

A STRATIFIED SAMPLING PLAN FOR BILLING ACCURACY IN HEALTHCARE SYSTEMS

Jirachai Buddhakulsomsiri^a

Parthana Parthanadee^b

Swatantra Kachhal^a

^a *Department of Industrial and Manufacturing Systems Engineering, University of Michigan-Dearborn*

^b *Department of Agro-Industry Technology Management, Kasetsart University, Thailand*

Abstract

This study presents a stratified sampling plan for estimating accuracy of billing performance for the claims submitted to third party payers in healthcare systems. The population consists of hospital claims with amounts ranging from zero, hundreds, thousands, to rare high million dollars. Accuracy of the billing process is estimated by auditing a sample of claims with two measurements: the overall percent accuracy and the total dollar accuracy. Difficulties in constructing the sampling plan arise when the number of strata and their boundaries are unknown, and when the two measurements require different sampling schemes. The proposed sampling plan is designed to perform effectively for estimating both measurements. It determines an overall sample size and tests various numbers of strata to find an appropriate stratification. The optimal stratum boundary points are found using the rectangular stratification method on the claim dollar amounts. The overall sample is then assigned to strata with a mixed strategy between the proportional and optimal allocations and finally the accuracy estimates and their precisions are obtained. The sampling plan is tested on an actual population obtained from insurance industry with simulated claim errors. The results show effectiveness of the plan for both accuracy measurements.

Introduction

In healthcare systems an interesting problem commonly encountered is the estimation of the accuracy performance in processing of the hospital claims submitted to third party payers. Important characteristics of the claim population include highly positive skewness of the claim dollar amounts and a relatively low processing error rate. Two measures of accuracy performance include the population percent accuracy and total dollar accuracy. The percent accuracy is the percentage of the claims which are processed correctly in the population. The total dollar accuracy measures the dollar amount that should have been claimed.

Statistical sampling is a widely used tool for this estimation. The sampling plan usually implemented for the population with such characteristics is stratified

random sampling. In this sampling method, all claims in the population are divided into several subpopulations, called strata, from which a simple random sample is independently selected. The sampled claims are then audited so that the accuracy measures can be estimated. Stratification can usually provide a higher precision in the estimation, if the population can be efficiently divided into homogeneous strata. However, under the situation where the population strata structure is unclear, two critical questions that must be addressed are the number of strata and their boundary points. In addition, when there are two different measures to be estimated simultaneously as in this case, the sample must be appropriately assigned to strata so that both measures can be estimated with high precision.

This paper proposes a stratified sampling procedure that can be used to estimate both accuracy measures from the same sample. The steps in constructing the sampling plan, including determining the overall sample size, forming the strata, choosing the appropriate number of strata, and estimating the accuracy, are presented. An example of implementing the proposed plan to an actual hospital claim population is also provided.

Literature Reviews

Designing a stratified random sampling plan for accounting data involves a number of decisions. When the population strata structure is unknown, the auditor must first select the stratifying variable for constructing strata. Then, the number of strata and their boundaries must be determined. For the sample allocation, the Neyman or optimum allocation has been proven to minimize the variance of the estimate for a fixed total sample size and is often used in practice. However, it can only be approximated, since the stratum variances, which are required in the computation, are typically not available. The number of strata is usually decided by the precision gain from stratification as the number of strata increases and the cost of stratified sampling (Cochran, 1977).

Homogeneous strata can be constructed by setting the boundaries, so that a minimum variability within every stratum can be attained. The set of equations to determine the optimum boundaries, under Neyman allocation, for a given number of strata was derived by Dalenius (1957). However, the equations are difficult to compute, thus, a

number of approximate methods for strata boundary construction have been proposed by many researchers.

Four such methods are briefly described next. Mahalanobis (1952), and Hansen et al. (1953) propose a method to construct the strata boundaries by making the product of the stratum weight and the stratum mean (or the aggregate value) of the stratifying variable equal for all strata. Dalenius and Hodges (1959) presented that the approximate optimum boundaries can be obtained by constructing equal intervals on the cumulative of the square roots of the frequency distribution of the stratifying variable. Ekman (1959) constructed the strata boundaries by equalizing the product of stratum weight and stratum range. Sethi (1963) constructed the tables of the optimum boundaries for some standard continuous distributions, under Neyman, equal, and proportional allocations. If the distribution of the study population resembles with the standard distributions, the boundaries are obtainable from the tables.

Hess et al. (1966) studied four sample allocation methods and four boundary construction methods on data for medical hospitals, which is a highly-positive-skewed population. It was found that, with optimum stratification, the equal allocation and the Neyman allocation perform the best in gaining variance reduction, while among the methods for constructing boundaries, Ekman's method, Sethi's method with adjustment, and Dalenius's and Hodges's methods with adjustment were found to perform the best.

The empirical study of sampling on accounting populations was explored by Neter and Loebbecke (1975). Four accounting populations were used to generate several study populations with various error rates. The behaviors of estimators on those generated populations under various sampling plans were reported. The study showed that satisfactory results were achieved by using the stratified sampling on the accounting data.

Proposed Stratified Sampling Plan

The study populations of interest consist of hospital claims with amounts ranging from zero, hundreds, thousands, to rare high million dollars. The summary of statistics of one of the population is provided in Table 1. The statistics shows that the population is highly positive-skewed, with a huge standard deviation. This characteristic is common for the population of accounting data (Neter and Loebbecke 1975). The objective of the sampling is to collect evidential information to fairly assess the accuracy of claim processing operations on a quarterly basis.

In this study, an error is defined as the difference between the processed amount and the audit amount of a claim. An overpaid claim implies positive error, whereas an underpaid claim implies negative error. From the past estimate of accuracy, it was found that the populations

have very low error rates (i.e. close to 2%), and there is no strong evidence that the error rate for high dollar claims is substantially different than that of low dollar claims. Based on this information and expert opinion, the following assumptions are made: (1) The error rate is statistically the same throughout the population, regardless of the claim amounts, (2) There is no significant correlation between the claim amount and the error amount, and (3) An overpaid error amount cannot exceed its processed claim amount, whereas an underpaid error amount may.

Table 1: Frequency distribution and summary statistics of the claim amount population in one quarter (3 months)

Claim amount (\$)	Number of claims
0	2,259,067
0.01 – 1,000	6,732,691
1,000.01 – 10,000	140,829
10,000.01 – 100,000	8,103
more than 100,000	111
Total	9,138,801
Total claim amount	\$ 806,400,496
Mean*	\$ 163.72
Standard deviation*	\$ 1,209.73
Skewness*	\$ 115.04
Maximum	\$ 672,796.59
Minimum	\$ 0

* These statistics are calculated only on non-zero dollar claims.

Two measurements of interest including the overall percent accuracy and the total dollar accuracy of the claims, are to be estimated, using the proposed stratified sampling plan. The purpose of using stratification is to gain more precision in estimating the total dollar accuracy measurement. Stratification can reduce the effect of skewness in claim amount population which will lead to reduction in the overall standard error of the estimates. Although the main reason for using stratified sampling is to better estimate the total dollar accuracy, the sampling plan will be used for both measurements. The procedure is described in the following subsections.

The variable chosen for stratification is the processed amount of hospital claims. There are two reasons justified for this: (1) the data are easy to obtain, and (2) the data can be used directly in calculating the measurements for assessing the accuracy of claim processing operations.

Determine the Overall Sample Size, n

The overall sample size consists of three sample components, which are the sample sizes for (1) zero dollar stratum, (2) non-zero dollar strata, and (3) rare high dollar stratum that will be 100% audited. The determinations of sample sizes are as follows. Table 2 lists the notation used in the calculation.

(1) Determine n_a : When processed claims are classified into two classes – correctly processed claims and processed claims with errors, the sampling procedure is called “attribute sampling”. The sample size for attribute sampling n_a can be calculated using Equations (1) and (2) as follows:

Table 2: Notation for the determination of sample sizes

$n_a, n_b,$	sample size for estimating population percent accuracy, total dollar accuracy, and 100% auditing for rare high dollar stratum
n_c	the population size
N	the initial sample size calculated without the finite population correction (fpc) factor
P	the expected percent accuracy (from past quarter performance)
d	the desired precision level
$Z_{\alpha/2}$	the standard normal variate associated with the level of confidence α
SD	the advance estimate of the standard deviation of the population
P_0	proportion of zero dollar claims in the population

$$n_{a0} = \left(\frac{Z_{\alpha/2}}{d} \right)^2 P(1-P) \quad (1)$$

$$n_a = \frac{n_{a0}}{1 + \frac{n_{a0} - 1}{N}} \quad (2)$$

In a situation where the population size is very large, n_{a0} can be used as an approximated sample size.

(2) Determine n_b : The sample size for total dollar accuracy is based on the use of interval estimation, as shown in Equation (3).

$$n_b = \left(\frac{NZ_{\alpha/2}SD}{d} \right)^2 \quad (3)$$

The desired precision parameter d in Equation (3) is defined to be an acceptable amount of dollar error that the auditor is willing to accept in the estimation process. The advance estimate of the population standard deviation, SD , can be obtained from the error and the claim audit amounts from past samples using the difference estimation method.

(3) Determine n_c : The sample size n_c is obtained by setting a cut-off point for rare high dollar stratum, which contains all claims with amounts higher than the cut-off point. In general, the auditor may choose the cut-off point depending on the amount of effort or resource allocated to the rare high dollar stratum. In this study, \$100,000 was used as suggested by experts in the field.

After the three sample sizes are determined, the overall sample size would consist of (1) the portion of n_a

that is assigned to the zero dollar stratum using proportional allocation, (2) the maximum between n_b and the remaining of n_a , and (3) n_c ; see Equation (4).

$$n = P_0 n_a + \max[n_b, (1 - P_0)n_a] + n_c \quad (4)$$

It is important to emphasize that in some situations where the desired precision is high, the calculated sample size using this procedure may require resources far more than what are available to the auditor. An alternative and more practical way is to use the auditor’s judgment to determine the overall sample size based on the amount of resources (e.g., time and man-hour) that are available. Nevertheless, it is strongly suggested that the overall sample size is at least $n_a + n_c$, so that at least the percent accuracy is estimated with the desired precision.

Design the Sampling Plan

Designing the sampling plan involves two major steps: forming the strata and allocating the overall sample. Two critical and interrelated decisions to be made in forming the strata are (1) the number of strata, and (2) the locations of the boundary points between strata.

(1) Number of Strata: Research studies in stratification suggest that the precision of the estimate is ordinarily not improved significantly by using beyond 20 (Arens and Loebbecke, 1981). Significant gains in the precision usually are obtained from the first few strata; hence, forming only a few strata (perhaps 5 to 10 strata) will typically yield most of the possible gains from stratification.

(2) Stratum Boundary Points: To determine the boundary points the approximate rectangular method¹, which implements the equal cumulative $\sqrt{f(y)}$ rule (i.e. equal cumulative square root of frequency), is used. The reason for choosing this rule is that it can be easily implemented and has been shown to work well on the population with strong positive skewness (Hess et al., 1966, and Neter and Loebbecke, 1975). The procedure is as follows.

Step 1: Arbitrarily choose a number of intervals L (e.g. 100, 200) The larger the L , the finer the scale for the boundary points would be.

Step 2: Set up L intervals of claim amounts, each with an associated interval width ω_i , where i denotes the interval index. For convenient purpose, the cut-off point for the rare high dollar stratum may be used in this step by setting each interval width to be equal to the cut-off point divided by L . Note that ω_i need not be of equal size.

Step 3: Count the number of claims in each interval N_i .

Step 4: Calculate the frequency, $f(y) = \omega_i \times N_i$, the square root of the frequency, $\sqrt{f(y)}$, and the cumulative of the square root of the frequency for each interval.

¹ First proposed by Dalenius and Hodges (1959), later modified by Cochran (1963), and tested by Hess et al. (1966)

Step 5: Determine the total value of the cumulative $\sqrt{f(y)}$.

Step 6: Divide the total cumulative $\sqrt{f(y)}$ by the desired number of strata H . The stratum boundary points are the boundary points of the intervals that are approximately equal in width on the cumulative $\sqrt{f(y)}$ scale.

The proposed allocation method is a mixed strategy between two common sampling allocations: proportional allocation and optimal allocation. In general, proportional allocation should be used when different parts of the population are proportionally represented in the sample. It is therefore appropriate for estimating the percent accuracy since it is assumed that the percent accuracy is approximately the same throughout the population. On the contrary, optimal allocation should be used when the variability of the measurement varies significantly across all strata, which is the case for dollar accuracy. Optimal allocation assigns different sample sizes to strata proportional to the strata variability. That is, the strata with larger claim amounts have more variability than the smaller ones; thus, to increase the overall precision of the estimate, the sampling fractions in those strata should be increased.

For each stratum h , the stratum standard deviation S_h and the number of claims in the stratum N_h are determined. The sample size for total dollar accuracy is then optimally allocated using Equation (5) if the remaining sample units equals n_b , or Equation (6) otherwise.

$$n_h = \frac{N_h S_h}{\sum_{h=1}^H N_h S_h} \times n_b \quad (5)$$

$$n_h = \frac{N_h S_h}{\sum_{h=1}^H N_h S_h} \times (1 - P_0) n_a \quad (6)$$

Calculate the Estimates of the Accuracy Performance and Their Precisions

Table 3 presents the notation used in the estimation formula. The population percent accuracy \hat{P} is estimated as follows. First, \hat{V} can be estimated using Equation (7),

$$\hat{V} = \sum_{h=0}^{H+1} \left(\sum_{i=1}^{n_b} v_{hi} \right) \quad (7)$$

\hat{P} can be estimated using the traditional method as in Equations (8), or using Agresti's method (Agresti and Coull, 1998) as in Equation (9), which adds some adjustment to the traditional estimation method. It is important to note that in this calculation the stratified sample is treated as if it were a simple random because

of the assumption that the error rate is constant throughout the population.

$$\hat{P} = \frac{\hat{V}}{n} \quad (8)$$

$$\hat{P} = \frac{\hat{V} + Z_{\alpha/2}^2/2}{n + Z_{\alpha/2}^2} \quad (9)$$

The $(1-\alpha) \times 100\%$ confidence interval of \hat{P} can be calculated using Equation (10),

Table 3: Notation for the estimation calculation

h	stratum index; $h = 0, 1, 2, \dots, H, H+1$, where 0 denotes zero dollar stratum, 1 to H denote the non-zero dollar strata, and $H+1$ denotes the rare high dollar stratum with 100% audit rate
x_{hi}	the processed amount of the i^{th} claim in the sample taken from stratum h
\bar{x}_h	the mean processed amount of claims in the sample taken from stratum h
y_{hi}	the audit amount of the i^{th} claim in the sample taken from stratum h
\bar{y}_h	the mean audit amount of claims in the sample taken from stratum h
s_h^2	the variance of the audit amount of claims in the sample taken from stratum h
e_{hi}	the error amount of the i^{th} claim in the sample taken from stratum h
v_{hi}	the binary variable indicating whether the i^{th} claim in the sample taken from stratum h is processed correctly
N_h	the size of stratum h
X_{hi}	the processed amount of the i^{th} claim in stratum h
X_h	the total processed amount of all claims in stratum h
X	the total processed amount of all claims in the population
\hat{V}	the total number of correct claims in all strata combined
\hat{P}	the estimate of the percent accuracy for the whole population
\hat{Y}	the estimate of the total dollar accuracy for the whole population
$s^2(\hat{Y})$	the estimate of the variance of \hat{Y}

$$\hat{P} \pm Z_{\alpha/2} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \quad (10)$$

2) The population total dollar accuracy is estimated using the difference estimation method. First, X is calculated using Equations (11),

$$X_h = \sum_{i=1}^{N_h} X_{hi} \quad (11)$$

\hat{Y} and $s^2(\hat{Y})$ can be estimated using Equations (12) and (13), respectively,

$$\hat{Y} = \sum_{h=1}^{H+1} [X_h + N_h(\bar{y}_h - \bar{x}_h)] \quad (12)$$

$$s^2(\hat{Y}) = \sum_{h=1}^{H+1} \left\{ N_h^2 \left(1 - \frac{n_h}{N_h} \right) \frac{1}{n_h(n_h - 1)} \left[\sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2 + \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 - 2 \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)(y_{hi} - \bar{y}_h) \right] \right\} \quad (13)$$

Finally, the $(1-\alpha) \times 100\%$ confidence interval of \hat{Y} can be calculated using Equation (14),

$$\hat{Y} \pm Z_{\alpha/2} \sqrt{s^2(\hat{Y})} \quad (14)$$

Computational Results

The Population

To test the proposed stratified sampling plan, an actual population of hospital claims is used. Since the true percent accuracy and the total dollar accuracy of the population are unknown, simulated errors were used to establish the target audit values for both measurements. Table 4 summarizes the population target values.

Table 4: Total audit values of the population

Percent error rate	3%
Total processed amount	\$ 806,400,496
Total simulated overpaid	\$ 11,525,104
Total simulated underpaid	\$ (3,231,768)
Total target audit amount	\$ 798,107,160

Table 5: Estimate of performance accuracy for the population with 3% target error rate

Plan	# of strata	Percent accuracy measure			Total dollar accuracy measure	
		Traditional	Agresti's	Precision	Relative Standard Error	% Off-Target
SRS	1	3.05%	3.35%	1.58%	0.74%	0.03%
Stratified sampling	3	2.62%	3.00%	1.50%	0.66%	0.16%
	4	2.77%	3.14%	1.53%	0.66%	0.05%
	5	2.66%	3.00%	1.50%	0.60%	-0.01%
	6	2.87%	3.23%	1.55%	0.61%	-0.03%
	7	2.72%	3.06%	1.51%	0.54%	0.01%
	8	2.73%	3.13%	1.53%	0.51%	0.10%
	9	2.83%	3.22%	1.55%	0.58%	-0.02%
	10	2.68%	3.06%	1.52%	0.54%	0.04%

The Tested Sampling Plans

For the plan testing purpose, the overall sample size was chosen arbitrarily at 500, according to the auditor's current practice at the insurance company that provided the population data. The two types of plan tested include simple random sampling (SRS) plan, and the proposed plan with non-zero dollar strata ranging from 1 to 20.

Overall Percent Accuracy

The test results are showed in Table 5, with all calculations done at 95% confidence level. The estimate of percent error is obtained using both traditional method and Agresti's Method. Note that the calculations for the precision of the percent error are based on Agresti's estimate. The bases for plan comparisons are the percent error estimates and their precision. Each result is an average of 100 estimates from 100 samples. Only the results from the proposed stratified sampling plans with 3 to 10 strata (i.e. 1 to 8 strata for non-zero dollar claims) are shown in the table. From the results, it was found that all plans perform statistically the same in terms of the estimate and its precision. The traditional method slightly underestimates the percent error, whereas Agresti's method slightly overestimates it. To be conservative in the estimation, the Agresti's method is therefore recommended.

Total Dollar Accuracy

The bases for plan comparison for total dollar accuracy include (1) the average percent off-target, calculated from the differences between the processed claim amounts and the estimated audit amounts from the samples, and (2) the relative standard error, which is the ratio of "the standard error of the estimate" to "the total audit value" for the population. From Table 5, all plans perform statistically the same with respect to the average percent off-target. However, it is clear that the stratified sampling plans perform better than the SRS plan in term of the relative standard error. Using the reduction in the relative standard error when the number of strata for the non-zero dollar strata

is ranging from 1 to 20, it was found that the appropriate number of strata for non-zero dollar claims (\$0.01 – \$100,000) should be approximately 8 (i.e. the total number of strata is 10). That is, adding more strata did not yield much more improvement in the precision. Finally, Table 6 presents the strata boundaries (from rectangular method) and the allocation of the sample size of 500 (from optimal allocation) for the stratified sampling plan with 10 strata.

Table 6: The ten-stratum sampling plan with $n = 500$

Stratum	Boundary	Stratum Sample Size
1	\$0	31
2	(\$0, \$40]	32
3	(\$40, \$110]	47
4	(\$110, \$250]	34
5	(\$250, \$650]	39
6	(\$650, \$1,570]	39
7	(\$1,570, \$3,960]	39
8	(\$3,960, \$10,430]	39
9	(\$10,430, \$100,000]	89
10	[\$100,000, ∞)	111

Conclusion

The problem of estimating the accuracy performance for the population of hospital claims is presented in this paper. When the population strata structure is unknown, the rectangular stratification method is applied on the claim amounts to determine the optimal boundary points between strata, and the precision gain in the estimation process (i.e. reduction in relative standard error) is used to identify the appropriate number of strata. The proposed sampling plan also implements a mixed strategy between proportional and optimal allocations so that both percent accuracy and total dollar accuracy can be estimated simultaneously. The plan is tested with an actual population obtained from insurance industry. The test results show effectiveness of the proposed plan in estimating both accuracy measures.

References

- Agresti, A. and Coull, B. A. (1998). “Approximate Is Better than Exact for Interval Estimation of Binomial Proportions”, *The American Statistician*, **52** (2), 119-126.
- Arens, A. A., and Loebbecke, J. K. (1981). *Applications of Statistical Sampling to Auditing*, Prentice Hall, Inc., New Jersey.
- Cochran, W. G. (1977). *Sampling Techniques*, 3rd edition. John Wiley and Sons, Inc. New York.

- Dalenius, T. (1957). *Sampling in Sweden*. Contributions to the Methods and Theories of Sample Survey Practice. Almqvist and Wiksell, Stockholm.
- Dalenius, T. and Hodges, J. L., Jr. (1959). “Minimum Variance Stratification”, *Journal of the American Statistical Association*, **54** (285), 88-101.
- Ekman, G. (1959). “An Approximation Useful in Univariate Stratification”, *The Analysis of Mathematical Statistics*, **30**, 219-229.
- Hansen, M. L., Hurwitz, W. N., and Madow, W. G. (1953). *Sample Survey Methods and Theory, Volume 1*. John Wiley and Sons, Inc. New York.
- Hess, I., Sethi, V. K., and Balakrishnan, T. R. (1966). “Stratification: A Practical Investigation”, *Journal of the American Statistical Association*, **61** (313), 74-90.
- Mahalanobis, P.C. (1952). “Some Aspects of the Design of Sample Surveys”, *Sankhyā*, **12**, 1-7.
- Neter, J., and Loebbecke, J. K. (1975). *Behavior of Major Statistical Estimators in Sampling Accounting Populations: An Empirical Study*. American Institute of Certified Public Accountants, Inc. New York.
- Sethi, V. K. (1963). “A Note on Optimum Stratification of Populations for Estimating the Population Means”, *The Australian Journal of Statistics*, **5**, 20-33.

Biographical Sketch

Jirachai Buddhakulsomsiri is an Assistant Professor in the Department of Industrial and Manufacturing Systems Engineering at the University of Michigan-Dearborn. He earned his Ph.D. degree in Industrial Engineering from Oregon State University in 2003. His research areas of interest include applied statistics, data mining, and project scheduling and management.

Parthana Parthanadee received her Ph.D. degree (2004) from the Department of Industrial and Manufacturing Engineering at Oregon State University, USA. Her research interests include applied operations research, vehicle routing, scheduling, and production planning. She is currently a faculty member in the Department of Agro-Industry Technology Management at Kasetsart University, Bangkok, Thailand.

Swatantra Kachhal, Ph.D., is a Professor and Chair of the Department of Industrial and Manufacturing Systems Engineering at the University of Michigan-Dearborn. He is a Fellow of the Institute of Industrial Engineers and a Fellow of the Healthcare Information and Management Systems Society. Dr. Kachhal is a past President of the Society for Health Systems and is currently serving as Technical Vice President for Society and Divisions for the Institute of Industrial Engineers.