SAMPLE SELECTION PROCEDURES FOR THE IACP UNIFORM CRIME REPORT AUDIT

David W. Chapman*, Westat, Inc.

1. Introduction

In 1974 the International Association of Chiefs of Police (IACP) began a project, sponsored by the Law Enforcement Assistance Administration, to investigate the quality of crime incidence reports that are submitted to the FBI by the Nation's police departments. This project, referred to as the IACP-UCR Audit/Evaluation Project, involves an audit of the processing of various types of information by police departments.

Twenty departments were audited during Phases II and III of the IACP-UCR Audit/Evaluation Project. Since there was no intent to make inferences from the 20 sample departments to all departments in the country, it was not necessary to select the sample on a probability basis. Consequently, the 20 departments were selected on a subjective basis. These test agencies were chosen by IACP personnel to be representative of the police departments across the country with respect to several characteristics.

For audit purposes the processing operation has been broken down into the following four stages:

Stage I - Telephone Tapes (Complaints)
Stage II - Complaint Control Cards
Stage III - Incident/Offense Reports
Stage IV - Clearance Data

Ideally it would be best to audit an agency by checking the accuracy of processing every piece of information at each of the four stages. However, this would be much too expensive and time consuming to do, especially in large departments. Therefore, a procedure was developed to sample the processing of information at the four stages for the audit check.

2. Sample Sizes and the Basic Selection Procedures

An initial decision had to be made between two possible basic selection procedures: (1) independent selection of cases at the four stages, and (2) selection of a sample of cases at Stage I to trace through the system.

Although it might have been useful to trace the processing of cases through the system, there would be a fundamental problem with this procedure. In order to have an adequate sample size for the latter stages (III and IV), a very large sample at Stage I would be required. Since sampling at Stage I (i.e., the telephone tapes) is the most time consuming phase of the audit, this procedure was not used.

Therefore, the first alternative, that of selecting independent samples of cases at each stage, was chosen for the audit procedures. In addition, it was decided to select records in such a way that would provide estimates of processing error rates with the same precision at each of the four stages.

In order to determine adequate sample sizes for selecting cases, the type of estimates to be made and the desired precision of such estimates had to be specified. The basic type of estimate calculated from the audit data is the estimated error rate at one of the four stages. This is defined as the estimated proportion of the cases processed at a stage that is incorrectly classified. The determination of whether or not a case was properly classified was a subjective judgment made by the IACP staff member doing the audit, based on specific guidelines.

For simple random sampling, estimates of the standard error of an estimated error rate can be made using the following well-known formula for the standard error of a sample proportion, p:

$$\sigma_p = \sqrt{\frac{N - n}{N - 1}\frac{PQ}{n}} \qquad (1)$$

where

N = the total number of cases processed at a particular stage during the last month,

n = the sample size,

P = the true error rate for that stage,

Q = 1 - P.

As indicated, the above equation applies to simple random sampling. Actually, systematic random sampling of cases with equal selection probabilities was used for the audit.[1] However, in this situation these two types of sampling procedures probably have about the same precision. For planning purposes the above formula should be adequate to estimate the standard error for an estimated error rate calculated from a systematic random sample.

Based on discussions with IACP personnel, it was agreed that a standard error of .02 for estimating a true rate of .10 and a standard error of .005 for estimating a true error rate of .01 would be adequate precision for the audit estimates. The sample size table that was used most often in the audit procedures, was based on this requirement.[2] This sample size table is Table 1.

3. Selection of the Samples of Cases

For sampling cases at Stages II - IV, the selection procedure was straightforward. The total number of cases processed (i.e., the group size) at each of these stages was usually easy to obtain since these cases were typically listed on cards or records in a file. From the group size, the required sample size was obtained from Table 1. The sample was then selected as a systematic random sample. The selection (or skip) interval used was obtained by dividing the group size (N) by the required sample size (n). (Tables of skip intervals and random digits to select random

*now at U.S. Bureau of the Census

starts were made available to simplify the selection procedures.)

The sampling of the telephone tapes (Stage I) was more complex than was the sampling at the other stages. Very few agencies have a record of the number of calls recorded on their tapes. Even when this is known, the number of these that are relevant to the audit (i.e., that involve at least some minimal crime) is not known.

Therefore, the first step in the sampling of the telephone tapes was to estimate the total number of relevant calls on the tapes for the month. This was done as a two-part procedure. First, the total number of calls in the month was approximated. Next, the ratio of relevant calls to total calls was estimated. From these two quantities an estimate of the total number of relevant cases, N, was calculated.[3] Reference to one of the sample size specification tables (i.e., Tables 1, 2, or 3, depending on the size of the department) provided the target sample size, n, for relevant telephone cases.

For a 30-day month, the number of hours, h, to be monitored was determined by multiplying the sampling rate, n/N, times the total number of hours in the month, 720. It was decided to monitor the tapes in terms of 15-minute segments throughout the month. Therefore, the total number of quarter-hour segments, q, to be monitored was calculated as four times the required number of hours (i.e., q = 4h).

The q segments to be monitored were selected systematically in two stages. First a sample of seven or eight days of the month was obtained by choosing every fourth day of the month, using a random start. (The selection interval of four was chosen to provide coverage of the different days of the week.) The number of 15-minute segments, S, in the days selected was then calculated (i.e., either 7 x 96 or 8 x 96). Finally, the segments to be monitored were selected systematically from the segments in the days chosen. The appropriate selection interval was, of course, S/q. An example of the selection of telephone tape segments is given below.

The calls monitored were all those that originated in any of the 15-minute segments selected for listening.[4] This procedure gave all calls on the tapes for the month an equal chance of selection (i.e., n/N).

The first 20 audits were carried out by IACP personnel with the cooperation and assistance of the police department personnel. It is intended that eventually the audits will be performed entirely by police department personnel. It may be difficult for them to carry out these selection procedures, especially those for the telephone tapes.

Example of the Selection of a Sample of Tape Segments

Estimated total number of calls on tape for February: 32,000

Estimated ratio of total calls to "meaningful" calls: 4:1

Therefore, $N \doteq (.25)(32,000) = 8,000$

Sample size from Table 1: $n = 250$

Sampling rate: $f = 250/8,000 = .01325$

Total number of hours in month: $(28)(24) = 672$

Number of hours to be sampled:
$h = (.01325)(672) = 21$

Number of quarter-hour segments to be sampled:
$q = 4(21) = 84$

Random start for the selection of days: 2

Select systematic random sample of every 4th day beginning with 2nd: 2nd, 6th, 10th, 14th, 18th, 22nd, 26th

Total number of segments in these days:
$S = (7)(96) = 672$

Selection interval for sampling time segments:
672/84 = 8, random start: 3

Obtain sample from time-interval table (Table 4)

Footnotes

[1]This method of selecting cases was chosen since it is a probability sampling procedure that is straightforward enough to eventually be carried out by police department personnel.

[2]Obtaining an adequate number of cases from the telephone tapes (Stage I) was so time consuming for smaller agencies that this precision requirement was relaxed somewhat for Stage I sampling in smaller agencies. The sample size tables used in such cases are Tables 2 and 3.

[3]In some cases the department personnel were not able to provide the estimates needed. In these instances IACP personnel listened to portions of the telephone tapes in order to make these estimates.

[4]In some departments it appeared that as the tape sampling progressed, the total number of meaningful calls selected in the sample segments would differ substantially from the target number. In such cases, the number of sample segments was either increased or decreased in an attempt to bring the sample size close to the target sample size.

Table 1 (.02 Standard Error - True Error Rate of .1 and a .005 Standard Error - True Error Rate of .01)

Group Size	Sample Size
1-60	all
61-80	50
81-120	70
121-200	90
201-500	120
501-1000	200
1,001-Over	250

Table 2 (.025 Standard Error - True Error Rate of .1)

Group Size	Sample Size
1-60	all
61-80	50
81-120	60
121-200	80
201-500	100
501-1000	125
1,001-Over	150

Table 3 (.03 Standard Error - True Error Rate of .1)

Group Size	Sample Size
1-60	all
61-80	40
81-120	50
121-200	60
201-500	80
501-1000	90
1,001-Over	100

Table 4 Stage I - Date/Time Segments

Example for Feburary

(Selection Interval = 8)

Time	2nd	6th	10th	14th	18th	22nd	26th
2400/0014			X				
0015/0029				X			
0030/0044					X		
0045/0059						X	
0100/0114							X
0115/0129							
0130/0144	X						
0145/0159		X					
0200/0214			X				
0215/0229				X			
0230/0244					X		
0245/0259						X	
0300/0314							X
0315/0329							
0330/0344	X						
0345/0359		X					
0400/0414			X				
0415/0429				X			
0430/0444					X		
0445/0459						X	
0500/0514							X
0515/0529							
0530/0544	X						
0545/0559		X					
0600/0614			X				
0615/0629				X			
0630/0644					X		
0645/0659						X	
0700/0714							X
0715/0729							
0730/0744	X						
0745/0759		X					
0800/0814			X				
0815/0829				X			
0830/0844					X		
0845/0859						X	
0900/0914							X
0915/0929							
0930/0944	X						
0945/0959		X					
1000/1014			X				
1015/1029				X			
1030/1044					X		
1045/1059						X	
1100/1114							X
1115/1129							
1130/1144	X						
1145/1159		X					

Time	2nd	6th	10th	14th	18th	22nd	26th
1200/1214			X				
1215/1229				X			
1230/1244					X		
1245/1259						X	
1300/1314							X
1315/1329							
1330/1344	X						
1345/1359		X					
1400/1414			X				
1415/1429				X			
1430/1444					X		
1445/1459						X	
1500/1514							X
1515/1529							
1530/1544	X						
1545/1559		X					
1600/1614			X				
1615/1629				X			
1630/1644					X		
1645/1659						X	
1700/1714							X
1715/1729							
1730/1744	X						
1745/1759		X					
1800/1814			X				
1815/1829				X			
1830/1844					X		
1845/1859						X	
1900/1914							X
1915/1929							
1930/1944	X						
1945/1959		X					
2000/2014			X				
2015/2029				X			
2030/2044					X		
2045/2059						X	
2100/2114							X
2115/2129							
2130/2144	X						
2145/2159		X					
2200/2214			X				
2215/2229				X			
2230/2244					X		
2245/2259						X	
2300/2314							X
2315/2329							
2330/2344	X						
2345/2359		X					