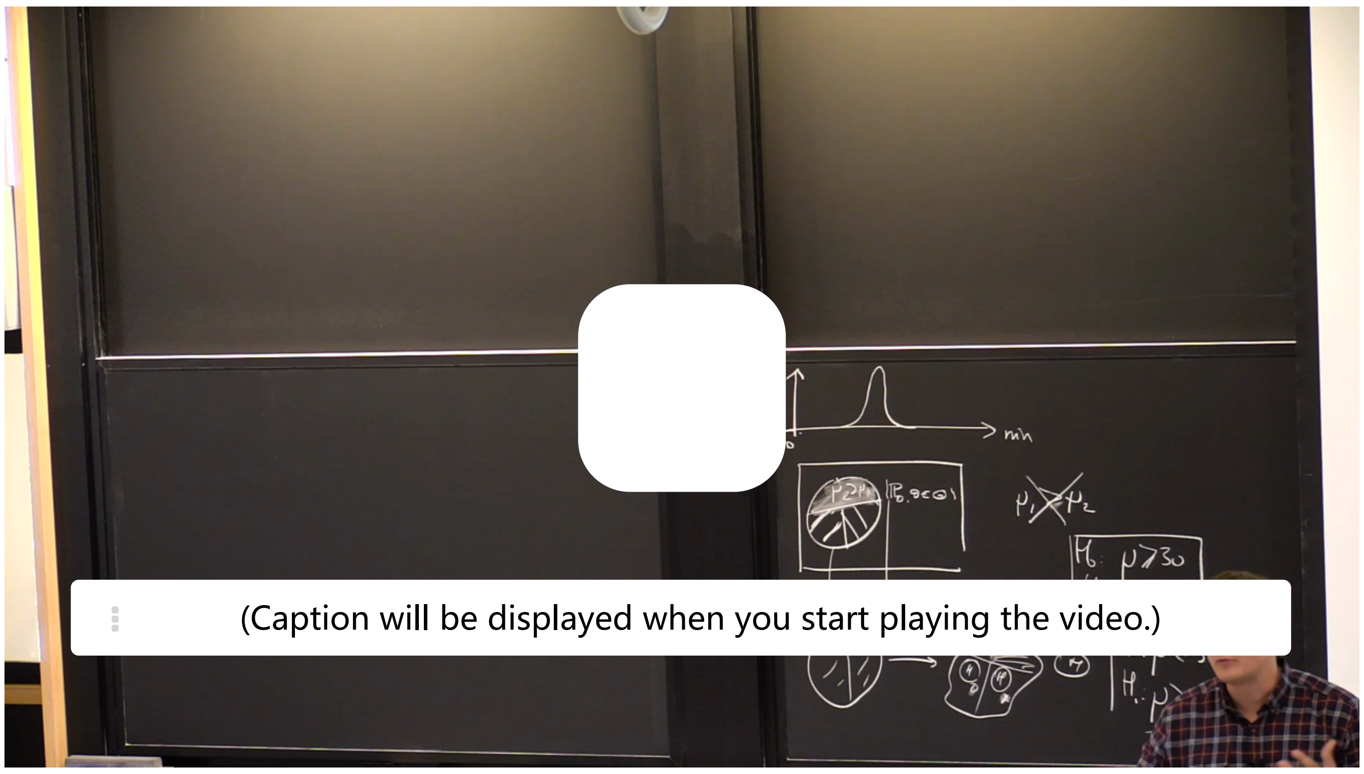
12. Statistical Tests

Statistical Tests



Start of transcript. Skip to the end.

OK, so now, let's do the same thing that we did before.

Remember when we talked about the kiss example,

we had this left right, left right,

and the first thing that we said is turn this

into a numerical thing.

I have a binary answer at the end of the day.

I might as well turn it into a 0 and 1.

视频
下载视频文件

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

Which Statistics are Tests?

1/1 point (graded)

Recall that a **statistic** is, intuitively speaking, a function that can be computed from the data.

A **(statistical) test** is an **statistic** whose output is **always** either 0 or 1, and like an estimator, does not depend explicitly on the value of true unknown parameter.

Let $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathrm{Ber}(\theta)$ for some unknown parameter $\theta \in (0,1)$. Which of the following statistics are also tests?

(Recall that $\mathbf{1}(A)$ is the indicator defined as follows: $\mathbf{1}(A) = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$.

(Choose all that apply.)

- [ ] $\overline{X}_n$
- [x] $\mathbf{1}(\overline{X}_n > 0.5)$ ✔
- [x] $\mathbf{1}(|\overline{X}_n - 0.5| > 0.01)$ ✔
- [ ] $\mathbf{1}(|\overline{X}_n - \theta| > 0.5)$
- [x] $\mathbf{1}(\overline{X}_n \text{ is a rational number})$ ✔

✔

Solution:

We examine the choices in order.

- $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ is **not** a statistical test, because the sample average is not **always** either 0 or 1.

- $\mathbf{1}\left(\overline{X}_n > 0.5\right)$ is a statistical test. Its expression only depends on the sample (and not the true parameter), and since it is an indicator, it takes values only in $\{0,1\}$.

- $\mathbf{1}\left(|\overline{X}_n - 0.5| > 0.01\right)$ is a statistical test. Its expression only depends on the sample (and not the true parameter), and since it is an indicator, it takes values only in $\{0,1\}$.

- $\mathbf{1}\left(|\overline{X}_n - \theta| > 0.5\right)$ is **not** a statistical test because it is not an estimator; *i.e.*, its expression depends on the unknown parameter θ.

- $\mathbf{1}\left(\overline{X}_n \textbf{is a rational number}\right)$ is a statistical test. Its expression only depends on the sample (and not the true parameter), and since it is an indicator, it takes values only in $\{0,1\}$. This is a rather bizarre test, but it does satisfy all required properties.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Applying a Statistical Test on a Data Set

1/1 point (graded)

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** $H_0 : \mu = 0$ and the **alternative hypothesis** $H_1 : \mu \neq 0$.

You design the test

$$\psi = \mathbf{1}\left(\sqrt{n}\left|\overline{X}_n\right| > 0.25\right).$$

If $\psi = 1$, you will **reject** the null hypothesis, and if $\psi = 0$, you will **fail to reject**. For simplicity, we will set the sample size to be $n = 7$.

On which of the following data sets would you reject the null hypothesis?
(Choose all that apply. Feel free to use computational tools.)

- [x] $-1.0, -0.8, -2.9, 1.4, 0.3, -0.8, 1.4$ ✓
- [x] $-1.7, -0.1, -0.2, 0.3, 0.3, -0.9, -0.03$ ✓
- [ ] $-0.2, 0.6, 1.1, -0.9, 0.1, -1.2, 1.1$

✓

Solution:

We examine the choices in order.

- The first choice is correct. For this data set, we compute $\sqrt{7}\,\overline{X}_7 \approx -0.9072$. Since $|-0.9072| > 0.25$, we reject.

- The second choice is correct. For this data set, we compute $\sqrt{7}\,\overline{X}_7 \approx -0.8768$. Since $|-0.8768| > 0.25$, we reject.

- The third choice is incorrect. For this data set, we compute $\sqrt{7}\,\overline{X}_7 \approx -0.2267$. Since $|0.2267| \leq 0.25$, we fail to reject.

Remark 1: It is useful to keep in mind the following mnemonic,

$$\psi = 0 \Rightarrow H_0$$
$$\psi = 1 \Rightarrow H_1.$$

Of course, the implications above are informal and should not be taken literally. To be precise, we say that if $\psi = 0$, we fail to reject H_0, and if $\psi = 1$, then we reject H_0 in favor of H_1.

Remark 2: If we assume the null hypothesis $H_0 : \mu = 0$, then since the variance is known to be 1, the CLT guarantees that

$$\sqrt{n}\,\overline{X}_n \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,1)\,.$$

The quantiles of $\mathcal{N}(0,1)$ can be understood using tables or computational software, so if n is very large, then we can approximate the probability of our test ψ **rejecting** or **failing to reject** under the null hypothesis. This concept will be further explored in the next page where we explore the “type 1" and “type 2 error" of a test.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 2 Foundation of Inference:Lecture 6: Introduction to Hypothesis Testing, and Type 1 and Type 2 Errors / 12. Statistical Tests

认证证书是什么?