2.2 Properties of the Distribution Function and the Cumulative Distribution Function

We remark some standard properties that the distribution function must follow. Firstly for any random variable we have,

$$\mathbb{P}(\{X = i\} \cup \{X = j\}) = \mathbb{P}(X = i) + \mathbb{P}(X = j).$$

This is because the events $\{X = i\}$ and $\{X = j\}$ are mutually exclusive. As a consequence we have the following: Suppose X is a discrete random variable with support in $\mathbb{N}_0$, then the distribution function $f_X(i)$ satisfies the following:

$$\sum_{i=0}^{\infty} f_X(i) = f_X(0) + f_X(1) + f_X(2) + \ldots = 1.$$

Essentially this is just a symbolic way of saying that all probabilities must sum to one. Indeed, the above sum just represents the fact that the random variable must pick one of values from $\mathbb{N}_0$. It should also be clear that $0 \leq f_X(i) \leq 1$ for all i, as each $f_X(i)$ represents a probability. We also introduce the *cumulative distribution function* as follows:

Definition 2.2.1. Suppose X is a discrete random variable with support in $\mathbb{N}_0$. Suppose $k \in \mathbb{N}_0$, the cumulative distribution function $F_X(k)$ is defined as follows:

$$F_X(k) := \mathbb{P}(X \leq k) = \sum_{i=0}^{k} \mathbb{P}(X = i).$$

We note that the cumulative distribution function just tells us the probability that a random variable takes a value which is at most k. We follow with an example to show how to check whether a function is indeed a distribution function, along with how to compute the cumulative distribution function.

Example 2.2.1. Suppose Z is a discrete random variable with the following distribution:

$$f_Z(i) = \begin{cases} \frac{1}{i^2} & \text{if } i = 2, 3; \\ \frac{23}{36} & \text{if } i = 5; \\ 0 & \text{otherwise.} \end{cases}$$

Is f_Z a distribution function? If so, what is F_Z?

To check if a function is a distribution function, we need to check two properties. Firstly that $0 \leq f_Z(i) \leq 1$ for every $i \in \mathbb{N}_0$. We also need to check that:

$$\sum_{i=0}^{\infty} f_Z(i) = 1.$$

The first condition can be seen immediately, if $i \notin \{2, 3, 5\}$ then $f_Z(i) = \mathbb{P}(X = i) = 0$. For $i \in \{2, 3, 5\}$ it can be seen from the definition of the function that $0 \leq f_Z(i) \leq 1$.

For the second condition we consider the sum:

$$\sum_{i=0}^{\infty} f_Z(i) = f_Z(0) + f_Z(1) + f_Z(2) + f_Z(3) + \dots$$

We know from above that if $i \notin \{2, 3, 5\}$ then $f_Z(i) = 0$. Therefore we have that:

$$\sum_{i=0}^{\infty} f_Z(i) = f_Z(2) + f_Z(3) + f_Z(5).$$

By replacing each of the terms by the definition in f_Z we have:

$$\sum_{i=0}^{\infty} f_Z(i) = \frac{1}{2^2} + \frac{1}{3^2} + \frac{23}{36} = 1.$$

So f_Z is indeed a distribution function. We now turn our attention to working out F_Z. We work term by term:

$$F_Z(0) = \mathbb{P}(Z \leq 0) = \mathbb{P}(Z = 0) = f_Z(0) = 0.$$
$$F_Z(1) = \mathbb{P}(Z \leq 1) = f_Z(1) + f_Z(0) = 0.$$

A useful property we can apply is that $\mathbb{P}(Z \leq i) = \mathbb{P}(Z \leq i - 1) + \mathbb{P}(Z = i)$, writing in terms of the distribution and cumulative distribution function we have that $F_Z(i) = F_Z(i-1) + f_Z(i)$. Therefore:

$$F_Z(2) = F_Z(1) + f_Z(2) = 0 + \frac{1}{2^2} = \frac{1}{4}.$$

Applying this fact again we have that:

$$F_Z(3) = F_Z(2) + f_Z(3) = \frac{1}{4} + \frac{1}{3^2} = \frac{13}{36}.$$

Again by the same argument:

$$F_Z(4) = F_Z(3) + f_Z(4) = \frac{13}{36} + 0 = \frac{13}{36}.$$

$$F_Z(5) = F_Z(4) + f_Z(5) = \frac{13}{36} + \frac{23}{36} = 1.$$

Now we note that $F_Z(i) = \mathbb{P}(Z \leq i)$. Therefore if $\mathbb{P}(Z \leq 5) = 1$, then for any $k > 5$ we have that $\mathbb{P}(Z = k) = 0$. Therefore, for all $k \geq 5$, we have that $F_Z(k) = 1$. Collating these values together we have the following:

$$F_Z(i) = \begin{cases} 0 & \text{if } 0 \leq i \leq 1; \\ \frac{1}{4} & \text{if } i = 2; \\ \frac{13}{36} & \text{if } 3 \leq i \leq 4; \\ 1 & \text{if } i \geq 5. \end{cases}$$