Review: Probability and Statistics

Basic Probability

A probability model consists of three components:

(1) A set Ω of elementary outcomes called the sample space.
(2) A set $\mathcal{G}$ of possible events (subsets of Ω).
(3) A probability function P that assigns probabilities (real numbers between 0 and 1) to events in $\mathcal{G}$.

The axioms of probability are:

(1) $P(A) \geq 0$ for all A, $\quad P(\Omega) = 1$.
(2) If A and B are disjoint events, then $P(A \cup B) = P(A) + P(B)$.

Important properties:

$$P(\Omega - A) = 1 - P(A), \qquad P(\emptyset) = 0, \qquad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The conditional probability of A given B is defined as follows:

$$P(A|B) = P(A \cap B)/P(B)$$

Important properties of conditional probabilities:

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B), \qquad P(A|B) = P(A)P(B|A)/P(B)$$

Two events are said to be independent if the following three (equivalent) conditions hold:

$$P(A \cap B) = P(A)P(B), \qquad P(A) = P(A|B), \qquad P(B) = P(B|A)$$

Discrete Random Variables

A random variable X is a real-valued function on the sample space Ω. If the range of X is finite or countable, X is called discrete, otherwise X is called continuous. For discrete random variables, the expectation value $\mathbb{E}(X)$ of X (often denoted as μ_X) , the variance $\mathrm{Var}(X)$, and the standard deviation σ_X are defined as follows:

$$\mu_X \equiv \mathbb{E}(X) = \sum_x xP(X = x), \qquad \sigma_X^2 = \mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \mathbb{E}[X^2] - (\mu_X)^2.$$

Important Discrete Distributions

(1) Binomial Distribution

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \mathbb{E}(X) = np, \qquad \mathrm{Var}(X) = np(1-p)$$

(2) Poisson Distribution

$$P(X = k) = \frac{\mu^k}{k!}\mathrm{e}^{-\mu}, \qquad \mathbb{E}(X) = \mu, \qquad \mathrm{Var}(X) = \mu$$

Continuous Random Variables

A probability distribution function f and the cumulative distribution function F have the following properties: (h is an integrable real function)

$$P(\{X \leq x\}) = \int_{-\infty}^{x} f(s)ds = F(x), \qquad \mathbb{E}(h(X)) = \int_{-\infty}^{\infty} h(x)f(x)dx$$

Important Continuous Distributions

(1) Gaussian (Normal) Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \qquad \mathbb{E}(X) = \mu, \qquad \sigma_X = \sigma, \qquad \mathbb{E}(\exp(\theta X)) = \exp\left(\theta\mu + \frac{1}{2}\theta^2\sigma^2\right)$$

(2) Log-normal Distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\frac{\exp\left(\frac{-(\ln x-\mu)^2}{2\sigma^2}\right)}{x}, \qquad \mathbb{E}(X) = \mathrm{e}^{\mu+\sigma^2/2}, \qquad \sigma_X^2 = \mathrm{e}^{\mu+\sigma^2/2}(\mathrm{e}^{\sigma^2} - 1)$$