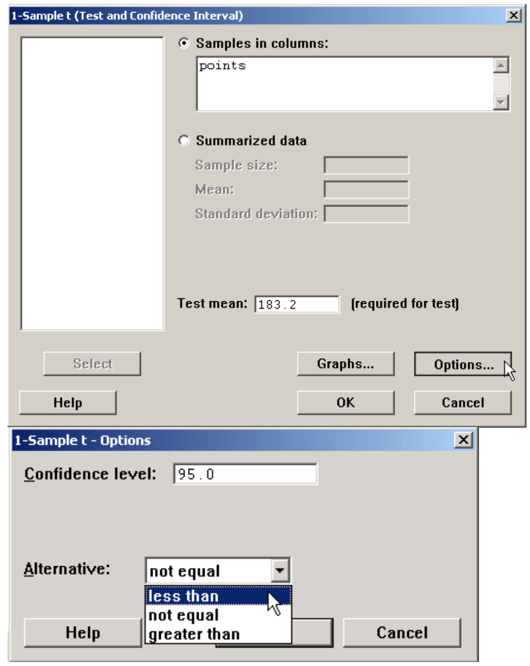
Using Minitab

Choose **Stat** ➢ **Basic Statistics** ➢ **1-Sample t**
Double click on the variable in the left window so it appears in the Variables box.
Specify the hypothesized value for the mean.

Under the Graphs button, you can also produce a histogram, boxplot, or dotplot of the sample.

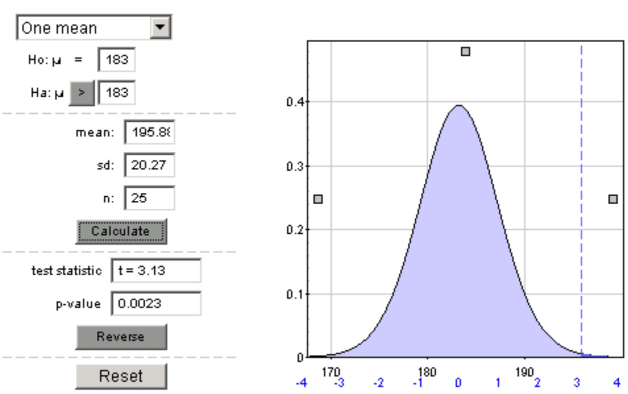
Click the Options button
Specify the direction of the alternative



Using Test of Significance Calculator Applet (when you have the *summary statistics*)
Choose "One mean" option

Specify hypothesized value and direction of alternative (>)

Specify the sample mean, sample standard deviation, and sample size and click Calculate.



Practice: Filling cola bottles

Bottles of a popular cola are supposed to contain 300 milliliters (ml). There is some variation from bottle to bottle because the filling machinery is not perfectly precise. However, the distribution of contents follow a normal distribution. An inspector who suspects that the bottler is underfilling measures the contents of six bottles. The results are:

299.4 297.7 301.0 298.9 300.2 297.0

Is this convincing evidence that the mean content of cola bottles is less than the advertised 300 ml?

Practice: Children's Television Viewing

Researchers at Stanford studied whether reducing children's television viewing might help to prevent obesity. Third and fourth grade students at two public elementary schools in San Jose were the subjects. One of the schools incorporated a curriculum designed to reduce watching television and playing video games, while the other school made no changes to its curriculum. At the beginning and end of the study a variety of variables were measured on each child. These included body mass index, triceps skinfold thickness, waist circumference, waist-to-hip ratio, weekly time spent watching television, and weekly time spent playing video games.

(a) Identify the observational units in this study.

(b) Specify one explanatory variable and one response variable in this study.

(c) Is this an observational study or an experiment?

(d) At the beginning of the study, children were asked to report how many hours of television they watch in a typical week. The 198 responses had a mean of 15.41 hours and a standard deviation of 14.16 hours. Do these data provide evidence at the .05 level for concluding that third and fourth graders watch more than two hours of television per day on average?

Solution

Filling cola bottles

Observational units = bottles, variable = amount of cola (quantitative)

So we are dealing with one sample mean

Given: population normal, plugging these 6 values into Minitab we get $\bar{x}$ = 299.03, s = 1.50 ml.

The sample is these 6 bottles, the population is presumed to be all bottles made by this machinery.

1. Let μ = mean about of soda in bottles *made by this process*

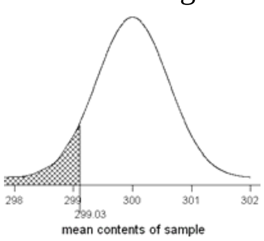
2. H_0: μ = 300 (they are not filled less than advertised)

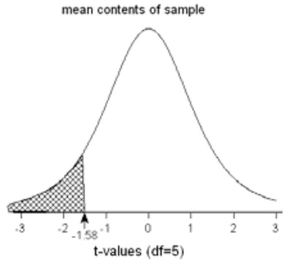
   H_a: μ < 300 (suspected the mean content of all cola bottles is less than advertised)

   Note, you can consider H_0 as saying $\mu \geq 300$, all that really matters is the "edge"

3. *Technical conditions*: We are not told that these 6 bottles were a simple random sample from some large population, but it seems reasonable to conclude that they are representative of bottles coming from this process. We should at least consider the point in time they were measured. Since we were told the population followed a normal distribution, we can use the one-sample *t*-test even with the small sample size. If we were considering the sample means, they would follow a normal distribution with mean 300 (assuming H_0 is true) and standard deviation $\sigma/\sqrt{6}$ (first sketch).

Since we don't know σ, we will approximate SD($\bar{x}$) by $s/\sqrt{n}$ = $1.50/\sqrt{6}$ =.612 and work with the *t*-distribution (second sketch) with df=6-1 = 5 degrees of freedom





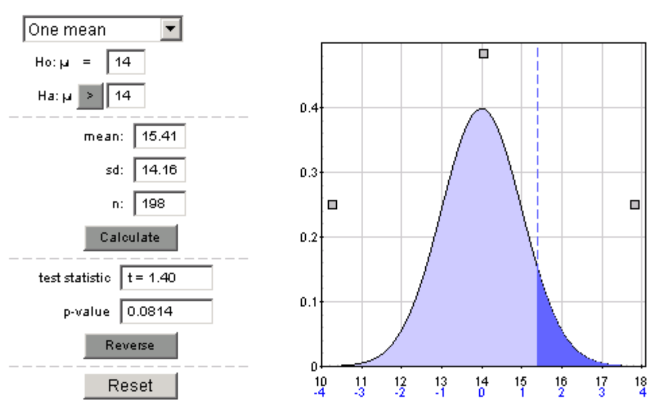
4. Test statistic: $t = \dfrac{299.03 - 300}{.612}$ =-1.58 5. p-value (Table III with 5 df): bewteen .05 and .10

6. While there is some evidence against H_0, I would not consider this strong evidence (.05≤p-value ≤.10). I choose to fail to reject H_0 at the 5% significance level. If the machine was not underfilling, it is not super surprising to get a sample mean of only 299.03 or less.

Conclusion: Based on a sample of 5 bottles, assuming they are representative of the overall process, I do not find overwhelming evidence that the mean content of all cola bottles is less than advertised.

Note: Since we are given the actual observations, we can use Minitab: t = -1.58, p-value=.088
Using the java applet with $\bar{x}$ and s, we get t=-1.58, p-value = .087

<u>Children's Television Viewing Habits</u>
(a) students
(b) EVs = weekly time spent watching television, weekly time spent playing videos
RVs = body mass index, skinfold thickness, waist circumference, waist-to-hip ratio
(c) This is an experiment since they actively changed the curriculum, though it doesn't appear that the individual students were randomized to the treatment groups (we can't send them to different schools)
(d) Using variable = how many hours of television (quantitative) we are dealing with means
Given: n=198, $\bar{x}$ =15.41 hours per week, s = 14.16, α = .05
The researchers state "third and fourth graders" but the 198 children were not randomly selected from this population, appearing to be more of a convenience sample, so we will have some caution in generalizing these results to all third and fourth graders, even to just those at public elementary schools in San Jose.
1. Let μ = mean amount of hours watched by all third and fourth graders (see note above)
2. H_0: μ=14 (2 hours per day translates to 14 hours per week)
 H_a: μ>14 (wanted to know if there was evidence they watched *more than* 2 hours per day on average)
3. *Technical conditions*: We have serious doubts about this being a random sample, but we can still ask how often we would get a sample mean this far from 14 just by chance. We don't know anything about the population distribution but since n=198 ≥ 30 we can consider that condition met. So we will apply the one-sample t procedure. Since we don't have the individual data values we cannot use Minitab to perform this calculation.



4. Test statistic $t = \dfrac{15.41 - 14}{14.16/\sqrt{198}}$ =1.401
5. p-value (Table III with df=100): between .05 and .10
6. We were told to use the 5% significance level, since our p-value .082>.05, we fail to reject H_0. We would get a sample mean at least this large in about 8% of samples when μ=14. This is not a "statically significant" difference at the 5% level.

7. We do not have strong evidence that 3rd and 4th grade students (from this population) watch more than 2 hours per day on average. However, we must be cautious in generalizing these results to a large population since they ere not a random sample.

Common oversights:

- Not realizing this is a test about "means." We are not saying that no children watch more than 14 hours per week, just that the *average amount watched* does not appear to be more than 14 hours.
- Keep in mind that the p-value is a probability statement about the sample data. We are not saying there is an 8.2% chance that H_0 is true, but are completing this calculation assuming H_0 is true.
- Forgetting to first decide whether you have quantitative or categorical data and trying to apply the wrong formula
    - Trying to plug in a round peg in a square hole… e.g., if you don't have a standard deviation to plug in, you might be using the wrong formula