COMP 182 HW 8

Quan Le

Thursday, April 23, 2020

Events and Their Probabilities [55 pts]

1. What is the probability of the following events when we randomly select a permutation of the 26 lowercase letters of the English alphabet? Show your work.

   (a) The first 13 letters of the permutation are in alphabetical order.

   If we select permutations at random, each permutation has the same probability of being selected, i.e., 1/26!. Now, we simply count the number of permutations such that the first 13 are in alphabetical order. This is equivalent to the number of combinations of choosing the first 13 letters, multiplied by the number of permutations of the remaining 13. Therefore, the probability is $\binom{26}{13} * 13!/26! = 1/13! \approx \boxed{1.6059 \times 10^{-10}}$

   (b) a is the first letter of the permutation and z is the last letter.

   Again, we need only to calculate the number of permutations of the remaining 24 letters. Therefore, our resusltant probability is $24!/26! = 1/(25 * 26) \approx \boxed{1.5385 \times 10^{-3}}$

   (c) a and z are next to each other in the permutation.

   Let us combine a and z as a singular letter to be permuted. Then, there are 25! ways to do so, and 2 ways to order a and z. Therefore the probability is $2 * 25!/26! = 1/13 \approx \boxed{0.076923}$

   (d) a and b are not next to each other in the permutation.

   Let us order the 25 letters not including b. The, we can place b in 24 distinct spaces between two letters, where b is not next to a. Therefore, there are $24 * 25!/26! = 12/13 \approx \boxed{0.92308}$. Note that the same result can be obtained by taking the complement of the previous answer.

   (e) a and z are separated by at least 23 letters in the permutation.

Let us consider the number of ways to permute 23 letters that are not a or z, i.e., $24!/(24-23)! = 24!$. For the final letter, there are 3 possible locations to insert it, on either end of a, z, or between them. Finally, we can still flip a, z. Therefore, the probability of such a case is $6 * 24!/26! = 3/(25 * 13) \approx \boxed{0.0092307}$.

(f) z precedes both a and b in the permutation.

For a given permutation of letters, there are 3!=6 ways to permute a, b, z. For 2 cases, z precedes a, b. Therefore, the probability is $2/6 = \boxed{1/3}$

2. Assume that all days of the week are equally likely. Show your work in each of the following parts.

(a) What is the probability that two people chosen at random were born on the same day of the week?

Since there are 7 days of the week, and we choose the days without replacement, there are 7 cases where they are born on the same day, out of 49. Thus, the probability they have the same birth-day-of-week is $\boxed{1/7}$

(b) What is the probability that in a group of n people chosen at random, there are at least two born on the same day of the week?

Let us consider the complement, that no one has the same birthday. Then, for $n \leq 7$, the number of ways is simply the number of ways you can choose n distinct days of the week. The first person has 7 choices, the second has 6, etc. Therefore, we have $\boxed{1 - \dfrac{7!}{(7-n)!7^n}}$. For $n > 7$, the pigeonhole principle gives that the probability must be $\boxed{1}$.

(c) How many people chosen at random are needed to make the probability greater than 1/2 that there are at least two people born on the same day of the week?

Consider $n = 3$. Then, $1 - \frac{7!}{(7-3)!7^3} = 1 - 7 * 6 * 5/7^3 = 0.3878$.
Consider $n = 4$. Then, $1 - \frac{7!}{(7-4)!7^4} = 1 - 7 * 6 * 5 * 4/7^4 = 0.6501$.
Therefore, 4 or more people need to be chosen at random.

3. Show that if E and F are independent events, then $\bar{E}$ and $\bar{F}$ are independent events.

We assume that E and F are independent. Thus $p(E)p(F) = p(E \cap F)$. We consider the following:

$$\begin{aligned}
p(\bar{E})p(\bar{F}) &= (1-p(E))(1-p(F)) \\
&= 1-p(E)-p(F)+p(E)p(F) \\
&= 1-(p(E)+p(F)-p(E\cap F)) \\
&= 1-p(E\cup F) \\
&= p(\bar{E}\cap\bar{F}) \quad \blacksquare
\end{aligned}$$

By the definition, then, $\bar{E}$ and $\bar{F}$ are independent.

Note that if the correctness of the above operations is not self-evident, then **Theorems 1 and 2** of **Section 7.1** demonstrate their validity.

4. Suppose that A and B are two events with probabilities $p(A) = 2/3$ and $p(B) = 1/2$.

   (a) What is the largest $p(A \cap B)$ can be? What is the smallest it can be? Give examples to show that both extremes for $p(A \cap B)$ are possible. Show your work.

   For two given events, the probability $p(A \cap B)$ is maximized if $B \subset A$, and thus would simply be $p(B) = 1/2$. We can consider a fair dice, and let $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3\}$ Likewise, to minimize $p(A \cap B)$, $A \cap B$ must be minimized. Thus, we can consider a fair dice, and let $A = \{1, 2, 3, 4\}$, $B = \{4, 5, 6\}$. Therefore, we reach a minimum probability of $1/6$.

   (b) What is the largest $p(A \cup B)$ can be? What is the smallest it can be? Give examples to show that both extremes for $p(A \cup B)$ are possible. Show your work.

   Let us allow $A \cup B$ to span the sample space. Thus, $p(A \cup B) = 1$. An easy example is rolling a fair dice, and let $A = \{1, 2, 3, 4\}$, $B = \{4, 5, 6\}$.
   For two given events, the probability $p(A \cup B)$ is minimized if $B \subset A$, and thus would simply be $\max(p(A), p(B)) = p(A) = 2/3$. We can consider a fair dice, and let $A = \{1, 2, 3, 4\}$, $B = \{1, 2, 3\}$.

5. Find each of the following probabilities when n independent Bernoulli trials are carried out with probability of success p. Show your work.

   (a) The probability of no success.

   The binomial distribution yields $\binom{n}{k}p^k(1-p)^{n-k}$. For $k = 0$, we thus have $\binom{n}{0}p^0(1-p)^n = \boxed{(1-p)^n}$

   (b) The probabilities of at least one success.

The probability of at least one success is the complement of the probability of no success. Thus, $\boxed{1-(1-p)^n}$

(c) The probability of at most one success.

At most one success implies $k = 0, 1$. Thus, $(1-p)^n + \binom{n}{1}p^1(1-p)^{n-1} = \boxed{(1-p)^n + np(1-p)^{n-1}}$

(d) The probability of at least two successes.

The probability of at least two successes is the complement of at most 1. Therefore, our previous result yields $\boxed{1-(1-p)^n - np(1-p)^{n-1}}$

Random Variables [45 pts]

1. Suppose that we roll a fair die until a 6 comes up.

   (a) What is the probability that we roll the die n times?

   The probability a 6 is not rolled is $5/6$. Thus, the probability a 6 is not rolled for $n-1$ times and rolled exactly once is $(5/6)^{n-1}(1/6) = \boxed{\frac{5^{n-1}}{6^n}}$

   (b) What is the expected number of times we roll the die?

   Using a geometric distribution, let X be equal to the number of times you roll the dice, and p be the probability a 6 is rolled. Thus, we have $\mathbb{E}(X) = 1/p = \boxed{6}$

2. A space probe near Neptune communicates with Earth using bit strings. Suppose that in its transmissions it sends a 1 one-third of the time and a 0 two-thirds of the time. When a 0 is sent, the probability that it is received correctly is 0.9, and the probability that it is received incorrectly(as a 1) is 0.1. When a 1 is sent, the probability that it is received correctly is 0.8, and the probability that it is received incorrectly (as a 0) is 0.2.

   (a) Find the probability that a 0 is received. Show your work.

   Let S_n be the event such that a bit n is sent, and R_n be the event where n is received. Thus, we are given the following:

   - $P(S_1) = 1/3$
   - $P(S_0) = 2/3$
   - $P(R_1|S_1) = 0.8$
   - $P(R_0|S_1) = 0.2$

- $P(R_1|S_0) = 0.1$
- $P(R_0|S_0) = 0.9$

We seek to find $P(R_0)$. We use the definition of conditional probability, and the law of total probability $P(R_0) = P(R_0 \cap S_0) + P(R_0 \cap S_1) = P(R_0|S_0)P(S_0) + P(R_0|S_1)P(S_1) = \boxed{2/3}$

(b) Find the probability that a 0 was transmitted, given that a 0 was received. Show your work.

Let us use Bayes' theorem

$$P(S_0|R_0) = \frac{P(R_0|S_0)P(S_0)}{P(R_0)} = \frac{0.9(2/3)}{2/3} = \boxed{0.9}$$

3. Suppose that X and Y are random variables and that X and Y are nonnegative for all points in a sample space S. Let Z be the random variable defined by $Z(s) = max(X(s), Y(s))$ for all elements $s \in S$. Show that $E(Z) \leq E(X) + E(Y)$.

Given the above, it is evident that Z is a nonnegative random variable, such that $Z(s) = max(X(s), Y(s)) \leq X(s) + Y(s)$. Now, let us consider $\mathbb{E}(Z) = \sum_{s \in S} p(s)Z(s)$. Since $p(s)$ is also nonnegative for all $s \in S$,the following must also be valid for all s:

$$p(s)Z(s) \leq p(s)[X(s) + Y(s)] = p(s)X(s) + p(s)Y(s)$$

Then, for such an expression, we are able to take the sum over all s:

$$\sum_{s \in S} p(s)Z(s) \leq \sum_{s \in S} p(s)X(s) + p(s)Y(s) \implies \mathbb{E}(Z) \leq \mathbb{E}(X) + \mathbb{E}(Y)$$

4. Let X and Y be the random variables that count the number of heads and the number of tails that come up when two fair coins are flipped. Are X and Y independent? Prove your answer.

For two flips, there are precisely four possible outcomes, which we will denote $S = \{HH, HT, TH, TT\}$. Thus, we have the following: $P(X = 0) = 1/4, P(Y = 0) = 1/4, P(X = 0 \wedge Y = 0) = 0$.

$$(1/4)(1/4) \neq 0$$

Therefore, by the definition of independence of random variables, X and Y are not independent. $\blacksquare$

5. What is the variance of the number of times a 6 appears when a fair die is rolled 10 times. Show your derivation.

Let us first consider the variance of a single Bernoulli random variable, X_i, the measure of whether the i-th roll is a 6 or not. Our expected values:

$$\mathbb{E}(X_i) = 1 * P(X_i = 1) + 0 * P(X_i = 0) = p = 1/6$$

$$\mathbb{E}(X_i^2) = 1^2 * P(X_i = 1) + 0^2 * P(X_i = 0) = p = 1/6$$

Thus, $V(X_i) = \mathbb{E}(X^2) - \mathbb{E}(X_i)^2 = p - p^2 = p(1-p) = 5/36$

Now, let it be demonstrated that X_i and X_j, the Bernoulli random variables corresponding to whether or not the i-th or j-th roll is a 6 or not, are independent. Distinct rolls of a dice are clearly independent. Nevertheless, a proof, by exhaustion, is offered below:

Our sample space: $S = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, with the pairs in S corresponding to the results of the i,j-th rolls.

$P(X_i = 1) = 1/6$, $P(X_j = 1) = 1/6$

- $P(X_i = 1 \wedge X_j = 1) = 1/36$,
- $P(X_i = 1 \wedge X_j = 0) = 5/36$, as we fix the first element in the pair to be 6, and let the second element not be 6
- $P(X_i = 0 \wedge X_j = 1) = 5/36$, likewise above
- $P(X_i = 0 \wedge X_j = 0) = 25/36$, as there are only 5^2 ways to choose non-6 elements in the pair

For each a, b, $P(X_i = a)P(X_j = b) = P(X_i = a \wedge X_j = b)$.

Since each roll has been proved to be independent, we can use Bienayme's Formula:

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = np(1-p) = 10 * 5/36 = \boxed{25/18}$$

Note that the same result can be found directly using the given formula of the variance of n Bernouilli trials: $V(X) = npq$.

6. Give an example that shows that the variance of the sum of two random variables is not necessarily equal to the sum of their variances when the random variables are not independent.

Consider the example posed in **Question 4**. The variance of X (number of heads, $\mathbb{E}(X) = 0 * 1/4 + 1 * 1/2 + 2 * 1/4 = 1$):

$$\mathbb{E}(X^2) = 0^2 * 1/4 + 1^2 * 1/2 + 2^2 * 1/4 = 3/2 \implies V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 1/2$$

Since the pdf of Y mirrors that of X, $V(Y) = 1/2$ by symmetry. Now, the variance of $X+Y$:

$$\mathbb{E}(X+Y) = 0*0 + 1*0 + 2*1 + 3*0 + 4*0 = 2$$

$$\mathbb{E}((X+Y)^2) = 0^2*0 + 1^2*0 + 2^2*1 + 3^2*0 + 4^2*0 = 4$$

$$V(X+Y) = \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 = 0$$

Thus, $V(X+Y) = 0 \neq 1 = V(X) + V(Y)$

7. Use Chebyshev's inequality to find an upper bound on the probability that the number of tails that come up when a biased coin with probability of heads equal to 0.6 is tossed n times deviates from the mean by more than $\sqrt{n}$.

   Chebyshev's inequality states:

$$P(|X - \mathbb{E}(X)| \geq a) \leq \frac{V(X)}{a^2}$$

   Let X, be the sum of n Bernouilli trials, the probability of success (tails) being 0.4. Then, $V(X) = p(1-p)n$. Thus, we have an upper bound

$$P(|X - \mathbb{E}(X)| \geq \sqrt{n}) \leq \frac{np(1-p)}{n} = p(1-p) = \boxed{0.24}$$

8. Suppose that the probability that x is the i-th element in a list of n distinct integers is $i/(n(n+1))$. Find the average number of comparisons used by the linear search algorithm (Algorithm 2 in Section 3.1 in the textbook) to find x or to determine that it is not in the list.

   Let us define X to be the random variable denoting the number of comparisons. Thus, the problem is equivalent to finding $\mathbb{E}(X)$.

   Let $X(i) = 2i + 1$, for $i \leq n$ (the while loop checks twice for each i, and the if statement checks once), and $X(i) = 2n + 2$ otherwise (here, the while loop checks one additional time for $i = n + 1$, at the first condition for the while loop).

$$\begin{aligned}
\mathbb{E}(X) &= \sum_i X(i)P(i) \\
&= (2n+2)\left(1 - \sum_{i=1}^{n} \frac{i}{n(n+1)}\right) + \sum_{i=1}^{n}(2i+1)\left(\frac{i}{n(n+1)}\right) \\
&= (2n+2)\left(1 - \frac{1}{n(n+1)}\sum_{i=1}^{n} i\right) + \frac{1}{n(n+1)}\sum_{i=1}^{n}(2i^2+i) \\
&= (2n+2)\left(1 - \frac{1}{n(n+1)}\frac{n(n+1)}{2}\right) + \frac{1}{n(n+1)}\left(\frac{2n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2}\right) \\
&= (n+1) + \left(\frac{(2n+1)}{3} + \frac{1}{2}\right) \\
&= \frac{6n+6+4n+2+3}{6} = \frac{10n+11}{6}
\end{aligned}$$

   Note that we assume the while loop will check the $(i \leq n)$ condition before it checks the $a_i = x$ condition, and that it does not check both conditions for each iteration of the while loop. If it does indeed check the both conditions, then for the case where x is not in the list, there are $(2n+3)$ comparisons. Therefore, our final answer would be: $\frac{10n+11}{6} + 1/2 = \frac{5n+7}{3}$.