Module 3a. Probability, The Central Limit Theorem, Confidence Intervals

Data Science & AI

Sabine De Vreese   Lieven Smits   Bert Van Vreckem
2023–2024

HO GENT

Contents

HO GENT

Learning Goals

- Probability distribution of a sample
- The normal and Student-*t* distribution
- Formulate the central limit theorem and explain its importance for statistical analysis
- Calculate confidence intervals

**HO
GENT**

Probability

HO
GENT

Probability

Informally, probabilities represent beliefs of how likely it is that a certain event will happen when performing a certain experiment.

In the frequentist view, **probability** represents the **relative frequency** of the occurrence of the event at hand (when performing a large number of independent experiments).

**HO
GENT**

Probability: (Simple) Examples

A fair six-sided die is thrown once.

- What is the probability of getting a "1"?
- What is the probability of throwing an even number?
- What is the probability of getting a number in the set {1, 2, 3, 4, 5, 6}?

HO GENT

Probability: (Simple) Examples

A fair six-sided die is thrown once.

- What is the probability of getting a "1"?
- What is the probability of throwing an even number?
- What is the probability of getting a number in the set {1, 2, 3, 4, 5, 6}?

Answers: 1/6, 1/2 and 1

HO GENT

Probability: Example

Two games of chance:

1. Bet on: at least one six when throwing a fair die 4 consecutive times.
   o You will *win* in the long run when playing this game.
2. Bet on: at least one "double six" when throwing two fair dice 24 times.
   o You will *lose* in the long run when playing this game, even though in some sense the two bets seem equivalent. (Probability divided by six but number of throws multiplied by six.)

HO GENT

Probability: Example

Win percentage for the first 100 games

Win percentage for the first 10000 games
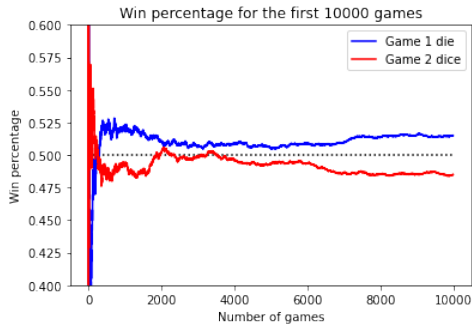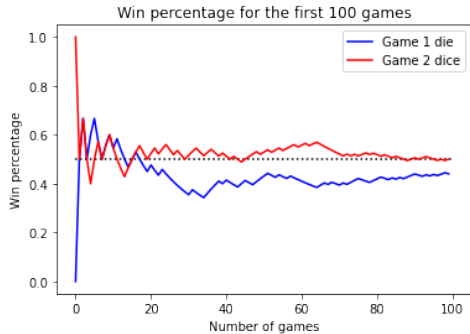
HO
GENT

Probability: Observations

- Probabilities are numbers assigned to sets.
- These sets are part of some all-encompassing set, the "universe", typically denoted Ω.
- The numbers (probabilities) assigned to the sets have to satisfy three simple rules (see later) in order to correspond to our intuition about how probabilities should behave.

**HO
GENT**

Example

For the first game: at least one six when throwing a fair die 4 consecutive times.

- What is the universe?

Example

For the first game: at least one six when throwing a fair die 4 consecutive times.

- What is the universe? Answer:

$$\Omega = \{(1, 1, 1, 1), (1, 1, 1, 2), (1, 1, 1, 3), \ldots, (6, 6, 6, 4), (6, 6, 6, 5), (6, 6, 6, 6)\}$$

**HO
GENT**

Example

For the first game: at least one six when throwing a fair die 4 consecutive times.

- What is the universe? Answer:

$$\Omega = \{(1,1,1,1),(1,1,1,2),(1,1,1,3),\ldots,(6,6,6,4),(6,6,6,5),(6,6,6,6)\}$$

- What set corresponds to the event: "you win" or "at least one 6"?

HO GENT

Example

For the first game: at least one six when throwing a fair die 4 consecutive times.

- What is the universe? Answer:

$$\Omega = \{(1,1,1,1),(1,1,1,2),(1,1,1,3),\dots,(6,6,6,4),(6,6,6,5),(6,6,6,6)\}$$

- What set corresponds to the event: "you win" or "at least one 6"? Answer

$$A = \{(6,1,1,1),(6,1,1,2),\dots,(5,5,5,6),(6,6,1,1),\dots$$
$$(6,6,5,5),(1,6,6,1),\dots,(6,6,6,6)\}$$

HO GENT

Axioms (Rules) of Probability

Axioms of Probability

1. **Probabilities are non-negative:** $P(A) \geq 0$ **for each** A.
2. **The universe has probability 1:** $P(\Omega) = 1$
3. *Sum rule*: **When** A **and** B **are** *disjoint* **events (i.e.** $A \cap B = \varnothing$**) then it holds that**

$$P(A \cup B) = P(A) + P(B)$$

HO
GENT

Properties of Probabilities

From the three axioms of probability, we can derive *all* properties of probabilities. Some important ones are listed below:

- **Complement rule**: for each *A* it holds that

$$P(\overline{A}) = 1 - P(A)$$

  where $\overline{A}$ represents the event "A does not occur".
- The impossible event has probability zero: $P(\emptyset) = 0$.
- The **general sum rule**:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

HO GENT

Examples

1. A coin is tampered with such that $P(\text{head}) = 0.6$. What is $P(\text{tail})$?
2. From the 250 students enrolled in this course, 200 own a PlayStation game console, while 100 own a Nintendo Switch. There are 75 students who own both game consoles. What is the probability, when you pick a random student, that
   o this student does not own a PlayStation
   o this student owns a PlayStation or a Nintendo Switch (or both)
   o this student owns neither a PlayStation nor a Nintendo Switch.

HO GENT

Examples

1. A coin is tampered with such that $P(\text{head}) = 0.6$. What is $P(\text{tail})$?
2. From the 250 students enrolled in this course, 200 own a PlayStation game console, while 100 own a Nintendo Switch. There are 75 students who own both game consoles. What is the probability, when you pick a random student, that
   o this student does not own a PlayStation
   o this student owns a PlayStation or a Nintendo Switch (or both)
   o this student owns neither a PlayStation nor a Nintendo Switch.

   Answers: $1 - 200/250 = 1/5$, $(200 + 100 - 75)/250 = 225/250 = 9/10$, $1 - 9/10 = 1/10$

**HO
GENT**

Independent Events: Example

When events are **independent** it means that the occurrence of one event (or knowing that the event occurred) does **not change** the probability of the other event occurring.

As an example: draw a card from a standard deck of cards and consider the following events:

$$A = \text{a} \spadesuit \text{ was drawn}$$

and

$$B = \text{an ace was drawn}$$

When I tell you that B occurred, this will not change your opinion about the probability of A occurring, this will still be 13/52 = 1/4.
(Note: this also works the other way around!)

Independent Events: Definition

Independent events

Mathematically, two events A and B are independent when

$$P(A \cap B) = P(A)P(B)$$

HO
GENT

Dependent Events: Example

Consider the events from before where students own either a PlayStation, a Nintendo Switch or both. Let *A* be the event that a student owns a PlayStation, while *B* represents the fact that a student owns a Nintendo Switch. From the numbers above we see that

$$P(A) = \frac{200}{250} = 0.8, \quad \text{and} \quad P(B) = \frac{100}{250} = 0.4$$

when we multiply these two probabilities we get

$$P(A)P(B) = 0.8 \times 0.4 = 0.32$$

**HO
GENT**

However, the probability that a student owns both a PlayStation and a Nintendo Switch, i.e. $A \cap B$ is given by

$$P(A \cap B) = \frac{75}{250} = 0.3 \neq 0.32$$

Hence,

$$P(A \cap B) \neq P(A)P(B)$$

and the two events are **dependent**.

HO GENT

Dependent Events: Example

In practical terms this means that e.g. knowing that a student owns a PlayStation gives you additional information about his or her ownership of a Nintendo Switch.

Before you know that the student owns a PlayStation the probability that this student owns a Nintendo Switch is

$$P(B) = \frac{100}{250} = 0.4$$

However, once you know that that the student owns a PlayStation you know that this students belongs to the group of 200 PlayStation owners, of which 75 also own a Nintendo Switch. With fancy notation we write this as

$$P(B \mid A) = \frac{75}{200} = 0.375$$

HO GENT

so the probability now is slightly lower.

Dependent Events: Example

On the other hand, when you know that a student owns a Nintendo Switch, what is the probability of this student owning a PlayStation?

HO GENT

Dependent Events: Example

On the other hand, when you know that a student owns a Nintendo Switch, what is the probability of this student owning a PlayStation?

Answer:

$$P(A \mid B) = \frac{75}{100} = 0.75$$

HO GENT

Discrete Random Variable

Example

A player draws a card from a standard deck of cards. She gets 1 € when drawing a jack, 2 € when drawing a queen and 3 € when drawing a king. In all other cases she receives nothing.
We can formalise this be defining a mapping X from Ω to $\mathbb{R}$:

$$X : \Omega \to \mathbb{R} : \omega \mapsto X(\omega) = \begin{cases} 1 & \text{when drawing a jack} \\ 2 & \text{when drawing a queen} \\ 3 & \text{when drawing a king} \\ 0 & \text{in all other cases.} \end{cases}$$

We have done nothing else but associate a *number* with different events.

Probability Mass Function

Using the rules of probability we can assign probabilities to the different numerical outcomes

$$f_X(3) = P(X = 3) = P(\text{king}) = \frac{4}{52} = \frac{1}{13},$$
$$f_X(2) = P(X = 2) = P(\text{queen}) = \frac{4}{52} = \frac{1}{13},$$
$$f_X(1) = P(X = 1) = P(\text{jack}) = \frac{4}{52} = \frac{1}{13},$$
$$f_X(0) = P(X = 0) = P(\text{any other card}) = \frac{40}{52} = \frac{10}{13}.$$

Note:

- It holds that $f_X(0) + f_X(1) + f_X(2) + f_X(3) = 1$.
- And also: $0 \leq f_X(x_i) \leq 1$.

HO GENT

Probability Mass Function

To make things more explicit we can put the probability mass function in a table:

x	0	1	2	3
$f_X(x) = P(X = x)$	$\frac{10}{13}$	$\frac{1}{13}$	$\frac{1}{13}$	$\frac{1}{13}$

We can also represent this probability mass function in a graph:
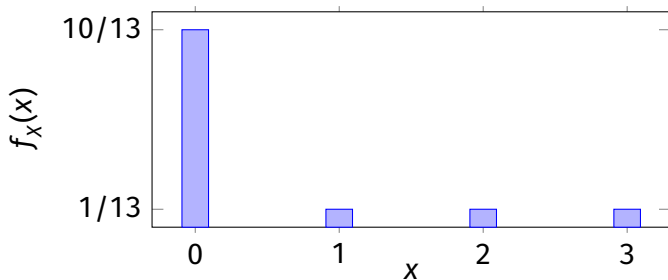
HO GENT

Probability Distribution: 1 die

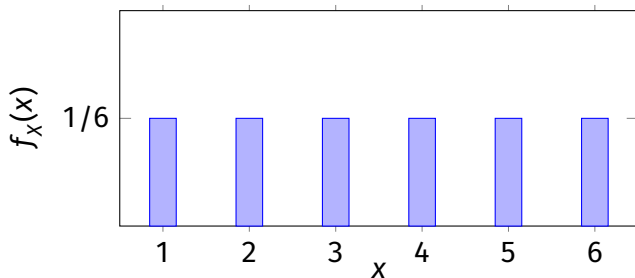What is the probability of each outcome when rolling a single die?

x	1	2	3	4	5	6

**HO
GENT**

Probability Distribution: 1 die

What is the probability of each outcome when rolling a single die?

x	1	2	3	4	5	6
$f_X(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

HO GENT

Probability Distribution: 2 dice

x	2	3	4	5	6	7	8	9	10	11	12

**HO
GENT**

Probability Distribution: 2 dice

x	2	3	4	5	6	7	8	9	10	11	12
$f_X(x) = P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
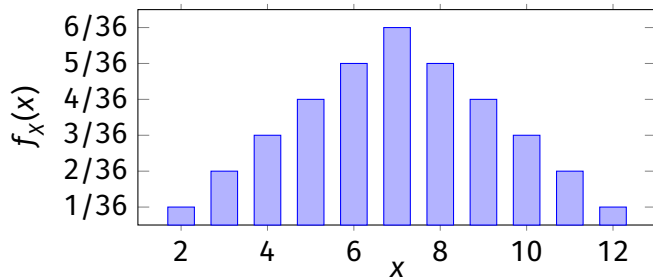
Using Probability Mass Func.

Once you have a probability mass function, you can compute other probabilities.

With the card game, what is the probability that you gain at most 2 EUR?

**HO
GENT**

Using Probability Mass Func.

Once you have a probability mass function, you can compute other probabilities.

With the card game, what is the probability that you gain at most 2 EUR? Answer:

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{10}{13} + \frac{1}{13} + \frac{1}{13} = \frac{12}{13}$$

We could also have used the complement rule in this case:

$$P(X \leq 2) = 1 - P(X > 2) = 1 - P(X = 3) = 1 - \frac{1}{13} = \frac{12}{13}.$$

HO GENT

Using Probability Mass Func.

When throwing two dice, what is the probability that the sum of the results is between 4 and 7, boundaries included

**HO
GENT**

Using Probability Mass Func.

When throwing two dice, what is the probability that the sum of the results is between 4 and 7, boundaries included
Answer:

$$P(4 \leq X \leq 7) = P(X = 4) + P(X = 5) + P(X = 6) + P(X = 7) = \frac{18}{36} = \frac{1}{2}$$

We could also compute this using left-tail probabilities only:

$$P(4 \leq X \leq 7) = P(X \leq 7) - P(X < 4) = P(X \leq 7) - P(X \leq 3) = \frac{21}{36} - \frac{3}{36} = \frac{1}{2}$$

HO GENT

Expectation of a R.V.

Consider the random variable associated with playing the card game with the following probability distribution function

$$f_X(3) = P(X = 3) = \frac{1}{13}, \qquad f_X(2) = P(X = 2) = \frac{1}{13},$$
$$f_X(1) = P(X = 1) = \frac{1}{13}, \qquad f_X(0) = P(X = 0) = \frac{10}{13}.$$

If somebody played this game a very large number of times, what would their average earnings be? I.e. what is the **expected value** of this random variable?

HO GENT

Expectation of a R.V.

Expectation of a random variable

The expectation of a random variable is denoted by μ_X or $E(X)$ and is given by

$$\mu_X = \sum_i x_i P(X = x_i) = \sum_i x_i f_X(x_i).$$

Note the similarity between this formula and the way you calculate the sample mean when a frequency table is given.
We often write μ instead of μ_X.

HO GENT

Expectation of a R.V.

The expected value of the random variable associated with the card game is given by:

$$\mu_X = \sum_i x_i f_X(x_i)$$

$$= 0 \cdot f_X(0) + 1 \cdot f_X(1) + 2 \cdot f_X(2) + 3 \cdot f_X(3)$$

$$= 0 \cdot \frac{10}{13} + 1 \cdot \frac{1}{13} + 2 \cdot \frac{1}{13} + 3 \cdot \frac{1}{13}$$

$$= \frac{6}{13}.$$

HO GENT

Variance of a R.V.

The variance of a R.V. is a measure of dispersion that resembles the sample variance very closely.

Variance of a random variable

The variance of a random variable is defined by:

$$\sigma_X^2 = \sum_i \left(x_i - \mu_X\right)^2 P(X = x_i) = \sum_i \left(x_i - \mu_X\right)^2 f_X(x_i).$$

Note: the **standard deviation** is the positive square root of the variance.

$$\sigma_X = \sqrt{\sigma_X^2}$$

HO GENT

Variance: Intuition

Consider two random variables variables *X* and *Y*.
The random variable *X* represents the number of received emails per day for a first office worker. This person receives either 48 or 52 emails per day, each with a 50 percent probability. The random variable *Y* represents the number of received emails of a second office worker. This second person receives no emails at all with 50 percent probability and with 50 percent chance receives 100 emails.

- Determine the probability mass function of *X* and of *Y*.
- Compute the expected value of *X* and *Y*. What do you notice?
- Who is experiencing a larger variability w.r.t. the number of emails received?
- Compute the variance of *X* and *Y*.

**HO
GENT**

Continuous Random Variable

- A continuous random variable takes on an uncountably infinite number of possible values.
- In this case it doesn't make a lot of sense to consider the probability that *X* equals some number *a* **exactly**, because this probability is always zero.
- What does make sense is to consider the probability that *X* takes on a value in some interval [*a*, *b*].
- This probability can be found be *integrating* (i.e. "summing up") the **probability density function** of the random variable.

HO GENT

Continuous Random Variable

Measurements of length, time, ... are examples of continuous random variables.

The leftmost histogram illustrates that most of the 5000 men were between 175cm and 180cm tall.

Heights of 5000 men
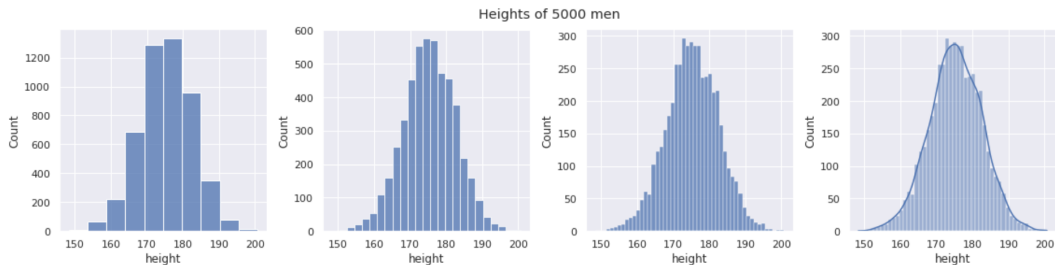
Continuous Random Variable

- From left to right the number of intervals is increased, but the total area of all the bars always equals 1.
- If the number of intervals is increased even more, the intervals would eventually get so small that we no longer have a histogram, but rather a curve (by connecting the 'dots' at the tops of the very tiny rectangles).
- The heights of people often follow an approximate **Normal Distribution**. The Normal Distribution is a type of **continuous probability distribution**.
- There are also other continuous probability distributions, e.g. the Exponential Distribution and the continuous Uniform Distribution.

Continuous Probability Distr.

The reaction speed *x* of Superman (in ms) can also be represented as a normal distribution:
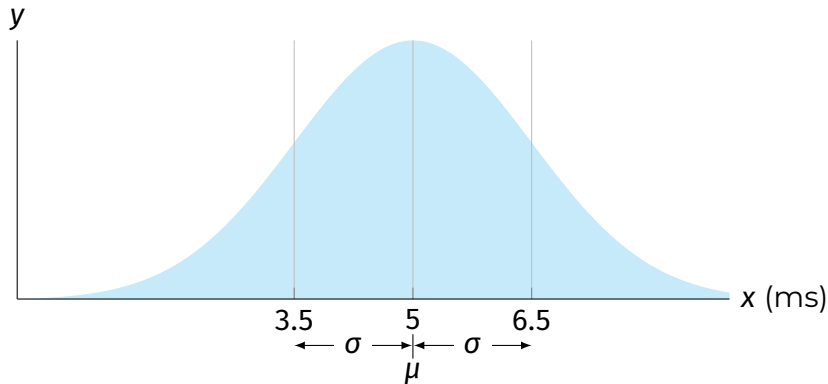
HO GENT

Expectation and Variance

Also for continuous R.V.'s there are formulas for expectation and variance. The expectation is given by

$$\mu_X = \int_{-\infty}^{+\infty} x \, f_X(x) \, dx,$$

while the variance σ_X^2 is given by

$$\sigma_X^2 = \int_{-\infty}^{+\infty} \left(x - \mu_X\right)^2 f_X(x) \, dx.$$

Note: these are essentially the same formulas as for discrete R.V.'s but with integrals instead of summations.

Note: you are **not** expected to know how to calculate integrals!

Standard Normal Distribution
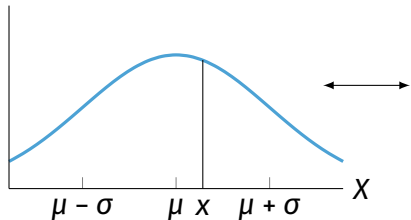
normal distribution $x \in X \sim Nor(\mu, \sigma)$
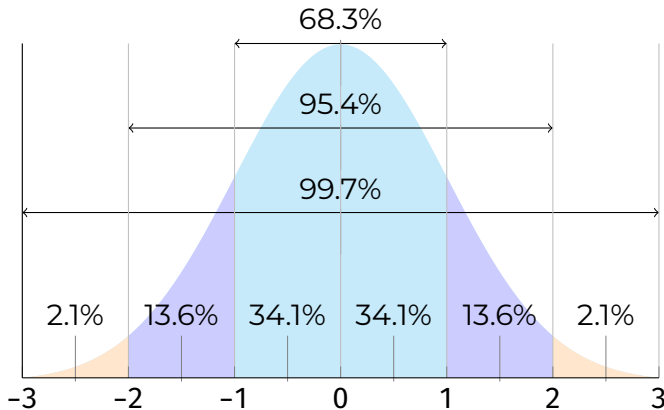
standard normal distribution
$z \in Z \sim Nor(0, 1)$
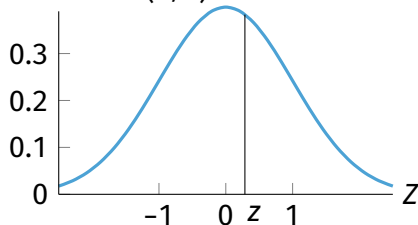
x and z have a similar position on the Gaussian bell curve.
What is the mathematical relationship between x and z?

**HO
GENT**

Standard Normal Distribution
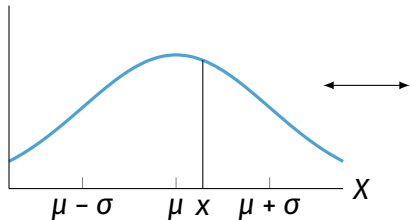
normal distribution $x \in X \sim Nor(\mu, \sigma)$

standard normal distribution
$z \in Z \sim Nor(0, 1)$
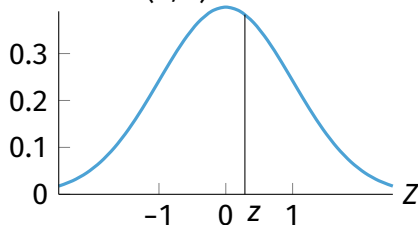
x and z have a similar position on the Gaussian bell curve.
What is the mathematical relationship between x and z?

$$X = \mu + Z.\sigma \quad \text{and} \quad Z = \frac{x - \mu}{\sigma}$$

HO GENT

Standard Normal Distribution

Python functions

Import `scipy.stats`
For a normal distribution with mean `m` and standard deviation `s`:

| Function	Purpose
`norm.pdf(x, loc=m, scale=s)`	Probability density at x
`norm.cdf(x, loc=m, scale=s)`	Left-tail probability $P(X < x)$
`norm.sf(x, loc=m, scale=s)`	Right-tail probability $P(X > x)$
`norm.isf(1-p, loc=m, scale=s)`	p% of observations are expected to be lower than result

HO GENT

Calculating Probabilities

What is the probability that Superman's observed reaction speed is over 6 ms?

Mathematical Notation:

$P(X > 6) = ?$ with $X \sim Nor(\mu = 5, \sigma = 1,5)$
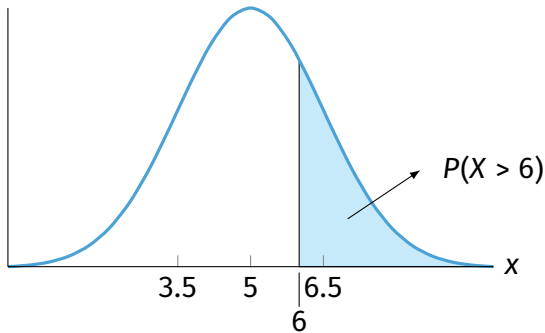
Calculating Probabilities: *z*-table

$P(X > 6) = ?$ with $X \sim Nor(\mu = 5, \sigma = 1,5)$

(Old) calculation method using z-table, e.g.
http://sixsigmastudyguide.com/wp-content/uploads/2014/04/
z-table.jpg

**HO
GENT**

Calculating Probabilities: *z*-table

$P(X > 6) = ?$ with $X \sim Nor(\mu = 5, \sigma = 1, 5)$

(Old) calculation method using z-table, e.g.
http://sixsigmastudyguide.com/wp-content/uploads/2014/04/
z-table.jpg

1. Calculate the *z*-score
   $z = \frac{6-5}{1,5} = 0,667$ so $P(X > 6) = P(Z > 0,667)$

2. Convert to a **left** tail probability
   o 100% probability rule: $P(Z > 0,667) = 1 - P(Z < 0,667)$
   o or symmetry rule: $P(Z > 0,667) = P(Z < -0,667)$

3. look for the corresponding value in the *z*-table

**HO
GENT**

Calculating Probabilities: using Python

$P(X > 6) = ?$ with $X \sim Nor(\mu = 5, \sigma = 1,5)$
Calculated directly using Python:

```
import scipy.stats as stats
stats.norm.sf(6, loc=5, scale=1.5)
```
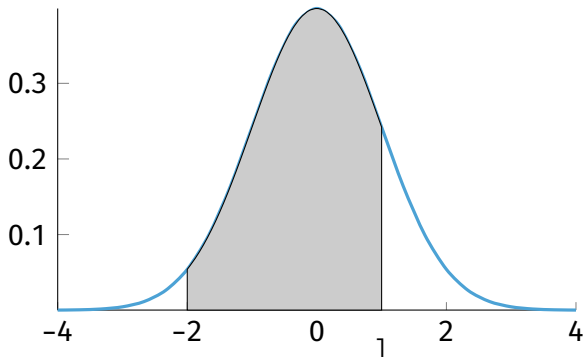
**HO
GENT**

Examples

1. What is the probability that Superman reacts in less than 4 ms?
2. What is the probability that he reacts in less than 7 ms?
3. What is the probability that he reacts in less than 3 ms?
4. What is the probability that he reacts between 2 en 6.5 ms?
5. What interval contains 80% of his reaction speed?
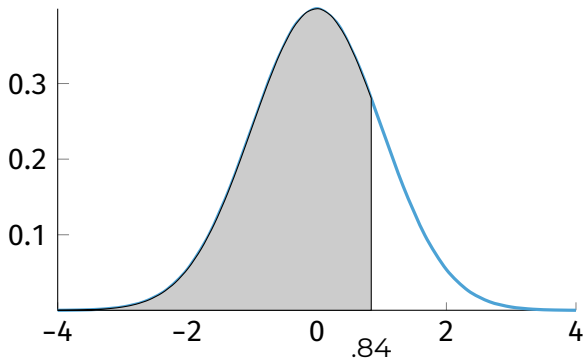
**HO
GENT**

Question 3: $P(X < 3)$
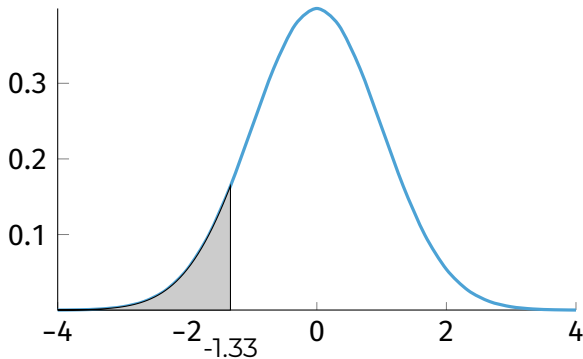
HO GENT

Question 5

For what value of x is $P(X < x) = 80\%$?

HO GENT

Exponentional Distribution

Besides the Normal Distribution, there are other often used continuous distributions, e.g. the **Exponentional Distribution**.
Values for an exponential random variable occur when there are fewer large values and more small values. For example, the amount of money customers spend in a trip to the supermarket follows an exponential distribution. There are more people who spend small amounts of money and fewer people who spend large amounts of money.
Another examples is the length, in minutes, of long distance business telephone calls.

HO GENT

Exponentional Distribution

Another example is the amount of time (in minutes) a postal clerk spends with his or her customers. In the image below, the time is known to have an exponential distribution with an average amount of time equal to four minutes.
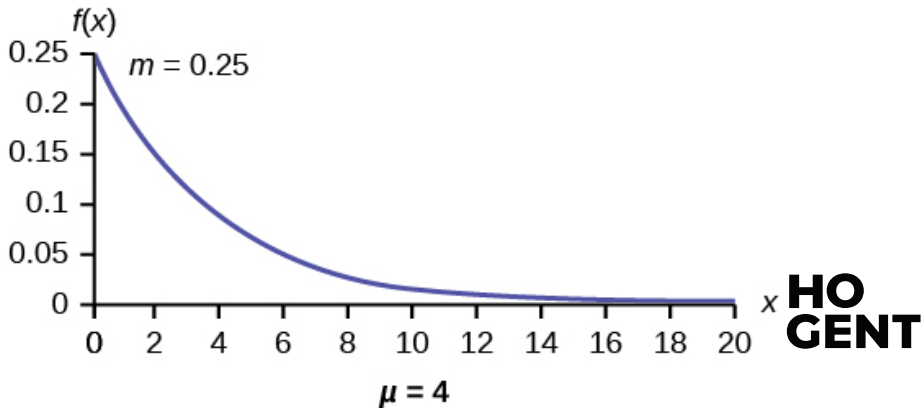
Continuous Uniform Distribution

The continuous **Uniform Distribution** describes an experiment where there is an arbitrary outcome that lies between certain bounds.
It is the simplest of all continuous probability distributions. The density function is constant where every value has an equal chance of occurring. Imagine you live in a building that has an elevator that will take you to your floor. From experience, once you push the button to call the elevator, it takes between ten and twenty seconds to arrive at your floor. This means the elevator arrival time is uniformly distributed between 10 and 20 seconds once you hit the button.

HO GENT

Continuous Uniform Distribution
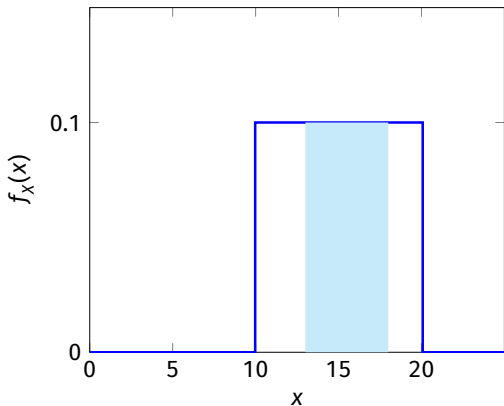
$P(13 < X < 18)$ = area under $f_X(x)$ between 13 & 18

$\quad = 5 \times 0.1$

$\quad = 0.5$

Note:

$P(13 < X < 18) = P(13 < X \leq 18) = P(13 \leq X < 18)$

$\quad = P(13 \leq X \leq 18)$

and it also holds that

$$P(10 \leq X \leq 20) = 1$$

so the total area under the probability density function is always equal to 1.

HO GENT

From Sample to Population

HO
GENT

The Central Limit Theorem

- 1 test
- 25 tests
- 100 tests

Demo:
`https://www.zoology.ubc.ca/~whitlock/Kingfisher/CLT.htm`

HO
GENT

The Central Limit Theorem

Consider a random sample of *n* observations drawn from a population with expected value *μ* and standard deviation *σ*. If *n* is sufficiently large, the probability distribution of the sample mean $\overline{x}$ will approximate a normal distribution with mean $\mu_{\overline{x}} = \mu$ and standard deviation $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$.

The larger the sample, the better the probability distribution of $\overline{x}$ will approximate the expected value of the population, *μ*.

HO GENT

Point Estimate

> **Point estimate**
>
> A **Point Estimate** for a population parameter is a formula or equation that allows us to calculate a value to estimate that parameter.

Example: sample variance/standard deviation, sample mean.

HO GENT

Confidence Interval

confidence interval

A **confidence interval** is an equation or formula that allows us to construct an interval that will contain the parameter to be estimated with a certain **level of confidence**.
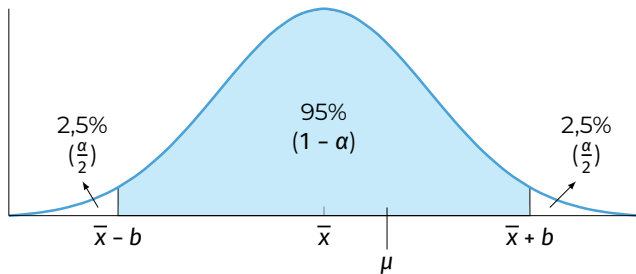
Demo:
https://www.zoology.ubc.ca/~whitlock/Kingfisher/CIMean.htm

HO GENT

Conf. Int. - Large Sample

Given a sample with mean $\overline{x}$.
We are looking for an interval $\left[\, \overline{x} - b \,,\, \overline{x} + b \,\right]$ for which we can say with a level of confidence $(1 - \alpha)$ of e.g. 95% that μ is inside this interval.

$$P\left(\overline{x} - b < \mu < \overline{x} + b\right) = 1 - \alpha = 0,95$$
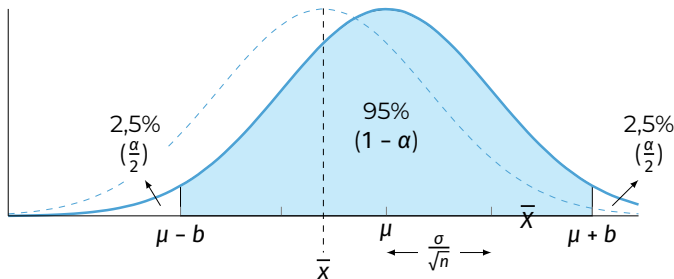

HO GENT

Conf. Int. - Large Sample

Because of the central limit theorem we know that: $\bar{x} \in \bar{X} \sim Nor(\mu, \frac{\sigma}{\sqrt{n}})$

And because of the symmetry we can say:

$$P\left(\bar{x} - b < \mu < \bar{x} + b\right) = P\left(\mu - b < \bar{X} < \mu + b\right)$$
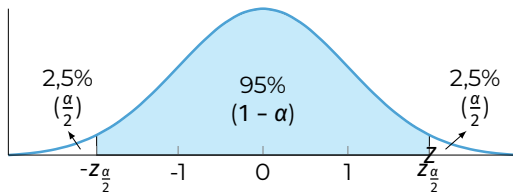
Conf. Int. - Large Sample

We now calculate the z-score for $\overline{x}$: $z = \dfrac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

We lookup (or calculate) the value $z_{\frac{\alpha}{2}}$ for which:

$$P\left(-z_{\frac{\alpha}{2}} < z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha = 0.95$$

$$P\left(z < z_{\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2} = 0.975$$

$z_{\frac{\alpha}{2}} = $ `stats.norm.isf(1-0.975)` ≈ 1.96

HO GENT

Conf. Int. - Large Sample

The result is:

$$P\left(-1,96 < \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < 1,96\right) = 0,95$$

$$P\left(\mu - 1,96\frac{\sigma}{\sqrt{n}} < \overline{x} < \mu + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

Because of symmetry we can swap μ and $\overline{x}$:

$$P\left(\overline{x} - 1,96\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + 1,96\frac{\sigma}{\sqrt{n}}\right) = 0,95$$

Now we can say with 95% confidence that:

$$\mu \in \left[\overline{x} - 1,96.\frac{\sigma}{\sqrt{n}} , \overline{x} + 1,96.\frac{\sigma}{\sqrt{n}}\right]$$

HO GENT

(in practice we will use s_{sample} as a point estimate for $\sigma_{population}$)

Conf. Int. - Small Sample

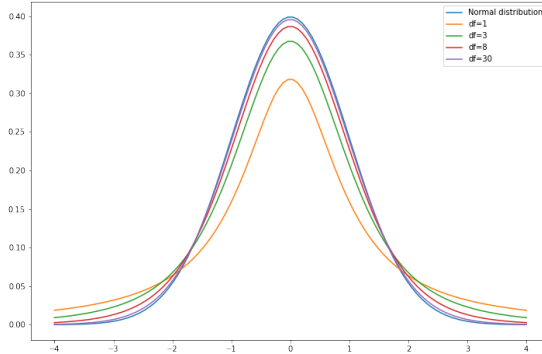For a *small* sample the central limit theorem is **no longer** valid.

Instead we can say:

If a population *X* has a normal distribution (*X ~ Nor(μ, σ)*) and you take a *small* sample with mean $\overline{x}$ and standard deviation *s*, then

$$t = \frac{\overline{x} - \mu}{\frac{s}{\sqrt{n}}}$$

will behave as a t-distribution with *n* – 1 degrees of freedom

HO GENT

Student t-distribution

Student t-distribution in Python

Import `scipy.stats`
For a *t*-distribution with `df` degrees of freedom:
`(df = degrees of freedom)`

Functie	Betekenis
`t.pdf(x, df=d)`	Probability density for x
`t.cdf(x, df=d)`	Left-tail probability *P*(*X* < x)
`t.sf(x, df=d)`	Right-tail probability *P*(*X* > x)
`t.isf(1-p, df=d)`	p% of observations is expected to be lower than this value

HO GENT

Conf. Int.: Small Sample

To determine the confidence interval for the mean μ of a population, based on a *small* sample, we first calculate $t_{\frac{\alpha}{2}, df}$.

With a confidence level of 95% we have $\frac{\alpha}{2} = 0,025$
Assume for example $n = 5$ (so $df=4$), then we have

$$t_{\frac{\alpha}{2}, df} = \texttt{stats.t.isf(1-0.975, df=4)} = 2.776$$

We can say with a confidence of 95% that:

$$\mu \in \left[\ \overline{x} - t_{\frac{\alpha}{2}, df} \cdot \frac{s}{\sqrt{n}}\ ,\ \overline{x} + t_{\frac{\alpha}{2}, df} \cdot \frac{s}{\sqrt{n}}\ \right]$$

HO GENT