Lecture IV:
Famous Continuous Distributions

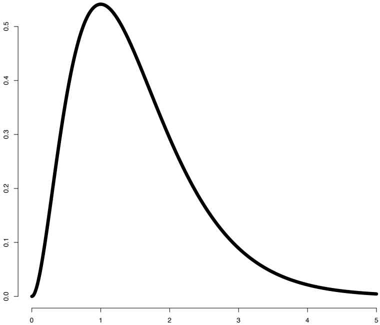
Carlo Cavicchia

carlo.cavicchia@uniroma1.it

Continuous Distributions:
a small recap

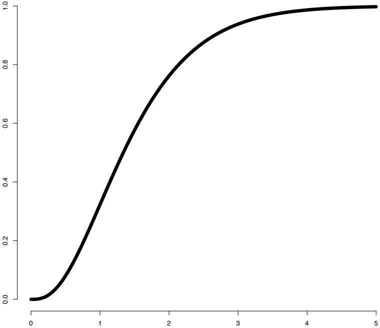
Probability density function

- $f_X(x) \geq 0$
- $f_X(x)$ **needs not** be ≤ 1
- $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$



Cumulative distribution function

- $0 \leq F(x) \leq 1$
- F is *non-decreasing*
- F is *right continuous*



Continuous Uniform Distribution

the intuition

The Continuous Uniform Distribution can be used to model phenomena that

- A random variable X is **uniformly distributed** between a and b, if X takes value in any interval of a given size with equal probability.

  Discrete case: it takes any value in the support with equal probability

- the probability of X being in an interval, is proportional to the length of the interval.

  Discrete case: probability of a set is proportional to its size

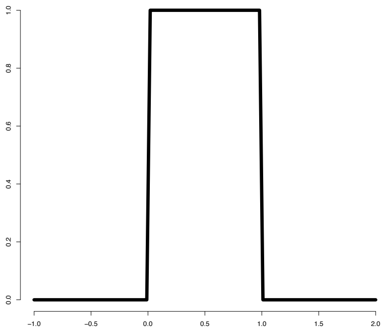
Example: the arrival of the bus 20 between the moment you get to the bus stop and midnight.

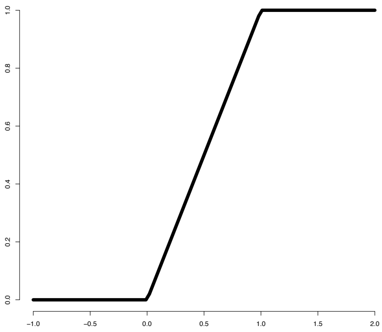
Continuous Uniform Distribution
the formalization

$X \sim \mathsf{Unif}(a, b)$

$$f_X(x) = \frac{1}{b-a}$$

$$F_X(x) = \frac{x-a}{b-a}$$





Uniform c.d.f

how to use it

- In the case of a Uniform random variable there is a closed form (& easy to derive) expression for the c.d.f.:

$$
\begin{aligned}
F_X(x) &= \int_a^x \frac{1}{b-a} dt = \frac{1}{b-a} \left(t|_a^x \right) \\
&= \frac{x-a}{b-a} \qquad a \leq x \leq b
\end{aligned}
$$

- It is trivial to see that the probability of a set only depends on its size:

$$
\begin{aligned}
P(X \in [a_1, b_1]) &= F_X(b_1) - F_X(a_1) \\
&= \frac{b_1 - a}{b-a} - \frac{a_1 - a}{b-a} \\
&= \frac{b_1 - a_1}{b-a}
\end{aligned}
$$

Mean and Expected Value

do try this at home

- Expected Value of $X \sim \text{Unif}(a, b)$

$$\mathbb{E}[X] = \frac{a+b}{2}$$

Since it is a *location/center* summary, the expected value depends on the specific values the random variable assumes.

- Variance of $X \sim \text{Unif}(a, b)$

$$\mathbb{V}[X] = \frac{(b-a)^2}{12}$$

Since it is a *scale/dispersion* summary, the variance depends only on the size of the support.

Example

As the name suggest, a pay-per-kilo clothes shop (something like Pifebo) charges the customer based on the weight of what they are buying.

Empirical evidence suggest that a client typically buys between 200 and 800 gr of clothes.

- Probability Density Function:

$$f_X(x) = \begin{cases} \frac{1}{600} & 200 \leq x \leq 800 \\ 0 & \text{otherwise} \end{cases}$$

Exercise
your turn!

- What is the average amount of clothes bought?
- What is its variance?
- What is the probability that a customer buys less than 300 gr of clothes?

Exponential Distribution

the intuition

A random variable X is said to have an **Exponential Distribution** with parameter $\lambda > 0$, if its probability distribution can be written as

$$f_X(x) = \lambda e^{-\lambda x} \qquad x \geq 0$$

The intuition behind an Exponential random variable is that the **larger** is a value, the **less likely** it is.

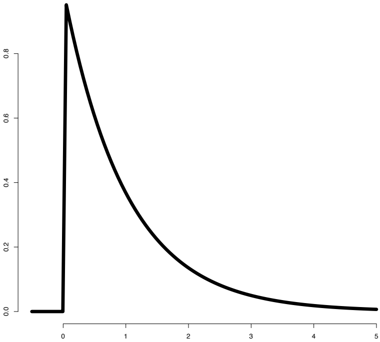
The Exponential is typically used to model **time until some specific event occurs**, and its parameter λ affects the mean time between events.

Examples: the amount of time until an earthquake occurs, the amount of money customers spend in one trip to the supermarket, the value of the change that you have in your pocket

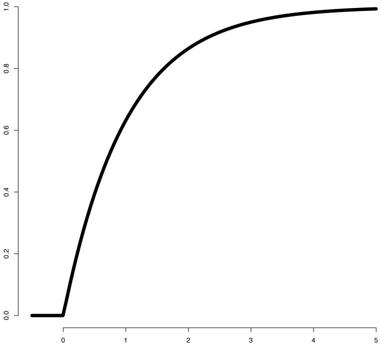
Exponential distribution
the formalization

$$X \sim \mathsf{Exp}(\lambda), \qquad \lambda > 0, \qquad x \geq 0$$

$$f_X(x) = \lambda e^{-\lambda x}$$



$$F_X(x) = 1 - e^{-\lambda x}$$



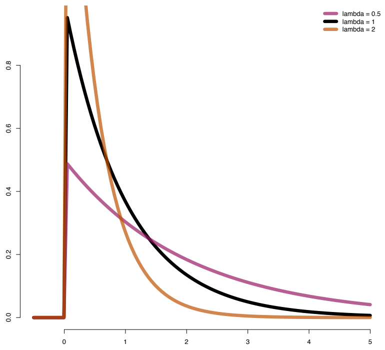
Exponential distribution

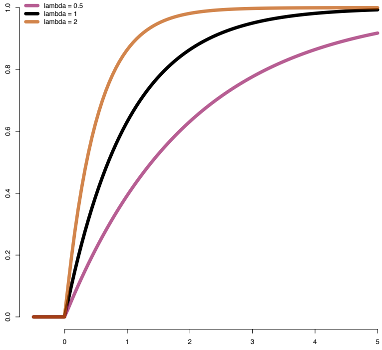
the formalization

$$X \sim \text{Exp}(\lambda), \qquad \lambda > 0, \qquad x \geq 0$$

$$f_X(x) = \lambda e^{-\lambda x}$$

$$F_X(x) = 1 - e^{-\lambda x}$$

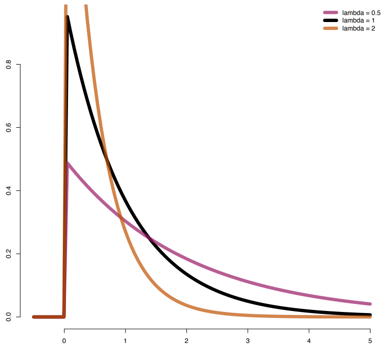


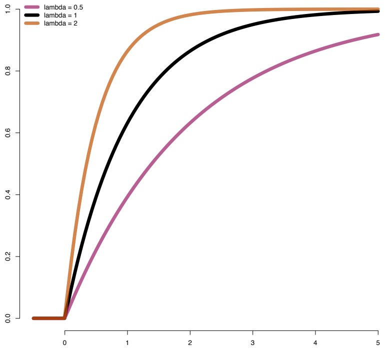


Expected Value and Variance

$$X \sim \mathsf{Exp}(\lambda), \qquad \lambda > 0$$

$$\mathbb{E}[X] = \frac{1}{\lambda} \qquad\qquad \mathbb{V}[X] = \frac{1}{\lambda^2}$$





Properties

- The exponential is memoryless

$$P(T \geq t) = P(T \geq t + s | T \geq s)$$

- The exponential represent the waiting time between two Poisson events

Example[1/2]

As the building entrance door closes behind, Bob glances at his post-it note. It has the directions and address of the car dealer. Bob is finally ready to buy his first (used) car. He walks to the nearby bus stop jubilantly thinking he will seldom use the bus again. Bob is tired of the waiting. Throughout these years the one thing he could establish is that the average wait time for his inbound 105 at the Cross St @ Main St is 15 minutes.

- waiting time $= X \sim Exp(\lambda)$
- $\mathbb{E}[X] = \frac{1}{\lambda} = 15 \rightarrow \lambda = 0.066$

Example[2/2]

Bob gets to the bus shelter, greets the person next to him and thinks to himself “Hope the wait will not exceed 10 minutes today.”

- $P(X > 10) = e^{\lambda x} = 0.513$

Bob is visibly anxious. He turns his hand and looks at his wristwatch. “10 minutes. The wait won’t be much longer.”

- memoryless property: The probability that he waits for another ten minutes, given he already waited 10 minutes is also 0.513.

Normal Distribution

The **Normal** or **Gaussian** Distribution is the *queen* of the random variables, and this is because:

- it represents many natural and economic phenomena
- it approximates other distributions
- it is key to inference in sampling

A random variable $X \sim \text{Norm}(\mu, \sigma^2)$ has an interpretable parametrization:

$$\mu = \mathbb{E}[X] \qquad \sigma^2 = \mathbb{V}[X]$$

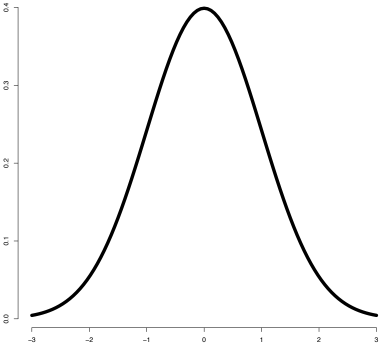
Normal Distribution

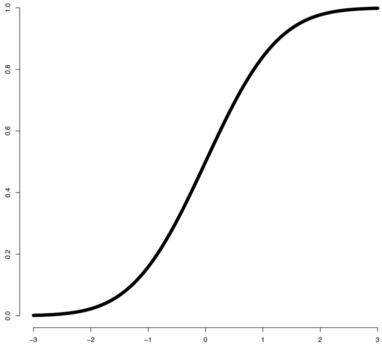
the formalization

$$X \sim \mathsf{Norm}(\mu, \sigma^2), \qquad \sigma^2 > 0, \mu \in \mathbb{R}$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt$$





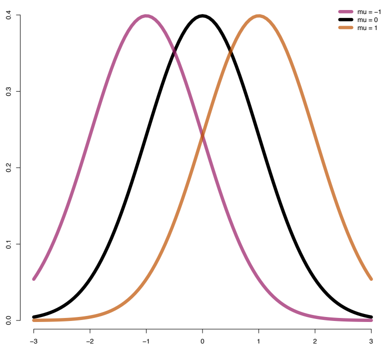
Normal Distribution

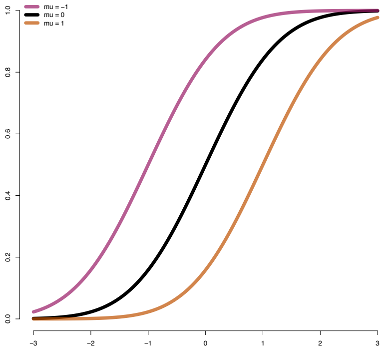
as the mean varies

$$X \sim \text{Norm}(\mu, \sigma^2), \qquad \sigma^2 > 0, \mu \in \mathbb{R}$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt$$





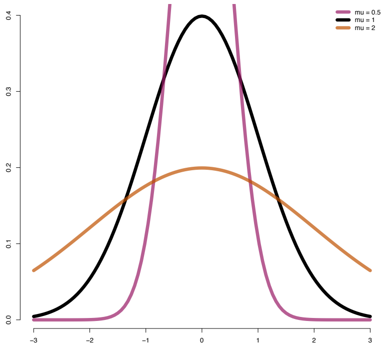
Normal Distribution

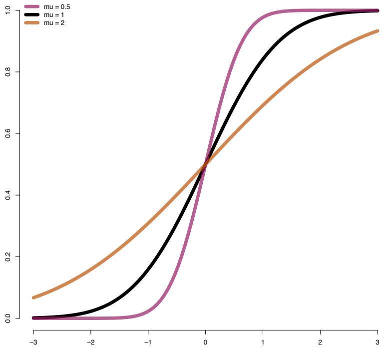
as the variance varies

$$X \sim \text{Norm}(\mu, \sigma^2), \qquad \sigma^2 > 0$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt$$





Properties

the Normal is a *stubborn* distribution

- a linear transformation of a Normal random variable is still a Normal random variable:

  $X \sim \text{Norm}(\mu, \sigma^2)$, if $Y = aX + b$, where $a, b \in \mathbb{R}$

$$Y \sim \text{Norm}(a\mu + b, a^2\sigma^2)$$

- a linear combination of Normal random variables is still a Normal random variable:

  $X_1, \dots, X_n$ independent random variables such that $X_i \sim N(\mu_i, \sigma_i^2)$ then

$$Y = \sum_{i=1}^{n} a_i X_i \sim \text{Norm}\left(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2\right),$$

Standard Normal

When $\mu = 0$ and $\sigma^2 = 1$, the random variable Norm$(0, 1)$ is called a **standard Normal** and it is denoted by Z.

Every Normal distribution can be turn into a standard Normal by means of **standardization**

If $X \sim$ Norm(μ, σ^2), then

$$Z = \frac{X - \mu}{\sigma} \sim \text{Norm}(0, 1)$$

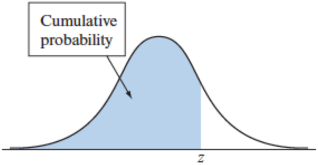
This is just a linear transformation of a Normal, so it is easy to show:

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{X - \mu}{\sigma}\right] = \frac{\mathbb{E}[X] - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0$$
$$\mathbb{V}[Z] = \mathbb{V}\left[\frac{X - \mu}{\sigma}\right] = \frac{\mathbb{V}[X]}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1$$

Tables of a standard Normal

what is the fuss about Standard Normal

Someone computed for you all the values of the cumulative distribution function of a Standard Normal and store them into **tables**.



Cumulative probability for *z* is the area under the standard normal curve to the left of *z*

Table A Standard Normal Cumulative Probabilities (*continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549

Toy Example

The time (in minutes) you need to solve the exercises I gave you, X, is Normally distributed with mean $\mu = 5$ and standard deviation $\sigma = 10$.

Formally $X \sim$ Norm$(5, (10)^2)$.

When I prepared this exercise at home, yesterday, it took me 6.2 minutes to solve it.

? What is the probability to find someone faster than me, i.e. $P(X \leq 6.2)$

$$
\begin{aligned}
P(X \leq 6.2) &= P\left(\frac{X-\mu}{\sigma} \leq \frac{6.2-5}{10}\right) \\
&= P(Z \leq 0.12) = 0.5478
\end{aligned}
$$

Exercise

The length of Black Mirror episodes (in minutes), is known to be Normally distributed with mean $\mu = 50$ and standard deviation $\sigma = 5$.

A new episode just got out:

- determine the probability that its length is exactly 50 minutes;
- determine the probability that its length is between 48 and 51 minutes;

A whole new season made of 8 episode is scheduled to be released next fall:

- determine the probability distribution of the length (in minutes) of the whole season;
- determine the expected length (in hours) of the whole season and its variance.

Central Limit Theorem

the intuition

Suppose you have $X_1, \dots, X_n$ random variables independent and with the same distribution.

Identical distribution implies that all the variables have the same expected value $\mu = \mathbb{E}[X_i]$ and variance $\sigma = \mathbb{V}[X_i]$

The average of this collection is also a random variable

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Even if we don't know the distribution of $\bar{X}$, the **Central Limit Theory** tell us that as $n \to \infty$

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \longrightarrow Z$$

Central Limit Theorem

the consequences

- if $X_1, \ldots, X_n$ are already Normals, then the result of the CLT is *exact*, that is, it works for any n

- even if we have no idea of what distribution generated the collection $X_1, \ldots, X_n$, we can *always* (albeit **asymptotically**) derive a distribution for its mean

- the CLT is very useful in statistical inference. We typically consider our data as realization of a collection of random variables $X_1, \ldots, X_n$ whose distribution we do not know; it is crucial to have a summary whose distribution we know in order to draw inferential conclusions.