Infinity in the Laws of Probability

While talking about discrete random variables, something was slipped in which actually requires a lot deeper understanding. We said that a discrete random variable X takes values in integers and put $p_i = P(X = i)$. We put the conditions

- $0 \leq p_i \leq 1$, and,
- $\sum_i p_i = 1$

The first condition is quite clear even when $p_i \neq 0$ for infinitely many i, but what does the second one mean? Let's look at an example which was introduced earlier. We flip a coin a number of times (independently) and let W denote the first occurence of head. Assume that the probability of a head on a single flip is p. As seen earlier this means that $P(W = n) = (1-p)^{n-1}p$. How do we understand the following statement?

$$1 = \sum_{n \geq 1} P(W = n) = \sum_{n \geq 1} (1-p)^{n-1}p$$

First of all, we note that $0 < p < 1$, so that $x = (1-p)$ has the following properties:

- $0 < x < 1$, and so,
- $x^n > x^{n+1}$ for all integers n.

In other words, we have a decreasing sequence $1 > x > x^2 > \cdots$. Does this mean that $\sum_n x^n$ makes sense? Not necessarily (as we shall see later)! However, in this case we have

$$s_n = (1 + x + x^2 + \cdots + x^n) = \frac{1 - x^{n+1}}{1 - x} \leq 1$$

(Note that the right-hand side makes sense since $x \neq 1$.) Now, this means that $s_n = \sum_{k=0}^{n} x^k$ (the *finite* sum) is always bounded by 1. It increases with n but cannot increase beyond 1. By the principle of Archimedes, there is a *least upper bound* of the sequence s_n and we *define* it to be the sum of *all* the x^n for $n \geq 1$. In other words

$$\sum_{n \geq 0} x^n := \sup_n \frac{1 - x^{n+1}}{1 - x} \leq 1$$

This still does not say *what* this limit is! We "know" the answer but like mathematicians, we will make a fuss about making sure that that is indeed the correct answer.

The limit is based on the following fundamental fact:

- if $0 < x < 1$ then as n goes to infinity x^n goes to 0.

In order to prove this statement we need to understand what it means. One way to understand it is to say that given *any* $M > 0$, there is an integer n_0 so that $|x^n| < M$ for *all* integers $n \geq n_0$.

Let us prove this somewhat indirectly as follows. By the least upper bound principle, the *is* some number $a \leq 1$ so that $\sum_{n\geq 0} x^n = a$. It folows that there is an n_0 so that $a \geq \sum_{k=0}^{n} > a - 1/M$ for $n \geq n_0$. It follows that $0 \leq x^n < 1/M$ for all such values of n.

It follows easily that $1 - x^{n+1}$ lies between 1 and $1 - 1/M$ for $n \geq n_1$ for a suitably chosen n_1. Thus, we obtain as a consequence that

$$\sum_{n\geq 0} x^n = \frac{1}{1-x} \text{ for } 0 < x < 1$$

This also gives us the identity

$$1 = \sum_{n\geq 1} P(W = n) = \sum_{n\geq 1} (1-p)^{n-1} p$$

That we were looking for earlier.

Probabilistic Interpretation

How do we interpret the above calculation in terms of Probability? Let E_n denote the event $W \leq n$. This is the event that we see a head in *at most* n coin flips. We have $E_n = \vee_{k\leq n}(W = k)$. Thus, $P(E_n) = \sum_{k\leq n} P(W = n)$. Clearly, $E_n \subset E_{n+1}$ since seeing a head in at most n flips means that we defnitely see a head in at most $n+1$ flips. Hence, $P(E_n) \leq P(E_{n+1})$ is a sequence of real numbers, all of which are bounded by 1 since the probability of any event is at most 1! By the Archimedean principle, this sequence of numbers has a least upper bound denoted as $\sup_n P(E_n)$.

Let $E = \vee_n E_n$; this is the event that we see at least one head. Now, it is clear that $P(E) \geq P(E_n)$. The Law of Infinity in Probability is that $P(E) = \sup_n P(E_n)$.

Let's re-state it in its full generality. Given a sequence of increasing events $E_n \subset E_{n+1}$. Their union $E = \vee_n E_n$ is also an event. Moreover, $P(E) = \sup_n P(E_n)$.

In the case of discrete probability this gives a meaning to $\sum_{i\geq 0} P(X = i) = 1$. The left-hand side is the probability of the union of the events $E_n : X \leq n$. The right-hand side is the assertion that this union exhausts the possibility. In particular, note that we have he assertion the the probability that we will see a head at least once is 1!

We can derive some natural consequences. Suppose we have a sequence of *decreasing* events $D_n \supset D_{n+1}$. Their intersection $D = \wedge_n D_n$ is also an event; in fact $D^c = \vee_n D_n^c$, so this follows from the previous case. Moreover, $P(D) =$

$\inf_n P(D_n)$; this too follows from the fact that $P(D^c) = \sup_n P(D_n^c)$ and the fact that $P(D^c) = 1 - P(D)$ and $P(D_n^c) = 1 - P(D_n)$.

Another application is to the case where A_n is a sequence of *mutually exclusive* events. In that case $B_n = \vee_{k \le n} A_k$ is an increasing sequence of events and $P(B_n) = \sum_{k \le n} P(A_k)$ is a non-decreasing sequenc of numbers. Hence $\sup_n P(B_n) = \sum_k P(A_k)$. On the other hand $B = \vee_n B_n = \vee_k A_k$. So we see that the probability of the union of mutually exclusive sequence of events is the sum of their probabilities.