CS 171
Lecture Outline
February 23, 2012

Example. Suppose that n people leave their hats at the hat check. If the hats are randomly returned what is the expected number of people that get their own hat back?

Solution. Let X be the random variable that denotes the number of people who get their own hat back. Let $X_i, 1 \leq i \leq n$, be the random variable that is 1 if the ith person gets his/her own hat back and 0 otherwise. Clearly,

$$X = X_1 + X_2 + X_3 + \ldots + X_n$$

By linearity of expectation we get

$$\mathbf{E}[X] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} \frac{(n-1)!}{n!} = n \times \frac{1}{n} = 1$$

Example. Suppose we throw n balls into n bins with the probability of a ball landing in each of the n bins being equal. What is the expected number of empty bins?

Solution. Let X be the random variable denoting 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

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

Example. The following pseudo-code computes the minimum of n distinct numbers that are stored in an array A. What is the expected number of times that the variable *min* is assigned a value if the array A is a random permutation of the n elements.

```
FindMin(A, n)
1    min ← A[1]
2    for i ← 2 to n do
3        if (A[i] < min) then
4            min = A[i]
5    return min
```

Solution. Let X be the random variable denoting the number of times that *min* is assigned a value. We want to calculate $\mathbf{E}[X]$. Let X_i be the random variable that is 1 if *min* is assigned $A[i]$ and 0 otherwise. Clearly,

$$X = X_1 + X_2 + X_3 + \cdots + X_n$$

Using the linearity of expectation we get

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

Note that $\Pr[X_i = 1]$ is the probability that $A[i]$ contains the smallest element among the elements $A[1], A[2], \ldots, A[i]$. Since the smallest of these elements is equally likely to be in any of the first i locations, we have $\Pr[X_i = 1] = \frac{1}{i}$. Thus equation (1) becomes

$$\mathbf{E}[X] = \sum_{i=1}^{n} \frac{1}{i} = H(n) \approx \ln n + c$$

where c is a constant less than 1.

Example. (Birthday Paradox) Suppose there are k people in a room and n days in a year. We are interested in the probability that there are at least two people in the room with the same birthday. What is the smallest value of k for which this probability is at least 1/2? Assume that it is equally likely for a person to be born on any of the n days of the year.

Solution. Let B be the event that at least two people in the room have the same birthday. We are interested in $\Pr[B]$.

$$\begin{aligned} \Pr[B] &= 1 - \Pr[\overline{B}] \\ &= 1 - \frac{P(n,k)}{n^k} \end{aligned}$$

For $n = 365$, the smallest value of k for which the RHS is at least $1/2$ is 23. If $m = 40$ then $\Pr[E] = 0.89$, and if $m = 60$ then $\Pr[E] = 0.994$. This means that if there are 60 people then it is almost certain that there exists two among them sharing the same birthday. To illustrate how good our model is, consider the set of presidents of the United States of America. Through Bill Clinton, 41 people belong to this set. The chances of two of them sharing the same birthday is at least 89%. Indeed, James Polk (11th president) and Warren Harding (29th president) are both born on Nov. 2.

Example. In the above example, on average how many pairs of people share the same birthday?

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, let X_{ij} be an indicator random variable that is 1 if i and j have the same birthday and is 0 otherwise. Clearly $X = \sum_{i,j} X_{ij}$. Using the linearity of expectation we get

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{i,j} \mathbf{E}[X_{ij}] \\
&= \sum_{i,j} \Pr[X_{ij} = 1] \\
&= \sum_{i,j} \Pr[X_{ij} = 1] \\
&= \sum_{i,j} \frac{1}{n} \\
&= \frac{\binom{k}{2}}{n} \\
&= \frac{k(k-1)}{2n}
\end{aligned}$$

Assuming $n = 365$, the smallest value of k for which the RHS is at least 1 is 28.