Welcome to BEBPA's

Sneak Peek Webinar

Confidence Intervals & Outliers

Begins at 5:00 PM PT

BEBPA's Asia Bioassay Sneak Peek Webinar

Welcome & Introduction

Nancy Niemuth, M.S.

Statistics Interest Group
Scientific Committee
BEBPA

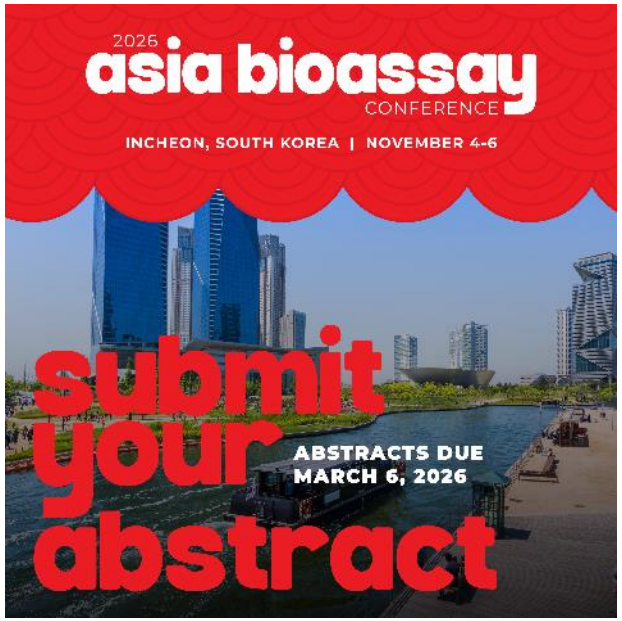
Ohio, USA

Who is BEBPA?

Welcome to the Biopharmaceutical Emerging Best Practices Association (BEBPA), a non-profit organization dedicated to advancing patient safety through the collaborative efforts of the biopharmaceutical scientific community.
www.bebpa.org

- 3+ Annual Hybrid Events
- 4+ Annual Virtual Events
- White Papers
- Free Scientific Webinars
- Technical Blog
- Journal Club
- Scientific Interest Groups
- Student Scholarships

BEBPA Social Media



Please **do not** take photos or screenshots of any BEBPA slides!

Thank You!

Asia Bioassay Conference

November 4-6, 2026 | Incheon, South Korea

$200 OFF Registration!

Discount Code: ABAWEBINAR

Interval Training: An Introduction to Confidence Intervals and other Statistical Intervals

Nancy Niemuth, M.S.

Act Two Consulting LLC

August 28, 2026

BEBPA Sneak Peak Webinar on Confidence Intervals and Outliers

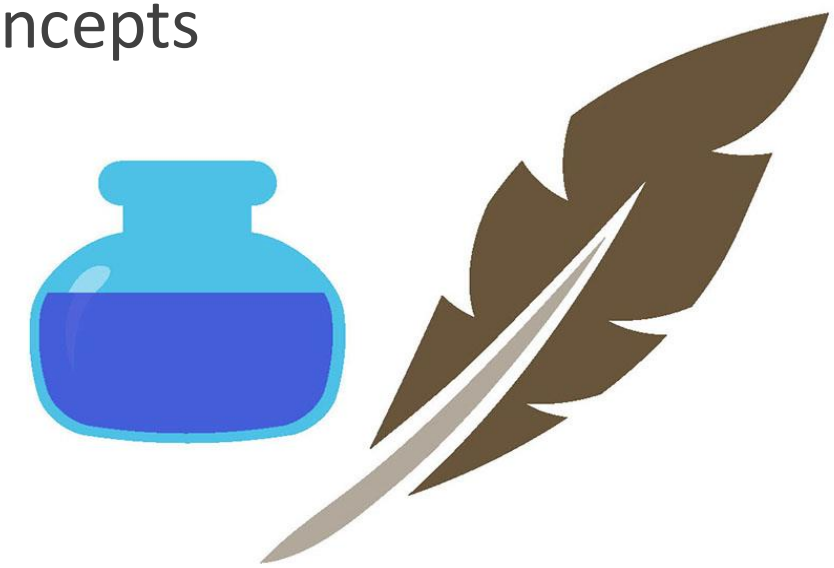
Statistics for Poets

Focus on understanding statistical concepts

No formulas

Lots of graphs

Samples of computer output



www.fastraincourses.com - Introduction to Statistics for Potency Bioassays

Key Questions

Where is the middle?

- Mean
- Median
- Geometric Mean (GM)

How much spread?

- Variance
- Standard Deviation (SD)
- Standard Error (SE)
- Coefficient of Variation (CV)
- Geometric CV (GCV)

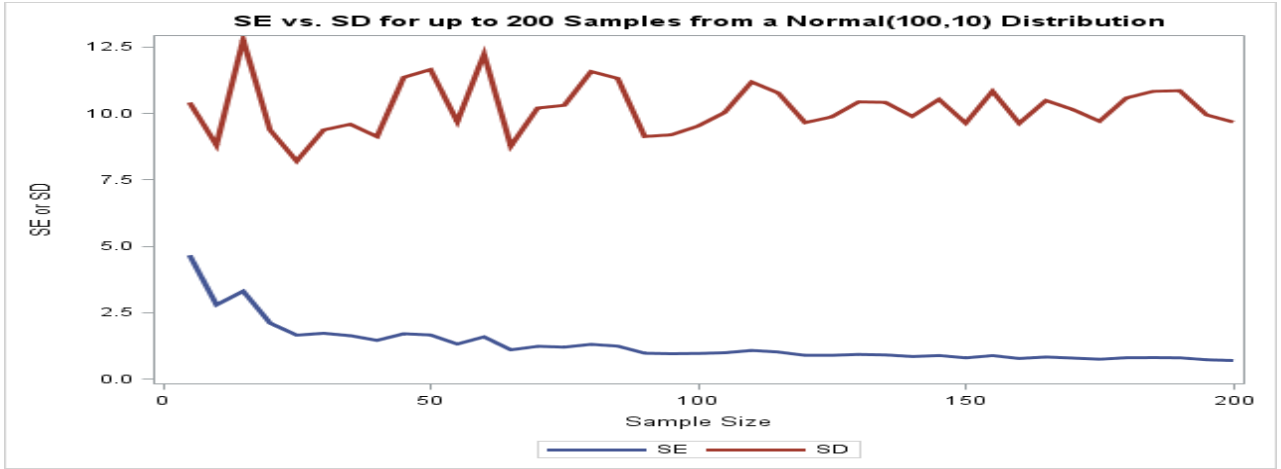
Standard Deviation vs. Standard Error

Standard Deviation

- SD describes the variability of the population
- Can be calculated for the population or for a random sample
- SD is similar for smaller & larger samples

Standard Error

- SE describes the variability of the sample mean
- SE is always smaller than SD (SE= SD/√n)
- SE decreases as sample size increases



Statistical Intervals

Confidence Intervals

- Where is the middle?
- Calculated using mean & SE
- Used for individual lots or for product (all lots)

Prediction Intervals & Tolerance Intervals

- Consider both middle and spread for the population
- Calculated using mean & SD
- Tolerance Intervals used to set system suitability criteria

Calculation is similar for CI, PI, TI

Basic calculation: mean ± constant x spread

Key difference is the estimate used for spread

- Confidence Interval: mean ± constant x SE
- Prediction Interval: mean ± constant x SD
- Tolerance Interval: mean ± constant x SD

Constant depends on the distribution, type of interval, confidence level and sample size (z, t, 2, TI value)

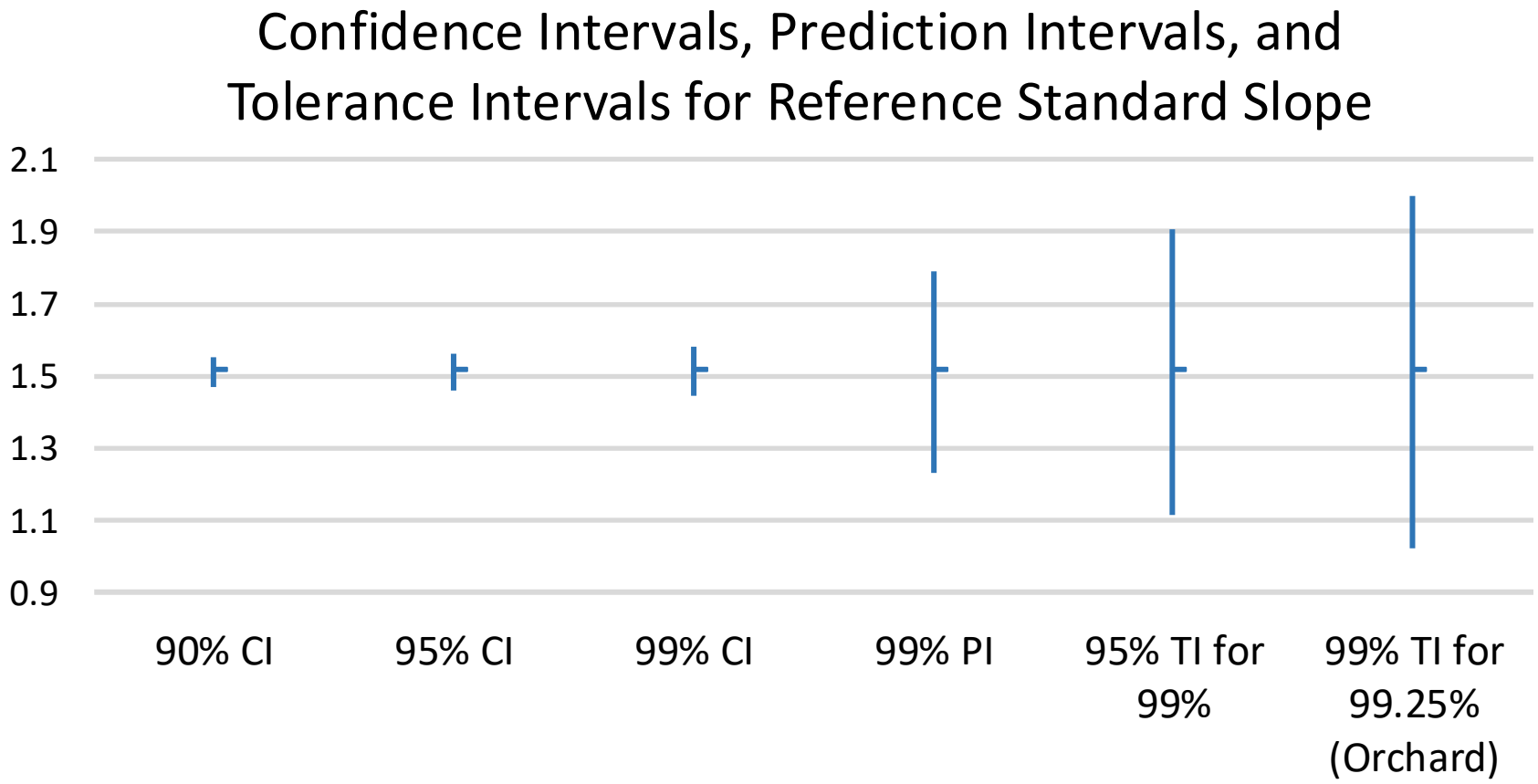
Example: Comparing Intervals for 4PL Steepness Parameter for Reference Standard

16 simulated curves
- Mean 1.5
- SD 0.10
- SE 0.025

CI based on SE

PI, TI based on SD

CI < PI < TI



Range of simulated slopes used for example (1.40, 1.71)

Interpreting Statistical Intervals

Confidence Interval for mean: We are 95% confident that the mean is in the interval.

Prediction Interval for new observation: We are 99% confident that the next observation will be in the interval.

Tolerance Interval for population: We are 95% confident that 99% of the observations will fall within the interval.

Confidence not Probability

Hypothetically,

If we collected many random samples from a population

And calculated the mean and SE for each sample

And calculated a 95% confidence interval for each sample

Then 95% of the intervals are expected to contain the population mean

Why such an odd definition for confidence interval?

Population mean is a fixed value

Any particular confidence interval either contains the mean or it doesn't

Can't make a probability statement on a fact, but we also don't know whether the interval contains the population mean

So the probability statement is on the hypothetical large number of intervals

Schrödinger's Cat

Just to be clear

We only collect one set of samples

Confidence interval is calculated using those samples

USP and Ph Eur Guidance on Combining Results from Multiple Assays

Good News: USP (Chapter 1034) and Ph Eur (Chapter 6.3) recommend the same methods

Bad News:

- They call them different things
- They use different endpoints
  - Ph Eur uses potency
  - USP uses relative potency
- Ph Eur recommends a test that USP doesn't use
- Disagree on how to implement one of the methods
  - Different formulas
  - USP version can give negative variance estimates (correction for this in draft guidance)

Methods for Combining Results from Multiple Assays

Method 1: Unweighted mean, only inter-assay variability

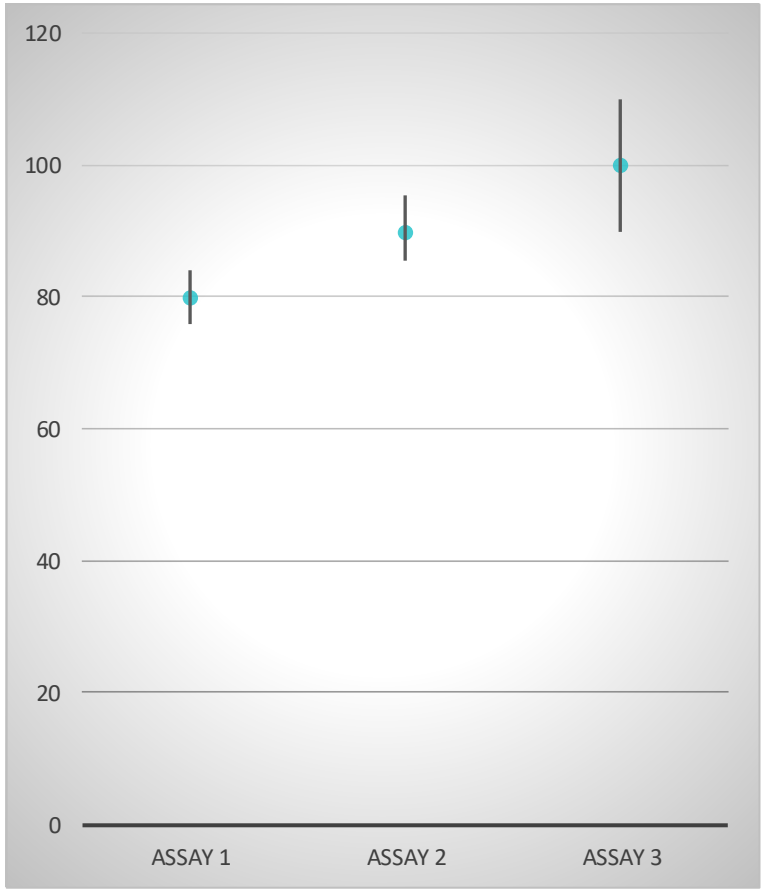
Method 2: Weighted mean, only intra-assay variability

Method 3: Weighted mean, inter-assay and intra-assay variability

Intra-assay and Inter-assay Variability

Intra-assay variability is within assay variability: width of confidence intervals

Inter-assay variability is between assay variability: distance between estimates



Unweighted and Weighted Means

Unweighted mean (average) gives equal weight to every value

Weighting adjusts the mean to emphasize values that are more important or more credible (lower variability)

Calculating a Weighted Mean

- Choose weights
- Multiply each value by the weight for that value
- Add all the products (value x weight)
- Add all the weights
- Divide the total of the products by the total of the weights

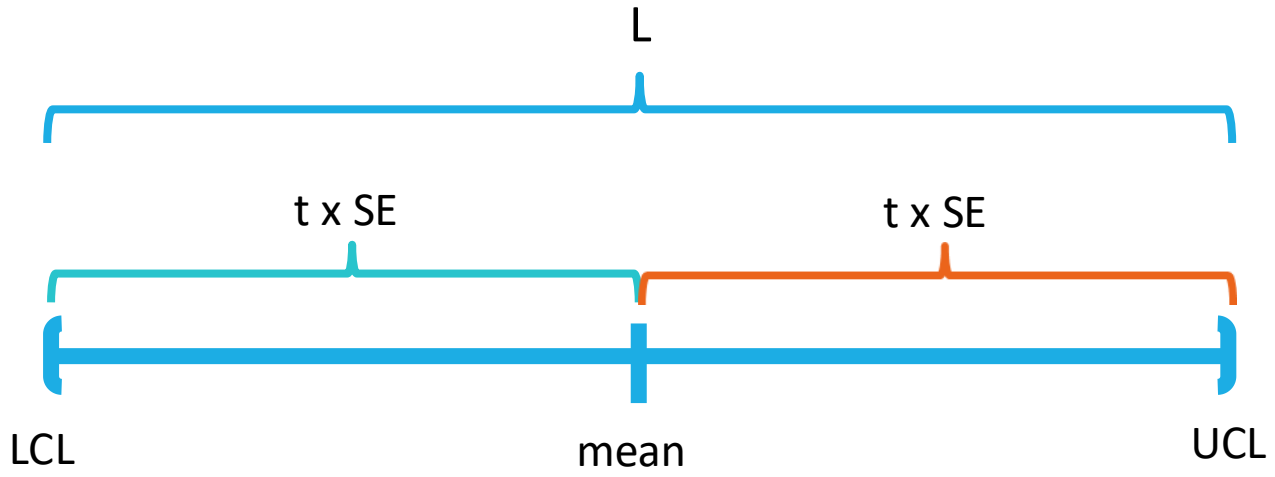
How to Calculate Weights in Bioassay

USP and Ph Eur both recommend weights of 1/variance of mean (1\SE^2)

Formula for weights based on length of the confidence interval $W=4t^2/L^2$

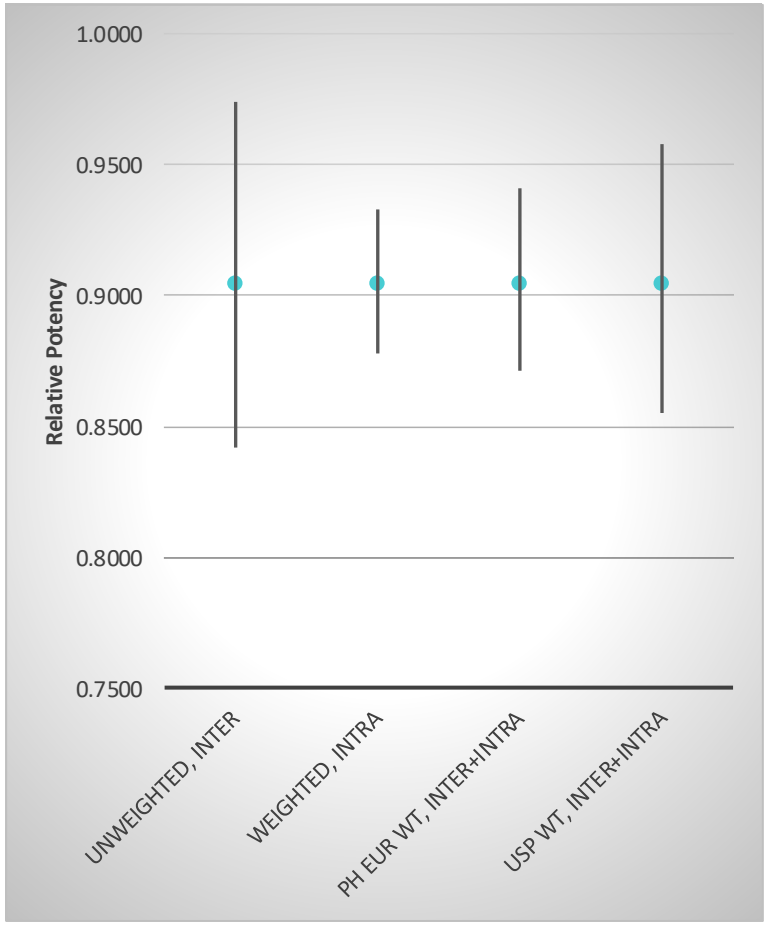
How Does That Work?

- CI is centered on mean
- Add & subtract t x SE
- Width of CI (L) is 2 x t x SE
- SE in terms of L and t is L/2t



Comparison of Weighting Methods on Mean Relative Potency when Inter-assay Variability is Large

- All methods gave similar means
- Method 1: “Unweighted, Inter” has wide CI; accounts for inter-assay variability but limited df
- Method 2: “Weighted, Intra” is narrow; doesn’t account for inter-assay variability – not a good choice for this example
- Method 3: “Weighted, Inter+Intra” methods account for within & between assay variability; Wider for USP than Ph Eur



When to Use Each Method

Method 3 accounts for both inter-assay and intra-assay variability; conservative choice

Method 1 usually not recommended, but it can be useful under some circumstances

- If inter-assay variability is large compared to intra-assay variability
- If no intra-assay variance estimate is available or not enough df to use other methods

Method 2 ignores inter-assay variability, should be used with caution; only when homogeneity between assays is verified

Using Method 3 when homogeneity assumption for Method 2 is met shouldn't hurt

Choose a method during assay development; verify during validation and use it consistently

Questions?

Thank You!

Asia Bioassay Conference

November 4-6, 2026 | Incheon, South Korea

$200 OFF Registration!

Discount Code: ABAWEBINAR