MATH 241 Fall 2014 Homework #9

Professor Adam Kapelner

Due 5PM in my office, Tues Nov 11, 2014

(this document last updated Thursday 6$^{\text{th}}$ November, 2014 at 8:56am)

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 the expectation and variance section of Chapter 2. Avoid the parts that deal with "moment generating functions" for now.

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 for *extra credit* which are also marked "[E.C.]." 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 may also use `writelatex.com` which is a web service (you don't have to install or configure anything on your local computer). 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. I STRONGLY recommend to write on a printout of this document since you will always have the questions handy to study from (and it is easier for me to grade accurately). 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): ______

Fundamentals of Random Variables We will go over expectation and variance again.

Problem 1

Imagine rolling two dice. Let X_1 be the r.v. corresponding to the first die and let X_2 be the r.v. corresponding to the second die. Let the outcomes be \$1 if you roll a 1, \$2 if you roll a 2, ..., and \$6 if you roll a six.

(a) [easy] What brand name r.v. is X_1 distributed as? Write $X \sim$ something and make sure the parameters are correct.

(b) [easy] Does $X_1 \overset{d}{=} X_2$?

(c) [easy] Are X_1 and X_2 independent?

(d) [easy] Find $\mathbb{E}[X_i]$ for $i \in \{1, 2\}$ from first principles.

(e) [easy] Find $\mathbb{V}\text{ar}[X_i]$ for $i \in \{1, 2\}$ from first principles.

(f) [easy] The standard deviation is also called "standard error" and it sometimes denoted "SE." Use your answer in (e) to find $\mathbb{SE}[X_i]$ for $i \in \{1, 2\}$. Please just use the square root and do not rederive the variance again from scratch.

(g) [easy] Draw the PMF for X_i for $i \in \{1, 2\}$ and mark $\mathbb{E}[X_i]$ and $\mathbb{SE}[X_i]$ on the graph similar to how we did in class.

(h) [easy] Imagine the game where you just double the winnings of a single roll. This would be equivalent to just multiplying the r.v. by a scale factor of 2. Calculate $\mathbb{E}[2X_i]$, $\mathbb{V}\text{ar}[2X_i]$ and $\mathbb{SE}[2X_i]$ from the formulas we learned in class.

(i) [easy] Draw the PMF for $2X_i$ for $i \in \{1, 2\}$ and mark $\mathbb{E}[2X_i]$ and $\mathbb{SE}[2X_i]$ that you calculated in (h) on the graph similar to how we did in class.

(j) [harder] Draw the PMF for $X_1 + X_2$. This involves taking a convolution. Since I don't want to focus on the convolution and it won't be on the midterm or final, I'm going to give a hint. There is 1 way to get 2 or 12, 2 ways to get 3 or 11, 3 ways to get 4 or 10, 4 ways to get 5 or 9, 5 ways to get 6 or 8 and 7 ways to get 7.

(k) [easy] Calculate $\mathbb{E}[X_1 + X_2]$, $\mathbb{V}\text{ar}[X_1 + X_2]$ and $\mathbb{S}\text{E}[X_1 + X_2]$ from the formulas we learned in class.

(l) [difficult] Why are the standard errors in (h) and (k) different and why is (h) larger? This involves a lot of thinking and I want a few sentences *in English*.

(m) [harder] Imagine n rolls of the same dice to produce n r.v.'s denoted $X_1, \ldots, X_n$ which of course are still $\overset{iid}{\sim}$. Calculate $\mathbb{E}[X_1 + \ldots + X_n]$, $\mathbb{V}\text{ar}[X_1 + \ldots + X_n]$ and $\mathbb{S}\text{E}[X_1 + \ldots + X_n]$.

(n) [harder] Calculate $\mathbb{E}\left[\bar{X}_n\right]$, $\mathbb{V}\text{ar}\left[\bar{X}_n\right]$ and $\mathbb{S}\text{E}\left[\bar{X}_n\right]$ using the definition of $\bar{X}_n$ we learned in class. Calculate means give the numeric answer, not leave in terms of n and μ and whatnot. I know you can copy the formulas from my notes.

(o) [difficult] What does it mean that $\mathbb{E}\left[\bar{X}_n\right]$ is an unbiased estimator for μ? Read about unbiased estimators in the book or online. It is also touched on in my lecture but I went through it quickly.

(p) [easy] If $n = 1000$, what is $\mathbb{SE}\left[\bar{X}_n\right]$? Does that mean it's getting really close to $\mathbb{E}\left[\bar{X}_n\right]$? Why or why not.

(q) [difficult] Now you have the choice between game A — where you roll n times and average the winnings (*i.e.* you collect $\bar{X}_n$ dollars at the end) or game B — where you roll one die and collect the amount you make on just one roll. Use your answers to the relevant previous questions (I won't tell you which ones explicitly) to explain why you would choose game A over B or vice versa. I want multiple sentences *in English*. You must convince me you understand the tradeoff that game A and B are making.

(r) [harder] Let Z be the standardized r.v. for $\bar{X}_n$. Prove from the formulas in class that $\mathbb{E}[Z] = 0$ and $\mathbb{V}\text{ar}[Z] = \mathbb{SE}[Z] = 1$.

(s) [easy] If $n = 1000$ and you made $\bar{x} = \$4.00$, what is the z-score of this $\bar{x}$? That is if $\bar{X}_n$ was standardized into the r.v. Z (as in the previous question), what would be the corresponding realization of z that corresponds to this $\bar{x}$.

(t) [easy] We will learn later in Math 241 that $z \notin [-3, 3]$ are very strange and smack of something being awry. Is something awry with making \$4.00 on average? Explain using a sentence *in English*.

Problem 2

More simple r.v. practice.

(a) [difficult] Let $T_n \sim \text{Binomial}\,(n,\, p)$. Prove that $\mathbb{V}\text{ar}\,[T_n] = np(1-p)$ from first principles. This involved two reindexing tricks in the notes.

(b) [harder] Let $X_1 \sim \text{Bernoulli}\,(p)$. Derive an expression for $\mathbb{V}\text{ar}\,[X_1]$ as a function of the parameter as we did in class.

(c) [easy] You know that T_n is the sum of $n \overset{iid}{\sim}$ bernoulli r.v.s with parameter p. Show that $\mathbb{V}\text{ar}\,[T_n]$ can be easily derived using the variance-sum formula we learned in class.

(d) [harder] If you had complete control of both parameters n and p, what would be the easiest manipulation to make variance as small as possible?

(e) [easy] In the limit of that manipulation in (d), what would the final r.v. be? The name and how it's distributed is all that's needed ($X \sim$ somthing).

(f) [easy] Prove $\mathbb{V}\mathrm{ar}\left[X\right] = \mathbb{E}\left[X^2\right] - \mu^2$ for any r.v. X.

(g) [easy] Show from the definition of variance that $\mathbb{V}\mathrm{ar}\left[X\right] \geq 0$ for any r.v. X.

(h) [difficult] Find a r.v. X where $\mathbb{V}\mathrm{ar}\left[X\right] = 0$.

(i) [difficult] Show for any two r.v.'s X and Y which are independent (you need the independence) that $\mathbb{V}\mathrm{ar}\left[XY\right] = \mu_X^2\sigma_Y^2 + \mu_Y^2\sigma_X^2 + \sigma_X^2\sigma_Y^2$.

(j) [easy] Let $a_1, a_2, \ldots, a_n$ be a sequence of constants. Let $X_1, \ldots, X_n$ be a sequence of r.v.'s which share the same mean. Create a simplified expression for $\mathbb{E}\left[a_1 X_1 + \ldots + a_n X_n\right]$.

(k) [harder] Let $a_1, a_2, \ldots, a_n$ be a sequence of constants. Assume $X_1, \ldots, X_n$ are a sequence of $\stackrel{iid}{\sim}$ r.v.'s. Create a simplified expression for $\mathbb{SE}\left[a_1 X_1 + \ldots + a_n X_n\right]$.

(l) [harder] Imagine a r.v. X with density $f(x) = c/x^2$ and $\text{Supp}\left[X\right] = \mathbb{N}$. What is the exact value of c which makes $f(x)$ a valid PMF? The answer can be found here.

(m) [easy] Show that $\mathbb{E}\left[X^2\right]$ does not exist (which of course means by (f) that σ^2 does not exist). This should be extremely easy from the definition of $\mathbb{E}\left[g(X)\right]$ which we've done in class many times.

(n) [harder] Show that $\mathbb{E}\left[X\right]$ does not exist. You will need a fact from that same wikipedia page that you visited in (m). And now you've learned what the harmonic series is too.

(o) [harder] Consider $X \sim \text{NegBin}\left(r,\, p\right)$. Prove that $\mathbb{V}\text{ar}\left[X\right] = r(1-p)/p^2$ assuming that the variance of a geometric r.v. with parameter p is $(1-p)/p^2$.