MATH 241 Spring 2015 FINAL Homework (#11)

Professor Adam Kapelner

Due *in class*, Thursday, May 14, 2015

(this document last updated Thursday 7$^{\text{th}}$ May, 2015 at 2:44pm)

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 confidence intervals and hypothesis tests for one proportion.

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): ______

The Power of the CLT In the last homework we proved the CLT, now we examine its power. We can do statistical inference and statistical testing!

Problem 1

These exercises will pick up after the proof of the CLT.

(a) [harder] If $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$ then what is the $\min\left\{\text{Supp}\left[\bar{X}_n\right]\right\}$?

(b) [harder] If $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}(p)$ then what is the $\max\left\{\text{Supp}\left[\bar{X}_n\right]\right\}$?

(c) [easy] If $\bar{x}$ is a realization from $\bar{X}_n$, what is the smallest value it can be?

(d) [easy] If $\bar{x}$ is a realization from $\bar{X}_n$, what is the largest value it can be?

(e) [easy] Because $\bar{X}$'s support is naturally bounded in the Bernoulli experiment case (see questions a and b of this problem), we give it a special notation: $\hat{P}$. Write the definition of $\hat{P}$ mathematically.

(f) [easy] What do we call $\hat{P}$?

(g) [easy] Find $\mathbb{E}\left[\hat{P}_n\right]$ if $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}\,(p)$.

(h) [harder] Find $\mathbb{SE}\left[\hat{P}_n\right]$ if $X_1, \ldots, X_n \overset{iid}{\sim} \text{Bernoulli}\,(p)$.

(i) [harder] Use your previous answers and assume n is big enough for the CLT to "kick in" as well as what you know about normal distributions shifted and scaled to find the distribution of $\hat{P}_n$.

(j) [difficult] Write the PDF of $\hat{P}_n$. This will force you to understand (1) the generalized density for $\mathcal{N}\left(\mu,\, \sigma^2\right)$ which is in the notes but has never been previously asked on a homework assignment, (2) substitutions for the mean and variance and (3) what the free variable is.

(k) [easy] Draw the PDF of $\hat{P}_n$ below. On the $\hat{p}$ axis, mark a tick for the mean, and ticks for three standard errors above the mean and ticks for three standard errors below the mean. I'm giving almost a full page for this because other problems depend on it. Use 1/3 of the space for your drawing of the PDF.

(l) [harder] Let's say $p = 0.2$ and $n = 100$. What is $\mathbb{P}\left(\hat{P}_n \geq 0.24\right)$? You will need to use shifts and scales to get it to look like $\mathbb{P}(Z > z)$ where Z is the standard normal and z is 0.24 standardized.

(m) [difficult] Let's say $p = 0.2$ and $n = 100$. What is the probability of realizing a $\hat{p}$ above 0.3? Without a computer, you cannot answer this exactly. Just using the empirical rule, find the lower and upper bound of this probability. Do not answer exactly. I will not ask you to answer exactly on the test but I will ask you to find a lower and upper bound using the empirical rule.

Problem 2

We have now used the CLT to understand the distribution of sample proportions. Here, we will assume that p, the true proportion of "success," is unknown and we will try to infer it. This is a problem of paramount importance. Imagine you're polling for a political candidate. You really do want to know p — the proportion of people who will vote for your candidate. Imagine you're investigating a drug with side effects. You really do want to know p — the proportion of people who suffer from that side effect.

(a) [easy] If you have realizations of $X_1, \ldots, X_n$, call them $x_1, \ldots, x_n$, what is your best guess of p? This is called a "point estimate." Make sure to use the special notation for this estimate we used in class and define it mathematically.

(b) [difficult] Write *in English* why $\hat{p} \neq p$. I marked this as hard because you need to really dredge up your knowledge about random variables, realizations and parameters and get all these concepts straight.

(c) [harder] Recall your answer to 1(i) and reference the picture in 1(k). Where are the "middle" 95% of $\hat{p}$'s realized? When I say "middle" I mean the 95% that are *closest* to p, the true expected value. Write the interval that corresponds to this set of numbers. Use the bracket set notation we learned in lecture 1.

(d) [harder] The interval you created in (c) left out 5% of the density of $\hat{P}_n$. Write the set that corresponds to this left out set. Use the union notation from lecture 1.

(e) [easy] What is the width of the interval you created in (c)?

(f) [easy] What is the half width of the interval you created in (c)?

(g) [easy] In your picture in 1(k), draw a dotted line from the center of the PDF, p down the length of the page. Create a realization of $\hat{P}$, call it $\hat{p}$ within the interval in (c) and mark it as a solid dot. Then carefully measure using a ruler, one half width (f) from the fot to the left and one half width from (f) to the right and draw these two lines and mark the ends with a square bracket. No need to write anything below this prompt.

(h) [easy] Did this interval *cover* p? That is, does it include p? You can see if it includes the dotted line drawn down from p.

(i) [easy] Would all $\hat{p}$'s realized within the interval in (c) cover p?

(j) [difficult] Prove that the probability that the interval covers p is 95%. We did this in class.

(k) [easy] Illustrate an example of a $\hat{p}$ with the same half-width to the left and right interval that does *not* cover p in your picture of 1(k). No need to write anything here.

(l) [harder] In what set do the $\hat{p}$'s such as in (k) come from that fail to cover p?

Problem 3

Now we'll just practice building CI's and ask some questions about the procedure.

(a) [harder] Why is it important to have a simple random sample when sampling from a population? Write a couple sentences *in English*.

(b) [easy] Write the mathematical definition of a two-sided $1-\alpha$ confidence interval below for a one-sample binomial proportion.

(c) [easy] In the notation above, the "CI" has two subscripts. What do these two subscripts mean?

(d) [difficult] The above CI in (b) does not give *exact* coverage, but only *approximate* coverage for two separate reasons. What are the two reasons? Assume the realizations are indeed sampled from $\stackrel{iid}{\sim}$ Bernoulli r.v.'s and it was a simple random sample.

(e) [easy] As in the example done in class, imagine we sampled 594 M&M's and found that 116 were blue. Compute a 95% CI for p, the true proportion of blue M&M's. Assume all of the CI assumptions are correct.

(f) [easy] Compute a 99% CI for p, the true proportion of blue M&M's using the same data as in (e). Assume all of the CI assumptions are correct. I gave you this $z_{\frac{\alpha}{2}}$ value in class.

(g) [difficult] The 99% CI is larger than the 95% CI. This is because $z_{\frac{\alpha}{2}}$ increased because we requested wider confidence which increases the margin of error (the margin of error is the $z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ term). Let's say we wanted a 99% CI with the same width as the 95% CI. How much more M&M's would we have to look at to do so?

(h) [harder] Let's say we wanted to be sure we captured p in the CI. What's the problem with just making our $1 - \alpha$ really large like 99.99999% (or α really small like 0.000001%)? Answer *in English*.

(i) [easy] For the interval you created in (e) provide two valid classical interpretations for it *in English*.

(j) [easy] For the interval you created in (e) write the non-classical interpretation *in English* — this is the interpretation you really want to be able to say...

(k) [easy] We were interested in political orientation at Queens College. We polled 100 students and 38 of them were registered democrats. Create a 95% confidence interval for the true proportion of registered democrats at Queens College.

(l) [easy] To get these 100 students, imagine we asked students who were waiting at the Q64 bus stop. Would there then be any problems with the CI we created in (k)? Answer *in English*.

(m) [harder] What would be the best way (or at least a better way) to sample 100 students in order to create the most accurate CI in (k)? Answer *in English*.

Problem 4

CI's are for inference but hypothesis testing is for assessing whether a certain theory is grounded. This will be the last topic covered on the Math 241 final. We will illustrate by using the human sex ratio example from class.

(a) [easy] We a priori assume an equal sex ratio. We'll define p as the probability of being male arbitrarily. What is the null hypothesis expressed using the mathematical notation we used in class?

(b) [easy] What would be the alternative hypothesis expressed using the mathematical notation we used in class?

(c) [harder] If $n = 200$, what is the null distribution? No symbols are acceptable as an answer, you must calculate exactly what the distribution is. Draw this null distribution how we did in class.

(d) [harder] Let's say failing to find a deviation from the hypothesized equal sex ratio meant you get fired from your boss (who believes in the unequal sex ratio theory). What kind of α would you pick for this test? Justify using a sentence *in English*.

(e) [easy] Regardless of the α chosen, choose $\alpha = 1\%$. What is the acceptance region under $\alpha = 1\%$? Denote it on the drawing in (c).

(f) [easy] What is the rejection region under $\alpha = 1\%$? Use the $\cup$ symbol to join the two sets. Denote it on the drawing in (c).

(g) [easy] What is a Type I error in general? Answer *in English*.

(h) [easy] What does a Type I error mean in this human sex ratio experiment? Answer *in English*.

(i) [easy] What is a Type II error in general? Answer *in English*.

(j) [easy] What does a Type II error mean in this human sex ratio experiment? Answer *in English*.

(k) [easy] You count there are 113 males. Calculate $\hat{p}$ and denote it on the drawing in (c).

(l) [easy] What is the result of the hypothesis test? Do you retain H_0 or reject H_0? Write a sentence in English interpreting what this means.

(m) [difficult] Now let $n = 4,247,000$ (the number of American newborns in the year 2008) and use the same $\alpha = 1\%$. You get 2,173,000 males. Write the hypotheses, find the acceptance region and find the result and interpret it.

Problem 5

This last example involved two-sided hypothesis tests of one proportion. Here, we will examine the case of firing the Uber driver where we only care about only testing one side.

The Uber driver takes 1,000 rides and then the company makes a decision about whether or not he's a good driver based on an arbitrary rule of having more than $p = 5\%$ true proportion / propensity of bad rides.

(a) [easy] Detecting him being a bad driver is sounding the alarm. Thus the status quo is that he's a good driver. Thus, $H_0 : p \leq 0.05$. What is H_a here?

(b) [harder] At $\alpha = 2.5\%$, what is the acceptance region?

(c) [easy] At $\alpha = 2.5\%$, what is the rejection region?

(d) [easy] The driver under investigation had 62 badly reviewed rides out of 1,000 rides. Calculate $\hat{p}$ and say if it's above the p cutoff Uber is using.

(e) [harder] If the driver under investigation had 62 badly reviewed rides out of 1,000 rides, would Uber fire him? Run the hypothesis test now and see. Hint: the answer to the (d) will mislead you.

(f) [difficult] You may be wondering why on Earth we defined H_a the way we did in (a). Wouldn't it seem more natural to switch the null and alternative hypothesis? Switch them here and run the test again.

(g) [difficult] Your answer changed! Why is it a bad idea to run the test like in (f) / why is it a better idea to run the test like in (e)?

(h) [E.C.] If the driver's true proportion / propensity of bad rides is 6%, what is the probability of a type II error when running the test in (e)?