Probability of Univariate Random Variables

- **Random Experiment and its Sample Space**

  A *random experiment* is a procedure that can be repeated an infinite number of times and has a set of possible outcomes. The *sample space* S of a certain random experiment is the totality of all its possible outcomes.

- **Random Event and its Probability**

  An *random event* A is a subset of S, $A \subset S$. A can be a null set $\emptyset$ (empty set), proper subset (e.g., a single outcome), or the entire S. Event A occurs if the outcome s is a member of A, $s \in A$. The probability of an event A, $P(A)$, is a real-valued function that maps A to a real number $0 \leq P(A) \leq 1$. In particular, $P(\emptyset) = 0$, $P(S) = 1$.

- **Probability Space**

  The triple of sample space, events and probability is called the probability space.

- **Conditional Probability**

  The *conditional probability* $P(A/B)$ is the probability that event A will occur given that event B has already occurred. If event A is *independent* of event B, then $P(A/B) = P(A)$.

- **Intersection and Union of Events**

  Let A and B be two events defined over S.

  - The *Intersection of A and B* is an event whose outcomes belong to both A and B, $A \cap B$.

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

    If A and B are *independent*, i.e.,

    $$P(A/B) = P(A), \quad \textit{and} \quad P(B/A) = P(B)$$

    then

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

- The *Union of A and B* is an event whose outcomes belong to either A or B, $A \cup B$.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

If A and B are *mutually exclusive*, i.e.,

$$A \cap B = \emptyset$$

then

$$P(A \cup B) = P(A) + P(B)$$

- **Bayes' Theorem**

Let $\{A_i, \ \ (i = 1, \cdots, n)\}$ be a set of n events that partition the sample space S in such a way that

$$\bigcup_{i=1}^{n} A_i = S$$

and

$$A_i \cap A_j = \emptyset \quad for\ any \ \ i \neq j$$

then for any event $B \subset S$, we have

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$

Proof

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B/A_i)P(A_i)}{P(B)}$$

but

$$P(B) = P(B \cap S) = P(B \cap (\bigcup_{i=1}^{n} A_i)) = P(\bigcup_{i=1}^{n}(B \cap A_i)) = \sum_{i=1}^{n} P(B \cap A_I)$$

Substituting this expression of $P(B)$ into the previous equation, we get the Bayes' theorem.

- **Random Variables**

  A random variable $X = X(s)$ is a real-valued function whose domain is the sample space S. In other words, this function maps every outcome $s \in S$ into a real number X. Random variables X can be either continuous or discrete.

- **Cumulative Distribution Function**

  The *cumulative distribution function* of a random variable X is defined as

  $$F_X(x) \stackrel{\triangle}{=} P(X < x)$$

  and we have

  $$P(a \le X < b) = F_x(b) - F_x(a)$$

- **Density Function**

  The density function $p(x)$ of a random variable X is defined by

  $$F_X(x) = \int_{-\infty}^{x} p(\xi) d\xi$$

  i.e.,

  $$p(x) = \frac{d}{dx} F_X(x)$$

  and we have

  $$P(a \le X < b) = F_X(b) - F_X(a) = \int_a^b p(\xi) d\xi$$

  and

  $$\int_{-\infty}^{\infty} p(\xi) d\xi = 1$$

- **Discrete Random Variables**

  If a random variable X can only take one of a set of finite number of discrete values $\{x_i \quad i = 1, \cdots, n\}$, then its *probability distribution* is

  $$P(X = x_i) \stackrel{\triangle}{=} p(x_i) = p_i \quad (i = 1, \cdots, n)$$

  and we have

  $$0 \le p_i \le 1$$

and

$$\sum_{i=1}^{n} p_i = 1$$

The cumulative distribution function is

$$F_X(x) = P(X < x) = \sum_{x_i < x} p(x_i) \stackrel{*}{=} \sum_{i=1}^{k} p_i$$

where the last equation assumes $x_1 < \cdots < x_n$ and $x = x_{k+1}$.

- **Expectation**

  The *expectation* or *mean value* of a random variable X is defined as

$$\mu = \overline{X} = E(X) \triangleq \int_{-\infty}^{\infty} xp(x)dx$$

  if X is continuous, or

$$\mu = \overline{X} = E(X) \triangleq \sum_{i=1}^{n} x_i p_i$$

  if X is discrete. ($E(X)$ is the weighted average in this case.)

- **Variance**

  The *variance* of a random variable X is defined as

$$\sigma^2 = Var(X) \triangleq E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$$

  if X is continuous, or

$$\sigma^2 = Var(X) \triangleq E[(X-\mu)^2] = \sum_{i=1}^{n} (x_i-\mu)^2 p_i$$

  if X is discrete.

  The *standard deviation* of X is defined as

$$\sigma = \sqrt{Var(X)}$$

  We have

$$\begin{aligned} \sigma^2 = Var(X) &= E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

- **Normal (Gaussian) Distribution**

  Random variable X has a *normal distribution* if its density function is

$$p(x) = N(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(\frac{x-\mu}{\sigma})^2}$$

  It can be shown that

$$\int_{-\infty}^{\infty} N(x, \mu, \sigma) dx = 1$$

$$E(X) = \int_{-\infty}^{\infty} x N(x, \mu, \sigma) dx = \mu$$

  and

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 N(x, \mu, \sigma) dx = \sigma^2$$

- **Function of a Random Variable**

  A function $Y = f(X)$ of a random variable X is also a random variable. If the density function of X is $p_X(\xi)$, the density function of Y can be found as shown below.

  Consider the cumulative distribution function of Y:

$$F_Y(y) = P(Y < y) = P(f(X) < y) \stackrel{*}{=} P(X < f^{-1}(y)) = \int_{-\infty}^{x} p_X(\xi) d\xi$$

  where $x \stackrel{\triangle}{=} f^{-1}(y)$ can be considered as a dummy variable. Then we can obtain the density function of Y as:

$$p_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{-\infty}^{x} p_X(\xi) d\xi = \left[p_X(x) \frac{dx}{dy} \right]_{x=f^{-1}(y)}$$

  The equation marked by * is based on the assumption that $Y = f(X)$ is a monotonically increasing function ($dy/dx > 0$). If it is monotonically decreasing ($dy/dx < 0$), we have

$$F_Y(y) = P(Y < y) = P(f(X) < y) = P(X > f^{-1}(y)) = \int_{x}^{\infty} p_X(\xi) d\xi$$

  and

$$p_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \int_{x}^{\infty} p_X(\xi) d\xi = \left[-p_X(x) \frac{dx}{dy} \right]_{x=f^{-1}(y)}$$

Considering both cases, if $Y = f(X)$ is monotonic, we have

$$p_Y(y) = \left[p_X(x) \left| \frac{dx}{dy} \right| \right]_{x=f^{-1}(y)}$$