MATH 241 Spring 2015 Homework #8

Professor Adam Kapelner

Due *in class*, Tuesday, Apr 14, 2015

(this document last updated Wednesday 8th April, 2015 at 10:43am)

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*. But for this homework set, read about expectation, variance and standard deviation.

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]", 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. If the problem asks you for a computation, **round to two or three decimals (do not answer in an exact fraction).**

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 50 points. Read more about this policy in the syllabus.

Between 1–15 points are arbitrarily given as a bonus (conditional on quality) if the homework is typed using LaTeX (15 points only if it is perfect LaTeX). Links to instaling LaTeX and program for compiling LaTeX are found on the syllabus (please read carefully).

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 this printout is NOT ACCEPTABLE.** There is also a redo policy which you can read about in the syllabus.

Keep this page printed for your records (if using LaTeX, this page will not show but a shortened header appears). 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): _____

Expectation and Variance   All in the title...

Imagine rolling two fair dice (no sorcery). 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 outcome results 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. Yes, "uniform" is a brand name r.v.

(b) [easy] Does $X_1 \overset{d}{=} X_2$? Yes or no is fine.

(c) [easy] Are X_1 and X_2 independent? Yes or no is fine.

(d) [easy] Compute $\mathbb{E}[X_1]$ from first principles.

(e) [easy] Compute $\mathbb{V}\text{ar}[X_2]$ 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 $\mathbb{E}[2X_i]$, $\mathbb{Var}[2X_i]$ 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.

(j) [difficult] Draw the PMF for $X_1 + X_2$. This involves taking a convolution. Since convolution 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 $\mathbb{SE}[X_1 + X_2]$ from the formulas we learned in class. Do not use the PMF from the last question; use the formulas from 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) [easy] Imagine the general case of $X_1, \ldots, X_n \overset{iid}{\sim}$ with mean μ and variance σ^2. Define $\bar{X}$ as we did in class. Redo the derivation of $\mathbb{E}\left[\bar{X}\right] = \mu$.

(n) [easy] Imagine the general case of $X_1, \ldots, X_n \overset{iid}{\sim}$ with mean μ and variance σ^2. Define $\bar{X}$ as we did in class. Redo the derivation of $\mathbb{SE}\left[\bar{X}\right] = \sigma/\sqrt{n}$.

(o) [easy] Calculate $\mathbb{E}\left[\bar{X}_n\right]$, $\mathbb{V}\text{ar}\left[\bar{X}_n\right]$ and $\mathbb{SE}\left[\bar{X}_n\right]$ using the definition of $\bar{X}_n$ we learned in class only as a function of n. Hint: use the formula you just derived.

(p) [difficult] Imagine n rolls of the dice to produce n r.v.'s denoted $X_1, \ldots, X_n$ which of course are still $\overset{iid}{\sim}$. Calculate $\mathbb{E}\left[X_1 + \ldots + X_n\right]$, $\mathbb{V}\text{ar}\left[X_1 + \ldots + X_n\right]$ and $\mathbb{SE}\left[X_1 + \ldots + X_n\right]$ only as a function of n.

(q) [difficult] What does it mean that $\mathbb{E}\left[\bar{X}_n\right]$ is an unbiased estimator for μ? Explain in a few sentences *in English* why this is a good thing.

(r) [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.

(s) [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.

(t) [easy] Redo the derivation of the rule $\mathbb{Var}\left[aX\right] = a^2\mathbb{Var}\left[X\right]$ for any discrete r.v. X and any constant $a \in \mathbb{R}$.

(u) [easy] Redo the derivation of the rule $\mathbb{Var}\left[X + c\right] = \mathbb{Var}\left[X\right]$ for any discrete r.v. X and any constant $c \in \mathbb{R}$.

(v) [harder] Let Z be the standardized r.v. for $\bar{X}_n$. Standardization of a r.v. X is defined as subtracting its mean and dividing by its standard error. For $\bar{X}$ this would be:

$$Z := \frac{\bar{X} - \mathbb{E}\left[\bar{X}\right]}{\mathbb{SE}\left[\bar{X}\right]} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Prove from the formulas in class that $\mathbb{E}\left[Z\right] = 0$ and $\mathbb{Var}\left[Z\right] = \mathbb{SE}\left[Z\right] = 1$. Hint: use those two rules about $\mathbb{Var}\left[aX\right]$ and $\mathbb{Var}\left[X + c\right]$ you just rederived.

(w) [easy] Why is "standardization" called "standardization"?

(x) [harder] 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 $\bar{x}$?

(y) [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*.

(z) [harder] This question was inspired by a student's question from last lecture. Returning to X, the dice game in the beginning of the problem (the outcome results being $\$1$ if you roll a 1, $\$2$ if you roll a 2, ..., and $\$6$ if you roll a six), you calculated variance using the

definition $\mathbb{V}\mathrm{ar}\,[X] := \mathbb{E}\left[(X - \mu)^2\right]$ assuming the classic squared error loss: $e(x, \mu) := (x - \mu)^2$. Imagine we defined a new variance metric using absolute loss, $\underbrace{e(x, \mu)} := |x - \mu|$. We'll denote this "new variance" with a big squiggly symbol, $\widetilde{\mathbb{V}\mathrm{ar}}\,[X] := \mathbb{E}\left[|X - \mu|\right]$ just to $\underbrace{\text{make}}$ sure you don't confuse it with the standard definition of $\mathbb{V}\mathrm{ar}\,[X]$. Calculate $\widetilde{\mathbb{V}\mathrm{ar}}\,[X]$ and include units.

(aa) [difficult] $\widetilde{\mathbb{V}\mathrm{ar}}\,[X]$ and $\mathbb{SE}\,[X]$ have the same units but why is $\widetilde{\mathbb{V}\mathrm{ar}}\,[X]$ *not the same* as $\mathbb{SE}\,[X]$?

Problem 2

More simple r.v. practice.

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

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

(c) [E.C.] On a separate sheet of paper, compute $\mathbb{V}\mathrm{ar}\,[T_n]$ as a function of n and p using the definition of variance.

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

(e) [easy] Prove $\mathbb{V}\mathrm{ar}\,[X] = \mathbb{E}\,[X^2] - \mu^2$ for any r.v. X (this is in your notes).

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

(g) [easy] Show that if $\mathbb{V}\mathrm{ar}\,[X] = 0$, then X must be degenerate.

(h) [E.C.] Show that if expected error for any $e(x, \mu)$ is 0, then X must be degenerate.

(i) [difficult] Show for any two r.v.'s X and Y which are independent that $\mathbb{V}\mathrm{ar}\,[X \times Y] = \mu_X^2 \sigma_Y^2 + \mu_Y^2 \sigma_X^2 + \sigma_X^2 \sigma_Y^2$. Remember, two r.v.'s multiplied together is a new r.v., $g(X, Y)$.

(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 thereby share the same mean μ. Create a simplified expression for $\mathbb{E}[a_1 X_1 + \ldots + a_n X_n]$. I want the simplest combination of symbols $a_1, a_2, \ldots, a_n$ and μ.

(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 $\overset{iid}{\sim}$ r.v.'s which thereby share the same variance σ^2. Create a simplified expression for $\mathbb{SE}[a_1 X_1 + \ldots + a_n X_n]$. I want the simplest combination of symbols $a_1, a_2, \ldots, a_n$ and μ.

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

(m) [difficult] X is the same as in the previous problem. Show that $\mathbb{V}\text{ar}[X]$ does not exist (which means it's not a real number). You will need a fact from that same wikipedia page that you visited in the last problem. And now you've learned what the harmonic series is too.

(n) [difficult] Let $Y := X^2$ where X is a r.v. with PMF unknown. Use the notation μ for $\mathbb{E}[X]$, σ^2 for $\mathrm{Var}[X]$, μ_3 for $\mathbb{E}[X^3]$ and μ_4 for $\mathbb{E}[X^4]$. Assume μ, σ^2, μ_3, μ_4 all exist in $\mathbb{R}$. Find $\mathrm{Var}[Y]$ as a function of μ, σ^2, μ_3 and μ_4. This was the extra credit on last semester's midterm, thus the solution is online. Try to do it yourself first.

(o) [difficult] Consider $X \sim \mathrm{NegBin}(r, p)$. Prove that $\mathrm{Var}[X] = r(1-p)/p^2$ assuming that the variance of a geometric r.v. with parameter p is $(1-p)/p^2$.

(p) [harder] Prove the memorylessness property of $X \sim \mathrm{Geometric}(p)$ by proving it holds for the definition of memorylessness i.e. $\mathbb{P}(X > x) = \mathbb{P}(X > x_0 + x \mid X > x_0)$.

(q) [easy] I'm waiting for a defective chip of RAM (similar to a previous HW). Each piece of RAM is independent has probability of being defective of 1 in 300. I've been doing this awhile and 100 non-defective chips have passed by. What is the probability I see a defective chip on the 117th item in the production run?