# Course Evaluations [10 points]

## Most preferences are ordinal data which should not be analyzed with standard parametric statistics. The attached file evaluations.csv contains survey results for two courses:

## subject: Course Subject (Calculus or Thermodynamics)

## instr\_eval: Instructor Evaluation

## course\_eval: Course Evaluation

## Instructor and course evaluations use the ordinal scale: Poor, Fair, Good, Excellent.

## Using the complete data set for both subjects, compute the Spearman’s Rank Correlation Coefficient between Instructor Evaluation and Course Evaluation.

Solution:

Spearman’s Rank = 0.5392

p-value = 0.0003

Small positive correlation between Instructor Evaluation and Course Evaluation

## Perform a Mann-Whitney *U* Test2 to evaluate whether the two subjects produce equivalent instructor evaluations and report:

## *U* statistic

Solution:

130

## *p*-value

Solution:

0.0095

Median 1 ≠ Median 2. Cannot accept H0.

## Assuming the samples are from related populations (i.e. ith X response and ith Y response are from the same student), perform a Wilcoxon Matched Pairs Test3 to evaluate whether the two subjects produce equivalent instructor evaluations and report:

## *T* statistic

Solution:

15

## *p*-value

Solution:

0.1655

Cannot accept H0. Median rank difference ≠ 0.

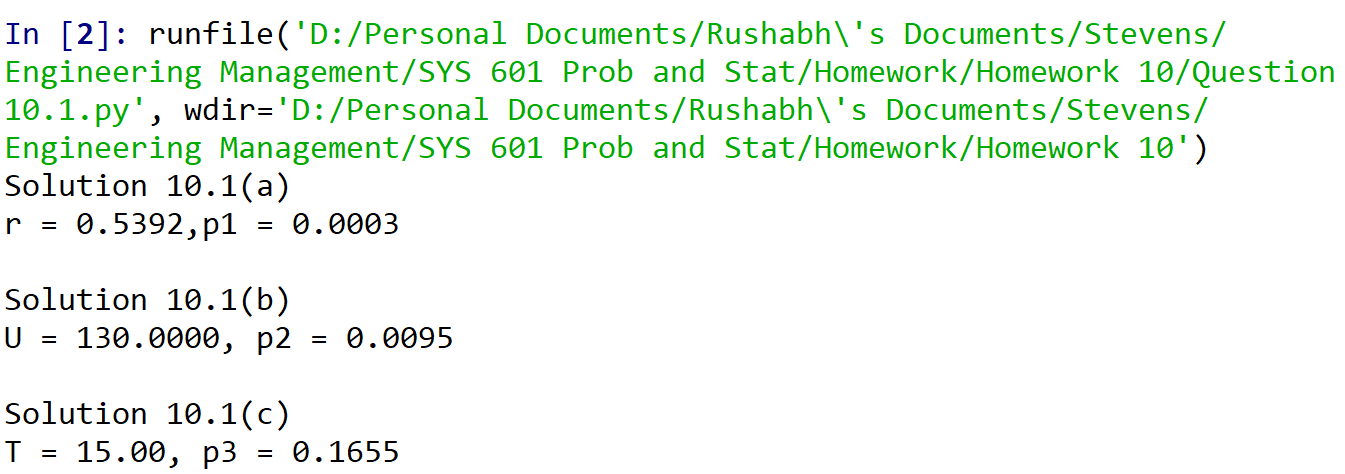


Figure 1 Screenshot from spyder

# Revisiting Admissions Data

## Recall the following dataset with 4486 college admission decisions in a frequency table with program, gender, and decision fields. Assume each person only applies to one program.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Program | Male (M) | | Female (F) | |
| Accepted (A) | Denied (D) | Accepted (A) | Denied (D) |
| P1 | 512 | 313 | 89 | 19 |
| P2 | 313 | 207 | 17 | 8 |
| P3 | 120 | 205 | 202 | 391 |
| P4 | 138 | 279 | 131 | 244 |
| P5 | 53 | 138 | 94 | 299 |
| P6 | 22 | 351 | 24 | 317 |

## (a–f) For each of the six programs P1–P6, perform a chi-squared test of independence for gender and acceptance decision and report:

Solution:

|  |  |  |  |
| --- | --- | --- | --- |
| Program | Chi2 | p-value | Comment |
| P1 | 16.37 | 0.000052 | Independent |
| P2 | 0.33 | 0.57 | Independent |
| P3 | 0.63 | 0.43 | Independent |
| P4 | 0.22 | 0.64 | Independent |
| P5 | 0.81 | 0.37 | Independent |
| P6 | 0.22 | 0.64 | Independent |

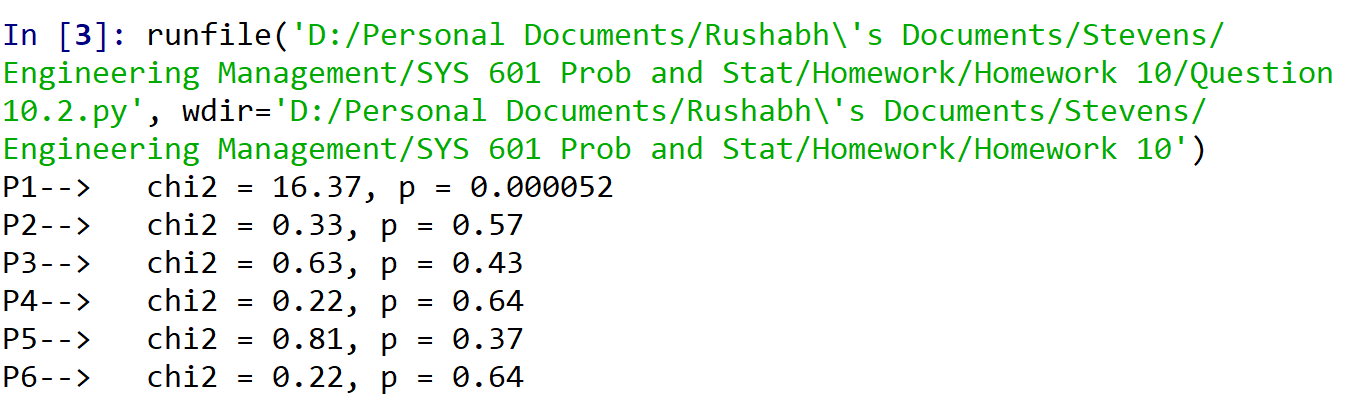


Figure 2 Output of 10.2

# Revisiting Café Java

## The following dataset collected by the manager at Café Java records 36 observations for the number of customers arriving the store in five minute intervals:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week 1 | 3 | 6 | 4 | 6 | 2 | 3 | 1 | 5 | 1 | 0 | 3 | 3 |
| Week 2 | 1 | 2 | 4 | 0 | 2 | 6 | 5 | 4 | 2 | 5 | 3 | 4 |
| Week 3 | 5 | 3 | 5 | 3 | 5 | 4 | 7 | 3 | 4 | 8 | 1 | 3 |

## Plot a histogram of the data, grouping the observations into bins such that no bin is empty and at least 80% of bins have 5 or more observations.

Solution:

## Perform a chi-square goodness-of-fit test to determine whether the data come from a Poisson distribution with rate *λ* (estimated from the data) and report:

## *χ*2 test statistic

Solutoin:

2.8317

## Number of degrees of freedom (note: *c* = 1)

Solution:

k = n-1-c

= 4

## *p*-value

Solution:

0.5864

Cannot reject H0. The data is consistent with distribution.