BASIC PROBABILITY : RANDOM VARIABLES

In the following we will consider a probability space $\left( \Omega, \mathcal{F}, P \right)$. We recall that: $\Omega$ is the sample space and enumerates all possibilities, $\mathcal{F}$ is a set of subsets of $\Omega$, they represent the events, and $P$ is a function on $\mathcal{F}$ which gives the probability of a certain event. An element $\omega$ of $\Omega$ is the conventional notation or an event.

1. Generalities about random variables

1.1. Basic definitions.

**Definition 1.1.** A random variable is a function $X : \Omega \to \mathbb{R}$ with the property that $\{ \omega \in \Omega, X(\omega) \leq x \} \in \mathcal{F}$ for each $x \in \mathbb{R}$.

**Remark 1.1.** We may pay no attention to that property for the moment. Moreover if $X$ is a random variable and $f$ is a real function then $f(X)$ is a random variable.

**Definition 1.2.** The distribution function of a random variable $X$ is the function $F(x) = P(X \leq x)$.

1.2. Properties of the distribution function.

**Lemma 1.1.** Any distribution function $F$ has the following properties

1. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$,
2. if $x < y$ then $F(x) < F(y)$,
3. $F$ is right-continuous, that is $F(x + h) \to F(x)$ when $h \to 0^+$.

Conversely any such function is the distribution function of a probability.

**Lemma 1.2.** Let $F$ be the distribution function of $X$. Then

1. $P(X > x) = 1 - F(x)$,
2. $P(x < X \leq y) = F(y) - F(x)$,
3. $P(X = x) = F(x) - \lim_{y \to x^-} F(y)$.

2. Discrete random variables

**Definition 2.1.** A random variable $X$ is called discrete if it takes values in some countable set.

The distribution function is determined by the mass function.
Definition 2.2. The mass function of a discrete random variable $X$ is the function $f(x) = P(X = x)$.

2.1. Independence vs dependence.

Definition 2.3. Discrete variables $X_1, X_2, \ldots, X_n$ are independent if the events $\{X_1 = x_1\}, \ldots, \{X_n = x_n\}$ are independent for all $x_1, \ldots, x_n$.

Definition 2.4. The joint distribution function $F : \mathbb{R}^2 \to [0, 1]$ of $X$ and $Y$ where $X$ and $Y$ are discrete random variables, is given by

$$F(x, y) = P(X \leq x \text{ and } Y \leq y).$$

Their joint mass function $f : \mathbb{R}^2 \to [0, 1]$ is given by

$$f(x, y) = P(X \leq x \text{ and } Y \leq y).$$

Theorem 2.1. If $X$ and $Y$ are independent and $g, h$ are two real functions, then $g(X)$ and $h(Y)$ are also independent.

Theorem 2.2. Two random variables $X$ and $Y$ are independent if, and only if, $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ for all $x, y$. More generally they are independent if $f_{X,Y}(x, y)$ can be factorized as a product $h(x)h(y)$.

2.2. Expectation.

Definition 2.5. The expectation, or expected value, or mean value of $X$ with mass function $f$ is defined by

$$E[X] = \sum_x xf(x).$$

whenever the sum is absolutely convergent.

Theorem 2.3. We have

1. If $X \geq 0$, then $E[X] \geq 0$,
2. if $a, b \in \mathbb{R}$ then $E[aX + bY] = aE[X] + bE[Y]$,
3. $E[1\{A\}] = P[A]$,

Definition 2.6. If $k$ is a positive integer, then the $k$-th moment of $X$ is $E[X^k]$. Moreover the variance $\operatorname{var}(X)$ of $X$ is


Proposition 2.1. If $X$ and $Y$ are independent, then

1. $E[XY] = E[X]E[Y]$,
2. $\operatorname{var}(X + Y) = \operatorname{var}(X) + \operatorname{var}(Y)$. 