Choosing the Best Device Sample Size for Verification and Validation

Steven Walfish
Staff Statistician, Corporate Quality
Agenda

- Validation scope
- Statistical tools
- MSA
- Design Verification and Validation
- Other Tools for validation
A manufacturer can assure through careful design of the device and packaging, careful design and validation of processes, and process controls, that there is a high probability that all manufactured units will meet specifications and have uniform quality.

Where process results cannot be fully verified during routine production by inspection and test, the process must be validated according to established procedures [820.75(a)].
When any of the conditions listed below exist, process validation is the only practical means for assuring that processes will consistently produce devices that meet their predetermined specifications:

- Routine end-product tests have insufficient sensitivity to verify the desired safety and efficacy of the finished devices;

- Clinical or destructive testing would be required to show that the manufacturing process has produced the desired result or product.

- Routine end-product tests do not reveal all variations in safety and efficacy that may occur in the finished devices.

- The process capability is unknown, or it is suspected that the process is barely capable of meeting the device specifications.
Why Statistics?

- Manufacturers should:
  - Understand the sources of variation
  - Detect the presence and degree of variation
  - Understand the impact of variation on the process and ultimately on product attributes
  - Control the variation in a manner commensurate with the risk it represents to the process and product.
**What is commensurate with the risk?**

<table>
<thead>
<tr>
<th>Decision Based on Sample</th>
<th>Reality in Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accept</td>
</tr>
<tr>
<td></td>
<td>Correct Decision</td>
</tr>
<tr>
<td>Accept</td>
<td>Correct Decision</td>
</tr>
<tr>
<td></td>
<td>Type II Error (β)</td>
</tr>
<tr>
<td></td>
<td>Consumer Risk</td>
</tr>
<tr>
<td>Reject</td>
<td>Type I Error (α)</td>
</tr>
<tr>
<td></td>
<td>Producer Risk</td>
</tr>
<tr>
<td></td>
<td>Correct Decision</td>
</tr>
</tbody>
</table>
Identify Critical Quality Attributes

- Critical quality attributes are usually the responses from the process (i.e. tensile strength)

- Do not want to select too many or too few.
  - Usually 3-5 is a representative set.

- Early studies help to determine the best set.
Identify Critical Process Factors

- Critical process factors are usually the inputs of the process (i.e. temperature)

- Use screening designs to reduce the trivial many to the vital few.

- Characterize the design space for these critical process parameters.
Focus of Characterization

- Need to know what happens to the process when conditions occur which “stress” the process

- Use “Worst-case” conditions to “challenge” the process and study how the output changes.

- This can help to identify measures that will protect the process as process changes occur.
Types of Designs

- **Plackett-Burman Screening Design**
  - Evaluate $n$ parameters in as little as $n+1$ batches.

- **Fractional Factorial Design**
  - Allow for main-effect only models and some two-way interactions.
Screening Design is based on a Risk Analysis

- Develop a list of all possible Critical Process Parameters by considering discrete and continuous parameters.

- Evaluate each parameter for potential impact on the entire process.

- Perform a detailed review of each parameter to assess relative importance.

- Assign ranges for parameter evaluation based on process Risk Analysis.
Randomization of treatments (runs) helps to distribute the random error over the entire experiment.

When doing a fractional factorial, you need to be careful to assign factors to A, B, etc to control for interactions that are under consideration for analysis.
When doing a fractional factorial interactions that are not tested directly are assumed to be negligible. This is required since these interactions are confounded with other interactions or main effects.

Should you spend more runs identifying main effects or interactions?
Why Analyze Measurement Systems?

**QUANTIFY** the amount of observed process variation due to error in the measurement system.

**EVALUATE:**
- new measurement devices
- suspected deficient measurement devices
- impact of repairs on measurement devices
- new operators involved in the measurement process

**ASSURE** no measurement system error will lead to false acceptance or rejection of product.
Design Verification confirms that Design Outputs meet Design Input requirements, Design Verification confirms “You Made it Right”

Design Validation confirms that the product fulfills the user needs and intended use. Design Validation confirms “You Made the Right Thing”
Utilization of statistics in the design verification and design validation has many different motivations.

- As a level of confidence that risks are appropriately mitigated as a result of the design and development process.
- As an input to risk management, statistics can be used to formulate probabilistic models.

The confidence level we must asserted through the use of statistics are influenced by several factors:

- Risk
- Variability
Statistical Justification

- Document if feature is statistical or non-statistical.
  - There should be written criteria for when a test is considered non-statistical versus it is considered statistical.
  - Such criteria generally consists of a list of the types of tests considered non-statistical.
  - Further, requirements that are statistical but for which recognized standards exist specifying a sample size and acceptance criteria, use of a recognized standard is a valid justification.
Sampling plans consist of both sample size and acceptance criterion.

- Associated with passing a sampling plan is a confidence statement one can make.
- A rationale must exist in writing for the sampling plan selected.
- This rationale should be based on risk to patient.

The phrase “95%/90%” means 95% confidence (Type I error rate of 5%) of having 90% reliability or conformance to requirements (Type II error rate is 10%).
Sample Size for Continuous Data

\[ n = \frac{\left( Z_\alpha + Z_\beta \right)^2 S^2}{\Delta^2} \]

- You can substitute the t-value for smaller sample sizes.
- Differs slightly from the method used by software such as JMP™.
Example

- Type I error rate is 5% (95% Confidence) $Z_\alpha = 1.96$

- Type II error rate is 1% (99% Power) $Z_\beta = 2.326$

- Unknown Standard Deviation and Delta, but want to detect a 0.5 standard deviation shift.

- 2-Sided Specification

$$n = \frac{(1.96 + 2.326)^2 (1)^2}{0.5^2} = 74$$
# Values of Alpha and Beta

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>$\beta$</th>
<th>1-Sided Specification</th>
<th>2-Sided Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>80%</td>
<td>0.842</td>
<td>0.842</td>
<td>1.282</td>
</tr>
<tr>
<td>90%</td>
<td>1.282</td>
<td>1.282</td>
<td>1.645</td>
</tr>
<tr>
<td>95%</td>
<td>1.645</td>
<td>1.645</td>
<td>1.960</td>
</tr>
<tr>
<td>98%</td>
<td>1.960</td>
<td>1.960</td>
<td>2.241</td>
</tr>
<tr>
<td>99%</td>
<td>2.326</td>
<td>2.326</td>
<td>2.576</td>
</tr>
<tr>
<td>100%</td>
<td>2.576</td>
<td>2.576</td>
<td>2.807</td>
</tr>
</tbody>
</table>
Using tolerance intervals

- Determine the Confidence Level
- Determine the percent of population coverage (Reliability)
- Use a ISO 16269-6 (ISO standard for tolerance intervals) to find the value of K

\[
\overline{X} + k \times S < U
\]

\[
\overline{X} - k \times S > L
\]
How large a sample do I need to have 95% confidence with 95% reliability for a process with a mean of 10.0 and standard deviation of 0.55?

The specification are 8.5 to 12.0

Since the mean is closest to the lower specification

From ISO 16269-6, a sample size is approximately 21

\[ 10 - k \times 0.55 > 8.5 \]
\[ k = 2.727 \]
Sample Size (based on all tests being acceptable)

At 95% Confidence and 90% Reliability

\[ n = \frac{\ln(1 - \text{Confidence})}{\ln(\text{Reliability})} = \frac{\ln(1 - 0.95)}{\ln(0.90)} = 28 \]

Does not mean 28 devices, but 28 tests (can be on 1 device depending on variability between devices)
Setting Validation Acceptance Criteria

- Confidence Intervals
- Tolerance Intervals
- Process Capability
A Confidence Interval represents a range of values that will most likely include the true population value.

The Confidence Interval is derived from the observed sample.

The width of the interval is based on:
- The error of our estimate
- A multiplier reflecting how sure we want the interval to include the true population value

The interpretation of the Confidence Interval is always the same:

We are X% confident that the interval includes the true population value
Tolerance Intervals

- Tolerance Intervals are Interval Estimates for individual values in the population.

- Statement about confidence of coverage.
  - i.e. 95% confident that 95% of the individual values will be within the interval.
Process Capability

- Can set a minimum Cp (Pp) or Cpk (Ppk) for the process.

- Usually need to use confidence intervals on these estimates to set acceptance criteria and sample size.
Example

- We collected 20 results for tensile strength of a new material used in a medical device.
  - Mean = 8.51N
  - Std Dev. = 0.697
  - LSL = 5N
  - Min = 7.4
  - Max = 9.5
Example

- Want to test if the **MEAN** tensile strength is greater than 5N.
  - The 95% Lower Bound on 8.51N is 8.24N. This implies if we sample 20 units from a population whose mean is 8.51N with SD=0.697, the average will be greater than 8.24N 95% of the time.

- Want to test if the **INDIVIDUAL** tensile strength is greater than 5N.
  - The 95% confidence with 99% coverage (reliability) is 6.21N. This implies if we sample 20 units from a population whose mean is 8.51N with SD=0.697, the 99% of the individual tensile strength values would be greater than 6.21N.
Other Tools for Validation

◆ The following is a non-exhaustive list of tools:

- Failure Mode Effects Analysis (FMEA)
- Failure Mode, Effects and Criticality Analysis (FMECA)
- Fault Tree Analysis (FTA)
- Hazard Analysis and Critical Control Points (HACCP)
- Hazard Operability Analysis (HAZOP)
- Preliminary Hazard Analysis (PHA)
References

- **GHTF SG3**: Quality Management Systems – Process Validation Guidance

- The FDA Quality Systems Regulation (QSR)
Conclusions

- Validation is required when a process cannot be fully verified.

- Statistical DOE and other methods are used to verify and validate processes.

- Sample size can be calculated that gives a high probability of meeting the specification.
Time for Questions