QR Prep

Vishal Raman

August 16, 2022

§1 Technical Portion: Trading

§1.1 Probability

Question 1.1. Given the probability of coin head is p. What is the expected number to get three heads in a row?

Proof. We use a standard Markov chain argument. Let μ_i denote the expected number of tosses to get 3 heads after the first i heads in a row. We have the system

$$\mu_0 = p(1 + \mu_1) + q(1 + \mu_0),$$
$$\mu_1 = p(1 + \mu_2) + q(1 + \mu_0),$$
$$\mu_2 = p(1) + q(1 + \mu_0).$$

Solving the system for μ_0, we obtain

$$\mu_0 = \frac{1 + p + p^2}{p^3} = \frac{1}{p^3} + \frac{1}{p^2} + \frac{1}{p}$$

Another way to solve the equation is as follows:

$$\mu_0 = \frac{1}{q}(\mu_0 + 1) + \frac{1}{pq}(\mu_0 + 2) + \frac{1}{p^2 q}(\mu_0 + 3) + \frac{1}{p^3}(3)$$

$\square$

Question 1.2. What is a martingale? List a few examples.

Question 1.3. You have r red balls, w white balls in a bag. If you keep drawing balls out of the bag until the bag now only contains balls of a single color (i.e. you run out of a color) what is the probability you run out of white balls first? (in terms of r and w).

Question 1.4. Given $\log X \sim N(0, 1)$, compute the expectation of X.

§1.2 Brain Teasers

Question 1.5. You are holding two eggs in a 100-story building. If an egg is thrown out of the window, it will not break if the floor number is less than X, and it will always break if the floor number is equal to or greater than X. What strategy would you use to determine X with the minimum number of drops in a worst case scenario?

Question 1.6. You are given two indistinguishable envelopes, each containing money. One contains twice as much as the other. You may pick one envelope and keep the money it contains. Having chosen an envelope at will, but before inspecting it, you are given the chance to switch envelopes. Should you switch?

Question 1.7. Suppose there are 10 lions and a meat. If anyone of the lions eats the meat, she falls asleep. While she is sleeping, any other lion can eat her and also fall asleep. And so on. The question is, what will happen at the beginning? Will any lion eat the meat?

Question 1.8. Given $3n$ people that the i-th person can pass a test with probability p_i, now you are required to divide them to n groups that each group has 3 people. The score of one group equals 1 if at least two people pass the test, 0 otherwise. In order to maximize the expectation of total score, how do you group them?

§2 Technical Portion: Statistics

§2.1 Covariance and Correlation

§2.2 Central Limit Theorem, Law of Large Numbers

§2.3 Hypothesis Testing

§2.4 Sampling Methods

Question 2.1. How do you sample points uniformly from a circle? How do you sample points uniformly from a sphere? Can you generalize to $\mathbb{S}^n$?

Question 2.2. How do you convert a uniform random variable to a normal random variable?

§2.5 MLE and MAP Estimation

Question 2.3. Derive an unbiased estimator for the German Tank Problem. Find the mean and variance of the estimator.

§3 Technical Portion: Regression Machine Learning

§3.1 Linear Algebra Fundamentals

Question 3.1. What is the time complexity of QR decomposition? What do you do if the dimension is too big?

$O(n^3)$.

Question 3.2. What is PCA? Implement an algorithm to solve the PCA problem for a given dataset.

§3.2 Linear Regression

Question 3.3. What is the difference between a linear regression of X on Y and of Y on X?

Question 3.4. What happens to the estimators of linear regression if you double the dataset?

Question 3.5. What is ridge regression? What is LASSO? When would you use one over the other?

Question 3.6. Find a closed form solution to the simple linear regression optimization problem.

Question 3.7. Find a closed form solution to the ridge regression optimization problem.

§3.3 Machine Learning

Question 3.8. What is a entropy in information theory? What is the significance of a bit?

Question 3.9. Implement the k-Nearest Neighbors Classification algorithm.

Question 3.10. What happens if you initialize the weights of a neural network to zero?

Question 3.11. Using stock data from last 5 years, how can we build a model to predict the price next day?

Question 3.12. What is the Bias-Variance Tradeoff? Derive the Bias-Variance decomposition.

§4 Technical Portion: Programming

§4.1 C++ Questions

Question 4.1. What is a virtual function? How do you use it?

Question 4.2. Describe all the C++ casts and their properties.

§4.2 Linear Regression in Python

§4.3 Algorithm Questions

Question 4.3.
- Given a series of real numbers, what's the optimal point in real line that minimize the distance between all these numbers?

- What's the best algorithm to calculate the median of a sequence of numbers?

- Design a algorithm able to update the median of sequence as there are new numbers coming in.

Question 4.4. Given an array of numbers: 1,5,3,2,4,2,3,1,2,3,2,4, return another array with duplicate numbers in consecutive positions and in order of their first appearance, in the example the result should be: 1,1,5,3,3,3,2,2,2,2,4,4.

Question 4.5. Given an array, find out three numbers such that their product is maximum.

Question 4.6. Find the rolling window correlation between two vectors $X, Y \in \mathbb{R}^n$.

Question 4.7. Given a list of stock prices indexed over time, find the way to maximize profit.

Question 4.8. Design a trading exchange in C++. Which data structures would you use to implement methods and what are their time complexity?