Part II - The Statistical Audit

Course Director: Edward J. Winslow

System Analyst

Auditmetrics

Third Party Administrator/Actuary –Ambulatory Health Services, Financial Auditor

info@auditmetrics.com

This section details with the statistical audit process as implemented and designed by Auditmetrics for the Massachusetts Department of Revenue. It was used for in-service training of sales and use tax auditors.

Text and Materials

- Audimetrics - AI v6.3 Software Recommended.
- Textbook: Statistical Audit - AI

Applying Artificial Intelligence techniques

- Slide Presentation
    - Part I Basic Principles
    - Part II the Statistical Audit

Sampling Risk

- Auditmetrics has developed a statistical sampling process to minimize the amount of sampling error due to random chance.
- At each step of the sampling process key indicators are monitored to provide immediate feedback.
- Statistical evaluation of sampling risk allows an objective quantified evidence of sample precision and representativeness.

Statistical Audit Two Stage Process

- Stage 1 – Random sample of an accounts ($X)
  - Based on ratio data quantitative variable sampling.
  - Sample criteria: 3% precision, 5% alpha error.
- Stage 2 – Review of sample to categorize transactions into an error/ non-error attribute or appropriate/not appropriate.
  - Nominal classification of all transactions into 1- in error and 0 not in error
  - Compile dollar volume for each classification and derive an error rate (p).
  - Use binomial distribution to determine post-audit margin of error and 90% confidence interval
  - 3% post audit precision or margin of error targeted

Calculate Sample Size
Simple Random Sample

Use Central Limit Theorem & Standard Normal Curve.

Standard Error = $\sigma / \sqrt{n}$

$Z_{.05} = 1.96$ ← **goal is 95% confidence**

$$1.96 = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

We want sample mean to be within 3% of population mean with 95% confidence or 5% alpha error. Proper sample size helps us define our sampling error and precision.

Calculate Sample Size (cont.)

Auditmetrics standard:

P (precision) = 3%

$Z_{.05}$ (95% Confidence) = 1.96

Sample size Calculation:

$$1.96 = \frac{P}{\sigma/\sqrt{n}}$$

$$\left(\frac{1.96 \times \sigma}{P} \right)^2 = n$$

Confidence Vs. Precision

- Confidence - Generally expressed as a percentage, confidence refers to the probability that a range around a sample estimate includes the true population value.

- Precision (margin of error) - is the size of the total range within which the population value would fall at a given confidence level.

For example we may target a precision to be within 3% of the true population mean with 95% confidence based on normal cure probabilities.

Simple Random Sample

Population:

Transactions	15,392
Total dollar Volume	$9,092,082
Mean Value	$2,380
Standard Deviation	$2,221

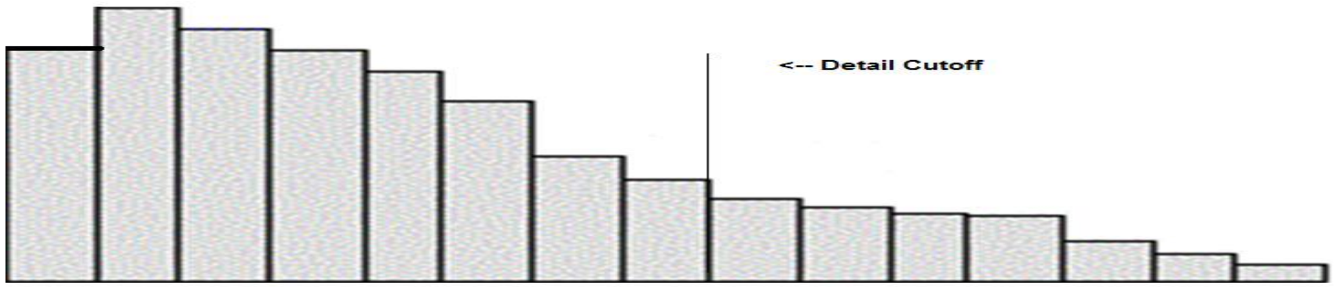
Precision .03 x $2,380 = $71.40

N = ((1.96 x $2,221)/ $71.40)^2

N= 3,717

Strategies for Efficiency

- What is the shape of the population?
- Large transactions evaluated at 100%



Our Sample Data Revisited

- Outliers (detail ≥ $4,000) Removed
- Keep margin of error the same, 3%
- Standard deviation is $975
- New mean will be $1,560
- Precision = $1,560 x .03 = $46.80
- Population =15,392 – 175 = 15,217

n = $((1.96 \times \$1{,}248)/\ \$46.8)^2$

n = 1,167 (started with 3,717)

Improve Sampling Efficiency

- What we have done previously is take very large transactions out of the population sampling pool.
- The rule of thumb is that the detail should be around 30% of total dollar volume.
- DOR has derived a distribution based criterion using the starting point to be the 97th percentile
- In this example 3% of the transactions and 27% of the dollar volume.
- There is special software to guide this process but can also be done with standard statistical software such as SAS, STATA and SPSS.

Improve Sampling Efficiency (cont.)

- The next step to improve sampling efficiency is to do a stratified random sample rather than a simple random sample.

- Why?
  - To gain sampling efficiency. A stratified random sample will usually yield more precise results by reducing random error.
  - More proportional representation of the population will more accurately reflect the population.
  - To offset the effect of extreme values (skewed distributions) which we have done when establishing the detail strata

Non-Stratified Standard Deviation

20

50

70

70

20

20

50

20

70

50

50

70

20

20

70

Mean = 44.7

Std. Dev. = 22.3 (error)

Let's now stratify these numbers

Strata Standard Deviations

The Impact of Stratification

strata	n	Mean	Std. Dev.
1 (20)	6	20	0
2 (50)	4	50	0
3 (70)	5	70	0

- In this example we went from SD of 22.3 to 0 using the stratification process.
- In sampling it spreads out the process and help’s assure better representation.

Strata Boundaries (Auditmetrics Criteria)

- The default number of strata used for sampling can vary from three to ten.

- The specific method used is “cumulative square root of the frequency method”. Most statistical packages do not generally focus on this specific method but its algorithm is pretty straight forward and DOR has developed its own software to accomplish this step.

- This method is well documented in many statistics textbooks and is a method well suited for determining strata boundaries for accounting records (variable sampling).

Stratify Transaction Data *(excluding Detail)*

	Sample (n)	Pop. Mean	Pop. Std. Dev.
0 -149.99	47	$44.04	$31.88
150-399.99	93	$230.93	$105.33
400-799.99	104	$435.50	$189.56
800-1524.99	130	$934.32	$378.80
1525-2624.99	150	$1,714.27	$596.87
2625-3999.99	174	$2,719.06	$789.95
Overall Std. Dev.	**$198**		
Sample Size	**698**		
Detail	**175**		
Total for Review	**873**		

Stratification Criteria

- Each Strata must be mutually exclusive: every element in the population must be assigned to only one stratum.
- The strata should also be collectively exhaustive: no population element can be excluded.
- Then simple random sampling is applied within each stratum.

Stratification Criteria (cont.)

- This improves the representativeness of the sample by reducing sampling error.
- It produces a weighted mean that has less variability than the arithmetic mean of a simple random sample of the population.

Sample Allocation to Each Strata

- Attribute Sampling Strata Allocation
  - Proportionate allocation- uses a sampling fraction in each of the strata that is proportional to that of the total population
  - For instance, if the population consists of 40% in the male stratum and 60% in the female stratum, the strata of the sample should reflect this proportion

- Variable Sampling Strata Allocation
  - Each stratum is proportionate to the standard deviation of the distribution of the variable.
  - Larger samples are taken in the strata with the greatest variability to generate the least possible sampling variance or uncertainty.

Neyman Allocation

Variable Sampling Stratification

Sample Size = 698

	N	Std. Dev.	N x Std. Dev.	Neyman Allo.	n
0 -149.99	5,996 (39.4%)	$31.88	191,152	6.7%	47
150-399.99	3,636 (23.9%)	$105.33	382,980	13.4%	93
400-799.99	2,250 (14.8%)	$189.56	426,510	14.9%	104
800-1524.99	1,404 (9.2%)	$378.80	531,835	18.6%	130
1525-2624.99	1,029 (6.8%)	$596.87	614,179	21.5%	150
2625-3999.99	902 (5.9%)	$789.95	712,535	24.9%	174
			2,859,192	**1**	**698**
Total N	**15,217**				
Detail	**175**				
Total Pop.	**15,392**				

Efficiency Factor

- Step 1 - We started with a simple random sample from a population with standard deviation of: $2,221
- Step 2 - Then by removing outliers standard deviation became: $975
- Step 3 - After stratification the *overall* standard deviation became: $198

The final stratified population for sampling represents an overall 91% reduction in standard deviation or an efficiency factor of .91

Efficiency Factor (Cont.)

- We started with a simple random sample:

  n = 3,717

- After creating the detail Strata

  n = 1,167 (excluding detail)

- After stratifying the remaining population

  n = 698 (excluding detail)

Sample Size (n) & Population Size (N)

- One important observation that one should note from statistical theory is that there is not a direct relationship between the audit population size and sample size.

- The amount of variation in the dollar amount of the invoices in the population has a greater impact on sample size than simply the size of the population.

- A sample from a population of records with a large range is going to require a larger sample than would be one with a smaller dollar range

Sample Submitted for Audit

	Sample (n)	Mean	Std. Dev.
0 -149.99	47	$44.04	$31.88
150-399.99	93	$230.93	$105.33
400-799.99	104	$435.50	$189.56
800-1524.99	130	$934.32	$378.80
1525-2624.99	150	$1,714.27	$596.87
2625-3999.99	174	$2,719.06	$789.95
Sample Size	**698**		
Detail	**175**	Targeted EF:	.80 or >
Total for Review	**873**	**Efficiency Factor .91**	

One Last Test -Sample Validity

- The calculation of an estimate of total tax liability is the end product of the audit.
- The statistical sample based on dollars is a cost effective first step of accomplishing that task.
- The next step is an assessment of the validity of the sample.
- In most situations the population and its parameters are unknown.
- In the statistical audit the population and its parameters are known.

Confidence Interval Test

- The stratified sample as originally derived can be further analyzed to reduce sampling error.
- Each strata random sample mean (statistic) can be compared to the mean of the population (parameter) of that stratum.
- Does the population parameter fall within a 95% confidence interval of the sample statistic.

Confidence Interval

Confidence Interval: An interval of values computed <u>from the sample,</u> that is almost sure to cover the true population value.

We make confidence intervals using values computed from the sample, not the known values from the population

Interpretation: In 95% of the samples we take, the true <u>population</u> proportion (or mean) will be in the interval.

This is also the same as saying we are 95% <u>confident</u> that the true population mean (or proportion) will be in that interval

95% CONFIDENCE INTERVAL

|-------------------------------|

strata	Sample Mean	Sample Std. Dev.	n	**Lower Bound**	**Upper Bound**	**Population mean**
0 -149	$44.30	$53.23	47	**$29.08**	**$59.51**	**$44.04**
150-399	$217.30	$139.84	93	**$188.88**	**$245.72**	**$230.93**
400-799	$555.65	$246.58	104	**$508.25**	**$603.04**	**$535.50**
800-1524	$972.79	$625.84	130	**$865.21**	**$1,080.38**	**$934.32**
1525-2624	$1,644.24	$1,258.32	150	**$1,442.87**	**$1,845.62**	**$1,714.27**
2625-3999	$2,426.48	$2,042.88	174	**$2,122.93**	**$2,730.02**	**$2,719.06**

698

Sample Precision and Accuracy
review of the process

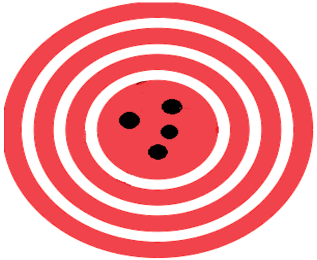
- We have just gone through the process of selecting a random sample to pass on to the auditor who completes the review.
- The steps of designing the sample size that meets DOR criteria were all selected to enhance projection precision:
  1. Determine outlier strata for review at 100%
  2. Stratify remaining population for sample size calculation using 3% precision and 95% confidence interval.
  3. Target efficiency factor to be .80 or better
  4. Perform validity check for each strata using 95% confidence interval test.

Sample Precision and Accuracy (cont.)

- All the steps of the process help reduce sampling error and therefore improve sampling precision

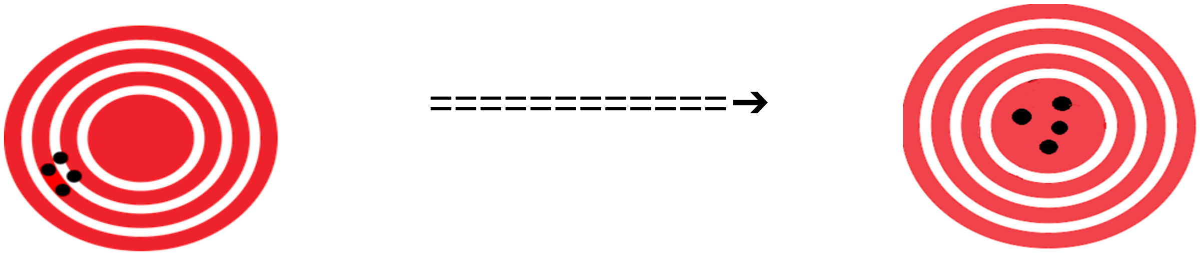
Simple random sample | stratified random sample

=============➔



Sample Precision and Accuracy (cont.)

- The validity check also helps in improving accuracy.
  - Since some strata sample sizes can be relatively small they are more susceptible to random fluctuations.
  - Validity check helps reign in extreme strata random fluctuations to assure a sample estimate within the targeted 95% confidence interval boundary.



Sample Precision and Accuracy (cont.)

- Accuracy can also be influenced by other factors other than random sampling chance, such as human error and other non-sampling risks:
  - Improper or imprecise audit population definition.
  - Improperly trained auditors
  - Unclear evaluation criteria
  - Improper or insufficient documentation
- Non-sampling risk is difficult to quantify, but can be minimized by implementing good quality control procedures.

Sampling & Non-Sampling Risk (Bias)

- Non-Sampling risk is the risk that the sample results do not represent the target population raising questions about validity.

- Sampling or random risk can be reduced by increasing sample size.

- Auditmetrics uses specific criteria help keep sampling risk within specified bounds.

- Both risks are always present and may interact in in unexpected ways

Uncertainty – Sampling Error

- There is an important distinction between uncertainty and bias.
- I have made the point that we are making decisions under conditions of uncertainty.
- Whenever a population has a lot of variation (high standard deviation) there is uncertainty about the extent to which the sample results mirror population.
- That is why the uncertainty or sampling error is contained by specific criteria for confidence interval and precision.

Non-Sampling Error-Selection Bias

- Selection bias occurs when sampling frame does not properly reflect the target population.
- The sampling frame may omit some category of transactions that should be included.
- Does the sampling frame include duplicates or some other form of over or under representation.
- These problems may be due to data retrieval difficulties .
- The auditor should be on guard for selection bias.

Non-Sampling Error - Measurement Bias

- Measurement bias is the incorrect analysis of items that appear in the sample.
- It may occur when the auditor is either not clear, incorrectly interprets or incorrectly described of the facts of the data elements.
- These incorrect interpretations can be prevented by better training and more openness by the taxpayer.

Balancing Risks and Cost

- Balancing sampling risk, non-sampling risk and sampling cost is a constant challenge.
- There is no “right answer”.
- Instead this is a process of mutual cooperation.
- It may include modifying sample size considerations or further population analysis and specifications before the sample is selected.

Is Sample Precision 3%

Using Same Separate Ratio Estimator

		Sample	Sample Projected	Actual
strata	Pop. N	Mean	Total $X	Total $X
0 -149.99	5,777	$44.30	$255,921	$254,419
150-399.99	3,415	$217.30	$742,080	$788,625
400-799.99	2,029	$555.65	$1,127,414	$883,629
800-1524.99	1,183	$972.79	$1,150,811	$1,105,300
1525-2624.99	1,470	$1,644.24	$2,417,033	$2,519,976
2625-3999.99	1,343	$2,426.48	$3,258,763	$3,651,697
	15,217		**$8,952,020**	**$9,203,649**

Statistic:	**Sample Projected Mean**	**$588.29**
Parameter:	**Actual Mean**	**$604.83**
Actual Precision		**2.7%**
Target Precision		**3%**

Steps in the Audit Process Revisited

- A process that starts with known parameters ($).
- Allows to test validity of sample selected.
- Statistically estimate an unknown parameter (error rate).

	Steps in the Audit Process	Dist	Parameters	Data Type
1	Define population in terms relevant $x.	Normal	known	Ratio
2	Properly sample the population of $x.	Normal	known	Ratio
3	From that sample determine those values of $x where an error was made.	Binomial	unknown	Nominal
4	From the sample make projections about the population error rate	Binomial	unknown	Nominal

Conducting the Audit

- Once a sample of the accounts is drawn that meets all Audit Bureau criteria the stage is set for the actual review of each sampled transaction.
- Each invoice has a known invoice amount.
- Each invoice is divided into one of two classes, one that the transaction is in error and one that is not in error.
- All transactions $x in error are summed for each stratum

Separate Ratio Estimator

- We assume that error invoices and non-error invoices are randomly distributed among the N population units.

- *Post Audit – error probability(p) is:*

  Sample: $p = \Sigma (\$X \text{ error}) / \Sigma (\$X \text{ total})$

- *This p is then applied to the corresponding stratum population total to estimate the stratum error in the population being tested.*

Sample Error Rate by Strata

Stratum Specific Error Rate

strata	n	Sample Mean	Sample Total $X	Sample Total $error	Strata Error Rate
0 -149	47	$44.30	$2,081.92	$193	9%
150-399	93	$217.30	$20,209.18	$1,256	6%
400-799	104	$555.65	$57,787.19	$1,775	3%
800-1524	130	$972.79	$126,463.00	$4,847	4%
1525-2624	150	$1,644.24	$246,636.21	$12,340	5%
2625-3999	174	$2,426.48	$422,207.43	$18,297	4%
	698				

Calculate Total Error Rate

Sample stratum specific error rates applied to population total dollar volume.

Separate Ratio Estimate

strata	N	Population Mean	Population Total $X	Sample Strata Error Rate	Projected Pop. Dollars in Error
0 -149	5777	$44.04	$254,410	9%	$23,162
150-399	3415	$230.93	$788,632	6%	$47,579
400-799	2029	$535.50	$1,086,527	3%	$32,846
800-1524	1183	$934.32	$1,105,304	4%	$44,474
1525-2624	1470	$1,714.27	$2,519,977	5%	$126,264
2625-3999	1343	$2,719.06	$651,698	4%	$146,333
Total	15217	$425.29	$6,471,679		$420,659

Detail	175	$14,973.73	$2,620,403	2%	$52,408

Population Error Rate/Amt. Taxable

Amount Error/Population Total

	N		Population Total	Amount in Error
Non-Detail	15,217		$6,471,679	$420,659
Detail	175		$2,620,403	$52,408
Total	15,392		$9,092,082	$471,476

Sampled Population Error Rate **6.5%** **(excluding Detail)**

Total Population Error Rate **5.2%** **(including Detail)**

Projected Amount Taxable

- Method One:

Excluding Detail	6.5%	$6,471,679	$420,659
Detail			$52,408
Total			**$473,067**

- Method Two **(used in multi-year projections):**

Including Detail	5.2%	$9,092,082	**$473,067**

Projected Taxes Due Boundaries

- In the previous table the audit estimated single year taxable amount is 6.5% X $6,471,679 = $420,659
- However we must also set up confidence intervals around that estimate.
- Along with confidence interval we must also assure the taxpayer that our precision or margin of error is within 3%.
- Do the criteria in setting up the sample hold true once the audit is completed?

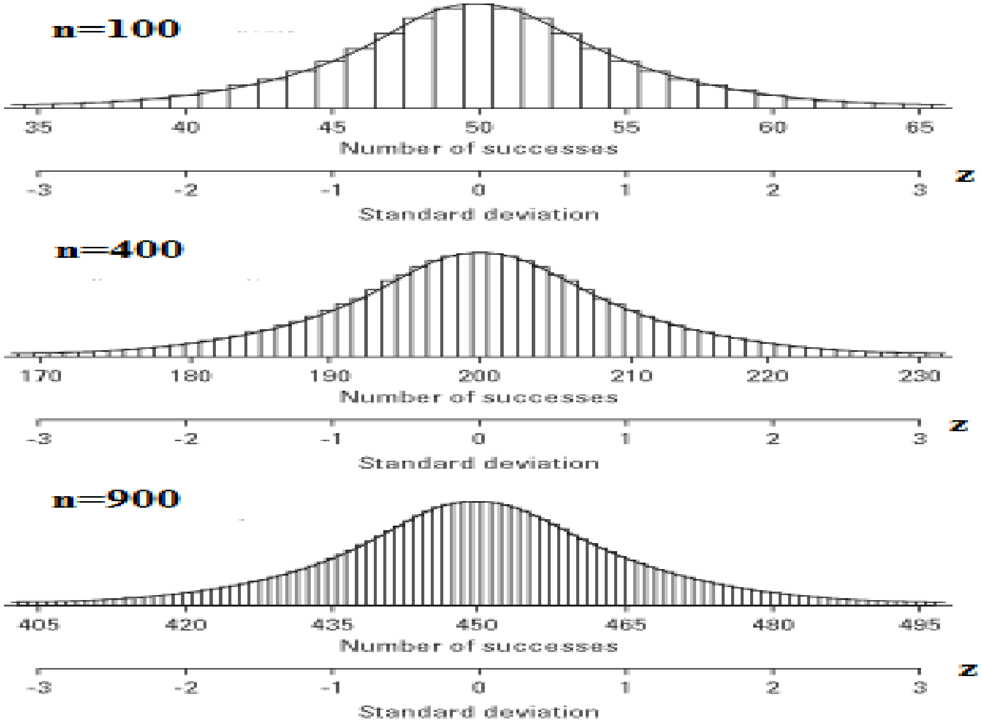
Our Binomial Revisited - (Toss of coin heads a success)

- The probabilities we calculated using the binomial formulas with sample size n=5; error rate = 50%

- The binomial is based on a random process that is symmetrical like the normal curve.
- When sample size increases then the binomial merges into the normal curve.

Binomial as n increases

- The z-score is the standard normal score.
- Between -1.96 z & 1.96 z = 95% of the area.



The Link Between Binomial and Normal

- P- error rate.
- P (1-P) = error rate variance = σ^2
- $\sqrt{P(1-P)}$ = error rate standard deviation = σ
- $\sqrt{P(1-P)}/\sqrt{n}$ = error rate standard error
- $\sqrt{P(1-P)}/\sqrt{n} \times Z$ = precision or margin of error

Z = +- 1.645 for 90% confidence interval

Z = +- 1.96 for 95% confidence interval

Completion of the Audit

Using the previous formulas:

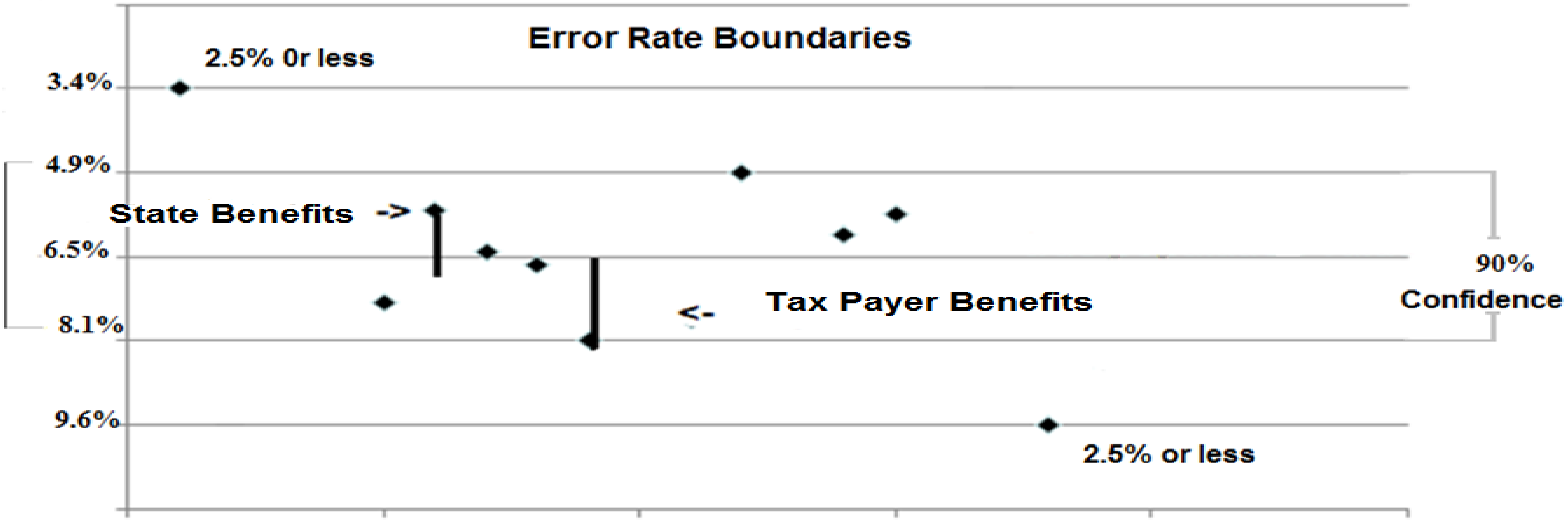
- P = 6.5% = $420,659 taxable
- Standard deviation =.246 = √.065 x .935
- standard error = 0.0094 = .246/√698
- margin of error = .0153 = .0094 x 1.645
- 90% lower bound = .049 = .065- .0153
- 90% upper bound = .081= .065 + .0153

Audit Projection Confidence Interval

- We calculated the margin of error = 1.53%
- DOR standard is margin of error is 3% or less
- This audit's error rate meets DOR standard
- Population total = $6,471,679

Amount Taxable-Midpoint:		$420,659	
90% confidence Interval	**Lower Bound**	**Upper Bound**	**Range**
Amount Taxable	$321,642	$519,676	$198,034

Final Step - Audit Boundaries



The distribution between the boundaries is not even – 68% of possible samples would be between 5.5% and 7.5%

In the long run over time all statistical errors will net out to zero