MATH 241 Fall 2014 Homework #2

Professor Adam Kapelner

Due 5PM in my office, Tues Sept 16, 2014

Instructions and Philosophy

Once again, 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, please read Chapter 1 of Hogg, Tanis and Zimmerman (H, T & Z) pages 1–18 as well as the first six pages of Donald Gillies "Philosophical Theories of Probability" chapter 1.[1]

The problems below are color coded: green problems are considered *easy*; yellow problems are considered *intermediate*, red problems are considered *difficult*; and purple problems are for *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. I have put more extra credit problems here by popular demand; they are VERY challenging.

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 up to a maximum of three days. After three days, it will receive a zero because I will post the solutions.

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. If you are handing in homework this way, read the comments in the code; there are two lines to comment out and place to put your name.

If you are not doing the homework in LaTeX, there are two options for hand-in formats: (1) you handwrite on this paper itself and (2) you handwrite on a separate paper. Thus, the compiled PDF of this document will be available on the course homepage in two forms: (1) with vertical spaces for your handwritten answers and (2) without spaces. You must show your work! At least enough of it so that I know you derived the result and not only copied the answer from someone else.

I STRONGLY recommend using strategy number 1. It is not only easy for me to grade your homework, but you always have the questions and answers in one place and it makes for easier referencing. Write your name on page 2 and do not hand in this page.

Write your name on the next page and do not print out this page. Mark your class section: section A is 9:15AM–10:30AM and section B is 12:15PM-1:30PM.

[1] `http://www.amazon.com/Philosophical-Theories-Probability-Issues-Science/dp/041518276X`

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) What is the probability of getting a blue card when drawing one card?

(b) What is the probability of drawing 3 red cards in a row *with replacement*?

(c) What is the probability of drawing 3 red cards in a row *without replacement*?

(d) Five cards are drawn. What is the probability of having 3 reds and 2 blues without regards to any order of the cards?

(e) 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.

(f) 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?

(g) 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?

(h) 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.

(i) Approximate $\ln(1000!)$. See your notes on using logs in conjuction with Stirling's Approximation.

(j) [E.C.] 500 cards are drawn at random from the same bag as problem (g) — 1,000 cards where 600 are blue and 400 are red. What is the probability of getting 250 blue cards and 250 red cards without regads to the order of the cards? The answer should be computed explicitly.

(k) [E.C.] Take the limit in problem (j) as $n \to \infty$ and compute the probability explicitly. Is the answer similar to f and g? Comment on the similarity of sampling *with replacement* and sampling *without replacement* when the bag is large.

Problem 2 Combinations are not only useful in probability problems. They come up all over mathematics.

(a) 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) Explain *in English* why the binomial expansion below is true.

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

Try to do this yourself. If you are having trouble, paraphrase the reasoning found on page 15, example 1.2-10.

(c) Explain *in English* why are the $\binom{n}{k}$ terms called “binomial coefficients.”

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

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

Problem 3
This problem involves using the multinomial coefficient to solve problems.

(a) Imagine you have 12 flowers: 4 red and 3 blue and 5 white. How many ways are there to arrange them in 12 flower pots.

(b) We add 2 orange flowers to collection in part (a). How many ways to arrange the flowers now?

(c) Imagine we have 5 flowers: one white, one blue, one red, one orange and one purple. How many ways to arrange them? Use the multinomial coefficient and show that it is equal the number you arrive at using the permutation concept from lecture 2.

Probability as Applied Set Theory Problems below are related to set theory and probability

Problem 4 We will get our feet wet with basic "axioms" and theorems. Assume all capital letters are sets. If the problem asks you to prove a fact, you may only use your knowledge of set theory and the definition of $\mathbb{P}(\cdot)$ given in the book / lecture. Some of the answers are in the book. Try to do them yourself and only use the book if you are having trouble. The extra credits are really, really difficult.

(a) List all assumptions prior to and the three conditions that make $\mathbb{P}(\cdot)$, the set function that returns a probability. These three conditions are also known as the "axioms of probability."

(b) Prove that if A_1 and A_2 are disjoint (mutually exclusive), $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$.

(c) Prove that $\mathbb{P}(\varnothing) = 0$.

(d) Prove that $\mathbb{P}(A) \in [0, 1]$.

(e) Prove that if $A \subseteq B$ then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(f) Prove that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A, B)$.

(g) [E.C.] Describe a sequence of sets $A_1, A_2, \ldots$ which are all non-empty where $\sum_{i=1}^{\infty} \mathbb{P}(A_i) = 1$. "Describe" means to explicitly state the elements in each of the sets. Hint: the sets do not have to be finite nor countable for that matter. I strongly suggest you also construct $A_1, A_2, \ldots$ as disjoint otherwise the sum of their probabilities may be greater than 1.

(h) [E.C.] Let $A_1 \subseteq A_2 \subseteq A_3, \subseteq \dots$ (this is called a sequence of "increasing events.") Prove that:

$$\lim_{n\to\infty} \mathbb{P}\left(A_n\right) = \mathbb{P}\left(\lim_{n\to\infty} A_n\right)$$

Philosophy of Probability Problems below are related to the readings in Gillies as well as the material we covered in class.

Problem 5 Answer the following questions by writing a paragraph or two *in English*.

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

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

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

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

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

(f) [E.C.] Who picks $\omega \in \Omega$ *i.e.* the outcome from the set of possible outcomes in the universe? Discuss your thoughts.

(g) [E.C.] Is probability an illusion or is it real? Is randomness a fundamental property of 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) 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) 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) 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

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

(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
```

The console should have popped up a plot.[2] 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) 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]`")

[2]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.