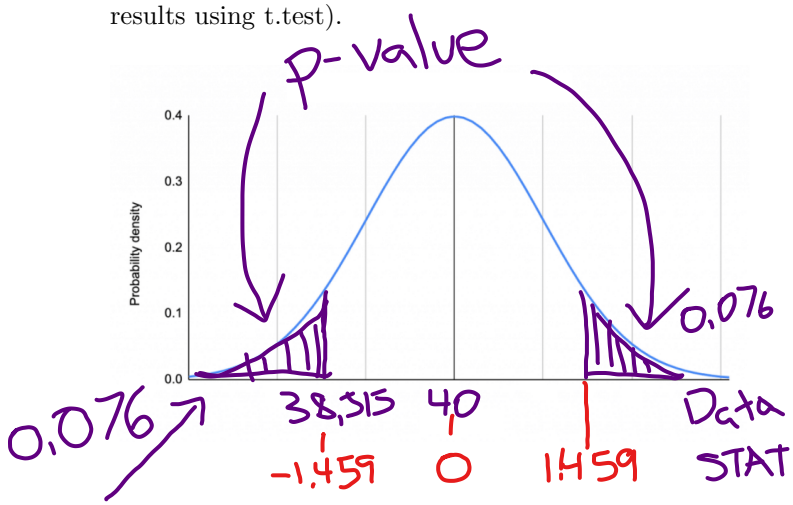
ACTIVITY 18

For each question where applicable 1) state your hypothesis, 2) state the calculated test statistic and/or p-value 3) shade the area corresponding to the p-value

Question 1

Toyota claims their fuel efficiency is superior and if you buy one of their cars your average weekly gas spending will be less than $40. You randomly survey 40 people with a Toyota to obtain the amount of money spent on gas weekly. The results are saved in 'data/gas_spending.csv' in RStudio Cloud. (Compute by hand but verify results using t.test).



$H_0: \mu = 40$

$H_A: \mu \neq 40$

$\bar{x} = 38.515 \quad SE = \frac{s}{\sqrt{n}} = 1.018$

$STAT = \frac{38.515 - 40}{1.018} = -1.459$

pt(q = -1.459, df = 39)

Assuming null is true in context of problem

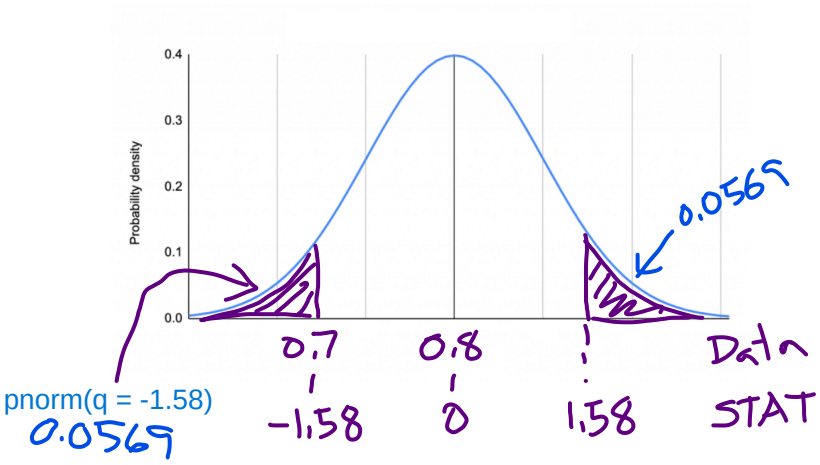
p-value = 0.076 + 0.076 = 0.152

Assuming the average weekly gas spending for a Toyota is equal to $40, there is a 15.2% chance of seeing data as extreme as our sample.

t.test(gas_spending$price, mu = 40)

Question 2

Trident Gum claims that 4 out of 5 dentists recommend sugarless gum for their patients who chew gum. You decide to test this popular claim and survey a random sample of 40 doctors and find that 28 recommended sugarless gum. (Compute by hand but verify results using prop.test)



$H_0: \pi = 0.8$

$H_A: \pi \neq 0.8$

sample

$\hat{\pi} = \frac{28}{40} = 0.7$

$STAT = \frac{0.7 - 0.8}{\sqrt{\frac{0.8(1-0.8)}{40}}} = -1.58$

π_0 because we assume null is true

pnorm(q = -1.58)

0.0569

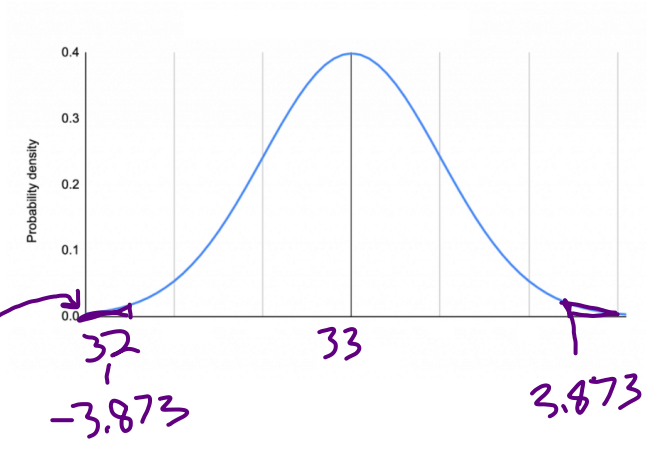
p-value = 0.0569 + 0.0569 = 0.1138

Assuming 4 out of 5 doctors recommend sugarless gum, there is a 11.38% chance of seeing data as extreme as our sample.

prop.test(x = 28, n = 40, p = 0.8, correct = FALSE)

Question 3

A top-tier University claims the average ACT score is 33. However, you suspect the university is inflating their score to make them look more appealing. You randomly survey 15 students at the University and calculate a mean ACT score of 32 with a standard deviation of 1. (No data so cannot use t.test to verify).



$H_0: \mu = 33$

$H_A: \mu \neq 33$

$\bar{x} = 32 \quad SE = \frac{1}{\sqrt{15}}$

$STAT = \frac{32-33}{1/\sqrt{15}} = -3.873$

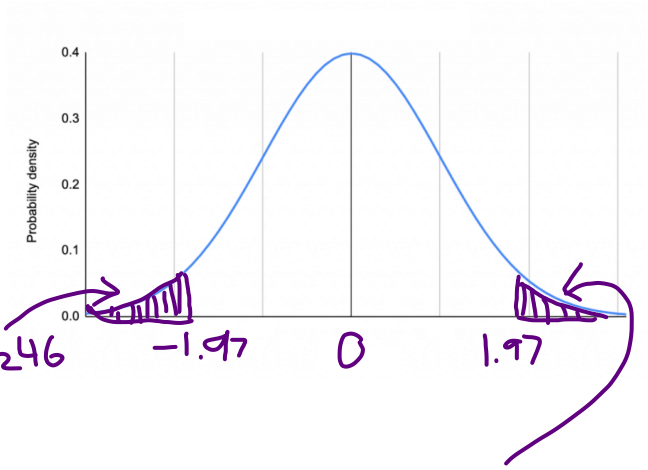
pt(-3.873, df = 14)
= 0.0008

p-value = 0.0008 + 0.0008 = 0.0016

Assuming the average ACT score at the top-tier University is equal to 33, there is a 0.16% chance of seeing data as extreme as our sample.

Question 4

You want to determine if alcohol preference changes as you age. You survey 120 people between the ages of 21-30 ('young') and ask if they prefer beer over wine. And you survey 120 people between the ages of 45-55 ('older') and ask if they prefer beer over wine. You save the results in the data frame 'data/alcohol_pref.csv'. An indicator of '1 = prefer beer' and '0 = prefer wine'. We want to know if there is a difference in alcohol preference between 'young' people and 'older' people.
Assume equal variance, compute by hand but verify results using prop.test.



$H_0: \pi_Y - \pi_{OLD} = 0$

$H_A: \pi_Y - \pi_{OLD} \neq 0$

$\hat{\pi}_Y = \frac{78}{120} = 0.65 \quad \hat{\pi}_{OLD} = \frac{63}{120} = 0.525$

pooled proportion: $\hat{\pi}_0 = \frac{78+63}{120+120} = 0.5875$

$STAT = \frac{(0.65-0.525)-0}{\sqrt{\frac{0.5875(1-0.5875)}{120} + \frac{0.5875(1-0.5875)}{120}}}$

$= 1.97$

pnorm(1.97, lower.tail = FALSE)
0.0246

p-value = 0.0246 + 0.0246 = 0.0492

Assuming there is no difference between the proportion of young people and proportion of older people that prefer beer over wine, there is a 4.92% chance of seeing data as extreme as our sample.

prop.test(x = c(78, 63), n = c(120, 120), correct = FALSE)