Probability

Chi-chao Chao

Department of Electrical Engineering
National Tsing Hua University
ccc@ee.nthu.edu.tw

Probability Space

Definition

A probability space consists of three parts $(\Omega, \mathcal{F}, P)$*:*

1. *A sample space* Ω*, which is the set of all possible outcomes.*
2. *A set of events* $\mathcal{F}$*, where each event is a subset of* Ω*.*
3. *A probability measure* P *for each* $A, B \in \mathcal{F}$ *such that*

$$P(\Omega) = 1$$

$$0 \leq P(A) \leq 1$$

$$P(A \cup B) = P(A) + P(B) \quad \textit{if} \quad A \cap B = \emptyset.$$

- The set of events $\mathcal{F}$ forms a σ-algebra (or σ-field):
  1. $\mathcal{F}$ is nonempty.
  2. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
  3. If $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Random Variable

Definition

A (real) random variable $X : \Omega \to \mathcal{R}$ is a function from the set of all possible outcomes Ω to the set of real numbers $\mathcal{R}$.

- The event $\{X \leq x\}$ is the set of outcomes $\{\xi \in \Omega : X(\xi) \leq x\}$.
- Cumulative distribution function (CDF):

$$F_X(x) = P(X \leq x) = P(\{\xi : X(\xi) \leq x\}).$$

- Probability density function (PDF):

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

where $f_X(x) \geq 0$, for all x.

Example

- Consider the experiment of rolling two fair dice.
- The set of all possible outcomes

$$\Omega = \{(d_1, d_2) : d_1, d_2 = 1, 2, 3, 4, 5, 6\}.$$

- The set of events $\mathcal{F} = 2^{\Omega}$ (the power set of Ω).
- The probability measure $P(\xi) = 1/36$, for each outcome $\xi \in \Omega$.
- Define the random variable X as the sum of two fair dice.
- The event

$$\{X \leq 4\} = \{(1,1), (1,2), (2,1), (1,3), (2,2), (3,1)\}$$

and

$$P(X \leq 4) = 6/36 = 1/6.$$

- We have

$$\begin{aligned} P(x_1 < X \le x_2) &= P(X \le x_2) - P(X \le x_1) \\ &= F_X(x_2) - F_X(x_1) \\ &= \int_{x_1}^{x_2} f_X(x)\, dx. \end{aligned}$$

- Also

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$$

and

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1.$$

Expectation

- Mean:

$$\mu_X = \mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx.$$

- If $Y = g(X)$,

$$\mathrm{E}[Y] = \mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

- Variance:

$$\mathrm{Var}[X] = \mathrm{E}[(X - \mu_X)^2] = \mathrm{E}[X^2] - \mu_X^2.$$

Gaussian (Normal) Random Variable

- Its probability density function is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad \text{for } -\infty < x < \infty.$$

- Its mean

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx = \mu$$

and variance

$$\mathrm{Var}[X] = \sigma^2.$$

- This random variable is usually denoted as $\mathcal{N}(\mu, \sigma^2)$.

Standard Gaussian (Normal) Random Variable

- If $\mu = 0$ and $\sigma^2 = 1$, then it is the standard Gaussian (normal) random variable $\mathcal{N}(0, 1)$ with probability density function

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$

- The cumulative distribution function

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt.$$

- The Q-function is defined as

$$Q(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt = 1 - F_X(x)$$

with $Q(-x) = 1 - Q(x)$.

- The complementary error function

$$\operatorname{erfc}(x) = \int_{x}^{\infty} \frac{2}{\sqrt{\pi}} \exp(-t^2)\, dt \quad \text{and} \quad Q(x) = \frac{1}{2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right).$$

Two Random Variables

- Given two random variables X and Y, the joint cumulative distribution function

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

and the joint probability density function

$$f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

with

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1.$$

- The marginal probability density functions

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx.$$

- The conditional probability density function

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

with

$$f_{Y|X}(y|x) \geq 0$$

and

$$\int_{-\infty}^{\infty} f_{Y|X}(y|x)\,dy = 1.$$

Transformation of Random Variables

- Let X and Y be continuous random variables with joint probability density function $f(x, y)$.
- Let $U = h_1(X, Y)$ and $V = h_2(X, Y)$.
- Suppose $u = h_1(x, y)$ and $v = h_2(x, y)$ define a one-to-one transformation, i.e., the inverse $x = w_1(u, v)$ and $y = w_2(u, v)$ exists.
- The joint probability density function of random variables U and V is given by

$$g(u, v) = f\left(w_1(u, v), w_2(u, v)\right)|J|$$

where J is the Jacobian of the transformation defined by

$$J = \begin{vmatrix} \frac{\partial w_1}{\partial u} & \frac{\partial w_1}{\partial v} \\ \frac{\partial w_2}{\partial u} & \frac{\partial w_2}{\partial v} \end{vmatrix} = \frac{\partial w_1}{\partial u}\frac{\partial w_2}{\partial v} - \frac{\partial w_1}{\partial v}\frac{\partial w_2}{\partial u}.$$

Independence and Uncorrelatedness

- Two random variables X and Y are independent if

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

  which implies

$$f_{Y|X}(y|x) = f_Y(y), \quad \text{for } f_X(x) > 0.$$

- Two random variables X and Y are uncorrelated if the covariance

$$\mathrm{Cov}[X,Y] = \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \mathrm{E}[XY] - \mu_X\mu_Y = 0.$$

- If X and Y are independent, then they are uncorrelated, but not vice versa.

Bivariate Gaussian Distribution

- If X and Y are jointly Gaussian random variables, then

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}.$$

- The marginal probability density functions

$$f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} \exp\left[-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right]$$

$$f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} \exp\left[-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right].$$

- The covariance

$$\mathrm{Cov}[X, Y] = \mathrm{E}[(X - \mu_X)(Y - \mu_Y)] = \rho\sigma_X\sigma_Y$$

where ρ is the correlation coefficient with $-1 \leq \rho \leq 1$.

- Since for two jointly Gaussian random variables X and Y

$$\rho = 0 \iff f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

they are uncorrelated if and only if they are independent.

Multivariate Gaussian Distribution

- If $X_1, X_2, \ldots, X_N$ are jointly Gaussian random variables, then the joint probability density function

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{C}}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{C}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right].$$

- The vectors

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_N \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdots \\ \mu_N \end{bmatrix}$$

  where $\mu_i = \mathrm{E}[X_i]$, for $i = 1, 2, \ldots, N$.

- The (i, j)-th entry of the covariance matrix $\boldsymbol{C}$ is

$$C_{i,j} = \mathrm{Cov}[X_i, X_j] = \mathrm{E}[(X_i - \mu_i)(X_j - \mu_j)].$$

- If X_1 and X_2 are jointly Gaussian random variables and

$$Y_1 = a_{11}X_1 + a_{12}X_2$$

$$Y_2 = a_{21}X_1 + a_{22}X_2$$

  then Y_1 and Y_2 are also jointly Gaussian random variables.
- Random variables resulting from linear combinations of jointly Gaussian random variables are still jointly Gaussian.