A Stratified Sampling Plan for Billing Accuracy in Healthcare Systems

Jirachai Buddhakulsomsiri
Parthana Parthanadee
Swatantra Kachhal

Department of Industrial and Manufacturing Systems Engineering
The University of Michigan-Dearborn

18th Annual Society for Health Systems Conference
February 2006 | San Diego, CA
Presentation Outline

- Background: Billing accuracy in healthcare systems
- Objectives of the study
- Characteristics of the populations
- Characteristics of the sampling units and claim errors.
- The proposed procedure in designing a sampling plan
- Computational testing, results, and discussion
- Conclusions
Billing Accuracy in Healthcare Systems

- Estimation of the accuracy performance in processing of the hospital claims submitted to third party payers.

- Two measures of accuracy performance:
  - The population percent accuracy
  - The population total dollar accuracy
Objectives of the Study

To the auditor...

- Samples provide evidential information to fairly assess the accuracy of claim processing operations on a quarterly basis.
- A sampling plan provides guideline to how samples are collected and how accuracy performance measures are estimated.
Objectives of the Study (Cont.)

- Design a stratified sampling procedure that can be used to estimate both accuracy measures from the same sample.
- Test the proposed sampling plan on actual population with generated errors.
Characteristics of the Populations

- A population consists of millions of hospital claims.
- Claim amounts are ranging from zero, hundreds, thousands, to rare high million dollars.
- Population is highly positive-skewed, with a very large standard deviation.
- Population error rate is low: 1% – 5%.
# Frequency Distribution and Summary Statistics of the Claim Amount Population in One Quarter

<table>
<thead>
<tr>
<th>Claim amount ($)</th>
<th>Number of claims</th>
<th>Total claim amount</th>
<th>Mean*</th>
<th>Standard deviation*</th>
<th>Skewness*</th>
<th>Maximum</th>
<th>Minimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2,259,067</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$ 0</td>
</tr>
<tr>
<td>0.01 – 1,000</td>
<td>6,732,691</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,000.01 – 10,000</td>
<td>140,829</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10,000.01 – 100,000</td>
<td>8,103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>more than 100,000</td>
<td>111</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9,138,801</td>
<td>$ 806,400,496</td>
<td></td>
<td>$ 163.72</td>
<td>$ 1,209.73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* These statistics are calculated only from non-zero dollar claims.
Characteristics of the Sampling Units and Claim Errors

- A sampling unit is simply one hospital claim.
- A claim is either correctly processed or processed with an error.
- 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.
Characteristics of the Sampling Units
and Claim Errors

Example of claim with error:

- A hospital claim is first submitted and processed for $1,000 (denote as “claim amount”, X).
- Later, this claim is sampled and audited that the correct amount should be $800 (denote as “audit amount”, Y).
- Thus, the error amount (E) is $1,000 - $800 = $200 overpaid, or $E = X - Y$. 
Key Assumptions

Based on expert knowledge and preliminary statistical test:

- The error rate is statistically the same throughout the population, regardless of the claim amounts.
- There is no significant correlation between the claim amount and the error amount.
- An overpaid error amount cannot exceed its processed claim amount, whereas an underpaid error amount has no restriction.
The proposed sampling plan is “stratified random sampling.”

Stratification provides precision gain in estimating the total dollar accuracy measurement.

Stratification reduces the effect of skewness in claim amount population, which 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.
Choosing Stratifying Variable

- Simply use the processed amount of hospital claims.
  - The data are easy to obtain.
  - The data can be used directly in calculating the accuracy measures.
The Overall Sample Size

- The overall sample size ($n$) consists of three components:
  - Sample size for zero dollar stratum ($n_a$)
  - Sample size for non-zero dollar strata ($n_b$)
  - Sample size for rare high dollar stratum ($n_c$) that will be 100% audited.
Sample Size for Zero Dollar Stratum

\( n_a \) is a function of:

- The expected percent accuracy (or error rate), \( P \).
- Level of confidence of the estimation process, \( \alpha \).
- Desired precision level, \( d \).

\[
\begin{align*}
   n_{a0} &= \left(\frac{Z_{\alpha/2}}{d}\right)^2 P(1-P) \\
   \Rightarrow \quad n_a &= \frac{n_{a0}}{1 + \frac{n_{a0} - 1}{N}}
\end{align*}
\]
Sample Size for Non-Zero Dollar Strata

\( n_b \) is a function of:

- Population size, \( N \).
- Level of confidence of the estimation process, \( \alpha \).
- Desired precision level, \( d \).
- Advance estimate of the population standard deviation, \( SD \).

\[
 n_b = \left( \frac{NZ_{\alpha/2}SD}{d} \right)^2
\]
Sample Size for Rare High Dollar Stratum

$n_c$ is obtained by setting a cut-off point for rare high dollar stratum

- This stratum contains all claims with amounts higher than the cut-off point.
- Cut-off point is chosen by the auditor.
- Trade-off between resource allocated to this stratum and total dollar accuracy.
- In this study, $100,000 was used as suggested by experts in the field.
Putting It All Together

- The portion of $n_a$ that is proportionally allocated to the zero dollar stratum
- The maximum between $n_b$ and the remaining of $n_a$
- $n_c$

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

- $P_0$ is the proportion of the zero dollar claims in the population.
Or Else...

- $n$ is too large.
- $n$ is the amount of resource available to the auditor.
- $n$ is recommend to be at least, $n_a + n_c$, so that one of the accuracy measure is estimated with the desired precision.
Design The Sampling Plan

Two major steps:

- **Forming the strata** – two decisions:
  - Number of strata, based on precision gain in each stratum added.
  - Strata boundary points, based on the Rectangular method.

- **Allocating the overall sample**
Forming the Strata

For the number of strata:

- Significant gains in the precision usually are obtained from the first few strata.
- Increasing the number of strata beyond 20 may not improve the precision of the estimation that much.
- In our study, the number of strata for non-zero dollar claims is increased one-by-one, and the gain in precision is assessed.
On strata boundary points:

- The Rectangular method is used.
- This method implements the equal cumulative $\sqrt{f(y)}$ rule (i.e. equal cumulative square root of frequency).
- The method is very easy to implement.
Determine The Boundary Points

Step 1: Arbitrarily choose a number of intervals \( L \) (The larger, the finer).

Step 2: Set up \( L \) intervals, each with an associated interval width, \( W_i \).

Step 3: Count the number of claims for each interval \( N_i \).

Step 4: Calculate the following.

- The frequency \( f(y) = W_i \times N_i \)
- The square root of the frequency \( \sqrt{f(y)} \)
- The cumulative of the square root of the frequency for each interval
Determine The Boundary Points (Cont.)

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 – appropriate for estimating the percent accuracy
- Optimal allocation – appropriate for estimating total dollar accuracy

Optimal allocation assigns different sample sizes to strata proportional to the strata variability.

The overall sample is allocated to the strata with larger claim amounts.
Optimal Allocation

For each stratum $h$, the sample size $n_h$ is a function of:

- Stratum size $N_h$
- Standard deviation of the stratum $S_h$
- Sample size for non-zero dollar strata $n_b$

$$n_h = \frac{\sum_{h=1}^{H} N_h S_h}{\sum_{h=1}^{H} N_h S_h} \times n_b$$
Example

- Suppose the overall sample size is arbitrarily chosen to be \( n = 500 \).
- Also, suppose the proportion of zero dollar claims in the population \((P_0)\) is estimated to be 0.2490.
- Then, for zero dollar stratum:
  - The sample size \((n_a)\) required to achieve 95% confidence and 3% precision is 122.
  - Proportional allocation is used to allocate \((n_aP_0)\) to the zero dollar stratum is \((122)(0.2490) = 31\).
Example (Cont.)

- For rare high dollar stratum:
  - A cut-off point is set at $100,000.
  - The 100% audit sample size $n_c = 111$.
- The samples left for non-zero strata is:
  \[500 - 31 - 111 = 358.\]
Example (Cont.)

- Suppose the desired number of non-zero strata is 8.
  To determine the strata boundaries, we use the
  "Rectangular" method.
  - The range $100,000 for non-zero dollars strata is
    divided into k (say, 10,000) equal intervals.
  - Construct the table as shown on the next slide.

- Then, allocate the rest of the sample (358) to each
  stratum using optimal allocation.
Example (Cont.)

<table>
<thead>
<tr>
<th>i</th>
<th>Interval of claim amount</th>
<th>W_i</th>
<th>N_i</th>
<th>f(y) = W_i × N_i</th>
<th>√f(y) = √W_i × N_i</th>
<th>Cumulative √f(y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01-10</td>
<td>10</td>
<td>837,095</td>
<td>8,370,950</td>
<td>2893.26</td>
<td>2893.26</td>
</tr>
<tr>
<td>2</td>
<td>10.01-20</td>
<td>10</td>
<td>970,557</td>
<td>9,705,570</td>
<td>3115.38</td>
<td>6,008.64</td>
</tr>
<tr>
<td>3</td>
<td>20.01-30</td>
<td>10</td>
<td>537,245</td>
<td>5,372,450</td>
<td>2317.85</td>
<td>8,326.49</td>
</tr>
<tr>
<td>4</td>
<td>30.01-40</td>
<td>10</td>
<td>571,766</td>
<td>5,717,660</td>
<td>2391.16</td>
<td>10,717.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10,000</td>
<td>99,000.01-100,000</td>
<td>10</td>
<td>2</td>
<td>20</td>
<td>4.47</td>
<td>91,146.04</td>
</tr>
</tbody>
</table>
## Boundary Points

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0</td>
</tr>
<tr>
<td>2</td>
<td>($0, $40]</td>
</tr>
<tr>
<td>3</td>
<td>($40, $110]</td>
</tr>
<tr>
<td>4</td>
<td>($110, $250]</td>
</tr>
<tr>
<td>5</td>
<td>($250, $650]</td>
</tr>
<tr>
<td>6</td>
<td>($650, $1,570]</td>
</tr>
<tr>
<td>7</td>
<td>($1,570, $3,960]</td>
</tr>
<tr>
<td>8</td>
<td>($3,960, $10,430]</td>
</tr>
<tr>
<td>9</td>
<td>($10,430, $100,000)</td>
</tr>
<tr>
<td>10</td>
<td>[$100,000, and above)</td>
</tr>
</tbody>
</table>
### 10-Stratum Sampling Plan with $n = 500$

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Boundary</th>
<th>Stratum Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0$</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>($0, $40]$</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>($40, $110]$</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>($110, $250]$</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>($250, $650]$</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>($650, $1,570]$</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>($1,570, $3,960]$</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>($3,960, $10,430]$</td>
<td>39</td>
</tr>
<tr>
<td>9</td>
<td>($10,430, $100,000)</td>
<td>89</td>
</tr>
<tr>
<td>10</td>
<td>[$100,000, and above)</td>
<td>111</td>
</tr>
</tbody>
</table>
Calculate the Estimates of the Accuracy Measures and Their Precisions

The population percent accuracy (P) is estimated by:

- **Method 1: Traditional method**  
  \[ \hat{P} = \frac{\hat{V}}{n} \]

  - Count the number of correct claims (V) in the sample
  - Then, divide by the overall sample size (n).

- **Method 2: Agresti’s method**  
  \[ \hat{P} = \frac{\hat{V} + Z_{\alpha/2}^2 / 2}{n + Z_{\alpha/2}^2} \]

  - Simply add 2 to the numerator and 4 to the denominator for 95% confidence level.
Calculate the Estimates of the Accuracy Measures and Their Precisions (Cont.)

Precision of the percent accuracy is expressed as,

- A 95% confidence interval (CI),

\[
\hat{P} \pm Z_{\alpha/2} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}
\]

- or, the half-width of the CI

\[
Z_{\alpha/2} \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}
\]
Calculate the Estimates of the Accuracy Measures and Their Precisions (Cont.)

The population total dollar accuracy is estimated by:

- Compute the total claim amount in each stratum $h$

\[ X_h = \sum_{i=1}^{N_h} X_{hi} \]

- Compute the mean claim amount $\overline{x}_h$ and the mean audit amount $\overline{y}_h$ of the sample in each stratum $h$
Calculate the Estimates of the Accuracy Measures and Their Precisions (Cont.)

- The total dollar accuracy of the population is the weighted average of the stratum means.

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

- The precision is expressed as \( s^2 (\hat{Y}) \)

\[ s^2 (\hat{Y}) = \sum_{h=1}^{H+1} \left\{ N_h^2 \left( 1 - \frac{n_h}{N_h} \right) \frac{1}{n_h (n_h - 1)} \left[ \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2 + \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2 - 2 \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)(y_{hi} - \bar{y}_h) \right] \right\} \]
The proposed sampling plan is tested on an actual population of hospital claims. The true accuracy performance of the population are unknown – thus, simulated errors were used to establish the target values for both measurements. Errors are simulated according to probability distributions (parametric and empirical) that are fit from past claim sample data.
Simulating Claim Errors

- Specify the error rate, say, 3%.
- Each actual claim is randomly assigned an error amount, where 97% of the times, error amounts are zeros.
- For the claims with errors, approximately 70% are overpaid error (positive error), and 30% are underpaid.
- Given an overpaid (or underpaid) error, the error amount is generated as a “percentage” of the claim amount.
- Different probability distributions are used depending on:
  - Overpaid or underpaid
  - Claims amount
### Total audit values of the population

<table>
<thead>
<tr>
<th>Description</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent error rate</td>
<td>3%</td>
</tr>
<tr>
<td>Total processed amount</td>
<td>$806,400,496</td>
</tr>
<tr>
<td>Total simulated overpaid</td>
<td>$11,525,104</td>
</tr>
<tr>
<td>Total simulated underpaid</td>
<td>$(3,231,768)</td>
</tr>
<tr>
<td>Total target audit amount</td>
<td>$798,107,160</td>
</tr>
</tbody>
</table>
The overall sample size was chosen arbitrarily at 500.

The two types of plan tested include:

- Simple random sampling (SRS) plan
- The proposed plan with non-zero dollar strata ranging from 1 to 20.

100 random samples are drawn for each plan.

Estimation of the accuracy measures are performed at the 95% confidence level.
# Estimate of Percent Accuracy for the Population with 3% Error Rate

<table>
<thead>
<tr>
<th>Plan</th>
<th># of strata</th>
<th>Traditional Estimate</th>
<th>Agresti’s Estimate</th>
<th>Precision</th>
<th>% Off-Target</th>
<th>Relative Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRS</td>
<td>1</td>
<td>3.05%</td>
<td>3.35%</td>
<td>1.58%</td>
<td>0.03%</td>
<td>0.74%</td>
</tr>
<tr>
<td>Stratified sampling</td>
<td>3</td>
<td>2.62%</td>
<td>3.00%</td>
<td>1.50%</td>
<td>0.16%</td>
<td>0.66%</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.77%</td>
<td>3.14%</td>
<td>1.53%</td>
<td>0.05%</td>
<td>0.66%</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2.66%</td>
<td>3.00%</td>
<td>1.50%</td>
<td>-0.01%</td>
<td>0.60%</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2.87%</td>
<td>3.23%</td>
<td>1.55%</td>
<td>-0.03%</td>
<td>0.61%</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2.72%</td>
<td>3.06%</td>
<td>1.51%</td>
<td>0.01%</td>
<td>0.54%</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>2.73%</td>
<td>3.13%</td>
<td>1.53%</td>
<td>0.10%</td>
<td>0.51%</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>2.83%</td>
<td>3.22%</td>
<td>1.55%</td>
<td>-0.02%</td>
<td>0.58%</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2.68%</td>
<td>3.06%</td>
<td>1.52%</td>
<td>0.04%</td>
<td>0.54%</td>
</tr>
</tbody>
</table>
Results and Discussion

On population percent accuracy:

- The target value of error rate is 3%.
- All plans perform statistically the same in terms of estimating the percent error and its precision.
- The traditional estimation method slightly underestimates the percent error, while Agresti’s method slightly overestimates it.
- Agresti’s method is therefore recommended.
Results and Discussion (Cont.)

On population total dollar accuracy:

- Plan comparisons are based on:
  - Average percent off-target – calculated from the differences between the processed claim amounts and the estimated audit amounts.
  - Relative standard error – the ratio of “the standard error of the estimate” to “the total audit value” for the population.
Results and Discussion (Cont.)

- All plans perform statistically the same with respect to the average percent off-target.
- The stratified sampling plans perform better than the SRS plan in terms of the relative standard error.
- Using the reduction in the relative standard error, it was found that the appropriate number of strata for non-zero dollar claims should be approximately 8 (i.e. total number of strata is 10).
# 10-Stratum Sampling Plan with $n = 500$

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Boundary</th>
<th>Stratum Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$0$</td>
<td>31</td>
</tr>
<tr>
<td>2</td>
<td>$(0, 40]$</td>
<td>32</td>
</tr>
<tr>
<td>3</td>
<td>$(40, 110]$</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>$(110, 250]$</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>$(250, 650]$</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>$(650, 1,570]$</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>$(1,570, 3,960]$</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>$(3,960, 10,430]$</td>
<td>39</td>
</tr>
<tr>
<td>9</td>
<td>$(10,430, 100,000]$</td>
<td>89</td>
</tr>
<tr>
<td>10</td>
<td>[$100,000, and above)</td>
<td>111</td>
</tr>
</tbody>
</table>
Conclusion

- The problem of estimating the accuracy performance for the population of hospital claims is presented.

- Stratified sampling – for situations where the population strata structure is unknown:
  - Stratifying variable is the claim amount.
  - The rectangular stratification method determines the optimal boundary points between strata.
  - The precision gain in the estimation process identifies the appropriate number of strata.
The proposed sampling plan implements a mixed strategy between proportional and optimal allocations.

Both percent accuracy and total dollar accuracy can be estimated simultaneously.

The plan is tested with an actual population obtained from insurance industry. Promising results are obtained.
Questions?