



POLITECNICO
MILANO 1863

Quality Data Analysis

Data modeling- part 1

Bianca Maria Colosimo – biancamaria.colosimo@polimi.it

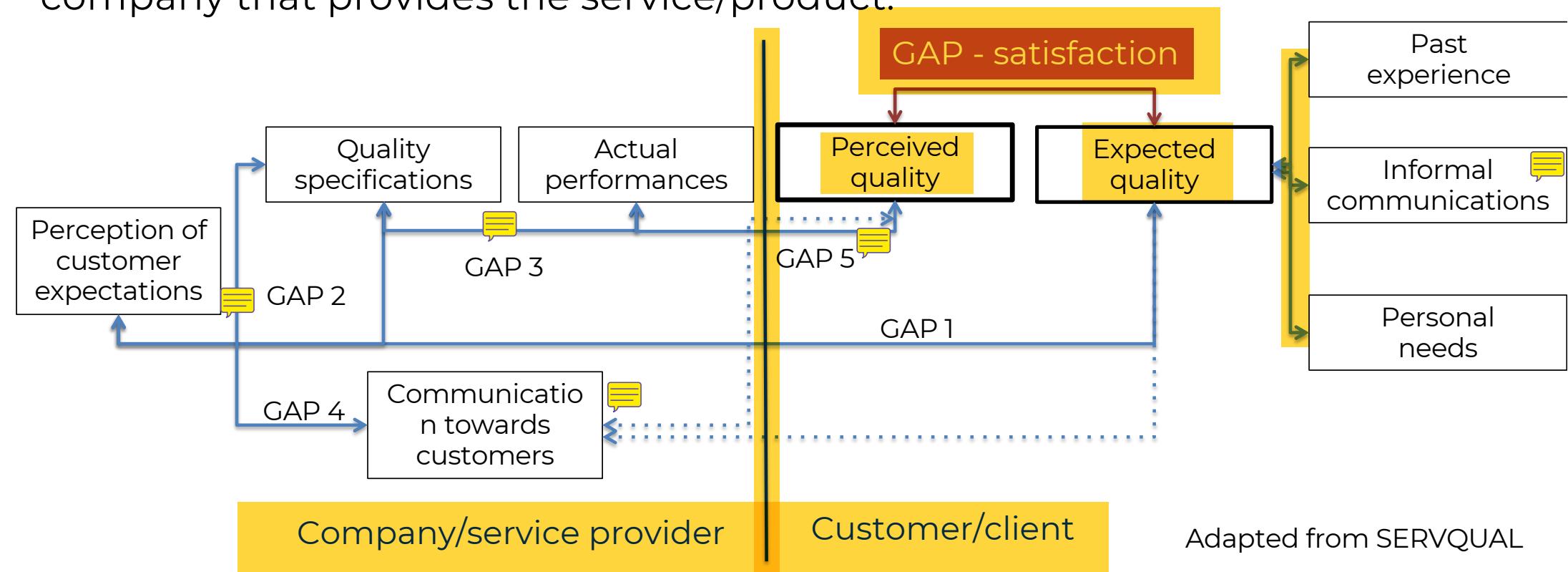
Quality of design
and
Quality of conformance 

the gap model.

customer is unsatisfied when the perceived is less than expected. (RED box)

Quality and customer satisfaction

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



The simplest (traditional model)



Company/service provider

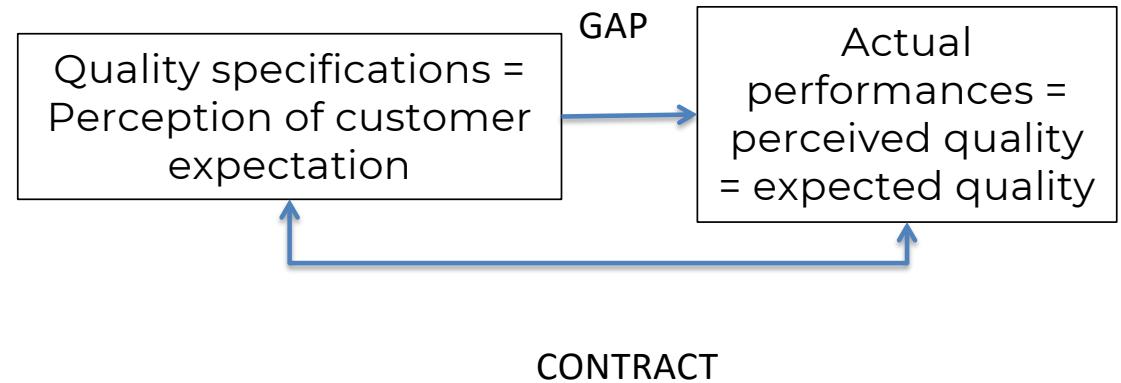


Customer/client

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 of design

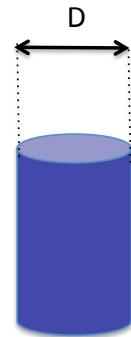
Quality of design (Specifications) is the quality which the producer or supplier is intending to offer to the customer.

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

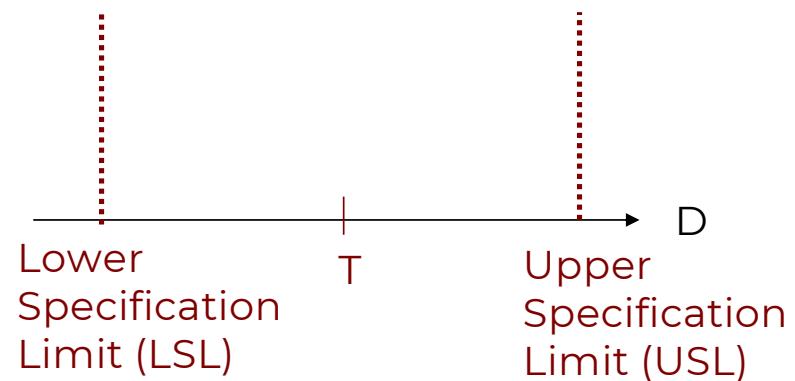
Specifications are **target** and **tolerances** determined by the designer of a product.

- 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

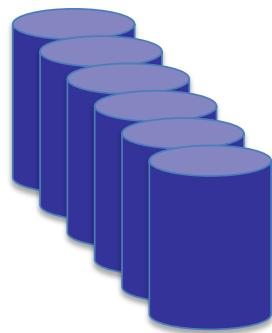


Quality of conformance

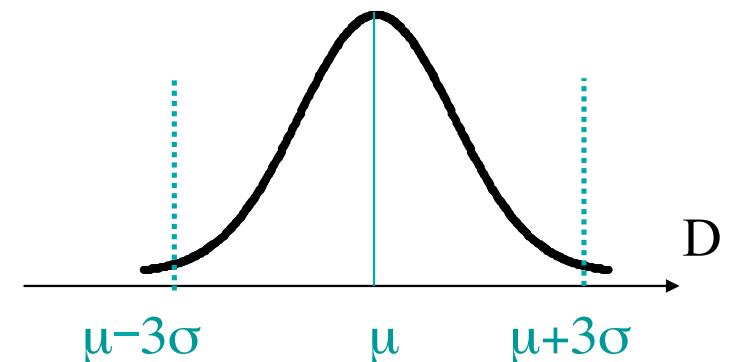
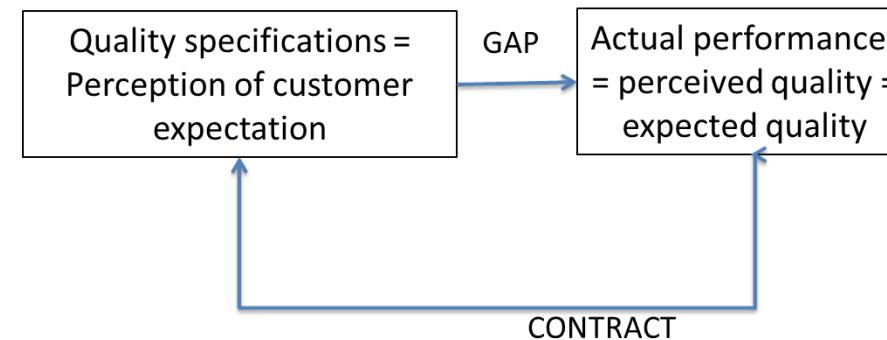
Quality of conformance is the level of the quality of product actually produced and delivered through the production or service process of the organization as per the specifications or design.

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

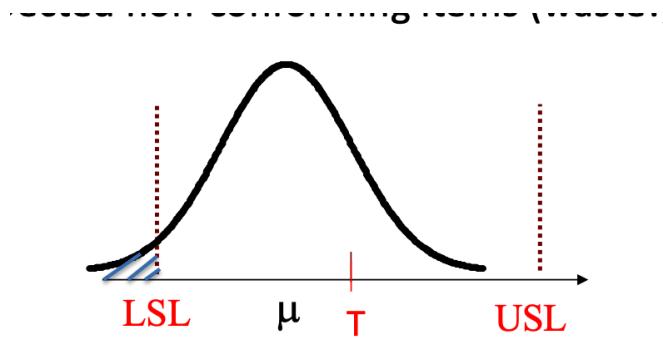


Quality of design and quality of conformance

Quality of conformance:

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

$$\begin{aligned}\gamma_L &= P(D \leq LSL | D \sim N(\mu, \sigma^2)) = \\&= P\left(\frac{D-\mu}{\sigma} \leq \frac{LSL-\mu}{\sigma} | D \sim N(\mu, \sigma^2)\right) = \\&= P\left(Z \leq \frac{LSL-\mu}{\sigma} | Z \sim N(0,1)\right) = \Phi\left(\frac{LSL-\mu}{\sigma}\right)\end{aligned}$$



$$\gamma_U = 1 - P(D \leq USL | D \sim N(\mu, \sigma^2)) =$$

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

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:

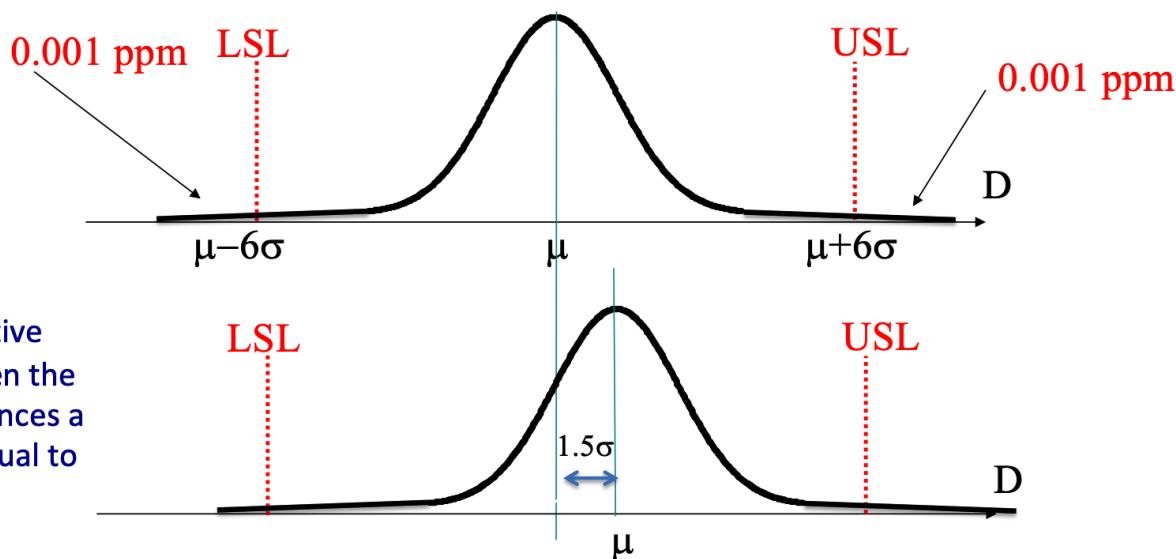
processes characterized by very low nonconforming rate 0.002 PPM (process has the mean on the target)

“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 expected defective percentage when the process experiences a process shift equal to 1.5σ ?

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



Quality of design and quality of conformance



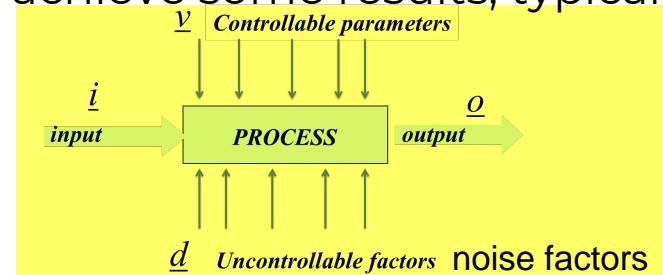
New definition (it includes the concept of conformance to requirements)

- ✓ 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)
- From quality inspection to quality evaluation

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).

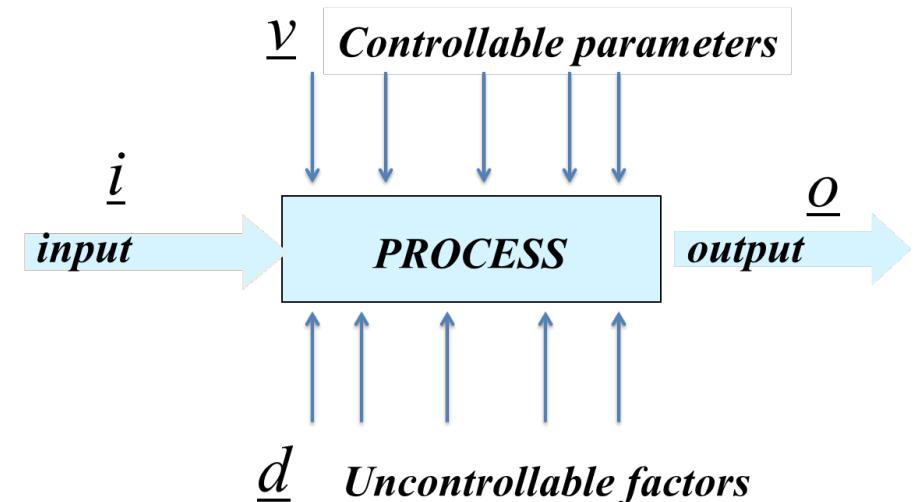
Most processes (e.g., purchasing or product design) are cross-functional, spanning various functions areas and organizational charts

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).

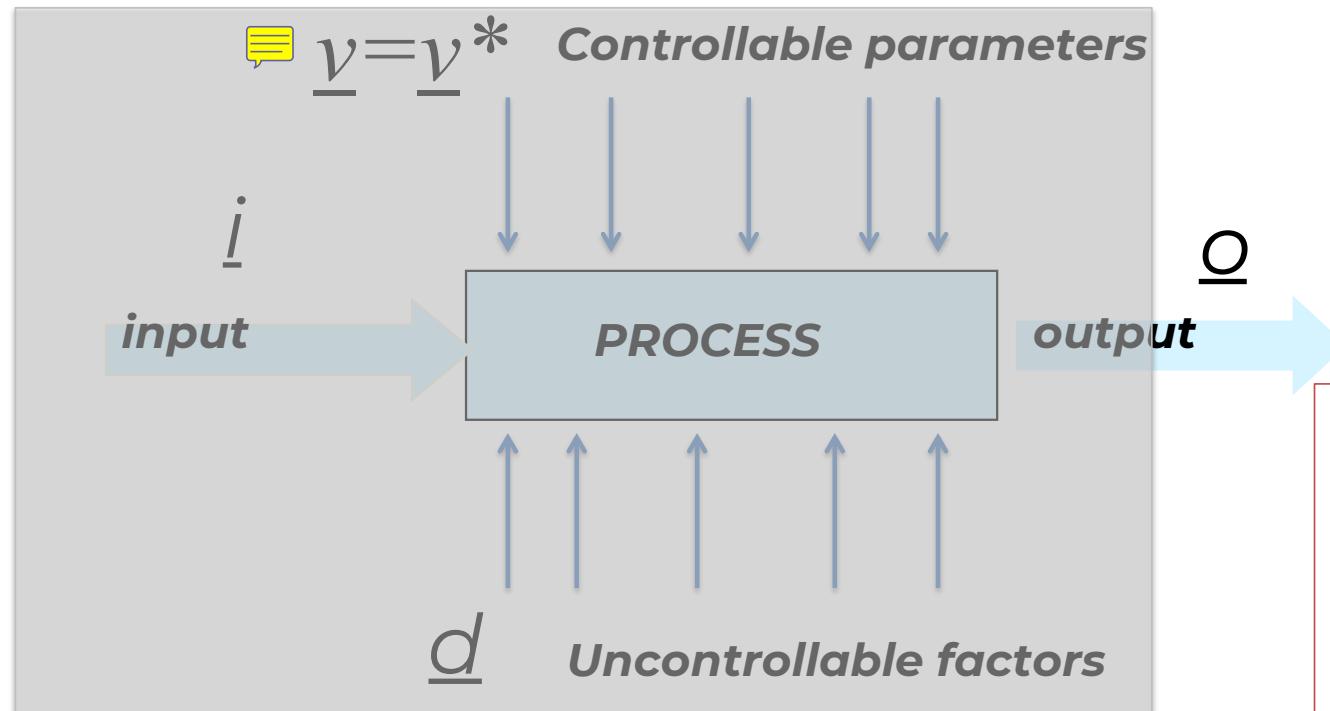


Design of Experiment – to decide the right set of experiments and analyse results to check the effect of controllable parameters to the quality output

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!



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

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

Let's do an experiment

Stopwatch data:

Take the stopwatch (on your smartphone or watch)

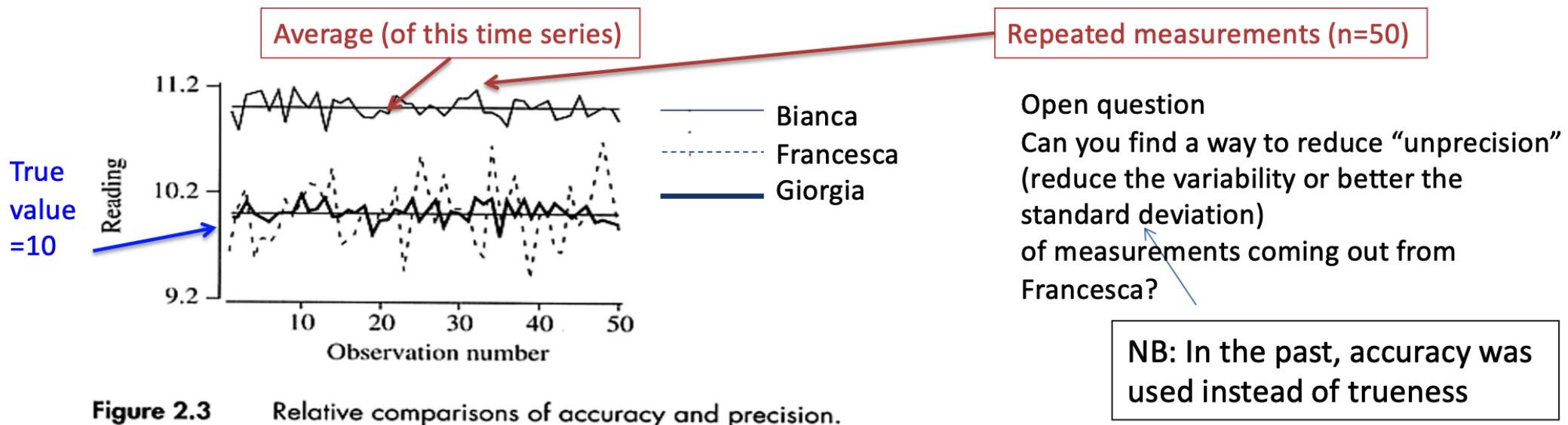
- Try to stop the stopwatch after 10 seconds for 20 times
- List the ten reading values.
- Plot it on a time series plot. Compute the average and the range (max – min)
- What can you observe?



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



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

- precise, but not true
- true, but not precise
- true and precise

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i \rightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{m})$$

TRY TO PROVE $E(\bar{X}) = \bar{Y}(\bar{X})$

1000 m \bar{X}

Vocabulary (VIM)

2.14 measurement trueness

trueness of measurement
trueness

closeness of agreement between the average of an infinite number of replicate measured quantity values and a reference quantity value

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 measured quantity value and a true quantity value of a measurand

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

0.00 can't observe the millisecond

smallest change in a **quantity** being measured that causes a perceptible change in the corresponding indication

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 standard deviation, numerically

closeness of agreement between indications or measured quantity values obtained by replicate **measurements** 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**.

Six sigma

Gauge R&R (Repeatability and Reproducibility)

technique

Repeatability 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.

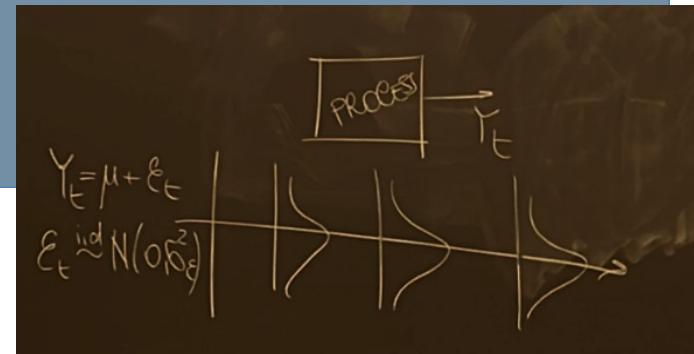


Quality data modeling: main assumptions

ASSUMPTION $(\mu_t = \mu)$ $\rightarrow (\varepsilon_t \stackrel{\text{iid}}{\sim} N(0, \sigma_\varepsilon^2))$

$$Y_t = \mu + \varepsilon_t$$

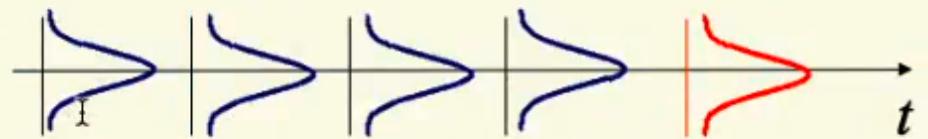
\uparrow \downarrow $\varepsilon_t \stackrel{\text{iid}}{\sim} N(0, \sigma_\varepsilon^2)$
DETERMINISTIC



Process model

Reference model (in basic statistics and traditional Statistical Process Monitoring):

$$Y_t = \mu_t + \varepsilon_t$$



assuming

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

$$-\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

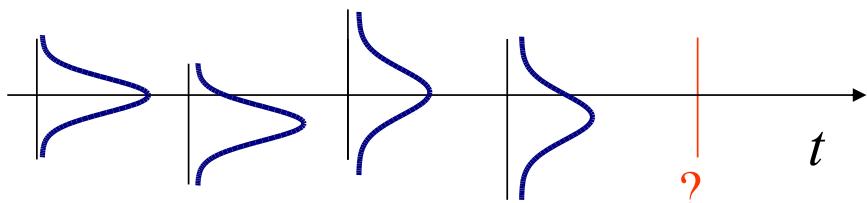
- normally distributed data
(many different causes of variability, none of them dominant)
- independence
- constant variance

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

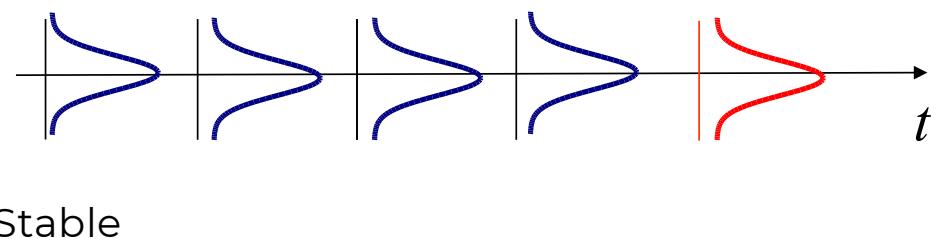
A mathematical model for the observed data



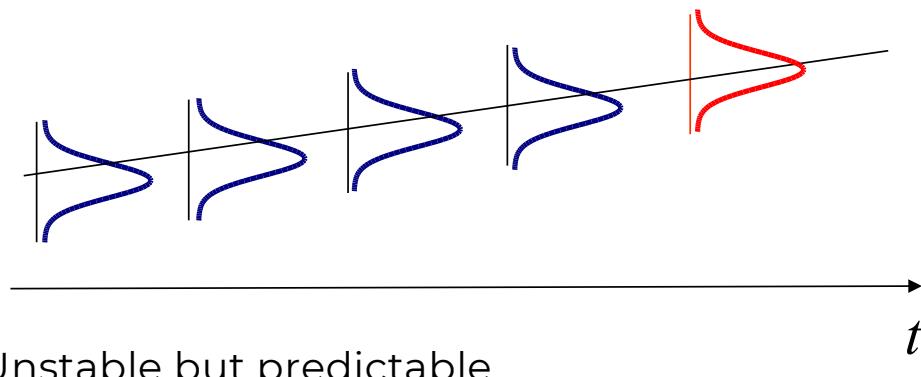
Model of data at time instant t : $Y_t \sim N(\mu_t, \sigma_t^2)$



Unstable



Stable



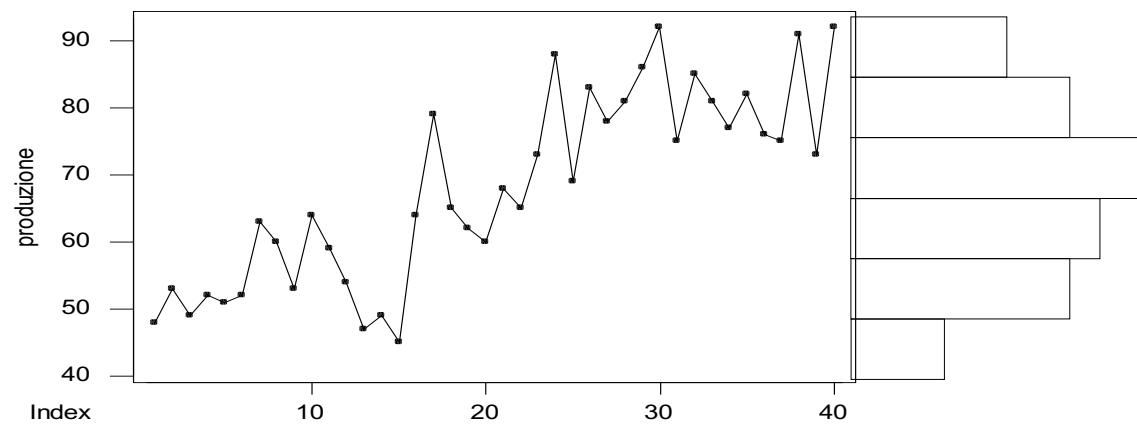
Unstable but predictable

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)	<ul style="list-style-type: none">- Runs test - Bartlett’s test- LBQ’s test	<ul style="list-style-type: none">-gapping-batching-(Linear) regression-Time series (ARIMA)
Normal distribution	Normality test	Transform data

Run chart (or time series plot) 🗣

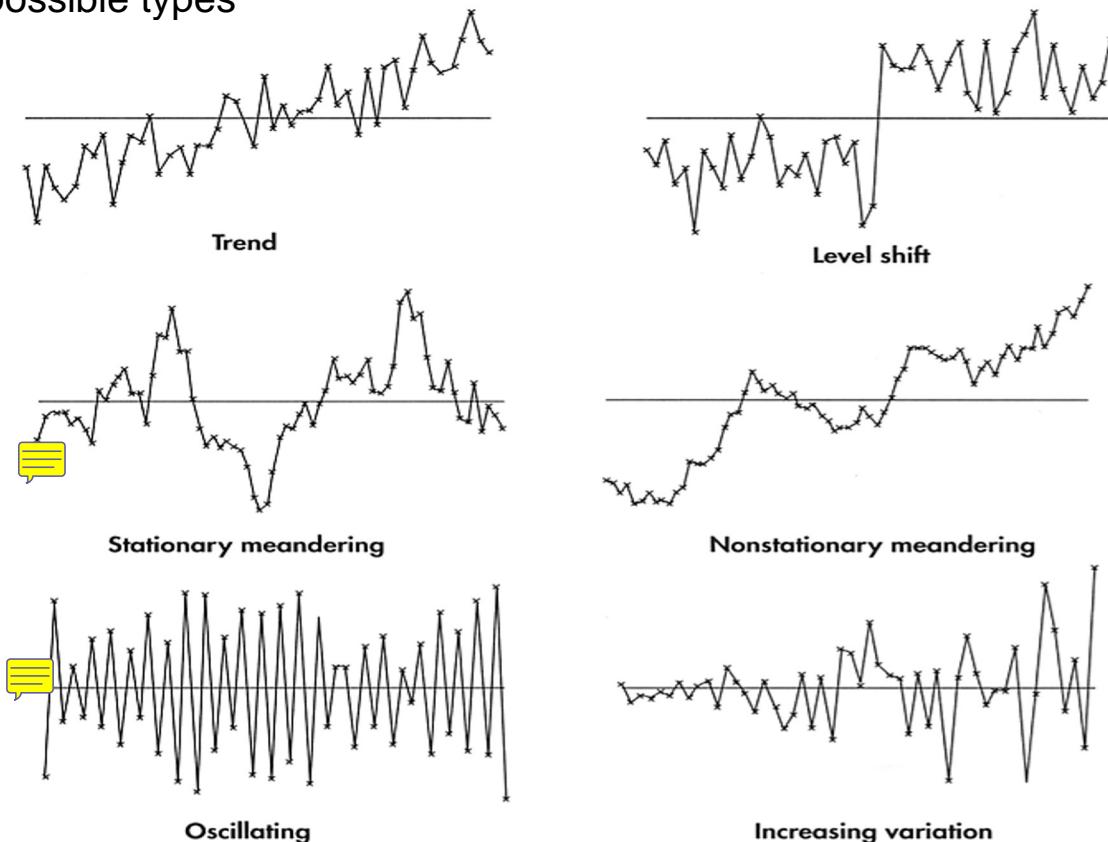


Using the histogram, we are losing relevant information about time

Runs test allow us to detect non random process
and it was not random, it means that there is a pattern in it so we'll go for autocorrelation

“Non random” processes

possible types



- 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

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

autocorrelated data => correlated with themselves, no independent data.
the data are similar but they are changing when you go in time

Autocorrelation refers to the degree of correlation of the same variables between two successive time intervals.

Independence



Does the process have a non random pattern?

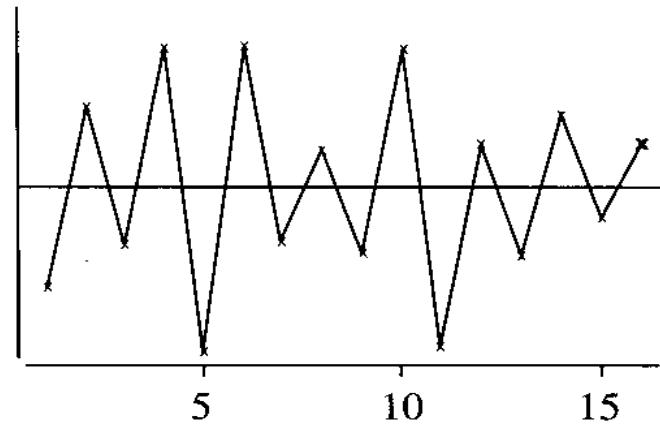
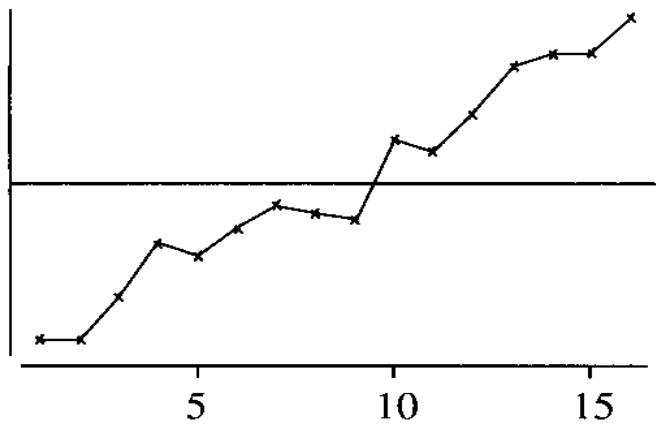
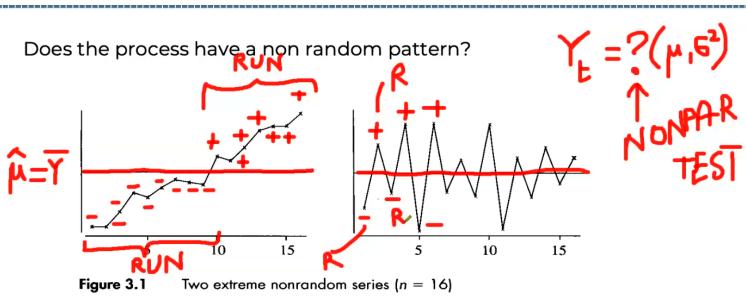
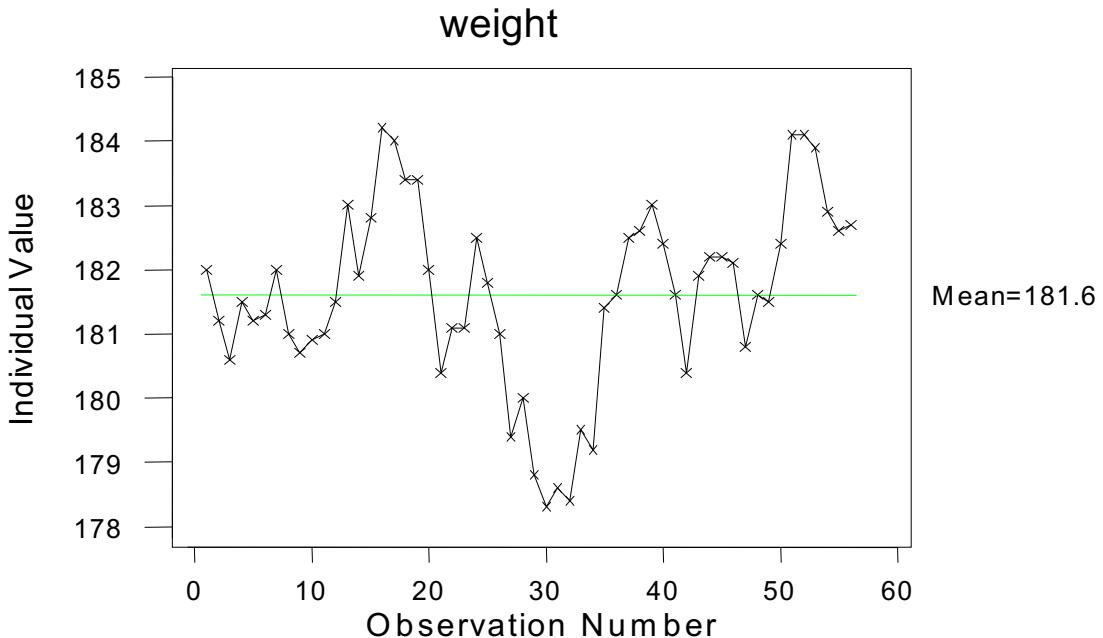


Figure 3.1 Two extreme nonrandom series ($n = 16$)



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 +

$$9 - e^7 +)$$

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

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

Too many or too few runs reflect nonrandom pattern

H_0 : data random

Y = number of runs

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

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

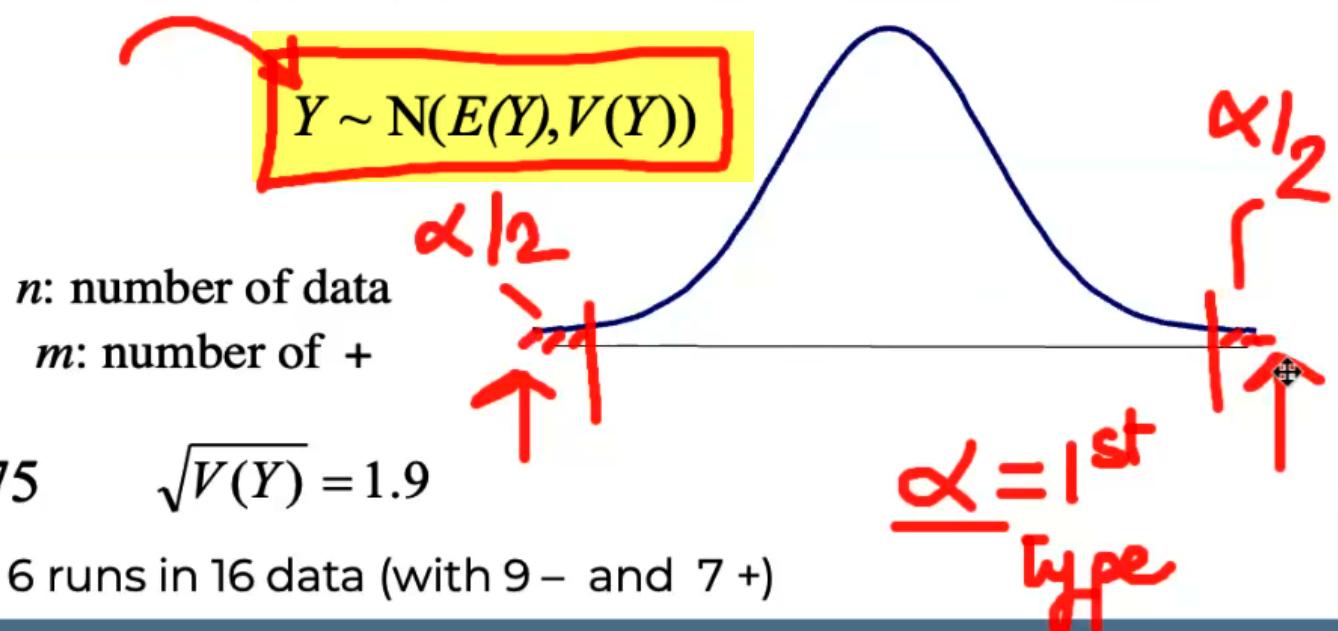
Pbm: when we can say that the number of observed runs (6 in the example) is “close enough” to 8.875?

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 :

$$Y = \text{number of runs}$$
$$E(Y) = \frac{2m(n-m)}{n} + 1$$
$$\sqrt{V(Y)} = \sqrt{\frac{2m(n-m)[2m(n-m)-n]}{n^2(n-1)}}$$

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	I type error (α)	correct decision

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

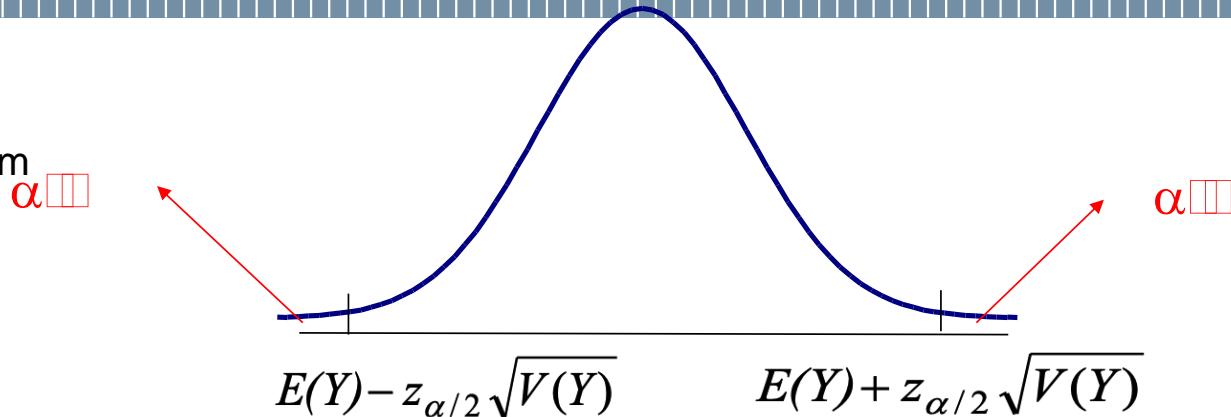
$$\beta = P[\text{second type error}] = P[\text{"accept"} H_0 \mid H_1 \text{ is true}]$$

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

Runs test

$$Y \sim N(E(Y), V(Y))$$

we assume that the Y is random



Once α is set, if H_0 is true

with $\alpha=5\%$

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

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

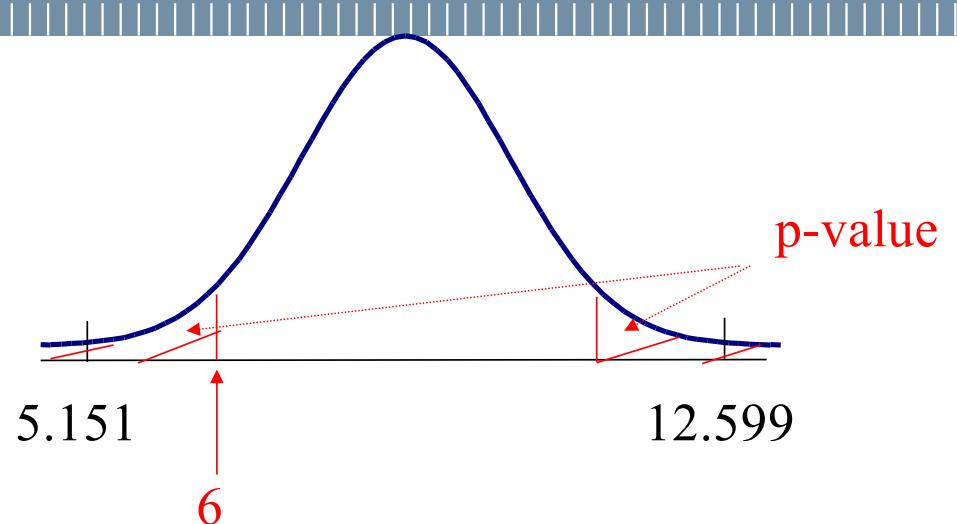
Runs test

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> α 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 α that let us reject the null hypothesis.

Example:
stopwatch
(timing.dat)

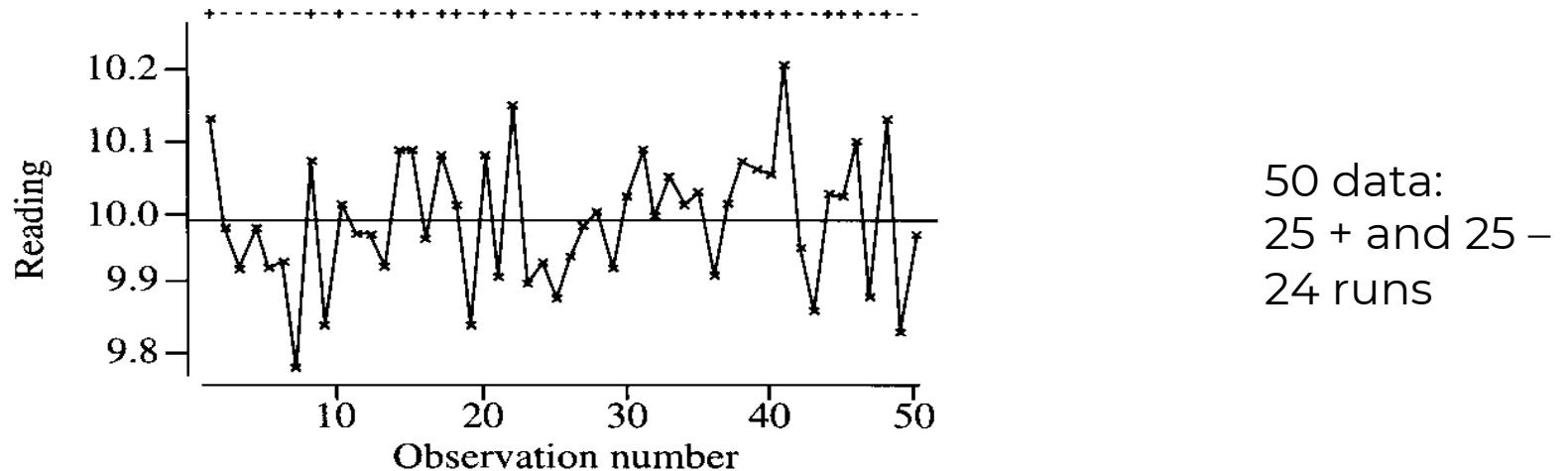
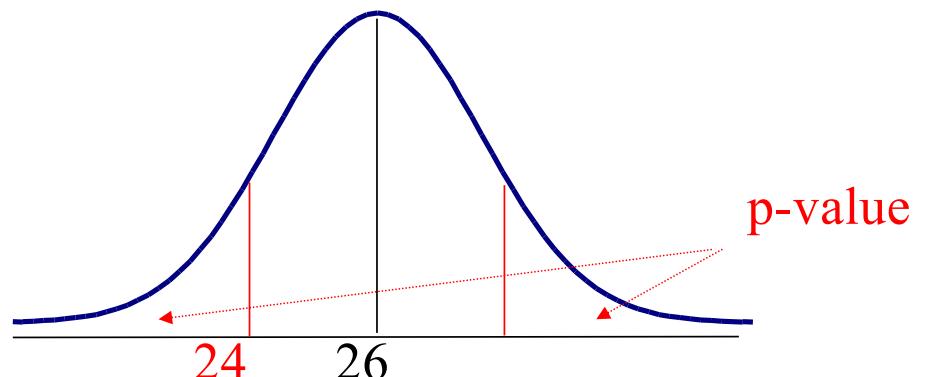


Figure 3.3 Counting number of runs in stopwatch data series.

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

$$p\text{-value} = 2(0.2838) = 0.5676$$

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 sw

```
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.

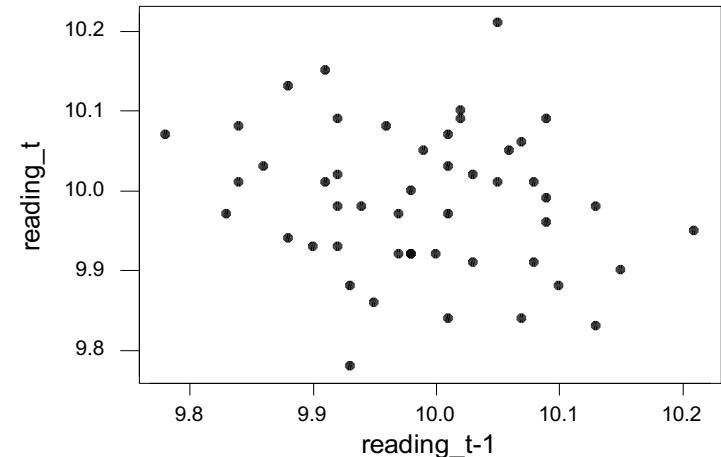
How can we decide whether the standard model is appropriate?

Assumptions	Hypothesis test (to check the assumption)	Remedy in case of violation
“independence” (random pattern)	<ul style="list-style-type: none">- Runs test- Bartlett’s test- LBQ’s test	<ul style="list-style-type: none">-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

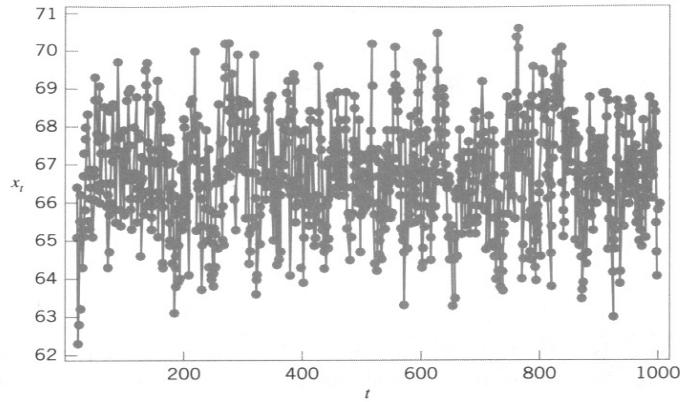


Figure 9-6 A process variable with autocorrelation.

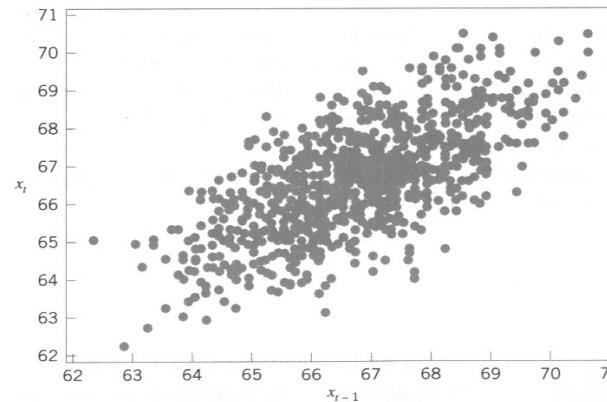
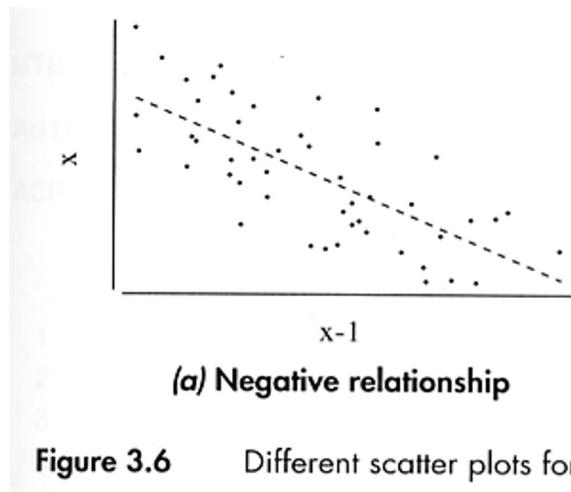
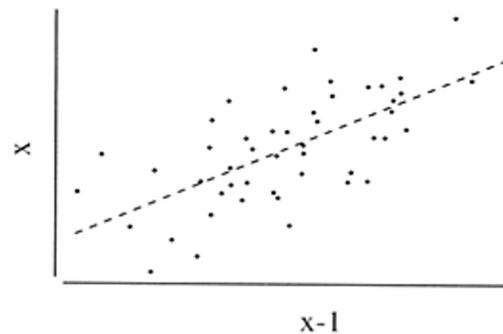


Figure 9-7 Scatter plot of x_t versus x_{t-1} .



(a) Negative relationship



(b) Positive relationship

LINEAR association

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

Autocorrelation

Autocovariance function*:

$$\gamma_{t,k} = \text{Cov}(X_t, X_{t-k}) = E[(X_t - \mu_t)(X_{t-k} - \mu_{t-k})] \quad k = 0, \pm 1, \pm 2, \dots$$

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

k=0: $\gamma_0 = E[(X_t - \mu)^2] = \text{Var}(X_t) = \sigma^2$

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

with

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

$$-1 \leq \rho_k \leq 1$$

$$\gamma_k = \gamma_{-k} \Rightarrow \rho_k = \rho_{-k} \quad \text{Plot of the function only for } k \geq 0$$

* Montgomery

Autocorrelation: Bartlett's test

In order to estimate the autocorrelation function:

Sample autocovariance and autocorrelation functions:

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

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

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

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

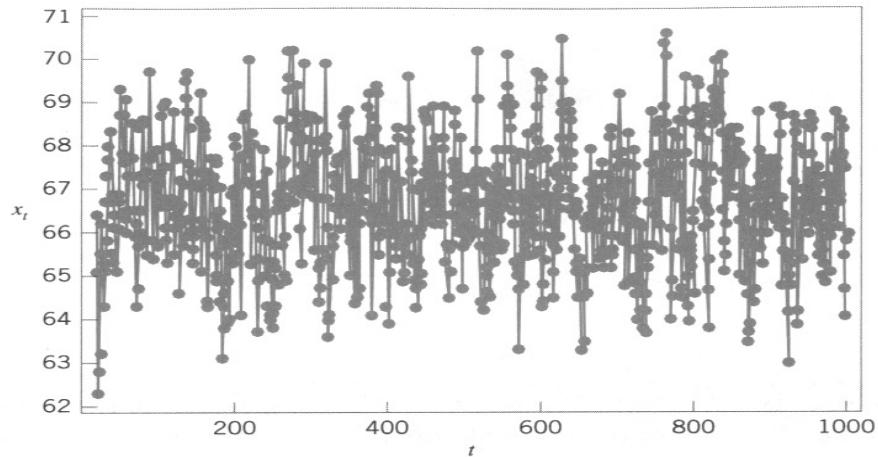


Figure 9-6 A process variable with autocorrelation.

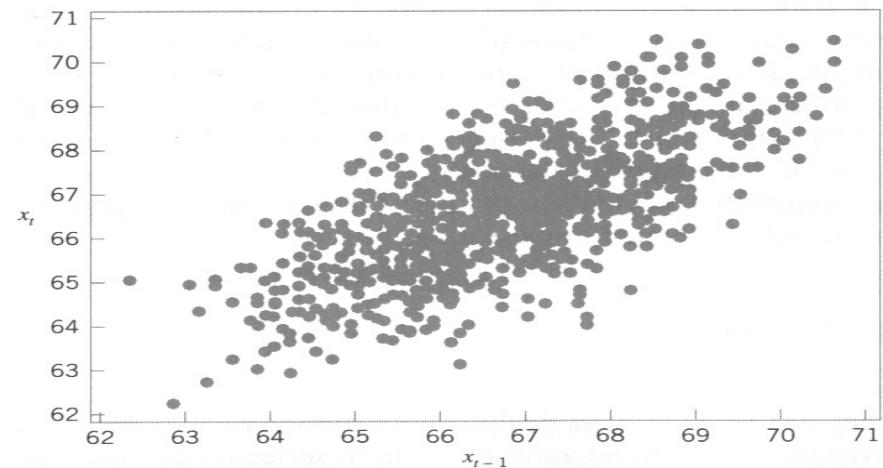


Figure 9-7 Scatter plot of x_t versus x_{t-1} .

$$\rho_k = \frac{\text{Cov}(x_t, x_{t-k})}{V(x_t)} \quad k = 0, 1, \dots$$

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

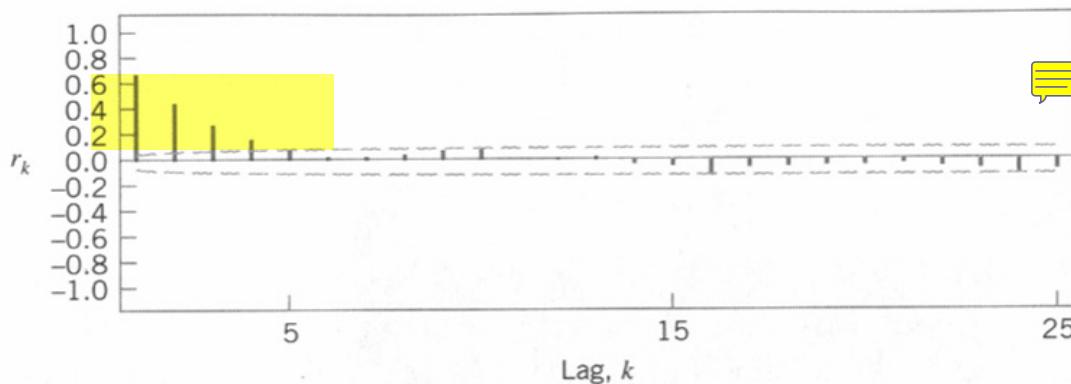
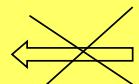


Figure 9-8 Sample autocorrelation function for the data in Fig. 9-6.

Correlation and independence

Furthermore:



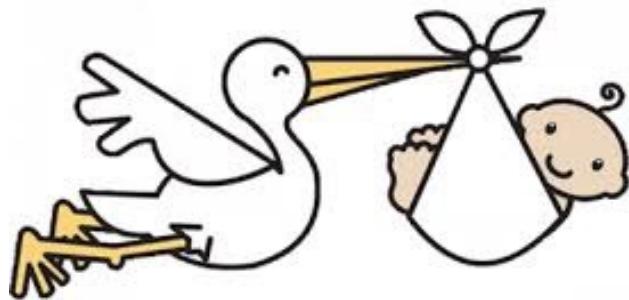
1. If X_t e X_{t+k} are independent, then they are uncorrelated ($\rho_k=0$ for any k)
2. If X_t e X_{t+k} are correlated ($\rho_k \neq 0$ for some k 's), then X_t and X_{t+k} are dependent

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

Correlation and causality

Do you find any example of correlation which does not imply causality?

Correlation and causality



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

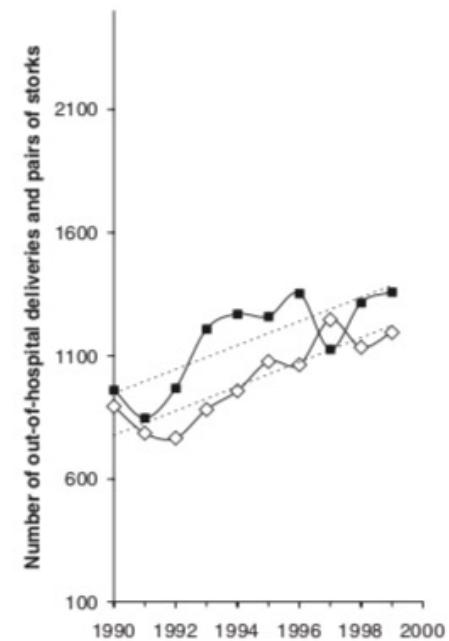


Figure 2. Storks in Brandenburg and the birthrates in Berlin, Germany (1990–99). Open triangles show number of clinical deliveries per year in Berlin. Open diamonds show number of out-of-hospital deliveries per year in Berlin. Number of pairs of storks are shown as full squares. Dotted lines represent 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.

<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 $\text{Var}(r_k)=1/n$ $\forall k$

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

It is based on the rejection region:

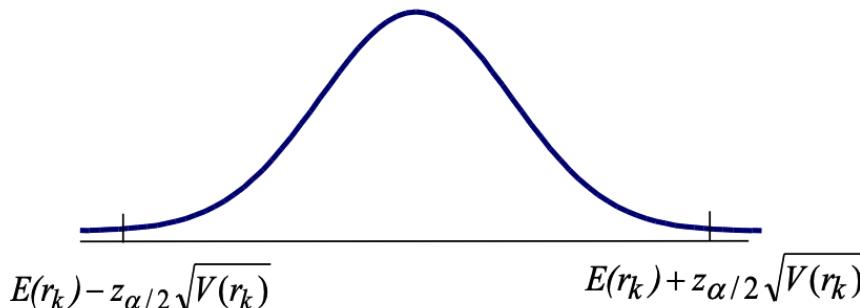
$$|r_k| > \frac{z_{\alpha/2}}{\sqrt{n}}$$

Rejection region:

$$r_k > 0 : r_k > \frac{z_{\alpha/2}}{\sqrt{n}} \Rightarrow |r_k| > \frac{z_{\alpha/2}}{\sqrt{n}}$$
$$r_k < 0 : r_k < -\frac{z_{\alpha/2}}{\sqrt{n}}$$

If H_0 is true:

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



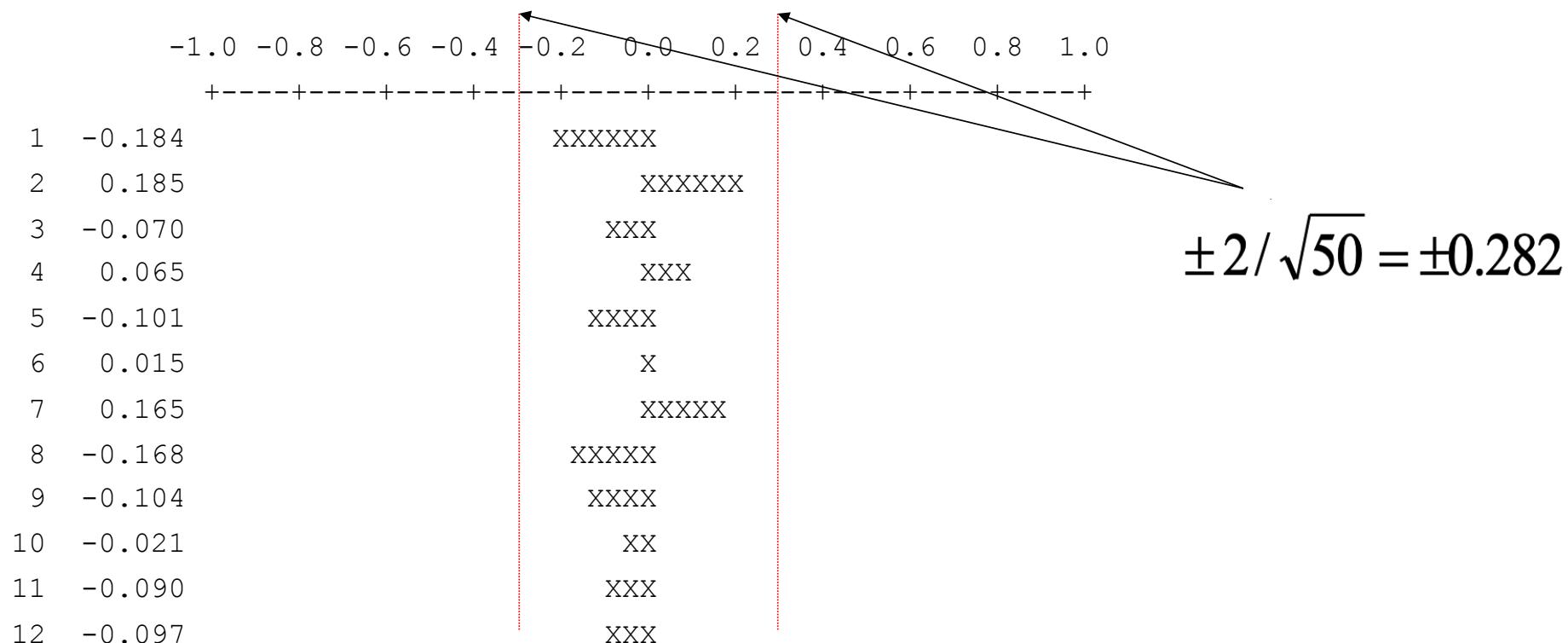
Value that is often used
 $Z_{\alpha/2} = 2 \rightarrow \alpha = 0.0456 \approx 0.05$

ACF (autocorrelation function)

```
MTB > acf c1
```

Autocorrelation Function: C1

ACF of C1



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.



Food for thoughts – check the proof:

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

Carlo Emilio Bonferroni
(1892 – 1960)

Bonferroni inequality

Assume we have N hypothesis tests ($i=1,\dots,N$)

- Each test has its own probability to reject H_{0i} 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**

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

Bonferroni inequality

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

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

If we set the same α for all the tests $\alpha_i = \alpha \quad \forall i = 1, \dots, N \Rightarrow \alpha' \leq 1 - (1 - \alpha)^N$
 $\alpha \leq 1 - (1 - \alpha')^{1/N}$

Bonferroni inequality

We 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

$$\alpha_i \leq \frac{\alpha'_{nom}}{N} \quad \forall i \in \{1, \dots, N\}$$

You'll have an overall first type error

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

False discovery rate

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

Other approaches

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):

$$|r_k| > \frac{z_{\alpha/2}}{\sqrt{n}}$$

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



Rejection region

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

Autocorrelation

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

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

“Portmanteau” test: $H_0 : \rho_i = 0, \quad i = 1, \dots, L$ $H_1 : \exists i \in [1, \dots, L] \ni \rho_i \neq 0$

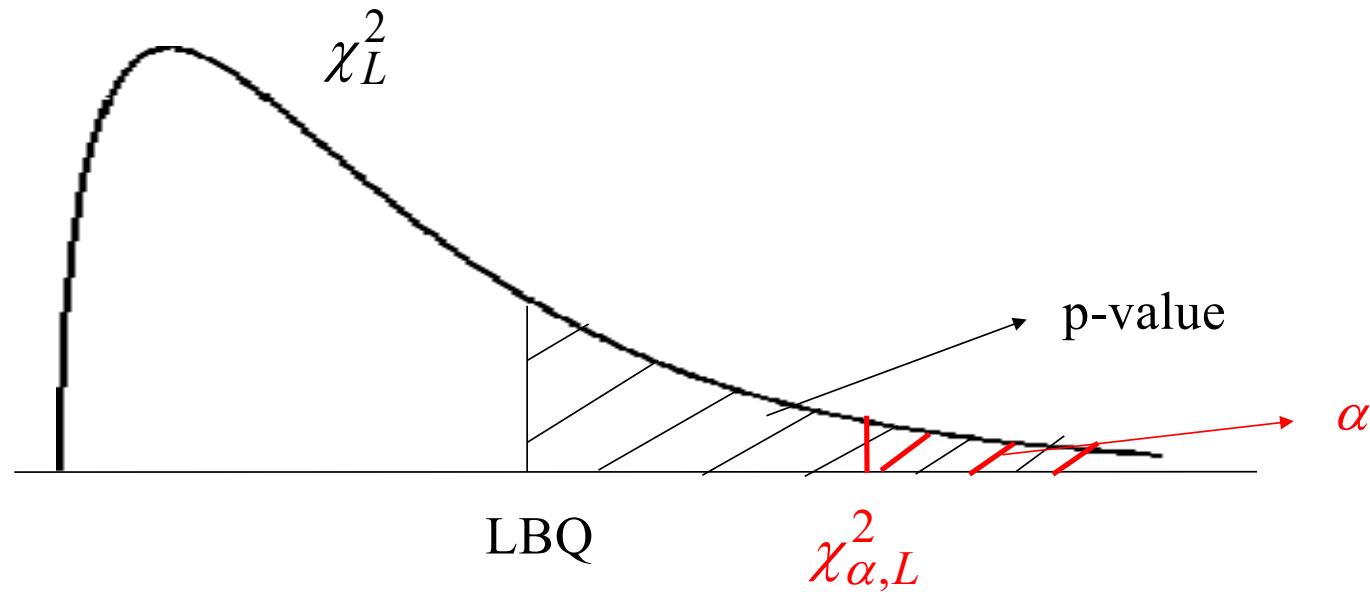
If H_0 true

$$Q \sim \chi_L^2 \Rightarrow \text{rejection region } Q > \chi_{\alpha, L}^2$$

LBQ test

$$Q \sim \chi_L^2 \Rightarrow \text{rejection region } Q > \chi_{\alpha,L}^2$$

If H_0 is true

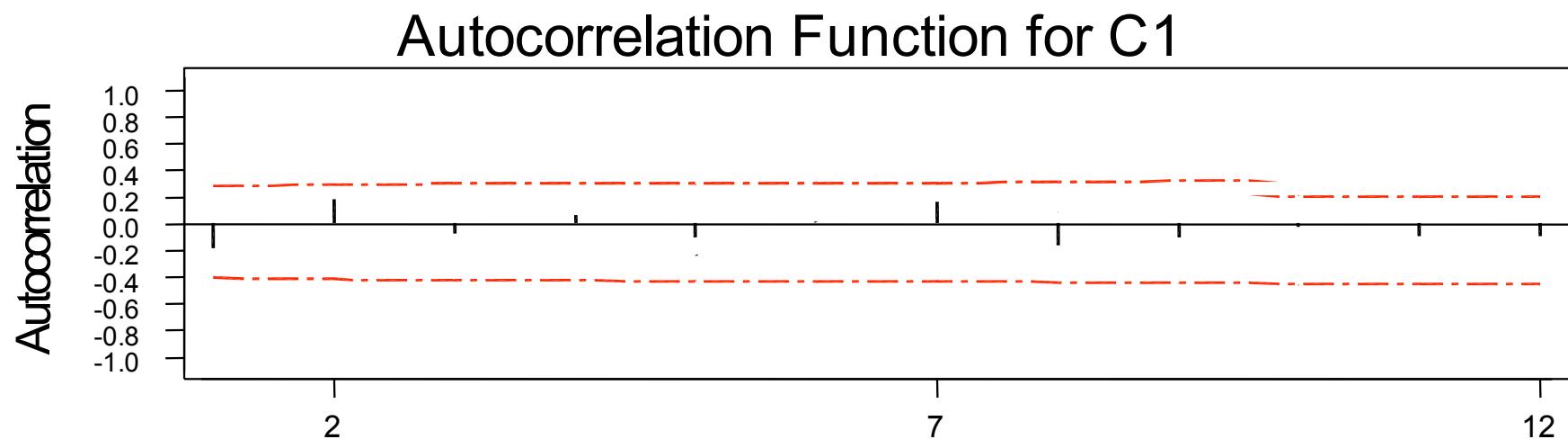


In general, when not specified

$$L \leq \sqrt{n}$$

LBQ test

In the example



$$L \leq \sqrt{50} \rightarrow L \leq 7$$

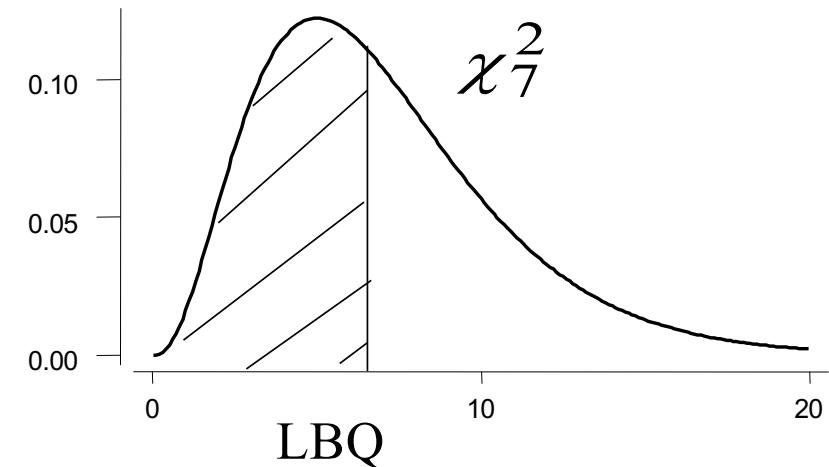
Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.18	-1.30	1.80	8	-0.17	-1.07	8.18
2	0.19	1.27	3.66	9	-0.10	-0.65	8.86
3	-0.07	-0.47	3.93	10	-0.02	-0.13	8.89
4	0.07	0.43	4.17	11	-0.09	-0.56	9.43
5	-0.10	-0.67	4.76	12	-0.10	-0.59	10.07
6	0.02	0.10	4.78				
7	0.17	1.08	6.43				

LBQ test

Cumulative Distribution Function

Chi-Square with 7 DF

x	P(X <= x)
6.4300	0.5095



$$p\text{-value} = 1 - 0.5095 = 0.4905 \approx 50\%$$

\Rightarrow with $\alpha=5\%$ I cannot reject the null hypothesis

Link Bartlett- LBQ

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

For $k=1=L$:
Bartlett:

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

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

LBQ:

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

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

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