Mathematics
for
Supply Chain
Msc Supply Chain

Guillaume Metzler

Université Lumière Lyon 2
Laboratoire ERIC, UR 3083, Lyon

guillaume.metzler@univ-lyon2.fr

Fall

Aim of this Course I

Present or recall some Mathematical tools

In different areas ...

- Statistics : description and inferential
- Analysis : study a function, derivation
- Linear Algebra : matrices (how to manipulate table of data)
- Basics in Optimization : finding the best solution to a problem (given constraints)
- Geometry : how to represent our data, *i.e.* having a graphical view of the abstract notions presented
- ...

Aim of this Course II

for different applications

- Behavior at the population scale
- Find correlations between observed phenomena and factors
- Make the appropriate decision when you are facing a problem
- Determine if a process is reliable or not

Examples : make decisions to avoid stock outs - evaluate the risks of a decision but also the costs (is it cost-effective to employ a person now or not ? do I need to open a new plant ?) - How should I label my products to avoid a costly lawsuit.

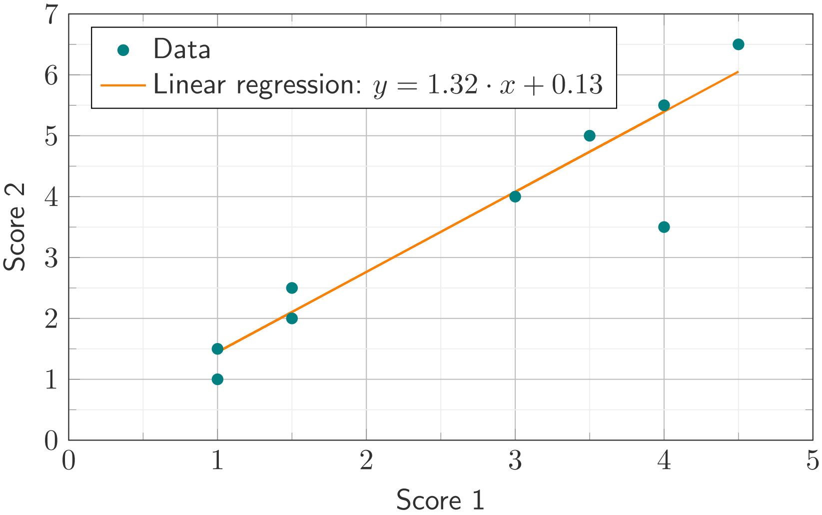
Develop tools to answer to these questions

Introduction I

Why do we study Statistics when can we use them ?

- Great tools to analyze the data and extract information (customer segmentation in Marketing - Decision Making - Risk and cost evaluations) : graphics - core values - other sophisticated methods...
- Important in decision making but also to make **predictions**, build model in **Machine Learning** to automatically provide an answer to a situation (easiest model you have ever seen linear regression)

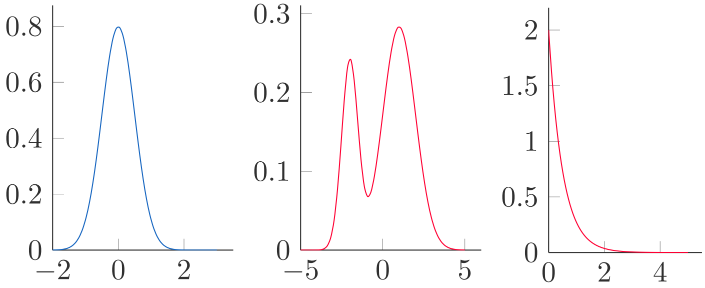
Introduction II



Introduction III

- Build some powertools to compute the risk of facing a problem and take a good decision (Expected value of the sum of two or more random variables, product of random variables)
- Extract information from the distribution of your data and have the best representation (everything is not normal !)
- Detect anomalies in the data, is the observed phenomenon normal … or is there a problem currently (dysfunction in an assembly line)

Introduction IV



Introduction V

But when can we use statistics ?

- To study stable and recurrent phenomena over time, when a stability is observed.
- In any case when the phenomena drift over time

Examples : Compare the sales of different stores in different locations - Evaluate the difference in salary between two populations - predict the sales of a product over time, depending on the period of the year (vacations, vacations,...)

But especially not to study events that are infrequent or punctual - or subject to strong fluctuations, *i.e.* **not to study random phenomenon** !

About the course I

How will it works ?

- we will be together for 8 sessions of 3 hours
- Mix between theory and pratice (using excel)
- No particular work to do at home (read the slides again but I can give exercises too if you ask for it)
- When it comes to the exam ... we will see during the second semester (not decided yet)

About the course II

The goal is not to make a course of Maths but to present you tools to help you make decisions so that you are a real force in your future company

We would like to introduce you to the tools and what they can be used for. Which tool can I use in a given situation ? What is the best method that I can use ?

It is therefore more a question of remembering the use of the tools presented (and understanding them) than of memorizing formulas and results that the solvers can find for you.

Let us go on with some basics in statistics

Basics in Statistics I

In few words

- you are not able to study all the population (huge cost if you have to check all the pieces on a line of production)
- you trying to get this information using a **sample** on which the value of interest is studied

$$x_1, x_2, x_3, \ldots.$$

It can be either **numerical** or **categorical**

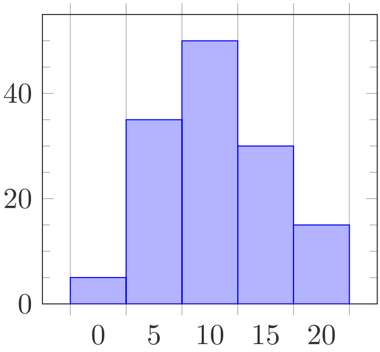
- to induce an information at the population scale

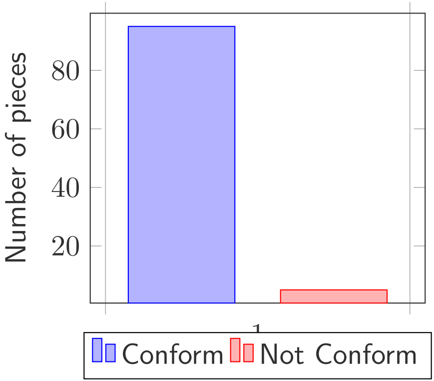
How to use this sample ? What can I do with these values ?

Basics in Statistics II

Draw some graphs to represent your data

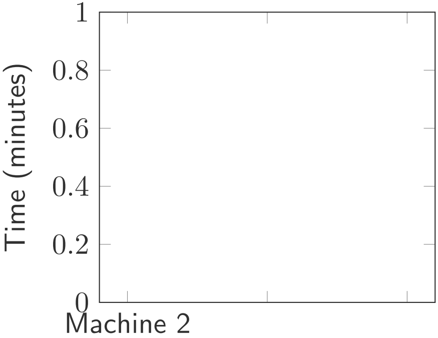
Plot a histogram just to have an idea of the repartition of the values (real values) or just count the number of occurences (categorical one)





Basics in Statistics III

Compare the effectiveness of two machines in production



Study the values distribution (or to detect some anomalies!!!

Basics in Statistics IV

How to get this graph ? How can I read them ?

We will need to extract some statistical figures to have more precise information on the values distribution and to provide answers to some questions

- Are my values located around a given one ? Are they concentrated ?
- What is the dispersion of the values ? Is the process stable or not ?
- How the distribution looks like ?

Basics in Statistics V

What is a central value ?

Basics in Statistics VI

You all have heard about the **mean** value on a sample, which is basically the average values of your sample.

Définition 3.1: Average

Given a collection of data $x_1, x_2, \ldots, x_n$, the mean value of the sample is defined by

$$\frac{1}{n}(x_1 + x_2 + \ldots + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

We can generalize this definition ... we will see that later.

Basics in Statistics VII

Example

We consider the following table which represents the salary of different employees of a company

Salary	2k	3k	2.5k	2.5k

The mean value is then equal to

$$\frac{1}{4}(2+3+2.5+2.5) = 2.5k$$

Okey a very easy example, let us consider another question that can raises when you are in a company.

Basics in Statistics VIII

Another Example

We are studying results of a company. The company generated 20% profitability in the first year, 15% in the second year, 33% in the third year and 25% in the fourth year.

Are you able to compute the mean generated profitability ? Do you even have an idea on how you can compute it ?

Basics in Statistics IX

Here the sample mean, *i.e.* the arithmetical one has completely no sense. We have the use the **geometric** one.

$$\sqrt[4]{x_1 \times x_2 \times x_3 \times x_4}$$

to have an idea of the real mean profitability over the four years. here we have an average profitability of

$$\sqrt[4]{1.2 \times 1.15 \times 1.33 \times 1.25} \simeq 1.23.$$

Check this fact assuming an initial profitability equal to 1000 at the beginning.

Basics in Statistics X

A last Example

We wish to estimate the average transport speed according to the routes taken in order to determine the fastest route.
A first trip allows to drive for 100km at 50km/h then at 90km/h for 100km and finally at 80km/h for 100km.

Are you able to compute the average speed using this way and determine how long does it take to ship the goods ?

Basics in Statistics XI

As in the previous example, the answer does not consist in computing the standard mean of the sample, but the **harmonic** one.

$$\frac{3}{\frac{1}{50}+\frac{1}{90}+\frac{1}{80}}.$$

When you compute such a mean, you see that high values do not contribute a lot in its evaluation (because the ratio will be really low).

In our case, we can sat that the average speed is approximately equal to 69km/h and ... it make sense !

Let us go back to central values

Basics in Statistics XII

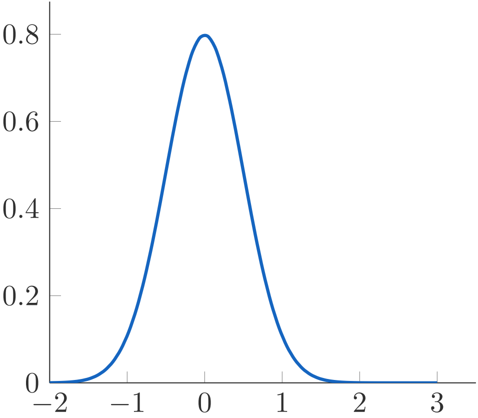
It is a value very often used for

- profile the average person in society (based on several characteristics)
- we can therefore identify a trend in the data
- this may reflect the majority thinking in the society

Basics in Statistics XIII

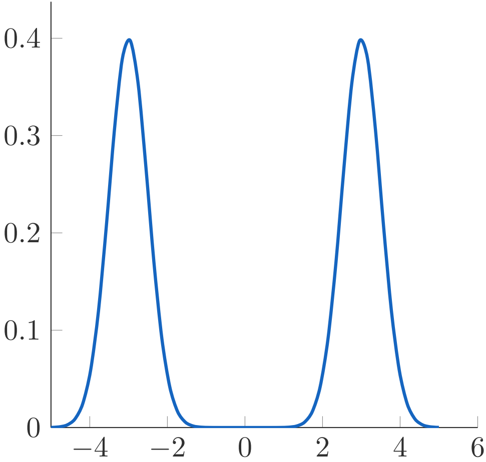
What if our sample is normally distributed ?

Why do we observe this bump ? How can interprete this curve ?



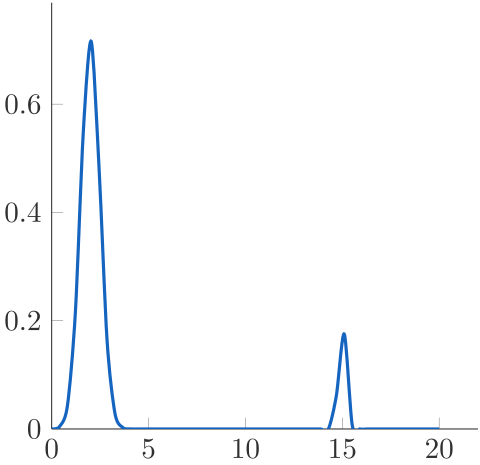
Basics in Statistics XIV

Does it always make sense to consider the mean ?



Basics in Statistics XV

Does it always make sense to consider the mean ?



Basics in Statistics XVI

- First example : the mean is exactly in the middle of the two groups (*i.e.* equal to 0 but does represent the mean behavior of the sample).
- Second example : the mean is **biased** by the large values that our data can take (imagine you want to compute the average the average salary of a sample in which there is a millionaire ... even a billionaire). The use of the median is more relevant in this case, more representative of a huge part of the population This little example also raises the question of how to build a good sample for statistical studies

Importance to study of the data are spread around the mean

Basics in Statistics XVII

How to measure dispersion ?

Basics in Statistics XVIII

Several measures can be used

- the standard deviation or the variance
- you can also compute the range of the values
- different quantiles

Basics in Statistics XIX

The Variance/Standard Deviation

Définition 3.2: Variance

Given a collection of data $x_1, x_2, \ldots, x_n$, the variance σ^2 our our data can be estimated by

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

In other words, it just compute a *quadratic distance*, *i.e.* the squared distance between each sample and the mean value and average them. The standard deviation σ is just the square root of the variance.

Basics in Statistics XX

The range

Définition 3.3: Range

Given a collection of data $x_1, x_2, \dots, x_n$, the range the statiscal series is defined by

$$\max(x_i) - \min(x_i)$$

In other words, we simply compute the difference between the largest sample value and the smallest one.

Basics in Statistics XXI

These two first measures provide several information on the stability of the process, and thus, its reliability.

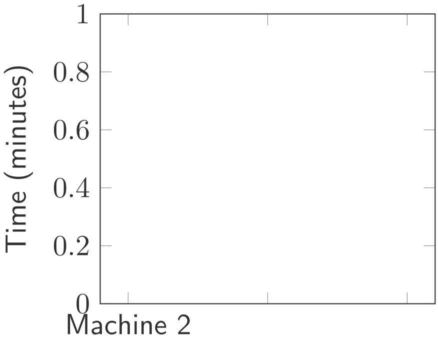
- a large variance means that the process is not stable and that most of values are located far from the mean
- if the range of values is high, it means that our is facing some anomalies

We usually use these facts to detect anomalies on an assembly line and perform maintenance.

Basics in Statistics XXII

The Quantiles

It gives a more precise information on the values partition of the statistical series.



Basics in Statistics XXIII

Définition 3.4: Quantiles

Let us consider $x_1, x_2, \ldots, x_n$ some observed values and let us denote $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ the same sample but **ordered** this time. Then the quantile of order α is the value $x_{(j)}$ such that a proportion α of the sample values are less or equal to $x_{(j)}$

Examples

- The quantile order 0.5 is also called the *median*
- The quantile of order 0.25 and 0.75 are called the 1st and 3rd *quartiles*
- The quantile of order 0.1 is called the first *decile*

Basics in Statistics XXIV

For now, we have seen how to study a single statistical series.
But most of the time, we want to know if there are links between two series of values and thus detect possible correlations.

For example, we could study the link between the turnover generated by a brand and its sales area.

Basics in Statistics XXV

Définition 3.5: Correlation

Let us consider $x_1, x_2, \dots, x_n$ and $y_1, y_2, \dots, y_n$ two series of observed values. The linear correlation between these two series of values is defined by :

$$\frac{1}{n} \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{s_x \times s_y},$$

where $\bar{x}$ and $\bar{y}$ are the mean values of x and y respectively and s_x and s_y their associated standard deviation.

Basics in Statistics XXVI

An application

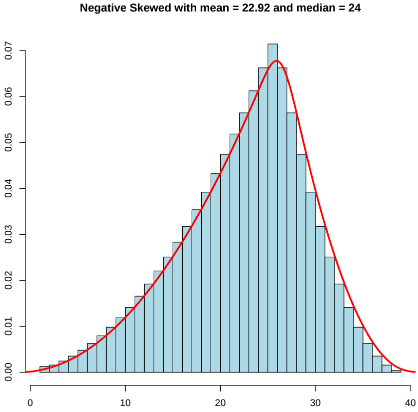
Let us consider the file providing on Brightspace associated to this first lecture. It consists in two columns with several values
For each of this sample, you will have to compute (with excel) :

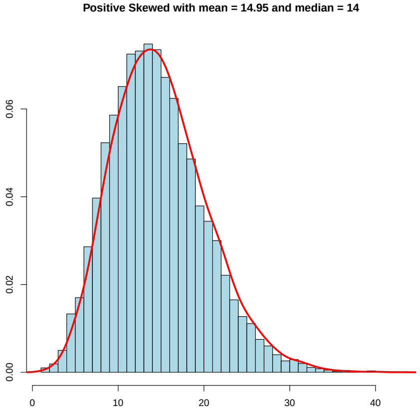
- the mean value
- the standard deviation
- the median
- the 1st and the 3rd quartile

You will finally compute the correlation coefficient between the two variables in order to determine if there is a link between the two random variables.

Basics in Statistics XXVII

We can further when it comes to study our data, testing the symmetry (or the *skewness*





Basics in Statistics XXVIII

Definition of the skewness :

$$\gamma = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s}\right)^3.$$

Interpretation :

- a **negative** coefficient indicates a distribution shifted to the right of the median, and thus a distribution tail spread to the left.
- a **positive** coefficient indicates a distribution shifted to the left of the median, and thus a tail of the distribution spread to the right

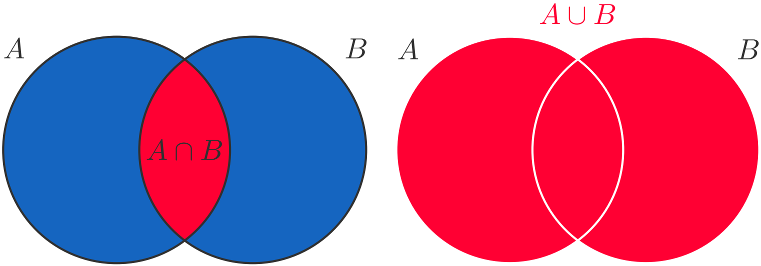
We can go further (see for instance the *kurtosis*)

Probabilities

Where these values come from ?

Probabilities I

Before talking about probabilities let us talk about logic and events.



Three operator *AND*, *OR* and *XOR*. Link with events ?

Probabilities II

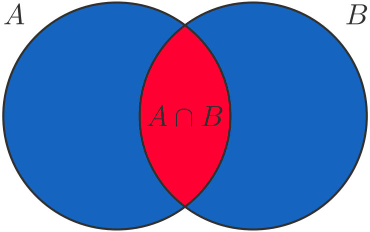
They appear naturally when we study several events at the same time and especially the possible links between the events.

This can be of interest when you want to evaluate the loss of turnover generated when a production line is stopped and according to the configuration of the latter.

Probabilities III

The AND operator

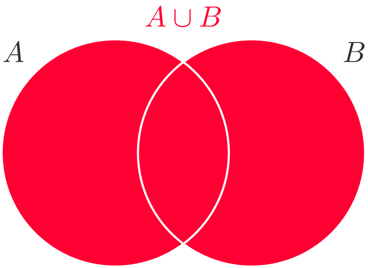
If a production line is in **parallel**, *e.g.* two lines are doing the same assembly process. Then the production will stop only if both lines are stopped. Thus, losses are generated if lines 1 **and** 2 are stopped.



Probabilities IV

The OR operator

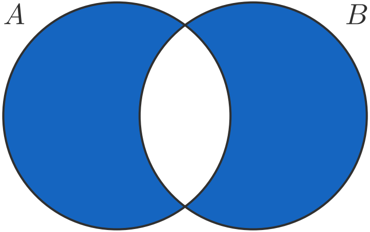
If a production line is in series and consists of several stages, then production will stop and result in a loss of turnover if at least one stage of the assembly process is stopped.



Probabilities V

The XOR operator

We can go back to the first example, but this time we can imagine that if the two production lines work at the same time, we also risk generating too much product, which would then have to be stored or which could degrade over time.
Use only one line, but not the two.



Probabilities VI

In fact, there would be one last case to study, the dependence between two events, which would also lead us to study the dependence between two random variables and thus conditional probabilities.

Example We could consider a random variable X that represents the summons and a random variable Y that represents the amount of compensation for an injury suffered.

Probabilities VII

What is a Random Variable ?

Probabilities VIII

Two types of variables

- **Discrete : Binomial** - **Bernoulli** - Geometric - Hyper-geometric - Rademacher - Poisson - Multinomial
- **Continuous : Normal** - log Normal - Exponential - Uniform - Fisher - $\mathcal{X}^2$ - Gamma - Beta - Dirichlet - Weibull - Laplace - ...

There exist more than a thousand of distributions ! We are not going to study all of them, the one in bold and the others will just be used.

Discrete Random Variable I

They take their values in a discrete set, for example the set of natural integers or the set of relative integers. This set can be **finite** or **infinite**.

Examples : toss of a coin, the number of customers per day in a supermarket, the number of rejected product on a line of production, ...

In general we estimate a probability

$$P(X = k) = p_i.$$

Discrete Random Variable II

A **probability distribution for a discrete variable** is a mutually exclusive list of all possible numerical outcomes for this variable and probability of occurrence associated with each outcome.

Nb Interruption/day in Network	Probability
0	0.67
1	0.1
2	0.09
3	0.08
4	0.06

The probabilities shall sum to 1 and we can represent them using a kind of histogram.

Discrete Random Variable III

The discrete law that you are probably most familiar with is the binomial law.

It is mainly used in the process of detecting anomalies in production lines when we know the probability that a machine generates a defective part. More precisely, it can be used to check that the machine is not out of order.
The way it works is quite easy because it is enough to count or enumerate the possible outcomes to get to find the probability of the studied event.

Discrete Random Variable IV

The most classical example, corresponds to the experiment of a coin toss that is performed several times independently.

Another example. Let's imagine that we are on a line where the proportion of defective parts is equal to 0.05. If we draw 5 pieces at random, what is the probability to observe at most 1 defective piece ?

Discrete Random Variable V

Given a random variable X which follows a Binomial distribution with parameters n (number of trials) and p (probability of success, observing a failure). The probability density function is defined by

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

where is the number of different ways to observe exactly k successes in n independent experiments.

Discrete Random Variable VI

Some quantities

As we have computed mean and variance for a sample, we can give a similar definition when we are studying random variables (they are exactly the same)

Définition 4.1: Expectation

Let us consider X a discrete random variable, then the expected value of X is defined by

$$\mathbb{E}[X] = \sum_k k \times P(X = k).$$

We can do the same thing with the variance.

Discrete Random Variable VII

Définition 4.2: Variance

Let us consider X a discrete random variable, then the Variance of X is defined by

$$\begin{aligned} Var[X] &= \mathbb{E}[(X - \mathbb{E}[X])^2], \\ &= \sum_k (k - \mathbb{E}[X])^2 \times P(X = k), \\ &= \sum_k k^2 \times P(X = k) - \mathbb{E}[X]^2 \end{aligned}$$

and the standard deviation σ is the square root of the variance.

Discrete Random Variable VIII

For the Binomial $\mathcal{B}(n, p)$ distribution, where n is number of trials and p the probability of "success", we have :

$$\mathbb{E}[X] = np \quad \text{and} Var[X] = np(1-p)$$

Discrete Random Variable IX

Définition 4.3: Covariance

Let us consider X and Y two discrete random variables which can take the same values. Then the covariance $Cov(X,Y)$ between the two random variable is defined by

$$Cov(X,Y) \quad = \quad \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

and the correlation is the the covariance divided by the standard deviation of X and the standard deviation of Y.

Discrete Random Variable X

Some properties

In general :

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] + Cov(X,Y).$$

$$Var[X+Y] = Var[X] + Var[Y] + 2Cov(X,Y).$$

$$Var[XY] = Var[Y]\mathbb{E}[X^2] + Var[X]\mathbb{E}[Y^2].$$

Discrete Random Variable XI

If X and Y are independent :

$$\mathbb{E}[X \times Y] = \mathbb{E}[X]\mathbb{E}[Y].$$

$$Var[X + Y] = Var[X] + Var[Y].$$

Discrete Random Variable XII

Poisson Distribution

You use the **Poisson distribution** when you are interested in the number of times an event occurs in a given area of opportunity.

An area of opportunity is a continuous unit or interval of time, volume, or such area in which more than one occurrence of an event can occur.

For example if we want to estimate the probability that 5 people are present in a shop knowing that on average 4 people enter the shop every 20 minutes

Discrete Random Variable XIII

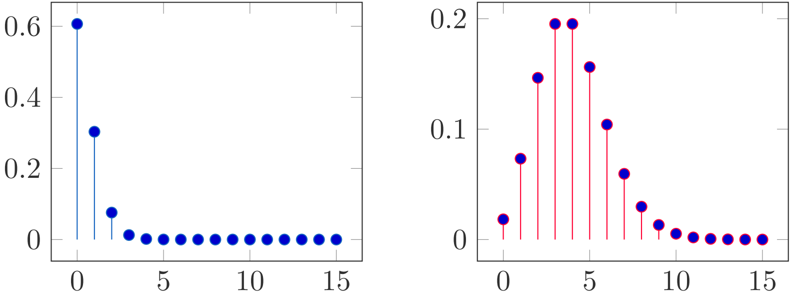
Poisson Distribution

If the average number of occurrences in a fixed time interval is λ, then the probability that there are exactly k occurrences (k being a natural number, $k = 0, 1, 2, \ldots$) is

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

For this distribution, the mean is equal to the variance which is equal to λ.

Discrete Random Variable XIV



Poisson distribution with $\lambda = 0.5$ and $\lambda = 0.4$.

Discrete Random Variable XV

It is used finance to model the probability of credit default, Yield Management (American Airlines, Lufthansa and SAS to estimate passenger demand).

A retailer is experiencing high traffic and wants to know if they need to hire.

In general, a customer leaves without paying when the store is too busy and the waiting time at the checkout is too long, which generates a loss of turnover.

It is therefore interesting to study the probability that this threshold is crossed and to deduce the loss of turnover (considering the average turnover per customer) in order to evaluate if it is profitable or not to hire another employee.

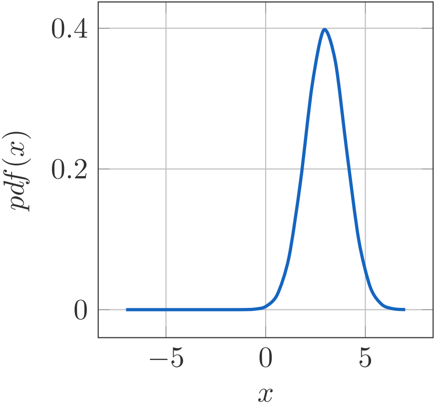
Discrete Random Variable XVI

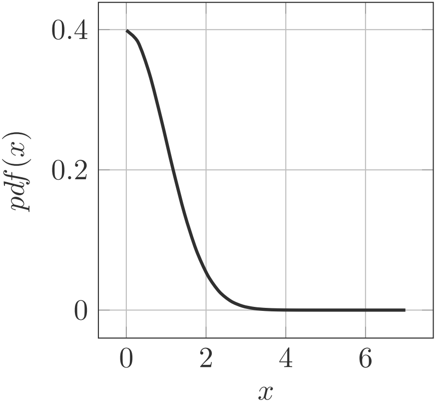
An example

Let's imagine that in a store the average number of customers is 7 per hour.

- Estimate the probability that there are exactly 7 people in the store
- Estimate the probability that the store has to handle more than 6 customers at the same time.

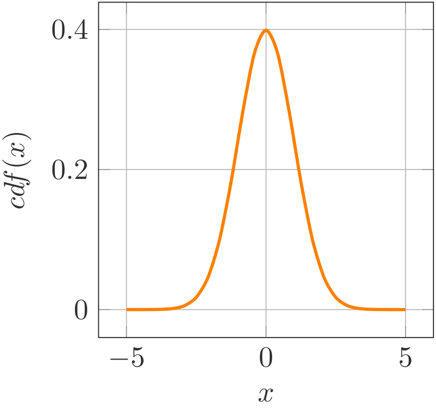
Continuous Random Variable I





The Normal Distribution I

About the Normal Distribution



The Normal Distribution II

It has the following properties

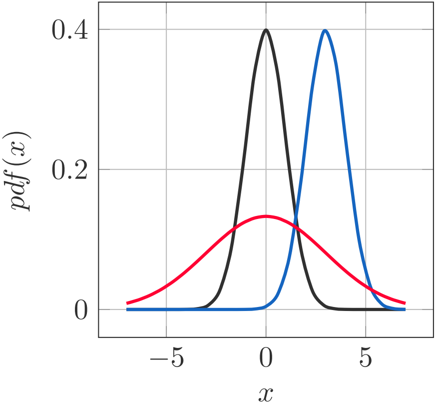
- completely defined by its **mean** and its **standard deviation**
- it can take any real values
- it is symmetric
- the mean parameter μ determines the location of the bell and σ how the values spread around the mean (diffusion)

The function is defined by :

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

The Normal Distribution III

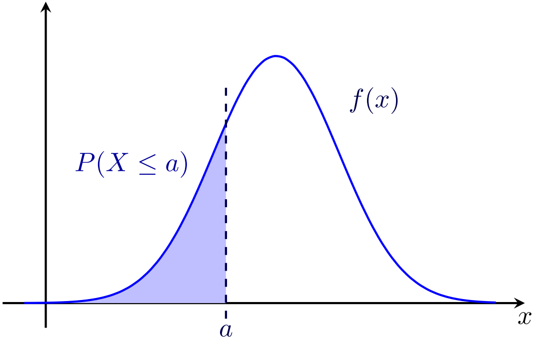
Changing the parameters



The Normal Distribution IV

Estimate probabilities

Suppose f is gaussian, then $F(a) = P(X \leq a) = \int_{-\infty}^{a} f(t)\, dt$



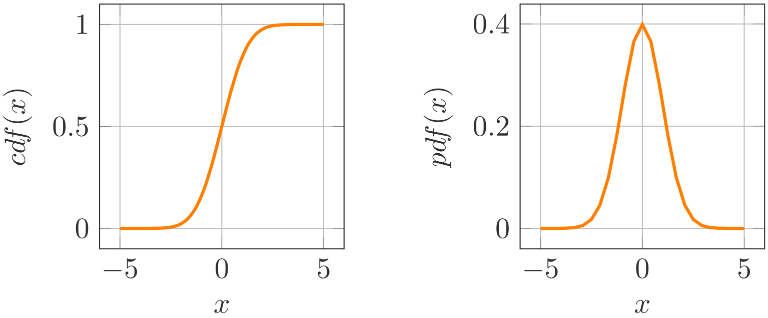
The Normal Distribution V

In other words, when you are studying density function for continuous random variable the probability that a random variable is taking values lower or equal than a the *area under curve* on left of the a value

It also shows that area under the curve of *probability density function* is always equal to 1.

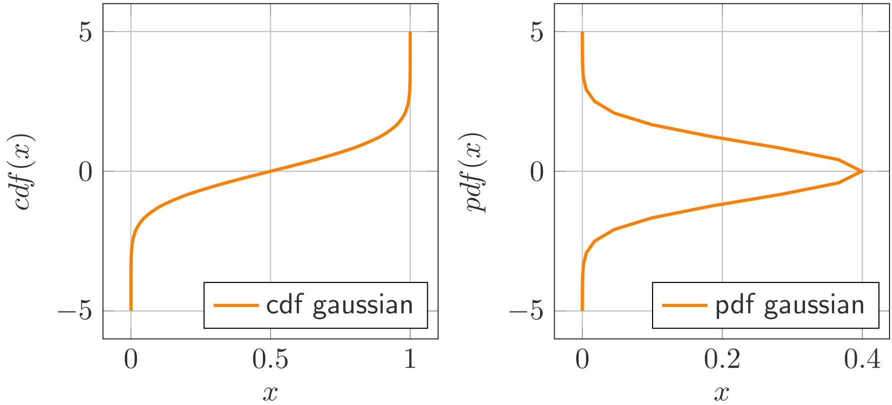
The Normal Distribution VI

The area under the curve can then be seen as function of value of a, *i.e.* to a *probability density function* we can associate another function which will describe the probability that a random variable having this distribution takes values lower or equal than a : **continuous density function**



The Normal Distribution VII

Link between cdf and pdf



The Normal Distribution VIII

Case

When we have access to the distribution of the process, we can easily compute, for instance, the probability that the object is rejected due to measured defect (use of the cdf function).

If this probability is to high ..., we aim to find the right settings of the machine, such that probability of being a reject is less than a given value. So the problem is exactly the reciprocal of the previous point :

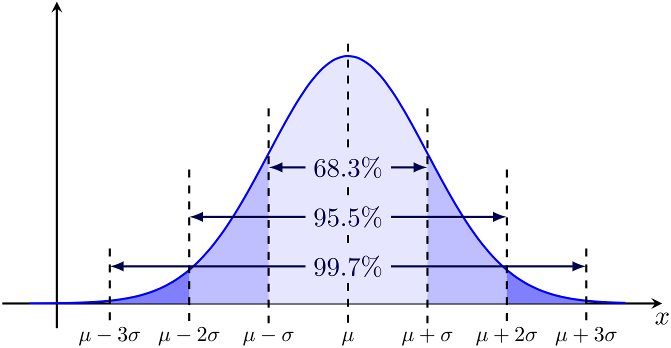
- we know the *probability*
- we aim to determine the associated *quantile*

The Normal Distribution IX

With the previous statement that the area under the curve is equal to 1, we can say that :

- $P(X \geq a) = 1 - P(X \leq a)$
- $P(a \leq X \leq b) = P(X \leq b) - P(X \leq a)$

The Normal Distribution X



The Normal Distribution XI

How to find these probabilities ?

The integral cannot be computed by a human

$$P(X \leq a) = \frac{1}{\sqrt{\sigma}2\pi} \int_{-\infty}^{a} e^{\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

However, it exists some tables where you can have access to such results but ... for only for a special normal distribution. The one with 0 mean and a variance (or standard deviation) equal to 1.

To go from any normal distribution to a centered and reduced normal distribution, we will use the standardization process.

The Normal Distribution XII

The standardization process

It consists in applying a transformation to any random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ in order to create, an equivalent random variable, with a mean equal to 0 and a variance equal to 1

$$Z = \frac{X - \mu}{\sigma}.$$

We can check that $\mathbb{E}[Z] = 0$ and $Var[Z] = 1$ using the properties of the expectation and variance.

The Normal Distribution XIII

Case

We consider the production of 85g cans. The process includes : filling in the cans, sealing, weighing and disposal if the weight is less than 84g (production waste), sorting, labeling. Even if the automatons are of very good quality, it is impossible to obtain cans weighing exactly 85g. There is a standard deviation in packing the cans related to the technical processes, which follows a normal distribution of the standard deviation equal to 0.8g.

You will have first to identify the studied random variable and its properties (mean and variance for instance) but also the target of this study !

The Normal Distribution XIV

Once it is done, answer to the following questions

1. estimate the probability that the weight of the can is less than $84.5g$
2. estimate the probability that the weight of the can is more than $85.5g$
3. estimate the probability that the weight of the can is between 84.7 and 85.3
4. estimate the probability of a can being rejected by the sorting process.
5. we aim to change the settings such that at most 1% of the product are rejected by the sorting process. How do we have to change the value of μ.

Normality Evaluation I

How to evaluate normality ?

As we have already seen, all the distribution are not normal, and there is no reason for all of them to be normal !
Evaluating normality can however be important, to use other statistical tools.

Empirically :

- build a boxplot of your data or an histogram
- check the symmetry but also how the values are distributed

Normality Evaluation II

Compute some figures :

- compute mean, median and find the mode
- are these three values equal ?
- is the range of the values between $\pm 3\sigma$
- are half of the values located in $\mu \pm 0.67\sigma$
- compute the skewness and other coefficients ...

But this is usually not enough ...

Normality Evaluation III

It exists an other graphical way to check normality which is called the quantile-quantile plot