

Chapter 8: Statistical Inference: Significance Tests about Hypothesis

Section 8.4: Decisions and Types of Errors in Significance Tests

1

Learning Objectives

1. Type I and Type II Errors
2. Significance Test Results
3. Type I Errors
4. Type II Errors
5. α , β , and Power

2

Learning Objective 1: Type I and Type II Errors

- When H_0 is true, a **Type I Error** occurs when H_0 is rejected
- When H_0 is false, a **Type II Error** occurs when H_0 is not rejected
- As $P(\text{Type I Error})$ goes **Down**, $P(\text{Type II Error})$ goes **Up**
 - The two probabilities are inversely related

3

Learning Objective 2: Significance Test Results

Type I and Type II errors are the two possible incorrect decisions. We make a correct decision if we do not reject H_0 when it is true or if we reject it when it is false.

Reality About H_0	Decision	
	Do not reject H_0	Reject H_0
H_0 true	Correct decision	Type I error
H_0 false	Type II error	Correct decision

← Type I error occurs if we reject H_0 when it is actually true.

4

Learning Objective 3: Decision Errors: Type I

- If we reject H_0 when in fact H_0 is true, this is a **Type I error**.
- If we *decide* there is a significant relationship in the population (reject the null hypothesis):
 - This is an incorrect decision only if H_0 is true.
 - The probability of this incorrect decision is equal to α .
- If we reject the null hypothesis when it *is true* and $\alpha = 0.05$:
 - There really is no relationship and the extremity of the test statistic is due to chance.
 - About 5% of all samples from this population will lead us to incorrectly reject the null hypothesis and conclude significance.

5

Learning Objective 3: $P(\text{Type I Error}) = \text{Significance Level } \alpha$

- Suppose H_0 is true. The probability of rejecting H_0 , thereby committing a Type I error, equals the significance level, α , for the test.
- We can control the probability of a Type I error by our choice of the significance level
- The more serious the consequences of a Type I error, the smaller α should be

6

Learning Objective 4: Decision Errors: Type II

- If we fail to reject H_0 when in fact H_a is true, this is a **Type II error**.
- If we *decide* not to reject the null hypothesis and thus allow for the plausibility of the null hypothesis
 - We make an incorrect decision only if H_a is true.
 - The probability of this incorrect decision is denoted by β

7

Learning Objective 5: α , β , and Power

- The probability that a fixed level α significance test will reject H_0 when a particular alternative value of the parameter is true is called the **power** of the test against that specific alternative value. Power = $1 - \beta$.
- While α gives the probability of wrongly rejecting H_0 when in fact H_0 is true, power gives *the probability of correctly rejecting H_0 when in fact H_0 should be rejected* (because the value of the parameter is some specific value satisfying the alternative hypothesis)
- When μ is close to μ_0 , the test will find it hard to distinguish between the two (low power); however, when μ is far from μ_0 , the test will find it easier to find a difference (high power).

8

Chapter 8: Statistical Inference: Significance Tests about Hypothesis

Section 8.5: Limitations of Significance Tests

9

Learning Objective

1. Statistical Significance vs. Practical Significance
2. Significance Tests Are Less Useful Than Confidence Intervals
3. Misinterpretations of Results of Significance Tests
4. Where Did the Data Come From?

10

Learning Objective 1: Statistical Significance Does Not Mean Practical Significance

- **When we conduct a significance test, its main relevance is studying whether the true parameter value is:**
 - Above, or below, the value in H_0 and
 - Sufficiently different from the value in H_0 to be of practical importance

11

Learning Objective 1: What the Significance Test Tells Us

- **The test gives us information about whether the parameter differs from the H_0 value and its direction from that value**
- **It does not tell us about the practical importance of the results**

12

Learning Objective 1: Statistical Significance vs. Practical Significance

- When the sample size is very large, tiny deviations from the null hypothesis (with little practical consequence) may be found to be statistically significant.
 - When the sample size is very small, large deviations from the null hypothesis (of great practical importance) might go undetected (statistically insignificant).
- Statistical significance is not the same thing as practical significance.*

13

Learning Objective 1: Statistical Significance vs. Practical Significance

- **A small P-value, such as 0.001, is highly statistically significant, but it does not imply an important finding in any practical sense**
- **In particular, whenever the sample size is large, small P-values can occur when the point estimate is near the parameter value in H_0**

14

Learning Objective 2: Significance Tests Are Less Useful Than Confidence Intervals

- **A significance test merely indicates whether the particular parameter value in H_0 is plausible**
- **When a P-value is small, the significance test indicates that the hypothesized value is not plausible, but it tells us little about which potential parameter values are plausible**

15

Learning Objective 3: Misinterpretations of Results of Significance Tests

- **Statistical significance does not mean practical significance**
 - **A small P-value does not tell us whether the parameter value differs by much in practical terms from the value in H_0**
- **The P-value cannot be interpreted as the probability that H_0 is true**

16

Learning Objective 3: Misinterpretations of Results of Significance Tests

- **It is misleading to report results only if they are “statistically significant”**
- **Some tests may be statistically significant just by chance**
- **True effects may not be as large as initial estimates reported by the media**

17

Learning Objective 3: Misinterpretations of Results of Significance Tests

- If H_0 represents an assumption that people have believed in for years, strong evidence (small P-value) will be needed to persuade them otherwise.
 - If the consequences of rejecting H_0 are great (such as making an expensive or difficult change from one procedure or type of product to another), then strong evidence as to the benefits of the change will be required.
- Although $\alpha = 0.05$ is a common cut-off for the P-value, there is no set border between “significant” and “insignificant,” only increasingly strong evidence against H_0 (in favor of H_a) as the P-value gets smaller.*

18

Learning Objective 4: Where Did the Data Come From?

- When you use statistical inference, you are acting as if your data are a probability sample or come from a randomized experiment.
- Statistical confidence intervals and tests cannot remedy basic flaws in producing data, such as voluntary response samples or uncontrolled experiments.
- If the data do not come from a probability sample or a randomized experiment, the conclusions may be open to challenge. To answer the challenge, ask whether the data can be trusted as a basis for the conclusions of the study.