COT 3100 Fall 2017 Homework 8 Solutions

1) In alpha testing a new software package, a software engineer finds that the number of defects per 100 lines of code is a random variable X with probability distribution:

x	1	2	3	4
Pr(X = x)	0.3	0.2	0.1	0.4

Find E(X) and Var(X).

Solution

$E(X) = 1(.3) + 2(.2) + 3(.1) + 4(.4) = \mathbf{2.6}$

$E(X^2) = 1(.3) + 4(.2) + 9(.1) + 16(.4) = 8.4$

$Var(X) = E(X^2) - [E(X)]^2 = 8.4 - 2.6^2 = \mathbf{1.64}$

2) You decide to start a new lottery. All winning combinations are 5 values chosen from the set {1, 2, 3, ..., 80}. When a player buys a ticket, she can buy a Pick-6, Pick-7 or Pick-8. As the names indicate, if you buy a Pick-k, you choose k numbers out of the 80. You win if each of the numbers from the winning combination are included in your choice of k numbers. The payout for winning is $100,000. How much should you be willing to pay for each ticket (Pick-6, Pick-7 and Pick-8) if you wish to break even (based on expected value)?

Solution

The probability of winning with a Pick-k where 5 ≤ k ≤ 80 is $\binom{k}{5} \Big/ \binom{80}{5}$. Thus, to get the desired value that you would want to pay so that on average, you break even, we would want the expected winnings to equal the price of the ticket. The expected winnings are this probability times $100,000. Thus, we have the following values one should be willing to pay to break even:

Willing to pay for pick-6 = $\$100000 \times \binom{6}{5} \Big/ \binom{80}{5} = \mathbf{2.50\ cents}$

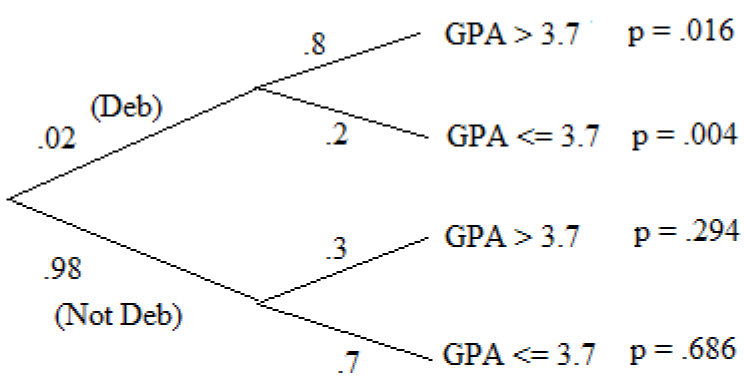
Willing to pay for pick-7 = $\$100000 \times \binom{7}{5} \Big/ \binom{80}{5} = \mathbf{8.74\ cents}$

Willing to pay for pick-7 = $\$100000 \times \binom{8}{5} \Big/ \binom{80}{5} = \mathbf{23.3\ cents}$

3) Two percent of students at a school are on the debate team. Given that a student is on the debate team at the school, the probability that her GPA is above 3.7 is .8. Given that a student is NOT on the debate team at the school, the probability that her GPA is above 3.7 is .3. Given that a student's GPA is above 3.7, what is the probability that she is on the debate team?

Solution

Draw the usual tree diagram:



The probability a student's GPA is above 3.7 is 0.016 + .294 = .31. Given that this is the case, the probability she is on the debate team is $\frac{.016}{.31} = \mathbf{5.2\%}$.

4) Sam's probability of getting A's on precisely 2 tests out of 4 is $\frac{8}{27}$. Assuming that Sam's probability of getting an A on each individual test is the same and independent from the other tests, and that her chance of getting an A on an individual test is greater than 50%, what is the probability that Sam gets an A on an individual test?

Solution

Let p be the probability of Sam getting an A on an individual test. Via the Binomial probability distribution, the chance that he gets As on exactly 2 tests out of 4 is $\binom{4}{2} p^2(1-p)^2$. Now, let's set this to $\frac{8}{27}$:

$$\binom{4}{2} p^2(1-p)^2 = \frac{8}{27}$$
$$(p(1-p))^2 = \frac{8}{162} = \frac{4}{81}$$

Taking the square root of both sides and noting that probabilities can't be negative we find:

$$p(1-p) = \frac{2}{9}$$
$$9p(1-p) = 2$$
$$9p - 9p^2 = 2$$

$$9p^2 - 9p + 2 = 0$$
$$(3p-2)(3p-1) = 0$$

It follows that $p = \frac{1}{3}$ or $p = \frac{2}{3}$. Since we are given that $p > \frac{1}{2}$, it follows that $p = \mathbf{\frac{2}{3}}$.

5) In Arup's Game of Dice 3, you roll a fair pair of six-sided dice and record the total. If this total is 3, 5 or 12, you win. If it's a 2 or 11, you lose. In all other cases, you roll the pair of dice again. If the sum of this second roll equals or exceeds the sum of your first roll, you win! Otherwise you lose. (For example, if you roll a 6 followed by a 6, you win, but if you roll a 10 followed by a 9, you lose.) What is the probability of winning Arup's Game of Dice?

<u>Solution</u>
Using the known dice roll probabilities (shown in class), create the following table showing the probability of each roll (p) and the probability that an independent roll of a pair of dice equals or beats that value (q):

1st roll	2	3	4	5	6	7	8	9	10	11	12
p	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36
q	36/36	35/36	33/36	30/36	26/36	21/36	15/36	10/36	6/36	3/36	1/36

Note: Each entry in the bottom row is just the sum of all entries on the second row from that column on forward. (For example, the 21/36 in column 7 corresponds to 6/36 + 5/36 + 4/36 + 3/36 + 2/36 + 1/36 for each term in row 2 in column 7 or later.)

To solve the question, separate out all winning opportunities and add up these disjoint probabilities:

$$p(W) = \frac{2}{36} + \frac{4}{36} + \frac{1}{36} + \frac{3}{36} \times \frac{33}{36} + \frac{5}{36} \times \frac{26}{36} + \frac{6}{36} \times \frac{21}{36} + \frac{5}{36} \times \frac{15}{36} + \frac{4}{36} \times \frac{10}{36} + \frac{3}{36} \times \frac{6}{36}$$

$$p(W) = \frac{72 + 144 + 36 + 99 + 130 + 126 + 75 + 40 + 18}{36 \times 36} = \frac{740}{1296} = \mathbf{\frac{185}{324}} \sim \mathbf{.571}$$

6) Suppose we roll a fair 4 sided die with the numbers [1,4] written on them. After the first die roll we roll the die k times where k is the number on the first die roll. The number of points you score is the sum of the face-values on all die rolls (including the first). What is the expected number of points you will score?

Solution

Let $E_r(k)$ be the expected score if you roll k dice. Suppose we have k dice. Let X_i be the discrete random variable representing the i^{th} die's score. By linearity of expectations $E(X_1 + X_2 + X_3 + X_4) = E(X_1) + E(X_2) + E(X_3) + E(X_4)$. As the die doesn't change we can calculate $E(X_i)$ using the standard formula for expected value: $E(X_i) = 0.25(1) + 0.25(2) + 0.25(3) + 0.25(4) = \frac{10}{4}$. Now we can calculate $E_r(k) = k \cdot \frac{10}{4}$.

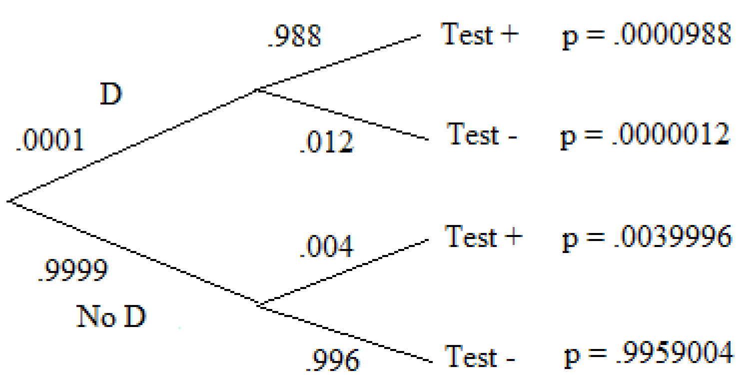
Let X be a discrete random variable representing the expected score total. We can calculate $E(X)$ be calculating each expected value after knowing the outcome of the first die. This value is $E(X) = 0.25\big(1 + E_r(1)\big) + 0.25\big(2 + E_r(2)\big) + 0.25\big(3 + E_r(3)\big) + 0.25\big(4 + E_r(4)\big)$

$$= 0.25\left(10 + \frac{10}{4}(1 + 2 + 3 + 4)\right) = 0.25\left(10\left(1 + \frac{10}{4}\right)\right) = \frac{35}{4} = \mathbf{8.75}$$

7) Suppose that one person in 10,000 people has a rare genetic disease. There is an excellent test for the disease; 98.8% of the people with the disease test positive and only 0.4% of the people who don't have it test positive. What is the probability that someone who tests positive has the disease? What is the probability that someone who tests negative does not have the disease?

Solution

Here is the tree diagram of the situation:



The probability of testing positive is .0039996 + .0000988 = .0040984. It follows that the probability you have the disease given you've tested positive is .0000988/.0040984 ~**2.4%.**

The probability of testing negative is .0000012 + .9959004 = .9959016. It follows that the probability you don't have the disease given that you tested negative is .9959004/.9959016 ~ **.0099%**

8) Consider the following algorithm to determine if an array, arr, of size n (indexes 0 through n-1, inclusive) is sorted from smallest to largest:

1) Repeat k times:
    i) Choose a random number, i, in between 0 and n-2, inclusive.
    ii) If arr[i] > arr[i+1], then answer that the array is out of order.
2) Answer that the array is in order.

Consider running this algorithm on the following array of size 10:

i	0	1	2	3	4	5	6	7	8	9
arr[i]	17	8	9	26	29	13	14	18	19	16

If we choose to run the algorithm with k = 5, what is the probability that the algorithm erroneously tells us that the array is sorted? Leave your answer as a fraction in lowest terms.

Solution

There are three locations out of nine where consecutive items are out of order in this array, indexes 0 and 1, indexes 4 and 5, and indexes 8 and 9. Thus, we have a $\frac{3}{9} = \frac{1}{3}$ chance on a single repetition of discovering that this array is not sorted and a $\frac{2}{3}$ chance on a single repetition that it is not discovered as being out of order. For the algorithm to fail with k = 5, the latter would have to happen 5 times in a row. Assuming independence of the random number chosen in step 1i, the probability that this algorithm tells us the array is sorted is $(\frac{2}{3})^5 = \mathbf{\frac{32}{243}}$.

9) Suppose E and F are events in a sample space and p(E) = 3/4, p(F) = 4/5, and p(F | E) =7/8. Find p(E | F).

Solution

$$\frac{7}{8} = p(F|E) = \frac{p(E \cap F)}{p(E)} = \frac{p(E \cap F)}{\frac{3}{4}}$$

$$p(E \cap F) = \frac{3}{4} \times \frac{7}{8} = \frac{21}{32}$$

Now, solve for $p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{\frac{21}{32}}{\frac{4}{5}} = \mathbf{\frac{105}{128}}$

10) This question deals with buying and using a pair of thumb drives.

(a) The probability a thumb drive you buy is defective is .002. Assume that the probability one thumb drive is defective does not affect the probability of another one being defective. Given that you've bought two thumb drives, what is the probability that both are defective?

(b) Luckily, both of the thumb drives you've bought work! Each has 1 GB of memory available for you to save files. You have four files you would like to save on the two thumb drives of the following sizes: .7 GB, .4 GB, .3 GB and .3 GB. Unfortunately, you've left your nimwit brother to copy the files from your desktop to the two thumb drives. For each of the four files, he randomly chooses one of the two thumb drives to copy the file, not checking if the copy was successful. (Assume a copy successful as long as the requisite space is available.) What is the probability that all four files get copied successfully? Leave your answer as a fraction in lowest terms.

Solution

(a) Since both events are independent, we can multiply the probabilities of each thumb drive being defective to obtain .002 x .002 = **.000004.**

(b) Label the four files as follows: A (.7 GB), B (.4 GB), C (.3 GB) and D (.3 GB)

Without loss of generality, assume that file A gets copied to the first thumb drive. Then we have 8 possible arrangements for the files on the first thumb drive, each equally likely:

1) A
2) A, B
3) A, C
4) A, D
5) A, B, C
6) A, B, D
7) A, C, D
8) A, B, C, D

Naturally, the second thumb drive will always contain the rest of the files not listed for the files on the first thumb drive. Note that the sizes of B, C and D are 1.0 GB. Thus, we are guaranteed that the second thumb drive will have its files properly copied. Thus, we simply need to count how many of these 8 arrangements are valid. The only valid arrangements are #1 (.7 GB), #3 (1.0 GB) and #4 (1.0 GB). Thus the corresponding probability is $\mathbf{\frac{3}{8}}$.

11) Give a summary of the life and mathematical contributions of Andrew Wiles. Please aim for a length of roughly 200 - 400 words. **Your summary must be typed.** Please state the sources you used in writing your summary.

Sample Summary

Andrew Wiles is a mathematician who was born in England in 1953, most noted for his proof of Fermat's Last Theorem. At the age of 10, Wiles is said to have read a book describing the theorem and was fascinated with it, since it was so simply stated yet eluded proof for so long. He soon realized that the problem was too difficult for him, but returned to it many years later after completing his Ph.D. Wiles's Ph.D. was on elliptic curves with complex multiplication and the direction of work on Fermat's Last Theorem in the early 80s heavily involved elliptic curves. Many contemporary mathematician's stayed away from the problem, even though several believed that a proof could be built upon the work of mathematicians Gerhard Frey, Jean-Pierre Serre and Ken Ribet. Wiles, after earning his Ph. D. from Clare College in Cambridge, England, came to the United States to work at the Institute for Advanced Study and then Princeton University. It is at Princeton when Wiles started working on Fermat's Last Theorem in near complete isolation, for over six years. When he emerged, he lectured for three days on his work over the past six years at a conference in Cambridge. His initial proof had a flaw, which took over a year to fix, but with the help of one of his former students, Richard Taylor, Wiles was able to successfully complete a proof of Fermat's Last Theorem. Recently, Wiles moved from Princeton to the University of Oxford in England where he is a research professor in mathematics.

Note: Fermat's Last Theorem is that there are no positive integer solutions (for a, b and c) to the equation

$$a^n + b^n = c^n$$

for any integers $n > 2$.

Source: https://en.wikipedia.org/wiki/Andrew_Wiles