Overview of Statistics used in QbD Throughout the Product Lifecycle

August 2014

The Windshire Group, LLC

Comprehensive CMC Consulting™

Presentation format and purpose

- Method name
- What it is used for and/or how
- Key points to consider (PTC)

Audience: those with limited knowledge of statistics.

The purpose is to establish a baseline of understanding to aid in communication and teamwork, and help increase appreciation of the importance of statistics

Overview

- QbD Introduction
- QbD Lifecycle- Major Statistical Approaches
  - Design, Qualification, Verification
- Statistical Overview
  - DOE
  - Chemometrics and Modelling
  - Small Data Sets
  - SPC (statistical process control)
  - Process Capability
  - Sampling/ Stability
- References

❖ + a few STORIES!

QbD Introduction

The Windshire Group, LLC

QbD

- Basis for using QbD:
  - Success in numerous other industries
  - ICH Q8, 9, 10, 11; FDA's "21st Century" initiative, PAT and Process Validation Guidance
- Risk management and statistical techniques form the basis for QbD implementation
- Rationale, focused design on quality attributes:
  - Reduces overall development and manufacturing costs
  - Increases understanding and enhances product quality
  - Reduces the cost of quality
  - Increases compliance

QbD Lifecycle

The Windshire Group, LLC

Lifecycle view

Design

- **Product and process characterization and optimization**: Initial scoping and screening; determine CPPs, develop functional relationships, and determine Design Space
- Major tools: DOE, chemometrics

Qualification

- **Scientific** evidence that a process is capable of **consistently** delivering quality products through collection and analysis of data
- Major Tool: Statistics of small data sets

Verification

- **Manufacturing and process control** through control systems, tracking and trending
- Major tools: SPC, sampling, process capability

Sampling of statistics in "Guidance for Industry. Process Validation: General Principles and Practices"

- "We recommend an **integrated team approach** to process validation that includes expertise from a variety of disciplines (e.g., process engineering, industrial pharmacy, analytical chemistry, microbiology, **statistics**, manufacturing, and quality assurance)."
- "Other CGMP regulations define the various aspects of validation. For example, …the CGMP regulations regarding sampling set forth a number of requirements for **validation: samples must represent the batch under analysis (§ 211.160(b)(3)); the sampling plan** must result in **statistical confidence** (§ 211.165(c)..."
- "…in-process **specifications** ". . . shall be derived from previous acceptable process average and process variability estimates where possible and determined by the **application of suitable statistical procedures** where appropriate."
- "In addition, we strongly recommend firms employ objective measures (e.g., **statistical metrics**) wherever feasible and meaningful to achieve **adequate assurance**."

Design

The Windshire Group, LLC

Basic definitions

- Level of a Factor- the **values** it takes on. Factor is an independent **variable manipulated** by the experimenter
- Replicate- A replicate is the **outcome** of an experiment or observation obtained in course of its replication. Replication is **repetition** of an experiment
- Treatment- something that researchers **administer to experimental units** in Levels for the Factors
- Blocking- A **schedule** for conducting treatment combinations such that any effects on the experimental results due to a known change in, e.g., raw materials, operators, machines, become **concentrated** in the levels of the **blocking variable**

- Blocking Variable- categorical variable which are **not experimental conditions** but are still **included** into the analysis as a **means of statistical control**
- Balanced Design- An experimental design where **treatment** combinations have the **same number of observations**
- Rotatability- A design is rotatable if the **variance** of the **predicted response** at any point x **depends only** on the **distance** of x **from** the **design center point**; design can be rotated around its center point without changing the prediction variance at x. *PTC*: Rotatability is a desirable property for response surface designs

- Effect- How **changing** the **settings** of a **Factor changes** the **response**. The effect of a **single factor** is also called a **main effect**
- Orthogonality: An experimental design whereby **effects of any Factor** balance out (**sum to zero**) **across** the **effects** of the **other Factors**.

DOE (Design of Experiments)

- A **systematic**, rigorous **approach** to that applies principles and techniques at the **data collection** stage so as to ensure the generation of **valid, defensible**, and **supportable conclusions** linking process parameters to quality attributes
  - *PTC*: Avoid false positives from "confirmation bias" or "p-hacking" by **clear statement upfront of methodology** and clear and **complete documentation**
  - Runs should be **randomized to reduce** the impact of **bias**. In practice, this can be a operational **challenge**

- **Identifies which** of many **factors** have a **significant effect** on the response or when it is **sufficient to consider factors at two levels** (e.g., process problem- let's look at pH, temperature, and rpm "a little higher and a little lower")
  - *PTC*: Typically **screening designs** have **more than 5 factors**
  - **Estimate experimental error** in two ways:
    - 1) **Replicate** the **entire experiment-** more common for small experiments
    - 2) For more than 5 factors, just **one replication** (e.g. **center point**)
  - Experimental runs in a factorial experiment should be **randomized** to **reduce** the impact that **bias** could have. In practice, this can be a challenge

Screening designs

- Full factorial designs consists of two or more factors, each factor with discrete possible levels, and whose variables take on **all possible combinations of levels across all factors**
  - *PTC*: **If** the number of **combinations** is **too high** (10 factors at 2 levels = 1024) to be logistically feasible, then **Fractional Factorial** should be **used** (usually ½ the number of combinations)
- Fractional factorial designs at 2 Levels (high/low)
  - **Reduced** number of **combinations** from Full Factorial
  - *PTC*: good for very large number of variables, **e.g. initial medium design**

- D-optimal designs
- Based on **computer generated** algorithms
- Used when:
  - **Resources** and time are **limited**: reduce the costs of experimentation with fewer experimental runs
  - The **design-space is constrained** (e.g., safety, biological limits)
  - Multiple types of factors, e.g., **process, mixture,** and **discrete data** need to be investigated
  - *PTC*: **Always an option** regardless of the type of model (i.e., first order, first order plus some interactions, full quadratic, cubic, etc.) or the objective (for example, screening, response surface, etc.)

Design Space- Response surface designs

- To develop functional relationships and a Design Space, factors that have more than 2 levels need to be examined. There are three steps:
  1. A **factorial** (perhaps fractional) **design** in the factors studied, each having two levels
  2. A set of **center points**, **experimental runs** whose values of each factor are the medians of the values used in the factorial portion (replicated in order to improve the precision)
  3. A set of **axial points**, experimental runs identical to the **center points except for one factor**, which will take on values both below and above the median of the two factorial levels, and typically both outside their range. **All factors** are **varied** in this way

- Central Composite Designs (CCDs)

Circumscribed (CCC)- The star points **establish new extremes** for the low and high settings for all factors. Can augment an existing factorial or fractional factorial design. 5 levels of each factor. *PTC*: Use to **establish AOR or failure limits**

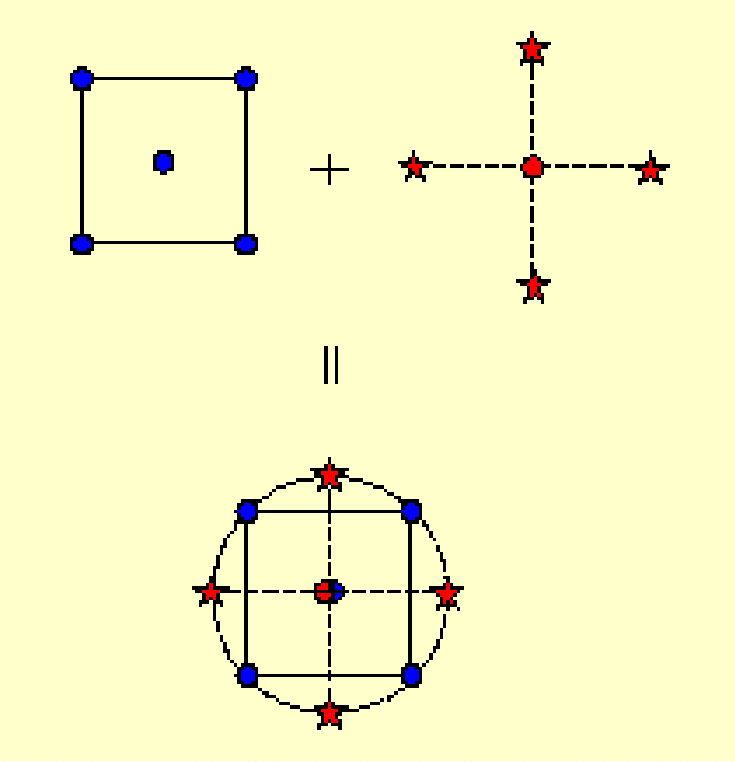
Inscribed (CCI)- The CCI design uses the factor settings as the star points and creates a factorial or fractional factorial design within those limits. Also requires 5 levels of each factor. **PTC**: For situations in which the **limits specified for factor settings are truly limits**.

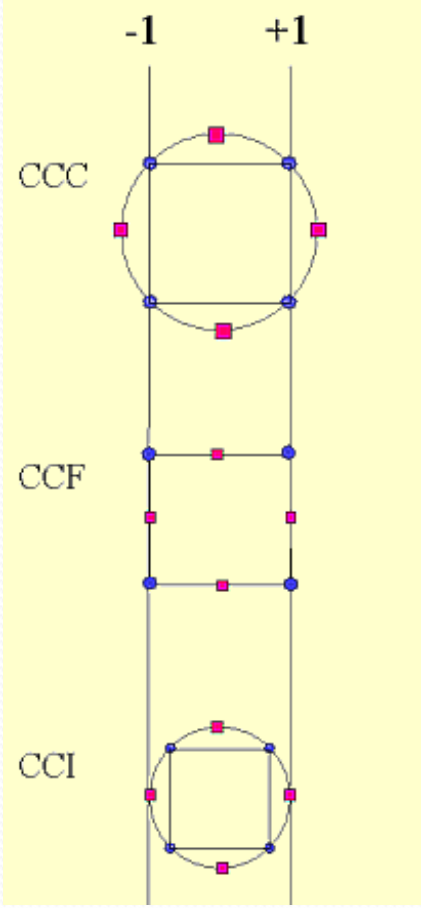
Face Centered (CCF)- The star points are at the center of each face of the factorial space. This variety requires 3 levels of each factor. *PTC*: **Augment an existing factorial or fractional design**

Central Composite Designs

PTC: generally should have 3-5 center points

CCC





Rotatable

Non-rotatable

Rotatable

- Box-Behnken

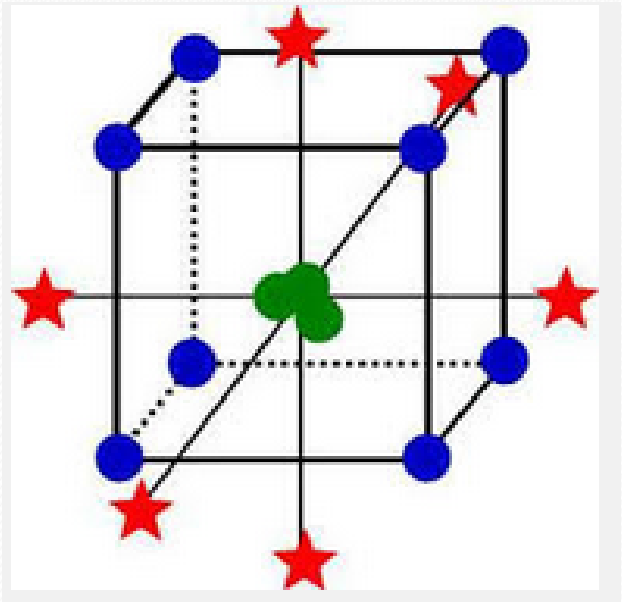
Independent design in that it **does not contain** an **embedded factorial or fractional factorial design**. Treatments are at the **midpoints of edges** of the process space and **at the center**. Requires 3 levels of each factor. Rotatable.

- *PTC*: **explore experimental boundaries** or **avoid** treatment combinations that are extreme (i.e., **corners/stars**). The design has **limited** capability for **orthogonal blocking** compared to the central composite designs,

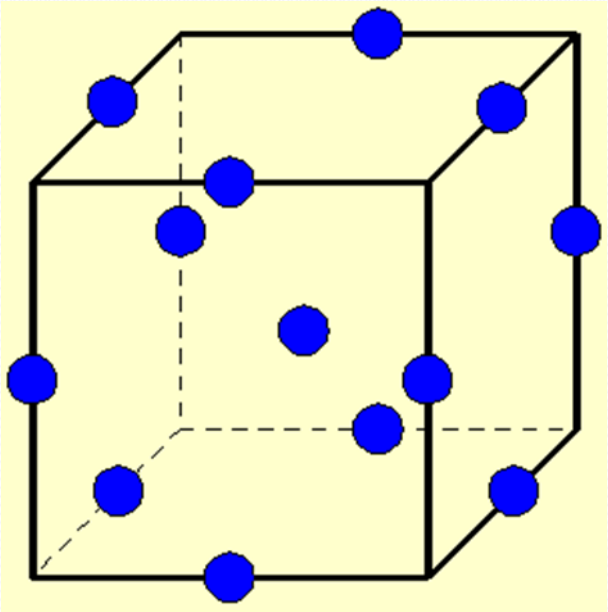
Blocked designs are better designs if they allows the ***estimation of individual and interaction factor effects independently*** *of the* ***block effects****. This condition is called* ***orthogonal blocking****. Blocks are assumed to have no impact on the nature and shape of the response surface*

Central Composite vs Box-Behnken

CC



BB



Chemometrics and modelling

- Chemometrics

Science of **relating measurements** made on a chemical system or process **to the state of the system** via application of mathematical or statistical methods

  - Analytical/Spectroscopic
  - Manufacturing

- Chemometric analysis:
  - Is **empirical**
  - Relates **multivariate data** to single or multiple responses
  - Utilizes **multiple linear regressions** to multivariate data

- Least squares regression analysis

The best fit **minimizes** the **sum of squared residuals**, a residual being the difference between an observed value and the fitted value provided by a model. p-value is statistical **significance** of fit; **R^2** ***coefficient of determination*****-** the "strength" between the two variables

❖ 3 STORIES

PTC: p-value is the most important to pay attention to. Models with R^2 <0.1 may be meaningful. **Think mechanistically**!

- Multiple linear regression
- Linear regression using more than one predictor variable

- Principal components regression (analysis)

Two or more of predictor **variables often have** a lot of **collinearity**. **PCR excludes** some of the related **low-variance** principal components in the regression step and regresses on only a subset of all the principal components. A lower effective number of parameters for the underlying (**simpler**) **model results**

PTC: Very useful for using data from **large or complex data sets,** e.g. manufacturing process data. **Sensitive to process changes and relationships** that traditional SPC is incapable of detecting

- Partial lease squares

Similar to PCR. It finds a **linear regression model** by **projecting** the predicted variables and the observable **variables to a new space to create simpler model**.

- Linear discriminant analysis

Similar to PCR. Finds a linear combination of features which **characterizes or separates two or more classes of objects or events**.

PTC: Example usage, a PAT application for analyte discrimination in real time

Qualification

The Windshire Group, LLC

Setting acceptance criteria*

- **Three basic scenarios** to establish: 1) small data sets around central point conditions; 2) larger data sets; 3) large data sets possible to accurately model the impact of process conditions on performance of the step
- Statistical approaches include: **mean ± 3SD, tolerance interval analysis, prediction profiler,** and **Monte Carlo simulation**. Strengths and shortcomings for each

PTC: **Selection** of the **right statistical approach** is the **first step** toward setting **appropriate acceptance criteria**

See Reference slide

Conformance run process consistency*

- Weisberg t-test outlier test

Key performance parameters, acceptance criteria for specific attributes must be defined and met. However, for **non-key performance parameters**, outlier test can be used to **claim data is *consistent*** for process validation purposes

PTC: Very good test for small data sets. For a given alpha (risk for false positive, e.g. 0.05), the **beta (risk for false negative)**, is **very desirable** and better than other tests

See Reference slide

Verification

The Windshire Group, LLC

Process monitoring

- Control charts

Control charts monitor data from measurements of variations at points on the process map. Control charts are a tool to help **differentiate "assignable" ("special")** sources of variation from **"common" sources**. "Common" sources are an expected part of the process

PTC: **Most useful for early detection and prevention of problems**, rather than the correction of problems after they have occurred. Needs support system and trained personnel for continuous monitoring and alerting

Western Electric rules for detecting instability

The Basic Four rules:

1. Any **single data** point falls **outside the 3σ** limit from the centerline
2. **Two out of three consecutive** points fall **beyond the 2σ** limit
3. **Four out of five** consecutive points fall **beyond the 1σ** limit
4. **Nine consecutive points** fall on the **same side** of the centerline

PTC: Even normal, **"common"** source variability **can trigger** rules. When control limits are not symmetrical and for small subgroups (e.g. qualification runs) use R charts or p-control charts

Process capability

- Measured by process capability index or process capability ratio- the **ability** of a process **to produce** output **within specification** limits.

$$\hat{C}_{pk} = \min\left[\frac{USL - \hat{\mu}}{3\hat{\sigma}}, \frac{\hat{\mu} - LSL}{3\hat{\sigma}}\right]$$

$\hat{\sigma}$ = standard deviation

$\hat{\mu}$ = mean

  - *PTC*: Equations exist for process capability **around a target** and can account for an off-center process
  - **Most useful on a stable process** to predict the ability of the process to produce "conforming product" in the future
  - ❖ STORY

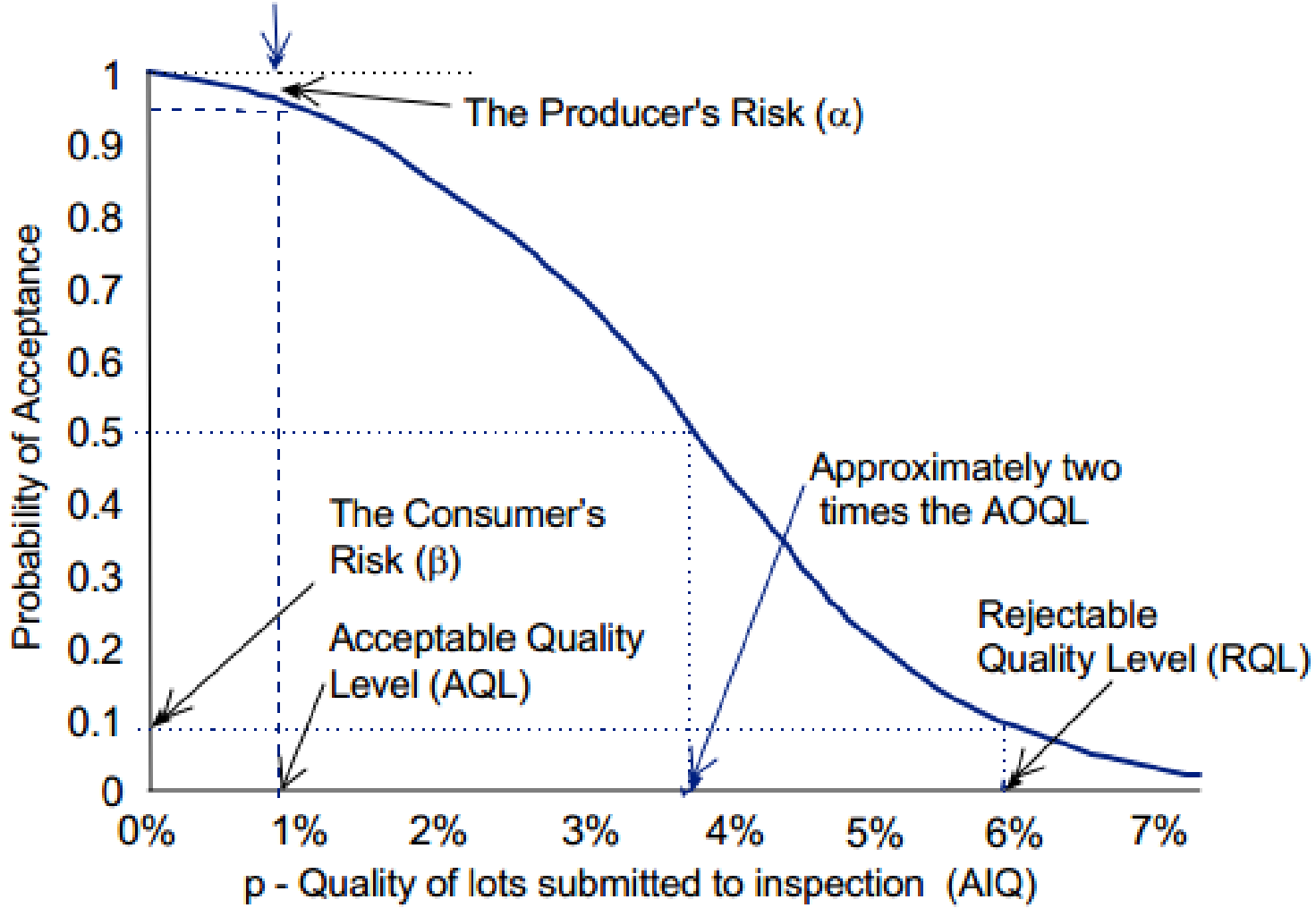
AQL Sampling

- AQL: 'Acceptance Quality Limit'

  Commonly used standards (there are others)

  - Inspection by variables for Percent Nonconforming, **ANSI**/ASQ Z1.9-2008; Inspection by attributes, ANSI/ASQ Z1.4-2008
  - The **producer's risk**, is alpha, for a given AQL and is the **probability** that a batch at AQL limit will be **rejected**. Or worst quality to accept on a regular basis
  - The **consumer's (patient's) risk** is denoted as beta. The beta risk is the **probability** that a batch at a 'Reject Quality Limit' RQL limit will be **accepted**. Or the quality that you want to reject on a regular basis

OC- Operating Characteristic curve



- *PTC*: The **FDA** and product companies focus primarily on **AQL** (alpha risk), however both kinds of risk need to be considered and understood

Generally, a steep OC curve is desired and **steeper** can only be achieved with **more samples inspected**

But, **> inspection** comes at a **cost** (100% inspection is vertical)

One way to reduce cost of inspection is a **Double Sampling Plan** consists of two sets of Acceptance Numbers, Rejection Numbers and Sample Sizes. If the lot was accepted or rejected from first inspection, the second inspection is not required. If it is, combine the total number of defective items (1st & 2nd). If equal to or greater than the Rejection Number, reject the lot.

- The Average Outgoing Quality Limit (AOQL) of a sampling plan is maximum value on the AOQ curve. The AOQ curve gives the average outgoing quality (y axis) as a function of the incoming quality (x axis). The **AOQL is the worst possible defect rate** for the **average outgoing quality regardless of the incoming quality**; the defect rate going to the customer should be no greater than the AOQL over an extended period of time
  - *PTC*: Assumes rejected lots are 100% inspected (i.e., only applicable for 100%) and inspections are effective

❖STORY

Drug Substance and Drug Product Stability

- Linear regression
- Used for **extrapolation of or verifying retest periods or shelf lives**, or potential future lot failures
  - Most common- determine the earliest time at which the **95 percent confidence limit** for the mean **intersects** the proposed or actual **acceptance criterion**
- Analysis of covariance (ANCOVA)
- Used to **test poolability of batches** to use in the regression analysis. Null hypothesis- slopes are equal. Use a **significance level of 0.25** to compensate for the expected low power of the design due to the relatively limited sample size (ICH Q1A(R2)

- *PTC*: Qualitative attributes and microbiological attributes are not amenable to this kind of statistical analysis
- Validated software should be used to support registration claims or to make ongoing lot acceptability decisions

References

- http://www.variation.com/ Dr. Wayne Taylor, good books, software, other info about statistics
- http://statpages.org/index.html Mega-web site, free software, other web resources, etc.
- http://www.itl.nist.gov/div898/handbook/index.htm Statistic handbook
- http://www.statistics.com/glossary/ Definitions
- Weisberg t-test: *BioPharm International, May 2003. Demonstrating the Consistency of Small Data Sets: Application of the Weisberg t-test for Outliers. Sealy R., et al.*
- Qualification (Validation) Acceptance Criteria*: Biotechnol Prog. 2007 Jan-Feb;23(1):55-60.Using statistical analysis for setting process validation acceptance criteria for biotech products. Wang X1, et al.*

Thanks!

Future questions?

The Windshire Group, LLC

James Blackwell, Ph.D. M.B.A.

President

Email: **info@windshire.org**

We would be delighted to hear from you!

The Windshire Group, LLC