Appendix B
Probability

Probability Spaces and Random Variables

A *probability space* is defined by a triple $(\Omega, \mathcal{F}, P)$, where Ω is a given set of elementary outcomes, $\mathcal{F}$ is a collection of subsets of Ω (each such subset B is called an *event*), and $P(\cdot)$ is a *probability measure* that assigns a nonnegative number $P(B)$ to each subset B in $\mathcal{F}$.

The collection of subsets $\mathcal{F}$ must satisfy

- If B is in $\mathcal{F}$, then so is its complement $\bar{B} = \{\omega \in \Omega : \omega \notin B\}$.
- If $B_1, B_2, \ldots$ are events in $\mathcal{F}$, then $\cup_k B_k$ and $\cap_k B_k$ are also in $\mathcal{F}$.

The probability measure must satisfy

- $P(B) \geq 0$ for all $B \in \mathcal{F}$.
- $P(\Omega) = 1$
- If $B_1, B_2, \ldots$ are disjoint events, the $P(\cup_k B_k) = \sum_k P(B_k)$.

A *random variable* is a function mapping elementary outcomes to real numbers, $X : \Omega \to \Re$ and is denoted $X(\omega)$—or simply X, where the dependence on ω is implicit. The *cumulative distribution function* (c.d.f.) of a random variable X—or just *distribution function* for short—is defined by

$$F(x) = P(X \leq x).$$

If X takes on only countable values, we define the *probability-mass function* (pmf) by the function

$$P(x) = P(X = x).$$

Such a random variable is said to be *discrete.* If F is differentiate, then the *probability-density function* is defined by

$$f(x) = \frac{\partial}{\partial x} F(x).$$

Such a random variable is said to be *continuous.*

Let $\mathbf{X} = (X_1, \ldots, X_n)$ denote a vector of random variables and $\mathbf{x} = (x_1, \ldots, x_n)$ a real vector. Then we define the *joint distribution* of X by

$$F(\mathbf{x}) = P(X_1 \leq x_1, \ldots, X_n \leq x_n).$$