

# MATH 241 Fall 2017 Homework #2

Professor Adam Kapelner

Due 5PM under my office door KY604, Monday, Sept 25, 2017

(this document last updated Sunday 17<sup>th</sup> September, 2017 at 3:32pm)

## Instructions and Philosophy

The path to success in this class is to do many problems. Unlike other courses, exclusively doing reading(s) will not help. Coming to lecture is akin to watching workout videos; thinking about and solving problems on your own is the actual “working out”. Feel free to “work out” with others; **I want you to work on this in groups.**

Reading is still *required*. For this homework set, read the section about sample spaces in Chapter 2 and relevant parts of Chapter 1 in Ross. Chapter references are from the 7th edition.

The problems below are color coded: **green** problems are considered *easy* and marked “[easy]”; **yellow** problems are considered *intermediate* and marked “[harder]”, **red** problems are considered *difficult* and marked “[difficult]” and **purple** problems are extra credit. The *easy* problems are intended to be “giveaways” if you went to class. Do as much as you can of the others; I expect you to at least attempt the *difficult* problems.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to 15 points are given as a bonus if the homework is typed using L<sup>A</sup>T<sub>E</sub>X. Links to installing L<sup>A</sup>T<sub>E</sub>X and program for compiling L<sup>A</sup>T<sub>E</sub>X is found on the syllabus. You are encouraged to use **overleaf.com**. If you are handing in homework this way, read the comments in the code; there are two lines to comment out and you should replace my name with yours and write your section. If you are asked to make drawings, you can take a picture of your handwritten drawing and insert them as figures or leave space using the “\vspace” command and draw them in after printing or attach them stapled.

The document is available with spaces for you to write your answers. If not using L<sup>A</sup>T<sub>E</sub>X, print this document and write in your answers. I do not accept homeworks not on this printout. Keep this first page printed for your records. Write your name and section below.

NAME: \_\_\_\_\_ SECTION (A, B or C): \_\_\_\_\_

**More counting** These counting questions will give you more practice in computing probabilities. Due to computations involving large factorials, we will also review Stirling's Approximation.

### Problem 1

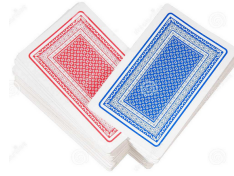
We have 4 blue marbles, 4 green marbles, 2 orange marbles, and 2 red marbles. For the following questions, if you are using “choose notation”, please write your choose notation, then write the formula using factorials, then write the actual number after you compute it.



- (a) [easy] Viewing all the marbles as *unique*, how many ways are there to order the marbles? Note that “order” is another way of saying “permute.”
  
  
  
  
  
  
  
  
  
  
- (b) [harder] Viewing all marbles of the same color as *interchangeable*, how many ways are there to order the marbles?
  
  
  
  
  
  
  
  
  
  
- (c) [E.C.] If I pick 4 marbles at random from the collection, how many ways are there to get two-of-a-kind *i.e.* two marbles of one color and two marbles of a different color.

## Problem 2

Imagine you have a bag of 10 cards where 6 are blue and 4 are red. A “draw” means one card is taken out of the bag at random and the color is revealed. If the problem asks “what is the probability,” this means an explicit computation is required unless otherwise stated.



- (a) [easy] What is the probability of getting a blue card when drawing one card?
- (b) [easy] What is the probability of drawing 3 red cards in a row *without replacement*?
- (c) [harder] Five cards are drawn. What is the probability of having 3 reds and 2 blues without regards to any order of the cards?
- (d) [difficult] Five cards are drawn. What is the probability of having 3 reds and 2 blues in that order? Think carefully about the numerator and denominator in this probability computation.

### Problem 3

Imagine you are putting together musical performances and you are employing musicians at random. There are many available for hire: 23 guitarists, 15 vocalists, 6 drummers, 14 bassists, 8 violinists, 9 violas, 6 cellists.



- (a) [easy] If we hire 4 musicians at random, what is the probability we get a rock band (a vocalist, a guitarist, a bassist and a drummer)?
- (b) [easy] If we hire 4 musicians at random, what is the probability we get a string quartet (two violinists, 1 cellist and one violist)?
- (c) [easy] If we hire 4 musicians at random, what is the probability we get a doowop group (four vocalists)?

- (d) [easy] We now move to a different city and the musicians for hire are different. Here, we have 10 guitarists, 10 vocalists, 10 drummers, 10 bassists. What is the probability we form a rock band when hiring four musicians at random?
- (e) [difficult] Given the same situation in part (d), what is the probability we get two pairs of musicians (e.g. two guitarists and two bassists or two drummers and two bassists)?
- (f) [difficult] Given the same situation in part (d), what is the probability we get all four musicians be the same type?

## Problem 4

Combinations are not only useful in probability problems. They come up all over mathematics.

- (a) [easy] Below is “Pascal’s Triangle” up to  $n = 4$ .

$$\begin{array}{rcccccc}
 n = 0: & & & & & 1 \\
 n = 1: & & & 1 & & 1 \\
 n = 2: & & 1 & & 2 & & 1 \\
 n = 3: & 1 & & 3 & & 3 & & 1 \\
 n = 4: & 1 & & 4 & & 6 & & 4 & & 1
 \end{array}$$

Explain why the 6 in the middle of the  $n = 4$  row is equivalent to  $\binom{4}{2}$  by using the fact proved in class.

(b) [easy] In the expansion  $(a + b)^{100}$ , how many terms are of the form  $a^2b^{98}$ ?

(c) [E.C.] Prove the binomial theorem for arbitrary  $n \in \mathbb{N}$ .

**Philosophy of Probability** We covered many definitions of probability and discussed some philosophy of probability.

### **Problem 5**

Answer the following questions by writing a paragraph or two *in English*.

(a) [easy] Previously we defined probability as  $\mathbb{P}(A) := \frac{|A|}{|\Omega|}$ . Describe a situation where this fails to produce the correct probability that is not the spinner used in lecture.

- (b) [easy] Which definition of probability does the book use and why do you think the authors chose this definition?
- (c) [easy] Give an example of an event whose probability cannot be approximated by the limiting frequency.
- (d) [E.C.] Who picks  $\omega \in \Omega$  *i.e.* the outcome from the set of possible outcomes in the universe? This is an issue we ignored. Discuss your thoughts.
- (e) [easy] What are some problems with the long run frequency definition of probability?
- (f) [harder] How did Chevalier de Mere in 1654 know that the  $\mathbb{P}(\text{one or more double sixes in 24 rolls of two dice}) < \frac{1}{2}$ ?

- (g) [easy] What are some problems with the propensity definition of probability?
- (h) [E.C.] What idea(s) inspired Karl Popper to invent the propensity definition?
- (i) [easy] What is the main problem with the subjective definition of probability?
- (j) [easy] According to Laplace (and his interpretation of Newton), if all information was known about physical systems including all laws and all initial conditions, would there be randomness? Yes/no and discuss.
- (k) [difficult] According to Laplace, what is randomness? I've uploaded Laplace's quote in lecture 5 on the course homepage. You can answer this in a few words.
- (l) [difficult] Knowing what we know in the 21st century, if all information was known about physical systems including all laws and all initial conditions, would there be randomness? If so, what theory has demonstrated evidence for this?
- (m) [difficult] What is the prevailing historical theory about why probability wasn't formalized using mathematics prior to the 1600's?



## Problem 6

Assume that the overall probability of contracting breast cancer in a 45 year old American woman is 0.1% on average (or one in a thousand). A typical diagnostic test is a mammographic scan. Assume also that a mammograph scan reading is 82% *sensitive* on average and 96% *specific* on average. Here, “sensitive” means among patients with cancer, the probability that the test is positive and “specific” means among patients without cancer, the probability that the test is negative.

- (a) [easy] Denote cancer as  $C$  and no cancer as  $C^C$  and mammography positive as  $T$  and mammography negative as  $T^C$ . What is  $\mathbb{P}(C)$ ,  $\mathbb{P}(T \mid C)$  and  $\mathbb{P}(T^C \mid C^C)$ ? These are readable from the problem statement above. You must use this notation going forward to get full credit.
  
- (b) [easy] Now solve for  $\mathbb{P}(C^C)$ ,  $\mathbb{P}(T^C \mid C)$  and  $\mathbb{P}(T \mid C^C)$  using the complement rule.
  
- (c) [easy] Draw a tree with two branches:  $C$  vs.  $C^C$  and then draw a second set of branches for  $T$  vs.  $T^C$  (four branches). Mark all four conditional probabilities in this tree’s configuration and all four marginal probabilities on the right. Check your answers by assuring that these four marginal probabilities form a partition of  $\mathbb{P}(\Omega) = 1$ .

- (d) [easy] Draw the “inverted” tree. It has two branches:  $T$  vs.  $T^C$  and then a second set of branches for  $C$  vs.  $C^C$  (four branches). Mark all four conditional probabilities in this tree’s configuration and all four marginal probabilities on the right. Check your answers by assuring that these four marginal probabilities form a partition of  $\mathbb{P}(\Omega) = 1$ .
- (e) [easy] Draw  $\Omega$  as a rectangle (that takes up the whole width of the page) with  $T$  and  $C$  inside. Try to draw to scale.
- (f) [harder] What is  $\mathbb{P}(T)$ ? Use the law of total probability here and explain what the law is and how exactly you’re using it to solve this problem.

(g) [difficult] What does  $\mathbb{P}(T)$  mean? Answer *in English*.

(h) [harder] Now the money question: if a woman is scanned and tests positive, what is the probability she has cancer? Use the notation I have provided and answer as a *percentage* so it is more viscerally interpretable to you. Do not be alarmed if the answer surprises you.

(i) [harder] You may have done the previous question over and over and gotten frustrated. Your answer is probably correct though. Can you explain why it's so low? Comment on the usefulness of mammography given the post test probability of cancer which you computed.

(j) [easy] Prove through the rules in class that  $\mathbb{P}(T) = \mathbb{P}(C, T) + \mathbb{P}(T \mid C^C) \mathbb{P}(C^C)$ . This is the denominator that is making part (g) so small.

(k) [harder] You have shown that  $\mathbb{P}(C, T)$  is trivially small but  $\mathbb{P}(C^C)$  is huge. So what is in actuality driving the low answer in (g)?

(l) [harder] If a woman is scanned and tests positive, what is the probability she does *not* have cancer?

(m) [harder] If a woman is scanned and tests negative, what is the probability she does *not* have cancer?

(n) [harder] What is the ratio of  $\frac{\mathbb{P}(C | T)}{\mathbb{P}(C | T^C)}$ ? What does this ratio mean? What does your answer suggest? Is it possible these scans aren't such a terrible diagnostic tool after all?