Probability 1

Chapter 05 : Continuous Random Variables - Part 1

Dr. Daniel Flores-Agreda,

(based on the notes of Prof. Davide La Vecchia)

Spring Semester 2021

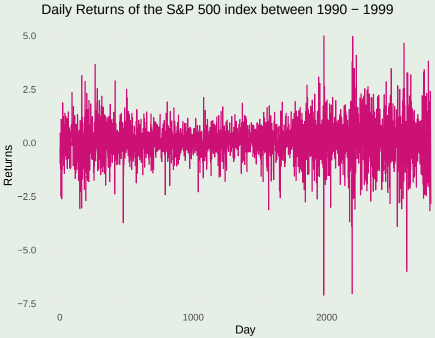
Objectives

- Illustrate Continuous Random Variables case uses.
- Characterise Continuous Distributions and overview their features (Density Function, Cumulative Distribution, Expectation, Variance, . . .)
- Explore the Uniform and Gaussian Distribution.

Outline

Outline

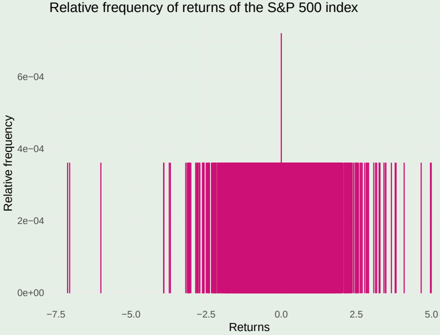
Example (Standard & Poors 500 returns)



Two Motivating Examples

Example (Standard & Poors 500 returns)

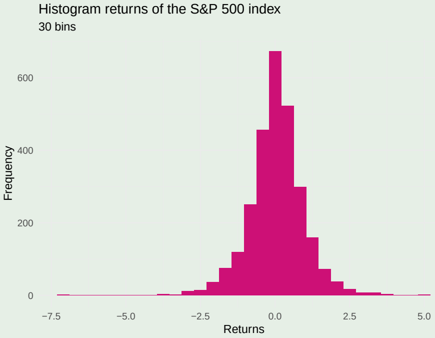
Let's try to count the relative frequency of each of these returns, to estimate the probability of each value of the return.



- Too many different returns. Low and "uniform" relative frequency
- However, there's some *concentration* of .
- Let's create a *histogram* by counting the number of observations within given intervals (or "bins").

Example (continued)

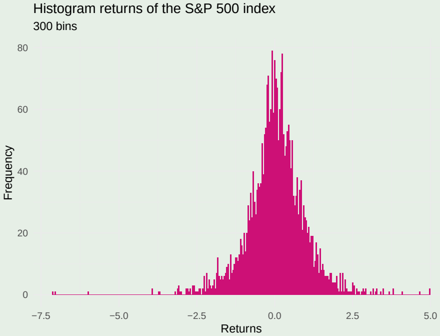
We can analyze the *distribution* (e.g. some returns are more likely than some others?) via the histogram



Notice how that *concentration* becomes more apparent.

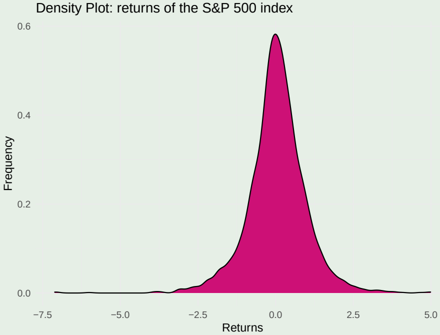
Example (continued)

Increasing the number of "bins" to 300



Example (continued)

Estimating the *density curve*



with an infinite number of bins (essentially estimating a curve).

Two Motivating Examples

Example (Cafeteria)

Let us consider a serious/significant issue: the arrivals to the cafeteria UniMail, from 10AM to 2PM

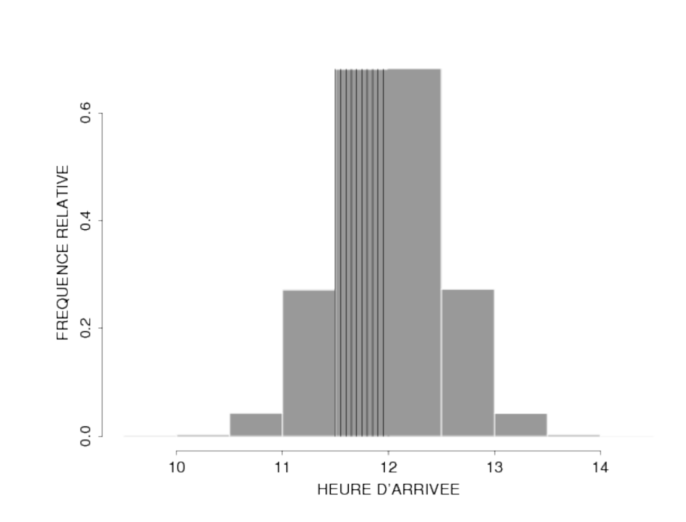
$$\text{relative freq} \quad = \quad \frac{\#\ \text{customers incoming}}{\#\ \text{total of customers}}$$

Aim & Scope

We want to study the distribution of this object over the considered time interval. E.g. we would like to know when the relative frequency has a peak...

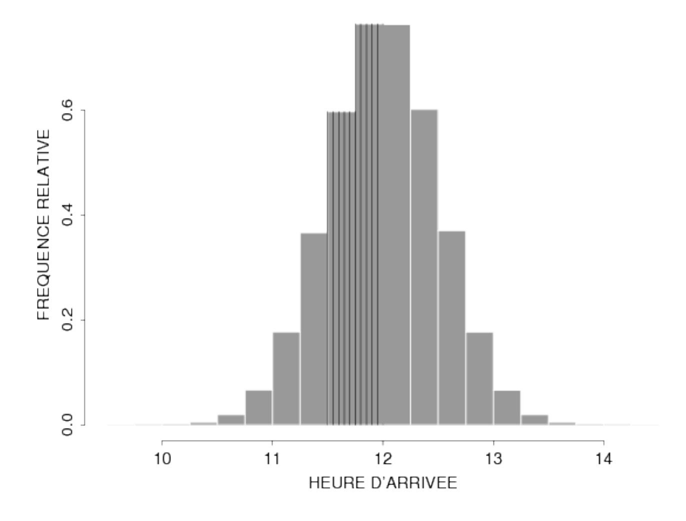
Two Motivating Examples

Example (continued)



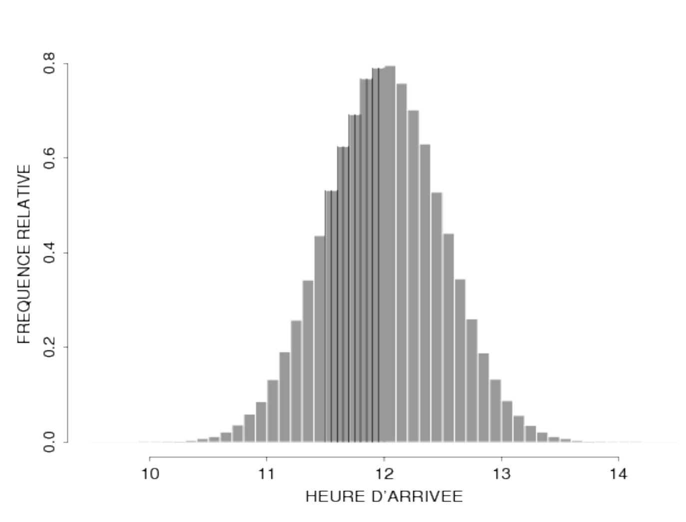
Two Motivating Examples

Example (continued)



Two Motivating Examples

Example (continued)



Two Motivating Examples

These phenomena constitute examples of phenomena that can be modeled with **continuous random variables** i.e. variables that **can take any value** in an interval or $\mathbb{R}$.

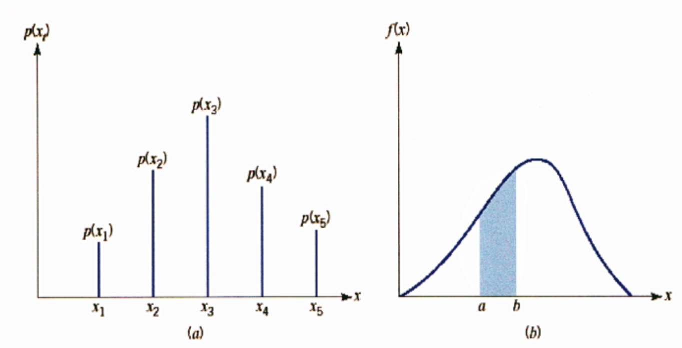
We cannot simply *list* all possible values of the random variable, because there are (infinitely many) an **uncountable number of possible outcomes** that might occur.

In this context, we characterise the probability distribution by **assigning a positive probability to each and every possible interval of values** that can occur.

This is done by defining the **Cumulative Distribution Function (CDF)**, also called "Probability Distribution Function" of the Random variable.

So, graphically, we have

Discrete   Continuous



Outline

Definition (Cumulative Density Function (CDF))

Let X be a continuous random variable and let $x \in \mathbb{R}$.

The CDF of X at the point x is a continuous function $F_X(x)$ such that:

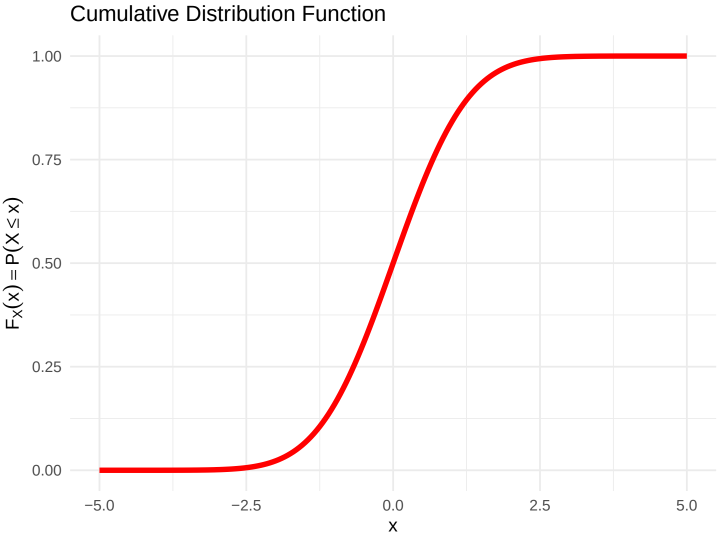
1. $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to +\infty} F_X(x) = 1$,
2. $0 \leq F_X(x) \leq 1$ for all $x \in \mathbb{R}$ and
3. the function is **monotonically non-decreasing in** x

$$F_X(x) \geq F_X(x') \quad \text{for all} \quad x > x',$$

and the value $F_X(x)$ yields **the probability that** X **lies in the interval** $(-\infty, x]$

$$F_X(x) = P(X \leq x)$$

Characterisations of Continuous Distributions



Definition (Probability Density Function (PDF))

Let X be a random variable taking values in the interval $(a, b]$. Since:

- $F_X(x)$ is zero for all $x < a$
- $0 < F_X(x) < 1$ for all x in (a, b) and
- $F_X(x) = 1$ for all $x \geq b$.

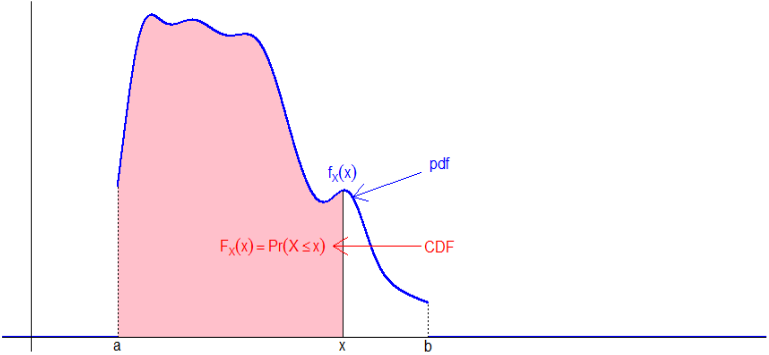
Then, we can define the Probability Density Function (pdf) of X at the point x as:

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

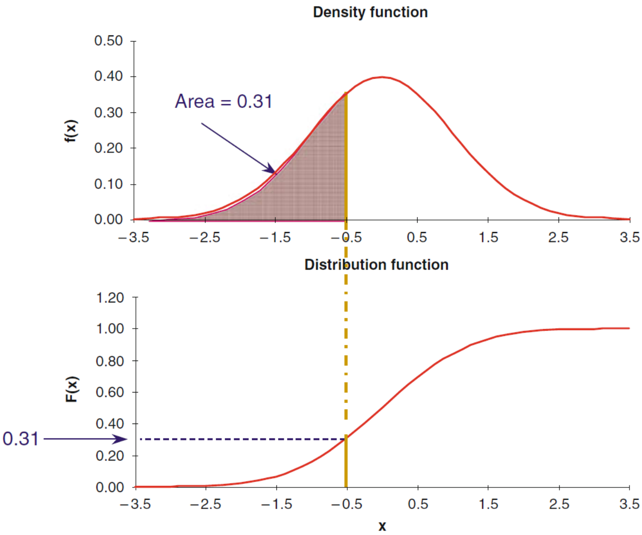
Remark

(VERY) roughly speaking, the Density measures the "concentration" of probability around a given point.

Characterisations of Continuous Distributions



Characterisations of Continuous Distributions



Characterisations of Continuous Distributions

In the illustration X is a random variable taking values in the interval $(a, b]$, and the pdf $f_X(x)$ is non-zero only in (a, b).

More generally we have, for a variable taking values on the whole real line ($\mathbb{R}$)

- the **fundamental theorem of integral calculus** yields

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt,$$

  the area under the CDF between $-\infty$ and x

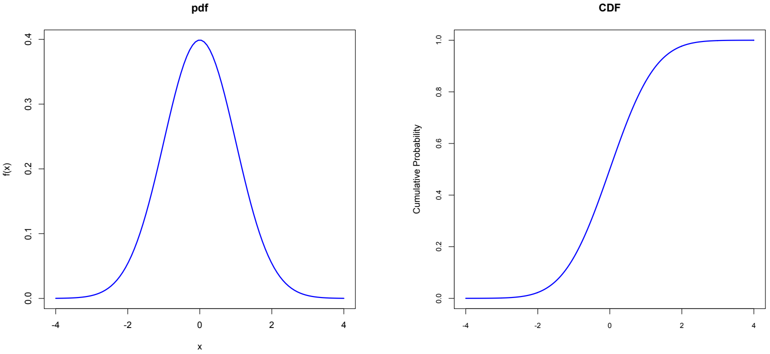
- or in terms of derivative

$$f_X(x) = \frac{dF_X(x)}{dx}$$

  for all x, the derivative of the CDF

Characterisations of Continuous Distributions

Most of the pdfs that we are going to consider are bell-shaped. So, typically, we will have



Outline

Recall that for **Discrete** random variables:

$$E[X] = \sum_i x_i p_i$$

The Expectation results from **summing** the **product of** x_i **and** $p_i = P(X = x_i)$, for all **possible values** x_i

Distributional Summaries

Expectation

Recall that for **Discrete** random variables:

$$E[X] = \sum_i x_i p_i$$

The Expectation results from **summing** the **product of** x_i **and** $p_i = P(X = x_i)$, for all **possible values** x_i

Definition

For **continuous** random variables, we obtain the Expectation via integration:

$$E[X] = \int_a^b x\, f_X(x)\, dx$$

The Expectation of X results from **integrating the product of** x **and its pdf** $f_X(x)$ over the **range of possible values of** x

Recall that, for **Discrete** random variables:

$$Var\left(X\right) = \sum_i \left(x_i - E\left[X\right]\right)^2 P\left(X = x_i\right)$$

Definition

Similarly, for **Continuous** random variables, we use integration[a]:

$$Var\left(X\right) = \int_a^b \left(x - E\left[X\right]\right)^2 \; f_X\left(x\right) dx$$

[a] Intuitively, we replace the sum ($\sum$) by its *continuous* counterpart, namely the integral ($\int$).

As with discrete random variables, the following properties hold when X is a continuous random variable and c is any real number (namely, $c \in \mathbb{R}$):

1. $E[cX] = cE[X]$

As with discrete random variables, the following properties hold when X is a continuous random variable and c is any real number (namely, $c \in \mathbb{R}$):

1. $E[cX] = cE[X]$
2. $E[c + X] = c + E[X]$

Distributional Summaries

As with discrete random variables, the following properties hold when X is a continuous random variable and c is any real number (namely, $c \in \mathbb{R}$):

1. $E[cX] = cE[X]$
2. $E[c + X] = c + E[X]$
3. $Var(cX) = c^2 Var(X)$

Distributional Summaries

As with discrete random variables, the following properties hold when X is a continuous random variable and c is any real number (namely, $c \in \mathbb{R}$):

1. $E[cX] = cE[X]$
2. $E[c + X] = c + E[X]$
3. $Var(cX) = c^2 Var(X)$
4. $Var(c + X) = Var(X)$

Let us consider, for instance, the following proofs for first two properties

$$\begin{aligned} E\left[cX\right] &= \int (cx)\, f_X(x)\, dx \\ &= c \int x f_X(x)\, dx \\ &= cE\left[X\right]. \end{aligned}$$

$$\begin{aligned} E\left[c+X\right] &= \int (c+x)\, f_X(x)\, dx \\ &= \int c f_X(x)\, dx + \int x f_X(x)\, dx \\ &= c \times 1 + E\left[X\right] \\ &= c + E\left[X\right]. \end{aligned}$$

Definition (Mode)

The Mode of a continuous random variable having density $f_X(x)$ is the **value of** x **for which** $f(x)$ **attains its maximum.**

$$\text{Mode}(X) = \text{argmax}_x\{f_X(x)\}$$

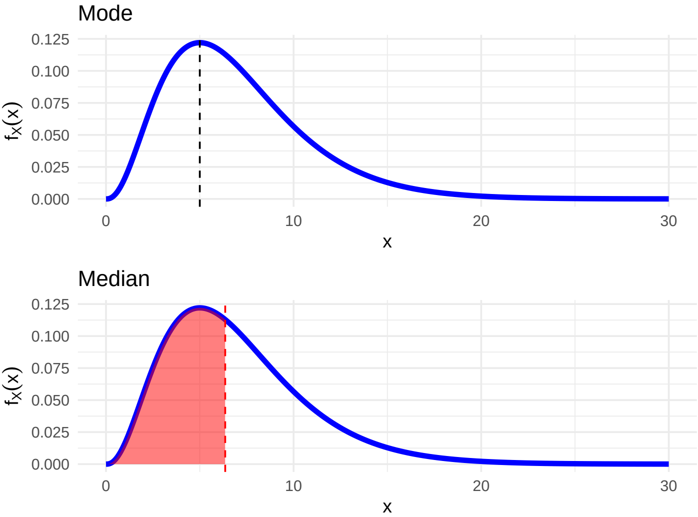
Definition (Median)

The Median of a continuous random variable having distribution function $F_X(x)$ is the value m such that $F(m) = 1/2$

$$\text{Median}(X) = \text{arg}_m \left\{ P(X \leq m) = F_X(m) = \frac{1}{2} \right\}$$

Distributional Summaries

Mode and Median



Outline

Some Important Continuous Distributions

- Continuous Uniform
- Normal
- Chi-squared
- Student's t
- F
- Lognormal
- Exponential
- ...and more

Outline

Continuous Uniform Distribution

Definition

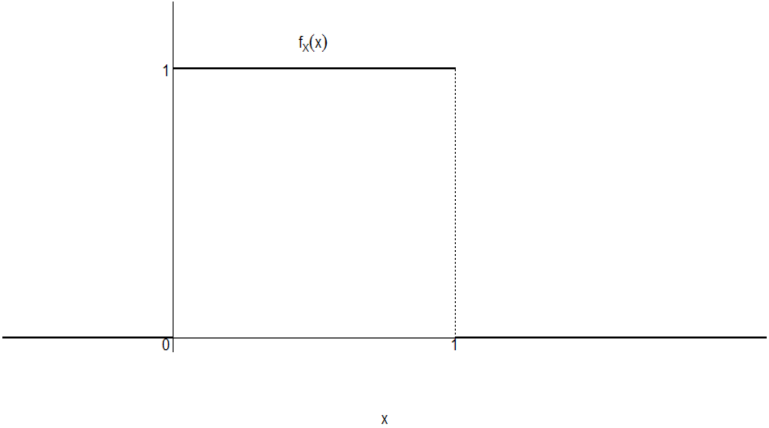
We say X has a continuous **uniform** distribution over the interval $[a, b]$, denoted as $X \sim \mathcal{U}(a, b)$, when the CDF and pdf are given by

$$F_X(x) = \begin{cases} 0, & x \leq a; \\ \frac{(x-a)}{(b-a)}, & a < x \leq b; \\ 1, & x > b. \end{cases} \quad \text{and } f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{for } a < x < b \\ 0, & \text{otherwise} \end{cases},$$

respectively.

Continuous Uniform Distribution

As a graphical illustration, let us consider the case when $a = 0$ and $b = 1$. So, we have:



The expected value of X is

$$
\begin{aligned}
E[X] &= \int_a^b \frac{x}{(b-a)} dx \\
&= \left. \frac{x^2}{2(b-a)} \right|_a^b \\
&= \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} \\
&= \frac{a+b}{2}
\end{aligned}
$$

Example

When $a = 0$ and $b = 1$, then $E[X] = \frac{1}{2}$.

Continuous Uniform Distribution

The variance of X is:

$$\begin{aligned} Var(X) &= \int_a^b \left(x - \left(\frac{a+b}{2}\right)\right)^2 \frac{1}{b-a} dx \\ &= E\left[X^2\right] - E\left[X\right]^2 \end{aligned}$$

We know the second term

$$E\left[X\right]^2 = \left(\frac{a+b}{2}\right)^2,$$

so we've just to work out

$$\begin{aligned} E\left[X^2\right] &= \int_a^b \frac{x^2}{b-a} dx = \left.\frac{x^3}{3(b-a)}\right|_a^b \\ &= \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)\left(ab + a^2 + b^2\right)}{3(b-a)} \\ &= \frac{\left(ab + a^2 + b^2\right)}{3}. \end{aligned}$$

Putting together, we get that the variance of X:

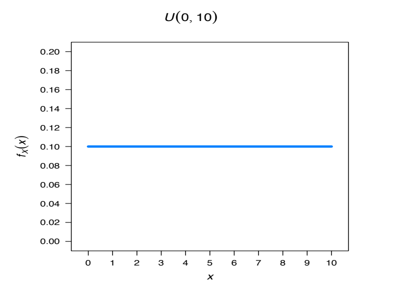
$$
\begin{aligned}
Var\left(X\right) &= \frac{\left(ab + a^2 + b^2\right)}{3} - \left(\frac{a+b}{2}\right)^2 \\
&= \frac{1}{12}\left(a-b\right)^2
\end{aligned}
$$

Example (continued)

When $a = 0$ and $b = 1$, then $Var\left(X\right) = \frac{1}{12}$.

Example

Let $X \sim \mathcal{U}(0, 10)$. Then its pdf is $f_X(x) = 1/10 = 0.1$ for $x \in [0, 10]$ and zero otherwise. The pdf plot is:

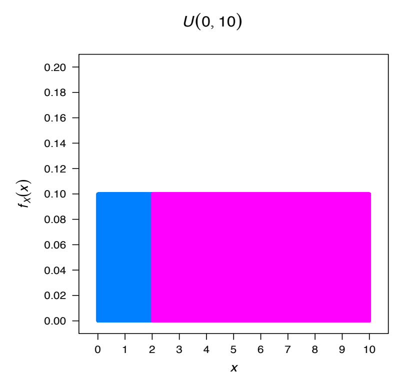


Example (continued)

$$
\begin{aligned}
P(0 \leq X \leq 1) &= \int_0^1 0.1dx & &= 0.1 \cdot x\Big|_{x=0}^{x=1} \\
&= 0.1(1-0) & &= 0.1 \\
\\
P(0 \leq X \leq 2) &= 2 \cdot 0.1 & &= 0.2 \\
\\
P(2 \leq X \leq 4) &= P(0 \leq X \leq 2) & &= 0.2 \\
\\
P(X \geq 2) &= P(2 \leq X \leq 10) & &= 8 \cdot 0.1 = 0.8
\end{aligned}
$$

Example (continued)

...and for $P(X \geq 2)$,



Outline

The Normal distribution was “discovered” in the eighteenth century when scientists observed an astonishing degree of regularity in the behavior of measurement errors.

They found that the patterns (distributions) that they observed, and which they attributed to chance, could be closely approximated by continuous curves which they christened the “normal curve of errors”.

The mathematical properties of these curves were first studied by

- Abraham de Moivre (1667-1745),

The Normal distribution was “discovered” in the eighteenth century when scientists observed an astonishing degree of regularity in the behavior of measurement errors.

They found that the patterns (distributions) that they observed, and which they attributed to chance, could be closely approximated by continuous curves which they christened the “normal curve of errors”.

The mathematical properties of these curves were first studied by

- Abraham de Moivre (1667-1745),
- Pierre Laplace (1749-1827), and then

The Normal distribution was "discovered" in the eighteenth century when scientists observed an astonishing degree of regularity in the behavior of measurement errors.

They found that the patterns (distributions) that they observed, and which they attributed to chance, could be closely approximated by continuous curves which they christened the "normal curve of errors".

The mathematical properties of these curves were first studied by

- Abraham de Moivre (1667-1745),
- Pierre Laplace (1749-1827), and then
- Karl Gauss (1777-1855), who also lent his name to the distribution.

Gaussian or "Normal" Distribution (Time Allowing)

Definition

A variable X is said to have a **Gaussian** or **normal** distribution, with mean μ and variance σ^2, if its pdf is given by

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

For simplicity we denote this by writing $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$.

Gaussian or "Normal" Distribution (Time Allowing)

Definition

A variable X is said to have a **Gaussian** or **normal** distribution, with mean μ and variance σ^2, if its pdf is given by

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

For simplicity we denote this by writing $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$.

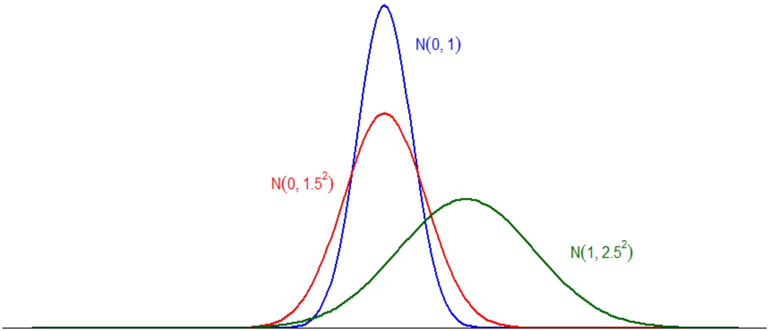
Remark

- *A normal distribution is completely characterised by its mean μ and its variance σ^2. Infinitely many different normal distributions are obtained by varying the parameters μ and σ^2.*
- *A normal random variable X can take any value $x \in \mathbb{R}$.*

Gaussian or “Normal” Distribution (Time Allowing)

The pdf of the normal distribution is

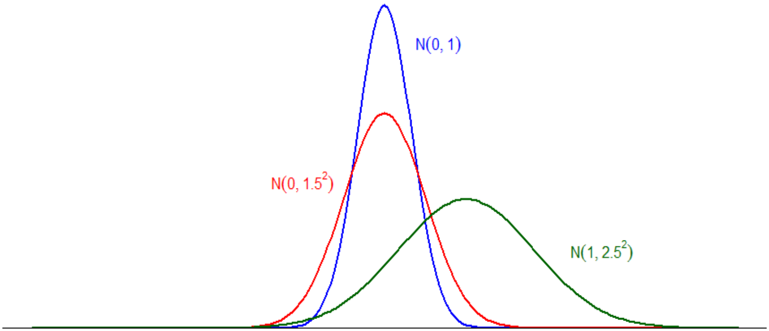
- ‘bell-shaped’



Gaussian or “Normal” Distribution (Time Allowing)

The pdf of the normal distribution is

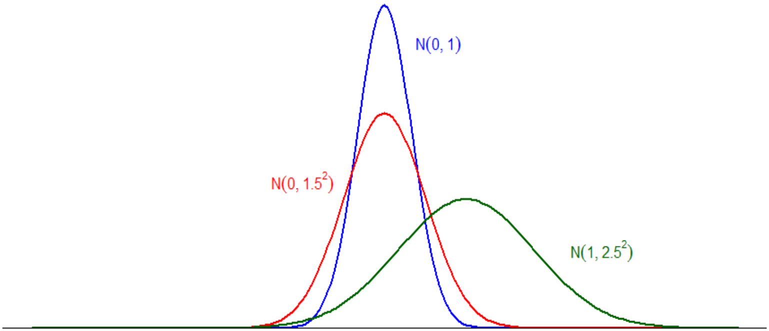
- ‘bell-shaped’
- symmetric



Gaussian or “Normal” Distribution (Time Allowing)

The pdf of the normal distribution is

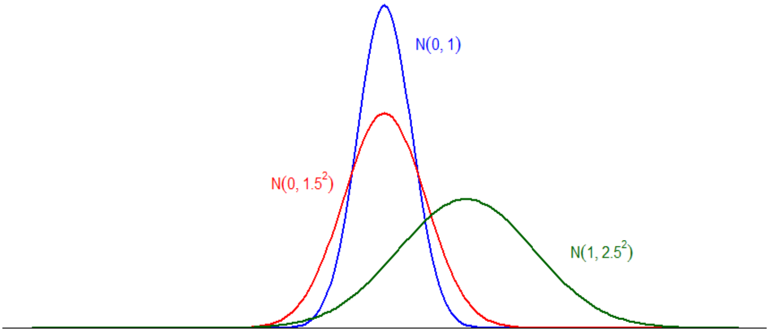
- ‘bell-shaped’
- symmetric
- unimodal



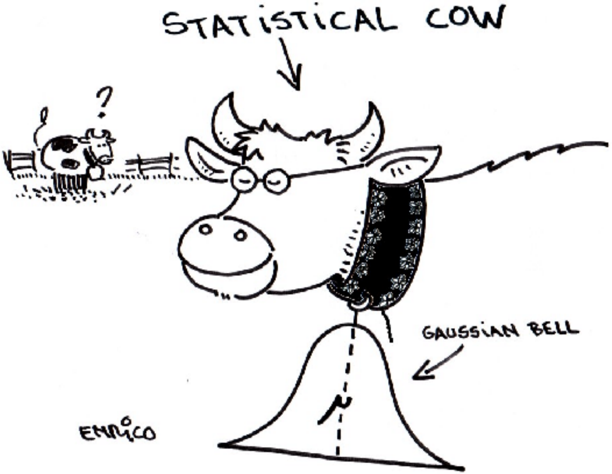
Gaussian or "Normal" Distribution (Time Allowing)

The pdf of the normal distribution is

- 'bell-shaped'
- symmetric
- unimodal
- the mean, median and mode are all equal.



Gaussian or “Normal” Distribution (Time Allowing)



Remark

First let us verify that $\phi_{(\mu,\sigma)}(x)$ can serve as a genuine density function. Integrating with respect to x using integration by substitution we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} \phi_{(\mu,\sigma)}(x)dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\}dz\end{aligned}$$

where $z = (x-\mu)/\sigma$. But the second integral on the right hand side equals

$$\frac{2}{\sqrt{2\pi}} \underbrace{\int_0^{\infty} \exp\{-z^2/2\}dz}_{=\sqrt{2\pi}/2}$$

which is **a known standard integral***.*

Gaussian or "Normal" Distribution (Time Allowing)

- The function $\phi_{(\mu,\sigma)}(x)$ does indeed define the pdf of a random variable with a mean of μ and a variance of σ^2.

Gaussian or "Normal" Distribution (Time Allowing)

- The function $\phi_{(\mu,\sigma)}(x)$ does indeed define the pdf of a random variable with a mean of μ and a variance of σ^2.
- This was established by **transforming from X to Z via the substitution** $Z = (X - \mu)/\sigma$. Such a variable is said to be **standardised**. Note also that the resulting integrand

$$\frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\} = \phi_{(0,1)}(z)\,,$$

is the pdf of a random variable $Z \sim \mathcal{N}(0, 1)$.

Gaussian or "Normal" Distribution (Time Allowing)

- The function $\phi_{(\mu,\sigma)}(x)$ does indeed define the pdf of a random variable with a mean of μ and a variance of σ^2.
- This was established by **transforming from X to Z via the substitution** $Z = (X - \mu)/\sigma$. Such a variable is said to be **standardised**. Note also that the resulting integrand

$$\frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\} = \phi_{(0,1)}(z)\,,$$

  is the pdf of a random variable $Z \sim \mathcal{N}(0,1)$.
- If $Z \sim \mathcal{N}(0,1)$ then Z is called a Standard Normal Random Variable, because $\mathrm{E}[Z] = 0$ and $\mathrm{Var}(Z) = 1$

Gaussian or "Normal" Distribution (Time Allowing)

- The function $\phi_{(\mu,\sigma)}(x)$ does indeed define the pdf of a random variable with a mean of μ and a variance of σ^2.
- This was established by **transforming from X to Z via the substitution** $Z = (X - \mu)/\sigma$. Such a variable is said to be **standardised**. Note also that the resulting integrand

$$\frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\} = \phi_{(0,1)}(z)\,,$$

is the pdf of a random variable $Z \sim \mathcal{N}(0, 1)$.

- If $Z \sim \mathcal{N}(0, 1)$ then Z is called a Standard Normal Random Variable, because $\mathrm{E}[Z] = 0$ and $\mathrm{Var}(Z) = 1$
- Because of the special role that the standard normal distribution has in calculations involving the normal distribution its pdf is given the special notation

$$\phi(z) = \phi_{(0,1)}(z).$$

Gaussian or "Normal" Distribution (Time Allowing)

The basic feature that underlies calculations involving the Normal distribution:

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

The basic feature that underlies calculations involving the Normal distribution:

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

  - We can always transform from X to Z by 'shifting' and 're-scaling':

$$Z = \frac{X-\mu}{\sigma} \text{ (for the random variable)} \quad \text{and} \quad z = \frac{x-\mu}{\sigma} \text{ (for its values)},$$

Gaussian or "Normal" Distribution (Time Allowing)

The basic feature that underlies calculations involving the Normal distribution:

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

  - We can always transform from X to Z by 'shifting' and 're-scaling':

    $$Z = \frac{X-\mu}{\sigma} \text{ (for the random variable)} \quad \text{and} \quad z = \frac{x-\mu}{\sigma} \text{ (for its values),}$$

  - and return back to X by a 're-scaling' and 'shifting':

    $$X = \sigma Z + \mu \text{ (for the random variable)} \quad \text{and} \quad x = \sigma z + \mu \text{ (for its values).}$$

The basic feature that underlies calculations involving the Normal distribution:

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

    - We can always transform from X to Z by 'shifting' and 're-scaling':

    $$Z = \frac{X-\mu}{\sigma} \text{ (for the random variable)} \quad \text{and} \quad z = \frac{x-\mu}{\sigma} \text{ (for its values)},$$

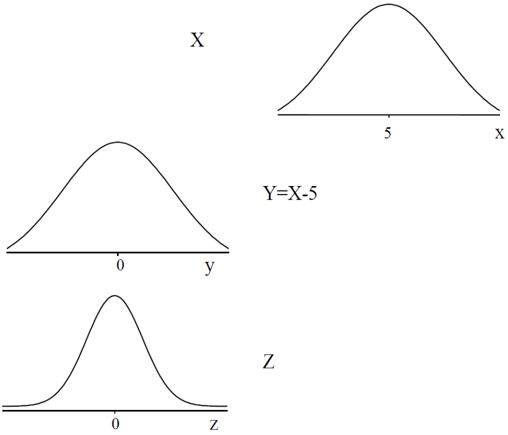
    - and return back to X by a 're-scaling' and 'shifting':

    $$X = \sigma Z + \mu \text{ (for the random variable)} \quad \text{and} \quad x = \sigma z + \mu \text{ (for its values)}.$$

- Thus statements about a Normal random variable can always be translated into equivalent statements about a standard Normal random variable, and vice versa.

Graphically

Start from $X \sim \mathcal{N}(5,3)$; then define $Y = X - 5$, which is a recentered/shifted X (it's centered at 0 and has the same variance as X); finally define Z, which is a recentered/shifted and rescaled X (it's centered at 0 and has unit variance).



Analitically

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, the CDF is given by

$$\Phi_{(\mu,\sigma)}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(t-\mu)^2\right\} dt$$

Analitically

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, the CDF is given by

$$\Phi_{(\mu,\sigma)}\left(x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(t-\mu\right)^2\right\} dt$$

- To calculate $\Phi_{(\mu,\sigma)}\left(x\right) = P(\{X \leq x\})$ we use integration by substitution, once again, to give

$$\begin{aligned} P(\{X \leq x\}) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt \\ &= \int_{-\infty}^{z} \phi(s) ds \\ &= P(\{Z \leq z\}) \end{aligned}$$

where $z = (x-\mu)/\sigma$, $s = (t-\mu)/\sigma$ and $ds = dt/\sigma$.

Analitically

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, the CDF is given by

$$\Phi_{(\mu,\sigma)}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(t-\mu)^2\right\} dt$$

- To calculate $\Phi_{(\mu,\sigma)}(x) = P(\{X \leq x\})$ we use integration by substitution, once again, to give

$$\begin{aligned} P(\{X \leq x\}) &= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} dt \\ &= \int_{-\infty}^{z} \phi(s) ds \\ &= P(\{Z \leq z\}) \end{aligned}$$

where $z = (x-\mu)/\sigma$, $s = (t-\mu)/\sigma$ and $ds = dt/\sigma$.

- The required probability has been mapped into a corresponding probability for a standard Normal random variable.

- We can evaluate the probabilities

$$P(\{Z \leq z\}) = \Phi(z) = \int_{-\infty}^{z} \phi(s)ds$$

either directly using a computer or indirectly via Standard Normal Tables.

- We can evaluate the probabilities

$$P(\{Z \leq z\}) = \Phi(z) = \int_{-\infty}^{z} \phi(s)ds$$

  either directly using a computer or indirectly via Standard Normal Tables.
- Standard Normal Tables give values of the integral $\Phi(z)$ for various values of $z \geq 0$. (The tables are themselves calculated using a computer, of course.)

- We can evaluate the probabilities

$$P(\{Z \leq z\}) = \Phi(z) = \int_{-\infty}^{z} \phi(s)ds$$

either directly using a computer or indirectly via Standard Normal Tables.

- Standard Normal Tables give values of the integral $\Phi(z)$ for various values of $z \geq 0$. (The tables are themselves calculated using a computer, of course.)
- For negative values of z the symmetry property of $\phi(z)$ (*i.e.* $\phi(z) = \phi(-z)$) tells us that

$$\Phi(-z) = 1 - \Phi(z)\,.$$

- We can evaluate the probabilities

$$P(\{Z \leq z\}) = \Phi(z) = \int_{-\infty}^{z} \phi(s)ds$$

either directly using a computer or indirectly via Standard Normal Tables.

- Standard Normal Tables give values of the integral $\Phi(z)$ for various values of $z \geq 0$. (The tables are themselves calculated using a computer, of course.)
- For negative values of z the symmetry property of $\phi(z)$ (*i.e.* $\phi(z) = \phi(-z)$) tells us that

$$\Phi(-z) = 1 - \Phi(z)\,.$$

- Similarly, if $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ then

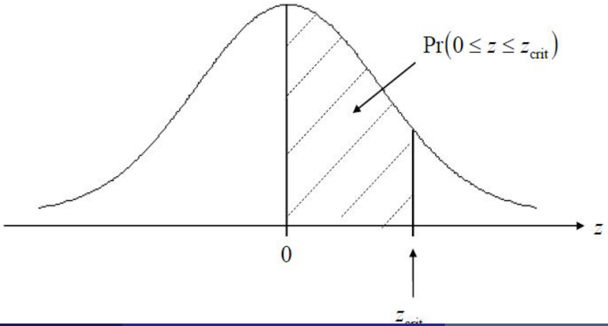
$$\begin{aligned} P(\{x_1 < X \leq x_2\}) &= P(\{z_1 < Z \leq z_2\}) \\ &= \Phi(z_2) - \Phi(z_1) \end{aligned}$$

where $z_1 = (x_1 - \mu)/\sigma$ and $z_2 = (x_2 - \mu)/\sigma$.

- Standard Normal Tables give values of the standard normal integral $\Phi(z)$ for various values of $z \geq 0$. Values for negative z are obtained via symmetry.

STATISTICAL TABLES

TABLE 1: AREAS UNDER THE STANDARDIZED NORMAL DISTRIBUTION

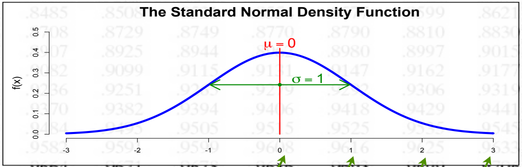


Gaussian or “Normal” Distribution (Time Allowing)

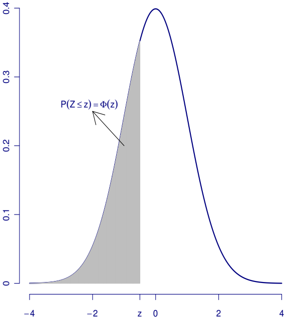
TABLE 5.1: AREA $\Phi(x)$ UNDER THE STANDARD NORMAL CURVE TO THE LEFT OF X

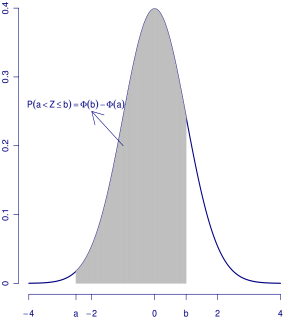
X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	[illegible]	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	[illegible]	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	[illegible]	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	[illegible]	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	.8621
1.1	.8643	.8665	.8686	[illegible]	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	[illegible]	.8925	.8944	[illegible]	.8980	.8997	.9015
1.3	.9032	.9049	.9066	[illegible]	.9099	.9115	[illegible]	[illegible]	.9162	.9177
1.4	.9192	.9207	.9222	[illegible]	.9251	[illegible]	[illegible]	[illegible]	.9306	.9319
1.5	.9332	.9345	.9357	[illegible]	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	[illegible]	[illegible]	.9505	[illegible]	.9525	[illegible]	.9545
1.7	.9554	.9564	.9573	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995

$\Phi(.46)$



.... and you can use these tables to compute integrals/probabilities of the type:





Example (Prob of Z)

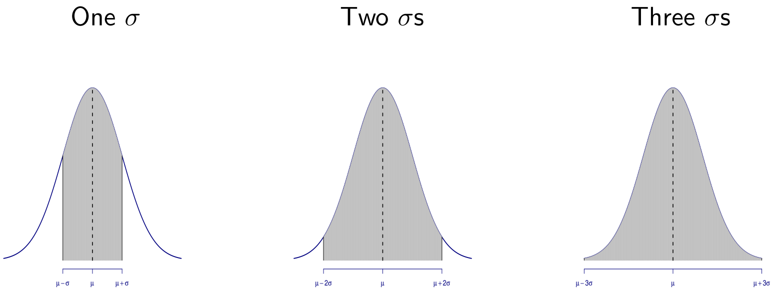
$$
\begin{aligned}
P(\{Z \leq 1\}) &\approx 0.8413 \\
P(\{Z \leq 1.96\}) &\approx 0.9750 \\
P(\{Z \geq 1.96\}) &= 1 - P(\{Z \leq 1.96\}) \approx 1 - 0.9750 = 0.0250 \\
P(\{Z \geq -1\}) &= P(\{Z \leq 1\}) \approx 0.8413 \\
P(\{Z \leq -1.5\}) &= P(\{Z \geq 1.5\}) = 1 - P(\{Z \leq 1.5\}) \approx 1 - 0.9332 = 0.0668
\end{aligned}
$$

Example (continued)

$$P(\{0.64 \leq Z \leq 1.96\}) =$$
$$P(\{Z \leq 1.96\}) - P(\{Z \leq 0.64\})$$
$$\approx 0.9750 - 0.7389 = 0.2361$$

$$P(\{-0.64 \leq Z \leq 1.96\})$$
$$= P(\{Z \leq 1.96\}) - P(\{Z \leq -0.64\})$$
$$= P(\{Z \leq 1.96\}) - (1 - P(\{Z \leq 0.64\}))$$
$$\approx 0.9750 - (1 - 0.7389) = 0.7139$$

$$P(\{-1.96 \leq Z \leq -0.64\})$$
$$= P(\{0.64 \leq Z \leq 1.96\})$$
$$\approx 0.2361$$



The shaded areas under the pdfs are (approximately) equivalent to 0.683, 0.954 and 0.997, respectively. So we state the following ....

... rule '68 – 95 – 99.7':

If X is a Normal random variable, $X \sim \mathcal{N}(\mu, \sigma^2)$, its realization has approximately a probability of

- 68 % of being in the interval $[\mu - \sigma,\ \mu + \sigma]$;
- 95 % of being in the interval $[\mu - 2\,\sigma,\ \mu + 2\,\sigma]$;
- 99.7 % of being in the interval $[\mu - 3\,\sigma,\ \mu + 3\,\sigma]$.

Gaussian or "Normal" Distribution (Time Allowing)

Some properties of the Normal distribution

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$

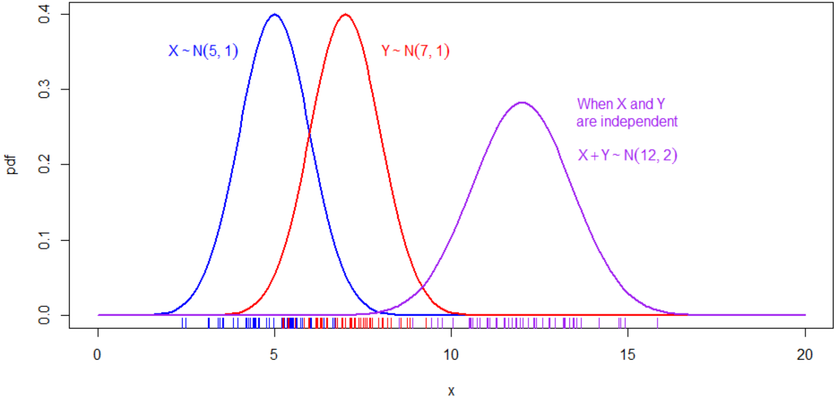
$$E\left[X\right] = \mu \text{ and } Var\left(X\right) = \sigma^2.$$

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$
$$E\left[X\right] = \mu \text{ and } Var\left(X\right) = \sigma^2.$$
- If a is a number, then
$$\begin{aligned} X + a &\sim \mathcal{N}\left(\mu + a, \sigma^2\right) \\ aX &\sim \mathcal{N}\left(a\mu, a^2\sigma^2\right). \end{aligned}$$

- For $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$
$$E[X] = \mu \text{ and } Var(X) = \sigma^2.$$
- If a is a number, then
$$\begin{aligned} X + a &\sim \mathcal{N}\left(\mu + a, \sigma^2\right) \\ aX &\sim \mathcal{N}\left(a\mu, a^2\sigma^2\right). \end{aligned}$$
- If $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ and $Y \sim \mathcal{N}\left(\alpha, \delta^2\right)$, and X and Y are **independent** then
$$X + Y \sim \mathcal{N}\left(\mu + \alpha, \sigma^2 + \delta^2\right).$$

Gaussian or "Normal" Distribution (Time Allowing)

The sum of two independent Normals



Locations of $n = 30$ sampled values of X, Y, and $X + Y$ shown as tick marks under each respective density.

Example

On the highway A2 (in the Luzern area), the speed is limited to 80 *km/h. A radar measures the speeds of all the cars. Assuming that the registered speeds are distributed according to a Normal law with mean* 72 *km/h and standard error* 8 *km/h:*

1. *what is the proportion of the drivers who will have to pay a penalty for high speed?*
2. *knowing that in addition to the penalty, a speed higher than* 30 *km/h (over the max allowed speed) implies a withdrawal of the driving license, what is the proportion of the drivers who will lose their driving license among those who will have a to pay a fine?*

Example (continued)

Let X be the random variable expressing the registered speed: $X \sim \mathcal{N}(72, 64)$.

1. Since a driver has to pay if its speed is above 80 km/h, the proportion of drivers paying a penalty is expressed through $P(X > 80)$:

$$P(X > 80) = P\left(Z > \frac{80 - 72}{8}\right) = 1 - \Phi(1) \simeq 16\%$$

where $Z \sim \mathcal{N}(0, 1)$.

Example (continued)

Let X be the random variable expressing the registered speed: $X \sim \mathcal{N}(72, 64)$.

1. Since a driver has to pay if its speed is above 80 km/h, the proportion of drivers paying a penalty is expressed through $P(X > 80)$:

$$P(X > 80) = P\left(Z > \frac{80 - 72}{8}\right) = 1 - \Phi(1) \simeq 16\%$$

   where $Z \sim \mathcal{N}(0, 1)$.

2. We are looking for the conditional probability of a recorded speed greater than 110 given that the driver has had already to pay a fine:

$$\begin{aligned} P(X > 110 | X > 80) &= \frac{P(\{X > 110\} \bigcap \{X > 80\})}{P(X > 80)} \\ &= \frac{P(X > 110)}{P(X > 80)} = \frac{1 - \Phi((110 - 72)/8)}{1 - \Phi(1)} \approx \frac{0}{16\%} \simeq 0. \end{aligned}$$

Wrap-up

It is important to keep in mind and review:

- The definition and interpretation of CDF and PDF.
- The relationship between CDF and PDF.
    - CDF results from integration of PDF.
    - PDF is the derivative of CDF.
- Computation of $F_X(x)$, $E[X]$ via integration (when possible).
- The properties of the Uniform distribution
- "Centering and scaling" to compute the CDF of the Normal.

Thank You for your Attention!

Thank You for your Attention!

“See you” Next Week