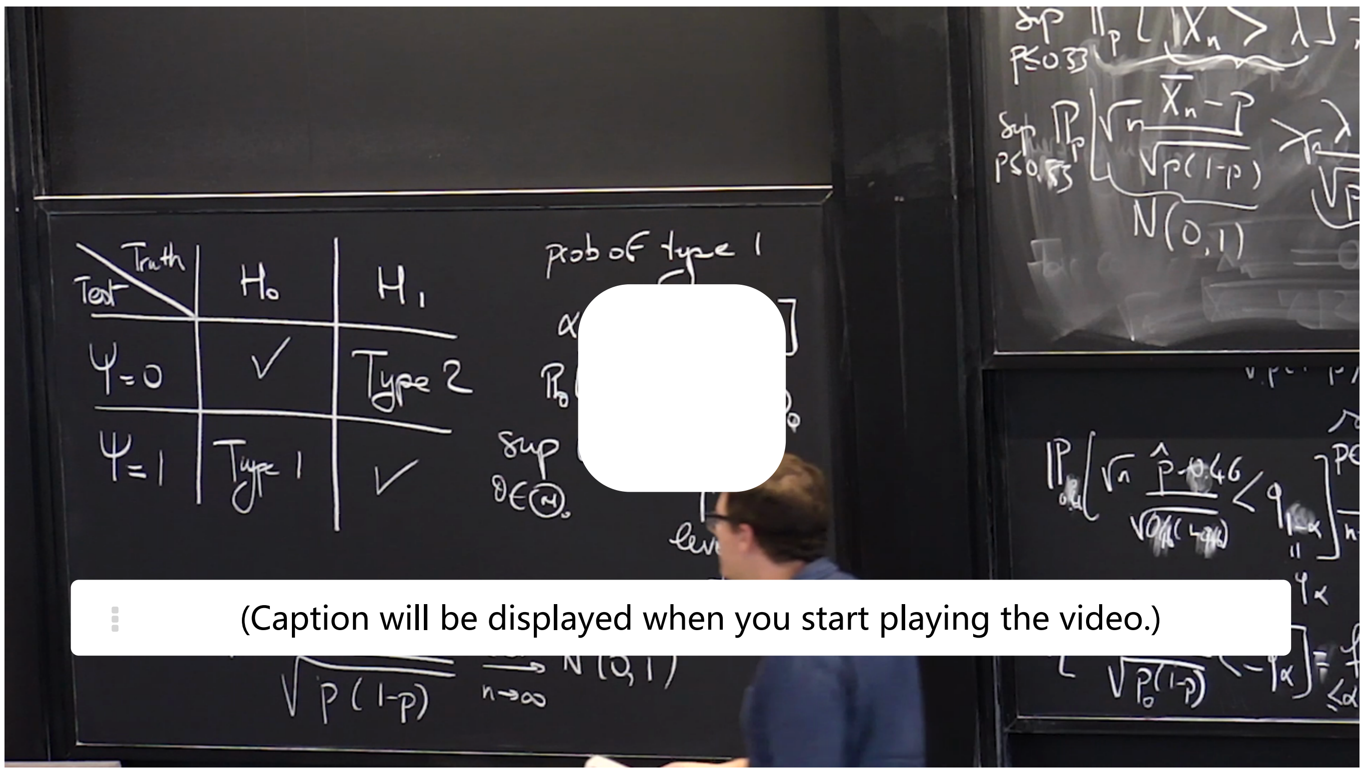
8. Worked Example: Find the P-value

Worked Example: The p-value of a Two-Sided Statistical Test



That's going to be the kind of exercises

you will have around tests.

So kiss example-- so here, we recorded x1 xn.

There were iid, Bernoulli p.

Let's write the statistical model, just for fun.

What was the sample space?

What is the sample space for the statistical model?

0, 1-- and the family of probability distributions

is Bernoulli p for p in 0, 1.

That does not have anything to do with hypothesis testing.

Now, the test we wanted to do was h0 p is equal to 1/2--

no preference-- versus h1.

p is not equal to 1/2.

And here, that's clearly the scientific discovery.

So everybody agrees why we go for h0

视频

下载视频文件

字幕

下载 SubRip (.srt) file

下载 Text (.txt) file

Motivating the p-value

3/3 points (graded)

Let us return to the test of fairness of a coin.

Setup:

We have a sample $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathrm{Ber}(p^*)$ and associated statistical model $\left(\{0,1\}, \{\mathrm{Ber}(p)\}_{p\in(0,1)}\right)$. The null and alternative hypotheses are

$$\begin{aligned} H_0 : p^* &= 1/2 \\ H_1 : p^* &\neq 1/2. \end{aligned}$$

Let

$$T_n = \sqrt{n}\left|\frac{\left(\overline{X}_n - 0.5\right)}{\sqrt{0.5\,(1-0.5)}}\right|$$

denote the test statistic and let

$$\psi = \mathbf{1}\left(T_n \geq q_{\eta/2}\right).$$

denote the test where q_η is the $1-\eta$ quantile of a standard Gaussian.

Questions:

In one run of the experiment, you obtain the data set consisting of 80 Heads, and evaluated test statistics T_n at this data set to be $T_n = 2.82842$ (as in the previous problem *Hypothesis Testing: A Sample Data Set of Coin Flips I*).

The **(asymptotic) p-value** for this data set is defined to be the smallest (asymptotic) level α such that ψ rejects H_0 on this data.

What is the asymptotic p-value for this data set?
(You are encouraged to use computational tools or tables.)

0.005 ✔ **Answer:** 0.0047

In another run of the experiment, you obtain the data set consisting of 106 Heads, and evaluated test statistics T_n at this data set to be $T_n = 0.8485$.

What is the asymptotic p-value for this second data set?
(You are encouraged to use computational tools or tables.)

0.4 ✔ **Answer:** 0.3962

Now let's generalize our findings above. In this two-sided test, as the test statistic T_n increases, the p-value ...

- ○ increases
- ◉ decreases ✔

Solution:

In the first experiment from the previous problem *Hypothesis Testing: A Sample Data Set of Coin Flips I*, we observed that $T_n = |-2.82842|$. For notational convenience, let $P_{1/2} = \text{Ber}(1/2)$. Recall that the asymptotic level is given by

$$\lim_{n\to\infty} P_{1/2}\left(T_n \geq q_{\eta/2}\right) = P\left(|Z| > q_{\eta/2}\right) = \eta$$

where $Z \sim N(0,1)$. Hence, we need to find the smallest level α such that ψ rejects, *i.e.*, such that

$$T_n \geq |-2.82842|.$$

Hence, we should set $q_{\eta/2} = 2.82842$ and solve for η. Using computational tools or a table of the standard Gaussian, we find that

$$\eta = 2P\left(Z \geq 2.82842\right) \approx 2\left(0.002339\right) = 0.00467.$$

In the second experiment, we observed that $T_n = 0.8485$. Following the same procedure as above, we set $q_{\eta/2} = 0.8485$, and using computational tools or a table of the standard Gaussian, we find that

$$\eta = 2P\left(Z \geq 0.8485\right) \approx 0.3961596$$

For the final question, as the test statistic increases, the p-value will decreases. Note that T_n measures (up to some rescaling) the deviation from the true mean under $H_0 : p^* = 0.5$. As this value grows, our observation moves further into the tails of the distribution $N(0,1)$. Since the asymptotic p-value for this problem is given by $1 - \Phi(T_n)$ where Φ is the cdf of $N(0,1)$, this implies that the asymptotic p-value decreases as T_n increases.

Remark 1: As a rule of thumb, a smaller p-value implies that one can more confidently reject the null hypothesis. Hence, in this scenario, we can more confidently reject the null for experiment I than the null from experiment II. You can think of a p-value as a measure of 'how surprised' you are to observe the given data set under the assumption that the null hypothesis holds. In particular, the smaller the p-value is, the more surprised you should be.

Remark 2: A very large value of T_n indicates a rare event under the null hypothesis, s we should be 'more surprised' at the data if we observe a very large value of T_n as opposed to a small one. The fact that the p-value decreases as T_n increases is consistent with that intuition, since our heuristic is to be more surprised at very small p-values than large ones 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

Computing p-values I: Kiss Example

1/1 point (graded)

Recall that in the kiss example, we record 1 if a couple prefers turning their head to the right and 0 otherwise. We modeled this as a Bernoulli statistical experiment $X_1, \ldots, X_n \overset{iid}{\sim} \mathrm{Ber}(p)$. For this question, we just want to test if couples as a whole have *some* preferred direction of turning their head; that is, we want to decide whether or not $p = 1/2$.

You set the null hypothesis to be $H_0 : p = 1/2$ and $H_1 : p \neq 1/2$. Your statistical test is given by

$$\mathbf{1}\left(\left|\sqrt{n}\frac{\overline{X}_n - 0.5}{\sqrt{0.5\,(1-0.5)}}\right| > q_{\eta/2}\right),$$

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

You observe that 75 out of 124 couples prefer turning their head to the right. What is the (asymptotic) p-value for this experiment? (You are encouraged to use computational tools or a table.)

0.02 ✔ **Answer:** 0.0196

Solution:

To solve for the asymptotic p-value, we find η such that

$$q_{\eta/2} = \left|\sqrt{n}\frac{\overline{X}_n - 0.5}{\sqrt{0.5\,(1-0.5)}}\right| = \left|\sqrt{124}\frac{\frac{75}{124} - 0.5}{\sqrt{0.5\,(1-0.5)}}\right| \approx 2.3340.$$

Indeed, if η is smaller than this, then ψ would fail to reject under observed sample mean $\frac{75}{124} \approx 0.6048$. To solve for η, we use computational tools or a table to find:

$$\eta = 2P\,(Z \geq 2.3340) \approx 2\,(0.0098) = 0.0196.$$

where $Z \sim N\,(0,1)$. Hence the p-value is around 1%, so it seems reasonable to reject the null hypothesis that couples, as a whole, do not have a preferred direction of turning their heads.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Concept Check: Interpreting the p-value

1/1 point (graded)

Consider a hypothesis test with null H_0 and alternative H_1 regarding an unknown parameter θ. You observe a sample $X_1, \ldots, X_n \overset{iid}{\sim} P_\theta$ and compute the p-value.

What is a correct interpretation of the p-value?

- [x] The smaller a p-value is, the more evidence that is suggested against H_0. ✔

- The larger a p-value is, the more evidence that is suggested against H_0.

Solution:

The rule of thumb is that the smaller the p-value is, the more confidently the null-hypothesis can be rejected. Hence, "A larger p-value suggests more evidence against H_1, while a smaller p-value suggests more evidence against H_0." is the correct choice.

Remark: Here is an explanation of this heuristic. As the p-value gets smaller, this means we can set the level of a test smaller and smaller and will still reject the null hypothesis based on the data. Since a smaller type 1 error tolerates rarer events under the null, this means that a small p-value lends evidence that the observation was a rare event under H_0. Therefore, a smaller p-value suggests more evidence against H_0.

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 2 Foundation of Inference:Lecture 7: Hypothesis Testing (Continued): Levels and P-values / 8. Worked Example: Find the P-value

认证证书是什么?