Chapter 4 Bivariate Distributions

- Bivariate Distributions of the Discrete Type
- Correlation Coefficient
- Conditional Distributions
- Bivariate Distributions of the Continuous Type
- Bivariate Normal Distributions

Examples of Bivariate Distributions

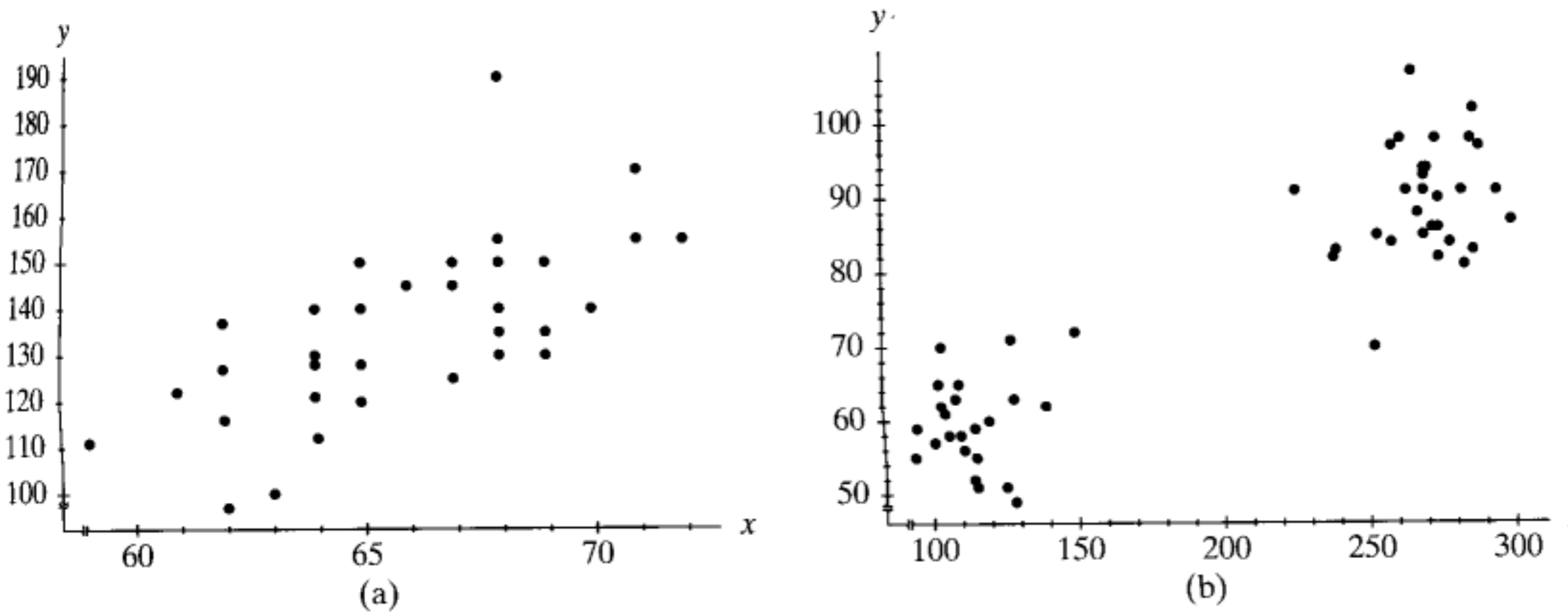


Figure 5.1-1 Plots of bivariate data: (a) (height, weight) for female college students, (b) (duration, time) for Old Faithful eruptions

1.5 Conditional Probability 條件機率

- **Definition 1.4-1**

  The conditional probability of an event A given that event B has occurred is defined by

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)},$$

provided that $P(B) > 0$.

REVIEW of Sec 1.5

Let A be the event $X_1 \geq 2$.

Let B denote the event $X_2 > X_1$.

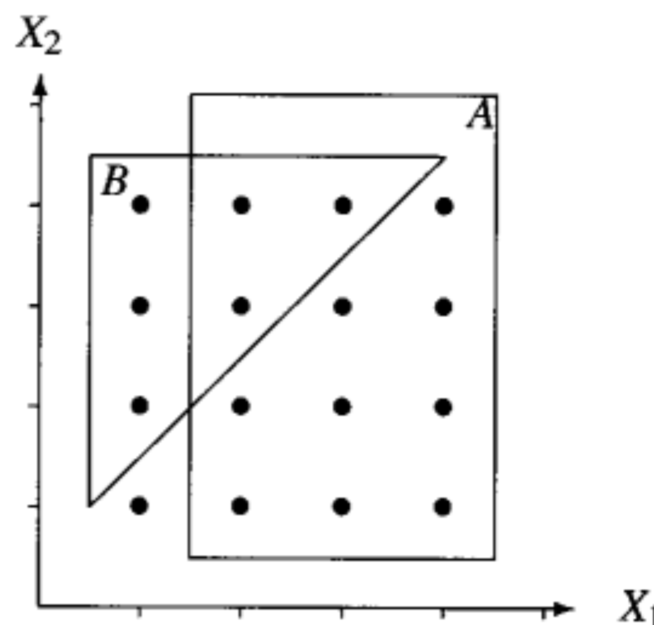
Example 1.18. *Roll two four-sided dice. Let X_1 and X_2 denote the number of dots that appear on die 1 and die 2, respectively. Draw the 4 by 4 sample space. Let A be the event $X_1 \geq 2$. What is $P[A]$? Let B denote the event $X_2 > X_1$. What is $P[B]$? What is $P[A|B]$?*

. . . .

Each outcome is a pair (X_1, X_2). To find $P[A]$, we add up the probabilities of the sample points in A.

From the sample space, we see that A has 12 points, each with probability $1/16$, so $P[A] = 12/16 = 3/4$. To find $P[B]$, we observe that B has 6 points and $P[B] = 6/16 = 3/8$. The compound event AB has exactly three points, $(2,3), (2,4), (3,4)$, so $P[AB] = 3/16$. From the definition of conditional probability, we write

$$P[A|B] = \frac{P[AB]}{P[B]} = 1/2$$



REVIEW of Sec 2.1

Properties of Probability Mass Function

$$P_X(k) \equiv Prob(X = k) = \sum_{q \in Q_k} Prob(q),$$

- The p.m.f. of a random variable X satisfies the following three properties:

$$(1)\ \ P_X(x) > 0\ ,\ \ x \in S\ :\ the\ \ space\ \ of\ \ X.$$

$$(2)\ \ \sum_{x_i \in S} P_X(x_i) = 1\ .$$

$P(X\ \varepsilon\ A)$

$$(3)\ \ Prob(A) = \sum_{x_j \in A} P_X(x_j)\ ,\ \ where\ \ A \subseteq S.$$

4.1 Bivariate Distributions of the Discrete Type

Joint Probability Mass Function

- Definition
  - Let X and Y be two random variables *defined on a discrete space.*
  - Let S denote the corresponding 2D space of X and Y *(the 2 discrete r.v.)*
  - The probability that $X=x$ and $Y=y$ is denoted by

$$f(x,y) \equiv Prob(X=x, Y=y)$$

    and is called the joint probability **mass** function (joint p.m.f.) of X and Y.
  - $f(x,y)$ satisfies the following 3 properties:

(a) $0 \le f(x,y) \le 1$.

(b) $\sum\sum_{(x,y)\in S} f(x,y) = 1$.

(c) $P[(X,Y) \in A] = \sum\sum_{(x,y)\in A} f(x,y)$, where A is a subset of the space S.

Example 4.1-1: joint p.m.f.

- Roll a pair of unbiased dice.
- For each of the 36 possible outcomes,
  - let X denote the smaller number
  - let Y denote the larger number.

→The joint p.m.f of X and Y is :

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

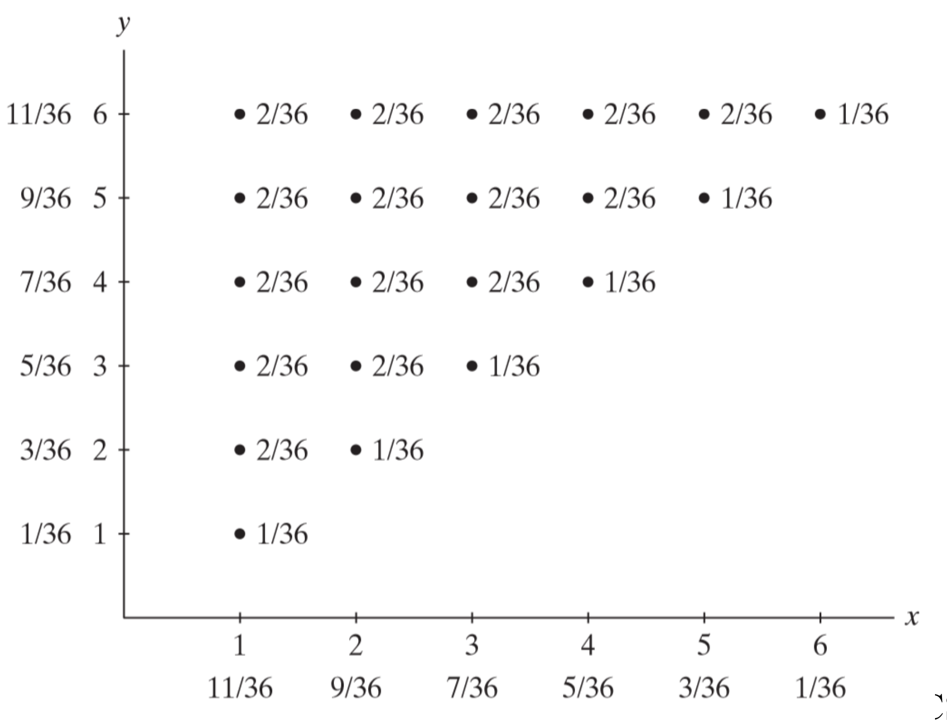
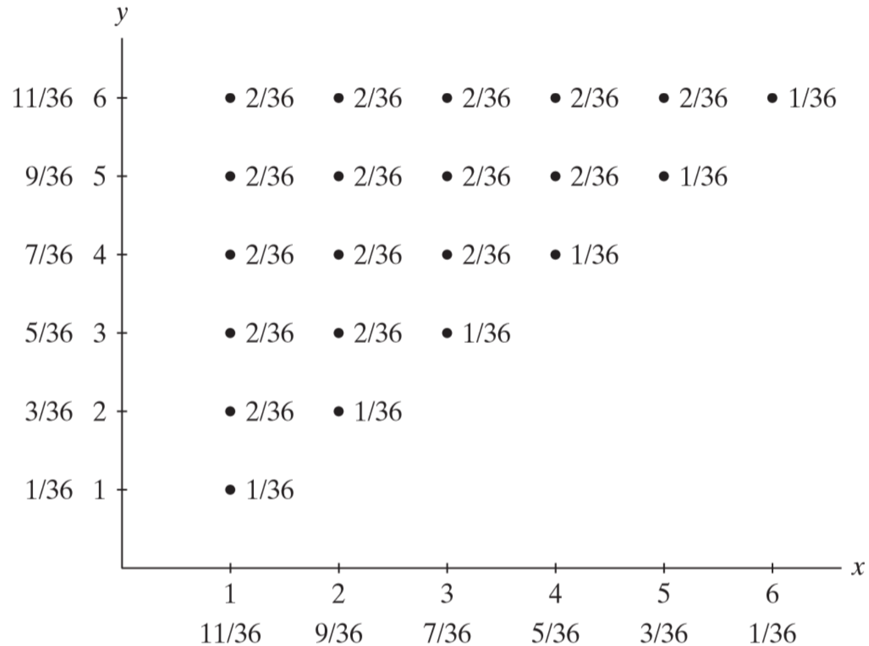


Figure 4.1-1 Discrete joint pmf



Notice that certain numbers have been recorded in the bottom and left-hand margins of Figure 4.1-1. These numbers are the respective column and row totals of the probabilities. The column totals are the respective probabilities that X will assume the values in the x space $S_X = \{1, 2, 3, 4, 5, 6\}$, and the row totals are the respective probabilities that Y will assume the values in the y space $S_Y = \{1, 2, 3, 4, 5, 6\}$. That is, the totals describe the probability mass functions of X and Y, respectively. Since each collection of these probabilities is frequently recorded in the margins and satisfies the properties of a pmf of one random variable, each is called a **marginal pmf.**

Definition 4.1-2

Let X and Y have the joint probability mass function $f(x, y)$ with space S. The probability mass function of X alone, which is called the **marginal probability mass function of** X, is defined by

$$f_X(x) = \sum_y f(x, y) = P(X = x), \qquad x \in S_X,$$

where the summation is taken over all possible y values for each given x in the x space S_X. That is, the summation is over all (x, y) in S with a given x value. Similarly, the **marginal probability mass function of** Y is defined by

$$f_Y(y) = \sum_x f(x, y) = P(Y = y), \qquad y \in S_Y,$$

where the summation is taken over all possible x values for each given y in the y space S_Y. The random variables X and Y are **independent** if and only if, for every $x \in S_X$ and every $y \in S_Y$,

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

or, equivalently,

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

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

Example 1 of Independent Random Variables

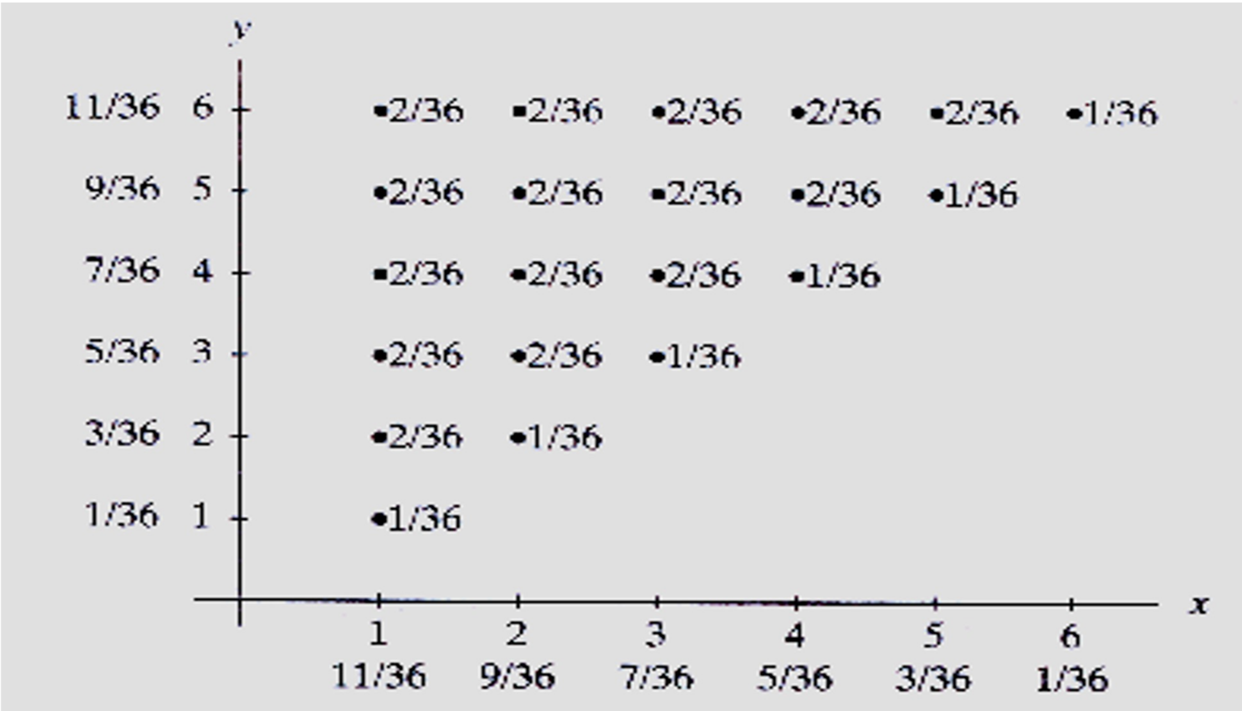
- Assume that we toss a coin two times.
- Let
  - random variable X corresponds to the outcome of the first tossing
  - random variable Y corresponds to the outcome of the second tossing.
- Then, X and Y are two independent random variables.

Example 2 of Independent Random Variables

- Assume that we randomly pick up a bridge card.
- Let
  - random variable X corresponds to the color of the card
  - a random variable Y corresponds to the number or figure on the card.
- Then, X and Y are independent.

Example 4.1-1 *Not Independent*

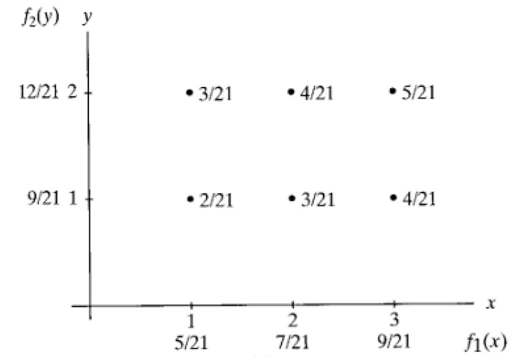
- Roll a pair of unbiased dice.
- For each of the 36 possible outcomes,
  - let X denote the smaller number
  - let Y denote the larger number



We note in Exampl 4.1-1 that X and Y are dependent because there are many x and y values for which $f(x, y) \neq f_1(x) f_2(y)$. For instance,

$$f_1(1) f_2(1) = \frac{11}{36} \cdot \frac{1}{36} \neq \frac{1}{36} = f(1, 1).$$

Example 4.1-2 *Not Independent*



Example 4.1-2 Let the joint p.m.f. of X and Y be defined by

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

Then

$$f_1(x) = \sum_y f(x, y) = \sum_{y=1}^{2} \frac{x + y}{21}$$

$$= \frac{x + 1}{21} + \frac{x + 2}{21} = \frac{2x + 3}{21}, \qquad x = 1, 2, 3;$$

and

$$f_2(y) = \sum_x f(x, y) = \sum_{x=1}^{3} \frac{x + y}{21} = \frac{6 + 3y}{21}, \qquad y = 1, 2.$$

Note that both $f_1(x)$ and $f_2(y)$ satisfy the properties of a probability mass function. Since $f(x, y) \neq f_1(x) f_2(y)$, X and Y are dependent. ▲

Example 4.1-3 *Independent*

Example 4.1-3 Let the joint p.m.f. of X and Y be

$$f(x, y) = \frac{xy^2}{30}, \qquad x = 1, 2, 3, \qquad y = 1, 2.$$

The marginal probability mass functions are

$$f_1(x) = \sum_{y=1}^{2} \frac{xy^2}{30} = \frac{x}{6}, \qquad x = 1, 2, 3,$$

and

$$f_2(y) = \sum_{x=1}^{3} \frac{xy^2}{30} = \frac{y^2}{5}, \qquad y = 1, 2.$$

Then $f(x, y) \equiv f_1(x) f_2(y)$ for $x = 1, 2, 3$, and $y = 1, 2$; thus X and Y are independent. See Figure 4.1-2 ▲

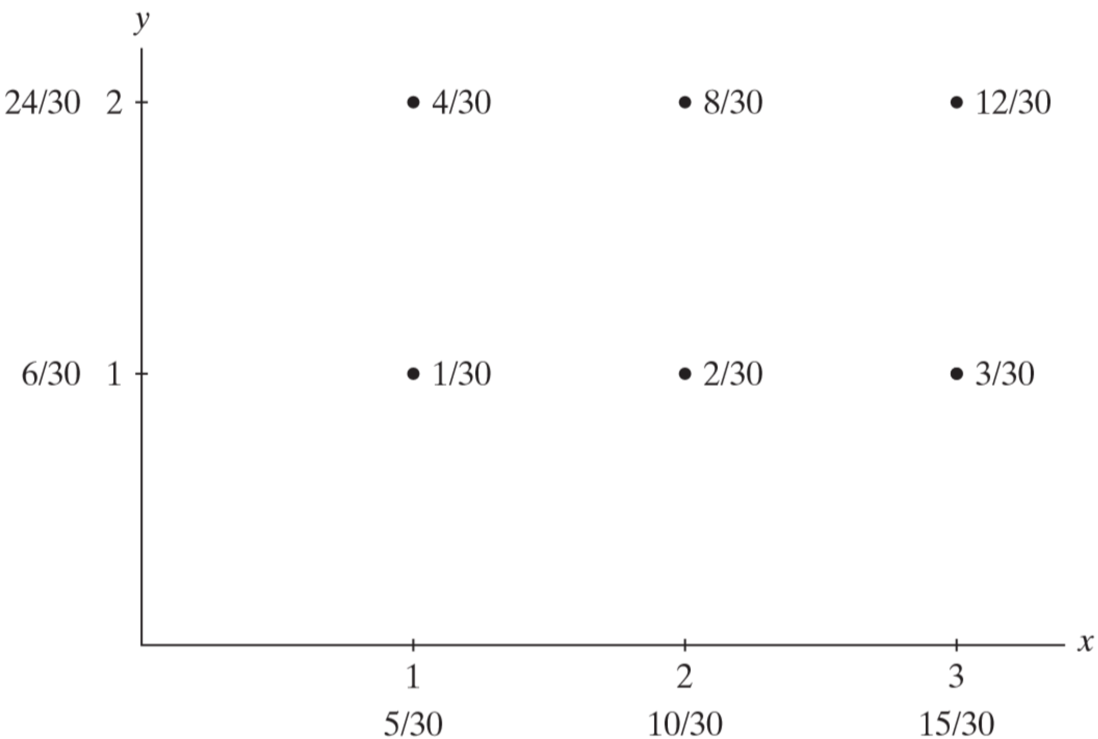


Figure 4.1-2 Joint pmf $f(x, y) = \dfrac{xy^2}{30}$, $x = 1, 2, 3$ and $y = 1, 2$

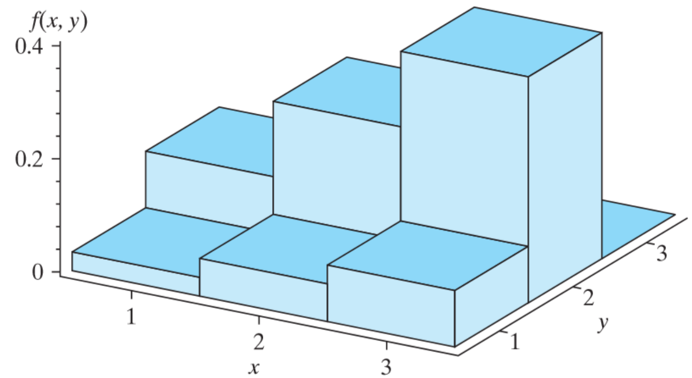


Figure 4.1-3 Joint pmf $f(x, y) = \dfrac{xy^2}{30}$, $x = 1, 2, 3$ and $y = 1, 2$

Example 4.1-4 Let the joint p.m.f. of X and Y be

$$f(x, y) = \frac{xy^2}{13}, \qquad (x, y) = (1, 1), (1, 2), (2, 2).$$

Then the p.m.f. of X is

$f(2.1) = 0$

$$f_1(x) = \begin{cases} \dfrac{5}{13}, & x = 1, \\[2ex] \dfrac{8}{13}, & x = 2, \end{cases}$$

and that of Y is

$$f_2(y) = \begin{cases} \dfrac{1}{13}, & y = 1, \\[2ex] \dfrac{12}{13}, & y = 2. \end{cases}$$

X and Y are dependent because $f(x, y) \not\equiv f_1(x) f_2(y)$ for $x = 1, 2$ and $y = 1, 2$. ▲

$0 = f(2,1) \neq f_1(2) \cdot f_2(1)$

2.2 Mathematical Expectation

Expected Value of a Discrete Random Variable

- Let X be a discrete random variable and S be its space. Then, the expected value of X is

$$E[X] \equiv \sum_{x_i \in S} x_i P_X(x_i)$$

- μ is a widely used symbol for expected value.

Expected Value of a Function of a Random Variable

- Let X be a random variable and $u(.)$ be a function. Then, the expected value of random variable $Y = u(X)$ is equal to

$$E[Y] = E[u(X)] \underset{?}{\equiv} \sum_{x_i \in S} u(x_i) P_X(x_i)$$

consistency

In textbook, $E[u(X)] \equiv \sum_{x \in S} u(x) f(x)$

Back to Chapter 4

Expected Value of a Function of Two Random Variables

- Let X_1, X_2 be discrete random variables with the joint p.m.f. $f(x_1, x_2)$ on the space S.
- The expected value (mathematical expectation) of the random variable $Y = u(X_1, X_2)$ is

$$E[Y] = E[u(X_1, X_2)] = \sum_{(x_1, x_2) \in S} \sum u(x_1, x_2) f(x_1, x_2)$$

The above expression can be a property, rather than a definition

Expectation: mean, variance, etc.

Let X_1 and X_2 be random variables of the discrete type with the joint pmf $f(x_1, x_2)$ on the space S. If $u(X_1, X_2)$ is a function of these two random variables, then

$$E[u(X_1, X_2)] = \sum\sum_{(x_1,x_2)\in S} u(x_1, x_2) f(x_1, x_2),$$

if it exists, is called the **mathematical expectation** (or **expected value**) of $u(X_1, X_2)$.

(a) If $u_1(X_1, X_2) = X_i$, then

$$E[u_1(X_1, X_2)] = E(X_i) = \mu_i$$

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

(b) If $u_2(X_1, X_2) = (X_i - \mu_i)^2$, then

$$E[u_2(X_1, X_2)] = E[(X_i - \mu_i)^2] = \sigma_i^2 = \mathrm{Var}(X_i)$$

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

Chapter 4 Bivariate Distributions

- Bivariate Distributions of the Discrete Type
- Correlation Coefficient
- Conditional Distributions
- Bivariate Distributions of the Continuous Type
- Bivariate Normal Distributions

$$E[u_1(X_1, X_2)] = E(X_i) = \mu_i$$

$$E[u_2(X_1, X_2)] = E[(X_i - \mu_i)^2] = \sigma_i^2 = \mathrm{Var}(X_i)$$

4.2 Correlation Coefficient

$$\mu_i = E(X_i) \quad \text{and} \quad \sigma_i^2 = E[(X_i - \mu_i)^2], \quad i = 1, 2.$$

We introduce two more special names:

(a) If $u_3(X_1, X_2) = (X_1 - \mu_1)(X_2 - \mu_2)$, then

$$E[u_3(X_1, X_2)] = E[(X_1 - \mu_1)(X_2 - \mu_2)] = \sigma_{12} = \mathrm{Cov}(X_1, X_2)$$

is called the **covariance** of X_1 and X_2.

non-zero

(b) If the standard deviations σ_1 and σ_2 are positive, then

$$\rho = \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$$

is called the **correlation coefficient** of X_1 and X_2.

$$E[u(X_1, X_2)] \equiv \sum \sum_{(x_1, x_2) \in S} u(x_1, x_2)\, f(x_1, x_2)$$

- Mean and variance of X_1 can be computed from either the joint p.m.f. or the marginal p.m.f. of X_1

For example,

$$\mu_1 = E(X_1) = \sum_{x_1} \sum_{x_2} x_1 f(x_1, x_2)$$

$$= \sum_{x_1} x_1 \left[\sum_{x_2} f(x_1, x_2) \right] = \sum_{x_1} x_1 f_1(x_1).$$

- Computation of Covariance needs the joint p.m.f.

$$E[(X_1 - \mu_1)(X_2 - \mu_2)] = E(X_1 X_2 - \mu_1 X_2 - \mu_2 X_1 + \mu_1 \mu_2)$$

$$= E(X_1 X_2) - \mu_1 E(X_2) - \mu_2 E(X_1) + \mu_1 \mu_2$$

$$\Rightarrow \mathrm{Cov}(X_1, X_2) = E(X_1 X_2) - \mu_1 \mu_2 - \mu_2 \mu_1 + \mu_1 \mu_2 = E(X_1 X_2) - \mu_1 \mu_2.$$

Example 4.2-1

- Given a joint p.m.f. of two r.v.s
- Compute
  - Marginal p.m.f.s of each r.v.
  - Mean and variance of each r.v.
  - Covariance of two r.v.s
  - Correlation coefficient

Example 4.2-1

Let X and Y have the joint pmf

$$f(x, y) = \frac{x + 2y}{18}, \qquad x = 1, 2, \qquad y = 1, 2.$$

The marginal probability mass functions are, respectively,

$$f_X(x) = \sum_{y=1}^{2} \frac{x + 2y}{18} = \frac{2x + 6}{18} = \frac{x + 3}{9}, \qquad x = 1, 2,$$

and

$$f_Y(y) = \sum_{x=1}^{2} \frac{x + 2y}{18} = \frac{3 + 4y}{18}, \qquad y = 1, 2.$$

Since $f(x, y) \not\equiv f_X(x) f_Y(y)$, X and Y are dependent.

Since $f(x, y) \neq f_X(x) f_Y(y)$, X and Y are dependent. The mean and the variance of X are, respectively,

$$\mu_X = \sum_{x=1}^{2} x \frac{x+3}{9} = (1)\left(\frac{4}{9}\right) + (2)\left(\frac{5}{9}\right) = \frac{14}{9}$$

and

$$\sigma_X^2 = \sum_{x=1}^{2} x^2 \frac{x+3}{9} - \left(\frac{14}{9}\right)^2 = \frac{24}{9} - \frac{196}{81} = \frac{20}{81}.$$

The mean and the variance of Y are, respectively,

$$\mu_Y = \sum_{y=1}^{2} y \frac{3+4y}{18} = (1)\left(\frac{7}{18}\right) + (2)\left(\frac{11}{18}\right) = \frac{29}{18}$$

and

$$\sigma_Y^2 = \sum_{y=1}^{2} y^2 \frac{3+4y}{18} - \left(\frac{29}{18}\right)^2 = \frac{51}{18} - \frac{841}{324} = \frac{77}{324}.$$

$$\boxed{\mathrm{Cov}(X_1, X_2) = E(X_1X_2) - \mu_1\mu_2 - \mu_2\mu_1 + \mu_1\mu_2 = E(X_1X_2) - \mu_1\mu_2.}$$

$$\boxed{f(x,y) = \frac{x+2y}{18}}$$

The covariance of X and Y is

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \sum_{x=1}^{2}\sum_{y=1}^{2} xy\,\frac{x+2y}{18} - \left(\frac{14}{9}\right)\left(\frac{29}{18}\right) \\
&= (1)(1)\left(\frac{3}{18}\right) + (2)(1)\left(\frac{4}{18}\right) + (1)(2)\left(\frac{5}{18}\right) \\
&\quad + (2)(2)\left(\frac{6}{18}\right) - \left(\frac{14}{9}\right)\left(\frac{29}{18}\right) \\
&= \frac{45}{18} - \frac{406}{162} = -\frac{1}{162}.
\end{aligned}$$

$$\boxed{\rho = \frac{\mathrm{Cov}(X_1, X_2)}{\sigma_1\sigma_2} = \frac{\sigma_{12}}{\sigma_1\sigma_2}}$$

Hence, the correlation coefficient is

$$\rho = \frac{-1/162}{\sqrt{(20/81)(77/324)}} = \frac{-1}{\sqrt{1540}} = -0.025.$$

Independent ➔ Uncorrelated (zero correlation)

Suppose that X and Y are independent so that $f(x, y) \equiv f_1(x) f_2(y)$ and we want to find the expected value of the product $u(X)v(Y)$. Subject to the existence of the expectations, we know that

$$\begin{aligned} E[u(X)v(Y)] &= \sum_{S_1}\sum_{S_2} u(x)v(y)f(x, y) \\ &= \sum_{S_1}\sum_{S_2} u(x)v(y)f_1(x)f_2(y) \\ &= \sum_{S_1} u(x)f_1(x)\sum_{S_2} v(y)f_2(y) \\ &= E[u(X)]E[v(Y)]. \end{aligned}$$

This can be used to show that the correlation coefficient of two independent variables is zero. For, in a standard notation, we have

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

Uncorrelated → Independent ??? **NO!**

Example 4.2-3

Let X and Y have the joint pmf

$$f(x, y) = \frac{1}{3}, \qquad (x, y) = (0, 1), (1, 0), (2, 1).$$

Since the support is not "rectangular," X and Y must be dependent. The means of X and Y are $\mu_X = 1$ and $\mu_Y = 2/3$, respectively. Hence,

$$\begin{aligned}\mathrm{Cov}(X, Y) &= E(XY) - \mu_X\mu_Y \\ &= (0)(1)\left(\frac{1}{3}\right) + (1)(0)\left(\frac{1}{3}\right) + (2)(1)\left(\frac{1}{3}\right) - (1)\left(\frac{2}{3}\right) = 0.\end{aligned}$$

That is, $\rho = 0$, but X and Y are dependent.