L17. Hypothesis Test on the Mean; Single Sample, Variance Unknown

Example 1

Consider a hypothesis test where $H_0 : \mu = 205$ and $H_1 : \mu > 205$. A random sample of 14 observations taken from a population that is normally distributed produced a sample mean of 212.37 and a standard deviation of 16.35. At the 1% level of significance, is there enough evidence to reject the null hypothesis?

Solution

We are given: $n = 14 \quad \overline{x} = 212.37 \quad s = 16.35$

Set up the null hypothesis

$$H_0 : \mu = 205$$

Set up the alternate hypothesis

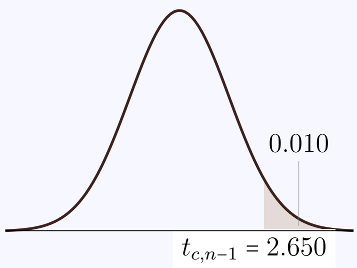
$$H_1 : \mu > 205$$ Right tail test.

Using the level of significance and H_1, draw the rejection zone, and find the critical values of t that sets the cut-off values for α

This is a right tail test. The level of significance is $\alpha = 1\%$, and the degrees of freedom is

$$df = n - 1 = 14 - 1 = 13$$

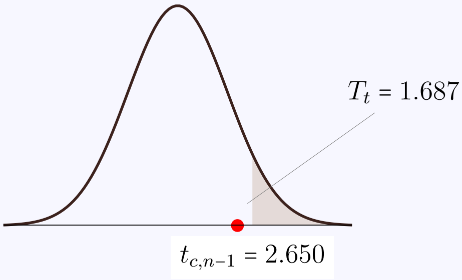
Thus, the the critical value associated with this test is $t_{0.01,13} = 2.650$.



Calculate the test statistic.

$$T_t = \frac{\overline{x} - k}{s/\sqrt{n}} = \frac{212.37 - 205}{16.35/\sqrt{14}} = 1.687$$

Make a statistical decision



$\because T_t = 1.687$ is **not** in the rejection zone $\Rightarrow$ Fail to reject H_0

To do this without drawing a graph, compare the $t-$ critical value against the T_t, the test statistic.

$$|T_t| = 1.687 < |t_c| = 2.650 \quad \Rightarrow \quad \text{Fail to reject } H_0.$$

Estimate the $P-$value to corroborate the statistical decision.

To do this first, be on the row for degrees of freedom that is associated with the hypothesis test. Next, find the two values for which the test-statistic is sandwiched between. Once you have isolated these two values, move up the column until you hit the row that says α. Read off the values; these will serve as the range for estimating the size of the $P-value$

$$0.05 < P-value < 0.075$$

Since the $P-value$ is somewhere 0.05 and 0.075, we can still conclude that it is larger than the level of significance (5%).

$$\alpha = 0.05 < P-value < 0.075$$

and fail to reject H_0

Conclusion in the context of the problem.

At the 1% level of significance, there is sufficient evidence to indicate that the mean for this population is greater than 205.

Example 2

Consider a hypothesis test where $H_0 : \mu = 50$ and $H_1 : \mu < 50$. A random sample of 8 observations taken from a population that is normally distributed produced a sample mean of 44.98 and a standard deviation of 6.77. At the 5% level of significance, is there enough evidence to reject the null hypothesis?

Solution

We are given: $n = 8 \quad \overline{x} = 44.98 \quad s = 6.77$

Null hypothesis

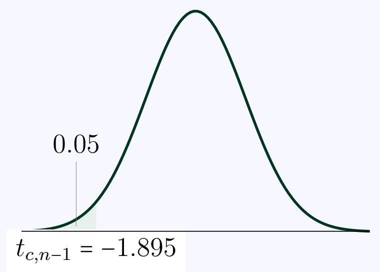
$$H_0 : \mu = 50$$

Alternate hypothesis

$$H_1 : \mu < 50$$ Left tail test.

Critical value of t for the test.

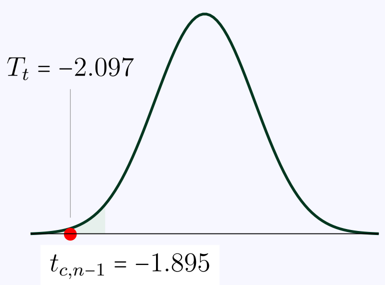
$$t_{0.05,7} = -1.895$$



Calculate the test statistic.

$$T_t = \frac{\overline{x} - k}{s/\sqrt{n}} = \frac{44.98 - 50}{6.77/\sqrt{8}} = -2.097$$

Statistical decision



$\because T_t = -2.097$ is in the rejection zone, $\Rightarrow$ reject H_0.

Alternate approach: compare the $t-$ critical value against the T_t, the test statistic.

$$|t_c| = 1.895 < |T_t| = 2.097 \quad \Rightarrow \quad \text{Reject } H_0.$$

Estimate the $P-$value to corroborate the statistical decision.

$T_t = 2.097$ is between the values of 2.365 and 2.998. This implies that

$$0.010 < P - value < 0.025 < \alpha = 0.05$$

Since the $P - value$ is smaller than the level of significance (5%) reject H_0

Conclusion:

At the 5% level of significance, there is sufficient evidence to indicate that the mean for this population is less than 50.

Example 3

Consider a hypothesis test where $H_0 : \mu = 10.70$ and $H_1 : \mu \neq 10.70$. A random sample of 47 observations taken from a population produced a sample mean of 12.025 and a standard deviation of 4.90. At the 1% level of significance, is there enough evidence to reject the null hypothesis?

Solution

We are given: $n = 47 \quad \overline{x} = 12.025 \quad s = 4.9$

$H_0 : \mu = 10.70$
$H_1 : \mu \neq 10.70$ Double tail test.

Critical value: $t_{0.005,45} = \pm 2.690$.

Test statistic:

$$T_t = \frac{\overline{x} - k}{s/\sqrt{n}} = \frac{12.025 - 10.70}{4.9/\sqrt{47}} = 1.854$$

$$\because \quad |t_c| = 2.690 > |T_t| = 1.854 \quad \Rightarrow \quad \text{Fail to reject } H_0.$$

$P-$value

The value $T_t = 1.854$, is between 1.465 and 1.690

$$0.050 < P - value < 0.075$$

Since this is a **double tail test**, the $P-value$ has to be **doubled**

$$2 \cdot (0.050) < P-value < 2 \cdot (0.075)$$
$$0.010 < P-value < 0.015$$

Comparing this against α, we have

$$\alpha = 0.01 < 0.010 < P-value < 0.015$$

Since the $P-value$ is greater than the level of significance (1%) we fail to reject H_0

Conclusion:

At the 1% level of significance, there is insufficient evidence to indicate that the mean for this population is different from 10.70

Example 4

The President of a university claims that the mean time spent partying by students at this university is less 11 hours per week. A random sample of 40 students taken from this university showed that they spent an average of 10.5 hours partying, with a standard deviation of 2.3 hours. Test the President's claim at the 2.5% level of significance.

Solution

We have: $n = 40$ $\quad \overline{x} = 10.5$ hrs/wk $\quad s = 2.3$ hrs/wk $\quad \alpha = 0.025$

$$H_0 : \mu = 11$$
$$H_1 : \mu < 11$$

Critical value: $t_{0.025,39} = 2.023$

Test statistic: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{10.5 - 11}{2.3/\sqrt{40}} = -1.375$

$$\because \quad |t_c| = 2.023 > |T_t| = 1.375 \quad \Rightarrow \quad \text{Fail to Reject } H_0$$

Estimated $P-$ value:

$$\alpha = 0.025 < P-value < 0.05$$

$\because \quad P-value > \alpha \quad \Rightarrow \quad$ Fail to reject H_0

Conclusion

At the 2.5% level of significance, there is insufficient evidence to indicate that the average amount of time spent partying by students at this university is less than 11 hours per week.

Example 5

Before Listerine was sold as a mouthwash, it was used as a surgeon's antispetic (in its undiluted form), advertised as a floor cleaner, and even touted as a cure for gonorrhea.

Listerine mouthwash is sold in 1.5 litre bottles. The machine responsible for filling the bottles is checked for regularly to ensure that it is properly calibrated. A random sample of 32 bottles were taken from a production run, and it was found that the average amount of mouthwash in the bottles was 1.507 litres, with a standard deviation of 0.03 litres.

At the 10% level of significance, does the data indicate that the machine is dispensing a volume that is different from 1.50 litres, and hence requires recalibration? What is the $P-$value for this test?

Solution

We have: $n = 32 \quad \overline{x} = 1.507\,L \quad s = 0.03\,L \quad \alpha = 0.10$

$$H_0 : \mu = 1.5$$
$$H_1 : \mu \neq 1.5$$

Critical value: $t_{0.05,31} = 1.696$

Test statistic: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{1.507 - 1.5}{0.03/\sqrt{32}} = 1.320$

$$\because \quad |t_c| = 1.696 > |T_t| = 1.320 \quad \Rightarrow \quad \text{Fail to Reject } H_0$$

Estimated $P-$ value:

$$0.15 < P - value < 0.20$$

$\because \quad \alpha = 0.10 < 0.15 < P - value < 0.20 \quad \Rightarrow \quad$ Fail to reject H_0

Conclusion

At the 10% level of significance, there is insufficient evidence to indicate that the average amount of Listerine inside the bottles is different from $1.5\,L$; the machine does not require recalibration.

Example 6

The mean balance of all checking accounts at a bank on December 31, 2011 was \$850. A random sample of 55 checking accounts taken recently from this bank gave a mean balance of \$880 with a standard deviation of \$75. Using a 1% significance level, can you conclude that the mean balance has increased during this period? What is the $P-value$ for this test?

Solution

We have: $n = 55 \quad \overline{x} = \$\,880 \quad s = \$\,75 \quad \alpha = 0.01$

$$H_0 : \mu = 850$$
$$H_1 : \mu > 850$$

Critical value: $t_{0.01,54} = 2.403$

Test statistic: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{880 - 850}{75/\sqrt{55}} = 2.966$

$$\because \quad |t_c| = 2.403 < |T_t| = 2.966 \quad \Rightarrow \quad \text{Reject } H_0$$

Estimated $P-$ value:

$$0.0005 < P - value < 0.005$$

$\because \quad 0.0005 < P - value < 0.005 < 0.01 = \alpha \quad \Rightarrow \quad$ Reject H_0

Conclusion

At the 1% level of significance, there is sufficient evidence to indicate that the average bank balance did increase during this period.

Example 7

The article *A Critical Appraisal of 98.6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich* reported body temperature, gender, and heart rate for a number of subjects. The body temperatures for 25 female subjects resulted in a sample average of $\overline{x} = 98.264°F$ and a standard deviation of $s = 0.4821°F$.

a. Assuming that the population is normally distributed, does the data indicate at the 0.05 level of significance that the average body temperature for women is different from $98.6°F$ What is the $P-value$?

b. Construct and explain how the question in (a) could be answered by constructing a two-sided confidence interval on the mean female body temperature.

c. State the Type I and Type II error in the context of the problem.

Solution

a. We have: $n = 25 \quad \overline{x} = 98.264°F \quad s = 0.04821°F \quad \alpha = 0.050$

$H_0 : \mu = 98.6$ Historical value of body temperature
$H_1 : \mu \neq 98.6$ Is different; double tail test

Critical value: $t_{c,n-1} = t_{0.025,24} = \pm 2.064$

Test statistic: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{98.264 - 98.6}{0.4821/\sqrt{25}} = -3.485$

$\because \quad |t_c| = 2.064 < |T_t| = 3.485 \quad \Rightarrow \quad$ Reject H_0

At the 5% level of significance, there is sufficient evidence to indicate average body temperature of women from this population is different from $98.6°F$

The $P-value$ for this test is

$$2(0.0005) < P-value < 2(0.0050)$$
$$0.001 < P-value < 0.010$$

b. The two-sided 95% confidence interval is

$$\overline{x} \pm t_{c,n-1}\frac{s}{\sqrt{n}} \quad \Rightarrow \quad 98.264 \pm 2.064\frac{0.4821}{\sqrt{25}}$$
$$\Rightarrow \quad 98.264 \pm 0.1990$$
$$\Rightarrow \quad 98.065 < \mu < 98.463$$

With repeated sampling, we are 95% confident that the true average body temperature for women from this population is between 98.065 and 98.463°F. Since 98.6 is not in the confidence interval, reject H_0.

c. **Type I Error:** Reject H_0 when it is true.

Conclude that the average body temperature for females from this population is different from 98.6°F when in fact it is not.

Type II Error: Fail to reject H_0 when it is false.

Conclude that the average body temperature for females from this population is 98.6°F when in fact it is not.

Example 8

A manufacturer of running shoes knows that the average lifetime for a particular model of shoes is 15 months. Someone in the research and development division of the shoe company claims to have developed a longer lasting product. This new product was worn by 36 individuals and lasted on average for 17 months. The variability of the original shoe is estimated based on the standard deviation of the new group which is 5.5 months.

a. At the 2.5% level of significance, test the designer's claim that he has developed a shoe which lasts longer than 15 months and write a conclusion in the context of the problem. What is the $P - value$ for this test?

b. Construct an appropriate one-sided confidence bound to support the conclusion obtained in (a)

c. State the Type I and Type II errors in the context of problem.

Solution

a. We have: $n = 36$ $\quad \overline{x} = 17$ months $\quad s = 5.5$ months

$H_0 : \mu = 15$ Designer's claim
$H_1 : \mu > 15$ Lasts longer; right tail test

If $\alpha = 0.025 \quad \Rightarrow \quad t_{c,n-1} = t_{0.025,35} = 2.030$

Test statistic: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{17 - 15}{5.5/\sqrt{36}} = 2.182$

$\because \quad |t_c| = 2.030 < |T_t| = 2.182 \quad \Rightarrow \quad$ Reject H_0

At the 2.5% level of significance, there is enough evidence to indicate that the shoe lasts longer than 15 months. $\quad 0.0005 < P - value < 0.0050$

b. H_0 was rejected.

$$\overline{x} - t_{c,n-1}\frac{s}{\sqrt{n}} \leq \mu \quad 17 - 2.030\frac{5.5}{\sqrt{36}} \leq \mu \quad \Rightarrow \quad 15.1392 \leq \mu$$

With repeated sampling, we are 97.5% confident the shoes last on average at least 15.1392 months.

In order to show that the alternate hypothesis ($\mu > 15$) is correct, we need to show that the lower bound for the one-sided confidence bound exceeds 15; which it does. So reject H_0.

c. **Type I Error:** Reject H_0 when it is true.

Conclude that the shoes on average last longer than 15 months, when they actually don't.

Type II Error: Fail to reject H_0 when it is false.

Conclude that the shoes lasts an average of 15 months when they do not.

Example 9

Sweethearts, or conversation hearts, are small heart shaped candies sold around Valentine's Day. First produced in 1901, the candies feature classic messages like: I ♡ You, Let's Kiss, and True Love. NECCO, the company which makes conversation hearts, adds 10 new messages each year. Recent additions include: Text Me, Hey You!, and Yeah Right.

The company claims that conversational hearts have a shelf life of 5 years. A random sample of 200 hearts taken from the warehouse found that the average shelf life of the sample was 58 months with a standard deviation of 4.5 months. Assume that the population is normally distributed.

a. At the 0.05 level of significance, does the data indicate that the average shelf life of these candies is less than 5 years? What is the $P-value$ for this test.

b. Construct and explain how an appropriate interval estimate could be used to support the conclusion obtained in (a).

Solution

a. We have: $n = 200$ $\quad \overline{x} = 58$ months $\quad s = 4.5$ months

$H_0 : \mu = 60$ Producer's claim
$H_1 : \mu < 60$ Less than; left tail test

If $\alpha = 0.050 \quad \Rightarrow \quad t_{c,n-1} = t_{0.050,100} = 1.660$

The test statistic is: $T_t = \dfrac{\overline{x} - k}{s/\sqrt{n}} = \dfrac{58 - 60}{4.5/\sqrt{200}} = -6.285$

$\because \quad |t_c| = 1.660 < |T_t| = 6.285 \quad \Rightarrow \quad \text{Reject } H_0$

At the 5% level of significance, there is enough evidence to indicate that the average shelf-life of these candies is less than 60 months (5 years).

Since the value $T_t = -6.285$ appears after the last column, we can conclude that the $P - value$ is:

$$0.0005 > P - value$$

b. H_0 was rejected, and since the hypothesis test was one-tailed, the appropriate interval estimate will be one-sided to match.

$$\mu \leq \overline{x} + t_{c,n-1}\frac{s}{\sqrt{n}} \quad \Rightarrow \quad \mu \leq 58 + 1.66\frac{4.5}{\sqrt{200}} \quad \Rightarrow \quad \mu \leq 58.5282$$

With repeated sampling, we are 95% confident the average shelf life of these candies is at most 58.5282 months.