MATH 241 Spring 2015 Homework #2

Professor Adam Kapelner

Due 5PM outside my office, Tuesday, Feb 17, 2015

(this document last updated Thursday 19$^{\text{th}}$ February, 2015 at 5:46pm)

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 permutations / combinations in Chapter 1 in Ross as well as the section about Stirling's approximation. Chapter references are from the 7th edition. You also need to read the first six pages of Chapter 1 of Donald Gillies "Philosophical Theories of Probability".

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]". There are no E.C. problems on this homework. 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. Late homework will be penalized 10 points per day.

15 points are given as a bonus if the homework is typed using LaTeX. Links to instaling LaTeX and program for compiling LaTeX 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.

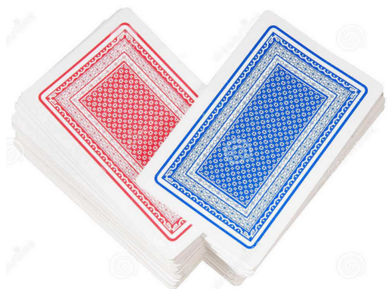
The document is available with spaces for you to write your answers. If not using LaTeX, print this document and write in your answers. **Handing it in without the printout incurs a penalty of 20 points.** Keep this page printed for your records. Write your name and section below where section A is if you're registered for the 9:15AM–10:30AM lecture and section B is if you're in the 12:15PM-1:30PM lecture.

NAME: ______________________________ SECTION (A or B): ______

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

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.

(e) [harder] Five cards are drawn from a new bag with 100 cards where 60 are blue and 40 are red. What is the probability of having 3 reds and 2 blues without regards to any order of the cards?

(f) [harder] Five cards are drawn from a new bag with 1000 cards where 600 are blue and 400 are red. What is the probability of having 3 reds and 2 blues without regards to any order of the cards?

(g) [harder] Five cards are drawn from a new bag with n cards where $0.6n$ are blue and $0.4n$ are red. What is the probability of having 3 reds and 2 blues without regards to any order of the cards? Do not compute the probability explicitly here; leave your solution as an algebraic expression *i.e.* as a function of n.

Problem 2

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 3

Computations of combinations and permutations is impossible for a computer due to the factorial computations. In this problem, we investigate Stirling's formula

(a) [easy] Use the natural log to derive an expression for $\binom{1500}{300}$ using sums.

(b) [easy] The logs in the previous problem would allow us to compute. What would be a downside?

(c) [easy] Show that Stirling's approximation is equivalent to the expression I wrote in class:

$$\ln(n!) \approx \frac{1}{2}\ln(2\pi) + \left(n + \frac{1}{2}\right)\ln(n) - n$$

(d) [difficult] Use the expression in (c) to approximate the probability of getting 300 Heads in 1500 coin flips.

Problem 4

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

(a) [harder] In the first lecture we mentioned that $\left|2^{\Omega}\right| = 2^{|\Omega|}$. (recall that the powerset contains all subsets of Ω *i.e.* $A \in 2^{\Omega} \quad \forall A \subseteq \Omega$). We reasoned that each $\omega \in \Omega$ can be either *in* or *out* of a subset. Thus on/off for the first outcome, on/off for the second outcome, etc. to make 2 raised to the number of elements. This will count every possibly subset. All "offs" would result in $\varnothing$ and all "ons" will result in Ω.

Assume $\Omega = \{a, b, c, d\}$. Explain *in English* how the following equation is true by explaining each element in the sum.

$$\left|2^{\Omega}\right| = \sum_{i=0}^{|\Omega|} \binom{|\Omega|}{i}$$

It may be helpful to draw out 2^{Ω} explicitly and write out the above equation in order to see the pattern. Each of the combination terms will correspond to a subset of 2^{Ω}.

(b) [easy] Recall the binomial theorem:

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$$

Explain *in English* why are the $\binom{n}{i}$ terms called "binomial coefficients."

(c) [E.C.] Prove the equality in part (a) for arbitrary but finite-sized Ω.

(d) [E.C.] Prove the binomial expansion in part (b) for arbitrary $n \in \mathbb{N}$.

Philosophy of Probability Problems below are related to the readings in Gillies as well as the material we covered in class. The link to Gillies' book is in on the first page of this PDF.

Problem 5

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

(a) [easy] Which definition of probability does the book use and why do you think the authors chose this definition?

(b) [easy] Give an example of an event whose probability cannot be approximated by the limiting frequency.

(c) [harder] Give an example of a random event involving an object's "propensity" and explain this definition of probability.

(d) [easy] Discuss the difference between the "logical" and the "subjective" definition of probability.

(e) [easy] Explain the difference between "objective" and "epistemic" interpretations of probability. Which definitions fall under these categories? Classify all four of Gillies' definitions in this way.

(f) [E.C.] For the objective definitions where experiments are conducted, who picks $\omega \in \Omega$ *i.e.* the outcome from the set of possible outcomes in the universe? Discuss your thoughts.

Problem 6

We will be looking into the long term frequency definition here. For this problem, you must have `R` installed. Please download it from `http://cran.r-project.org/` (there are links for Windows, MAC and Linux) and then double-click to open an `R` console.

(a) [easy] To calculate combinations, use the `choose(n,k)` function. Calculate the number of five-card hands from a standard deck by copying the following code into `R` and then pressing enter:

```
choose(52, 5)
```

Please write down the answer. Is the answer the same as we computed in class?

(b) [easy] Verify the probability in class of a "full house" by copying the following code into `R` and then pressing enter:

```
choose(13, 1) * choose(4, 3) * choose(12, 1) * choose(4, 2) /
  choose(52, 5)
```

Write down the answer as a *percentage*.

(c) [harder] We are going to do a little experiment to explore the definition of probability as a limiting frequency. We will be looking at the context of flipping a coin and getting heads. Remember the definition was

$$\mathbb{P}\left(\{H\}\right) = \lim_{n \to \infty} \frac{\sum_{i=1}^{n} \mathbb{1}_{\omega_i \in \{H\}}}{n}$$

(where $\mathbb{1}_T$ is the "indicator function" which equals 1 when the expression T is true and 0 if the expression T is false). We will run a simulation with large values of n. Copy and paste the following code into your `R` terminal:

```
N = 30000
sims = sample(0:1, N, replace = T)
freqs_by_n = array(NA, N)
for (n in 1 : N){
  freqs_by_n[n] = sum(sims[1:n]) / n
}
plot(10:N,
  freqs_by_n[10:N],
  xlim = c(10, N),
  ylim = c(0.40, 0.60),
  pch = ".",
  xlab = "number of samples",
  ylab = "frequency of heads",
  main = "P(H) as a limiting frequency: 30,000 samples")
abline(h = 0.5, col = "blue")
freqs_by_n[N]
#last line placeholder
```

If the code throws an error, download the PDF to your computer and try copying and pasting again. It does NOT work from a browser window. The PDF must be downloaded.

The console should have popped up a plot.[1] Print this out and attach it to your homework. If you are using LaTeX, you can include the figure into the PDF.

From the title of the plot and the x and y axes, tell a story about what is going on here *in English*.

(d) [easy] What is the limiting frequency of heads after 30,000 coin flips to 3 decimals based on the simulation in the previous problem? (that is the number that appears in the console directly after "> `freqs_by_n[N]`")

[1] This is a *real* statistical simulation. Each time you run this code it will be different. You can compare plots with your friends but take note that they will not look exactly the same.