Lecture #20: Random Variables and Expectation
Lecture Presentation

A Problem You Might Find in a Textbook

Suppose you roll a die (whose surfaces show the numbers 1, 2, 3, 4, 5 and 6) over and over again — but this, time you always roll the die exactly ten times. You are interested in the "expected" number of times that you roll a number that is the same as the number you rolled immediately before this.

Sample Space:

The Function You are Interested In:

Let us suppose — that is, ***assume*** — that this is a fair die. That is, you roll each number between 1 and 6 with the same probability, and the number that you see (each time) does not depend on any of the numbers you rolled before.

Corresponding Probability Distribution:

Notice that there nine ways for you to "roll the same number as you did immediately before this":

- The *second* number that you roll might be the same as the *first*.
- The *third* number that you roll might be the same as the *second*.
- The *fourth* number that you might be the same as the *third*.

...and so on.

Taking Advantage of This:

Using a Result from the Preparatory Reading to Finish:

A Problem Involving Hash Tables with Chaining

Once again, consider a ***hash table with chaining***. As before, suppose that the hash table stores a set

$$S = \{k_1, k_2, \dots, k_n\}$$

of n values from some "universe" $\mathcal{U}$" — where n is a positive integer. Suppose, as well, that the table size is m, so that the hash table is constructed, and accessed, using some ***hash function***

$$h : \mathcal{U} \to \{0, 1, 2, \dots, m-1\}.$$

For $0 \le i \le m-1$, let S_i be the set of integers j such that $h(k_j) = i$:

$$S_i = \{j \in \mathbb{Z} \mid 1 \le j \le n \text{ and } h(k_j) = i\}.$$

Suppose that we now model the problem using the same kind of ***sample space*** as in recent lectures, and that we make the same ***assumptions*** about hash functions, as well.

What This Means:

Let us now consider the "expected" number of integers j such that the j^{th} key, k_j, is "hashed" to location 0:

How is This Value Represented?

Calculating This Value:

Something Worth Noting:

Proving a Result from the Required Reading

Consider the following ***exercise*** from the preparatory reading:

> ***Exercise:*** Prove that if Ω is a sample space with probability distribution $\mathsf{P} : \Omega \to \mathbb{R}$, $B \subseteq \Omega$ is an event such that $\mathsf{P}(B) > 0$, and X is a random variable, then
>
> $$\mathsf{E}[X \mid B] = \frac{1}{\mathsf{P}(B)} \times \sum_{\sigma \in B} (\mathsf{P}(\sigma) \times X(\sigma)).$$

What Do You Need To Do, To Get Started?

Proving the Claim: