Introduction to Statistics - Young Researchers Fellowship Program

Lecture 6 - Foundations of Hypothesis Testing

Daniel Sánchez Pazmiño

Laboratorio de Investigación para el Desarrollo del Ecuador

October 2024

Introduction to Hypothesis Testing
●○○○

Errors
○○○

Hypothesis testing
○○○○

One sample tests about the mean
○○○○○○○○○○○○○○

Introduction to Hypothesis Testing

Null and Alternative Hypotheses

- We use a sample to make inferences about the population
- A hypothesis test helps us determine if there's enough evidence in the sample to support a statement about the population
- **Null hypothesis**: the baseline assumption
- **Alternative hypothesis**: the opposite of the null
    - Typically what we "want", the phenomenon being studied

Defining the null

- We define the null by setting our baseline scenario: what we believe might be true in "normal circumstances"

- For instance, we may think that typically the sample mean is of a certain historical value:

$$H_0 : \mu = \mu_0$$

- μ_0 is typically called "mu naught" or hypothesized mean
    - The hypothetical baseline value

Introduction to Hypothesis Testing
OOO●

Errors
OOO

Hypothesis testing
OOOO

One sample tests about the mean
OOOOOOOOOOOOOO

Defining the alternative

- We define the alternative as the contrary to the null

- Usually what we want to look out for (i.e. is the treatment effective?)

- For example, that the historical value is no longer the baseline and the population mean changed.

$$H_0 : \mu \neq \mu_0$$

Introduction to Hypothesis Testing
OOOO

Errors
●OO

Hypothesis testing
OOOO

One sample tests about the mean
OOOOOOOOOOOOOOO

Errors

Type I and Type II Errors

Decision	H_0 is True	H_0 is False
Reject H_0	Type I Error	Correct Decision
Fail to reject H_0	Correct Decision	Type II Error

Introduction to Hypothesis Testing
0000

Errors
00●

Hypothesis testing
0000

One sample tests about the mean
00000000000000

Type I and Type II Errors

- **Type I Error (α)**: Rejecting the null hypothesis when it's true
  - Example: Null: $\mu \geq 3.0$, we conclude $\mu < 3.0$
  - The probability of making this error is the significance level α
- **Type II Error (β)**: Failing to reject the null hypothesis when it's false

Hypothesis testing

Introduction to Hypothesis Testing
0000

Errors
000

Hypothesis testing
0●00

One sample tests about the mean
00000000000000

General idea of a hypothesis test

- We are asking: "if the sampling distribution is truly distributed according to the null, what is the most extreme value of a sampling exercise I can get to not doubt the null?"

- Means computing a number of standard errors that the sample statistic is located relative to the hypothesized mean.

    - The number of standard errors is called the "test statistic"

- If the null is true, this shouldn't be too far from μ_0

    - If it is far, we have to reject the null
    - Otherwise, "fail to reject" H_0.

- How far will we tolerate? Based on α, our significance level (related to confidence we want).

Introduction to Hypothesis Testing
0000

Errors
000

Hypothesis testing
00●0

One sample tests about the mean
0000000000000

p-values

- The p-value is the probability associated to the test statistic.

- Answers the question: how likely is it to get this test statistic if the null were true?

- The general rejection rule is:

$$p < \alpha$$

because it means that the p-value shows that it is more likely that we have gotten an extreme value, which does not come from the H_0 scenario, than the tolerance we've set to commit a type-I error.

Introduction to Hypothesis Testing
0000

Errors
000

Hypothesis testing
0000

One sample tests about the mean
0000000000000

General type of tests

Case 1: Left-tailed Test (One-side)

- Null hypothesis: population parameter $\mu \geq \mu_0$
- Alternative hypothesis: $\mu < \mu_0$
- Example: Average GPA of students is less than 3.0

Case 2: Right-tailed Test (One-side)

- Null hypothesis: population parameter $\mu \leq \mu_0$
- Alternative hypothesis: $\mu > \mu_0$
- Example: Average GPA of students is greater than 3.0

Case 3: Two-tailed Test (Two side)

- Null hypothesis: population parameter $\mu = \mu_0$
- Alternative hypothesis: $\mu \neq \mu_0$
- Example: Average GPA is different from 3.0

One sample tests about the mean

One Sample Tests?

- This means we're only working with one sample, not several.

- Comparing a mean against a numerical value.

- Later we will work with many samples.

The z-test (one sample)

- Tests whether the population mean μ is equal to a given value
- Used when population standard deviation is known
- We can work with a normal distribution for computing probabilities

Z-Test Types

- **Left-tailed test**: $H_0 : \mu \geq \mu_0$, $H_1 : \mu < \mu_0$
- **Right-tailed test**: $H_0 : \mu \leq \mu_0$, $H_1 : \mu > \mu_0$
- **Two-tailed test**: $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$

General procedure to do a hypothesis test

1. Compute sample mean to be used or use the given one.

2. Compute the *test statistic*. In the case of a Z-test, the test statistic is Z:

$$Z = \frac{x - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

3. Compute the p-value associated with the test statistic and given α

4. If the p-value is smaller than α, reject H_0.

Calculating p-values

■ We need to know how to calculate p-values based on each type of test, to accurately reject based on available information.

Test Type	Null Hypothesis (H_0)	Alternative Hypothesis (H_1)	Formula for p-value	R Code Example
Two-tailed	$H_0 : \mu = \mu_0$	$H_1 : \mu \neq \mu_0$	$2 \times P(Z \geq |z_{\text{score}}|)$	`2 * (1 - pnorm(abs(z_score))`
Left-tailed	$H_0 : \mu \geq \mu_0$	$H_1 : \mu < \mu_0$	$P(Z \leq z_{\text{score}})$	`pnorm(z_score)`
Right-tailed	$H_0 : \mu \leq \mu_0$	$H_1 : \mu > \mu_0$	$P(Z \geq z_{\text{score}})$	`1 - pnorm(z_score)`

Example of Left-Tailed Z-Test

- Test whether population mean is **greater than or equal** to 3.0
  - $H_0 : \mu \geq 3.0$
  - $H_1 : \mu < 3.0$
  - Sample size = 100, sample mean = 2.8, $\sigma = 0.3$
  - $z = \frac{2.8 - 3.0}{0.03} = -6.67$
- Use p-value approach or critical value approach for rejecting

Introduction to Hypothesis Testing
○○○○

Errors
○○○

Hypothesis testing
○○○○

One sample tests about the mean
○○○○○○○●○○○○○○

R implementation

- Base R does not offer a built-in package for this test, however, we may easily calculate probabilities using pnorm()

- For this, we should know how to calculate the p-value as per the table above.

R Implementation for the Example

```r
# Given values
mu_0 <- 3.0     # Hypothesized population mean
x_bar <- 2.8    # Sample mean
sigma <- 0.3    # Population standard deviation
n <- 100        # Sample size

# Calculate the standard error
se <- sigma / sqrt(n)

# Calculate the Z-score
z_score <- (x_bar - mu_0) / se

# Calculate the p-value for a left-tailed test
p_value <- pnorm(z_score)
```

Proportions

- Proportions work just like means.

- Need only to redefine the Z test statistic to follow the modified standard error for $\hat{p}$

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

where p_0 is the hypothesized population proportion.
- Apply the same rules you would for Z-tests.

R Implementation for Proportion Z-tests

```r
# Given values
p_hat <- 0.55   # Sample proportion
p_0 <- 0.50     # Hypothesized population proportion (null hypoth
n <- 100        # Sample size

# Calculate the Z-score
z_score <- (p_hat - p_0) / sqrt((p_0 * (1 - p_0)) / n)

# P-value
p_value <- 2 * (1 - pnorm(abs(z_score)))
```

Errors and Significance Levels, different rejection rules

- **Significance level (α)**: Probability of making a Type I error
    - Example: $\alpha = 0.05$ means 95% confidence level
- Compare p-value to α or use critical value
- Critical values are the number of standard errors associated with the α probability under a specific type of test.
    - Rejection rules with these vary by type of test

Introduction to Hypothesis Testing
OOOO

Errors
OOO

Hypothesis testing
OOOO

One sample tests about the mean
OOOOOOOOOOOOOO●O

Common Critical Values

α	Left-Tailed	Right-Tailed	Two-Tailed
0.05	-1.645	1.645	$\pm$ 1.96
0.01	-2.33	2.33	$\pm$ 2.58
0.10	-1.28	1.28	$\pm$ 1.64

Rejection rules for critical values, one sample tests

Test Type	Hypothesis (H_0, H_1)	Critical Value (Z)	Rejection Rule
Left-tailed	$H_0 : \mu \geq \mu_0$ vs $H_1 : \mu < \mu_0$	z_α (negative)	Reject H_0 if $z < z_\alpha$
Right-tailed	$H_0 : \mu \leq \mu_0$ vs $H_1 : \mu > \mu_0$	z_α (positive)	Reject H_0 if $z > z_\alpha$
Two-tailed	$H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$	$z_{\alpha/2}$ (positive and negative)	Reject H_0 if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$