Homework 6 Solutions

Due Monday, August 4 at 11:59pm

CS 70: Discrete Mathematics and Probability Theory, Summer 2014

1. Pandemic! (25 points)

   You are creating a test for a zombism, a disease only 1 in 1000 people have. If a person has the disease, there is a 95% chance that your test will be positive. But if the person does not have the disease, there is only an 85% chance the test will be negative.

   (a) Let D be the event that you have the disease, and H be the event that you are healthy. Let A be the event that the test comes out positive, and B be the event that it comes out negative. Write an expression for $\Pr[D \mid A]$ (the probability you have the disease given a positive test result) in terms of the probabilities given above. Plug in the given probabilities into your expression and calculate the numerical value of $\Pr[D \mid A]$.

   Answer: We know from Bayes rule that

   $$P(D|A) = \frac{P(D)P(A|D)}{P(A)}.$$

   We know that $P(D) = 0.001$ and $P(A|D) = 0.95$, but it is unclear how to compute $P(A)$. However, we do know that the probability of a positive test result should be the probability of a positive test result and disease, plus the probability of a positive test result and healthy. This gives us

   $$P(A) = P(A, D) + P(A, H) = P(D)P(A|D) + P(H)P(A|H).$$

   which we have expanded above using Bayes rule. Now we have $P(A)$ in term of quantities we know, and can write

   $$P(D|A) = \frac{P(D)P(A|D)}{P(D)P(A|D) + P(H)P(A|H)} = \frac{0.001 \cdot 0.95}{0.001 \cdot 0.95 + 0.999 \cdot 0.15} = 0.62997\%,$$

   so even given a positive test result, we get a very very small probability of disease.

   (b) Write an expression for $\Pr[H \mid B]$, the probability you are healthy given a negative test result. Evaluate the numerical result of this expression as well.

   Answer: Similarly, we get

   $$P(H|B) = \frac{P(H)P(B|H)}{P(B)} = \frac{P(H)P(B|H)}{P(H)P(B|H) + P(D)P(B|D)} = \frac{0.999 \cdot 0.85}{0.999 \cdot 0.85 + 0.001 \cdot 0.05} = 99.9941\%$$

(c) For your test to gain approval, the chance of disease given a positive test result must be above 90% (we don't want to quarentine any normal humans). What would the accuracy of the test have to be to ensure this result? (You may now assume the accuracy of the test is the same whether you have the disease or not.)

Answer: Solve for x:

$$P(D|A) = \frac{0.001 \cdot x}{0.001 \cdot x + 0.999 \cdot (1-x)} = 0.9 \quad \Rightarrow \quad x = 99.9889\%$$

2. Holey Socks! (15 points)

A drawer contains 10 socks, where 6 of them have holes and 4 of them do not. Suppose you pull two random socks out of the drawer, look at them, and then put them back. If you do this 5 times, what is the probability that you pull out a pair with no holes precisely 4 out of 5 times?

Answer: This is just binomial! Each trial, we have a chance to pull out a pair of socks with no holes. That probability is $\frac{\binom{4}{2}}{\binom{10}{2}} = \frac{2}{15}$. Thus, $n = 5$ and $p = \frac{2}{15}$, and

$$Pr[\text{4 pairs have no holes}] = \binom{5}{4}(\frac{2}{15})^4(\frac{13}{15})$$

3. Kids (25 points)

Mr. and Mrs. Tripathi decide to continue having children until they either have their first girl or until they have five children. Assume that each child is equally likely to be a boy or a girl, independent of all other children, and that there are no multiple births. Let B and G denote the numbers of boys and girls (respectively) that the Tripathis have.

(a) Write down the sample space together with the probability of each outcome.

(b) Write down the distributions of the random variables B and G.

(c) Compute the expectations of B and of G using a direct calculation.

Solution:

(a) Let b mean they had a boy and g mean they had a girl.

Outcome	Pr
g	$\frac{1}{2}$
bg	$\frac{1}{4}$
bbg	$\frac{1}{8}$
$bbbg$	$\frac{1}{16}$
$bbbbg$	$\frac{1}{32}$
$bbbbb$	$\frac{1}{32}$

(b) Looking at the table in (a.), we get the following.

i	Pr[$B=i$]	Pr[$G=i$]
0	$\frac{1}{2}$	$\frac{1}{32}$
1	$\frac{1}{4}$	$\frac{31}{32}$
2	$\frac{1}{8}$	0
3	$\frac{1}{16}$	0
4	$\frac{1}{32}$	0
5	$\frac{1}{32}$	0

(c) $E[B] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{16} + 4 \cdot \frac{1}{32} + 5 \cdot \frac{1}{32} = \frac{31}{32}$. $E[G] = 0 \cdot \frac{1}{32} + 1 \cdot \frac{31}{32} = \frac{31}{32}$. So oddly enough, the expected number of boys the get is the same as the expected number of girls.

4. Monkies, Buildings, Coins, and Arcades (35 points)

   Solve each of the following problems using linearity of expectation. Clearly explain your methods.

   (a) A monkey types at a 26-letter keyboard with one key corresponding to each of the lower-case English letters. Each keystroke is chosen independently and uniformly at random from the 26 possibilities. If the monkey types 1 million letters, what is the expected number of times the sequence "book" appears?

   (b) A building has n floors numbered $1, 2, \ldots, n$, plus a ground floor G. At the ground floor, m people get on the elevator together, and each gets off at a uniformly random one of the n floors (independently of everybody else). What is the expected number of floors the elevator stops at (not counting the ground floor)?

   (c) A coin with Heads probability p is flipped n times. A "run" is a maximal sequence of consecutive flips that are all the same. (Thus, for example, the sequence HTHHHTTH with $n = 8$ has five runs.) Show that the expected number of runs is $1+2(n-1)p(1-p)$. Justify your calculation carefully.

   (d) In an arcade, you play game A 10 times and game B 20 times. Each time you play game A, you win with probability 1/3 (independently of the other times), and if you win you get 3 tickets (redeemable for prizes), and if you lose you get 0 tickets. Game B is similar, but you win with probability 1/5, and if you win you get 4 tickets. What is the expected total number of tickets you receive?

Solution:

(a) There are $1,000,000-4+1 = 999,997$ places where "book" can appear, each with a (non-independent) probability of $\frac{1}{26^4}$ of happening. If A is the random variable that tells how many times "book" appears, and A_i is the indicator variable that is 1 if "book" appears starting at the i^{th} letter, then $E[A] = E[A_1 + \cdots + A_{999,996}] = E[A_1] + \cdots + E[A_{999,997}] = \boxed{\mathbf{\frac{999,997}{26^4}}} \approx 2.19$ times.

(b) Let A_i be the indicator that the elevator stopped at floor i. $\Pr[A_i = 1] = 1 - \Pr[\text{no one gets off at } i] = 1 - \left(\frac{n-1}{n}\right)^m$. If A is the number of floors the elevator stops at, them $E[A] = E[A_1 + \cdots + A_n] = E[A_1] + \cdots + E[A_n] = \boxed{\mathbf{n \cdot \left(1 - \left(\frac{n-1}{n}\right)^m\right)}}$

(c) Let A_i be the indicator for the event that a run starts at the i toss. Let $A = A_1 + \cdots + A_n$ be the random variable for the number of runs total. $E[A_1] = 1$ obviously. For $i \neq 1$, $E[A_i] = \Pr[A_i = 1] = \Pr[i = H \mid i-1 = T] \cdot \Pr[i-1 = T] + \Pr[i = T \mid i-1 = H] \cdot \Pr[i-1 = H] = p \cdot (1-p) + (1-p) \cdot p = 2p \cdot (1-p)$. This gives $E[A] = E[A_1 + A_2 + \cdots + A_n] = E[A_1] + E[A_2] + \cdots + E[A_n] = 1 + 2(n-1)p(1-p)$.

(d) Let A_i be the indicator you win the i^{th} time you play game A and B_i be the same for game B. Let T_A be the random variable for the number of tickets you win in game A, and T_B be the number of tickets you win in game B. $E[T_A + T_B] = 3E[A_1] + \cdots + 3E[A_{10}] + 4E[B_1] + \cdots + 4E[B_{20}] = 10 + \left(\frac{4}{5}\right) \cdot 20 = \boxed{\mathbf{26}}$