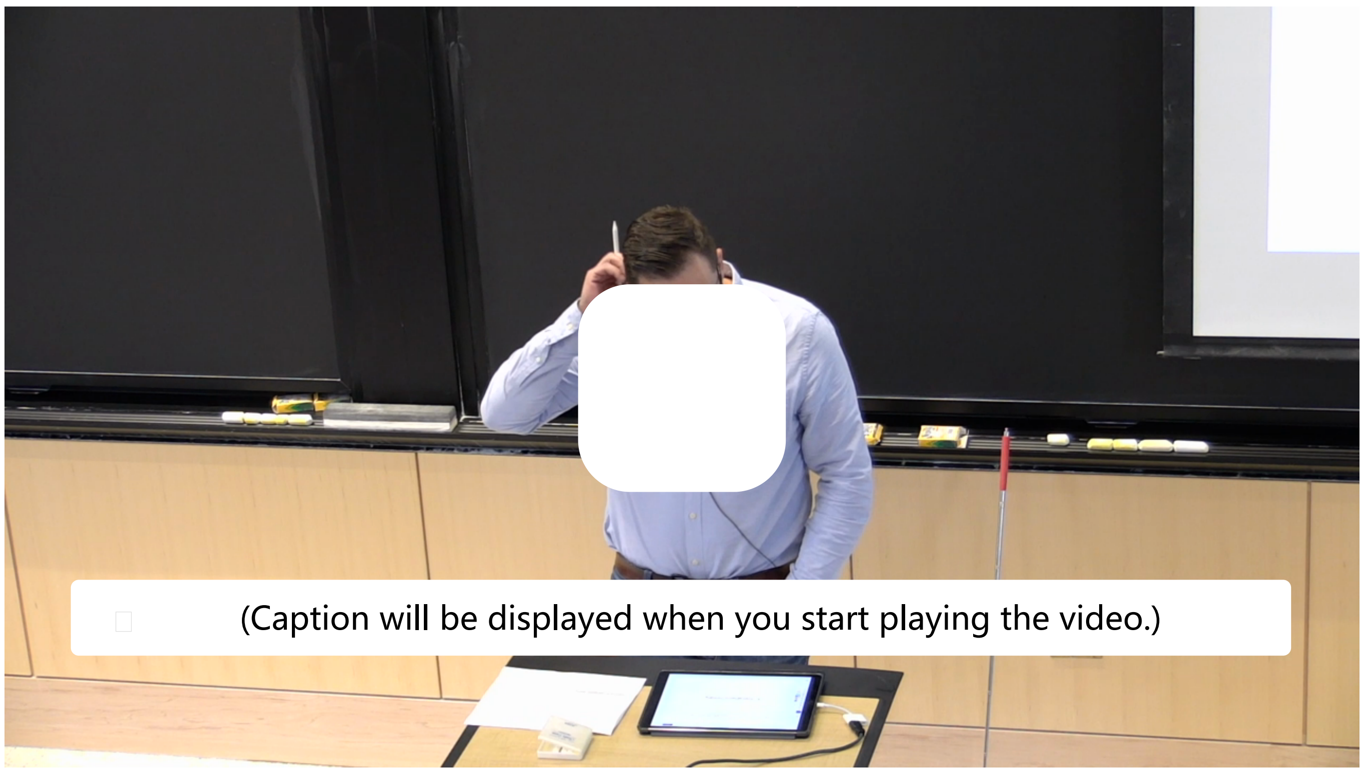
3. Unit Overview

What We Have Seen in Hypothesis Testing So Far...



Start of transcript. Skip to the end.

So we're going to talk to you about a hypothesis testing.

And if you look at this title, you should be like, well,

we've already seen hypothesis testing, right?

So basically, we know some of the language that's

around hypothesis testing, we know the basic ideas,

how to build a test, and, here, we're

视频
下载视频文件

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

Now would be a good time to review Hypothesis Testing and its related terminology as we have seen in Lecture 6, Lecture 7, and Homework 3.

Hypothesis Testing Review I

1/1 point (graded)
In the next few problems, we use a similar (but not identical) set-up to Problem 6 in Homework 3.

The National Assessment of Educational Progress tested a simple random sample of n thirteen year old students in both 2004 and 2008 and recorded each student's score. The standard deviation in 2004 was 39. In 2008, the standard deviation was 38.

Your goal as a statistician is to assess whether or not there were statistically significant changes in the average test scores of students from 2004 to 2008. To do so, you make the following modeling assumptions regarding the test scores:

- $X_1, \ldots, X_n$ represent the scores in 2004.
- $X_1, \ldots, X_n$ are iid Gaussians with standard deviation 39.
- $\mathbb{E}[X_1] = \mu_1$, which is an unknown parameter.
- $Y_1, \ldots, Y_n$ represent the scores in 2008.
- $Y_1, \ldots, Y_n$ are iid Gaussians with standard deviation 38.
- $\mathbb{E}[Y_1] = \mu_2$, which is an unknown parameter.
- $X_1, \ldots, X_n$ are independent of $Y_1, \ldots, Y_n$.

You define your hypothesis test in terms of the null $H_0 : \mu_1 = \mu_2$ (signifying that there were not significant changes in test scores) and $H_1 : \mu_1 \neq \mu_2$.

The test given above is a

- One-sided, two-sample test.
- Two-sided, two-sample test.
- One-sided, one-sample test.
- Two-sided, one-sample test.

Solution:

The test above is a **Two-sided, two-sample test.** This is because there are two samples: $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$. This is also a two-sided test, because we are testing whether or not $\mu_1 = \mu_2$ or $\mu_1 \neq \mu_2$.

It would be a one-sided test if, for instance, H_1 were defined to be $\mu_1 > \mu_2$. In this situation, we would only be testing whether or not μ_1 is larger than μ_2.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

Hypothesis Testing Review II

2/2 points (graded)

Let us use the same set-up as in the previous problem. Assuming that the null hypothesis $H_0 : \mu_1 = \mu_2$ holds, which of the following are true about the distribution of the statistic

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}}?$$

(Choose all that apply.)

- [x] The distribution is standard Gaussian.
- [ ] The distribution is known, but changes depending on n.
- [ ] The distribution cannot be determined.

Suppose now that the variance σ_X^2 of $X_1, \ldots, X_n$ and the variance σ_Y^2 of $Y_1, \ldots, Y_n$ are unknown. Let $\widehat{\sigma_X}^2$ denote the sample variance of $X_1, \ldots, X_n$, and let $\widehat{\sigma_Y}^2$ denote the sample variance of $Y_1, \ldots, Y_n$. Replacing the true variances with the sample variances, we construct the statistic

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{\widehat{\sigma_X}^2 + \widehat{\sigma_Y}^2}}.$$

Still assuming that $H_0 : \mu_1 = \mu_2$ holds, which of the following are true about the distribution of this statistic? (Choose all that apply.)

- [ ] The distribution is a standard Gaussian for $n = 10$.

- [ ] The distribution is a standard Gaussian for all $n \in \mathbb{N}$.

- [x] By Slutsky's theorem, as $n \to \infty$, its distribution converges to standard Gaussian.

Solution:

几个高斯分布的线性组合也是高斯分布

Recall that a linear combination of independent Gaussian random variables is again a Gaussian. Therefore, it suffices to determine the mean and variance of the given statistic. (Already, we see that the response **The distribution cannot be determined.** is incorrect.)

By linearity of expectation,

$$\boxed{\mathbb{E}\left[\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}}\right] = \sqrt{\frac{n}{38^2 + 39^2}}\left(\mathbb{E}\left[\overline{X}_n\right] - \mathbb{E}\left[\overline{Y}_n\right]\right) = 0,}$$

since $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are centered.

Next, by independence, the variance is additive:

$$\boxed{\mathrm{Var}\left(\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}}\right) = \frac{1}{n\left(38^2 + 39^2\right)}\left(\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right) + \sum_{i=1}^{n}\mathrm{Var}\left(Y_i\right)\right) = 1.}$$

Hence, the first response **The distribution is standard Gaussian.** is correct. Note that the distribution does not depend on the sample size n.

For the next question, we consider the statistic

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{\widehat{\sigma_X}^2 + \widehat{\sigma_Y}^2}}.$$

Since $\widehat{\sigma_X}$ and $\widehat{\sigma_Y}$ are random variables, the distribution of the above cannot be standard Gaussian for any fixed n. However, we know that

$$\widehat{\sigma_X} \xrightarrow{n\to\infty} \sigma_X, \quad \widehat{\sigma_Y} \xrightarrow{n\to\infty} \sigma_Y.$$

Therefore, Slutsky's theorem applies because

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \sim \mathcal{N}\left(0, 1\right).$$

We conclude that

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{\widehat{\sigma_X}^2 + \widehat{\sigma_Y}^2}} \xrightarrow[(d)]{n\to\infty} \mathcal{N}\left(0, 1\right).$$

Hence, the third choice **By Slutsky's theorem, as $n \to \infty$, its distribution converges to standard Gaussian.** is correct.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Hypothesis Testing Review III

2/2 points (graded)

We use the same statistical set-up as in the previous two problems, and now we assume that the true variances are known: $\sigma_X^2 = 39^2$ and $\sigma_Y^2 = 38^2$. Accordingly, you design the test

$$\psi = \mathbf{1}\left(\sqrt{n}\left|\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}}\right| \geq q_{\eta/2}\right).$$

where q_η represents the $1-\eta$ quantile of a standard Gaussian.

The level of this test is... (Choose all that apply.)

- [x] Asymptotic
- [x] Non-asymptotic

The level of this test is ... (Choose all that apply.)

- [x] η
- [ ] $1-\eta$
- [x] The probability of making a type 1 error under H_0.
- [ ] The probability of making a type 2 error under H_0.

Solution:

From the previous problems, we know that if $H_0 : \mu_1 = \mu_2$ holds, then

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}} \sim \mathcal{N}(0,1)$$

本身就是正态分布

for all $n \geq 1$. Since we know the distribution of the test statistic for all values of n, the test given has a **non-asymptotic** level η, and therefore also has an **asymptotic** level η.

. If $x_n \leq \alpha$ for all n, then $\lim x_n \leq \alpha$ if the limit exists.

CLT以后才是正态分布

Recall that the level of a test is the maximum probability of error assuming the null hypothesis. If the null hypothesis is true, then $\mu_1 = \mu_2$, and we conclude that

$$\sqrt{n}\frac{\overline{X}_n - \overline{Y}_n}{\sqrt{38^2 + 39^2}} \sim \mathcal{N}(0,1).$$

Therefore, the level is given by

$$P\left(|Z| > q_{\eta/2}\right) = \eta$$

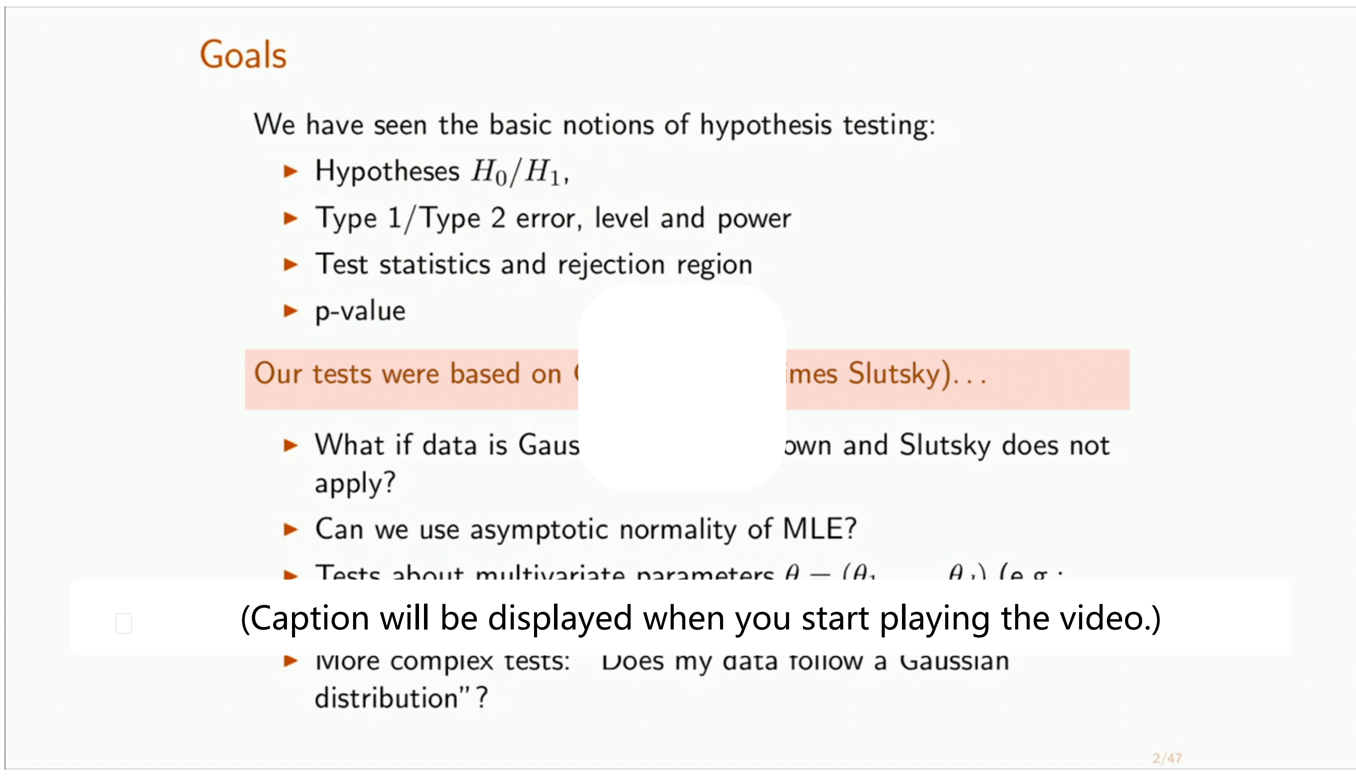
where $Z \sim \mathcal{N}(0,1)$. Equality holds above by the symmetry of the distribution $\mathcal{N}(0,1)$.

Hence, for the second question, the correct responses are η and **The probability of making a type 1 error under H_0**.

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem

What We Will See in Hypothesis Testing this Chapter...



Start of transcript. Skip to the end.

So our tests, sorry, were mostly asymptotic, right,

because we wanted to be able to use a central limit

theorem, sometimes Slutsky, so that we could actually

have asymptotic Gaussian distribution.

Why did we want to do this?

Well, because once we have a Gaussian distribution,

视频
下载视频文件

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

讨论

显示讨论

主题： Unit 4 Hypothesis testing:Lecture 13: Chi Squared Distribution, T-Test / 3. Unit Overview