USING THE BOUNDING INTERVIEWS IN NCS

Ha H. Nguyen[1] and Charles Alexander[1], US Bureau of the Census
Ha H. Nguyen, US Bureau of the Census, Washington DC 20233

1. Introduction

The National Crime Survey (NCS), conducted by the Bureau of the Census for the Bureau of Justice Statistics, began in July 1972. In this survey, the members (12 years or older) of a sampled housing unit are interviewed seven times at six-month intervals. During each interview, the respondents are asked to report all incidents of certain type of crimes that they have experienced during the 6 months preceding the interview month. At any given time, the NCS sample consists of one incoming rotation group that is interviewed for the first time (first time in sample) and six returning rotation groups (second to seventh time in sample).

Previous studies have shown that, covering the same time periods, the number of crimes reported in the first interview is substantially higher than the average for the second through seventh interviews (Murphy and Cowan, 1976). This is attributed in part to "telescoping" into the reference period of crimes which occured before it. Because of this "external telescoping", the initial unbounded interview results have not been used in NCS estimates. The first interview is only used to provide a bound to avoid telescoping into the reference period for the second interview.

If a way could be found to adjust the initial interviews so as to remove the effect of external telescoping, then these interviews could be added to the effective NCS sample. Such adjustments have not been made previously, primarily because of uncertainty about what level to adjust to. One problem is that even the bounded interviews are though to be affected by forgetting of crimes and other response errors. Further, the different interviews (two through seven) may have different expected values due to increasing experience with the survey. Discussions at the Census Bureau have centered around adjustment of the first interviews results to the average level of the other interviews, which corresponds to the present survey results. Lavange and Folsom (1985) suggested adjustment to the level of the second interview, which arguably is the bounded interview least affected by the respondent's fatigue. They also propose making adjustments for other recall effects using an additive model.

The present research is intended to consider a variety of NCS estimators which incorporate data from the first interview in different ways. This paper presents some preliminary results of the initial portions of the research. The paper considers some extremely simple models for the relationship of the different interviews, in order to illustrate some basic issues involved in deciding how to use the bounding interviews. Asymptotic expressions can be derived for the variances of some of the estimators, under these models. These expressions may prove to be useful in deriving variances for these estimators from the usual NCS variances.

This paper will concentrate on the problem of adjusting the bounding interview to the average level of interviews 2,...,7. Thus the goal is to reduce the variance by including additional data in the estimates, rather than change the bias which may be present in the current data. To simplify the notation in our initial investigation, it will be assumed that the estimators based on each of interviews 2,...,7 have the same expected values. This assumption will be weakened in subsequent work.

This paper will also ignore the problem of unbounded interviews which occur in visits 2,...,7. This occurs when the occupants of a household move and are replaced, or when some person in the household could not be interviewd at the previous visit. Approximately 15% of the "bounded" interviews are actually unbounded in this way. Adjustment for this effect is complicated by the fact that these households are "self-selected" based on their decision to move, rather than a random sample of all households.

The analysis in this paper could be applied to any single type of crime. Different parameters might apply to different types of crime.

2. Model for the telescoping effect

Let $X_r(y)$ be the estimate of the number of crimes in year y, X(y), based on the r^{th} inteview only (r=1,...,7). Assume that $X_r(y)$ (r=2,...,7) is an unbiased estimate of X(y).

$$E(X_r(y)) = X(y) \qquad r=2,...7. \qquad (1)$$

Since the number of crimes reported in the 1^{st} interview, $X_1(y)$, is higher than in the other interviews, we need an additional term, B, for its expected value. We can make the additional term additive, i.e.,

$$E(X_1(y)) = X(y) + B, \qquad (B>0) \qquad (2)$$

or multiplicative

$$E(X_1(y)) = B\,X(y). \qquad (B>1) \qquad (3)$$

For simplicity, we will assume that B is constant from year to year, and from rotation to rotation. We note that (3) can be rewritten as

$$E(X_1(y)) = X(y) + (B-1)X(y).$$

If X(y) is relatively constant from one year to the next, then this has approximately the same form as (2), i.e., the multiplicative and additive models will be approximately the same, as long as crime rates do not change too much.

If crime rates were to change drastically, the additive model would probably not continue to apply. For example, suppose that first interviews would lead to an estimate of about 8,000,000 burglaries, compared to an unbiased (under the model) estimate of about 6,000,000 based on interviews 2,...,7. In this illustration, unbounded interviews would give results about 33% too high. If the actual number of burglaries were to drop over time to about 1,000,000 burglaries, the additive model would require unbounded interviews to give an expected 3,000,000 burglaries, or 200% too high. This would require a major change in the recall process for those households which have burglaries. The multiplicative model would assume that unbounded interviews would continue to be 33% too high. Thus the authors feel that the multiplicative model is more plausible in principal. In practice, since crime rates seem to change slowly, there is probably little difference between the models.

In this paper, we will primarily use the multiplicative model for the telescoping effect. The estimate of the number of crimes reported in NCS which are based on bounded interviews will be denoted by $\hat{X}(y)$:

$$\hat{X}(y) = \frac{1}{6} \sum_{r=2}^{7} X_r(y). \qquad (4)$$

Under assumption (1), we remark that

$$E(\hat{X}(y)) = \frac{1}{6} \sum_{r=2}^{7} E(X_r(y)) = X(y).$$

The estimators proposed in this paper will be classified into 2 groups:

1/ Those that use one year of data

2/ Those that use two or more consecutive years of data.

3. Using one year of data

Assume that $X_r(y)$ (r=1,...7) is the number of crimes reported in calendar year y in the r^{th} interview, and we would like to use all these interviews to construct an estimate for X(y), the true number of crimes in year y. First, we note that the mean of the seven $X_r(y)$ (r=1,...,7) overestimates X(y): Let

$$\hat{X}^*(y) = \frac{1}{7} \sum_{r=1}^{7} X_r(y)$$

then,

$$E\left[\hat{X}^*(y)\right] = \frac{1}{7} \sum_{r=1}^{7} E\left[X_r(y)\right] = \frac{1}{7} (B+6)\ X(y). \qquad (5)$$

To correct this bias, it is then necessary to get a good estimate of B.

3.1. Method of moments.

Using the method of moments, an estimate of B which uses all the seven $X_r(y)$ can be found by noting that:

$$B = \frac{E(\ X_1(y)\ )}{X(y)} = \frac{E(\ X_1(y)\ )}{E(\ \frac{1}{6}\ \Sigma_2^7\ X_r(y))}$$

$$\text{or}\quad est(B) = \tilde{B} = \frac{X_1(y)}{\frac{1}{6} \sum_{r=2}^{7} X_r(y)} \qquad (6)$$

Using (6) in (5) we obtain a new estimate for X(y) :

$$X(y) = \frac{\sum_{r=1}^{7} X_r(y)}{\tilde{B}+6} = \hat{X}(y).$$

The new estimate of X(y) is the same as the one currently used in NCS! A closer inspection of $\tilde{B}$ in (6) indicates that this estimate downweights $X_1(y)$ to the average of $X_r(y)$, (r=2,...,7). This explains why the bounding interview did not contribute to the final estimate. Based on one year of data, an additive model will lead to the same conclusion, using the usual linear model parameter estimates.

3.2. Maximum likelihood.

In order to derive a maximum likelihood estimate for X(y), we need assumptions about the distribution of $X_r(y)$. A distribution frequently used for discrete random variable that counts the number of occurences of an event is the Poisson distribution. Let's assume that

$$X_1(y) \sim P(\ BX(y)\ )$$

$$X_r(y) \sim P(\ X(y)\ ) \qquad r=2,\ldots,7.$$

It can be shown that the Maximum Likelihood estimators are :

$$\tilde{B} = \frac{X_1(y)}{\hat{X}(y)}$$

and $$\tilde{X}(y) = \hat{X}(y).$$

Again, the new estimator coincides with $\hat{X}(y)$, the estimator based on bounded interviews. It can also be shown that if we use the additive model for the telescoping effect, and the Poisson assumption

$$X_1(y) \sim P\ (X(y)+B\ )$$

$$X_r(y) \sim P\ (\ X(y)\ ) \qquad r = 2,\ldots,7$$

then the Maximum Likelihood estimator is also $\hat{X}(y)$.

Although these results may seem obvious, they imply an important lesson. Using the unbounded interviews will not necessarily reduce the variance, even though more sample cases are used in deriving the estimator. Whether there is a variance reduction depends on the form of the estimator. That remark also holds for other assumptions about the expected value of $X_r(y)$. For example, adjustment to the level of the second interview (*i.e.*, $E(X_r(y))=B_rX(y)$ for $r\neq 2$ and $E(X_2(y))= X(y)$) would lead to an estimator based on a single year's data which is equal to $X_2(y)$. This is based on data from only one-sixth of the bounded interviews and thus would have roughly six times the variance of the present NCS estimator, even though in theory the estimator could be said to be based on more observations. Similarly adjustment to the average of the second through seventh interview (*i.e.*, $E(X_r(y))=B_rX(y)$ and $E(\frac{1}{6}\sum_2^7 X_r(y))=X(y)$) would lead to $\hat{X}(y)$, the same estimate as the one currently used in NCS.

Since using ad hoc estimators based on one year of data does not change the estimate of X(y), it is thus natural to try using two years of data.

4. Using two years of data

4.1. Maximum Likelihood

A simple extension of the Poisson assumption for 2 years of data can be made as follows:

$$
\begin{array}{lll}
X_1(y-1) & \sim & P(\ BX(y-1)\ ) \\
X_r(y-1) & \sim & P(\ X(y-1)\ ) \qquad r=2,\dots,7 \\
X_1(y) & \sim & P(\ BX(y)\ ) \\
X_r(y) & \sim & P(\ X(y)\ ) \qquad r=2,\dots,7
\end{array}
$$

It can then be shown that MLE for B, X(y-1), X(y) are:

$$\tilde{B} = \frac{X_1(y-1) + X_1(y)}{\hat{X}(y-1) + \hat{X}(y)}$$

$$\tilde{X}(y-1) = \frac{\hat{X}(y-1) + \hat{X}(y)}{\hat{X}^*(y-1)+\hat{X}^*(y)}\ \hat{X}^*(y-1)$$

$$\tilde{X}(y) = \frac{\hat{X}(y-1) + \hat{X}(y)}{\hat{X}^*(y-1)+\hat{X}^*(y)}\ \hat{X}^*(y).$$

This MLE is basically $\hat{X}^*(y)$ (the adjusted estimator based on all interviews) downweighted by the ratio of the average of estimates of X(y) and X(y-1) based on bounded and all interviews. We also note that $\tilde{X}(y-1)$, $\tilde{X}(y)$, $\hat{X}(y-1)$, $\hat{X}(y)$ are related by

$$\tilde{X}(y-1) + \tilde{X}(y) = \hat{X}(y-1) + \hat{X}(y).$$

This MLE looks intuitively acceptable, but more study must be done before we can comfortably use it. How can we justify the Poisson assumption? Should there be terms in the joint distribution that represent the correlation between $X_r(y)$ and $X_{r-1}(y-1)$? At year y, we have 2 MLE for X(y-1): one based on (y,y-1) data, one based on (y-1,y-2) data. How should we handle this?

More than two years data may be used in a similar fashion, with further reduction in variance. However, this requires assuming that B remains constant for an extended period of time. The optimal number of years to use for estimators under this model remains to be determined.

4.2. Estimator based on relative year to year change.

Thus far, all proposed estimates of X(y) require first an estimate of B. In this section we will introduce estimates that do not require an estimate of B. For that let's define R, the relative year to year change, to be:

$$R = \frac{X(y) - X(y-1)}{X(y-1)}.$$

Then it is observed that the estimate of R based on all interviews is asymptotically unbiased. Let

$$\hat{R}(y) = \frac{\hat{X}^*(y) - \hat{X}^*(y-1)}{\hat{X}^*(y-1)}$$

then,

$$E\left[\hat{R}(y)\right] \cong \frac{E\left[\hat{X}^*(y) - \hat{X}^*(y-1)\right]}{E\left[\hat{X}^*(y-1)\right]} \cong R$$

And we can use $\hat{R}$ to construct new estimates of X(y) without having to compute B. Under some conditions, $\hat{R}$ may have smaller variance than the corresponding estimator based on interviews 2,...,7. However, there is an immediate problem of how to obtain an annual level estimate for X(y) such that the estimates for X(y) and X(y-1) are consistent with $\hat{R}$.

A simple estimate of X(y) based on $\hat{R}(y)$ will have the form

$$\tilde{X}(y) = \Big[\text{ Estimate of } X(y-1)\Big]_1 + \hat{R}(y)\Big[\text{ Estimate of } X(y-1)\Big]_2 . \quad (7)$$

If the 2 estimates of X(y-1) in (7) are equal, then (7) will be reduced to

$$\tilde{X}(y) = \frac{\hat{X}^*(y)}{\hat{X}^*(y-1)}\Big[\text{Estimate of } X(y-1)\Big].$$

Up to now we only have 2 good estimates of X(y-1) ($\hat{X}(y-1)$ and $\tilde{X}(y-1)$); hence, it is natural to consider the 4 following $\tilde{X}(y)$:

$$\tilde{X}(y) = \tilde{X}(y-1) + \hat{R}(y)\tilde{X}(y-1) = \frac{\hat{X}^*(y)}{\hat{X}^*(y-1)}\tilde{X}(y-1) \quad \text{(a)}$$

$$\tilde{X}(y) = \tilde{X}(y-1) + \hat{R}(y)\hat{X}(y-1) \quad \text{(b)}$$

$$\tilde{X}(y) = \hat{X}(y-1) + \hat{R}(y)\hat{X}(y-1) = \frac{\hat{X}^*(y)}{\hat{X}^*(y-1)}\hat{X}(y-1) \quad \text{(c)}$$

$$\tilde{X}(y) = \hat{X}(y-1) + \hat{R}(y)\tilde{X}(y-1). \quad \text{(d)}$$

Among those 4 estimates, (b) and (d) are more attractive than the 2 others: We expect $\tilde{X}(y)$ to be better than $\hat{X}(y)$ and (c) only contains $\hat{X}(y)$. Also it can be shown that using (a) recursively the estimate in (a) will be reduced to

$$\tilde{X}(y) = \frac{\hat{X}^*(y)}{\hat{X}^*(y_0)} \tilde{X}(y_0)$$

(where y_0 is the first year of available data). In other words, $\tilde{X}(y)$ in (a) depends on $\hat{X}^*(y)$ and $\hat{X}^*(y_0)$, $\tilde{X}(y_0)$ only. Consequently the estimate in year y will always be adjusted by the ratio $\tilde{X}(y_0)/\hat{X}^*(y_0)$ from the first year of the survey. This is clearly undesirable since it ignores data from intervening years, and is heavily dependent on the assumption that B is constant over a long period of time. This is especially questionable for the first year of the survey, when a new sample was still being "rotated in".

It can also be shown that if we want $\tilde{X}(y)$ to yield new estimate of R(y) (based on $\tilde{X}(y)$) consistent with $\hat{R}(y)$ then we must use the estimate (a) which is not very desirable.

Expressions for the variance of $\tilde{X}(y)$ in (b) and (d) can be derived using a Taylor expansion.

Weakening the requirement that $\tilde{R}(y)$ must be consistent with $\hat{R}(y)$ raises other possibilities. For example, let

$$\approx{X}(y) = \beta\hat{X}(y) + (1-\beta)\ (\hat{R}(y) + 1)\ \hat{X}(y-1) \quad (8),$$

where $0<\beta<1$.

Thus $\overset{\approx}{X}(y)$ is a weighted average of the usual NCS estimator and a "projection" from the previous year's usual estimator, using the full-sample estimate $\hat{R}(y)$ in the projection. Another alternative would be a recursive definition

$$\overset{\approx}{X}(y) = \beta\hat{X}(y) + (1-\beta)\ (\hat{R}(y) + 1)\ \overset{\approx}{X}(y-1). \quad (9)$$

This would avoid the excessive dependence on the year y_0, since

$$\overset{\approx}{X}(y) = \beta\left[\ \hat{X}(y) + (1-\beta)\hat{R}(y)\ \hat{X}(y-1)\ \right] + (1-\beta)^2\ \hat{R}(y)\hat{R}(y-1)\ \overset{\approx}{X}(y-2),$$

and, continuing the recurrence, the effect of the distant past is "damped out" by increasing powers of $(1-\beta)$.

4.3. Method of moments

If we have data for two consecutive years,

$X_1(y), X_2(y), \quad X_3(y), \ldots\ldots, X_7(y)$

$X_1(y-1), X_2(y-1), \ldots., X_6(y-1), X_7(y-1)$

then we can use the rotations common to both years y and y-1 to construct an estimate for B. We remark that:

$$B = \frac{E(X_1(y-1))}{E(X_2(y))} \frac{E(X_3(y)+\ldots..+X_7(y))}{E(X_2(y-1)+\ldots+X_6(y-1))}$$

and an estimate of B, based on the method of moments, is:

$$\tilde{B} = \frac{X_1(y-1)}{X_2(y)} \frac{\sum_{r=3}^{7} X_r(y)}{\sum_{r=2}^{6} X_r(y-1)};$$

a new estimate of X(Y) can be obtained by substituting the above $\tilde{B}$ in

$$\tilde{X}(y) = \frac{\sum_{r=1}^{7} X_r(y)}{\tilde{B} + 6}.$$

Using a Taylor expansion, an expression for the variance of $\tilde{X}(y)$ can be obtained.

Note that this method relies heavily on the assumption

$$E(\hat{X}_2(y)) = E(\hat{X}_3(y)) = \ldots\ldots. = E(\hat{X}_7(y))$$

thus it may be difficult to generalize when this assumption is weakened.

5. Conclusion

From the above preliminary study, we conclude that in order to use effectively the bounding interview, we need to construct new estimators based on <u>at least</u> two years of data. A few elementary estimators considered, based on one year of data, either reduce to $\hat{X}(y)$ or $X_2(y)$ depending on the assumption made on $E(X_r(y))$. Some elementary estimators based on two consecutive years of data have been proposed, and approximations for their variances can be derived using a Taylor expansion. (The variance formulas are quite complicated.) A comparison of these new estimators with the one currently used in NCS, $\hat{X}(y)$, based on empirical results, will be the topic of future study. For estimators based on an estimate of R(y), the relative change from year y-1 to y, there is a particular problem with finding a good estimator $\tilde{X}(y)$ which has the corresponding $\tilde{R}(y)$ consistent with $\hat{R}(y)$ (the estimate of R(y) based on all interviews).

Also of interest for further work is the generalization of the above multiplicative model to

$$E(X_r(y)) = B_r X(y) \qquad r=2,\ldots 7$$

which is more realistic and allows us to include the panel bias effect. Additionally, as discussed in section 4, we will consider composite estimators (Wolter, 1979) for the estimate of X(y). Since, for each year, we have 2 estimators, $\hat{X}(y)$ (based on bounded interviews) and $\tilde{X}(y)$ (based on all interviews), a composite estimator, $\tilde{X}^*(y)$, can be constructed :

$$\tilde{X}^*(y) = \beta\hat{X}(y) + (1-\beta)\tilde{X}(y)$$

and this is often known to have smaller variance than both $\hat{X}(y)$ and $\tilde{X}(y)$ for a good choice of β.

References

Lavange, Lisa M. & Folsom, Ralph E. (1985) "Development of NCS error adjustment models". Reseach Triangle Institute Technical Report.

Murphy , Linda R. & Cowan, Charles D. (1976) "Effects of Bounding on Telescoping in the National Crime Survey," Paper presented at the annual meetings of the American Statistical Association.

Wolter, Kirk M. (1979) "Composite Estimation in Finite Populations," *Journal of the American Statistical Association*, 74, 604-613.

Footnote

(1) This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.