

MATH 369/650 Fall 2021 Homework #2

Professor Adam Kapelner

Due by email 11:59PM, Saturday, October 2, 2021

(this document last updated Wednesday 15th September, 2021 at 11:03am)

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*. For this homework set, read about one and two sided hypothesis testing, the binomial exact test, the one proportion z test, size, level, Fisher’s p-values.

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 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. “[MA]” are for those registered for 621 and extra credit otherwise.

This homework is worth 100 points but the point distribution will not be determined until after the due date. See syllabus for the policy on late homework.

Up to 7 points are given as a bonus if the homework is typed using L^AT_EX. Links to installing L^AT_EX and program for compiling L^AT_EX is found on the syllabus. You are encouraged to use **overleaf.com**. 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. The easiest way to use overleaf is to copy the raw text from hwxx.tex and preamble.tex into two new overleaf tex files with the same name. 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. If not using L^AT_EX, print this document and write in your answers. I do not accept homeworks which are *not* on this printout. Keep this first page printed for your records.

NAME: _____

Problem 1

Here we will review theory testing from a conceptual point of view. For each question, state whether the theory under consideration should become a null hypothesis or alternative hypothesis. If null, also write the alternative; if alternative also write the null.

- (a) [easy] A new grand unified theory of physics.
- (b) [easy] The latest conspiracy theory about the president.
- (c) [easy] You are a shareholder in a pharmaceutical company. Your new drug cures cancer.
- (d) [difficult] Assume that no virus in any pandemic in history has spread due to asymptomatic transmission. The virus in the current pandemic spreads through asymptomatic transmission.

Problem 2

We will revisit the concept of a degenerate point estimator. Let $X_1, \dots, X_n \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$, our DGP and we are focused on point estimation for θ . We then choose the point estimator $\hat{\theta}_{\text{bad}} = 0.1989$.

- (a) [harder] Graph the bias of $\hat{\theta}_{\text{bad}}$ over all θ . Label your axes.

- (b) [harder] Graph the risk of $\hat{\theta}_{\text{bad}}$ under squared error loss. Label your axes.
- (c) [difficult] Compare $\hat{\theta}_{\text{bad}}$ to \bar{X} using the sup risk under squared error loss. How much better is \bar{X} ? With n getting larger does it get even better?
- (d) [difficult] Assume that θ is drawn from $\Theta = [0, 1]$ uniformly. This breaks all of our rules about how θ is one fixed unknown value. But ignore that rule. This allows you to compare both MSE curves by computing $\int_0^1 MSE(\theta)d\theta$. How much better is \bar{X} ? With n getting larger does it get even better?

For those of you who are curious, this is a metric used in Bayesian statistics called “average risk”. We won’t do any Bayesian concepts in this class. You’ll have to take 341 for that.

Problem 3

Here we will do a binomial exact test. We want to demonstrate that the iphone users in our class is greater than the national average (which is 52.4%). Recall that our data was as follows: for $n = 20$, the $\hat{\theta} = 0.65$ where the estimator we chose was the sample proportion.

- (a) [easy] Write down H_a then H_0 .
- (b) [easy] Declare your α level desired for this test. You do not need to justify it. It is what you are comfortable with.
- (c) [easy] Regardless of if you understood the previous question or not, what is the exact sampling distribution given the null hypothesis? Marked easy because you can copy from class.
- (d) [easy] Draw the PMF of the sampling distribution. Label all axes carefully and provide sufficient tick marks. Marked easy because you can copy from class.
- (e) [easy] Indicate the RET and the rejection region in the above illustration. Use your α . Everyone's answer may be different!

(f) [harder] Were you able to create a rejection region at your exact level of α ? Yes / no and why?

(g) [easy] Run the test. Write your conclusion in English.

Problem 4

Now we will test the same theory (i.e. that the iPhone percentage in our class is different than the national average of 52.4%), but use the approximate one-proportion z test to do it. Recall that our data was as follows: for $n = 20$, the $\hat{\theta} = 0.65$ where the estimator we chose was the sample proportion.

(a) [easy] Write the alternative and null hypotheses again (from lecture).

(b) [easy] Write the asymptotic distribution of our estimator under the null hypothesis which we denote $\hat{\theta} \mid H_0$. Answer in terms of standard error and round it to three decimal places. The answer is in lecture 4. Then illustrate the sampling distribution. Label the x-axis and provide tick marks on the x-axis.

- (c) [easy] What theorem did you use to get the asymptotic distribution of the estimator? State the conditions of the theorem and the theorem's result.
- (d) [easy] If we employ this asymptotic distribution of the estimator to test our theory, why is this no longer an *exact test* but instead an *approximate test*?
- (e) [harder] I want to make an apples-apples comparison with the two-sided binomial exact test from lecture. There $\alpha = 2.3\%$ so I want to use that same Type I error setting here. Compute the retainment region and rejection region (remember $\Theta = [0, 1]$) and denote these two regions in your illustration in (b). To compute these regions, I'll provide you with the following fact: $\Phi(-2.27) = 2.3\%/2 = 1.15\%$ where Φ is the CDF of the standard normal rv. (These are the kind of facts that will be provided to you on exams).
- (f) [easy] How does the retainment region compare here to the retainment region in the binomial exact test from lecture? Is the normal approximation to the binomial a good approximation in this case?
- (g) [difficult] Write about a scenario where this approximation to the exactness will provide the wrong outcome. Would practical advice does this scenario teach you?

(h) [easy] Why is this test named the *two-sided one-proportion z test*?

(i) [easy] Run the test and write your conclusion using an English sentence.

(j) [easy] What type of error could you have made?

(k) [easy] State the definition of Fisher's p-value.

(l) [harder] Find the p-value of our estimate as a function of Φ . Illustrate the p-value in the illustration in (b).

(m) [easy] Without computing the p-value explicitly, would it be above or below $\alpha = 2.3\%$?
Is the estimate *statistically significant*?

Problem 5

We will revisit the concept of a degenerate point estimator as in the first problem but this time let the DGP be $X_1, \dots, X_n \stackrel{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ where σ^2 is considered known. We are focused on point estimation for θ and the range of possible values is all numbers i.e. $\Theta = \mathbb{R}$. Of course we use $\hat{\theta} = \bar{X}$. But we also consider the “bad” point estimator $\hat{\theta}_{\text{bad}} = 1$.

- (a) [harder] Graph the bias of $\hat{\theta}_{\text{bad}}$ over all θ . Label your axes.
- (b) [harder] Graph the risk of $\hat{\theta}$ and $\hat{\theta}_{\text{bad}}$ under squared error loss. Label your axes and provide appropriate tick marks on the axes.
- (c) [difficult] Compare $\hat{\theta}_{\text{bad}}$ to $\hat{\theta}$ using the sup risk under squared error loss. How much better is \bar{X} ? With n getting larger does it get even better?
- (d) [E.C.] Assume that θ is drawn from $\Theta = \mathbb{R}$ uniformly. This breaks all of our rules about how θ is one fixed unknown value. But ignore that rule. How much better is \bar{X} ? With n getting larger does it get even better?

Problem 6

In lecture we did a *two-sided one sample z test*. Using the same data: $n = 10$ and $\bar{x} = 70$, we will now test if the population that this sample was drawn from has a *greater* mean than female height at $\alpha = 5\%$. According to an article ([click here](#)), female height is $\overset{iid}{\sim} \mathcal{N}(65, 3.5^2)$.

- (a) [easy] Write the alternative and null hypotheses. Remember our theory we want to prove is: our population mean is greater than the mean female height.
- (b) [easy] Write the distribution of our estimator under the null hypothesis which we denote $\hat{\theta} \mid H_0$. Assume the standard deviation is the same in our population as in the female height population. You will then calculate the standard error of \bar{X} . Provide your final answer in terms of standard error and round it to two decimal places. Then illustrate the sampling distribution. Label the x-axis and provide tick marks on the x-axis.
- (c) [harder] Will this test be an *exact test* or instead an *approximate test*? Explain.

- (d) [harder] Compute the retainment region and rejection region (remember $\Theta = \mathbb{R}$) and denote these two regions in your illustration in (b). To compute these regions, I'll provide you with the following fact: $\Phi(-1.645) = 5\%$.
- (e) [easy] Why is this test called a *one-sided one-sample z test*?
- (f) [easy] Run the test and write your conclusion using an English sentence.
- (g) [easy] What type of error could you have made?
- (h) [harder] Find the p-value of our estimate as a function of Φ . Illustrate the p-value in the illustration in (b).
- (i) [easy] Without computing the p-value explicitly, would it be above or below $\alpha = 5\%$? Is the estimate *statistically significant*?

Problem 7

In the previous problem we did a *one-sided one-sample z test*. We will now do a *one-sided one-sample t test* Using the same data: $n = 10$ and $\bar{x} = 70.5$, we will again test if the population that this sample was drawn from has a *greater* mean than female height at $\alpha = 5\%$. According to this article, female height is $\sim \mathcal{N}(65, 3.5^2)$. In this test, you need to compute $s = 2.07$, the sample standard deviation.

- (a) [easy] Write the alternative and null hypotheses. Same as 2(a).
- (b) [easy] Write the distribution of the *standardized* estimator under the null hypothesis. Standardized means the mean is subtracted and you divide by the standard error. In contrast to 2(b), we do *not assume* the standard deviation is the same in our population as in the female height population. Illustrate the sampling distribution. Label the x-axis. You do *not* need to provide tick marks on the x-axis as this would require using a special calculator.
- (c) [harder] Will this test be an *exact test* or instead an *approximate test*? Explain.

- (d) [harder] Compute the retainment region and rejection region (remember $\Theta = \mathbb{R}$) and denote these two regions in your illustration in (b). To compute these regions, I'll provide you with the following fact: $F_{T_9}(-1.833) = 5\%$ where F_{T_9} is the CDF of Student's T distribution with 9 degree of freedom. (These are the kind of facts that will be provided to you on exams).
- (e) [easy] Why is this test called a *one-sided one sample t test*?
- (f) [easy] Why is σ^2 called a *nuisance parameter*? Explain.
- (g) [easy] Run the test and write your conclusion using an English sentence.
- (h) [easy] What type of error could you have made?
- (i) [harder] Find the p-value of our estimate as a function of F_{t_9} . Illustrate the p-value in the illustration in (b).

- (j) [easy] Without computing the p-value explicitly, would it be above or below $\alpha = 5\%$? Is the estimate *statistically significant*?

Problem 8

In the previous problem, we needed to estimate the nuisance parameter σ^2 in order to obtain Student's T distribution for the sampling distribution of the estimator, the sample average. An intuitive estimator for σ^2 is $\hat{\sigma}^2 := \frac{1}{n} \sum_{i=1}^n (x_i - \theta)^2$.

- (a) [easy] Given only the data x_1, \dots, x_n , what is the practical problem with $\hat{\sigma}^2$ as defined by $\hat{\sigma}^2$ above?
- (b) [easy] Solving this practical problem, how would you define $\hat{\sigma}^2$?
- (c) [easy] [MA] Let $S^2 = \frac{n}{n-1} \hat{\sigma}^2$. The constant $\frac{n}{n-1}$ is known as Bessel's correction. Write the formula for S^2 only as a function of the data (the quantity n is considered a function of the data, because it is the sample size which is the data's dimensionality).
- (d) [harder] Prove S^2 is unbiased to directly without computing $\mathbb{E}[\hat{\sigma}^2]$ first (which is in the lecture). This is remarkably similar to the computation of $\mathbb{E}[\hat{\sigma}^2]$ but different enough to make you think. Somewhere in this proof, you need to use the assumption that the DGP is $\stackrel{iid}{\sim}$. Mark that location clearly.

- (e) [easy] [MA] S^2 is an estimator for σ^2 . A natural estimator for σ then is $S := \sqrt{S^2}$. Write the formula for S below.

- (f) [harder] [MA] Even though S^2 is unbiased, it turns out that S is biased due to Jensen's Inequality (which is taught in 368). Bessel's correction was a "one-size fits all" correction to the bias for sample variance. Unfortunately, there's no analogous "one-size fits all" correction to the bias for sample standard deviation; for every DGP it's different. Let the DGP be $\overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ where neither parameter are known. In 368, an extra credit problem shows that

$$\mathbb{E}[S] = \sqrt{\frac{2}{n-1}} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)} \sigma$$

where $\Gamma()$ is called the *gamma function*. Create a new estimator S_* that's unbiased for the DGP $\overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ and write its formula as a function of the data only (the quantity n is considered a function of the data, because it is the sample size which is the data's dimensionality).

This looks harder than it is! You just need to mirror the trick used in class when we derived Bessel's Correction. You don't even need to know anything about the Gamma function.

- (g) [difficult] [MA] Prove that S is asymptotically unbiased (a calculus exercise).

- (h) [E.C.] Assume squared error loss and compute $R(S^2, \sigma^2)$. Do this on a separate page.

Problem 9

We will now explore the concept of power and Type II errors. Let the DGP be $\overset{iid}{\sim}$ Bernoulli (θ).

- (a) [easy] Let $H_0 : \theta = .524$ and $H_a : \theta = .284$ and let $n = 20$. Find the approximate sampling distribution of $\hat{\theta} = \bar{X}$ using the CLT for both the null and alternative hypotheses. Leave in terms of standard error and compute to 3 decimal places.

- (b) [easy] Would this be a left-sided test, a right-sided test or a two sided test? Explain.

- (c) [easy] Illustrate the sampling distributions from the previous problem on one plot. Label the x-axis and provide tick marks.

- (d) [easy] For $\alpha = 5\%$, compute the retainment region and rejection region (remember $\Theta = [0, 1]$) and denote these two regions in your illustration in (c). To compute these regions, I'll provide you with the following fact: $\Phi(-1.645) = 5\%$.
- (e) [easy] Write the definition of power (POW).
- (f) [easy] Explain why, in general, you want POW to be large and close to 100%.
- (g) [harder] Calculate POW as a function of Φ .
- (h) [easy] Assuming H_a , in the illustration in (c), shade in the area corresponding to the integral that computes POW using vertical stripes and shade in the area corresponding to the integral that computes Type II error using horizontal stripes. Don't write anything to answer this question.

- (i) [harder] Let's do this generally. Let $H_0 : \theta \geq \theta_0$ and $H_a : \theta = \theta_a < \theta_0$. Derive the power function $POW(\theta_0, \theta_a, n, \alpha)$ which uses the Φ function.

- (j) [harder] Prove that $\lim_{n \rightarrow \infty} POW(\theta_0, \theta_a, n, \alpha) = 1$.

- (k) [harder] Assume the DGP $\overset{iid}{\sim} \mathcal{N}(\theta, \sigma^2)$ where σ^2 is known. In a right-tailed test, we derived the power function in lecture as

$$POW(\theta_0, \theta_a, n, \sigma, \alpha) = 1 - \Phi\left(-\frac{\sqrt{n}}{\sigma}(\theta_a - \theta_0) + z_{1-\alpha}\right).$$

Prove that $\lim_{\sigma \rightarrow 0} POW(\theta_0, \theta_a, n, \sigma, \alpha) = 1$.

- (l) [E.C.] Let $H_0 : \theta \geq \theta_0$ and $H_a : \theta = \theta_a < \theta_0$. Calculate the power function for the binomial exact test at size α . Do this on a separate page.