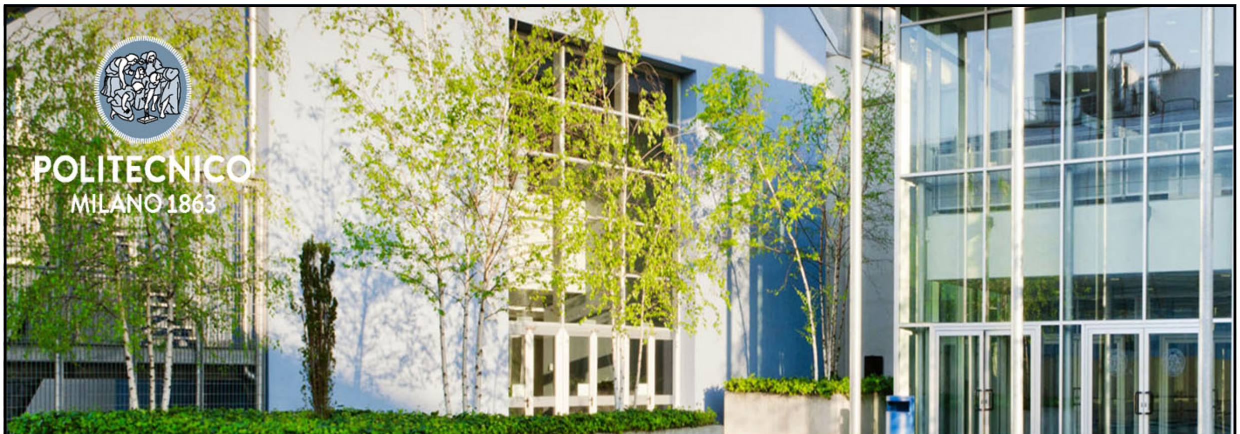


Quality Data Analysis

Statistical Quality Monitoring - Introduction

Bianca Maria Colosimo – biancamaria.colosimo@polimi.it

Reference:
Introduction to Statistical Quality Control – D.C. Montgomery

1

Process monitoring – Statistical Process Monitoring

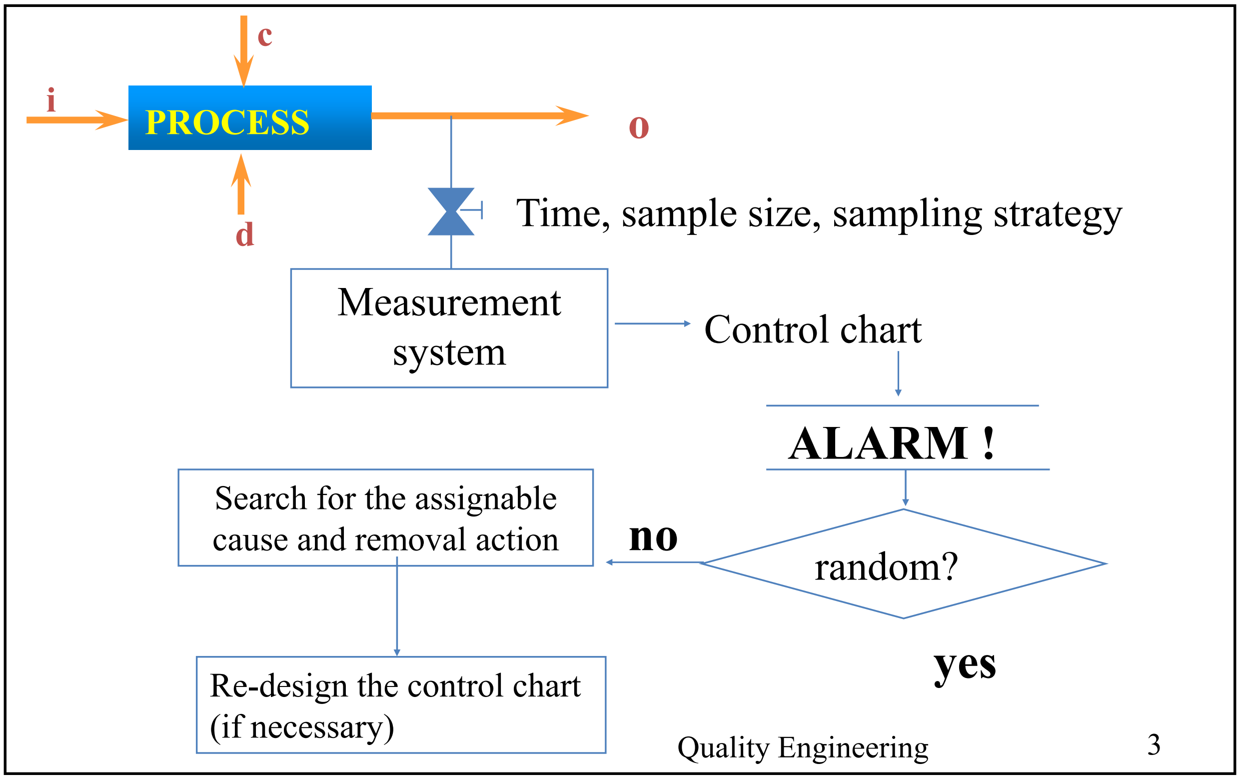
Model of the process output

$$\boxed{Y_t \overset{iid}{\sim} (\mu, \sigma_\varepsilon^2)}$$

Basic model
(in-control process)

Is the process stable with time?
- Is the mean μ stable with time?
- Is the variance σ_ε^2 (standard deviation σ_ε) stable with time?
- (is the distribution stable with time)

2



3

Control chart

statistical procedure –a systematic analysis of the process output is used to identify out of controls as effect of special causes

1: success depends on the detection of the out-of-control states but especially on the identification of the assignable cause (even this is an *improvement step*)

2: Schonberger (1986): People collecting data are the ones that can analyze those data and think to possible solutions. The success depends on the ability to collect the right data at the right time. Operators tend to collect data after a problem has been detected. Usually, the relevant part of data are is related to the time before the problem detection.

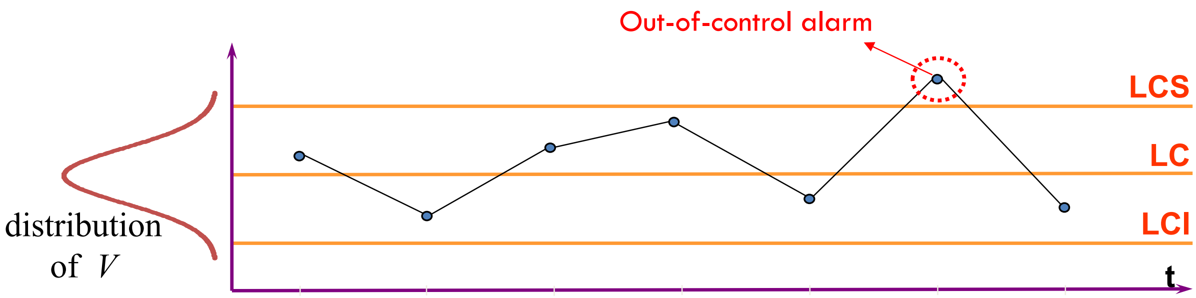
→ Data gathering should be systematic

4

Traditional control charts (Shewhart)

1. Process output O: take a sample (*subgroup*) of data (n) at regular intervals
2. From all the samples: sample statistic V (eg. Sample mean, sample standard deviation,, *sample* range)
3. If O **is caused by a** random IID process, any sample statistic computed from these data will be random IID

(Pay attention: distribution of the individual measurements is usually different from the distribution of the sample statistic)



Basic scheme Shewhart's control chart

$$UCL = \mu_V + K\sigma_V$$
$$CL = \mu_V$$
$$LCL = \mu_V - K\sigma_V$$

where
- $\mu_V = E(V)$
- $\sigma_V = \sqrt{V(V)}$
- $K =$ positive coeff (usually $K = 3$)

Note:

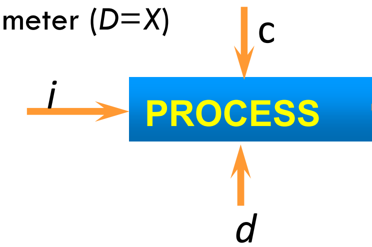
- Symmetry
- (quantitative) out-of-control rule: alarm if points outside the limits
- (qualitative) out-of-control rule: look for strange data pattern for V (trend, cicles, etc.)

Example: Shewhart control chart

Assume that the process produces shafts:

we are interested to the outer diameter (D=X)

Stable process:



$$o = X \sim N(\mu_X, \sigma_X^2)$$

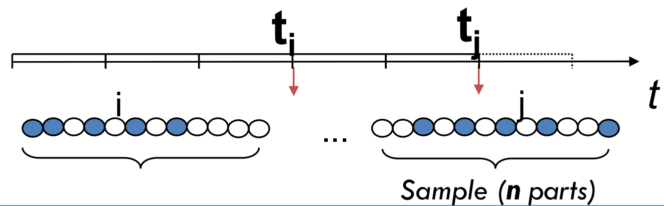
$$\text{con } \mu_X = 74.000 \text{ mm}$$

$$\sigma_X = 0.010 \text{ mm}$$

Monitoring the process mean:

1. Take at regular interval

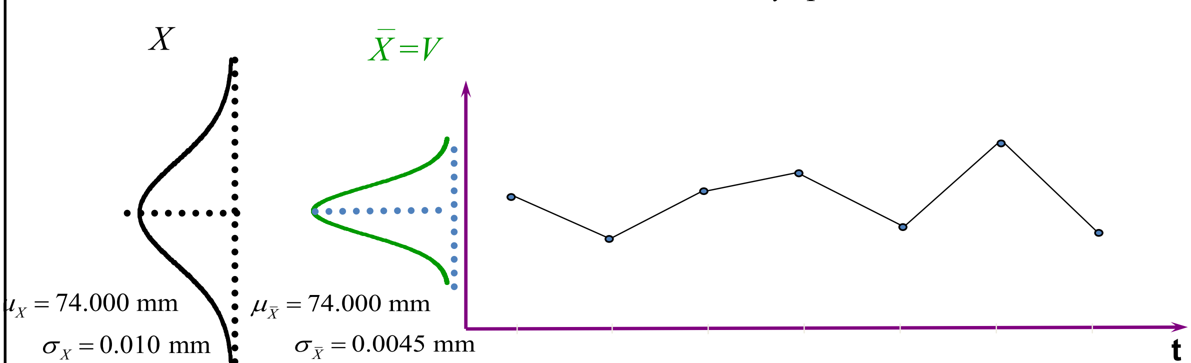
Samples n=5



Example: Shewhart control chart

2. Compute

$$\Rightarrow \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu_X, \frac{\sigma_X^2}{n})$$



Example: Shewhart control chart

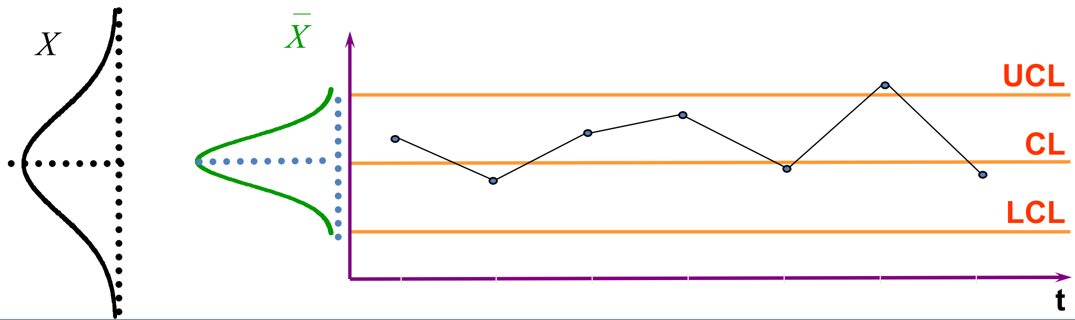
3. Control limits

$$K = 3$$

$$UCL = \mu_{\bar{X}} + K\sigma_{\bar{X}} = 74.000 + 3(0.0045) = 74.0135$$

$$CL = \mu_{\bar{X}} = \mu_{\bar{X}} = 74.000$$

$$LCL = \mu_{\bar{X}} - K\sigma_{\bar{X}} = 74.000 - 3(0.0045) = 73.9865$$



Quality Data Analysis- BM Colosimo

POLITECNICO MILANO 1863

9

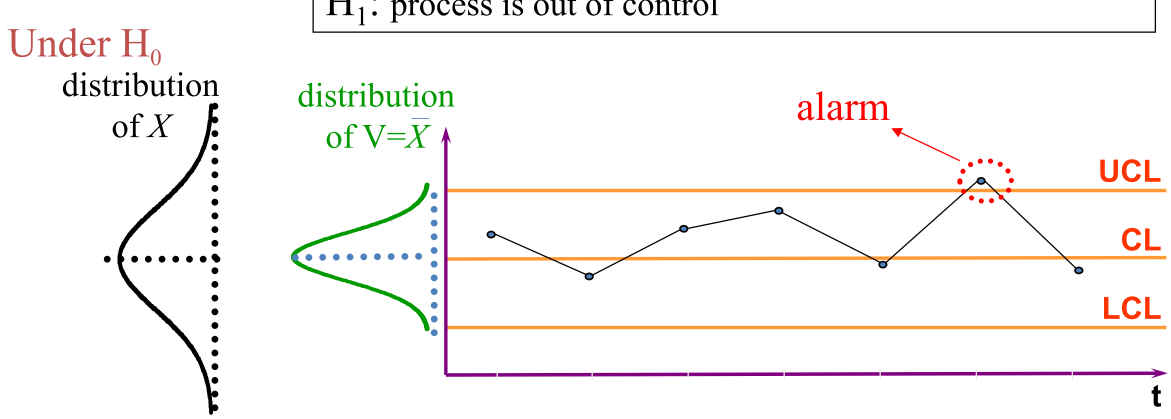
9

Connection with hypothesis testing

test

H_0: process is in control (IID)

H_1: process is out of control



Quality Data Analysis- BM Colosimo

POLITECNICO MILANO 1863

10

10

Sources of variability

Process variability = (1) Common causes + (2) Special causes

(1): due to many different sources (material, machines, tools, operators, measurement systems, …). NB: anche in Business process

(2): due to some specific issues that can be easily identified (batch of material of poor quality, operator error, wrong setup, wear, etc.). Sometimes positive effect (serendipity)

Common causes (natural variability): management is responsible but contributions from all the people in the organization are needed to improve… "
Management may be the only ones who can really do something about the opportunity for improvement, but they cannot act if they really do not know about it" (Scherkenbach, 1986)

Special causes: identification is often linked to the way in which data are gathered and analized - they can be removed by the operator (setup) or from a decision at a higher level (input material)

Connecton with hypothesis testing

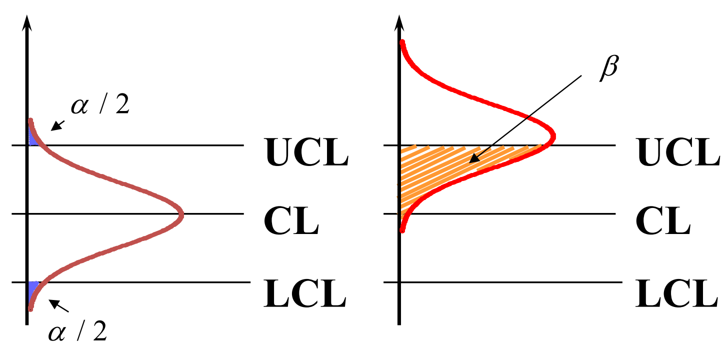
Out of control: it can be due to an assignable cause … but it can be also due to the chance

In control: natural variability… but it can also be that an assignable cause is present

(Dynamic) test

H_0: process is in control

H_1: process is out of control



In the previous example:

in-control state: H_0 $\quad \mu_X = 74.000 \text{ mm}, \ \sigma_X = 0.010 \text{ mm}$

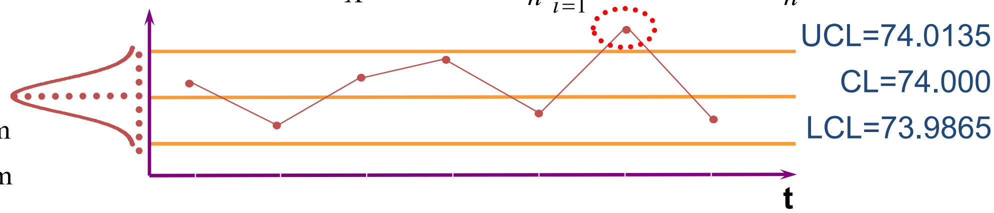
out-of-control state H_1 $\quad \mu'_X = 74.020 \text{ mm}, \ \sigma_X = 0.010 \text{ mm}$

$$X_i \overset{iid}{\sim} N(\mu_X, \sigma_X^2) \Rightarrow \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu_X, \frac{\sigma_X^2}{n})$$

H_0 true

$\mu_{\bar{X}} = 74.000$ mm

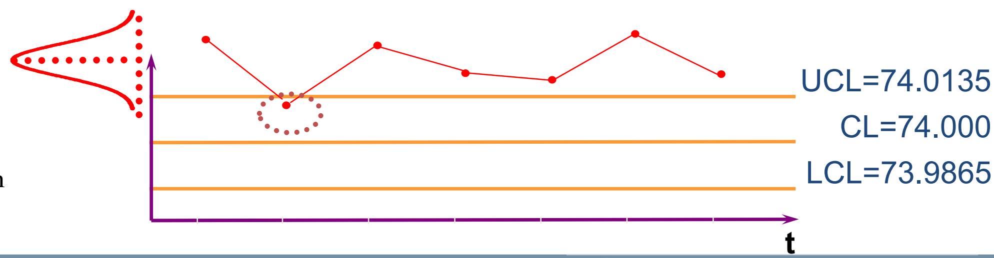
$\sigma_{\bar{X}} = 0.0045$ mm



H_1 true

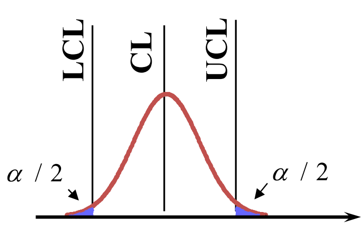
$\mu_{\bar{X}} = 74.020$ mm

$\sigma_{\bar{X}} = 0.0045$ mm



Control chart: K

Assume $\boxed{V \sim NID(\mu_V, \sigma_V^2)}$



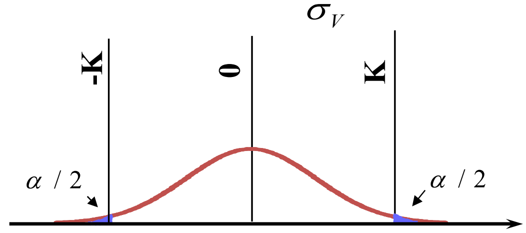
standardize

$$\frac{V - \mu_V}{\sigma_V} = Z \sim N(0,1)$$

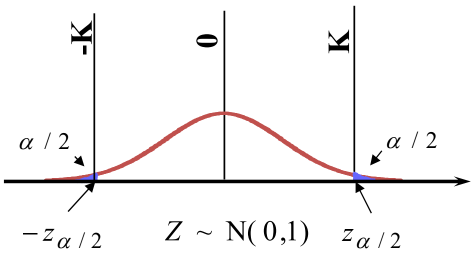
$$\frac{LCL - \mu_V}{\sigma_V} = \frac{\mu_V - K\sigma_V - \mu_V}{\sigma_V} = -K$$

$$\frac{UCL - \mu_V}{\sigma_V} = K$$

$Z \sim N(0,1)$



K



$$K = z_{\alpha/2}$$

$$V \sim \mathrm{N}(\mu_V, \sigma_V^2) \quad \alpha/2 = \Pr(V \le LCI) = \Pr(V \ge LCS)$$

$$\alpha/2 = \Pr\left(\frac{V - \mu_V}{\sigma_V} \le \frac{LCI - \mu_V}{\sigma_V}\right) = \Pr(Z \le -\mathrm{K}) = \Pr(Z \ge \mathrm{K}) \Rightarrow \mathrm{K} = z_{\alpha/2}$$

Cumulative Distribution Function

```
Normal with mean = 0 and standard
deviation = 1.00000

         x       P( X <= x )

   -3.0000          0.00135
```

$$\alpha = 0.0027 = 2.7\ {}^{0}/_{00}$$

15

Link K-α

Assumptions: V is approximately distributed as a normal variable **(from IID to NID)**

- If variability due to a lot of causes
- Central limit Theorem

If $V \sim \cancel{\mathrm{N}}\mathrm{ID}(\mu_V, \sigma_V^2)$?

1. ***Tchebyscheff inequality***: X follows any distribution

$$\Pr\left(|X - E(X)| > K\sqrt{V(X)}\right) \le 1/K^2$$

$\boxed{\text{se } \mathrm{K} = 3}$ $\alpha \le 1/9 = 11\%$

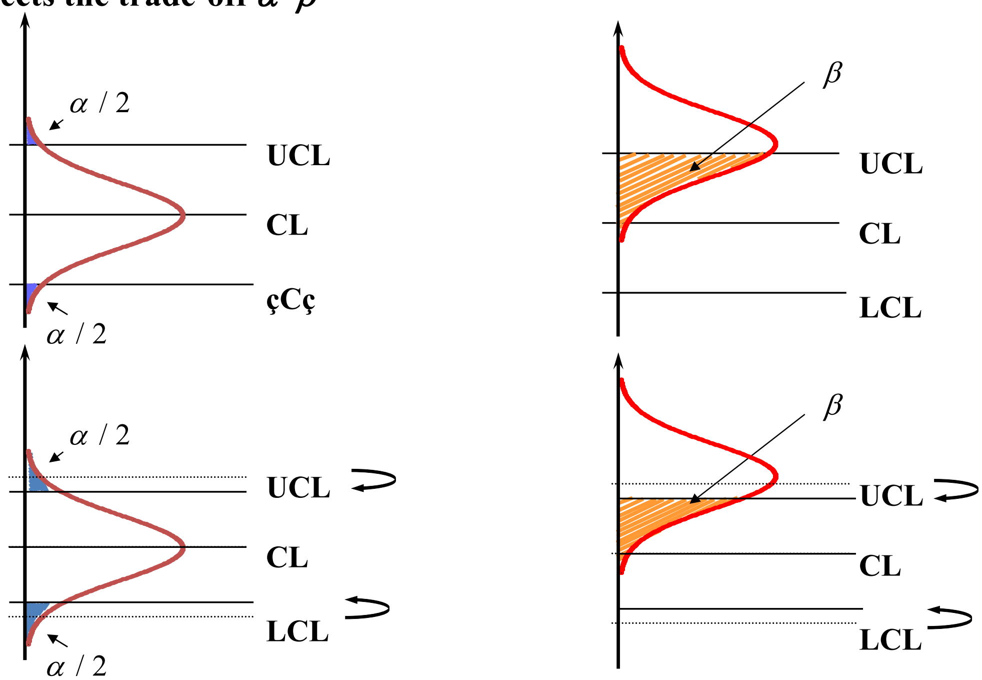
16

2. ***Camp-Meidell*** *inequality* (X follows a unimodal distribution in μ_0):

$$\Pr\left(|X - E(X)| > K\sqrt{V(X)}\right) \le \frac{4}{9}\frac{1+f^2}{(K-f)^2} \qquad K > f \qquad f = \frac{|E(X)-\mu_0|}{\sqrt{V(X)}}$$

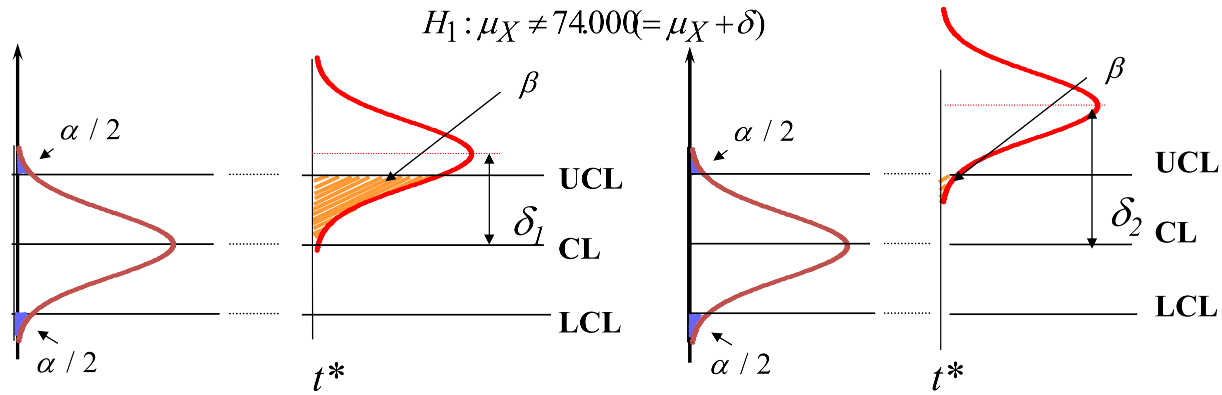
$$if \quad E(X)=\mu_0 \quad f=0 \qquad \text{and if } K=3 \qquad \boxed{\alpha \le 4/81 = 4.9\%}$$

K affects the trade-off α–β



Second-type error β:

$$H_0 : \mu_X = 74.000$$
$$H_1 : \mu_X \neq 74.000 (= \mu_X + \delta)$$



$$\boxed{\beta = \beta(\delta)}$$

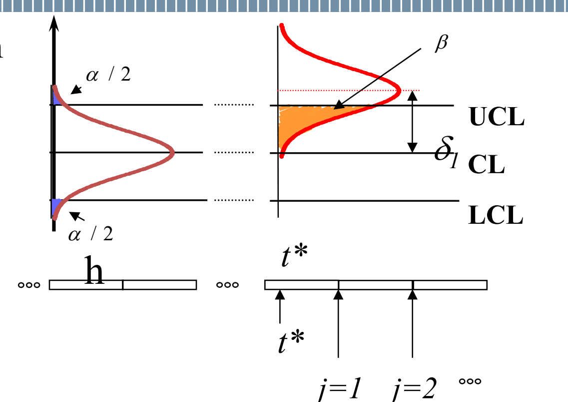
$$\boxed{\beta(\delta = 0) = 1 - \alpha}$$

Average Run Length (ARL)

ARL= expected value of points to be plotted on the control chart before an alarm

Assume shift (H_0->H_1)

$$\boxed{ARL_1 = \sum_{j=1}^{\infty} j \, p(x = j)}$$



but:

$$p(x = 1) = 1 - \beta$$
$$p(x = 2) = \beta(1 - \beta)$$
$$\ldots$$
$$p(x = j) = \beta^{j-1}(1 - \beta)$$

$$\sum_{k=0}^{\infty} \beta^k = \frac{1}{1-\beta} \qquad \sum_{k=0}^{\infty} k\beta^k = \frac{\beta}{(1-\beta)^2}$$

J-1=k

$$ARL\,(H_1) = (1-\beta)\sum_{J=1}^{\infty} J\beta^{J-1} = (1-\beta)\sum_{k=0}^{\infty}(k+1)\beta^k =$$

$$= (1-\beta)\left[\sum_{k=0}^{\infty} k\beta^k + \sum_{k=0}^{\infty}\beta^k\right] = (1-\beta)\left[\frac{\beta}{(1-\beta)^2} + \frac{1}{1-\beta}\right]$$

$$= (1-\beta)\left[\frac{\beta + 1 - \beta}{(1-\beta)^2}\right] = \frac{1}{1-\beta}$$

$$\boxed{\text{ARL}_1 = \sum_{j=1}^{\infty} j\,\beta^{\text{j-1}}(1-\beta) = \frac{1}{1-\beta}}$$

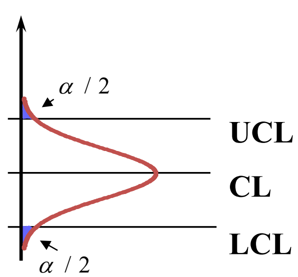
Shewhart control charts:

$$\text{ARL}_1 = \sum_{j=1}^{\infty} j\,\beta^{\text{j-1}}(1-\beta) = \frac{1}{1-\beta}$$

In-control process

$$\text{ARL}_0 = 1/\alpha$$

$$\alpha = 0.0027 \Rightarrow \text{ARL}_0 = 370.4$$

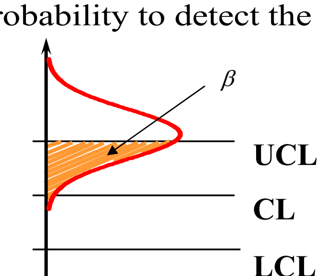


- Out-of-control process (shift)

$$\text{ARL}_1 = 1/1-\beta$$

con $1\text{-}\beta$ = power =

= probability to detect the shift



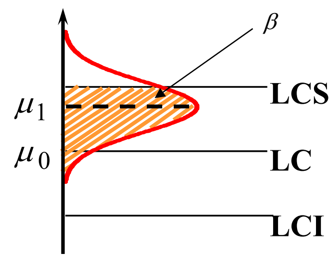
Example

$$X \overset{iid}{\sim} N(\mu_0, \sigma) \qquad \mu_0 \rightarrow \mu_1 = \mu_0 + 2\sigma$$

$$LCS = \mu_0 + 3\sigma$$

$$LC = \mu_0$$

$$LCI = \mu_0 - 3\sigma$$



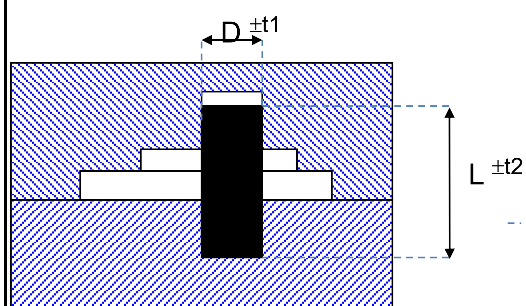
$$1-\beta = \Pr(X < LCI) + \Pr(X > LCS) =$$

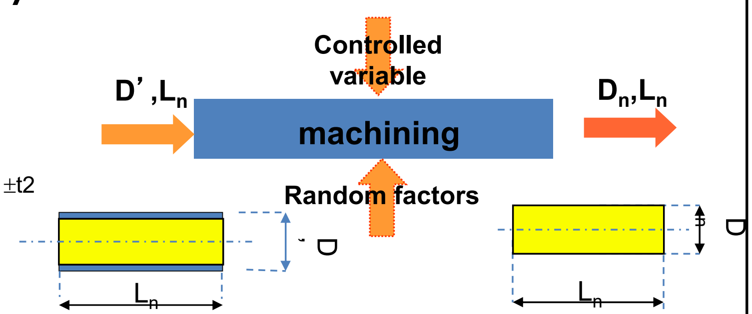
$$= \Pr\left(\frac{X-\mu_1}{\sigma} < \frac{\mu_0 - 3\sigma - \mu_1}{\sigma}\right) + \Pr\left(\frac{X-\mu_1}{\sigma} > \frac{\mu_0 + 3\sigma - \mu_1}{\sigma}\right) =$$

$$= \Pr(Z < -5) + \Pr(Z > 1) = 0.000000 \ + 0.158655 \ = 0.158655$$

$$ARL_1 = 1/0.158655 = 6.3$$

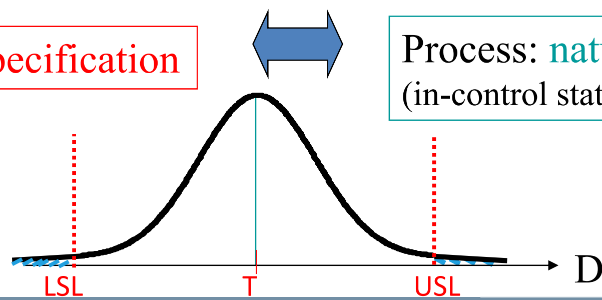
1. SPC and process capability





Design quality: specification

Process: natural tolerance
(in-control state)



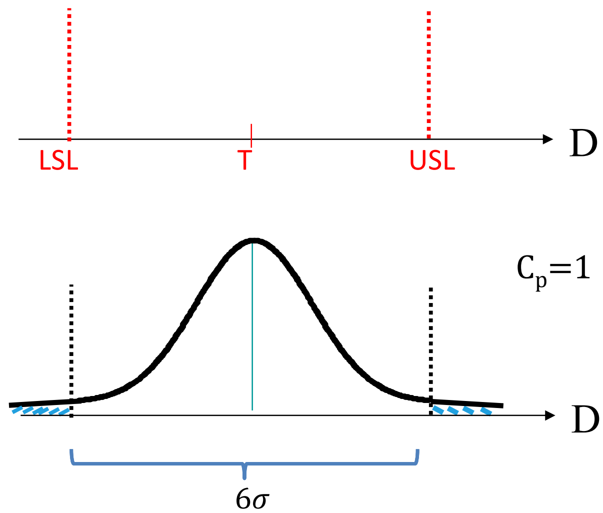
Expected
nonconforming rate

Process capability

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

$$C_{pu} = \frac{USL - \mu}{3\sigma} \qquad C_{pl} = \frac{\mu - LSL}{3\sigma}$$

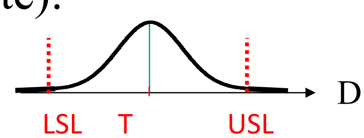
$$C_{pk} = \min\left(C_{pu}; C_{pl}\right)$$



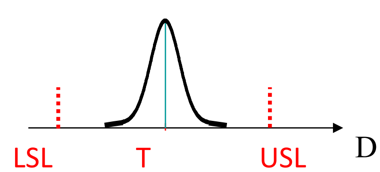
1. SPC and process capability

Process capability (expected nonforming rate):

1. Process capability - std

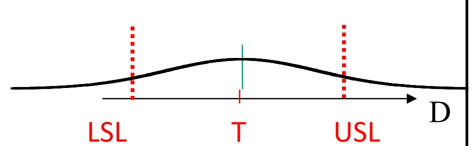


2. process capability - high
   - I can accept process changes

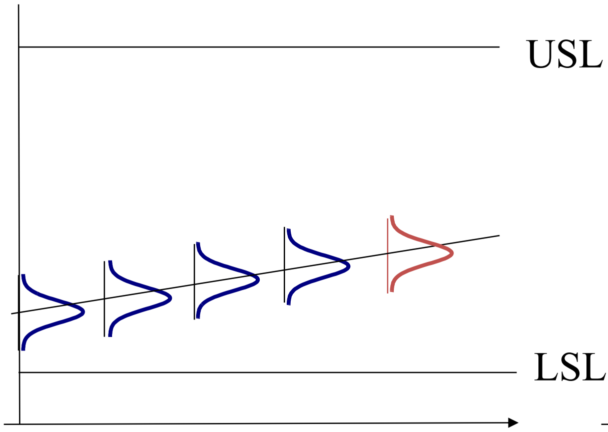


3. processo capability - low

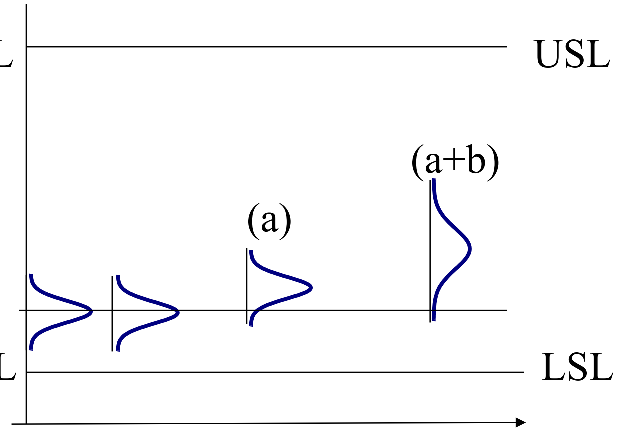
   Reduce variability due to common causes:
   - Process improvement
   - Change the process



Case 2 – high capability



Special cause: trend



Special cause:- shift mean (a)
- shift dispersion (b)