Bernoulli Distribution Derivation

The Bernoulli distribution is useful in situations where the events have two results, either *true* or *false.* To derive the formula we will first look at one practical example: *a dice roll problem, you have five throws and we want to measure the probability of getting two sixes.*

1. What's the probability of getting a six? And the probability of getting any other number?
2. What's the amount of events that have 2 sixes in them?

The first question is simple, we have a probability of $p = 1/6$ of getting a six if the dice is good.

Then the second probability is $q = 5/6$ which is $q = 1 - p$. Why? well, getting a positive event is one number in six, whilst getting any other number is the sum of probabilities $\sum^{n-k} p$, excluding k possible events, this applies because p is the same for all results.

So, for example, if you throw a dice you get that, the probability of getting 2 sixes are given by $\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)$. How can one look at this intuitively? This can be seen as a Cartesian product like $[1, 6] \times [1, 6] \in \mathbb{N}$, which give 36 different combinations, and just one combination is the point $(6, 6)$. So you are basically taking one point of the samples that are possible. Then for the other part, because you only want to get two sixes, you want to measure the probability of *not* getting a six. This probability is $5/6$ and is easy to see by knowing the probability of the positive event.

In general, the probability of the five throws is given by

$$p_{event}(x) = p^x (1 - p)^{n-x}$$

where x is the number of positive events and n the total events.

Now, we have to count the total possible combinations of dices, we know that the amount of combinations of a subset of k elements in a set of n elements is given by $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. So, the probability distribution of a random discrete variable X is given by

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

Where n is the number of total events. This means that $P(X = k)$ is the probability of having k successful events in n throws.