Nestlé | Research

How many data for "process capability"?

Scott Hindle

Manufacturing Department

Nestlé PTC Konolfingen

March 24th, 2016

Behind the smile… In the dry runs yesterday I was consistently above 15 minutes… Am I now capable of hitting this 15 minute target?

The points below hope to hold your attention!

- The reason for collecting and analysing data
- Being a statistician isn't always easy
- Some fundamentals surrounding process capability
- Degrees of freedom for a measure of dispersion
- Giving degrees of freedom an understandable and communicable meaning to non-statisticians
- Some examples to better understand the "*How many data…?*" question
- Conclusions

The reason for collecting and analysing data

Only statisticians collect and analyse data for fun

For the rest of the world:

The Only Reason to Collect Data is to Take Action!

As so well stated by William Scherkenbach, former Corporate Director of Statistical Methods at Ford Motor Company

Only statisticians collect and analyse data for fun

For the rest of the world:

The Only Reason to Collect Data is to Take Action!

Some key points

- *Action is unlikely to happen unless the data represent something important*
- *Action may not happen (when it should) if data are analysed ineffectively*
- *Ineffective actions may result if data are analysed ineffectively*
- *The expected benefits of the action (improvements in quality, productivity, other cost element etc.) should be understood and communicable*
- *etc.*

Being a statistician isn't always easy

Setting the scene

One of the most asked questions:

How many data do I need to for my process capability study?

Setting the scene

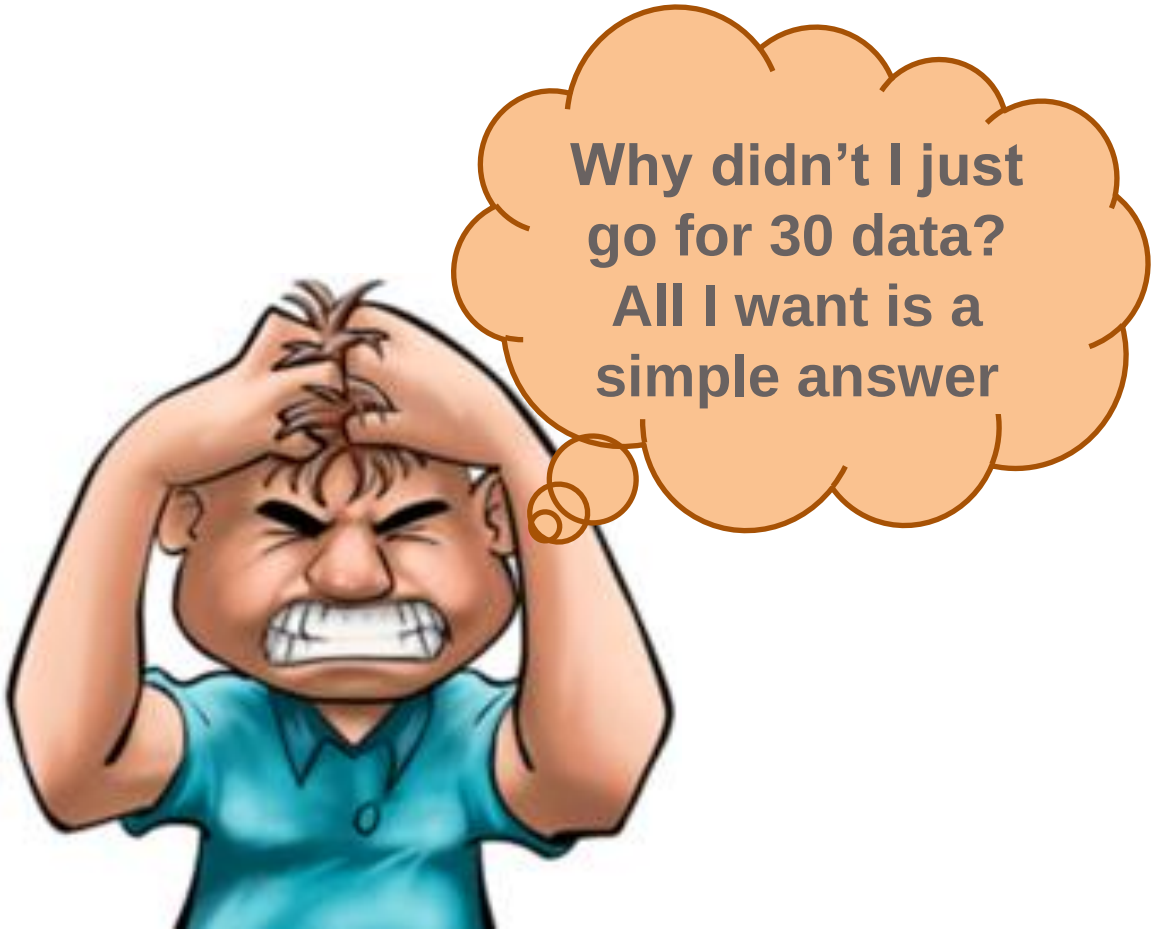
Probably the most frequent answer to the "*How many data?*" question:

Setting the scene

Probably followed by a question like…

Setting the scene

Might the reaction be something like this?



Setting the scene

The statistician might now be in “recovery mode”

Setting the scene

To which the colleague (searching assistance!) might think

Some fundamentals surrounding process capability

Getting started…

- A manufacturer's question: *How good is a production process?*
  - This important question needs answering
  - But, as a rather vague question it needs to be made more specific

- How about this: "*Is fully conforming process output expected?*"
  - "Process capability" can help, but only if the collected data form a basis for action (remember Scherkenbach's message)

- So, data are needed, but how many?

Process Capability

- *Definition*: Uses actual data from the process to assess if the process output is acceptable or not based on the customer's requirements and expectations (i.e. specifications)
- *Definition*: Quantifies the relationship between the Voice of the Customer and the Voice of the Process

A more personal definition:

- "Process capability" provides a basis for action on the process if:
  - *A predictable process is characterised as not capable*
  - *A predictable process is operating off-target*
  - *A process is characterised as unpredictable*

  (The only other possible outcome is predictable, capable and on-target)

Predictable process → Statistically controlled, or "in control", process (within 3-sigma limits)
Unpredictable process → Not in statistical control ("out of control")

Putting “process capability” into numbers

- Many capability indexes are in circulation (see Bothe, 1997)
- We will focus on the commonly used C_p and C_{pk}

$$C_p = \frac{USL - LSL}{6 \times SD_{within}}$$

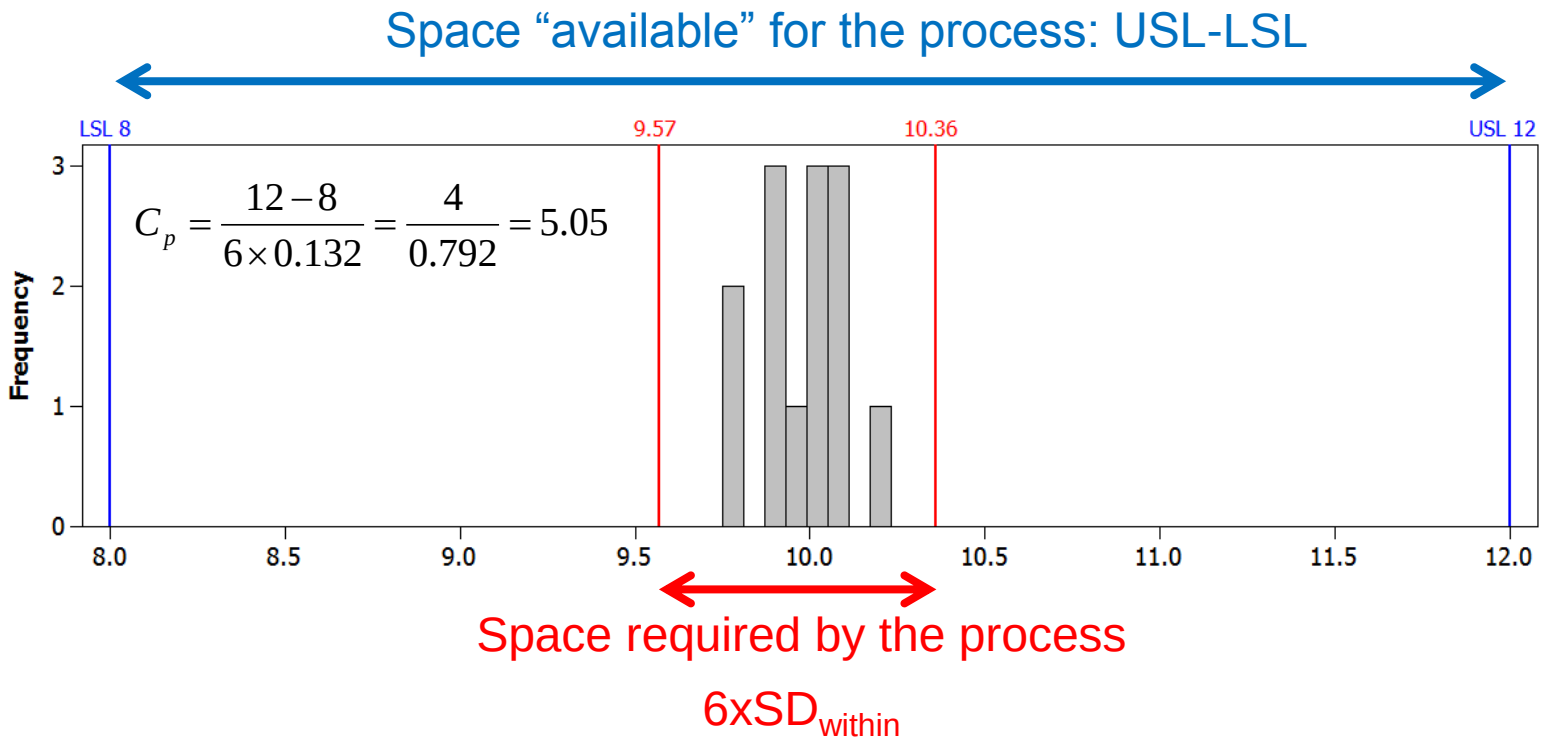
$$C_{pk} = \min\left\{ \frac{\bar{X} - LSL}{3 \times SD_{within}} ; \frac{USL - \bar{X}}{3 \times SD_{within}} \right\}$$

- To compute these statistics we need:
  - Voice of the Customer – the specifications (e.g. LSL and USL)
  - Voice of the Process – in the formulas represented by summary statistics for location and dispersion
  - SD_{within} is a *within*-subgroup dispersion statistic (or an average or median dispersion statistic) → not a global measure of dispersion

Putting “process capability” into a picture

- Let us visualise a “good” capability in terms of C_p

$$C_p = \frac{USL - LSL}{6 \times SD_{within}}$$

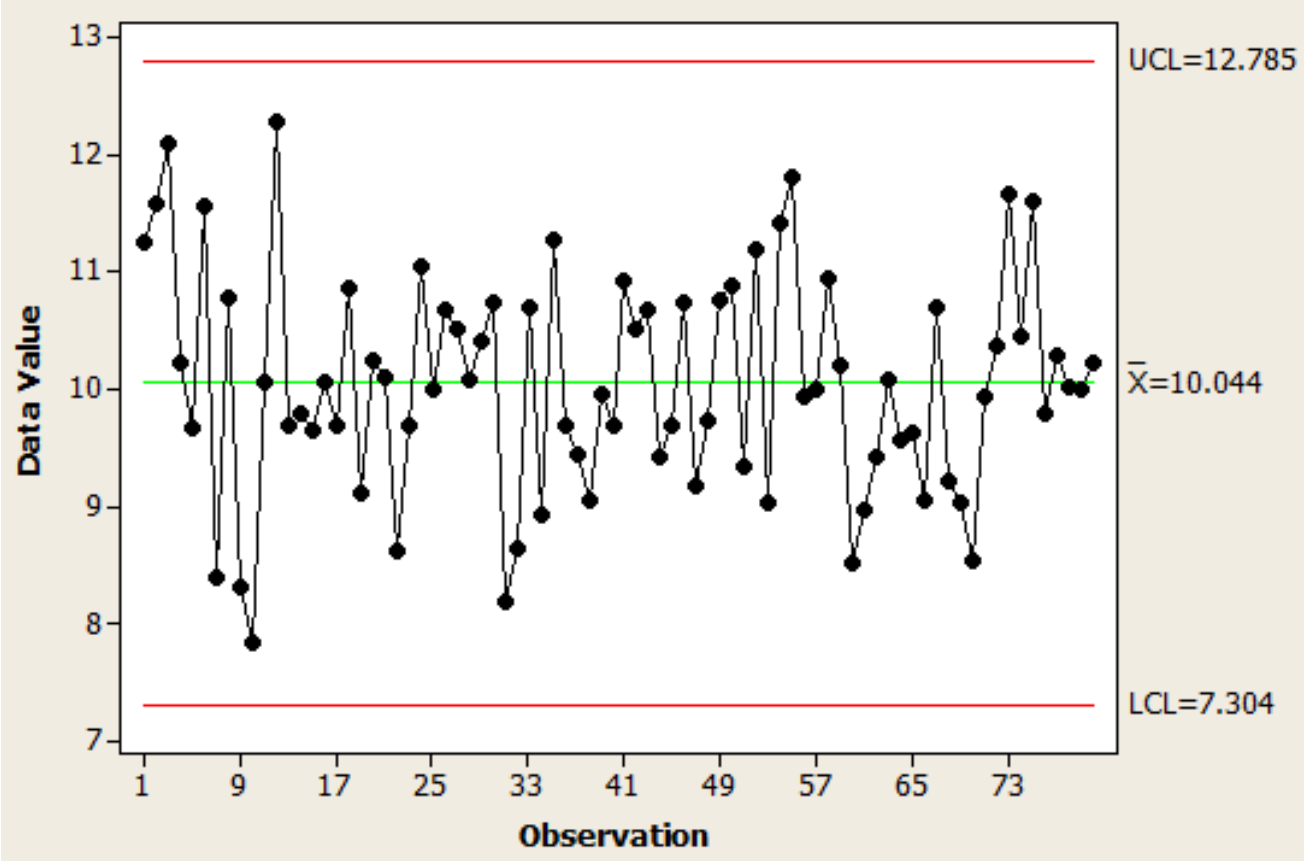


- Here the space available for the process is ~500% wider than the space required by the process

When the Voice of the Process is well-defined

A process in “statistical control” is said to speak with one voice

Process data that allow for a characterisation of predictable process behaviour (“in statistical control”) will look something like this:

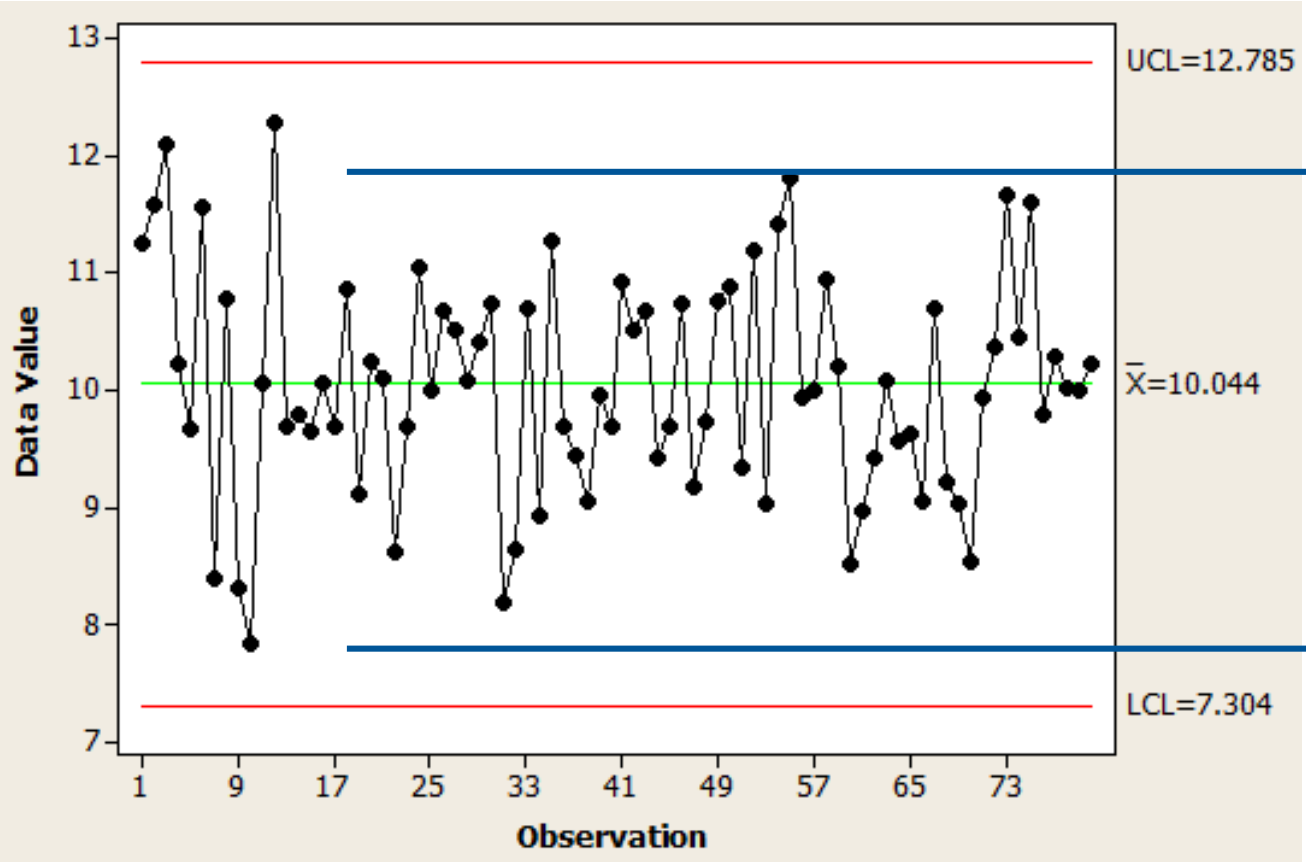


Future “routine” process output predicted to fall within these limits

Nestlé Research

A well-defined Voice of the Process is not enough

What if the specifications fall inside the natural limits of the process – the “red lines” (3-sigma limits)?



Customer accepts product in this range

Capable process?

Control chart for individual values

Nestlé Research

Degrees of freedom for a measure of dispersion

Standard deviation and degrees of freedom

- The sample standard deviation statistic, *s*, has a well-defined number of degrees of freedom (d.f.) equal to *n* minus one

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}}$$

- d.f. can help to an answer to the "*How many data?*" question

- But, are d.f. understandable and communicable in the workplace?

Giving degrees of freedom an understandable and communicable meaning to non-statisticians

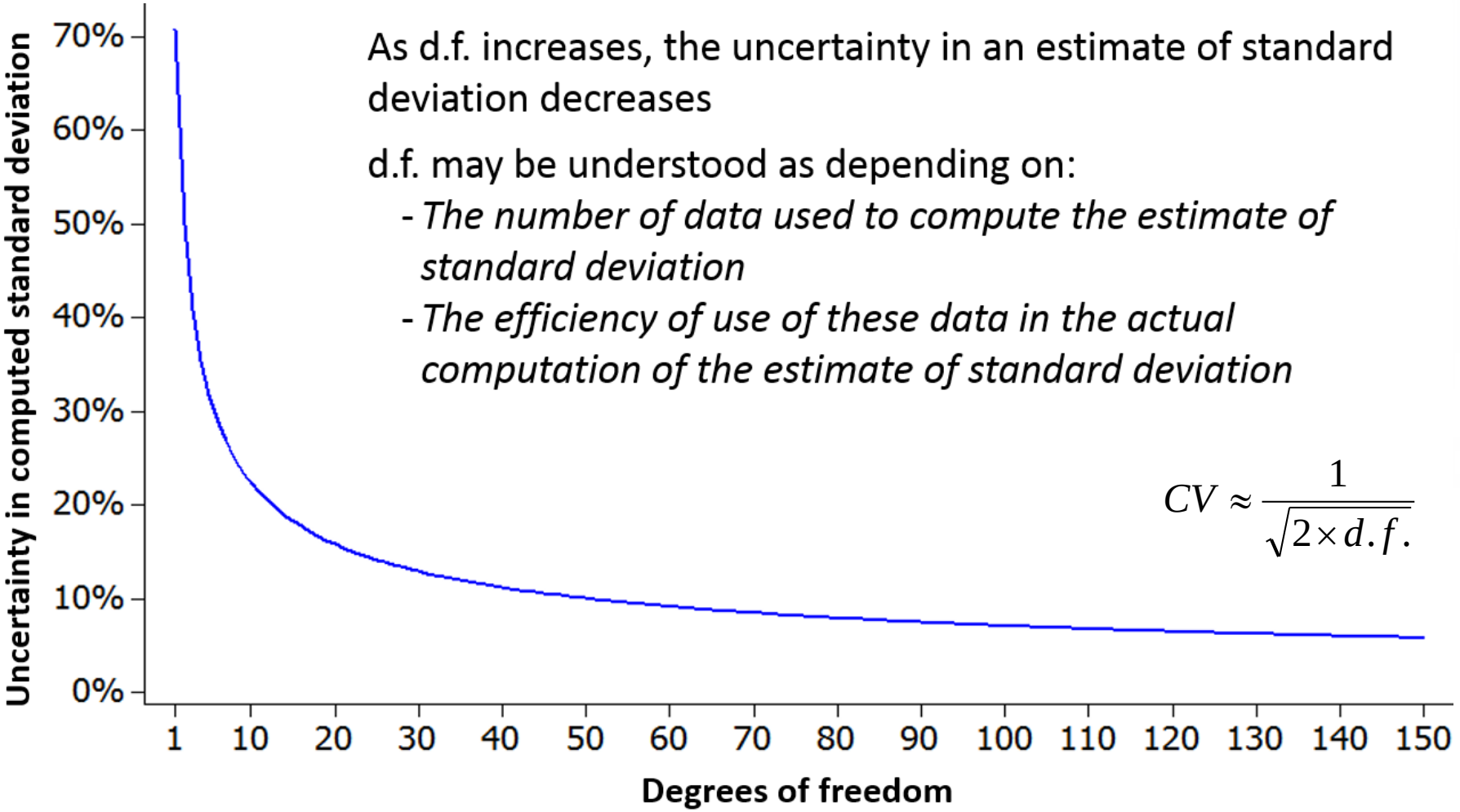
Degrees of freedom and uncertainty

- Using d.f., we can speak of the uncertainty associated with an estimate of standard deviation by using its CV (Coefficient of Variation):
  - *CV is obtained from the ratio of the standard deviation of a variable to the mean of the variable*

- It can be shown that (Wheeler, 2004) $CV \approx \frac{1}{\sqrt{2 \times d.f.}}$

- The relationship, on the next slide, between CV and d.f. is non-linear

Degrees of freedom and uncertainty CV



As d.f. increases, the uncertainty in an estimate of standard deviation decreases

d.f. may be understood as depending on:

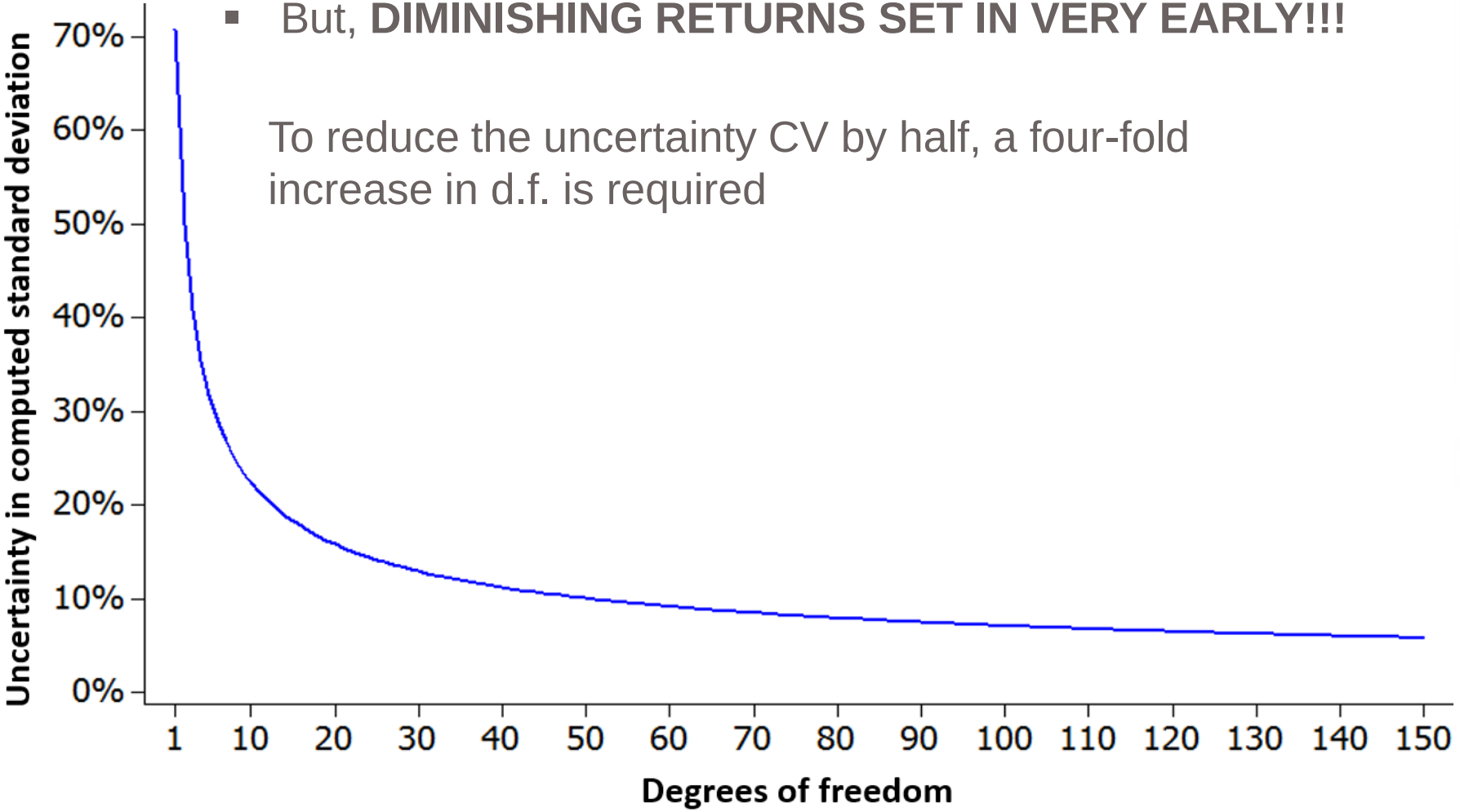
- *The number of data used to compute the estimate of standard deviation*
- *The efficiency of use of these data in the actual computation of the estimate of standard deviation*

$$CV \approx \frac{1}{\sqrt{2 \times d.f.}}$$

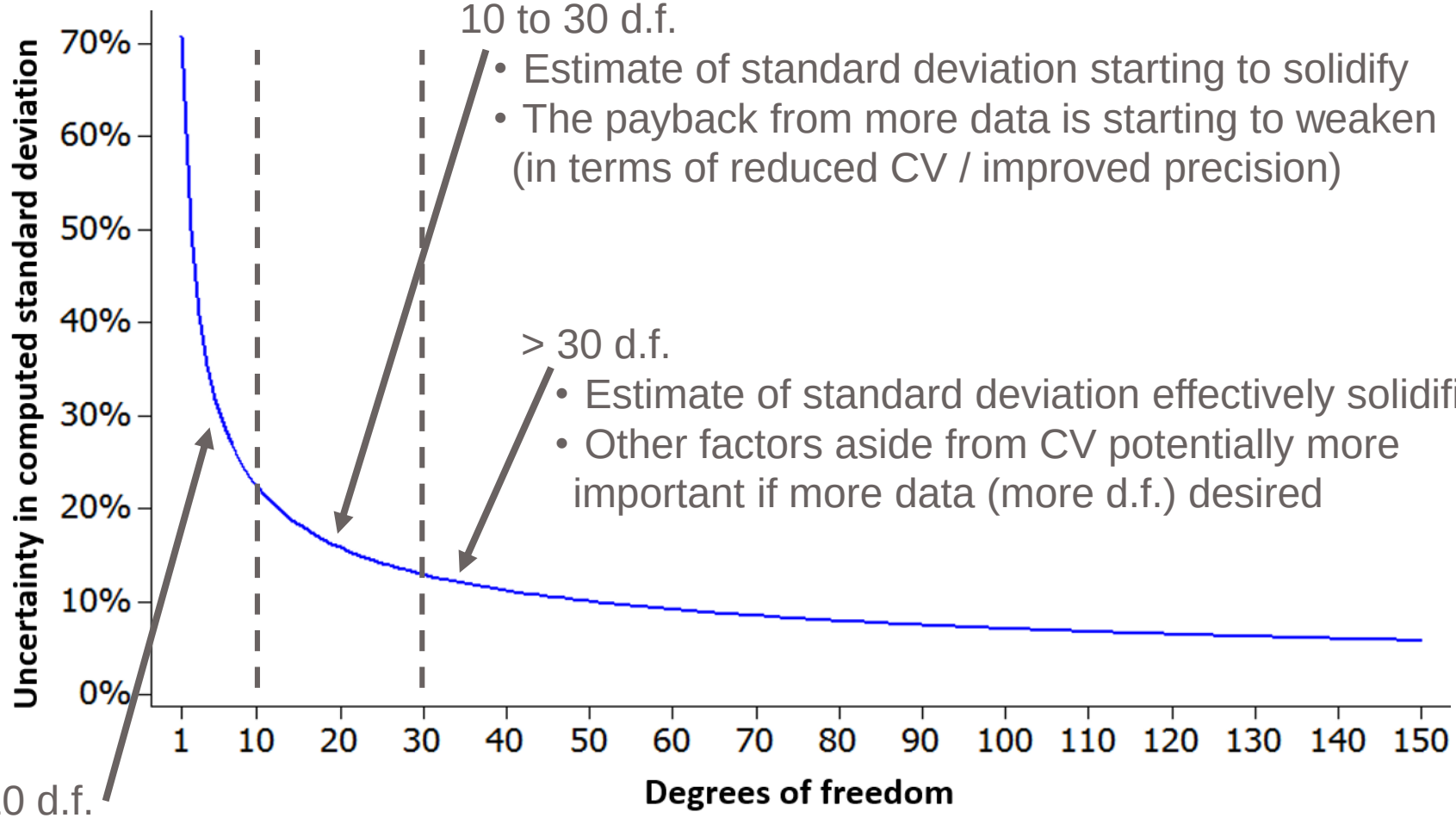
An important learning from this graph

- The first few d.f. are very important, perhaps critical
- But, **DIMINISHING RETURNS SET IN VERY EARLY!!!**

To reduce the uncertainty CV by half, a four-fold increase in d.f. is required



Interpreting this relationship in practice:
How "good" is your estimate of standard deviation?



10 to 30 d.f.
- Estimate of standard deviation starting to solidify
- The payback from more data is starting to weaken (in terms of reduced CV / improved precision)

> 30 d.f.
- Estimate of standard deviation effectively solidified
- Other factors aside from CV potentially more important if more data (more d.f.) desired

< 10 d.f.
- High uncertainty in estimated standard deviation
- More data likely to be highly desired (in most/many cases)

Process capability and SD_{within}

- *s* as a global estimator of dispersion, is not suitable for process capability applications (or for control charts which are always needed to make sense of process capability statistics)

- The examples herein are based on the use of individual data

- To estimate SD_{within} when using individual data – no subgrouping – we can use the "average moving range" method

$$\overline{mR} = \sum_{i=1}^{n-1} \left| X_{i+1} - X_i \right| \Big/ (n-1)$$

$$SD_{within} = \overline{mR} / d_2 = \overline{mR} / 1.128$$

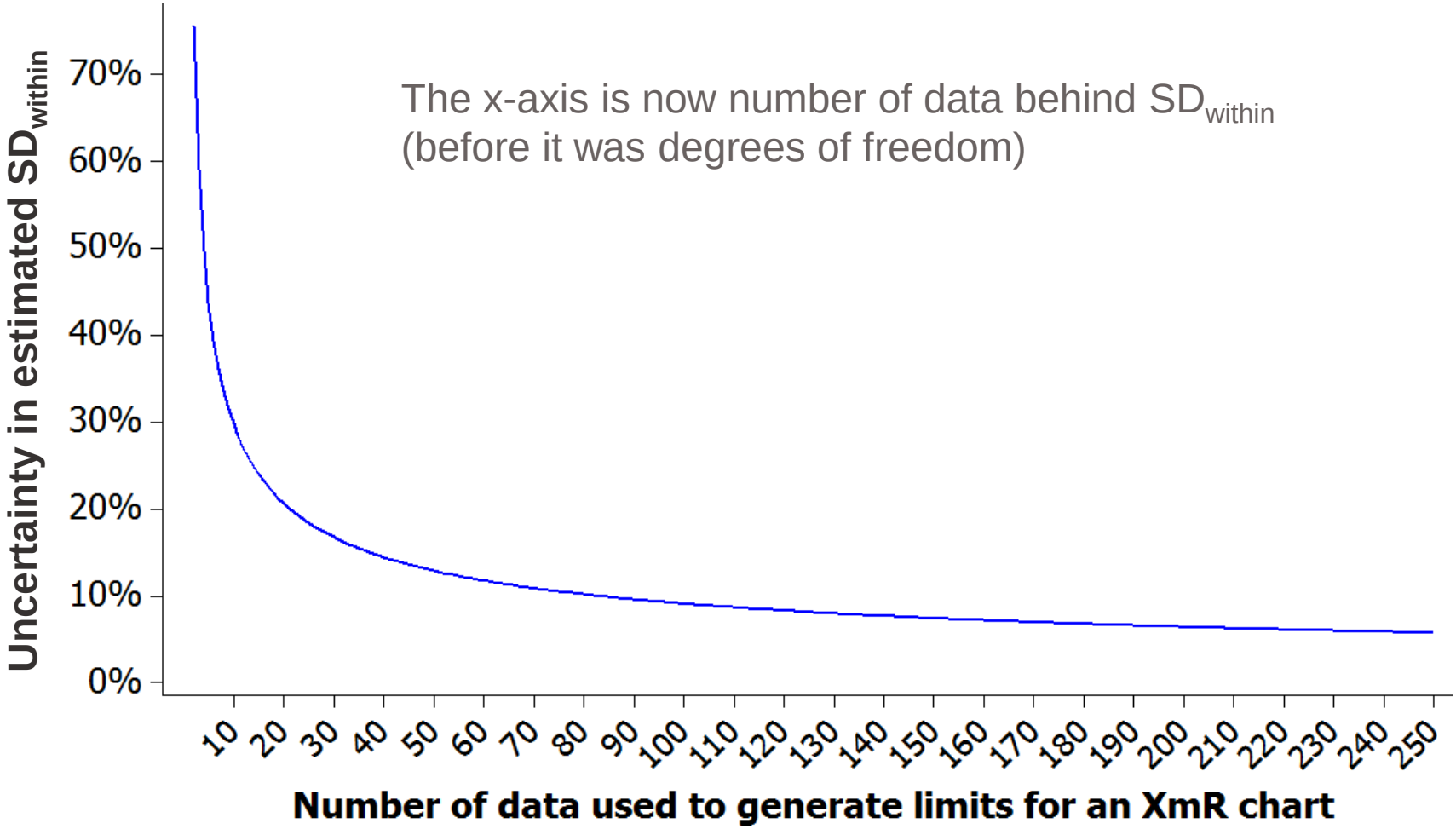
- But, SD_{within} does not have d.f.=*n*-1 (where *n* means the total number of data used)

Effective number of degrees of freedom

- As shown by Wheeler, 2004, for the average moving range method: *Effective* d.f. ≈ 0.62 x (*n*-1)
  - Where *n* is the total number of individual data values we have

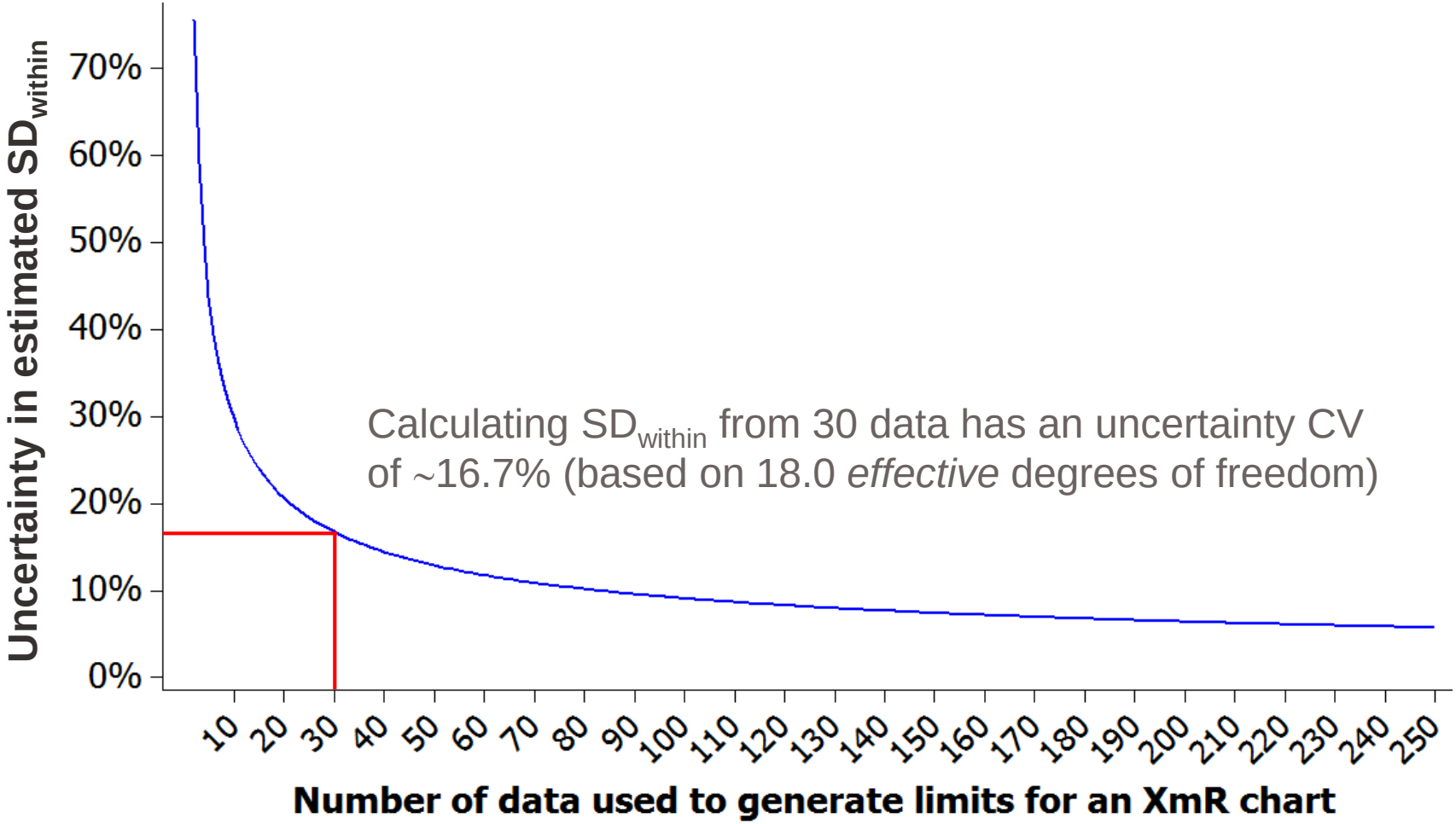
- Example:
  - 30 individual data have an *effective* number of d.f. 18.0 and an associated uncertainty of ~16.7% in the calculated value for SD_{within} (Wheeler, 2004).

CV vs. number of data using the average moving range

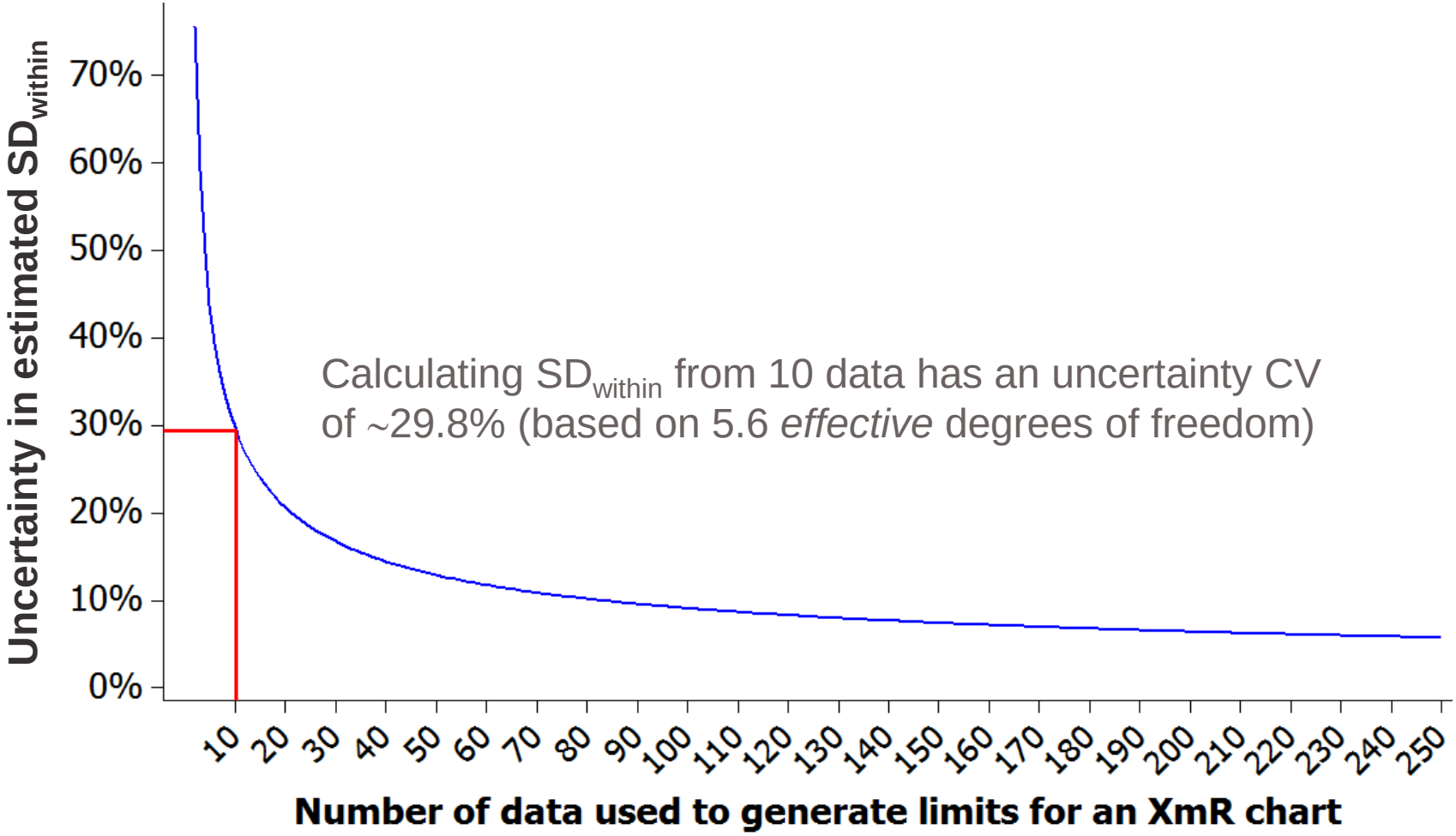


The x-axis is now number of data behind SD_{within} (before it was degrees of freedom)

CV vs. number of data using the average moving range



CV vs. number of data using the average moving range



Summary to this point

- Process capability compares the Voice of the Process with the Voice of the Customer
  - *A capable process is one where the Voice of the Process fits the Voice of the Customer ("good news" for a manufacturer)*
- The Voice of the Process is based on a *within*-subgroup estimate of dispersion (hence the name SD_{within})
- The Voice of the Process is only well-defined if the process is characterised as predictable ("statistical control" on a control chart)
- Statistical theory allows us to estimate the uncertainty in SD_{within} :
  - *How "solid", or "soft", is SD_{within}?*
  - *The uncertainty CV is well-defined (as an estimate of a standard deviation parameter) only if the process is characterised as predictable*

Some examples to better understand the "*How many data…?*" question

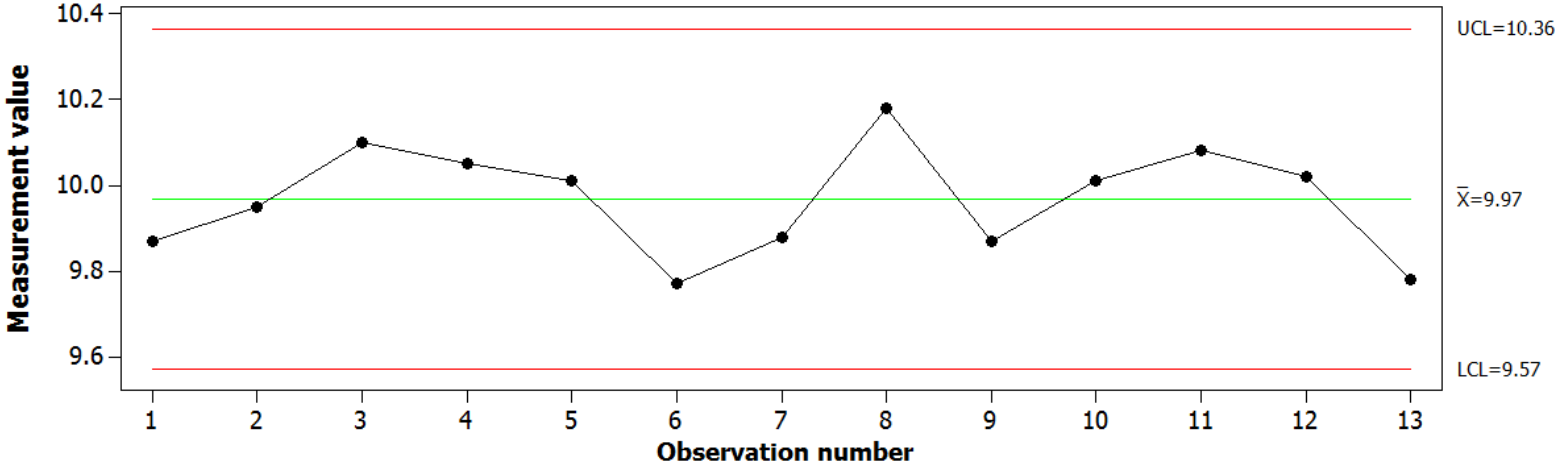
Example: 13 data values

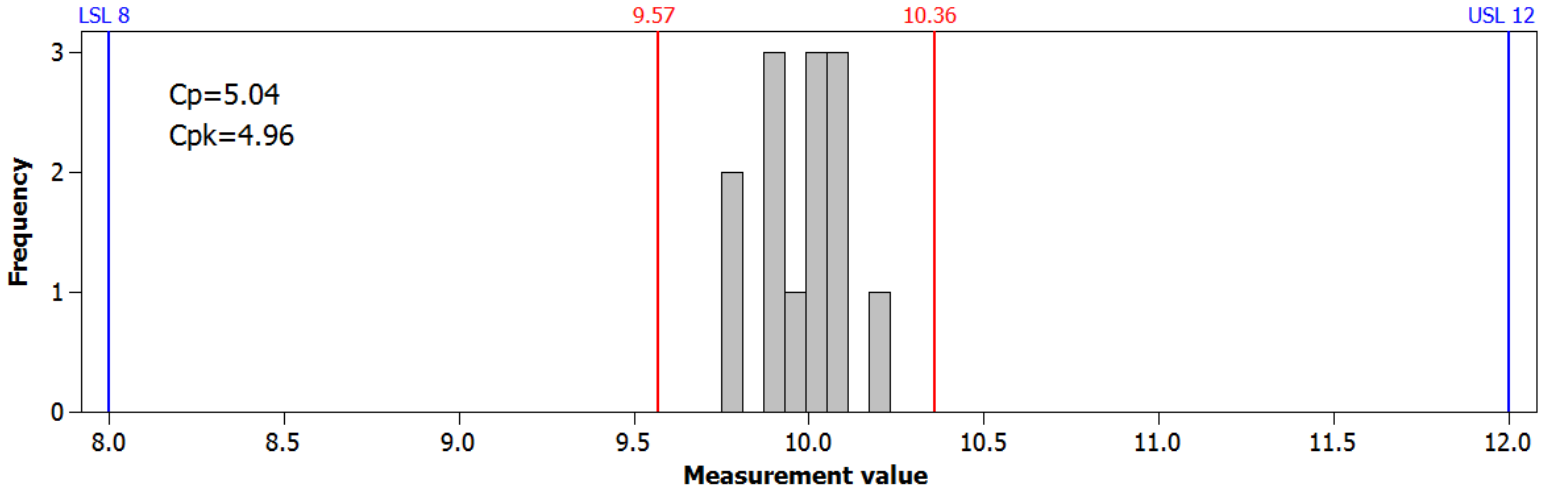
The context

- The production process is in operation every three to four weeks, and one data value per production run is judged appropriate
- Specifications are LSL=8 and USL=12
- Process target is 10, the midpoint of the specifications
- After some 9 months you have 13 data values
- 13 values see a high uncertainty in SD_{within} at ~25.5%

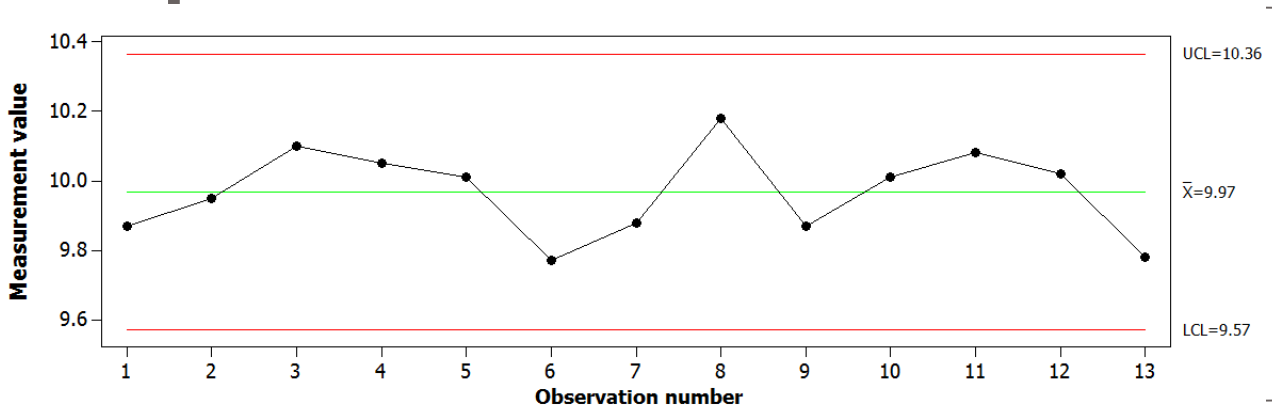
Example: 13 data values

- Below an X chart and histogram with specs and process limits



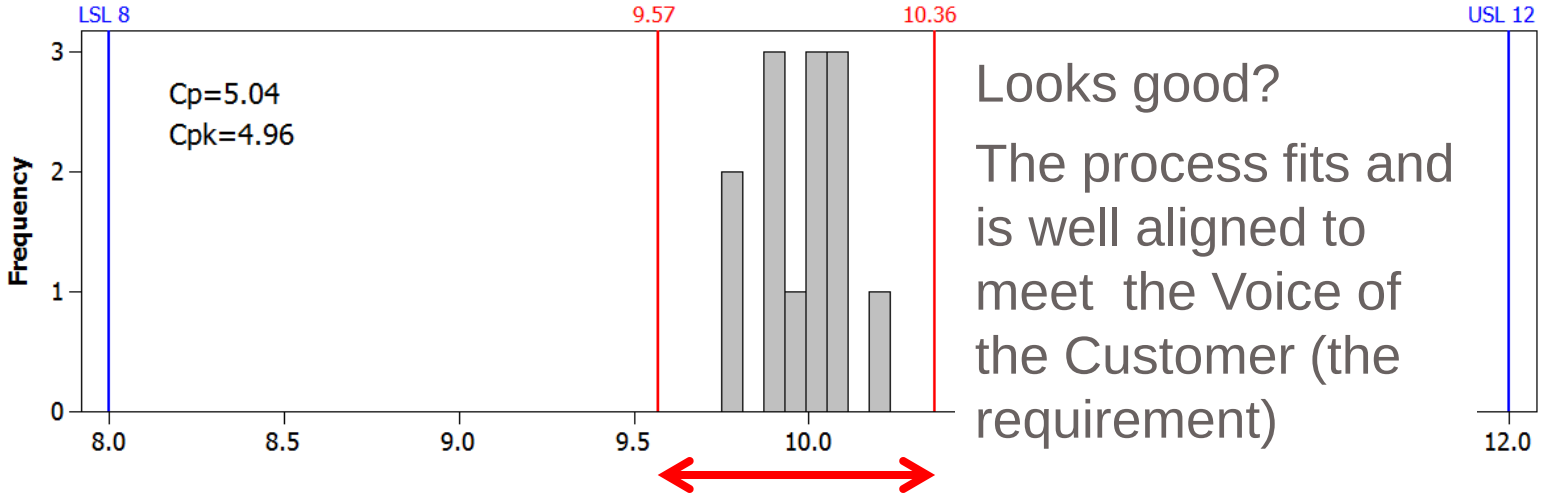


Example: 13 data values



This chart is consistent with a predictable process

Space “available” for the process



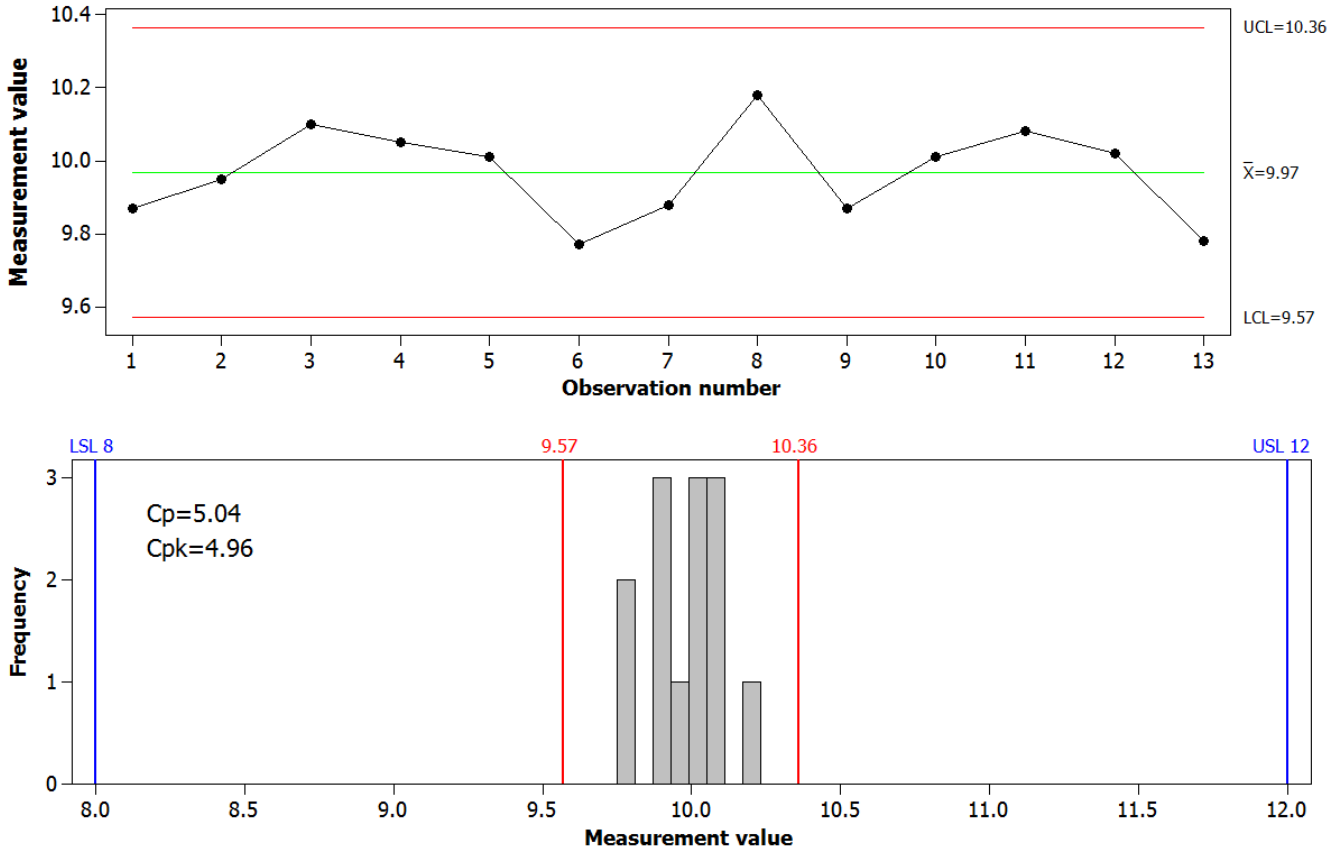
Looks good?

The process fits and is well aligned to meet the Voice of the Customer (the requirement)

Space required by the process

Nestlé Research

Example: 13 data values



- Question: *Do we have enough data?*
- *Reminder: It's taken around 9 months to get these 13 values!*

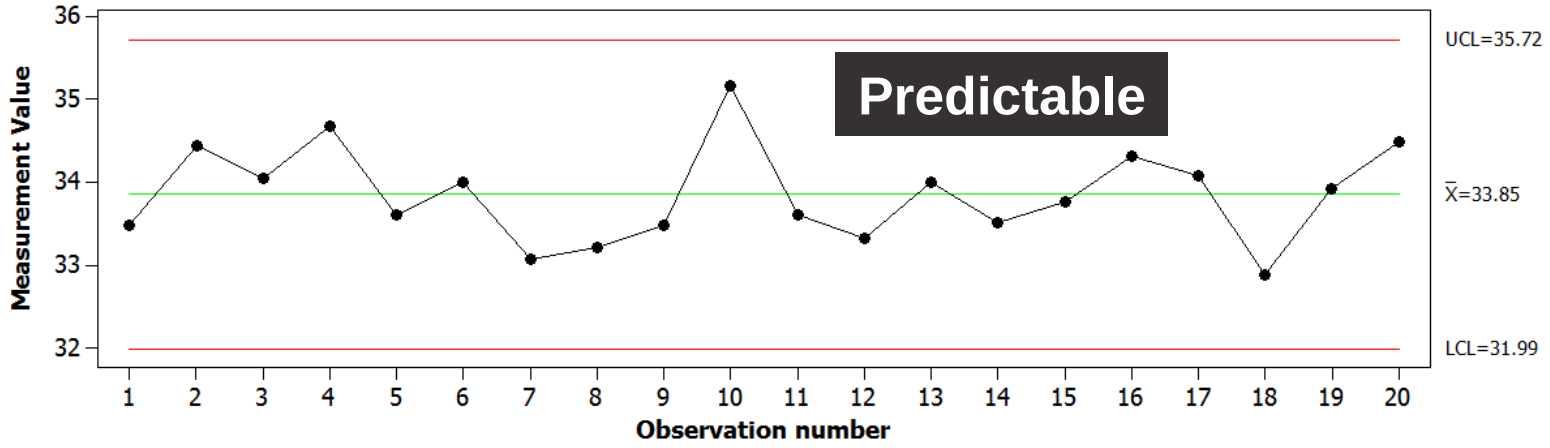
Example: 20 data values

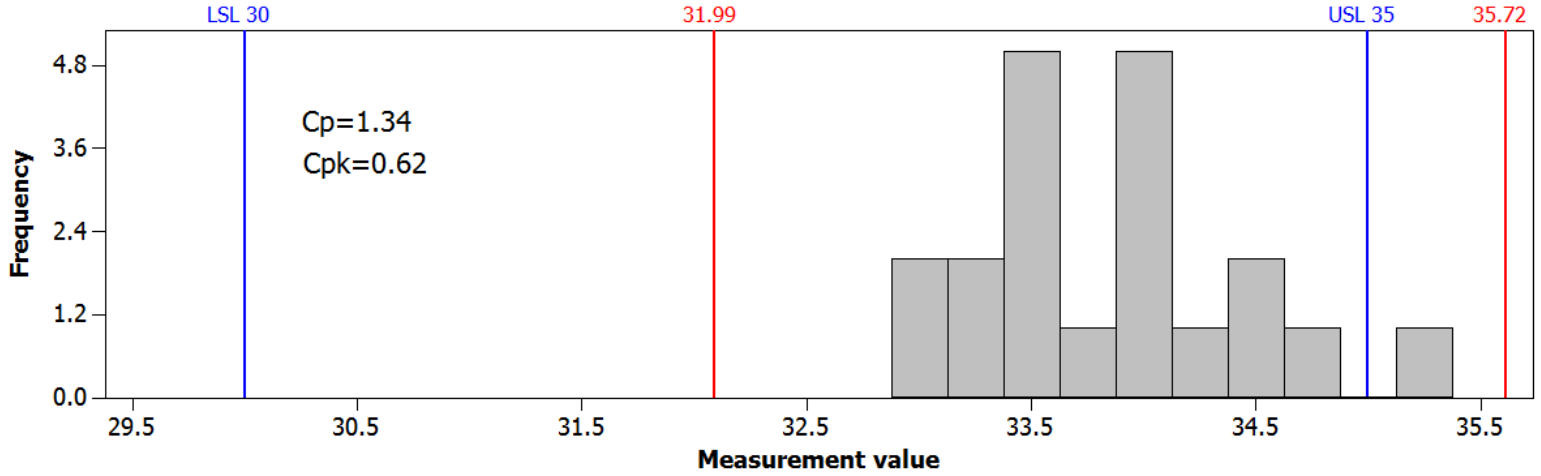
The context

- A production process has been operated over four days, and five values per day have been obtained
- Twenty data are available to be analysed
- Specifications are LSL=30 and USL=35
- Process target is 32.5, the midpoint of the specifications
- 20 values have an associated uncertainty in SD_{within} at ~20.5%

Example: 20 data values

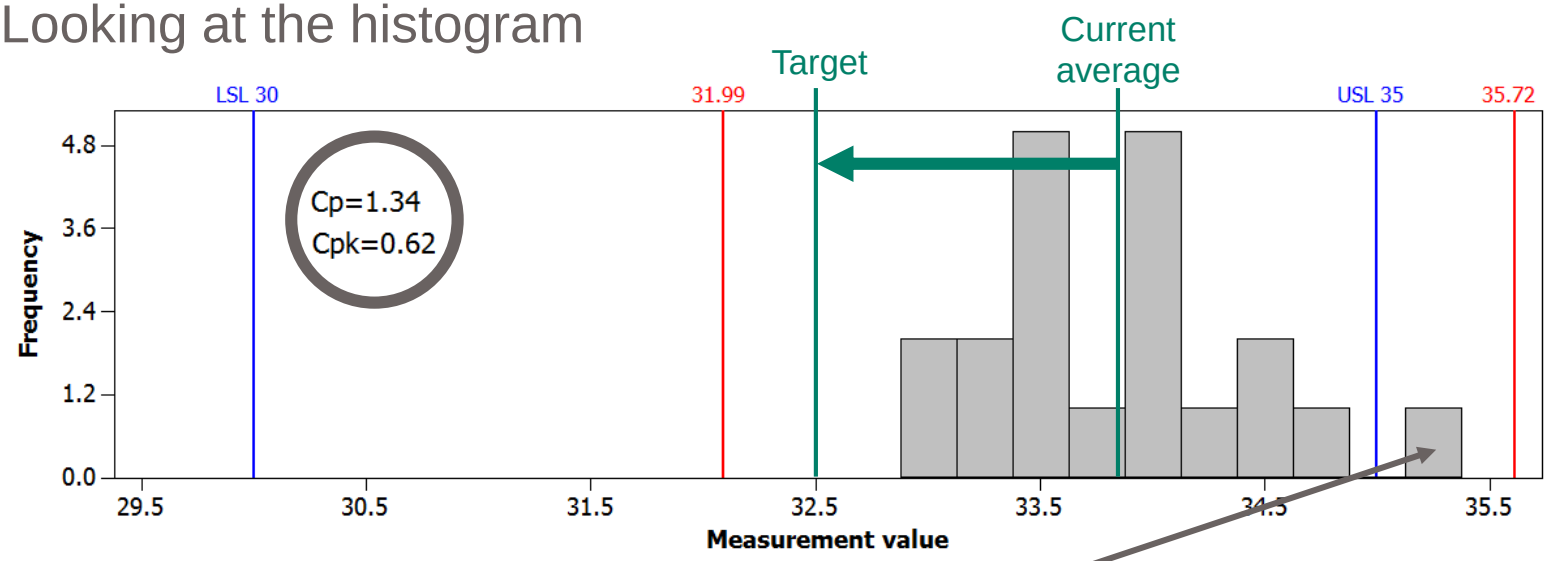
- Below an X chart and histogram with specs and process limits





Example: 20 data values

- Looking at the histogram



- One non-conforming unit has been found
- The process is off-target (relocate the average towards 32.5)
- If centred (on-target), we have some rationale to expect that the process would be characterised as capable (because C_p=1.34)

Questions to consider:

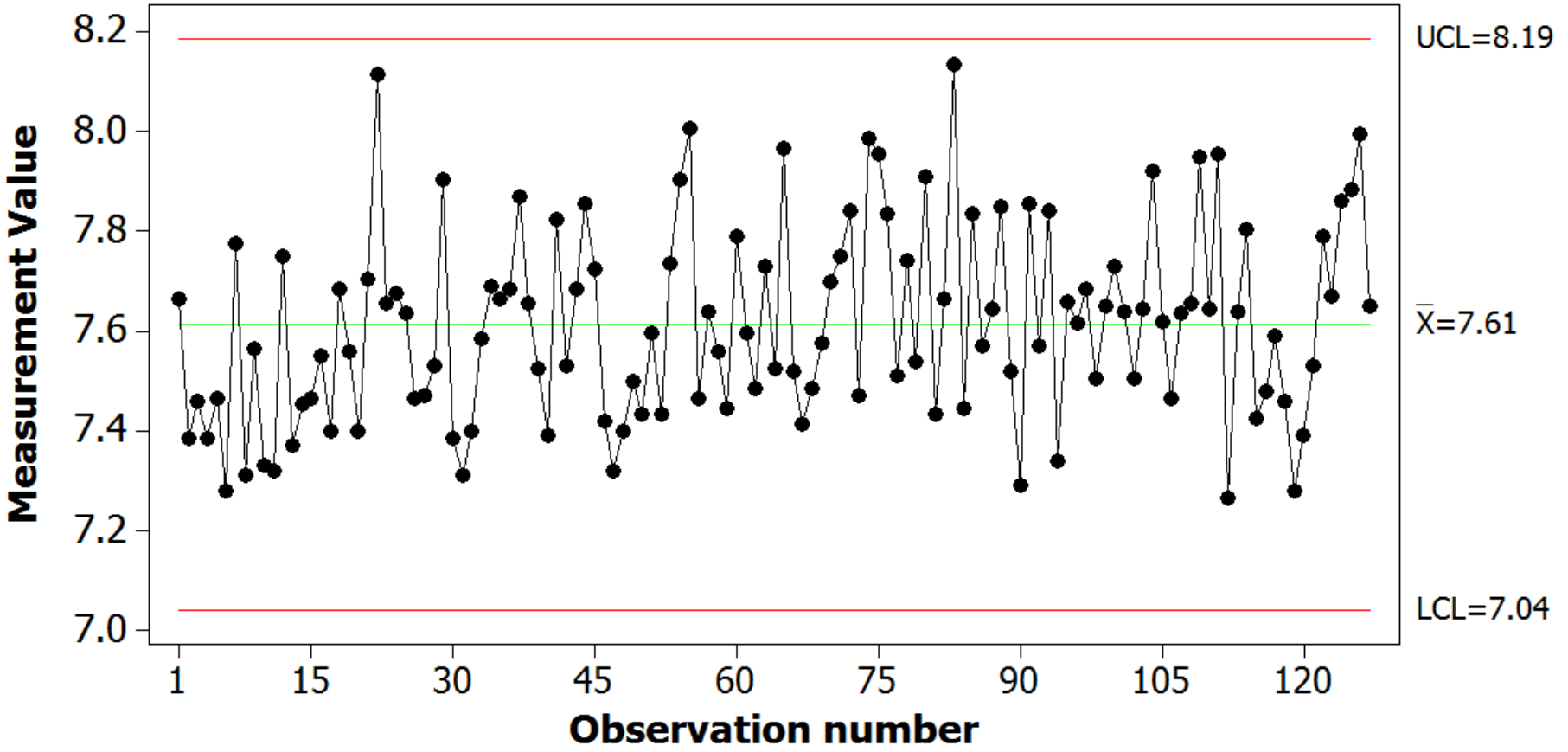
- *Do we not have enough data to justify action on the process?*
- *What action would we take?*

Example: 127 data values

The context

- A total of 127 values were obtained over one long production run
- The frequency at which data were collected was based on good process understanding ("rational sampling" in SPC)
- 127 values have an associated uncertainty in SD_{within} at ~8.1%
- To have much impact on reducing the CV of 8.1% many more data would be needed
  - *Example*: To reduce by half, to ~4%, some 500 or so data values would be needed!
- Are 127 values sufficient data?

Example: 127 data values



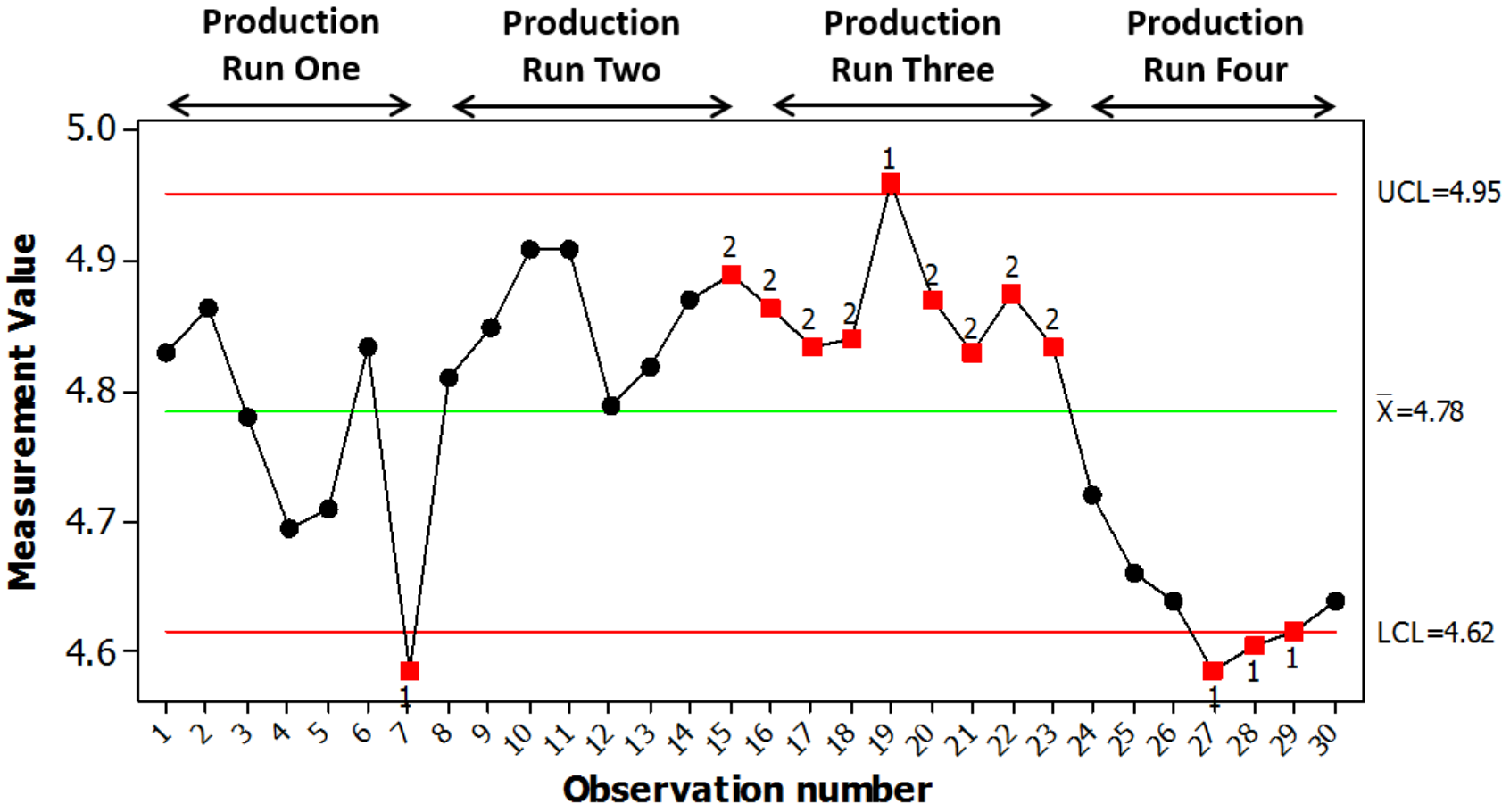
- This process demonstrates a reasonable degree of consistency
- Characterising the process as predictable ("in statistical control") seems reasonable

Example: 127 data values

- We have 127 data, so we are not short of data… reducing the uncertainty in SD_{within} in any meaningful manner would require a lot of extra data:
- Is 127 data enough in this case?
- If we'd like two production runs worth of data, then no
- In the data already obtained:
  - Do we have information (data) on the consistency of re-setting up the line (i.e. *between*-production run variability)?
  - Do we know of different shift influences, if any?
  - Do we know of inconsistencies in raw materials, if any?
  - etc.

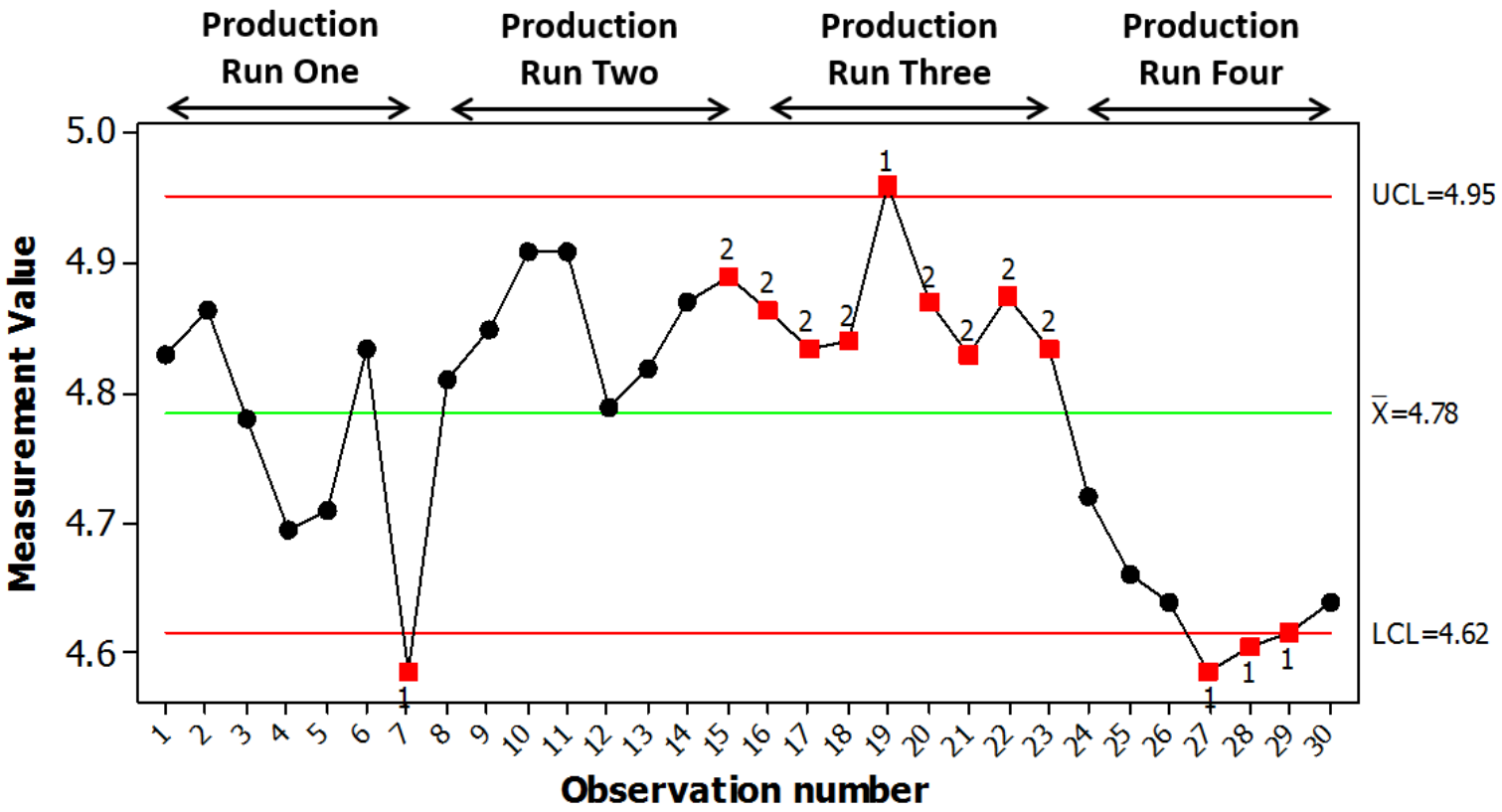
Example: 30 data

- Thirty data – shown below on an X chart – were recommended as a minimum to safeguard an analysis of process capability
- All data were collected before the analysis started



Example: 30 data

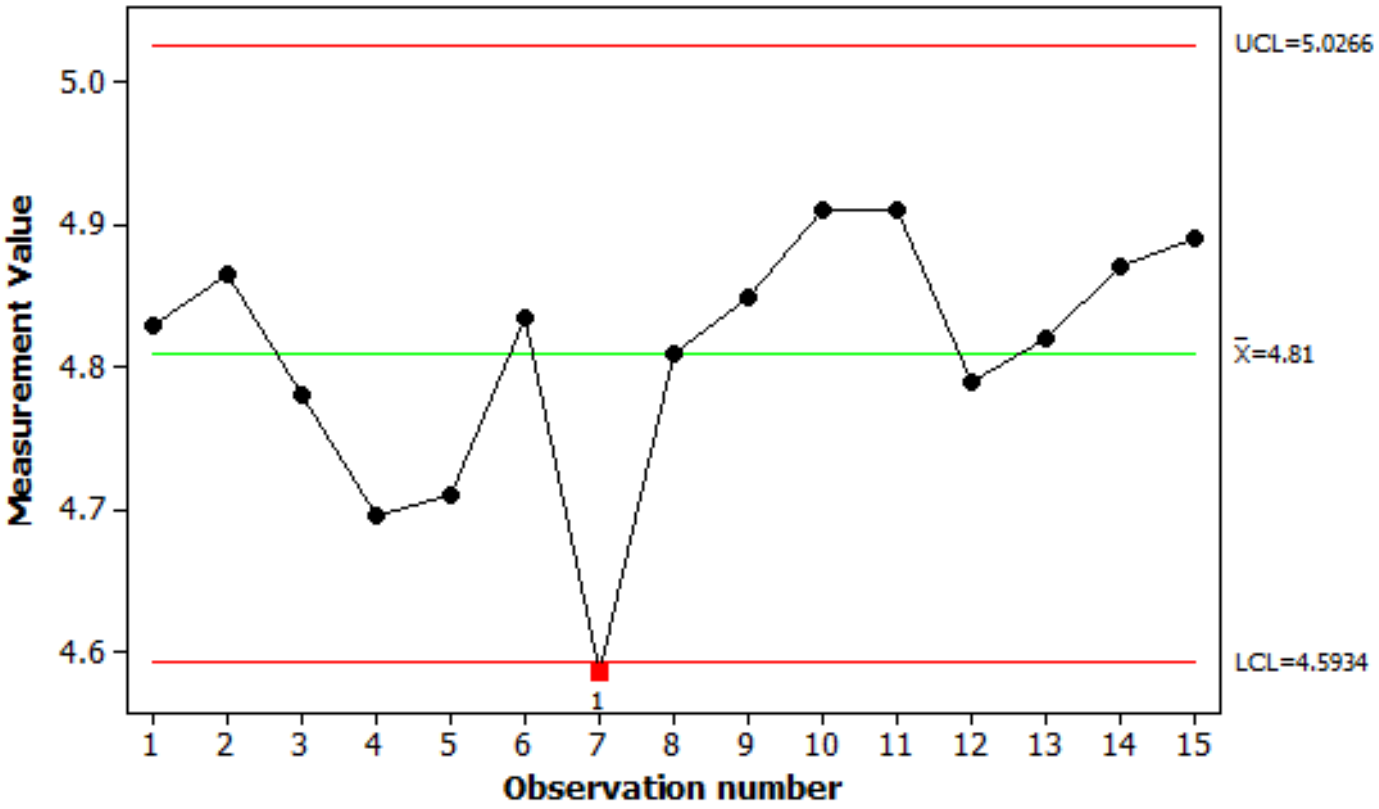
- For “process capability” to make sense, *wanting* predictability is implicit
  - If the below chart doesn’t lead to action, “process capability” is not the “way of working”
  - If action is an aim, should action have started earlier?



Nestlé Research

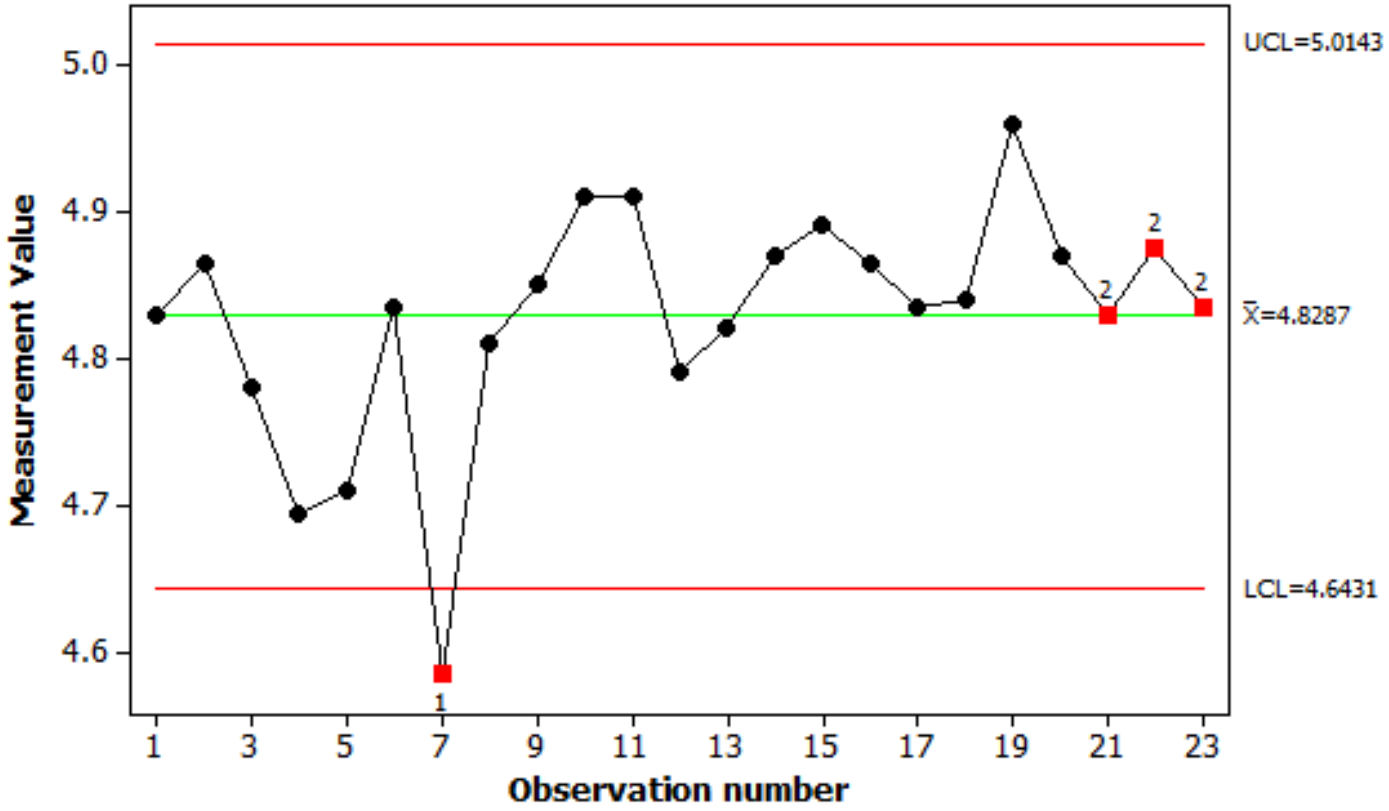
Example: 30 data

- Production runs one and two only
- One signal is present, giving a licence to investigate a *detected* process change



Example: 30 data

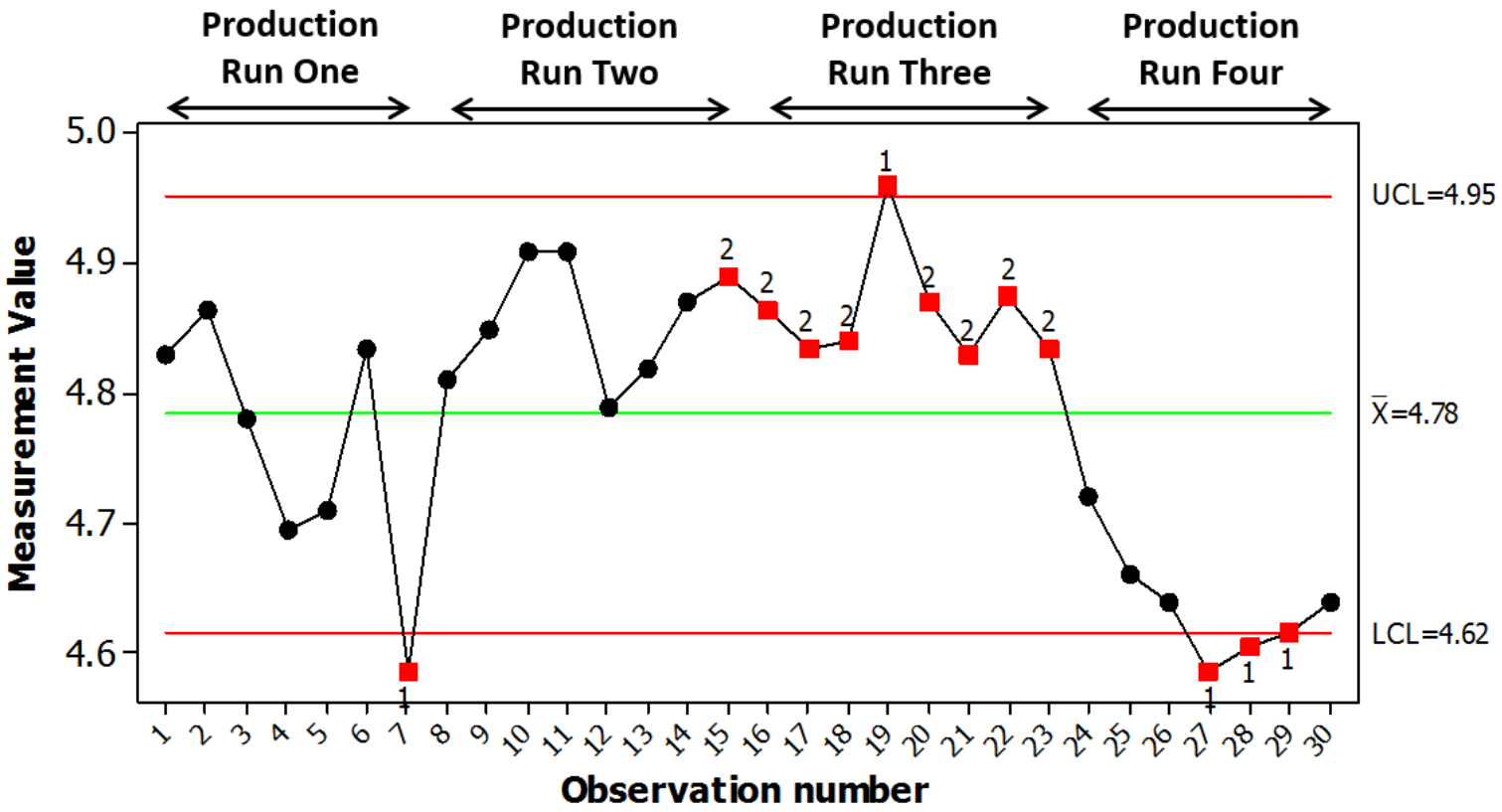
- Production runs one and two and three: The previous chart showed the process to be unpredictable, we now have more evidence of this
- The User has a choice:
  - Identify the assignable causes and better control them to improve the process
  - Do nothing and suffer the waste associated with this excessive variability



Nestlé Research

Example: 30 data

- You can only learn something if you look at the data, no matter how many, or how few, you have
- With 30 data we learn that the process is unpredictable, but that we could have learnt looking earlier…



Conclusions

How many data for capability?

Conclusions: *How many data for capability?*

Is this answer good enough?

Conclusions: *How many data for capability?*

Is this a better answer?

This answer will only make sense if the rationale behind it is understood (e.g. that timely action is important and that capability is about action)

In summary

- Statistical theory helps to answer the "*How many data?*" question
- But, statistical theory alone is not enough
- Judgement, based on context and understanding, is also needed
- A "standard answer" might not be the best answer because each case is different
- "*How many data?*" will be best answered only if the problem at hand is understood

References

- Bothe, D. R. (1997). Measuring Process Capability: Techniques and Calculations for Quality and Manufacturing Engineers. McGraw-Hill Inc., US.
- Deming, W. E. (1975). On Probability As A Basis For Action. *The American Statistician*, 29(4), 146-152.
- Deming, W.E. (1982). Quality, Productivity and Competitive Position, MIT.
- Shewhart, W. A. (1931). Economic Control of Quality of Manufactured Product. (pp. 302). New York: D. van Nostrand Company, Inc.
- Wheeler, D. J. (2004). Advanced Topics in Statistical Process Control: The Power of Shewhart's Charts. (pp. 80-83, 180-186 & 446). Second Edition, SPC Press, Knoxville, Tennessee.
- Wheeler, D.J. (2000). Normality and the Process Behavior Chart. SPC Press, Knoxville, Tennessee.