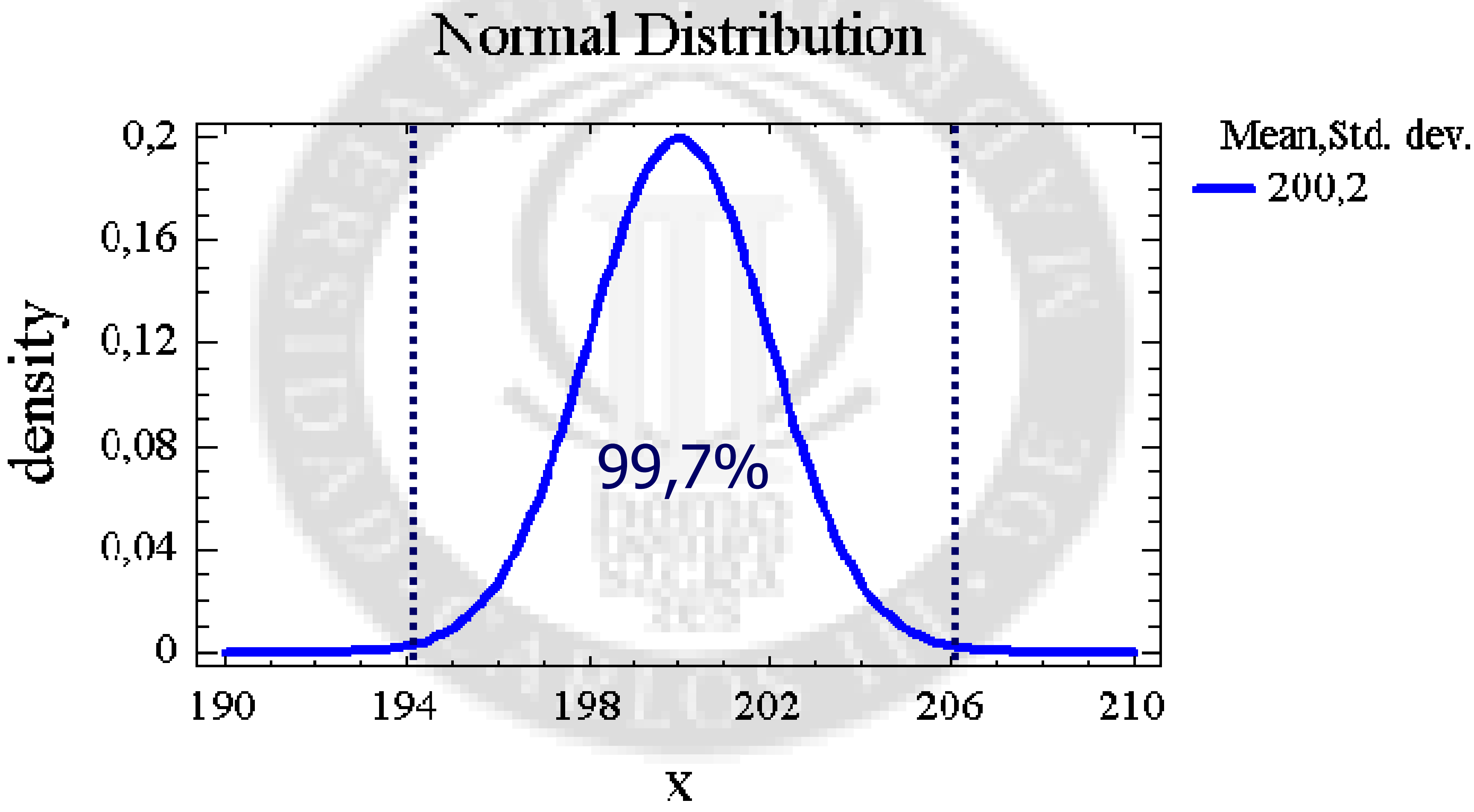
V. Probability models



$$\mu \pm 3\sigma$$

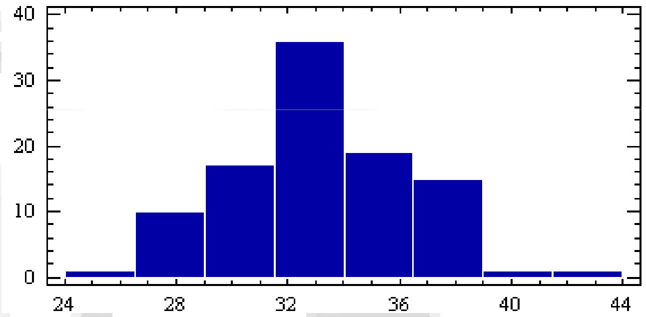
Chapter 5: Probability models

1. Introduction
2. Bernoulli process
3. Poisson process
4. Normal distribution
5. Relationship between Normal, Binomial and Poisson distributions
6. Simple regression model

1. Introduction

Aim: To list a small set of probability models that are useful to describe the random variables that is common to find in real engineering problems

We have real data...

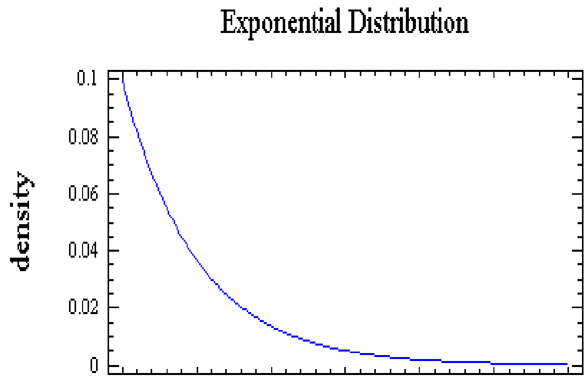


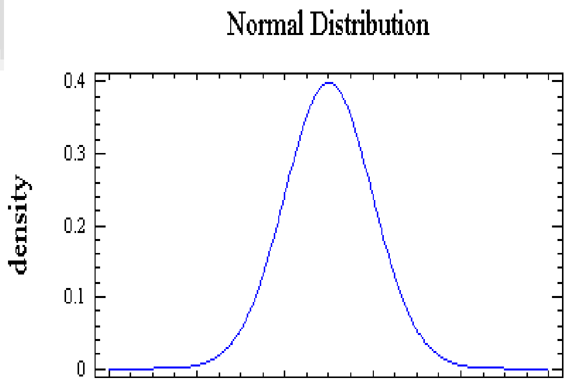
...we choose the model that looks like the data

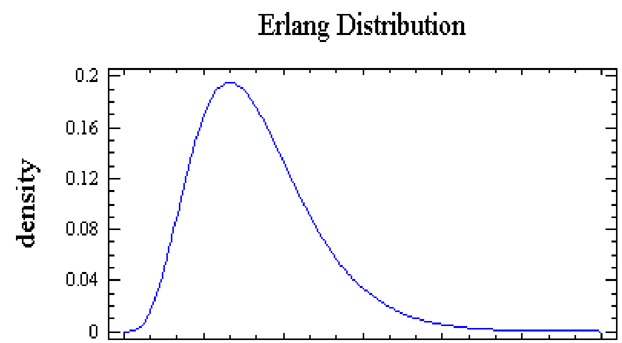
¿f(x)?

¿f(x)?

¿f(x)?







Chapter 5: Probability models

1. Introduction
2. **Bernoulli process**
3. Poisson process
4. Normal distribution
5. Relationship between Normal, Binomial and Poisson distributions
6. Simple regression model

2. Bernoulli process

We assume that our experiment only may give two possible results:

- Success/Failure
- Observe/ Not observe an given attribute

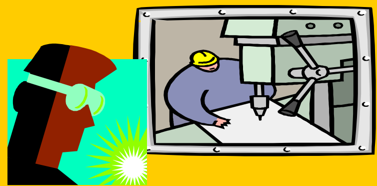
Examples:

- A piece is acceptable/defective
- Satisfied/Unsatisfied customers
- Connection: right or wrong
- Head/Tail
- Vote/Not vote
- Buy/Not buy

2. Bernoulli process

As example we assume to have a productive process which is producing pieces (acceptable or defective)

Productive process



Each piece can be defective or acceptable

We assume a random variable X_i to each piece...

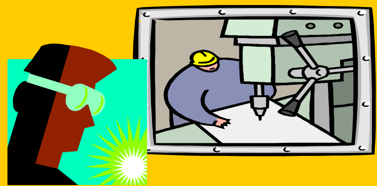
$$X_i = \begin{cases} 1 & \text{if the piece } i \text{ is defective} \\ 0 & \text{if the piece } i \text{ is not defective} \end{cases}$$

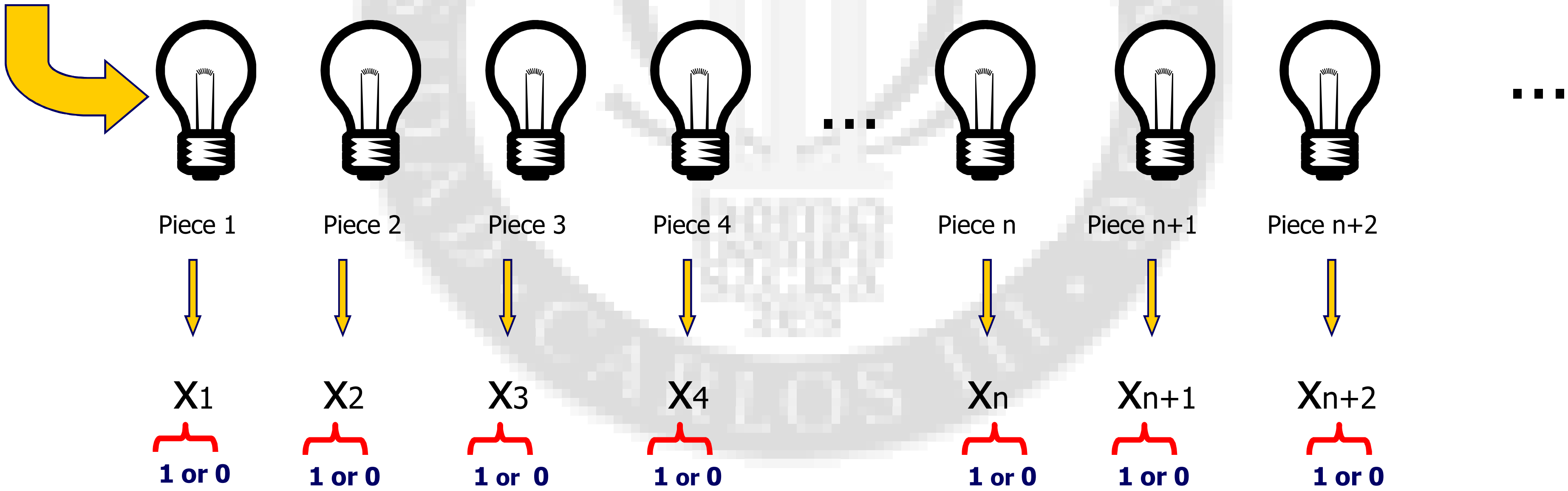
→ **Bernoulli** random variable

2. Bernoulli process

As example we assume to have a productive process which is producing pieces (acceptable or defective)

Productive process



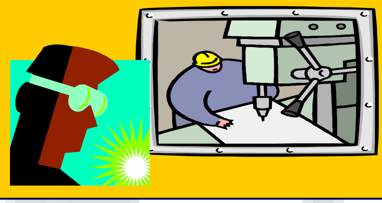


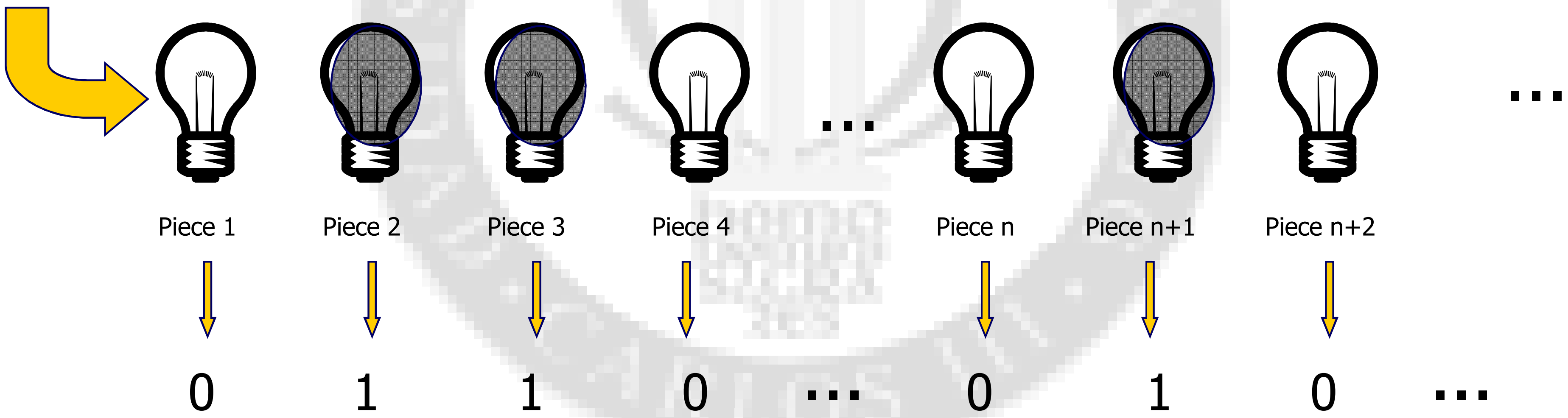
The sequence of these random variables constitutes a: $\longrightarrow$ BERNOULLI PROCESS

$X_1, X_2, X_3, \ldots$

As example we assume to have a productive process which is producing pieces (acceptable or defective)

Productive process





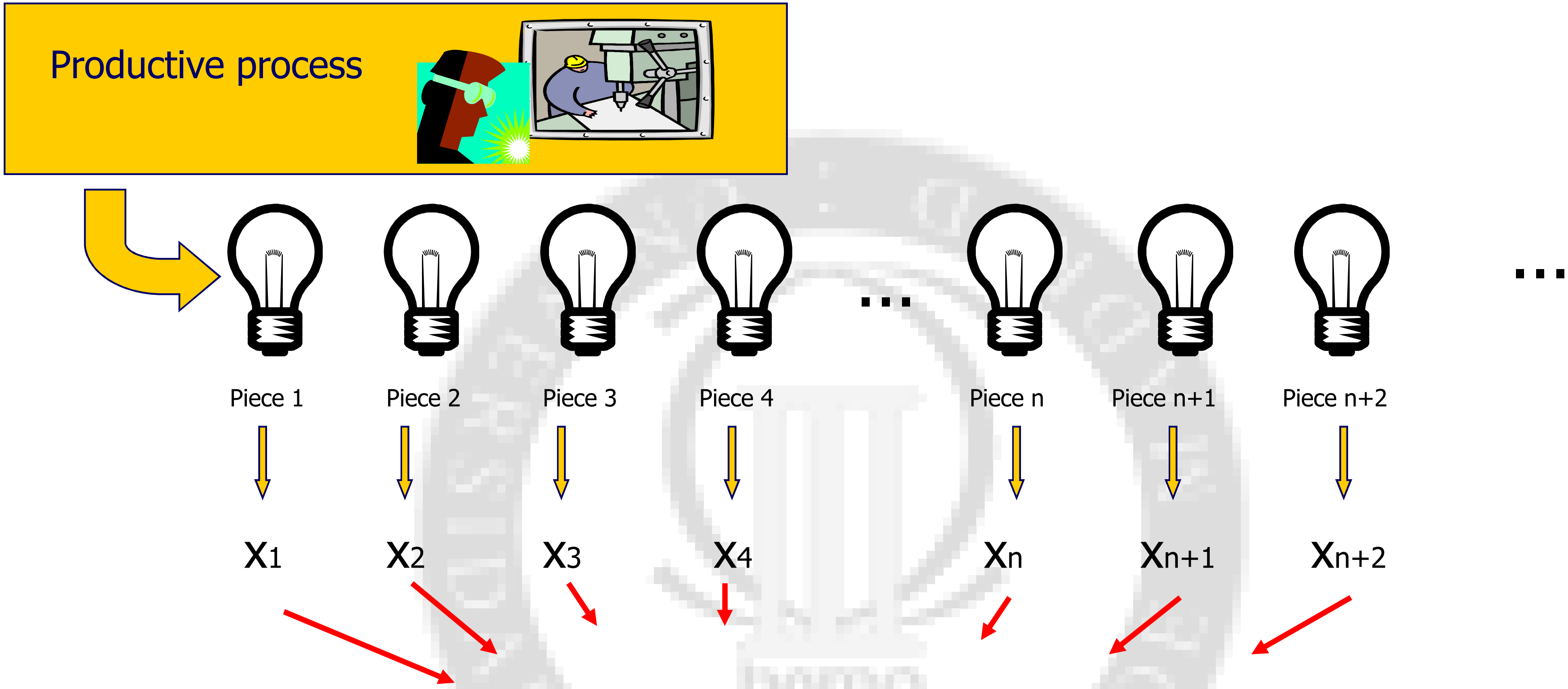
Hypothesis:

Stability: the probability that a piece is defective is always the same, i.e. the system neither get worse nor get better along the time.

Independence: the probability that a piece is defective does not depend on the on the fact that the previous piece was defective or acceptable. A piece turns to be defective because of fortuitous causes, and so independently by the state of the other pieces. THERE IS NO MEMORY.

The probability that a piece is defective is always the same = p

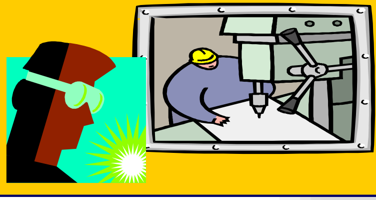
$$P(X_i=1)=p; \quad i=1,2,\ldots$$



All X_i have the same properties

$X_1, X_2, X_3, \ldots$ is a sequence of a identical Bernoulli's

Productive process



Piece 1 … Piece 2 … Piece 3 … Piece 4 … … Piece n … Piece n+1 … Piece n+2 …

$X_1 \quad X_2 \quad X_3 \quad X_4 \quad \ldots \quad X_n \quad X_{n+1} \quad X_{n+2} \quad \ldots$

$$X_i = \begin{cases} 1 \text{ if the piece i is defective .... } P(X_i = 1) = p \\ 0 \text{ if the piece i is not defective ..... } P(X_i = 0) = 1-p \end{cases}$$

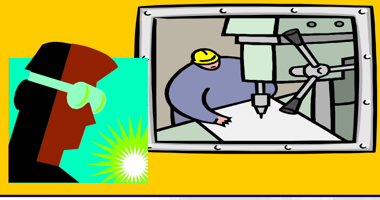
$$\boxed{E(x_i) = p; \qquad Var(x_i) = p(1-p)}$$

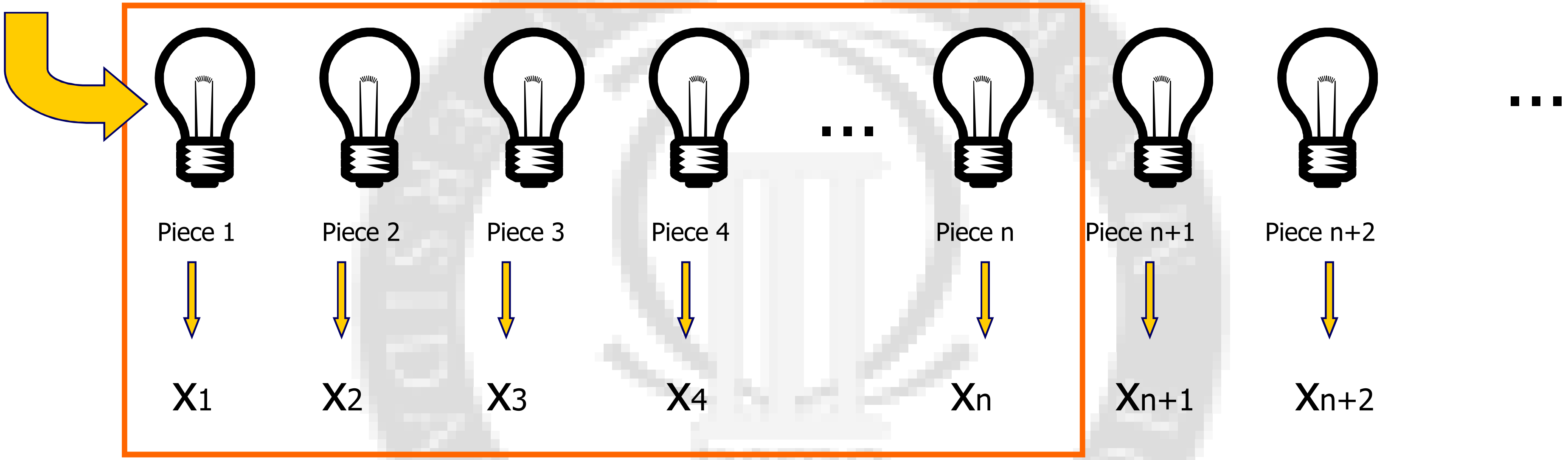
Chapter 5: Probability models

1. **Introduction**
2. **Bernoulli process**
   2.1 Binomial distribution
3. **Poisson process**
4. **Normal distribution**
5. **Relationship between Normal, Binomial and Poisson distributions**
6. **Simple regression model**

2.1 Binomial distribution

Productive process





We take a random sample of n pieces

X= number of defective pieces in a set of n pieces

$X=?\quad X\in\{0,1,2,3,4,\ldots,n\}$

X= BINOMIAL random variable

$X \sim B(n,p)$

X= BINOMIAL random variable

X~B(n,p)

- Sample space: x={0,1,2,...,n}
- Probability function

$$p(r) = P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}; r = 0, 1, \ldots, n;$$

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Example:

We have 4 pieces. The probability of each piece is defective is p

Let X=number of defective pieces:

¿$P(X=3)$?

Option 1: DDDA → $P(\text{Option 1})=P(DDDA)=P(D)P(D)P(D)P(A)$

$=p^3(1-p)^1$

First piece: Defective (D)

Fourth piece: Acceptable (A)

Option 2: DDAD → $P(\text{Option 2})=P(DDAD)=P(D)P(D)P(A)P(D)$

$=p^3(1-p)^1$

Other options: DADD, ADDD each of them with probability $p^3(1-p)^1$

$$P(X=3)=P[(\text{Option 1}) \cup (\text{Option 2}) \cup \ldots \cup (\text{Option 4})]$$

$$= P(\text{Option 1}) + P(\text{Option 2}) + \ldots + P(\text{Option 4})]=4\,p^3(1-p)^1$$

in general...

r with attribute (probability p)

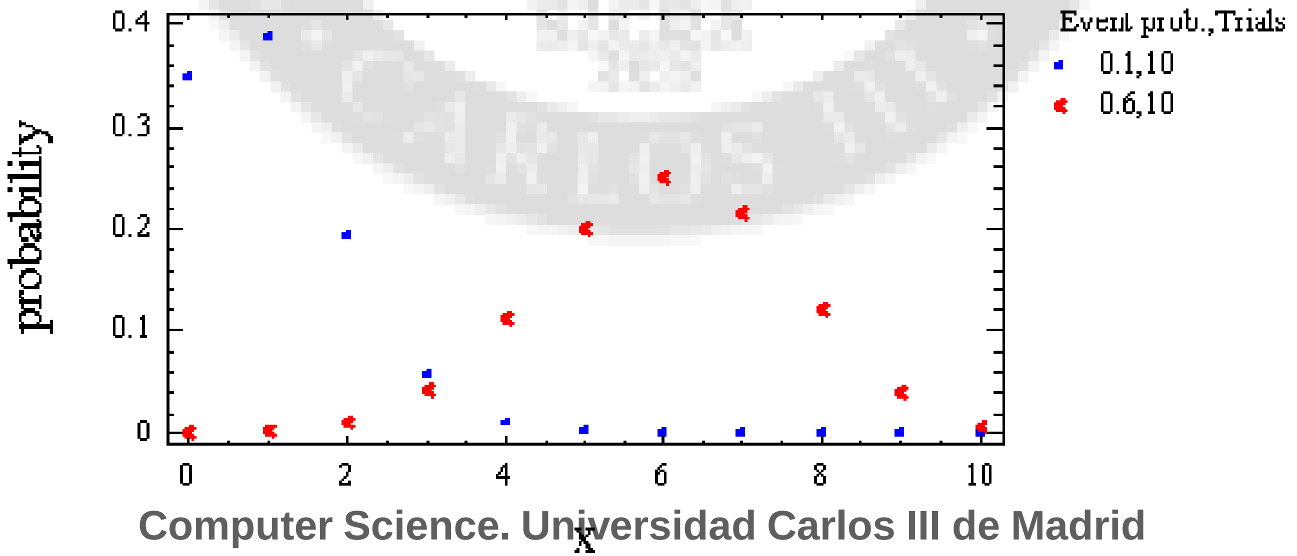
n-r without attribute (probability 1-p)

$$p(r) = P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}; r = 0, 1, \ldots, n;$$

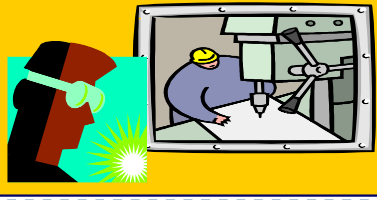
Number of options: combinations with r defectives and (n-r) acceptable

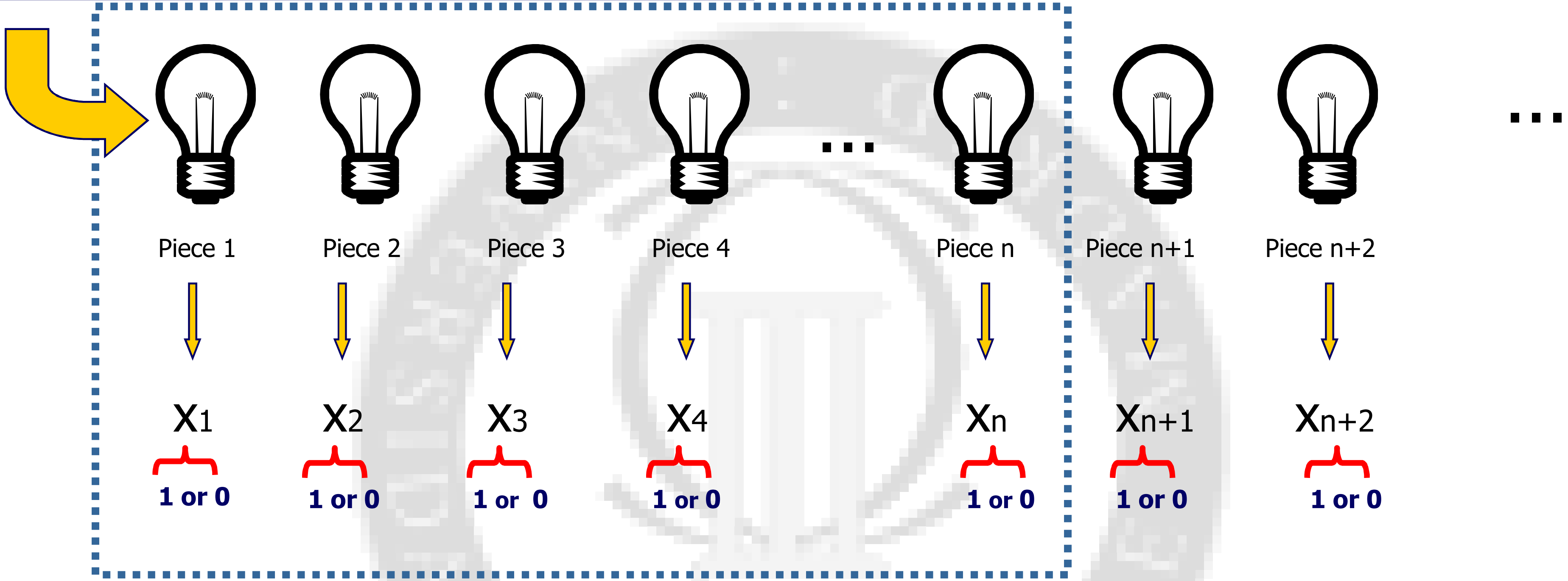
Probability of each of the options

Binomial Distribution



Productive process





X= number of defective pieces in n

Bernoulli's (1 ó 0)

$$X = X_1 + X_2 + \cdots + X_n$$

X= sum of n Bernoulli's

Chapter 5: Probability models

1. Introduction
2. Bernoulli process
3. **Poisson process**
4. Normal distribution
5. Relationship between Normal, Binomial and Poisson distributions
6. Simple regression model

3. Poisson process

Bernoulli process: sequence of elements where we (DO/DO NOT) observe an event

↓

The events are **fortuity** over discrete support

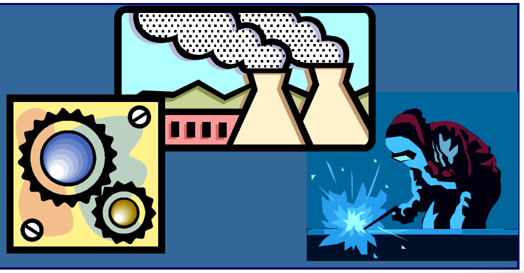
↓

Some elements have the required attribute and others not

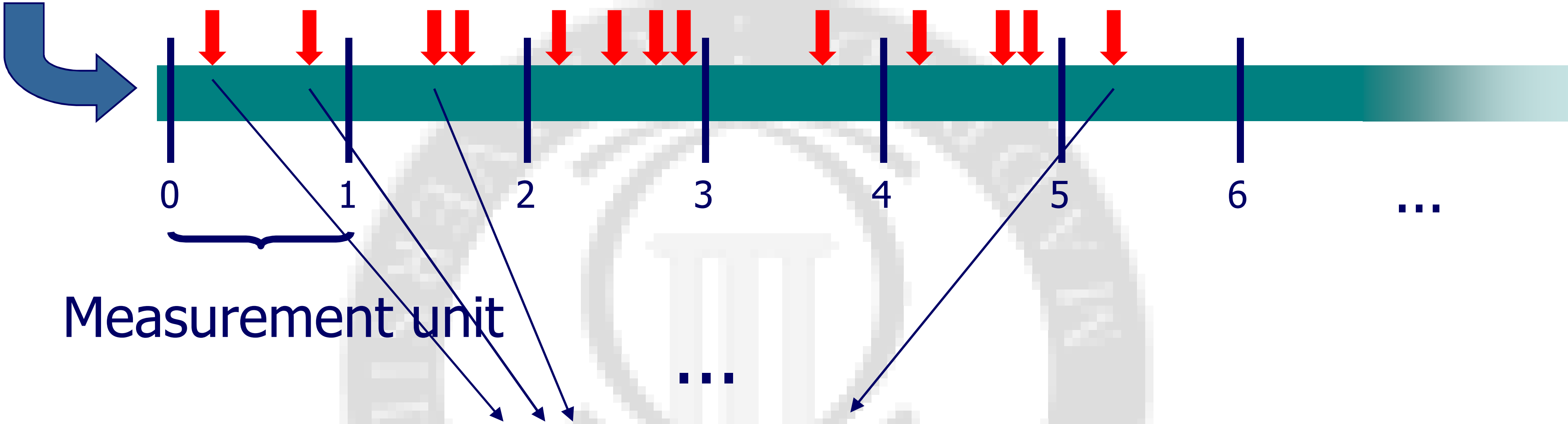
Poisson process: the events are **fortuity**...

over continuous support (length, time, surface, etc)

Productive process



The support over which we observe the events: time, length, surface ...



X=observed events (i.e. defects) in each measurement unit

Poisson process: the events appear independently and with constant frequency over continuous support:

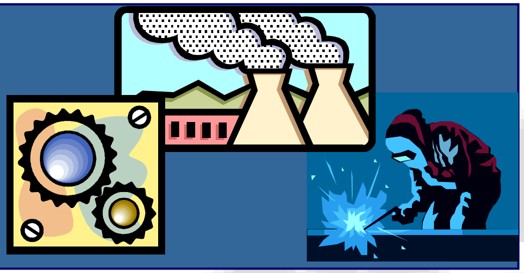
there is no memory

Chapter 5: Probability models

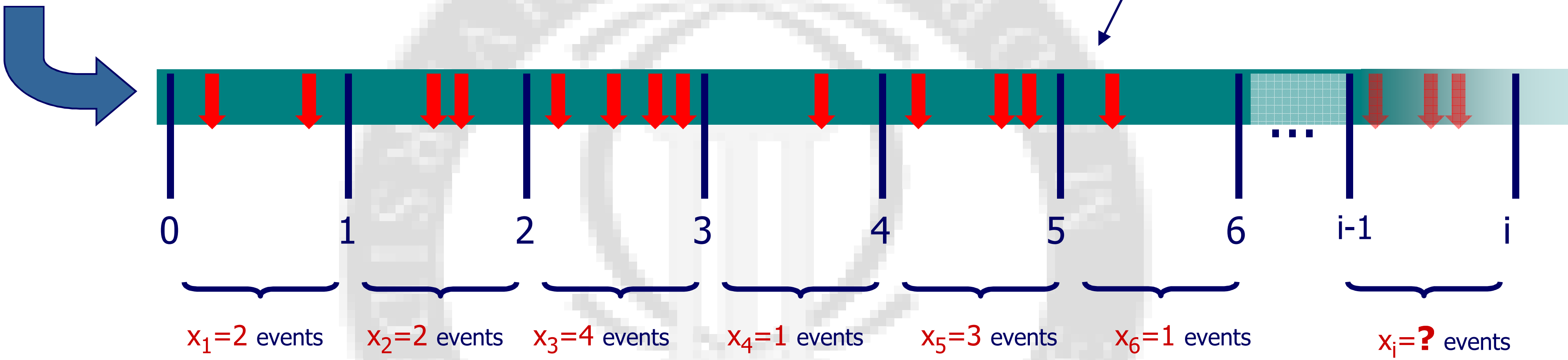
1. **Introduction**
2. **Bernoulli process**
3. **Poisson process**
   - 3.1 Poisson distribution
   - 3.2 Exponential distribution
4. **Normal distribution**
5. **Relationship between Normal, Binomial and Poisson distributions**
6. **Simple regression model**

3.1 Poisson distribution

Productive process



The support over which we observe the events: time, length, surface ...



$X=$ observed events in each measurement unit

$x=\{0,1,2,3,\ldots\}$

X is a Poisson random variable

$X \sim P(\lambda)$

$X \sim Poi(\lambda)$

3.1 Poisson distribution

X is a Poisson random variable

$X \sim P(\lambda)$ $X \sim Poi(\lambda)$

Average number of events per measurement unit

- Probability function:

$$P(X = r) = \frac{\lambda^r}{r!} e^{-\lambda}; r = 0, 1, \ldots$$

- Characteristics moments:

$$E(X) = \lambda,$$

$$\mathrm{Var}(X) = \lambda.$$

- Additivity:

Let X_1 y X_2 be independent Poisson random variables:

$X_1 \sim Poi(\lambda_1)$; $X_2 \sim Poi(\lambda_2)$

Then, $Y = X_1 + X_2$ is Poisson too

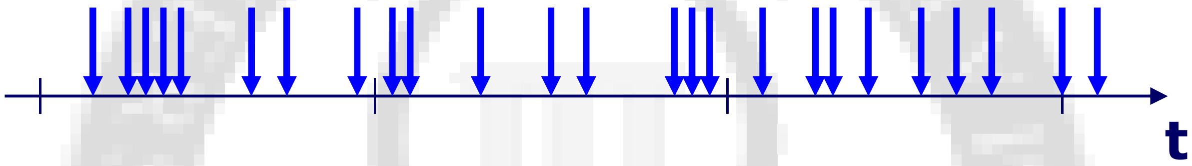
$Y \sim Poi(\lambda_1 + \lambda_2)$

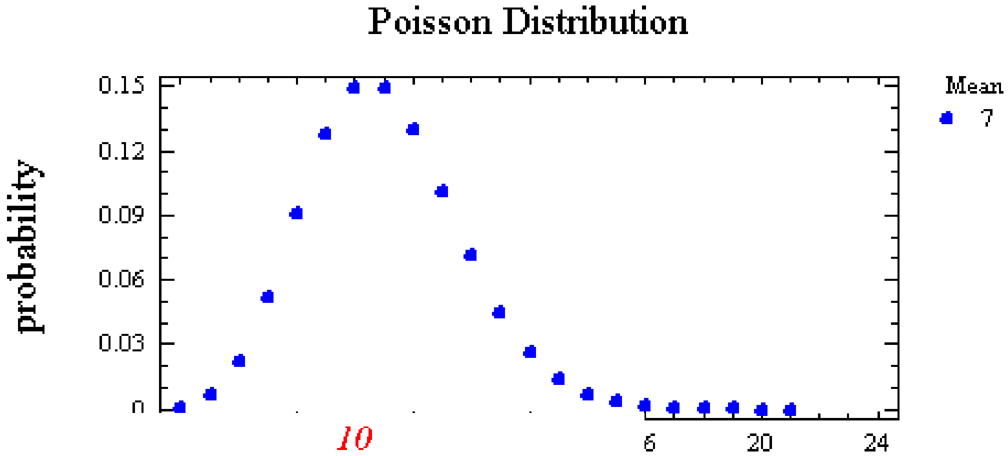
Example:

A network server receives on average 7 accesses per minute. We assuming that the accesses show up independently and at constant rate.

We want to calculate the probability that the server receives more than 10 access in a minute, that is the number of accesses at which the server starts to show a low performance.

X= number of access in a minute; X~Poi(7)





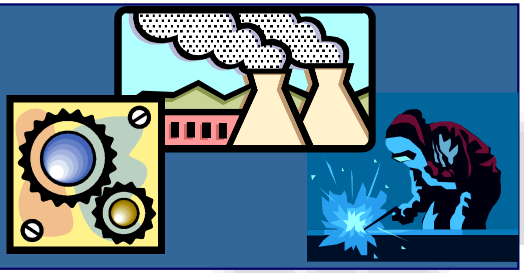
$$P(X=r)=\frac{\lambda^r}{r!}e^{-\lambda}; r=0,1,\ldots$$

$$P(X>10)=1-P(X\leq 10)=1-\sum_{r=0}^{10}P(X=r)$$

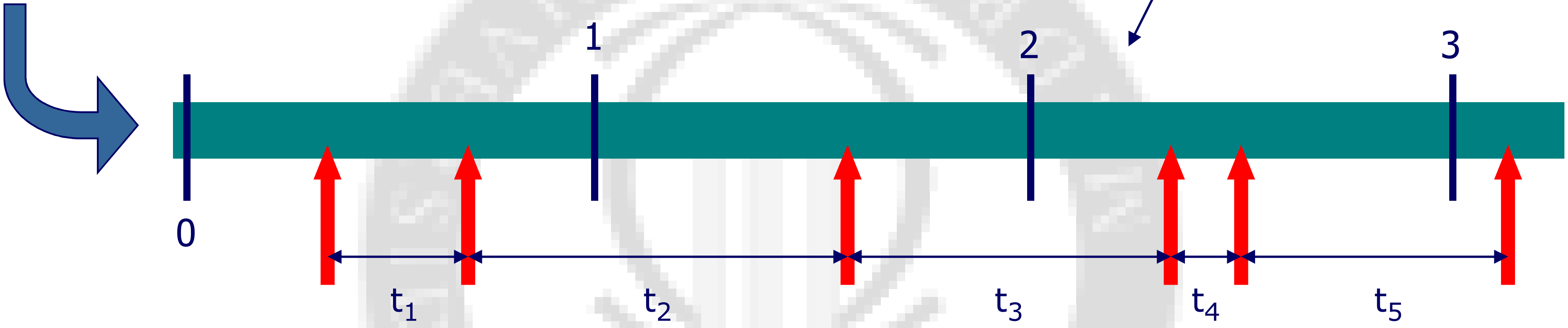
$$=1-\sum_{r=0}^{10}\frac{7^r}{r!}e^{-7}=0.099,$$

3.2 Exponential distribution

Productive process



The support over which we observe the events: time, length, surface ...



- T=interval between two consecutive events of a Poisson process
- Intervals: time units, length, etc
- Continuous random variable. Its values are positive: $t_i>0$

T is named exponential random variable

$T \sim Exp(\lambda)$

The mean of the Poisson process. Average number of events per measurement unit

3.2 Exponential distribution

T=interval between two consecutive events of a Poisson process

T~Exp(λ)

- Density function:

$$f(t) = \lambda e^{-\lambda t}; \lambda, t > 0;$$

- Probability function

$$F(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

$$P(T > t) = e^{-\lambda t}$$

- Characteristics moments

$$E(T) = \frac{1}{\lambda}$$

$$\mathrm{Var}(T) = \frac{1}{\lambda^2}$$

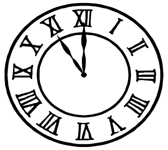
Example:

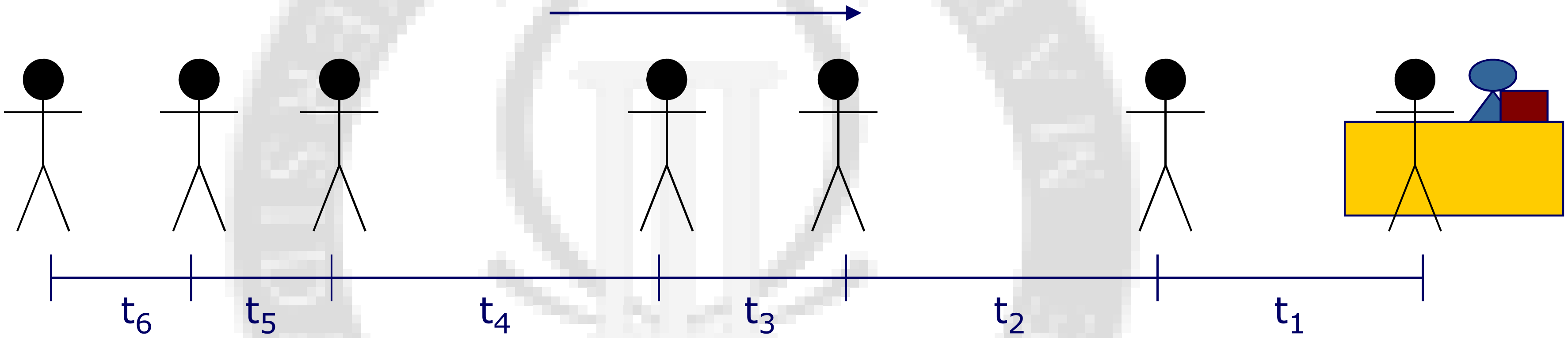
The arrivals of clients at a service point are stable and independent, being the **time between the arrival of two consecutive clients an exponential random variable.**

On average there is only one arrival per minute.

(a) What is the probability that no client arrives for 3 minutes?

(b) What is the probability that the interval of time between two consecutive arrivals is less than a minute?





T=interval (minutes) between two consecutive clients

T~Exp(λ); λ=1 client/minute

(a) $P(T > t) = e^{-\lambda t} \longrightarrow P(T > 3) = e^{-1\times 3} = 0.05$

(b) $P(T \leq 1) = \int_0^1 \lambda e^{-\lambda t} dt = 1 - e^{-\lambda} = 1 - e^{-1} = 0.63$

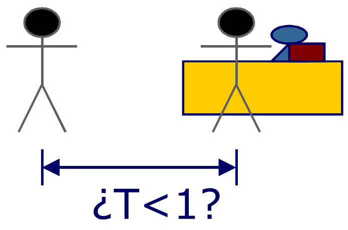
Lack of memory

Poisson process has no memory.

It means that the process does not remember how much time has elapsed from the last observed event

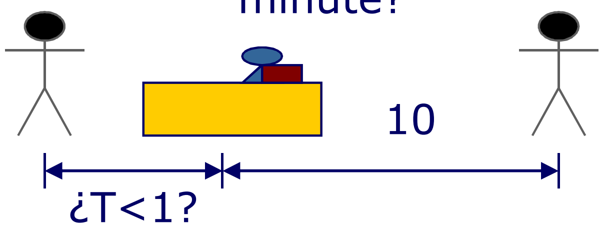
Probability that the next client arrive during the next minute?

$$P(T<1)=\int_0^1 f(t)dt=\int_0^1 \lambda e^{-\lambda t}dt=1-e^{-\lambda\times 1}$$



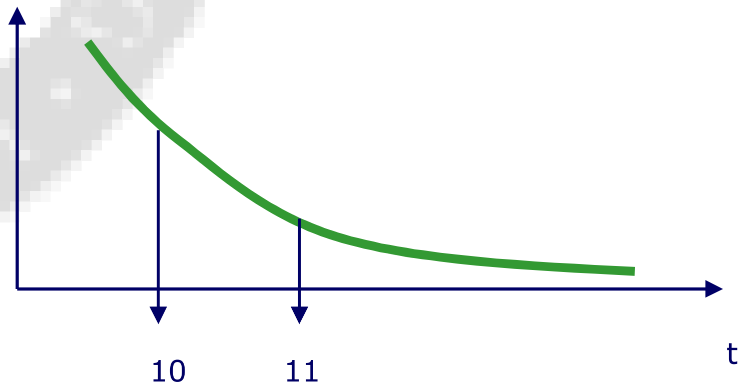
After waiting 10 minutes without receiving clients...

what is the probability that the next client arrives in the next minute?



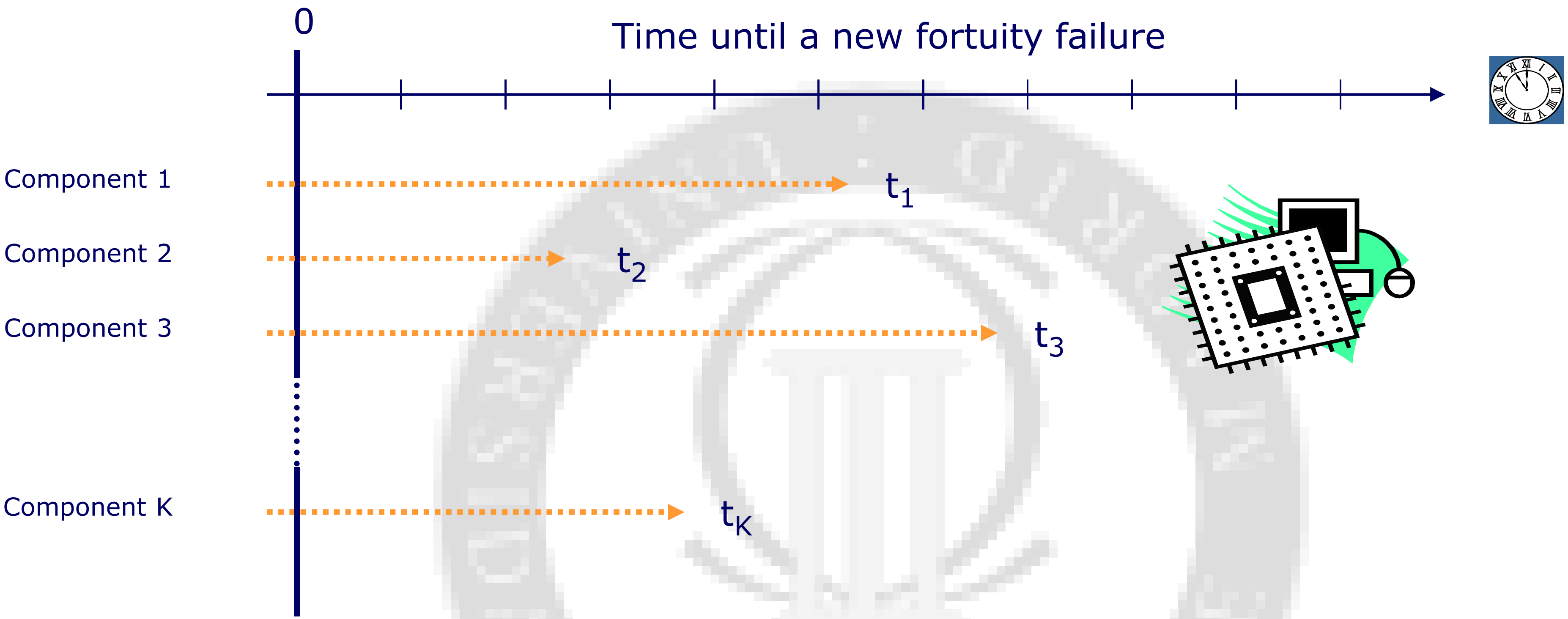
$$P(T<11|T>10)=\frac{P[(T<11)\cap(T>10)]}{P(T>10)}=\frac{P(10<T<11)}{P(T>10)}$$

$$=\frac{P(T<11)-P(T<10)}{P(T>10)}$$



$$=\frac{\left(1-e^{-\lambda\times 11}\right)-\left(1-e^{-\lambda\times 10}\right)}{e^{-\lambda\times 10}}=1-e^{-\lambda\times 1}=P(T<1)$$

The exponential distribution is useful to model the time elapsed from a given instant until the moment a fortuity failure occurs



T=time until a component (a communication channel, a machine, etc) fortuitously fails

$T \sim Exp(\lambda)$

Why fortuitously? Because THERE IS NO MEMORY

There is neither ageing nor improvement,

$E(T)=1/\lambda$ is constant

Example: The duration of a component is an exponential random variable of mean equal to 5000 h.

(a) What is the probability that the component lasts (more than) 6000 hours?

(b) What is the probability that the component lasts (more than) 6000 hours knowing that it was working during 1000 hours?

T= duration: time until FAILURE T~Exp(λ)

E(T)=1/λ=5000 hours $\longrightarrow$ λ=(1/5000) failure/hour

(a) $$P(T>6000)=\int_{6000}^{\infty}\lambda e^{-\lambda t}dt=e^{-\lambda\times 6000}=e^{-\frac{6000}{5000}}=0.30$$

(b) $$P(T>7000|T>1000)=\frac{P[(T>7000)\cap(T>1000)]}{P(T>1000)}$$

$$=\frac{P(T>7000)}{P(T>1000)}=\frac{e^{-\lambda\times 7000}}{e^{-\lambda\times 1000}}=e^{-\lambda\times 6000}=e^{-\frac{6000}{5000}}=0.30$$

There is no ageing. Neither ‘infantile’ failures (by manufacturing defects)

Is this always a hypothesis reasonable?

Chapter 5: Probability models

1. Introduction
2. Bernoulli process
3. Poisson process
4. **Normal distribution**
5. Relationship between Normal, Binomial and Poisson distributions
6. Simple regression model

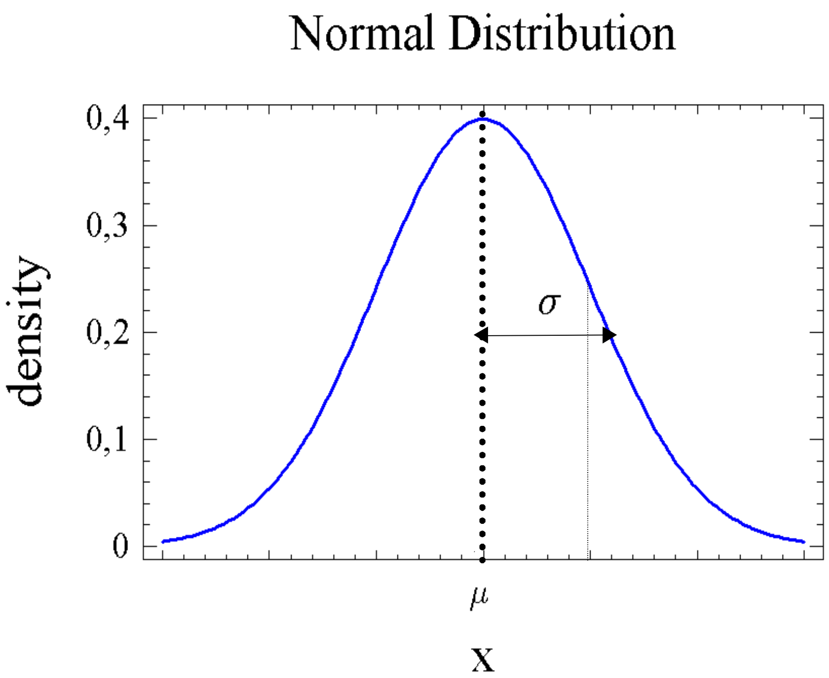
4. Normal random variable

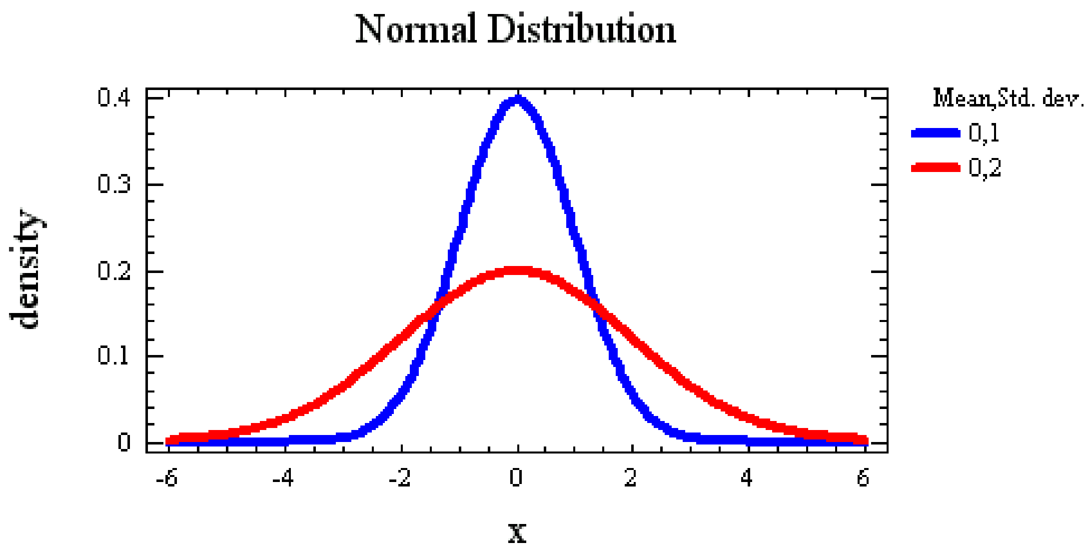
Let X be a continuous random variable with con $E(X)=\mu$ and $var(X)=\sigma^2$

X is a Normal or Gaussian distribution if it has the density function:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty.$$

$$\boxed{X \sim N(\mu, \sigma^2)}$$

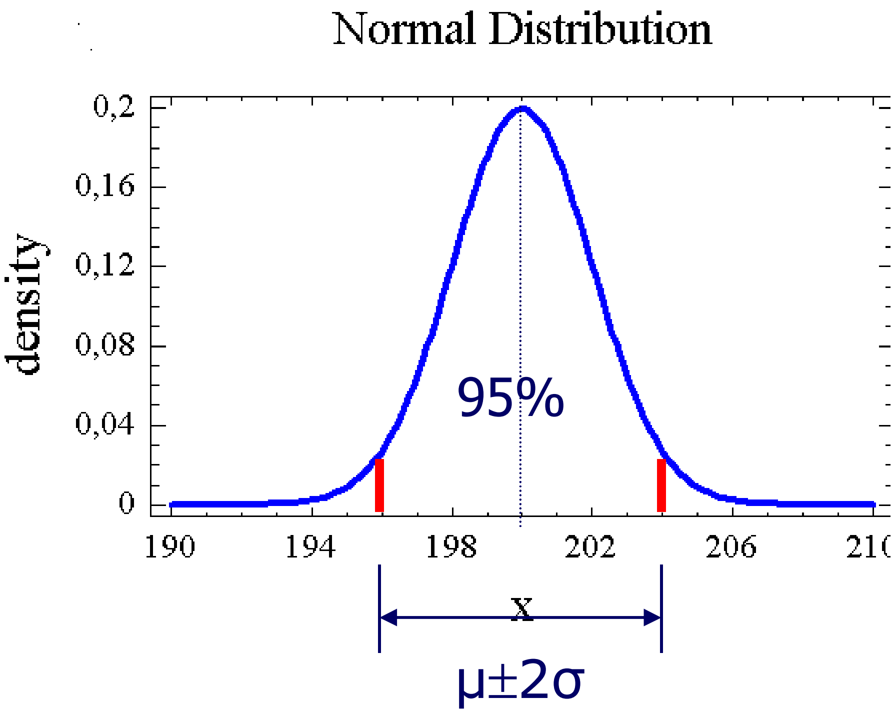


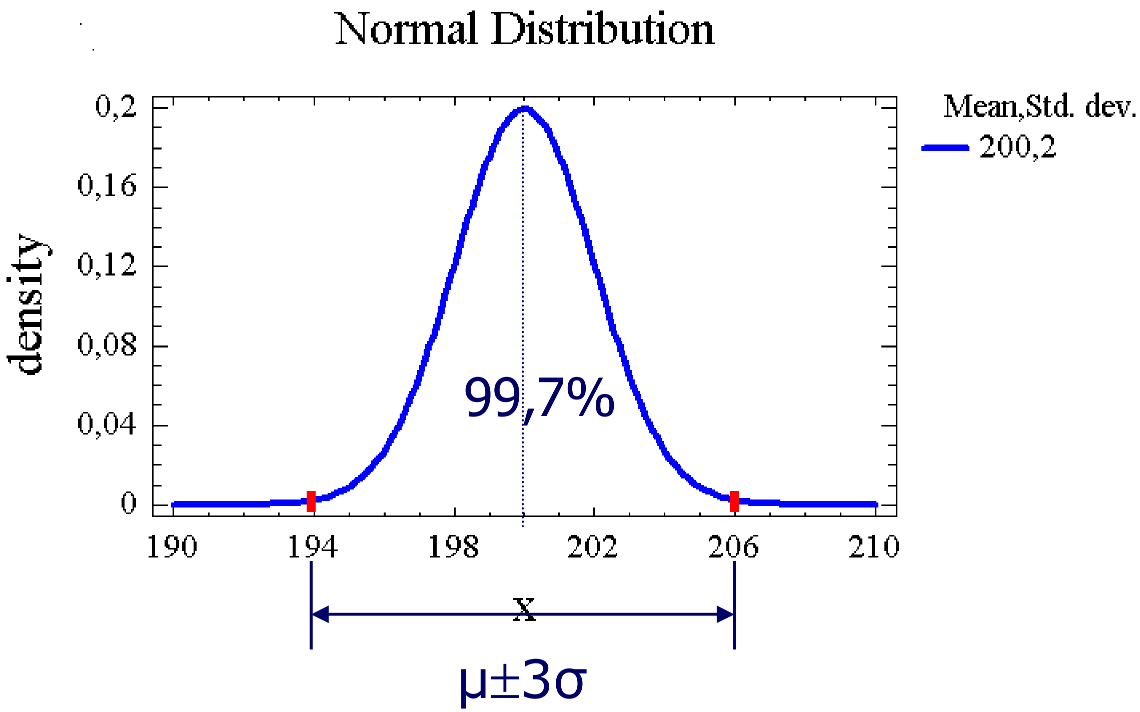


4. Normal random variable

Important properties:

- Symmetric and bell-shaped (‘Gauss’ Bell’)
- Mean and variance are independent characteristics
- It is concentrated around the mean (mode=mean)





4. Normal random variable

Other important properties:

- If we make a linear transformation of a Norman random variable, we obtain a new Normal random variable.

- A linear combinations of Normal random variables is a Normal random variable.

Example:

Let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$, two independent Normal random variables.

Then $Y = aX_1 + bX_2$ is a Normal random variable too.

Moreover we know that

$$E(Y) = E(aX_1 + bX_2) = a\mu_1 + b\mu_2$$

And that

$$Var(Y) = Var(aX_1 + bX_2) = a^2\sigma_1^2 + b^2\sigma_2^2,$$

because the initial random variables are independents.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right), -\infty < x < \infty.$$

It is difficult to integrate

- We can integrate using a computer
- We can integrate using Tables
  - There are only tables for the N(0,1), which is the so-called Standard Normal or Z
  - We can transform a problem in $X \sim N(\mu, \sigma^2)$ into one in $Z \sim N(0,1)$. This transformation is called standardization.

$$X \sim N(\mu, \sigma^2) \longrightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma} E(X-\mu) = \frac{1}{\sigma}[E(X)-\mu] = 0$$

$$\mathrm{Var}(Z) = \mathrm{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\mathrm{Var}(X-\mu) = 1.$$

Then:

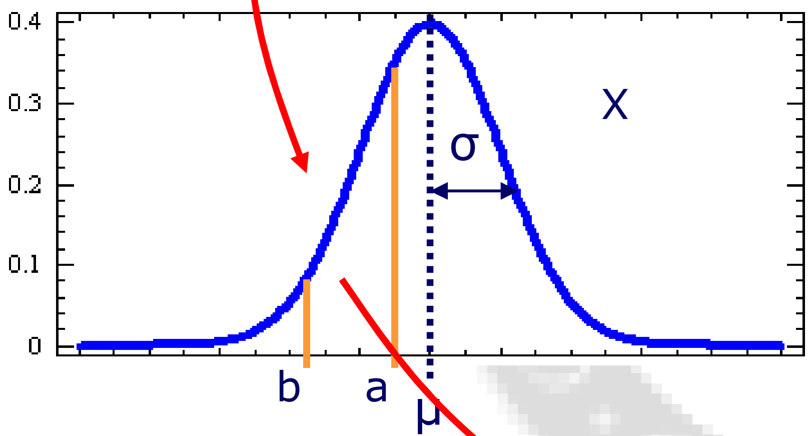
$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$

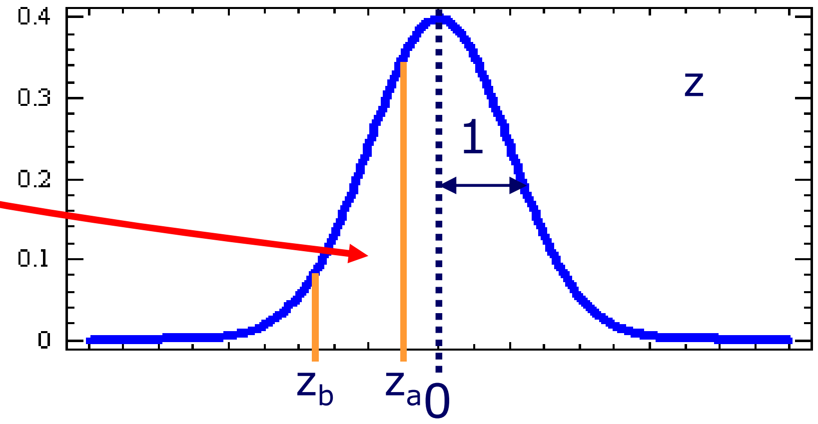
$$= P\left(Z < \frac{b-\mu}{\sigma}\right) - P\left(Z < \frac{a-\mu}{\sigma}\right)$$

$$= P(Z < z_b) - P(Z < z_a)$$

$$= \Phi(z_b) - \Phi(z_a),$$

Distribution function of a Standard Normal: $P(z<z_0)$





<u>The area is the same</u>

Example:

The length L (in millimetres) of the pieces produced in a process is distributed as a Normal random variable $N(32, 0.3^2)$. If we consider acceptable those pieces whose length is between (31.1,32.6), what is the probability that a randomly chosen piece is acceptable?

$$P(31.1 < L < 32.6)$$

$$P\left(\frac{31.1 - 32}{0.3} < Z < \frac{32.6 - 32}{0.3} \right) = 0.976$$

Example:

Some intelligence tests revealed scores whose distribution can be approximated as Normal random variable with mean 100 and standard deviation 15.

(a) Calculate the percentage of population that obtains a score between 95 and 100.

(b) What interval centred in 100 contains up to 50% of the population?

(a) $P(95 \leq X \leq 100) = P\left(\frac{95-100}{15} \leq \frac{X-100}{15} \leq \frac{100-100}{15} \right) = P\left(\frac{-1}{3} \leq Z \leq 0 \right) = 0.13$

(b) $P(0 \leq Z \leq z) = 0.25 \Rightarrow z = 0.675$

$$\longrightarrow \quad \frac{X-100}{15} = 0.675 \Leftrightarrow X \simeq 110 \quad \longrightarrow \quad [90, 110]$$

4. Normal random variable

Central Limit Theorem (CLT)

- Let $X_1,X_2,...,X_n$ be a collection of **any** random variables such as

$$E(X_i) = \mu_i < \infty$$

$$\mathrm{Var}(X_i) = \sigma_i^2 < \infty$$

- Let $Y= X_1+X_2+...+X_n$

Then, if $n\to\infty$

$$\boxed{Y \sim N[E(Y),Var(Y)]}$$

If we sum a lot of random variables, the result is a Normal random variable, although the original random variables are not Normal.

Chapter 5: Probability models

1. Introduction
2. Bernoulli process
3. Poisson process
4. Normal distribution
5. **Relationship between Normal, Binomial and Poisson distributions**
6. Simple regression model

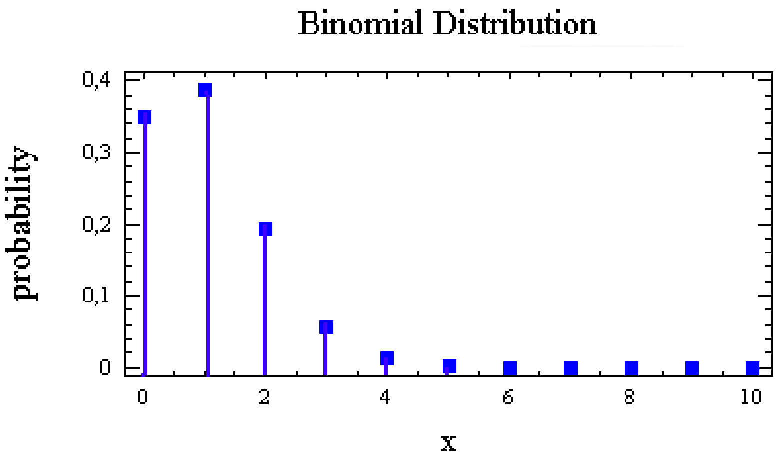
CLT application to Binomial

Let $X \sim B(n,p)$. Then $X = X_1 + X_2 + \ldots + X_n$,

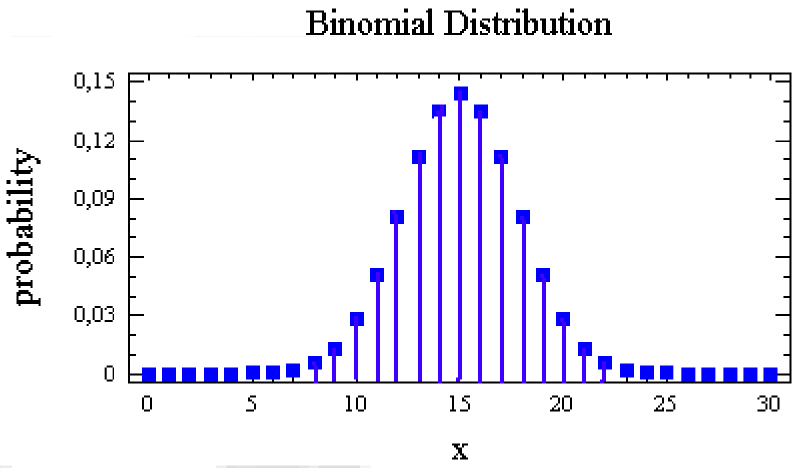
where X_i are Bernoulli random variables

If n is sufficiently large such as $np(1-p) > 5$

$X \sim N(np, np(1-p))$



B(10,0.1)



B(30,0.5)

Example: We have a batch of 5000 pieces produced by the same machine M. We know that in normal conditions the machine produces 1% of defective pieces. What is the probability that in that batch there are more than 50 defective pieces?

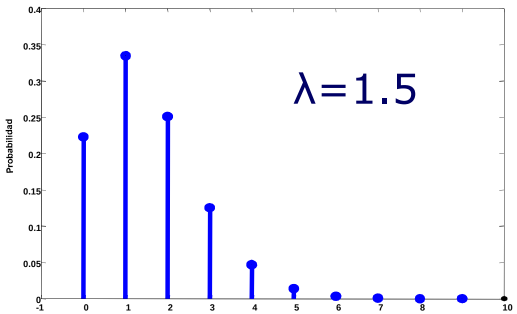
$X \sim B(5000, 0.01)$. As $np(1-p) = 5000 \times 0.01 \times 0.99 = 49.5$ we have $X \sim N(50, 49.5)$, $P(X > 50) = 0.5$

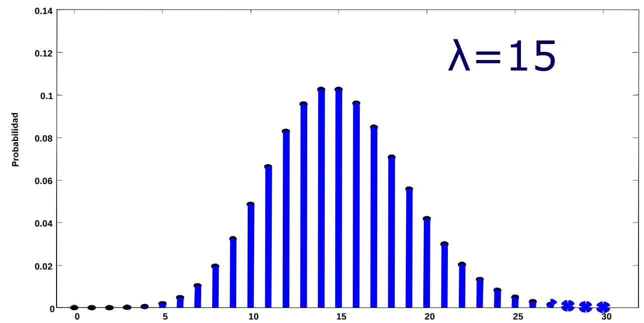
CLT application to Poisson

Let $Y_i \sim Poi(\lambda_i)$ $\longrightarrow$ Then $Y = Y_1 + Y_2 + \ldots + Y_n \sim Poi(\lambda)$ with $\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_n$

If n is sufficiently large such as $\lambda > 5$, then

$$Y \sim N(\lambda, \lambda)$$





Example:

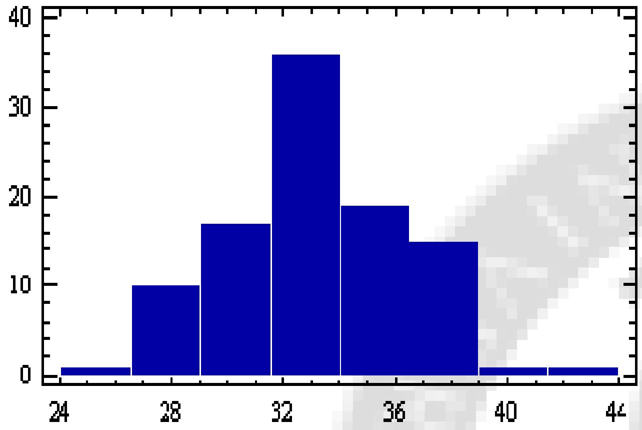
A wire of optical fiber has on average 2 defects per 10 meters. What measurement unit should be used to approximate the number of defects per unit of measure by a Normal distribution?

We need $\lambda > 5$ defects per measurement unit. Then we must use 25 meters as minimum measurement unit.

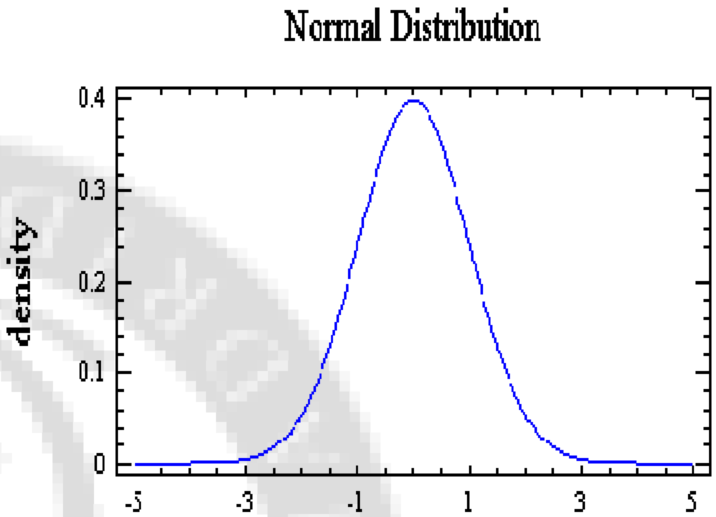
Chapter 5: Probability models

1. Introduction
2. Bernoulli process
3. Poisson process
4. Normal distribution
5. Relationship between Normal, Binomial and Poisson distributions
6. **Simple regression model**

With univariate information:

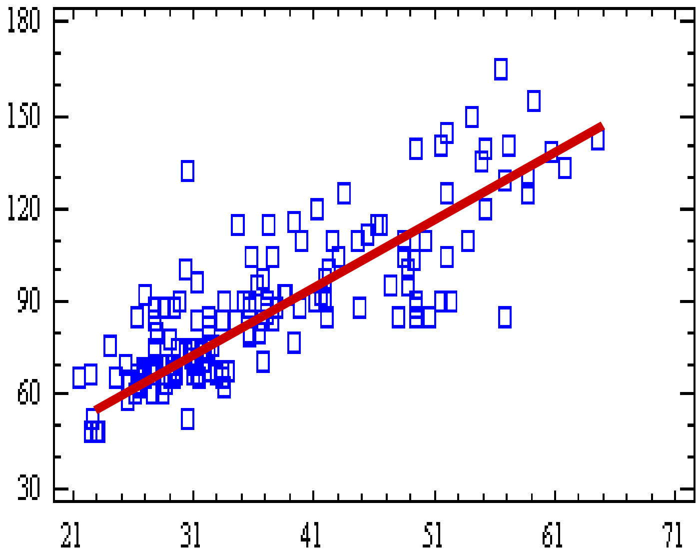


Sample

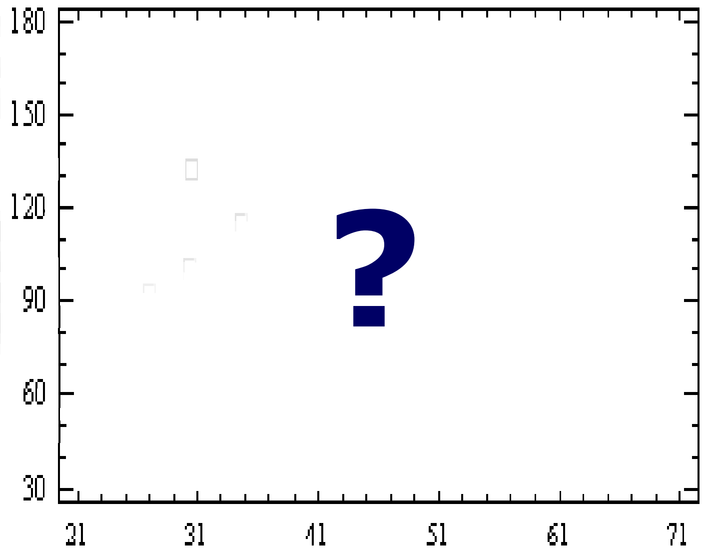


Population

What about two linearly related variables?



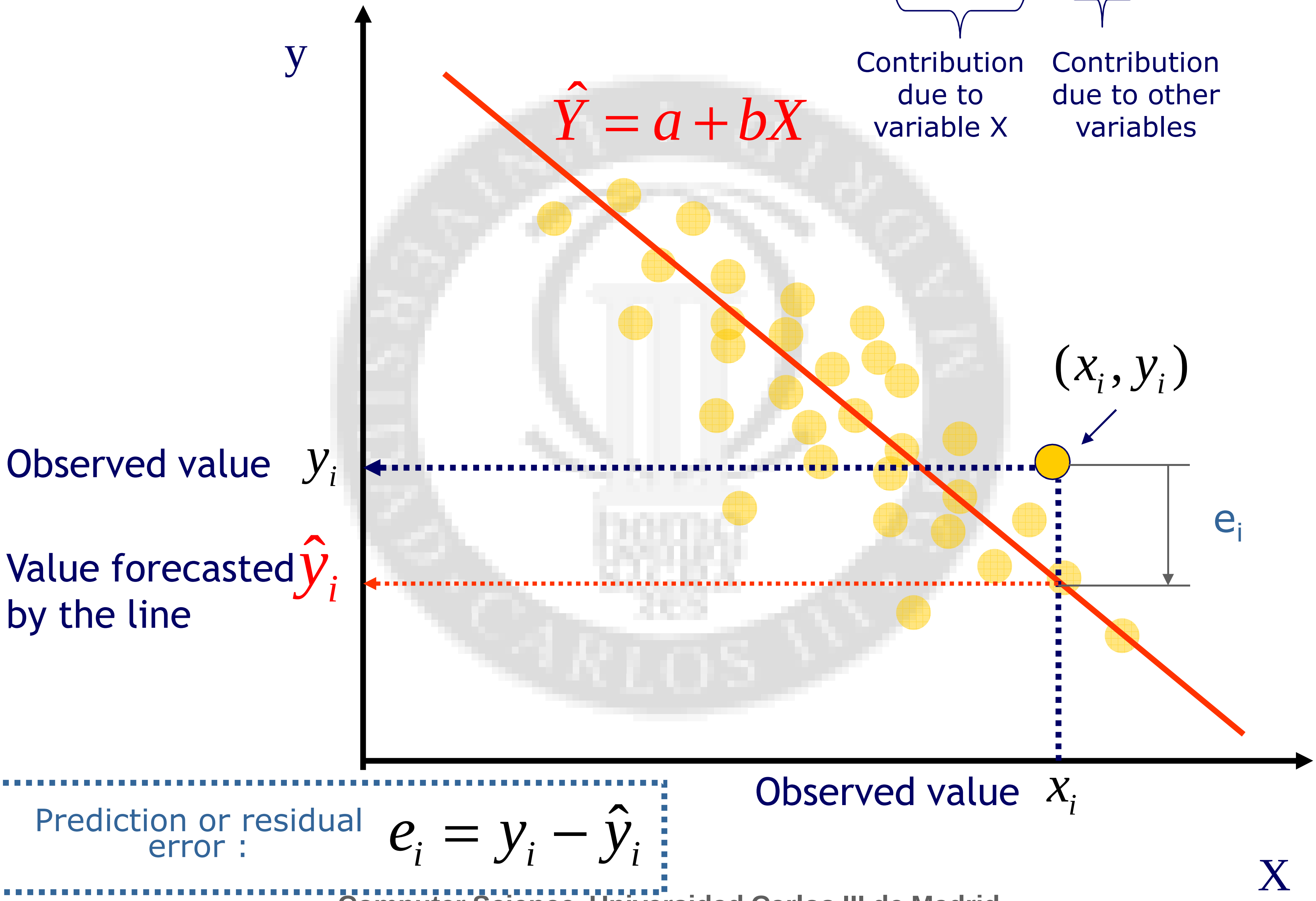
Sample



Population

Regression line

$$\underbrace{Y = a + bX}_{\text{Contribution due to variable X}} \underbrace{+\,e}_{\text{Contribution due to other variables}}$$



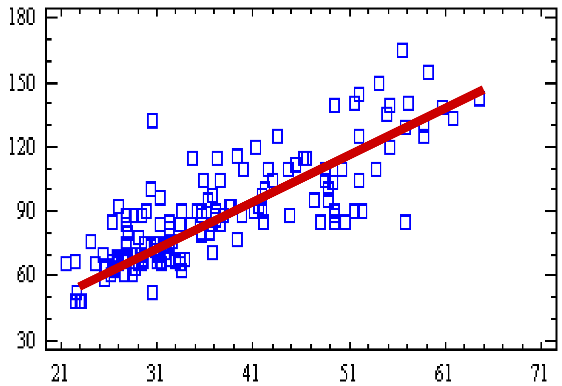
Prediction or residual error : $e_i = y_i - \hat{y}_i$

The "General Linear Model" of simple regression

$$Y = a + \underbrace{bX}_{\text{Contribution due to variable X}} + \underbrace{e}_{\text{Contribution due to other variables}}$$

$$Y = \underbrace{a + bX + \overbrace{c_1 Z_1 + c_2 Z_2 + c_3 Z_3 + \cdots}^{\text{Contribution due to other variables}}}_{\text{We assume that all the variables affect Y in an additive (or linear) form}}$$

We assume that all the variables affect Y in an additive (or linear) form



If we fix the value $X=x_i$, the value of Y varies depending on the values of the variables $Z_1, Z_2, \ldots$ that we cannot control

e is a random variable

Y is a random variable

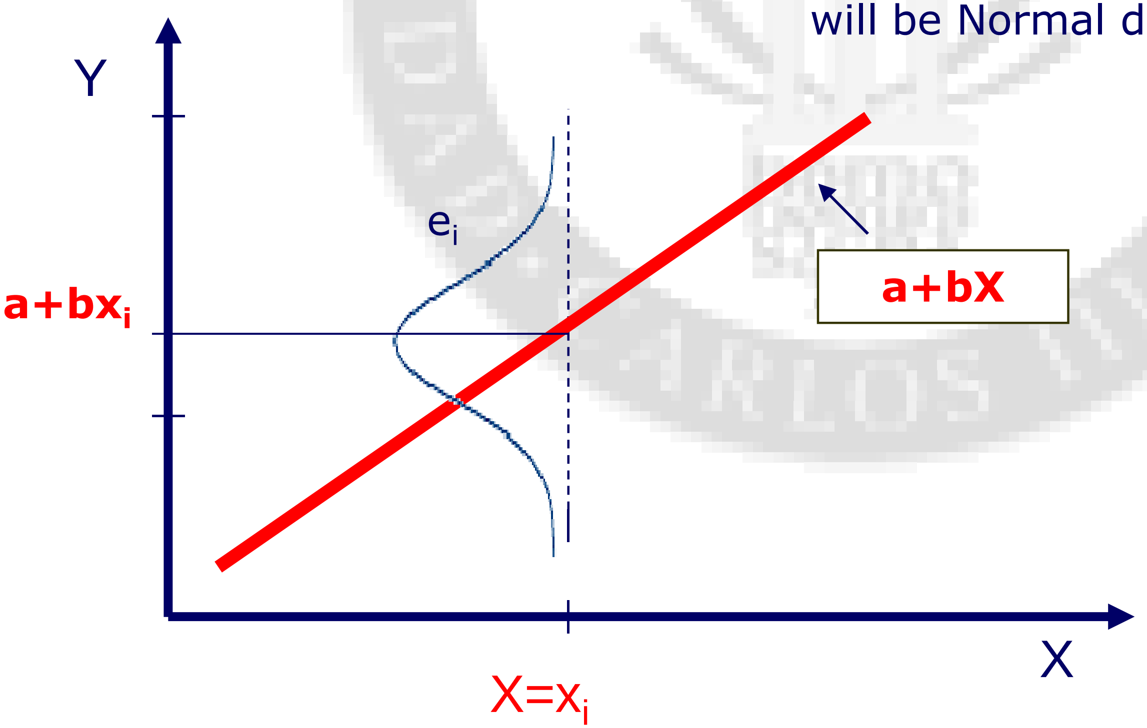
$$Y = \underbrace{a + bX}_{\text{constant}} + \underbrace{e}_{\text{random}}$$

Contribution due to variable X that assumes value x_i

Contribution due to other variables that makes the linear prediction not exact: $E(e)=0$

$$Y = a + bX + \underbrace{c_1 Z_1 + c_2 Z_2 + c_3 Z_3 + \cdots}$$

By the Central Limit Theorem this sum will be Normal distributed

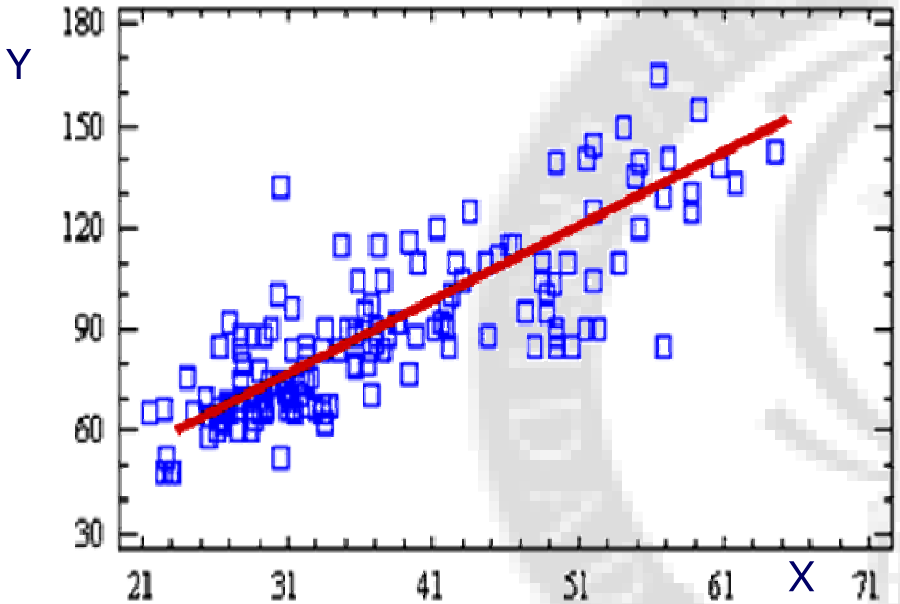


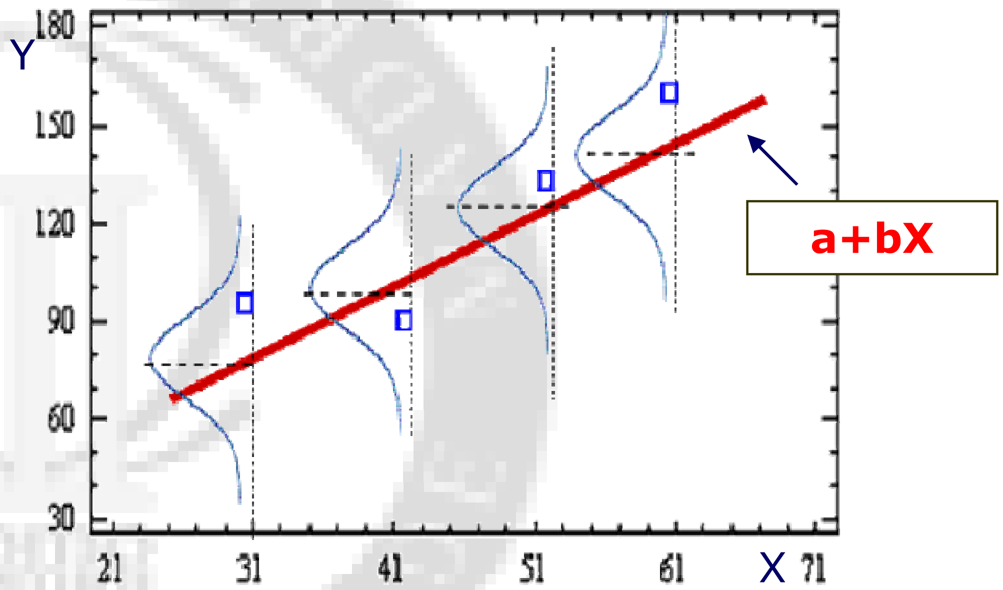
$$e_i \sim N(0, \sigma^2)$$

$$\downarrow$$

$$Y_i \sim N(a + bx_i, \sigma^2)$$

$$\underbrace{Y = a + bX}_{\text{constant}} + \underbrace{e}_{\text{random}}$$





Each observed point y_i can be interpreted as a random value normal distributed around the line, i.e.

$$Y_i \sim N(a + bx_i, \sigma^2)$$

We assume that the 'noise' is homogeneous along all the liner:
i.e. the variance is constant (homoelasticity assumption)

Example

Let Y=50+2X+e be a simple regression model with **e** distributed as an N(0,σ²=25). Compute the probability that Y is greater than 160 in the following cases:

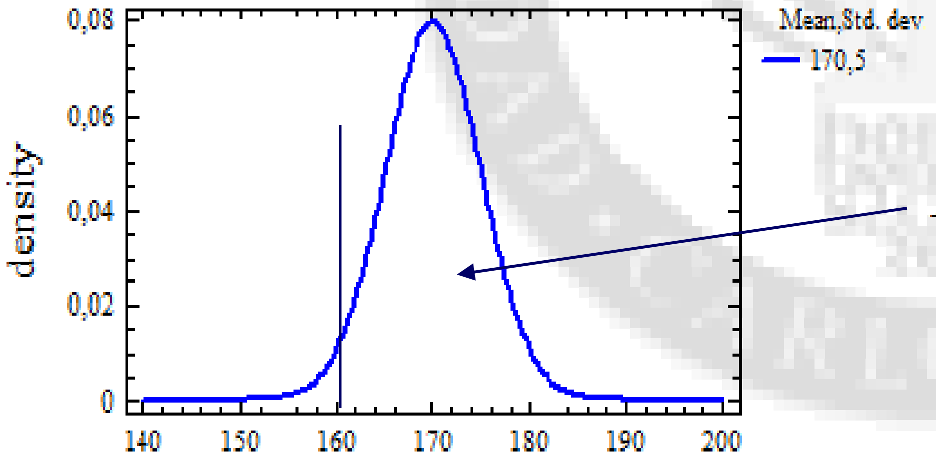
a) X=60,

b) X=40.

Case(a) $$Y = 50 + 2 \times 60 + e = 170 + e; \;\; e \sim N(0, \sigma^2 = 25)$$

$$Y \sim N(170, 25)$$

Normal distribution



$$P(Y > 160 | Y \sim N(170, 25)) = 0{,}977$$

Example

Let Y=50+2X+e be a simple regression model with **e** distributed as an N(0,σ^2=25). Compute the probability that Y is greater than 160 in the following cases:

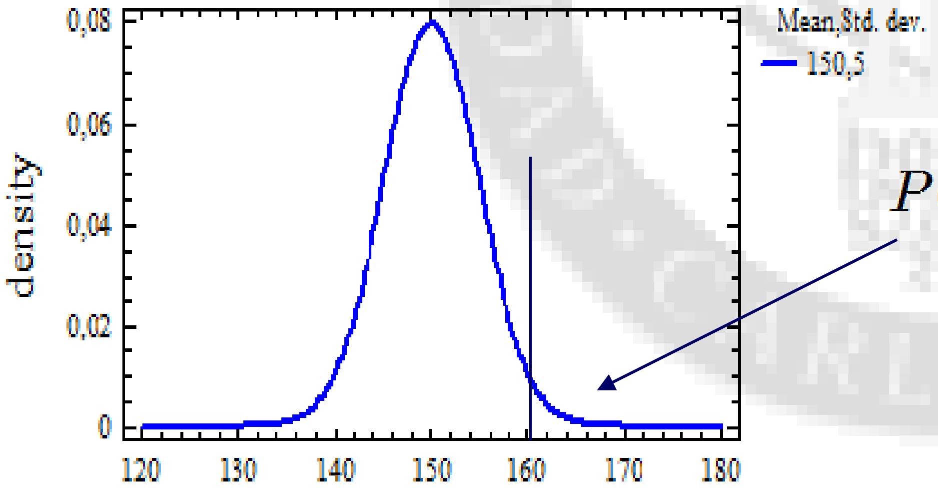
a) X=60,

b) X=40.

Case (b) $$Y = 50 + 2 \times 50 + e = 150 + e;\ \ e \sim N(0, \sigma^2 = 25)$$

$$Y \sim N(150, 25)$$

Normal distribution



$$P(Y > 160 | Y \sim N(150, 25)) = 0{,}023$$

Example

Let $Y=50+2X+e$ be a simple regression model with **e** distributed as an $N(0,\sigma^2=25)$. Compute the probability that Y is greater than 160 in the following cases:

a) $X=60$,

b) $X=40$.