**BUS 354**  
**Spring 2025**  
**Assignment 2 - Probability**

**DUE DATE: Friday February 28th at 11:59 PM**

**Upload your submission as one document (word or PDF)**

**Put your name at the top of your document**

Many of these questions ask you to create a figure and discuss. When you create tables, charts, and graphs you need to think about formatting. Axes should be labeled, charts should have titles, and so on. You do not have to format in one specific way but work on making these neat and easy to read. Be careful when copying and pasting documents into word (or similar programs). You may want to paste as a picture. Make sure I can see the chart/table in the document. Please ask me if you have any questions.

**Section 1: Bayes’ Theorem and Conditional Probability (22 points)**

For this set of problems, assume you are consulting with a health insurance company called “The Ninth Circle of Health”. They do swab testing for tobacco use and want to know how to interpret the results. From the CDC we know that 15% of adults use tobacco regularly, another 10% use it occasionally, and 75% are non-users. The test the company uses has the following historic conditional probabilities: 92% positive results among regular users (in probability terms: P(+ | Regular user)=0.92), 20% positive results among occasional users, and 10% positive results among non-users. The company plans to charge higher health insurance rates to regular users; occasional and non-users will be charged the same (lower) rates.

1. (4 points) Describing possible outcomes.
   * Create a two-way (contingency) table that shows the probability of the 6 distinct joint outcomes (level of use and test outcome). The “insides” of your table should be joint probabilities (intersections) and you should also show the total (marginal) probabilities for each category. Place the usage categories in the rows, and test results in the columns. A probability tree might help you organize the information.
   * Based on this, assuming those being tested are a random sample of the population, what percentage of tests should the company expect to come back positive?
2. (6 points) Assessing test results.
   * Create a new two-way table that shows the conditional probabilities of usage level given a certain test result. In other words: what percentage of those testing positive should the company expect are regular users? Occasional users? Non-users? What percentage of those testing negative are in each category?
   * Create a chart (or set of charts) in Excel that displays the probability distributions across the test results.
   * Discuss these results. How confident should the company be in charging different rates based on the test results? Remember, the company only cares about regular usage in terms of prices.
3. (6 points) Should we do more testing? The company is considering doing a second round of testing to improve the accuracy of the results. This will be expensive, so they want to know how much the additional information would help. Assume the second round of testing is independent, so the probabilities of testing positive based on type of user are the same as in the first round. The second round of testing only happens when the first test result is positive.
   * Calculate and report the conditional probabilities of user type based on testing positive twice. What is the P(regular user | 2 positive tests)? Do the same for occasional and non-users.
   * Create a chart that compares the probabilities that a person that tests positive twice is a regular, occasional, or a non-user to these probabilities for those that test positive when we only do one round of testing.
   * Based on the probabilities, do you think it would be worth it to test everyone with a positive result twice? If they do, how should they handle the results of those that test positive one out of two times?
4. (6 points) Consider the earlier example (Q1 and Q2) where we only have one round of testing. Suppose the CDC tells us that for 18 to 24-year-olds the distribution is: 10% regular users, 10% occasional users, 80% non-users. For those 55 to 64 years of age, the distribution is: 25% regular 15% occasional, 60% non-users. The conditional probabilities of testing positive based on usage type are the same as in the initial description of the situation and do not vary by age.
   * What % do you expect to test positive in each age group?
   * Compare the conditional probabilities to our test results. What is the probability of regular, occasional, and non-user given a positive test result for those that are 18 to 24? For those that are 55 to 64? Make a table to display these results.
   * Summarize the differences in the probability distributions between the two age groups. What does this mean for the company?
   * In your opinion, is it ethical for the company to consider an individual’s age when assessing how to treat test results? Discuss your reasoning.

**Section 2: Binomial Distribution** **(18 points)**

Suppose you own a romantic restaurant called “Meat and Greet” that specializes in serving fresh game. Your restaurant is very popular, only open two nights per week, and reservations are required. You have 50 tables available and are trying to decide how many reservations to accept. You know from your data over the last 5 years that 86% of parties that make reservations show up (14% no-shows). These are high-end meats we are talking about, so you make an average of $40 in profit for each table (after all costs are considered). The maximum number of parties that can be served on a night is 50. If a party shows up and there is no table for them (you are overbooked), you will give them a $60 gift certificate for the inconvenience (you lose $60).

1. (7 points) You are considering taking anywhere between 50 and 65 reservations.
   * Make a table that has each possible number of reservations taken (50, 51, …,65) in the rows.
   * Your table will have four calculated columns: (1) the expected number of guests (E(X)), (2) the probability you overbook (more than 50 show up), (3) the probability you are less than 90% full, and (4) the probability exactly 50 parties show up. Include this table in your submitted document.
   * Discuss the findings of the table. Based on this information, what number of reservations do you think is ideal?
   * Does your opinion change if you live in a relatively small area and believe that most of your customers are repeat guests? Explain.
2. (7 points). Suppose the ownership team is down to their final decision: accept either 55 or 57 reservations each night (even numbers carry the mark of the beast).
   * Make a table that shows a probability distribution for X (the number of parties that show up) for each number of reservations. This should not include all possible individual values of X (P(X=0), P(X=1), P(X=2), etc.) by themselves, but should cover the range of possible values. Group into reasonable categories so we can easily make the comparison.
   * Make a chart (column, stacked column, pie chart etc.) that visualizes the information in your table.
   * Discuss the distributions and what you feel is the better choice between these two options (accept 55 or 57 reservations).
3. (4 points). Calculate the expected profit for each of the final options (55 or 57 reservations) from question 6. This will involve making the full probability distribution for each (P(X=0), P(X=1), …) and calculating the profit for each possible outcome in the distribution. An IF function could be helpful in making the profit column.
   * Make a simple table that displays the expected profits (expected value), the probability of being overbooked, and the probability of being less than 90% full for each of the two reservation options (55 and 57). Discuss whether the optimal choice using expected profit matches what you thought was the better choice in question 6.
   * Given all the information presented in this series of questions (and your business sense), discuss how many reservations you would recommend should be accepted on a nightly basis.