Statistical Testing of Installed Meters

Background, Development, Application, and Implementation
Background

- Statistical Meter Sampling & Testing Concepts
- Why Statistical Meter Testing Programs?
- New York Program Case Study
- Results & Lessons Learned
Advent Design is primarily a custom automation designer & integrator and a manufacturing consulting firm. Over the years, Advent had become involved with the design and automation of meter shops. With deregulation in many energy markets, the needs of customers shifted towards cost saving measures, and the use of statistical testing, a common tool in many manufacturing operations, was an obvious way to reduce costs while improving the quality of the information gained from meter testing.
What is Statistical Testing?

- Statistical testing is the testing of a population or group for specific characteristics or parameters using a valid statistically-derived sampling plan.
Features of a Statistical Sampling/Testing Plan

- Homogeneous Population(s)
- Sample(s) of a Suitable Size for the Plan
- Random Sample Selection of Items to Be Tested
- Expectation that the Group or Population Being Tested Fits the Statistical Model
Homogeneous Population(s)

- The groups or populations being sampled and tested are made up of the same or similar items, items which operate in the same way and were made in the same manner.

- For gas meters, this has traditionally been interpreted as being meters of a specific meter type from a manufacturer (i.e. AC250TC, R175, SP250, etc.).
Suitably Sized Samples

- The sample size for each group must be large enough to provide a statistically valid sample for the group’s population.

- The larger the group’s population, then the larger the sample will be up to a certain point.
Random Sample Selection

- Every item within the group or population has an equal chance of being selected as part of the sample for testing.

- Random sample selection is critical to providing for a statistically valid sample.
The statistical model being used for the sampling/testing plan needs to match the actual distribution of the population.

In most circumstances, one is looking at a normal or Gaussian distribution (i.e. a Bell curve).

This can be checked using a histogram plot or a correlation analysis.
For diaphragm meters, a normal distribution fits actual data fairly well.

For rotary and turbine meters, there is some question due to the failure modes of these meters.

With the typically small populations of these meter types, it is difficult in most situations to verify a normal distribution.
Why Use a Statistical Sampling/Testing Plan?

- Focuses testing on the proper meters.

- Minimizes number of meters to be tested; usually requires less than 30% of what a periodic testing plan would require.

- Can provide data and analysis tools for use in understanding what is happening with meters installed in the field or for use in the purchasing of new meters.
Case Study History

- New York State Gas Measurement Committee (GOSURE) formed the Regulatory Sub-Committee in 1997 to address regulation reform issues related to gas measurement on a state-wide basis.

- Existing Alternate In-Test Program (AIP) was targeted by the group for possible reform in 1998.

- Advent Design was selected to complete an analysis of the existing meter change programs, compile state-wide meter test data, and work with GOSURE to develop a new statistically-based meter test and replacement program for the state of New York.
Program Background

- Previous Alternate In-Test Program (AIP)
  - Formula-based meter change rate
  - Change rate begins at 3.5%
  - Rate increased based upon previous year’s meter test results
  - Annual test rate historically averaged between 4.5% and 5.5% of the meter population
  - 90% of meter tests found within +/- 2% limits
Statistical Meter Testing Evaluation Project Goals

- To Maintain or Improve the Accuracy of the Installed Meter Population, While Reducing the Annual Meter Change Rate

- Separate "Analytic Testing Process" from the "Meter Change Out" Process
- Utilize Accepted Sampling Techniques to Evaluate Meter Group Performance
- Reduce Number of Required Meter Tests
- Develop Group Remediation Process which Targets Poorest Performing Meter Types
- No Change in Tested Parameter (in-test check rate)
“Meters shall be divided into homogeneous groups such as year set, manufacturer, case type, diaphragm material, etc. However, any group of meters maybe combined with another group with similar operating characteristics to provide an adequate sample size for analysis and control. Meters in any given group maybe further subdivided according to location, age or other factors that maybe disclosed by test data to have an effect on the performance of the meters. Subsequently, groupings maybe modified or combined as justified by the performance records.”
Statistical Meter Testing Program

KEY CONCEPTS

Assumptions

Gas Meter performance degrades over time and level of usage

Meter Sampling Criteria

Meters are separated into Homogenous Groups which have similar performance characteristics using ANSI B109 Series

Acceptable Quality Levels (AQL)

Determine acceptable level of meter performance

Analysis of Failed Groups & Corrective Action

Develop meter removal and replacement plan
Random Sampling Plan Primer

- Allows one to draw a statistically-valid sample from a population to determine conclusions about specific characteristic(s).

- Establishes a limit which defines the maximum allowable number of defects or the maximum percentage defective that will be considered acceptable as a process average.

- Operating Characteristic (OC) Curves - probability of acceptance at all levels of percentage defective

- Types of Sampling Plans

- Classification of Sampling Plans: Attributes (ANSI Z1.4) or Variables (ANSI Z1.9)
NY METER SAMPLING SURVEY STRUCTURE & Random Sampling Plan Development

STATISTICAL RANDOM SAMPLE PLAN

- Understand Meter Populations
- Develop Sampling Concepts
- Develop Method to Group Meters
- Develop Sampling Method (AQL, Inspection Level)
- Calculate Sample Size
- Determine Group Pass/Fail Criteria
- Consider Present Meter Selection Plans
- Develop Optimum Sample Plan
- Develop Sample Pass/Fail Criteria and Evaluation Methods

SURVEY STRUCTURE

- Meter Population
- Meter Information and Tracking
- Selection of Meter Samples
- Test Data Tracking Methods
- Reporting Procedures & Data

- Numbers
  - Types
  - Groups
  - Serial Numbers
- Information
  - Consumption
  - Historical Data
  - Maintenance
- Methods
  - Approaches
  - Retirement Plan
- Records
- Failures
- PSC Reports
- Comment
Developing the Meter Sampling Program

The Random Sample Program uses:

1. **SAMPLING BY VARIABLES**
   - ANSI/ASQ Z1.9 OR MILSTD 414

2. **SAMPLING BY ATTRIBUTES**
   - ANSI/ASQC Z1.4

3. **VARIABILITY UNKNOWN - STANDARD DEVIATION METHOD**

**SELECT SPEC LIMIT**

- **SINGLE SPECIFICATION LIMIT**
- **DOUBLE SPECIFICATION LIMIT**

**DETERMINE TESTING LEVEL**

- Special S3, Special S4, General I, General II, OR General III.
  - ANSI/ASQC states that unless otherwise suggested, General II be used.

**DEFINE AQL**

- AQL may be dictated by Utility Commission based on confidence requirements.

- **10% and 20% AQLs**

**CHOOSE SAMPLING PLAN**

- SAMPLING BY VARIABLES
- SAMPLING BY ATTRIBUTES
- VARIABILITY UNKNOWN - STANDARD DEVIATION METHOD
AQL Determination

- Based upon existing state-wide meter accuracy
- Agreed to by all New York utilities participating in the plan and the NY PSC
- Single value for all companies for a particular meter type
- New plan must not degrade current level of metering accuracy
How the Statistical Testing Program Works

1. Identify reject meter groups
2. Analyze test data for bad data, outliers, etc.
3. Determine meter subgroups that should be retired based on set date, purchase date, etc.

- **STEP 1**: DETERMINE METER GROUPS
  - Should consist of meter units of a single attribute and manufactured under similar conditions

- **STEP 2**: OBTAIN RANDOM SAMPLE SIZE FOR EACH METER GROUP
  - Obtain Sample Size

- **STEP 3**: GENERATE RANDOM SAMPLE FROM EACH GROUP
  - A small percentage of extra meters could be pulled to allow for meters not able to be tested.

- **STEP 4**: PERFORM TESTING ON SELECTED METERS
  - As required depending on type of meter (gas, electric, water)

- **STEP 5**: PERFORM CALCULATIONS OF METER TEST RESULTS
  - Calculate Standard Deviation, Calculate Mean, Define Specification Limits, Calculate Quality Index, Look up Estimates of Lot Pct Ncf, Test for Acceptability Criteria

- **STEP 6**: TEST RESULTS ARE REVIEWED
  - 1. Identify reject meter groups
  - 2. Analyze test data for bad data, outliers, etc.
  - 3. Determine meter subgroups that should be retired based on set date, purchase date, etc.
Both use variables (a measured parameter or characteristic) as the basis for its analysis. In this case, in-test check rate was used.

- Variety of special and general inspection levels
- Selection of Acceptable Quality Levels (AQL’s)
The selected method was the Variability Unknown - Standard Deviation Method with Double Specification Limits.

General Inspection Level II for normal inspection (MIL-STD Inspection Level IV) was selected.

The Inspection Level is used in conjunction with group size in Table A-2 to determine sample size code letters.
### Table A-2

<table>
<thead>
<tr>
<th>Lot Size</th>
<th>Inspection Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>3 to 6</td>
<td>B</td>
</tr>
<tr>
<td>9 to 15</td>
<td>B</td>
</tr>
<tr>
<td>16 to 25</td>
<td>B</td>
</tr>
<tr>
<td>26 to 40</td>
<td>B</td>
</tr>
<tr>
<td>41 to 65</td>
<td>B</td>
</tr>
<tr>
<td>66 to 110</td>
<td>B</td>
</tr>
<tr>
<td>111 to 180</td>
<td>C</td>
</tr>
<tr>
<td>181 to 300</td>
<td>D</td>
</tr>
<tr>
<td>301 to 500</td>
<td>E</td>
</tr>
<tr>
<td>501 to 800</td>
<td>F</td>
</tr>
<tr>
<td>801 to 1,300</td>
<td>G</td>
</tr>
<tr>
<td>1,301 to 3,200</td>
<td>H</td>
</tr>
<tr>
<td>3,201 to 8,000</td>
<td>I</td>
</tr>
<tr>
<td>8,001 to 22,000</td>
<td>J</td>
</tr>
<tr>
<td>22,001 to 110,000</td>
<td>K</td>
</tr>
<tr>
<td>110,001 to 550,000</td>
<td>L</td>
</tr>
<tr>
<td>550,001 and over</td>
<td>M</td>
</tr>
</tbody>
</table>

1Sample size code letters given in body of table are applicable when the indicated inspection levels are to be used.
ANSI/ASQC Z1.9 and MIL-STD 414E
Acceptable Quality Level (AQL)

- AQL is the maximum percent nonconforming that, for purposes of sampling inspection, can be considered satisfactory as a process average.

- For ANSI/ASQC Z1.9, AQL’s vary from 0.10 to 10.00 and from 0.04 to 15.00 for MIL-STD 414E.

- For use with in-service gas meter testing, AQL’s from 4.0 to 20.00 are normally utilized.
## MIL-STD 414E - Table B-3

### TABLE B-3

Master Table for Normal and Tightened Inspection for Plans Based on Variability Unknown

(Double Specification Limit and Form 2 - Single Specification Limit)

<table>
<thead>
<tr>
<th>Sample code letter</th>
<th>Sample size</th>
<th>Acceptable Quality Levels (Normal Inspection)</th>
<th>Acceptability Quality Levels (Tightened Inspection)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.04</td>
<td>.065</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Notes:
- All AQL and table values are in percent defective.
- Use first sampling plan below arrow. When sample size equals or exceeds lot size, every item in the lot must be inspected.
Determine the mean and the standard deviation for the sample results.

Determine Quality Indexes
Qu = (Upper Limit - mean)/standard deviation
Ql = (mean - Lower Limit)/standard deviation
Upper Limit is normally 102, and Lower Limit is normally 98.

Use Qu and Ql to determine estimate of percent nonconformance above the Upper Limit (Pu) and below the Lower Limit (Pl) using Table B-5.
<table>
<thead>
<tr>
<th>(Q_0) or (Q_1)</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
<th>50</th>
<th>75</th>
<th>100</th>
<th>150</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>.72</td>
<td>28.57</td>
<td>26.00</td>
<td>25.09</td>
<td>24.39</td>
<td>24.03</td>
<td>23.83</td>
<td>23.75</td>
<td>23.71</td>
<td>23.68</td>
<td>23.67</td>
<td>23.64</td>
<td>23.61</td>
<td>23.60</td>
<td>23.59</td>
<td>23.59</td>
</tr>
<tr>
<td>.73</td>
<td>28.22</td>
<td>25.67</td>
<td>24.76</td>
<td>24.07</td>
<td>23.72</td>
<td>23.52</td>
<td>23.44</td>
<td>23.40</td>
<td>23.37</td>
<td>23.36</td>
<td>23.33</td>
<td>23.31</td>
<td>23.30</td>
<td>23.29</td>
<td>23.28</td>
</tr>
<tr>
<td>.74</td>
<td>27.86</td>
<td>25.33</td>
<td>24.44</td>
<td>23.75</td>
<td>23.41</td>
<td>23.21</td>
<td>23.13</td>
<td>23.09</td>
<td>23.07</td>
<td>23.05</td>
<td>23.02</td>
<td>23.00</td>
<td>22.99</td>
<td>22.98</td>
<td>22.98</td>
</tr>
<tr>
<td>.75</td>
<td>27.50</td>
<td>25.00</td>
<td>24.11</td>
<td>23.44</td>
<td>23.10</td>
<td>22.90</td>
<td>22.83</td>
<td>22.79</td>
<td>22.76</td>
<td>22.75</td>
<td>22.72</td>
<td>22.70</td>
<td>22.69</td>
<td>22.68</td>
<td>22.68</td>
</tr>
<tr>
<td>.76</td>
<td>27.13</td>
<td>24.67</td>
<td>23.79</td>
<td>23.12</td>
<td>22.79</td>
<td>22.60</td>
<td>22.52</td>
<td>22.48</td>
<td>22.46</td>
<td>22.44</td>
<td>22.42</td>
<td>22.40</td>
<td>22.39</td>
<td>22.38</td>
<td>22.38</td>
</tr>
<tr>
<td>.77</td>
<td>26.76</td>
<td>24.33</td>
<td>23.47</td>
<td>22.81</td>
<td>22.48</td>
<td>22.30</td>
<td>22.22</td>
<td>22.18</td>
<td>22.16</td>
<td>22.14</td>
<td>22.12</td>
<td>22.10</td>
<td>22.09</td>
<td>22.08</td>
<td>22.08</td>
</tr>
<tr>
<td>.88</td>
<td>22.42</td>
<td>20.67</td>
<td>20.00</td>
<td>19.48</td>
<td>19.23</td>
<td>19.10</td>
<td>19.05</td>
<td>19.02</td>
<td>19.00</td>
<td>18.99</td>
<td>18.98</td>
<td>18.96</td>
<td>18.95</td>
<td>18.95</td>
<td>18.95</td>
</tr>
<tr>
<td>.89</td>
<td>21.99</td>
<td>20.33</td>
<td>19.69</td>
<td>19.19</td>
<td>18.95</td>
<td>18.82</td>
<td>18.77</td>
<td>18.74</td>
<td>18.73</td>
<td>18.72</td>
<td>18.70</td>
<td>18.69</td>
<td>18.68</td>
<td>18.68</td>
<td>18.68</td>
</tr>
</tbody>
</table>
ANSI/ASQC Z1.9 and MIL-STD 414E Calculations for Standard Deviation Method

- With the values of Pu and Pl determined from Table B-5 using Qu and Ql, estimated percent nonconformance equals to Pu plus Pl. (% ncf = Pu + Pl)

- Acceptance is based on whether the estimated percent nonconformance is below the allowed percent nonconformance given in Table B-3.
Failed Group Remediation Options

- Failed meter groups may be retired from service within a period not to exceed 8 years
  - *No Further sampling Required*

- Meter group may be sub-divided to determine if group failure can be traced to common trait using a root cause analysis.
  - Retire sub-set of entire group
  - Continue to sample remaining population
  - *If remaining population meets AQL performance requirements, remediation is discontinued*
Current NY Program Status
Based on Representative Utility Data

- 2001 and 2002 testing completed
- 7 of 8 major utilities now participating

Summary Analysis for AQL Selection
- Residential Dia. Meters AQL = 10%
- C/I Diaphragm Meters AQL = 20%
- Rotary and Turbine Meter AQL = 4.0 or 6.5%

2001 Data Summary

<table>
<thead>
<tr>
<th>Type</th>
<th>Population</th>
<th>Failed Estimate</th>
<th>%NCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Diaphragm Meters (&lt; 500 CFH)</td>
<td>1,504,325</td>
<td>143,852</td>
<td>9.6%</td>
</tr>
<tr>
<td>Large Diaphragm Meters (&gt; 500 CFH)</td>
<td>29,021</td>
<td>7,397</td>
<td>25.5%</td>
</tr>
</tbody>
</table>
### Old AIP vs Statistical Meter Testing

The chart below compares meter change and test requirements between the old AIP and the new Statistical Sampling Plan. All calculations using 2001 test data.

<table>
<thead>
<tr>
<th>Meter Type</th>
<th>Sample Quantity</th>
<th>Remediation</th>
<th>Total Random Plan</th>
<th>AIP Quantity</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Diaphragm Meters</td>
<td>1,408</td>
<td>19,229</td>
<td>20,637</td>
<td>79,957</td>
<td>59,320</td>
</tr>
<tr>
<td>Large Diaphragm Meters</td>
<td>275</td>
<td>1,699</td>
<td>1,974</td>
<td>2,150</td>
<td>176</td>
</tr>
<tr>
<td>Total</td>
<td>1,683</td>
<td>20,928</td>
<td>22,611</td>
<td>82,106</td>
<td>59,496</td>
</tr>
</tbody>
</table>

- **Random Sample Plan**
  - 59,320 fewer residential meter changes
  - 176 fewer C/I diaphragm meter changes
Failed Meter Performance

Percentage of “Fast” and “Slow” meters by meter type

2001 Data Summary

<table>
<thead>
<tr>
<th>Meter Type</th>
<th>Total Fail</th>
<th>Percent Slow</th>
<th>Percent Fast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sm. Dia. Meters for Removal</td>
<td>23.4%</td>
<td>26.5%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Acceptable Sm. Dia. Meters</td>
<td>6.7%</td>
<td>7.7%</td>
<td>1%</td>
</tr>
<tr>
<td>Lg. Dia. Meters for Removal</td>
<td>33.9%</td>
<td>38.5%</td>
<td>4.6%</td>
</tr>
<tr>
<td>Acceptable Lg. Dia. Meters</td>
<td>7.5%</td>
<td>13.6%</td>
<td>6.1%</td>
</tr>
</tbody>
</table>

• Meters for removal are the poorest performers
• Poor performing groups are overwhelmingly “Slow”
Measurement Accuracy Gain

- Removed meter accuracy comparison between the old AIP and the new Statistical Sampling Plan based on 2001 and earlier test data

<table>
<thead>
<tr>
<th>Meter Type</th>
<th>Total Random Plan</th>
<th>Random Plan Weighted Avg. Meter Accuracy</th>
<th>AIP Qty.</th>
<th>AIP Weighted Avg. Meter Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Diaphragm Meters (&lt;500cfh)</td>
<td>20,637</td>
<td>101.0</td>
<td>79,957</td>
<td>100.5</td>
</tr>
<tr>
<td>Large Diaphragm Meters (&gt;500cfh)</td>
<td>1,974</td>
<td>101.1</td>
<td>2,150</td>
<td>100.6</td>
</tr>
<tr>
<td>Total</td>
<td>22,611</td>
<td></td>
<td>82,106</td>
<td></td>
</tr>
</tbody>
</table>

Removed meters under the Statistical Sampling Plan are 0.5% slower than those being removed under the old AIP Plan. So, the new program focuses remediation efforts on poor performing groups.
Meter Remediation for 2002

- Meter change quantities based on the results from the new AIP Plan in 2001:

<table>
<thead>
<tr>
<th>Small Dia Summary</th>
<th>Total failed Meters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Fail Population</td>
<td>153,832</td>
</tr>
<tr>
<td>Annual Removals (8 yr)</td>
<td>19,229</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Large Dia Summary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Fail Population</td>
<td>13,590</td>
</tr>
<tr>
<td>Annual Removals (8 yr)</td>
<td>1,699</td>
</tr>
</tbody>
</table>

| Total Projected 2002 Remolutions for Cause | 20,928 |
Management of Random Sample Program

- Random sample selection targets specific customer account locations.
  - Select more candidates than required for sample (2 to 5 times more) to accommodate incorrect meter type, access problems, etc.
  - Track progress in obtaining random sample quantity

- Some changes in Meter Groups may be required.

- Ability to analyze test data is important for corrective action.
  - Use of a computer program to do the analysis and to review test data & statistics is very helpful.
Management of Random Sample Program

- **Meter handling issues**
  - Meter transportation and handling procedures following removal
  - Strong connection to “did not register” (DNR) meters
  - Use of third-party testing increases handling issues dramatically

- **Test procedures to ensure accurate test results**
  - Meter acclimation
  - Test equipment setup and maintenance
  - *A few erroneous tests (“outliers”) can dramatically effect statistical results due to reduced sample sizes!*
Data Usage in Operations

- Proactively remove poor performers as “tag-along” jobs where possible
  - Reduce risk of future Meter Group failure and remediation

- Use statistics in meter purchasing decisions to factor in meter life cycle cost and accuracy
Open Issues for the New AIP Plan

- Final determination on the program
  - Two year evaluation window ended in March 2003

- Finalization of AQL values
  - Probably 10% for small diaphragm meters and 20% for large diaphragm meters

- Handling of “outliers”, test results well outside of the normal range of test results

- Keeping rotary and turbine meters in the AIP versus putting them back on a periodic testing plan

- Sampling from small meter groups (population <100)
  - For groups with fewer than 100 meters, sampling under the new AIP Plan is actually greater than that for a periodic program.
Statistical Meter Testing Program Summary

◆ The new AIP Plan has accomplished the following:
  
  ◆ Identified meter groups not performing to accepted test standards with confidence level of 95% or better
  
  ◆ Provided a method to analyze failed meter groups
  
  ◆ Developed a remediation process which targets the poorest performing meter types
  
  ◆ Improved meter population performance while reducing the number of required tests nearly 80%!