

ANLY511_Quiz 1

Dr.Purna Gamage

2022-09-16

Problem: Please read this case study and answer the following questions.

Apple introduced iPhone 14 and iPhone 14 Plus last week, featuring two sizes — 6.1 inches and 6.7 inches — in a sophisticated design with impressive camera upgrades featuring a new dual-camera system and groundbreaking new safety capabilities; including Crash Detection, a smartphone industry-first safety service with Emergency SOS via satellite, and the best battery life on iPhone. (~ Apple.com)

(*This part is fictitious.*)

Assume Apple Inc. is researching for a new feature for next year's iPhone upgrade and you are one of the Data Scientists working in this research team. Your team wants to survey people in the most populous states in the US; California, Texas, Florida and New York. Suppose Your team has decided to have a big event outdoors for this research purpose. Your team randomly selects people to survey and asks them a series of questions about iPhones and new features.

Such as:

Do they have an iPhone? If they do not have an iPhone do they plan to switch to an iPhone? Do they like the new features regardless of whether or not they have an iPhone?

In New York; From those people you randomly picked, there were 355 people who are currently using an iPhone, 181 people using a different phone and don't want to switch phones ever and 238 people are currently using another phone but hoping to switch to an iPhone in the future.

Of those who are currently using an iPhone, 82% like the new feature. From those who are currently using a different phone and don't want to switch phones ever, 51% liked the new feature. From those who are currently using another phone but hoping to switch to an iPhone in the future, 49% liked the new feature.

1. (4 points) Notations.
 - a. Use capital letters to denote different events in this question. For example: I - currently uses an iPhone
 - b. Specify all probabilities mentioned in the problem using conditional probability notations and general probability notations. For example $P(I) = ??$, $P(L|I) = ??$
2. (20 points) You can see that these probabilities are only for the survey conducted in New York. But you need to easily find the probabilities for any given state. Therefore, Write a function(*using simulation*) that will find the probability that if a person liked the phone then they are currently using an iPhone. (Hint: In other words your function will calculate the probability indirectly like in the past exam question or the lab question about the flu and positive tests rather than calculating it directly with Bayes).
3. (5 points) Using the function you wrote above and using the New York data; find the probability that if a person liked the phone then they are currently using an iPhone.
4. (15 points) Write a function to Calculate the probability directly without using a simulation, use that function to compute the probability and compare your answer to part 3.

5. (20 points) Now use Monte Carlo simulation and compare your answer to the probability found in part 4. Here, write a function that can calculate the probability in part 3(the probability that if a person liked the phone then they are currently using an iPhone) for many different number of iterations (as in Monte Carlo simulation).
6. (8 points) What would be a rough minimum “large enough” number of iterations that you need to get a good probability. Use the New York Data to answer this question. (Hint: You can use a graph to answer this question).
7. (17 points) Assuming you do this survey again in California, Find the probability that this person is currently using an iPhone given that they likes the feature. See the data for California below.
 - a.(10 points) Use your function from part 5 and the function you wrote in part 2 to find the above probability.
 - b. (5 points) Use the function you wrote in part 4 to calculate the above probability without using a simulation.
 - c. (2 points) Compare these probabilities you got from the simulation with the probability you calculate by hand(Q7: part a and b). What can you conclude?

Data: From those people you randomly picked, there were 613 people who are currently using an iPhone, 579 people using a different phone and don't want to switch phones ever and 335 people are currently using another phone but hoping to switch to an iPhone in the future.

Of those who are currently using an iPhone, 93% like the new feature. From those who are currently using a different phone and don't want to switch phones ever, 47% liked the new feature. From those who are currently using another phone but hoping to switch to an iPhone in the future, 66% liked the new feature.

8. (5 points)What can you say about using Monte Carlo simulation for this problem?
9. (6 points) Based on the data from these 2 states, as the Data Scientist in this research team, what would be your final (non-technical) conclusion? (BONUS: You will get 2 bonus points for using more probabilities to answer this question)