Audit Sampling: Steering in the Right Direction

Jason McGlamery
Director – Audit Sampling
Ryan, LLC
Dallas, TX
Jason.McGlamery@ryan.com

Brad Tomlinson
Senior Manager
(non-attorney professional)
Zaino Hall & Farrin LLC
Columbus, OH
btomlinson@zhftaxlaw.com
Learning Objectives

• Identify where sampling is applicable and where it is not

• Understand the various types of sampling methods

• Understand the pros and cons of various sampling methods
Learning Objectives

• Identify which sampling method to use

• Extrapolate sample results using various extrapolation methods

• Discuss why statistical sampling is the preferred method used by most taxing jurisdictions
Audit Sampling

“Audit sampling is the application of an audit procedure to less than 100 percent of the items within an account balance or class of transactions for the purpose of evaluation some characteristic of the balance or class.”

Statement on Auditing Standards No. 39 (SAS 39)
AICPA Professional Standards AU § 350.01
Sample Applications

• What can be sampled?
  – A/P transactions – purchases.
  – A/R transactions – sales data, joint interest billings.
  – G/L transactions – journal entries.
  – Property data, inventory records.
  – Unclaimed property records.

Almost anything can be sampled!
Statistical versus Non-statistical

• Statistical Methods:
  – Provide objective quantifiable estimates of sample risk based on scientific statistical theory.
  – Statistical sampling requires explicit consideration of sampling risk and probability theory during all phases including planning, selection, evaluation, and projection.
Statistical versus Non-statistical

• Non-statistical Methods
  – Provide no objective measure of sampling risk.
  – Evaluation of sampling risk depends on the auditor’s judgment and subjective opinion.
  – Often, no evaluation of sampling risk occurs.
Statistical versus Non-statistical

• Statistical Example:
  – The audit is planned to achieve a 95 percent probability that the estimated tax adjustment is within one percent (1%) of the true tax adjustment in the population.
  – If the sample does not achieve a one percent (1%) level of precision, the sample is expanded.
Statistical versus Non-statistical

• Non-statistical Example:
  – Block sampling depends on the auditor’s judgment that the selected items (i.e. months) are representative of the entire audit period.
Sampling Methods

• Block Sampling:
  – Selection of several items in sequence.
  – If too few blocks are selected, the probability of selecting a non-representative sample is too great.
  – “Representativeness” cannot be measured objectively.
Sampling Methods

• Cluster Sampling:
  – Special form of sampling, using either the random or systematic approach.
  – A group or group of sample items is selected. This group becomes one sample unit.
  – Since each group represents one unit, some sample efficiency may be lost.
Sampling Methods

• Systematic Sampling with Random Start:
  – Involves selecting samples at given intervals, after establishing a random starting place.
  – The random start is essential to allow each unit in the population an equal chance of being selected.
  – Advantage of systematic sampling is its ease of use – the sample can be drawn quickly, items are automatically in sequence, and documentation is easy.
Sampling Methods

• Unrestricted Random Sampling:
  – Used to draw individual sample items from the population, without segregating or separating any portion of the population.
  – Every item in the population has an equal chance of being selected as a sample unit.
  – One of the most commonly used sample selection techniques.
Sampling Methods

• Stratified Random Sampling:
  – This method segregates the population into similar homogeneous groups, which improves the efficiency of the sample and allows more attention to be focused on certain segments of the population.
  – Stratify (a.k.a. grouping) by dollar amount, attribute, or periods pre and post change (law, accounting system, operations).
Sampling Risk

• Sampling Risk Defined:
  – The risk that the projected tax assessment or refund based on the sample is significantly different from the amount that would have been determined if the audit procedures were applied to each item in the population.
  – Sampling risk can be reduced but not eliminated whenever a sample rather than a complete detail exam (census) is performed.
Non-Sampling Risk

• Non-Sampling Risk Defined:
  – The risk that the auditor reaches a conclusion different from the true overpayment or underpayment in the population for any reason other than the sampling procedures.
  – Non-Sampling risk would exist even if the auditor examined every item in the population.
Non-Sampling Risk

• Examples:
  – Data file does not properly represent population.
    • File contains duplicates or omissions.
    • File contains irrelevant adjusting JEs.
  – Tax errors in the sample incorrectly determined.
    • Auditor incorrectly applies law to sampled item.
    • Auditor does not consider changes in law during audit period.
    • Taxpayer fails to detect auditor’s errors.
Sample Planning Account Selection

- Judgmental Sampling Bias in Selection of Accounts:
  - Auditor’s need judgmental sampling to reduce the accounts they need to review:
    - Focus is to maximize tax underpayment recoveries for the Agency in a minimal amount of time.
    - This generally creates a sample bias in favor of the Agency.
Sample Planning Account Selection

• Judgmental Sampling Bias in Selection of Accounts:
  – Short tests may not be reliable if the 25% review is not performed in the exact order that the sample population was selected.
  • It is common for auditors to sort the sample in the order in which the taxpayer’s records are stored prior to reviewing the sample records.
  • May result in erroneous decisions to include or not include certain accounts.
Sample Planning Sample Size

• Required sample size increases for the following reasons:
  – The desired probability that the confidence interval will contain the true amount in the population is increased.
  – Greater variance exists in the population items.
  – The preferred tolerable error amount is reduced.

• Increasing sample size increases sample cost.
Sample Projection

• Estimation Methods:
  – Projected from Dollars in Sample and Population.
    • Separate Ratio.
    • Combined Ratio.
  – Projected from Number of Items in Sample and Population.
    • Mean Error.
    • Regression Estimation.
  – Best method varies depending on distribution of data.
Evaluate Sampling Risk Example

- **Goal:** The projection from the sample should have a 95% probability of being within 2% of the true tax underpayment in the population.

- **Result:** The projected assessment is $1,000,000 with a 95% confidence interval of $500,000 to $1,500,000 ($1,000,000 plus or minus 50%).

- **Conclusion:** Expand the sample until confidence intervals is within 2% of the mean (such as $1,000,000 plus or minus 2%, or $500,000 plus or minus 2%).
Estimation Methods

- Sample Design and Estimation Matrix:

<table>
<thead>
<tr>
<th>Type</th>
<th>Statistical Design</th>
<th>Random</th>
<th>Projection Methodology</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Separate</td>
<td>Combined</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ratio</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Difference</td>
<td>Regression</td>
</tr>
<tr>
<td>Block</td>
<td>N</td>
<td>Y/N</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Simple</td>
<td>N</td>
<td>Y</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>Stratified</td>
<td>N</td>
<td>Y</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>!</td>
<td>!</td>
</tr>
</tbody>
</table>

! - Denotes mathematically intense
Estimation Methods

• Ratio Projection Method
  – Basic idea: The ratio of the total sample error to the total sample base amount should be approximately equal to the ratio of the total population error to the total population base amount.

\[
\frac{\text{Sample Total Error}}{\text{Sample Base Amount}} \approx \frac{\text{Population Total Error}}{\text{Population Base Amount}}
\]
Estimation Methods

• Ratio Projection Method
  – Solving the equation for the population total error:

\[
\text{Pop Base Amount} \times \frac{\text{Sample Total Error}}{\text{Sample Base Amount}} \approx \text{Pop Total Error}
\]

– Example:

<table>
<thead>
<tr>
<th>Population Base</th>
<th>Sample Base</th>
<th>Audited Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>Count</td>
<td>Amount</td>
</tr>
<tr>
<td>$1,300,000</td>
<td>50,000</td>
<td>$10,000</td>
</tr>
<tr>
<td>$10,000</td>
<td>500</td>
<td>$300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error Rate Formula</th>
<th>Projecting the Error</th>
<th>Converting the Projection to Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Amount / Sample Amount</td>
<td>Error Rate</td>
<td>Projected Error</td>
</tr>
<tr>
<td>$300 / $10,000 = 3%</td>
<td>3% X $1,300,000 = $39,000</td>
<td>$39,000 X 5% = $1,950</td>
</tr>
</tbody>
</table>

2017 IPT Annual Conference – Charlotte, North Carolina
Estimation Methods

- Mean Projection Method
  - Basic idea: The ratio of the total sample error to the total sample base count should be approximately equal to the ratio of the total population error to the total population base count.
Estimation Methods

• Mean Projection Method
  – Solving the equation for the population total error:

\[
\text{Pop Base Count} \times \frac{\text{Sample Total Error}}{\text{Sample Base Count}} \approx \text{Pop Total Error}
\]

– Example:

<table>
<thead>
<tr>
<th>Population Base</th>
<th>Sample Base</th>
<th>Audited Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>Amount</td>
<td>Amount</td>
</tr>
<tr>
<td>$1,300,000</td>
<td>$10,000</td>
<td>$300</td>
</tr>
<tr>
<td>Count</td>
<td>Count</td>
<td>Count</td>
</tr>
<tr>
<td>50,000</td>
<td>500</td>
<td>12</td>
</tr>
</tbody>
</table>

- Error Rate Formula
  - Error Amount / Sample Count = Average Error
  - Mean Method
    - Error Amount = $300
    - Sample Count = 500
    - Average Error = $0.60

- Projecting the Error
  - Average Error = $0.60
  - Population Count = 50,000
  - Projected Error = $30,000

- Converting the Projection to Tax
  - Taxable Error = $30,000
  - Tax Rate = 5%
  - Projected Tax Error = $1,500
Estimation Methods

• Allocating the Error
  – Most states get it right, a few don’t
  – Allocating the error improperly is not only inaccurate, it can be costly
  – Essentially an apportionment calculation
  – Should take into account seasonal, cyclical, and/or economical circumstances
Estimation Methods

- Allocating the Error
  - Example: Using a straight line approach

<table>
<thead>
<tr>
<th>Simple Interest Rate</th>
<th>10%</th>
<th>Monthly Interest</th>
<th>0.833%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>Population</td>
<td>Monthly Interest</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>Amount</td>
<td>Assessment</td>
<td></td>
</tr>
<tr>
<td>J '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>5.00%</td>
</tr>
<tr>
<td>A '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>4.17%</td>
</tr>
<tr>
<td>S '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>3.33%</td>
</tr>
<tr>
<td>O '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>2.50%</td>
</tr>
<tr>
<td>N '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>1.67%</td>
</tr>
<tr>
<td>D '15</td>
<td>216,666.67</td>
<td>250.00</td>
<td>0.83%</td>
</tr>
<tr>
<td></td>
<td>1,300,000.00</td>
<td>1,500.00</td>
<td>43.75</td>
</tr>
</tbody>
</table>
Estimation Methods

• Allocating the Error
  – Example: Using the population characteristics to allocate the error

<table>
<thead>
<tr>
<th>Simple Interest Rate</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly Interest</td>
<td>0.833%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Month</th>
<th>Actual Spend of Total Pop</th>
<th>Monthly %</th>
<th>Monthly Interest Rate</th>
<th>Interest Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>J '15</td>
<td>50,000.00</td>
<td>3.85%</td>
<td>57.69</td>
<td>5.00%</td>
</tr>
<tr>
<td>A '15</td>
<td>75,000.00</td>
<td>5.77%</td>
<td>86.54</td>
<td>4.17%</td>
</tr>
<tr>
<td>S '15</td>
<td>150,000.00</td>
<td>11.54%</td>
<td>173.08</td>
<td>3.33%</td>
</tr>
<tr>
<td>O '15</td>
<td>275,000.00</td>
<td>21.15%</td>
<td>317.31</td>
<td>2.50%</td>
</tr>
<tr>
<td>N '15</td>
<td>350,000.00</td>
<td>26.92%</td>
<td>403.85</td>
<td>1.67%</td>
</tr>
<tr>
<td>D '15</td>
<td>400,000.00</td>
<td>30.77%</td>
<td>461.54</td>
<td>0.83%</td>
</tr>
<tr>
<td></td>
<td>1,300,000.00</td>
<td>100.00%</td>
<td>1,500.00</td>
<td>30.77</td>
</tr>
</tbody>
</table>
Estimation Methods

• Point Estimate versus Confidence Bound
  – Most common “form” of estimation utilizes the *Point Estimate*
  – A handful of states use the *Confidence Bound* (lower for net assessments or upper for net refunds)
    ▪ IA, KS, TN, UT, WA (soon to be MN)
    ▪ Difficult to explain and compute
    ▪ Each error’s individual projected value changes with identification of each additional error
    ▪ Total audit can’t be computed accurately until error schedule is finalized
Estimation Methods

• Point Estimate versus Confidence Bound

The Normal Curve

- $\pm 1\sigma \approx 68.3\%$
- $\pm 2\sigma \approx 95.4\%$
- $\pm 3\sigma \approx 99.7\%$
Estimation Methods

• Point Estimate versus Confidence Bound

The Normal Curve with 75% confidence interval

- $30,000 P.E.
- $18,263 LCL
- $11,737 Precision

$11,737

39% % Precision

LCL $18,263
Point Estimate $30,000
UCL $41,737
Estimation Methods

• Point Estimate versus Confidence Bound

The Normal Curve with 75% confidence bound

P.E. $30,000

LCL $23,121

Precision $6,879

$30,000

$23,121

$6,879

23% % Precision
Number of Errors Required for Projection

- The sample projection is based on the theory that the sample is representative of the population. A low error rate encountered in the sample may leave doubt as to whether the sample represents the population acceptably well.
  - The auditor’s judgment based on his or her knowledge of the business under audit must determine whether the sample should be used to project the measure of tax.
Advantages of Computer-Assisted Sample Audit

• Can be applied to a very large number of transactions.
• Efficient for state auditor and for taxpayer – saves money and time.
• Only small portion of transactions/invoices pulled.
• Minimizes on-site presence of auditor.
• Minimizes business disruption.
Computer-Assisted Audits: Important Considerations

- If your audit occurs in a state that conducts computer-assisted audits, research the state’s requirements for these audits.
- Prepare and participate in an opening conference on working out an agreement for a computer-assisted audit.
- Expect an analysis of the integrity of your electronic systems/files to be conducted.
- Agree on the statistical methods used.
Computer-Assisted Audits: Important Considerations

• Overview of sampling guidelines and projection methods used
  – Various statistical estimation (projection) methods are used in the stratified random samples:
    • Mean estimator
    • Difference estimator
    • Combined ratio estimator
    • Combined regression estimator
Computer-Assisted Audits: Important Considerations

• Overview of sampling guidelines and projection methods used
  – “Statistical vs. Non-statistical Sampling Methods”
  • Statistical sampling methods produce a measure of the reliability of the projection (estimate) usually expressed as a confidence interval estimate. Often, statistical procedures are used to determine the sample sizes in the stratified random sample.
  • Non-statistical sampling methods do not produce a measure of the reliability of the estimator.
Computer-Assisted Audits: Summary

• Computer-assisted audits require a considerable amount of preparation on the part of the taxpayer (e.g., meeting the electronic data and file requirements, negotiating with the auditor on the sampling plan). In return, the computer-assisted audit process saves time and money, and is less disruptive to business processes than “paper-based” audits. The use of the computer-assisted audits is certain to increase due to technological advances in data file software and electronic imaging of documents.
Stratification

- Stratification has two purposes:
  - If the error rate differs from low-dollar to high-dollar items, then stratifying by dollar value will create strata with the error rates more similar within a stratum than between different strata.
  - Stratification and sample size can control the impact that one error has on the total projection.
Stratification

- Stratification Methods:
  - Cumulative Square Root of the Frequency Distribution Method.
  - Cumulative Square Root of the Interval Width Method.
  - Dollar Frequency Method.
  - Judgmental Selection.
Stratification

• Cumulative Square Root of the Frequency Method:
  – After the number of sampling strata is decided, there are several methods to set the strata boundaries.
  – A generally recognized technique for setting strata boundaries is the cumulative square root of the frequency method [Dalenius and Hodges, “Minimum Variance Stratification”, 54 Journal of the American Statistical Association 88 (1959)].
Statistical Confidence Intervals

- Central Limit Theorem
  - The Central Limit Theorem (CLT) predicts that for a sufficiently large sample, the distribution of the sample mean is approximately normally distributed even if the population is skewed.
  - The normal distribution is the “bell curve”
Statistical Confidence Intervals

• Central Limit Theorem:
  – The difficulty in applying the CLT is to know how large a sample is needed for the distribution to be approximately normal.
  – Larger samples are needed when sampling for rare events or when working with severely skewed population.
Statistical Confidence Intervals

• Central Limit Theorem Example:
  Sample Mean = $100,000
  Standard Error of Sample Mean = $5,000

  – 68% of samples will have a sample mean with ± 1.00 standard errors.
  – 68% of sample means within interval ($95,000; $105,000)

  – 95% of samples will have a sample mean within ± 1.96 standard errors.
  – 95% of sample means within interval ($90,200; $109,000)
Statistical Confidence Intervals

• Sampling Distribution of the Sample Mean

\[ \mu_X = \mu \]

\[ \sigma_X = \frac{\sigma}{\sqrt{n}} \]

\[ \mu_X - 1.96\sigma_X \] to \[ \mu_X + 1.96\sigma_X \] corresponds to a 95% confidence interval.
Statistical Confidence Intervals

• Confidence is the probability that the designated interval contains the population value that would result from a 100% actual basis examination.

  – For example, suppose the auditor concludes the 95% confidence interval for sample mean error is ($400; $600)

  – This implies that if 100 different samples of that size were drawn from the population, then 95 of those sample means would be within ($400; $600), and 5 of those sample means would be outside that interval.
Statistical Confidence Intervals

• Precision is the amount from one end of the interval to the mean. The mean is usually at the midpoint of the confidence interval.
  – For example, if the mean is $500 and the precision is $100, then the confidence interval is $500 ± $100 or the interval is ($400; $600).
  – Precision could be stated as a percentage or as an absolute dollar amount.
  – Precision could be around the amount of tax error (“difference”) or around the amount of tax remitted.
Statistical Confidence Intervals

• Picking the confidence level for a lower bound is an important policy decision if the following apply:
  – (1) statistical evaluation is used to determine whether a sample is adequate evidence for projection, and
  – (2) assessment or refund will be made at the lower bound of the confidence interval
    • 90% two-sided confidence interval has lower bound (LB).
    • 5% probability that the mean of the population is below LB.
    • 5% above LB versus 5% below LB.
    • 19:1 odds ratio that mean in population is above LB.
Statistical Confidence Intervals

• Impact of Sample Size:
  – Assuming the sample mean is an unbiased estimate of the population, the sample mean is just as likely to increase or decrease from expanding the sample.
  – Expanding the sample will decrease the standard error of the sample mean, tighten the confidence interval, and very likely raise the lower bound of the confidence interval.