Hypothesis Testing

Lecture 21

03/07/2012

Turn on your clickers!

Which one is larger a 95% confidence interval or a 99% confidence interval (given that everything else is the same)?

(a) 95%

(b) 99%

(c) We cannot compare them.

Null & Alternative Hypothesis

$$H_0 : \mu = \mu_0$$
$$H_\alpha : \mu \neq \mu_0$$

- A test of significance is designed to access the strength of evidence **against** a baseline hypothesis that is called the *null hypothesis* and in favor of another hypothesis that is called *alternative hypothesis*.
- Usually, the null hypothesis means that there is "no effect" or "no difference".
- The alternative hypothesis is the statement that we hope or suspect to be true instead of the null.

Interpretation of the Null Hypothesis

- Default
- Uninteresting
- Nothing is happening
- There is no difference
- We want to disprove the Null!

Interpretation of the Alternative Hypothesis

- What might be happening
- The interesting thing
- Opposite of Null
- We want evidence for the alternative!

Alternative Hypothesis

Three types for the Alternative Hypothesis

- $H_\alpha : \mu > \mu_0$: One-sided Test
- $H_\alpha : \mu \neq \mu_0$: Two-sided Test
- $H_\alpha : \mu < \mu_0$: One-sided Test

Testing a New Diet

- Record how much weight lost by each subject
- Null
  - The diet has no effect
  - $\mu = 0$ because no weight was lost

$$H_0 : \mu = 0$$

- Alternative
  - The diet leads to people losing weight

$$H_\alpha : \mu < 0$$

Test Statistic

- A significance test is based on a *test statistic* that shows whether or not the data provide *evidence against the null hypothesis*.
- When H_0 is true, we expect the estimate to take a value **near** the parameter value specified by H_0.
- Values of the estimate **far** from the parameter value specified by H_0 give evidence against H_0. The alternative hypothesis determines which direction (or directions) counts against H_0.

P-Value

- P-value is the probability, assuming that H_0 is true, that the test statistic will take a value at least as extreme as that actually observed.
- The smaller the p-value, the stronger the evidence against H_0 provided by the data.

P-Value

- Reject the Null or Not??
  - Compare the **p-value** with a fixed value that we regard as decisive.
  - That value is the significance level α.
  - If α=0.05, we require that the data give evidence against the null so strong that it would happen no more than 5% of the time when H_0 is true.
- So, if the p-value is as small or smaller than α, we say that the data are **statistically significant** at level α.

Test Statistic for the Population Mean

i. **Normal Population, Known σ, Small sample size.**

i.e. $\mathcal{N}(\mu, \sigma)$

Test Statistic: standardized sample mean (**z-test statistic**)

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

Distribution of z under H_0

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Test Statistic for the Population Mean

ii. **IID Population, Large sample size.**

Test Statistic: standardized sample mean (**z-test statistic**)

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

Distribution of z under H_0

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim \mathcal{N}(0, 1)$$

Test Statistic for the Population Mean

iii. **Normal Population, Unknown σ, Small sample size.**

i.e. $\mathcal{N}(\mu, \sigma)$, σ unknown

Test Statistic: **t-test statistic**

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

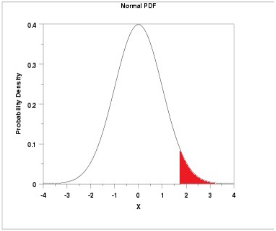
In this case the t-test follows approximately the **t-distribution.**

(We are not going to discuss this case in detail.)

Computing p-values for the population mean

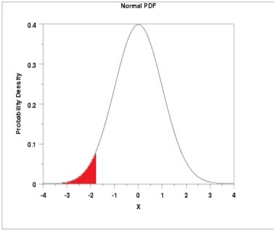
H_α	p-value
$\mu > \mu_0$	$p = \mathbb{P}(Z \geq z) = 1 - \mathbb{P}(Z \leq z)$
$\mu < \mu_0$	$p = \mathbb{P}(Z \leq z)$
$\mu \neq \mu_0$	$p = \mathbb{P}(\lvert Z\rvert \geq \lvert z\rvert) = \mathbb{P}(Z \geq \lvert z\rvert) + \mathbb{P}(Z \leq -\lvert z\rvert)$

$H_\alpha \colon \mu > \mu_0$



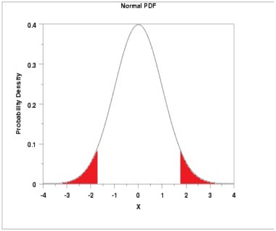
$$\mathbb{P}(Z \geq z)$$

$H_\alpha \colon \mu < \mu_0$



$$\mathbb{P}(Z \leq z)$$

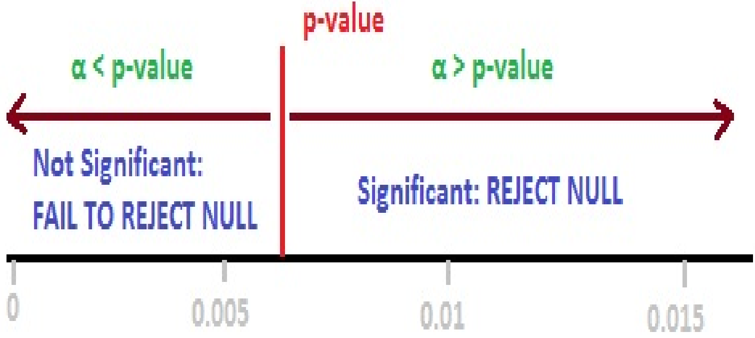
$H_\alpha \colon \mu \neq \mu_0$



$$\mathbb{P}(Z \geq |z|) + \mathbb{P}(Z \leq -|z|$$

Decision

- $p < \alpha$: there is enough evidence to reject the null hypothesis
- $p > \alpha$: there is not enough evidence to reject the null hypothesis. (We *fail to reject* the null hypothesis)


p-value
α < p-value
α > p-value
Not Significant:
FAIL TO REJECT NULL
Significant: REJECT NULL
0
0.005
0.01
0.015

Steps to perform a Hypothesis Test

- **Step 1:** State the null (H_0) and the alternative hypothesis (H_α).
- **Step 2:** Calculate the value of the test statistic.
- **Step 3:** Find the p-value for the observed data.
- **Step 4:** Compare the p-value with the desired level of significance α.

Turn on your clickers

Is it easier to reject the alternative hypothesis with significance level 1% or 5%?

(a) 1%

(b) 5%

(c) We cannot compare them.

Elite distance runners

- An exercise researcher believes that the mean weight of competitive runners is about 140 pounds.
- A sample of 35 elite distance runners has mean weight $\bar{x}$=136 pounds.
- Suppose that these runners are a simple random sample from this population with standard deviation σ=11 pounds.
- Is there enough evidence that the mean weight of elite distance runners is less than 140 pounds with 5% level of significance?

Elite distance runners

Which is the correct combination of null and alternative hypotheses?

(a) $H_0 : \mu = 140, H_\alpha : \mu < 140$

(b) $H_0 : \mu = 140, H_\alpha : \mu > 140$

(c) $H_0 : \mu = 140, H_\alpha : \mu \neq 140$

Elite distance runners

- The first step is to identify the two hypotheses. Since we want to test whether the mean weight of the runners is less than 140 pounds we have

$$H_0 : \qquad \mu = 140$$
$$H_\alpha : \qquad \mu < 140$$

- Since we have a large sample size (n=35 > 30) and known standard deviation for the population, we will use the z-test

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{136 - 140}{11/\sqrt{35}} = -2.15$$

Elite distance runners

- We are going to compute the corresponding p-value, that is $p = \mathbb{P}(Z \leq -2.15)$. Since Z follows the standard normal distribution

$$p = \mathbb{P}(Z \leq -2.15) = 0.0158$$

- We compare this p-value with the significance level α:

$$p = 0.0158 < 0.05 = \alpha,$$

  which means that the data are statistically significant at level α and therefore we **reject the null hypothesis.**

Random numbers

- A computer has a random number generator that is supposed to produce random numbers that are uniformly distributed within the interval from 0 to 1.
- If this is true, the numbers generated come from a population with mean $\mu = 0.5$.
- A command to generate 100 random numbers gives outcomes with sample mean $\bar{x} = 0.4817$ and sample standard deviation $s = 0.0833$.
- Is there enough evidence that the random number generator is working properly with 1% level of significance?

Random numbers

Which is the correct combination of null and alternative hypotheses?

(a) $H_0 : \mu = 0.5, H_\alpha : \mu < 0.5$

(b) $H_0 : \mu = 0.5, H_\alpha : \mu > 0.5$

(c) $H_0 : \mu = 0.5, H_\alpha : \mu \neq 0.5$

Random numbers

- The first step is to identify the two hypotheses. Since we want to test whether the mean is different than 0.5

$$H_0 : \qquad \mu = 0.5$$
$$H_\alpha : \qquad \mu \neq 0.5$$

- Since we have a large sample size (n=100 > 30) and sample standard deviation 0.0833, we will use the z-test

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{0.4817 - 0.5}{0.0833/\sqrt{100}} = -2.2$$

Random numbers

- The p-value is

$$p = \mathbb{P}(|Z| \geq |2.2|) = 2\mathbb{P}(Z \leq -2.2) = 2 \cdot 0.0139 = 0.0278$$

- Compare this p-value with the significance level α:

$$p = 0.0278 > 0.01,$$

  which means that the data are **not** statistically significant at level α and therefore we **fail to reject the null hypothesis.**

Hypothesis Testing for the Population Proportion

Setup

- Null Hypothesis

$$H_0 : p = p_0$$

- Alternative Hypothesis

$$\begin{aligned} H_\alpha : &\quad p > p_0 \\ H_\alpha : &\quad p < p_0 \\ H_\alpha : &\quad p \neq p_0 \end{aligned}$$

Test Statistic & p-value

- z-test

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}},$$

where $\hat{p}$ is the proportion of the sample and p_0 is the proportion under the null hypothesis.

- p-value

H_α	p-value
$p > p_0$	$\mathbb{P}(Z \geq z) = 1 - \mathbb{P}(Z \leq z)$
$p < p_0$	$\mathbb{P}(Z \leq z)$
$p \neq p_0$	$\mathbb{P}(\lvert Z \rvert \geq \lvert z \rvert) = \mathbb{P}(Z \geq z) + \mathbb{P}(Z \leq -z)$

Example: *Coffee!*

- Each subject was presented with two unmarked cups of coffee, one freshly brewed and one instant.
- Of the 50 subjects who participated in the study 19 stated that they preferred the cup containing the instant coffee.
- Let p denote the probability that a randomly chosen individual selects instant coffee in preference to freshly brewed coffee.
- Is there enough evidence to support the claim that the instant coffee tastes just as good as the freshly brewed, for a significance level of 0.05.

Example: *Coffee!*

- Hypothesis

$$H_0 : p = 0.5, \qquad H_\alpha : p \neq 0.5$$

- z-test

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

  - $\hat{p} = \frac{19}{50} = 0.38$
  - $p_0 = 0.5$
  - $z = \frac{0.38-0.5}{\sqrt{\frac{0.5(1-0.5)}{50}}} = -1.71$
- p-value

$$\mathbb{P}(|Z| \geq |z|) = \mathbb{P}(Z \geq 1.71) + \mathbb{P}(Z \leq -1.71) = 0.0436 + (1-0.9564) =$$

- $0.0872 > 0.05$

Turn on your clickers!

- Do we have enough evidence to reject the null hypothesis?

(a) Yes, since $0.0872 > 0.05$.

(b) No, since $0.0872 > 0.05$.