

Chapter 18. Inference About a Population Mean

Note. In this chapter, we finally liberate ourselves from the requirement of having the population standard deviation, as is required in the z procedures. We will again find confidence intervals and perform hypothesis tests.

Conditions for Inference

Note. The conditions for inference about a mean include:

- We can regard our data as a **simple random sample** (SRS) from the population. This condition is very important.
- Observations from the population have a **normal distribution** with mean μ and standard deviation σ . In practice, it is enough that the distribution be symmetric and single-peaked unless the sample is very small. Both μ and σ are unknown parameters.

Definition. When the standard deviation of a statistic is estimated from data, the result is called the **standard error** of the statistic. The standard error of the sample mean $\bar{x}$ is $s/\sqrt{n}$.

The t -Distributions

Definition. Draw a SRS of size n from a large population that has the normal distribution with mean μ and standard deviation σ . The **one-sample t statistic**

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

has the t **distribution** with $n - 1$ degrees of freedom.

Note. Some properties of the t distributions are:

- The density curves of the t distributions are similar in shape to the standard normal curve.
- The spread of the t distributions is a bit greater than that of the standard normal distribution.
- As the degrees of freedom increase, the t density curve approaches the standard normal curve.

Figure 18.1 on page 436 illustrates these properties.

The One-Sample t Confidence Interval

Definition. Draw a SRS of size n from a large population having unknown mean μ . A level C **confidence interval for μ** is

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

where t^* is the critical value for the $t(n - 1)$ density curve with area C between $-t^*$ and t^* . The interval is exact when the population distribution is normal and is approximately correct for large n in other cases.

Example S.18.1. Stooge Confidence Intervals.

A stoogeologist wants a confidence interval of the number of times Moe is violent to the third stooge per film. A sample of $n = 10$ Three Stooges films with Curly as the third stooge reveals a sample mean of $\bar{x} = 14.3$ violent acts per film and sample standard deviation of $s = 10.46$ violent acts per film. A sample of $n = 10$ Three Stooges films with Shemp as the third stooge reveals a sample mean of $\bar{x} = 10.3$ violent acts per film and sample standard deviation of $s = 6.22$ violent acts per film. Give two 95% confidence intervals for the mean number of times Moe is violent to the third stooge (Curly and Shemp) per film. This data is from “Hypothesis Testing

Using the Films of the Three Stooges” by R. Gardner and R. Davidson, to appear in *Teaching Statistics*.

Partial Solution. The degree of freedom is $n - 1 = 9$, so we use $t^* = 2.262$.

The One-Sample t Test

Note. The one-sample t test consists of the following steps: Draw a SRS of size n from a large population having unknown mean μ . To **test the hypothesis** $H_0 : \mu = \mu_0$, compute the **one-sample t statistic**

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

In terms of a variable T having the $t(n - 1)$ distribution, the P -value for a test of H_0 against

$$H_a : \mu > \mu_0 \text{ is } P(T \geq t)$$

$$H_a : \mu < \mu_0 \text{ is } P(T \leq t)$$

$$H_a : \mu \neq \mu_0 \text{ is } 2P(T \geq |t|)$$

These P -values are exact if the population distribution is normal and are approximately correct for large n in other cases.

Example S.18.2. t -Testing Stooges.

The Stoogeologist watches all of the Three Stooges films with Joe Besser as the third stooge. There are 16 such films and the stoogeologist finds that Moe is violent towards Joe an average of $\mu_0 = 2.94$ times per film (this data is from “Hypothesis Testing Using the Films of the Three Stooges” by R. Gardner and R. Davidson, to appear in *Teaching Statistics*). NOTICE THAT THIS IS NOT A SAMPLE, BUT IN FACT A POPULATION PARAMETER!!! SINCE THE POPULATION OF MOE-LARRY-JOE FILMS IS SMALL (SIZE 16), WE CAN SIMPLY WATCH ALL OF THE FILMS AND WE DO NOT NEED TO STATISTICALLY ESTIMATE THIS PARAMETER!!! It is widely held among Stooge fans that the Joe Besser episodes were the least dynamic. To motivate a hypothesis along these lines, consider the following two quotes concerning Joe Besser’s Stooges years. An article “Pure Slap Shtick” by Richard von Busack (1997) includes the quote:

“To Bessers eternal shame, he had it written into his contract that he would not be subject to slapping or bodily harm.”

In addition, the Mackinac Media DVD *The Men Behind the Mayhem* (2004) has an interview of Joe in which he says:

“Moe and Larry—they were the best. I enjoyed every minute of it with them. In fact, to show you how wonderful they were, I never liked to be hit with anything. And Larry used to say to me ‘dont worry Joe, I’ll take it.’ Now thats the kind of guys they were.”

It appears that Joe was the recipient of less violence than the other Stooges. Perform a one-sample t test on the null hypotheses “ H_0 : Curly was the recipient of violence by Moe per film *the same number of times* as Joe was” against the alternate hypothesis “ H_a : Curly was the recipient of violence by Moe per film *more times* than Joe was.” With μ_C as the mean number of times Moe was violent to Curly per film, we need to test:

$$H_0 : \mu_C = 2.94 \text{ and } H_a : \mu_C > 2.94.$$

Use the sample data from Example S.18.1. Do the same for the Shemp data.

Partial Solution. We can either use Table C and do this by hand, or we can input the data into Minitab. For the Curly question, in Minitab click on the **Stat** tab, select **Basic Statistics**, and **1-Sample t....** In the pop-up window,

select **Summarized data**, enter the **Sample size** (10), **Mean** (14.3), **Standard deviation** (10.46), and **Test mean** (2.94). You are interested in $H_0 : \mu_C > 2.94$, so click on the **Options** tab and in the **Alternative** pull-down menu, select **greater than**. Click **OK** twice and the results will be displayed. A P value of 0.004 is displayed.

Matched Pairs t Procedures

Note. To compare the responses to the two treatments in a matched pairs design, find the difference between the responses within each pair. Then apply the one-sample t procedures to these differences. Minitab can perform computations for matched pairs.

Example. Exercise 18.11, pages 446 and 447.

Robustness of t Procedures

Note. A confidence interval or significance test is called **robust** if the confidence level or P -value does not change very much when the conditions for use of the procedure are violated.

Note. The conditions under which a t test should be used include:

- Except in the case of small samples, the condition that the data are a SRS from the population of interest is more important than the condition that the population distribution is normal.
- *Sample size less than 15:* Use t procedures if the data appear close to normal (roughly symmetric, single peak, no outliers). If the data are skewed or if outliers are present, do not use t .
- *Sample size at least 15:* The t procedures can be used except in the presence of outliers or strong skewness.
- *Large samples:* The t procedures can be used even for clearly skewed distributions when the sample is large, roughly $n \geq 40$.

Example S.18.3. Stooge Assumptions.

We used a t test in Example S.18.3. Do you have any concerns about the use of a t test in that setting? How would you explore the validity of the use of a t test? How might you modify the test or data collection to compensate for your concerns?

Note. We cannot use the one-sample t test to perform a hypothesis test which compares the mean number of times per film Moe was violent to Curly versus the mean number of times per film Moe was violent to Shemp (since we have sample *twice*; i.e., we have sampled both the Curly and Shemp populations of films). However, there is a similar procedure called a two-sample t test (available in Minitab). It requires that two populations are sampled and the means of the populations are compared. The two-sample test statistic is:

$$r = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

For more details, see Chapter 19 (especially page 465).

rbg-3-28-2009