統計學（一）

第六章 雙變量機率函數
(Bivariate Probability Distributions)

授課教師：唐麗英教授

國立交通大學 工業工程與管理學系
聯絡電話：(03)5731896
e-mail：litong@cc.nctu.edu.tw

2013

本課程內容參考書目

- **教科書**
  - Mendenhall, W., & Sincich, T. (2007). *Statistics for engineering and the sciences*, 5th Edition. Prentice Hall.
- **參考書目**
  - Berenson, M. L., Levine, D. M., & Krehbiel, T. C. (2009). *Basic business statistics: Concepts and applications*, 11th Edition. Upper Saddle River, N.J: Pearson Prentice Hall.
  - Larson, H. J. (1982). *Introduction to probability theory and statistical inference*, 3rd Edition. New York: Wiley.
  - Miller, I., Freund, J. E., & Johnson, R. A. (2000). *Miller and Freund's Probability and statistics for engineers*, 6th Edition. Upper Saddle River, NJ: Prentice Hall.
  - Montgomery, D. C., & Runger, G. C. (2011). *Applied statistics and probability for engineers*, 5th Edition. Hoboken, NJ: Wiley.
  - Watson, C. J. (1997). *Statistics for management and economics*, 5th Edition. Englewood Cliffs, N.J: Prentice Hall.
  - 林惠玲、陳正倉（2009），「統計學：方法與應用」，第四版，雙葉書廊有限公司。
  - 唐麗英、王春和（2013），「從範例學MINITAB統計分析與應用」，博碩文化公司。
  - 唐麗英、王春和（2008），「SPSS 統計分析 14.0 中文版」，儒林圖書公司。
  - 唐麗英、王春和（2007），「Excel 2007統計分析」，第二版，儒林圖書公司。
  - 唐麗英、王春和（2005），「STATISTICA6.0與基礎統計分析」，儒林圖書公司。
  - 陳順宇（2004），「統計學」，第四版，華泰書局。
  - 彭昭英、唐麗英（2010），「SAS123」，第七版，儒林圖書公司。

間斷型與連續型雙變量機率函數
(Bivariate Probability Distributions for Discrete & Continuous Random Variables)

Bivariate Probability Distributions

- **Def: Random Vector**
  - Suppose that $\boldsymbol{S}$ is the sample space associated with an experiment.
  - Let $\mathbf{X} = \mathbf{X}(\boldsymbol{\omega})$ and $\mathbf{Y} = \mathbf{Y}(\boldsymbol{\omega})$ be two functions each assigning a real number to every point $\boldsymbol{\omega}$ of $\mathbf{S}$. Then $(\mathbf{X}, \mathbf{Y})$ is called a two-dimensional random vector (or we say that $\mathbf{X}, \mathbf{Y}$ are **jointly distributed random variables**).

- **Remark**
  - If $\mathbf{X_1} = \mathbf{X_1}(\boldsymbol{\omega}), \mathbf{X_2} = \mathbf{X_2}(\boldsymbol{\omega}), \dots, \mathbf{X_n} = \mathbf{X_n}(\boldsymbol{\omega})$ are **n** functions each assigning a real number to every outcome, we call $(\mathbf{X_1}, \mathbf{X_2}, \dots, \mathbf{X_n})$ an **n-dimensional random vector** (or we say $\mathbf{X_1}, \mathbf{X_2}, \dots, \mathbf{X_n}$ are **jointly distributed random variables**).

Bivariate Probability Distributions

- **Def: Joint Probability Mass Function for Discrete Random Variables**
  - Suppose that X and Y are discrete random variables defined on the same probability space and that they take on values $x_1, x_2, \dots$, and $y_1, y_2, \dots$, respectively.
  - Their **joint probability mass function** $\mathbf{P(X, Y)}$ is

$$\boxed{\mathbf{P(x, y) = P(X = x, Y = y)}}$$

  - The joint probability mass function must satisfy the following conditions:

    1) $\mathbf{P(X = x, Y = y) \geq 0 . \forall (x, y) \in R}$

    2) $\mathbf{\sum_{all\ x} \sum_{all\ y} P(X = x, Y = y) = 1}$

Bivariate Probability Distributions

- **How to find the joint probability mass function for the discrete X and Y ?**
  - Construct a table listing each value that the R.V. $\mathbf{X}$ and $\mathbf{Y}$ can assume. Then find $\mathbf{p}(\mathbf{x},\mathbf{y})$ for each combination of $\mathbf{P}(\mathbf{X},\mathbf{Y})$.

- **Example 1 :**
  - Toss a fair coin 3 times. Let **X** be the number of **heads** on the **first** toss and **Y** the **total number of heads** observed for the **three** tosses.
  - What is the joint probability mass function of $(\mathbf{X},\mathbf{Y})$?

Bivariate Probability Distributions

		Y				P(X=x)
		0	1	2	3	
X	0	1/8	2/8	1/8	0	4/8
	1	0	1/8	2/8	1/8	4/8
P(Y=y)		1/8	3/8	3/8	1/8	1

$S = \{(HHH), (HHT), \dots, (TTT)\}$

Note:

Each entry represents a P(x, y). e.g.,

- P(0, 2) = P(X=0, Y=2) = 1/8
- P(1, 0) = P(X=1, Y=0) = 0

Sometimes we are interested in only the probability mass function for X or for Y.

i.e., $P_X(x) = P(X = x)$ or $P_Y(y) = P(Y = y)$

For instance, in example 1,

$$P_Y(0) = P(Y = 0) = P(X = 0, Y = 0) + P(X = 1, Y = 0) = \frac{1}{8} + 0 = \frac{1}{8}$$

$$P_Y(1) = P(Y = 1) = P(X = 0, Y = 1) + P(X = 1, Y = 1) = \frac{2}{8} + \frac{1}{8} = \frac{3}{8}$$

In general, the marginal functions can be found by

$$P_X(x) = P(X = x) = \sum_y P(X = x, Y = y)$$

$$P_Y(y) = P(Y = y) = \sum_x P(X = x, Y = y)$$

Bivariate Probability Distributions

- **How to find the Marginal Probability function from the table?**
  1) To find $P_Y(y)$, sum down the appropriate column of the table.
  2) To find $P_X(x)$, sum across the appropriate row of the table.

- **Note**
  – Since $P_Y(y)$ and $P_X(x)$ are located in the row and column **"margins"**, these distributions are called **marginal probability functions.**

Bivariate Probability Distributions

- **Example 2 :**
  - In example1, find the marginal probability functions for X and Y.

		Y				P(X=x)
		0	1	2	3	
X	0	1/8	2/8	1/8	0	4/8
	1	0	1/8	2/8	1/8	4/8
P(Y=y)		1/8	3/8	3/8	1/8	1

$$P_X(x) = \frac{4}{8}\text{, for } x = 0,1$$

$$P_Y(0) = P(Y = 0) = \frac{1}{8}$$

$$P_Y(1) = P(Y = 1) = \frac{3}{8}$$

$$P_Y(2) = P(Y = 2) = \frac{3}{8}$$

$$P_Y(3) = P(Y = 3) = \frac{1}{8}$$

Bivariate Probability Distributions

- **Def: Joint Probability Density Function for Continuous Random Variables**
  - Suppose that **X** and **Y** are jointly distributed continuous random variables. Their joint density function is a piecewise continuous function of two variables, f(x, y), such that for any " reasonable " two-dimensional set A

$$\boxed{P\big((\mathbf{X},\mathbf{Y}) \in \mathbf{A}\big) = \int_A \int f(x,y)\,dydx}$$

- **The joint probability density function must satisfy the following conditions**

1) $f(x,y) \geq \mathbf{0}\,.\,\forall(\mathbf{x},\mathbf{y}) \in \mathbf{R}$

2) $\int_{\text{all } y}\int_{\text{all } x} f(x,y)\,dxdy = \mathbf{1}$

Bivariate Probability Distributions

- **The marginal density function for X and Y are**

  1) $f_X(x) = \int_{\text{all } y} f(x,y)\,dy$

  2) $f_Y(y) = \int_{\text{all } x} f(x,y)\,dx$

- **Note**
  - In the discrete case, the marginal mass function was found by summing the joint probability mass function over the other variable; in the **continuous** case, it is found by **integration**.

Bivariate Probability Distributions

- **Example 3 :**
  - Assume (X, Y) is a 2-dimensional continuous random vector with the following joint density function

$$\boxed{\begin{aligned} f_{XY}(x, y) &= c \quad \text{if } 5 < x < 10, \qquad 4 < y < 9 \\ &= 0 \quad \text{otherwise} \end{aligned}}$$

a) Find c.

b) Find the marginal density functions for X and Y.

Bivariate Probability Distributions

- **Solution**

a) Because $\int_{\text{all } y}\int_{\text{all } x} f(x,y)\,dxdy = 1$, it follows that

$$=> \quad \int_4^9\int_5^{10}(c)\,dxdy = 1 = c\int_4^9\left[\int_5^{10} dx\right]dy = c\int_4^9 [x]_5^{10}\,dy$$

$$= c\int_4^9 5dy = 5c\int_4^9 dy = [5cy]_4^9 = 5c(9-4) = 25c = 1$$

$$\Rightarrow \quad c = \frac{1}{25}$$

b) $f_X(x) = \int_4^9 \frac{1}{25}dy = \frac{1}{25}[y]_4^9 = \frac{1}{5}$ for $5 < x < 10$

$f_Y(y) = \int_5^{10} \frac{1}{25}dx = \frac{1}{25}[x]_5^{10} = \frac{1}{5}$ for $4 < y < 9$

Bivariate Probability Distributions

- **Example 4 :**

$$\text{if } f(x,y) = 2 \quad \text{if } 0 < x < y, \qquad 0 < y < 1$$
$$= 0 \quad \text{otherwise}$$

a) Find $f_X(x)$ b) Find $f_Y(y)$ c) P(X<1, Y<1) d) P(X < Y)

- **Solution**

a) $f_X(x) = \int_{\text{all } y} f(x,y)dy = \int_x^1 2dy = 2[y]_x^1 = 2 - 2x, \;\; 0 < x < 1.$

b) $f_Y(y) = \int_{\text{all } x} f(x,y)dx = \int_0^y 2dx = 2[x]_0^y = 2y, \;\; 0 < y < 1.$

c) $P(\, X < 1, Y < 1 \,) = \int_0^1 \int_0^y 2dxdy$

d) $P(\, X < Y \,) = \int_0^1 \int_0^y 2dxdy = \int_0^1 [2x]_0^y dy = \int_0^1 [2y] \; dy = [\frac{2y^2}{2}]_0^1 = 1$

Bivariate Probability Distributions

- **Example 5 :**
  - Suppose that the joint p.d.f. of X and Y is specified as follows:

$$\boxed{\begin{aligned} f(x,y) &= 2e^{-x}e^{-y} \quad \text{for } 0 < x < \infty,\ \ 0 < y < \infty \\ &= 0 \quad \text{otherwise} \end{aligned}}$$

Find $P(X > 1, Y < 1)$.

- **Solution**

$$P(X > 1, Y < 1) = \int_0^1 \int_1^\infty 2e^{-x}e^{-y}dxdy = \int_0^1 2e^{-y}(-e^{-x}|_1^\infty)dy$$

$$= \int_0^1 2e^{-y}(0 + e^{-1})dy = e^{-1}\int_0^1 2e^{-y}dy = 2e^{-1}\int_0^1 e^{-y}dy$$

$$= 2e^{-1}(-e^{-y})|_0^1 = (1 - e^{-1})2e^{-1}$$

Bivariate Probability Distributions

- **Example 6 :** The density function of X and Y is given by

$$f_{XY}(x,y) = 2e^{-x}e^{-2y} \quad \text{for } 0 < x < \infty,\ \ 0 < y < \infty$$
$$= 0 \quad \text{otherwise}$$

a) Find $f_X(x)$ b) $P(X > 1, Y < 1)$ c) $P(X < a)$, $a > 0$

- **Solution**

a) Let $u = 2y$

$du = 2dy \quad \Rightarrow \quad y = 0 \rightarrow u = 0,\ y \approx \infty \rightarrow u \approx \infty$

$\frac{1}{2}du = dy$

$$f_X(x) = \int_{\text{all } y} f(x,y)dy = \int_0^\infty 2e^{-x}e^{-2y}dy = 2e^{-x}\int_0^\infty e^{-2y}dy$$

$$= 2e^{-x}\int_0^\infty e^{-u}\frac{1}{2}du = e^{-x}(-e^{-u}|_0^\infty)$$

$$= e^{-x}\left(0-(-1)\right) = e^{-x},\quad 0 < x < \infty$$

Bivariate Probability Distributions

- **Solution**

b) $P(X > 1, Y < 1) = \int_0^1 \int_1^\infty 2e^{-x}e^{-2y}dxdy = \int_0^1 2e^{-2y}[\int_1^\infty e^{-x}dx]dy$

$= \int_0^1 2e^{-2y}(-e^{-x}|_1^\infty)dy = \int_0^1 2e^{-2y}\big(0-(-e^{-1})\big)dy$

$= e^{-1}\int_0^1 2e^{-2y}dy = e^{-1}(1-e^{-2})$

c) $P(X < a) = \int_0^a f_X(x)dx = \int_0^a e^{-x}dx = -e^{-x}|_0^a = -e^{-a}-(-1) = 1-e^{-a}$

$(a > 0)$

雙變量機率函數之期望值與共變異數
(The Expected values and Covariance for Jointly Distribution Random Variables)

Expected values and Covariance

- **Recall: The Expected value of a Random variable**
  - $E(X) = \sum_X x \cdot P(x)$ if X is a discreate random variable.
  - $E(X) = \int_X x \cdot f(x)dx$ if X is a continuous R. V.

- **Remark: In general**
  - $E[g(x)] = \sum_X g(x) \cdot P(x)$ if X is a **discreate** random variable.

    $= \int_X g(x) \cdot f(x)dx$ if X is a **continuous** random variable.

Expected values and Covariance

- **Theorem: If X and Y are independent , then**

$$\boxed{\begin{array}{l} E[g(x)\cdot h(y)]= E[g(x)]\cdot E[h(y)] \\ E[XY]= E[X]\cdot E[Y] \end{array}}$$

<u>Proof:</u> (see the textbook for discrete case)

Let X,Y be continuous with joint density function $f_{XY}(x,y)$

$$E[g(x)\cdot h(y)]=\int_y \int_x g(x)h(y)\cdot \underline{f(x,y)}dxdy=\int_y \int_x g(x)h(y)\cdot \underline{f_X(x)\, f_Y(y)}dxdy$$

$$=\int_X g(x)\, f(x)dx \cdot \int_y h(y)\, f(y)dy$$

$$= E[g(x)]\cdot E[h(y)]$$

Since X,Y are independent,

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

(see Ch6.4)

Expected values and Covariance

- **Recall**

$$Var(X)=E[(X-\mu)^2]=E(X^2)-[E(X)]^2=E(X^2)-\mu^2$$

- **Def: The Covariance of any two variable X and Y is**

$$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=\underline{E(XY)-E(X)E(Y)\quad \text{“short-cut” formular}}$$

- **Example 1: Show that Cov(X,Y)=E(XY)－E(X)E(Y)**

$$Cov(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E\{XY-X\,E(Y)-Y\,E(X)+E(X)E(Y)\}$$

$$=E(XY)-E(X)E(Y)-E(Y)E(X)+E(X)E(Y)$$

$$=E(XY)-E(X)E(Y)$$

Expected values and Covariance

- **Theorem**
  - If X and Y are **independent** , then **Cov(X,Y)=0**

Proof: If X and Y are independent ,then we know E(XY)= E[X] • E[Y]
Cov(X,Y)=E(XY)－E(X)E(Y)= E(XY)－E(XY)=0

- **Question**
  - If Cov(X,Y)=0, X,Y are independent?? **NO!!**
- **Remark**
  - If Cov(X,Y)=0, X and Y may **not** be independent.

Expected values and Covariance

- **Example 2: (Counterexample)**
  - Let X,Y be discrete random variable with P(X=x, Y=y) as follows:

X \ Y	-1	0	1	$P_X(x)$
-1	1/8	1/8	1/8	3/8
0	1/8	0	1/8	2/8
1	1/8	1/8	1/8	3/8
$P_Y(y)$	3/8	2/8	3/8	1

a) Find Cov(X,Y)

b) Are X and Y independent?

Expected values and Covariance

- **Solution**

a) $$E(X)=\sum_x x\cdot P_X(x)=(-1)(3/8)+0(2/8)+1(3/8)=0$$

$$E(Y)=\sum_y y\cdot P_Y(y)=(-1)(3/8)+0(2/8)+1(3/8)=0$$

$$E(XY)=\sum_x\sum_y xy\,P(X=x,Y=y)$$

$$=(-1)(-1)(1/8)+0(-1)(1/8)+(1)(-1)(1/8)+(-1)(0)(1/8)+0\cdot 0\cdot 0+1\cdot 0(1/8)+(-1)(1)(1/8)$$
$$+0(1)(1/8)+1(1)(1/8)=1/8-1/8-1/8+1/8=0$$

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

b) But $P(X=0, Y=0)\overset{?}{=}P(X=0)\,P(Y=0)$?

$$0\overset{?}{=}2/8\cdot 2/8 \implies P(X=0, Y=0)\neq P(X=0)\,P(Y=0)$$

$\therefore$NO! X,Y are **<u>not</u>** independent.

雙變量機率函數之獨立性與相關性 (Independence and Conditional Distributions)

Independence and conditional Distributions

- **Def: Independent random variables**
  - X and Y are independent random variables if and only if

$$P(X\in A,\ Y\in B)=P(X\in A)\cdot P(Y\in B)$$

- **Theorem**
  - X and Y are **independent** random variables if and only if

$$F_{XY}(X,Y)=F_X(X)\cdot F_Y(Y)$$

- **Corollary**

1) If X and Y are **independent <u>discrete</u>** random variables, then
$P(X=x,\ Y=y)=P(X=x)\cdot P(Y=y)=P_X(x)\cdot P_Y(y)$

2) If X and Y are **independent <u>continuous</u>** random variables, then
$f(x,y)=f_X(x)\cdot f_Y(y)$

Independence and conditional Distributions

- **Example 1 :**

Given $P(x,y)=\frac{1}{n^2}$, x=1,2,…,n , y=1,2,…,n

=0 otherwise.

Are X and Y are independent?

- **Solution**

Check: $P(x,y)\overset{?}{=}P_X(x)\cdot P_Y(y)$

The marginal function for X is

$$P_X(x)=\sum_{y=1}^{n}P(x,y)=\sum_{y=1}^{n}\frac{1}{n^2}=\frac{1}{n^2}\sum_{y=1}^{n}1=\frac{n}{n^2}=\frac{1}{n}\ ,X=1,2,\ldots,n$$

The marginal function for X is

$$P_Y(y)=\sum_{x=1}^{n}P(x,y)=\sum_{x=1}^{n}\frac{1}{n^2}=\frac{1}{n}\ ,Y=1,2,\ldots,n$$

Check: $P(x,y)\overset{?}{=}P_X(x)\cdot P_Y(y)$ for all (x,y)? $\frac{1}{n^2}\overset{?}{=}\frac{1}{n}\cdot\frac{1}{n}$

Yes! X and Y are independent.

Independence and conditional Distributions

- **Example 2 :** Let X and Y be continuous random variables with

$$f(x,y) = \begin{cases} e^{-x-y} & x > 0, y > 0, \\ =0 & \text{otherwise.} \end{cases}$$

Are X and Y are independent?

- **Solution:**

$$f_X(x) = \int_0^\infty e^{-x-y}dy = \int_0^\infty e^{-x} \cdot e^{-y}dy = e^{-x}\int_0^\infty e^{-y}dy$$

$$= e^{-x}\left(-e^{-y}\Big|_0^\infty\right) = e^{-x}\big(0-(-1)\big) = e^{-x},\ 0 < x < \infty$$

$$f_Y(y) = \int_0^\infty e^{-x-y}dx = e^{-y}, 0 < y < \infty$$

Check: $f(x,y) \overset{?}{=} f_X(x) \cdot f_Y(y)$

$e^{-x-y} = e^{-x} \cdot e^{-y}$ Yes! X and Y are independent.

Independence and conditional Distributions

- **Example 3 :** Let X and Y be continuous random variables with

$$\boxed{\begin{array}{ll} f(x,y)=2 & 0<x<y,\ 0<y<1 \\ \quad =0 & \text{otherwise.} \end{array}}$$

  Are X and Y are independent?

- **Solution:**

$$f_X(x) = \int_x^1 2\,dy = 2 - 2x, \quad \mathbf{0 < x < 1}$$

$$f_Y(y) = \int_0^y 2\,dx = 2y$$

Check: $f(x,y) \overset{?}{=} f_X(x) \cdot f_Y(y)$

$2 \neq (2-2x) \cdot 2y$ No! X and Y are **not** independent.

Independence and conditional Distributions

Recall: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

- **Conditional Distributions**

1) If X and Y are **discrete** random variables, then

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{P_{XY}(x, y)}{P_Y(y)}, \qquad P_Y(y) \neq 0$$

Note:

i) $P_{XY}(x, y) = P(x|y)P_Y(y)$ ii) $P_X(x) = \sum_{\text{all } y} P(x|y)P_Y(y)$

2) If X and Y are **continuous** random variables, then

$$f(x|y) = \frac{f_{XY}(x, y)}{f_Y(y)}, \qquad f_Y(y) \neq 0$$

Note:

i) $f_{XY}(x, y) = f(x|y)f_Y(y)$ ii) $f_X(x) = \int_{-\infty}^{\infty} f(x|y)f_Y(y)\,dy$

Independence and conditional Distributions

- **Example 4 :**
  - Let X,Y be discrete random variable with P(X=x, Y=y) as follows, find the conditional probability function of X given Y=1.

X \ Y	0	1	2	3	$P_X(x)$
0	1/8	2/8	1/8	0	
1	0	1/8	2/8	1/8	
		3/8			

- **Solution**

$$P(X=0|Y=1)=\frac{P(X=0,Y=1)}{P(Y=1)}=\frac{(\frac{2}{8})}{(\frac{3}{8})}=\frac{2}{3}$$

$$P(X=1|Y=1)=\frac{P(X=1,Y=1)}{P(Y=1)}=\frac{(\frac{1}{8})}{(\frac{3}{8})}=\frac{1}{3}$$

Independence and conditional Distributions

- **Example 5 :** The density function of X and Y is given by

$$f_{XY}(x,y) = \lambda^2 e^{-\lambda y} \qquad 0 \leqq x \leqq y \qquad y \geqq 0$$
$$= 0 \qquad \text{otherwise}$$

We find $f_X(x) = \lambda e^{-\lambda x}$, $x \geqq 0$ and $f_Y(y) = \lambda^2 y e^{-\lambda y}$, $y \geqq 0$

Find the conditional density $f(y|x)$.

- **Solution**

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{\lambda^2 e^{-\lambda y}}{\lambda e^{-\lambda x}} = \lambda e^{-\lambda(y-x)} \qquad \textbf{for } \underline{x \leqq y} < \infty$$

$$f(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{\lambda^2 e^{-\lambda y}}{\lambda^2 y e^{-\lambda y}} = \frac{1}{y} \quad , \qquad 0 \leqq x \leqq y$$

雙變量機率函數之共變異數與相係數
(Covariance and Correlation)

Covariance and Correlation

Two measures of association between two random variables:

1) Covariance

Recall: $\mathrm{Var}(X) = E[(X-\mu)^2] = E(X^2) - \mu^2$

Def: The **covariance** of any two random variables **X** and **Y** is

$$\mathbf{Cov(X,Y) = E\{[X - E(X)][Y - E(Y)]\} = E[(X - \mu_X)(Y - \mu_Y)]}$$

Note: The value of **Cov(X, Y)** can be **positive**, **negative** or **zero**.

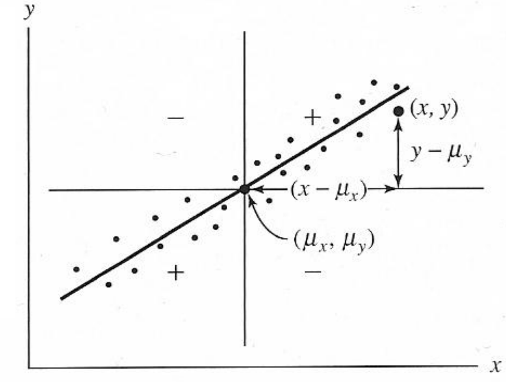

FIGURE 6.5
Signs of the cross-products $(x-\mu_x)(y-\mu_y)$

Theorem 1 : $\mathbf{Cov(X,Y) = E(XY) - \mu_X\mu_Y}$

Proof: $\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY - \mu_Y X - \mu_X Y + \mu_X\mu_Y]$

$$= E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y$$

$$= E[XY] - \mu_Y\mu_X - \cancel{\mu_X\mu_Y} + \cancel{\mu_X\mu_Y}$$

$$= E[XY] - \mu_X\mu_Y$$

Covariance and Correlation

2) Correlation

Def: The **correlation** of any two random variables **X** and **Y** is

$$\boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) = \frac{\mathbf{Cov}(\mathbf{X},\mathbf{Y})}{\boldsymbol{\sigma}_{\mathbf{X}}\boldsymbol{\sigma}_{\mathbf{Y}}}, \qquad \text{provide that } \sigma_X < \infty \textit{ and } \sigma_Y < \infty$$

Note:

1) $-1 \leq \boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) \leq \mathbf{1}$
2) **X** and **Y** are said to be **positively** correlated if $\boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) > 0$.
3) **X** and **Y** are said to be **negatively** correlated if $\boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) < 0$.
4) **X** and **Y** are said to be **uncorrelated** if $\boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) = \mathbf{0}$.

Covariance and Correlation

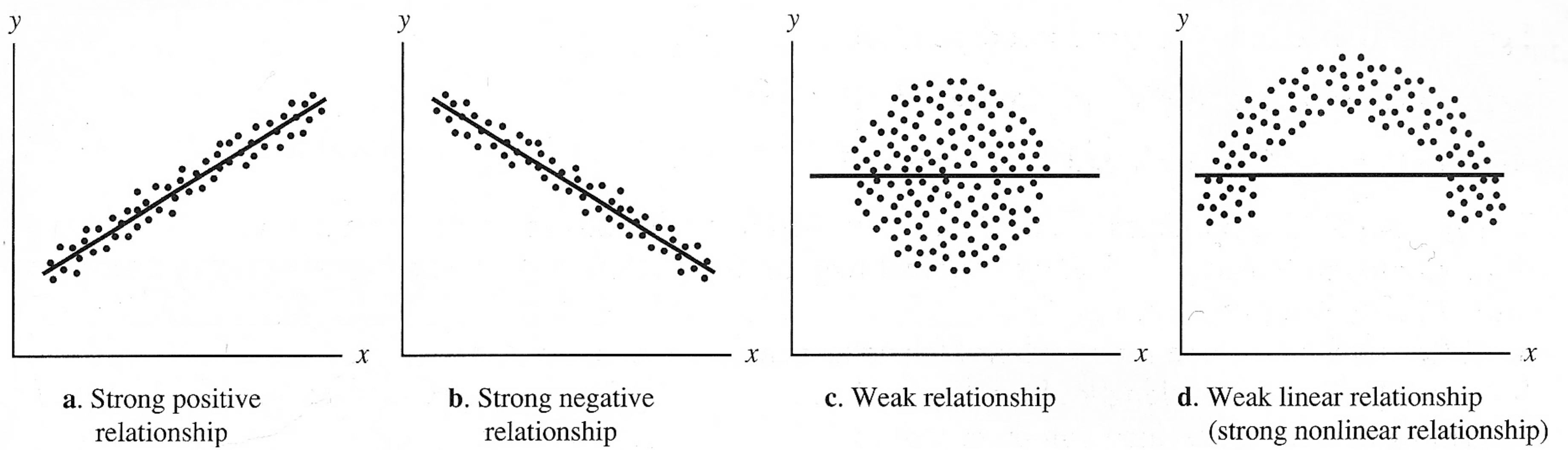

FIGURE 6.4
Linear relationships between X and Y

Covariance and Correlation

- **Theorem 2**
  - If X and Y are independent, then $\mathbf{Cov}(\mathbf{X},\mathbf{Y}) = \boldsymbol{\rho}(\mathbf{X},\mathbf{Y}) = \mathbf{0}$.
- **Proof**

$$\mathrm{Cov}(X,Y) = E(XY) - \mu_X\mu_Y = E(X)\cdot E(Y) - \mu_X\mu_Y$$

$$= \mu_X\mu_Y - \mu_X\mu_Y = 0$$

Since X and Y are <u>independent</u>.

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y} = \frac{0}{\sigma_X\sigma_Y} = 0$$

<u>Remark:</u> If $\mathrm{Cov}(X,Y) = 0$, X and Y may **<u>not</u>** be independent (i.e., X and Y may be dependent). **(The reverse of Thm2 is <u>not</u> true.)**

Covariance and Correlation

- **Example 1: Dependent but Uncorrelated Random Variables**
  - Suppose that the random variable X can take only the three values -1, 0 and 1, and that each of these three values has the same probability. Also, let $Y = X^2$.

a) Find COV(X, Y)

b) Are X and Y independent?

Covariance and Correlation

- **Solution**

X	-1	0	1
f(X)	1/3	1/3	1/3

a) $\mathrm{Cov}(X,Y) = E[XY] - \mu_x u_Y = 0$

$$E[XY] = E[X \cdot X^2] = E[X^3] = \sum_{x=-1}^{1} X^3 \cdot f(x) = (-1)^3 \cdot \frac{1}{3} + 0 + 1^3 \cdot \frac{1}{3} = 0$$

$$\mu_X = E[X] = \sum_{x=-1}^{1} X \cdot f(x) = (-1) \cdot \frac{1}{3} + 0 + 1 \cdot \frac{1}{3} = 0$$

b) $\therefore$ X and Y are uncorrelated.

Since $Y = X^2$ the value of Y is completely determined by the value of X, X and Y are clearly dependent.

Covariance and Correlation

- **Theorem 3:**
  - Suppose that X is a random variable such that $0 \leq \sigma_X^2 \leq \infty$, and that $\mathbf{Y = aX + b}$ for some constants a and b, where $a \neq 0$. If $a > 0$, then $\rho(X, Y) = +1$. If $a < 0$, then $\rho(X, Y) = -1$.

  (That is, if **Y** is a linear function of **X**, then **X** and **Y** must be **<u>correlated</u>** and $|\boldsymbol{\rho}(\mathbf{X}, \mathbf{Y})| = \mathbf{1}$)

<u>Proof:</u>

If $Y = aX + b$, then $\mu_Y = E(Y) = E(aX + b) = aE(X) + b = a\mu_X + b$

$$\therefore \ Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E\{(X - \mu_X) \cdot [(aX + b) - (a\mu_X + b)]\}$$
$$= E[(X - \mu_X) \cdot a(X - \mu_X)] = E[a(X - \mu_X)^2] = aE[(X - \mu_X)^2]$$
$$= a\sigma_X^2$$

$$\rho(X, Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{a\sigma_X^2}{2\sigma_X \cdot |a| \sigma_X} = \frac{a}{|a|} = \begin{cases} 1 & \text{if } a > 0 \\ -1 & \text{if } a < 0 \end{cases}$$

$$\sigma_Y{}^2 = Var(Y) = Var(aX + b) = a^2 Var(X) = a^2 \sigma_x{}^2, \quad \therefore \sigma_Y = |a| \sigma_X$$

Covariance and Correlation

- **Theorem 4: If X and Y are random variables, then**
  - $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$

    (provide that $\mathrm{Var}(X) < \infty$ *and* $Var(Y) < \infty$)

Proof:

$$\begin{aligned}\mathrm{Var}(X+Y) &= E[(X+Y)-(\mu_x+\mu_Y)]^2 = E[(X-\mu_x)+(Y-\mu_Y)]^2 \\ &= E[(X-\mu_X)^2 + (Y-\mu_Y)^2 + 2(X-\mu_X)(Y-\mu_Y)] \\ &= E[(X-\mu_X)]^2 + E[(Y-\mu_Y)]^2 + 2E[(X-\mu_X)(Y-\mu_Y)] \\ &= \sigma_X^2 + \sigma_Y^2 + 2\mathrm{Cov}(X,Y)\end{aligned}$$

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

Covariance and Correlation

- **Remark**
  1) $\mathrm{Cov}(aX, bY) = ab \cdot \mathrm{Cov}(X, Y)$.
  2) $\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab \cdot \mathrm{Cov}(X, Y)$.
  3) $\mathrm{Var}(X - Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X, Y)$

- **Theorem 5**
  - If $X_1, X_2, \cdots, X_n$ are random variables and $\mathrm{Var}(X_i) < \infty$, for $i = 1, \cdots, n$, then

$$\mathrm{Var}\left(\sum X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X) + 2\sum\sum_{i<j} \mathrm{Cov}\left(X_i, Y_j\right)$$

本單元結束