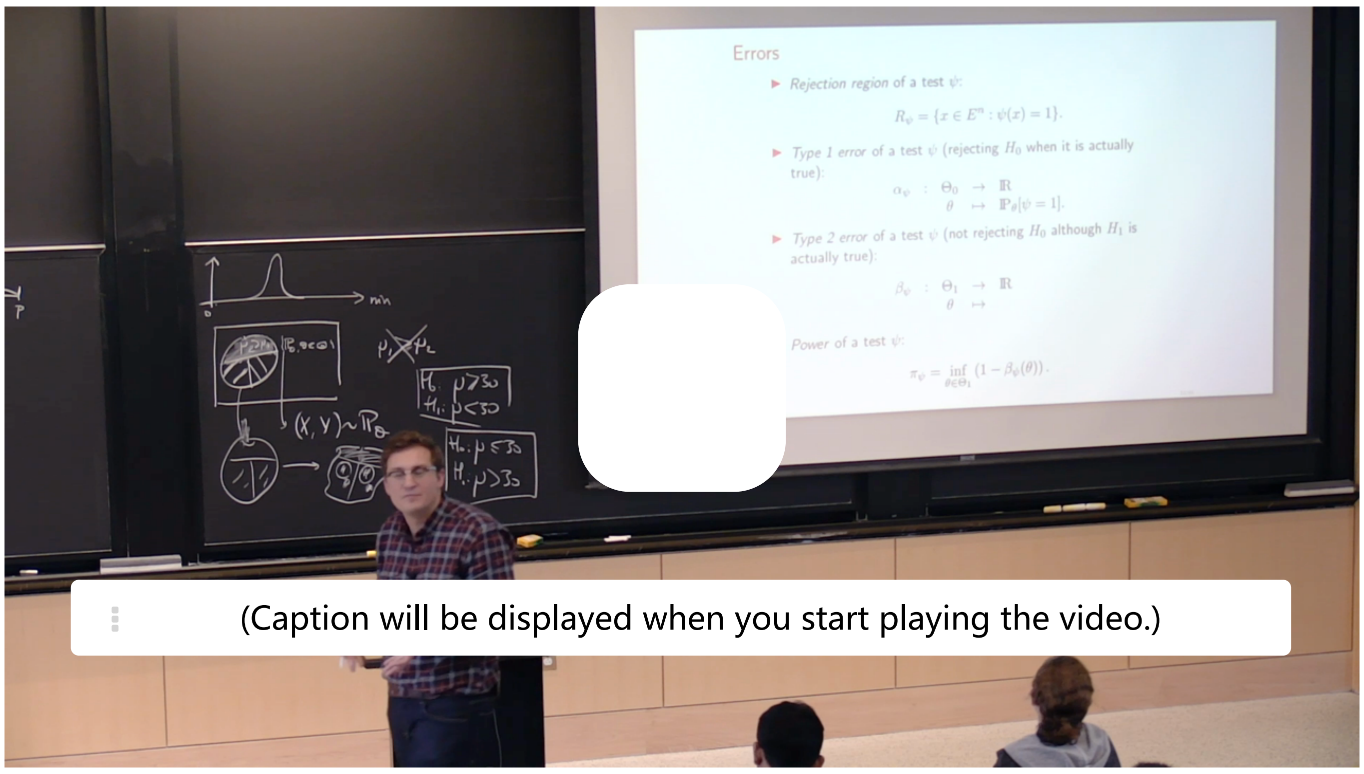
13. Type 1 Error of a Statistical Test

Type 1 Error of a Statistical Test



Start of transcript. Skip to the end.

So to do that, we need to say, well, I
want to have a good test, right?
I mean, I could pick anything I want.
Then, I would have a test.
Because right now, I've just defined a test
as being something that says value of 0 1.
So I could put C is equal to square root of pi over 2
if I wanted

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

An Analogy to the U.S. Justice System:Type 1 and Type 2 Errors

3/3 points (graded)

In a criminal court in the U.S., the goal is to decide between the following null and alternative hypotheses:

$$\begin{aligned} H_0 &: \text{The defendant is innocent.} \\ H_1 &: \text{The defendant is guilty.} \end{aligned}$$

In the U.S. criminal justice system, the informal principle "innocent until proven guilty" is the status quo, so this is the rationale for the choice of null hypothesis above. While this example is not, strictly speaking, a statistical hypothesis test, it provides some intuition about the meaning of type 1 and type 2 errors.

Suppose we have a defendant X who will be tried by a jury in the U.S. If guilty, X will go to jail, and otherwise is free to go.

In this example, let's say that the jury makes a **type 1 error** if the suspect satisfies H_0 while the jury rules in favor of H_1. Let's say the jury makes a **type 2 error** if the suspect satisfies H_1 while the jury rules in favor of H_0.

If the jury commits a type 1 error, the defendant is...

- ○ Innocent in reality, and will walk away free.
- ○ Guilty in reality, and will go to jail.
- ◉ Innocent in reality, but still will go to jail. ✔
- ○ Guilty in reality, but will walk away free.

If the jury commits a type 2 error, the defendant is...

- ○ Innocent in reality, and will walk away free.
- ○ Guilty in reality, and will go to jail.
- ○ Innocent in reality, but still will go to jail.
- ◉ Guilty in reality, but will walk away free. ✔

What strategy could the jurors follow if they wanted to never commit a type 2 error?

- ○ Always acquit– *i.e.*, always decide that the defendant is innocent.
- ◉ Always convict– *i.e.*, always decide that the defendant is guilty. ✔

Solution:

Let's examine the questions in order.

1. Since the null hypothesis is that X is innocent, in a type 1 error, the jury will convict X even though the defendant is innocent. Hence, the correct choice is that the defendant is "Innocent in reality, but still will go to jail."

2. Similary, since the alternative hypothesis is that X is guilty, in a type 2 error, the jury deems that X is innocent even though the defendant committed the crime. Hence the correct choice is that the defendant is "Guilty in reality, but will walk away free."

3. The correct response is "Always convict." If the jury always convicts, then there will never be a case where a guilty defendant walks away free: this strategy minimizes the type 2 error. However, it also maximizes the type 1 error. Every defendant who is innocent will be convicted, so practically speaking, this is a very questionable strategy.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

The Threshold for a Statistical Test

1/1 point (graded)

Continuing from problem on the previous page, let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$ where μ is an unknown parameter. You are interested in answering the **question of interest**: "**Does** $\mu = 0$ **?**".

To do so, you construct

- the **null hypothesis** $H_0 : \mu = 0$;
- the **alternative hypothesis** $H_1 : \mu \neq 0$.

Motivated by the central limit theorem, you decide to use a test of the form

$$\psi_C = \mathbf{1}\left(\sqrt{n}\,|\overline{X}_n| > C\right)$$

where $C > 0$ is a constant known as the **threshold** that you will choose in designing the test. (In the previous problem, C was chosen to be $0.25.$) On observing the data set, if $\psi = 1,$ you will **reject** $H_0.$ If $\psi = 0,$ then you will **fail to reject** $H_0.$

Suppose that indeed $\mu = 0$. Then $\mathbf{P}\left(\psi_C = 1\right)$, the probability of rejecting H_0, quantifies how likely we are to make the error of rejecting H_0 even though H_0 holds.

Under the assumption that $H_0 : \mu = 0$, for which value of C is $\mathbf{P}(\psi_C = 1)$ likely the largest?

- ◉ $C = 0.01$ ✔
- ○ $C = 0.1$
- ○ $C = 0.5$
- ○ $C = 1.0$

Solution:

The probability $\mathbf{P}(\mathbf{1}(|\overline{X}_n| > 0.01)$ is the largest.

Consider the events A_1, A_2, A_3, A_4 defined by

$$\begin{aligned} A_1 &: |\overline{X}_n| > 0.01, \quad A_2 : |\overline{X}_n| > 0.1 \\ A_3 &: |\overline{X}_n| > 0.5, \quad A_4 : |\overline{X}_n| > 1. \end{aligned}$$

Observe that $A_4 \subset A_3 \subset A_2 \subset A_1$, hence, by basic probability, $P(A_4) \leq P(A_3) \leq P(A_2) \leq P(A_1)$. Indeed, A_1 has the highest probability, so $P(\psi_1 = 1) = P(\mathbf{1}(|\overline{X}_n| > 0.01) = 1)$ is the largest out of $\psi_1, \ldots, \psi_4$. Thus, the test where $C = 0.01$ has the highest probability of rejection.
Remark: We did not need to know the shape of the distribution of X_n to make this conclusion; so in particular, we did not rely on the CLT.

提交 你已经尝试了2次（总共可以尝试2次）

ℹ Answers are displayed within the problem

Compute the Type 1 Error

1/1 point (graded)
As above, let $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$ where μ is an unknown parameter. You are interested in answering the **question of interest**: "**Does $\mu = 0$?**".

To do so, you construct

- the **null hypothesis** $H_0 : \mu = 0$;
- the **alternative hypothesis** $H_1 : \mu \neq 0$.

Motivated by the central limit theorem, you decide to use a test of the form

$$\psi_C = \mathbf{1}(\sqrt{n}|\overline{X}_n| > C).$$

Recall from lecture that the **type 1 error** (also known as **type 1 error rate**) of a test ψ is the **function**

$$\begin{aligned} \alpha_\psi : \Theta_0 &\to [0,1] \\ \theta &\mapsto \mathbf{P}_\theta(\psi = 1) \end{aligned}$$

If you choose the threshold $C = q_{0.05}$, what is the type 1 error α_ψ?
(In this case, since H_0 only consists of one point, the function α_ψ is defined only at one point, and we loosely use the termininology "type 1 error" to mean the value of α_ψ at that point.)

Type 1 Error α_ψ : 0.1 ✔ **Answer:** 0.1

Solution:

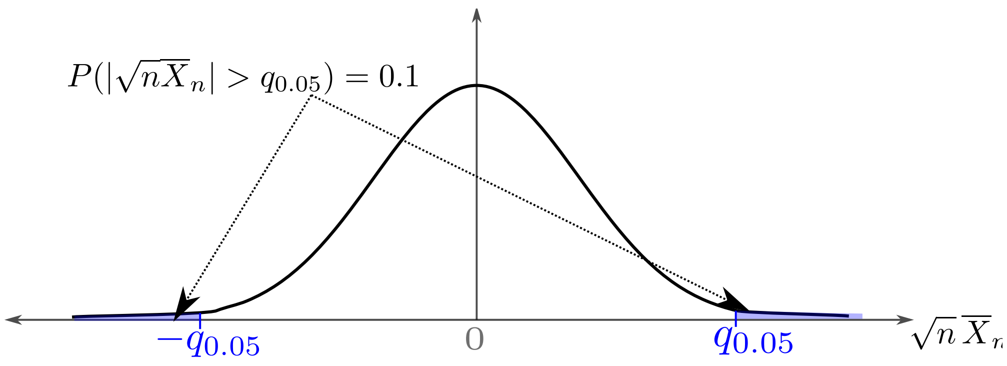
If we assume the null hypothesis $H_0 : \mu = 0$, and since the variance is known to be 1, the CLT gives

$$\sqrt{n}\,\overline{X}_n \sim \mathcal{N}(0,1) \qquad \text{for large } n.$$

The probability of a type 1 error is

$$\alpha_\psi(0) \;=\; \mathbf{P}_0(\psi_C = 1) \;=\; \mathbf{P}_0\left(\sqrt{n}\,|\overline{X}_n| > q_{0.05}\right) \;=\; 0.1.$$

as depicted in the figure below:



If H_0 is true, i.e. $\mu = 0$, then $\sqrt{n}\,\overline{X}_n$ is asymptotically normal. Hence, the total area of the two shaded regions is $\mathbf{P}_0(\psi_C = 1) \;=\; \mathbf{P}_0\left(\sqrt{n}\,|\overline{X}_n| > q_{0.05}\right)$, the probability that H_0 is rejected even though it is true.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

讨论

主题： Unit 2 Foundation of Inference:Lecture 6: Introduction to Hypothesis Testing, and Type 1 and Type 2 Errors / 13. Type 1 Error of a Statistical Test

显示讨论

认证证书是什么?