Bayes' Rule / Discrete Random Variables – Solutions

STAT-UB.0103 – Statistics for Business Control and Regression Models

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
>
> $$\begin{aligned} \mathrm{P}(\mathrm{Pass}) &= .45 \\ \mathrm{P}(\mathrm{Bio}) &= .30 \\ \mathrm{P}(\mathrm{Bio} \mid \mathrm{Pass}) &= .60 \end{aligned}$$
>
> The problem is asking us to compute the quantity $\mathrm{P}(\mathrm{Pass} \mid \mathrm{Bio})$. Using Bayes' rule,
>
> $$\begin{aligned} \mathrm{P}(\mathrm{Pass} \mid \mathrm{Bio}) &= \mathrm{P}(\mathrm{Bio} \mid \mathrm{Pass}) \cdot \frac{\mathrm{P}(\mathrm{Pass})}{\mathrm{P}(\mathrm{Bio})} \\ &= (.60) \cdot \frac{(.45)}{(.30)} \\ &= .90. \end{aligned}$$
>
> 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: $\mathrm{P}(\mathrm{Coupon}) = 0.20$.

Solution: The information in the problem is

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

(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,

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

Lie Detector

3. Through accounting procedures, it is known that 10% of the employees of a store are stealing. To find out who is stealing, the manager decides to make all employees to take a lie detector test. The lie detector is accurate 80% of the time: if an employee is a thief, then he or she will fail the test with probability 0.8; if an employee is honest, then he or she will pass the test with probability 0.8. In this activity we will simulate the results of the manager's investigation.

   (a) Use your smartphone (or your neighbor's smartphone) to go to `http://random.org`. Click the "Generate" button to draw a random number between 1 and 100. Write down your number. Everyone in the class should generate his or her own number. If your number is in the range 1–10, write "Thief"; if your number is in the range 11–100, write "Honest".

   (b) Click "Generate" again to generate a new random number. Write down the number. If the number is in the range 1–80, then lie detector gives the correct answer ("Fail" for thief, "Pass" for honest). If the number is in the range 81–100, then the lie detector gives the wrong answer and records "Pass" for thief and "Fail" for honest. Write down the result of the test.

Solution: In class, we computed the probability that someone who fails the test is a thief. The problem tells us that $\mathrm{P}(\text{Thief}) = 0.1$ and $\mathrm{P}(\text{Fail} \mid \text{Thief}) = 0.8$. To use Bayes' rule, we need to compute $\mathrm{P}(\text{Fail})$. This is given by:

$$\mathrm{P}(\text{Fail}) = (0.1)(0.8) + (0.9)(0.2) = 0.26;$$

to compute this probability, we have noted that to Fail, an employee is either a Thief and they Fail, or they are Honest and they Fail. Thus,

$$\begin{aligned} \mathrm{P}(\text{Thief} \mid \text{Fail}) &= \mathrm{P}(\text{Fail} \mid \text{Thief}) \frac{\mathrm{P}(\text{Thief})}{\mathrm{P}(\text{Fail})} \\ &= (0.80) \frac{(0.10)}{(0.26)} \\ &= 0.31. \end{aligned}$$

Thus, approximately 31% of the people who fail the lie detector test are thieves.

Discrete random variables: PDF & Expectation

4. 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

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

$$\boxed{\text{On averate, we lose \$1 every time we play the game.}}$$

5. 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.