Discrete Random Variables and the Binomial Distribution

Consider a dice with the following information:

X = Output 1 with the probability of 1/2

X = Output 2 with the probability of 1/3

X = Output 3 with the probability of 1/6

Hence, $E(X) = \sum_{x \in X} x.P(x)$

$$E(X) = 1.(1/2) + 2.(1/3) + 3(1/6) = 1 + 2/3 = 5/3$$

Thus, any variable that has probabilities of equaling different values is a discrete random variable. We calculate the average or expected value of that discrete random variable using the formula above. As an exercise, let Y be the discrete random variable equal to the sum of the faces after rolling two standard six-sided dice. Show that E(Y) = 7.

In addition to expectation, we define a term called variance for discrete random variables. (This calculation is rarely made in computer science for algorithm analysis, but I am including it for completeness sake with respect to the topic of discrete random variables.) Variance is simply a definition which roughly gauges, "how spread out" the distribution of the discrete random variable is. Here is the formula:

$$Var(X) = \sum_{x \in X} (x - E(X)^2 p_x$$

For our example above, we have

$$Var(X) = (1-5/3)^2(1/2) + (2-5/3)^2(1/3) + (3-5/3)^2(1/6) = 5/9$$

Also, the standard deviation of a discrete random variable is simply defined as the square root of its variance. An alternate way to calculate variance is as follows:

$$Var(X) = E(X^2) - [E(X)]^2$$

We define $E(X^2)$ as follows: $E(X^2) = \sum_{x \in X} x^2 p_x$

In the text, a proof is given for the result for the alternate calculation of variance. This is Theorem 3.12.

Binomial Probability Distribution

If we run n trials, where the probability of success for each single trial is p, what is the probability of exactly k successes?

$$\frac{1-p}{1} \quad \frac{p}{2} \quad \frac{p}{3} \quad \frac{1-p}{4} \quad \frac{p}{5} \quad \ldots\ldots \quad \frac{}{n}$$

k slots where prob. success is p, n-k slots where prob. failure is $1-p$

Thus, the probability of obtaining a specific configuration as denoted above is $p^k(1-p)^{n-k}$. From here, we must ask ourselves, how many configurations lead to exactly k successes. The answer to this question is simply, "the number of ways to choose k slots out of the n slots above. This is $\binom{n}{k}$. Thus, we must add $p^k(1-p)^{n-k}$ with itself exactly $\binom{n}{k}$ times.

This leads to the following answer to the given question:

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

We can also define a discrete random variable based on a binomial distribution. We can simply allow the variable to equal the number of successes of running a binomial trial n times. We then separately calculate the probability of obtaining 0 successes, 1 success, etc. , n successes. Here is a concrete example with n = 3 and p = 1/3:

X = 0, with probability $\binom{3}{0}(\frac{1}{3})^0(\frac{2}{3})^3 = \frac{8}{27}$

X = 1, with probability $\binom{3}{1}(\frac{1}{3})^1(\frac{2}{3})^2 = \frac{12}{27}$

X = 2, with probability $\binom{3}{2}(\frac{1}{3})^2(\frac{2}{3})^1 = \frac{6}{27}$

X = 3, with probability $\binom{3}{3}(\frac{1}{3})^3(\frac{2}{3})^0 = \frac{1}{27}$

We can calculate that E(X) = $1(\frac{12}{27}) + 2(\frac{6}{27}) + 3(\frac{1}{27}) = 1$.

Why can we leave at the term when X = 0? Also, why is this value in tune with our intuitive idea of what we should expect? We can formally prove this intuitive notion, namely that for a binomial distribution X, E(X) = np. The proof is Theorem 3.11 in the text.