Chapter 7

Hypothesis Testing with One Sample

Objectives:

- Sections 7-1, 7-2: Basics of Hypothesis Testing
- Section 7-3: Testing a Claim Asbout a Proportion
- Section 7-4: Testing a Claim About a Mean: σ Known.
- Section 7-5: Testing a Claim About a Mean: σ Unknown.

A quality control engineer draws a random sample of 100 cans of soft drink from the assembly line and find out that the average soda amount is 11.98 ounces. The cans are stated to contain 12 ounces of soft drink. Based on the sample mean, can we conclude that the cans are on average underfilled?

Motivating Example Continued

- Scenario 1: The mean fill amount is 12 ounces and 11.98 is due to chance variation.
- Scenario 2: The mean fill amount of 11.98 ounces signifies a true difference between stated fill amount on the cans and the process mean fill amount on the production line.
- To distinguish between Scenario 1 and Scenario 2, we need to conduct a test of **statistical hypothesis**.

Chapter 7: Sections 7-1, 7-2

Hypothesis Testing Steps:

In general, the steps to conducting a statistical test of hypothesis are as follows.

1. Write down the parameter(s) of interest.
2. Write down the *null* and *alternative* hypotheses.
3. Calculate the value of the *test statistic*.
4. Based on the test statistic, either
   (i) Calculate the *p-value* of the test and decide if it is “small". Or,
   (ii) Check to see if the test statistic fall into the *rejection region*; this is the classical approach.
5. Make a decision as to whether the null hypothesis ought to be rejected in favour of the alternative hypothesis.
6. Write a conclusion in the words of the problem.

Chapter 7: Sections 7-1, 7-2

Null & Alternative Hypotheses

Null Hypothesis

The **null hypothesis** is a statement about the population parameter of interest that is assumed to be true until it is declared false. i.e. it is the status-quo belief about the parameter. It is denoted using the notation H_0.

Null Hypothesis

The **null hypothesis** is a statement about the population parameter of interest that is assumed to be true until it is declared false. i.e. it is the status-quo belief about the parameter. It is denoted using the notation H_0.

Alternative Hypothesis

- The **alternative hypothesis** is a statement about the population parameter that is the complement of the null hypothesis. It is denoted using the notation H_1 or H_a.
- H_1 can be one-sided (e.g. $H_1 : \mu > \mu_0$, $H_1 : \mu < \mu_0$), or two-sided (e.g. $H_1 : \mu \neq \mu_0$.)

- We never prove or disprove H_0 or H_1. Instead, we either find sufficient evidence against H_0 in favour of H_a, or we fail find sufficient evidence against H_0 in favour of H_a.

- We never prove or disprove H_0 or H_1. Instead, we either find sufficient evidence against H_0 in favour of H_a, or we fail find sufficient evidence against H_0 in favour of H_a.
- In our motivating example, the parameter is μ, the mean fill amount of all soda cans in the production process.
- $H_0 : \mu = 12$.
- $H_1 : \mu < 12$.
- Technically, H_0 ought to be the compliment of H_1. However, in this course always use equality in H_0. E.g. $H_0 : \mu = \mu_0$, $H_0 : p = p_0$.

- How do we decide if there is sufficient evidence to reject H_0 in favour of H_a?
- In the classical approach, one computes a *test statistic*; this is a summary of the data in a single number.

- How do we decide if there is sufficient evidence to reject H_0 in favour of H_a?
- In the classical approach, one computes a *test statistic*; this is a summary of the data in a single number.
- One then constructs a *rejection region* using a critical value.

- If the value of the test statistic falls into the rejection region, then we reject H_0 in favour of H_a. Otherwise, we fail to reject H_0.
- In our course, for a critical value C, rejection regions have one of the following general forms:
  (i) $\{T : T > C\}$,
  (ii) $\{T : T < C\}$,
  (iii) $\{T : T < -C \text{ or } T > C\}$
- The critical value is obtained based on the *level of significance*, denoted by α.

Chapter 7: Sections 7-1, 7-2

Committing Errors

In practice, when conducting a hypothesis test, we either make the correct decision, or we might end up committing one of two types of errors. See Table 7-1 on p. 329 of textbook.

Chapter 7: Sections 7-1, 7-2

Committing Errors

In practice, when conducting a hypothesis test, we either make the correct decision, or we might end up committing one of two types of errors. See Table 7-1 on p. 329 of textbook.

Type I error

The mistake of rejecting H_0 when it is actually true.

Chapter 7: Sections 7-1, 7-2

Committing Errors

In practice, when conducting a hypothesis test, we either make the correct decision, or we might end up committing one of two types of errors. See Table 7-1 on p. 329 of textbook.

Type I error

The mistake of rejecting H_0 when it is actually true.

Type II error

The mistake of failing to reject H_0 when it is actually false.

Chapter 7: Sections 7-1, 7-2

Committing Errors

In practice, when conducting a hypothesis test, we either make the correct decision, or we might end up committing one of two types of errors. See Table 7-1 on p. 329 of textbook.

Type I error

The mistake of rejecting H_0 when it is actually true.

Type II error

The mistake of failing to reject H_0 when it is actually false.

level of significance (α)

$\alpha = P(\text{Type I error})$

Chapter 7: Sections 7-1, 7-2

Committing Errors

In practice, when conducting a hypothesis test, we either make the correct decision, or we might end up committing one of two types of errors. See Table 7-1 on p. 329 of textbook.

Type I error

The mistake of rejecting H_0 when it is actually true.

Type II error

The mistake of failing to reject H_0 when it is actually false.

level of significance (α)

$\alpha = P(\text{Type I error})$

β

$\beta = P(\text{Type II error})$

- The researcher selects α prior to data collection/decision making.
- One selects small α values such as 0.01, 0.05, 0.10.
- Ideally, we would like to control α and β simultaneously. But this is not possible.
- Traditionally, H_0 and H_1 are framed in such a way so that committing a Type I error is viewed as more serious.

Suppose we select $\alpha = 0.05$. Since α is a probability, this means that based on the relative frequency definition of probability, if we were to conduct a large number of tests, say 100, we would expect 5% of our tests to reject H_0 when it is in fact true. Or, we would expect 95% of our test to reject H_0 when H_1 is true. Our particular test for a given random sample, may be one of the 5% incorrect decisions, or it may be one of the 95% correct decisions.

P-value

The **p-value** is the probability of getting the data you got as summarized by the test statistic, or something more extreme (determined by H_1) *assuming* H_0 is true.

Decision Using P-value

Reject H_0 in favour of H_1 if the P-value $< \alpha$. Otherwise, fail to reject H_0.