

CM2104: Computational Mathematics Laboratory Worksheet (Week 10 and 11)

Prof. D. Marshall

Aims and Objectives

After working through this worksheet you should be familiar with:

- Basic Discrete Probability Theory: Probability, Conditional Probability, Bayes' Theorem
- Discrete Random Variables: Probability distributions, Expectation and Variance, Uniform, Binomial, Geometric and Poisson Distributions.
- Estimators: Maximum likelihood estimation, Bias, Bayesian Inference
- Implementation of the above theory in MATLAB.

This is a combined worksheet for all the Statistics/Probability lectures, as week 11 will be used to demonstrate the CM2104 coursework. Some of the problems here may not be covered until a lecture after the week 10 lab class.

This lab sheet focuses more on MATLAB programming examples with a couple theoretical questions. A lot more theoretical questions are available in an additional work sheet available online:

http://www.cs.cf.ac.uk/Dave/CM2104/Labs/Week_10/LabWeek10Additional.pdf

Solutions for both worksheets will be made available on Learning Central after the lab class.

None of the work here is part of the assessed coursework for this module.

MATLAB Statistics/Discrete Probability Programming Problems

Write MATLAB code to solve the following problems:

1. Output the sample space for rolling two n sided dice.
2. Output the sample space for rolling m, n sided dice.
3. Eight students need to complete a group assignment. There are four students needed for programming, three students to do a requirement analysis, and one student to write the report. In how many ways can the eight students be assigned a task (programming / requirement analysis / report)?
4. Write MATLAB code to output the distinguishable permutations of a string. (Hint: The solution should be one line of MATLAB, see `doc/help unique`.)
5. Given a dice biased such that

$$p(1) = p(2) = p(3) = p(4) = p(5) = \frac{1}{10} \quad p(6) = \frac{1}{2}$$

- (a) When throwing 3 dice, What is the probability of getting 1 one, 1 three, and 1 six.
 - (b) When throwing 6 dice, What is the probability of getting 3 ones, 1 three, 1 four and 1 six.
6. The probability that a person supports Cardiff City football club in Cardiff is 0.6. Find the probability that out of 8 randomly selected people in Cardiff there are
 - (a) *Exactly three* supporters of Cardiff City
 - (b) *More than 5* who support Cardiff City
 7. A coin is biased so that the probability of obtaining a head is 0.6. If X is the random variable for the number of tosses up to and including the first head find:
 - (a) $P(X \leq 4)$
 - (b) $P(X > 5)$
 8. Find the probability that at least two double sixes are obtained when two dice are thrown 90 times.
 9. Using the Spam Filter code from the lecture (http://www.cs.cf.ac.uk/Dave/CM2104/MATLAB/Statistics/Spam_Filter.zip) augment the functionality of the Spam filter to

- (a) account for *stop words*. Amend the sentence generating code to generate stop words as well as spam/ham words to test you code.
- (b) accommodate n -gram sequence of words, for a variable number of sequence lengths. Experiment with the classification success rate of different length n -grams on the test data.

Discrete Random Variables

1. Prove $E[a] = a$ where a is any constant.
2. Prove $E[aX] = aE[X]$ where a is any constant.
3. Prove $Var[a] = 0$ where a is any constant.
4. Prove that the variance of the geometric distribution is:

$$Var[X] = \frac{1 - q}{q^2}$$

Additional Exercises

Additional Exercises are available at:

http://www.cs.cf.ac.uk/Dave/CM2104/Labs/Week_10/LabWeek10Additional.pdf