STAT 614 - HW 4

The food-frequency questionnaire (FFQ) is an instrument often used in dietary epidemiology to assess consumption of specific foods. A person is asked to write down the number of servings per day typically eaten in the past year of over 100 individual food items. A food-composition table is then used to compute nutrient intakes (protein, fat, etc.) based on aggregating responses for individual foods. The FFQ is inexpensive to administer but is considered less accurate than the diet record (DR) (the gold standard of dietary epidemiology). For the DR, a participant writes down the amount of each specific food eaten over the past week in a food diary and a nutritionist, using a special computer program, computes nutrient intakes from the food diaries. This is a much more expensive method of dietary recording. To validate the FFQ, 173 nurses participating in the Nurses' Health Study completed 4 weeks of diet recording about equally spaced over a 12-month period and an FFQ at the end of diet recording. Data are in Blackboard in the file valid.txt.

Consider the data on total alcohol consumption for both the DR and FFQ, alco_dr and alco_ffq, respectively. You are to assess whether the two methods, diet record and the food-frequency questionnaire, are comparable for total alcohol consumption. In particular, is there evidence that FFQ *underestimates* total alcohol consumption, in general? Estimate by how much the FFQ generally underestimates total alcohol consumption.

1. Explain why the initial model needed to address these research goals is a matched-pairs t-procedure.

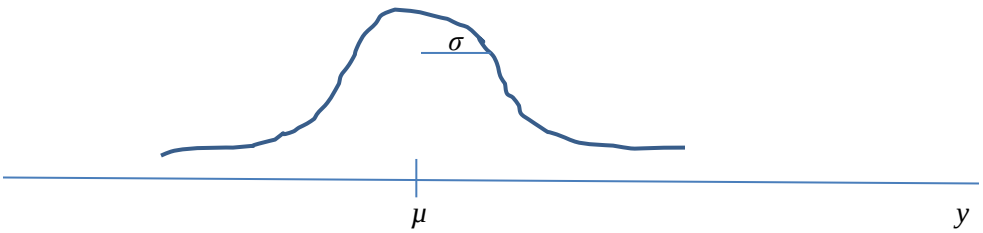
Two measurements are taken on each of the 173 nurses, alco_dr and alco_ffq (alcohol consumption based on dietary record and alcohol consumption based on the food frequency questionnaire). These are both taken on each individual, so these observations are *paired* (i.e. dr is not independent of ffq). Thus, the initial method to be used is the matched-pairs t-procedure to compare the difference in mean alcohol consumption as measured by DR and by FFQ.

2. Use both the model notation we developed in class and a brief written description of the model (you may also use pictures) to illustrate the model. (Be careful! The matched-pairs procedure works on the *difference* in the two measures on each individual. Start with y = alco_dr – alco_ffq and describe the model for y!)

The initial method to be used is the matched-pairs t-procedure to compare the difference in mean alcohol consumption as measured by DR and by FFQ. The population MODEL for this method is that the population distribution of the difference of alcohol consumption as measured by the FFQ versus as measured by the DR follows a normal distribution with some mean, μ, and variance, σ^2.

Using model notation, let y_i = the difference in alcohol consumption measures (DR – FFQ) for person i = 0,…, 173. Then, $y_i = \mu + \varepsilon_i$ where $\varepsilon_i \sim N(0, \sigma^2)$

This can also be written as $y_i \sim N(\mu, \sigma^2)$. And here is my attempt at a Word art Normal density curve!

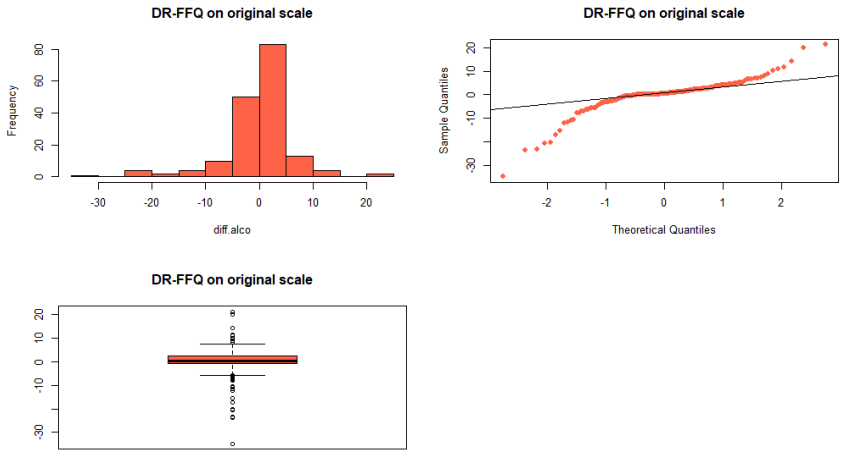


3. What are the model assumptions?

The assumptions of the model are (1) individual measures are independent and (2) the population distribution of the difference of alcohol consumption as measured by the FFQ versus as measured by the DR follows a normal distribution. We also need to assess (3) that there are no outliers influencing our conclusions.

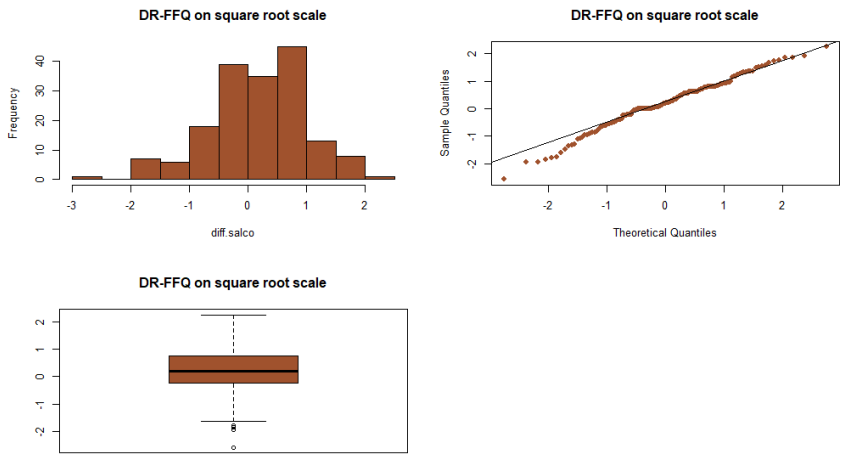
4. Which of the model assumptions are not met? Give and refer to specific output.

Graphical summaries for the difference in alcohol consumption as measured by DR minus FFQ are given on the next page. From the QQ Plot we see there are major deviations from normality evidenced in both tails of the distribution. The boxplot highlights *many* outliers in both tails. While the sample size of n = 173 makes me less concerned about the nonnormal the large number of outliers is an issue. We will consider both a transformation and a nonparametric method to address these issues.



5. Consider a square root transformation of the alcohol data: salcoDR = $\sqrt{\text{alco\_dr}}$ and salcoFFQ = $\sqrt{\text{alco\_ffq}}$. Are the model assumptions met for the transformed data? Give and refer to specific output.

An examination of the difference in the square root of alcohol consumption (DR-FFQ) suggests that there are far fewer issues with the assumptions. If y_i = the difference in the square root of alcohol consumption measures ($\sqrt{DR} - \sqrt{FFQ}$) for person i = 0,..,173 then the model claims that this measure follows a normal distribution. Examination of the histogram, QQ Plot, and boxplot suggest this is reasonably met with only a single outlier in the lower tail.



6. Conduct the appropriate test on the square root transformed data and interpret the results. Be sure to address the research questions stated above.

Based on the study goals, we are interested in testing Ho: μ = 0 vs. Ha: μ > 0 where μ is the population mean of the difference in the square root of alcohol consumption between DR and FFQ (awkward, I know.) That is, the population mean of $\sqrt{DR} - \sqrt{FFQ}$. We would also like to estimate this. (Note: a two-sided can also be used with the confidence interval providing evidence about the direction of the effect.)

The one-sided test has a p-value of 0.00218 (two-sided is 0.00436). Thus, there is evidence that the **average** alcohol consumption, on the square root scale, is reported higher on the DR than using the FFQ (also on the square root scale). We estimate, with 95% confidence, that the mean square of DR is 0.06 to 0.31 more than the mean square root of FFQ. (Note: To me, this is an awkward scale. We were told the units for which alcohol consumption was measured. Likely ounces, grams, or even drinks so assuming that DR and FFQ are measured as grams of alcohol consumed, were are say that the mean square root of alcohol consumption as measure by dietary record is 0.06 square root of grams to 0.31 square root of grams more, on average, than the mean square root alcohol consumption using FFQ.)

7. Consider a nonparametric method for addressing the research questions. What null and alternative hypotheses are addressed by the appropriate nonparametric method? Carry out and interpret the results of the nonparametric method. Include and interpret the confidence interval estimate.

I will use the signed-rank test as this is a matched pairs design. This tests that the distribution is symmetric around zero, so that the negative differences distribution (when DR < FFQ) is the same as the positive differences' distribution (when DR > FFQ). The alternative hypothesis of interest is that there is a shift in the positive direction, so that DR > FFQ is predominant. The sample sizes is large and there are only 16 observations where DR = FFQ so I will use the exact=F option in the test. I will also use the data on the original scale (not transformed) as that first histogram and boxplot suggest the distribution of the difference in alcohol consumption is reasonably symmetric (just "heavy tailed" with lots of outliers).

The one-sided p-value for a test of the distribution of DR being shifted higher than the distribution of FFQ alcohol consumptions measures is p = 0.013 (two-sided p-value is 0.026). Thus, there is evidence that the distribution of DR measured alcohol consumption is above the distribution of FFQ measured alcohol consumption suggesting. In other words, in general, the FFQ underestimates alcohol consumption.

A 95% confidence interval estimates that the distribution of alcohol consumption as measured by DR is 0.10 to 1.32 units (drinks? Ounces? Gram? we weren't given the units) above the distribution of consumption as measured by FFQ. R also gives the estimated median for the difference as 0.735 units so our "best" estimate is that the median alcohol consumption using DR is 0.735 above the median using FFQ (0.10 to 1.32 with 95% confidence).

8. Which of the results in (6) or (7) do you prefer to use to draw conclusions for this study and why?

Both methods give significant results and suggest FFQ underestimates DR, in general. Both methods give confidence intervals that are a little awkward to interpret as they aren't directly about the mean (of the difference in amount of alcohol consumed). I prefer the nonparametric method in this case only because the confidence interval does translate to the original scale (which I presume is drinks, but it might be ounces or grams or some other unit – we weren't provided this information). It is not so straightforward with the square root transformation to interpret the confidence interval on the original scale. (But that is ultimately fine, if alcohol consumption is underestimated on the square root scale, it is still underestimated!) **You may have other reasons to prefer one method over the other!**

(Note: What would happen if we didn't check the assumptions and used the t-procedure on the data as is? We would get a p-value of 0.49 and fail to reject the null hypothesis!)