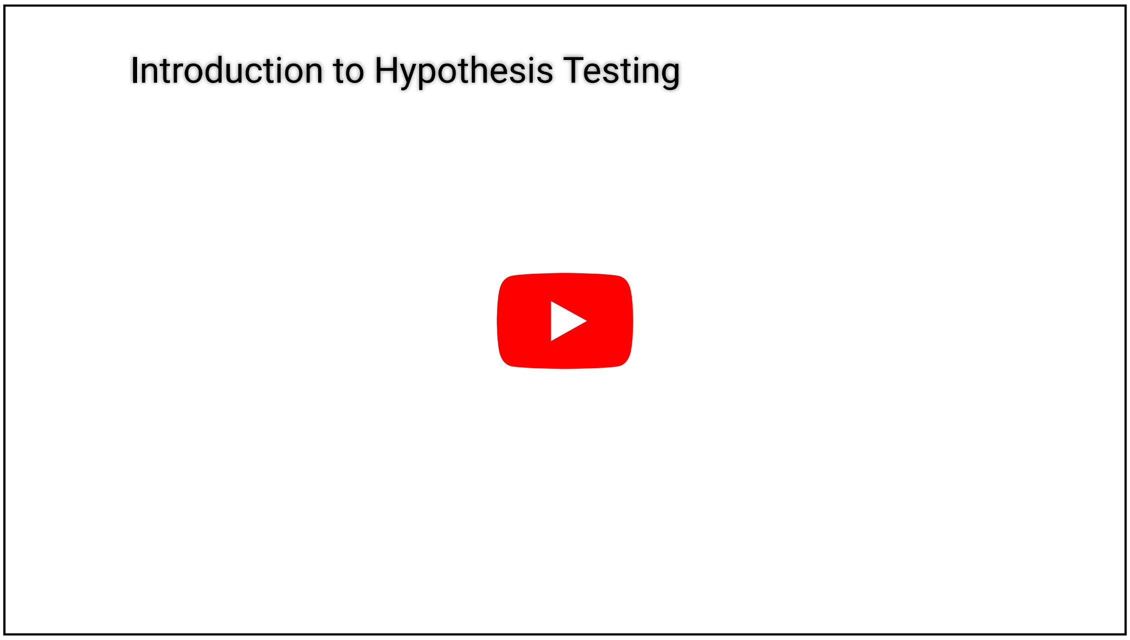
Name: ____________________ **Class:**
Class #: ____________________ **Section #:** ____________________
Instructor: Ralf Becker **Assignment:** Hypothesis Testing

Question 1: (0 points)

Introduction



Parameters in a population of interest are usually unknown. Say the average house price in Greater Manchester 2021, or the median, or the 90th percentile, or the variance of house prices, or the proportion of property transactions that happened at a price larger than £ 1m. All of these are population parameters which typically are unknown.

You have learned that we can use the information in a sample, for instance the sample mean, to estimate such unknown population parameters. So we now know how to estimate population means, variances and proportions. In the section on sampling you learned that sample means $\bar{X}$ are random variables which have a distribution. In particular you learned from the CLT that the sample mean is, for a big enough sample, normally distributed.

While we talked about the sample mean $\bar{X}$ as an estimate for the population mean (μ), there are different population parameters we may be interested in and their sample estimators.

Population Parameter	Estimator	Estimate
mean: μ	$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$	$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$
variance: σ^2	$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$	$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$
covariance: σ_{XY}	$S_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$	$s_{XY} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$
correlation: ρ	$R = \frac{S_{XY}}{S_X S_Y}$	$r = \frac{s_{XY}}{s_X s_Y}$
proportion: π	P	p

You will recognise all of these statistics from the descriptive statistics section. The estimators (in the middle column) are expressed as functions of random variables (random draws from the population, X_i). The estimates are the actual values you calculate once you have a sample (formulated on the basis of sample values x_i).

In this and the next section we will demonstrate how you can use a sample estimates (i.e. the realisation of a sample mean, $\bar{x}$) to learn something about an unknown population mean (μ). In particular you will learn about two commonly used techniques, the hypothesis test (in this section) and the confidence interval (next section). Both inference techniques are only possible if you know what distribution your sample estimator follows. This is why the CLT is so crucial!

Question 2: (1 point)

Hypothesis Testing

Hypothesis testing is a widely used concept. In fact it is a very subtle concept and often applications (or communications of applications) do not pay significant tribute to the subtleness.

We start by reviewing what we were able to do following the introduction of the CLT.

Consider a random variable M which is known to be normally distributed $M \sim N(3, 25)$.
What is the probability that an individual draw of M delivers a value which is smaller than 0?
$Pr(M < 0) = Pr\left(Z < \frac{0-3}{5}\right) = Pr\left(Z < -0.6\right) = 0.2743$
What is the distribution of $\bar{M}$ if $\bar{M}$ is the sample average calculated for a sample of size $n = 16$?
$\bar{M} \sim N\left(3, \frac{25}{16}\right)$
Calculate the probability that a particular sample mean (of a sample of size $n = 16$) is smaller than 0.
$Pr(\bar{M} < 0) = Pr\left(Z < \frac{0-3}{\sqrt{25/16}}\right) = Pr\left(Z < -2.4\right) = 0.0082$

What you learned from this example is, that, if you know the distribution of the original random variable to be normal and if you know the population parameters, then you can calculate probabilities for certain outcomes of the sample mean random variable, here $\bar{M}$.

Let's look at another example.

Consider a random variable R. which you know to have expected value $E[R] = 10$ and population variance $Var[R] = 16$. While you do not know how R is distributed, you do have a sample of size $n = 100$ which you are confident to be large enough to apply the CLT.

With that information you can derive that

$$\bar{R} \sim N\left(10, \frac{16}{100}\right)$$

Calculate the probability that a particular sample mean (of a sample of size $n = 100$) is larger than 11.2?

$$\begin{aligned} Pr(\bar{R} > 11.2) &= Pr\left(Z > \frac{11.2 - 10}{\sqrt{16/100}}\right) \\ &= Pr\left(Z > \frac{1.2}{0.4}\right) = Pr\left(Z > 3.0\right) \\ &= 1 - Pr\left(Z \leq 3.0\right) \\ &= 1 - 0.9987 = 0.0013 \text{ from the Normal distribution table} \end{aligned}$$

Let's be clear about the interpretation. If we know that the population mean is 10 (and the population variance is 16) then the probability to draw a sample of size $n = 100$ with a sample mean ($\bar{R}$) larger than 11.2 is 0.0013, or 0.13%. So that is rather unlikely. We knew things about the population and calculated a probability about a sample outcome.

Now we turn the table. The reality of an applied researcher is that she has information about a sample but we need to find out something about the population. Here comes the trick in hypothesis testing. Let's say you do have a sample of size $n = 100$ and in that sample the sample mean was $\bar{r} = 11.2$ (note now the small $\bar{r}$ as we are now talking about an actual realisation, one particular sample with that sample mean.). As we turned the table (we know sample information but we do not know the population mean μ) we now hypothesise that the population mean takes a certain value.

Let's hypothesise that the sample mean was 10. Do you think that having obtained the actual sample mean of $\bar{r} = 11.2$ is consistent with that hypothesised population mean of $\mu = 10$?

(a) Yes

(b) No

So, we conclude that the obtained sample mean of 11.2 is not consistent with the hypothesised population mean of 10, as there is too small a probability to actually obtain such a sample mean (or larger) if the true population mean was actually 10. What we have done here is basically, what we call a hypothesis test. We stated a hypothesis, looked at the evidence in light of this hypothesis and decided that the evidence is too unlikely to have occurred if the hypothesis was true. The probability we calculated earlier, $Pr(\bar{x} \geq 11.2) = 0.0013$, now takes a very different interpretation: If the population mean was $\mu = 10$, then the probability of obtaining a sample mean ($\bar{r}$) of 11.2 (or larger) is 0.13%. This is an incredibly important interpretation of the above probability. In fact it has its own name, it is called a **p-value**.

Hypothesis testing in "real life"

Here is a "real life" equivalent of the above strategy.

You return home from a long overseas holiday. You are jet-lagged and you sleep on the couch in your darkened flat for a long time. You wake up and just as you wonder whether the sun is shining (you hypothesise that the sun is shining) your flatmate enters the flat with soaking wet clothes (your flatmate is the sample information).

Is the evidence consistent with your hypothesis? No, it isn't. You judge that it is unlikely that your flatmate entered the flat with wet clothes if the sun was shining.

There are some wrinkles to work out below. For instance, we said that we did not know the population mean, and therefore had to hypothesise a value for μ, but we happily relied on calculations which also did require a known value for the population variance. As it turns out, as we perform a formal hypothesis test, we will have to solve the following issues.

- What is the hypothesis we should choose
- What is the sampling distribution of the sample statistic (as we need the distribution to calculate the p-value)?
- How do we get the variance of the population (as we will need that to calculate the p-value)?
- When is the p-value low enough for us to reject our hypothesis?

Question 3: (0 points)

Introduction

From the above example it should be obvious that the value which we hypothesised the population mean μ to be was crucial. Had we changed it, to say 11, the calculation of the p-value would have changed.

Following on from the above example, there is a random variable R with a population variance of $Var[R] = 16$. You have a sample ($n = 100$) with sample mean of $\bar{r} = 11.2$. You hypothesis that the population mean is $\mu = 11$. What is the probability of obtaining a sample mean of 11.2 (or larger) if the population mean was $\mu = 11$?

$$\begin{aligned} Pr(\bar{R} > 11.2) &= Pr\left(Z > \frac{11.2 - 11}{\sqrt{16/100}}\right) \\ &= Pr\left(Z > \frac{0.2}{0.4}\right) = Pr\left(Z > 0.5\right) \\ &= 1 - Pr\left(Z \leq 0.5\right) \\ &= 1 - 0.6915 = 0.3085 \text{ from the Normal distribution table} \end{aligned}$$

So clearly, how we set the hypothesis matters. If the population mean was hypothesised to be $\mu = 11$ we would have a probability of around 30% of obtaining a sample mean of 11.2 or larger. That is a sizeable probability and we would not easily dismiss the hypothesis. It is apparent from this that the p-value is a property of the combination of the sample and the hypothesised population value.

In fact, when we perform a hypothesis test we need to think about a null and an alternative hypothesis.

Question 4: (0 points)

Null Hypotheses

A null hypothesis is a claim about the value of some population parameter, for example $\mu = 20, \pi = 0.45$ or $\sigma^2 = 1$. Recall that we have only used the population mean as an example but we could form hypotheses about other population parameters. For now we will stick with the population mean.

A null hypothesis can also be much more general than this, and refer to functions of population parameters. For example, we may obtain a random sample from a population with population mean μ_1 ,and then a random sample from another population with mean μ_2 and ask whether $\mu_1 = \mu_2$ or equivalently $\mu_1 - \mu_2 = 0$. This type of test will be discussed in a later section.

The last example shows the link between the adjective "null" and the numerical value 0. The earlier examples can also be rephrased to emphasise this as $\mu - 20 = 0, \pi - 0.45 = 0$ and $\sigma^2 - 1 = 0$.

A null hypothesis expresses the preconception or expectation about the value of the population parameter. There is a standard, compact, notation for this:

$$H_0 : \mu = \mu_0$$

expresses the null hypothesis that $\mu = \mu_0$ where μ_0 represents a particular value like 20 above.

Deciding whether or not $\mu = \mu_0$ on the basis of sample information is usually called testing the hypothesis.
As you saw from the earlier example we may end up in the situation where we wish to dismiss (or formally reject) the null hypothesis. In fact, it turns out that when we perform a hypothesis test we are basically making a binary decision, we either reject, or not reject the null hypothesis.

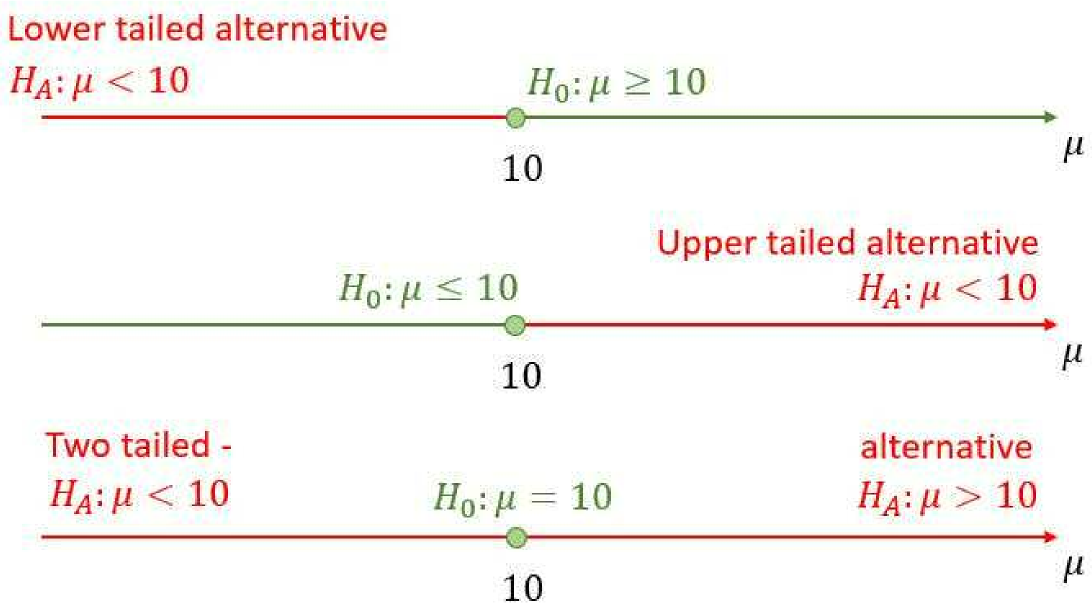
It is important to note that even if we do not reject a null hypothesis, we are not saying that the null hypothesis is correct or the truth. For instance, in the earlier example, we possibly would not have rejected the null hypothesis that $\mu = 11$. But equally we would have not rejected the null hypothesis that $\mu = 11.4$. There would be many null hypotheses we would not want to reject.

Question 5: (1 point)

Alternative Hypotheses

What happens if we reject a null hypothesis, like the above hypothesis that $\mu = 10$? The strategy when performing a hypothesis is to divide the space of possibilities into two areas, the area covered by the null hypothesis and the rest, which we then call the alternative hypothesis. If we reject the null hypothesis we "adopt" the alternative hypothesis. You will see below that the alternative hypothesis is always a range of possible values. It does not point to any specific value. So, in rejecting the null hypothesis, we are not deciding in favour of another specific value of μ.

There are three different ways in which the alternative hypothesis can be set up (and the null hypothesis will be the complement). These are illustrated in the following image:



You should always think of the null and alternative hypothesis as pairs. The value μ_0 is the value we "hypothesise".

$$H_0 : \mu \geq \mu_0; H_A : \mu < \mu_0$$
$$H_0 : \mu \leq \mu_0; H_A : \mu > \mu_0$$
$$H_0 : \mu = \mu_0; H_A : \mu \neq \mu_0$$

Here, H_A stands for the alternative hypothesis. In some textbooks, this is denoted H_1, but is still called the alternative hypothesis. Notice that the null and alternative hypotheses are expected to be mutually exclusive: it would make no sense to reject H_0 if it was also contained in H_A.

Note that for the one-sided tests, we sometimes state $H_0 : \mu = \mu_0$ rather than $H_0 : \mu \geq \mu_0$ or $H_0 : \mu \leq \mu_0$. This makes no substantial difference as we will see later. The alternative hypothesis will always reveal what sort of set-up you are looking at.

Which sort of alternative hypothesis should be chosen? This all depends on context and perspective. Suppose that a random sample of jars of jam coming off a packing line is obtained. The jars are supposed to weigh, on average, 454g. You have to decide, on the basis of the sample evidence, whether or not this is true. So, the null hypothesis here is

$$H_0 : \mu = 454.$$

As a jam manufacturer, you may be happy to get away with selling, on average, lighter jars, since this gives more profit, but be unhappy at selling, on average, overweight jars, owing to the loss of profit. So, the manufacturer might choose the alternative hypothesis

$$H_A : \mu > 454.$$

A consumer, or a trading standards conscious jam manufacturer, might be more concerned about underweight jars, so the alternative might be

$$H_A : \mu < 454.$$

A mechanic from the packing machine company might simply want evidence whether or not the machine is working to specification, and would choose

$$H_A : \mu \neq 454$$

as the alternative.

You are working for a car manufacturer and you are testing whether the speedometers in your cars are working accurately. On your test track you are driving cars exactly at a speed of 110 km/h. You then take a sample of the speeds which are shown on the cars speedometers.

Which of the following hypotheses should you be testing?

(a) $H_0 : \mu \geq 110; H_A : \mu < 110$

(b) $H_0 : \mu \leq 110; H_A : \mu > 110$

(c) $H_0 : \mu = 110; H_A : \mu \neq 110$

Question 6: (0 points)

Types of Error

After discussing the set-up of the hypotheses, we are now getting closer to understanding for what sort of p-values you should be rejecting the null hypothesis. As it turns out that is a crucial question and understanding this point will lead you to a very good understanding of what hypothesis testing does.

By setting the null and alternative hypothesis we basically divided all the possibilities for the population parameter into two categories. One called the null hypothesis and the other the alternative hypothesis. When we use a two-tailed (or two-sided) alternative, then the null hypothesis actually only consists of one particular value.

The unknown truth will be in either of the two categories and we will eventually make a decision of either rejecting or not rejecting the null hypothesis. This means there are four possible outcomes as illustrated in the following table.

The (unknown) truth	The tests' decision H_0	H_A
H_0	Correct decision	Rejecting a correct H_0 Type 1 Error
H_A	Accepting an incorrect H_0 Type 2 Error	Correct decision

To explain this, let us return to the example of the jam manufacturer who investigates whether the filling machine works as it should. Let's say they do worry about jars being filled with less than the advertised 454g, wanting to avoid unfavourable consumer feedback. They therefore design the following hypotheses:

$$H_0 : \mu \geq 454$$
$$H_A : \mu < 454$$

If in truth, the population parameter is correctly described by the null hypotheses, i.e. the mean filling weight of the machine is 454g or higher, then performing a hypothesis test and coming to the conclusion that H_0 should not be rejected would be a correct decision. Equally, if in truth the filling machine had a mean filling weight of less than 454g and the hypothesis test made us conclude that we should reject H_0 would again be a correct decision.

However, if in truth the mean filling weight is larger or equal to 454g (H_0) and after looking at a sample and performing a hypothesis test we conclude that we should reject H_0 then we would have made a mistake. We call this type of mistake a Type 1 error. And if in truth the average filling weight is below 454g but after performing a hypothesis test using sample information we decide to not reject H_0 then we have again made an incorrect decision, a Type 2 error.

Of course, the problem is that we do not know the truth! All we know is what we decided after performing a hypothesis test on the basis of sample information. We have to accept that, when we make a decision then there will be a probability that we are right and another probability that we are wrong.

The objective in designing a procedure to test an H_0 against an H_A - i.e. decide whether to accept H_0 or to reject H_0, is to ensure that a Type 1 error does not occur too often. More precisely, the objective is to design a procedure which fixes the probability of a Type 1 error occurring at a prespecified level, α, which is small and therefore presumably tolerable.

Question 7: (1 point)

P-Values

At this stage it is worth thinking a little more carefully about p-values.

You measure the weight of the jam (W) in a random sample of 50 jam jars. You obtain a sample mean of $\bar{w} = 453.78$. You know the standard deviation of the filling machine to be 10g and indeed you know that the distribution of the weight coming from the filling machine is a normal distribution.

What is the p-value when testing the following hypotheses?

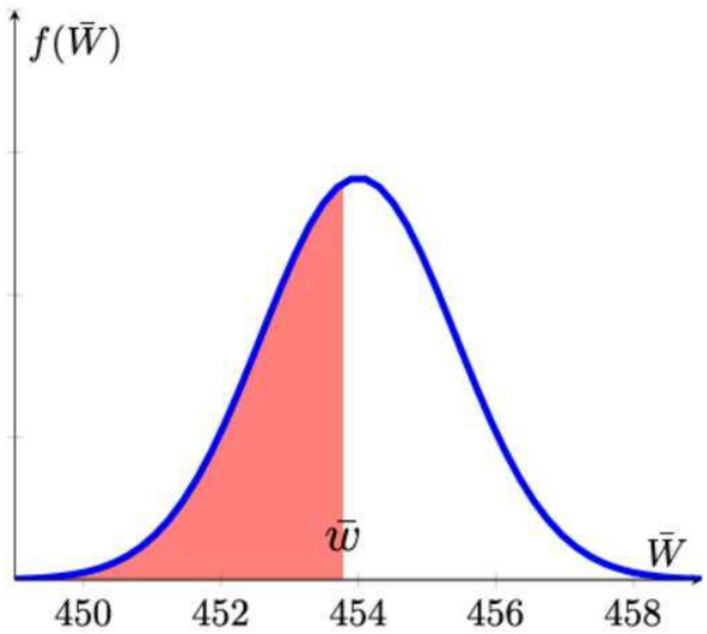
$$H_0 : \mu \geq 454$$
$$H_A : \mu < 454$$

The p-value is the probability to obtain a sample mean of $\bar{w} = 453.78$ (or more extreme) if in truth the population mean was 454. There are two elements of implementation we need to explain. First, what does "or more extreme" mean in this context. We are thinking from the alternative hypothesis here. The evidence that would make us reject H_0 is very low sample means. So more extreme than the actual sample mean in this example would be sample means even lower than $\bar{w}$. Second, in order to calculate a p-value we need to fix a hypothesised population mean which comes from the null hypothesis. But H_0 is $\mu \geq 454$ which includes a range of potential population means. We shall always fix to the value at the margin, i.e. in this example $\mu = 454$.

With this in mind we now know that we need to calculate $Pr(\bar{W} < 453.78)$.

$$Pr(\bar{W} < 453.78) = Pr\left(Z < \frac{453.78 - 454}{\sqrt{100/50}}\right)$$

To calculate the probability we refer to the sampling distribution of $\bar{W}$. Graphically this is represented by the red area in the following picture:



We can complete the calculation.

$$
\begin{aligned}
Pr(\bar{W} < 453.78) &= Pr\left(Z < \frac{453.78 - 454}{\sqrt{100/50}}\right) \\
&= Pr\left(Z < -\frac{0.22}{1.4142}\right) = Pr(Z < -0.1556) \\
&= 0.4364 \text{ from the Normal distribution table}
\end{aligned}
$$

The probability to obtain a sample mean of $\bar{w} = 453.78$ (or more extreme, i.e. lower) if in truth the population mean was 454 is 43.65%.

In the above example we calculated the p-value to be almost 44%. What do we conclude from here about H_0? Well, we would have to conclude that it is indeed quite likely to have obtained a sample mean of 453.78 or lower if the null hypothesis was true. This means that this sample did not deliver strong evidence against the null hypothesis. And this is indeed how you should think about p-values. The lower the p-value, the less likely it is that you should have obtained the sample you have if the null hypothesis was true.

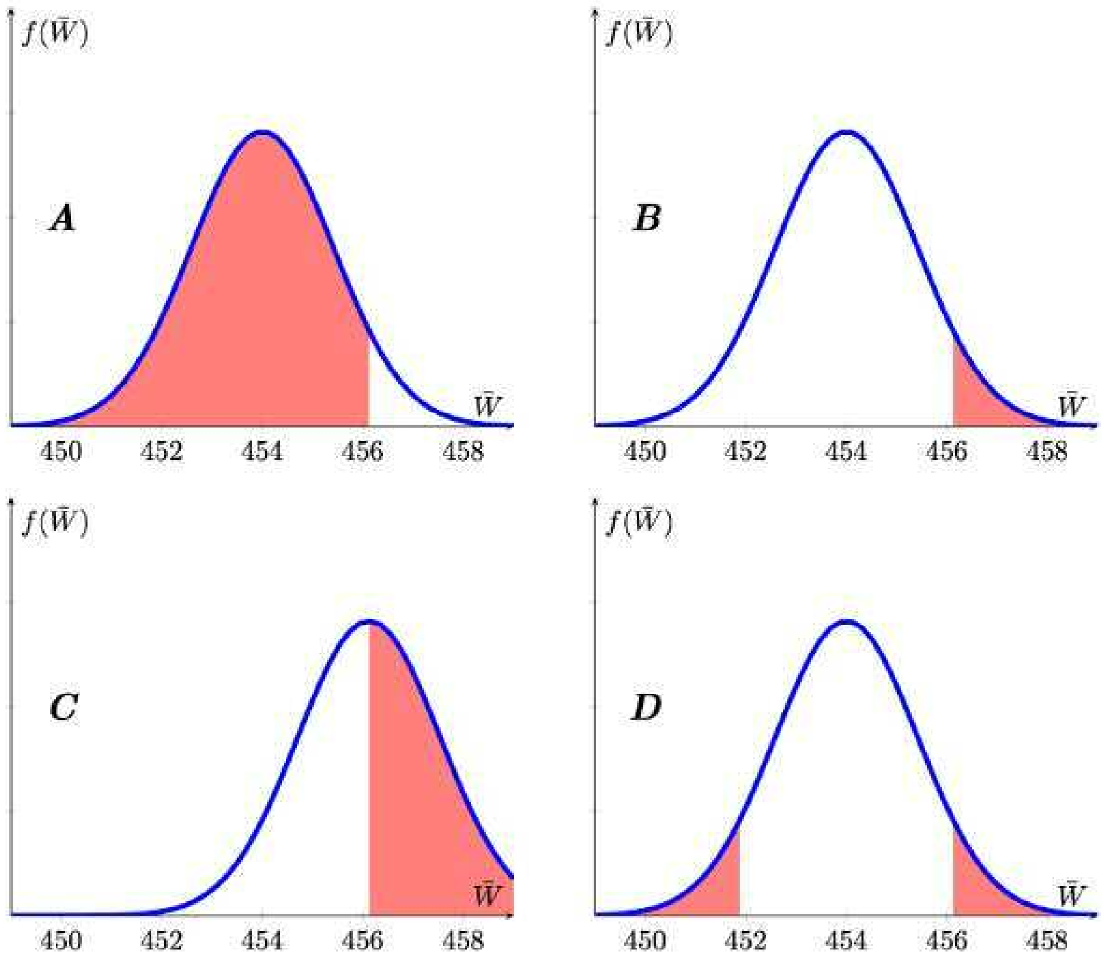
In order to calculate this p-value we had to assume that the null hypothesis was true (here $\mu = 454$). This is very important to remember as it means that the p-value is not the probability that H_0 is true. It would actually be nice if we could calculate that probability of the null hypothesis being true. Unfortunately, just on the basis of the sample information we cannot calculate that probability.

You measure the weight of the jam (W) in a random sample of 40 jam jars. You obtain a sample mean of $\bar{w} = 456.12$. You know the standard deviation of the filling machine to be 10g and indeed you know that the distribution of the weight coming from the filling machine is a normal distribution.

What is the p-value when testing the following hypotheses?

$$
\begin{aligned}
H_0 &: \mu \leq 454 \\
H_A &: \mu > 454
\end{aligned}
$$

(1) Which of the following pictures best reflects the situation you are confronted with (red areas representing p-values).



(a) Picture **A**

(b) Picture **B**

(c) Picture **C**

(d) Picture **D**

(2) What is the p-value?
$Pr(\bar{W} > 456.12) =$ ___________

(3) What is the correct interpretation of the p-value?

(a) The p-value represents the probability that H_0 is correct.

(b) The p-value represents the probability that H_A is correct.

(c) The p-value represents the probability that assuming H_0 is correct we should obtain a sample (or more extreme) as we did.

(d) The p-value represents the probability that the sample mean is larger than 0.

Question 8: (1 point)

In this example you obtained a p-value of 9%. This means that, if the null hypothesis was true, there is a probability of 9% to get a sample as extreme (meaning as high or higher than the sample mean). What does that mean? Would you now reject the null hypothesis or not? Is this p-value low enough for you to reject H_0? All we can say at this stage is that, if the null hypothesis was true we should expect to see sample as extreme as this around 1 in 10 times.

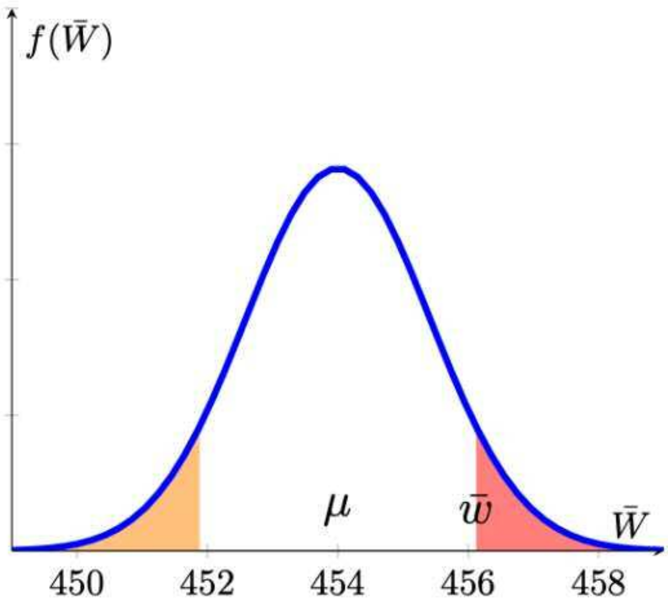
In the previous two examples we explored how to calculate p-values for one-sided hypothesis tests. In the following example we explore how to calculate p-values for a two-sided alternative. We keep working with the same information as the previous example but change the hypothesis.

You measure the weight of the jam (W) in a random sample of 40 jam jars. You obtain a sample mean of $\bar{w} = 456.12$. You know the standard deviation of the filling machine to be 10g and indeed you know that the distribution of the weight coming from the filling machine is a normal distribution.

What is the p-value when testing the following hypotheses?

$$H_0 : \mu = 454$$
$$H_A : \mu \neq 454$$

Here the alternative is a two sided alternative. Evidence against the null hypothesis would now be collected from sample means which are either sufficiently smaller or larger than the hypothesised population mean ($\mu = 545$). Recall that the p-value looks for the probability to observe a sample with a sample mean at least as extreme as the one observed. Let's look at the image which illustrates this situation. The obtained sample mean is $\bar{w} = 456.12$ and at least as extreme as that would immediately be associated with values larger or equal than that value (the probability indicated by the red area). But we would have also rejected the null hypothesis for sample means that are at least as far away from the hypothesised population mean but smaller (the probability indicated by the orange area).



$$\underbrace{Pr(\bar{W} < 451.88)}_{orange} + \underbrace{Pr(\bar{W} > 456.12)}_{red}$$
$$= Pr\left(Z < \frac{451.88 - 454}{\sqrt{100/40}}\right) + Pr\left(Z > \frac{456.12 - 454}{\sqrt{100/40}}\right)$$
$$= Pr\left(Z < -\frac{2.12}{1.5811}\right) + Pr\left(Z > \frac{2.12}{1.5811}\right)$$
$$= Pr\,(Z < -1.3418) + Pr\,(Z > 1.3418)$$
$$= 2 \cdot Pr\,(Z < 1.3418) \text{ by symmetry of N dist}$$
$$= 2 \cdot 0.0901 = 0.1802 \text{ from the Normal distribution table}$$

We conclude that the probability of getting a sample outcome as extreme as the one we got, assuming that the null hypothesis is true, is 18%. This is a sizeable probability and we should expect a sample as extreme as the one we had (deviating from the hypothesised population mean as much as our sample) once in every five samples (as the p-value is approximately 20%).

Calculate p-values for the following examples. In all examples assume that the sampling distribution is a normal distribution.

What is the p-value when testing the following hypotheses?

(1)

Random Variable: X: $\sigma_X = 4$
Sample: $n = 20, \bar{x} = 7.3$

$$H_0 : \mu \geq 8$$
$$H_A : \mu < 8$$

$Pr(\bar{X} < 7.3) =$ ___________

(2)

Random Variable: Z: $\sigma_Z = 6$
Sample: $n = 40, \bar{x} = 11.9$

$$H_0 : \mu \leq 9$$
$$H_A : \mu > 9$$

$Pr(\bar{Z} > 11.9) =$ ___________

(3)

Random Variable: Q: $\sigma_Q = 7$
Sample: $n = 100, \bar{x} = -0.372$

$$H_0 : \mu = 1$$
$$H_A : \mu \neq 1$$

$Pr(\bar{Q} < -0.372) + Pr(\bar{Q} > 2.372)$ ___________

So you calculated a range of different p-values. The smaller the p-value the stronger is the evidence from the sample against the null hypothesis.

Question 9: (0 points)

Introduction

In the previous section we discussed how to calculate p-values (always assuming you know the population variance and the distribution of the sampling distribution). We discussed that smaller p-values indicate stronger evidence (from the sample) against H_0. In some sense that is all there is to know from the methodology of hypothesis testing. However, hypothesis testing results in a binary decision, either reject or do not reject the null hypothesis. If we want to move from a continuous p-value to a binary decision we need to provide a decision rule.

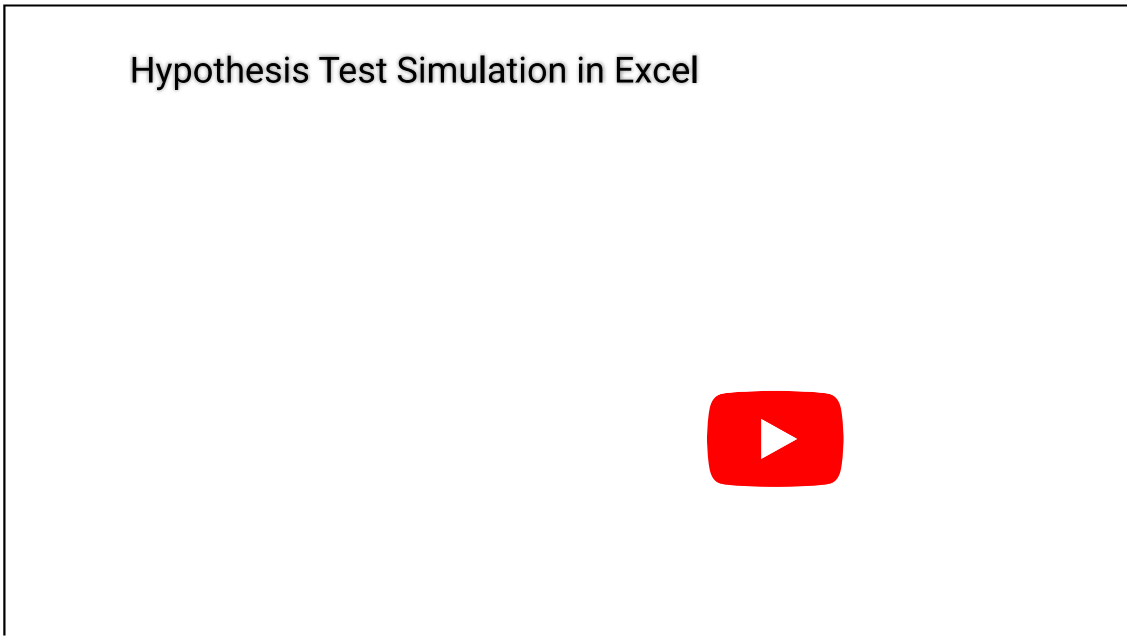
Let's consider for a moment you would always reject the null hypothesis if the p-value was smaller than 0.05. What that implies is that if you were to perform 100 tests for which the null hypothesis is correct, then you should expect to reject 5 (5%) of these. So in other words, by setting such a decision rule you are controlling the probability for a Type 1 error.

Example decision rule: **Reject H_0 if p-value is smaller than 0.05 (significance level, α).**

What about the Type 2 error? A Type 2 error is when we fail to reject an incorrect null hypothesis. It is the situation where we hypothesise that μ is the population mean but in actual fact the population mean is $\mu + \delta$ where $\delta \neq 0$. For very small values of $|\delta|$ it is very unlikely that we reject the incorrect null hypothesis as the difference between the hypothesised population mean and the actual population mean is very small. In other words, the probability of making a Type 2 error is large. The larger the discrepancy (larger $|\delta|$), the smaller is the probability that we make a Type 2 error. However, once you set the decision rule like above you have to accept a certain probability for the Type 2 error. It turns out that you could decrease the probability of a Type 1 error, i.e. reject H_0 only if the p-value, say, is smaller than $\alpha = 0.01$, but if you do that, then you will be increasing the probability of a Type 2 error.

For a given sample size you can either control the Type 1 error probability (and that is what we usually do) or the Type 2 error probability but not both. However, increasing the sample size can reduce the probability of a Type 2 error while keeping the Type 1 error probability constant.

In this video we use a simulation to explain how sample size and significance level relate to the probabilities of making Type I or Type II errors.



This video uses the following Excel file (but it is not important to understand the details of what is done in the file, focus on the message):
https://manchester.mobius.cloud/web/Econ1007011/Public_Html/Simulated%20hypothesis%20test.xlsx
(/web/Econ1007011/Public_Html/Simulated%20hypothesis%20test.xlsx)

Question 10: (0 points)

An Alternative Decision Rule Design

As stated above, the decision rule for a hypothesis test can be formulated as follows:

Reject H_0 if p-value is smaller than α (the significance level).

The significance level (α) is to be set by the researcher in advance and the calculation of the p-value will depend on the sample evidence and on the hypotheses formulated. As the p-value has a very clear (if somewhat awkward) interpretation (The probability of getting an outcome which is at least as extreme as the sample result **assuming** that the null hypothesis is true), it is very useful to think of decision rules in this manner.

However, you may come across alternative decision rules (which, if formulated correctly) will always give you the same result as using the p-value decision rule. Consider the following situation: You have exactly the same information as we assumed previously, you know your sample size n, you know that the sampling distribution of the sample mean is a normal distribution and you have set a significance level α.

With this information we can return to a previous example and as a slightly different question.

Consider a random variable R, which you **hypothesise** to have expected value $E[R] = \mu = 10$ and you want to test this against the alternative that the population mean is larger than that:

$$H_0 : \mu \leq 10; H_A : \mu > 10$$

Further, you know that the population variance is $Var[R] = \sigma^2 = 16$. While you do not know how R is distributed, you do have a sample of size $n = 100$ which you are confident to be large enough to apply the CLT.

With that information and assuming that $\mu = 10$, you can derive the sampling distribution:

$$\begin{aligned} \bar{R} &\sim N\left(\mu, \sigma_{\bar{R}}^2 = \frac{\sigma^2}{n}\right) \\ &\sim N\left(10, \frac{16}{100}\right) \end{aligned}$$

or in standardised form

$$\frac{\bar{R} - \mu}{\sigma_{\bar{R}}} \sim N(0, 1)$$

Before looking at the actual sample mean, what type of sample evidence would make you reject H_0 if you set the significance level at $\alpha = 0.05$? In other words, we want to find a value $\bar{r}^{high}$ which will trigger a rejection of the null hypothesis. This should happen if $Pr(\bar{R} > \bar{r}^{high} | \mu = 10) < \alpha = 0.05$, where the conditioning on the assumption that $\mu = 10$ is made explicit.

$$Pr\left(\bar{R} > \bar{r}^{high} | \mu = 10\right) = Pr\left(Z > \frac{\bar{r}^{high} - 10}{\sqrt{\frac{16}{100}}} | \mu = 10\right) = 0.05$$

For ease of notation we will not show the conditioning ($|\mu = 10$) any longer, but you should always keep in mind that we are only able to make these calculations as we are setting μ to the hypothesised value.

We reformulated the problem stated in terms of the random variable $\bar{R}$ into one formulated in terms of the standard normally distributed random variable Z. The reason for doing so is that we can tell from a standard normal distribution table that $Pr(Z > 1.645) = 1 - Pr(Z \leq 1.645) = 1 - 0.95 = 0.05$.

Cumulative Area Under the Standard Normal Distribution

Example: P(Z<-2.54) = 0.0055, P(Z>=-2.54) = 1 - P(Z<-2.54) = 0.9945

Z	0	1	2	3	4	5	6	7	8	9
-3.00	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.90	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.80	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
...										
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706

Extract from the standard normal distribution table, identifying that $Pr(Z > 1.645) = 1 - Pr(Z \leq 1.645) = 1 - 0.95 = 0.05$. The required Z value is somewhere between 1.64 and 1.65. We average to 1.645.

From here we can infer that

$$\frac{\bar{r}^{high} - 10}{\sqrt{\frac{16}{100}}} = 1.645$$

$$\bar{r}^{high} - 10 = 1.645 \cdot \sqrt{\frac{16}{100}} = 1.645 \cdot 0.4$$

$$\bar{r}^{high} = 10 + 1.645 \cdot 0.4 = 10.658$$

Therefore, **if the sample mean turns out to be higher than 10.658, we shall reject $H_0 : \mu \leq 10$ at a 5% significance level**. This can be seen as a decision rule.

This video applies the same thinking as that in the above example.

Hypothesis Testing using various decision rules

<\p>

Question 11: (0 points)

We have used all the knowledge we accumulated previously to design a decision rule prior to actually observing the sample mean. We needed knowledge of the test statistic (the sample mean here) its sampling distribution, the sample size (n) and the significance level (α).

Right-tailed Hypothesis

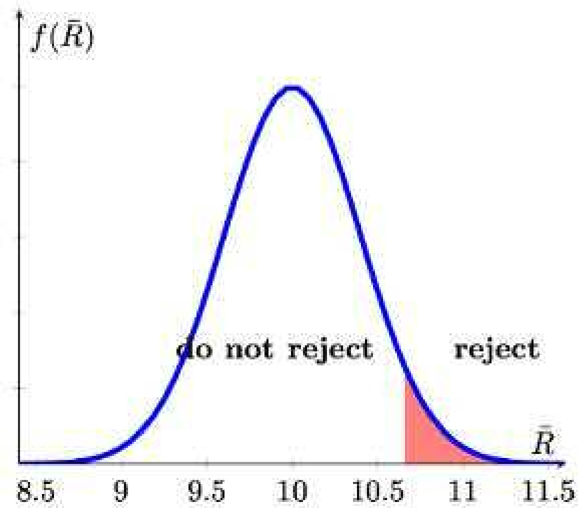
With all that knowledge we were able to derive the following decision rule.

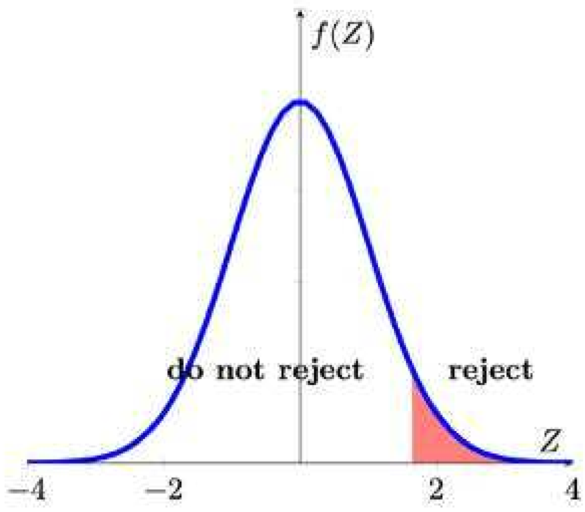
- For testing $H_0 : \mu \leq 10; H_A : \mu > 10$ the decision rule is

Reject H_0, at a significance level $\alpha = 0.05$, if the sample mean exceeds 10.658. Which is equivalent to saying **Reject H_0, at a significance level $\alpha = 0.05$, if the test statistic $Z = \frac{\bar{r} - 10}{\sqrt{\frac{16}{100}}}$ exceeds 1.645.**

It most common to see these decision rules formulated in terms of the Z statistic rather than the sample mean. The values at which we would reject H_0 are also called **critical values**.

Graphically we can represent the situation as follows:



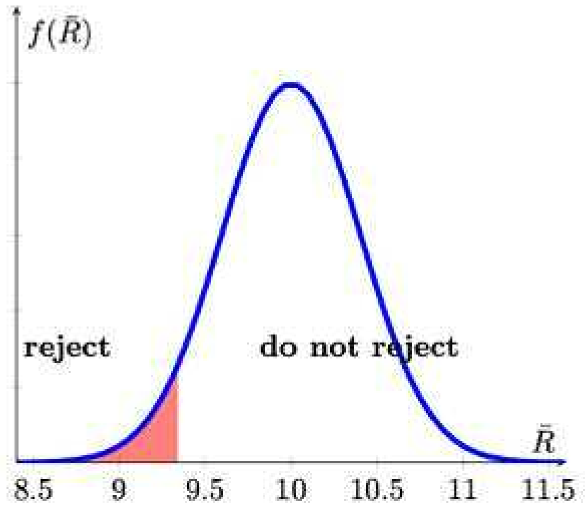


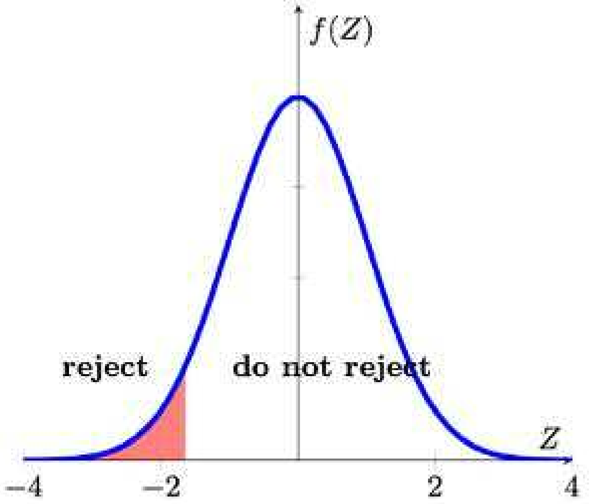
Left-tailed Hypothesis

If we had wanted to to test a left-tailed hypothesis, the decision rule would have been

- For testing $H_0 : \mu \geq 10; H_A : \mu < 10$ the decision rule is

Reject H_0, at a significance level $\alpha = 0.05$, if the sample mean is smaller than 9.342 (=10-0.658). Which is equivalent to saying **Reject H_0, at a significance level $\alpha = 0.05$, if the test statistic $Z = \frac{\bar{r} - 10}{\sqrt{\frac{16}{100}}}$ is smaller than -1.645.**



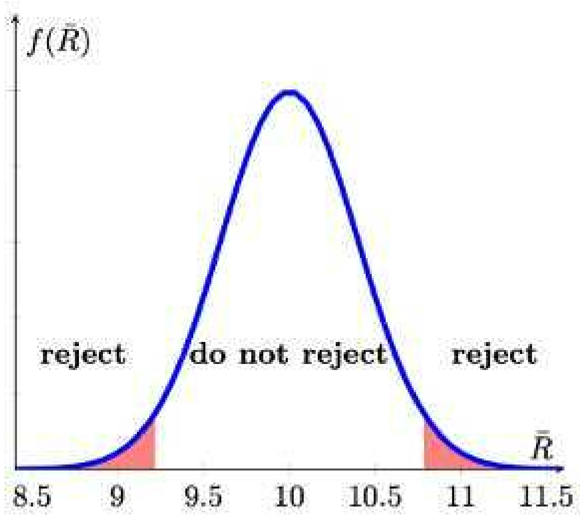


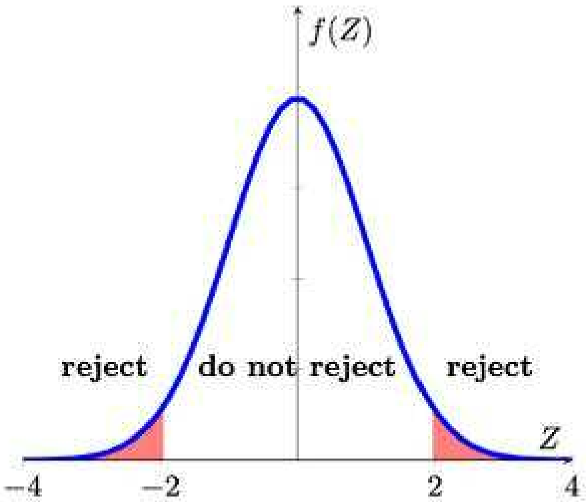
Two-tailed Hypothesis

Lastly we have to establish how this type of decision rule would look if we were to test a two-sided hypothesis. Let is first state the hypothesis and the resulting decision rule in our example and then explain it graphically.

- For testing $H_0 : \mu = 10; H_A : \mu \neq 10$ the decision rule is

Reject H_0, at a significance level $\alpha = 0.05$, if the sample mean exceeds 10.784 or is smaller than 9.216. Which is equivalent to saying **Reject H_0, at a significance level $\alpha = 0.05$, if the test statistic $Z = \frac{\bar{r} - 10}{\sqrt{\frac{16}{100}}}$ exceeds 1.96 or is smaller than -1.96.**





The 5% probability (α) is now divided into two 2.5% regions in each tail (red areas). So where do the Z values of -1.96 and 1.96 come from? They do come from the standard normal distribution table as the values which cut off 0.025 in the left and right tail respectively.

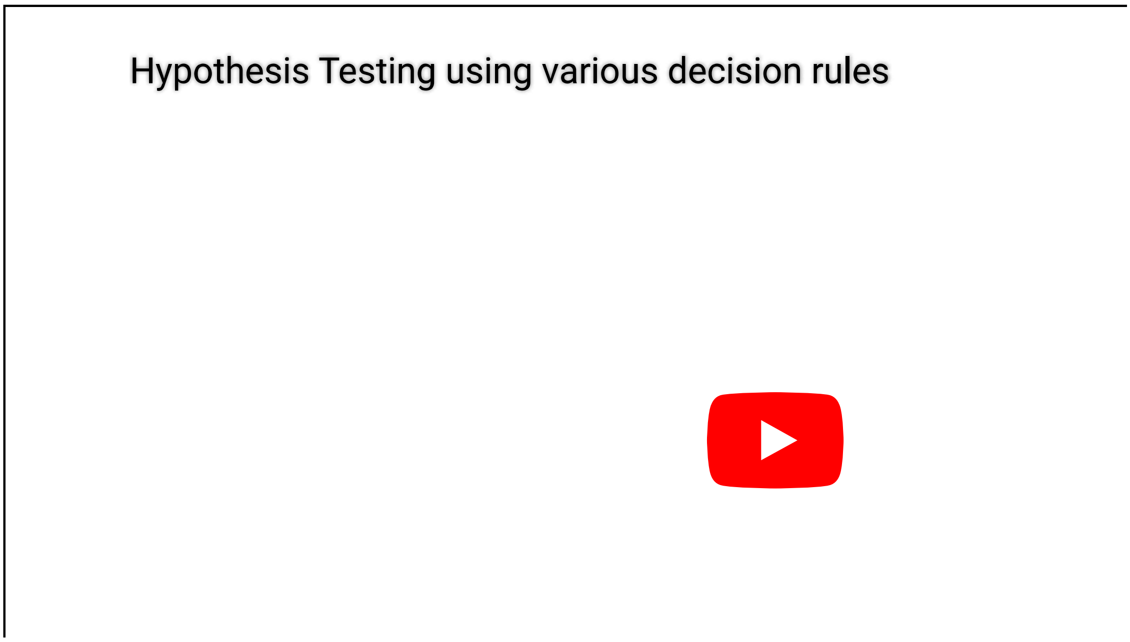
Cumulative Area Under the Standard Normal Distribution

Example: P(Z<-2.54) = 0.0055, P(Z>=-2.54) = 1 - P(Z<-2.54) = 0.9945

Z	0	1	2	3	4	5	6	7	8	9
-3.00	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.90	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.80	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
				...						
-2.10	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.00	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.90	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.80	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
				...						
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857

Extract from the standard normal distribution table, identifying that $Pr(Z > 1.96) = 1 - Pr(Z \leq 1.96) = 1 - 0.975 = 0.025$. Also $Pr(Z < -1.96) = 0.025$.

In this video we show how the application of different decision rules all lead to the same decision in a hypothesis test.



Question 12: (1 point)

Summary and Example

Looking back at the previous calculations it should be obvious that the decision rules could be calculated/formulated before you observe the sample value. The decision rule was dependent on

- the hypothesised population value μ,
- the sample size n,
- the sampling distribution, and
- the chosen significance level.

Once you have formulated the decision rule you merely need to observe the sample mean and calculate $Z = \frac{\bar{r}-10}{0.4}$. Once you have done that you can easily see whether your calculated Z value falls in the rejection region or not.

Why hypothesis testing?

After the previous discussion you may wonder why the discipline has developped the tradition of hypothesis testing. One explanation is that it has been used as a decision tool. This means that the information coming from the hypothesis test has been used to decide whether a certain course of action should be followed or not.

Indeed, if the information from the hypothesis test is the only information used to decide on a course of action, then using a formal hypothesis testing framework (which means presenting a significance level α, a pair of hypothesis and a resulting decision rule) will force you to consider the probabilities of committing a Type 1 and 2 error. The probability of making a Type 1 error is explicitly encoded as $Pr(\text{Type 1 error}) = \alpha$. The probability of the Type 2 error is not explicit, but the applied researcher will be aware that its probability is inversely related to the probability of a Type 1 error.

In the following exercises we ask you to formulate decision rules.

In all examples assume that the sampling distribution is a normal distribution.
Formulate decision rules.

(1)
Random Variable: X: $\sigma_X = 4$
Sample: $n = 20$
Hypothesis: $H_0 : \mu \leq 25; H_A : \mu > 25$
Significance level: $\alpha = 0.05$

Reject H_0 if p-value is < 0.05.
Reject H_0 if Z is > 1.645.
Reject H_0 if $\bar{x}$ is > 26.4713.

If the sample mean is 26.8. What is your decision?

(a) Reject

(b) Do not reject

(2)
Random Variable: Y: $\sigma_Y = 6$
Sample: $n = 40$
Hypothesis: $H_0 : \mu \geq 0; H_A : \mu < 0$
Significance level: $\alpha = 0.10$

Reject H_0 if p-value is < 0.1.
Reject H_0 if Y is < -1.28.
Reject H_0 if $\bar{y}$ is < -1.2143.

If the sample mean is -0.8. What is your decision?

(a) Reject

(b) Do not reject

(3)

Random Variable: Q: $\sigma_Q = 7$
Sample: $n = 100$
Hypothesis: $H_0 : \mu = 12; H_A : \mu \neq 12$
Significance level: $\alpha = 0.01$

Reject H_0 if p-value is < 0.01.
Reject H_0 if Z is < -2.575 or > 2.575.
Reject H_0 if $\bar{q}$ is < 10.1975 or > 13.8025.

If the sample mean is 13.5. What is your decision?

(a) Reject

(b) Do not reject

As you can see from the prior examples, the sample size plays a crucial role in the calculation of the critical values in a hypothesis test. By increasing the sample size, you can actually decrease the probability of making a Type 2 error (not rejecting an incorrect H_0). The mechanism at work is that the standard error of your sample statistic (here mainly the sample mean but it could be any other sample statistic) is an inverse function to the sample size n:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

In an earlier example we worked with the following information:

Random Variable: Q: $\sigma_Q = 7$
Sample: $n = 100$
Hypothesis: $H_0 : \mu = 12; H_A : \mu \neq 12$
Significance level: $\alpha = 0.01$

The resulting decision rule was:

Reject H_0 if $\bar{q}$ is < 10.1975 or > 13.8025.

How large would the sample size have to be such that a sample mean of 13.5 will trigger a rejection of H_0?
The critical value, in this example is calculated according to (using $Pr(Z > 2.575) = \alpha/2$).

$$13.8025 = 12 + 2.575 \cdot (7/\sqrt{100})$$

We now wonder what should be the n in

$$13.5 = 12 + 2.575 \cdot (7/\sqrt{n})$$

All we need to do is to solve the above equation for n

$$\begin{aligned} 13.5 &= 12 + 2.575 \cdot (7/\sqrt{n}) \\ 13.5 - 12 &= 2.575 \cdot (7/\sqrt{n}) \\ \sqrt{n} &= \frac{2.575 \cdot 7}{1.5} = 12.0167 \\ n &= 12.0167^2 = 144.4003 \end{aligned}$$

This means that the sample size required is 144.4003, or in practice 145.

Consider the following knowledge

Random Variable: X: $\sigma_X = 4$

Hypothesis: $H_0 : \mu \leq 25; H_A : \mu > 25$

Significance level: $\alpha = 0.05$

How large a sample would you need to make sure that you can reject the null hypothesis (at $\alpha = 0.05$) if the sample mean is 0.01 larger than the hypothesised population mean?

The sample size required is ____________ , or in practice ____________.

Question 13: (0 points)

t-distribution

So far we only dealt with the situation in which the sampling distribution was a normal distribution;

$$\bar{X} \sim N(\mu, \sigma_{\bar{X}})$$

There are two situations in which we established that the sampling distribution of the sample mean is a normal distribution.

1. The population distribution is a normal distribution itself and we know the population variance, σ_X, $X \sim N(?, \sigma^2)$. The population is unknown (if it was known there was no need for hypothesis testing of the mean). In this situation $\bar{X} \sim N(\mu, \sigma^2_{\bar{X}})$ at any sample size.
2. We know the population variance, σ_X but don't know the distribution the population random variable X follows, but the sample size is sufficiently large for $\bar{X}$ to be normally distributed (by the power of the Central Limit Theorem, CLT).

It is worth repeating the main conclusion of the CLT here:

If $\bar{X}$ is obtained from a random sample of size n from a population with mean μ and variance σ^2, then, irrespective of the distribution sampled,

$$\frac{\bar{X} - \mu}{SE\left[\bar{X}\right]} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} \to N(0, 1) \quad \text{as } n \to \infty.$$

That is, the probability distribution of $\frac{\bar{X} - \mu}{SE\left[\bar{X}\right]}$ approaches the standard normal distribution as $n \to \infty$.

So, both of the earlier results rely on us "knowing" the standard deviation σ of the random variable X. This is the reason why, throughout this section, when we presented examples we always gave you information on this standard deviation. This created somewhat awkward situations where we said that the population mean μ was not known but the population standard deviation σ was known. That is, of course, a situation which often is unrealistic.

We therefore need to consider the case where σ is also unknown. When that is the case (almost always!) then we will replace the unknown σ with a sample estimate s, which we learned how to calculate earlier. This means that we will be operating with the following term

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

This term is commonly called a t statistic for reasons that will soon become obvious. Fortunately, all the principles discussed previously continue to apply when we do so. Here we will point out what details will change.

In what follows we will discuss four realistic cases (all of which assume that σ is unknown and is estimated by s). For each of these four cases we will want to establish how t is distributed. The complication that arises is that the t statistic, now, incorporates two random variables, $\bar{X}$ the sample mean and S the sample standard deviation (which is why it is written as a capitalised letter in the above t-statistic). This makes deriving the distribution of t less straightforward.

If the result is that $t \sim N(0,1)$ then all the details discussed previously will apply, in particular how you can calculate p-values or critical values and therefore make decisions on hypotheses. If the answer is anything else but $t \sim N(0,1)$ we will have to discuss the changed procedures. The following table outlines the four cases (all assuming that σ needs to be estimated with s):

	Sample Size (n)	
Dist of X	small	large
Normal	$t \sim ?$	$t \sim ?$
not Normal	$t \sim ?$	$t \sim ?$

We shall briefly discuss these four cases in turn.

- $X \sim N$, n small. In this case $t \sim t - distribution$ with $n-1$ degrees of freedom (t_{n-1}). (You will soon learn about this distribution).
- $X \sim N$, n large. In this case $t \sim t - distribution$ with $n-1$ degrees of freedom (t_{n-1}) but you will soon learn that for large n this implies $t \sim N(0,1)$.
- $X \sim not\ N$, n small. In this case $t \sim ?$. In other words we cannot specify the type of distribution, as it is hugely data dependent. If you wanted to make inference in such a case you will have to apply data-driven methods which are beyond the scope of this course unit (like bootstrapping).
- $X \sim not\ N$, n large. In this case $t \sim N(0,1)$. This is thanks to (another) CLT (one which allows for unknown σ).

We can therefore complete the above table:

	Sample Size (n)	
Dist of X	small	large
Normal	$t \sim t_{n-1}$	$t \sim N(0,1)$
not Normal	$t \sim ?$	$t \sim N(0,1)$

This implies that, whenever we have a large sample size the procedures detailed earlier apply exactly, just with the sample standard deviation, s, replacing the role of the known σ.

Details of why the t statistic follows a t distribution are not for this course unit. But it is worth noting that a t distribution arises when you combine a normally distributed random variable, $\bar{X}$, with a χ^2 distributed random variable, S^2. Details are not important other than to realise that t combines two random variables. You will soon encounter the chi^2 distribution in another context, which follows from a different combination of multiple random variables.

Question 14: (1 point)

Let us work through an example that represents a situation where we are confident that the CLT applies. As discussed previously, an often quoted "rule" is that $n > 30$ generally allows you to make this judgement. I would certainly say that with $n > 100$ you are pretty much on the safe site. If you have any doubt, you should just state your doubt.

We have the following information:

Random Variable: X distribution unknown, note that σ is unknown
Sample: $n = 80$, $s = 4$, assume that CLT applies
Hypothesis: $H_0 : \mu \leq 25; H_A : \mu > 25$
Significance level: $\alpha = 0.05$

With that information we should be able to formulate decision rules. We will do that in three versions, in terms of the p-vale, the test statistic (here a t-test) and the sample statistic (here $\bar{x}$). In applications one is sufficient as they all give the same result.

- Reject H_0 if p-value is $< 0.05.$

  This is merely the statement of the general p-value decision rule using the chosen significance level.
- Reject H_0 if t is > 1.645. The value from 1.645 comes from the normal distribution table as in this situation $t \sim N(0, 1)$.
- Reject H_0 if $\bar{x}$ is > 25.7357. The critical value is calculated as $25 + 1.645 * (4/\sqrt{80})$.

Note that we could formulate these decision rules before we know the value of the sample mean. In fact the first two rules can be formulated without knowing anything about the sample other than the sample size, n. For the third we also needed to know the sample standard deviation, s.

If the actual sample mean was 26.12, what would be your decision? From here it is easiest to apply the last decision rule. We would reject H_0 as $26.12 > 25.7357$. We could also calculate the t-statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{26.12 - 25}{4/\sqrt{80}} = 2.5044$$

According to the 2nd decision rule we would also reject. And on the basis of this t-statistics we could also calculate the p-value, which is 0.0062. This is smaller than the significance level of 0.05 and leads us (as it should) to the same rejection decision.

For both examples you can assume that the sample size is large enough for a CLT to be applicable.

(1)

Random Variable: Y is normally distributed
Sample: $n = 40, s = 6$
Hypothesis: $H_0 : \mu \geq 0; H_A : \mu < 0$
Significance level: $\alpha = 0.10$

Reject H_0 if p-value is < 0.1.
Reject H_0 if the t-statistic is < -1.28.
Reject H_0 if $\bar{y}$ is < -1.2162.

If the sample mean is -0.8. What is your decision?

(a) Reject

(b) Do not reject

(2)

Random Variable: Q distribution is unknown

Sample: $n = 100, s = 10$

Hypothesis: $H_0 : \mu = 13; H_A : \mu \neq 13$

Significance level: $\alpha = 0.01$

Reject H_0 if p-value is < 0.01.

Reject H_0 if t-statistic is < -2.575 or > 2.575.

Reject H_0 if $\bar{q}$ is $<$ ____________ or $>$ ____________ .

The sample mean is 13.1.

What is the value of the test statistic? $t =$ ____________

What is your decision?

(a) Reject

(b) Do not reject

Question 15: (1 point)

Working with the t-distribution

This now leaves us to discuss how to work the case where we know that the random variable from which we sample is normally distributed, the population standard deviation is unknown and the sample size is small (such that we are not confident that the CLT applies). In the above table it was indicated that, in such a situation,

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

This is a table with propabilities from the t-distribution (/web/Econ1007011/Public_Html/tTable.pdf).

Let's say we want to perform a hypothesis test in such a setting. Let's work through an example.

We have the following information:

Random Variable: X distribution is normal with σ unknown
Sample: $n = 20$, $s = 4$
Hypothesis: $H_0 : \mu \leq 25; H_A : \mu > 25$
Significance level: $\alpha = 0.05$

With that information we should be able to formulate decision rules. We will do that in three versions, in terms of the p-vale, the test statistic (here a t-test) and the sample statistic (here $\bar{x}$). In applications one is sufficient as they all give the same result.

- Reject H_0 if p-value is < 0.05.

  This is merely the statement of the general p-value decision rule using the chosen significance level.
- Reject H_0 if t is > 1.729. The value from 1.725 comes from the t distribution with 19 ($= n - 1 = 20 - 1$) degrees of freedom as, in this situation, we know that $t \sim t_{n-1}$.
- Reject H_0 if $\bar{x}$ is > 26.5465. The critical value is calculated as $25 + 1.729 * (4/\sqrt{20})$.

If the actual sample mean was 26.12, what would be your decision? From here it is easiest to apply the last decision rule. We would not reject H_0 as $26.12 < 26.5465$. We could also calculate the t-statistic:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{26.12 - 25}{4/\sqrt{20}} = \frac{1.12}{0.8944} = 1.2522$$

According to the 2nd decision rule we would also not reject as the t-statistic does not exceed its critical value of 1.729. And on the basis of this t-statistics we could also calculate the p-value. From the t-distribution table we know that the critical value for a one sided t test at $\alpha = 0.1$ and 19 degrees of freedom (dof) is 1.328. As the t-statistic is 1.2522, we can say that the p-value has to be larger than 0.10. We cannot give any more precise p-value using the table. You can use Excel to give you a more precise p-value. Use the formula "=1-T.DIST(1.2522,19,TRUE)" which delivers the precise p-value of 0.1128. In any case, as our significance level is 0.05 we fail to reject the null hypothesis. As it should be, all decision rules give the same result.

This is a work through a t-test with a two-tailed H_A

Hypothesis test for a population mean using the t-distribution

The University's welfare manager is concerned about students spending too much on unhealthy food items every week (let S be the weekly spend on unhealthy food measured in pounds). He asks a random sample of $n = 20$ students to keep a weekly food shopping diary. From previous studies the manager is confident that S is normally distributed. Five years ago the average spend was 61 pounds and she is concerned that this value has increased since.

From the sample the welfare manager learns that $\bar{s} = 79$ and $s = 10$. Test the following hypotheses, using $\alpha = 0.05$:

$$H_0 : \mu = 61$$
$$H_A : \mu > 61$$

What is the distribution of the test-statistic $t = \frac{\bar{S} - \mu}{s/\sqrt{n}}$?

(a) $N(0, 1)$

(b) $N(0, n)$

(c) t_{10}

(d) t_{20}

(e) t_{19}

What is the Decision Rule?

Reject H_0 if t __________ ____________

What is the value of the test-statistic?

$$t = \frac{\bar{S} - \mu}{s/\sqrt{n}} = \text{\_\_\_\_\_\_\_\_\_\_\_\_}$$

What is the value of the test's p-value?

$p - value <$ ____________

What is the most accurate conclusion?

(a) There is evidence that students spend more on unhealthy foods than 5 years ago.

(b) There is insufficient to suggest that students spend more on unhealthy foods than 5 years ago.

After presenting these results to a colleague working in the economics department she recognises that she has to redo the test acknowledging that in general food prices were affected by inflation over the last five years. In fact food prices increased by 20% over that period.

What should change in the test's setup to acknowledge that food prices increased?

(a) The test should be changed to a left tailed test.

(b) The population mean in the hypotheses should change to $\mu = 73.2.$

(c) The sample size should be increased by 20%.

(d) The significance level should be increased.

(e) The population mean in the hypotheses should change to $\mu = 50.84.$

(f) The significance level should be decreased.

What is the most accurate conclusion after re-doing the test with the appropriate adjustment?

(a) There is evidence that students spend more on unhealthy foods than 5 years ago.

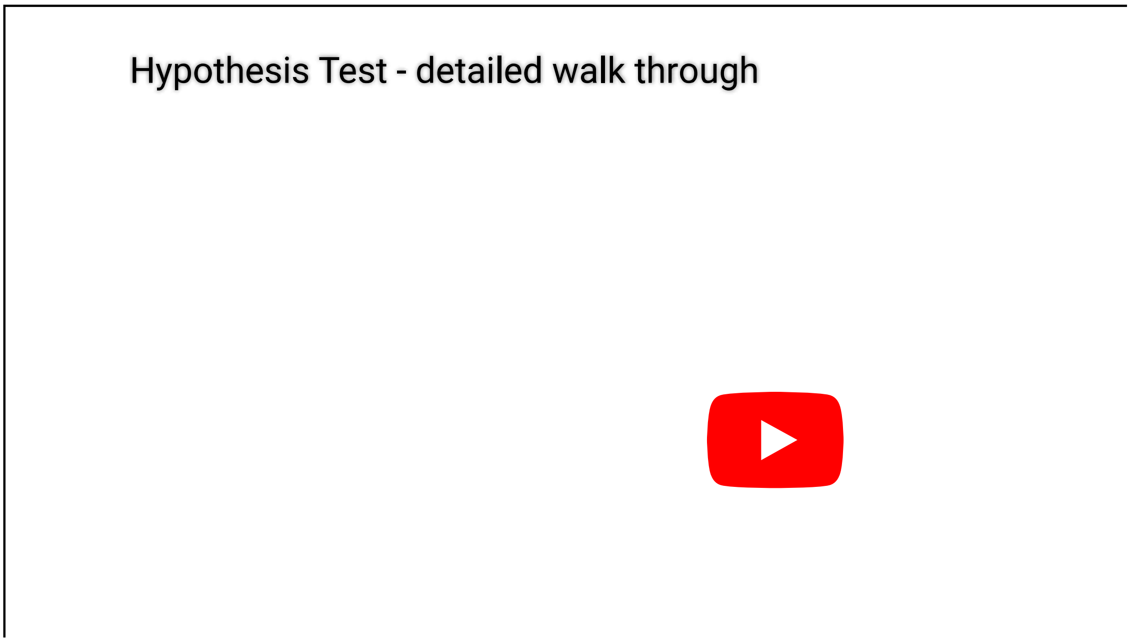
(b) There is insufficient to suggest that students spend more on unhealthy foods than 5 years ago.

Question 16: (0 points)

Summary

The sections above introduced the concept of hypothesis testing. It used your knowledge of sampling distribution to allow you to make (probabilistic statements) about unknown population characteristics, based on just one sample. The example used here is that where we are interested in an unknown population mean.

In this video I walk through two examples and lao explain in detail how you decide what distribution the test statistic follows (assuming the null hypothesis is true). The video uses this scheme (/web/Econ1007011/Public_Html/HT_for_mean_Overview_YT.pdf).



In the next lesson you will learn about other characteristics we may want to investigate using the hypothesis testing principle.

Extra Reading

If you want to know more about the subtleties of p-values, then this is a good read: Lew, M.J. (2020) A Reckless Guide to P-values (https://link.springer.com/chapter/10.1007/164_2019_286): Local Evidence, Global Errors, Part of the Handbook of Experimental Pharmacology book series (HEP,volume 257).