Proceedings of Statistics Canada Symposium 2001
Achieving Data Quality in a Statistical Agency: A Methodological Perspective

IMPROVING MEASURES OF CRIME: SAMPLE ADJUSTMENTS TO POLICE CRIME DATA

James J. Nolan, III[1], Yoshio Akiyama[2] and James A. Woods[3]

ABSTRACT

Police records are the source of national crime statistics that are collected by the FBI through the Uniform Crime Reporting (UCR) Program. Audits to correct UCR records have recently raised concerns as to how to handle the errors that are discovered. Concerns center on the methodology used to detect errors and procedures for correcting errors once they are discovered. The authors explore this, focusing on sampling methodology, establishment of a statistical-adjustment factor, and alternative solutions. They distinguish between sample adjustment and sample estimates of an agency's data, and recommend sample adjustment as the most accurate way of dealing with errors.

KEY WORDS: Sample-adjustment; Statistical-adjustment factor; Uniform Crime Reports.

1. INTRODUCTION

1.1 Introduction

Police records are the source of the U.S. national crime statistics and are collected by the FBI through the Uniform Crime Reporting (UCR) Program. The UCR Program is a nationwide, cooperative statistical effort of over 17,000 city, county, and state law enforcement agencies voluntarily reporting data on crimes reported to them (USDOJ, 2000). The primary objective of the Program is to generate a reliable set of criminal statistics for use in law enforcement administration, operation, and management; however, its data have over the years become one of the country's leading social indicators. The public looks to Uniform Crime Reports for information on fluctuations in the level of crime, and criminologists, sociologists, legislators, municipal planners, the media, and other students of criminal justice use the statistics for varied research and planning purposes.

Although the quality of police records has often come into question (Merton, 1957; Kitsuse and Cicourel, 1963; Black, 1970; Skogan, 1975; Gove, Hughes, & Geerken, 1985), the UCR is a popular source of crime data in the United States (Regoli & Hewitt, 2000; Schmalleger, 1999; Siegal & Senna, 2000). Although UCR has been around since the 1930s, the auditing of those records to confirm the degree to which errors occur is a recent phenomenon. Recent attempts to audit and correct UCR crime records in large U.S. cities, e.g., Baltimore, Maryland and Philadelphia, Pennsylvania, have raised concerns as to how to handle the errors that

[1]James J. Nolan, III, Division of Sociology and Anthropology, West Virginia University, 316 Knapp Hall, Morgantown, WV, USA 26506-6326.

[2]Yoshio Akiyama, Federal Bureau of Investigation, J. Edgar Hoover Building, Room 11194, 935 Pennsylvania Avenue, Washington, DC, USA 20535.

[3]James A. Woods, Crime Analysis, Research, and Development Unit, Criminal Justice Information Services Division, Federal Bureau of Investigation, 1000 Custer Hollow Road, Clarksburg, WV, USA 26306.

are discovered. Some of the concerns center on the particular methodology employed to detect the error, while others pertain to procedures for correcting errors once they are discovered. Since the question of handling errors through auditing is new, no policy has yet been established by the UCR Program. Currently, when a city submits no data for a particular year, these missing data are imputed as the mean crime rate of the agency's population stratum, i.e., $C_{estimate} = RP_i/100{,}000$, where R = mean crime rate of the stratum, and P_i = agency population.

This paper presents an alternative method for adjusting data based on the review of a sample. The authors focus particularly on sampling methodology, the establishment of a statistical-adjustment factor, and alternative solutions to dealing with the data errors.

2. BASIC TERMINOLOGY

2.1 Classification Errors and Recording Errors

There are two types of errors that can be detected in a review of crime data: classification errors and recording errors.

2.1.1 Classification Errors

Classification errors occur when an officer correctly records the facts of a case, but mis-classifies the crime type or when the computer mis-classifies a crime in an automated conversion process. For example, an officer classifies a crime as a “simple assault,” but records in his/her report that the victim was assaulted with a club. The use of the club makes this an "aggravated assault" based on UCR definitions. Classification errors are detected by reading police reports. These are the types of errors usually uncovered by the Quality Assurance Review (QAR) process and can often involve system-generated errors in the conversion of state statutes to UCR codes[4].

2.1.2 Recording Errors

Recording errors occur when the officer misinterprets what the victim initially reported and, therefore, records the wrong crime type. For example, a police report may describe the theft of a wallet from a vehicle. However, it could be determined in a follow-up call to the victim that during the theft, the suspect threatened her with a gun, making the crime a robbery. Recording errors are detected in direct follow-up calls to the victims or by listening to the complaints telephone logs (in the case that they are still available).

2.2 Record Accuracy and Statistical Accuracy

There are two types of accuracy to consider: record accuracy and statistical accuracy.

2.2.1 Record Accuracy

Record accuracy refers to the percent of errors found in the collection of a particular crime type. For example, a check of simple assault records may reveal that 3 percent of the reports examined were mis-classified (i.e., should have been classified as some other crime). More important, there are many types of record errors that do not lead to re-classification of UCR crime. For example, a $400 larceny loss mistakenly recorded as $440

[4] System-generated errors in the conversion of state statutes to UCR code should be corrected through computer software modifications rather than through sample adjustment.

does not change the fact that larceny was committed. Likewise, if the correct date (5/14/00) of the theft was recorded as 5/15/00, this is a record error even though it does not affect UCR statistics. One record may contain multiple errors. What constitutes a record error is up to the agency. As such, record accuracy can be defined differently for various purposes.

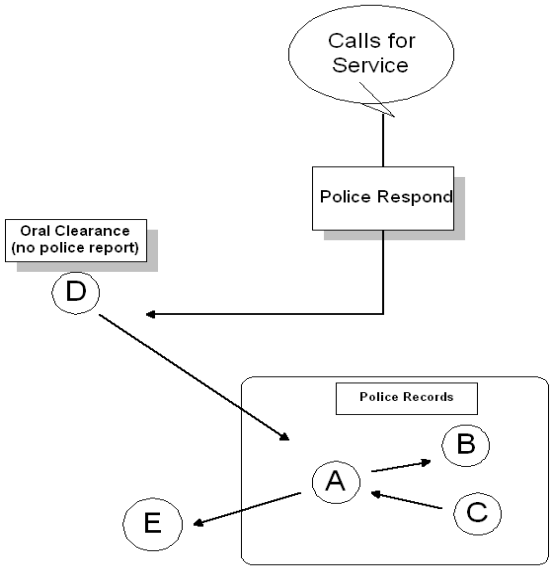
2.2.2 Statistical Accuracy

Statistical accuracy refers to the percentage of error found in the totals after all crime types have been examined and offsetting mis-classifications have been considered. Since some mis-classifications lead to over-reporting (over-counting) of UCR crimes while others lead to under-reporting (under-counting), the correct UCR number can be obtained by considering the canceling effect of the two types of errors--over-counting errors and under-counting errors. Statistical accuracy should have a uniform meaning from one agency to another. Sample-adjustment arises from this.

According to Figure 1, a sample adjustment of the crime type “simple assault” would be calculated as follows (using a random sample of 100 simple assaults as an example):

A = 100 (a random sample of 100 reported simple assaults)
B = 10 (crimes reported as “simple assault (A)” but were “aggravated assaults (B)”
C = 10 (crimes reported as “intimidation (C)” that were actually “simple assault”
D = 20 (crimes cleared orally “no report (D)” by the dispatched officers, but which actually should have been recorded as “simple assault”
E = 10 crimes recorded as “simple assault” that were actually not crimes at all (E).
F = 100 - (10+10) + (10+20) = 110 (actual simple assaults were found in the random sample or F = A - (B + E) + (C + D)).

Figure 1. Illustration of the Calculation of a Sample Adjustment of the Crime Type "Simple Assault."



The population of simple assaults (all simple assault records for the year under study) could be adjusted by multiplying the total recorded simple assaults by 1.1 (i.e. 110/100 = 1.1). In the meantime, there were 20 over-classifications (B + E = 20) and 30 under-classifications (C + D = 30). Therefore, for example, an agency can state that the error ratio was 50% (based on the argument that there were 50 errors committed while counting 100 crimes).

2.3 Estimation and Statistical Adjustment

There are two types of statistical procedures to consider: estimation and statistical adjustment.

2.3.1 Estimation

Estimation in this context refers to a statistical procedure used to estimate or impute crimes as opposed to counting the total number of records. Crime estimation is used by the UCR Program to account for agencies for which no crime data are submitted; incomplete data are submitted, i.e., there are missing months; or reported data are found unacceptable by the national UCR Program. The term "estimation" also applies in the reporting process, when an agency submits data that are estimated by such procedures as sampling, forecasting from past data, etc. UCR does not accept sample-estimated reporting.

2.3.2 Statistical Adjustment

Statistical adjustment refers to a procedure for adjusting the total reported crime based on the errors identified in a sample. Statistical adjustments are for those agencies that have complete crime reports (counted from their total files), but have found substantial errors in these reports that require statistical adjustment.

2.4 Sample-Counted Data and Sample-Adjusted Data

An important distinction must be made between sample-adjusted data and sample-counted data.

2.4.1 Sample-Counted Data

Sample-counted data would allow agencies to compute their total number of offenses in their records by taking a sample of their data and statistically computing their total numbers. For example, if an agency has 20,000 records they would look at, for example, 2,000 to determine the number of aggravated assaults in the 20,000 records.

2.4.2. Sample-Adjusted Data

Sample-adjusted data allows an agency to correct the number and types of offenses in its records by counting all offenses and then computing a statistical correction based on the errors detected. For example, if an agency has 20,000 records they count the 20,000 records and determine that they have 4,000 aggravated assaults. They then sample 2,000 records and determine that of 200 aggravated assaults in the sample, 20 should have been classified as simple assaults. This 10 percent error rate would determine that 400 of the 4,000 total aggravated assaults should have been be reclassified as simple assaults.

3. METHODOLOGY

3.1 Selecting the Sample Size

We can apply the following formula **[1]** to select the sample size for each crime type.

$$n = \frac{\tilde{k}^2 N \tilde{P}\tilde{Q}}{\tilde{k}^2 \tilde{P}\tilde{Q} + N \tilde{E}^2}, \qquad \textbf{[1]}$$

where

n = the sample size to be computed

$\tilde{k}$ *= multiple of standard error (e.g.,* $\tilde{k}$ *= 1.96 for 95 % confidence)*

$\tilde{E}$ *= admissible limit of error for estimated mean (e.g.,* $\tilde{E}$ *= .05)*

N = population size (total number of actual reports in each crime category under consideration)

$\tilde{P}$ *= pre-estimated error rate.*

$\tilde{Q} = 1 - \tilde{P}$ *= estimated rate of accuracy.*

The factors $\tilde{k}$ and $\tilde{E}$ express the level of accuracy that are determined by the researcher/agency and should be applied uniformly to all crime categories.

3.2 Determining the Actual Error Rate and Confidence Interval

For each crime category in which the researcher/agency is interested, two types of errors should be examined: classification errors and recording errors.

3.2.1 Classification Error Calculation

The classification errors are calculated by reading incident reports for a given period and determining how many crimes are mis-classified.

3.2.2 Recording Error Calculation

Recording errors are calculated using the following notations

N = the total number of records for a particular crime category.

e = the number of errors found in the initial reading of reports.

$P = \frac{e}{N}$ = the proportion of classification errors **[2]**

$\rho = \frac{m}{M}$ = the estimated proportion of record errors given X = 0 **[3]**

M = the sample size of records for which victims were contacted

m = the number of errors found in the contacted sample

$E = e + (N - e)\rho =$ the estimated number of all errors, *and* **[4]**

$\xi = \frac{E}{N} = P + \rho(1 - P)$ = the estimated proportion of all errors, **[5]**

the variable ρ has the following variance:

$$\sigma 2\,(\rho) = \left[\frac{N - e - M}{N - e} \right] \frac{\rho(1-\rho)}{M}. \qquad \textbf{[6]}$$

Since P is a constant, the variance of ξ is given by

$$\sigma^2(\xi) = (1-P)^2 \cdot \sigma^2(\rho), \qquad \textbf{[7]}$$

i.e., the standard deviation of ξ is

$$\sigma(\xi) = (1-P)\sigma(\rho) = (1-P)\sqrt{\left[\frac{N - e - M}{N - e - 1} \right] \frac{\rho(1-\rho)}{M}}. \qquad \textbf{[8]}$$

Once the estimated total error rate, ξ, is computed, its confidence interval is given by

$$\xi \pm \tilde{k}(1-P)\sqrt{\left[\frac{N - e - M}{N - e - 1} \right] \frac{\rho(1-\rho)}{M}}, \qquad \textbf{[9]}$$

where $\tilde{k}$ expresses the desired accuracy earlier chosen in the sample size formula. For a 95% confidence level, $\tilde{k}$ = 1.96. The confidence interval formula is a standard formula.

4. DATA AND RESULTS

4.1 Data

To demonstrate this procedure we present hypothetical data on a crime that we argue could have been obtained from an audit carried out by the police department in a large U.S. city. In our example, one thousand six hundred and fourteen reports were reviewed by the auditors. This represents all of the theft reports in the agency under study for a one year period. Sixty reports were found to be incorrectly classified. Using **[2]** above and dividing the number of changed reports (60) by the total number of reports reviewed (1,614), gives a proportion of classification errors (P) of .0372.

Reports reviewed (N):	1614
Reports read correctly:	1554
Reports changed (e):	60
Proportion of Classification Errors	$P = \frac{e}{N} = \frac{60}{1614} = .0372$

The police department then sampled the reports using **[1]** and interviewed 310 victims. Of these, 231 reports were correct and 79 were incorrect. Using **[3]** we get .2548 as the estimated proportion of recording error,

Victims sampled (M):	3	1	0
Reports unchanged:	2	3	1
Reports changed (m):	79		

Estimated Proportion of Recording Error $(\rho) = \frac{m}{M} = \frac{79}{310} = .2548$

The estimated proportion of all errors in the study is found by applying **[5]**:

Total error rate $(\xi) = \frac{456}{1614} = .2825 = P + \rho(1 - P).$

The standard deviations, found by using **[6]** and **[8]**, are:

$$\sigma(\rho) = \sqrt{\frac{N - e - M}{N - e - 1}\,\frac{\rho(1-\rho)}{M}} = \sqrt{\frac{1614 - 79 - 310}{1614 - 79 - 1}\,\frac{(.2548)(1-.2548)}{310}} = .0219.$$

and, $\sigma(\xi) = (1 - P) \cdot \sigma(\rho) = (.9628)(.0219) = .0211$

Therefore, when $\tilde{k}$ ' 1.96 (95% confidence level), the confidence interval for ξ is found by using **[9]**:

$\xi \pm 2\sigma(\xi) = .2825 \pm 2(.0211) = .2825 \pm .0422$, i.e., the interval is from .2403 to .3247.

5. DISCUSSION AND CONCLUSION

In the example presented above, the FBI UCR Program has two choices. It could (1) reject the sample-adjusted data and use an estimation procedure to impute the agency's crime data or (2) accept the sample-adjusted data and use it under certain terms and conditions.

It has been the FBI UCR Program policy to estimate crime data for law enforcement agencies that have supplied no data or have submitted incomplete data (i.e., there are missing data for certain months of the year). A limitation in the cross-sectional method used to calculate missing data is that it is not agency-specific. Instead, it estimates crime totals for a specific agency based on the mean crime rate of similar sized agencies in the same geographic area. The method does not assert that imputing the stratum average to a particular agency yields an accurate estimate for that agency. Rather, it is based on the statistical theory that the process yields an unbiased estimate for the stratum if the agencies to be imputed are randomly selected members of the stratum. Therefore, individual estimates are never published, but are used to calculate the state, regional, and national crime estimates. Moreover, these estimates are never corroborated by the police agency for which the estimate was generated.

If a law enforcement agency declares that its sample-adjusted data are accurate and reliable and that the sampling was conducted appropriately, it could be argued that these data are better than the estimated data calculated by the FBI and should be accepted under certain terms and conditions.

Sample adjustments to UCR crime data are applicable under the following general conditions: (1) when

requested by a law enforcement agency to correct for substantial errors uncovered during a methodologically sound audit (e.g., the audit uncovers a significant number of classification and recording errors); (2) when the agency is making "statistical adjustments," not "record adjustments" of individual crime types; (3) when the agency makes the adjustment for a specified period of time (e.g., one year) and has outlined the sample design (including study scope, proposed measurements, sample frames, sample sizes, logistics to carry out the design, etc.); (4) when the agency can verify and affirm the reliability and accuracy of the sample-adjusted data; (5) when the State UCR Program agrees to accept the agency's sample-adjusted reports; (6) when adjusting summary data only, not incident-based data; (7) when the reasonableness of the agency's adjusted data can be confirmed by the UCR Program; (8) when post-sampling documentation is provided by the agency, including error estimates (by type) in the file, associated confidence intervals and confidence levels.

The decision by the FBI to accept sample-adjusted data applies to both **recording errors** and to **classification errors.**[5] Further, it applies to the determination of **statistical accuracy,** not **record accuracy**. This policy pertains to an agency's **sample adjustments** of crime data, not the agency's **crime estimation**. Finally, the sample-adjusted data are not used in the calculation of strata averages for data quality control checks (i.e., outlier detection) or data imputation (i.e., estimates to fill in for missing data).

REFERENCES

Black, D. (1970), "Production of Crime Rates," *American Sociological Review, 35* (4), pp. 733-748.

Gove, W., Hughes, M., & Geerken, M. (1985), " Are Uniform Crime Reports a Valid Indicator of Index Crimes? An Affirmative Answer with Minor Qualifications," *Criminology, 23*, pp. 451-501.

Kitsuse, J. & Cicourel, A. (1963), "A Note on the Use of Official Statistics," *Social Problems 11*, pp. 131-139

Merton, R. (1957), *Social Theory and Social Structure, Revised*, Glencoe: The Free Press.

Regoli, R. & Hewitt, J. (2000), *Delinquency in Society*, Boston: McGraw Hill.

Schmalleger, F. (1999), *Criminal Justice Today*, Upper Saddle River, NJ: Prentice Hall.

Siegel, L. & Senna, J. (2000), *Juvenile Delinquency: Theory, Practice, and Law*, (7th ed.). Belmont, CA: Wadsworth/Thompson Learning.

Skogan, W. G. (1975), "Measurement Problems in Official and Survey Crime Rates, *Journal of Criminal Justice, 3*, pp. 17-32.

United States Department of Justice, Federal Bureau of Investigation (2000). *Crime in the United States, 1999*. Washington, DC: U.S. Government Printing Office.

[5] This is true only in situations where the classification errors are not system-generated state code to UCR code errors, but are randomly distributed errors made by officers when classifying and scoring the crime category.