ST2334 - Two-Dimensional Random Variables and Conditional Probability Distributions

3.1 Two-Dimensional Random Variables

3.2 Joint Probability Density Functions

3.3 Marginal and Conditional Probability Distributions

3.4 Independent Random Variables

3.5 Expectation

3.1 Two-Dimensional Random Variables

- **Definition 3.1** : Let X and Y be functions each assigning a real no. to each $s \in S$.

  Then (X, Y) : two-dimensional random variable / random vector

- Range space : $R_{X,Y} = \{(x, y) \mid x = X(s),\ y = Y(s),\ s \in S\}$

- **Definition 3.2** : Let $X_1, X_2, \ldots, X_n$ be n functions each assigning a real no. to every outcome $s \in S$.

  Then $(X_1, X_2, \ldots, X_n)$ is an n-dimensional random variable or random vector

- **Definition 3.3** : The possible values of $(X(s), Y(s))$

  ① are finite or countable infinite

  $\longrightarrow$ (X, Y) is a two-dimensional discrete random variable

  ② can assume all values in some region of the Euclidean plane $\mathbb{R}^2$

  $\longrightarrow$ (X, Y) is a two-dimensional continuous random variable

3.2 Joint Probability Density Functions

3.2.1 Joint Probability Function for Discrete RVs

- **Definition 3.4** : Let (X, Y) be a 2-dimensional discrete random variable defined on the sample space of an experiment.

  $\forall (x_i, y_j)$ we have $f_{X,Y}(x_i, y_j) = Pr(X = x_i, Y = y_j)$ and satisfying :

  ① $f_{X,Y}(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{X,Y}$

  ② $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} Pr(X = x_i, Y = y_j) = 1$

- Joint probability function of (X, Y) : $f_{X,Y}(x,y)$ defined $\forall (x_i, y_j) \in R_{X,Y}$
- Let A be any set consisting of (x, y) values.

  Then $Pr((X,Y) \in A) = \underbrace{\sum\sum}_{(x,y) \in A} f_{X,Y}(x,y)$

3.2.2 Joint Probability Density Function for Continuous RVs

- Let (X, Y) be a 2-dimensional continuous random variable assuming all values in some region R of the Euclidean plane, $\mathbb{R}^2$.

  $\forall (x_i, y_j)$ we have joint pdf $f_{X,Y}(x_i, y_j)$ if it satisfies :

  ① $f_{X,Y}(x_i, y_j) \geq 0 \quad \forall (x_i, y_j) \in R_{X,Y}$

  ② $\iint_{(x,y) \in R_{X,Y}} f_{X,Y}(x,y)\, dx\, dy = 1$

  OR

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

3.3 Marginal and Conditional Probability Distributions

3.3.1 Marginal Probability Distributions

- **Definition 3.6** : Let (X, Y) be a 2-dimensional discrete /continuous random variable with joint probability (density) function $f_{X,Y}(x,y)$.

  Marginal probability distributions of X and Y:

  ① Discrete : $f_X(x) = \sum_y f_{X,Y}(x,y)$ and $f_Y(y) = \sum_x f_{X,Y}(x,y)$

  ② Continuous : $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$ and $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

3.3.2 Conditional Distribution

- **Definition 3.7** : Let (X, Y) be a 2-dimensional discrete /continuous random variable with joint probability (density) function $f_{X,Y}(x,y)$.

  Let $f_X(x)$ and $f_Y(y)$ be the marginal probability functions of X and Y respectively.

  Then :

  ① **Conditional distribution of Y given that $X = x$:**

  $$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad \text{if } f_X(x) > 0$$

  for each x within the range of X.

  ② **Conditional distribution of X given that $Y = y$:**

  $$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}, \quad \text{if } f_Y(y) > 0$$

  for each y within the range of Y.

3.4 Independent Random Variables

3.4.1 Definition of Independent RVs

- Definition : Random variables X and Y are independent

  $\longleftrightarrow$ $f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \forall x,y$.

- Extension : Random variables $X_1, X_2, \cdots, X_n$ are independent

  $\longleftrightarrow$ $f_{X_1,X_2,\cdots,X_n}(x_1,x_2,\cdots,x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n) \quad \forall x_i, i=1,\cdots,n$.

3.5 Expectation

- **Definition 3.5.1** : Expectation of $g(X,Y)$:

$$E(g(X,Y)) = \begin{cases} \sum_x \sum_y g(x,y) f_{X,Y}(x,y) & \text{for discrete RVs} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy & \text{for continuous RVs} \end{cases}$$

- **Definition 3.5.2** : Let (X,Y) be a bivariate random vector with joint p.f./p.d.f. $f_{X,Y}(x,y)$.

Then covariance of (X,Y) : $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

① Discrete : $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

$$= \sum_x \sum_y (x-\mu_x)(y-\mu_y) f_{X,Y}(x,y)$$

② Continuous : $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

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

- Remarks : ① $Cov(X,Y) = E(XY) - \mu_X\mu_Y$

② X and Y are independent $\longrightarrow$ $Cov(X,Y) = 0$, but

$Cov(X,Y) = 0$ $\not\longrightarrow$ X and Y are independent

③ $Cov(aX+b, cY+d) = ac\, Cov(X,Y)$

④ $V(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\, Cov(X,Y)$

- **Definition 3.5.3** : Correlation coefficient of X and Y, denoted by $Cor(X,Y)$, $\rho_{X,Y}$ or ρ :

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

- Remarks : ① $-1 \leq \rho_{X,Y} \leq 1$

② $\rho_{X,Y}$: measure of the degree of linear relationship between X and Y

③ X and Y are independent $\longrightarrow$ $\rho_{X,Y} = 0$, but

$\rho_{X,Y} = 0$ $\not\longrightarrow$ X and Y are independent