Random variables

Math 477

October 2, 2014

1 Random variable as a way to quantify random events

In an experiment, we have (random) outcomes. We can give them names (for example tossing a coin twice, we can get $HH, TT \cdots$). Each of these have some weight attached to them, i.e. their probability (in the coin toss example, 1/4 for each). However, we cannot do computations with these outcomes unless we give them some numerical values. A *random variable* is a way to *quantify* the random outcomes in a meaningful manner. We use capital letters at the end of the alphabet: X, Y, Z, to denote random variables. We will also use lowercase letter x, y, z to denote deterministic numbers. You should take care to distinguish between these two.

Because we assign numerical values to a random event, the set of the form $\{X = k\}$ or more generally $\{X \leq x\}$ are random events. The meaningful manner referred to above is that we need to be able to assign probability to events of the type $\{X \leq x\}$, from which we can deduce probability of events of the type $\{X = x\}$ or $\{x_1 \leq X \leq x_2\}$.

Example 1.1. *Suppose we play a game where you win 3 dollars if the toss is H and lose 2 dollars if the toss is T. Assuming the coin is fair and let X be the random variable denoting your win / loss. Then*

$$\begin{aligned} P(X = 3) &= 1/2 \\ P(X = -2) &= 1/2. \end{aligned}$$

1.1 Examples

Once we have the notion of random variables, there are many interesting random events that we want to quantify. Consider the following examples.

Example 1.2. *Suppose we toss a coin until a H shows up. Suppose the probability of the coin showing a H is p and the tosses are independent. Let X denotes the number of tosses until a H shows, including the toss that shows H. What is the distribution of X?*

Ans: By independence

$$P(X = n) = (1 - p)^{n-1} p.$$

Note that we can also compute probability of events such as $\{X \le n\}$:

$$P(X \le n) = \sum_{k=1}^{n} (1-p)^{k-1} p = p \frac{1 - (1-p)^n}{p} = 1 - (1-p)^n.$$

The above is the probability of having to wait at most n for the first H, which is the same as the probability of not getting all T in the first n tosses.

Example 1.3. *Coupon collection Suppose therea re N distinct types of coupons, and the chance of getting any coupon is equally likely. Also each time we collect a coupon it is independent of our previous results. Let T be the random variable that denotes the number of trials before we obtain a complete set of at least one type of each coupon. What is the distribution of T?*

Ans: We want to compute $P(T = n)$. But this can be difficult. Rather we will compute $P(T > n)$ and deduce $P(T = n)$ afterward. For a fixed n, let $E_1, E_2, \cdots, E_N$ denote the events that the coupon of type $i, i = 1, \cdots, N$ did not show up in the first n trials. Then

$$\begin{aligned} P(T > n) &= P(\cup_{i=1}^{N} E_i) \\ &= \sum_{i=1}^{N} P(E_i) - \sum_{i<j}^{N} P(E_i E_j) + \sum_{i<j<k}^{N} P(E_i E_j E_k) + \cdots \\ &\quad + (-1)^{N+1} P(E_1 E_2 \cdots E_N). \end{aligned}$$

Actually note that $P(E_1 E_2 \cdots E_N) = 0$ since you have to get some coupon during the trials, you cannot miss all of them. Now by independence

$$\begin{aligned} P(E_i) &= \left(\frac{N-1}{N}\right)^n \\ P(E_i E_j) &= \left(\frac{N-2}{N}\right)^n \\ P(E_i E_j E_k) &= \left(\frac{N-3}{N}\right)^n \\ \cdots \end{aligned}$$

Thus

$$\begin{aligned} P(T > n) &= N\left(\frac{N-1}{N}\right)^n - \binom{N}{2}\left(\frac{N-2}{N}\right)^n + \cdots + \\ &\quad (-1)^N \binom{N}{N-1}\left(\frac{1}{N}\right)^n \\ &= \sum_{i=1}^{N-1} (-1)^{i+1} \binom{N}{i}\left(\frac{N-i}{N}\right)^n . \end{aligned}$$

Now we can compute $P(T = n)$ as:

$$P(T = n) = P(T > n-1) - P(T > n).$$

2 Discrete random variables

2.1 Definition

Definition 2.1. *A discrete random variable is a RV that can take at most countably many values with positive probability.*

Remark: a. If a RV takes on finitely many values, then it is automatically a discrete RV. An example of a non-discrete RV would be a RV that is used to describe the waiting time for some event (for a bus to arrive, for a machine to break down etc.)

b. We will most often deal with RV that takes on values in the natural number set: $0, 1, 2, 3, \cdots$. Just keep in mind that this needs not be the case: we have seen RV taking negative value in example (1.1). If your winning is a fractional amount of dollar, then we can also have a discrete RV that takes on a fractional value.

c. When investigating a RV, it is useful to look for its range: the set of values that it takes on with positive probability. For example, the range of X in example (1.1) is $\{3, -2\}$, and the range of X in example (1.3) is $\{N, N+1, N+2, \cdots\}$.

d. To specify a discrete random variable, we describe its probability mass function: $P(X = k)$ for all k such that $P(X = k) > 0$. From a RV point of view, it is completely specified when its probability mass distribution is known. There are two ways to do this: we either describe an experiment where X represents some quantity from that experiment; or we just abstractly specify the distribution of X and look for some examples in reality that fit the distribution. An example of the second approach would be to specify

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, \cdots$$

e. One should always check when given a probability mass function that it is valid. There are two conditions:

$$0 \leq P(X = k) \leq 1, \forall k$$
$$\sum_k P(X = k) = 1.$$

2.2 Some elementary probability identity

Here we assume that X is a discrete RV taking values on $0, 1, 2, \cdots$. Then for all integers $a \leq b$:

$$\begin{aligned}
P(X < b) &= P(X \leq b + 1) \\
P(X > a) &= P(X \geq a - 1) \\
P(a < X < b) &= P(X < b) - P(X \leq a) \\
P(a \leq X < b) &= P(X < b) - P(X < a) \\
P(a \leq X \leq b) &= P(X \leq b) - P(X < a).
\end{aligned}$$

2.3 Function of random variables

Let $g : \mathbb{R} \to \mathbb{R}$ and X a real valued RV. Then $g(X)$ is also a RV. What is the distribution of g? If X is a discrete RV then we also have $g(X)$ is a discrete RV. Moreover

$$P(g(X) = k) = \sum_{j:g(j)=k} P(X = j).$$

Example 2.2. *Let X be a RV with distribution $P(X = 3) = 1/3$ and $P(X = -2) = 2/3$. Then X^2 is a discrete RV with distribution $P(X^2 = 9) = 1/3$ and $P(X = -2) = 2/3$.*

Example 2.3. *Let X be a RV with distribution $P(X = 1) = 1/6, P(X = -1) = 1/3, P(X = 2) = 1/2$. Then X^2 is a discrete RV with distribution $P(X^2 = 1) = 1/2$ and $P(X^2 = 4) = 1/2$.*