

Probability and Statistics Practical List

Sr.	Practical	Hours
1	Basics of Python and Numpy/R/Matlab	2
2	matplotlib, seaborn, scipy and statsmodels/ Library for Probability and Statistics in R/Matlab	4
3	Write a program which scans value of k where k indicates number of mutually exclusive and exhaustive events (E_1, E_2, \dots, E_k). Assume any another event B . Implement Bayes' theorem assuming these k events and event B to estimate the probability of each of the k events assuming that the event B has already occurred (i.e. $P(E_1 B), P(E_2 B), \dots, P(E_k B)$). Your program should also scan (i) probability of each of the k (or $k-1$) events and (ii) probability of occurrence of event B assuming that each of the k events has occurred (i.e. $P(B E_1), P(B E_2), \dots, P(B E_k)$).	2
4	Write a program which reads m (scan m from the user) real values in 2 variables ($m/2$ in each) and computes and displays their (i) mean (ii) variance (iii) standard deviation. The program should also output z-score for each value. The program should furthermore display (i) scatter plot (ii) histograms for these variables for n bins where n is read from the user and (iii) density plot for these variables assuming Gaussian as the probability density function.	2
5	<p>(a) Write a program which scans marks of n (scan n from the user) students in P & S. Assuming that the students are to be graded on 5 grade-scale (A, B, C, D, and F), display marks and grade of each student. Fit Gaussian distribution to the data. Repeat the exercise if students are to be graded on 7 grade-scale (A, B, C+, C, C-, D, F). Also display bar chart for both the cases.</p> <p>(b) Change the above program further in 2 different ways: (1) Scan number of grades and z-score range for each of the grades to be used to grade students (2) Scan number of grades and number of students instructor wants to put in each grade.</p> <p>In (b), apart from the modifications suggested, other requirements and instructions are same as in (a).</p>	2
6	Assume that you are a professional fighter and in the past, you had fights with Undertaker. There were total of m (scan m from the user) fights between you and Undertaker. Assume that you set up a fresh fight tournament of n (scan n from the user) fights between you and Undertaker. Write a program which scans result of m earlier fights and computes the probability of (i) you winning x number of fights (ii) you winning more than y number of fights in the new tournament. Scan x and y from the user. Assume that fights are independent and each fight can either result in you or him winning. Assume that probability of you winning in each fight remains constant.	2
7	Scan an integer n from the user. Scan n sentences (no special character, punctuation, all lower case letters). Each sentence's sentiment is either <i>positive</i> or <i>negative</i> . Scan 2 different words y and z from the user. y and z should be words from the set of distinct words constructed from n	2

	sentences. Write a program to estimate (i) the probability of y being present in the positive sentences (ii) the joint probability of y and z being present in positive sentences assuming that the presence of y and z in any sentence are independent events. What would be this probability if these events are not independent?	
8	Scan an integer n from the user. Scan n sentences (no special character, punctuation, all lower case letters). Each sentence's sentiment is either <i>positive</i> or <i>negative</i> . Scan 5 words- a, b, c, d and e from the user. a, b, c, d , and e should be words from the set of distinct words constructed from n sentences. a, b, c, d , and e need not to be distinct words. Write a program to estimate (i) the joint probability of sampling a, b, c, d , and e from positive sentences. Assume that events of sampling a word are independent and the probability of sampling any word remains constant over different trials of sampling events.	2
9	Write a program to implement simple and multiple linear regression. Use the model parameters to make prediction for new data. Evaluate the model on standard evaluation measures. Your program should compute and display important model statistics (all the ones which are reported by statsmodels) and critically comment on these statistics.	8
10	Write a program to implement (i) Pearson correlation and display heatmap (ii) chi-square test	4