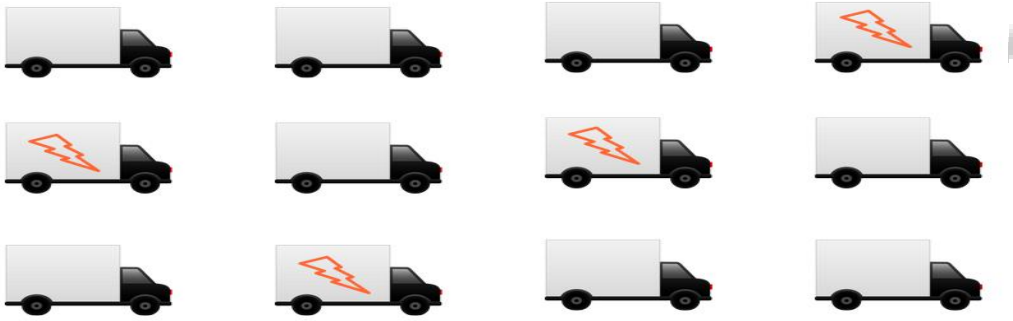
Practical Data Science/Analytics (Probability & Random Variable)

Write R Scripts or use R to perform any mathematical operations while solving the following problems.

Problem 1: Faulty brakes in Delivery trucks

A delivery company has 12 trucks, of which 4 have faulty brakes. If an inspector randomly chooses two of the trucks for brake check, what is the probability that neither one has faulty brakes?



Problem 2: Conditional Probability

A total of 500 married working couples were polled about their annual salaries, with the following information resulting. 212 couples reported both the husband and wife earned less than $25,000, 198 couples reported only the husband earned more than $25,000, 36 couples reported that only the wife earned more than $25,000, and 54 couples reported that both the husband and wife earned more than $25,000. If one of the couples is randomly chosen, what is

a. the probability that the husband earns less than $25,000?
b. the conditional probability that the wife earns more than $25,000 given that the husband earns more than this amount?
c. the conditional probability that the wife earns more than $25,000 given that the husband earns less than this amount?

Problem 3: Chain Rule

A recent college graduate is planning to take the first three data science examinations in the coming summer. She will take the first actuarial exam in June. If she passes that exam, then she with take the second exam in July, and if she also passes that one, then she will take the third exam in September. If she fails an exam, then she is not allowed to take any others. The probability that she passes the first exam is 0.9. If she passes the first exam, then the conditional probability that she passes the second one is 0.8, and if she

passes both the first and the second exams, then the conditional probability that she passes the third exam is 0.7. What is the probability that she passes all three exams?

Problem 4: Bayes Rule

a. Two dice are tossed, one green and one red. What is the conditional probability that the number on the green die is 6, given that the sum on the two dice is 7?

b. Suppose that 5% of men and 0:25% of women are color-blind. A color-blind person is chosen at random. What is the probability of this person being male? Assume that there are an equal number of males and females.

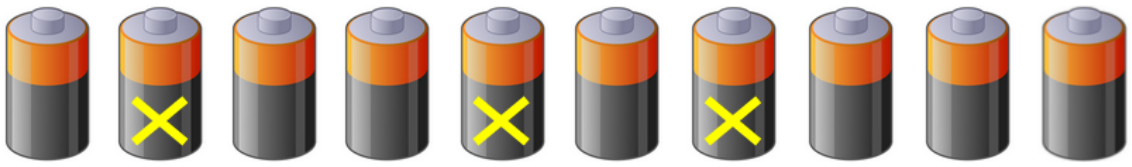
Problem 5: Rain forecast for Marie's Marriage

Marie is getting married tomorrow at an outdoor ceremony in the desert and the weatherman is trying to predict whether it will rain tomorrow or not. In recent years, it has rained only 5 days each year (5/365 = 0.014). When it actually rains, the weatherman correctly forecasts rain 90% of the time. When it doesn't rain, he incorrectly forecasts rain 10% of the time.

a) What is the probability that it will rain on Marie's wedding?
b) What is the best hypothesis weatherman can conclude?

Problem 6: P.D.F of defective batteries

Assume that you have conducted an experiment of choosing 2 batteries randomly from the following group of batteries. If a random variable X represents the number of defective batteries in the experiment, find the probability distribution function of X?



10 batteries: 7 good, 3 defective

Problem 7: Coin Toss & Fruit Fly Lifetimes

I. The expected value and variance of a coin toss are?
a.50, .50 b.50, .25 c.25, .50 d.25, .25

Practical Data Science/Analytics
(Probability & Random Variable)

II. Suppose that the probability function shown below reflects the possible lifetimes (in months after emergence) for fruit flies:

x	1	2	3	4	5	6
p(x)	0.30	?	0.20	0.15	0.10	0.05

a) What proportion of fruit flies dies in their second month?
b) What is the probability that a fruit fly lives more than four months?
c) What is the mean lifetime for a fruit fly?
d) What is the standard deviation of fruit fly lifetimes?

Problem 8: Gambling Insight

In casino, a roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet $1 that an odd number comes up, you win or lose $1 according to whether or not that event occurs.

a) Find the expected gain/loss for casino for each game?
b) What are the expected gain/loss for casino if stakes of bet are increased from $1 to $10? What will be total expected gain/loss for 10000 games after the increased stake?
c) What is the standard deviation and how do you interpret it?

Problem 9: Gold Investment

You have $10000, and a commodity like gold presently sells for $20 per gram. Suppose that after one week the commodity will sell for either $10 or $40 per gram, with these two possibilities being equally likely.

a. If your objective is to maximize the expected amount of money that you possess at the end of the next week, what strategy should you employ?
b. If you objective is to maximize the expected amount of the gold that you possess at the end of the week, what strategy should you employ?

Problem 10: Maximum Likelihood Estimation

Suppose H is a set of possible hypotheses and D is a set of training data. We would like our program to output the most probable hypothesis h from H, given the data D. Under what conditions does the following hold?

$$\arg\max_{h \in H} P(H \mid D) = \arg\max_{h \in H} P(D \mid H);$$