Ch2 - Testing

Statistics For Business and Economics - II

Shaikh Tanvir Hossain

East West University, Dhaka

Last Updated October 29, 2023

Outline

Outline

Hypothesis Tetsing - Key Ideas

Hypothesis Tetsing - Key Ideas

What is hypothesis testing?

What is Hypothesis Tetsing?

- After Estimation, now we will learn another major tool for inferential statistic, this is known as *Hypothesis Testing*, in short *Testing*.
- Hypothesis Testing or Testing is slightly different than estimation. Here we initially start with two (competing) hypotheses about the population parameter (One is called "Null Hypothesis" and the other is called "Alternative Hypothesis").
- Then we use the sample data to reject or accept one of two hypotheses. Question is - What is a hypothesis Ans: It's simply a conjecture about the population parameter. Let's see an example.
- Suppose somehow we have an information that the true population mean of income is less than 24, this means we know $\mu < 24$. Then we can form two hypotheses

$$H_0 : \mu \geq 24$$
$$H_a : \mu < 24$$

- Where H_0 is what we call a *Null Hypothesis* and H_a is what we call an *Alternative Hypothesis*.
- After we are done with constructing the hypotheses, we use the random sample (data) to either *reject Null Hypothesis* or *accept the Null Hypothesis* (sometimes it is written fail to reject). Notice! Everything is happening around Null (Why? We will come back to the answer later!)

Hypothesis Tetsing - Key Ideas

Types of Errors

Types of Errors

- Note that we are making conclusion about the Population using a sample. This seems to be a very difficult task, right? So obviously there is a possibility of making an error. Here are the possible errors.

		Population Reality	
		H_0 True	H_a True
Conclusion	Accept H_0	Correct Conclusion	Type II Error
	Reject H_0	Type I Error	Correct Conclusion

- Before interpreting the table, first of all recall that, we do not know what is the truth. Now the table says, if the hypothesized Null is actually true and after the testing we accept Null, then there is no error,
- But if the hypothesized Null is true and we reject the Null then we will make an error and the error is called Type-I error. Can you interpret other cells of the table?
- Ideally we want to construct a test that minimizes both of these errors, but actually for a fixed sample size, this is almost impossible. So the idea is we fix the Type-I error and look for a test which minimizes the Type-II error.
- So *for a given Type-I error, all of the tests that we will do will minimize Type-II error in theory*.

Hypothesis Tetsing - Key Ideas

Different Formations of Testing

Different Formations of Testing

- Below we will write μ_0 for the hypothesized population mean (rather than 24), and μ is our unknown Population mean
- There are other ways we can form Hypotheses, in particular for all problems in this course, we will use either one of the following three formations,
  - *Two Tail test*

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

  - *Upper Tail Test*

$$H_{0:} : \mu \leq \mu_0$$
$$H_a : \mu > \mu_0$$

  - *Lower Tail Test*

$$H_{0:} : \mu \geq \mu_0$$
$$H_a : \mu < \mu_0$$

- Together the last two tests are called - *One tail tests*. Why the word "tail" and why this naming? you will see in a minute!

Hypothesis Tetsing - Key Ideas

Test Statistic and Sampling Distribution in Testing

Test Statistic and Sampling Distribution in Testing

- In the hypothesis testing again we will use the *sampling distribution of $\bar{X}_n$*.
- Recall before we called $\bar{X}_n$ an estimator of μ.
- In the Hypothesis testing we won't call this an estimator, rather we will call it a *Test Statistic*. In general often a *Test Statistic* is same or very similar to a point estimator, but its a *Statistic*.
- Recall in general, *any* function of the random sample is called a "*Statistic*". When a *Statistic* is used for estimation we call it an *Estimator*. Similarly, when a *Statistic* is used for *Testing* we call it a *Test Statistic*. These are just some naming conventions that you need to know.
- For example $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is a function of random sample, so this is a statistic. When we use it for point estimation, we call this point estimator, but if we use it for hypothesis testing we call it a *Test Statistic*.
- And we already know that the distribution of a *Statistic* is called *Sampling Distribution*.
- So the distribution of $\bar{X}_n$ is also called a sampling distribution (old stuff!)

Tests - σ^2 known (z-test)

Tests - σ^2 known (z-test)

Basic Steps of a Two-Tail Test

Let's see the basic steps of Hypothesis testing. First we will do an example of a two tail test.

- step-1: Write down the Null and Alternative hypotheses. Again this step might be different depending upon the question. For example we can start with

  $$H_0 : \mu = \mu_0$$
  $$H_a : \mu \neq \mu_0$$

  This is called *Two-Tail* test (We will see more details on one-tail test later!)

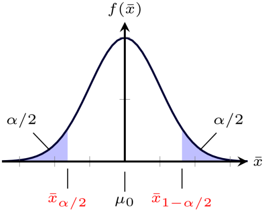
- step-2: Using a sample calculate $\bar{x}$.

- step-3: Fix Type - I error. Conventionally Type - I error is denoted by α, where $0 < \alpha < 1$. This will be essentially a probability, e.g., .05, 0.025, etc. There are different names of *Type I Error*, sometimes we call it also *rejection region*, or *level of significance* or *significance level*. But all are same thing. Divide it equally with an area of $\alpha/2$ in two-tails of the distribution. Essentially this is the possible mistake we can make.

- step-4: In this step we need to use the *sampling distribution* of the test statistic under the Null (this means the sampling distribution of the test statistic assuming the Null Hypothesis is true) and find *two critical values*. For example, if the test statistic is $\bar{X}_n$, then maybe we can use $\bar{X}_n \sim \mathcal{N}(\mu_0, \sigma^2/n)$ and use it for testing, where μ_0 is coming from the Null Hypothesis (for now assume σ^2 is known). Now *critical values* are $\bar{x}_{\frac{\alpha}{2}}$ and $\bar{x}_{1-\frac{\alpha}{2}}$ such that

  $$\mathbb{P}(\bar{X}_n \leq \bar{x}_{\frac{\alpha}{2}}) = \alpha/2 \quad \text{and} \quad \mathbb{P}(\bar{X}_n \leq \bar{x}_{1-\frac{\alpha}{2}}) = 1 - \alpha/2$$

  Visually this means



For example if you set $\alpha = 0.05$ and use R to find critical values, the function is `qnorm(0.025,` μ_0`,` $\sigma/\sqrt{n}$`)` and `qnorm(0.975,` μ_0`,` $\sigma/\sqrt{n}$`)`. So these two numbers are essentially $\alpha/2$ quantile and $1-\alpha/2$ quantile of the sampling distribution under the Null.

- step-5: Reject the Null if

$$\bar{x} \geq \bar{x}_{1-\frac{\alpha}{2}} \text{ or } \bar{x} \leq \bar{x}_{\frac{\alpha}{2}}$$

So just reject the Null if any one of the two happens.

- step-6: Finally conclude that with α probability of error we are rejecting the Null Hypothesis (or more philosophically correct statement - *we failed to accept the Null*).

Questions

- Why did we use the sampling distribution under the Null? Ans: Because we need to use type - I error, and type - I error is in the sampling distribution under the Null.
- What is the intuition behind the test ? (Ans: Explanation on the board?)

Tests - σ^2 known (z-test)

Distributional results of $\bar{X}_n$ that we will use

Distributional Results for $\bar{X}_n$

- Question: Note that in the basic steps, we used the sampling distribution $\bar{X}_n \sim \mathcal{N}(\mu_0, \sigma^2/n)$, why?
- Ans: We know from Theorem 1.3 and Theorem 1.4 (Chapter 1) that the sampling distribution of sample means is Normal either when the population data is Normally distributed or sample size is very large so that we can apply CLT.
- Let's exaplin these two points, the story is the same as the confidence interval.
- **Normality Assumption of the Data:** If the population data is notmal and we assume independence, this means $X_1, X_2, \ldots, X_n$ all follow $\mathcal{N}(\mu, \sigma^2)$. In this case from Theorem 1.3 (iii) of Chapter 1 tells us that $\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$. This holds for any sample size. In this case we have two options,
  - 1) σ^2 **known** - Somehow we know σ^2 (not very practical). Then we can use $\mathcal{N}(\mu, \frac{\sigma^2}{n})$ distribution for the test. Or if we transform and use Z_n statistic (rather than $\bar{X}_n$) we can use $\mathcal{N}(0, 1)$.
  - 2) σ^2 **unknown** - We do not know σ^2 so we need to use an estimator S^2. Then we can use t_{n-1} distribution
- **Without Normality Assumption:** If the population data is not normal, we cannot assume they are Normal. If $X_1, X_2, \ldots, X_n$ follows *any arbitrary distribution* with mean μ and variance σ^2, then $\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ as sample size n is very large. In this case even if σ^2 is not known, we can still use S^2 and $\mathcal{N}(0, 1)$. The benefit of having large sample size is we can drop the Normality assumption of the random variables.

First we will see the case when σ^2 is known.

Tests - σ^2 known (z-test)

Two-Tail Test: First Example

The following example is slightly modified from Anderson et al. (2020)). Here is the story

Application Problem

The Bangladesh Golf Federation (BGF) establishes rules that manufactures of golf equipment must meet if their products are to be acceptable for use in BGF events. Company ABC uses a high-technology manufacturing process to *produce golf balls* with a mean driving distance of 295 yards. This is what it advertises to sell its balls. Sometimes, however, the process gets out of adjustment and produces golf balls with a mean driving distance different from 295 yards. When the mean distance *falls below* 295 yards, the company worries about losing sales because the golf balls do not provide as much distance as advertised. When the mean distance *goes above* than 295 yards, ABC also worries because golf balls may be rejected by the BGF for exceeding the overall distance standard concerning carry and roll. ABC's quality control department takes time to time a sample of 50 golf balls to monitor the manufacturing process. For each sample, the department conducts hypothesis testing to determine whether the process has fallen out of adjustment. Develop the null and alternative hypotheses and do the testing at 5% level of significance. Suppose somehow the quality control team knows that population standard deviation $\sigma = 12$.

- How do we develop the Null and Alternative Hypothesis?
- Note that, both below and above of 295 is problematic for the company. So maybe we could form the hypotheses like this.

$$H_0 : \mu = 295$$
$$H_a : \mu \neq 295$$

- Notice, here $\mu_0 = 295$
- What does the company want? It's clear that the company would like to accept the Null. This is because if the company rejects the Null then it's costly, why? maybe because it has to change its entire production process.
- From the story, we know
  - $n = 50$
  - $\bar{x} = 297.6$
  - $\alpha = 0.05$
  - $\mu_0 = 295$
  - $\sigma = 12$
- This means the standard error is $\sigma/\sqrt{n} = 12/\sqrt{50} = 1.697$.
- Recall that to do this testing we need $\alpha/2$ quantile and $1 - \alpha/2$ quantile of the sampling distribution under the Null. In this case the sampling distribution under the Null is $\mathcal{N}(295, 1.697)$.

Two-Tail Test

First Example Using Direct Sampling Distribution of $\bar{X}_n$

- Since we know $\alpha = .05$, then this means we need to find the following two values (*these are called the critical values*
  - $\bar{x}_{\alpha/2} = \bar{x}_{.025}$ This is called .025th quantile or 2.5th percentile of the sampling distribution.
  - $\bar{x}_{1-\alpha/2} = \bar{x}_{.975}$ This is called .975th quantile or 97.5th percentile of the sampling distribution.
- Let's do the calculations in R

R code - sigma known, two-tail (using direct sampling distribution)

```
# First give the data
n <- 50
xbar <- 297.6
alpha <- 0.05
mu0 <- 295
sigma <- 12

# (alpha/2) quantile
qnorm(alpha/2, mu0, sigma/sqrt(n))
#> [1] 291.6738

# (1 - alpha/2) quantile
qnorm(1 - alpha/2, mu0, sigma/sqrt(n))
#> [1] 298.3262
```

Two-Tail Test

First Example Using Direct Sampling Distribution of $\bar{X}_n$

- So we found our two critical values,
  - $\bar{x}_{\alpha/2} = \bar{x}_{.025} = 291.6738$
  - $\bar{x}_{1-\alpha/2} = \bar{x}_{.975} = 298.3262$
- Now note

$$291.67 < 297 < 298.32$$

- So our sample mean 297 does not fall in the rejection region, so we accept the Null or we can say we fail to reject the Null.
- Is it good news for ABC company. Yes it is, since it does not have to adjust the production process.

Two-Tail Test

First Example Using Standard Normal Distribution

- We can also use standard normal distribution, which we denoted with $\mathcal{N}(0,1)$ to do the testing of the ABC company.
- In this case we need find the $\alpha/2$ and $1-\alpha/2$ quantiles from the standard normal distribution, we will write these two values with $z_{\alpha/2}$ and $z_{1-\alpha/2}$. These are now the *critical values*!
- *Ques:* Can we compare $\bar{x}$ with $z_{\alpha/2}$ and $z_{1-\alpha/2}$ (you can find these quantiles from the table), so how can we use the table?
- The answer is NO! We cannot directly compare!
- But good news! we know the trick, that is we can use *standardization* (or calculate standardized sample mean or the so called *z-score*) with the formula, we will call it z_{calc}

$$z_{calc} = \frac{\bar{x}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

- And then we can compare this z_{calc} with $z_{\alpha/2}$ and $z_{1-\alpha/2}$, here "calc" means calculated.
- If we do the testing using z_{calc} and $z_{\alpha/2}$ and $z_{1-\alpha/2}$, we will reach to exactly same conclusion.
- We need to check whether $z_{calc} > z_{1-\alpha/2}$ and $z_{calc} < z_{\alpha/2}$

Two-Tail Test

First Example Using Standard Normal Distribution

- Here are steps using standard normal distribution.
- step - 1: Set Null and Alternative hypotheses $H_0 : \mu = \mu_0$ and $H_a : \mu \neq \mu_0$
- step - 2: Using the random sample calculate $\bar{x}$. Transform the $\bar{x}$, call the new estimate z_{calc} (Notice it's a z value or value of the standard normal distribution)

$$z_{calc} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- step - 3: Fix Type - I error and divide it equally with an area of $\alpha/2$ in two-tails.
- step - 4: For two-tail test, find two critical values $z_{\frac{\alpha}{2}}$ and $z_{1-\frac{\alpha}{2}}$ using $\mathcal{N}(0, 1)$ such that

$$\mathbb{P}(Z_n \leq z_{\frac{\alpha}{2}}) = .025 \quad \text{and} \quad \mathbb{P}(Z_n \leq z_{1-\frac{\alpha}{2}}) = .975$$

- step - 5: Reject the Null if $z_{calc} \geq z_{1-\frac{\alpha}{2}}$ or $z_{calc} \leq z_{\frac{\alpha}{2}}$ (either one)
- step - 6: Finally conclude that with α probability of error we are rejecting the Null Hypothesis (or more philosophically correct statement - *we failed to accept the Null*).
- So this means the hypothesis testing procedure is same. The only change is we need to transform everything to standard normal distribution.
- Now let's see how to use standard normal in R and do the same problem.

Two-Tail Test

First Example Using Standard Normal Distribution

R code - sigma known, two-tail (using standard normal)

```
# First give the data and calculate zcalc
n <- 50
xbar <- 297.6
alpha <- 0.05
mu0 <- 295
sigma <- 12

zcalc <- (xbar - mu0)/(sigma/sqrt(n))

# check the value
zcalc
# [1] 1.532065

# (alpha/2) quantile of the standard normal
qnorm(alpha/2)
# [1] -1.959964


# (1 - alpha/2) quantile of the standard normal
qnorm(1 - alpha/2)
# [1] 1.959964
```

Two-Tail Test

First Example Using Standard Normal Distribution

- So we found the critical values,
  - $z_{\alpha/2} = z_{.025} = -1.96$, this is the .025th quantile or 2.5th percentile of the standard normal distribution.
  - $z_{1-\alpha/2} = z_{.975} = 1.96$, this is the .975th quantile or 97.5th percentile of the standard normal distribution.
- We found these two values using R (but you can also use the table in Anderson et al. (2020) to get these two values).
- So we can see that

$$-1.96 < 1.532 < 1.96$$

- So our transformed sample mean 1.532 does not fall in the rejection region, and we get the same result as before that is *we fail to reject the Null*.

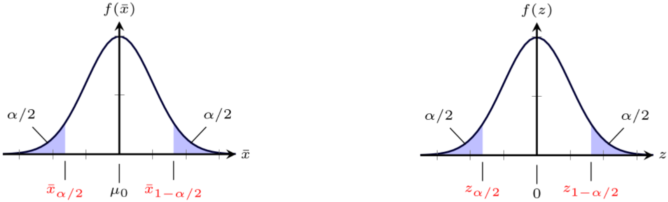
- Theoretical Discussion - *Why can we do this? or why we can transfer the problem to standard normal?*
- We already know why, the answer is in the Theorem 1.3 (iii) Chapter 1.
- Recall that if $\bar{X}_n \sim \mathcal{N}(\mu_0, \sigma^2/n)$, then we can do standardization (subtract mean and divide by standard deviation), this will give us a new (transformed) random variable Z_n

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

- where $Z_n \sim \mathcal{N}(0, 1)$.
- Again recall, both $\bar{X}_n$ and Z_n are normal random variables, but Z_n is called *standard normal* random variable because its distribution is standard normal
- We can always go from one to other if we know μ (in this case μ_0) and σ
- Following picture will help to see this

Two-Tail Test

First Example Using Standard Normal Distribution



- Left: This is the sampling distribution of sample means $\bar{x}$. The left figure shows that this distribution is Normal with mean μ_0 and variance σ^2/n, and $\bar{x}_{1-\alpha/2}$ and $\bar{x}_{\alpha/2}$ are the two quantiles of this distribution. This is the sampling distribution under the Null.
- Right: This is standard Normal distribution (we did the transformation of all values $\bar{x}$ with $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$) where critical values are denoted by $z_{1-\alpha/2}$ and $z_{\alpha/2}$
- When we use the standard normal distribution to do the test, this is known as *z-test*.
- In exam, if σ^2 is given and if nothing is specified you should generally do *z-test*, rather than directly using sampling distribution, this is because this is more standard.
- However if specifically mentioned then you should use direct sampling distribution. Note that the idea of using direct sampling distribution is more transparent and it helps to understand testing.

Do it now,

- Do the same z-test using the same information but for $\alpha = 10\%$ and $\alpha = 1\%$. In other words, Type - I error will be $\alpha = 10\%$ and $\alpha = 1\%$.

Tests - σ^2 known (z-test)

One-Tail Tests: Concepts

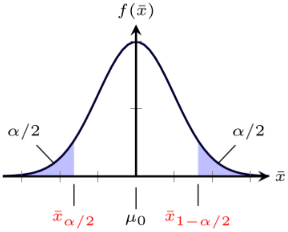
One-Tail Tests

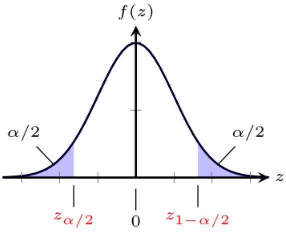
Concepts

- So far we have seen Hypothesis Testing with following type of Null and Alternative,

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

- We call this *two-tail* test.
- This is because the rejection regions for Null (or critical regions) are in two tails of the distribution and we are testing in both tails.
- Recall the picture (both the distribution of $\bar{x}$ and z)





- But the methods that we have learned can be extended to *one-tail tests*, where the rejection probability will be in one tail of the probability distribution.
- Following formation is what we call a *lower-tail* test

$$H_0 : \mu \geq \mu_0$$
$$H_a : \mu < \mu_0$$

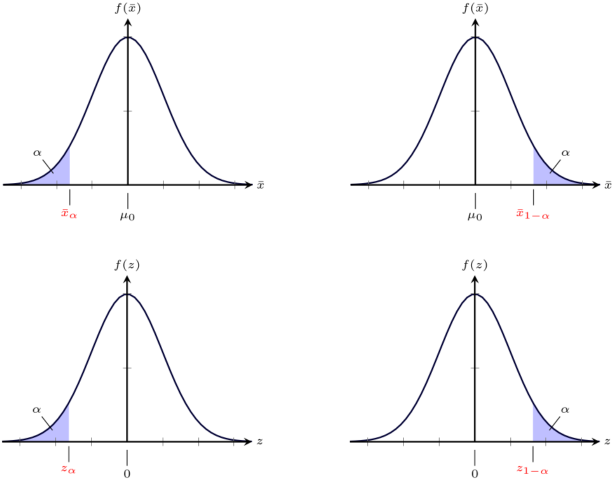
- This is called the *lower tail* test because the rejection region for Null is in the lower-tail.
- Similarly, following is the *upper-tail* test, the rejection probability is in the upper tail of the distribution.

$$H_0 : \mu \leq \mu_0$$
$$H_a : \mu > \mu_0$$

- Trick - an easy way to remember is to see the alternative hypothesis.
- Following picture might be helpful to understand why the names, you should always draw the picture when you solve the problems.

One-Tail Tests

Concepts



- On the left we have the figures for *lower-tail* test (Left-Top: the distribution of $\bar{x}$ and Left-Bottom: the standard normal distribution). The rejection areas are colored.
- On the right we have the figures for *upper-tail* test (Right-Top: the distribution of $\bar{x}$ and Right-Bottom: the standard normal distribution)

- Together they are called *one-tail tests*.
- The testing procedure for the *one-tail tests* is same except now we will *NOT divide α in two tails*.
- rather keep it one tail and do testing.
- We will see some examples now...

Tests - σ^2 known (z-test)

Lower-Tail Test: Steps and Example

Lower-Tail Test

Basic Steps

- The testing procedure is very similar.
- Now we will do the test using standard normal distribution $\mathcal{N}(0,1)$, but you will always get equivalent result if you directly use the sampling distribution of $\bar{X}_n$ under the Null.
- Here are the steps,

- step-1: Set Null and Alternative hypotheses,

$$H_0 : \mu \geq \mu_0$$
$$H_a : \mu < \mu_0$$

- step-2: Using the random sample calculate $\bar{x}$, and then transform $\bar{x}$, call the new estimate z_{calc}

$$z_{calc} = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

- step-3: Fix Type - I error α, and put in the left tail of the standard normal.
- step-4: Find critical value z_{α}. Note that z_{α} is such a value such that,

$$\mathbb{P}(Z_n \leq z_{\alpha}) = .05$$

- step-5: Reject the Null if $z_{calc} \leq z_{\alpha}$
- step-6: Finally conclude that with α probability of error we are rejecting the Null Hypothesis (or more philosophically correct statement - *we failed to accept the Null*).

- Let's see a concrete example of a lower tail test.
- We can modify the story of ABC company so that we can apply the lower tail test.

Application Problem Continued... (modified for lower-tail test)

Suppose ABC now only check whether the average distance is *at least* 295

- This means 295 or more is fine, but the company worries if the average distance is less than 295.
- In this case we can form following test, which is a lower tail test,

$$H_0 : \mu \geq 295$$
$$H_a : \mu < 295$$

- Here is how we can do the lower-tail test in R, we will use standard normal.
- Again before doing the problem please draw the picture.

Lower-Tail Test

Example

R code - sigma known, lower tail (using standard normal)

```
# First give the data and calculate zcalc
n <- 50
xbar <- 297.6
alpha <- 0.05
mu0 <- 295
sigma <- 12

zcalc <- (xbar - mu0)/(sigma/sqrt(n))

# check the value
zcalc
# [1] 1.532065

# (alpha) quantile of the standard normal
qnorm(alpha)
# [1] -1.644854
```

- So here we have $z_{\alpha} = -1.64$, clearly $z_{calc} > z_{\alpha}$. So again we accept the Null or *fail to reject the Null*.

Lower-Tail Test

Example

- So again ABC would like to accept the Null.
- Ques: For this same situation, can you also do

$$H_0 : \mu \leq 295$$
$$H_a : \mu > 295$$

- Which is the upper tail test.
- Ans: NO, why? Null is problematic here, what the company wants to check is not solely on the Null... can you see why....

Tests - σ^2 known (z-test)

Upper-Tail Test: Steps and Example

Upper-Tail Test
Steps

- For the upper-tail test, the steps are very similar to the lower-tail test, there are slight modifications,
- Here are the steps in brief,
  - step-1 will be adjusted to upper-tail test formation
  - step-2 is same as lower-tail test, calculate z_{calc}
  - step-3 Fix α and put α in the upper-tail of the standard normal
  - step-4 Find $z_{1-\alpha}$
  - step-5 Reject the Null if $z_{calc} \geq z_{1-\alpha}$

Upper-Tail Test

Example

- Let's again modify the story of ABC company.

Application Problem Continued... (modified for upper-tail test)

Now suppose ABC now only checks whether the average distance is *at max* 295, but worries if the average distance exceeds 295.

- In this case, we can form following test, which is an upper tail test,

$$H_0 : \mu \leq 295$$
$$H_a : \mu > 295$$

- Now clearly the company here doesn't like the alternative, and wants to accept the Null.
- Here is the R code. Note everything is same as the lower tail, but we will use the $1 - \alpha$ quantile, since the rejection region is in the right tail.

Upper-Tail Test

Example

R code - sigma known, upper tail (using standard normal)

```
# First give the data and calculate zcalc
n <- 50
xbar <- 297.6
alpha <- 0.05
mu0 <- 295
sigma <- 12

zcalc <- (xbar - mu0)/(sigma/sqrt(n))

# check the value
zcalc
# [1] 1.532065

# (alpha) quantile of the standard normal
qnorm(1 - alpha)
# [1] 1.644854
```

- So here we have $z_{1-\alpha} = 1.64$, clearly $z_{calc} < z_{1-\alpha}$. So again we accept the Null or *fail to reject the Null*.

Tests - σ^2 known (z-test)

Summary of Testing with known σ^2

Testing Summary

σ^2 known

Here is the summary of the procedures for the z-test

	Lower-Tail Test	Upper-Tail Test	Two-Tail Test
Hypotheses	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Value of the Test Statistic, Here we use z-statistic	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$
Rejection Rule (Reject H_0) (Critical Value Approach)	if $z_{calc} \leq z_{\alpha}$	if $z_{calc} \geq z_{1-\alpha}$	if $z_{calc} \leq z_{\alpha/2}$ or $z_{calc} \geq z_{1-\alpha/2}$

Table: z-test summary critical value approach

Tests - σ^2 unknown (t-test)

- So far we have assumed that we know σ^2, then we used z-statistic, which is

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

- and did a z-test.
- If we don't know σ^2, then we need to use an estimator of σ^2, which is S^2.
- In that case, we can use t-statistic, which is

$$T_n = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

- Here T_n is a random variable which follows t_{n-1} distribution (recall this means it follows t distribution with $n-1$ degrees of freedom!)
- Using this we can calculate t_{calc} and do the t-test, here is the formual for t_{calc} (notice looks very similar to z_{calc}, except σ now is replaced by s.)

$$t_{calc} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- In the t-test, after calculating t_{calc}, we will compare with the critical values from t_{n-1} distribution, following table should give you the summary

	Lower-Tail Test	Upper-Tail Test	Two-Tail Test
Hypotheses	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$	$H_0 : \mu = \mu_0$
	$H_a : \mu < \mu_0$	$H_a : \mu > \mu_0$	$H_a : \mu \neq \mu_0$
Value of the Test Statistic,	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$
Here we use t-statistic			
Rejection Rule (Reject H_0)	if $t_{calc} \leq t_\alpha$	if $t_{calc} \geq t_{1-\alpha}$	if $t_{calc} \leq t_{\alpha/2}$
(Critical Value Approach)			or $t_{calc} \geq t_{1-\alpha/2}$

Table: t-test summary critical value approach

- suppose we have slightly changed the BGF problem (the two tail test) problem, and we now assume the company don't know σ, rather the company only known $s = 10.8$. In this case we will do the t-test. Here is the R code

R code - sigma unknown, t-test, two-tail (using t_{n-1} distribution)

```
# First give the data and calculate tcalc
n <- 50
xbar <- 297.6
alpha <- 0.05
mu0 <- 295
s <- 10.8

tcalc <- (xbar - mu0)/(s/sqrt(n))

# check the value
tcalc
# [1] 1.532065

# (alpha/2) quantile of the t distribution with n-1 df
qt(alpha/2, n - 1)
# [1] -2.009575


# (1 - alpha/2) quantile of the t distribution with n-1 df
qt(1 - alpha/2, n - 1)
# [1] 2.009575
```

- since $-2.00 < 1.5 < 2.00$, this means $t_{\alpha/2} < t_{calc} < t_{1-\alpha/2}$, we will accept the Null.

- The methods we used to do hypothesis testing (both for z-test and t-test) is known as *critical value approach*.
- There is another approach of doing the test, which is called ***p** value approach*.
- It's important that both way of doing the tests will give you same answer, but p value approach has some advantages.
- *Question: what is a p-value?* In simple words, *p-value is the lowest probability at which H_0 can be rejected*. Or we can say it is the *smallest significance level at which we can reject the Null*.
- For example, if we are testing a given hypothesis with $\alpha = 0.05$ and we calculated the p value equal to 0.03, then in the $p-$value approach we will compare α with p-value and come to a conclusion.

- The calculation of p-value is quite easy.
- For *z-test* (or when we know σ), if we have calculated z_{calc}, then we need to find following probabilities and these are the p values for a z-test.
  - 1. p value for the both upper and lower tail test: $\mathbb{P}(Z_n > |z_{calc}|)$
  - 2. p value for a two-tail test: $2 \times \mathbb{P}(Z_n > |z_{calc}|)$.
- So you should think p value is like a probability and it should always be between 0 and 1.
- The rejection rule using p-values is always same that is

$$\text{reject the Nulll } H_0 \text{ if } p - \text{value } < \alpha$$

- In the p-value approach we are comparing probabilities vs. probabilities, where in critical value approach, we are comparing z values vs. z values.
- For the *t-test* the calculation is similar
  - 1. p value for both the upper and lower tail test: $\mathbb{P}(T_n > |t_{calc}|)$
  - 2. p value for a two-tail test: $2 \times \mathbb{P}(T_n > |t_{calc}|)$
- The rejection rule using p-values is always same that is

$$\text{reject the Nulll } H_0 \text{ if } p - \text{value } < \alpha$$

Testing Summary

Testing Summary

Here is the summary of the two procedures for the z-test

	Lower-Tail Test	Upper-Tail Test	Two-Tail Test
Hypotheses	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Value of the Test Statistic, Here we use z-statistic	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$	$z_{calc} = \frac{\bar{x}_n - \mu_0}{\sigma/\sqrt{n}}$
Rejection Rule (Reject H_0) (Critical Value Approach)	if $z_{calc} \leq z_\alpha$	if $z_{calc} \geq z_{1-\alpha}$	if $z_{calc} \leq z_{\alpha/2}$ or $z_{calc} \geq z_{1-\alpha/2}$
Rejection Rule (Reject H_0) (p Value Approach)	if $p \leq \alpha$	if $p \leq \alpha$	if $p \leq \alpha$

Table: z-test summary of the two approaches

Testing Summary

Here is the summary of the two procedures for the t-test

	Lower-Tail Test	Upper-Tail Test	Two-Tail Test
Hypotheses	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Value of the Test Statistic, Here we use t-statistic	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$	$t_{calc} = \frac{\bar{x}_n - \mu_0}{s/\sqrt{n}}$
Rejection Rule (Reject H_0) (Critical Value Approach)	if $t_{calc} \leq t_\alpha$	if $t_{calc} \geq t_{1-\alpha}$	if $t_{calc} \leq t_{\alpha/2}$ or $t_{calc} \geq t_{1-\alpha/2}$
Rejection Rule (Reject H_0) (p Value Approach)	if $p \leq \alpha$	if $p \leq \alpha$	if $p \leq \alpha$

Table: t-test summary of the two approaches

Testing Summary

- There is one clear benefit of the p value approach over the critical value approach, that is if we are asked to do the same test for different α, for example we are doing the same test for $\alpha = .10$, $\alpha = .05$ and $\alpha = .01$, in the p-value approach we can compute the p-value only once and the can compare it different α values. Only one calculation is enough. This is very convenient.
- In the critical value approach, every time we change α, we need to calculate the critical values again to do the test. For example if we change α, then we need to calculate z_{α} or $z_{1-\alpha}$ or $z_{\alpha/2}$ or $z_{1=\alpha/2}$ again to do the test. This is very cumbersome....

References

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., Cochran, J. J., Fry, M. J. and Ohlmann, J. W. (2020), *Statistics for Business & Economics*, 14th edn, Cengage, Boston, MA.

Bertsekas, D. and Tsitsiklis, J. N. (2008), *Introduction to probability*, 2nd edn, Athena Scientific.

Blitzstein, J. K. and Hwang, J. (2015), *Introduction to Probability*.

Casella, G. and Berger, R. L. (2002), *Statistical Inference*, 2nd edn, Thomson Learning, Australia ; Pacific Grove, CA.

DeGroot, M. H. and Schervish, M. J. (2012), *Probability and Statistics*, 4th edn, Addison-Wesley, Boston.

Hansen, B. (2022), *Econometrics*, Princeton University Press, Princeton.

Newbold, P., Carlson, W. L. and Thorne, B. M. (2020), *Statistics for Business and Economics*, 9th, global edn, Pearson, Harlow, England.

Pishro-Nik, H. (2016), *Introduction to probability, statistics, and random processes*.

Ramachandran, K. M. and Tsokos, C. P. (2020), *Mathematical Statistics with Applications in R*, 3rd edn, Elsevier, Philadelphia.

Rice, J. A. (2007), *Mathematical Statistics and Data Analysis*, Duxbury advanced series, 3rd edn, Thomson/Brooks/Cole, Belmont, CA.