THE SCIENCE OF BEER

SPC Theory and Tools

Statistical Process Control

- Goal = Reduce variability
- On-going study of a process
- Utilizes statistical techniques to signal unusual events

SPC Tools – The Magnificent Seven

- **Histograms**
- Check sheet
- **Pareto chart**
- Cause and effect diagram
- Defect concentration diagram
- Scatter diagram
- **Control Chart**

Population vs. Sample

Population

- **Parameter**
  - Population characteristic
  - Calculated from all possible measurements
- Parameters don't vary

Sample

- **Statistic**
  - Sample characteristic
  - Calculated from only members of the sample.
- Samples vary

Summarizing Data

- Center -- Mean, Median, Mode

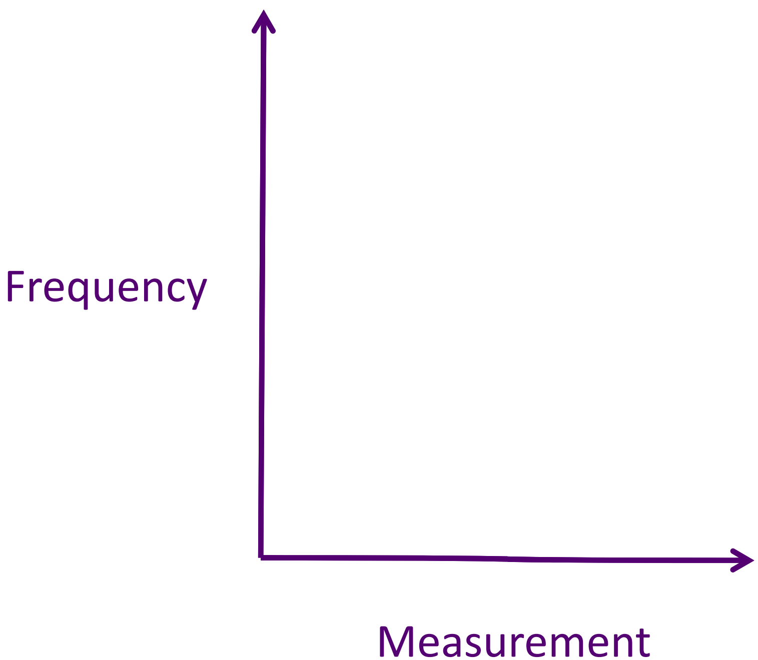
$$\mu_{Pop} = \frac{\sum x_i}{N} \qquad \bar{x}_{sample} = \frac{\sum x_i}{n}$$

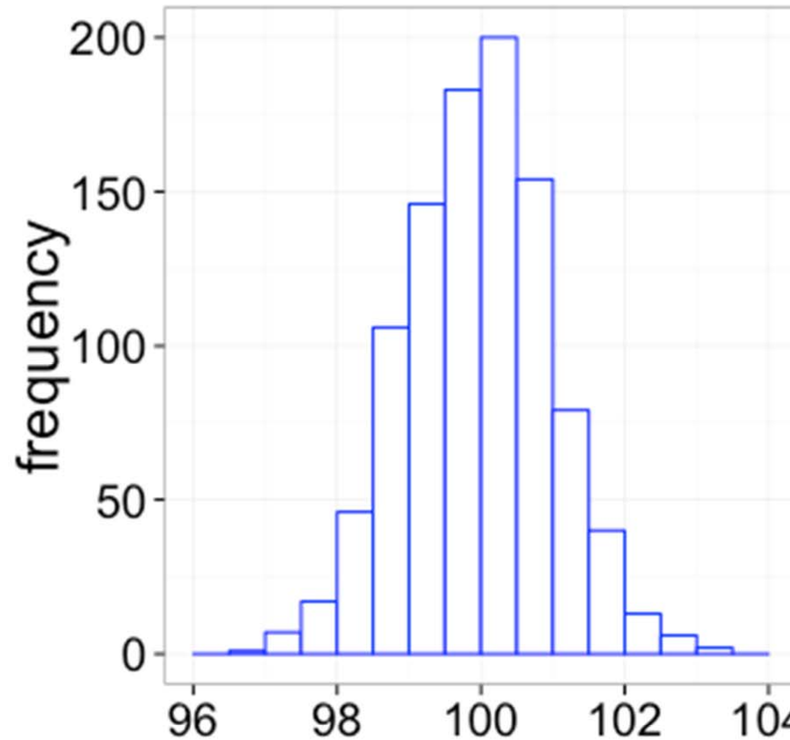
- Spread – Range, Standard Deviation

$$\sigma_{Pop} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} \qquad s_{sample} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

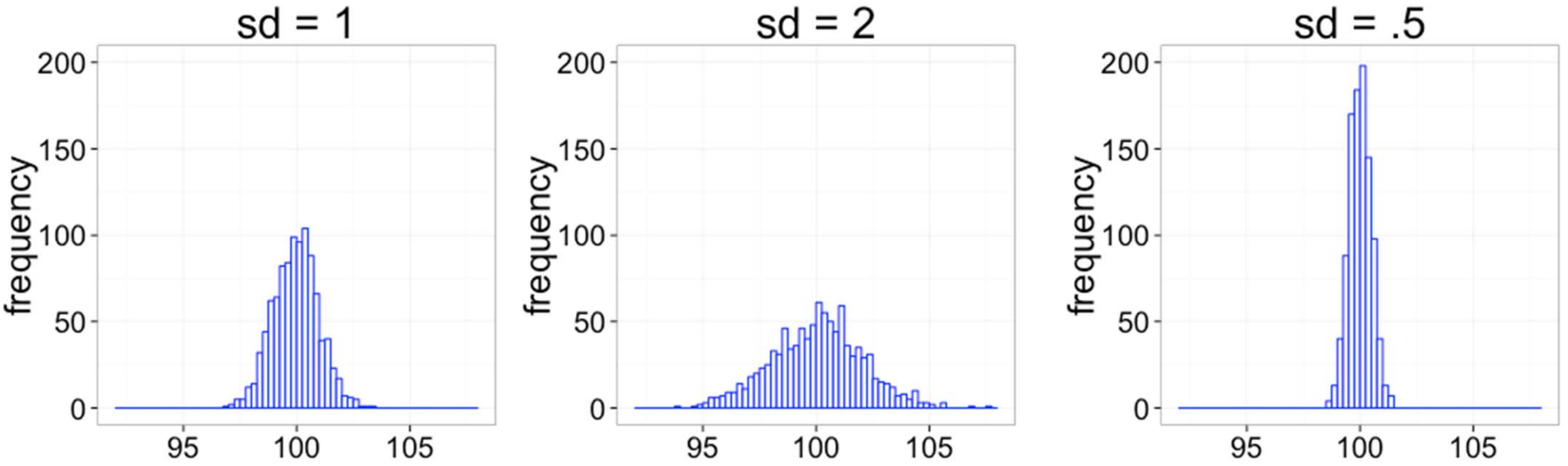
Visualizing Data

- Histograms – **Distributions** of data

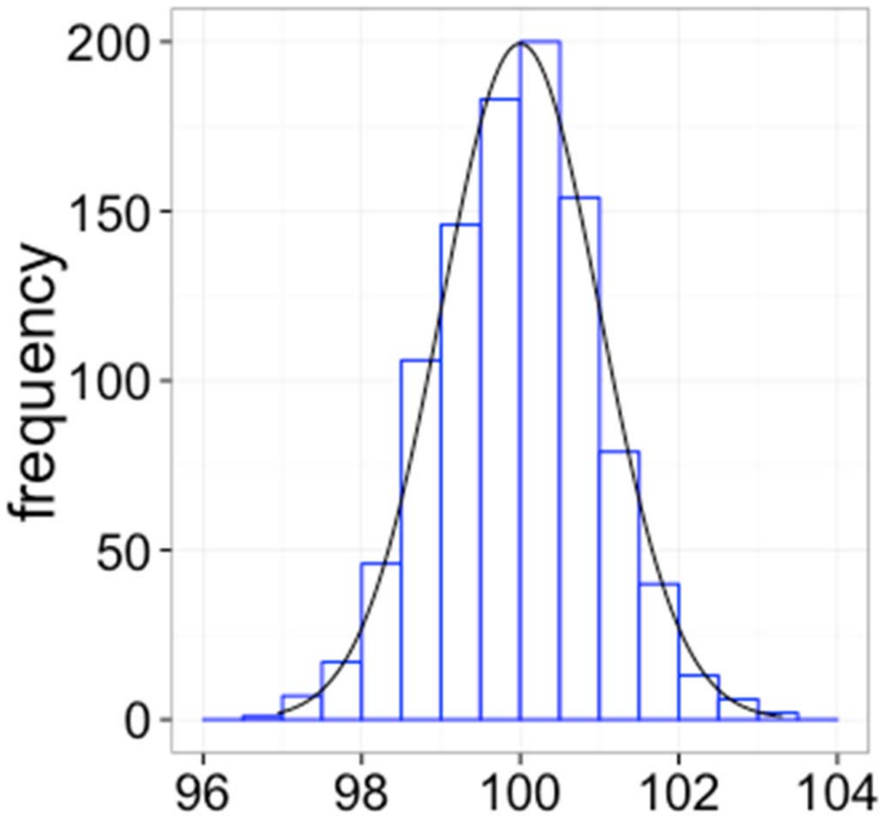




Variability

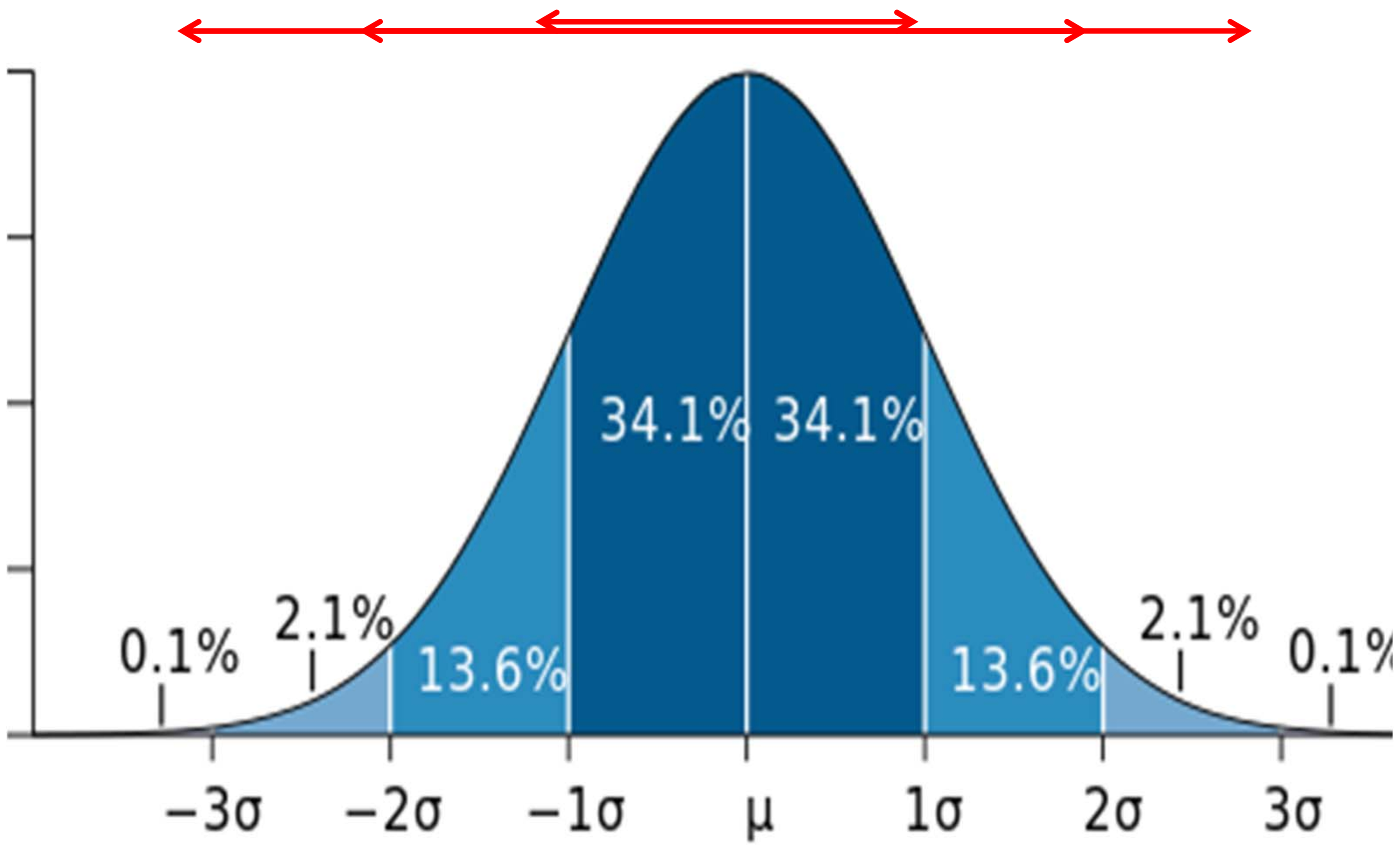


Normal Distribution as Model for Data

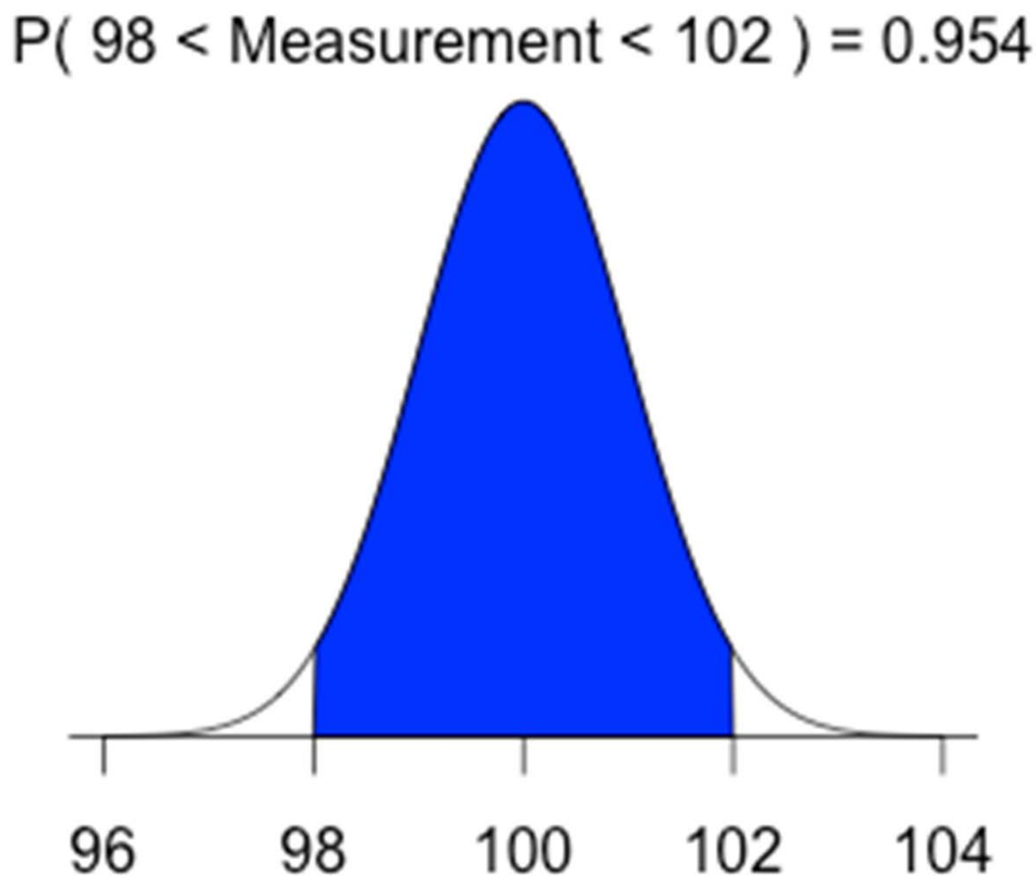


Normal Distribution Properties

± 3 SD ≈ 99.7%



Normal Probability Distribution



Central Limit Theorem

Sampling Demo

Take home message --

1. The mean of the sampling distribution is the mean of the population distribution

   So…
   We can use the sample mean for an estimate of the population mean

Central Limit Theorem

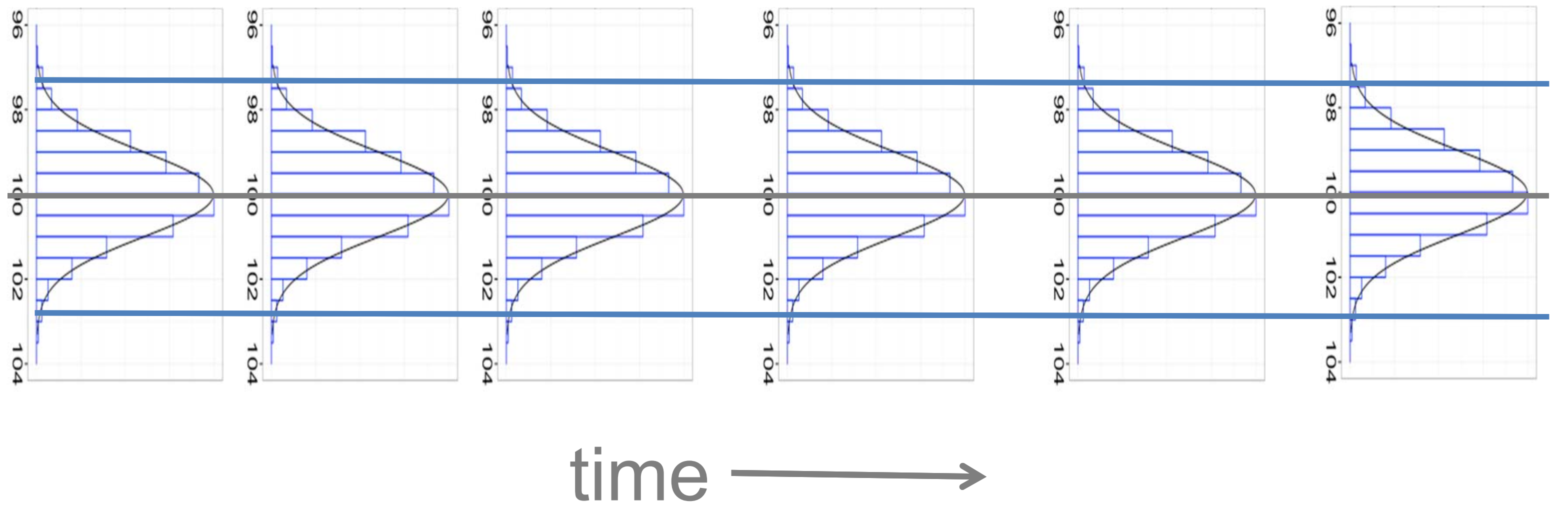
2. The standard deviation of the sampling distribution is the **standard error**

$$= \frac{\sigma}{\sqrt{n}}$$

3. The sampling distribution is **normally distributed** *regardless* of the shape of the population distribution.

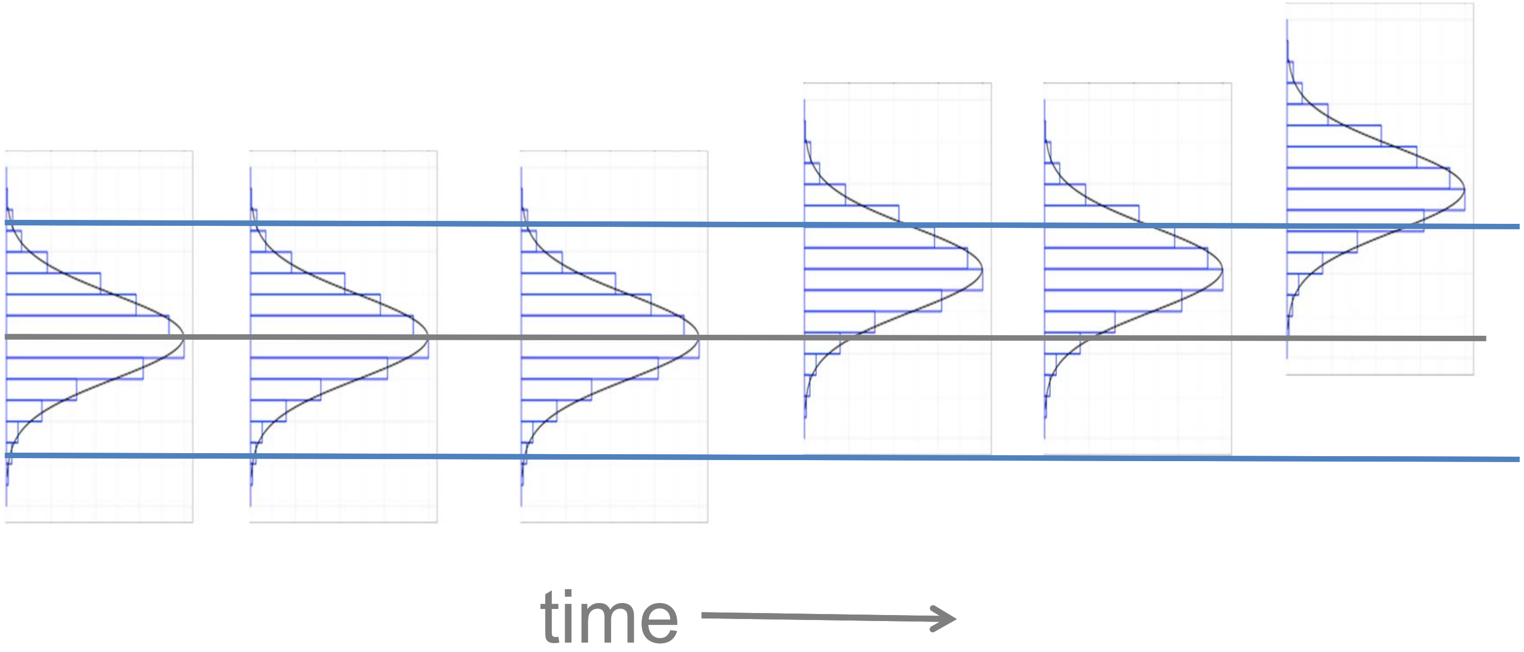
So what has this got to do with SPC?

- A variable that continues to be described by the ***same*** distribution over time = ***In control***



Disturbances Change the Distribution

- Shift in mean

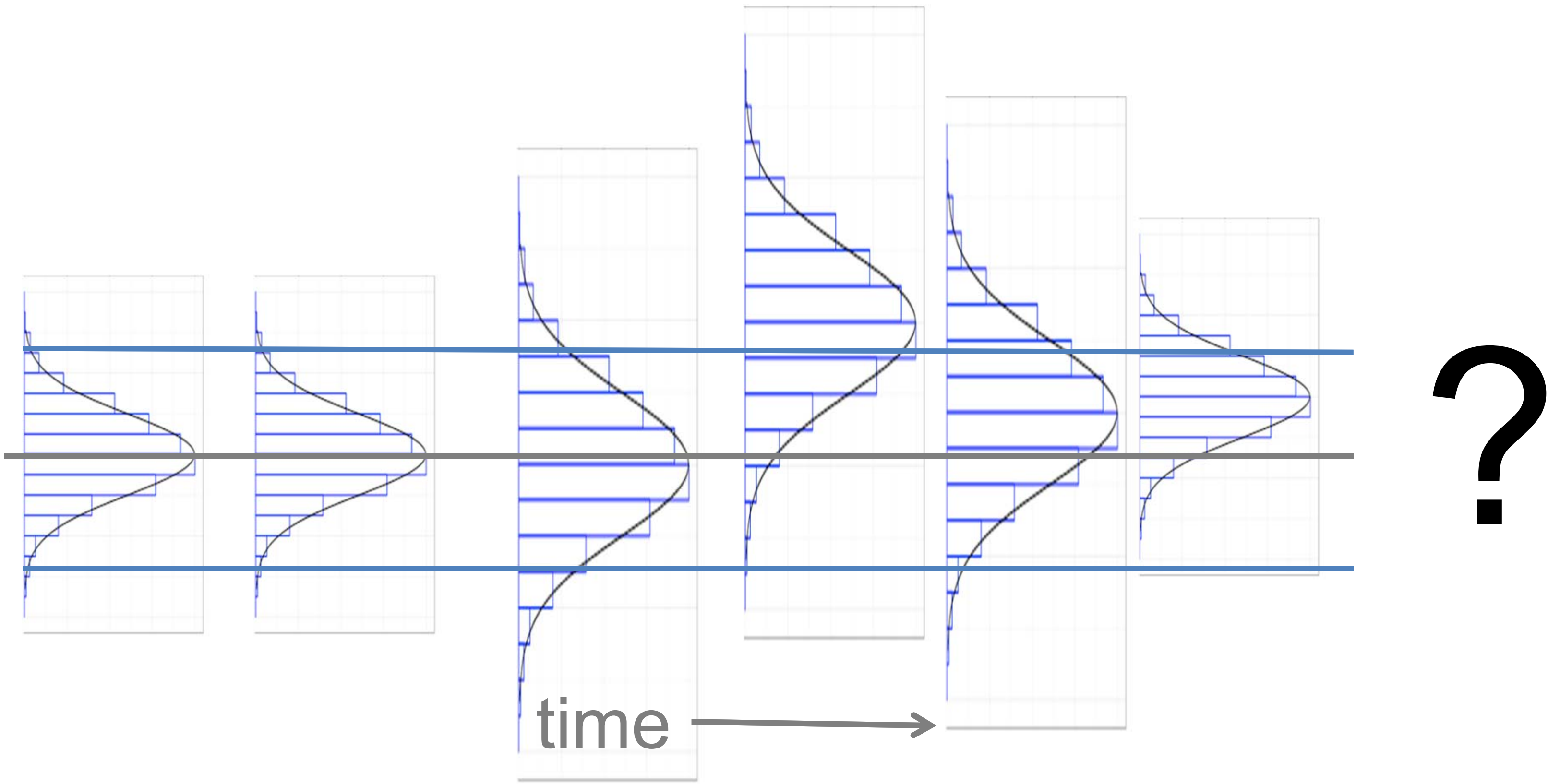


Disturbances

- Add Variability
  - **Common** Cause vs. **Assignable** Cause
  - **Assignable** Cause variation adds to **Common** Cause variation.

A process with **assignable variation** is said to be operating **“out of control**”.

Added Variability



Control Charts in a Nut Shell

If we have a **stable** process,

- **Sample** (25-30 or more)
  - Size? Frequency?

- **Model** our process with a normal curve
  - Control Limits (+/- 3 SD), x-bar

- **Compare** new measurements to the curve

Sampling – "Rational Subgroups"

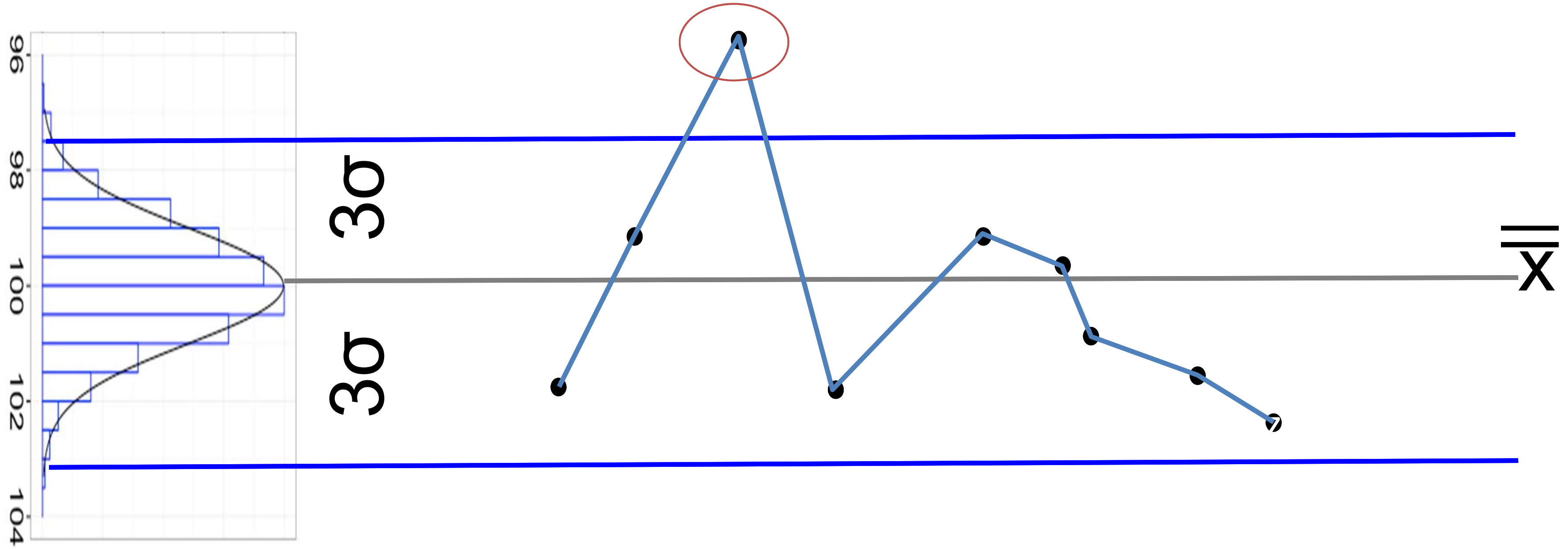
- Each sample contains only common cause variation from a stable process
  - Assignable cause in sample ~ CLs too wide
- Observations within a sample are independent
  - Autocorrelation~ CLs too narrow
- Spacing of samples (frequency) minimize within sample variation and maximize between sample variation

Phases of Control Charting

- Phase I – Setting up control limits
  - Start with process you believe is in control
  - Sample, Model (calculate CLs), Graph
  - Engineering Statistics Handbook

- Phase II – Monitoring process

Phase I – Control Charting



Variable Charts

- **Continuous scale**
- **X-bar** chart monitors *mean* of a process
- **R**-chart/ **S**-chart monitors *variation*
  - **R** measures range of values (2 < n < 10)
  - **S** measures standard deviation (n > 10)

Xbar and R/S charts

- Used together

- If the sample variability is ***not*** in control, then the entire process is judged to be not in statistical control, regardless of what the **xbar** chart indicates.

Attribute Charts

Attribute charts are charts for non-continuous data.

- **p** monitors proportion nonconforming, constant n
- **np** monitors proportion, varying n
- Based on binomial distribution

- **c** monitors counts, constant n
- **u** monitors counts, n varies
- Based on Poison Distribution

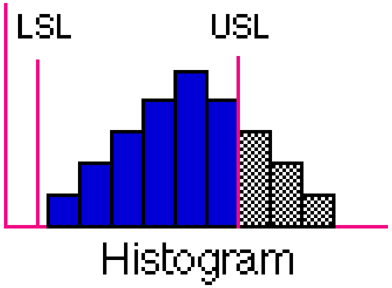
Process Capability Analysis

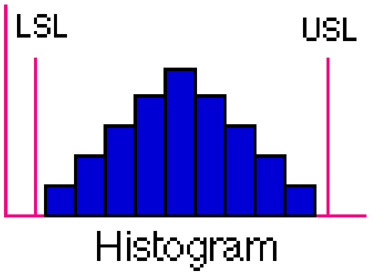
- Cp = Process Capability

$$C_p = \frac{USL - LSL}{6\sigma}$$

Tolerance limits
____
Process Range

- Values must be > 1 to have process capacity.
- $C_p < 1$ $\quad$ $C_p = 1$





C_{pk} – index of how centered process is

$$C_pU = \frac{USL - \bar{x}}{3\hat{\sigma}}$$

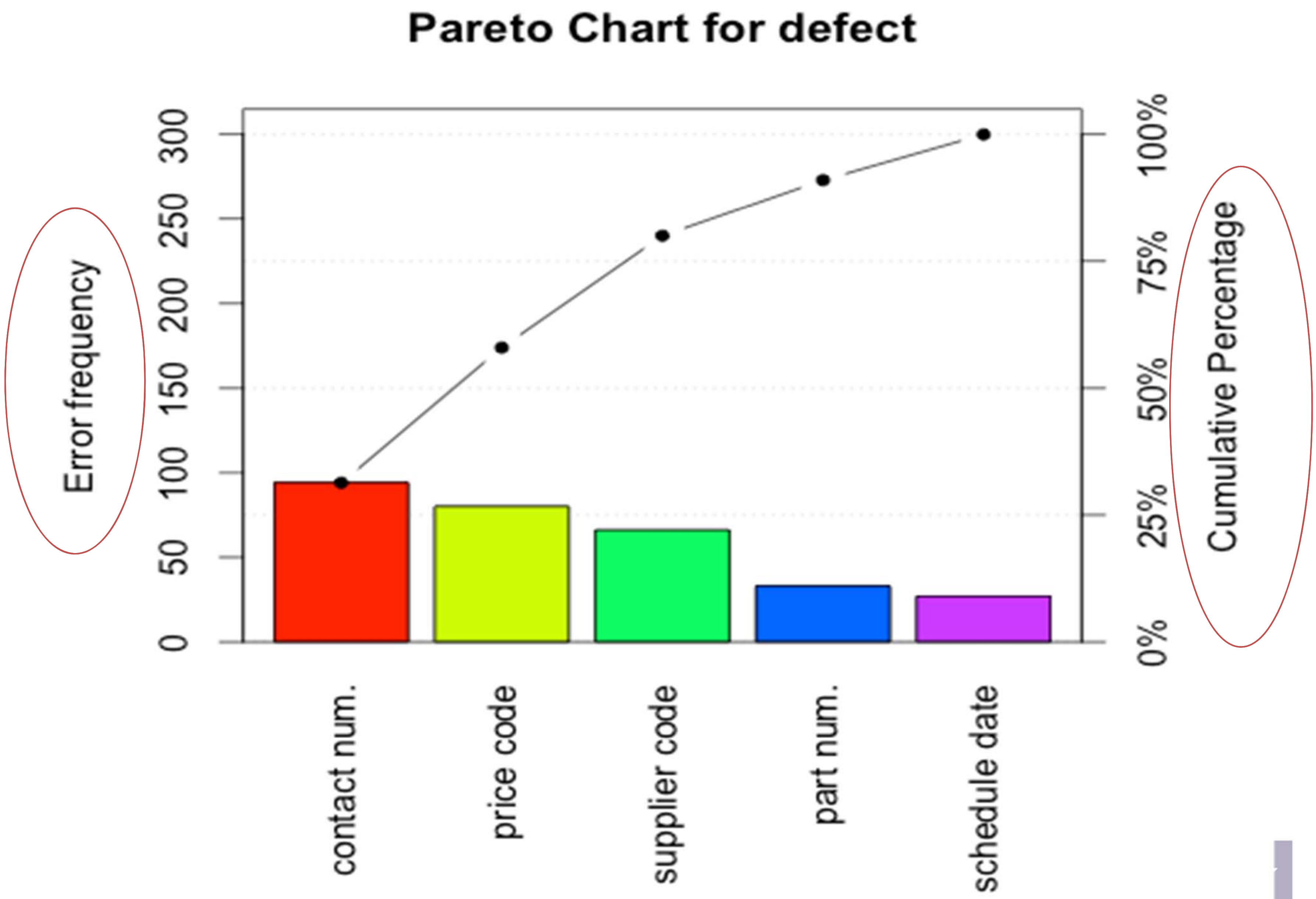
$$C_pL = \frac{LSL - \bar{x}}{3\hat{\sigma}}$$

$$C_{pk} = \min(C_pU, C_pL)$$

If $C_{pk} < C_p$, process is not centered

Pareto Chart Steps

1. Develop a list of causes to be compared.
2. Develop a standard measure for comparison
   - frequency
   - cost
   - total time it took
3. Collect data over a set timeframe.
4. Create an ordered histogram
   - Count on left axis
   - Cumulative % on right axis


Pareto Chart for defect
Error frequency
0
50
100
150
200
250
300
0%
25%
50%
75%
100%
Cumulative Percentage
contact num.
price code
supplier code
part num.
schedule date