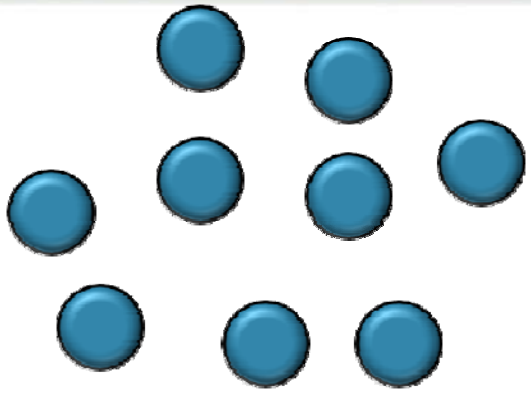
Setting Specifications

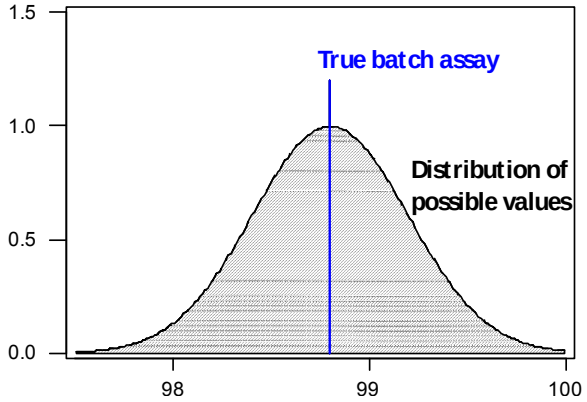
Statistical considerations

Enda Moran – Senior Director, Development, Pfizer
Melvyn Perry – Manager, Statistics, Pfizer

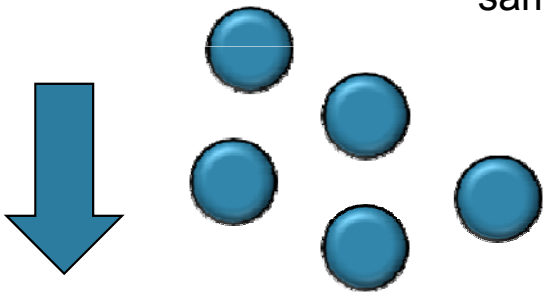
Basic Statistics

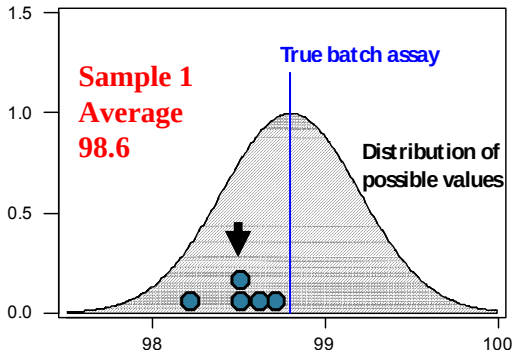


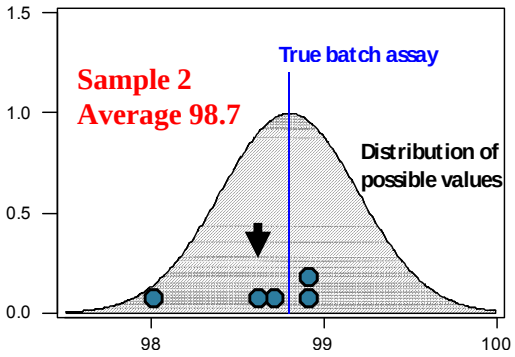
Population distribution (usually unknown). Normal distribution described by μ and σ.

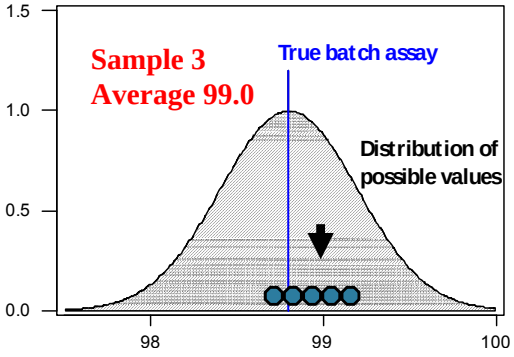


We infer the population from samples by calculating $\bar{x}$ and s.

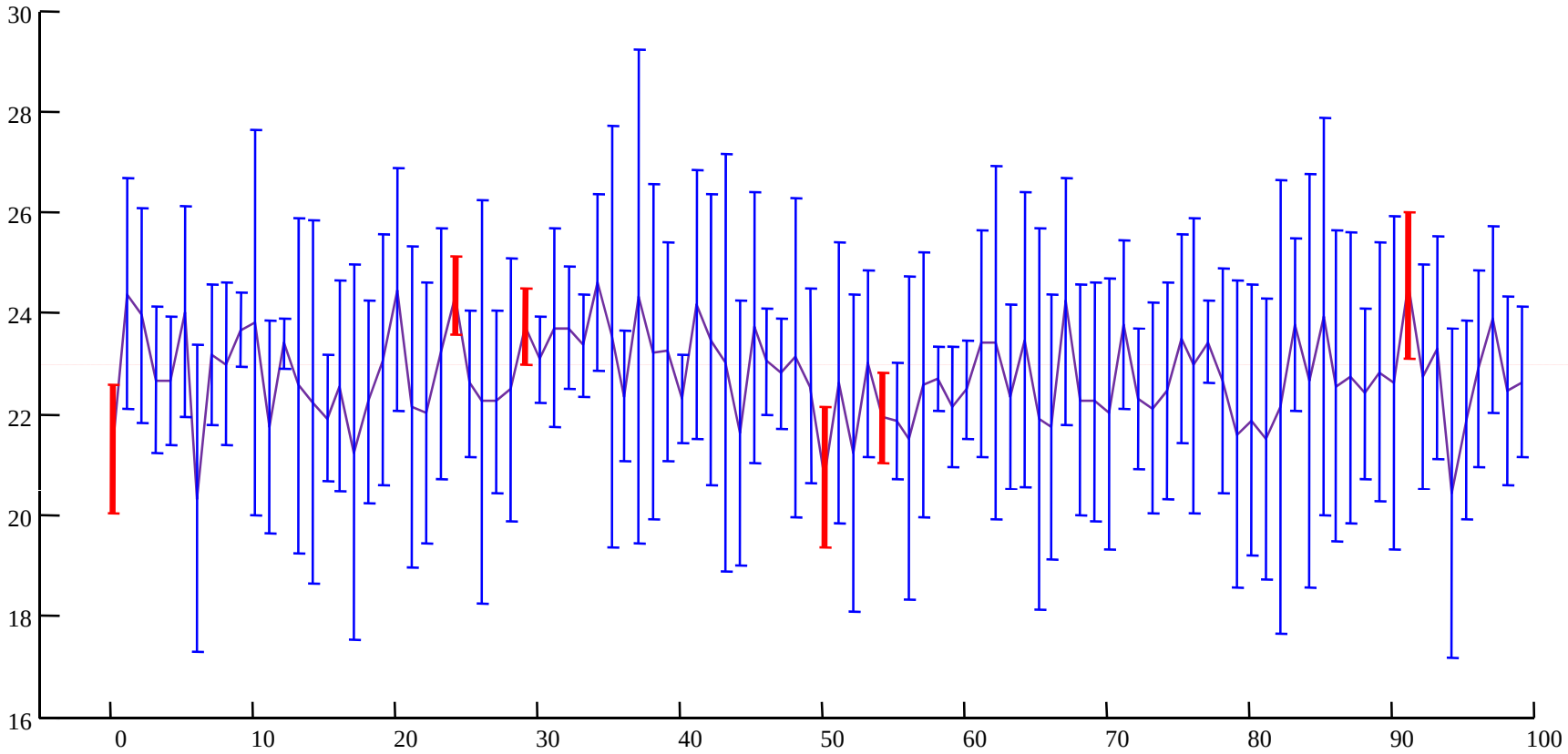








Intervals



- 100 samples of size 5 taken from a population with an average of 23.0 and a standard deviation of 2.0.
- The highlighted intervals do not include the population average (there are 6 of them).
- For a 95% confidence level expect 5 in 100 intervals to NOT include the population average.
- Usually we calculate just one interval and then act as if the population mean falls within this interval.

Intervals

Point Estimation

The best estimate; eg MEAN

Interval Estimation

A range which contains the true population parameter or a future observation to a certain degree of confidence.

Confidence Interval

- The interval to estimate the true population parameter (e.g. the population mean).

Prediction Interval

- The interval containing the next single response.

Tolerance Interval

- The interval which contains at least a given proportion of the population.

Formulae for Intervals

Intervals are defined as: $\bar{x} \pm ks$

Assuming a normal distribution

- Confidence (1-α) interval

$$CI = \bar{x} \pm \left(\frac{1}{n}\right)^{0.5} t_{1-\frac{\alpha}{2},n-1} s$$

- Prediction (1-α) interval for m future observations

$$PI = \bar{x} \pm \left(1+\frac{1}{n}\right)^{0.5} t_{1-\frac{\alpha}{2m},n-1} s$$

- Tolerance interval for confidence (1-α) that proportion (p) is covered

$$TI = \bar{x} \pm \sqrt{\frac{(n-1)\left(1+\frac{1}{n}\right) z^2_{\frac{(1-p)}{2}}}{\chi^2_{\alpha,n-1}}} s$$

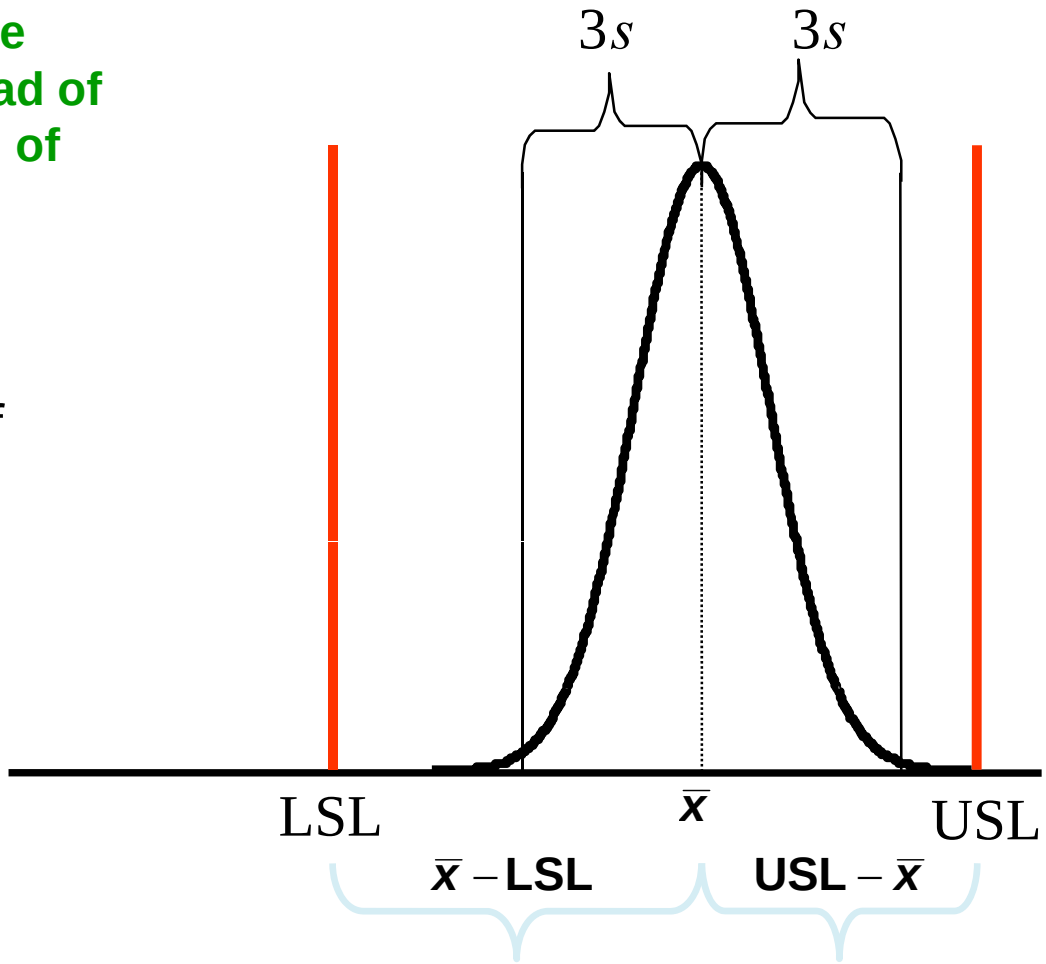
Process Capability

Process capability is a measure of the risk of failing specification. The spread of the data are compared with the width of the specifications.

The distance from the mean to the nearest specification relative to half the process width (3s).

The index measures actual performance. Which may or may not be on target i.e., centred.

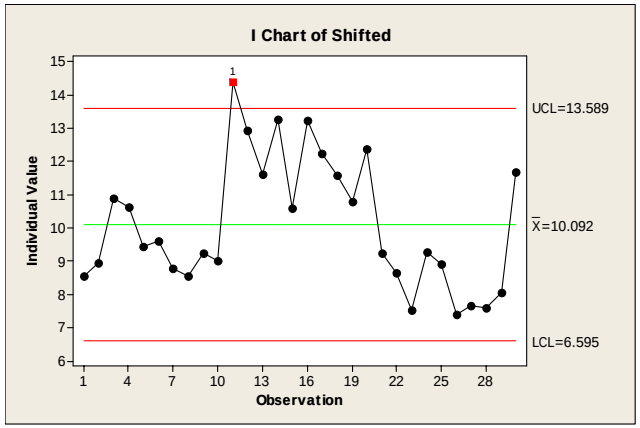
$$P_{pk} = \min\left\{\frac{USL - \bar{x}}{3s}, \frac{\bar{x} - LSL}{3s}\right\}$$

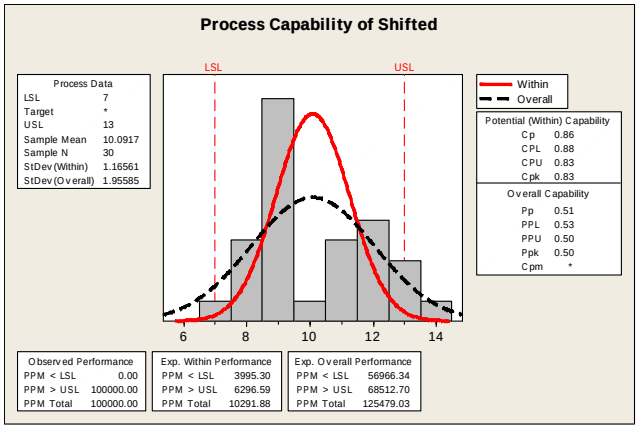


Process Capability – Ppk and Cpk

- Ppk should be used as this is the actual risk of failing specification.
- Cpk is the potential capability for the process when free of shifts and drifts.

Random data of mean 10 and SD 1, thus natural span 7 to 13.
Added shifts to simulate trends around common average.
With specs at 7 and 13 process capability should be unity.

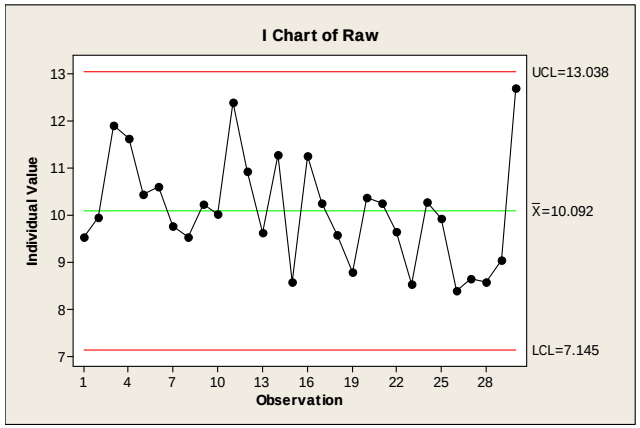


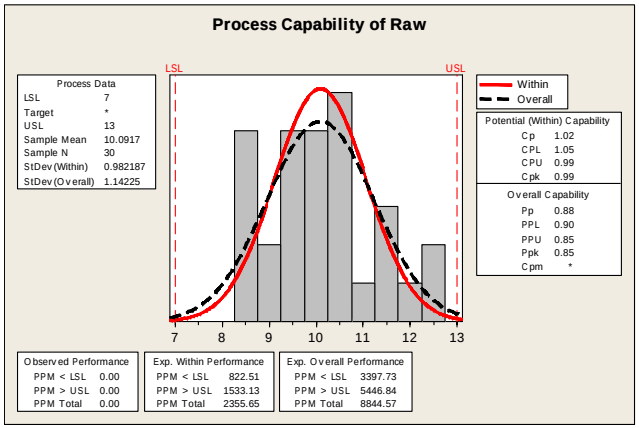


When data is with trend Ppk less than Cpk due to method of calculation of std dev.

Ppk uses sample SD. Ppk less than 1 at 0.5.

Cpk uses average moving range SD (same as for control chart limits).
Cpk is close to 1 at 0.83.



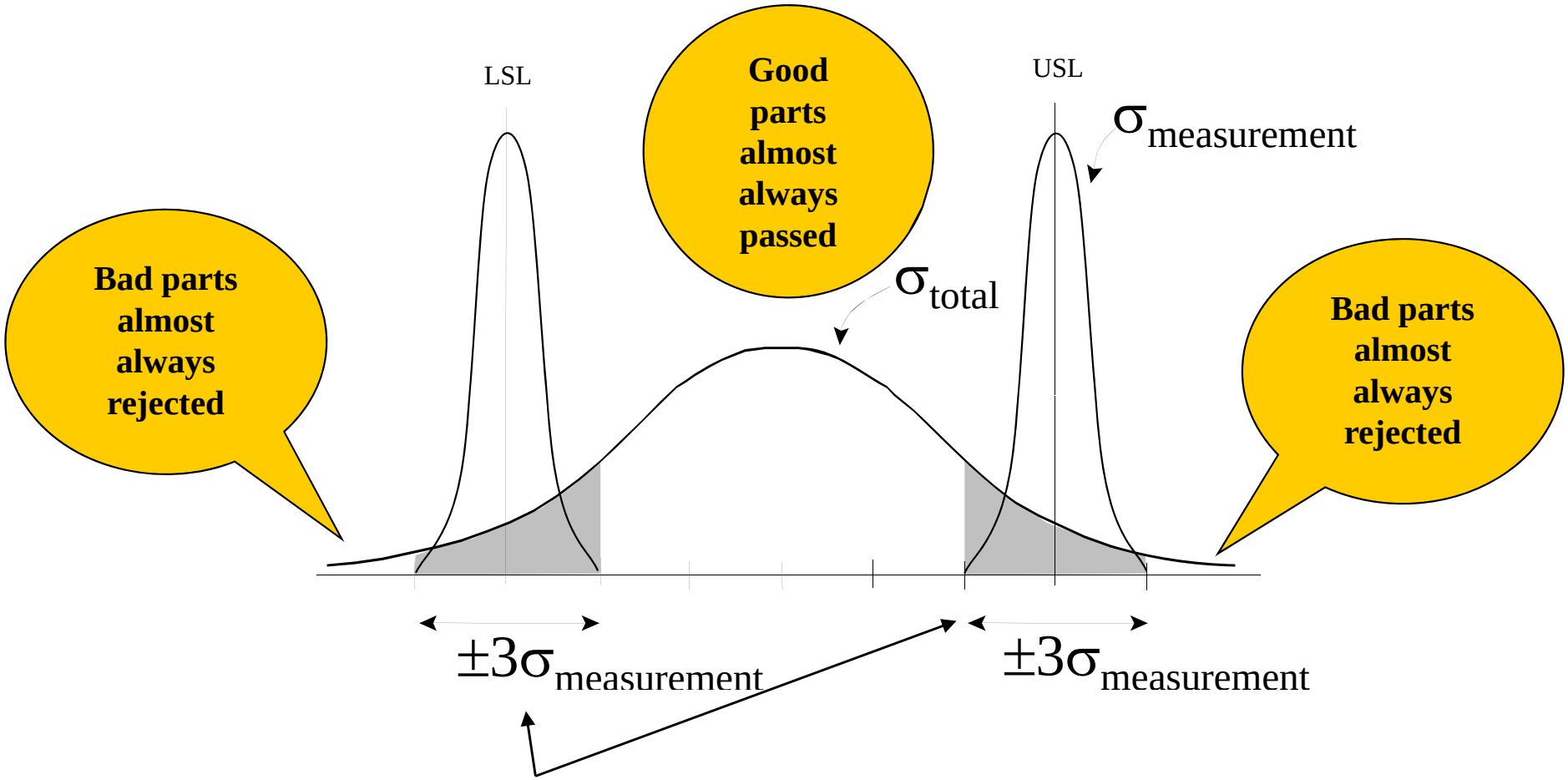


When data is without trend Ppk is same as Cpk. Only small differences are seen.

Cpk effectively 1 at 0.99.

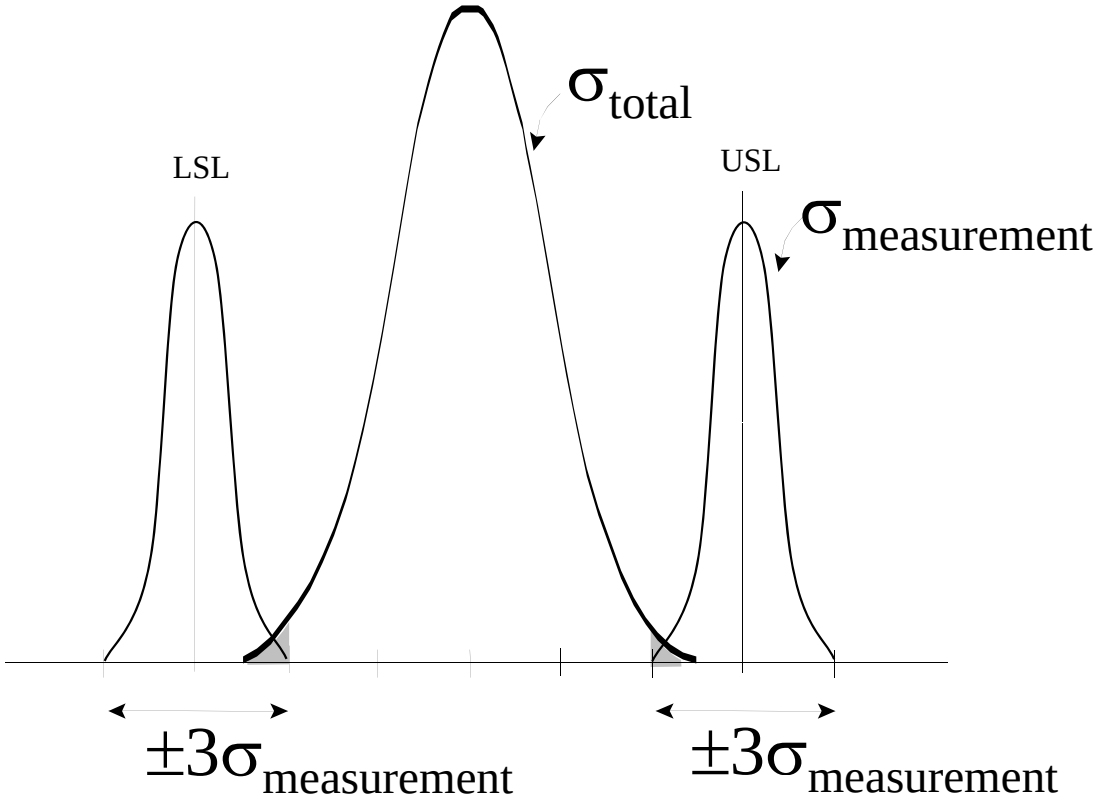
Ppk close to 1 at 0.85.

Measurement Uncertainty



The grey areas highlighted represent those parts of the curve with the potential for wrong decisions, or mis-classification.

Misclassification with less variable process



Measurement Uncertainty

Precision to Tolerance Ratio:

$$P/T = \frac{6 \times \sigma_{MS}}{Tolerance}$$

How much of the tolerance is taken up by measurement error.

This estimate may be appropriate for evaluating how well the measurement system can perform with respect to specifications.

$Tolerance = USL - LSL$

$USL = \text{Upper Spec Limit}$

$LSL = \text{Lower Spec Limit}$

$\sigma_{MS} = \text{Std. Dev. of Measurement Sys.}$

Gage R&R (or GRR%)

$$\%R\ \&\ R = \frac{\sigma_{MS}}{\sigma_{Total}} \times 100$$

What percent of the total variation is taken up by measurement error (as SD and thus not additive).

Use Measurement Systems Analysis to assess if the assay method is fit for purpose. It is unwise to have a method where the specification interval is consumed by the measurement variation alone.

Specification example – Tolerance Interval

Data from three sites used to set specifications.

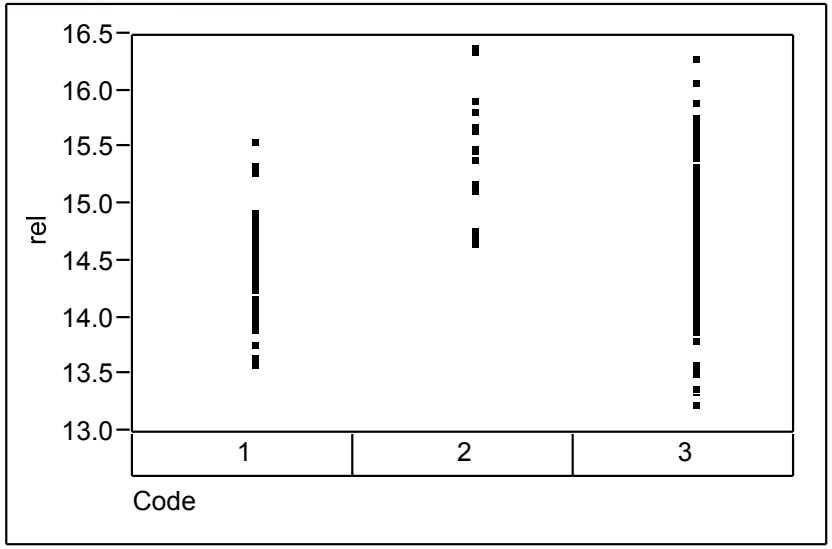
Tolerance interval found from pooled data of 253 batches.

Tolerance interval chosen as 95% probability that mean ± 3 standard deviations are contained.

Sample size: n=253	Mean ± 3s	Tolerance interval
Mean = 14.77 s=0.58		(95% / 99.7%)
Range	13.03 - 16.51	12.89 - 16.65

If sample size was smaller, difference between these calculations increases.

As the sample size approaches infinity the TI approaches mean ± 3s.

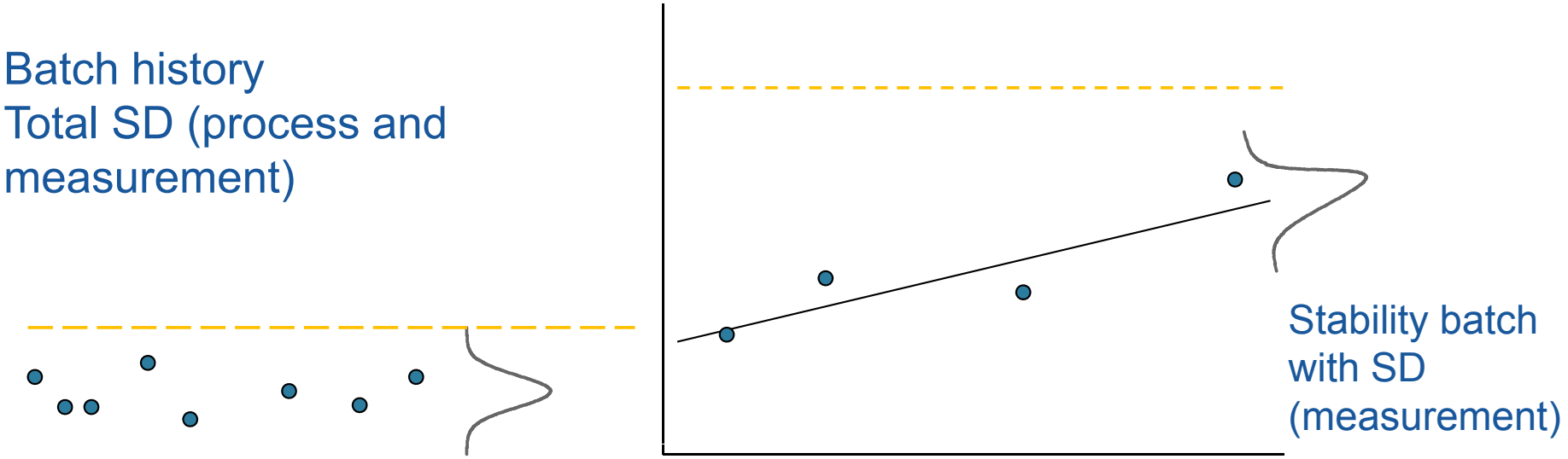


N	k multiplier
5	6.60
10	4.44
15	3.89
30	3.35
∞	2.58

Table of values for 95% probability of interval containing 99% of population values.

Note at ∞ value is 2.58 which is the z value for 99% coverage of a normal population

Specification Example - Stability



Shelf life set from 95% CI on slope from three clinical batches.

Need to find release criteria for high probability of production batches meeting shelf life based on individual results being less than specification.

Review of batch history will lead to a process capability statement against release limit.

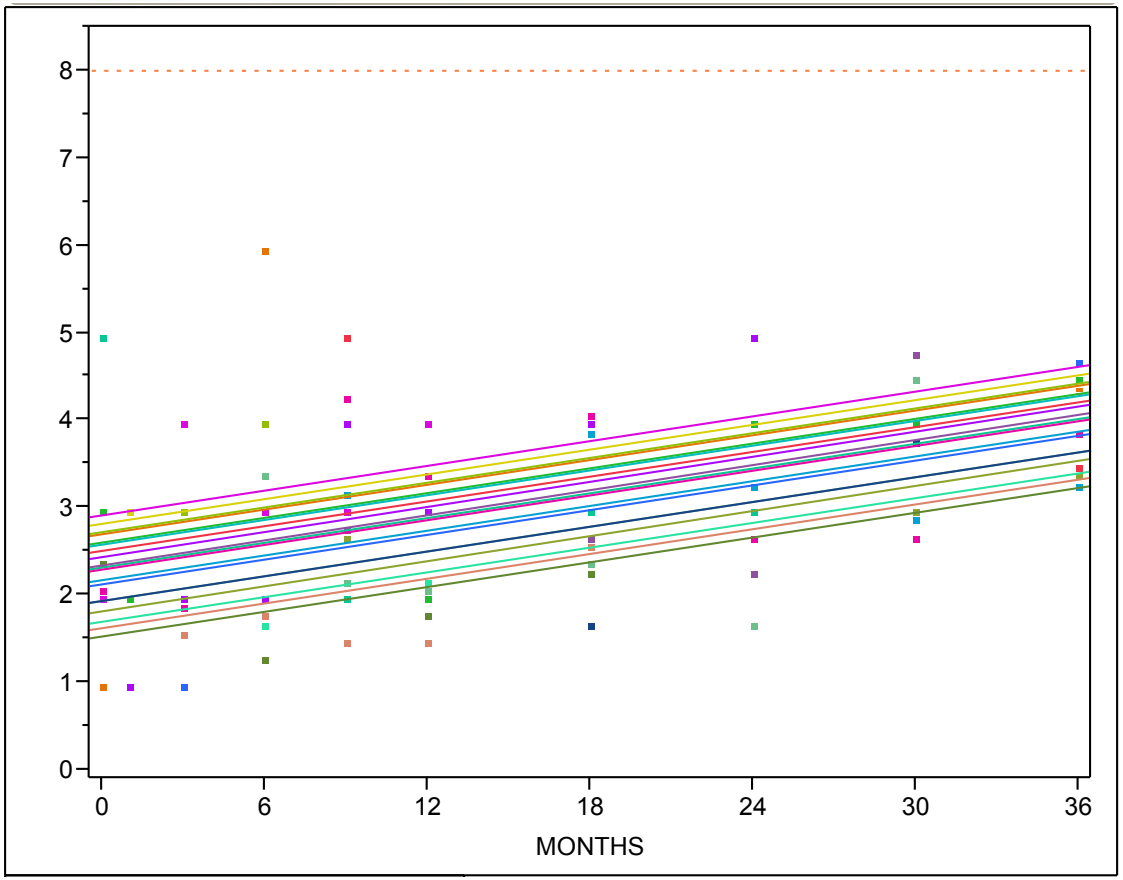
Specification Example - Stability

Release limit is calculated from the rate of change and includes the uncertainty in the slope estimate (rate of change of parameter) and the measurement variation).

$$RL = Spec - \left(Tm + t\sqrt{T^2 s_m^2 + \frac{s_a^2}{n}} \right)$$

Specification = 8
Shelf life (T) = 24
Parameter slope (m) = 0.04726
Variation in slope (s_m) = 0.00722
Variation around line (s_a) = 0.86812
t on 126 df approx. = 2
N= 1 for single determination

Release limit = 5.1



A similar approach can be taken in more complex situations; e.g,, after reconstitution of lyophilised product. Product might slowly change during cold storage and then rapidly change on reconstitution.

*Allen et al., (1991) Pharmaceutical Research, **8**, 9, p1210-1213*

Coping with limited data

Problems with n<30.

- Difficult to calculate a reliable estimate of the SD.
- Difficult to assess distribution.
- Difficult to assess process stability.

With very limited data the specifications will be wide due to the large multiplier.

- This reflects the uncertainty in the estimates of the mean and SD.

Is there a small scale model that matches full production scale?

How variable is the measurement system?

Alternatives are:

- Min / max values based on scientific rationalisation
- Mean ± 3s
- Mean ± 4s or minimum Ppk = 1.33.
- Tolerance interval approach

Example with 4 values

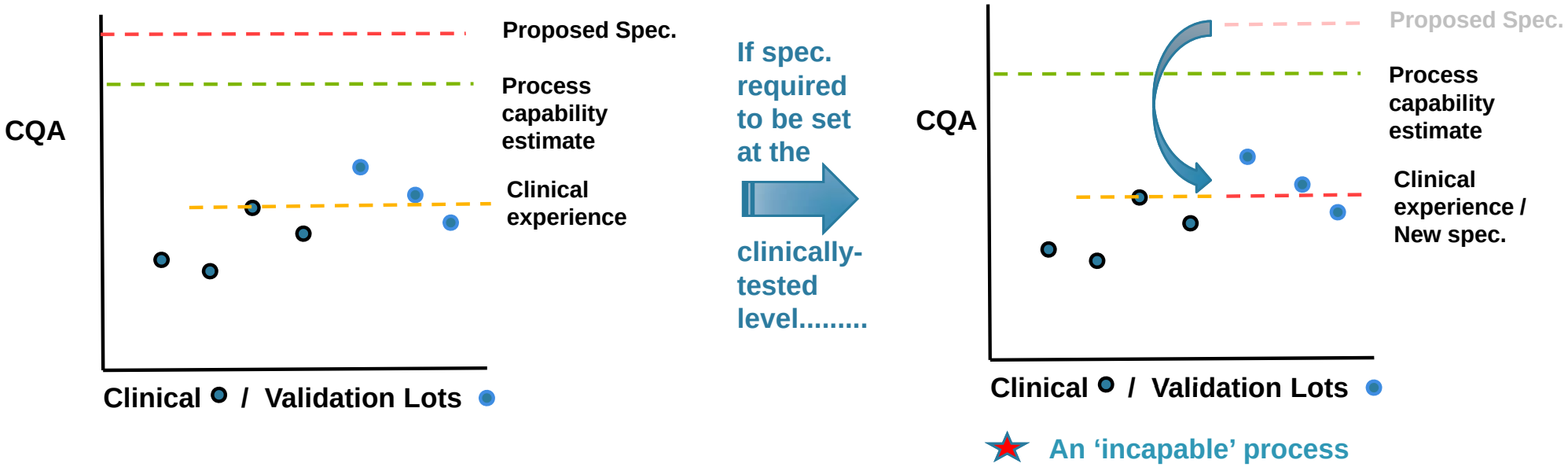
N(20,4) – 19.9, 16.9, 22.1, 21.3

Mean ± 3s	13.2	27.0
Mean ± 4s	10.9	29.3
TI (95/99.7)	-1.5	41.7

Updating Specifications

When, why, and a scenario for consideration

- Specifications for CQAs might require revision during the product lifecycle. For example, new clinical data and intelligence on continued use of a product might drive a change to a CQA specification
- Scenario:
  What if a process producing a new product approaching regulatory approval is 'incapable' of meeting the clinically-tested CQA in the long-term? For example.....



- How could we manage this situation? Complex. Many factors need to be considered if setting specifications at clinically tested level *(will be discussed through this workshop)*
- IF it must be done for risk reduction purposes, and if there is sufficient uncertainty that the Proposed Spec. for the CQA could cause harm, a possible approach may be to move towards the clinically-tested level in stages.
  - Accrue approx. n=25-30 batches/lots to consider estimates of future process capability acceptable. Adjust spec. to a revised process capability estimate.
  - Accrue further lots of the stabilised process to approx. n=85. Make FINAL spec. adjustment to a revised process capability estimate.

Summary

- Process capability & measurement uncertainty
- Setting spec.s - large number of datapoints
- Setting spec.s – small number of datapoints