10.1 Course Evaluations [10 points]

Most preferences are ordinal data which should not be analyzed with standard parametric statistics. The attached file evaluations.csvcontains 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.

(a) 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

(b) Perform a Mann-Whitney *U* Test[2] to evaluate whether the two subjects produce equivalent instructor evaluations and report:

(i) *U* statistic

Solution:

130

(ii) *p*-value

Solution:

0.0095
Median 1 ≠ Median 2. Cannot accept H_0.

(c) Assuming the samples are from related populations (i.e. i^{th} X response and i^{th} Y response are from the same student), perform a Wilcoxon Matched Pairs Test[3] to evaluate whether the two subjects produce equivalent instructor evaluations and report:

(i) *T* statistic

Solution:

15

(ii) *p*-value

Solution:

0.1655
Cannot accept H_0. Median rank difference ≠ 0.

```
In [2]: runfile('D:/Personal Documents/Rushabh\'s Documents/Stevens/
Engineering Management/SYS 601 Prob and Stat/Homework/Homework 10/Question
10.1.py', wdir='D:/Personal Documents/Rushabh\'s Documents/Stevens/
Engineering Management/SYS 601 Prob and Stat/Homework/Homework 10')
Solution 10.1(a)
r = 0.5392,p1 = 0.0003

Solution 10.1(b)
U = 130.0000, p2 = 0.0095

Solution 10.1(c)
T = 15.00, p3 = 0.1655
```

Figure 1 Screenshot from spyder

10.2 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

```
In [3]: runfile('D:/Personal Documents/Rushabh\'s Documents/Stevens/
Engineering Management/SYS 601 Prob and Stat/Homework/Homework 10/Question
10.2.py', wdir='D:/Personal Documents/Rushabh\'s Documents/Stevens/
Engineering Management/SYS 601 Prob and Stat/Homework/Homework 10')
P1-->   chi2 = 16.37, p = 0.000052
P2-->   chi2 = 0.33, p = 0.57
P3-->   chi2 = 0.63, p = 0.43
P4-->   chi2 = 0.22, p = 0.64
P5-->   chi2 = 0.81, p = 0.37
P6-->   chi2 = 0.22, p = 0.64
```

Figure 2 Output of 10.2

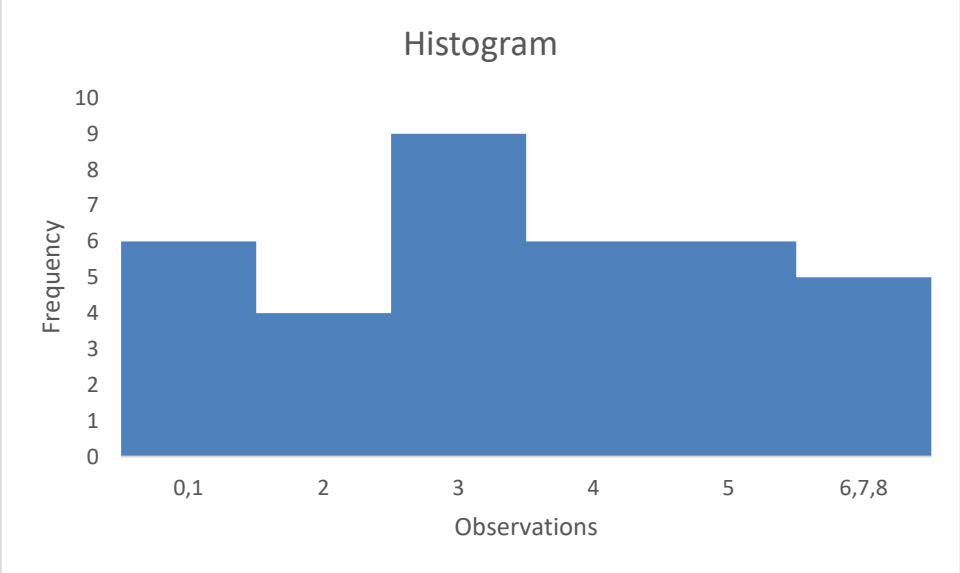
10.3 Revisiting Café Java

The following dataset collected by the manager at CaféJava 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

(a) 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:



(b) 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:

(i) χ^2 test statistic

Solutoin:

2.8317

(ii) Number of degrees of freedom (note: $c = 1$)

Solution:

k = n-1-c

= 4

(iii) p-value

Solution:

0.5864

Cannot reject H_0. The data is consistent with distribution.