CHAPTER 9 HYPOTHESIS TESTING WITH ONE SAMPLE

HYPOTHESIS TESTING

In this chapter, you will conduct hypothesis tests on single means and proportions. You will also learn about the errors associated with these tests. Hypothesis testing consists of two contradictory hypotheses or statements, a decision based on the data, and a conclusion.

To perform a hypothesis test, a statistician will:

1. Set up two contradictory hypotheses.

2. Collect sample data (in homework problems, the data or summary statistics will be given to you).

3. Determine the correct distribution to perform the hypothesis test.

4.Analyze sample data by performing the calculations that ultimately will allow you to reject or decline to reject the null hypothesis.

5. Make a decision and write a meaningful conclusion.

CHAPTER 9.1 NULL AND ALTERNATIVE HYPOTHESES

Two hypothesis: Null (H_o) and Alternative (H_a)

The null hypothesis H_o states that the parameter is equal to a specific value.

The alternative hypothesis H_a states that the parameter differs from the value specified by the null hypothesis.

The three types of alternative hypotheses are:

Left-tailed (the parameter is less than (<) the value given in the null hypothesis.

Right-tailed (the parameter is greater than (>) the value given in the null hypothesis.

Two-tailed (the parameter is not equal (≠) to the value given in the null hypothesis.

ONE-TAILED AND TWO TAILED TESTS FOR MEANS

Ho always has a symbol with an equal in it. Ha never has a symbol with an equal in it.

Left-tailed and Right-tailed are referred to as one-tailed test.

Left-tailed test	Right-tailed test	Both One-tailed and Two-tailed test
H_0: μ ≥ population mean H_a:μ < population mean	H_0: μ ≤ population mean H_a:μ > population mean	H_0: μ = population mean H_a:μ ≠ population mean or μ > population mean or μ < population mean

EXAMPLE OF ONE MEAN HYPOTHESES

The mean number of years Americans work before retiring is 34.

The null and alternative hypotheses are:

Ho: $\mu = 34$

Ha: $\mu \neq 34$

The mean number of cars a person owns in their lifetime is not more than ten.

The null and alternative hypotheses are:

Ho: $\mu \leq 10$

Ha: $\mu > 10$

EXAMPLE OF ONE MEAN HYPOTHESES

Europeans have a mean paid vacation each year of six weeks.

The null and alternative hypotheses are:

Ho: $\mu = 6$

Ha: $\mu \neq 6$

Private universities' mean tuition cost is more than $20,000 per year.

The null and alternative hypotheses are:

Ho: $\mu \leq \$20{,}000$

Ha: $\mu > \$20{,}000$

ONE-TAILED AND TWO TAILED TESTS FOR PROPORTIONS

Left-tailed test	Right-tailed test	Both One-tailed and Two-tailed test
H_0: p ≥ # H_a:p < #	H_0: p ≤ # H_a:p > #	H_0: p = # H_a:p ≠ # or p > # or p < #

EXAMPLE OF ONE PROPORTION HYPOTHESES

At most 60% of Americans vote in the presidential election.

The null and alternative hypotheses are:

Ho: $p \leq 0.60$

Ha: $p > 0.60$

Twenty-nine percent of seniors drink alcohol during the weekend.

The null and alternative hypotheses are:

Ho: $p = 0.29$

Ha: $p \neq 0.29$

EXAMPLE OF ONE PROPORTION HYPOTHESES

About half of Americans prefer to live away from cities, given the choice.

The null and alternative hypotheses are:

Ho: $p = 0.50$

Ha: $p \neq 0.50$

The chance of developing breast cancer is under 11% for women.

The null and alternative hypotheses are:

Ho: $p \geq 0.11$

Ha: $p < 0.11$

9.2 OUTCOMES AND THE TYPE I AND TYPE II ERRORS

	REALITY	
Decision	Null hypothesis is true	Null hypothesis is false
Reject null hypothesis	Type I Error	Correct Decision
Don't reject null hypothesis	Correct Decision	Type II Error

THE TYPE I AND TYPE II ERRORS

Each of the errors occurs with a particular probability.

The Greek letters **α** and **β** represent the probabilities.

α = probability of a Type I error = P(Type I error) = probability of rejecting the null hypothesis when the null hypothesis is true.

β = probability of a Type II error = P(Type II error) = probability of not rejecting the null hypothesis when the null hypothesis is false.

α and **β** should be as small as possible because they are probabilities of errors. They are rarely zero.

The Power of the Test is 1 – β. Ideally, we want a high power that is as close to one as possible. Increasing the sample size can increase the Power of the Test.

EXAMPLES OF TYPE OF ERRORS

Suppose the null hypothesis, Ho, is: Frank's rock climbing equipment is safe.

Type I error: Frank thinks that his rock climbing equipment may not be safe when, in fact, it really is safe.

Type II error: Frank thinks that his rock climbing equipment may be safe when, in fact, it is not safe.

α=probability that Frank thinks his rock climbing equipment may not be safe when, in fact, it really is safe.

β= probability that Frank thinks his rock climbing equipment may be safe when, in fact, it is not safe.

Notice that, in this case, the error with the greater consequence is the Type II error. (If Frank thinks his rock climbing equipment is safe, he will go ahead and use it.)

9.3 DISTRIBUTION NEEDED FOR HYPOTHESIS TESTING

DISTRIBUTIONS

If you are testing a single population mean, the distribution for the test is for means:

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma X}{\sqrt{n}}\right)$$

or t_{df}

If you are testing a single population proportion, the distribution for the test is for proportions or percentages:

$$P' \sim N\left(p, \frac{\sqrt{p*q}}{n}\right)$$

DISTRIBUTION NEEDED FOR HYPOTHESIS TESTING

Particular distributions are associated with hypothesis testing. Perform tests of a population mean using

a **normal distribution** or a **Student's t-distribution**.

The **normal distribution** is used when the population standard deviation is known.

The **Student's t-distribution** is used when the population standard deviation is unknown.

HYPOTHESIS TESTING USING NORMAL DISTRIBUTION

It is called a z-test

- Population mean (μ) is the mean that is used in the hypotheses
- Sample mean ($\bar{x}$) is the mean of the sample
- The sample size is n and it has to be ≥ 30 to use this test
- The population standard deviation (σ) has to be given
- The Level of significance (α) or Level of confidence (c)

HYPOTHESIS TESTING USING THE STUDENT'S T DISTRIBUTION

It is called a t-test

- A simple random sample is used
- Population standard deviation (σ) is unknown so use the sample standard deviation
- Sample size (n) is 30 or greater or normally distributed
- Note that if the sample size is sufficiently large, a t-test will work even if the population is not approximately normally distributed.

HYPOTHESIS TEST OF A SINGLE POPULATION PROPORTION P

You take a simple random sample from the population.

You must meet the conditions for a binomial distribution which are:

There are a certain number n of independent trials, the outcomes of any trial are successor failure, and each trial has the same probability of a success (p).

The shape of the binomial distribution needs to be similar to the shape of the normal distribution. To ensure this, the quantities (np) and (nq) must both be greater than five ($np>5$and $nq>5$).

Remember $q = 1 – p$ or q = non-success

9.4 RARE EVENTS, THE SAMPLE, DECISION AND CONCLUSION

RARE EVENTS

Suppose you make an assumption about a property of the population

(this assumption is the null hypothesis).

Then you gather sample data randomly. If the sample has properties that would be very unlikely to occur if the assumption is true, then you would conclude that your assumption about the population is probably incorrect.

(Remember that your assumption is just an assumption—it is not a fact and it may or may not be true. But your sample data are real and the data are showing you a fact that seems to contradict your assumption.)

EXAMPLE OF RARE EVENTS

For example, Didi and Aliar eat a birthday party of a very wealthy friend. They hurry to be first in line to grab a prize from a tall basket that they cannot see inside because they will be blindfolded. There are 200 plastic bubbles in the basket and Didi and Ali have been told that there is only one with a $100 bill. Didi is the first person to reach into the basket and pull out a bubble.

Her bubble contains a $100 bill. The probability of this happening is 1/200 =0.005. Because this is so unlikely, Ali is hoping that what the two of them were told is wrong and there are more $100 bills in the basket. A "rare event" has occurred (Didi getting the $100 bill) so Ali doubts the assumption about only one $100 bill being in the basket.

P-VALUE

Use the sample data to calculate the actual probability of getting the test result, called the **p-value**.

The **p-value** is the probability that, if the null hypothesis is true, the results from another randomly selected sample will be as extreme or more extreme as the results obtained from the given sample.

$p \leq \alpha$; Reject Null Hypothesis

P-VALUE

A large **p-value** calculated from the data indicates that we should not reject the null hypothesis.

The smaller the **p-value**, the more unlikely the outcome, and the stronger the evidence is against the null hypothesis. We would reject the null hypothesis if the evidence is strongly against it.

DECISION AND CONCLUSION

$p \leq \alpha$; Reject Null Hypothesis

This means the null hypothesis is false and your alternative hypothesis is true.

$p > \alpha$; Fail to Reject Null Hypothesis

This means the null hypothesis is true and your alternative hypothesis is false.

9.5 ADDITIONAL INFORMATION AND FULL HYPOTHESIS TEST EXAMPLES

MORE INFORMATION

- In a hypothesis test problem, you may see words such as "the level of significance is 1%." The "1%" is the preconceived or preset α.
- The statistician setting up the hypothesis test selects the value of αto use before collecting the sample data.
- If no level of significance is given, a common standard to use is α= 0.05.
- When you calculate the p-value and draw the picture, the p-value is the area in the left tail, the right tail, or split evenly between the two tails. For this reason, we call the hypothesis test left, right, or two tailed.
- The alternative hypothesis, Ha, tells you if the test is left, right, or two-tailed. It is the key to conducting the appropriate test.
- Ha never has a symbol that contains an equal sign.

THINKING ABOUT THE MEANING OF THE P-VALUE:

A data analyst (and anyone else) should have more confidence that he made the correct decision to reject the null hypothesis with a smaller p-value (for example, 0.001 as opposed to 0.04) even if using the 0.05 level for alpha.

Similarly, for a large p-value such as 0.4, as opposed to a p-value of 0.056 (alpha = 0.05 is less than either number), a data analyst should have more confidence that she made the correct decision in not rejecting the null hypothesis.

This makes the data analyst use judgment rather than mindlessly applying rules.

Z-TEST WITH EXAMPLES

Z-TEST IN CALCULATOR IF ENTERING DATA

Enter STAT, Arrow over to TESTS, chose Z-Test

The next screen will show the following for entering information:

Inpt: Data (choose if you have to enter data)

μ_0: Population mean (the mean in the hypotheses)

σ: Population standard deviation

List: L1 (where the data is)

Freq: 1

μ: choose the alternative hypothesis

Scroll to Calculate and it will go to a new screen

OUTPUT OF Z-TEST IN CALCULATOR

The output screen is:

The first line is the alternative hypothesis

z: z- test statistic

p: p-value

$\bar{x}$: sample mean

n: sample size

EXAMPLE OF HYPOTHESIS TESTING

A college football coach thought that his players could bench press a mean weight of 275 pounds. It is known that the standard deviation is 55 pounds. Three of his players thought that the mean weight was more than that amount. They asked 30 of their teammates for their estimated maximum lift on the bench press exercise. The data ranged from 205 pounds to 385 pounds.

The actual different weights were (frequencies are in parentheses) 205(3); 215(3); 225(1); 241(2); 252(2); 265(2); 275(2); 313(2); 316(5); 338(2); 341(1); 345(2); 368(2); 385(1).

Conduct a hypothesis test using a 2.5% level of significance to determine if the bench press mean is more than 275 pounds.

EXAMPLE OF HYPOTHESIS TESTING

a. Is this a test of one mean or proportion? **One mean**

b. State the null and alternative hypotheses.

H0: μ = 275 pounds

Ha : μ > 275 pounds

c. Is this a right-tailed, left-tailed, or two-tailed test? **Right tailed because it is >**

d. What symbol represents the random variable for this test? $\bar{X}$

e. In words, define the random variable for this test. **The mean weight in pounds lifted.**

f. Is the population standard deviation known and, if so, what is it? **55 pounds**

EXAMPLE OF HYPOTHESIS TESTING

g. Calculate the following:

i. $\bar{x}$= **286.16 (sample mean) from calculator**

ii. s= **55.90 (sample standard deviation) from calculator**

iii. n= **30**

h. Which test should be used? **Z-test because the population standard deviation was known.**

i. State the distribution to use for the hypothesis test.

$$\bar{X} \sim \mathrm{N}\ (\mu_X, \frac{\sigma X}{\sqrt{n}}) \ = \bar{X} \sim \mathrm{N}\ (275, \frac{55}{\sqrt{30}})$$

EXAMPLE OF HYPOTHESIS TESTING

j. Find the p-value. **0.1331**

k. At a pre-conceived α= 0.025, what is your: **p versus α 0.1331 > 0.025**

i. Decision: **fail to reject the null hypothesis**

ii. Reason for the decision: **p was larger than α**

iii. Conclusion (write out in a complete sentence):

At the 2.5% level of significance, there is not sufficient evidence to conclude that the true mean weight lifted in more than 275 pounds.

Since I failed to reject it, I need to still conclude based on not supporting the alternative hypothesis.

Z-TEST IN CALCULATOR IF NOT ENTERING DATA

Enter STAT, Arrow over to TESTS, chose Z-Test

The next screen will show the following for entering information:

Inpt: Stats (choose if you are not entering data)

μ_0: Population mean (the mean in the hypotheses)

σ: Population standard deviation

$\bar{x}$: sample mean

n: sample size

μ: choose the alternative hypothesis

Scroll to Calculate and it will go to a new screen

OUTPUT OF Z-TEST IN CALCULATOR

The output screen is:

The first line is the alternative hypothesis

z: z- test statistic

p: p-value

$\bar{x}$: sample mean

n: sample size

EXAMPLE OF HYPOTHESIS TESTING

Jeffrey, as an eight-year old, established a mean time of 16.43 seconds for swimming the 25-yard free style , with a standard deviation of 0.8 seconds. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for 15 25-yard freestyle swims. For the 15 swims, Jeffrey's mean time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the16.43 seconds.

Conduct a hypothesis test using a preset $\alpha=0.05$. Assume that the swim times for the 25-yard freestyle are normal.

EXAMPLE OF HYPOTHESIS TESTING

a. Is this a test of one mean or proportion? **One mean**

b. State the null and alternative hypotheses.

H0: $\mu = 16.43$

Ha : $\mu < 16.43$

c. Is this a right-tailed, left-tailed, or two-tailed test? **Left-tailed**

d. What symbol represents the random variable for this test? $\bar{X}$

e. In words, define the random variable for this test. **The mean time to swim the 25-yard freestyle**

f. Is the population standard deviation known and, if so, what is it?

The population standard deviation is 0.8

EXAMPLE OF HYPOTHESIS TESTING

g. Calculate the following:

i. $\bar{x}$= **16 seconds (sample mean) from problem**

ii. n= **15 from problem**

h. Which test should be used? **Z-test because the population standard deviation was known.**

i. State the distribution to use for the hypothesis test.

$$\bar{X} \sim N\left(\mu_X, \frac{\sigma X}{\sqrt{n}}\right) = \bar{X} \sim N\left(16.43, \frac{0.8}{\sqrt{15}}\right)$$

EXAMPLE OF HYPOTHESIS TESTING

j. Find the p-value. **0.0187**

k. At a pre-conceived $\alpha = 0.05$, what is your: **p versus α 0.0187 < 0.05**

i. Decision: **reject the null hypothesis**

ii. Reason for the decision: **p was smaller than α**

iii. Conclusion (write out in a complete sentence):

At the 5% level of significance, we conclude that Jeffrey swims faster with the new googles. OR

There is sufficient evidence to show that Jeffrey's mean time to swim the 25-yard freestyle is less than 16.43 seconds.

T-TEST WITH EXAMPLES

T-TEST IN CALCULATOR IF ENTERING DATA

Enter STAT, Arrow over to TESTS, chose T-Test

The next screen will show the following for entering information:

Inpt: Data (choose if you have to enter data)

μ_0: Population mean (the mean in the hypotheses)

List: L1 (where the data is)

Freq: 1

μ: choose the alternative hypothesis

Scroll to Calculate and it will go to a new screen

OUTPUT OF T-TEST IN CALCULATOR

The first line is the alternative hypothesis

t: t- test statistic

p: p-value

$\bar{x}$: sample mean

sx: sample standard deviation

n: sample size

T-TEST IN CALCULATOR IF NOT ENTERING DATA

Enter STAT, Arrow over to TESTS, chose T-Test

The next screen will show the following for entering information:

Inpt: Stats (choose if you are not entering data)

μ_0: Population mean (the mean in the hypotheses)

$\bar{x}$: sample mean

sx: sample standard deviation

n: sample size

μ: choose the alternative hypothesis

Scroll to Calculate and it will go to a new screen

OUTPUT OF T-TEST IN CALCULATOR

The first line is the alternative hypothesis

t: t- test statistic

p: p-value

$\bar{x}$: sample mean

sx: sample standard deviation

n: sample size

T-TEST EXAMPLE

Statistics students believe that the mean score on the first statistics test is 65. A statistics instructor thinks the mean score is higher than 65.He samples ten statistics students and obtains the scores 65; 65; 70; 67; 66; 63; 63; 68; 72; 71. He performs a hypothesis test using a 5% level of significance. The data are assumed to be from a normal distribution.

EXAMPLE OF HYPOTHESIS TESTING

a. Is this a test of one mean or proportion? **One mean**

b. State the null and alternative hypotheses.

H0: μ = 65

Ha : μ > 65

c. Is this a right-tailed, left-tailed, or two-tailed test? **Right-tailed**

d. What symbol represents the random variable for this test? $\overline{X}$

e. In words, define the random variable for this test. **Average score on the first test**

f. Is the population standard deviation known and, if so, what is it?

The population standard deviation is unknown so I have to run a T-test.

EXAMPLE OF HYPOTHESIS TESTING

g. Calculate the following:

i. $\bar{x}$= **67 (sample mean) from calculator**

ii. s = **3.20 sample standard deviation**

ii. n= **10 from problem**

h. Which test should be used? **T-test because the population standard deviation was known.**

i. State the distribution to use for the hypothesis test.

$t_{df} = t_9$ (since df is n – 1)

EXAMPLE OF HYPOTHESIS TESTING

j. Find the p-value. **0.0396**

k. At a pre-conceived $\alpha = 0.05$, what is your: **p versus α 0.0396 < 0.05**

i. Decision: **reject the null hypothesis**

ii. Reason for the decision: **p was smaller than α**

iii. Conclusion (write out in a complete sentence):

At the 5% level of significance, there is sufficient evidence to show that the mean test score is more than 65.

ONE PROPORTION TESTING IN CALCULATOR

Enter STAT, Arrow over to TESTS, chose 1-PropZTest

The next screen will show the following for entering information:

p_0: Percentage given in problem (the percentage in the hypotheses)

x: portion of sample affected by percentage

n: sample size

prop: choose the alternative hypothesis

Scroll to Calculate and it will go to a new screen

OUTPUT OF ONE PROPORTION TESTING

The first line is the alternative hypothesis

z: z test statistic

p: p-value

$\hat{p}$: $\frac{\boldsymbol{x}}{\boldsymbol{n}}$

n: sample size

EXAMPLE OF HYPOTHESIS TESTING

Joon believes that 50% of first-time brides in the United States are younger than their grooms. She performs a hypothesis test to determine if the percentage is the same or different from 50%. Joon samples 100 first-time brides and 53 reply that they are younger than their grooms. For the hypothesis test, she uses a 1% level of significance.

EXAMPLE OF HYPOTHESIS TESTING

a. Is this a test of one mean or proportion? **proportion**

b. State the null and alternative hypotheses.

Ho: **$p = 0.50$**

Ha: **$p \neq 0.50$**

c. Is this a right-tailed, left-tailed, or two-tailed test? **Two-tailed (hint: ≠)**

d. What symbol represents the random variable for this test? **P'**

e. In words, define the random variable for this test. **The percentage of first-time brides younger than their grooms.**

EXAMPLE OF HYPOTHESIS TESTING

g. Calculate the following:

i. x= **53 (number of brides younger than grooms)**

ii. p'**=** $\boldsymbol{\frac{x}{n}}$ **=** $\boldsymbol{\frac{53}{100}}$ **= 0.53 (in calculator)**

iii. n= **100**

h. Which test should be used? **Since no means are mentioned, use the estimated proportion distribution**

i. State the distribution to use for the hypothesis test.

$\boldsymbol{P' \sim N\ (p, \frac{\sqrt{p*q}}{n})} \quad = \boldsymbol{P' \sim N\ (0.5, \frac{\sqrt{0.5 * 0.5}}{100})}$

EXAMPLE OF HYPOTHESIS TESTING

j. Find the p-value. **0.5485**

k. At a pre-conceived α= 0.01, what is your: **0.5485 > 0.01**

i. Decision: **fail to reject the null hypothesis**

ii. Reason for the decision: **p is greater than α**

iii. Conclusion (write out in a complete sentence):

At the 5% level of significance, there is not sufficient evidence to conclude that the percentage of first-time brides who are younger than their grooms is different from 50%.