STAT 641
Homework 7

Keegan Smith

July 29, 2025

Problem 1

1. independent samples
2. matched pairs
3. matched pairs
4. independent samples
5. matched pairs
6. independent samples

Problem 2

First we will apply the two sample t-test to this problem:
we have $n = m = 15$. The average before is 120.8667 and the average after is 110.8. let $\bar{X}$ = sample mean before and $\bar{Y}$ = sample mean after. Let $\hat{\theta} = \bar{Y} - \bar{X}$ and $\theta = \mu_Y - \mu_X$.
Our null hypothesis is $\theta \geq 0$ and our alternative hypothesis is $\theta < 0$.
The pooled standard deviation is:

$$\begin{aligned}
s_p &= \sqrt{\frac{(n-1)s_1^2 + (m-1)s_2^2}{n+m-2}} \\
&= \sqrt{\frac{(14)4.240395^2 + (14)5.427443^2}{28}} \\
&= 4.8702200809755
\end{aligned}$$

We then have our test statistic:

$$T = \frac{th\hat{e}ta - \theta_0}{\sqrt{\frac{s^2}{n} + \frac{s^2}{m}}} \\ = \frac{-10.067 - 0}{\sqrt{\frac{4.87^2}{15} + \frac{4.87^2}{15}}} \\ = -5.6611$$

our degrees of freedom is $15 + 15 - 2 = 28$, so from the table we get that the:

$$P(T < T_o bs) = tcdf(-5.6611, 28) \approx 2.29 \cdot 10^{-6}$$

this p value is much less than $\alpha = .05$, so we reject the null hypothesis.
Using the wilcoxon rank sum test, we will let $W_1 =$ sum of before ranks and $W_2 =$ sum of after ranks.
We get $W_1 = 516$ and $W_2 = 216$. So our p value is:

$$p = G(W_1) = pwilcox(w_1 - n(n+1)/2, n, m) = 0.004937411$$

Again the p-value is much less than $\alpha = .05$ so we reject the null hypothesis.
For a paired t-test to be used, we must have:

1. D is normally distributed where $D_i = X_i - Y_i$: using the shapiro wilkes test we get a p value of 0.5105, suggesting that indeed D is normally distributed.

2. D_is are independent: this depends on how exactly the study was carried out, but it would seem that one patient's blood sugar results would not impact another patient's blood sugar results, so I think it is safe to assume that this condition holds.

given the above, it would seem that the paired t-test is suitable for this problem.

Problem 3

1. We have the null hypothesis:

$$H_o : p_1 - p_2 = 0$$

where p_1 is the proportion of population 1 which has the trait, and p_2 is the proportion of population 2 which has the trait. We have the alternative hypothesis:

$$H_A : p_1 - p_2 \neq 0$$

2. We have the test statistic:

$$Z = \frac{\hat{p_1} - \hat{p_2}}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

where:

$$\hat{p} = \frac{n_1 \cdot \hat{p_1} + n_2 \cdot \hat{p_2}}{n_1 + n_2} = \frac{100 \cdot .4 + 100 \cdot .3}{200} = .35$$