Random Variables – Solutions

COR1-GB.1305 – Statistics and Data Analysis

Bayes' Rule

1. Every year in March there is a standardized exam for people who want to be licensed sheep herders. It happens that, with probability 0.45, a person will pass this exam. In the process of screening people, it turns out that among those who passed the exam, 60% had taken college courses in biology. It happens also that 30% of all those who take the exam had college courses in biology. Find the probability that a person with college courses in biology will pass the exam.

> **Solution:** The information in the problem is
>
> $$\begin{aligned} \text{P(Pass)} &= .45 \\ \text{P(Bio)} &= .30 \\ \text{P(Bio} \mid \text{Pass)} &= .60 \end{aligned}$$
>
> The problem is asking us to compute the quantity P(Pass | Bio). Using Bayes' rule,
>
> $$\begin{aligned} \text{P(Pass} \mid \text{Bio)} &= \text{P(Bio} \mid \text{Pass)} \cdot \frac{\text{P(Pass)}}{\text{P(Bio)}} \\ &= (.60) \cdot \frac{(.45)}{(.30)} \\ &= .90. \end{aligned}$$
>
> That is, there is a 90% chance that a person with college courses in biology will pass the exam.

2. Amazon.com maintains a list of all registered customers, along with their email addresses. During July, they sent coupons to 20% of their customers. They recorded that 5% of their customers made purchases in July, and 40% of all purchases were made with coupons. In this problem we will compute the proportion of customers sent a coupon in July who made a purchase in that month. For simplicity, we will assume that customers either make 0 or 1 purchases in July.

   (a) Consider a random customer, and define two events:

   $$\begin{aligned} \text{Coupon} &= \text{the customer received a coupon in July,} \\ \text{Purchase} &= \text{the customer made a purchase in July.} \end{aligned}$$

   Express all percentages given in the problem statement as probabilities or conditional probabilities of these two events. Example: P(Coupon) = 0.20.

Solution: The information in the problem is

$$\begin{aligned} P(\text{Coupon}) &= .20 \\ P(\text{Purchase}) &= .05 \\ P(\text{Coupon} \mid \text{Purchase}) &= .40 \end{aligned}$$

(b) Use Bayes' rule to compute the proportion of custoers sent a coupon in July who made a purchase that month.

Solution: The problem is asking us to compute the quantity P(Purchase | Coupon). Using Bayes' rule,

$$\begin{aligned} P(\text{Purchase} \mid \text{Coupon}) &= P(\text{Coupon} \mid \text{Purchase}) \cdot \frac{P(\text{Purchase})}{P(\text{Coupon})} \\ &= (.40) \cdot \frac{(.05)}{(.20)} \\ &= .10. \end{aligned}$$

Probability Distribution Function and Expectation

3. Consider the following game:

   1. You pay \$6 to flip a coin.
   2. If the coin lands heads, you get \$10; otherwise, you get nothing.

   (a) Would you play this game? Why or why not?

   Solution: It will usually be beneficial to play the game when our expected winnings are positive. In this situation, if we play the game many times, then in the long run we will make a profit.

   (b) What is the random experiment involved in the game? What are the sample space? What are the probabilities of the sample points?

   Solution: The random experiment is the coin flip. The sample points for the coin flip are H and T; each of these has probability $\frac{1}{2}$.

   (c) Let W be the random variable equal to the amount of money you win from playing the game. If you lose money, W will be negative. Find the value of W for each of the sample points.

   Solution:
   The values of the random variable corresponding to the sample points are as follow:

Outcome	W
H	4
T	-6

   (d) Describe W in terms of its probability distribution function (PDF).

   Solution: The PDF of W is given by the table:

w	-6	4
$p(w)$	0.5	0.5

   (e) What are your expected winnings? That is, what is μ, the expectation of W?

   Solution:
   Using the PDF computed in part (c), the expected value of W is

$$\begin{aligned}\mu &= (.5)(-6) + (.5)(4) \\ &= -1.\end{aligned}$$

On averate, we lose $1 every time we play the game.

4. Suppose you flip two coins. Let X be the random variable which counts the number of heads on the two tosses.

(a) List all of the sample points of the experiment, along with the corresponding values of X.

Solution:

Outcome	X
HH	2
HT	1
TH	1
TT	0

(b) Compute the probability distribution function of X.

Solution:

x	0	1	2
$p(x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

(c) Compute the expectation of X.

Solution: Using the PDF we computed in part (b), the expectation of X is

$$\begin{aligned} \mathrm{E}(X) &= (\tfrac{1}{4})(0) + (\tfrac{2}{4})(1) + (\tfrac{1}{4})(2) \\ &= 1. \end{aligned}$$

(d) What is the interpretation of the expectation of X?

Solution: If we were to repeat the experiment many times, getting a different value of X each time, then the average value of X will be close to 1.

5. Let X be a random variable describing the number of cups of coffee a randomly-chosen member of the class drinks on a typical day. There is a 22% chance that the student has one cup, a 16% chance that the student has two cups, a 16% chance that the student has three cups, an 11% chance that the student has four cups, and a 3% chance that the student has five cups. Also, there is a 32% chance that the student doesn't drink any coffee.

(a) Let $p(x)$ be the probability distribution function of X. Fill in the following table:

x	0	1	2	3	4	5
$p(x)$						

Solution:

x	0	1	2	3	4	5
$p(x)$	.32	.22	.16	.16	.11	.03

(b) Find $\mathrm{E}(X)$, the expectation of X.

Solution:

$$\mathrm{E}(X) = (.32)(0) + (.22)(1) + (.16)(2) + (.16)(3) + (.11)(4) + (.03)(5) \quad = 1.61.$$

(c) What is the interpretation of the expectation of X?

Solution: The long-run sample mean. If we performed the random experiment upon which X is measured many times, getting a different value of X each time, then the sample mean would be very close to the expectation of X. In particular, if every day we sample a student from the class and measure how many cups of coffee they drink, then after awhile, the average number of cups of coffee will be close to the expectation (1.61).

Variance and Standard Deviation

6. This is a continuation of problem 5.

   (a) Find var(X) and sd(X), the variance and standard deviation of X, the number of cups of coffee that a random student from the class drinks on a typical day.

   Solution:

   $$\begin{aligned}\text{var}(X) &= (.32)(0-1.61)^2 + (.22)(1-1.61)^2 + (.16)(2-1.61)^2 + (.16)(3-1.61)^2 \\ &\quad + (.11)(4-1.61)^2 + (.03)(5-1.61)^2 \\ &= 2.2179\end{aligned}$$

   The standard deviation of X is given by

   $$\text{sd}(X) = \sqrt{\text{var}(X)} = \sqrt{2.2179} = 1.48.$$

   (b) What is the interpretation of the standard deviation of X?

   Solution: The long-run sample standard deviation. If we performed the random experiment upon which X is measured many times, getting a different value of X each time, then the sample standard deviation would be very close to the standard deviation of X. In particular, if every day we sample a student from the class and measure how many cups of coffee they drink, then after awhile, the sample standard deviation of the number of cups of coffee will be close to the 1.48.

7. Consider the following game:

   1. You pay \$6 to pick a card from a standard 52-card deck.
   2. If the card is a diamond ($\diamondsuit$), you get \$22; if the card is a heart ($\heartsuit$), you get \$6; otherwise, you get nothing.

   Perform the following calculations to decide whether or not you would play this game.

   (a) Let W be the random variable equal to the amount of money you win from playing the game. If you lose money, W will be negative. Find the PDF of W.

   Solution: The sample points corresponding to the suit of the card are ♠, $\heartsuit$, ♣, and $\diamondsuit$; each of these has probability $\frac{1}{4}$. The values of the random variable W corresponding to the sample points are as follow:

Outcome	W
♠	-6
$\heartsuit$	0
♣	-6
$\diamondsuit$	16

Thus, the PDF of W is given by the table:

w	-6	0	16
$p(w)$	0.50	0.25	0.25

(b) What are your expected winnings? That is, what is μ, the expectation of W?

Solution:
Using the PDF computed in part (a), the expected value of W is

$$\begin{aligned}\mu &= (.50)(-6) + (.25)(0) + (.25)(16) \\ &= 1.\end{aligned}$$

On average, we win \$1 every time we play the game.

(c) What is the standard deviation of W?

Solution:
Using the PDF computed in part (a), and the expected value computed in part (b), we compute the variance of W as

$$\begin{aligned}\sigma^2 &= (.50)(-6-1)^2 + (.25)(0-1)^2 + (.25)(16-1)^2 \\ &= 81.\end{aligned}$$

Thus, the standard deviation of W is

$$\sigma = \sqrt{81} = 9.$$

(d) What are the interpretations of the expectation and standard deviation of W?

Solution: If we played the game many many times, then the average of our winnings over all times we played would be close to the \$1, and the standard deviations of our winnings over all times we played would be close to \$9.