CMSC 250: Discrete Structures

Summer 2017

Lecture 23 - Outline

July 13, 2017

Problem: Using LOE, calculate the expected value of the sum of the value from three fair dice rolls.

Solution: Let X_1, X_2, and X_3 be the random variables that are the value of roll 1, 2, and 3, respectively. Note that $X = X_1 + X_2 + X_3$. By linearity of expectation:

$$\mathbf{E}[X] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \mathbf{E}[X_3]$$

Note that each of $\mathbf{E}[X_1]$, $\mathbf{E}[X_2]$, $\mathbf{E}[X_3]$ is 3.5, from last class:

$$\begin{aligned} &= 3.5 + 3.5 + 3.5 \\ &= 10.5 \end{aligned}$$

Indicator Random Variables

An **indicator random variable** is an random variable that takes value 1 if and only if some event occurs, and is 0 otherwise.

Let X_A be the indicator random variable for the event A. Note that:

$$\begin{aligned} \mathbf{E}[X_A] &= 0 \cdot \Pr[X_A = 0] + 1 \cdot \Pr[X_A = 1] \\ &= \Pr[X_A = 1] \\ &= \Pr[A] \end{aligned}$$

Problem: Suppose we throw n balls into n bins, such that for any particular ball, it lands in any one of the n bins with equal probability. What is the expected number of empty bins?

Solution: Let X be the RV that is the number of empty bins. Let X_i be a random variable that is 1 if the ith bin is empty and is 0 otherwise. Clearly

$$X = \sum_{i=1}^{n} X_i$$

By linearity of expectation, we have

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{i=1}^{n} \mathbf{E}[X_i] \\
&= \sum_{i=1}^{n} \Pr[X_i = 1] \\
&= \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^n \\
&= \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^n
\end{aligned}$$

As $n \to \infty$, $(1 - \frac{1}{n})^n \to \frac{1}{e}$. Hence, for large enough values of n we have

$$\mathbf{E}[X] = \frac{n}{e}$$

Problem: Suppose there are k people in a room and n days in a year. What is the expected number of pairs of people that share the same birthday?

Assume that for any particular person, they have a birthday chosen from the n days with equal probability, independent from other people.

Solution: Let X be the random variable denoting the number of pairs of people sharing the same birthday. For any two people i and j, where $i < j$, let X_{ij} be an indicator random variable for the event that person i and j have the same birthday. Note that $X = \sum_{i,j} X_{ij}$.

Using the linearity of expectation we get

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{i<j} \mathbf{E}[X_{ij}] \\
&= \sum_{i<j} \Pr[X_{ij} = 1]
\end{aligned}$$

Note that for all i and j, $\Pr[X_{ij} = 1]$ should be the same – there is no reason to expect that any two people are more likely to have the same birthday.

Let us determine $\Pr[X_{ij} = 1]$, for some arbitrary i and j. Since the probability space is uniform, we just need to find $|X_{ij} = 1|$ and $|\Omega|$.

Note that $|\Omega| = n^k$ (since each of the k people have n options for their birthday). We can construct outcomes in $X_{ij} = 1$ in the following way:

Step 1: Choose a birthday for person i – n ways

Step 2: Fix the birthday for person j – 1 ways

Step 3: Choose a birthday for each of the remaining $k-2$ people – n^{k-2}

By the multiplication rule, $|X_{ij} = 1| = n^{k-1}$. So, $\Pr[X_{ij} = 1] = \frac{n^{k-1}}{n^k} = \frac{1}{n}$.

Plugging this in:

$$\begin{aligned}\mathbf{E}[X] &= \sum_{i<j} \Pr[X_{ij} = 1] \\ &= \sum_{i<j} \frac{1}{n}\end{aligned}$$

Note that we are simply summing $\frac{1}{n}$ for $\binom{k}{2}$ times:

$$= \binom{k}{2} \cdot \frac{1}{n}$$

Problem: Suppose we construct an n length string from lower-case alphabet letters, such that the letter for each position in the string is chosen uniformly at random from the 26 letters.

What is the expected number of different letters that are used?

Solution: Let X be the random variable that denotes the number of different letters that are used. Let X_i be the indicator random variable for the event that letter i is used in the string. Note that $X = \sum_i X_i$.

By the Linearity of Expectation:

$$\begin{aligned}\mathbf{E}[X] &= \sum_{i} \mathbf{E}[X_i] \\ &= \sum_{i} \Pr[X_i = 1]\end{aligned}$$

Note again that for all i, $\Pr[X_i = 1]$ should be the same. From the problem definition, there is no reason to expect that any particular letter is more likely to appear in the string.

Let us determine $\Pr[X_i = 1]$ for an arbitrary i. Since the probability space is uniform, we just need to find $|X_i = 1|$ and $|\Omega|$.

Note that $|\Omega| = 26^n$ (since we are choosing letters for n positions total, with 26 options per letter).

Let us determine $|X_i = 1|$ complementarily. Note that $|X_i = 1| = |\Omega| - |\overline{X_i = 1}|$. But what is $\overline{X_i = 1}$? This is the event that the letter i is not used in the string. Note that $|\overline{X_i = 1}| = 25^n$, since we are still choosing letters for n positions total, but each position only has 25 options (we cannot choose letter i). So we have that $|X_i = 1| = 26^n - 25^n$.

Putting this together, we have that $\Pr[X_i = 1] = \frac{26^n - 25^n}{26^n}$.

Plugging this in:

$$\begin{aligned}\mathbf{E}[X] &= \sum_i \Pr[X_i = 1] \\ &= \sum_i \frac{26^n - 25^n}{26^n} \\ &= 26 \cdot \frac{26^n - 25^n}{26^n}\end{aligned}$$

Problem: Recall the problem of goats, cars and doors. Suppose now that there are n doors, but still only 1 door has a car behind it. Suppose further that the contestant selects doors uniformly at random and opens them until he finds the door with the car behind it. What is the expected number of doors that the contestant will have to open until he opens the door with the car behind it? (You should include the opening of the door with the car in this calculation)

Solution: We did not complete this in class. We will see this again next time.