Role of Statistics in Process Validation

Harry Yang, Ph.D.
Senior Director
MedImmune, LLC

June 23, 2015
Outline

◆ **Process Validation Fundamentals**
  - Life-cycle approach to process validation
  - Build quality into process through design
  - Statistics – enabler of regulatory compliance of process validation

◆ **A Statistics-based Approach to Process Validation**
  - How to establish acceptance criteria
  - How to determine number of validation batches
  - How to estimate inter- and intra-batch variations
  - How to develop effective sampling plans
  - How to trend data to ensure continued process verification
FDA Process Validation Guidance Ushers in A Life-Cycle Approach

*Process Validation:*

The collection and evaluation of data, from the process design stage through commercial production, which establishes scientific evidence that a process is capable of consistently delivering quality product.
Quality by Design

◆ A significant shift of regulatory requirements from the traditional “test to compliance” at the end of process development to the current “quality by design” throughout the life cycle of the product and process.
Three-Stage Validation

- **Stage 1 – Process design**: The commercial process is defined during this stage based on knowledge gained through development and scale-up activities.

- **Stage 2 – Process qualification**: The Process Design is evaluated to determine if the process is capable of reproducible commercial manufacturing.

- **Stage 3 – Continued process verification**: Ongoing assurance is gained during routine production that the process remains in a state of control.
The Golden Three-Batch-Validation

- No longer the norm
- Manufacturer has the responsibilities to provide assurance that the process is adequately qualified
- The use of statistical methods is strongly recommended.
Statistics in Process Validation

◆ **Stage 1 – Process Design**
  – Identification of critical quality attributes thru risk analysis
  – Determination of design space based on statistically designed experiments (DOE)

◆ **Stage 2 – Process Qualification**
  – Acceptance criteria
  – Number of validation batches
  – Sampling plan
  – Estimation of inter- intra-batch variations

◆ **Stage 3 – Continued Process Verification (CPV)**
  – Data trending, SPC, process capability analysis
Importance of Statistics in Process Validation

◆ Statistics mention 15 times
  – “statistical”
  – “statistics”
  – “statistically”
  – “statistician” – as a suggested team member

◆ Clear that FDA expects more statistical thinking in validation
High Degree of Assurance

◆ Phrase “high degree of assurance” mentioned four times
◆ “…the PPQ study needs to be completed successfully and a high degree of assurance in the process achieved before commercial distribution of a product.” – FDA Guidance on PV

◆ “A documented program that provides a high degree of assurance that a specific process, method, or system will consistently produce a result meeting pre-determined acceptance criteria.” – ICH Q7A GMP for APIs:

![Image of a cat and a mirror]

MedImmune
Statistical Acceptance Criteria for Validation

Provide X% confidence that the requirement has been met

Requirements: Process performance to consistently meet acceptance criteria related to identity, strength, quality, purity, and potency

Statistical confidence required may be based on…

- Risk analysis
- Scientific knowledge
- Criticality of attribute
- Prior / historical knowledge
Three Common Situations

Provide statistical confidence that...

- A high percent of the population is within specification
- A population parameter is within specification
  - Mean; Standard Deviation; RSD; Cpk/Ppk
- A standard test (UDU, Dissolution, etc.) will pass
Common Statistical Tools for Setting Specifications

- Confidence, prediction and tolerance interval
Choice of Interval Method to Set Spec Needs to Reflect Quality Attributes Being Tested and Reported

◆ Confidence Interval
  – When reportable value is average

◆ Prediction interval
  – When reportable value is single observation

◆ Tolerance interval
  – When reportable value is % of conformance
Interval Approach

- When data are normally distributed…
Tolerance Interval

- **Confidence and coverage statement**
  - We are 95% confident that 95% of batches will have potency between A and B
What Is Tolerance Interval?

- Example of a tolerance interval statement:
  - “I am 90% confident that at least .95 of my data are in the interval (93, 106)”
Tolerance Intervals Are of the Form
(\( \bar{X} - k*s, \bar{X} + k*s \))

- Assuming a normal distribution
- \( K \) is based on 3 parameters
  - \( c \) (proportion or coverage)
  - \( p \) level of confidence
  - \( n \) sample size
- \( \bar{X} \) is sample mean
- \( s \) is sample standard deviation

\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}
\]

\[
s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}
\]
Table 1: Factors $k$ for Two-Sided Tolerance Intervals for Normal Distributions (Based on an Approximation in Introduction to Statistical Tolerance Analysis, McGraw-Hill, New York. Dixon and Massey (1969))

<table>
<thead>
<tr>
<th>Percent Contained</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
<th>90%</th>
<th>95%</th>
<th>99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>2.625</td>
<td>3.128</td>
<td>4.111</td>
<td>2.967</td>
<td>3.555</td>
<td>4.646</td>
<td>3.822</td>
<td>4.554</td>
<td>5.985</td>
</tr>
<tr>
<td>11</td>
<td>2.463</td>
<td>2.935</td>
<td>3.857</td>
<td>2.737</td>
<td>3.261</td>
<td>4.286</td>
<td>3.397</td>
<td>4.047</td>
<td>5.319</td>
</tr>
<tr>
<td>16</td>
<td>2.246</td>
<td>2.676</td>
<td>3.517</td>
<td>2.437</td>
<td>2.904</td>
<td>3.816</td>
<td>2.872</td>
<td>3.422</td>
<td>4.497</td>
</tr>
<tr>
<td>17</td>
<td>2.219</td>
<td>2.644</td>
<td>3.474</td>
<td>2.399</td>
<td>2.859</td>
<td>3.757</td>
<td>2.808</td>
<td>3.346</td>
<td>4.398</td>
</tr>
<tr>
<td>19</td>
<td>2.172</td>
<td>2.588</td>
<td>3.402</td>
<td>2.336</td>
<td>2.784</td>
<td>3.659</td>
<td>2.703</td>
<td>3.221</td>
<td>4.233</td>
</tr>
<tr>
<td>21</td>
<td>2.134</td>
<td>2.543</td>
<td>3.343</td>
<td>2.286</td>
<td>2.724</td>
<td>3.579</td>
<td>2.620</td>
<td>3.122</td>
<td>4.102</td>
</tr>
<tr>
<td>22</td>
<td>2.118</td>
<td>2.524</td>
<td>3.317</td>
<td>2.264</td>
<td>2.707</td>
<td>3.545</td>
<td>2.584</td>
<td>3.079</td>
<td>4.046</td>
</tr>
<tr>
<td>23</td>
<td>2.103</td>
<td>2.506</td>
<td>3.294</td>
<td>2.244</td>
<td>2.673</td>
<td>3.514</td>
<td>2.551</td>
<td>3.040</td>
<td>3.995</td>
</tr>
<tr>
<td>25</td>
<td>2.077</td>
<td>2.475</td>
<td>3.252</td>
<td>2.208</td>
<td>2.631</td>
<td>3.458</td>
<td>2.494</td>
<td>2.972</td>
<td>3.906</td>
</tr>
<tr>
<td>30</td>
<td>2.025</td>
<td>2.413</td>
<td>3.171</td>
<td>2.140</td>
<td>2.550</td>
<td>3.351</td>
<td>2.385</td>
<td>2.842</td>
<td>3.734</td>
</tr>
<tr>
<td>50</td>
<td>1.916</td>
<td>2.284</td>
<td>3.001</td>
<td>1.996</td>
<td>2.379</td>
<td>3.126</td>
<td>2.162</td>
<td>2.576</td>
<td>3.385</td>
</tr>
<tr>
<td>60</td>
<td>1.887</td>
<td>2.248</td>
<td>2.955</td>
<td>1.958</td>
<td>2.333</td>
<td>3.086</td>
<td>2.103</td>
<td>2.506</td>
<td>3.293</td>
</tr>
<tr>
<td>70</td>
<td>1.865</td>
<td>2.222</td>
<td>2.921</td>
<td>1.929</td>
<td>2.299</td>
<td>3.021</td>
<td>2.060</td>
<td>2.454</td>
<td>3.225</td>
</tr>
<tr>
<td>80</td>
<td>1.848</td>
<td>2.202</td>
<td>2.894</td>
<td>1.907</td>
<td>2.272</td>
<td>2.986</td>
<td>2.026</td>
<td>2.414</td>
<td>3.173</td>
</tr>
<tr>
<td>90</td>
<td>1.834</td>
<td>2.185</td>
<td>2.872</td>
<td>1.889</td>
<td>2.251</td>
<td>2.958</td>
<td>1.999</td>
<td>2.382</td>
<td>3.130</td>
</tr>
<tr>
<td>100</td>
<td>1.822</td>
<td>2.172</td>
<td>2.854</td>
<td>1.874</td>
<td>2.233</td>
<td>2.934</td>
<td>1.977</td>
<td>2.355</td>
<td>3.096</td>
</tr>
<tr>
<td>∞</td>
<td>1.645</td>
<td>1.960</td>
<td>2.576</td>
<td>1.645</td>
<td>1.960</td>
<td>2.576</td>
<td>1.645</td>
<td>1.960</td>
<td>2.576</td>
</tr>
</tbody>
</table>
Opportunities to Reduce Risk

- Lower variability yields lower risk

- Variability is controlled by design
Interpretation of Tolerance Intervals

• Tolerance intervals are about **future results**.
  – Not past means or performance estimates

• Two sets of elements drive the location and the width of the Tolerance intervals
  1. The performance estimates (mean, SD, ...)
  2. The **uncertainty** of those estimates!!
Existing Information

• There is valuable knowledge and data to reduce the uncertainty of the estimates of performance at PPQ.
  – Clinical Batches
  – Development Batches
  – Validation analytical methods

• Scientific/technical Knowledge

Can the number of required PPQ batches be reduced by taking advantage of existing information?
Bayesian Principle

“PRIOR DISTRIBUTION” from previous studies, expert opinion, literature,...

Available Data + Observed Data = Total Data

“LIKELIHOOD” data coming from the experiment

“POSTERIOR DISTRIBUTION” combination of information collected before the experiment and what comes from the experimental data
Bayesian Principle (Cont’d)

• PPQ batches are produced to collect evidence of the quality of the process
  
  – *Frequentist analysis*:
    • point estimate and confidence intervals as summaries of process (mean and sd)
      ➔ what do PPQ batches tell us about the process?
  
  – *Bayesian analysis*:
    • Before the PPQ: a priori opinion on the process
      ➔ how should those PPQ batches *change our opinion* about the process?
      ➔ how should those PPQ batches *provide assurance* about future batches?
Comparison between Bayesian and Frequentist Methods

Use the Predictive distribution to compute the probability to be in specifications.

Bayesian statistics allows computing a probability instead of a Tolerance Interval only.

⇒ What’s the risk?
Bayesian Approach to Determining Number of Validation Batches

- Traditionally 3 batches
- Life-cycle approach
  - The more confidence one has based on stage 1 data, the fewer validation batches needed

<table>
<thead>
<tr>
<th>Process</th>
<th>Model Parameters $\alpha, \beta$</th>
<th>Number of Validation Batches Needed $(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(1.5, 25)$</td>
<td>480</td>
</tr>
<tr>
<td>2</td>
<td>$(25, 1.5)$</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>$(1, 1)$</td>
<td>18</td>
</tr>
</tbody>
</table>
Important Distinction – Number of Validation Batches ≠ Sample Size

Variance Components

Total process

Within batch

Between batch

Total Variation
How to Estimate Inter- and Intra-batch Variations

Variance Components Model

Total process variation

Within batch

Between batch

Batch 1  Batch 2  Batch 3  Batch 4  Batch 5

Total Variation
Variance Component Analysis (VCA)

- VCA is most commonly used to determine inter- and intra-batch variations
- A typical experiment might select several batches, several samples from each batch, and then run replicates tests on each sample
How to Establish Effective Acceptance Sampling Plans

“*The increased level of scrutiny, testing, and sampling should continue through the process verification stage as appropriate, to establish levels and frequency of routine sampling and monitoring for the particular product and process.*“
What is Acceptance Sampling?

- A form of inspection conducted to determine if a lot or batch of items after a process conforms to pre-specified standards
- A lot is defined as a quantity of product accumulated under uniform conditions.
Attributes and Variables

Sampling for Attributes vs. Variables.

- **Attribute** sampling is a Pass/Fail Decision (ex. container bursts when filled or does not burst).
- **Variable** sampling is a continuous measurement of a quality characteristic (ex. pH, weights, viscosity).
Acceptance sampling is useful

◆ When:
  – Testing is destructive
  – The cost of 100% inspection is very high
  – 100% inspection takes too long
Single Sampling

Select Sample

Inspect Sample

Make Decision

The decision to accept or reject the lot is made based on the results of a single sample.
Single Sampling (Cont’d)

- Items are selected at random from a lot
- Disposition of the lot is determined, based on
  - Resulting information
  - Pre-specified standards
- Operationally simple but not most efficient
Double Sampling

- Select Sample
- Inspect Sample
- Is sample a clear accept or reject?
  - Yes: Make Decision
  - No: Resample
    - Inspect Resample

If there is not a clear accept/reject decision from inspection of the 1\textsuperscript{st} sample, a 2\textsuperscript{nd} sample is taken and the decision is made based in the results of the combined samples.
Double Sampling Plans

- After the first sample is tested, there are three possibilities:
  - Accept the lot
  - Reject the lot
  - No decision

- In case of “No decision”, a second sample is taken, and a final decision is made, using combined test results
Sampling Plan Based on Decision Theory

<table>
<thead>
<tr>
<th></th>
<th>Good Lot</th>
<th>Bad Lot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept</td>
<td>L0</td>
<td>L1</td>
</tr>
<tr>
<td>Reject</td>
<td>L2</td>
<td>L3</td>
</tr>
</tbody>
</table>

Choose plan to minimize the overall risk
Which plan to use depends on

- Cost of test
- Time
- Sample size
- Balance between reduced cost and operational complexity
Definitions

◆ AQL – quality level that is the worst tolerable process average when a continuing series of lots is submitted for acceptance sampling. Value refers to having a 95% or 99% confidence that lots above the specific percent nonconforming will be rejected.

◆ Producers Risk – probability of rejecting a good lot

◆ Acceptance # - maximum number of defective units of defects in a sample that will permit acceptance of the inspection lot.

◆ Rejection # - minimum number of defects or defective units in the sample that will cause the lot or batch to be designated as unacceptable.
Definitions

◆ $P_a$ – probability that a lot will be accepted under a given sampling plan.

◆ LTPD – level of % nonconforming such that there is only a 5% or 10% probability of accepting the lot or lots when the % nonconforming is that great or greater.

◆ Operating Characteristic curve – a curve showing, for a given sampling plan, the probability of accepting a lot as a function of the lot quality. For a continuous stream of lots, a curve showing the probability of accepting a lot as a function of the process average.
# Representative Sampling Error Matrix

<table>
<thead>
<tr>
<th>The Decision Made</th>
<th>Lot Quality</th>
<th>Good</th>
<th>Bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Called Good</td>
<td>1-α, Producer’s Confidence</td>
<td>B Type II Error</td>
<td></td>
</tr>
<tr>
<td>Called Bad</td>
<td>A, Type I Error</td>
<td>1-β, Consumer’s Confidence</td>
<td></td>
</tr>
</tbody>
</table>
Representative Sampling Plans

◆ For New Process
  – Based on Risk Assessment
    • Identify AQL and LTPD
    • Set up sampling plans and associated acceptance criteria accordingly.

◆ Sampling risks should be known and compatible with Consumers risks.

◆ Quality Index Chosen (AQL, LTPD, etc.) should reflect respective needs of both producer and consumer.
Operating Characteristic Curves (OC)

- Sample Size and Accept/Reject number uniquely determine attributes sampling plan.
- OC curve is developed by determining the probability of acceptance for each of several values of lot quality.
- Several distributions used to determine OC curves.
  - Binomial, Poisson, Hypergeometric.
Operating Characteristic (OC) Curve

- Ideal OC curve
- Typical OC curve

Probability of acceptance

- $\alpha$
- $\beta$

AQL               CQL

Proportion defective
Ideal OC Curve

◆ Discriminates perfectly between good and bad lots
  – $\alpha = \beta = 0$
  – Corresponds to 100% inspection
Operating Characteristic Curve

![Operating Characteristic Curve Image]

- Probability of acceptance
- Proportion defective (hundredths)
- (AQL) = 0.122
- (CQL) = 0.126

Probability of acceptance vs. Proportion defective (hundredths) graph showing points at 0.878, 0.663, 0.463, 0.308, 0.199, 0.126, 0.078, 0.048, 0.029, 0.017.
Constructing an OC Curve

<table>
<thead>
<tr>
<th>Proportion defective (p)</th>
<th>np</th>
<th>Probability of c or less defects (Pa)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01 (AQL)</td>
<td>0.6</td>
<td>0.878</td>
<td>α = 1.000 – 0.878 = 0.122</td>
</tr>
<tr>
<td>0.02</td>
<td>1.2</td>
<td>0.663</td>
<td></td>
</tr>
<tr>
<td>0.03</td>
<td>1.8</td>
<td>0.463</td>
<td></td>
</tr>
<tr>
<td>0.04</td>
<td>2.4</td>
<td>0.308</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>3.0</td>
<td>0.199</td>
<td></td>
</tr>
<tr>
<td>0.06 (CQL)</td>
<td>3.6</td>
<td>0.126</td>
<td>β = 0.126</td>
</tr>
<tr>
<td>0.07</td>
<td>4.2</td>
<td>0.078</td>
<td></td>
</tr>
<tr>
<td>0.08</td>
<td>4.8</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>0.09</td>
<td>5.4</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>6.0</td>
<td>0.017</td>
<td></td>
</tr>
</tbody>
</table>

n = 60

C = 1

Proportion defective (hundredths)

Probability of acceptance
How do we find $\alpha$ and $\beta$ using an OC curve?

- AQL = 0.01
- CQL = 0.05

Then $\alpha = 1 - P_a(p=0.01) = 1 - 0.978 = 0.122$

And $\beta = P_a(p=0.06) = 0.20$
How to Trend Data to Ensure Continued Process Verification
The Two Kinds of Variation

◆ Variation in industrial processes falls into one of two categories:

◆ Common cause variation
  – Random in nature
  – Happens all the time
  – Relatively small in magnitude

◆ Special Cause variation
  – Has an assignable cause
  – Usually rare
  – Relatively large in magnitude
Common Cause Variation

- This type of variation occurs in all processes - it cannot be eliminated without making fundamental changes in the process.

- Attempting to adjust the process to reduce common cause variation is tampering - this actually increases variation from target.
Popularity of Control Charts

◆ A proven technique for improving productivity
◆ Effective in defect prevention
◆ Prevent unnecessary process adjustment
◆ Provide diagnostic information
◆ Provide information about process capability
Types of Control Charts

◆ **Variables control chart**
  - These charts are applied to data that follow a continuous distribution.
  - I-MR chart
  - X-bar and R control charts
  - X-bar and S control charts

◆ **Attributes control chart**
  - These charts are applied to data that follow a discrete distribution.
  - P chart and nP chart
  - U chart and C chart
Types of Control Chart

Source: Control Charts for Minitab 15
I-MR Chart

◆ The I-MR chart (x-chart) monitors the mean and the variation of a process. Individual or x-charts should be used when there is only one data point to represent a situation at a given time.
X-bar Range (X-bar-R) Chart

- X-bar-R is the averages of a set of measurements taken at the same time
- It uses the range to estimate the process variation
X-bar-Standard Deviation (X-bar-S) Chart

- As the subgroup size increases, the standard deviation is an increasingly better estimator of the process variation than the range.
- The Xbar-S chart monitors the mean and the variation of a process.
Attribute Control Chart

◆ Attribute control charts are for variables which follow discrete distribution
  – defective units or defects per unit
Defects And Defective Units

◆ Defects are flaws, such as scratches, dents or bumps on the surface of a car. A part may have more than one defect, and the defects do not necessarily make the part unacceptable. You can count the defects over a length of time, over an area, or over a set number of items.

◆ A defective unit is a part with a flaw so severe that it is unacceptable for use, such as a broken light bulb or cracked bolt.
Defective Units and Defects per Unit

- Defective units are the counts items that are classified into one of two categories such as pass/fail or go/no-go. Often to calculate as a proportion (%defective).

- Defects per unit are counts defects or the presents of undesired characteristics or activities for each unit. Often used to determine an occurrence rate (defects per unit).
The P chart monitors the proportion of defective units.

Exactly like an np chart, but the number of units sampled varies.

The number of units sampled should not depend on the number of defective units observed.
NP Chart

- The NP chart monitors the number of defective units per subgroup when the subgroup size is the same.
- You can also use P chart to minor the proportion of defective units when the subgroup size is constant.
C Chart

- The C chart monitors the total number of defects per subgroup when subgroups are the same size.
U Chart

- The U chart monitors the average defects per unit when subgroups are the same or subgroups of different sizes.
## Summary of Control Charts

<table>
<thead>
<tr>
<th>Control Chart Type</th>
<th>Centerline</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>The average of the single (individual) data values for the baseline chart</td>
<td>x (x bar)</td>
</tr>
<tr>
<td>Moving Range</td>
<td>The average of the moving (or rolling) ranges calculated from the chart data</td>
<td>MR</td>
</tr>
<tr>
<td>X-bar</td>
<td>The average of the subgroup averages calculated from the baseline chart data</td>
<td>x (x double bar)</td>
</tr>
<tr>
<td>Range</td>
<td>The average of the subgroup ranges calculated from the baseline chart data</td>
<td></td>
</tr>
<tr>
<td>np</td>
<td>The average of the np's (number good/bad) in each sample group collected for the chart</td>
<td>np (np bar)</td>
</tr>
<tr>
<td>p</td>
<td>The average of the p's (proportion good/bad) calculated for each sample group in the chart</td>
<td>p (p bar)</td>
</tr>
<tr>
<td>c</td>
<td>The average of the c's (total number of flaws, defects, occurrences, etc.) in each sample group collected for the chart</td>
<td>c (c bar)</td>
</tr>
<tr>
<td>u</td>
<td>The average of the u's (average number of flaws, defects, occurrences, etc., per unit) calculated for each sample group</td>
<td>u (u bar)</td>
</tr>
</tbody>
</table>