Chapter 4 Bivariate Distributions (二元分布)

Section 4.1 Bivariate Distributions with the **discrete** type

➢ Motivation

very often, the outcome of a random experiment is a **tuple** of several things of interests:

- Observe female college students to obtain information such as ***height x***, and ***weight y***.
- Observe high school students to obtain information such as ***rank x***, and ***score of college entrance examination y***.

➢ In order to define joint probability mass function (joint *pmf*):

* **Complete way :**

①identify the Sample Space S;

②Define a RV $Z=\begin{bmatrix} X \\ Y \end{bmatrix}: S \to Z(S)$;

③Define a pmf for Z, $f(z): Z(S) \to [0,1]$.

* **Simplified way :**

①Ignore the Sample Space S;

②Specify $Z(S)$ directly and denote it by D;

③Define the pmf for Z, $f(z): D \to [0,1]$;

equivalently, for $\begin{bmatrix} X \\ Y \end{bmatrix}$, $f(x,y): D \to [0,1]$.

Definition 4.1-1 [joint probability mass function (joint *pmf*)]

Let X and Y be 2 RVs. The probability that $X = x$ and $Y = y$ is denoted by $f(x, y) = P(X = x,\ Y = y)$.

The function $f(x, y)$: D → [0,1] is called the **joint probability mass function** (**joint *pmf***) of X *if*:

① $0 \le f(x, y) \le 1$;

② $\sum_{(x,y)\in D} f(x, y) = 1$;

③ $P[(X,Y) \in A] \triangleq P(\{(x, y) \in A\}) = \sum_{(x,y)\in A} f(x, y), \quad A \subseteq D.$

Example 1 [Page 134]

Roll a pair of fair dice. The sample Space contains 36 outcomes. And let X denote the smaller outcome and Y the larger outcome on the die.

For instance, if the outcome is (3,2), then X=2, Y=3.

Obviously, $P(\{X = 2, Y = 3\}) = 1/36 + 1/36 = 2/36.$

$P(\{X = 2, Y = 2\}) = 1/36.$

Furthermore, the *joint pmf* of X and Y is: $f(x, y) = \begin{cases} 1/36, & 1 \le x = y \le 6 \\ 2/36, & 1 \le x < y \le 6 \end{cases}$

Definition 4.1-2 [Marginal *pmf*]

Let X and Y have the joint probability mass function $f(x, y) : D \to [0,1]$. Sometimes we are interested in the pmf of X *or* Y alone, which is called the **marginal probability mass function of** X or Y and defined by

$$f_X(x) = \sum_{y \in D_Y} f(x, y) = P(X = x), \qquad x \in D_X = \{\text{all possible values of } X \text{ in } D\}.$$

$$f_Y(y) = \sum_{x \in D_X} f(x, y) = P(Y = y), \qquad y \in D_Y = \{\text{all possible values of } Y \text{ in } D\}.$$

Definition 4.1-3 [independent Random Variables]

The random variables X and Y are **independent** if and only if, for every $x \in D_X$ and $y \in D_Y$,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

$A \cap B$ Event A Event B

or equivalently,

$$f(x, y) = f_X(x) f_Y(y).$$

otherwise, X and Y are said to be **dependent**.

Example 2 [Page 135]

Let the joint *pmf* of X and Y be defined by

$$f(x,y)=\frac{x+y}{21},\qquad x=1,2,3,\qquad y=1,2.$$

Check if *RV X* and *Y* are independent.

Solution :

$$f_X(x)=\sum_{y\in D_Y} f(x,y)=\sum_{y=1}^{2}\frac{x+y}{21}=\frac{2x+3}{21},\qquad x=1,2,3.$$

$$f_Y(y)=\sum_{x\in D_X} f(x,y)=\sum_{x=1}^{3}\frac{x+y}{21}=\frac{3y+6}{21},\qquad y=1,2.$$

$$f(x,y)=\frac{x+y}{21}\neq\frac{2x+3}{21}\cdot\frac{3y+6}{21}=f_X(x)f_Y(y)\Rightarrow X \text{ and } Y \text{ are dependent.}$$

What's the interpretation of $f_X(x)$ and $f_Y(x)$ and independence?

Consider the *conditional pmf* :

$$f(y|x)=P(Y=y|X=x)=\frac{f(x,y)}{f_X(x)};\ f(x|y)=P(X=x|Y=y)=\frac{f(x,y)}{f_Y(x)}$$

We will learn it formally in section 4.3

➢ Expectation and Value

Let X_1 and X_2 be discrete *RV* with their joint *pmf* $f(x_1, x_2) : \mathrm{D} \to [0,1]$. Consider a function $u(x_1, x_2)$ of x_1 *and* x_2. Then:

Expectations of functions of bivariate RVs are computed just as with univariate RVs.

(a) The **mathematical expectation** of $u(X_1, X_2)$, if exists, is given by

$$E\left[u(X_1, X_2)\right] = \sum_{(x_1, x_2) \in D} u(x_1, x_2) f(x_1, x_2).$$

$$E(X) = \sum_{(x,y)\in \bar{S}} xf(x,y) = \sum_{x \in \bar{S}_X} xf_X(x).$$

(b) If $u_i(X_1, X_2) = X_i$ for $i = 1, 2$, then

$$E\left[u_i(X_1, X_2)\right] = E(X_i) = u_i$$

is called the **mean** of X_i for $i = 1, 2$.

(c) If $u_i(X_1, X_2) = (X_i - u_i)^2$ for $i = 1, 2$, then

$$E\left[u_i(X_1, X_2)\right] = E\left[(X_i - u_i)^2\right] = \sigma_i^2 = Var(X_i)$$

is called the **variance** of X_i for $i = 1, 2$.

Example 1 [Page 134]—revisited

Recall that X and Y are discrete *RVs* with joint *pmf* $f(X,Y): \mathrm{D} \to [0,1]$ with $D_X = D_Y = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$

$$f(x,y) = \begin{cases} 2/36, & 1 \le x < y \le 6 \\ 1/36, & 1 \le x = y \le 6 \end{cases}$$

Compute $E(X+Y)$:

Solution :

$$E(X+Y) = \sum_{(x,y)\in D} (x+y) f(x,y) = \sum_{1\le x=y\le 6} (x+y)\cdot\frac{1}{36} + \sum_{1\le x<y\le 6} (x+y)\frac{2}{36}$$

$$= \sum_{x=1}^{6} 2x\cdot\frac{1}{36} + \sum_{x=1}^{6}\sum_{y=x+1}^{6} (x+y)\cdot\frac{2}{36} = \frac{252}{36}.$$

Work it by yourself!

Chapter 4 Bivariate Distributions (二元分布)

Section 4.2 The correlation coefficient

Let X_1 and X_2 be discrete *RV* with their joint *pmf* $f(x_1, x_2)$: D → [0,1].

recall that for $u(X,Y)$, its expectation $E\left[u(X,Y)\right] = \sum_{(x,y)\in D} u(x,y) f(x,y)$.

Definition 4.2-1 [Covariance of X and Y]

Take $u(X,Y) = [X - E(X)][Y - E(Y)]$

$$E\left[(X - E(X))(Y - E(Y))\right] = Cov(X,Y),$$

which is called the covariance of X and Y.

- Motivation: To study the relation between X and Y.

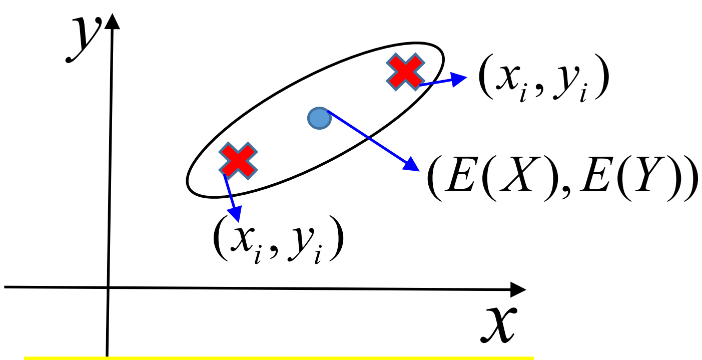
- $Cov(X,Y) = E(XY) - E(X)E(Y)$ → Verify it by yourself!

When $Cov(X,Y) = 0$, we say X *and* Y are ***uncorrelated***.

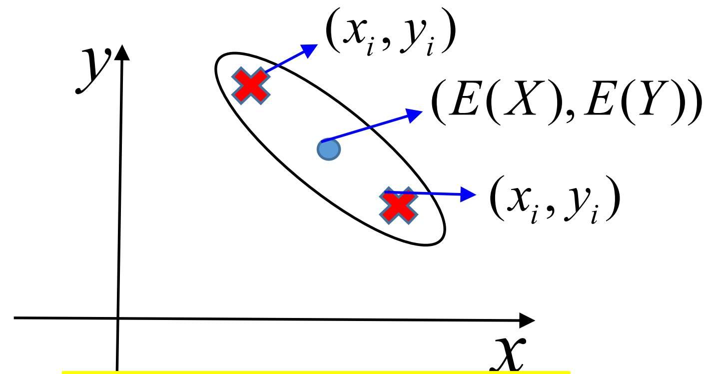
- Interpretation: Roughly speaking, a positive or negative covariance indicates that the values of $X - E(X)$ and $Y - E(Y)$ obtained in a single experiment ‘tend’ to have the same or the opposite sign.

Example 1: Demonstration of positively correlated and negatively correlated RVs

Assume that X and Y are uniformly distributed over the ellipses.



positively correlated



negatively correlated

Independence of X and Y could imply the **uncorrelation** of X and Y.

Consider the case that X and Y are independent:

$$E(XY) = \sum_{(x,y)\in D} xyf(x,y) = \sum_{x\in D_X}\sum_{y\in D_Y} xyf_X(x)f_Y(y)$$

$$= \sum_{x\in D_X} xf_X(x)\left[\sum_{y\in D_Y} yf_Y(y)\right] = E(X)E(Y).$$

$$f(x,y) = f_X(x)f_Y(y)$$
$$\Rightarrow D = D_X D_Y$$

Therefore, $cov(X,Y) = E(XY) - E(X)E(Y) = 0.$

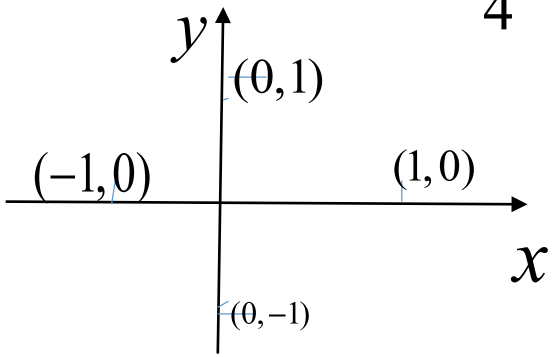
Independent of 2 RVs $\Rightarrow$ uncorrelation of 2 RVs.

However, the converse is not true, that is to say, there exists X and Y which are ***uncorrelated*** but ***not independent***.

Example 2 (uncorrelation **doesn't imply** independence)

Let X and Y be RVs that take values (1,0), (0,1), (-1,0), (0,-1)

and with probability $\frac{1}{4}$, as shown in the figure below.



Q1 : what are the **marginal pmf** of X and Y?

Q2 : what is $Cov(X,Y)$?

Q3 : Are X and Y *independent*?

Solution : To find marginal pmf of X and Y, $D_X = D_Y = \{-1, 0, -1\}$.

$$f_X(x) = \begin{cases} 1/4, & x = 1 \\ 1/2, & x = 0 \\ 1/4, & x = -1 \end{cases}, \qquad f_Y(y) = \begin{cases} 1/4, & y = 1 \\ 1/2, & y = 0 \\ 1/4, & y = -1 \end{cases}$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - 0 \cdot 0 = 0.$$

$$f_X(0) f_Y(1) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \neq f(0,1) = \frac{1}{4} \Rightarrow X \text{ and } Y \text{ are } \textit{not independent}!$$

Definition 4.2-2 [correlation coefficients]

The correlation coefficients of X and Y that have nonzero variance is defined as

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

- It is a normalized version of $Cov(X,Y)$ and in fact $-1 \leq \rho \leq 1$
- Interpretation: $\rho > 0$ $(or\ \rho < 0)$ indicate the values of $X - E(X)$ and $Y - E(Y)$ 'tend' to have the same(or opposite, respectively) sign.
- $\rho > 0$ $(or\ \rho < 0)$ have the same interpretation as $Cov(X,Y) > 0$ (or $Cov(X,Y) < 0$)
- The size of $|\rho|$ provides a normalized measure of the extent to which this is true.
- $\rho = 1$ or $\rho = -1$ if and only if there exists a positive (or negative, respectively) constant c such that

$$Y-E(Y)=c\left[X-E(X)\right]$$

Example 3

Consider n **independent** tosses of a coin with probability of a head equal to p. Let X and Y be the number if heads and of tails, respectively. Calculate the correlation coefficient of $\boldsymbol{X}$ and $\boldsymbol{Y}$**.**

Solution :

$$X + Y = n \Rightarrow E(X) + E(Y) = n \Rightarrow X - E(X) = -\left[Y - E(Y)\right]$$

$$Cov(X,Y) = E\left[(X - E(X))(Y - E(Y))\right] = -E\left[(Y - E(Y))^2\right] = -Var(Y)$$

$$Var(X) = E\left[(X - E(X))^2\right] = E\left[(Y - E(Y))^2\right] = Var(Y)$$

$$\Rightarrow \rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \frac{-Var(Y)}{\sqrt{Var(Y)}\sqrt{Var(Y)}} = -1.$$