Lab 1 Part ?

Part 1: Foundational Exercises

1.1 Professional Magic

1. What is the type 1 error rate of your test?

A type I error is a false positive and occurs when a true null hypothesis is rejected incorrectly.

In this case, the type I error rate can be calculated by the probability of the test statistic(u) equals to 0 or 6 while null hypothesis is true($p = \frac{1}{2}$).

	$X_i = 0$	$X_i = 1$
$Y_i = 0$	1/4	1/4
$Y_i = 1$	1/4	1/4

since $u = X_1 + Y_1 + X_2 + Y_2 + X_3 + Y_3$, we can get $u = 0$ only when all 6 variables are 0, and $u = 6$ only when all 6 variables are 1. Then

$$P(u = 0) = P(X_1 = Y_1 = 0) \cdot P(X_2 = Y_2 = 0) \cdot (X_3 = Y_3 = 0) = (\frac{1}{4})^3 = \frac{1}{64}$$

$$P(u = 6) = P(X_1 = Y_1 = 1) \cdot P(X_2 = Y_2 = 1) \cdot (X_3 = Y_3 = 1) = (\frac{1}{4})^3 = \frac{1}{64}$$

So we can get the type I error rate α:

$$\alpha = P(u = 0) + P(u = 6) = \frac{1}{64} \cdot 2 = \frac{1}{32}$$

2.What is the power of your test for the alternate hypothesis that $p = 3/4$?

For the alternate test that $p = 3/4$, we have

	$X_i = 0$	$X_i = 1$
$Y_i = 0$	3/8	1/8
$Y_i = 1$	1/8	3/8

So we can get the power of the alternate test by calculating the probability of the test statistic(u) equals to 0 or 6.

$$P(u = 0) = P(X_1 = Y_1 = 0) \cdot P(X_2 = Y_2 = 0) \cdot (X_3 = Y_3 = 0) = (\frac{3}{8})^3 = \frac{27}{64}$$

$$P(u = 6) = P(X_1 = Y_1 = 1) \cdot P(X_2 = Y_2 = 1) \cdot (X_3 = Y_3 = 1) = (\frac{3}{8})^3 = \frac{27}{64}$$

So we can get the power of the alternate test $1 - \beta$:

$$1 - \beta = P(u = 0) + P(u = 6) = \frac{27}{64} \cdot 2 = \frac{27}{32}$$

1.3 Test Assumptions

1.3.4 Attitudes toward the religious

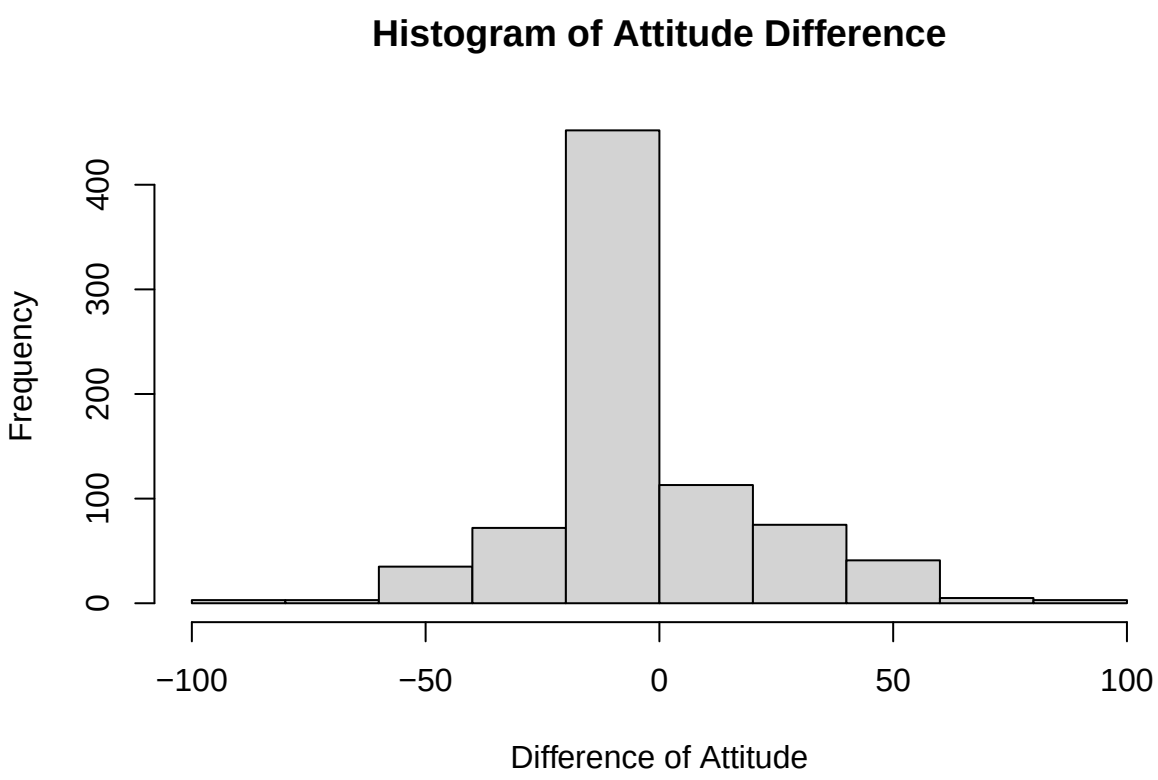
Assumptions for paired t-test:

1. A and B have a metric scale with the same units.
2. There is a natural pairing between observations for A and for B.
3. Each pair (A_i, B_i) is drawn i.i.d.
4. The distribution of $A - B$ is sufficiently normal given the sample size.

Evaluation:

1. Do attitudes toward the religious have a metric scale with the same units?
   No.
   From the description of the GSS questions, it shows that each respondent has their own scales of the "feeling thermometer". The score cannot be summarized statistically in a meaningful way.
2. Is there a natural pairing between observations for attitudes toward the two religious?
   Yes. The principles of their faith are different. There are painful divisions between Protestants and Catholics. For one who has a positive attitude toward Catholics is more likely to feel negative abouth Protestants.
3. Is person drawn i.i.d ?
   From the official info, the GSS sample is drawn using an area probability design that randomly selects respondents in households across the United States to take part in the survey. Respondents that become part of the GSS sample are from a mix of urban, suburban, and rural geographic areas. Participation in the study is strictly voluntary.
   Therefore, we could say that i.i.d is not strongly violated for the GSS sampling.
4. Is the distribution of attitude difference sufficiently normal given the sample size?

```
df <- read.csv("GSS_religion.csv")
dff <- mutate(df, temp_diff = prottemp - cathtemp)
hist(dff$temp_diff,
     main = "Histogram of Attitude Difference",
     xlab = "Difference of Attitude"
     )
```



From the histogram of the "Difference of Attitude",we can see a rough normal distribution.

In summary, the samples in GSS are drawn without violation of i.i.d. The observation of each have a natural paring attribute and a normal distribution of the difference of the paired observation. However, the observations are nonparametric which is not appropriate for the application of t test.