ROBUST MORTALITY ESTIMATION

Presented at the Fourteenth Actuarial Research Conference

Iowa City, Iowa, September 6, 1979

Stuart Klugman

The University of Iowa

ABSTRACT

There are three commonly used methods for compiling exposure and death data for use in obtaining crude mortality rates. They are lives, policies and amounts of insurance. Most major actuarial studies use amounts of insurance in an attempt to reflect the financial impact of the deaths. When using amounts of insurance it is necessary to recognize the greater possibility of fluctuations due to claims for large amounts. Robust procedures, with their automatic reduction of the contribution of outliers, may provide some relief. This paper investigates the properties of estimators based on lives, amounts and robust alternatives. It is concluded that lives provide the best estimator, even when there is a moderate dependence of mortality rates on amounts of insurance.

1. INTRODUCTION

Most modern mortality studies tabulate data by using amounts of insurance. This has several benefits, one being that the problem of multiple policies on a single individual is eliminated. A second reason is based on the evidence that mortality rates decrease as the amount of insurance increases. The reported mortality rates should reflect the financial loss (especially if the rates are to be used for gross premium and reserve determinations) and the use of amounts provides an automatic weighting.

As a preliminary step toward evaluating robust estimators, it is first necessary to determine the properties of the standard estimators. The determination of crude mortality rates can be modelled as follows. Let n be the number of observed lives. Define independent and identically distributed random variables $X_1, X_2, \cdots, X_n$ with X_i = amount of insurance on life i. Let $\theta_1, \theta_2, \cdots, \theta_n$ be independent random variables with θ_i possibly dependent upon X_i but independent of $X_1, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n$. Further assume that the θ_i have identical distributions. Let $\theta_i = 0$ if life i survives and $\theta_i = 1$ if life i dies. The standard estimators are $\hat{q}_L = \Sigma\theta_i/n$ and $\hat{q}_A = \Sigma\theta_i X_i/\Sigma X_i$, the estimators based on lives and amounts respectively. In this paper all sums run from $i = 1$ to $i = n$ and so indication of the index and limits on sums will be omitted.

Let $F(x)$ be the distribution function of the random variable X which has the same distribution as each X_i. Define $\alpha = E[X]$ and $\beta = E[X^2]$. Further define $q = E[\theta_i]$ and $q(x) = E[\theta_i | X_i = x]$. Then $q = E[E[\theta_i | X_i]] = E[q(X)] = \int q(x)dF(x)$. Since q is the probability of death of a randomly selected life, $\hat{q}_L$ is likely to be a superior estimator of q. Of more interest for insurance purposes is the probability of death of a randomly selected dollar of death benefit. This quantity is $q_A = E[\theta_i X_i]/E[X_i] = \int xq(x)dF(x)/\int xdF(x)$. It has been accepted that $\hat{q}_A$ is a good estimator of q_A.

The criteria to be used to evaluate the merits of the estimators are bias and mean squared error. In all cases it will be assumed that q_A is the quantity to be estimated. The bias of $\hat{q}$ is defined as $b(\hat{q}) = E[\hat{q}] - q_A$ and the mean squared error is $MSE(\hat{q}) = E[(\hat{q}-q_A)^2] = Var(\hat{q}) + b(\hat{q})^2$.

In the following sections only two forms of the function $q(x)$ are investigated. The first is $q(x) = q = q_A$, the case in which mortality does not depend on amount. This assumption is referred to as Case 1. Case 2 is $q(x) = b + a\,\ell n(x)$. In this case $q = b + aE[\ell n\ X]$ and $q_A = b + aE[X\ \ell n\ X]/E[X]$.

2. EVALUATION OF $\hat{q}_L$ AND $\hat{q}_A$ FOR CASE 1

For Case 1, $E[\hat{q}_L] = \Sigma E[\theta_i]/n = q$ and $Var(\hat{q}_L) = \Sigma\ Var(\theta_i)/n^2 = q(1-q)/n$. The estimator $\hat{q}_L$ is unbiased with $MSE(\hat{q}_L) = q(1-q)/n$.

For $\hat{q}_A$, $E[q_A] = E[E[\Sigma\theta_i X_i/\Sigma X_i \mid X_1,\cdots,X_n]] = E[q\Sigma X_i/\Sigma X_i] = q$ and $Var(\hat{q}_A) = E[Var(\Sigma\theta_i X_i/\Sigma X_i \mid X_1,\cdots,X_n)] + Var(E[\Sigma\theta_i X_i/\Sigma X_i \mid X_1,$ $\cdots,X_n]) = E[q(1-q)\Sigma X_i^2/(\Sigma X_i)^2] + Var(q) = q(1-q)E[\Sigma X_i^2/(\Sigma X_i)^2]$. Thus $\hat{q}_A$ is also unbiased while its mean squared error depends on the distribution of X. If the amounts of insurance are considered to be nonrandom, the mean squared error can be obtained by removing the expectation and using the actual values of the amounts. A numerical value for the expectation may be difficult to obtain; the following theorem provides a method of obtaining its asymptotic value.

Theorem: Let $X_1,X_2,\cdots$ be independent and identically distributed random variables with means μ_x and variances σ_x^2. Let $Y_1,Y_2,\cdots$ be independent and identically distributed random variables with means μ_y and variances σ_y^2. Let $\sigma_{xy} = Cov(X_i,Y_i)$ and assume X_i and Y_j are independent for $i \neq j$. Let $W_n = \Sigma X_i/n$ and $Z_n = \Sigma Y_i/n$. Let $f(w,z)$ be any real-valued function with first and second derivatives existing in a neighborhood of (μ_x,μ_y). Then $\sqrt{n}(f(W_n,Z_n) - f(\mu_x,\mu_y))$ $\xrightarrow{D} V$ where V has a normal distribution with mean zero and variance

$$\left(\frac{\partial f}{\partial w}\Big|(\mu_x,\mu_y)\right)^2\sigma_x^2 + 2\left(\frac{\partial f}{\partial w}\Big|(\mu_x,\mu_y)\right)\left(\frac{\partial f}{\partial z}\Big|(\mu_x,\mu_y)\right)\sigma_{xy} + \left(\frac{\partial f}{\partial z}\Big|(\mu_x,\mu_y)\right)^2\sigma_y^2.$$

Proof: See Theorems 4.2.3 and 4.2.5 of An Introduction to Multivariate Statistical Analysis, T. W. Anderson, Wiley, 1958.

To use the Theorem, let $X_i = X_i^2$, $Y_i = X_i$ and $f(w,z) = w/z^2$. Then $n\ \mathrm{Var}(\hat{q}_A) = q(1-q)E[W_n/Z_n^2] \rightarrow q(1-q)E[X^2]/(E[X])^2 = q(1-q)\beta/\alpha^2$. Since $\beta \geq \alpha^2$ with equality only if X has all its mass at one point, $\mathrm{MSE}(\hat{q}) \geq \mathrm{MSE}(\hat{q}_2)$ (asymptotically).

3. ROBUST ALTERNATIVES

It was not expected that estimation by amounts would be superior to lives when the mortality rate is independent of the amount of insurance. The additional variability due to the amounts is not offset by the weighting of the deaths by amounts. In his text, Mortality Table Construction (Prentice-Hall, 1978), Batten observes that "complications arise when this method [use of amounts] is employed, notably when several huge claims in a single cell distort the resulting mortality rate. The effects of such fluctuations can be somewhat muted by the elimination, in both exposure values and deaths, of any protection in a single policy in excess of some predetermined level" (page 217).

The method suggested by Batten is similar to several robust methods. The method developed by Huber ("Robust Estimation of a Location Parameter," Annals of Math. Stat., Vol. 35, pp. 73-101, 1964) estimates the mean as the solution $\tilde{\mu}$ to $n\tilde{\mu} = \Sigma z_i x_i + \Sigma(1-z_i)(\tilde{\mu} + c\sigma\ \mathrm{sgn}(x_i - \tilde{\mu}))$ where $z_i = 1$ if $|x_i - \tilde{\mu}| \leq c\sigma$ and $z_i = 0$ otherwise. The function $\mathrm{sgn}(x)$ is

1 if $x > 0$ and -1 if $x < 0$ and σ is the standard deviation of X. The value of c is selected in advance; Huber recommends $1 \le c \le 2$. This estimator has minimax variance over the class of variables containing symmetrically contaminated normal random variables. The greater the contamination, the smaller the value of c should be. Huber's method is essentially a two-sided version of Batten's suggestion with the cutoffs set at $\tilde{\mu} \pm c\sigma$. In most cases the value of σ will also need to be estimated.

A second robust estimator is trimming. In this case, all amounts above or below the cutoff are eliminated from consideration. In general, these alternative estimators can be defined as $\hat{q} = \Sigma\theta_i h(X_i)/\Sigma h(X_i)$. The three methods described above use the following $h(x)$.

(i) One-sided reduction

$$\hat{q}_R : h(x) = \begin{cases} x & x \le m+k \\ m+k & x \ge m+k \end{cases}$$

(ii) Two-sided reduction

$$\hat{q}_H : h(x) = \begin{cases} m-k & x \le m-k \\ x & |x-m| \le k \\ m+k & x \ge m+k \end{cases}$$

(iii) One-sided trimming

$$\hat{q}_T : h(x) = \begin{cases} x & x \le m+k \\ 0 & x > m+k. \end{cases}$$

For any of these estimators, if $q(x) = q$ then $E[\hat{q}] = q$ and $n \operatorname{Var}(\hat{q}) \to q(1-q)E[h(X)^2]/(E[h(X)])^2$.

4. AN ILLUSTRATION

Mortality data was obtained from The Equitable Life Insurance Company of Iowa. A sample of 2090 policies was fitted to a lognormal distribution. The sample mean and variance for the logarithms of the amounts were $\hat{\mu} = .79$ and $\hat{\sigma}^2 = .81$. Assuming these as the exact values leads to $E[X] = 4044$ and $E[X^2] = 36{,}762{,}909$. The cutoff points are based on the corresponding normal distribution. They are $\exp(7.9 \pm .9c)$. The following table gives asymptotic values of $n\,MSE(\hat{q})/q(1-q)$ with the expectations computed by using the lognormal distribution.

	$\hat{q}_T$	$\hat{q}_R$	$\hat{q}_H$	
c = 2	1.743	1.830	1.827	
c = 1.5	1.631	1.615	1.604	For $\hat{q}_A$: 2.248
c = 1	1.634	1.411	1.371	For $\hat{q}_L$: 1
minimum	1.613	1	1	
c at min	1.25	$-\infty$	0	

For Case 1 it is clear that the estimator based on lives is superior. It does appear that $\hat{q}_A$ can be improved by using one of the robust modifications. When Case 2 is investigated it is hoped that the dependency of mortality on amount will close the gap between $\hat{q}_L$ and the robust estimators.

5. EVALUATION OF THE ESTIMATORS FOR CASE 2

It is now assumed that $q(x) = b + a \ln x$. It is expected that the estimators will be biased for q_A and that the variances will be in the same relationship as in the example in the previous section. The combination of bias and variance may lead to mean squared errors that are different from those in Case 1 with the relationships depending on sample size.

A general result may be obtained for the situation in which $E[\theta_i | X_i] = q(X_i)$, $Y_i = h(X_i)$ and $\hat{q} = \Sigma\theta_i Y_i/\Sigma Y_i$. All the estimators discussed in this paper can be placed in this framework. For computing the bias, $E[\hat{q}] = E[E[\Sigma\theta_i Y_i/\Sigma Y_i | X_1,\cdots,X_n]] = E[\Sigma q(X_i)h(X_i)/\Sigma h(X_i)] \to E[q(X)h(X)]/E[h(X)]$. For the variance,

$$\begin{aligned}\mathrm{Var}(\hat{q}) &= E[\mathrm{Var}(\Sigma\theta_i Y_i/\Sigma Y_i | X_1,\cdots,X_n)] + \mathrm{Var}(E[\Sigma\theta_i Y_i/\Sigma Y_i | X_1,\cdots,X_n]) \\ &= E[\Sigma q(X_i)(1-q(X_i))h(X_i)^2/(\Sigma h(X_i))^2] \\ &\qquad + \mathrm{Var}(\Sigma q(X_i)h(X_i)/\Sigma h(X_i)).\end{aligned}$$

Let $T_n = \Sigma q(X_i)h(X_i)/n$, $U_n = \Sigma q(X_i)h(X_i)^2/n$, $V_n = \Sigma q(X_i)^2 h(X_i)^2$ and $Z_n = \Sigma h(X_i)/n$. Then $n\,\mathrm{Var}(\hat{q}) = E[(U_n - V_n)/Z_n^2] + n\,\mathrm{Var}(T_n/Z_n)$. If $\gamma_1 = E[q(X)h(X)]$, $\gamma_2 = E[q(X)h(X)^2]$, $\gamma_3 = E[q(X)^2 h(X)^2]$, $\alpha_h = E[h(X)]$ and $\beta_h = E[h(X)^2]$ then, from the Theorem,

$$E[(U_n - V_n)/Z_n^2] \to (\gamma_2 - \gamma_3)/\alpha_h^2 \quad \text{and}$$

$$n\ \mathrm{Var}(T_n/Z_n) \rightarrow (\gamma_3-\gamma_1^2)/\alpha_h^2 - 2\gamma_1(\gamma_2-\alpha_h\gamma_1)/\alpha_h^3+\gamma_1^2(\beta_h-\alpha_h^2)/\alpha_h^4.$$

Finally, $n\ \mathrm{Var}(\hat{q}) \rightarrow \gamma_2/\alpha_h^2 + \gamma_1^2\beta_h/\alpha_h^4 - 2\gamma_1\gamma_2/\alpha_h^3$.

To evaluate $\hat{q}_L$, use $h(x)=1$. Then $\gamma_1=\gamma_2=E[q(X)]=q$ and $\alpha_h=\beta_h=1$. This yields $E[\hat{q}_L]=\gamma_1/\alpha_h=q$ and $n\ \mathrm{Var}(\hat{q}) = q(1-q)$. As expected, the mean and variance of $\hat{q}_L$ are unchanged from Case 1. However, the bias is now $q-q_A$ and the mean squared error is $q(1-q)/n+(q-q_A)^2$. If X is assumed to have a lognormal distribution with $\mu=E[\ell n\ X]$ and $\sigma^2=\mathrm{Var}(\ell n\ X)$ then $q=b+a\mu$ and $q_A=b+a(\mu+\sigma^2)$. The measures of performance become $b(\hat{q}_L)=-a\sigma^2$ and $\mathrm{MSE}(\hat{q}_L) = (b+a\mu)(1-b-a\mu)/n+a^2\sigma^4$.

For $\hat{q}_A$, use $h(x)=x$ to obtain $E[\hat{q}_A] \rightarrow E[q(X)X]/E[X] = q_A$ and $n\,\mathrm{MSE}(\hat{q}_A) = n\,\mathrm{Var}(\hat{q}_A) \rightarrow (\beta/\alpha^2)[q_A(1-q_A)+a\sigma^2(1-2q_A)]$ using the lognormal assumption. Using the values of α, β, μ and σ^2 as in the illustration in Section 4, $\mathrm{MSE}(\hat{q}_A)<\mathrm{MSE}(\hat{q}_L)$ if $.6561a^2(n-1) - 2.6308a + (5.2617a - 1.2480)q_A + 1.2480q_A^2 > 0$. To complete the illustration, the most recent Large Amount Study (<u>TSA Reports</u>, 1975) was used. That study found mortality on policies for more than \$50,000 to be 90% of overall mortality. To apply this result to the Equitable data, the cutoff was reduced to \$15,000 and q_A set at .02. This produces $a=-.0016$ and $b=.034$. When placed in the above formula, it is seen that $\hat{q}_A$ will have the smaller mean squared error if $n \geq 12{,}205$. This is an extremely large number of lives to observe for a single cell.

As a second example, consider the cell with the largest exposure ($98,367,065 at ages 45-49) in the Basic Male Table used in constructing the proposed New Minimum Mortality Standard for the Valuation of Standard Individual Ordinary Life Insurance. Assuming an average size policy of $7000 (from the Life Insurance Fact Book 1978) a sample size of about 14,000 results. The mortality rate was .00235 for this cell, leading to $a = -.000122$ and $b = .00348$ using the same principles as in the first illustration but with $\mu = 8.45$ and $\sigma^2 = .81$ to provide for an average policy size of $7000. With these parameters, a sample size of 265,352 is required before $\hat{q}_A$ will have the smaller mean squared error. In addition, when a is negative, the bias in $\hat{q}_L$ will be positive and thus it can be expected that the error in $\hat{q}_L$ will be on the conservative side.

It is reasonable to conjecture that the robust approaches will provide a suitable compromise between amounts and lives. A smaller bias than $\hat{q}_L$ and a smaller variance than $\hat{q}_A$ may yield an estimator that is optimal for moderate sample sizes. For the estimator $\hat{q}_R$ define $h(x) = x$ if $x \leq \exp(\mu + c\sigma) = \delta$ and $h(x) = \delta$ otherwise. Then

$$\alpha_h = \alpha\Phi(c-\sigma) + \delta(1-\Phi(c))$$

$$\beta_h = \beta\Phi(c-2\sigma) + \delta^2(1-\Phi(c))$$

$$\gamma_1 = b\alpha_h + a\alpha[(\mu+\sigma^2)\Phi(c-\sigma) - \sigma\ \exp(-(c-\sigma)^2/2)/\sqrt{2\pi}]$$
$$+ a\delta[\mu(1-\Phi(c)) + \sigma\ \exp(-c^2/c)/\sqrt{2\pi}]$$

$$\gamma_2 = b\beta_h + a\beta[(\mu+2\sigma^2)\Phi(c-2\sigma) - \sigma\ \exp(-(c-2\sigma)^2/2)/\sqrt{2\pi}]$$
$$+ a\delta^2[\mu(1-\Phi(c)) + \sigma\ \exp(-c^2/2)/\sqrt{2\pi}]$$

where $\alpha = \exp(\mu+\sigma^2/2)$, $\beta = \exp(2\mu+2\sigma^2)$ and Φ is the standard normal cumulative distribution function. Using the values from the Equitable illustration, the calculations for $\hat{q}_R$ are presented in the following table.

	c	$b(\hat{q}_R)$	$n\ Var(\hat{q}_R)$	Prefer for
$(\hat{q}_L)$	$-\infty$	.00130	.020902	$n < 2046$
	0	.00084	.022913	$2046 \le n < 9118$
	1	.00043	.027645	$9118 \le n < 29{,}717$
	1.5	.00025	.031107	$29{,}717 \le n < 78{,}026$
	2	.00013	.034704	$78{,}026 \le n < 434{,}935$
$(\hat{q}_A)$	∞	0	.041402	$n \ge 434{,}935$

It turns out that the formulas for $\hat{q}_T$ are identical to those for $\hat{q}_R$ with $\delta = 0$ used throughout. For this example, when selecting the value of c optimal for a given sample size, $\hat{q}_R$ will always have a smaller mean squared error than $\hat{q}_T$. The formulas for $\hat{q}_H$ will be two-sided versions of those for $\hat{q}_R$. The mean squared errors are about the same, but there could appear to be no practical motivation for adjusting the small amount policies.

6. APPLYING $\hat{q}_R$ TO PRACTICAL SITUATIONS

In applications, it is necessary to estimate μ and σ^2 in order to use $\hat{q}_R$ with a given value of c. Instead of using the sample mean and variance of the $\ell n(X_i)$, it would be in keeping with the spirit of robust estimation to use robust estimators for these quantities. Huber suggests a method for use when both location and scale are unknown. There are many other proposals that use similar reasoning (see the paper by Lenth in this volume); none has overwhelming support. The method uses iteration to find the solution to

$$\frac{1}{n}\sum \Psi\left(\frac{x_i-\tilde{\mu}}{\tilde{\sigma}}\right) = 0 \quad \text{and}$$

$$\frac{1}{n}\sum \Psi^2\left(\frac{x_i-\tilde{\mu}}{\tilde{\sigma}}\right) = E[\Psi(X)^2]$$

where $\Psi(x) = x$ for $|x| \le c$ and $\Psi(x) = c$ for $|x| \ge c$. $E[\Psi(X)^2] = 2\Phi(c) - 1 + 2c^2(1-\Phi(c)) - 2c\exp(-c^2/2)/\sqrt{2\pi}$. Values for various c are tabled below. An effective method of

c	.5	1	1.5	2	2.5	∞
$E[\Psi(X)^2]$	.1851	.5161	.7785	.9205	.9776	1

solving the two equations is to define weights $w_i = \Psi\left(\frac{x_i-\tilde{\mu}}{\tilde{\sigma}}\right)\Big/\left(\frac{x_i-\tilde{\mu}}{\tilde{\sigma}}\right)$. The equations are then easily solved for $\tilde{\mu} = \Sigma w_i x_i/\Sigma w_i$ and $\tilde{\sigma}^2 = \Sigma w_i^2(x_i-\tilde{\mu})^2/nE[\Psi(X)^2]$. Noting that the w_i depend on $\tilde{\mu}$ and $\tilde{\sigma}$ suggests starting with all $w_i = 1$, obtaining $\tilde{\mu}$ and $\tilde{\sigma}^2$, resetting the w_i with these values and then obtaining new $\tilde{\mu}$ and $\tilde{\sigma}$. This process is to be continued until the values of $\tilde{\mu}$ and $\tilde{\sigma}$ stabilize.

As an illustration, the methods studied were applied to selected cells from the Equitable data. Robust estimators were used for each cell with $\tilde{\mu}$ and $\tilde{\sigma}$ obtained from the $\ell n(x_i)$ and $\hat{q}_R$ obtained using $c = 1.5$.

Ages	Lives	Deaths	$\hat{q}_L$	$SD(\hat{q}_L)$	Exposures	Deaths	$\hat{q}_A$	$SD(\hat{q}_A)$
48-52	261	1	.00383	.0024	1,711,459	2,500	.00146	.0042
58-62	383	4	.01044	.0066	1,648,600	24,000	.01456	.0101
68-72	300	6	.02000	.0085	1,106,666	19,500	.01762	.0114
78-82	142	6	.04225	.0192	506,629	24,500	.04836	.0246
83-87	76	9	.11842	.0420	217,300	31,000	.14266	.0560

Ages	$\tilde{\mu}$	$\tilde{\sigma}$	δ	# Truncated	Exposures	Deaths	$\hat{q}_R$	$SD(\hat{q}_R)$
48-52	8.370	1.0396	20,525	5	1,631,084	2,500	.00153	.0032
58-62	7.865	.9216	10,378	25	1,417,750	24,000	.01693	.0084
68-72	7.656	.7589	6,597	27	795,785	17,694	.02223	.0102
78-82	7.578	.6976	5,566	12	337,171	18,632	.05526	.0225
83-87	7.622	.7597	6,384	3	194,452	31,000	.15942	.0505

The standard deviations were calculated assuming $q(x) = q = \hat{q}_R$, $\mu = \tilde{\mu}$ and $\sigma = \tilde{\sigma}$. Then $Var(\hat{q}_L) = q(1-q)/n$, $Var(\hat{q}_A) \doteq q(1-q)\beta/n\alpha^2$

and

$$\mathrm{Var}(\hat{q}_R) \doteq q(1-q)[\beta\Phi(c-2\sigma) + \delta^2(1-\Phi(c))]/n[\alpha\Phi(c-\sigma) + \delta(1-\Phi(c))]^2.$$

7. OTHER CONSIDERATIONS

Robust methods are designed to be superior when the observations are either contaminated or are from a heavy-tailed distribution. The analysis in this paper assumed a normal distribution (after the logarithm transformation). It would be reasonable to conjecture that the presence of outliers would improve the superiority of $\hat{q}_R$ over $\hat{q}_A$. The derivation of asymptotic variances under such contamination is only slightly more complex. A greater difficulty is that the methods of obtaining mean squared errors were all based on assuming μ and σ^2 known. In practice, the use of estimates of μ and σ^2 will lead to errors in the selection of the cutoff point δ and in the mean squared error measurement.

Perhaps the most useful result from this investigation is the attention focused on the variance of mortality estimates. In particular, when using amounts of insurance (whether adjusted or not) it is common to think of confidence intervals in terms of $\sqrt{\hat{q}(1-\hat{q})/n}$. It is clear that the coefficient $\sqrt{\beta}/\alpha$ must also be considered when obtaining the standard deviation and its effect can be of significant magnitude.

8. A NOTE ON USING POLICIES

The comparisons made in this paper were between the ideal, use of lives, and the accepted, use of amounts, methods of estimation. If amounts are to be discarded, the next most practical alternative is the use of policies, a readily available quantity. If P_i is the number of policies held by life i and θ_i is as before, define $\hat{q}_P = \Sigma\theta_i P_i/\Sigma P_i$. If θ_i and P_i are independent, $q = E[\theta_i] = E[\hat{q}_P]$ and $n\ \mathrm{Var}(\hat{q}_P) \longrightarrow q(1-q)E[P^2]/E[P]^2$. Suppose $P-1$ has a Poisson distribution with mean λ. Then $E[P] = \lambda+1$ and $E[P^2] = (1+\lambda)^2+\lambda$ and therefore $n\ \mathrm{Var}(\hat{q}_P) \longrightarrow q(1-q)(1+\lambda/(1+\lambda)^2)$. Estimating by policies would be preferable to amounts if $1+\lambda/(1+\lambda)^2 \leq \beta/\alpha^2$. This is equivalent to $\lambda/(1+\lambda)^2 \leq \mathrm{Var}(X)/E[X]^2$. Since $\lambda/(1+\lambda)^2 \leq 1/4$ for all $\lambda \geq 0$, $\hat{q}_P$ will always be preferred when $E[X] \leq 2SD(X)$. This relationship held for all five cells used in the Equitable example.

9. ACKNOWLEDGEMENT

My thanks to Don Iverson at The Equitable Insurance Company of Iowa for providing the sample from one of their mortality studies. Due to the nature of the sample and the small sample sizes, the values of the estimators in no way reflect the actual mortality experience of the company.