slutsky's theorem (see [5], p. 255) indicates that $\hat{q}_G$ is a consistent estimator of q_G, since

$$\Sigma\,\theta_i X_i/n \xrightarrow{P} E[Xq(X)] , \qquad \Sigma\, X_i/n \xrightarrow{P} E(X) .$$

Therefore, $\hat{q}_G$ is asymptotically unbiased and $\mathrm{MSE}(\hat{q}_G) \doteq \mathrm{Var}\,(\hat{q}_G)$.

If the values of X_i are fixed and known, then the second term of formula (2) is zero and the first term becomes $\Sigma\, X_i^2 q(X_i)[1 - q(X_i)]/(\Sigma\, X_i)^2$. This formula for the variance appeared in [4].

REFERENCES

1. ANDERSON, T. W. *An Introduction to Multivariate Statistical Analysis.* New York: John Wiley & Sons, 1958.
2. BATTEN, R. W. *Mortality Table Construction.* Englewood Cliffs, N.J.: Prentice-Hall, 1978.
3. BRESLOW, N., and CROWLEY, J. "A Large Sample Study of the Life Table and Product Limit Estimates under Random Censorship," *Annals of Statistics*, II, No. 3 (1974), 437–53.

4. CODY, D. D. "Actuarial Note: The Standard Deviation of the Rate of Mortality by Amounts," *TASA*, XLII (1941), 69–73.
5. CRAMÉR, H. *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton University Press, 1946.
6. ELVEBACK, L. L. "Estimation of Survivorship in Chronic Disease: The 'Actuarial' Method," *Journal of the American Statistical Association*, LIII (1958), 420–40.
7. "Experience under Certain Ordinary Waiver of Premium Benefits between 1959 and 1974 Anniversaries," *TSA, 1978 Reports of Mortality and Morbidity Experience*, pp. 85–216.
8. GREVILLE, T. N. E. *Part 5 Study Notes: Graduation*. Chicago: Society of Actuaries, 1974.
9. HICKMAN, J. C., and MILLER, R. B. "Notes on Bayesian Graduation," *TSA*, XXIX (1977), 7–22.
10. KLUGMAN, S. "Robust Mortality Estimation," *ARCH* 1979.3, pp. 61–75.
11. "Mortality on Policies for Large Amounts," *TSA, 1975 Reports of Mortality and Morbidity Experience*, pp. 57–120.
12. "Mortality under Standard Ordinary Insurance Issues between 1976 and 1977 Anniversaries," *TSA, 1978 Reports of Mortality and Morbidity Experience*, pp. 1–68.
13. SEAL, H. L. "Deaths among Prospective 'Existings,' " *PCAPP*, XI (1960–61), 308–10.
14. TENENBEIN, A., and VANDERHOOF, I. T. "New Mathematical Laws of Select and Ultimate Mortality," *TSA*, XXXII (1980), 119–58.

DISCUSSION OF PRECEDING PAPER

KENNETH S. AVNER:

Mr. Klugman rightly notes that for statistical estimators, any well-conceived report must include a statement about the probable error of the estimators. Further, it has become common in actuarial literature to see the notion of risk quantified in terms of the variance of an unknown quantity. Toward that end, an investigation of the variance of decrement estimators—the estimators that are fundamental to actuarial work—is always timely in this publication. I should like to raise some questions related to those discussed by Mr. Klugman and clarify some inaccuracies in his otherwise clear and concise paper.

The paper discusses estimators of decrement probabilities that can be expressed in the form $\hat{q} = D/G$, where G relates to the number of observations and D bears a similar relationship to the number of successes. It states that if the observations are independent with probability of success q, then $G\hat{q}$ will have a binomial distribution with parameters G and q, from which it follows that $E(\hat{q}) = q$ and $\text{Var}(\hat{q}) = q(1-q)/G$. This is certainly a common scheme to actuaries; in fact, Mr. Klugman cites the Society's Part 5 text by Batten. But he sidesteps a logical error in Batten and goes on to claim that if we take $1 - q \simeq 1$, then $\text{Var}(\hat{q}) \simeq q/G$, the formula recommended by Batten.

The first point I consider is this derivation of Var ($\hat{q}$). Certainly there are conditions that yield a binomially distributed $G\hat{q}$; in practice, however, they rarely arise. In fact, this approach serves only to obscure the underlying situation, which, when $1 - q \simeq 1$, can give directly the estimate $\text{Var}(\hat{q}) \simeq q/G$.

What ruins the binomial distribution is the inequality of the increments that sum to form G. Consider an unequally weighted sum of independent Bernoulli random variables (i.e., each takes on only the values 0 and +1) with parameters q:

$$S = w_1X_1 + w_2X_2 + \ldots + w_nX_n .$$

Clearly, $E(S) = q(w_1 + \ldots + w_n)$, and

$$\text{Var}(S) = q(1-q)(w_1^2 + \ldots + w_n^2) ,$$

which is quite similar to binomial distribution results. However, unless the weights are equal, the distribution is not binomial. Specifically, in cases where the observations are weighted by amounts of insurance or where there is censoring, $G\hat{q}$ would be better estimated as normally distributed. This does not affect the formulas derived by Batten (for confidence intervals, etc.), since he uses a normal approximation to his assumed binomial.

Now consider a direct analysis when $1 - q \simeq 1$. There are two approaches. First, there is the well-known approximation to such a normal or binomial distribution by the Poisson distribution. Using the method of moments, we match the means and are led to a Poisson with parameter qG. And that is my point. A Poisson with parameter qG also has variance qG, and we return to the standard approximation!

The second approach uses the same calculations to match moments and derive variances, but involves a different philosophy. Here we dispense with the normal approximation to sums of independent random variables. Recall that that approximation is based on the notion that the individual variables are (probabilistically) sufficiently small so that no single one contributes significantly to the overall sum. An alternative situation is the summing of independent random variables, but because the probability of a nonzero value is sufficiently small, each such value may contribute appreciably to the overall sum. In such a case one would expect direct convergence to a Poisson-like variate. This might justify a match of moments to derive the formulas above.

For the basic derivation in Section II, Mr. Klugman posits that X_i and θ_i are mutually independent. (The X_i's and θ_i's are always sets of independent variables.) Under this assumption $E(\hat{q}) = q$, and Var $(\hat{q}) = q(1 - q)$ $E[\Sigma X_i^2/(\Sigma X_i)^2]$.

Then he considers two situations: where the X_i's are known in advance, and where there is no a priori knowledge of the X_i's because the calculations are "being done in advance of data collection" or because "the data set is so large as to make evaluation of ΣX_i and ΣX_i^2 impractical." In both cases we wish to compute $E[\Sigma X_i^2/(\Sigma X_i)^2]$.

When the X_i's are not random, the expectation is superfluous, so in the former case Var $(\hat{q})$ is easy to compute. This is not true in the latter case. There, Mr. Klugman suggests the expedient of an asymptotic estimate,

$$E(X^2)/n[E(X)]^2 \simeq E[\Sigma X_i^2/(\Sigma X_i)^2] ,$$

derived by the so-called delta method.

His use of the delta method is in itself worthy of note. This, the most popular and one of the most useful methods of advanced applied statistics, is too little known in actuarial circles, although it is now beginning to appear in our journals (see also Hickman and Miller's 1977 paper—reference [9] in Mr. Klugman's paper—which uses a variance stabilizing transformation supplied by the delta method). To be applied in the present case, however, it does require that the X_i's be independent and identically distributed.

When the calculations are being done in anticipation of the experiment, independent and identically distributed X_i's are an acceptable assumption, and in keeping with the spirit of this paper, we probably should pay more attention to the variance of the asymptotic normal. But when the X_i's merely constitute a data set that is too large for computations, the assumption is false. It is not surprising, then, that the expedient is useless in practice. For it requires a computation of $E(X)$ and $E(X^2)$, which presumably would be based on ΣX_i and ΣX_i^2—quantities assumed to be impractical to evaluate. We are left where we started.

Finally, recall the inequalities among the various estimates of Var $(\hat{q})$:

$$q(1 - q)/G \leq q(1 - q)/n \leq q(1 - q)E(X^2)/n[E(X)]^2 .$$

The first estimate is a naive approach based on the binomial distribution where the exposure itself is used in the denominator. It is commonly recognized as inadequate and is often multiplied by the average exposure (e.g., policy size) per life. This approximates the middle estimate. That the quantity is still inadequate is the content of the second inequality, and this may explain some of the arbitrary reduction in observed exposures in practice (cf. ref. [9] of the paper).

(AUTHOR'S REVIEW OF DISCUSSION)

STUART A. KLUGMAN:

I will confine my remarks to three items mentioned in Mr. Avner's discussion. First, the sum he labels S is identical with the one I label D, and from there our results are identical. My reference to the binomial distribution in the first paragraph of Section I was meant only to apply to the case in which G counts lives, or, equivalently, where $X_1 = X_2 = \ldots = X_n$. This does not invalidate Mr. Avner's second point, that the approximation Var $(\hat{q}) = qG$ is best viewed as being derived from the Poisson distribution.

Finally, I must agree with Mr. Avner's remark that neither formula (1) nor the approximation $q(1 - q)E(X^2)/n[E(X)]^2$ is of direct use when one is dealing with a large data set. I might suggest an alternative, namely, estimating either of them by sampling from the available data.

I would like to thank Mr. Avner for taking the time to comment on my paper.