

# Z-Scores and $t$ -tests - Solutions to Examples 5 and 6

Ideas from Chapters 5, 17, and 20

## Example 5

Suppose I take a sample of 100 babies who were born in December 1998 and record their gestation times. Conduct a hypothesis test of the claim that the mean gestation time is 40 weeks, at the  $\alpha = 0.05$  significance level. As part of your work, show how the p-value is calculated. No need to check the conditions for inference for this example. You may use the following R output:

```
babies_sample %>%
  summarize(
    mean_gestation = mean(gestation),
    sd_gestation = sd(gestation)
  )

## # A tibble: 1 x 2
##   mean_gestation sd_gestation
##           <dbl>         <dbl>
## 1           38.83         2.238574

t.test(~gestation, mu = 40, data = babies_sample)

## ~gestation
##
## One Sample t-test
##
## data:  gestation
## t = -5.2265, df = 99, p-value = 9.611e-07
## alternative hypothesis: true mean is not equal to 40
## 95 percent confidence interval:
##  38.38582 39.27418
## sample estimates:
## mean of x
##      38.83

pt(-5.226542, df = 99)

## [1] 4.805429e-07

2 * pt(-5.226542, df = 99)

## [1] 9.610858e-07
```

I'll check the conditions despite not having asked you to do that for this problem:

1. Is the sample representative of the population? I haven't said anything about how I took the sample, but I did take a random sample and there was no bias in the way I selected the babies for the sample - so the sample is representative of the population.
2. Are the gestation times of the babies in my sample independent? This condition is also satisfied, again because I took a random sample of babies from the population.
3. Are we thinking about a quantitative variable for each observational unit? Yes, we measured the gestation time (a number) for each baby.

4. Is it reasonable to summarize the center of that variable with the mean? Yes - we have seen that the distribution of gestation times is a little skewed, but not too badly. The mean will be representative of the center of this distribution.

The test statistic is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{38.83 - 40}{2.238574/\sqrt{100}} = -5.226542$$

The interpretation of this test statistic is that the sample mean for this sample is below the hypothesized mean of 40, by an amount about equal to 5.23 times our best estimate of the standard deviation of the sample mean.

The p-value is the probability of obtaining a test statistic at least as extreme as what we observed in this sample - in other words, the probability of obtaining a sample mean that is at least 5.22654 standard deviations away from the hypothesized mean.

To find this probability, we use the fact that if the null hypothesis is true and  $\mu = 40$ , then the test statistic  $t$  follows a  $t_{99}$  distribution. The degrees of freedom for the  $t$  distribution is the sample size ( $n = 100$ ) minus 1. We can calculate probabilities involving the  $t$  distribution with the `pt` function in R.

Specifically, from the output of the command `pt(-5.22654, df = 99)` in the R output above, we know that  $P(t < -5.22654) = 4.805429e - 07$ .

This is half of the p-value for the test. The full p-value is

$$\begin{aligned} P(t < -5.22654 \text{ or } t > 5.22654) &= P(t < -5.22654) + P(t > 5.22654) \\ &= 2 * P(t < -5.22654) \\ &= 2 * 4.805429e - 07 \\ &= 9.610858e - 07 \end{aligned}$$

**Example 6 (also see Lab 15 on Gryn; adapted from SDM4 20.44)**

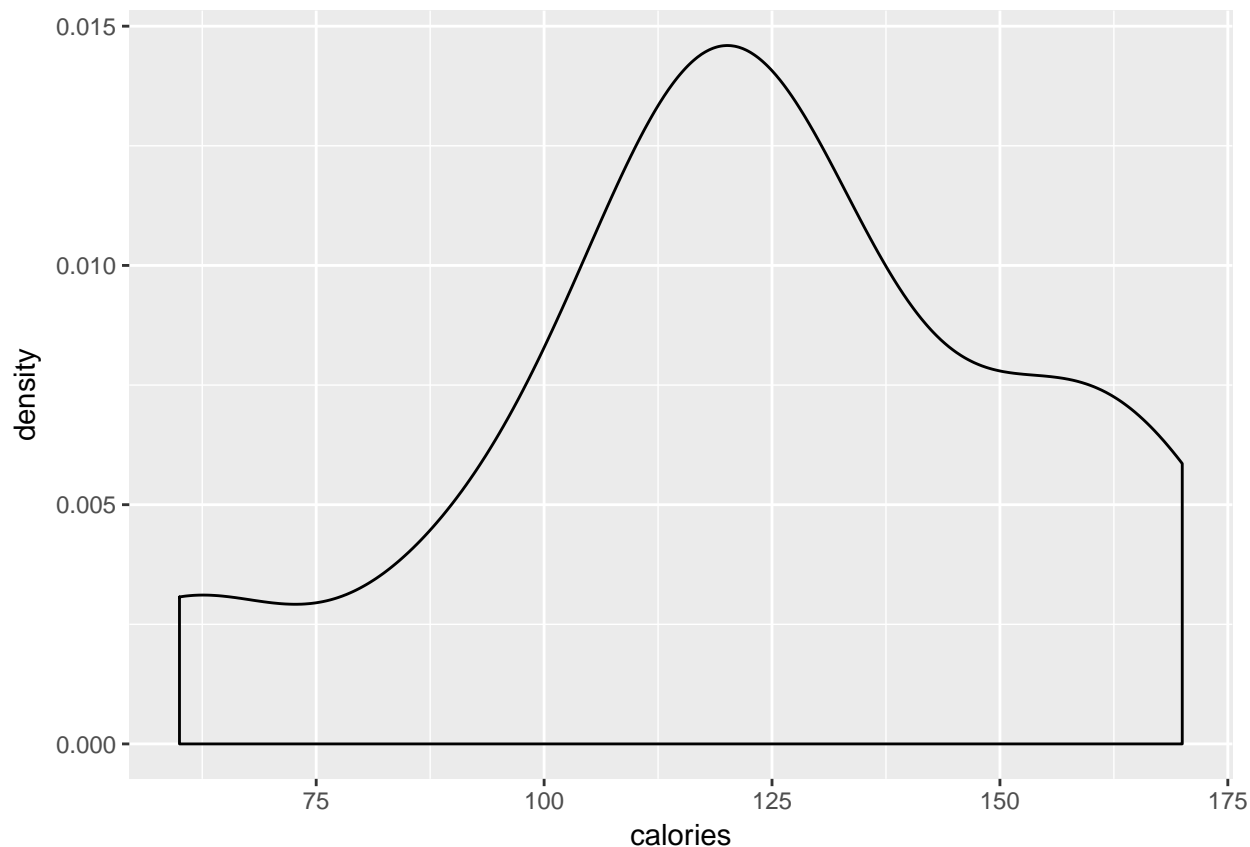
*Consumer Reports* tested 11 brands of vanilla yogurt and found the following numbers of calories per serving:

```
##      calories
## 1         130
## 2          60
## 3         150
## 4         120
## 5         120
## 6         110
## 7         170
## 8         160
## 9         110
## 10        130
## 11         90
```

**(a) Check the conditions for inference.**

1. Is the sample representative of the population? We don't know how Consumer Reports selected the yogurts for their sample. To claim that the sample was representative of the population, we would need to know that the yogurts in the sample were randomly selected and that there were no forms of bias in the way they chose these yogurts.
2. Are the yogurts in my sample independent? Again, we would need to know that the yogurts in the sample were randomly selected.
3. Are we thinking about a quantitative variable for each observational unit? Yes, we measured the number of calories per serving (a number) for each yogurt.
4. Is it reasonable to summarize the center of that variable with the mean? In order to check this condition, we need to make a plot of the data to verify that it is symmetric and there are no outliers.

```
ggplot(data = yogurt, mapping = aes(x = calories)) +  
  geom_density()
```



With a sample size of only 11, distributions rarely look perfectly symmetric. This distribution isn't too far off of being symmetric. There is one mode, and no major outliers. Overall, using the mean to summarize the center is reasonable here.

(b) A nutrition guide claims that you will get an average of 120 calories from a serving of vanilla yogurt. Conduct an appropriate hypothesis test and state your conclusion. As part of the hypothesis test, find the p-value for the test “manually”. When you're doing this, draw a picture of a  $t$  distribution, show where the test statistic is, and shade in the region corresponding to the p-value.

**Note:** Lab 15 on Gryd sets up data and a place where you can run the necessary R commands.

The sample mean and sample standard deviation are

```
yogurt %>%
  summarize(
    mean_calories = mean(calories),
    sd_calories = sd(calories)
  )
```

```
##   mean_calories sd_calories
## 1      122.7273    31.33398
```

For a hypothesis test of whether the mean is equal to 120, the hypotheses are:

Null hypothesis:  $\mu = 120$

Alternative hypothesis:  $\mu \neq 120$

The test statistic is:

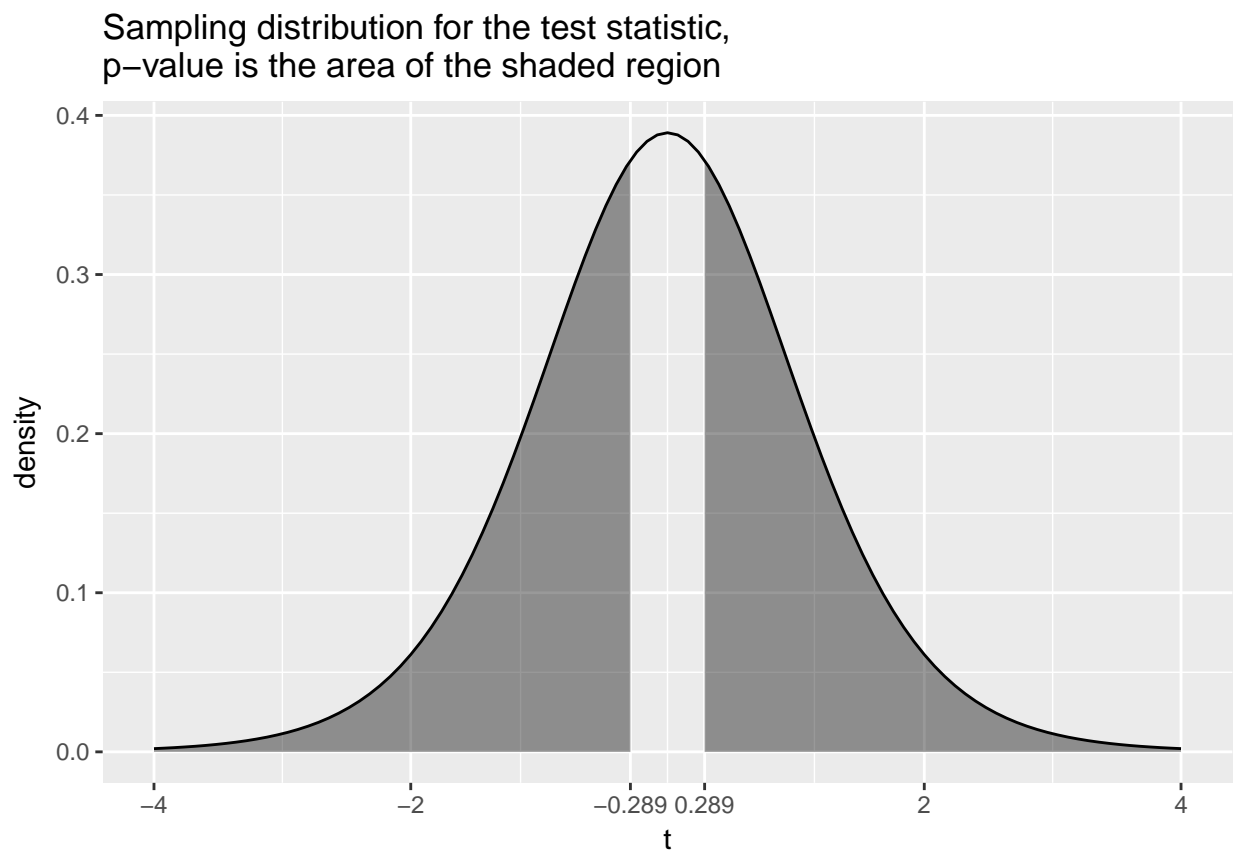
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{122.727 - 120}{31.334/\sqrt{11}} = 0.289$$

The interpretation of this test statistic is that the sample mean for this sample is above the hypothesized mean of 120, by an amount about equal to 0.289 times our best estimate of the standard deviation of the sample mean.

The p-value is the probability of obtaining a test statistic at least as extreme as what we observed in this sample - in other words, the probability of obtaining a sample mean that is at least 0.289 standard deviations away from the hypothesized mean.

To find this probability, we use the fact that if the null hypothesis is true and  $\mu = 120$ , then the test statistic  $t$  follows a  $t_{10}$  distribution. The degrees of freedom for the  $t$  distribution is the sample size ( $n = 11$ ) minus 1. We can calculate probabilities involving the  $t$  distribution with the `pt` function in R.

Here is a picture of the  $t_{10}$  distribution, with the area corresponding to the p-value shaded in:



Since the  $t$  distribution is symmetric, the p-value can be calculated in R as 2 times  $P(t < -0.289)$ :

```
2 * pt(-0.289, df = 10)
```

```
## [1] 0.7784839
```