

CS410 HW4

Russell Miller and Thomas Schreiber

November 12, 2011

2.14 The geometric distribution arises as the distribution of the number of times we flip a coin until it comes up heads. Consider now the distribution of the number of flips X until the k th head appears, where each coin flip comes up heads independently with probability p . Prove that this distribution is given by

$$\Pr(X = n) = \binom{n-1}{k-1} p^k (1-p)^{n-k} \quad \text{for } n \geq k.$$

The distribution of a binomial random variable is

$$\Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The difference between this and the aforementioned distribution is that this gives the probability that there are k heads. What we want to find is the probability that after finding k heads, we have only flipped n coins.

Assume that on the n th flip, we achieve our k th heads. Now we know that on the previous $n-1$ flips, we had exactly $k-1$ heads. So the probability that we tossed $k-1$ heads in $n-1$ flips is the same as the probability that the k th heads was achieved on the n th flip. However, p and $p-1$ are not raised to the $k-1$, and that's because there are in fact k total heads. ■

3.6 For a coin that comes up heads independently with probability p on each flip, what is the variance in the number of flips until the k th head appears?

The distribution of the number of flips of a coin until the k th heads could be viewed as a sum of the distributions of geometric random variables representing the previous $k-1$ heads. Let X be the random variable for the number of coin flips until the k th heads, and each X_i be a geometric random variable for the number of flips to get the i th heads. For example, X_1 is the number of coin flips until the first heads.

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

We know that if each X_i is mutually independent (which we're told each coin flip is)

$$\text{Var}\left[\sum_{i=1}^k X_i\right] = \sum_{i=1}^k \text{Var}[X_i]$$

We need the variance of X_i , which we derived in class.

$$\text{Var}[X_i] = \frac{1-p}{p^2}$$

Plugging this back into our distribution of X to find its variance

$$\begin{aligned} \text{Var}[X] &= \sum_{i=1}^k \frac{1-p}{p^2} \\ &= \boxed{\frac{k(1-p)}{p^2}} \end{aligned}$$

2.21 Let a_1, a_2, \dots, a_n be a random permutation of $\{1, 2, \dots, n\}$, equally likely to be any of the $n!$ possible permutations. When sorting the list a_1, a_2, \dots, a_n , the element a_i must move a distance of $|a_i - i|$ places from its current position to reach its position in the sorted order. Find

$$E \left[\sum_{i=1}^n |a_i - i| \right],$$

the expected total distance that elements will have to be moved.

We will be applying the linearity of expectation, so we need to consider $E[|a_i - i|]$. $\Pr(a_i = j)$, where j is any of the n positions, is $1/n$ since any of them is equally likely.

$$\begin{aligned} E[|a_i - i|] &= \sum_{i=1}^n |a_i - i| \Pr(a_i = i) \\ &= \frac{1}{n} \sum_{i=1}^n |a_i - i| \\ &= \frac{1}{n} \left| \sum_{i=1}^n a_i - \sum_{i=1}^n i \right| \\ &= \frac{1}{n} \left| \frac{n(n+1)}{2} - i \right| \\ &= \left| \frac{n+1}{2} - \frac{i}{n} \right| \end{aligned}$$

Now we are ready to apply linearity of expectation.

$$\begin{aligned} E \left[\sum_{i=1}^n |a_i - i| \right] &= \sum_{i=1}^n E[|a_i - i|] \\ &= \sum_{i=1}^n \left| \frac{n+1}{2} - \frac{i}{n} \right| \\ &= \left| \sum_{i=1}^n \frac{n+1}{2} - \frac{1}{n} \sum_{i=1}^n i \right| \\ &= \left| \frac{n(n+1)}{2} - \frac{1}{n} \frac{n(n+1)}{2} \right| \\ &= \left| \frac{n^2 - 1}{2} \right| \end{aligned}$$

3.3 Suppose that we roll a standard fair die 100 times. Let X be the sum of the numbers that appear over the 100 rolls. Use Chebyshev's inequality to bound $\Pr(|X - 350| \geq 50)$. Chebyshev's tells us that:

$$\Pr(|X - 350| \geq 50) \leq \frac{\text{Var}[X]}{50^2}$$

So, next we will find $\text{Var}[X]$:

$$\begin{aligned}
Var[X] &= E[X^2] - E[X]^2 \\
&= E\left[\sum_{i=1}^{100} \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2)\right] - E\left[\sum_{i=1}^{100} \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6)\right]^2 \\
&= \sum_{i=1}^{100} (E[\frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2)] - E[\frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6)]^2) \\
&= \sum_{i=1}^{100} (\frac{91}{6} - (\frac{21}{6})^2) \\
&= \sum_{i=1}^{100} (\frac{91}{6} - \frac{73.5}{6}) \\
&= \sum_{i=1}^{100} (\frac{17.5}{6}) \\
&= \frac{1750}{6} \\
&= 291.6\bar{6}
\end{aligned}$$

Finally, we plug this into Chebyshev's:

$$\begin{aligned}
Pr(|X - 350| \geq 50) &\leq \frac{Var[X]}{50^2} \\
&= \frac{291.6\bar{6}}{50^2}
\end{aligned}$$

$$\boxed{= 0.116\bar{6}}$$

paragraph 3.21 A fixed point of a permutation $\pi : [1, n] \Rightarrow [1, n]$ is a value for which $\pi(x) = x$. Find the variance in the number of fixed points of a permutation chosen uniformly at random from all permutations.

Let X be the number of fixed points.

We know that: $Var[X] = E[X^2] - E[X]^2$

First, we find $E[X]$:

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

Now we will find $E[X^2]$:

$$\begin{aligned}
E[X^2] &= E[(X_1 + X_2 + \cdots + X_n)^2] \\
&= E\left[\left(\sum_{i=1}^n X_i^2 + \sum_{i \neq j} X_i X_j\right)\right], \text{ the square of the polynomial is rewritten as the sum of each term} \\
&= \sum_{i=1}^n E[X_i^2] + \sum_{i \neq j} E[X_i X_j] \quad \text{squared plus that term multiplied by every other term.} \\
&= \sum_{i=1}^n Pr(X_i = 1) + \sum_{i \neq j} (Pr(X_i = 1)Pr(X_j = 1)) \\
&= n\left(\frac{1}{n}\right) + (n(n-1))\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right) \\
&= 1 + 1 \\
&= 2
\end{aligned}$$

Now, we plug these in to get the solution.

$Var[X] = 2 - 1^2 = 1$
