Math 241 Spring 2015
Midterm Examination Two   *Solutions*

Professor Adam Kapelner

April 16, 2015

Full Name _____ Section (A or B) _____

Code of Academic Integrity

Since the college is an academic community, its fundamental purpose is the pursuit of knowledge. Essential to the success of this educational mission is a commitment to the principles of academic integrity. Every member of the college community is responsible for upholding the highest standards of honesty at all times. Students, as members of the community, are also responsible for adhering to the principles and spirit of the following Code of Academic Integrity.

Activities that have the effect or intention of interfering with education, pursuit of knowledge, or fair evaluation of a student's performance are prohibited. Examples of such activities include but are not limited to the following definitions:

Cheating Using or attempting to use unauthorized assistance, material, or study aids in examinations or other academic work or preventing, or attempting to prevent, another from using authorized assistance, material, or study aids. Example: using a cheat sheet in a quiz or exam, altering a graded exam and resubmitting it for a better grade, etc.

I acknowledge and agree to uphold this Code of Academic Integrity.

_____   _____
              signature   date

Instructions

This exam is seventy five minutes and closed-book. You are allowed one page (front and back) of a "cheat sheet." You may use a graphing calculator of your choice. Please read the questions carefully. If I say "compute," this means the solution will be a number. **Please round to two decimal places.** I advise you to skip problems marked "[Extra Credit]" until you have finished the other questions on the exam, then loop back and plug in all the holes. I also advise you to use pencil.

The exam is 100 points total plus extra credit. Partial credit will be granted for incomplete answers on most of the questions. $\boxed{\text{Box}}$ in your final answers. Good luck!

Problem 1

Consider the coin toss. Assume all coins are fair i.e. probability of heads is one half. If Heads comes up, we pay \$1; if Tails comes up, we pay \$0.

(a) [3 pt / 3 pts]   Create a random variable model for the winnings of one coin flip. You must write "$X \sim$ something" below.

$$X \sim \text{Bernoulli}\left(\tfrac{1}{2}\right)$$

(b) [4 pt / 7 pts]   Now, three coins are tossed. Create a random variable model T for the winnings on three coin flips. You must write "$T \sim$ something" below.

$$T \sim \text{Binomial}\left(3, \tfrac{1}{2}\right)$$

(c) [5 pt / 12 pts]   In the previous question you created a r.v. T. Enumerate every element in the domain of T.

$$T : \Omega \to \mathbb{R} \qquad \Omega = \{H,T\} \times \{H,T\} \times \{H,T\}$$
$$= \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$$

Problem 2

The site `qcgrades.com` where you check your grades has five-character passwords. Each password is generated by drawing five characters from the set $\{a, b, \ldots, z, 0, 1, \ldots, 9\}$ five times (with replacement). That's 36 possible characters. There are 46 students in this class (sections A and B combined).

We are interested in the specific password "4mh8t" where the characters are in that order. The probability of this password being generated for a single student is an exercise for Midterm I so I have done it for you below. We will denote this value as p:

$$p := \mathbb{P}(4\text{mh8t}) = \mathbb{P}(4)\,\mathbb{P}(m)\,\mathbb{P}(h)\,\mathbb{P}(8)\,\mathbb{P}(t) = \left(\frac{1}{36}\right)^5 \approx 1.65 \times 10^{-8}$$

(a) [3 pt / 15 pts]   What is the probability in the class of 46 students that one person has this password? Answer as a function of p and numerical constants only.

$$T \sim \text{Binomial}\left(46, p\right) \qquad P(T=1) = \binom{46}{1} p^1 (1-p)^{45} = \boxed{46p(1-p)^{45}}$$

2

(b) [4 pt / 19 pts]   What is the probability that this password is shared by two or more students among the 46 students? Answer as a function of p and numerical constants only.

$$P(T \geq 2) = 1 - P(T = 0) - P(T = 1)$$

$$= 1 - \binom{46}{0} p^0 (1-p)^{46} - (\text{answer from a}) = \boxed{1 - (1-p)^{46} - 46p(1-p)^{45}}$$

(c) [4 pt / 23 pts]   I generate one password for each student. How many students do I expect to assign passwords for until this password ("4mh8t") is assigned? Answer as a function of p and numerical constants only.

$$X \sim \text{Geom}(p) \quad, \quad E(X) = \boxed{\frac{1}{p}}$$

(d) [4 pt / 27 pts]   So far, there have been 97 students who have been assigned qcgrades passwords but none of them were assigned "4mh8t". How many students do I expect to assign passwords until this password ("4mh8t") is assigned *conditional* on knowing 97 passed without being assigned this password? Answer as a function of p and numerical constants only.

$$E\left[X \mid X > 97\right] = E(X) = \boxed{\frac{1}{p}}$$

memorylessness of the geometric r.v

(e) [5 pt / 32 pts]   If I assign 1,000,000 passwords, what is the probability I have 50 duplicates of "4mh8t" if the 1,000,000th password is "4mh8t?" Answer as a function of p and numerical constants only.

$$X \sim \text{Neg Bin}(50, p) \qquad P(X = 1,000,000) = \boxed{\binom{999,999}{49}(1-p)^{999,950} \, p^{50}}$$

Problem 3

There are 13 female and 33 male students between sections A and B.

(a) [5 pt / 37 pts]   If I call on three different students at random, find the probability two of them are female. No need to compute explicitly.

$K = 13$

$N = 13 + 33 = 46$

$h = 3$

$$X \sim \text{Hyper}(3, 13, 46) \quad, \quad P(X = 2) = \boxed{\frac{\binom{13}{2}\binom{33}{1}}{\binom{46}{3}}}$$

3

(b) [4 pt / 41 pts]   Given that I call on three different students at random every day, what would the average number of females I call on per day be? Of course this is a long-run approximation. Compute to two decimal places.

$$E(X) = n\frac{K}{N} = 3 \cdot \frac{13}{46} \approx \boxed{.85}$$

(c) [4 pt / 45 pts]   If I call on 35 different students at random, how many females can I possibly call on? List all the possibilities below as a set.

$$n = 35$$
$$K = 13$$
$$N = 46$$

$$X \sim Hyper(35, 13, 46)$$

$$Supp(X) = \boxed{\{2, 3, \dots, 13\}}$$

(d) [4 pt / 49 pts]   Consider the proportion of males to females to be the same in all math classes across all of CUNY as it is in our Math 241 class. Enrollment is about 10^5. We are going to redo the previous question in this new scenario. If a lottery was done which selected a random sample of 35 math students in total, find the probability two of them will be female. Of course, your answer will be approximate. No need to compute to decimal but your answer must be computable without log approximations.

$$Y = \lim_{N \to \infty} Hypergeom\left(35, p = \frac{13}{46}, N\right) \sim Bin\left(35, \frac{13}{46}\right)$$

$$P(Y=2) = \boxed{\binom{35}{2}\left(\frac{13}{46}\right)^2 \left(\frac{33}{46}\right)^{33}}$$

Problem 4

We play a casino game with two independent colored spinners pictured below.
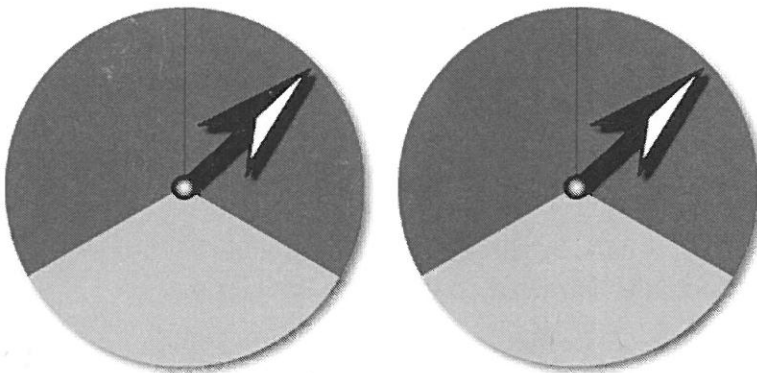
(a) [3 pt / 52 pts]   Assume the spinners are fair (meaning that the probability of red is $1/3$, the probability of blue is $1/3$ and the probability of green is $1/3$). Assume the following payouts: Red wins \$2, Blue wins \$5 and Green loses \$4. Create a r.v. X for the payout of the game of one spinner. Write the PMF using brace notation below.

$$X \sim \begin{cases} \$2 & \text{wp } \frac{1}{3} \\ \$5 & \text{wp } \frac{1}{3} \\ -\$4 & \text{wp } \frac{1}{3} \end{cases}$$

(b) [4 pt / 56 pts]   If you played this game over and over again, what would your average winnings be close to? Compute to the nearest cent.

$$E[X] = \frac{1}{3}(\$2) + \frac{1}{3}(\$5) + \frac{1}{3}(-\$4) = \frac{1}{3}(2+5-4) = \frac{1}{3} \cdot 3 = \boxed{\$1}$$

(c) [1 pt / 57 pts]   What law did you invoke to justify the approximation in the last question?

$$\text{Law of Large Numbers (LLN)}$$

(d) [4 pt / 61 pts]   Compute $SE[X]$ to the nearest cent.

$$Var(X) = E(X-\mu)^2 = \frac{1}{3}(\$2-\$1)^2 + \frac{1}{3}(\$5-\$1)^2 + \frac{1}{3}(-\$4-\$1)^2$$
$$= \frac{1}{3}(1^2 + 4^2 + 5^2) = \frac{1}{3} \cdot 42 = 14\$^2$$
$$SE[X] = \sqrt{14\$^2} = \boxed{\$3.74}$$

(e) [1 pt / 62 pts]   Now we play with two spinners X_1 and X_2. Is X_1 and $X_2 \overset{iid}{\sim}$? Yes/no.

$$\text{Yes}$$

(f) [3 pt / 65 pts]   Playing with two spinners, we are interested in our total winnings. That is, we now add our winnings from the two spinners together. Call this new r.v. $T := X_1 + X_2$. Compute $\mathbb{E}[T]$ to the nearest cent.

$$E[T] = E[X_1 + X_2] = E[X_1] + E[X_2] = 2E[X] = 2 \cdot \$1 = \boxed{\$2}$$
$$\text{rule for expectation} \quad \text{by id.}$$
$$5 \quad \text{sum}$$
$$\text{r.v.'s}$$

(g) [5 pt / 70 pts] Compute $\mathbb{SE}[T]$.

$$Var(T) = Var[X_1 + X_2] = Var(X_1) + Var(X_2) = 2Var(X)$$

rule for summing variances
of ind. r.v.'s   by i.d.

$$\Rightarrow SE(T) = \sqrt{2}\, SE(X) = \sqrt{2} \cdot \$3.74 = \boxed{\$5.29}$$

(h) [3 pt / 73 pts] Suppose you are now playing in Europe and the payout is in Euros (as of April 6th, 2015, 1 € = \$1.10). Call the payour of two spinners T' and compute $\mathbb{E}[T']$ to the nearest Euro-cent. If you did not solve part (f), consider the answer to part (f) to be \$1.

$$T' = \frac{1\,€}{\$1.10}\, T = 0.91\, €/\$\, T$$

$$\mathbb{E}(T') = \mathbb{E}\left[0.91\, €/\$\, T\right] = 0.91\, €/\$\, \mathbb{E}(T) = 0.91\, €/\$\,(\$2) = \boxed{1.82\,€}$$

scalar multiple rule

(i) [5 pt / 78 pts] Suppose the European casino charges you 3 Euros to play. Call your net winnings T'. Compute $\mathbb{SE}[T']$ rounded to the nearest Euro-cent. Note that this r.v. model T now includes this entrance fee. If you did not solve part (g), consider the answer to part (g) to be \$1.

$$SE(T') = SE\left[3€ + 0.91\, €/\$\, T\right] = \left|0.91\, €/\$\right| SE(T) = 0.91\, €/\$ \cdot \$5.29 = \boxed{4.81\,€}$$

(j) [3 pt / 81 pts] Now imagine I wire the spinners such that if the first spinner goes red, the second spinner goes blue and if the first spinner goes blue the second spinner goes green and if the first spinner goes green, the second spinner goes red. Are X_1 and X_2 independent? Prove your answer.

$$1 = P\underbrace{\left(X_1 = \$2\right.}_{Red} \Big| \underbrace{\left.X_2 = \$5\right)}_{Blue} \overset{?}{=} \underbrace{P\left(X_1 = \$2\right)}_{Red} = \frac{1}{3} \Rightarrow \boxed{No}$$

(k) [2 pt / 83 pts] Given the manipulative wiring in the previous question, are $X_1 \overset{d}{=} X_2$? Yes/no is fine.

Yes, despite being dependent

6

(l) [3 pt / 86 pts]   Given the manipulative wiring in the previous question, would $\mathbb{SE}[T]$ change from your answer in (g)? Answer yes / no and why.

Yes, $\mathbb{SE}[T] = \sqrt{Var[T]} = \sqrt{Var[X_1 + X_2]}$

and $Var[X_1 + X_2] \neq Var(X_1) + Var[X_2]$ since X_1, X_2 are no longer dependent!

(m) [6 pt / 92 pts]   [Extra Credit] Given the manipulative wiring in the previous question, calculate $\mathbb{SE}[T]$ in dollars rounded to the nearest cent.

— or —

$E[T] = \$2$

$T = X_1 + X_2 \sim \begin{cases} 2+5 & \text{wp } \frac{1}{3} \\ 5+-4 & \text{wp } \frac{1}{3} \\ -4+2 & \text{wp } \frac{1}{3} \end{cases} = \begin{cases} \$7 & \text{wp } \frac{1}{3} \\ \$1 & \text{wp } \frac{1}{3} \\ \$-2 & \text{wp } \frac{1}{3} \end{cases}$

$Var[X_1 + X_2] = Var[X_1] + Var[X_2] + 2Cov[X_1, X_2]$
$= \$4^2 \cdot \$3^2 14$
$+ 2(E[X_1 \cdot X_2] - E[X_1]E[X_2])$
$= \$28^2 + 2(E[X_1 X_2] - (\$1)^2)$
$= \$28^2 - 2.\$1^2 + 2 E[X_1 X_2]$
$= \$26^2 + 2 E[X_1 X_2] = \$26^2 + -12\$^2$
$= \$14 \$^2$

$Var[T] = E(T - E[T])^2 = (7-2)^2 \frac{1}{3} + (1-2)^2 \frac{1}{3} + (-2-2)^2 \frac{1}{3}$

$= \frac{1}{3}(25 + 1 + 16) = \frac{1}{3} 42 = 14 \$^2$

$\Rightarrow \mathbb{SE}[T] = \sqrt{14 \$^2} = \boxed{\$ 3.74}$

$E[X_2] = \frac{1}{3}(10)$
$+\frac{1}{3}(-20)$
$+\frac{1}{3}(8)$
$= \frac{1}{3}(10-20-8)$
$= \frac{1}{3}(-18) = -6$

Problem 5

Some theoretical questions are below. The subparts are all independent unless otherwise indicated.

(a) [3 pt / 95 pts]   Give an example of a brand-name r.v. X where $\mu \notin \text{Supp}[X]$. You must give the parameter values but no need to describe $\text{Supp}[X]$ nor compute μ for your r.v.

$X \sim \text{Bernoulli}(3, \frac{1}{2})$, $E[X] = \frac{3}{2} \notin \{0,1,2,3\}$

(b) [3 pt / 98 pts]   Give an example of a brand-name non-degenerate r.v. X where $\mu \in \text{Supp}[X]$. You must give the parameter values but no need to describe $\text{Supp}[X]$ nor compute μ for your r.v.

$X \sim \text{Bernoulli}(4, \frac{1}{2})$, $E[X] = 2 \in \{0,1,2,3,4\}$

7

(c) [2 pt / 100 pts] If $X_1, \ldots, X_n \overset{iid}{\sim} \text{NegBin}(r, p)$. How is $T = X_1 + \ldots + X_n$ distributed?

Since each NegBin is the sum of r geometrics, this is the sum of sum of r geometrics

$$\Rightarrow T \sim \text{NegBin}\left(\underbrace{r + r + \ldots + r}_{n}, p\right) = \boxed{\text{NegBin}(nr, p)}$$

(d) [2 pt / 102 pts] I am trying to guess my friend's number. I will be penalized using an error function $e(x, y)$ where x is the correct value and y is my guess. If I guess the *correct value*, what is the value of the error function?

0 by definition of error functions

(e) [4 pt / 106 pts] My serum glucose level this morning was 82mg/dL. Could this measurement have been different? Yes / no and discuss. Use terminology, concepts and definitions from homework and lectures.

$X = 82$mg/dL is a draw/realization from X, the r.v. or data generating process. Since 82mg/dL is not the only element in the supp(X), my serum glucose level could have been different due to randomness.

(f) [4 pt / 110 pts] [Extra Credit] If $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$, show $\mathbb{SE}\left[\frac{\bar{X}-p}{\sqrt{p(1-p)/n}}\right] = 1$.

$$\mathbb{SE}[\bar{X}] = \frac{\sigma}{\sqrt{n}}, \quad \sigma \text{ for a Bernoulli is } \sqrt{p(1-p)} \Rightarrow \mathbb{SE}[\bar{X}] = \sqrt{\frac{p(1-p)}{n}}$$

$$\mathbb{SE}\left[\frac{\bar{X}-p}{\sqrt{\frac{p(1-p)}{n}}}\right] = \frac{1}{\sqrt{\frac{p(1-p)}{n}}} \mathbb{SE}[\bar{X}-p] = \frac{1}{\sqrt{\frac{p(1-p)}{n}}} \mathbb{SE}[\bar{X}] = \boxed{1}$$

by scalar multiple rule   by scalar addition rule

(g) [4 pt / 114 pts] [Extra Credit] Give an example of a discrete r.v. X where μ does not exist.

From the homework, $X \sim p(x) = \frac{6}{\pi^2} \frac{1}{x^2}$, supp$(X) = \mathbb{N}$

$$\mu := \mathbb{E}[X] = \sum_{x \in \text{supp}(X)} x \times p(x) = \sum_{x=1}^{\infty} x \times \frac{6}{\pi^2} \frac{1}{x^2} = \frac{6}{\pi^2} \underbrace{\sum_{x=1}^{\infty} \frac{1}{x}}_{} = \infty \notin \mathbb{R}$$

diverges! $\Rightarrow$ does not exist