Tolerance Intervals – an Adaptive Approach for Specification Setting

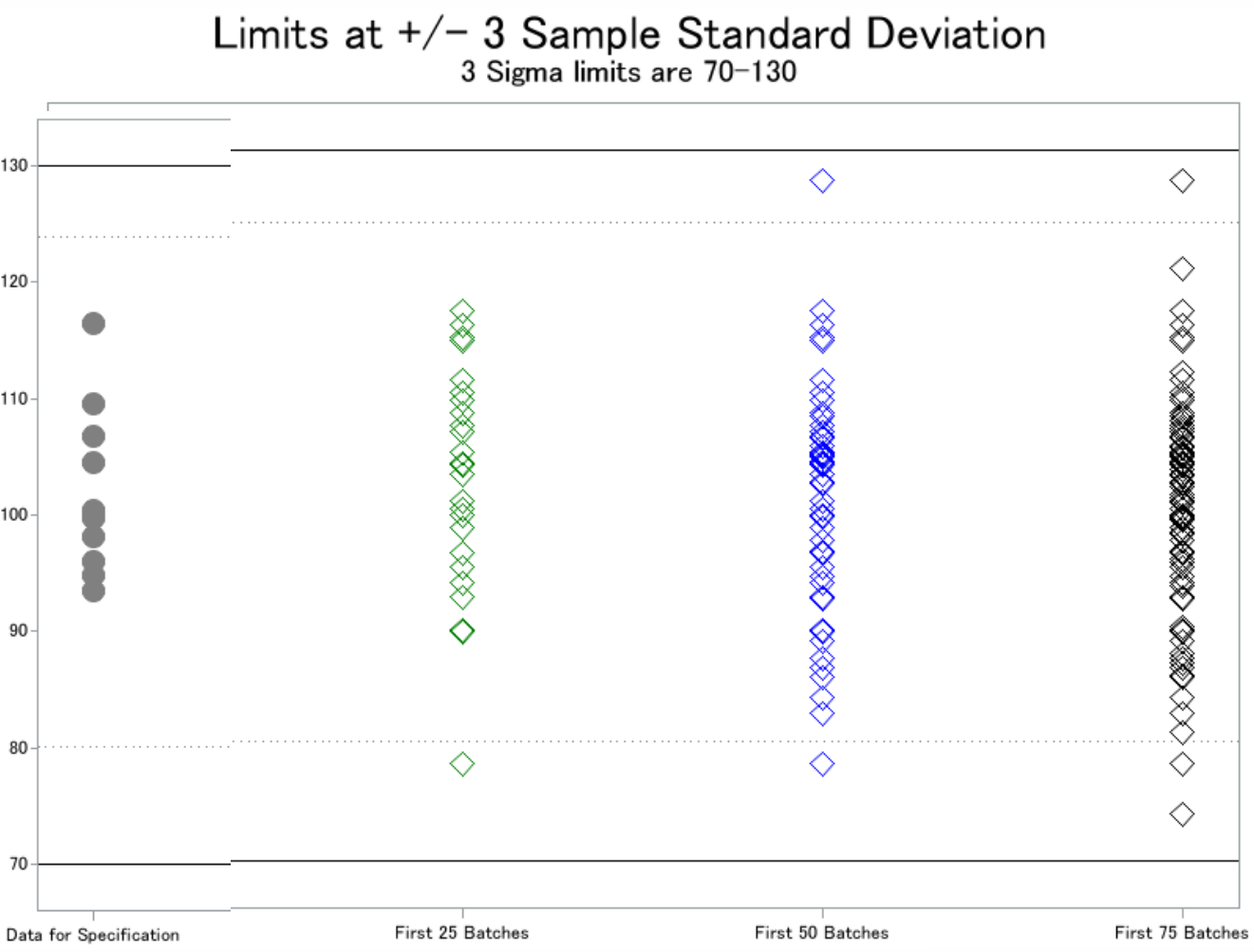
Brad Evans
Associate Director, Pharm Sci & PGS Statistics
Pfizer R&D
Fall Technical Conference 2018

Specifications…

- Ideally defined ahead of time
  - Clinically relevant
  - Build the process to meet
  - Indices to measure performance, stability, etc. (previous talk)
- Ensure safety and efficacy
- But we DON'T always know these ahead of time, so we need

- **"data driven specifications"**

Setting Specifications … with limited data



Ideally, specifications cover the vast majority of results, so true 3 sigma limits might be acceptable

Very wide = unacceptable?

First 10 batches much narrower than true +/- 3 sigma

The **sample** limits are narrower than the process limits, so future values are out of Specification

Why isn't +/- 3SD enough?

In the (very long) run, mean +/- 3 SD will cover 99.73% of the population if we have a "Normal" or bell-shaped data distribution.

A data-driven specification is set based on 5-15 batches (typically)

Mean +/- 3 SD almost always does the cover the data *we have in hand.*

But it needs to cover **future data, from the same manufacturing process.**

The Tolerance Interval is built to account for sampling variation.

In practice, the following variability will **also** occur:

- New batches of raw materials
- Changes to the assay / method transfers / site transfers.
- Process improvements

Contrast to SPC, which uses +/- 3SD

Control chart (typically)

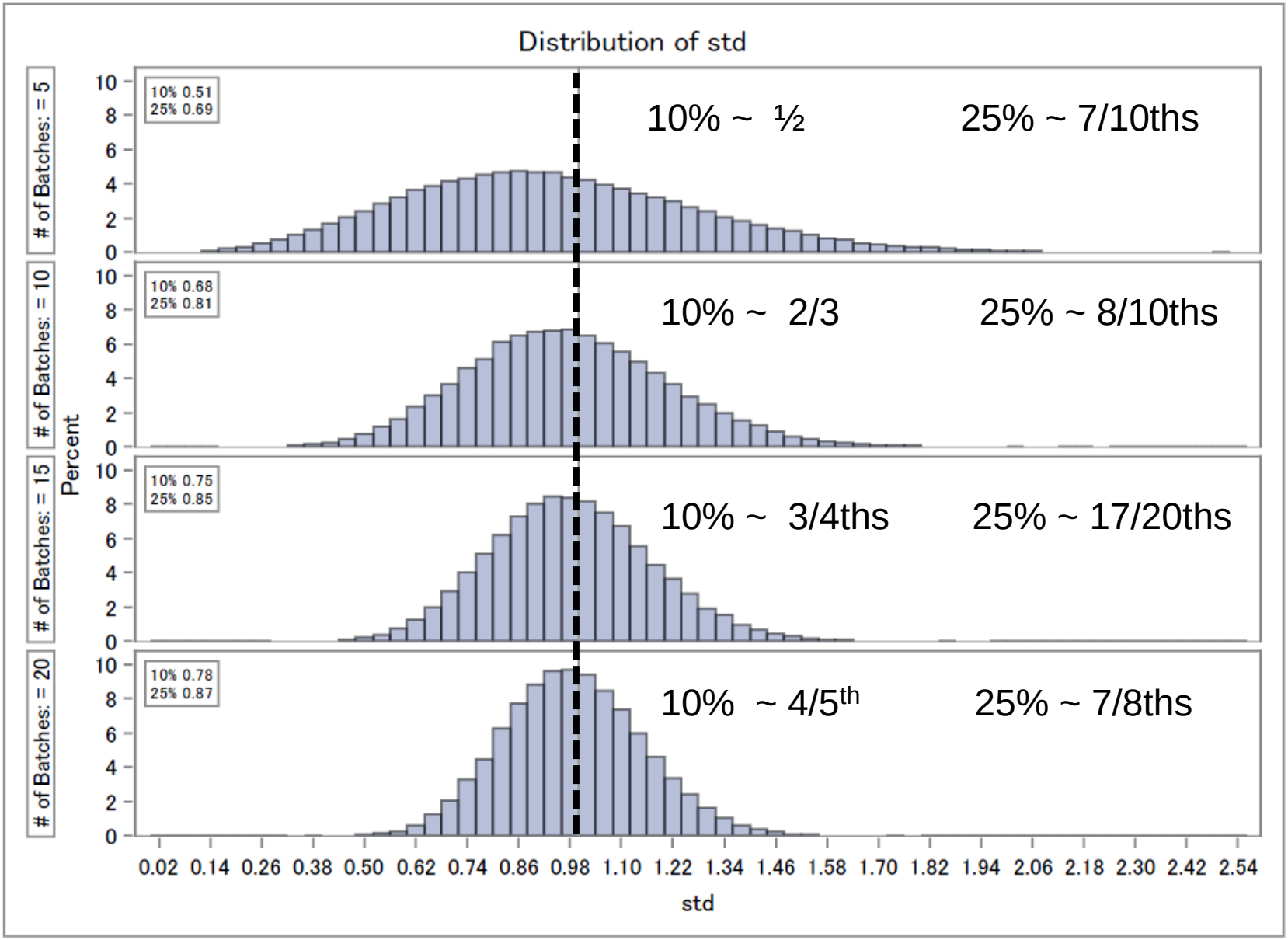
- **n >= 25**
- Continuous process, with subsamples (xbar, r), (xbar, s)
- A precise measurement system
- **Spec is pre-specified, so Cpk >> 1 means "highly capable process"**
- Out of *Control* -> go investigate

Pharmaceutical Data-driven Specifications (typically):

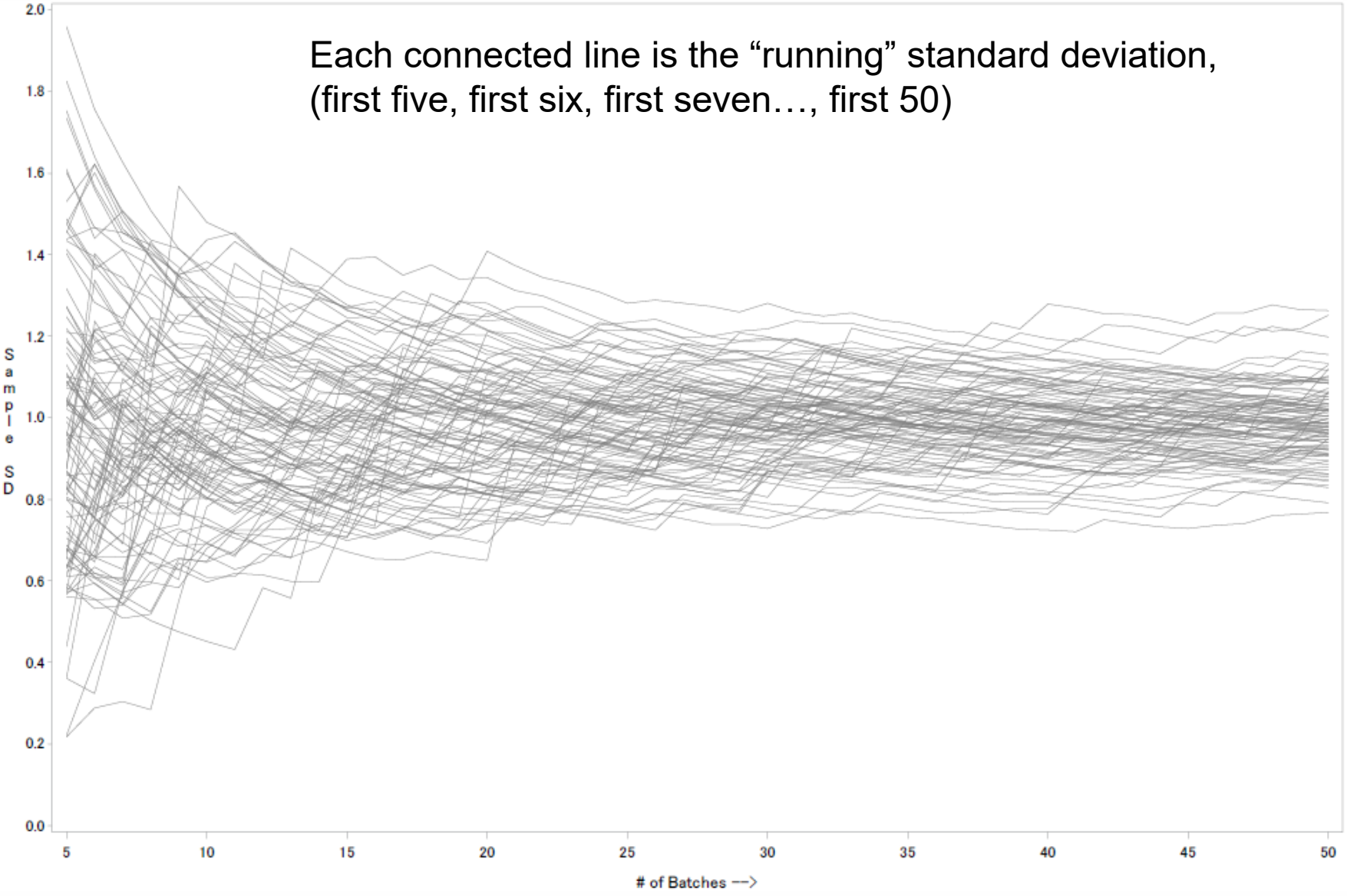
- **n ~ 5-15**
- Batch process, n=1 measurement/batch
- Measurement device (assay) may be a significant source of variability
- **Spec is data driven, so Cpk >> 1 amounts to "specs too wide"**
- Out of *Specification* -> (potentially) dispose of the batch

Sample SD relative to true Sigma

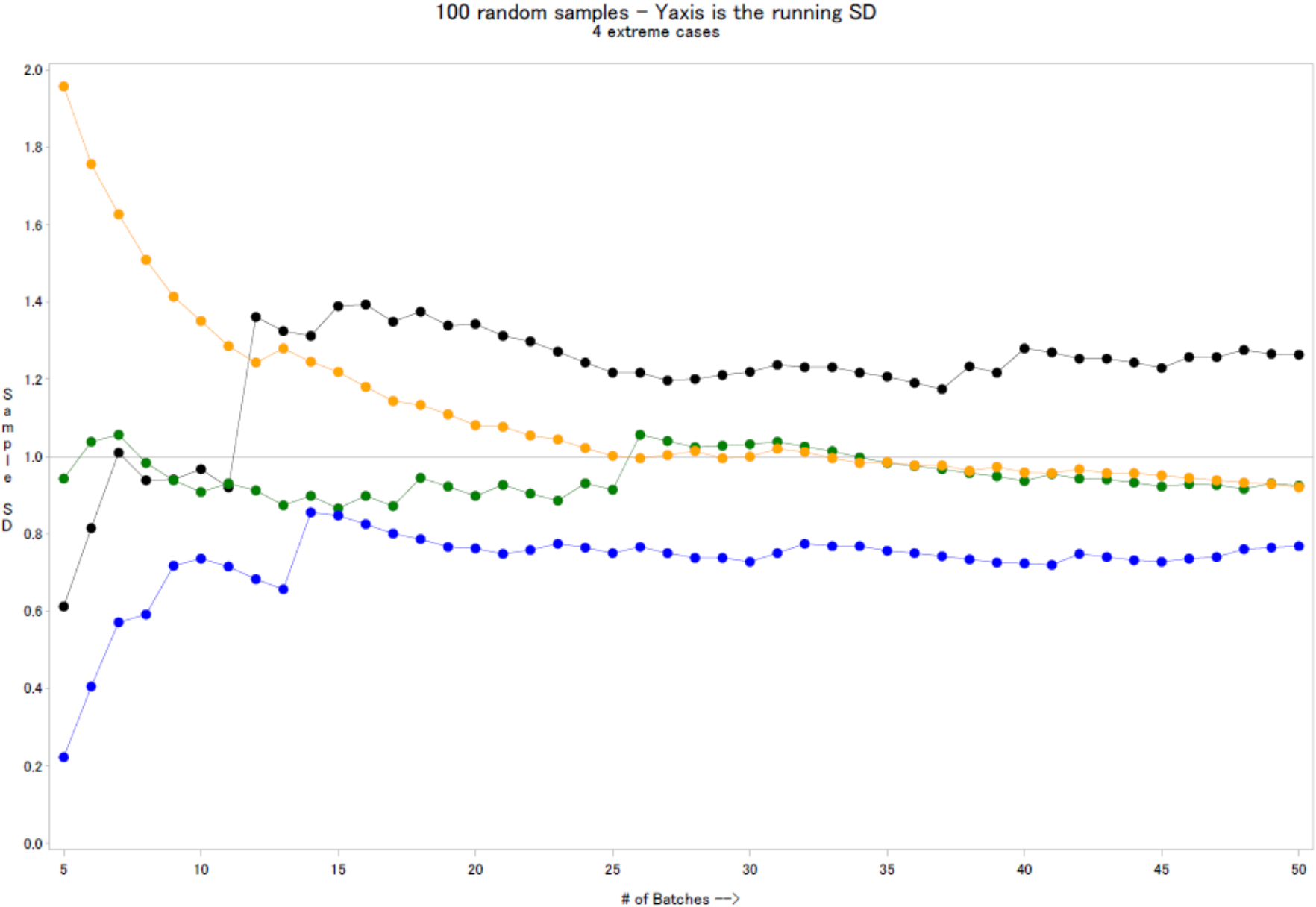
(sample variance is scaled Chi-sq)



In the long run….



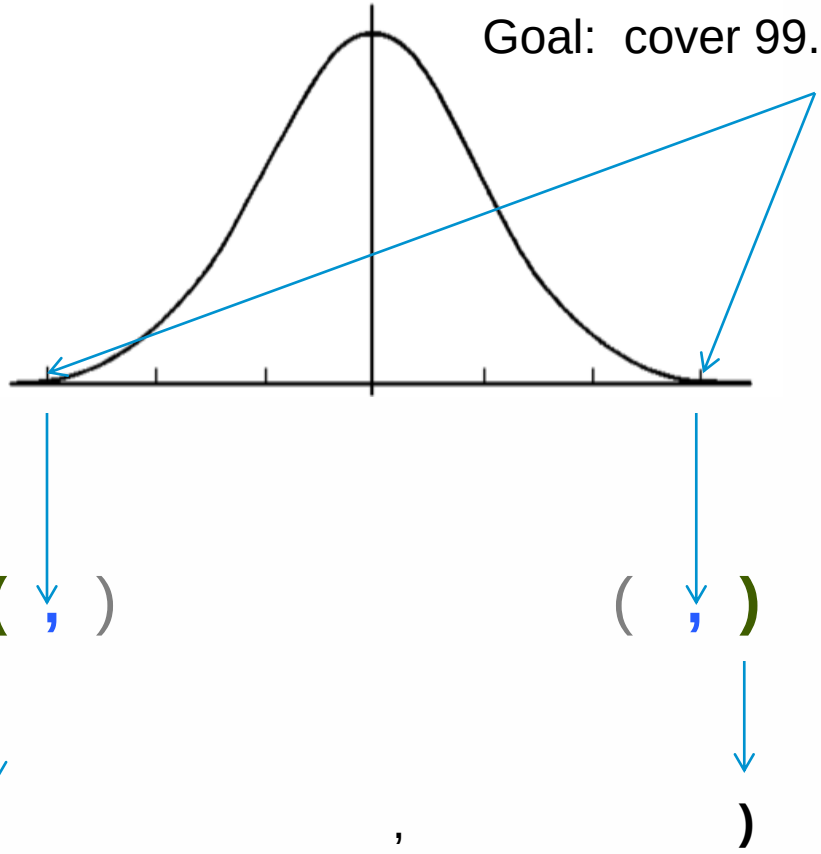
But there are extremes…



Why a Tolerance Interval is Proposed

- Specifications are a commitment that *future* batches will land in the specified window
- “Data based Specification Setting” – the situation where a specification is NOT known ahead of time
- A Tolerance Interval (TI) is one way to calculate a range intended to include a fixed % of the population (coverage) with some specified confidence, and depends on:
    - What data is selected
    - Confidence and coverage used to determine the multiplier

Tolerance Interval



Goal: cover 99.73% of the distribution (+/- 3 sigma)

Point estimates for **coverage**

Confidence limits around our estimates

For greater **coverage** the estimates (blue commas) would be farther apart.

Greater **confidence** → wider **( )**

Sample mean and sample SD both will bounce around. Formula = mean +/- k*sd, k = k(n, cov, conf)

Interplay of (n, coverage, confidence)

Each column is a Coverage / Confidence combination, with color coding on the multiplier

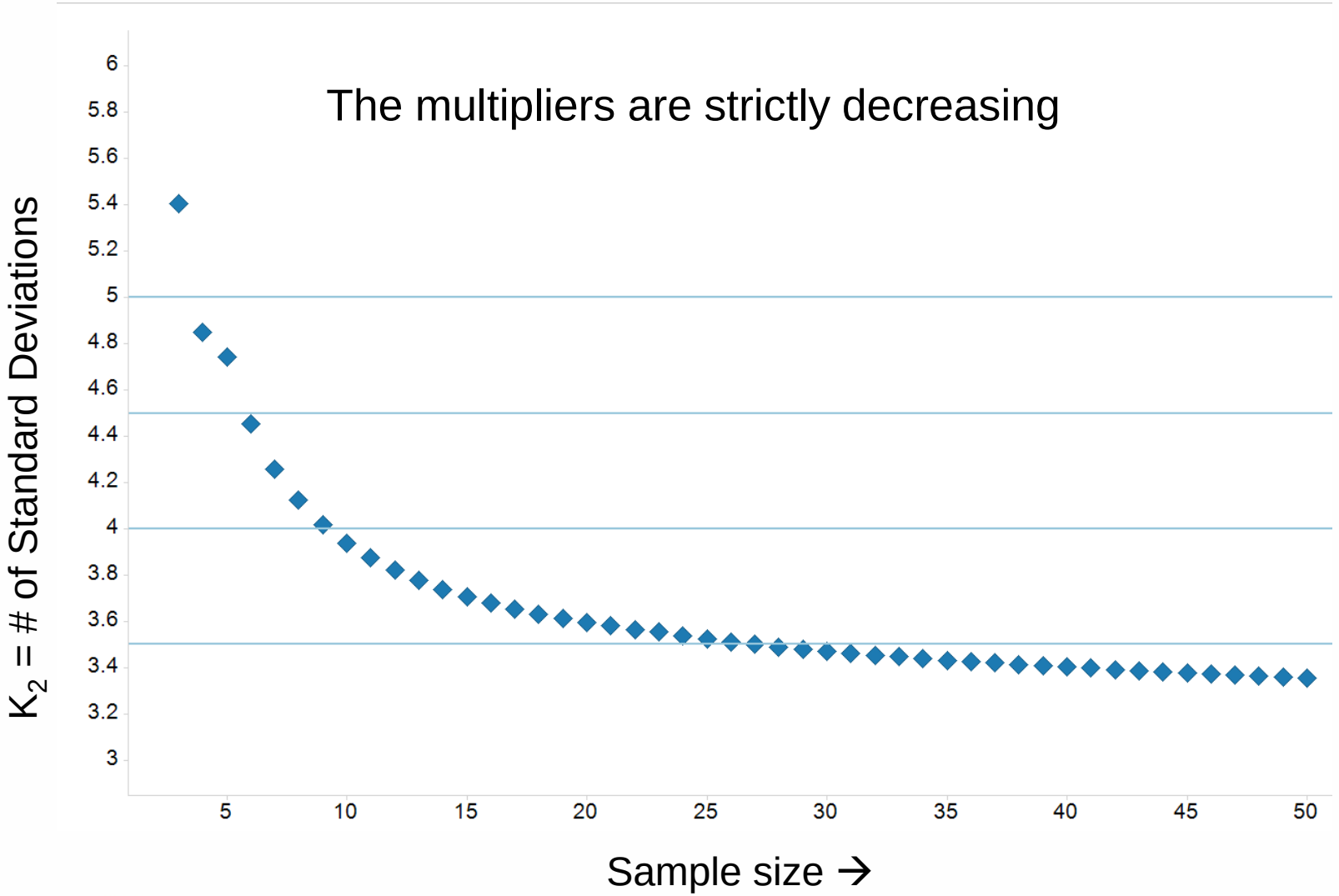
High confidence, high coverage *but small n* we get multipliers of 5 or 6 or 12+ (upper right)

If we back off coverage and confidence *but large n* we get multipliers < 3 (lower left)

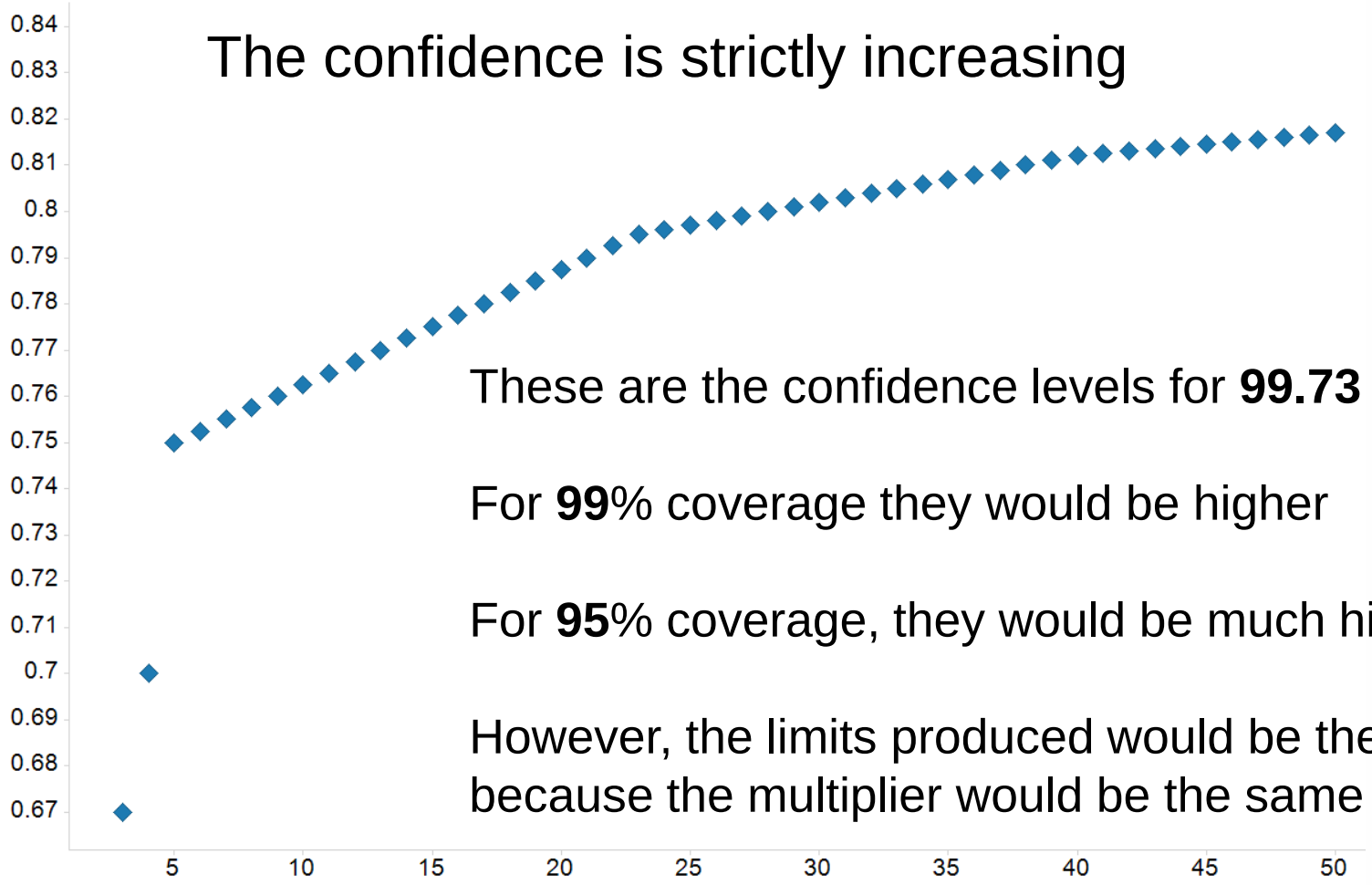
Cov	0.9	0.9	0.95	0.95	0.9	0.99	0.95	0.99	0.9973	0.9973	0.99	0.9973
Conf	0.9	0.95	0.9	0.95	0.99	0.9	0.99	0.95	0.9	0.95	0.99	0.99
N												
5	3.49	4.27	4.16	**5.09**	6.61	5.47	7.88	6.69	6.37	7.80	10.35	12.06
6	3.13	3.71	3.73	4.42	5.34	4.90	6.36	5.81	5.71	6.77	8.36	9.73
7	2.90	3.37	3.46	4.01	4.61	4.54	5.50	5.27	5.29	6.14	7.22	8.41
8	2.74	3.14	3.27	3.74	4.15	4.29	4.94	4.91	5.00	5.72	6.49	7.56
9	2.63	2.97	3.13	3.53	3.82	4.11	4.55	4.65	4.79	5.41	5.98	6.97
10	2.54	2.84	3.02	3.38	3.58	3.97	4.27	4.44	4.62	5.18	5.61	6.53
11	2.46	2.74	2.93	3.26	3.40	3.86	4.05	4.29	4.49	4.99	5.32	6.20
12	2.40	2.65	2.86	3.16	3.25	3.77	3.87	4.16	4.38	4.84	5.09	5.93
13	2.36	2.59	2.81	3.08	3.13	3.69	3.73	4.05	4.30	4.72	4.90	5.71
14	2.31	2.53	2.76	3.01	3.03	3.62	3.61	**3.96**	4.22	4.61	4.74	5.52
15	2.28	2.48	2.71	2.95	2.94	3.57	3.51	3.88	4.15	4.52	4.61	5.37
16	2.25	2.44	2.68	2.90	2.87	3.52	3.42	3.82	4.10	4.44	4.50	5.24
17	2.22	2.40	2.64	2.86	2.81	3.47	3.35	3.76	4.05	4.38	4.40	5.12
18	2.19	2.37	2.61	2.82	2.75	3.44	3.28	3.71	4.00	4.32	4.31	5.02
19	2.17	2.34	2.59	2.78	2.70	3.40	3.22	3.66	3.96	4.26	4.23	4.93
20	2.15	2.31	2.56	2.75	2.66	3.37	3.17	3.62	3.93	4.21	4.16	4.85

Propose: Confidence increases with N Multiplier decreases with N

Two sided multipliers vs. Sample Size



Confidence vs. Sample Size



The confidence is strictly increasing

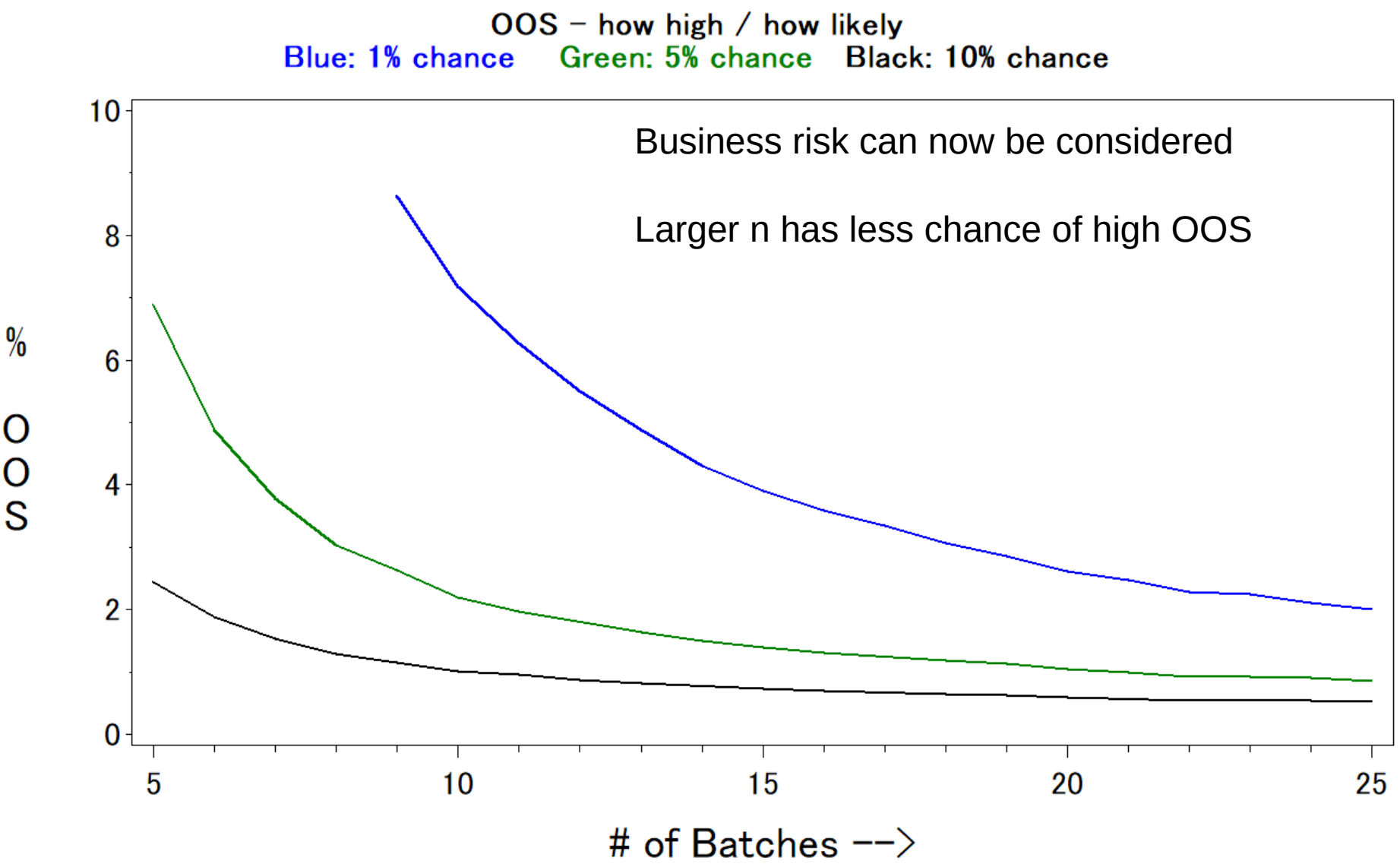
These are the confidence levels for **99.73** coverage

For **99**% coverage they would be higher

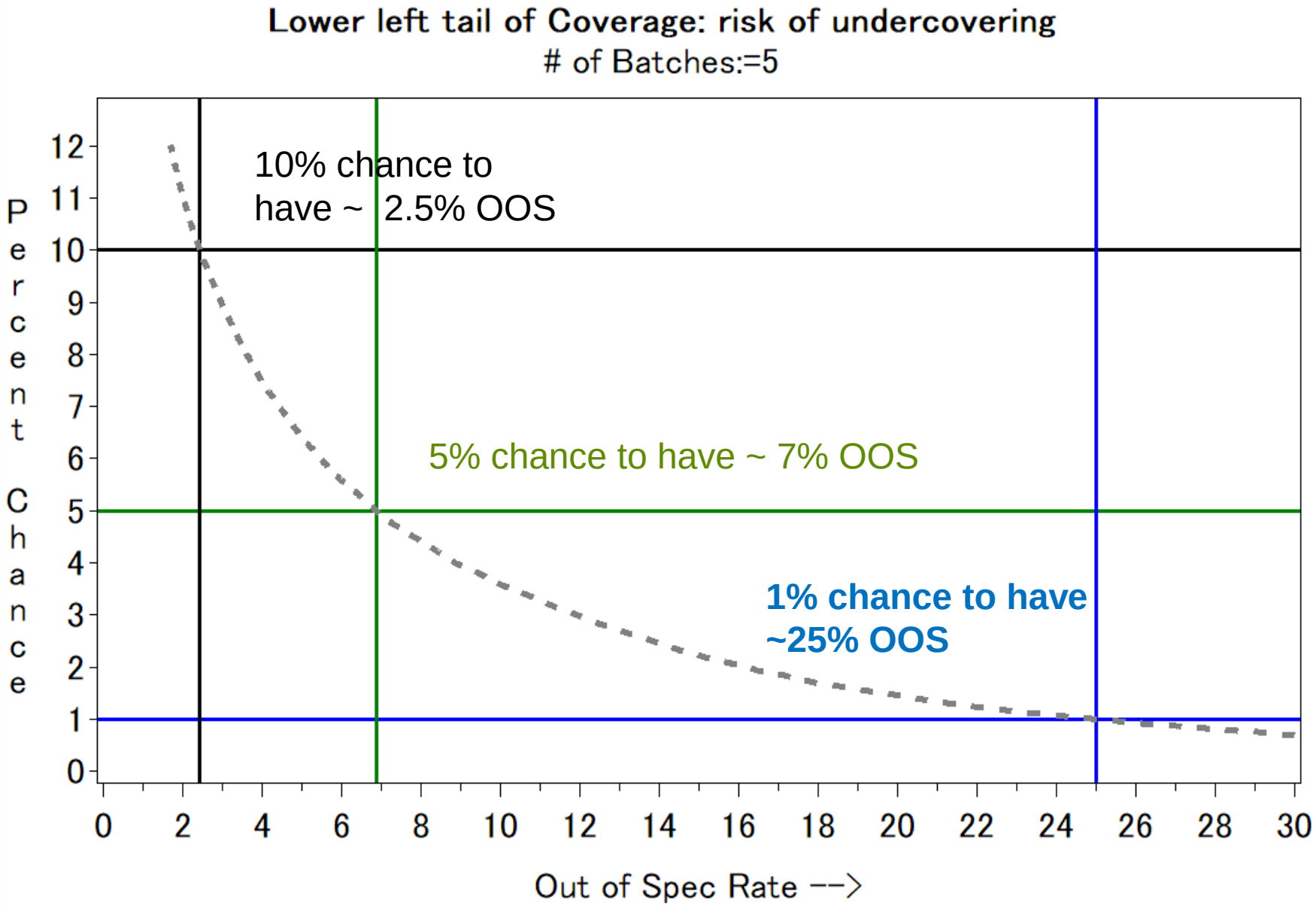
For **95**% coverage, they would be much higher

However, the limits produced would be the same because the multiplier would be the same

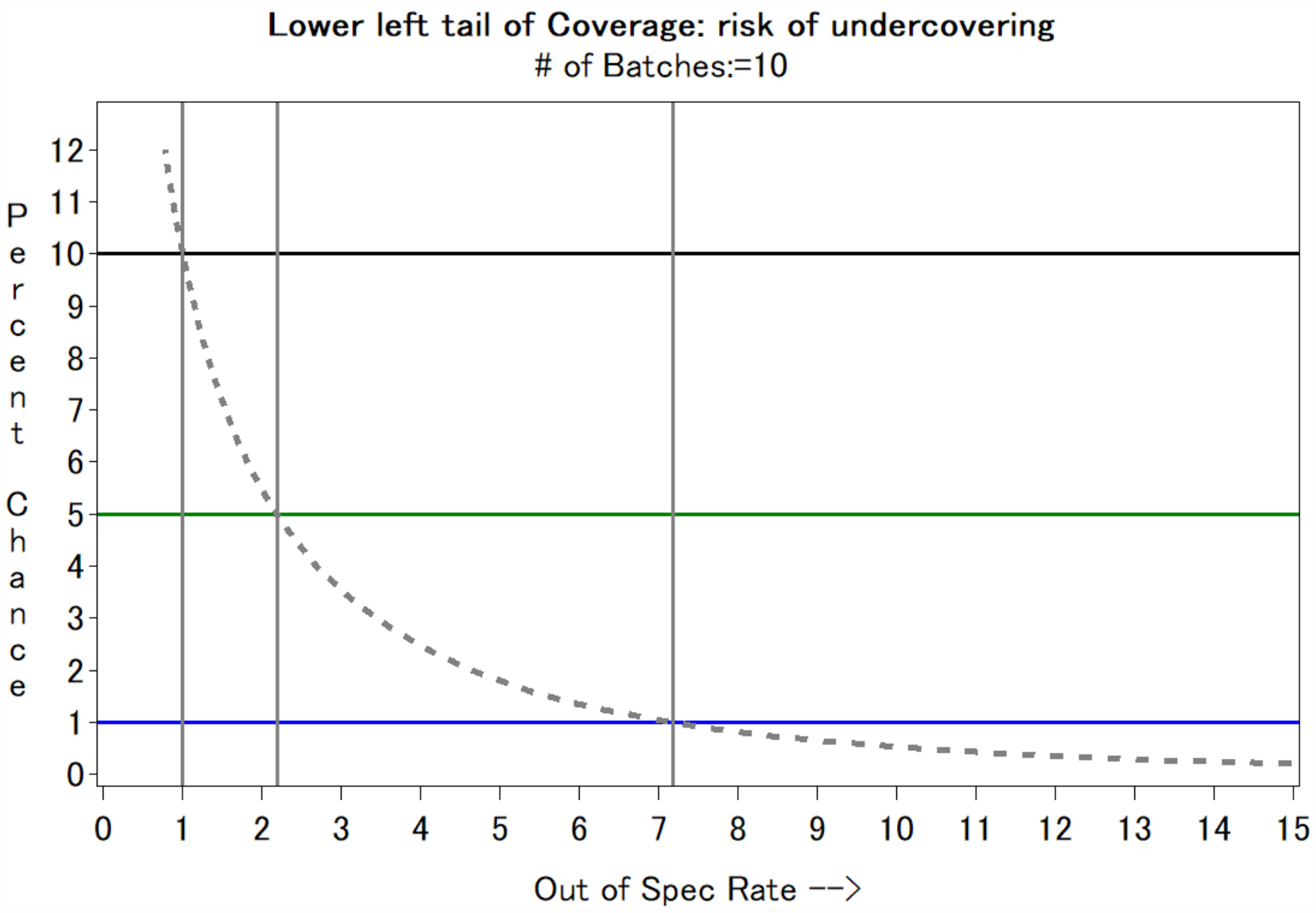
TI gives us a *distribution* of expected risk…



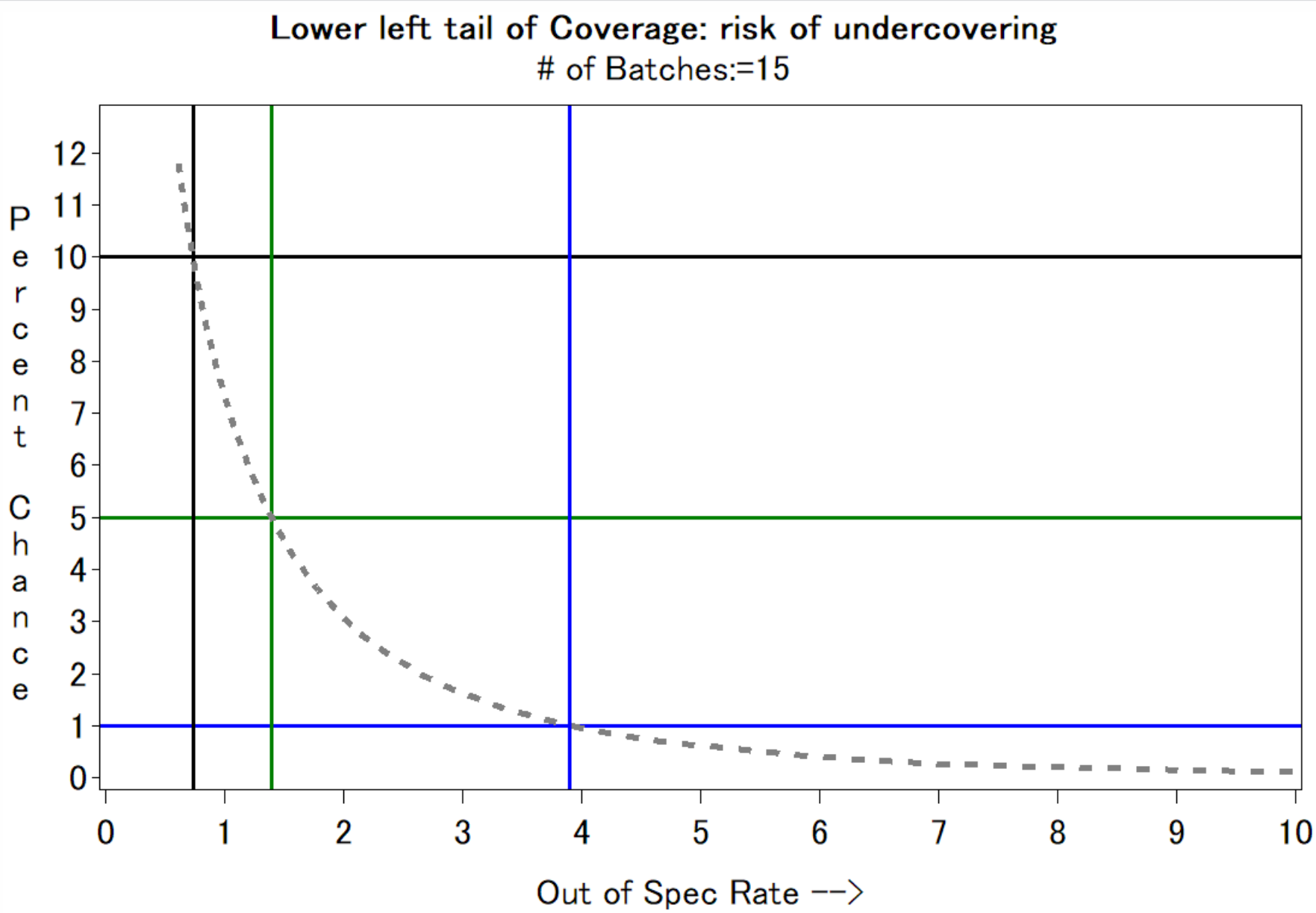
N=5, multiplier = 4.74, Confidence = 75%



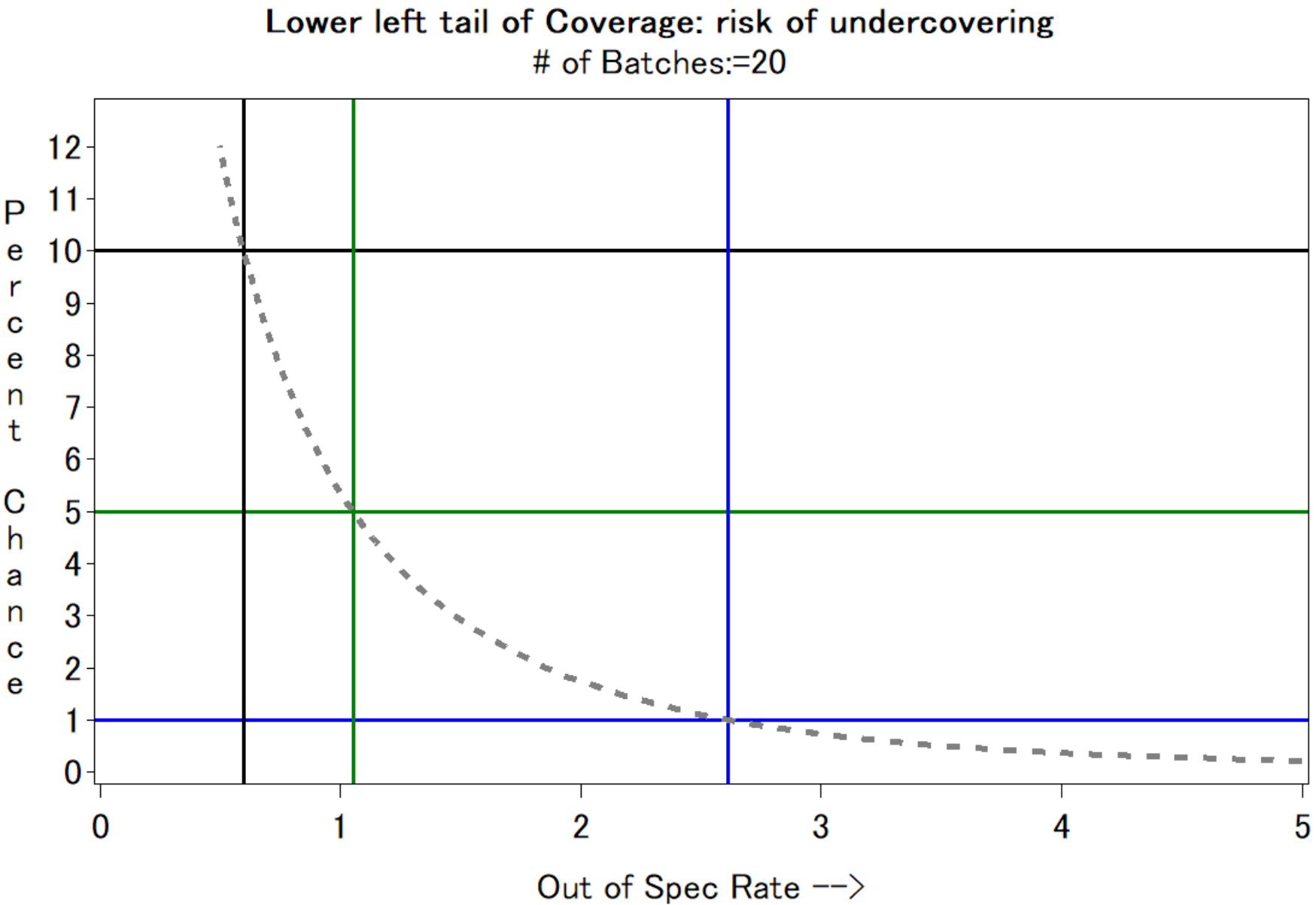
N=10, multiplier = 3.94, Confidence = 76.3%



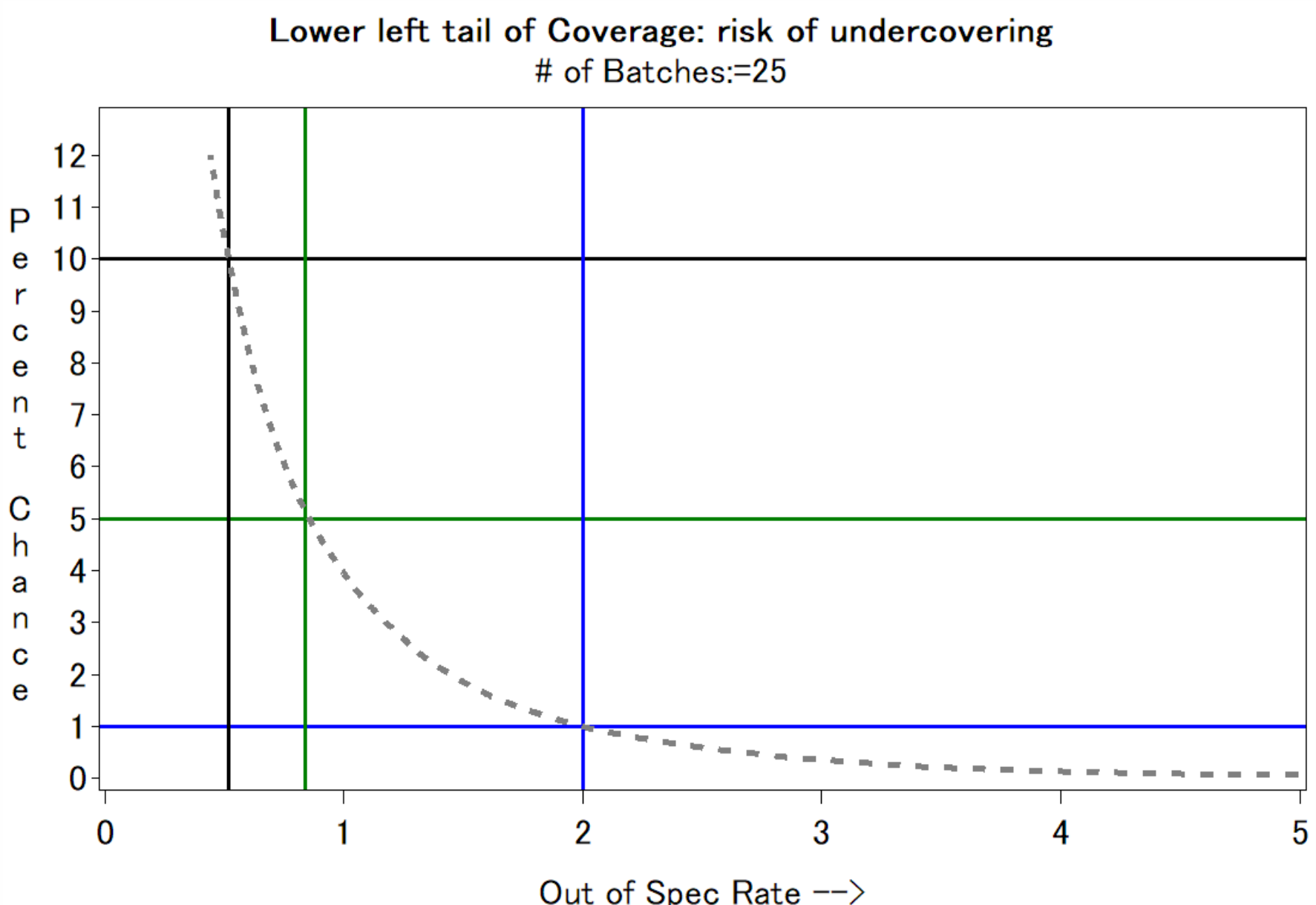
N=15, multiplier = 3.71, confidence = 77.5%



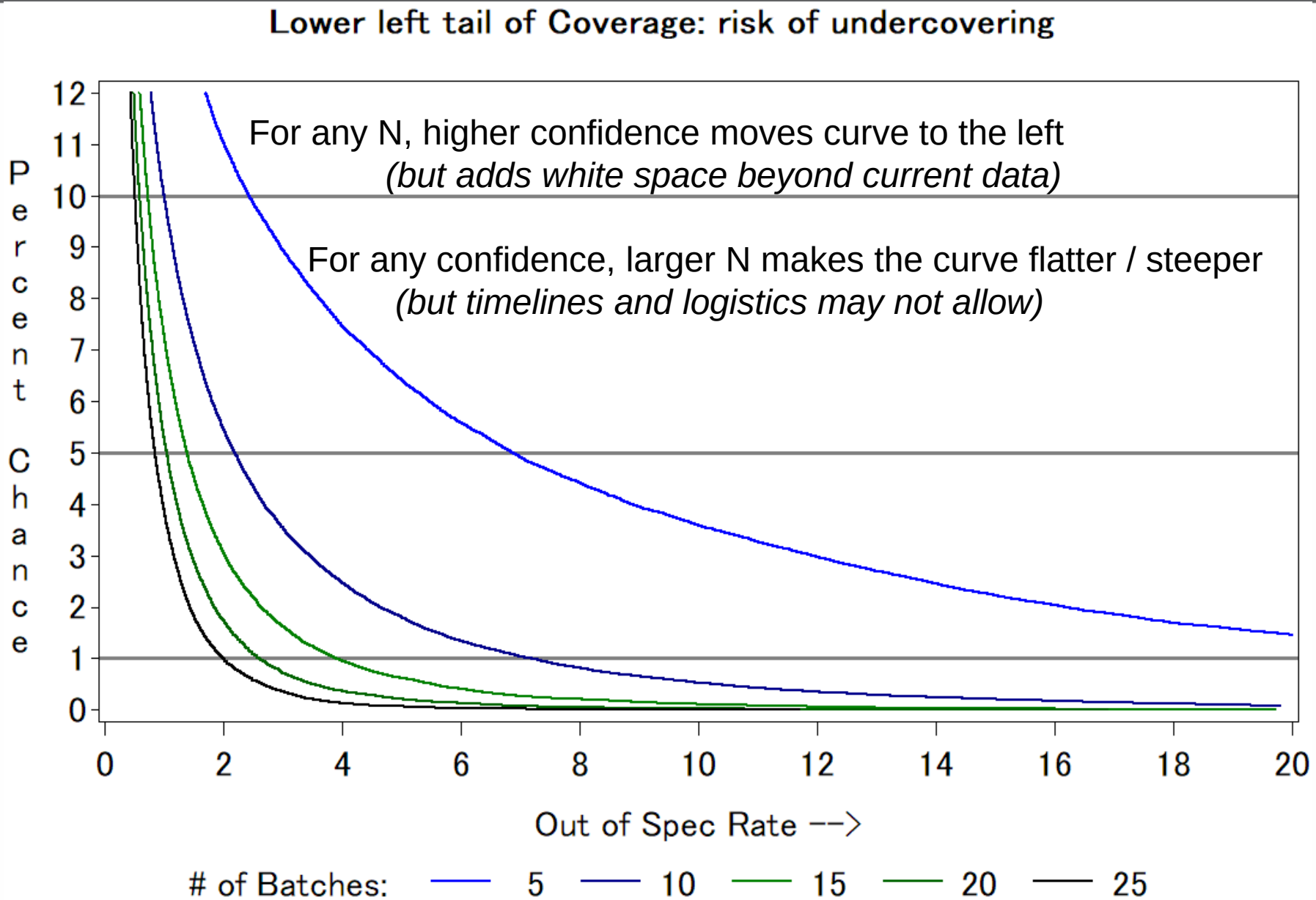
N=20, multiplier = 3.6, confidence = 78.8



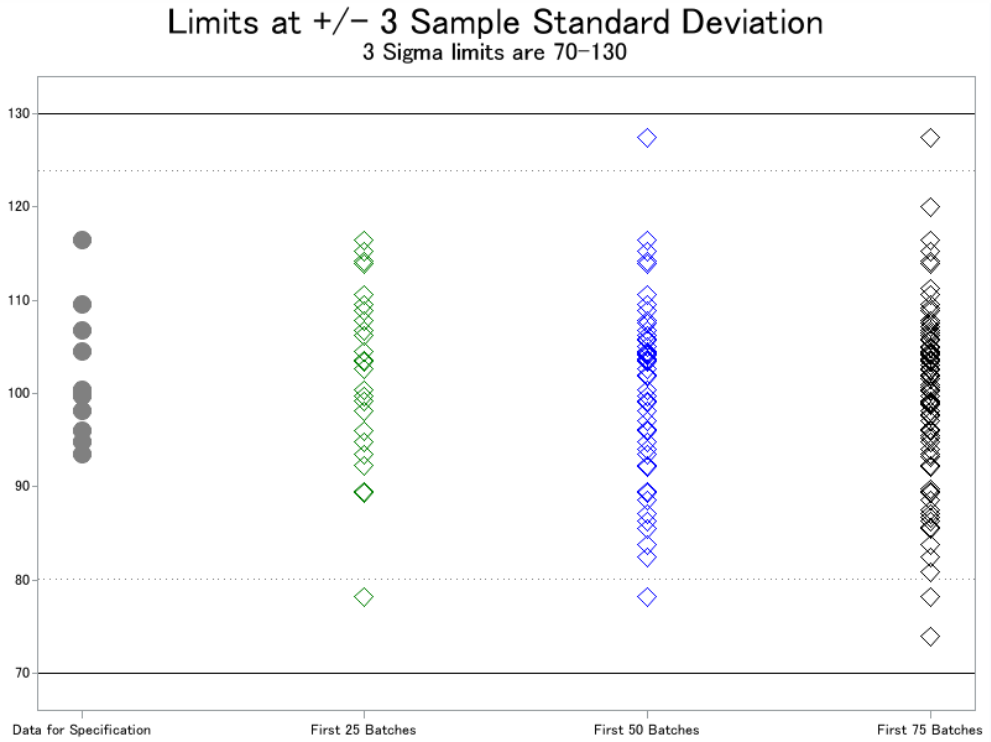
N=25, multiplier = 3.53, confidence = 79.7%



All combined…



Back to the example: TI Proposed Limits



Conclusion: setting +/- 3 SD limits will small n very likely leads to high Out of Spec rates

This proposed approach dampens the risk

Specification Set at N=10	Limits	Cpk at 25 Lots	Cpk at 50 Lots	Cpk at 75 Lots
Mean +/- 3SD	(80.2,1 23.8)	0.77	0.70	0.67
Proposed TI	(73.8, 130.2)	0.91	0.91	0.89

Cpk – a standard metric for how well a process can meet a Specification. Higher is better

But what about Stability data…..

- All Batches get tested upon manufacture
- Many batches are put “on stability”
  - Samples pulled, analyzed at pre-defined intervals
  - Some attributes change over time
  - Many do not change

Stability data – no trend

If no trend exists, we can model the stability data as:

Attribute = Batch effect + Assay variation

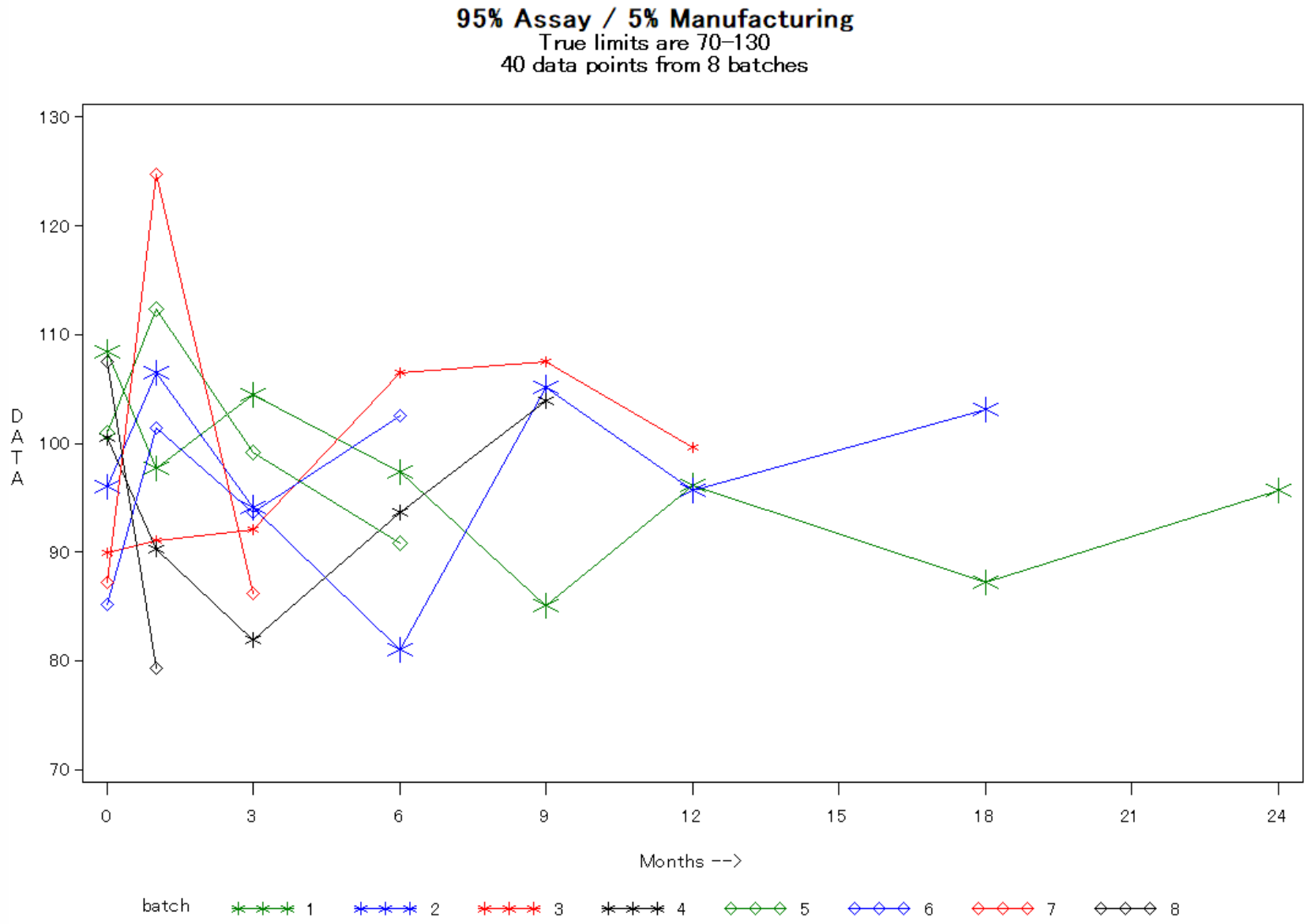
By doing a variance components breakdown, we can get

Var(Total)= Var(Manufacturing) + Var(Assay)

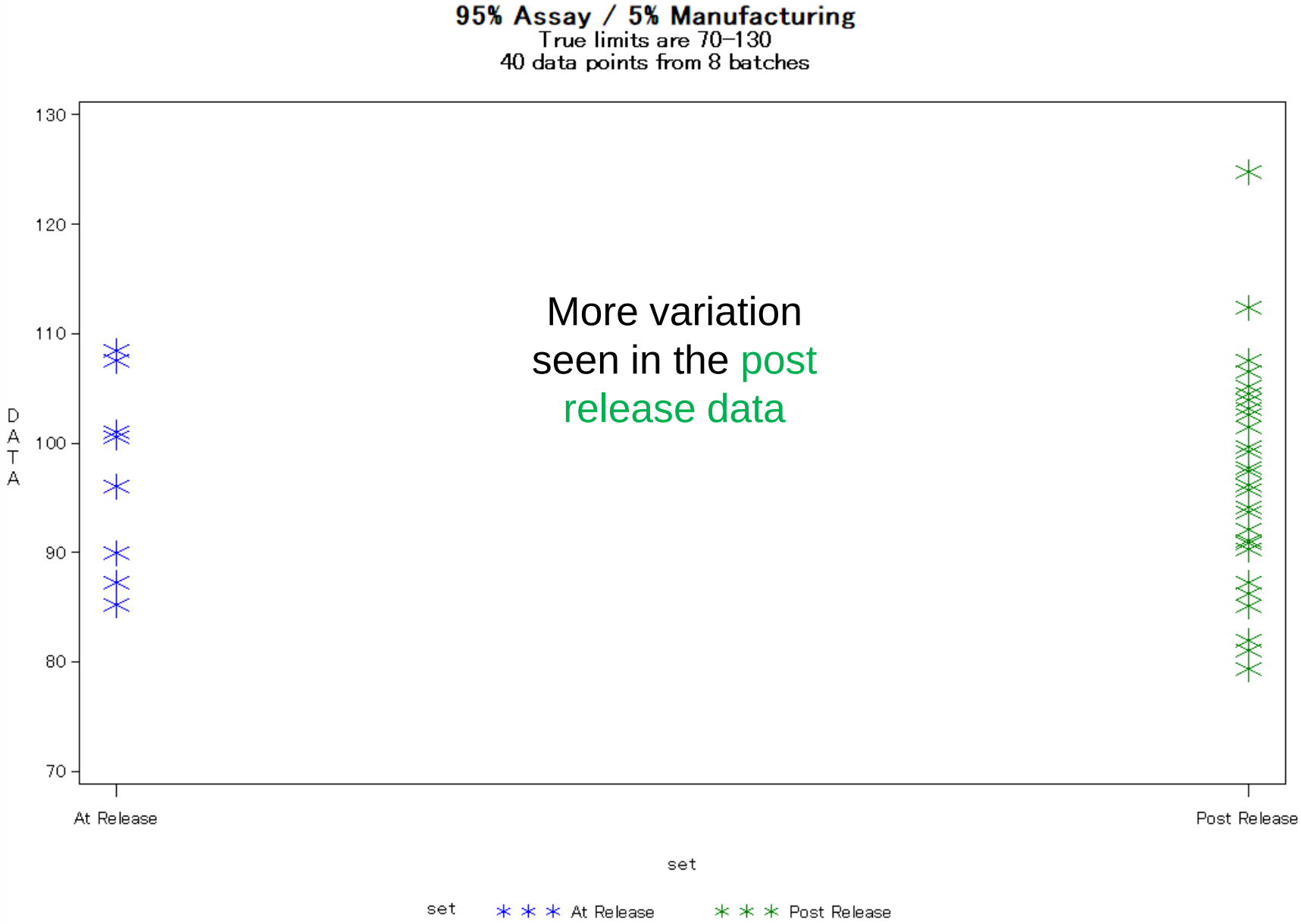
Then we create a second Tolerance Interval using SD(Total)

It can (and does) happen that the stability data indicates a larger Var(Total) than the release data alone

Stability data – no trend

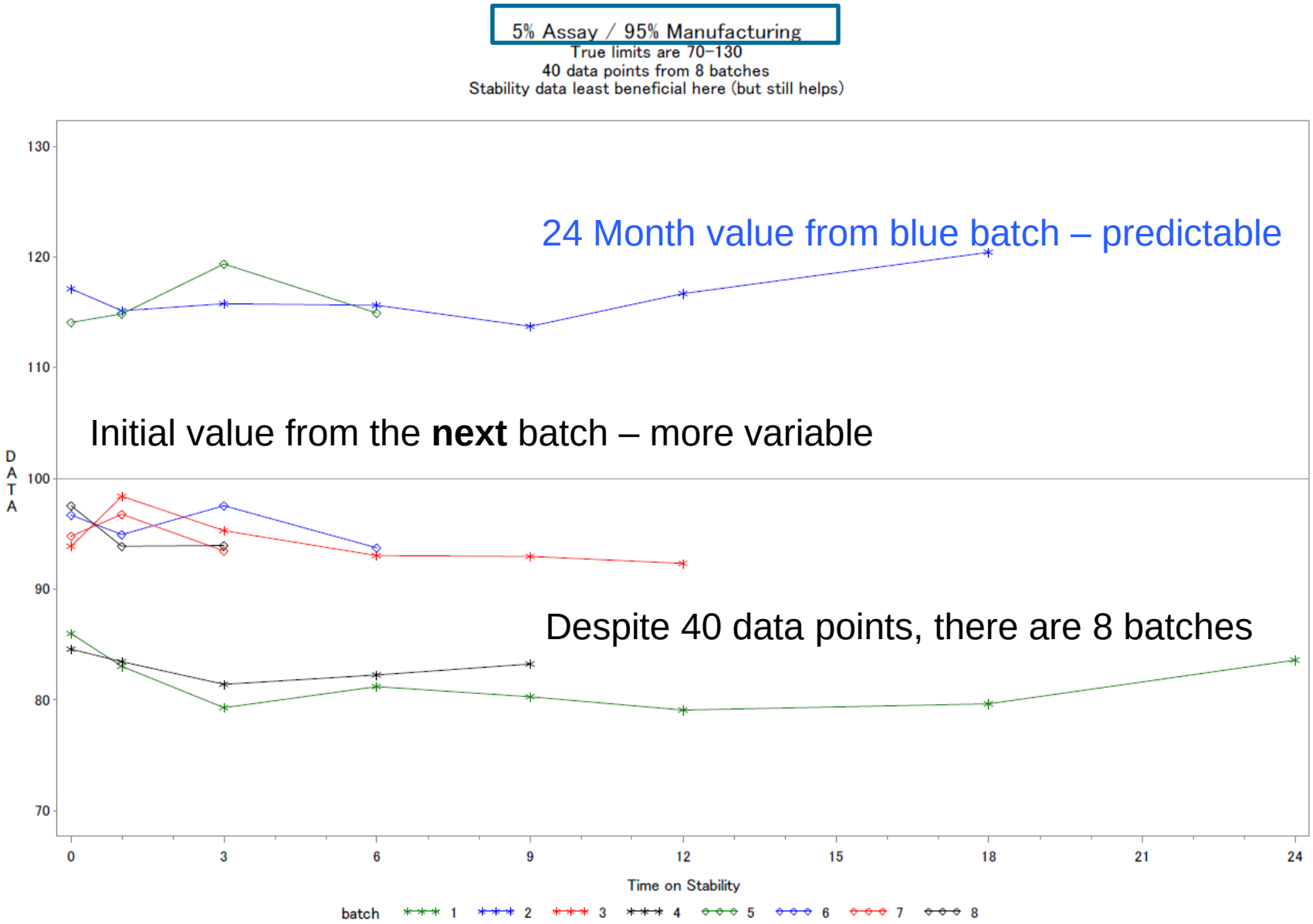


Release / Post Release



More variation seen in the post release data

Stability data (no trend)



Stability data – with trend

When a trend exists, the proposal is:

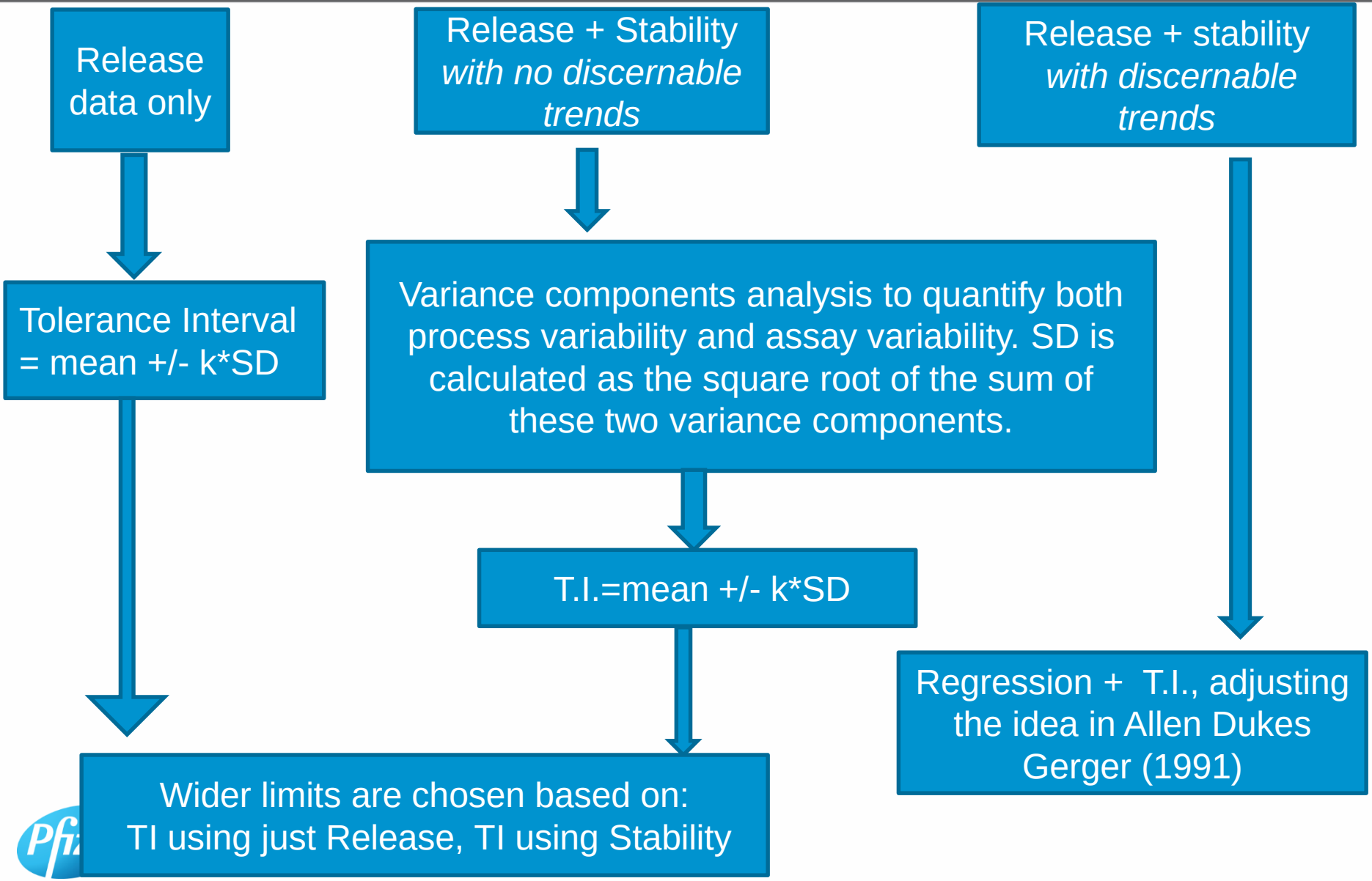
TI(Release) +

Estimated total change over time +

Uncertainty allowance that goes with total change

$$TI(Release) + T_{months} * \hat{\beta} + k * se(T_{months} * \hat{\beta})$$

General workflow



Other considerations

How to chose the slope?

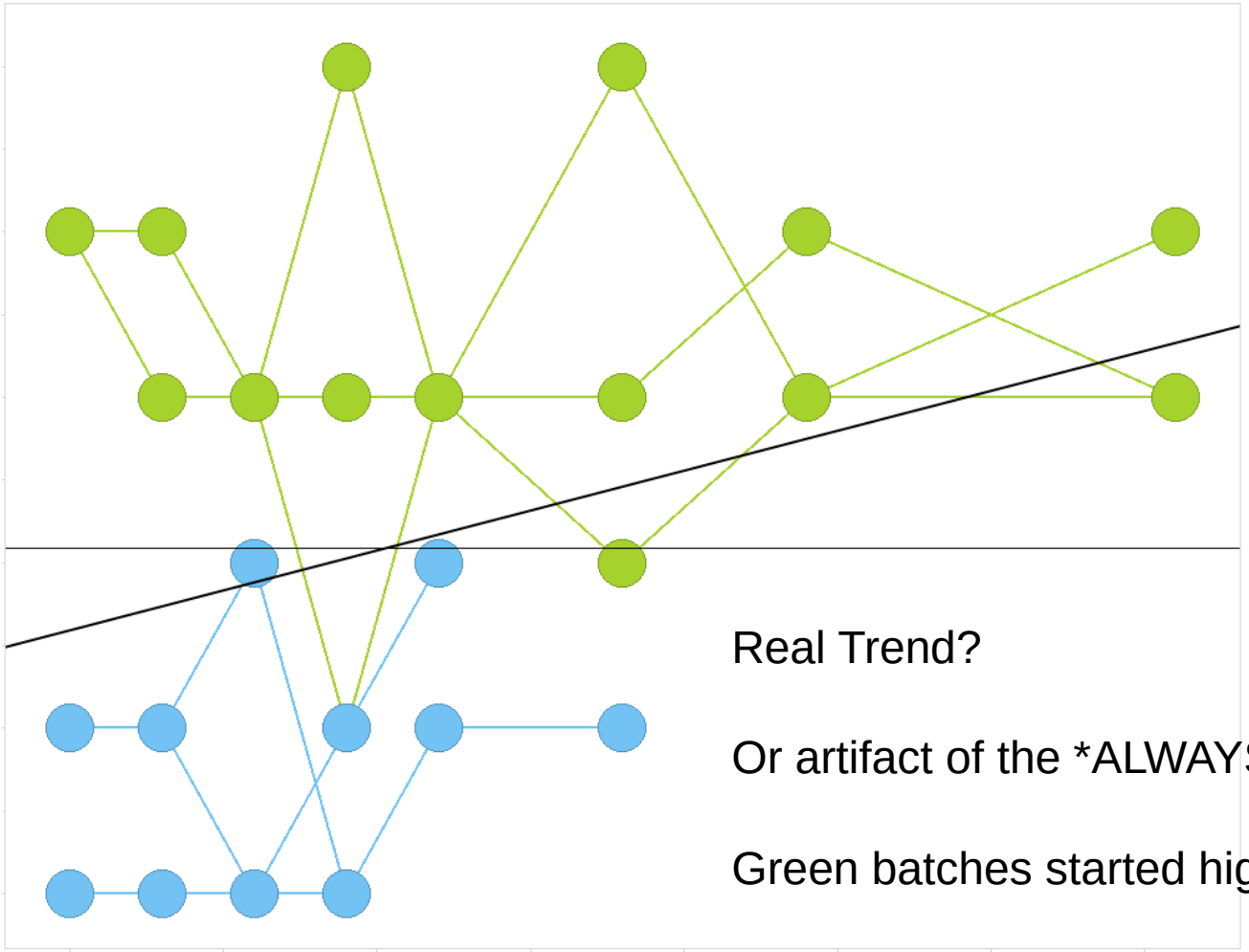
- One slope for all batches?
- Separate slope for each batch?

Random effects models?

Bayesian perspective:

- Similar compounds
- Assay variation prior knowledge

Considerations: over-ride the default model?



Time →

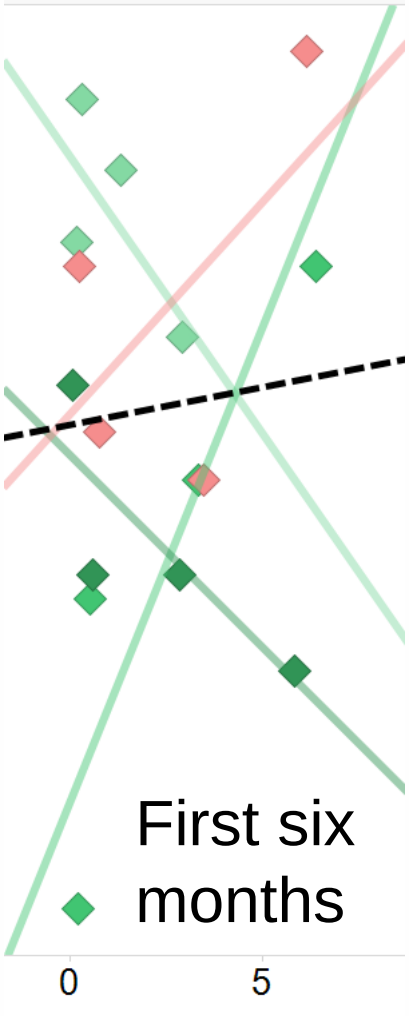
Real Trend?

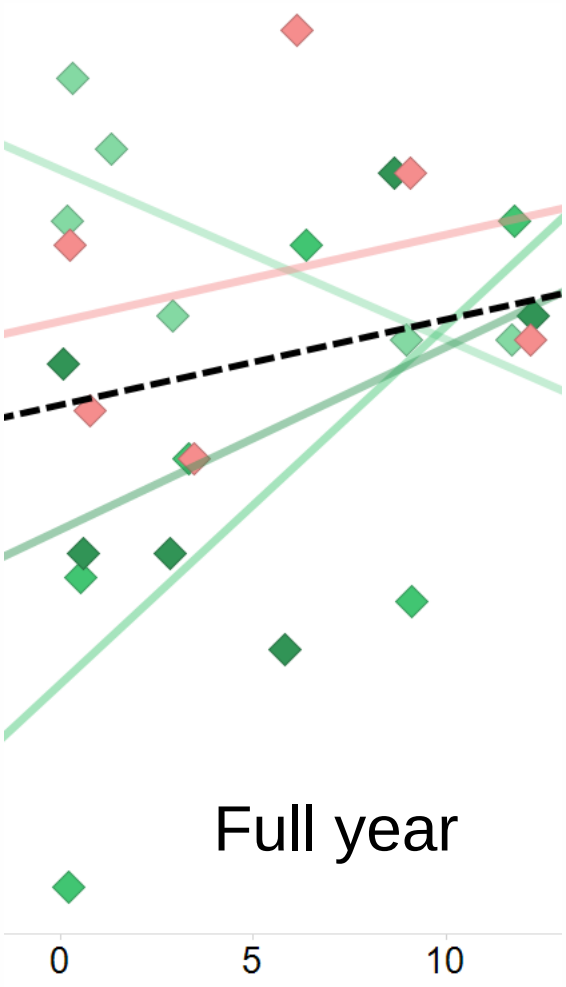
Or artifact of the *ALWAYS* unbalanced design?

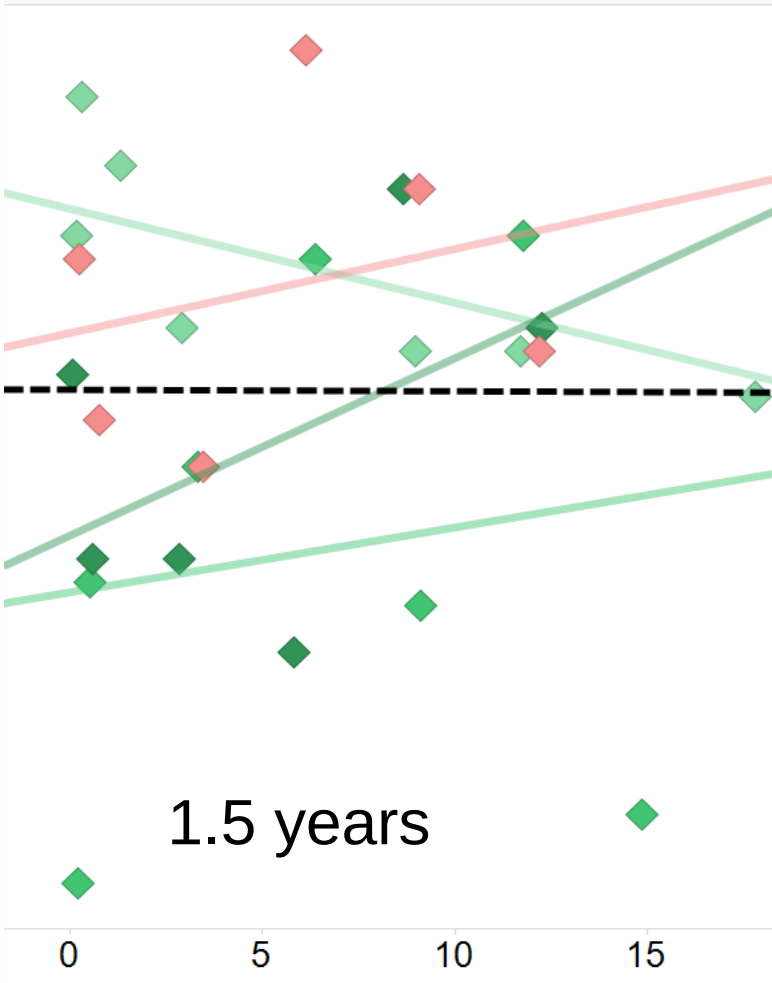
Green batches started higher (for known reasons).

Blue started lower (known reason)

Over-ride the default model? (noisy slopes)







Summary

- Product Specific Specifications are set when possible, but there are times when Data Driven Specifications are needed

- A solution is needed, balancing:
  - Want a high chance to cover future data,
  - Limiting the “white space” beyond current data

- A tolerance interval approach, with increasing confidence and decreasing multipliers as sample size increases

- All data is considered, with or without trend

- Suggestions / comments?

Acknowledgements

Al Annamalia
Mohamed Badreddine
Anthony Carella
Aili Cheng
Brent Harrington
Charlie Kish
Greg Larner
Richard Montes
Vicki Morris
Tony Okinczyc
Vaibhav Patil
Mel Perry
Andrew Rugaiganisa
Les Van Alstine
Kim Vukovinsky
Ke Wang
Jenna Zhang

References

Allen, P. V. (1991) Determination of Release Limits: A General Methodology. Pharmaceutical Research, Vol 8, No. 9

Hahn GJ, Meeker WQ. Statistical intervals: A guide for practitioners (Wiley series in probability and mathematical statistics - applied probability and statistics section). New York: John Wiley & Sons; 1991

ICH Topic Q 6 A Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances

Krishnamoorthy K, Thomas Mathew. Statistical Tolerance Regions: Theory, Applications, and Computations (Wiley series in probability and mathematical statistics). New Jersey: John Wiley & Sons; 2009

https://www.ebe-biopharma.eu/wp-content/uploads/2013/03/28MAR2013-Specifications_final.pdf