

AP Statistics – Unit 5 Additional Review

- #1. A researcher conducted a medical study to investigate whether taking a low-dose aspirin reduces the chance of developing colon cancer. As part of the study, 1,000 adult volunteers were randomly assigned to one of two groups. Half of the volunteers were assigned to the experimental group that took a low-dose aspirin each day, and the other half were assigned to the control group that took a placebo each day. At the end of six years, 15 of the people who took the low-dose aspirin had developed colon cancer and 26 of the people who took the placebo had developed colon cancer. At the significance level $\alpha = 0.05$, do the data provide convincing statistical evidence that taking a low-dose aspirin each day would reduce the chance of developing colon cancer among all people similar to the volunteers?

a)

b) Interpret the p-value

#2. Each person in a random sample of 1,026 adults in the United States was asked the following question.

“Based on what you know about the Social Security system today, what would you like Congress and the President to do during this next year?”

The response choices and the percentages selecting them are shown below.

Completely overhaul the system	19%
Make some major changes	39%
Make some minor adjustments	30%
Leave the system the way it is now	11%
No opinion	1%

(a) Find a 95% confidence interval for the proportion of all United States adults who would respond “Make some major changes” to the question. Give an interpretation of the confidence interval and give an interpretation of the confidence level.

(b) An advocate for leaving the system as it is now commented, “Based on this poll, only 39% of adults in the sample responded that they want some major changes made to the system, while 41% responded that they want only minor changes or no changes at all. Therefore, we should not change the system.” Explain why this statement, while technically correct, is misleading.

c) If we were happy with a margin of error of .04, how many adults should we sample?

#3. To increase business, the owner of a restaurant is running a promotion in which a customer's bill can be randomly selected to receive a discount. When a customer's bill is printed, a program in the cash register randomly determines whether the customer will receive a discount on the bill. The program was written to generate a discount with a probability of 0.2, that is, giving 20 percent of the bills a discount in the long run. However, the owner is concerned that the program has a mistake that results in the program not generating the intended long-run proportion of 0.2.

The owner selected a random sample of bills and found that only 15 percent of them received discounts. A confidence interval for p , the proportion of bills that will receive a discount in the long run, is 0.15 ± 0.06 . All conditions for inference were met.

(a) Consider the confidence interval 0.15 ± 0.06 .

- (i) Does the confidence interval provide convincing statistical evidence that the program is not working as intended? Justify your answer.
- (ii) Does the confidence interval provide convincing statistical evidence that the program generates the discount with a probability of 0.2? Justify your answer.

A second random sample of bills was taken that was four times the size of the original sample. In the second sample 15 percent of the bills received the discount.

- (b) Determine the value of the margin of error based on the second sample of bills that would be used to compute an interval for p with the same confidence level as that of the original interval.
- (c) Based on the margin of error in part (b) that was obtained from the second sample, what do you conclude about whether the program is working as intended? Justify your answer.

- # 4. Psychologists interested in the relationship between meditation and health conducted a study with a random sample of 28 men who live in a large retirement community. Of the men in the sample, 11 reported that they participate in daily meditation and 17 reported that they do not participate in daily meditation.

The researchers wanted to perform a hypothesis test of

$$H_0 : p_m - p_c = 0$$

$$H_a : p_m - p_c < 0,$$

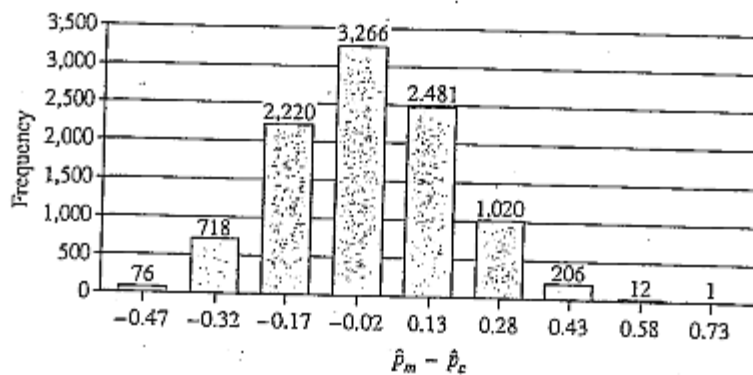
where p_m is the proportion of men with high blood pressure among all the men in the retirement community who participate in daily meditation and p_c is the proportion of men with high blood pressure among all the men in the retirement community who do not participate in daily meditation.

- (a) If the study were to provide significant evidence against H_0 in favor of H_a , would it be reasonable for the psychologists to conclude that daily meditation causes a reduction in blood pressure for men in the retirement community? Explain why or why not.

The psychologists found that of the 11 men in the study who participate in daily meditation, 0 had high blood pressure. Of the 17 men who do not participate in daily meditation, 8 had high blood pressure.

- (b) Let \hat{p}_m represent the proportion of men with high blood pressure among those in a random sample of 11 who meditate daily, and let \hat{p}_c represent the proportion of men with high blood pressure among those in a random sample of 17 who do not meditate daily. Why is it not reasonable to use a normal approximation for the sampling distribution of $\hat{p}_m - \hat{p}_c$?

Although a normal approximation cannot be used, it is possible to simulate the distribution of $\hat{p}_m - \hat{p}_c$. Under the assumption that the null hypothesis is true, 10,000 values of $\hat{p}_m - \hat{p}_c$ were simulated. The histogram below shows the results of the simulation.



- (c) Based on the results of the simulation, what can be concluded about the relationship between blood pressure and meditation among men in the retirement community?

#5. As lab partners, Sally and Betty collected data for a significance test. Both calculated the same Z statistics, but Sally found the results were significant at the 0.05 level while Betty found that the results were not. When checking their results, the women found that the only difference in their work was that Sally had used a two-sided test, while Betty used a one-sided test.

Which of the following could have been their test statistic?

- a) -1.980
- b) -1.690
- c) 1.340
- d) 1.690
- e) 1.780