COMPSCI 230: Discrete Mathematics for Computer Science 4/17/2020

Recitation 13: Probability

Created By: David Fischer; adapted from Erin Taylor, Kevin Sun, Alex Steiger

1. Rahul owns a pizza parlor that offers 9 distinct toppings.

   (a) Consider the process of choosing a pizza with exactly three toppings uniformly at random. What is the size of the state space of this random process?
   Solution: This question is equivalently asking how many different pizzas can Rahul make with exactly three toppings? There are $\binom{9}{3}$ pizzas since there are that many distinct combinations of exactly 3 toppings out of the 9 available. We assume some familiarity with this notation; if not, there are different ways you can think of this result. One way is as having a bag with 9 items in it, and drawing 3 from it without replacement. There are nine options of what the first item can be, 8 for the second and 7 for the third, so there are 9*8*7 total possible ways to draw things out of the bag. Then we need to divide by the number of different permutations of three items, to ensure we do not double count any options (a drawing of pepperoni, olives, and ham is the same as a drawing of olives, ham and pepperoni). This gives the classic combination formula $\frac{n!}{k!(n-k!)} = \frac{9*8*7}{3*2}$.

   (b) How many different pizzas can be made with any number of distinct toppings?
   Solution: There are 2^9 pizzas with any set of distinct toppings. There are many different ways to see this, two of which are below.
   Count the number of pizzas with exactly k distinct toppings, for $k \in \{1, .., 9\}$ and sum them up. We must be sure we do not double count pizzas, but note that a pizza does not have both exactly k and ℓ distinct toppings for $\ell \neq k$. From the argument in part (a), there are then $\sum_{k=0}^{9} \binom{9}{k} = 2^9$ pizzas. This was covered in lecture last year; it should be taken as a fact that for any $n \in \mathbb{N}$, $\sum_{k=0}^{n} \binom{n}{k} = 2^n$. For those interested in the math we include the derivation below.
   Another way to see this, is to note that the set of all distinct pizzas that can be made with any number of toppings is the powerset of the toppings. We know the size of the powerset of a set with size n is 2^n, so there must be 2^n such pizzas.
   Here is one proof for $\sum_{k=0}^{n} \binom{n}{k} = 2^n$: By induction on n. For $n = 0$ we know $\binom{0}{0} = 1 = 2^0$. Suppose for some t we know that $\sum_{k=0}^{t} \binom{t}{k} = 2^t$. Then note the following facts $\binom{t+1}{0} = \binom{t}{0}$ and $\binom{t+1}{t+1} = \binom{t}{t}$. As well, recall Pascal's identity: $\binom{n-1}{k} + \binom{n-1}{k-1} = \binom{n}{k}$.

$$\begin{aligned}\sum_{k=0}^{t+1} \binom{t}{k} &= \binom{t+1}{0} + \binom{t+1}{1} + \binom{t+1}{2} + ... + \binom{t+1}{t} + \binom{t+1}{t+1} \\ &= \binom{t}{0} + \binom{t}{0} + \binom{t}{1} + \binom{t}{1} + \binom{t}{2} + \binom{t}{2} + \binom{t}{3} + ... + \binom{t}{t-1} + \binom{t}{t} + \binom{t}{t} \\ &= 2 * \sum_{k=0}^{t} \binom{t}{k} = 2 * 2^t = 2^{t+1}\end{aligned}$$

Just to be complete, let's show pascal's identity too

$$\begin{aligned}
\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-k-1)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\
&= \frac{(n-1)!(n-k)}{k!(n-k-1)!(n-k)} + \frac{(n-1)!k}{(k-1)!(n-k)!k} \\
&= \frac{(n-1)!(n-k)}{k!(n-k)!} + \frac{(n-1)!k}{(n-k)!k!} \\
&= \frac{(n-1)!(n-k) + (n-1)!k}{k!(n-k)!} \\
&= \frac{(n-1)!(n-k+k)}{k!(n-k)!} = \frac{n!}{k!(n-k)!} \\
&= \binom{n}{k}
\end{aligned}$$

(c) What is the total number of pizzas with any number of distinct toppings between 3 and 6?
Solution: From the arguments of the previous parts, the total number of pizzas is $\binom{9}{3} + \binom{9}{4} + \binom{9}{5} + \binom{9}{6}$. Using Pascal's identity above and the fact that $\binom{n}{k} = \binom{n}{n-k}$, this can be wriiten as $2\binom{10}{4} = 2\binom{10}{6}$.

2. George Martin bought an assortment of 17 plot items, 6 action sequences, 4 love sequences, and 7 superstitious or magic sequences, to divide up amongst his three characters Jon Snow, Bran Stark, and the Night King.

(a) How many ways can George divide the plot points?
Solution: Let's denote the division of the plot points by a sequence $(J_a, B_a, N_a, J_l, B_l, N_l, J_s, B_s, N_s)$, where J_a denotes the number of action sequences Jon is given, B_l represents the number of love sequences Bran is given, and N_s represents the number of superstitious or magic sequences the Night King is given, and so on. Let D represent the set of all possible sequences. Intuitively, it seems we can individually count the number of ways to divide each treat separately, then multiply these three quantities to determine the total number of divisions of treats. In fact this holds. Let D_a be the set of sequences of the form (J_a, B_a, N_a) and define D_l and D_s similarly. Then D is the cross product of D_a, D_l and D_s: $D = D_a \times D_l \times D_s$, and if we recall the product rule we know $|D| = |D_a| * |D_l| * |D_s|$. So let's focus on finding those quantities.
First let's count the number of ways to divide the action sequences among the characters. The stars and bars method gives that there are $\binom{n+k-1}{k-1}$ ways to distribute n things among k entites, so there are $\binom{\binom{4+3-1}{3-1}=6}{2}$ ways to divide the action sequences among the characters. Similarly there are $\binom{8}{2}$ and $\binom{9}{2}$ ways to divide the 6 love sequences and 7 superstitious or magic sequences, respectively. In total, then, there are $\binom{6}{2}\binom{9}{2}\binom{8}{2}$ ways to divide all 17 plot points. Note that since we care about the type of plot point, we cannot simply apply the stars and stripes method to all 17 plot points and 3 characters.
For those who don't know, there is a way of conceptualizing the number of ways to distribute n things among k people by imagining a string of n stars $* * * *$ ..., and inserting k bars into that string. Here the stars represent each item, and each bar represents each entity. The stars to the left of a bar represent the items the

entity represented by the bar is given. Since we are distributing all n items, there must be a bar on the far right of this string, and so there are $k-1$ bars left to distribute. So, there are $n+k-1$ possible "spots" in the string to put the bars, and we choose $k-1$ of those spots, and so there are $\binom{n+k-1}{k-1}$ ways to distribute the n items among the k entities.

(b) To divide them up, George rolls a fair "three-sided die" independently for each plot point, where, if the outcome of the die roll is 1,2, or 3 then Jon, Bran or the NK receives the plot point. What is the probability that:

i. Jon gets no plot points?
Solution: The event that Jon receives nothing is where his number, 1, is never rolled on any of the 17 rolls. The probability any roll shows a specific number is $\frac{1}{3}$, so the probability that a 1 is *not* rolled is $1 - \frac{1}{3} = \frac{2}{3}$. Since each dice roll is independent, the probability of the event that no die roll is a 1 is the probability of the intersection of the events that each die roll is not 1, which is the product of the probabilities that each die roll is not 1 which is $\frac{2}{3}^{17}$.

ii. Jon gets exactly 2 plot points?
Solution: The event that Jon gets exactly two of the plot points is where his number is rolled exactly twice. There are $\binom{17}{2}$ ways for this to occur, and each has a probability of $(1/3)^2(2/3)^{15}$, so the total probability is $\binom{17}{2}(1/3)^2(2/3)^{15}$

iii. The Night King gets exactly one or three love plot points?
Solution: The event that the Night King gets exactly one or three love plot points is where his number is rolled in exactly one or three rolls of the four dice rolls for the love plot points. Similar to the above argument, the probability he gets exactly one love sequences is $\binom{4}{1}(1/3)^1(2/3)^3$, and the probability of exactly three love sequences is $\binom{4}{3}(1/3)^3(2/3)^1$. Note that these events are disjoint. That is, he cannot simultaneously get exactly 1 and 3 love plot points. So, the total probability of getting either 1 or 3 is the sum of the probabilities above: $\binom{4}{1}(1/3)^1(2/3)^3 + \binom{4}{3}(1/3)^3(2/3)^1$

iv. Bran gets exactly 2 action or magic plot points and one love plot point.
Solution: The event that Bran gets exactly two action or magic plot points and one love plot points is where his number is rolled in exactly one of the four rolls for the love plot points, and 2 of the 6+7 = 13 rolls for the action and magic plot points. The probability of the exactly one role in the love plot point rolls is $\binom{4}{1}(1/3)^1(2/3)^3$, and the probability for the action and magic plot points is $\binom{13}{2}(1/3)^2(2/3)^{11}$. Since once again the rolls are independent, we have that the probability that both occur is their product: $\binom{4}{1}(1/3)^1(2/3)^3\binom{13}{2}(1/3)^2(2/3)^{11} = 4\binom{13}{2}(1/3)^3(2/3)^{14}$.

3. Union Bound - A certain model of generating random graphs creates a class of graphs called Erdos-Renyi random graphs. They have two parameters, n and p, where n is the number of vertices, and p is the probability that two nodes are connected. In other words, we create a disconnected n vertex graph and then add every edge with probability p. We denote this graph by $G(n,p)$ One question that arises is whether or not there exists some isolated node, where an isolated node is a node which is not connected to any other vertex in the graph. To be formal, let B_n be the event that a graph randomly generated according to this model $G(n,p)$ has at least one

isolated node. Show that

$$P(B_n) \leq n(1-p)^{n-1}$$

Then, conlcude that for any $\epsilon > 0$, if $p = p_n = (1+\epsilon)(\log(n)/n)$ then there is an isolated node with low probability; in other words, that,

$$\lim_{n\to\infty} P(B_n) = 0$$

Solution: There are n nodes in the network; let's label them with the integers $1...n$. Let A_i be the event the i^{th} node is isolated. Then we have

$$B_n = \bigcup i = 1^n A_i$$

We can apply the union bound, which says $P(\bigcup_{i\in S} E_i) \leq \sum_{i\in S} P(E_i)$, where S is the state space and E_i is an event in the state space. This gives

$$P(B_m) = P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$$

Since all nodes are conceptually the same, we have $P(A_i) = P(A_j)$ for all i, j, so we can simplify the sum to

$$P(B_m) \leq nP(A_1)$$

Event A_1 occurs if Node 1 is not connected to any of the other $n-1$ nodes. Since each edge occurs with probability p, each edge (i, j) does not occur with probability $1-p$. Since each edge is independently created or not, we conclude that the probability the first node is not connected is the intersection of the independent events that the first node is not connected to a particular other node, or

$$P(A_1) = (1-p)^{n-1}$$

Thus, we have

$$P(B_n) \leq n(1-p)^{n-1}$$

as desired.

To prove the limit result, we use the following identity

$$\lim_{x\to\infty}(1+\frac{c}{x})^x = e^c, \text{ for any constant } c \in \mathbb{R}$$

So we have that

$$
\begin{aligned}
\lim_{\to\infty} P(B_n) &\leq \lim_{n\to\infty} n(1-p_n)^{n-1} \\
&= \lim_{n\to\infty} n\left[1-(1+\epsilon)\frac{\log n}{n}\right]^{n-1} \\
&= \lim_{n\to\infty} n\left[1-\frac{1+\epsilon}{\frac{n}{\log n}}\right]^{n-1} \\
&= \lim_{n\to\infty} n\left(\left[1-\frac{1+\epsilon}{\frac{n}{\log n}}\right]^{\frac{n}{\log n}}\right)^{\frac{(n-1)\log n}{n}} \\
&= \lim_{n\to\infty} ne^{-(1+\epsilon)\log n} \\
&= \lim_{n\to\infty} \frac{1}{n^{\epsilon}} \\
&= 0
\end{aligned}
$$

Finally, since $P(B_n)$ is a valid probability, we have that $P(B_n) \geq 0$, so $\lim_{n\to\infty} P(B_n) = 0$, as desired. Note that neither knowledge of this identity, nor the ability to do this derivation on your own is expected. We include it simply because the process of having an event, bounding the probability of that event and then doing math manipulations to force that bound to fit a known identity in the hopes of simplifying things is extremely common, and gathering some exposure to it early on does not hurt!

4. As an exercise for the reader, use the inclusion-exclusion principle to calculate $P(B_n)$ exactly.