COT 3100 Homework #9 Solutions (by Ivette Carreras)
Probability

1) (5 pts) 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)	0.4	0.3	0.2	0.1

Find (a) E(X); and (b) Var(X)

<u>Solution</u>
Use the equation for expected value $E(X) = \sum_{x \in X} x * p(x)$ and the equation for variance $Var(X) = \sum_{x \in X}\left(x - E(X)\right)^2 * p(x)$ and plug in the values given in the table.

a) $E(X) = \sum_{x=1}^{4} x * p(x)$

$E(X) = 1(0.4) + 2(0.3) + 3(0.2) + 4(0.1)$
$E(X) = 2$ **(2 pts)**

b) $Var(X) = \sum_{x=1}^{4}\left(x - E(X)\right)^2 * p(x)$

$$Var(X) = \sum_{x=1}^{4} (x-2)^2 * p(x)$$

$Var(X) = 1(0.4) + (0)(0.3) + 1(0.2) + 4(0.1)$

$Var(X) = 1.0$ **(3 pts)**

2) (7 pts) A lottery allows a player to choose 5 values out of 40. The goal of the lottery is to break even exactly, making its cash prizes equal to the amount of money spent by the contestants. Each ticket costs $1 to buy and players receive winnings if they match 3, 4 or all 5 numbers. If the payout for matching all 5 numbers is $250,000 and the payout for matching 4 numbers is $1000, how much should the payout for matching exactly 3 numbers be, to the nearest penny?

<u>Solution</u>
Since we know that the lottery breaks even all the money must be taken by the winners (3, 4, and 5 numbers correct). The payment for 5 and 4 winning numbers are known, $250,000 and $1000. The expected value for each lottery ticket is $1. So we can use the expected value equation to find the third value of x (the payment for getting 3 of the numbers).

p(n) -> denotes the probability of getting n numbers correct in the lottery **(4 pts)**

n	3	4	5
p(n)	$\frac{\binom{5}{3}*\binom{35}{2}}{\binom{40}{5}}$	$\frac{\binom{5}{4}*\binom{35}{1}}{\binom{40}{5}}$	$\frac{1}{\binom{40}{5}}$

Let's call the unknown payments y.

$$E(X) = (250{,}000) * p(1) + (1000) * p(4) + y * p(3)$$
$$E(X) = 1$$
$$1 = (250000) * \left(\frac{1}{658008}\right) + (1000) * \left(\frac{175}{658008}\right) + y * \left(\frac{5950}{658008}\right)$$

$$1 = (0.37993459) + (0.265954) + y * (0.009042)$$

$$1 - 0.6458859 = y * (0.009042)$$
$$y = \$39.16 \textbf{ (3 pts)}$$

3) (5 pts) The probability that it rains during a summer's day in a certain town is 0.2. In this town, the probability that the daily maximum temperature exceeds 25 degrees Celsius is 0.3 when it rains and 0.6 when it does not rain. Given that the maximum daily temperature exceeded 25 degrees Celsius on a particular summer's day, find the probability that it rained on that day.

Solution

First, determine the events: let A be the event that it rains, and B the event that the temperature exceeds 25 degrees Celsius. From the problem we know:
p(A) = 0.2
p(B|A) = 0.3
p(B|~A) = 0.6

We want to find the probability of A given B, p(A|B). Using Bayes theorem we have:
$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$
Since we don't know p(B) we must calculate it, use the information given:
$$p(B) = p(A) * p(B|A) + p(\sim A) * p(B|\sim A)$$
$$p(B) = (0.2) * (0.3) + (1 - 0.2) * (0.6)$$
$$p(B) = 0.54 \textbf{ (3 pts)}$$
Now we can solve for p(A|B)
$$p(A|B) = \frac{0.3 * 0.2}{0.54}$$
$$p(A|B) = \frac{1}{9} \textbf{ (2 pts)}$$

4) (4 pts) When John throws a stone at a target, the probability that he hits the target is 0.4 He throws a stone 6 times. (a) Find the probability that he hits the target exactly 4 times. (b) Find the probability that he hits the target for the first time on his third throw.

Solution

Let A be the event that John hits the target; p(A) = 0.4.

a) Since we don't care the order in which John hits the target the four times, we can simply use the Binomial probability distribution. n is the total number of trials, k is the number of successes we want, and p is the probability of one success. In this case n = 6 throws, and k = 4 the number of times we want to successfully hit the target.

$$\binom{n}{k} p^k (1-p)^{n-k}$$ **(2 pts)**

$$= \binom{6}{4} p^4 (1-p)^{6-4}$$

$$= \binom{6}{4} (0.4)^4 (0.6)^{6-4}$$

$$= 1\,5 * 0.0256 * 0.36$$

$$= 0.13824$$ **(1 pt)**

b) In this case we want the first 2 throws to be a miss (1-p), and the third to be a hit (p). This is simply: $(1-p)^2 * p$

$= (0.6)^2(0.4) = 0.144$ **(1 pt)**

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

Solution

In this game we can either win in the first round or the second round. The possible pairs to win in the first round are (5, 6), (6, 5), and (6, 6). The probability of each of these is $\frac{1}{6} * \frac{1}{6} = \frac{1}{36}$, and add up to $\frac{3}{36}$. **(2 pts)**
In order to move on to the second round we must roll a sum between 3 and 10 inclusive. And to win we must roll a sum greater than the previous (or not smaller or equal to the previous). We need to consider the probability of every possible value from 3 to 10, and for each we need to consider the probability of getting a sum smaller or equal to that particular previous sum.

$$\sum_{i=3}^{10} p(i) * (\,1 - \sum_{j=2}^{i} p(j))$$

Since we don't know the probability for every value between 3 and 10, we need to calculate it.
p(2) = 1/36
p(3) = 2/36

p(4) = 3/36
p(5) = 4/36
p(6) = 5/36
p(7) = 6/36
p(8) = 5/36
p(9) = 4/36
p(10) = 3/36

$$\sum_{i=3}^{10}(\,p(i) * (\,1 - \sum_{j=2}^{i} p(j)))$$

$$= \frac{2}{36} \times \frac{33}{36} + \frac{3}{36} \times \frac{30}{36} + \frac{4}{36} \times \frac{26}{36} + \frac{5}{36} \times \frac{21}{26} + \frac{6}{36} \times \frac{15}{36} + \frac{5}{36} \times \frac{10}{36} + \frac{4}{36} \times \frac{6}{36} + \frac{3}{36} \times \frac{3}{36} = \frac{538}{1296}$$

Finally, we add this to $\frac{3}{36}$ to get our final answer: $\frac{3}{36} + \frac{538}{1296} = \frac{646}{1296} = \frac{323}{648} \sim .4985$ **(8 pts, 1 per term)**

Note: Due to the tedium of this sum, it's acceptable to write a program to calculate it.

6) (5 pts) There are 792 cards in a Topps Set of baseball cards. Imagine buying a pack of 40 of these cards, each of which are randomly chosen. What is the probability that at least two copies of some card will be in the pack?

Solution

Finding the probability of having no repeated cards in a pack is much simpler than the original question, and we can use the result to calculate the answer to the original question. Now the problem is just choosing one card with no repetition 40 times.

$$\prod_{k=0}^{39} \frac{792-k}{792}$$ **(4 pts)**

$$= 0.3673$$

The answer we are looking for is having at least two cards repeated. So the answer is 1 – 0.3673

$$= 0.6327$$ **(1 pt)**

Note: You can use the posted bday.c and put in the appropriate constants to get this numerical value.

7) (7 pts) Suppose that one person in 10,000 people has a rare genetic disease. There is an excellent test for the disease; 99.9% of the people with the disease test positive and only 0.02% 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

First, determine the events: let A be the event that someone has the disease, and B the event that someone's test is positive. From the problem we know:

p(A) = 1/10,000

p(B|A) = 0.999

p(B|~A) = 0.0002

What is the probability that someone who tests positive has the disease? We want to find the probability of A given B, p(A|B). Using Bayes theorem we have:

$$p(A|B) = \frac{p(B|A) * p(A)}{p(B)}$$

$$p(B) = p(A) * p(B|A) + p(\sim A) * p(B|\sim A)$$

$$p(B) = (\frac{1}{10000}) * (0.999) + (1 - \frac{1}{10000}) * (0.0002)$$

$$p(B) = 0.00029988$$

Now we can solve for p(A|B)

$$p(A|B) = \frac{0.999 * 0.0001}{0.00029988}$$

$p(A|B) = 0.333$ **(4 pts)**

What is the probability that someone who tests negative does not have the disease? We want to find the probability of ~A given ~B, p(~A|~B). Using the definition of conditional probability we have:

$$p(\sim A|\sim B) = \frac{p(\sim A \cap \sim B)}{p(\sim B)}$$

Note: $p(\sim B) = p(\sim A \cap \sim B) + p(A \cap \sim B) = (.9999)(.9998) + (.0001)(.001) = .99970012$

Thus, we have $p(\sim A|\sim B) = \frac{p(\sim A \cap \sim B)}{p(\sim B)} = \frac{.9999 \times .9998}{.99970012} = .9999999$ **(3 pts)**

As we can see here, someone who tests negative in this particular instance can be quite sure that they don't have the disease.

8) (7 pts) Suppose we flip a fair coin until it either comes up tails twice in a row of we have flipped it seven times. What is the expected number of times we flip the coin?

We must find what the probability is of getting two tails in a row for each time we flip the coin. The number of flips is represented by n, and the probability of having tails twice in a row is p(n). We will consider stopping before the seventh time a success.

Two tails are only possible starting with n = 2, since we need two tails. T represents tails and H represents head in the following table. Staring at n = 3 we know that the last 3 flips must be HTT in order to have a success

n	**Possible coin flip sequences**	**p(n)**
2	TT	**1/4**
3	HTT	**1/8**
4	(T/H)HTT	**2/16 = 1/8**
5	(HH/HT/TH)HTT	**3/32**
6	(HHH/HHT/HTH/THH/THT)HTT	**5/64**
7	rest	**1-1/4-1/8-1/8-3/32-5/64 = 21/64**

Now that we have the probabilities of each possible X value, use the equation for expected value $E(X) = \sum_{x \in X} x * p(x)$

a) $E(X) = \sum_{n=2}^{7} n * p(n)$

$$E(X) = 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{8}\right) + 4\left(\frac{1}{8}\right) + 5\left(\frac{3}{32}\right) + 6\left(\frac{5}{64}\right) + 7\left(\frac{21}{64}\right) = \frac{32 + 24 + 32 + 30 + 30 + 147}{64} = \frac{295}{64}$$

$$E(X) = 4.609375$$

Grading: 1 pt for each row of the chart, 1 pt for the final calculation.