

STEVENS INSTITUTE OF TECHNOLOGY

SYS-601 Homework #10

Due Apr. 23 2018

Submit the following using the online submission system: 1) Completed assignment cover sheet, 2) Written responses in PDF format, 3) All saved models (e.g. .xlsx or .py files).

10.1 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.

- 2 PT Using the complete data set for both subjects, compute the Spearman's Rank Correlation Coefficient¹ between Instructor Evaluation and Course Evaluation.
- 4 PTS Perform a Mann-Whitney U Test² to evaluate whether the two subjects produce equivalent instructor evaluations and report:
 - U statistic
 - p -value
- 4 PTS 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 Test³ to evaluate whether the two subjects produce equivalent instructor evaluations and report:
 - T statistic
 - p -value

Hint: When assigning ranks, remember that any samples with equivalent (tied) values receive a rank equal to the average of the spanned ranks. For example, if the first 10 samples all have the lowest values, they all are assigned rank $(1 + 10)/2 = 5.5$.

¹Consider using the `scipy.stats` function `spearmanr` to quickly solve this problem.

²Consider using the `scipy.stats` function `mannwhitneyu` to quickly solve this problem.

³Consider using the `scipy.stats` function `wilcoxon` to quickly solve this problem.

10.2 Revisiting Admissions Data [6 points]

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

(i) χ^2 test statistic

(ii) p -value

10.3 Revisiting Café Java [4 points]

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

(a) 1 PT 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.

(b) 3 PTS 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

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

(iii) p -value