I N T R O D U C T I O N T O P R O B A B I L I T Y

by

Dimitri P. Bertsekas and John N. Tsitsiklis

C H A P T E R 2 : A D D I T I O N A L P R O B L E M S

SECTION 2.2. Probability Mass Functions

Problem 1. The probability of a royal flush in poker is $p = 1/649,740$. Show that approximately $649,740$ hands would have to be dealt in order that the probability of getting at least one royal flush is above $1 - 1/e$.

Problem 2. The annual premium of a special kind of insurance starts at \$1000 and is reduced by 10% after each year where no claim has been filed. The probability that a claim is filed in a given year is 0.05, independently of preceding years. What is the PMF of the total premium paid up to and including the year when the first claim is filed?

SECTION 2.3. Functions of Random Variables

Problem 3. Let X be a discrete random variable that is uniformly distributed over the set of integers in the range $[a, b]$, where a and b are integers with $a < 0 < b$. Find the PMF of the random variables $\max\{0, X\}$ and $\min\{0, X\}$.

Problem 4. Let X be a discrete random variable, and let $Y = |X|$.

(a) Assume that the PMF of X is

$$p_X(x) = \begin{cases} Kx^2 & \text{if } x = -3, -2, -1, 0, 1, 2, 3, \\ 0 & \text{otherwise,} \end{cases}$$

where K is a suitable constant. Determine the value of K.

(b) For the PMF of X given in part (a) calculate the PMF of Y.

(c) Give a general formula for the PMF of Y in terms of the PMF of X.

SECTION 2.4. Expectation, Mean, and Variance

Problem 5. Let X be a random variable that takes integer values and is symmetric, that is, $\mathbf{P}(X = k) = \mathbf{P}(X = -k)$ for all integers k. What is the expected value of $Y = \cos(X\pi)$ and $Y = \sin(X\pi)$?

Problem 6. You are visiting the rainforrest, but unfortunately your insect repellent has run out. As a result, at each second, a mosquito lands on your neck with probability

0.5. If a mosquito lands, it will bite you with probability 0.2, and it will never bother you with probability 0.8, independently of other mosquitoes. What is the expected time between successive bites?

Problem 7. Fischer and Spassky play a sudden-death chess match whereby the first player to win a game wins the match. Each game is won by Fischer with probability p, by Spassky with probability q, and is a draw with probability $1 - p - q$.

(a) What is the probability that Fischer wins the match?

(b) What is the PMF, the mean, and the variance of the duration of the match?

Problem 8. A particular binary data transmission and reception device is prone to some error when receiving data. Suppose that each bit is read correctly with probability p. Find a value of p such that when 10,000 bits are received, the expected number of errors is at most 10.

Problem 9. Imagine a TV game show where each contestant i spins an infinitely calibrated wheel of fortune, which assigns him/her with some real number with a value between 1 and 100. All values are equally likely and the value obtained by each contestant is independent of the value obtained by any other contestant.

(a) Find $\mathbf{P}(X_1 < X_2)$.

(b) Find $\mathbf{P}(X_1 < X_2, X_1 < X_3)$.

(c) Let N be the integer-valued random variable whose value is the index of the first contestant who is assigned a smaller number than contestant 1. As an illustration, if contestant 1 obtains a smaller value than contestants 2 and 3 but contestant 4 has a smaller value than contestant 1 ($X_4 < X_1$), then $N = 4$. Find $\mathbf{P}(N > n)$ as a function of n.

(d) Find $\mathbf{E}[N]$, assuming an infinite number of contestants.

Problem 10. Let N be a nonnegative integer-valued random variable. Show that

$$\mathbf{E}[N] = \sum_{i=1}^{\infty} \mathbf{P}(N \geq i).$$

Problem 11. Let $X_1, \ldots, X_n$ be independent, identically distributed random variables with common mean and variance. Find the values of c and d that will make the following formula true:

$$\mathbf{E}\big[(X_1 + \cdots + X_n)^2\big] = c\mathbf{E}[X_1]^2 + d\big(\mathbf{E}[X_1]\big)^2.$$

SECTION 2.5. Joint PMFs of Multiple Random Variables

Problem 12. The MIT football team wins any one game with probability p, and loses it with probability $1 - p$. Its performance in each game is independent of its performance in other games. Let L_1 be the number of losses before its first win, and let L_2 be the number of losses after its first win and before its second win. Find the joint PMF of L_1 and L_2.

Problem 13. A class of n students takes a test in which each student gets an A with probability p, a B with probability q, and a grade below B with probability $1 - p - q$, independently of any other student. If X and Y are the numbers of students that get an A and a B, respectively, calculate the joint PMF $p_{X,Y}$.

SECTION 2.6. Conditioning

Problem 14. Your probability class has 250 undergraduate students and 50 graduate students. The probability of an undergraduate (or graduate) student getting an A is 1/3 (or 1/2, respectively). Let X be the number of students that get an A in your class.

(a) Calculate $\mathbf{E}[X]$ by first finding the PMF of X.

(b) Calculate $\mathbf{E}[X]$ by viewing X as a sum of random variables, whose mean is easily calculated.

Problem 15. A scalper is considering buying tickets for a particular game. The price of the tickets is \$75, and the scalper will sell them at \$150. However, if she can't sell them at \$150, she won't sell them at all. Given that the demand for tickets is a binomial random variable with parameters $n = 10$ and $p = 1/2$, how many tickets should she buy in order to maximize her expected profit?

Problem 16. Suppose that X and Y are independent discrete random variables with the same geometric PMF:

$$p_X(k) = p_Y(k) = p(1-p)^{k-1}, \qquad k = 1, 2, \ldots,$$

where p is a scalar with $0 < p < 1$. Show that for any integer $n \geq 2$, the conditional PMF

$$\mathbf{P}(X = k \,|\, X + Y = n)$$

is uniform.

Problem 17. Let X, Y, and Z be independent geometric random variables with the same PMF:

$$p_X(k) = p_Y(k) = p_Z(k) = p(1-p)^{k-1},$$

where p is a scalar with $0 < p < 1$. Find $\mathbf{P}(X = k \,|\, X + Y + Z = n)$. *Hint:* Try thinking in terms of coin tosses.

Problem 18. Alvin shops for probability books for K hours, where K is a random variable that is equally likely to be 1, 2, 3, or 4. The number of books N that he buys is random and depends on how long he shops according to the conditional PMF

$$p_{N\,|\,K}(n \,|\, k) = \frac{1}{k}, \qquad \text{for } n = 1, \ldots, k.$$

(a) Find the joint PMF of K and N.

(b) Find the marginal PMF of N.

(c) Find the conditional PMF of K given that $N = 2$.

(d) Find the conditional mean and variance of K, given that he bought at least 2 but no more than 3 books.

(e) The cost of each book is a random variable with mean \$30. What is the expected value of his total expenditure? *Hint:* Condition on the events $\{N = 1\}, \ldots, \{N = 4\}$, and use the total expectation theorem.

SECTION 2.7. Independence

Problem 19. At his workplace, the first thing Oscar does every morning is to go to the supply room and pick up one, two, or three pens with equal probability 1/3. If he picks up three pens, he does not return to the supply room again that day. If he picks up one or two pens, he will make one additional trip to the supply room, where he again will pick up one, two, or three pens with equal probability 1/3. (The number of pens taken in one trip will not affect the number of pens taken in any other trip.) Calculate the following:

(a) The probability that Oscar gets a total of three pens on any particular day.

(b) The conditional probability that he visited the supply room twice on a given day, given that it is a day in which he got a total of three pens.

(c) $\mathbf{E}[N]$ and $\mathbf{E}[N \,|\, C]$, where $\mathbf{E}[N]$ is the unconditional expectation of N, the total number of pens Oscar gets on any given day, and $\mathbf{E}[N \,|\, C]$ is the conditional expectation of N given the event $C = \{N > 3\}$.

(d) $\sigma_{N \,|\, C}$, the conditional standard deviation of the total number of pens Oscar gets on a particular day, where N and C are as in part (c).

(e) The probability that he gets more than three pens on each of the next 16 days.

(f) The conditional standard deviation of the total number of pens he gets in the next 16 days given that he gets more than three pens on each of those days.

Problem 20. Your computer has been acting very strangely lately, and you suspect that it might have a virus on it. Unfortunately, all 12 of the different virus detection programs you own are outdated. You know that if your computer does have a virus, each of the programs, independently of the others, has a 0.8 chance of believing that your computer as infected, and a 0.2 chance of thinking your computer is fine. On the other hand, if your computer does not have a virus, each program has a 0.9 chance of believing that your computer is fine, and a 0.1 chance of wrongly thinking your computer is infected. Given that your computer has a 0.65 chance of being infected with some virus, and given that you will believe your virus protection programs only if 9 or more of them agree, find the probability that your detection programs will lead you to the right answer.

Problem 21. Joe Lucky plays the lottery on any given week with probability p, independently of whether he played on any other week. Each time he plays, he has a probability q of winning, again independently of everything else. During a fixed time period of n weeks, let X be the number of weeks that he played the lottery and Y be the number of weeks that he won.

(a) What is the probability that he played the lottery on any particular week, given that he did not win on that week?

(b) Find the conditional PMF $p_{Y \mid X}(y \mid x)$.

(c) Find the joint PMF $p_{X,Y}(x, y)$.

(d) Find the marginal PMF $p_Y(y)$. *Hint:* One possibility is to start with the answer to part (c), but the algebra can be messy. But if you think intuitively about the procedure that generates Y, you may be able to guess the answer.

(e) Find the conditional PMF $p_{X \mid Y}(x \mid y)$. Do this algebraically using the preceding answers.

(f) Rederive the answer to part (e) by thinking as follows: for each one of the $n - Y$ weeks that he did not win, the answer to part (a) should tell you something.