

Andrew Nalundasan
OMSBA 5067, Seattle University
Week 1 Assignment
April 11, 2021

"I have not received unauthorized aid on this assignment. I understand the answers that I have submitted. The answers submitted have not been directly copied from another source, but instead are written in my own words."

Question 1:

A box contains three cards. One card is red on both sides, one card is green on both sides, and one card is red on one side and green on the other. We randomly select one card from this box, and we can know the color of the selected card's upper side. If this side is green, what is the probability that the other side of the card is also green? (10 points)

| Card | Top | Bottom |
|------|-------|--------|
| 1 | Red | Red |
| 2 | Green | Green |
| 3 | Red | Green |

nobs = 6

$P(\text{Top}=\text{Green and Bottom}=\text{Green}) = 1/3$

$P(\text{Bottom}=\text{Green}) = 1/2$

$$(1/3) / (1/2) = 2/3$$

Question 2:

Suppose that X is a random variable. If $E[X] = \mu$ and $\text{Var}(X) = \sigma^2$, then what is the value $E[X(X - 1)]$?

$$E[X(X - 1)]$$

$$E[X^2 - X]$$

$$E[X^2] - E[X]$$

$$\sigma^2 = E[X^2] - (E[X])^2$$

$$E[X^2] = \sigma^2 + (E[X])^2 = \sigma^2 + \mu^2$$

$$(\sigma + \mu)^2 - \mu$$

$$\sigma^2 + \mu^2 - \mu$$

$$X = \sigma^2 + \mu(\mu - 1)$$

Question 3:

Suppose that the PDF of a random variable X is as follows

a) What is the value of c? (5 points)

$$\int_1^2 Cx^2 dx = \frac{1}{3}Cx^3 = 1/3C(2)^3 - 1/3C(1)^3 = \frac{8}{3C} - \frac{1}{3C} = \frac{7}{3C}$$

$$1 = \frac{7}{3}C$$

$$\frac{3}{7} = C$$

b) What is the probability of X being larger than 1.5? (5 points)

$$\int_{1.5}^2 Cx^2 dx = \frac{1}{3}Cx^3 = 1/3C(2)^3 - 1/3C(1.5)^3 = \frac{8}{3}C - \frac{27}{24}C = \frac{64}{24}C - \frac{27}{24}C = \frac{37}{24}C$$

$$\frac{37}{24} * \frac{3}{7} = 0.66$$

Question 4: We have used a machine learning algorithm for a binary classification problem and the test results are as follows:

- The total number of test data points is 200
- There are 95 negative data points and 105 positive data points in the test data
- Out of 95 negative data points, 55 are predicted as negative and the rest are predicted as positive
- Total number of predicted positive data points is 140

Draw the confusion matrix and calculate the accuracy, precision, and recall. (15 points)

Confusion Matrix

| | | Predicted Class | | |
|--------------|-------------|-----------------|------------|-----|
| | | Class = Yes | Class = No | |
| Actual Class | Class = Yes | 100 | 5 | 105 |
| | Class = No | 40 | 55 | 95 |
| Total | | 140 | | |

Accuracy:

$$\frac{TP + TN}{TP + TN + FP + FN}$$

$$\frac{100 + 55}{100 + 55 + 40 + 5} = \frac{155}{200}$$

$$\text{Accuracy} = 0.775$$

Precision:

$$\frac{TP}{TP + FP}$$

$$\frac{100}{100 + 400} = \frac{100}{500}$$

$$\text{Precision} = 0.25$$

Recall:

$$\frac{TP}{TP + FN}$$

$$\frac{100}{100 + 5} = \frac{100}{105}$$

$$\text{Recall} = 0.952$$