DATA 1220 Standard Test

Name:

"**LungCapData**" data describe the lung capacity (**LungCap**) of a sample of 725 US citizens. The data also include the following variables: **Age**, **Height**, **Smoke** (smoking status), **Gender**, and whether or not a person was born by **Caesarean**. The dataset can be downloaded from Canvas. Import the data into RStudio and complete the following questions. Use the level of significance α = 0.05.

2' (Drop 1' if they need help import data)

1. What are the values of quartiles of lung capacity in this sample?

Q1 = 6.15 Q2=8 Q3=9.8

3' (1' each; drop 1' if provide more than asked)

2. Construct a 95% confidence interval for smoking rate of all US citizens (keep 4 decimal places). And write one sentence to interpret the interval.

$\hat{p} = \frac{77}{725} = 0.1062$

3' for the interval (break down: 1' for phat, 1' for 1.96, and 1' for SE)

SE = 0.0114, then 0.1062 ± 1.96*SE = (0.0838, 0.1286)

From R: 95 percent confidence interval: (0.0851959, 0.1314957)

I'm 95% confident that smoking rate of all US citizens is from 8.38% to 12.86%.

3' for interpretation (break down: 1' for 95%, 1' for parameter, and 1' for the interval)

3. Suppose you want to investigate whether smokers have larger lung capacity than non-smokers, on average.
   (a) Which statistical methodology is appropriate?

two-sample t test

2' (1' if a relevant but wrong method)

   (b) Conduct the statistical methodology in part (a).

4' for the test (break down: 1' for Ho&Ha, 1' for T, 1' for p-val, and 1' for decision) Both df are right.

Ho: $\mu_{smoke} - \mu_{non} = 0$

Ha: $\mu_{smoke} - \mu_{non} > 0$

$\bar{x}_{smoke} - \bar{x}_{non} = 8.6455 - 7.7702 = 0.8753$, $S_{smoke} = 1.8829$, $S_{non} = 2.7261$. Then SE = 0.2398

T = 3.6499, df = 76, p-value = 0.00024

From R: t = 3.6498, df = 117.72, p-value = 0.0003927

2' if conduct a relevant but wrong method correctly

Reject Ho. There is enough evidence to show that smokers have larger lung capacity than non-smokers, on average.

   (c) A researcher concludes that smoking causes larger lung capacity. Do you agree with this conclusion or not? Please briefly explain your rationale.

No. Association doesn't mean causation.

2' (break down: 1' for "No" and 1' for rationale)