# Problem Set 5: POLS 3316 Statistics for Political Science University of Houston

Fall 2023

## Due November 15, 2023, 11:59 PM

.

**For** **both** **problems,** **be** **sure** **to** **show** **work** **as** **instructed** **and** **to** **fill** **in** **the** **score,** **p-value,** **and** **whether** **you** **retain** **or** **reject** **the** **null** **hypothesis.** **Answer** **any** **specific** **questions.**

**1.** **Complete** **a** *X* 2 **(Chi-square)** **hypothesis** **test** **of** **independence** **on** **the** **following** **data.** The extra tables are for your convenience, but you must set up the initial table, figure the totals and sample sizes, and expected values. You can get the p-value using the table or the R function for chi-square probability distribution:

**Treated** **patients,** **no** **improvement:** 50

**Treated** **patients,** **with** **improvement:** 75

**Untreated** **patients,** **no** **improvement:** 60

**Untreated** **patients,** **with** **improvement:** 75

**H0:** The variables are independent (treatment is unrelated to improvement.)

**H1:** The variables are dependent (treatment affects improvement.)

Table 1: The data

|  |  |  |
| --- | --- | --- |
|  | TREATED | UNTREATED |
| IMPROVEMENT | 75 | 75 |
| NO IMPROVEMENT | 50 | 60 |

Table 2: with Marginal Frequencies (totals)

|  |  |  |  |
| --- | --- | --- | --- |
|  | TREATED | UNTREATED | TOTAL |
| IMPROVEMENT | 75 | 75 | 150 |
| NO IMPROVEMENT | 50 | 60 | 110 |
| TOTAL | 125 | 135 | 260 |

Table 3: expected outcomes

|  |  |  |  |
| --- | --- | --- | --- |
|  | TREATED | UNTREATED | TOTAL |
| IMPROVEMENT |  |  | 150 |
| NO IMPROVEMENT |  |  | 110 |
| TOTAL | 125 | 135 | 260 |

Extra tables for work:

|  |  |  |  |
| --- | --- | --- | --- |
|  | TREATED | UNTREATED | TOTAL |
| IMPROVEMENT |  |  |  |
| NO IMPROVEMENT |  |  |  |
| TOTAL |  |  | 0.5252525 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | TREATED | UNTREATED | TOTAL |
| IMPROVEMENT | 75 | 75 | 150 |
| NO IMPROVEMENT | 50 | 60 | 110 |
| TOTAL | 125 | 135 | 260 |

Chi-Square Score *X*2: 0.5252525

p-value: 0.4686097

Retain or reject the null: reject the retain null hypothesis – treatment is unrelated to improvement

**2. Complete a paired samples T-test on the following sample.**

**Show your work**: you may either copy and paste all your R code into a document or show the math steps.

**H0:** The true mean difference between the paired samples is zero.

**H1:** The true mean difference between the paired samples is not equal to zero. **Is this a one or two tailed test?** Two Tailed test

|  |  |  |  |
| --- | --- | --- | --- |
| Student | Score 1 | Score 2 | Differences |
| 1 | 3 | 15 | -12 |
| 2 | 2 | 13 | -11 |
| 3 | 3 | 14 | -11 |
| 4 | 12 | 21 | -9 |
| 5 | 15 | 30 | -15 |
| 6 | 16 | 31 | -15 |
| 7 | 17 | 23 | -6 |
| 8 | 19 | 21 | -2 |
| 9 | 23 | 25 | -2 |

t-score:

p-value: 0.0005127166

Retain or reject the null: reject the null hypothesis for

> diff <- c(-12,-11,-11,-9,-15,-15,-6,-2,-2)

> mean(diff)

[1] -9.222222

> sd(diff)

[1] 4.944132

> mean(diff)/(sd(diff)/3)

[1] -5.595859

> 2\*pt(q=5.595859, df=8, lower.tail=FALSE)

[1] 0.0005127166

> score1 <- c(3,2,3,12,15,16,17,19,23)

> score2 <- c(15,13,14,21,30,31,23,21,25)

> t.test(score1, score2, paired = TRUE, alternative = "two.sided")

Paired t-test

data: score1 and score2

t = -5.5959, df = 8, p-value = 0.0005127

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:

-13.022619 -5.421826

sample estimates:

mean difference

-9.222222