MODULE 16

1-SAMPLE T-TEST

Objectives:

1. Identify when a 1-Sample t-Test is appropriate.
2. Perform the 11 steps of a significance test in a 1-Sample t-Test situation.

Contents

PRIOR TO THIS MODULE, hypothesis testing methods required knowing σ, which is a parameter that is seldom. When σ is replaced by its estimator, s, the test statistic follows a Student's t rather than a standard normal (Z) distribution. In this module, the t-distribution is described and a 1-Sample t-Test for testing that the mean from one population equals a specific value is discussed. A 2-sample t-Test for comparing means from two populations is in Module 17.

16.1 t-distribution

A t-distribution is similar to a standard normal distribution (i.e., N(0,1)) in that it is centered on 0 and is bell shaped (Figure 16.1). The t-distribution differs from the standard normal distribution in that it is heavier in the tails, flatter near the center, and its exact dispersion is dictated by a quantity called the degrees-of-freedom (df). The t-distribution is "flatter and fatter" because of the uncertainty surrounding the use of s rather than σ in the standard error calculation.[1] The degrees-of-freedom are related to n and generally come from the denominator in the standard deviation calculation. As the degrees-of-freedom increase, the t-distribution becomes narrower, taller, and approaches the standard normal distribution (Figure 16.1).

[1] Recall that the sample standard deviation is a statistic and is thus subject to sampling variability.

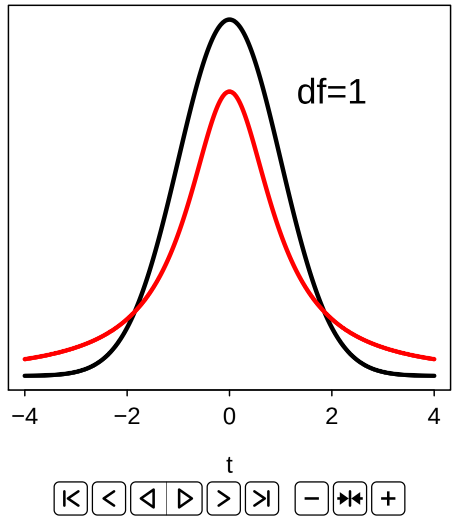

Figure 16.1. Standard normal (black) and t-distributions (red) with varying degrees-of-freedom.

> ◇ **A t-distribution is "wider" than a Z-distribution because of the extra uncertainty from using s rather than σ in the standard error.**

Proportional areas on a t-distribution are computed using `distrib()` similar to what was described for a normal distribution in Modules 7 and 12. The major exceptions for using `distrib()` with a t-distribution is that `distrib="t"` must be used and the degrees-of-freedom must be given in `df=` (how to find the df will be discussed in subsequent sections). For example, the area right of $t = -1.456$ on a t-distribution with 9 df is 0.9103 (Figure 16.2).

```
> ( distrib(-1.456,distrib="t",df=9,lower.tail=FALSE) )
[1] 0.9103137
```

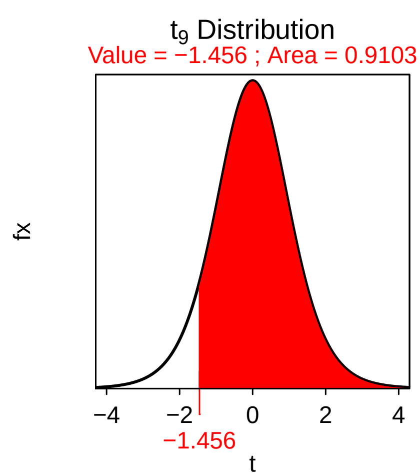

Figure 16.2. Depiction of the area to the right of $t = -1.456$ on a t-distribution with 9 df.

Similarly, the t with an upper-tail area of 0.95 on a t-distribution with 19 df is -1.729 (Figure 16.3).[2]

```
> ( distrib(0.95,distrib="t",type="q",df=19,lower.tail=FALSE) )
[1] -1.729133
```

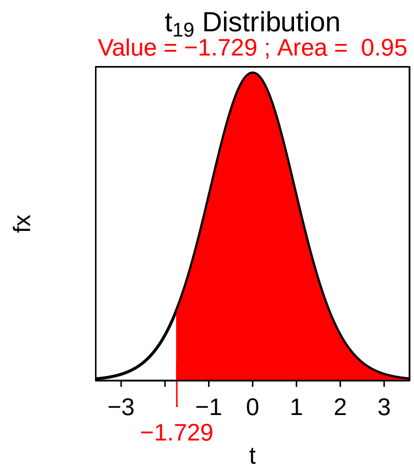

Figure 16.3. Depiction of the value of t with an area to the right of 0.95 on a t-distribution with 19 df.

Review Exercises

16.1 What is the p-value if $H_A : \mu < 125$, $t = -2.178$, and $df = 35$? *Answer*

16.2 What is t^* for the previous question if $\alpha = 0.05$? *Answer*

16.3 What is the p-value if $H_A : \mu > 125$, $t = 1.856$, and $df = 81$? *Answer*

16.4 What is t^* for the previous question if $\alpha = 0.01$? *Answer*

16.5 What is the p-value if $H_A : \mu \neq 125$, $t = -2.178$, and $df = 99$? *Answer*

16.6 What is t^* for the previous question if $\alpha = 0.10$? *Answer*

16.2 1-Sample t-Test Specifics

A 1-Sample t-Test is similar to a 1-Sample Z-test in that both test the same H_0. The difference, as discussed above, is that when σ is replaced by s, the test statistic becomes t and the scaling factor for confidence regions becomes a t^*. Other aspects are similar between the two tests as shown in Table 16.1.[3]

[2] This last "reverse" calculation would be t^* for a 95% lower confidence bound.
[3] Compare Table 16.1 to Table 15.1.

Table 16.1. Characteristics of a 1-Sample t-Test.

- **Hypothesis:** $H_0 : \mu = \mu_0$
- **Statistic:** $\bar{x}$
- **Test Statistic:** $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}}$
- **Confidence Region:** $\bar{x} + t^* \frac{s}{\sqrt{n}}$
- **df**: $n-1$
- **Assumptions:**
  1. σ is UNknown
  2. $n > 40$, $n > 15$ and the **sample** (i.e., histogram) is not strongly skewed, OR the **sample** is normally distributed.
- **When to Use:** Quantitative response, one population, σ is UNknown.

16.2.1 Example - Purchase Catch of Salmon?

Consider the following situation,

> *A prospective buyer will buy a catch of several thousand salmong if the mean weight of all salmon in the catch is at least 19.9 lbs. A random selection of 50 salmon had a mean of 20.1 and a standard deviation of 0.76 lbs. Should the buyer accept the catch at the 5% level?*

The 11-steps (Section 15.1) for completing a full hypothesis test for this example are as follows:

1. α=0.05.
2. $H_0 : \mu = 19.9$ lbs vs. $H_A : \mu > 19.9$ lbs where μ is the mean weight of ALL salmon in the catch.
3. A 1-Sample t-Test is required because (1) a quantitative variable (weight) was measured, (ii) individuals from one population were sampled (this catch of salmon), and (iii) σ is **UN**known.[4]
4. The data appear to be part of an observational study with random selection.
5. (i) $n = 50 > 40$ and (ii) σ is unknown.
6. $\bar{x}$=20.1 lbs (and s=0.76 lbs).
7. t=$\frac{20.1-19.9}{\frac{0.76}{\sqrt{50}}}$=$\frac{0.2}{0.107}$=1.87 with df=$50-1=49$.
8. p-value=0.0337.
9. H_0 is rejected because the p-value $< \alpha = 0.05$.
10. The average weight of ALL salmon in this catch appears to be greater than 19.9 lbs; thus, the buyer should accept this catch of salmon.
11. I am 95% confident that the mean weight of ALL salmon in the catch is greater than 19.92 lbs (i.e., $20.1 - 1.677\frac{0.76}{\sqrt{50}}$=$20.1-0.18$=19.92).

R Appendix:

```
( distrib(1.87,distrib="t",df=49,lower.tail=FALSE) )
( distrib(0.95,distrib="t",type="q",df=49,lower.tail=FALSE) )
```

[4] If σ is given, then it will appear in the background information to the question and will be in a sentence that uses the words "population", "assume that", or "suppose that."

Review Exercises

16.7 A general achievement test is standardized so that students should average 80 with a standard deviation of 5 (this is for the entire population not the population of students at the school described below). The superintendent at a school in a large district would like to show that her students averaged better than the 80 points. To test this, she had the test given to 32 randomly selected students from her school. The summary statistics for those 32 students are: mean=83.2, median=82.5, standard deviation=5.5, and IQR=7. Perform the appropriate hypothesis test for this superintendent at the 0.05 level. Answer

16.8 The Northwestern University Placement center conducts random surveys on starting salaries of college graduates and publishes the results every year. The Dean of the College of Liberal Arts suggested to prospective students that graduates from the College would earn more than $32000 as a starting salary on average. The results in the table below are from a part of the Placement Center's results for graduates of the College of Liberal Arts for the year just prior to the Dean's statements [Note that the measurements are in 1000s of dollars.]. Use these results at the 10% level to determine the correctness of the Dean's statement. Answer

```
 n   Min.  1st Qu. Median 3rd Qu.   Max.   Mean  StDev
42  29.30   31.30   32.50   33.80  36.80 32.511  1.713
```

16.3 1-Sample t-Test in R

If raw data exist, the calculations for a 1-Sample t-test can be efficiently computed with `t.test()`. The arguments to `t.test()` are the same as those for `z.test()`, with the exception that `sd=` is not used with `t.test`. Thus, `t.test()` requires the vector of quantitative data as the first argument, the hypothesized value for μ in `mu=`, the type of alternative hypothesis in `alt=` (again, can be `alt="two.sided"` (the default), `alt="less"`, or `alt="greater"`), and the level of confidence as a proportion in `conf.level=` (defaults to 0.95). The use of `t.test()` is illustrated in the following example.

16.3.1 Example - Crab Body Temperature

Consider the following situation,

> *A marine biologist wanted to determine if the body temperature of crabs exposed to ambient air temperature would be different than the ambient air temperature. The biologist exposed a sample of 25 crabs to an air temperature of 24.3^{0}C for several minutes and then measured the body temperature of each crab. The body temperatures for individual crabs is shown below. Perform a hypothesis test (at the $\alpha = 0.01$) level to answer the biologist's question.*

```
22.9,22.9,23.3,23.5,23.9,23.9,24.0,24.3,24.5,24.6,24.6,24.8,24.8,
25.1,25.4,25.4,25.5,25.5,25.8,26.1,26.2,26.3,27.0,27.3,28.1
```

The 11-steps (Section 15.1) of a hypothesis test for this example are as follows:

1. α=0.01.
2. $H_0 : \mu = 24.3^0$C vs. $H_A : \mu \neq 24.3^0$C, where μ is the mean body temperature of ALL crabs.
3. A 1-Sample t-Test is required because (1) a quantitative variable (temperature) was measured, (ii) individuals from one population were sampled (an ill-defined population of crabs), and (iii) σ is **UN**known.
4. The data appear to be part of an experimental study (the temperature was controlled) with no suggestion of random selection of individuals.
5. (i) $n = 25 > 15$ and the sample distribution of crab temperatures appears to be only slightly right-skewed (Figure 16.4) and (ii) σ is **UN**known.
6. $\bar{x}$=25.03^{0}C (Table 16.2).
7. t=2.713 with 24 df (Table 16.2).
8. p-value=0.0121 (Table 16.2).
9. H_0 is not rejected because the p-value $> \alpha = 0.01$.
10. It appears that the average body temperature of ALL crabs is not different than the ambient temperature of 24.3^{0}C.
11. A confidence interval is not required as H_0 was not rejected. [*However, this confidence interval shows that the mean body temperature of ALL crabs is likely between 24.28^{0}C and 25.78^{0}C (Table 16.2). Note that this interval contains 24.3^{0}C which is why H_0 was not rejected.*]

R Appendix:

```
df <- read.csv("data/CrabTemps.csv")
hist(~ct,data=df,xlab="Crab Body Temp (C)")
( ct.t <- t.test(df$ct,mu=24.3,conf.level=0.99) )
plot(ct.t)
```

Table 16.2. Results from 1-Sample t-Test for body temperature of crabs.

```
t = 2.7128, df = 24, p-value = 0.01215
99 percent confidence interval:
 24.27741 25.77859
sample estimates:
mean of x
   25.028
```

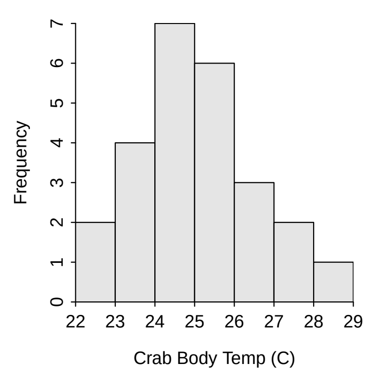

Figure 16.4. Histogram of the body temperatures of n=25 crabs exposed to an ambient temperature of 24.3^{0}C.

Review Exercises

16.9 Fishing line is graded by the pounds (lbs) of pressure that it can withstand before breaking. For example, line that is rated as 6-lbs should not break for pressures under 6 lbs. Two physics students developed an apparatus for testing the breaking point of 2-foot sections of line to test the manufacturer's claim (i.e., determine if line rated at 6-lbs broke, on average, at pressures below 6 lbs). To test this, they measured the pounds of pressure it took 20 randomly selected 2-foot sections of line to break. Use their results shown below to test their hypothesis at the 10% level. Answer

```
6.1 5.3 5.5 4.9 6.2 6.5 5.7 5.5 4.7 6.2
6.8 5.9 5.8 6.7 6.3 6.2 5.4 5.5 6.7 5.9
```

16.10 Last year I planted 400 everbearing strawberry plants in my garden. The company I bought the plants from claimed that in the year following planting, each plant would produce an average of 12 berries. I was surprised by this claim and hypothesized that the plants would actually produce less than what the company said, on average. To test this claim, I counted the number of ripe berries produced for the entire season on 50 randomly selected plants. Use the data in Strawberries.csv to test the company's claim at the 10% level. Answer

16.11 The toy industry rates toys regarding their ease for being put together in three categories: easy, moderate, and difficult. A toy is placed into the easy category if it takes 10 minutes or less to put the toy together, in the moderate category if it takes 20 minutes or less (and more than 10 minutes), and in the difficult category if it takes more than 20 minutes. A randomly selected group of 34 adults were asked to put together a new toy to determine which rating the toy should receive. The results from these 34 individuals are in ToyTime.csv. Conduct a hypothesis test, at the 10% level, to determine whether the toy should receive the difficult rating. Answer

16.12 One of the dominant uses of Madison area lakes is for boating. To develop a long-term data set on the temporal fluctuations and trends in such activity, the Long Term Ecological Research (LTER) project obtained records of boat traffic that passes through the locks at the head of the Yahara River on its stretch between Lake Mendota and Lake Monona. These data in Yahara.csv) have been collected nearly daily from April through October since 1976. Use these data to determine, at the 5% level, if the mean total number of boats passing through the locks during the months of June, July, and August of 2005 is greater than 75. HINT: create a new data frame that contains just the data for this period (i.e., the data file contains more data than is needed for this question). I suggest that you do this in three separate steps – isolate 2005 data, isolate data for months after May (5), and then isolate data for months before September (9). Answer

16.13 The golden rectangle is a rectangle with a length-to-width ratio of 1:1.618, or equivalently, a width-to-length ratio of 0.618:1 (See a description of the golden rectangle here). The golden rectangle is evident in several works by ancient Greeks and Egyptians. Anthropologists measured the width-to-length ratios of beaded rectangles used by the Shoshoni Indians of America to decorate their leather goods. Use their data[5] in Shoshoni.csv to determine, at the 5% level, if the golden rectangle is evident in the beadwork of the Shoshonis. Answer

[5] This question and these data originated at OzDASL.