Mathematics for Supply Chain

Msc Supply Chain & Purchasing

Guillaume Metzler
Institut de Communication (ICOM)
Université de Lyon, Université Lumière Lyon 2
Laboratoire ERIC UR 3083, Lyon, France
guillaume.metzler@univ-lyon2.fr

1 Course summary: Tests on one Population

In this document, we are dealing with the topic of Hypothesis Testing. The aim is to use the previous work on confidence estimation in order to provide more concrete answers to questions on some processes. More precisely, we are going to deal with:

- testing the mean on a population, when the variance is known or not
- testing a proportion
- compare the means of two different populations
- compare the variance of two different populations

1.1 Introduction

A first example . Before presenting the formalism of testing, let us take the case of setting up a automaton on an assembly line, where the aim is to find out whether or not the automaton is correctly set up. Let us imagine that we know the correct setting of the machine μ_0 and that we have a sample on which we measure a certain value $\bar{x}$ and that we know the variability for the execution of the task of our automaton, *i.e.* we know σ^2.

Our aim is to find out whether our automaton's current setting μ (a priori unknown!) is the right one or not, *i.e.* whether its value is equal to our reference value μ_0. Knowing whether or not our automaton is right is like **testing**:

$$H_0 : \mu = \mu_0 \quad \text{contre} \quad H_1 : \mu \neq \mu_0.$$

Saying differently, is the machine's current setting close to the reference setting μ_0?

In the above statement, H_0 and H_1 are called **hypotheses** and **our objective is to find out which hypothesis is the most realistic, in the probabilistic and statistical sense of the term**.

The "verification" process is very similar to what we have seen for establishing our confidence intervals. Still considering the example of our automaton, we have seen that under the assumption H_0, *i.e.* when we suppose that the machine is correctly set up, then the random variable Z defined by

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

follows a normal distribution (we'll see later that this random variable Z is also called **Test statistic**).

To find out whether our hypothesis H_0 is true, we need to ensure that our expectation μ is not too far from the ideal (or reference) value μ_0. To do this, the values taken by our samples would have to lie within an interval of values centered around the reference value μ_0 with a certain probability $1-\alpha$. If we were to rephrase this, the values obtained by sampling would have to lie in the interval 95% of the time:

$$\left[\mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} ; \mu_0 + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

In practice, we only have a limited number of samples, if not just one. So, in the event that the estimate $\bar{x}$ on our sample does not belong to this interval, we decide to **reject the hypothesis** H_0 at the risk of error α. This means that the risk of wrongly rejecting H_0 is equal to α. Otherwise, H_0 is not rejected.

Important note: the hypothesis H_0 is not rejected if the value obtained by sampling is well within the calculated confidence interval.

1.2 Assumptions and Tests

A statistical test is therefore a process that enables us to judge the validity of a hypothesis given the observations made on a sample. In practice, we will consider two **hypotheses**

Definition 1.1: Hypothesis

A hypothesis is an assumption made about the values taken by the parameters of a law (when parametric tests are performed). We always consider two hypotheses: the H_0 hypothesis, also called **null hypothesis**, which is the hypothesis used to perform the test. It is opposed to the H_1 hypothesis, called **alternative hypothesis**.

When we want to draw a conclusion about the expectation of a population, the various tests can take the following forms:

- $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$, this type of test will be called a **bilateral** test, we will try to estimate the difference between μ_0 and the value obtained by sampling, *i.e.* the value of $|\bar{x} - \mu_0|$.
- $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$, this type of test will be called a **unilateral** test, we want to check that μ does not take values greater than μ_0.
- $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$, this time we want to check that μ does not take values smaller than μ_0.

When performing a hypothesis test, there are generally 4 possible outcomes, summarized in the Table **??**. Each decision taken can also be associated with a confidence score or risk. For example, when a test is carried out under the hypothesis H_0 and this is true, there are two possible outcomes.

- we do not reject the H_0 hypothesis, in which case we can have $1 - \alpha$ confidence in the test performed
- we reject the hypothesis H_0 even though it is true; this is an error of the first kind. This happens with a risk of α, which we call **first class risk**.

Exemple 1.1. *A researcher wants to compare the efficacy of two drugs by testing different people. The null hypothesis H_0 is that the drugs are equally effective, whereas H_1 is that they are not equally effective. In this case, an error of the first kind occurs if the researcher rejects the null hypothesis and concludes that the two drugs are different when, in fact, they are not. This error is basically not very serious, as in reality the two drugs are equally effective, so there is no risk to the patient. On the other hand, if an error of the second kind occurs, i.e. if the researcher keeps H_0 when H_1 is true, this can have serious consequences for certain patients. This error is therefore more serious.*

Decision \ Truth	H_0	H_1
H_0	Correct	Second class risk
H_1	First class risk	Correct

Decision \ Truth	H_0	H_1
H_0	Confidence $1-\alpha$	Risk β
H_1	Risk α	Power $1-\beta$

Table 1: Name the different outcomes according to the decision made and the truth. The second table shows the levels of confidence, risk and power for each situation.

Remark: when we perform a statistical test, we are not trying to verify the validity of a hypothesis, *i.e.* the result of a statistical test allows us to conclude whether or not we reject the hypothesis H_0, but we are not saying that we **accept the hypothesis** H_0.

As we saw in the introductory example (and as we did in the construction of confidence intervals), we will always have to study the values taken by a reference random variable, constructed according to the context (whether or not we know the expectation and variance of our population).

1.3 Link with Confidence Intervals

We'll focus on the case of testing the mean μ when we know the variance σ^2 associated with a population, to illustrate the two types of tests: **bilateral** and **unilateral**. We remind you of the two methods available to us for rejecting or not rejecting the H_0 hypothesis, based on confidence intervals

- **Two tailed test :**
  We perform this test when we carry out the hypothesis test

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu \neq \mu_0.$$

  Recall that in this case, a confidence interval of level $1 - \alpha$, under hypothesis H_0 is given by

$$\left[\mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \mu_0 + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

  We want to check that the values taken by our samples are not too far s from the reference value μ_0. If this is not the case, we reject the hypothesis H_0. Figure 1 (left) illustrates the areas of rejection and non-rejection of the H_0 hypothesis as a function of the $\bar{x}$ values taken by our sample.

  Our hypothesis H_0 is rejected if it lies in one of the two red zones. It is retained in the opposite case, *i.e.* if the value taken by our test statistic is well within the two quantiles defined by the confidence level $1 - \alpha$ (or by the risk of error α) of our test.

$$\bar{x}_n \in \left[\mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}; \mu_0 + z_{1-\alpha_2}\frac{\sigma}{\sqrt{n}}\right]$$

  or, equivalently, if

$$\frac{\bar{x}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \in \left[z_{\alpha/2}; z_{1-\alpha/2}\right].$$

- **One tailed test on the left :**
  This test is performed when we carry out the hypothesis test

$$H_0 : \mu = \mu_0; \text{or } \mu \geq \mu_0 \quad \text{contrary} \quad H_1 : \mu < \mu_0.$$

  In this case, we want to check that the values obtained by sampling **are not lower** than the reference value μ_0. In the case of a one-sided test **to the left**, our confidence interval of level $1 - \alpha$ is defined by

$$\left[\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}; +\infty\right].$$

  Figure reffigure:test (middle) illustrates the areas of rejection and non-rejection of the H_0 hypothesis as a function of the $\bar{x}_n$ values taken by our sample.

  Our hypothesis H_0 is rejected if it lies in the red zone. It is retained in the opposite case, *i.e.* if the value taken by our test statistic is well within the two quantiles defined by the confidence level $1 - \alpha$ (or by the risk of error α) of our test.

$$\bar{x}_n \in \left[\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}; +\infty\right]$$

  or, equivalently, if

$$\frac{\bar{x}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \in [z_\alpha; +\infty].$$

- **One-tailed test on the right:**
  We perform this test when we carry out the hypothesis test

$$H_0 : \mu = \mu_0; \text{or } \mu \leq \mu_0 \quad \text{contrary} \quad H_1 : \mu > \mu_0.$$

  In this case, we want to check that the values obtained by sampling **are not greater** than the reference value μ_0.

  In the case of a one-sided test **right**, our confidence interval of level $1-\alpha$ is defined by

$$\left[-\infty; \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}}\right].$$

  Figure 1 (right) illustrates the rejection and non-rejection zones of the H_0 hypothesis as a function of the $\bar{x}_n$ values taken by our sample.

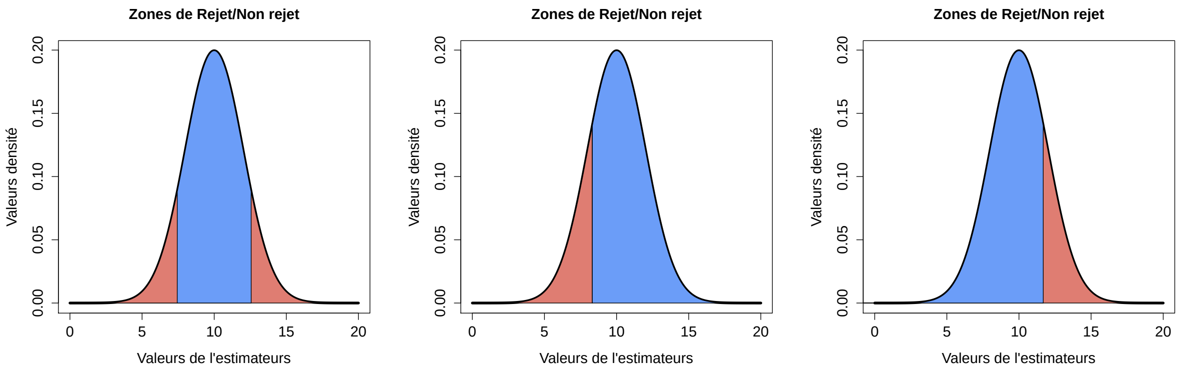

Figure 1: Graph showing the areas of rejection (for a two-tailed or one-tailed test on the left or right) or non-rejection of the H_0 hypothesis as a function of the values taken by our estimator on a sample with a $\mathcal{N}(10, 4)$. distribution. The blue zone corresponds to a zone of non-rejection of the H_0 hypothesis, while the red zone corresponds to zones of rejection of the H_0 hypothesis at a risk of error $\alpha = 0.2$.

Our hypothesis H_0 is rejected if it lies in the red zone. It is retained in the opposite case, *i.e.* if the value taken by our test statistic is well within the two quantiles defined by the confidence level $1 - \alpha$ (or by the risk of error α) of our test.

$$\bar{x}_n \in \left[-\infty; \mu_0 + z_{1-\alpha} \frac{\sigma}{\sqrt{n}} \right]$$

or, equivalently, if

$$\frac{\bar{x}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}} \in \left[-\infty; z_{1-\alpha} \right].$$

1.4 Test and p-value

Up to now, we have simply used tools based on confidence intervals to conclude whether or not to reject a hypothesis H_0 at a fixed risk threshold α. However, there is a method that provides more precise information about the risk of error we are likely to make if we reject the H_0 hypothesis, known as **the p-value**.

Definition 1.2: p-value

Let U be a test statistic distributed according to a $\mathcal{L}$ distribution and let $\bar{u}$ be the value taken by our test statistic for a given sample (e.g.: $\bar{u} = \dfrac{\bar{x}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ in the case of a test on the mean when σ is known). The p-value is then the probability that our random variable U (*i.e.* our statistic, takes on a more "improbable" value), *i.e.*

- for a two-tailed test:

$$2 \times \mathbb{P}\left[U \geq |\bar{u}|\right],$$

- for a one-tailed left-hand test:

$$\mathbb{P}\left[U \leq \bar{u}\right],$$

- for a one-tailed test on the right:

$$\mathbb{P}\left[U \geq \bar{u}\right].$$

This p-value can be deduced from the quantile table (or probability table) associated with the distribution of U. It is then compared with the risk of error α. If the p-value is smaller than the risk of error α, this means that the risk taken in rejecting H_0 is less than the threshold risk α that we had set, so we can reject H_0 *a priori* without fear. Otherwise, we keep H_0.

The p-value therefore gives us more precise information than previous methods, which rely on the use of confidence intervals. **In effect, it tells us how likely it is to reject a hypothesis incorrectly, given the observation made**.

Steps for Hypothesis Testing To summarize the different steps for hypothesis testing.

1) Define the null assumption H_0 and its alternative one H_1.
2) Choose an error rate α, which will be used to take decisions
3) Determine the statistical distribution of U under H_0, *i.e.* the random variable and its distribution in order to study the possible rejection of H_0.
4) Compute the value of the statistical test using the provided information.
5) Conclude to the rejection of H_0 using one of the following methods. We will focus on the particular case of a two tailed test, but it works similarly for the other types

of test.

If u does not belong in the confidence interval with a confidence rate equal to $1-\alpha$ associate to the distribution of the random variable U

$$\left[u_{\alpha/2};\ u_{1-\alpha/2}\right]$$

OR if $\bar{x}_n$ does not belong to the confidence interval, which is:

$$\left[\mu_0 + u_{\alpha/2}\sigma_{\bar{x}};\ \mu_0 + u_{1-\alpha/2}\sigma_{\bar{x}}\right]$$

OR compute the p-value associated to our statistical test u and compare it to α. If the p-value is lower than α, then reject H_0.

1.5 Some Statistical Quantities for Tests

This last section present the different tests you are going to meet.

- **Test on the mean μ when σ is known**

  This the case we have previously met. The used statistic is defined by:

  $$U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}},$$

  where $\bar{X}_n$ is the estimator of the mean on our sample. This statistic U is normally distributed $\mathcal{N}(0,1)$, or, equivalently, $\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

- **Test on the mean μ when σ is unkown**

  To build our statistic, we need to have a estimation of the variance on our sample s^2. The used statistical test U is defined by:

  $$U = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}},$$

  In this case, U follows a **Student distribution with $n-1$ degree of freedom** where n is the sample size.

- **Test on the proportion** p

  The statistical test U used here is defined by:

$$U = \frac{\bar{P} - p}{\sqrt{\dfrac{p(1-p)}{n}}},$$

  where $\bar{P}$ is the estimator of the proportion on the sample of size n which verifies, $n\bar{P} \sim \mathcal{B}(n, p)$. p is unknown, the variance of the distribution is estimated using the one of $\bar{P}$ on the sample, *i.e.* we replace $\sqrt{\dfrac{p(1-p)}{n}}$ by $\sqrt{\dfrac{\bar{p}(1-\bar{p})}{n}}$.

  The random variable U **is normally distributed** $\mathcal{N}(0,1)$, or equivalently, $\bar{P} \sim \mathcal{N}\left(p, \dfrac{p(1-p)}{n}\right)$ when n is sufficiently large.

Exemple 1.2. *Let us go come back to the automaton used in the filling process. We aim to verify if the machine is well adjusted with a margin of error equal to* 5%*, i.e., if the filling process is normally distributed* $\mathcal{N}(\mu, \sigma^2)$*, we aim to check if* $\mu = \mu_0 = 86.5g$*.*
We also make the assumption that the variance σ^2 *is unknown.*

100 *boxes are randomly selected to test this assumption. On the selected boxes, we measure a mean weight equal to* $\bar{x}_n = 86g$ *and the variance is equal to* $s^2 = 6.25g$*.*

To answer to the questions, there are several possiblities. First you need to set the assumptions

- $H_0 : \mu = \mu_0 = 86.5$
- $H_1 : \mu \neq \mu_0$

We are in the situation when we test a mean value and we do not have access to the standard deviation of the population. The statistical test U *will then follow a Student distribution with* $n-1$ *degree of freedom, i.e.,* 99 *degree of freedom.*

Two possibilities to answer to the question: (*i*) *Build the confidence interval as we have done before and check if the value* μ_0 *belongs in this interval* ***or*** *(ii) compute the p-value associated to the statistical test and compare it to the error rate* $\alpha = 0.05$*.*

(i) *We have previously seen that the confidence interval is given by:*

$$\left[\bar{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}; \bar{x} + t_{1-\alpha/2}\frac{s}{\sqrt{n}}\right].$$

In our case, we have $t_{1-\alpha/2} = t_{0.975} = 1.984$*, and thus have the following confidence interval:*

$$[85.504; 86.496].$$

We can thus reject the assumption H_0 *with a probability equal to* 5% *of doing a mistake. We are now ready to look at the second solution we can perform in order to have a more precise answer.*

(ii) *Our statistical test* U *follows a Student Distribution with* 99% *degree of freedom. The value* u *taken by the statistic on the sample is equal to:*

$$u = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{86 - 86.5}{\sqrt{\frac{6.25}{100}}} = -2.$$

We perform a ***two tailed test****, so the p-value is defined by:*

$$\mathbb{P}\left[|U| \geq |u|\right] = 2\mathbb{P}\left[U \geq |u|\right] \quad \textit{using the symmetry of the distribution.}$$

We need to estimate $2\mathbb{P}[U \geq 2] = 2 \times (1 - \mathbb{P}[U \leq 53.03]) = 0.048$*. The p-value is lower than* 0.05*, we can reject the assumption* H_0 *with an error rate of* 0.048.

You can test with other values of s^2*, for instance, for* $s^2 = 10$*, can we reject* H_0 *? What is the associated p-value?*

2 Course Summary: Tests on two Populations

Up to now, we have only had information on a single sample, but it can sometimes be useful to compare several populations to see whether or not they share the same characteristics. For example, we could study cell counts according to two different experimental protocols, and see if the number of cells observed differs significantly from one protocol to another. From a marketing point of view, we could, for example, look at whether people spend the same average amount of money on summer vacations as they do on winter vacations. We're therefore interested in **comparing the average of two populations** represented by two samples, but we have two possible cases :

- in the case of cell counts, our two different cultures imply that we are taking measurements on different populations or **independent**, *i.e.* our measurements come from two different populations.

- In the case of our marketing analysis, we could survey the same group of people who go on vacation in summer **and** in winter. In this case, the measurements are taken from the same sample of people, but at two different times. This is known as **dependent or matched samples**.

In this part, we will make the assumption that our data are normally distributed or that the sample size is at least greater than 30.

2.1 Independent Samples

THis is the case where the measures or the collected data come from tow different and independent populations, *i.e.* we do not collect the information on the same persons. Another example could consist in comparing the means of the students from two groups to a given course and see if there is statistical difference between the two. Same thing when studying two different strategies in Supply Chain.

When we have averages from two independent samples and we want to test whether or not the averages are identical, we are no longer going to compare the values obtained from the samples to a theoretical value, as we did in the previous section. This time, we're going to study the difference between the two averages and therefore formulate the following hypothesis test:

- $H_0 : \mu_1 - \mu_2 = 0$,means are equal
  v.s.

- $H_1 : \mu_1 \neq \mu_0$ for a two tails test
  $H_1 : \mu_1 < \mu_0$ or $\mu_1 > \mu_2$ for a one tail test

In studying the difference in means, this means assuming that the difference under study follows a centered Normal distribution, i.e. with mean zero.

Let us consider two different samples:

- one of size n_1 coming from a population with a mean equal to μ_1
- one of size n_2 coming from a population with a mean equal to μ_2. This second sample is independant from the first one

We must now distinguish between two cases, as we did for confidence intervals, (i) the case where variances are known and (ii) the case where variances are unknown.

(i) Case where the variance of the distributions of the two populations σ_1^2 and σ_2^2 are known

The variance associated with the difference-in-means estimator is equal to the sum of the variances on each sample., *i.e.*

$$\sigma_{\bar{x}_1-\bar{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2},$$

where σ_1^2 and σ_2^2 are the variances of the two samples respectively.

The statistic that is used for the test is:

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

This random variable Z is normally distributed, $\mathcal{N}(0,1)$. To test equality of means, for a two tails test under $H_0 : \mu_1 = \mu_2$, lour statistic is equal to:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

In fine, we reject H_0 with an error α if

$$z \notin \left[z_{\alpha/2}; z_{1-\alpha/2}\right].$$

(ii) Case where σ_1^2 and σ_2^2 are unknown.

We will consider also two different cases:

1. **Variances are equal**

   In this first situation, we need to compute the pooled variance of our sample. We can see it as the weighted mean variance of our two samples. ‘ It is defined as:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2},$$

   where s_1^2 and s_2^2 are the estimated variances on the samples.
   Then the used statistic is defined by

$$T = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

   which follows a Student distribution with $n_1 + n_2 - 2$ degrees of freedom.

2. **Variances are different**

   This case is much more complicate to study, we are going to make an example in class and see how Excel is able to solve this problem.

Dependent or paired samples. This is the case when measurements are taken on the same population, but at two different times. Another example we could cite is the study of the impact of a diet on a person's body mass. Two samples are studied, one consisting of measurements taken before the diet and the other of measurements taken after or during the diet.

In this case, we have two samples of the same size n (since we're studying the same people at two different times):

$$(x_1, x_2, x_3, ..., x_{n-1}, x_n) \quad \text{and} \quad (y_1, y_2, y_3, ..., y_{n-1}, y_n).$$

The test will be performed on the single sample defined as the difference between the measurements taken on the two samples, *i.e.*

$$(x_1 - y_1, x_2 - y_2, x_3 - y_3, ..., x_{n-1} - y_{n-1}, x_n - y_n).$$

If we do not know the variance of the distribution of our population, which will always be the case for us, we consider $\bar{x}$ to be the mean estimated on the "difference" sample and we note s the standard deviation associated with this same "difference" sample.

The hypothesis test consists of investigating the nullity of the mean of our population, *i.e.* we formulate the following hypotheses

- $H_0 : \mu = 0$, the mean of D is equal to 0.
  v.s.
- $H_1 : \mu \neq 0$ for a two tails test
  $H_1 : \mu < 0$ or $\mu > 0$ one tail test.

The statistic is defined by:

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{s^2}{n}}}$$

it follows a Student distribution with $n - 1$ degrees of freedom. In the case of a two-tailed test, the null hypothesis H_0 is rejected if the value of the test statistic, at the risk of error α under hypothesis H_0, defined by

$$t = \frac{\bar{x}}{\sqrt{\frac{s^2}{n}}}$$

is outside the range:

$$\left[t_{\alpha/2}; t_{1-\alpha/2}\right].$$

Exemple 2.1. *We are interested in the height of young children. Taking a sample from an elementary school,* 41 *boys and* 61 *girls in first grade classes were measured. The average height of the boys in this sample is* $\bar{x}_1 = 107$ *cm with a standard deviation* $v_1 = 8$ *cm. The mean height of the girls is* $\bar{x}_2 = 104$ *cm with a standard deviation* $v_2 = 9$ *cm. Both samples are assumed to be Gaussian. With a risk of error of* $\alpha = 0.05$*, can we say that boys are taller than girls?*

This is a case where we want to compare means for which we have two estimates obtained on independent samples. In this case, we don't know the variance associated with each population, so the statistical test we will use will be based on Student's law.

We are going to test the hypothesis H_0: boys and girls are the same height, i.e. $\mu_1 = \mu_2$ where μ_1 designates the average height of boys and μ_2 designates the average height of girls, against the alternative hypothesis H_1: boys are on average taller than girls, i.e. $\mu_1 > \mu_2$. We are going to do ***one tail test on the right (upper tail test)****.*

The statistic is given by:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{s\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}},$$

where s^2 is the pooled variance which is equal to:

$$s^2 = \frac{40 \times 64 + 60 \times 81}{100} = 74.2$$

The test statistic T is distributed according to a Student's t distribution with 100 degrees of freedom.

La valeur de la statistique de test t nous est donnée par

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s^2}{n_1} + \dfrac{s^2}{n_2}}} = \frac{107 - 104}{\sqrt{\dfrac{74.2}{41} + \dfrac{74.2}{61}}} = 1.706.$$

Our p-value is defined by:

$$p\text{-}value = \mathbb{P}[T > 1.706] = 1 - \mathbb{P}[T < 1.765] = 1 - 0.954 = 0.046.$$

Our p-value is smaller than the risk α considered, so we can reject the hypothesis H_0 and can therefore assert that, on average, boys are taller than girls.

3 Exercises

Exercice 3.1 (Internet Service).
The technical operations department wants to ensure that the mean target upload speed for all Internet service subscribers is at least 0.97 *on a standard scale in which the target value is* 1.0*. Each day, upload speed was measured 50 times that are stored in the UploadSpeed file.*

1. *Compute the sample statistics and determine whether there is evidence that the population mean upload speed is less than* 0.97*.*

2. *Write a memo to management that summarizes your conclusions.*

Exercice 3.2 (Coffee Shop).
The owner of a specialty coffee shop wants to study coffee purchasing habits of customers at her shop. She selects a random sample of 60 *customers during a certain week, with the following results:*

- *The amount spent was* $\bar{x} = 7.25$ *and* $s = 1.75$*.*
- *Thirty-one customers say they "definitely will" recommend the specialty coffee shop to family and friends.*

1. *At the* $\alpha = 0.05$ *level of significance, is there evidence that the population mean amount spent was different from* 6.50\$*?*
2. *Determine the p-value*
3. *At the* 0.05 *level of significance, is there evidence that more than* 50% *of all the customers say they "definitely will" recommend the specialty coffee shop to family and friends?*

4. *What is your answer to* 1. *if the sample mean equals* 6.25\$*?*

5. *What is your answer to* 3. *if* 39 *customers say they "definitely will" recommend the specialty coffee shop to family and friends?*

Exercice 3.3 (Credit Risk).

A credit risk engineer, employed in a company specialized in consumer credit, wants to verify the hypothesis that the average value of the monthly payments of his customers is the hypothesis that the average monthly payment of the clients in his portfolio is 200 *euros. A random sample of* 144 *customers, taken at random from the database, gives an empirical average* $\bar{x} = 193.74$ *and an unbiased estimate of the standard deviation* $s = 48.24$.

1. *What are the statistical hypotheses associated with the engineer's problem and what type of test should be implemented to help him make a statistically correct decision?*
2. *Can he conclude, at the risk of* 5%, *that the postulated average value of repayments is correct?*

Exercice 3.4 (Test).

In a study of adolescents' learning ability, a sample of 30 *adolescents is recruited for a series of tests. In order to have a relatively homogeneous sample in terms of IQ (Intelligence Quotient), the statistical protocol requires that the standard deviation of the IQ not exceed* 20 *points.*

							QI							
131	*108*	*85*	*96*	*86*	*126*	*128*	*107*	*119*	*87*	*103*	*110*	*125*	*77*	*90*
109	*109*	*129*	*95*	*117*	*107*	*102*	*83*	*114*	*72*	*99*	*103*	*97*	*109*	*97*

1. *Estimate the mean and the standard deviation of the sample*
2. *Can we say that, on average, the mean QI is higher than* 115 *at a significance level of* $\alpha = 0.05$*?*

Exercice 3.5 (Mail Services, part I).

*AMS communicates with customers who subscribe to telecommunications services through a special secured email system that sends messages about service changes, new features, and billing information to in-home digital set- top boxes for later display. To enhance customer service, the operations department established the business objective of reducing the amount of time to fully update each subscriber's set of messages. The department selected two candidate messaging systems and conducted an experiment in which 30 randomly chosen cable subscribers were assigned one of the two systems (*15 *assigned to each system). Update times were measured and are provided in the UpdateTime file.*

1. *Analyze the data and write a report to the computer operations department that indicates your findings.*
2. *Suppose that instead of the research design described in the case, there were only 15 subscribers sampled, and the update process for each subscriber email was measured for each of the two messaging systems.*