Quality Data Analysis

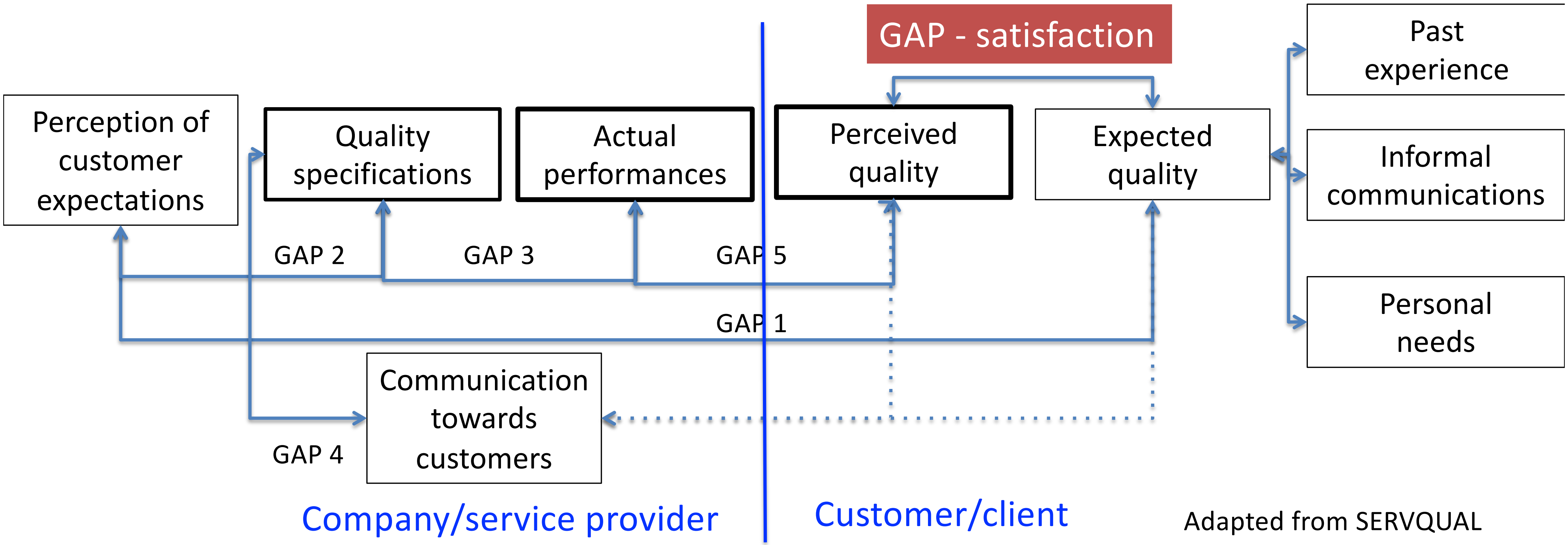
Data modeling- part 1

Bianca Maria Colosimo – biancamaria.colosimo@polimi.it

Quality of design and Quality of conformance

Quality and customer satisfaction

The customer lack-of-satisfaction may originate from different «gaps» between the customer expectations and their interpration made by the company that provides the service/product.



Adapted from SERVQUAL

The simplest (traditional model)

Product / Service

Contract

Company/service provider

Customer/client

Traditional definition of quality:

- ✓ Quality = conformance to requirements (Crosby)
- ✓ Quality = fitness to use (Juran)

•Focus on the first part (producer)

•Assumption: second part is “qualified” (able to define what she/he wants in the contract)

Quality specifications = Perception of customer expectation

GAP

Actual performances = perceived quality = expected quality

CONTRACT

Quality of design

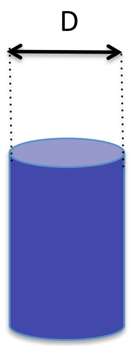
Quality of design (Specifications) is the quality which the producer or supplier is intending to offer to the customer.

Designer should take into consideration the customer's requirements in order to satisfy them with **fitness for use** of the product.

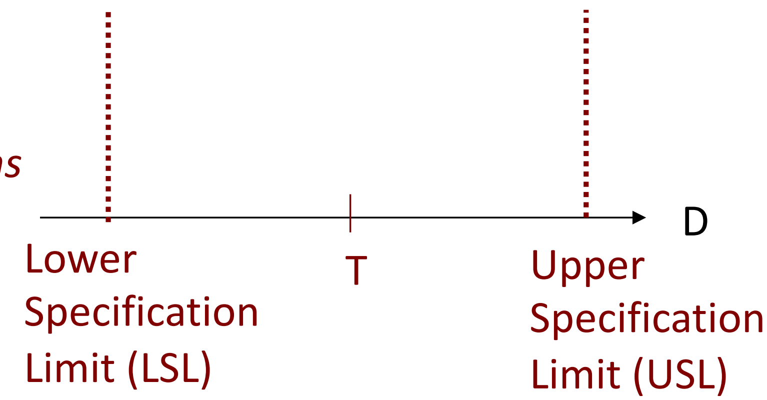
Specifications are **target** and **tolerances** determined by the designer of a product.

- Targets are the ideal values for which production is expected to strive;
- tolerances are acceptable deviations from these ideal values as it is difficult to meet the exact targets all the time due to variability in material, machine, men and process

Example



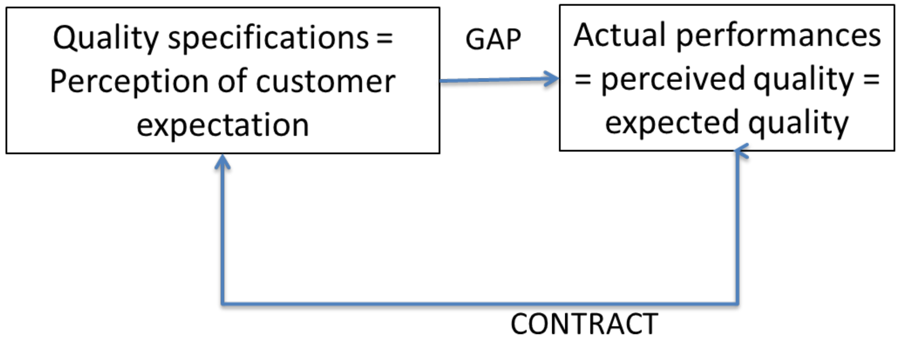
Quality of design: *specifications*



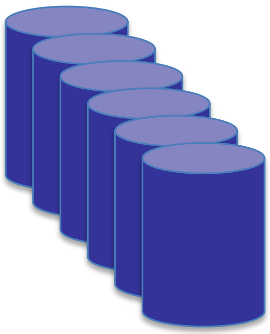
Quality of conformance

Quality of conformance is the level of the quality of product actually produced and delivered through the production or service process of the organization as per the specifications or design.

When the quality of a product entirely conforms to the specification (design), the quality of conformance is deemed excellent.

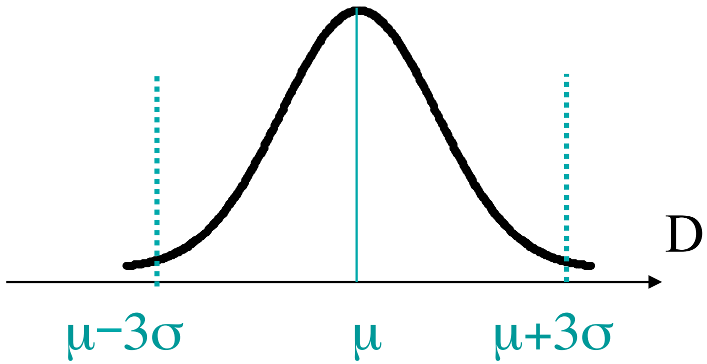


Example



$$D \sim N(\mu, \sigma^2)$$

Model of diameter data



Quality of design and quality of conformance

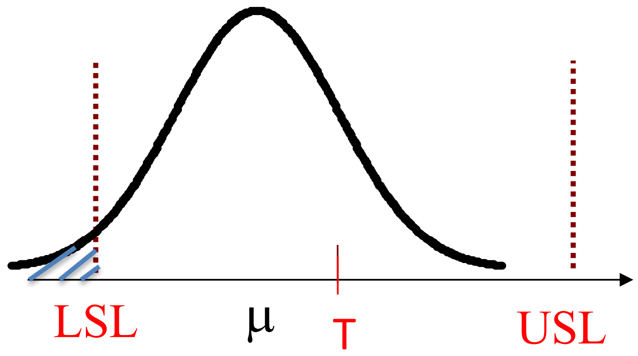
Quality of conformance:

A possible (and useful) model – expected non-conforming items (waste!)

$$\gamma_L = P\left(D \le LSL \mid D \sim N(\mu,\sigma^2)\right) =$$

$$= P\left(\frac{D-\mu}{\sigma} \le \frac{LSL-\mu}{\sigma} \mid D \sim N(\mu,\sigma^2)\right) =$$

$$= P\left(Z \le \frac{LSL-\mu}{\sigma} \mid Z \sim N(0,1)\right) = \Phi\left(\frac{LSL-\mu}{\sigma}\right)$$



$$\gamma_U = 1 - P\left(D \le USL \mid D \sim N(\mu,\sigma^2)\right) =$$

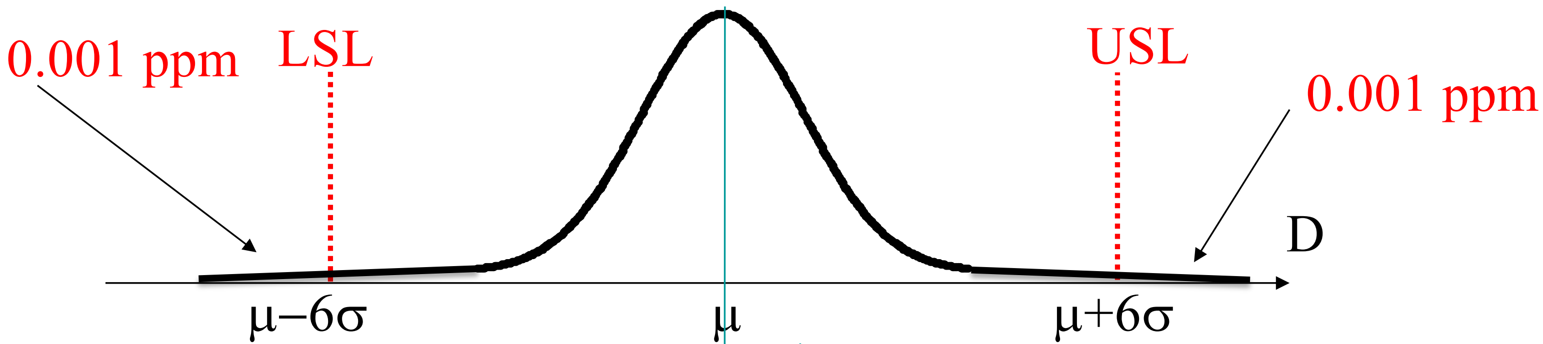
$$= 1 - \Phi\left(\frac{USL-\mu}{\sigma}\right)$$

Q1: what is the target location minimizing $\gamma=\gamma_L+\gamma_U$?

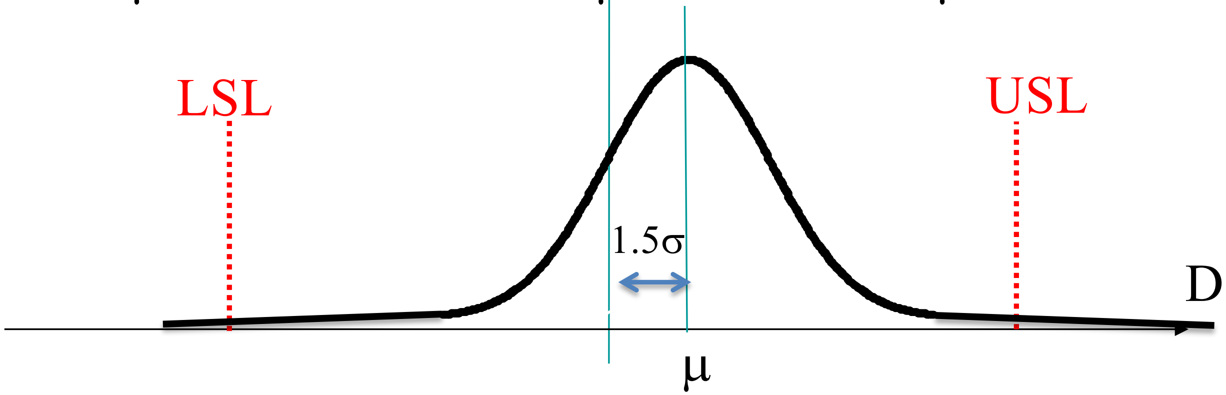
Quality of design and quality of conformance

"Six sigma quality performance": (Motorola '80s – in GE '90s) in US company:

processes characterized by very low nonconforming rate 0.002 PPM (process has the mean on the target)



Q2: what is the expected defective percentage when the process experiences a process shift equal to 1.5 σ?



Product / Service

Contract

Company/service provider

Customer/client

Needs

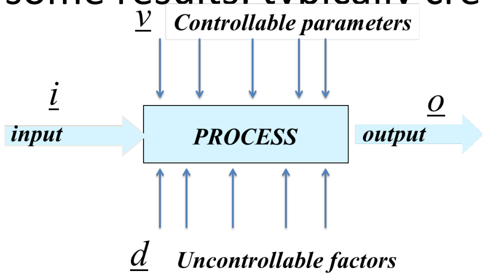
New definition (it includes the concept of conformance to requirements)

- ✓ Quality should be oriented to current and future needs of customers
- ✓ Quality = (Deming) capability to satisfy (expressed and unexpressed) customer needs
- Focus on the customer (II part)
- Form quality inspection to quality evaluation
- Services

Quality data: measurement system

Process

Any collection of activities to achieve some results, typically creating added value for the customer



Examples: manufacturing or assembly of a product, writing computer code, purchasing, billing, and treating of patients in a hospital; employee training, etc.

- Industrial process: "any process that comes in physical contact with the hardware or software that will be delivered to an external customer, up to the point the product is packaged"
- Business or administrative process: all service processes and processes supporting the industrial ones (billing, payroll, etc.)

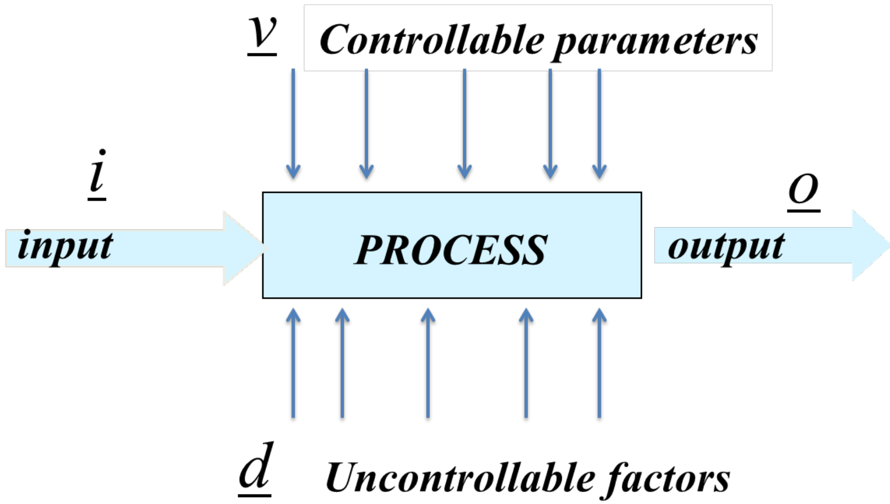
NB: customers/users are 5 times more likely to abandon a company for its poor business processes than for its poor product (Harrington, 1991).

Most processes (e.g., purchasing or product design) are cross-functional, spanning various functions areas and organizational charts

Process

Modeling quality data as process output. Why?

- To better understand the past behaviour
- To predict future behaviour
- To decide appropriate actions (to improve the process).



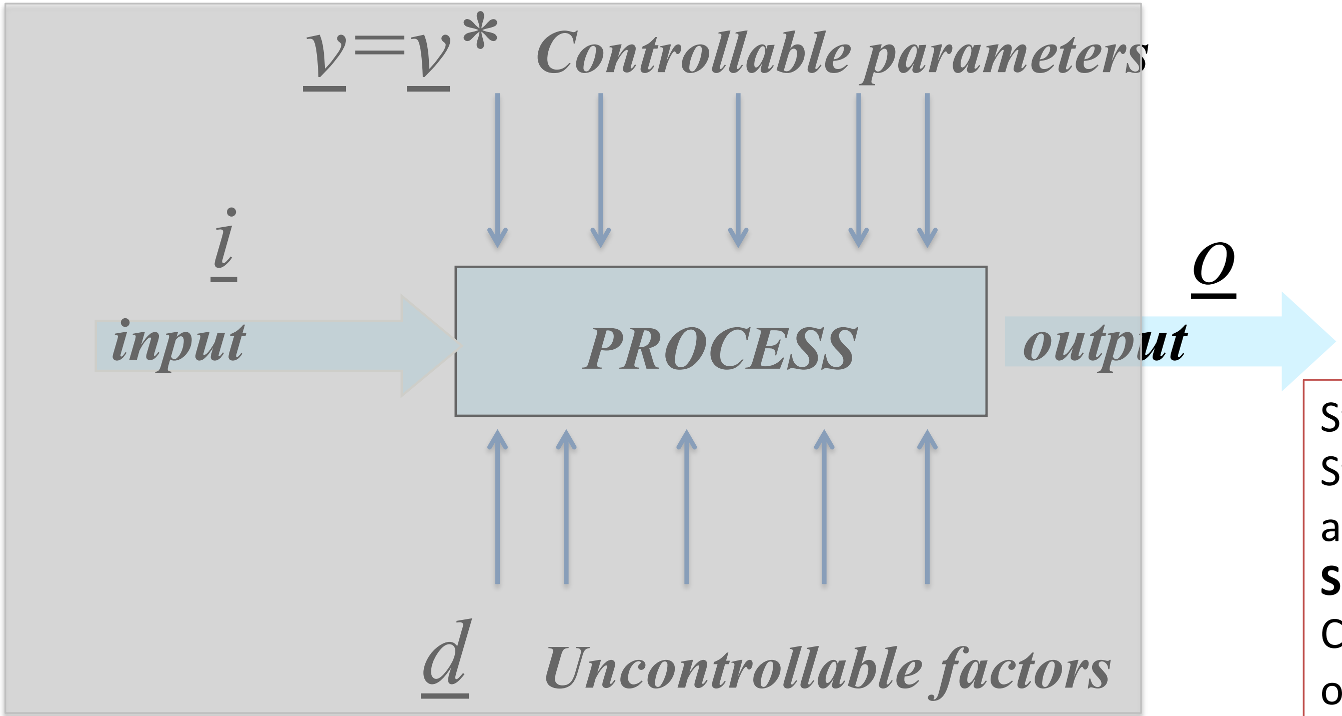
Design of Experiment – to decide the right set of experiments and analyse results to check the effect of controllable parameters to the quality output

Process optimization – to select the "best" setting of the controllable parameters to optimize the quality data

Process monitoring

Let us assume to set the controllable inputs at their “optimal” or “target” level

(---: often the operators use to change controllable parameters without any warning!)



Statistical quality monitoring,
Statistical process monitoring
also known as
Statistical process control (SPC):
Check for stability of the output over time

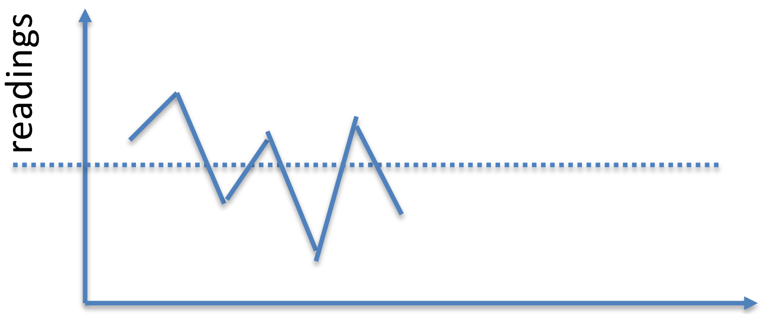
Let us take a sample of quality data observed as output o=*Y*

Let’s do an experiment

Stopwatch data:

Take the stopwatch (on your smartphone or watch)

- Try to stop the stopwatch after 10 seconds for 20 times
- List the ten reading values.
- Plot it on a time series plot. Compute the average and the rang (max – min)
- What can you observe?



Measurement system

Measurement system performance:

- Trueness: it is the difference between the average of repeated measurements and the true value of the measured feature (if the difference exists: systematic error or bias)
- Precision: ability of the measurement system to replicate the reading of the same item

Average (of this time series)

Repeated measurements (n=50)

True value =10

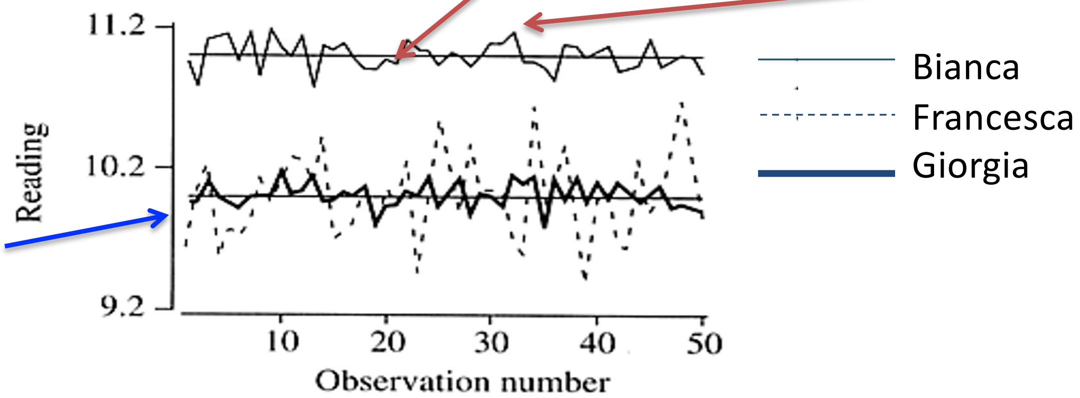


Figure 2.3 Relative comparisons of accuracy and precision.

Open question
Can you find a way to reduce "unprecision" (reduce the variability or better the standard deviation) of measurements coming out from Francesca?

NB: In the past, accuracy was used instead of trueness

Stopwatch data set: experiment is to sequentially start and stop a electronic stopwatch as close as possible to 10 seconds. Measurements are taken to the nearest 0.01 s

Vocabulary (VIM)

2.14

measurement trueness

trueness of measurement

trueness

closeness of agreement between the average of an infinite number of replicate <u>**measured quantity values**</u> and a <u>**reference quantity value**</u>

NOTE 1 Measurement trueness is not a **quantity** and thus cannot be expressed numerically, but measures for closeness of agreement are given in ISO 5725.

NOTE 2 Measurement trueness is inversely related to **systematic measurement error**, but is not related to **random measurement error**.

NOTE 3 **Measurement accuracy** should not be used for 'measurement trueness' and vice versa.

2.13 (3.5)

measurement accuracy

accuracy of measurement

accuracy

closeness of agreement between a <u>**measured quantity value**</u> and a <u>**true quantity value**</u> of a <u>**measurand**</u>

NOTE 1 The concept 'measurement accuracy' is not a **quantity** and is not given a **numerical quantity value**. A **measurement** is said to be more accurate when it offers a smaller **measurement error**.

NOTE 2 The term "measurement accuracy" should not be used for **measurement trueness** and the term **measurement precision** should not be used for 'measurement accuracy', which, however, is related to both these concepts.

4.14

resolution

smallest change in a <u>**quantity**</u> being measured that causes a perceptible change in the corresponding <u>**indication**</u>

NOTE Resolution can depend on, for example, noise (internal or external) or friction. It may also depend on the **value** of a quantity being measured.

2.15

measurement precision

precision

closeness of agreement between <u>**indications**</u> or <u>**measured quantity values**</u> obtained by replicate <u>**measurements**</u> on the same or similar objects under specified conditions

NOTE 1 Measurement precision is usually expressed numerically by measures of imprecision, such as standard deviation, variance, or coefficient of variation under the specified conditions of measurement.

NOTE 2 The 'specified conditions' can be, for example, **repeatability conditions of measurement**, **intermediate precision conditions of measurement**, or **reproducibility conditions of measurement** (see ISO 5725-3:1994).

NOTE 3 Measurement precision is used to define **measurement repeatability**, **intermediate measurement precision**, and **measurement reproducibility**.

NOTE 4 Sometimes "measurement precision" is erroneously used to mean **measurement accuracy**.

https://www.iso.org/sites/JCGM/VIM/JCGM_200e_FILES/MAIN_JCGM_200e/02_e.html

Gauge R&R (Repeatibility and Reproducibility)

Repeatibility is the dispersion of measurements of the **same measurand**, when measurements are acquired in *the same conditions* (the same method, the same operator, the same instrument, the same place, the same usage conditions, measurements taken at short interval time)

Reproducibility: it measures the dispersion when **one or more of the usage conditions are changing**.

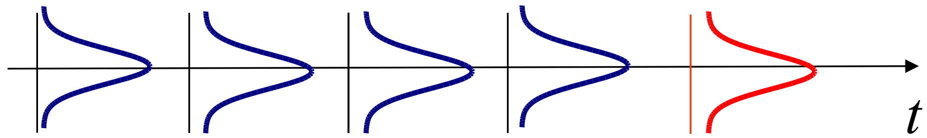
R&R study is required to trust the measurement instrument.

Quality data modeling: main assumptions

Process model

Reference model (in basic statistics and traditional SPC):

$$Y_t = \mu_t + \varepsilon_t$$



assuming

$-\mu_t = \mu \quad \text{(cost)}$

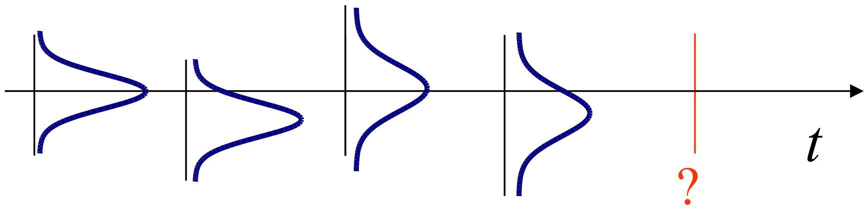
$-\varepsilon_{\mathrm{t}} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$

- normally distributed data
  (many different causes of variability, none of them dominant)
- independence
- constant variance

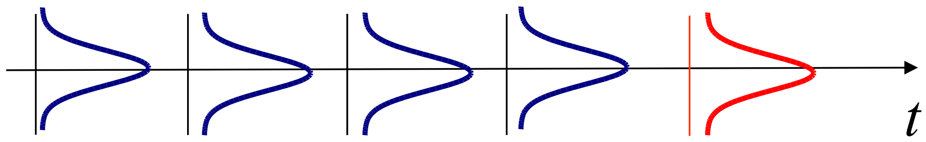
Best prediction $\hat{y}_t = \mu$

A mathematical model for the observed data

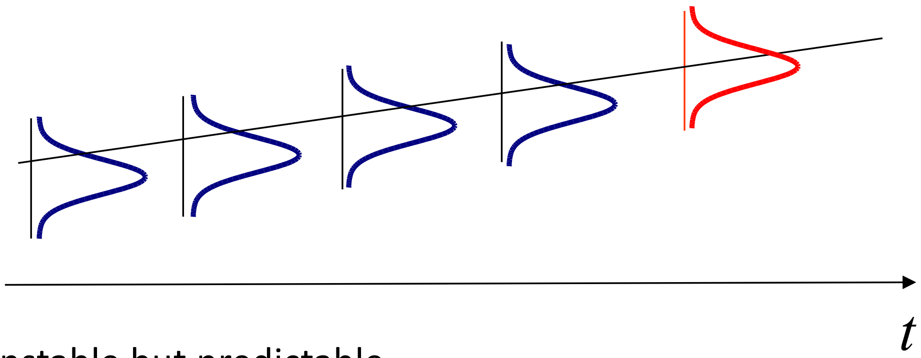
Model of data at time instant *t*: $Y_t \sim \mathrm{N}(\mu_t, \sigma_t^2)$



Unstable



Stable



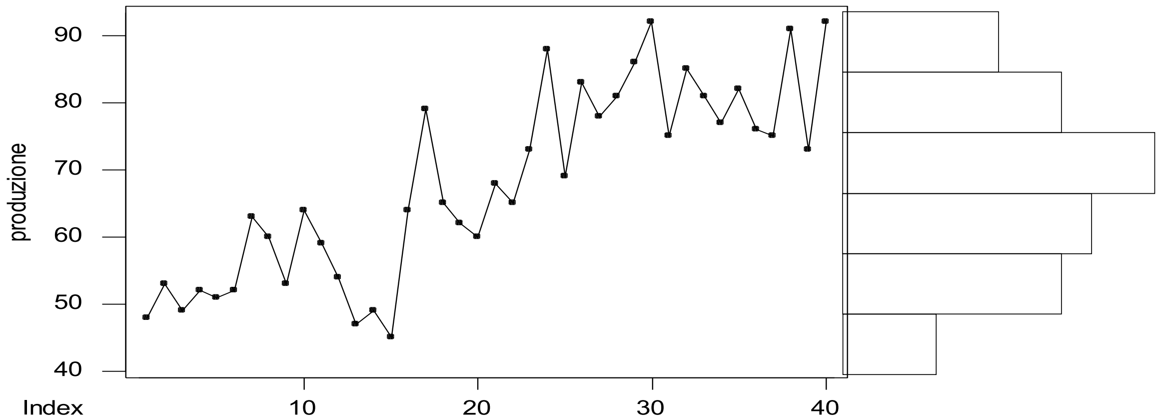
Unstable but predictable

Why do we need a model?
(- To analyze the past)
- To predict the future

How can we decide whether the standard model is appropriate?

Assumptions	**Hypothesis test (to check the assumption)**	**Remedy in case of violation**
"independence" (random pattern)	- Runs test - Bartlett's test - LBQ's test	-gapping -batching -(Linear) regression -Time series (ARIMA)
Normal distribution	Normality test	Transform data

Run chart (or time series plot)



Using the histogram, we are loosing relevant information about time

“Non random” processes

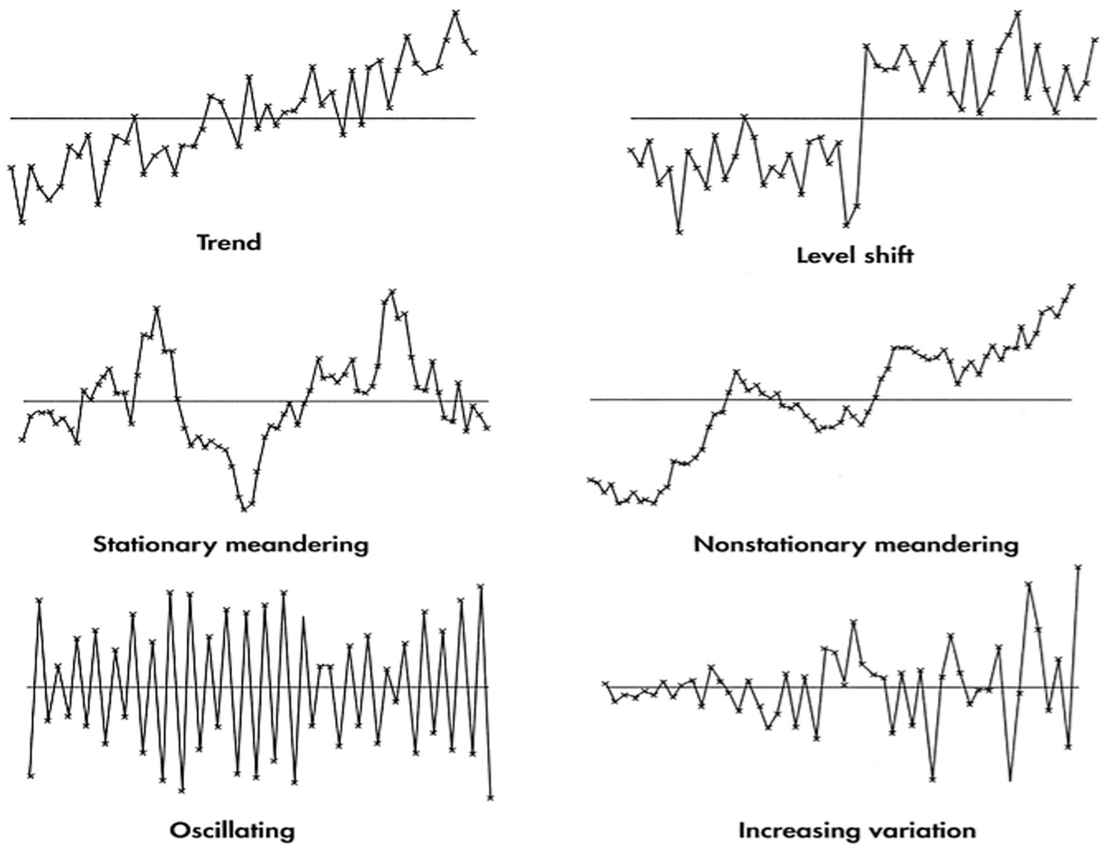


Figure 2.4 Time-series plots of different types of nonrandom process behavior.

– The mean of the process is not constant

– There is a systematic pattern (and the process mean is not the better prediction for future data)

– Dispersion around mean value is not constant

Independence

Does the process have a non random pattern?

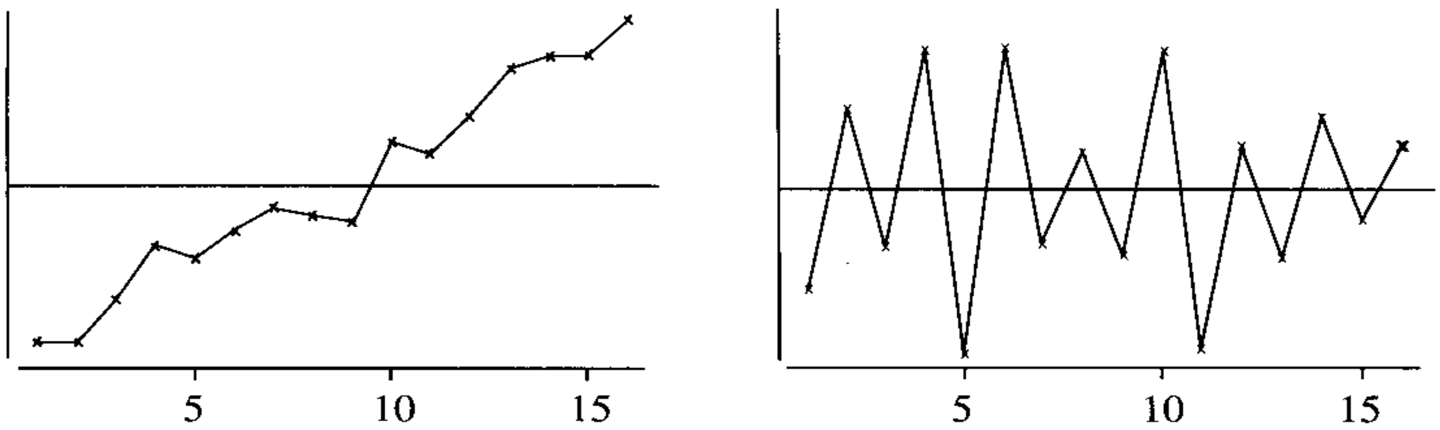
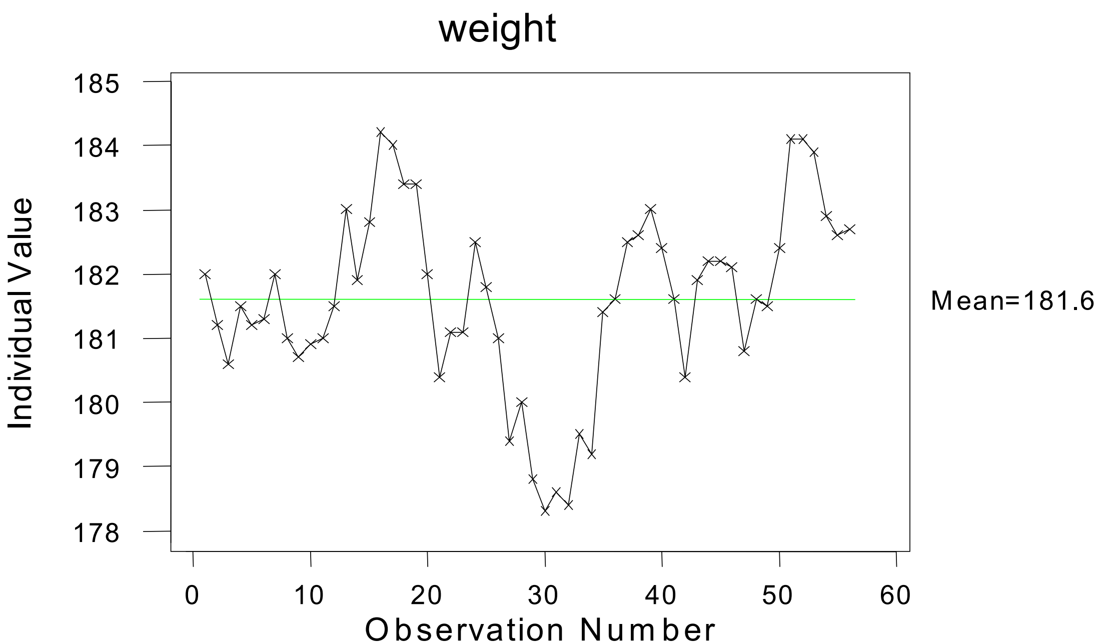


Figure 3.1 Two extreme nonrandom series ($n = 16$)

Random process: Runs test*



- Are the data random?
- If the answer were yes, we'll expect data randomly going up and down with respect to the central line (average)

* *Alwan, p.118*

The test classifies the data as lying above (+) and below (-) a reference line. In general, the reference line is the overall mean of the observed data.

Run: is a sequence of successive and equal symbols that precedes a different symbol.

No specific assumption of distribution is required (non-parametric test)

Runs test: motivation

Sequence ---++----+++--++

6 run in 16 observations (con 9 – e 7 +)

Two extreme situations: ---------+++++++ : 2 runs in 16 observations

-+-+-+-+-+-+-+--: 15 runs in 16 observations(*)

Too many or too few runs reflect nonrandom pattern

Y = number of runs $\quad E(Y) = \dfrac{2m(n-m)}{n} + 1$

n: number of observations

m: number of +

In the text case: 2*7*9/16+1=8.875 (about one half of the number of runs characterizing the two extreme behaviours) (*))

(*) there is typo in Alwan's book

Runs test

Pbm: when we can say that the number of observed runs (6 in the example) is “close enough” to 8.875?

If the process is random, the number of runs observed on a large number of samples will be (approximately) distributed as a normal with mean *E(Y)* and standard deviation :

Y = number of runs

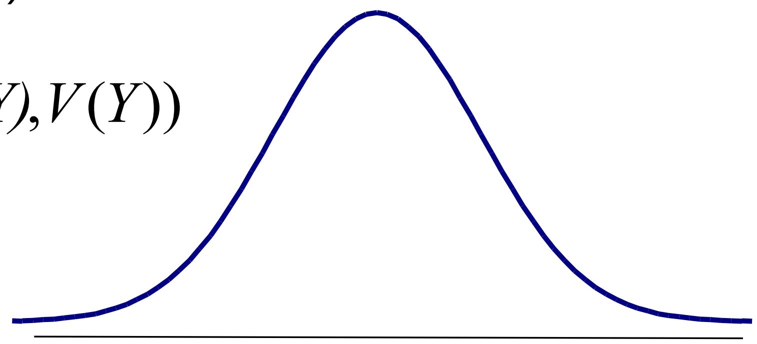
$$E(Y) = \frac{2m(n-m)}{n} + 1$$

$$\sqrt{V(Y)} = \sqrt{\frac{2m(n-m)[2m(n-m)-n]}{n^2(n-1)}}$$

n: number of data

m: number of +

$$Y \sim \mathrm{N}(E(Y), V(Y))$$



In the example: $E(Y) = 8.875$ $\sqrt{V(Y)} = 1.9$

Sequence ---++----+++--++ 6 runs in 16 data (with 9 – and 7 +)

Runs test

H_0: process is random

H_1: process is not random

Decision	State of the nature	
	H_0 is true	H_1 is true
"Accept" H_0	correct decision	II type error (β)
Reject H_0	1° type error (α)	correct decision

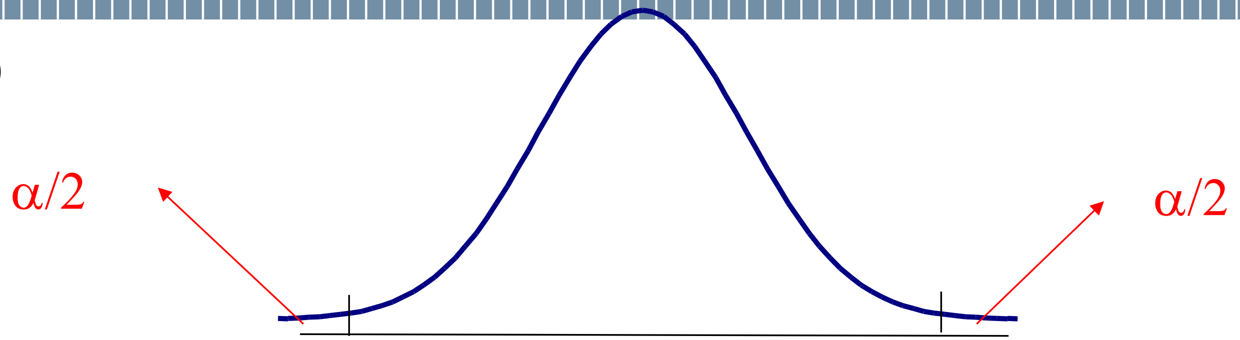
$$\alpha = P[\text{first type error}] = \text{P}[\text{reject } H_0 \mid H_0 \text{ is true}]$$

$$\beta = P[\text{second type error}] = \text{P}[\text{"accept" } H_0 \mid H_1 \text{ is true}]$$

For a given sample size *n*: trade-off between α e β

Runs test

$$Y \sim \mathrm{N}(E(Y), V(Y))$$



$$E(Y) - z_{\alpha/2}\sqrt{V(Y)} \qquad E(Y) + z_{\alpha/2}\sqrt{V(Y)}$$

Once α is set, if H_0 is true

with α=5%

$$z_{\alpha/2} = 1.96 \Longrightarrow E(Y) \mp z_{\alpha/2}\sqrt{V(Y)} = (5.151, 12.599)$$

The observed value (y=6) is included in the interval: we can conclude that the associate process is random.

Runs test

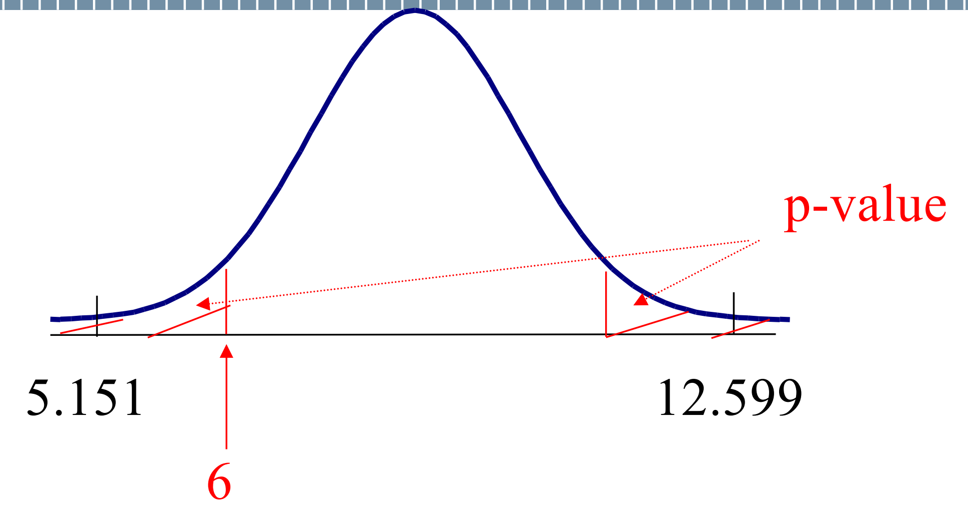
To conclude: p-value

Cumulative Distribution Function

```
Normal with mean =
8.87500 and standard
deviation = 1.90000

x      P( X <= x )

6.0000 0.0651
```



p-value=2(0.0651)=**0.1302**$>\alpha$ We cannot reject H_0

It means that **13% of the times we will see a difference between the number of runs actually observed and the expected value which is equal or greater than the value observed this time.** p-value is a measure of how “unusual” or “surprising” are data observed when the null hypothesis is true. The p-value can be also seen as the minimum level of *a* that let us reject the null hypothesis.

Example: stopwatch (timing.dat)

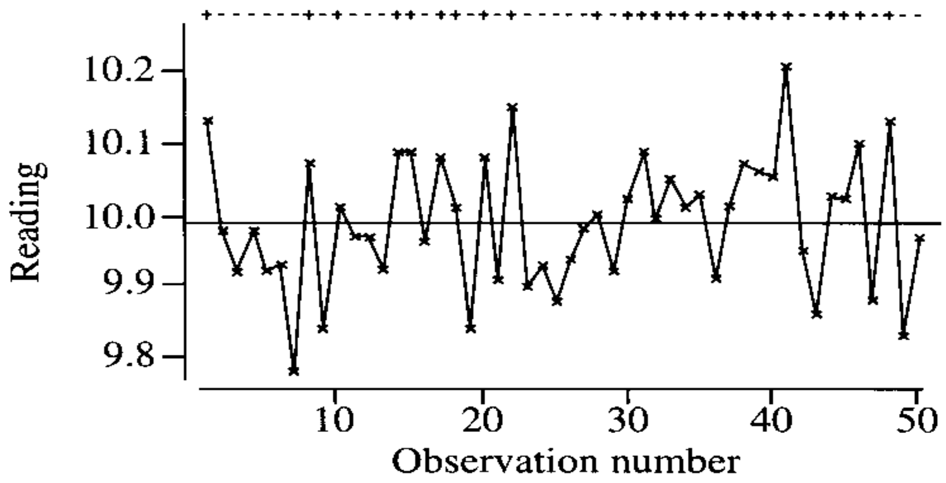


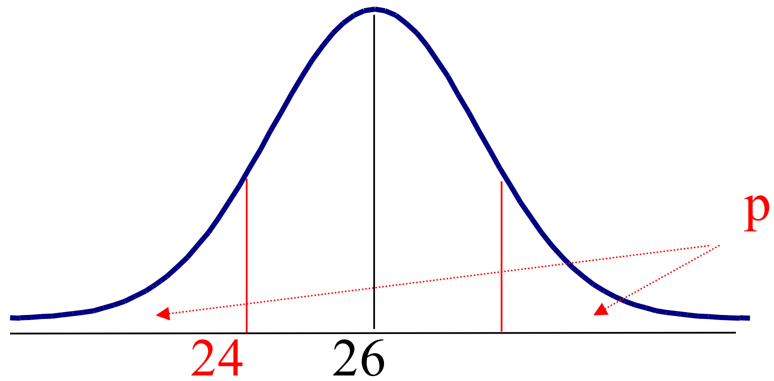
Figure 3.3 Counting number of runs in stopwatch data series.

50 data:
25 + and 25 –
24 runs

if the process is random (H_0):

$$E(\text{number of runs}) = 26$$

$$\sqrt{V(\text{number of runs})} = 3.499$$



Cumulative Distribution Function

```
Normal with mean = 26.0000 and standard deviation = 3.49900

        x     P( X <= x )
   24.0000         0.2838
```

p-value=2(0.2838)=0.5676

For a random process, 56% of the times we'll see a distance from the expected value that is equal or greater than the one observed in this case (p-value & unusual observation)

Using minitab:

```
MTB > runs 'reading'

Runs Test

    reading

    K =    9.9890

    The observed number of runs =  24
    The expected number of runs =  26.0000
    25 Observations above K   25 below
              The test is significant at  0.5676
              Cannot reject at alpha = 0.05
```

Figure 3.4 Runs test for stopwatch data series.

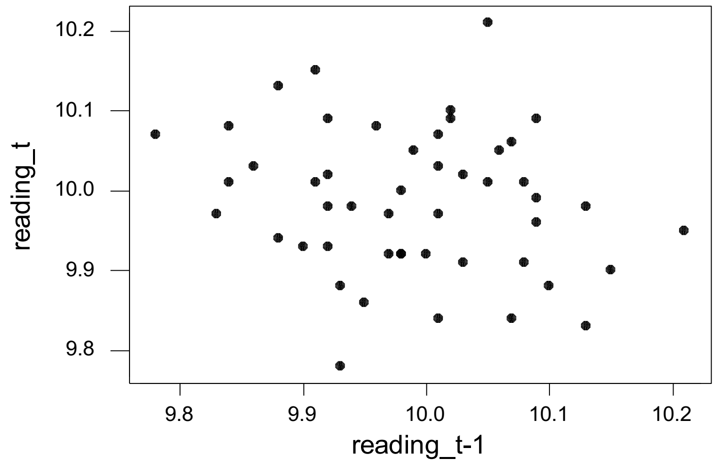
How can we decide whether the standard model is appropriate?

Assumptions	Hypothesis test (to check the assumption)	Remedy in case of violation
“independence” (random pattern)	- Runs test - Bartlett’s test - LBQ’s test	-gapping -batching -(Linear) regression -Time series (ARIMA)
Normal distribution	Normality test	Transform data

Independence 2: Autocorrelation

Lagging of one variable: create a second variable such that the observation at time *t* is close to the observation of the same time series at time t-k (lag k)

t	*reading t*	*reading t-1*	*reading t-2*	…	*reading t-4*
1	10.13	*	*		*
2	9.98	10.13	*		*
3	9.92	9.98	10.13		*
4	9.98	9.92	9.98		*
5	9.92	9.98	9.92		10.13
6	9.93	9.92	9.98		9.98
:					
46	10.1	10.02	10.03		9.95
47	9.88	10.1	10.02		9.86
48	10.13	9.88	10.1		10.03
49	9.83	10.13	9.88		10.02
50	9.97	9.83	10.13		10.1



Scatter plot: reading *t* vs. reading *t-1*

Autocorrelation

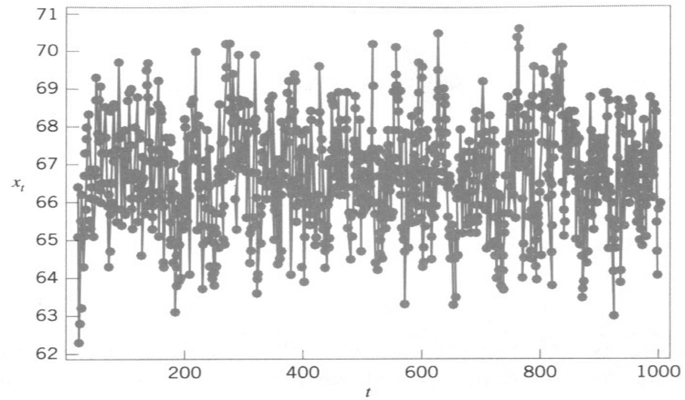


Figure 9-6 A process variable with autocorrelation.

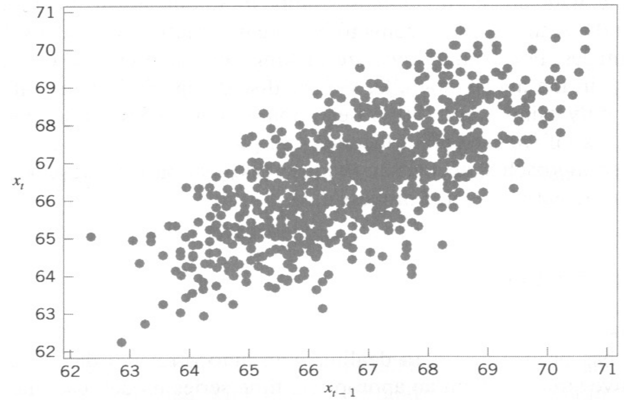
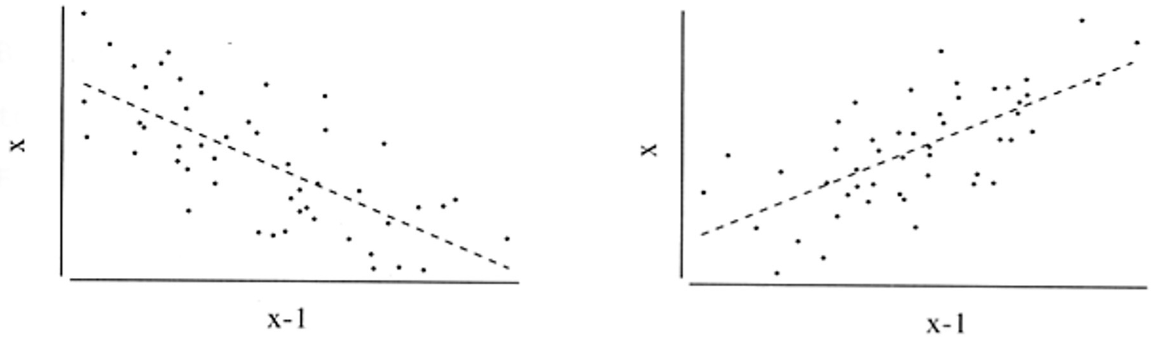


Figure 9-7 Scatter plot of x_t versus x_{t-1}.



(a) **Negative relationship**

(b) **Positive relationship**

Figure 3.6 Different scatter plots for x_t versus x_{t-1}.

LINEAR association

Autocorrelation

Autocovariance function*: $\gamma_{t,k} = Cov(X_t, X_{t-k}) = E\left[\left(X_t - \mu_t\right)\left(X_{t-k} - \mu_{t-k}\right)\right] \quad \mathrm{k} = 0, \pm 1, \pm 2, \ldots$

For stationary process (with mean μ_t constant in time): $\gamma_{t,k} = \gamma_k$

k=0: $\gamma_0 = E\left[\left(X_t - \mu\right)^2\right] = Var(X_t) = \sigma^2$

Generally, it is an a-dimensional value: autocorrelation function

with

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\mathrm{Cov}(X_t, X_{t\text{-}k})}{\mathrm{Var}(X_t)} \qquad k = 0, \pm 1, \pm 2, \ldots$$

$$-1 \le \rho_k \le 1$$

$$\gamma_k = \gamma_{-k} \Rightarrow \rho_k = \rho_{-k}$$

Plot of the function only for k≥0

* Montgomery

Autocorrelation: Bartlett's test

In order to estimate the autocorrelation function:

Sample autocovariance and autocorrelation functions:

$$\gamma_{t,k} = Cov(X_t, X_{t-k}) = E\left[\left(X_t - \mu_t\right)\left(X_{t-k} - \mu_{t-k}\right)\right]$$

$$\boxed{c_k = \hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}$$

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\mathrm{Cov}(X_t, X_{t+k})}{\mathrm{Var}(X_t)}$$

$$\boxed{r_k = \hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \qquad k \le n/4}$$

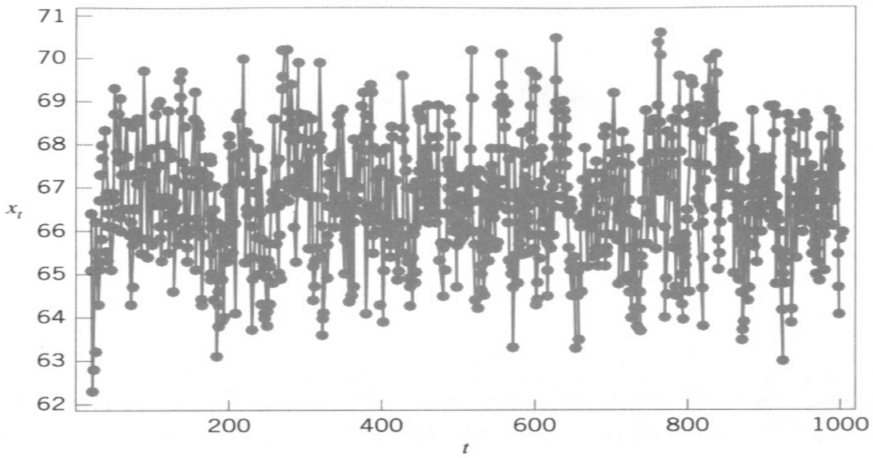


Figure 9-6 A process variable with autocorrelation.

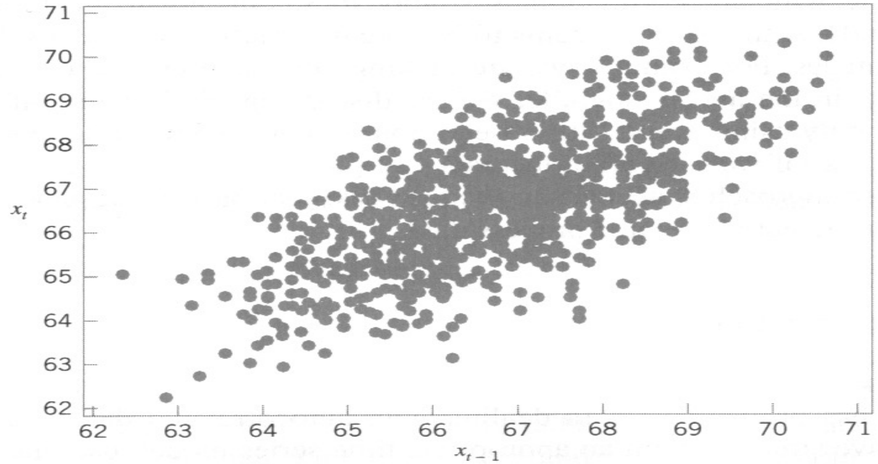


Figure 9-7 Scatter plot of x_t versus x_{t-1}.

$$\rho_k = \frac{\mathrm{Cov}(x_t, x_{t-k})}{V(x_t)} \qquad k = 0, 1, \ldots$$

$$r_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} \qquad k = 0, 1, \ldots, K$$

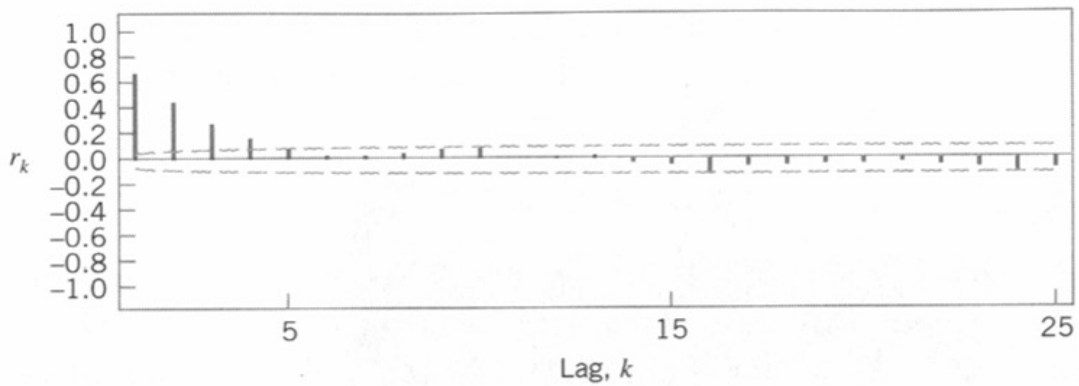


Figure 9-8 Sample autocorrelation function for the data in Fig. 9-6.

Correlation and indenpendence

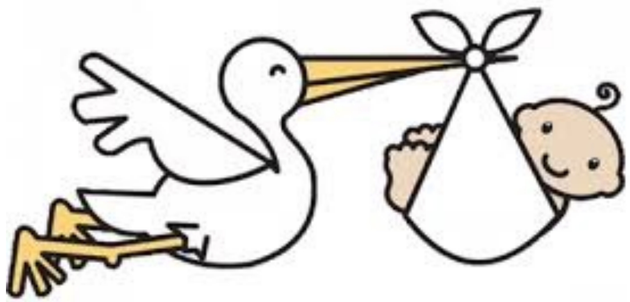
Furthermore:

$\not\Leftarrow$

1. If X_t e X_{t+k} are independent, then they are uncorrelated (ρ_k=0 for any k)
2. If X_t e X_{t+k} are correlated ($\rho_k \neq 0$ for some k's), then X_t and X_{t+k} are dependent

Pay attention to the direction of the arrows (correlation is a measure of **linear** association)

Correlation and causality



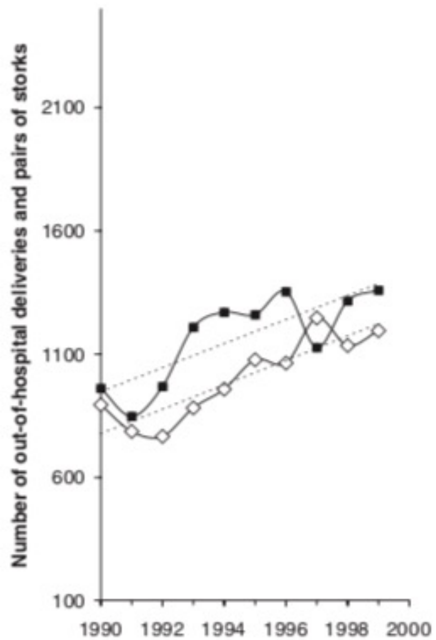


Figure 2. Storks in Brandenburg and the birthrates in Berlin, Germany (1990–99). Open triangles show number of clinical deliveries per year in Berlin. Open diamonds show number of out-of-hospital deliveries per year in Berlin. Number of pairs of storks are shown as full squares. Dotted lines represent linear regression trend (y = mx + b). For the convenience of the readers, two figures are presented. Left graph shows clinical deliveries against pairs of storks using two scalings, right graph shows numbers of out-of-hospital deliveries and pairs of storks both on the same scale. In both figures, data are from the years 1990–2000.

Correlation between the number of pairs of storks and out-of-hospital deliveries per year in Berlin

http://web.stanford.edu/class/hrp259/2007/regression/storke.pdf

Correlation and causality

fallacy: is the use of invalid or otherwise faulty reasoning, or "wrong moves"in the construction of an argument.

cum hoc, ergo propter hoc: with this, therefore because of this

Autocorrelation: Bartlett's test

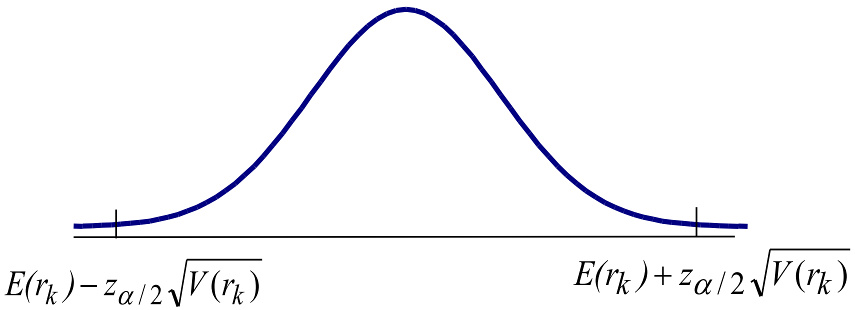
Bartlett showed that for a random process (iid), the sample autocorrelation coefficient r_k approximately follows a normal distribution with: $E(r_k)=0$ and $Var(r_k)=1/n$ $\forall k$

A useful test with $H_0 : \rho_k = 0$ vs $H_1 : \rho_k \neq 0$

It is based on the rejection region: $\boxed{| r_k | > \dfrac{z_{\alpha/2}}{\sqrt{n}}}$

If H_0 is true:

$r_k \sim N(0,1/n)$



Rejection region:

$$\left.\begin{array}{l} r_k > 0 : r_k > \dfrac{z_{\alpha/2}}{\sqrt{n}} \\ r_k < 0 : r_k < -\dfrac{z_{\alpha/2}}{\sqrt{n}} \end{array}\right\} \Rightarrow | r_k | > \dfrac{z_{\alpha/2}}{\sqrt{n}}$$

Value that is often used

$$Z_{\alpha/2} = 2 \rightarrow \alpha = 0.0456 \approx 0.05$$

ACF (autocorrelation function)

```
MTB > acf c1
```

Autocorrelation Function: C1

```
ACF of C1

             -1.0 -0.8 -0.6 -0.4 -0.2  0.0  0.2  0.4  0.6  0.8  1.0
               +----+----+----+----+----+----+----+----+----+----+
  1  -0.184                          XXXXXX
  2   0.185                               XXXXXX
  3  -0.070                             XXX
  4   0.065                               XXX
  5  -0.101                            XXXX
  6   0.015                               X
  7   0.165                               XXXXX
  8  -0.168                           XXXXX
  9  -0.104                            XXXX
 10  -0.021                             XX
 11  -0.090                            XXX
 12  -0.097                            XXX
```

$$\pm 2/\sqrt{50} = \pm 0.282$$

Bonferroni inequality

Attention not to use the test for many different lags *k's* (different r_k) at the same time.

When conducting multiple analyses on the same dependent variable, the chance of committing a Type I error increases, thus increasing the likelihood of coming about a significant result by pure chance.
To correct for this, or protect from Type I error, a **Bonferroni correction** is conducted.

Carlo Emilio Bonferroni
(1892 – 1960)

Food for thoughts – check the proof:

https://www.youtube.com/watch?v=0xuRh3dz_Nc

Bonferroni inequality

Assume we have N hypothesis tests (i=1,…,N)

- Each test has its own probability to reject $H0_i$ when it is true - α_i
- Family-wise "first type" error: α'

the probability of **rejecting at least one** null hypothesis when **they are all true**

$$\boxed{\alpha' \leq \sum_{i=1,\dots,N} \alpha_i}$$

Bonferroni inequality

For **independent** tests, it can be shown that

$$\boxed{1-\alpha' = \prod_{i=1,..N} (1-\alpha_i)}$$

If we set the same α for all the tests

$$\alpha_i = \alpha \quad \forall i = 1,\dots,N \Rightarrow \alpha' = 1-(1-\alpha)^N$$

$$\alpha = 1-(1-\alpha')^{1/N}$$

Bonferroni inequality

I can build intervals to “constrain” the family error rate *α*

(the overall first type error)

Choose the nominal family error rate alpha α'_{nom}

For each of the N tests to be performed (using the same set of data) choose

$$\alpha_i = \frac{\alpha'_{nom}}{N} \qquad \forall i = 1, \ldots, N$$

You’ll have an overall first type error $\alpha' \leq \sum_{i=1,..,N} \alpha_i = \alpha'_{nom}$

False discovery rate

Other approaches

J. R. Statist. Soc. B (1995)
57, *No.* 1, *pp.* 289–300

Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing

By YOAV BENJAMINI† and YOSEF HOCHBERG

Tel Aviv University, Israel

[Received January 1993. Revised March 1994]

SUMMARY

The common approach to the multiplicity problem calls for controlling the familywise error rate (FWER). This approach, though, has faults, and we point out a few. A different approach to problems of multiple significance testing is presented. It calls for controlling the expected proportion of falsely rejected hypotheses—the false discovery rate. This error rate is equivalent to the FWER when all hypotheses are true but is smaller otherwise. Therefore, in problems where the control of the false discovery rate rather than that of the FWER is desired, there is potential for a gain in power. A simple sequential Bonferroni-

Example: Bartlett test for autocorrelation

-Bartlett test one lag: $\boxed{|r_k| > \dfrac{z_{\alpha/2}}{\sqrt{n}}}$

-Bartlett test for *L different* lags (k=1,…,L): $$\alpha_i = \frac{\alpha'_{nom}}{L}$$

Rejection region $$\boxed{|r_k| > \frac{z_{\alpha'_{nom}/(2L)}}{\sqrt{n}} \qquad \forall k = 1,\dots,L}$$

As an alternative solution, we can use the **Test di Ljung Box Pierce** (in Minitab LBQ) based on the statistic:

$$Q = n(n+2)\sum_{k=1}^{L}\frac{{r_k}^2}{n-k}$$

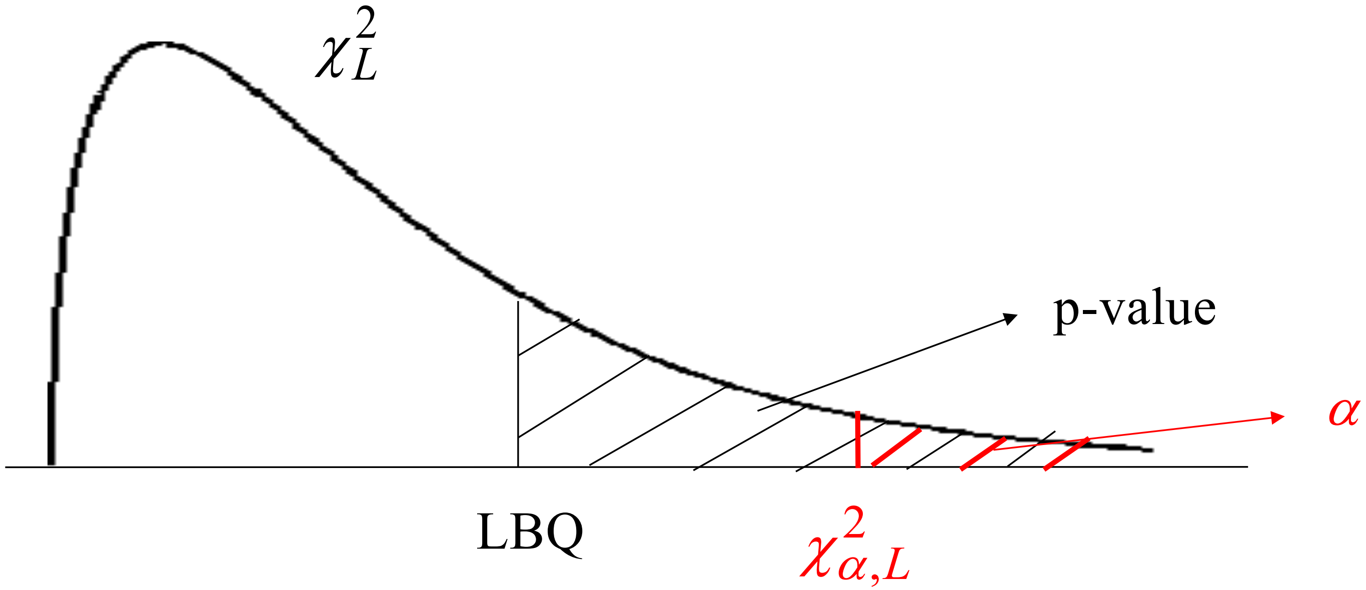
"Portmanteau" test: $H_0 : \rho_i = 0, \quad i = 1,\ldots,L$ $\qquad H_1 : \exists i \in [1,\ldots,L] \ni \rho_i \neq 0$

If H_0 true $\qquad Q \sim \chi^2_L \Rightarrow$ rejection region $Q > \chi^2_{\alpha,L}$

LBQ test

$$Q \sim \chi^2_L \Rightarrow \text{rejection region } Q > \chi^2_{\alpha,L}$$

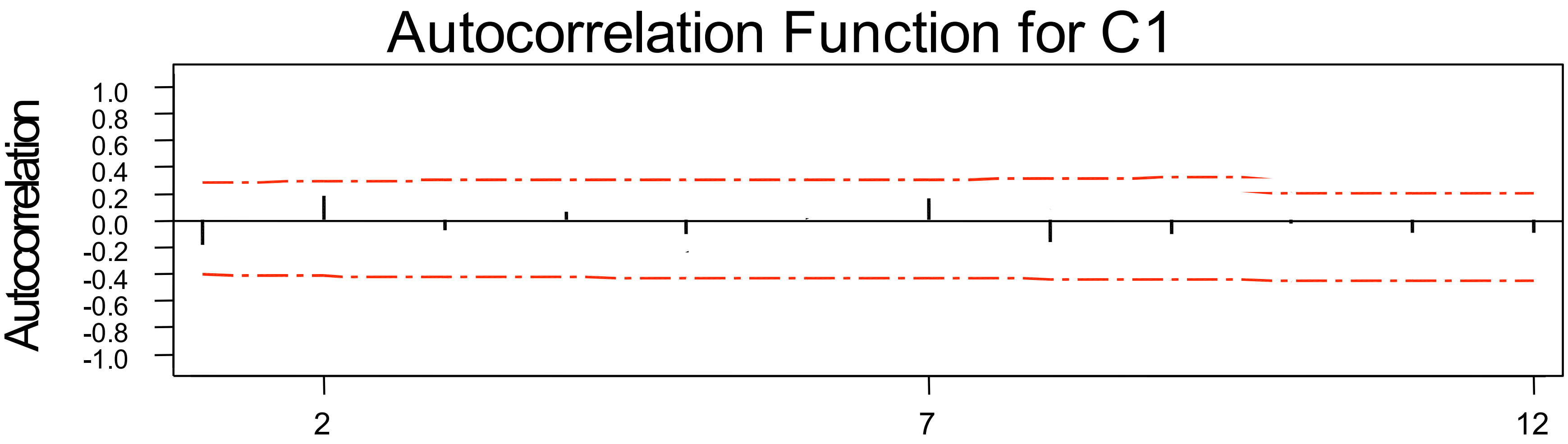
If H_0 is true



In general, when not specified $\quad L \leq \sqrt{n}$

LBQ test

In the example



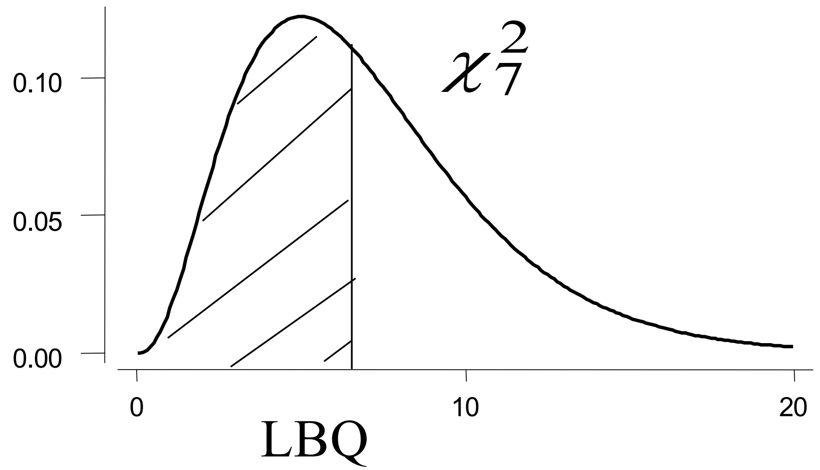
$$L \le \sqrt{50} \rightarrow L = 7$$

Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.18	-1.30	1.80	8	-0.17	-1.07	8.18
2	0.19	1.27	3.66	9	-0.10	-0.65	8.86
3	-0.07	-0.47	3.93	10	-0.02	-0.13	8.89
4	0.07	0.43	4.17	11	-0.09	-0.56	9.43
5	-0.10	-0.67	4.76	12	-0.10	-0.59	10.07
6	0.02	0.10	4.78				
7	0.17	1.08	$\boxed{6.43}$				

LBQ test

Cumulative Distribution Function

```
Chi-Square with 7 DF
          x     P( X <= x )
     6.4300          0.5095
```



p-value=1-0.5095=0.4905≅50% ⇒ with α=5% I cannot reject the null hypothesis

Link Bartlett- LBQ

$$H_0 : \rho_1 = 0 \quad \text{vs} \quad H_1 : \rho_1 \neq 0$$

For *k=1=L:*

Bartlett:

$$\boxed{r_1 \sim \mathrm{N}(0, 1/\mathrm{n}) \Rightarrow \mathrm{Z} = \frac{r_1 - 0}{\sqrt{1/n}} \sim \mathrm{N}(0,1)}$$

$$H_0 : \rho_i = 0, \quad i = 1 \qquad H_1 : \exists i \in [1] \ni \rho_i \neq 0$$

LBQ:

$$\boxed{Q = n(n+2)\sum_{k=1}^{1} \frac{r_k^{\;2}}{n-k} = \frac{r_1^{\;2}}{1/n}\,\frac{n+2}{n-1} \sim \chi_1}$$

$L=1$

$$\text{then if } \frac{n+2}{n-1} \rightarrow 1 (n \text{ large}) \Rightarrow Q = \frac{r_1^2}{1/n} = (\mathrm{Z})^2 \sim \chi_1$$

for large n, the two tests for $k = 1$ are the same