Section 3.7 – Hypothesis Tests for a Population Mean

Objective

Test a claim about a population mean (with σ known) by using a formal method of hypothesis testing.

Notation

n = Sample size
$\overline{x}$ = sample mean
$\mu_{\overline{x}}$ = population mean of all sample means from samples of size n
σ = known value of the population standard deviation

Requirements for Testing Claims about a Population Mean (with σ Known)

1. The sample is a simple random sample.
2. The value of the population standard deviation σ is known.
3. Either or both of these conditions is satisfied: The population is normally distributed or $n > 30$.

Test Statistic for Testing a Claim About a Mean (with σ Known)

To test hypotheses regarding the population mean assuming the population standard deviation is unknown, we use the *t*-distribution rather than the *Z*–distribution. When we replace σ with s

$$z = \frac{\overline{x} - \mu_{\overline{x}}}{\frac{\sigma}{\sqrt{n}}}$$

follows **Student's *t*-distribution** with $n - 1$ degrees of freedom.

Properties of the *t*–Distribution

1. The *t*-distribution is different for different degrees of freedom.
2. The *t*-distribution is centered at 0 and is symmetric about 0.
3. The area under the curve is 1. Because of the symmetry, the area under the curve to the right of 0 equals the area under the curve to the left of 0 equals 1/2.
4. As t increases (or decreases) without bound, the graph approaches, but never equals, 0.
5. The area in the tails of the *t*-distribution is a little greater than the area in the tails of the standard normal distribution because using s as an estimate of σ introduces more variability to the *t*-statistic.
6. As the sample size n increases, the density curve of t gets closer to the standard normal density curve. This result occurs because as the sample size increases, the values of s get closer to the values of σ by the Law of Large Numbers.

Testing Hypotheses Regarding a Population Mean

To test hypotheses regarding the population mean, we use the following steps, provided that:

- ✓ The sample is obtained using simple random sampling.
- ✓ The sample has no outliers, and the population from which the sample is drawn is normally distributed or the sample size is large ($n \geq 30$).
- ✓ The sampled values are independent of each other.

***Step* 1**: Determine the null and alternative hypotheses. The hypotheses can be structured in one of three ways:

Two–Tailed	***Left–Tailed***	***Right–Tailed***
$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$	$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$

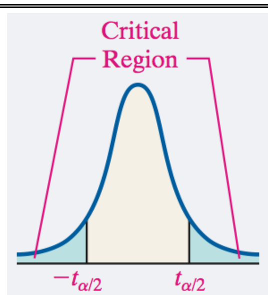
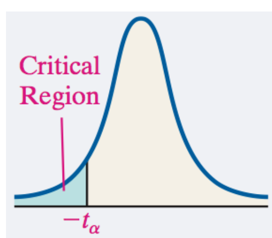
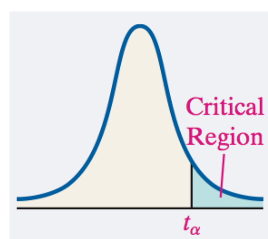
μ_0 is assumed value of the population mean.

***Step* 2:** Select a level of significance, α, based on the seriousness of making a Type I error.

Classical Approach

Step* 3**: Compute the ***test statistic $t_0 = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

which follows the Student's t-distribution with $n - 1$ degrees of freedom.

Two–Tailed	***Left–Tailed***	***Right–Tailed***
		

P-Value Approach

Step* 3:** Compute the ***test statistic:

***Step* 4:** If the P-value $< \alpha$, reject the null hypothesis.

***Step* 5:** State the conclusion.

The procedure is robust, which means that minor departures from normality will not adversely affect the results of the test. However, for small samples, if the data have outliers, the procedure should not be used.

Example

People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: $n = 40$ and $\bar{x} = 172.55$ lb. Research from several other sources suggests that the population of weights of men has a standard deviation given by $\sigma = 26$ lb. Use these results to test the claim that men have a mean weight greater than 166.3 lb, which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and use the *P*-value method outlined in Figure 8-8.

Solution

Requirements are satisfied: simple random sample, σ is known (26 lb), sample size is 40 ($n > 30$)

Step 1: Express claim as $\mu > 166.3$ lb

Step 2: alternative to claim is $\mu \leq 166.3$ lb

Step 3: $\mu > 166.3$ lb does not contain equality, it is the alternative hypothesis:

H_0 : $\mu = 166.3$ lb. null hypothesis

H_1 : $\mu > 166.3$ lb. alternative hypothesis and original claim

Step 4: significance level is $\alpha = 0.50$

Step 5: claim is about the population mean, so the relevant statistic is the sample mean (172.55 lb), σ is known (26 lb), sample size greater than 30

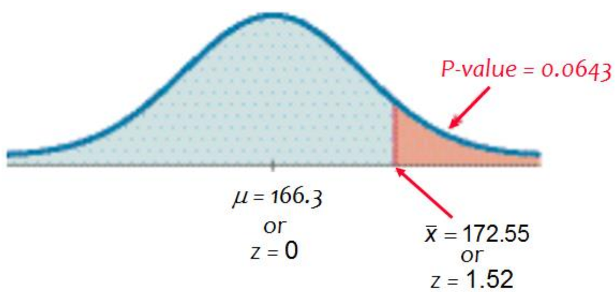
Step 6: calculate z:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}} = \frac{172.55 - 166.3}{\frac{26}{\sqrt{40}}} = 1.52$$

Right--tailed test, so *P*-value is the area is to the right of $z = 1.52$;

From the Normal Distribution Table; area to the left of $z = 1.52$ is 0.9357, so the area to the right is $1 - 0.9357 = 0.0643$. The *P*-value is 0.0643

***Step* 7:** The *P*-value of 0.0643 is greater than the significance level of $\alpha = 0.05$, we fail to reject the null hypothesis.



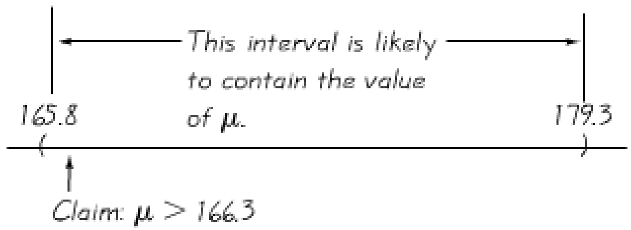
- ✓ The *P*-value of 0.0643 tells us that if men have a mean weight given by $\mu = 166.3$ lb, there is a good chance (0.0643) of getting a sample mean of 172.55 lb. A sample mean such as 172.55 lb could easily occur by chance. There is not sufficient evidence to support a conclusion that the population mean is greater than 166.3 lb, as in the National Transportation and Safety Board's recommendation.

The ***traditional method***: Use $z = 1.645$ instead of finding the P-value. Since $z = 1.52$ does not fall in the critical region, again fail to reject the null hypothesis.

Confidence Interval method: Use a one-tailed test with a = 0.05, so construct a 90% confidence interval:

$165.8 < \mu < 179.3$

The confidence interval contains 166.3 lb, we cannot support a claim that μ is greater than 166.3. Again, fail to reject the null hypothesis.



Underlying Rationale of Hypothesis Testing

If, under a given assumption, there is an extremely small probability of getting sample results at least as extreme as the results that were obtained, we conclude that the assumption is probably not correct. When testing a claim, we make an assumption (null hypothesis) of equality. We then compare the assumption and the sample results and we form one of the following conclusions:

- If the sample results (or more extreme results) can easily occur when the assumption (null hypothesis) is true, we attribute the relatively small discrepancy between the assumption and the sample results to chance.
- If the sample results cannot easily occur when that assumption (null hypothesis) is true, we explain the relatively large discrepancy between the assumption and the sample results by concluding that the assumption is not true, so we reject the assumption.

Distinguish between *Statistical* Significance and *Practical* Significance

When a large sample size is used in a hypothesis test, the results could be statistically significant even though the difference between the sample statistic and mean stated in the null hypothesis may have no ***practical significance***.

> ***Practical significance*** refers to the idea that, while small differences between the statistic and parameter stated in the null hypothesis are statistically significant, the difference may not be large enough to cause concern or be considered important.

Large sample sizes can lead to statistically significant results while the difference between the statistic and parameter is not enough to be considered practically significant.

Exercises *Section* 3.7 – Hypothesis Tests for a Population Mean

1. Because the amounts of nicotine in king size cigarettes listed below

1.1	1.7	1.7	1.1	1.1	1.4	1.1	1.4	1	1.2	1.1	1.1	1.1
1.1	1.1	1.8	1.6	1.1	1.2	1.5	1.3	1.1	1.3	1.1	1.1	

We must satisfy the requirement that the population is normally distributed. How do we verify that a population is normally distributed?

2. If you want to construct a confidence interval to be used for testing the claim that college students have a mean IQ score that is greater than 100, and you want the test conducted with a 0.01 significance level, what confidence level should be used for the confidence interval?

3. A jewelry designer claims that women have wrist breadths with a mean equal to 5 cm. A simple random sample of the wrist breadths of 40 women has a mean of 5.07 cm. Assume that the population standard deviation is 0.33 cm. Use the accompanying TI display to test the designer's claim.

```
Z-Test
 μ≠5
 z=1.341572341
 p=.1797348219
 x̄=5.07
 n=40
```

Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

4. The U.S. Mint has a specification that pennies have a mean weight of 2.5 g. Assume that weights of pennies have a standard deviation of 0.0165 g and use the accompanying Minitab display to test the claim that the sample is from a population with a mean that is less than 2.5 g. These Minitab results were obtained using the 37 weights of post 1983 pennies.

```
Test of mu = 2.5 vs < 2.5. Assumed s.d. = 0.0165
                               95% Upper
 N      Mean     StDev     Bound          Z        P
37   2.49910   0.01648   2.50356      -0.33    0.370
```

5. In the manual "How long to have a Number One the Easy Way," by KLF Publications, it is stated that a song "must be no longer than 3 minutes and 30 seconds" (or 210 seconds). A simple random sample of 40 current hit songs results in a mean length of 252.5 sec. Assume that the standard deviation of song lengths is 54.5 sec. Use a 0.05 significance level to test the claim that the sample is from a population of songs with a mean greater than 210 sec. What do these results suggest about the advice given in the manual?

6. A simple random sample of 50 adults is obtained, and each person's red blood cell count (in cells per microliter) is measured. The sample mean is 5.23. The population standard deviation for red blood cell counts is 0.54. Use a 0.01 significance level to test the claim that the sample is from population with a mean less than 5.4, which is value often used for the upper limit of the range of normal values. What do the results suggest about the sample group?

7. A simple random sample of 106 body temperature with a mean of 98.20 °F. Assume that σ is known to be 0.62 °F. Use a 0.05 significance level to test the claim that the mean body temperature of the population is equal to 98.6 °F, as is commonly believed. Is there sufficient evidence to conclude that the common belief is wrong?

8. When 40 people used the Weight Watchers diet for one year, their mean weight loss was 3.0 lb. Assume that the standard deviation of all such weight changes is $\sigma = 4.9$ lb. and use a 0.01 significance level to test the claim that the mean weight loss is greater than 0. Based on these results, does the diet appear to be effective? Does the diet appear to have a practical significance?

9. The health of the bear population in Yellowstone National Park is monitored by periodic measurements taken from anesthetized bears. A sample of 54 bears has a mean weight of 182.9 lb. Assuming that σ is known to be 121.8 lb. use a 0.05 significance level to test the claim that the population mean of all such bear weights is greater than 150 lb.

10. A simple random sample of 401 salaries of NCAA football coaches in the NCAA has a mean of $415,953. The standard deviation of all salaries of NCAA football coaches is $463,364. Use a 0.05 significance level to test the claim that the mean salary of a football coach in the NCAA is less than $500,000.

11. A simple random sample of 36 cans of regular Coke has a mean volume of 12.19 oz. Assume that the standard deviation of all cans of regular Coke is 0.11 oz. Use a 0.01 significance level to test the claim that cans if regular Coke have volumes with a mean of 12 oz., as stated on the label. If there is a difference, is it substantial?

12. A simple random sample of FICO credit rating scores is obtained, and the scores are listed below.

714	751	664	789	818	779	698	836	753	834	693	802

As the writing, the mean FICO score was reported to be 678. Assuming the standard deviation of all FICO scores is known to be 58.3, use a 0.05 significance level to test the claim that these sample FICO scores come from a population with a mean equal to 678.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

13. Listed below are recorded speeds (in mi/h) of randomly selected cars traveling on a section of Highway 405 in Los Angeles.

68	68	72	73	65	74	73	72	68	65	65	73	66	71	68	74	66	71	65	73
59	75	70	56	66	75	68	75	62	72	60	73	61	75	58	74	60	73	58	75

That part of the highway has posted speed limit of 65 mi/h. Assume that the standard deviation od speeds is 5.7 mi/h and use a 0.01 significance level to test the claim that the sample is from a population with a mean that is greater than 65 *mi/h*.
Identify the null hypothesis, alternative hypothesis, test statistic, *P*-value or critical value(s), conclusion about the null hypothesis, and final conclusion that address the original claim.

14. Given a simple random sample of speeds of cars on Highway in CA, you want to test the claim that the sample that the sample values are from a population with a mean greater than the posted speed limit of 65 *mi/hr*. Is it necessary to determine whether the sample is from a normally distributed population? If so, what methods can be used to make that determination?

15. In statistics, what does ***df*** denote. If a simple random sample of 20 speeds of cars is to be used to test the claim that the sample values are from a population with a mean greater than the posted speed limit of 65 mi/h, what is the specific value of *df*?

16. Claim about IQ scores of statistics instructors: $\mu > 100$, sample data: $n = 15, \quad \bar{x} = 118, \quad s = 11$. The sample data appear to come from a normally distributed population with unknown μ and σ. Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

17. Claim about FICO credit scores of adults: $\mu = 678$, sample data: $n = 12, \quad \bar{x} = 719, \quad s = 92$. The sample data appear to come from a population with a distribution that is not normal, and σ Is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

18. Claim about daily rainfall amounts in Boston: $\mu < 0.20\, in$, sample data: $n = 19, \bar{x} = 0.10\, in, s = 0.26\, in.$

19. The sample data appear to come from a population with a distribution that is very far from normal, and σ is unknown Determine whether the hypotheses test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither.

20. Testing a claim about the mean weight of M&M's: Right-tailed test with n = 25 and test statistic t = 0.430. Find the *P*-value and find a range of values for the *P*-value.

21. Test a claim about the mean body temperature of healthy adults: left-tailed test with n = 11 and test statistic t = −3.158. Find the *P*-value and find a range of values for the *P*-value.

22. Two-tailed test with n = 15 and test statistic t = 1.495. Find the *P*-value and find a range of values for the *P*-value.

23. In an analysis investigating the usefulness of pennies, the cents portions of 100 randomly selected checks are recorded. The sample has mean of 23.8 *cents* and a standard deviation of 32.0 *cents*. If the amounts from 0 *cents* to 99 *cents* are all equally likely, the mean is expected to be 49.5 *cents*. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 49.5 *cents*. What does the result suggest about the cents portions of the checks?

24. A simple random sample of 40 recorded speeds (in *mi/h*) is obtained from cars traveling on a section of Highway 405 in Los Angeles. The sample has a mean of 68.4 *mi/h* and a standard deviation of 5.7 mi/h. Use a 0.05 significance level to test the claim that the mean speed of all cars is greater than the posted speed limit of 65 *mi/h*.

25. The heights are measured for the simple random sample of supermodels. They have mean height of 70.0 *in*. and a standard deviation of 1.5 *in*. Use a 0.01 significance level to test the claim that supermodels have heights with a mean that is greater than the mean heights of 63.6 *in*. for women in general population. Given that there are only nine heights represented, can we really conclude that supermodels are taller than the typical woman?

26. The National Highway Traffic Safety Administration conducted crash tests of child booster seats for cars. Listed below are results from those tests, with measurement given in hic (standard *head injury condition* units). The safety requirement is that the hic measurement should be less than 1000 *hic*. Use a 0.01 significance level to test the claim that the sample is from a population with a mean less than 1000 *hic*.

774 649 1210 546 431 612

Do the results suggest that all of the child booster seats meet the specified requirement?

27. The trend of thinner Miss America winners has generated charges that the contest encourages unhealthy diet habits among young women. Listed below are body mass indexes (BMI) for recent Miss America winners. Use a 0.01 significance level to test the claim that recent Miss America winners are from a population with a mean BMI less than 20.16, which was the BMI for winners from the 1920s and 1930s.

19.5 20.3 19.6 20.2 17.8 17.9 19.1 18.8 17.6 16.8

Do recent winners appear to be significantly different from those in the 1920s and 1930s?

28. The list measured voltage amounts supplied directly to the author's home

123.8	123.9	123.9	123.3	123.4	123.3	123.3	123.6	123.5	123.5	123.5	123.7
123.6	123.7	123.9	124.0	124.2	123.9	123.8	123.8	124.0	123.9	123.6	123.5
123.4	123.4	123.4	123.4	123.3	123.3	123.5	123.6	123.8	123.9	123.9	123.8
123.9	123.7	123.8	123.8								

The Central Hudson power supply company states that it has a target power supply of 120 volts. Using those home voltage amounts, test the claim that the mean is 120 volts. Use a 0.01 significance level.

29. When testing a claim about a population mean with a simple random sample selected from a normally distributed population with unknown σ, the student t distribution should be used for finding critical values and/or a *P*–value. If the standard normal distribution is incorrectly used instead, does that mistake make you more or less likely to reject the null hypothesis, or does it not make a difference? Explain.

30. The list measured human body temperature.

98.6	98.6	98.0	98.0	99.0	98.4	98.4	98.4	98.4	98.6	98.6	98.8	98.6	97.0	97.0
98.8	97.6	97.7	98.8	98.0	98.0	98.3	98.5	97.3	98.7	97.4	98.9	98.6	99.5	97.5
97.3	97.6	98.2	99.6	98.7	99.4	98.2	98.0	98.6	98.6	97.2	98.4	98.6	98.2	98.0
97.8	98.0	98.4	98.6	98.6	97.8	99.0	96.5	97.6	98.0	96.9	97.6	97.1	97.9	98.4
97.3	98.0	97.5	97.6	98.2	98.5	98.8	98.7	97.8	98.0	97.1	97.4	99.4	98.4	98.6
98.4	98.5	98.6	98.3	98.7	98.8	99.1	98.6	97.9	98.8	98.0	98.7	98.5	98.9	98.4

98.6	97.1	97.9	98.8	98.7	97.6	98.2	99.2	97.8	98.0	98.4	97.8	98.4	97.4	98.0
97.0														

Use the temperatures listed for 12 AM on day 2 to test the common belief that the mean body temperature is 98.6 °F. Does that common belief appear to be wrong?

31. Determine whether the hypothesis test involves a sampling distribution of means that is a normal distribution, student *t* distribution, or neither

a) Claim $\mu = 981$. Sample data: $n = 20, \quad \bar{x} = 946, \quad s = 27$. The sample data appear to come from a normally distributed population with $\sigma = 30$.

b) Claim $\mu = 105$. Sample data: $n = 16, \quad \bar{x} = 101, \quad s = 15.1$. The sample data appear to come from a normally distributed population with unknown μ *and* σ.

32. Assume the resting metabolic rate (RMR) of healthy males in complete silence is 5710 kJ/day. Researchers measured the RMR of 45 healthy males who were listening to calm classical music and found their mean RMR to be 5708.07 with a standard deviation of 992.05.
At the α = 0.05 level of significance, is there evidence to conclude that the mean RMR of males listening to calm classical music is different than 5710 kJ/day?

33. People have died in boat accidents because an obsolete estimate of the mean weight of men was used. Using the weights of the simple random sample of men from Data Set 1 in Appendix B, we obtain these sample statistics: n = 40 and $\bar{x} = 172.55$ *lb.*, and σ = 26.33 *lb*. Do not assume that the value of σ is known. Use these results to test the claim that men have a mean weight greater than 166.3 *lb*., which was the weight in the National Transportation and Safety Board's recommendation M-04-04. Use a 0.05 significance level, and the traditional method.