

1 Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

2 Quadratic Regression

In this question, we will find the best quadratic estimator of Y given X . First, some notation: let μ_i be the i th moment of X , i.e. $\mu_i = \mathbf{E}[X^i]$. Also, define $\beta_1 = \mathbf{E}[XY]$ and $\beta_2 = \mathbf{E}[X^2Y]$. For simplicity, we will assume that $\mathbf{E}[X] = \mathbf{E}[Y] = 0$ and $\mathbf{E}[X^2] = \mathbf{E}[Y^2] = 1$. (Note that this poses no loss of generality, because we can always transform the random variables by subtracting their means and dividing by their standard deviations.) We claim that the best quadratic estimator of Y given X is

$$\hat{Y} = \frac{1}{\mu_3^2 - \mu_4 + 1} (aX^2 + bX + c)$$

where

$$\begin{aligned} a &= \mu_3\beta_1 - \beta_2, \\ b &= (1 - \mu_4)\beta_1 + \mu_3\beta_2, \\ c &= -\mu_3\beta_1 + \beta_2. \end{aligned}$$

Your task is to prove the Projection Property for \hat{Y} .

(a) Prove that $\mathbf{E}[Y - \hat{Y}] = 0$.

(b) Prove that $\mathbf{E}[(Y - \hat{Y})X] = 0$.

(c) Prove that $\mathbf{E}[(Y - \hat{Y})X^2] = 0$.

Any quadratic function of X is a linear combination of 1, X , and X^2 . Hence, these equations together imply that $Y - \hat{Y}$ is orthogonal to any quadratic function of X , and so \hat{Y} is the best quadratic estimator of Y .

3 Projection Property

Use the Projection Property to answer the following questions.

(a) Prove or disprove: for any function ϕ , $\mathbf{E}[\mathbf{E}[Y | X]\phi(X)] = 0$.

(b) Prove or disprove: $\mathbf{E}[(Y - \mathbf{E}[Y | X])L[Y | X]] = 0$.

(c) Prove the following: $\mathbf{E}[X^2 | Y] = \mathbf{E}[(X - \mathbf{E}[X | Y])^2 | Y] + \mathbf{E}[X | Y]^2$. [*Hint*: In the expression $\mathbf{E}[X^2 | Y]$, try replacing X with $(X - \mathbf{E}[X | Y]) + \mathbf{E}[X | Y]$.]

(d) We have already shown that $\mathbf{E}[\mathbf{E}[Y | X]] = \mathbf{E}[Y]$. Prove that $\mathbf{E}[L[Y | X]] = \mathbf{E}[Y]$.

(e) Prove the following property of conditional expectation:

$$\mathbf{E}[\mathbf{E}[Z | X, Y] | X] = \mathbf{E}[Z | X].$$

[*Hint*: Take a closer look at the method by which we prove properties of conditional expectation in Note 26.]

4 Balls in Bins Estimation

We throw $n > 0$ balls into $m \geq 2$ bins. Let X and Y represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbf{E}[Y | X]$. [*Hint*: Your intuition may be more useful than formal calculations.]

(b) What are $L[Y | X]$ and $Q[Y | X]$ (where $Q[Y | X]$ is the best quadratic estimator of Y given X)? [*Hint*: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the MMSE.]

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbf{E}[X]$ and $\mathbf{E}[Y]$.

(d) Compute $\text{var}(X)$.

(e) Compute $\text{cov}(X, Y)$.

(f) Compute $L[Y | X]$ using the formula. Ensure that your answer is the same as your answer to part (b).

5 Swimsuit Season

In the swimsuit industry, it is well-known that there is a “swimsuit season”. During this time, swimsuit sales skyrocket!

We will model this with a random variable X which is either λ_L or λ_H with equal probability; λ_L represents the mean number of customers in a day when swimsuits are not in season, and λ_H represents the mean number of customers during swimsuit season. So, λ_L is the “low rate” and λ_H is the “high rate”. The number of customer arrivals Y on a particular day is modeled as a Poisson random variable with mean X .

You observe Y customers on a certain day, and the task is to estimate X .

- (a) What is $L[X | Y]$?
- (b) What is $\mathbf{E}[X | Y]$?

6 Political War

Initially, there are d Democrats and r Republicans in a room. They begin to argue. On each day, a random person in the room leaves and returns with an additional member of his or her political party; that is, either a Democrat will leave and return with a Democrat friend, or a Republican will leave and return with a Republican friend. Let D_n denote the number of democrats in the room at the end of the n th day. Let $D_0 = d$.

- (a) Find $\mathbf{E}[D_n | D_{n-1}]$.
- (b) Find $\mathbf{E}[D_n]$ using the law of iterated expectation.
- (c) What is the expected fraction of Democrats in the room at the end of day n ?

7 Optimal Gambling

In even-money gambling games, you bet a fixed amount of money. If you win the game, you are given back the money that you bet, and you receive an additional amount of money equal to your original bet. If you lose the game, you lose the amount of money you bet.

- (a) You are gambling and your probability of winning, on each round, is $1/2 < p < 1$: the game is in your favor! You use the following strategy: on each round, you will bet a fraction q of the money you have at the start of the round. Let X_n denote the amount of money you have on round n . X_0 represents your initial assets and is a constant value. What is $\mathbf{E}[X_n]$?
- (b) What value of q will maximize $\mathbf{E}[X_n]$? For this value of q , what is the distribution of X_n ? Can you predict what will happen as $n \rightarrow \infty$? [*Hint*: Under this betting strategy, what happens if you ever lose a round?]

- (c) The problem with the previous approach is that we were too concerned about expected value, so our gambling strategy was too extreme. Let's start over: again we will use a gambling strategy in which we bet a fraction q of our money at each round. Express X_n in terms of n , q , X_0 , and W_n , where W_n is the number of rounds you have won up until round n . [Hint: Does the order in which you win the games affect your profit?]
- (d) By the law of large numbers, $W_n/n \rightarrow p$ as $n \rightarrow \infty$. Using this fact, what does $(\log X_n)/n$ converge to as $n \rightarrow \infty$?
- (e) The rationale behind $(\log X_n)/n$ is that if $(\log X_n)/n \rightarrow c$, where c is a constant, then that means for large n , X_n is roughly e^{cn} . Therefore, c is the asymptotic growth rate of your fortune! Find the value of q that maximizes your asymptotic growth rate.
- (f) Using the value of q you found in the previous part, compute $\mathbf{E}[X_n]$.