Chapter 7

Bivariate random variables

Wei-Yang Lin
Department of Computer Science & Information Engineering
mailto:wylin@cs.ccu.edu.tw
• 7.1 Joint and marginal probabilities
• 7.2 Jointly continuous random variables
• 7.3 Conditional probability and expectation
• 7.4 The bivariate normal
• 7.5 Extension to three or more random variables
• The main focus of this chapter is the study of pairs of continuous random variables that are not independent.

• Consider the following functions of two random variables $X$ and $Y$, $X + Y$, $XY$, $\max(X, Y)$, $\min(X, Y)$.

• Show that the cdfs of these four functions of $X$ and $Y$ can be expressed in the form $P((X, Y) \in A)$ for various sets $A \subset \mathbb{R}^2$. 
Example 7.1

- A random signal $X$ is transmitted over a channel subject to additive noise $Y$.
- The received signal is $Z = X + Y$.
- Express the cdf of $Z$ in the form $P((X, Y) \in A_z)$ for some set $A_z$. 
Solution

• Write

\[ F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = P((X, Y) \in A_z), \]

where

\[ A_z = \{(x, y) : (x + y) \leq z\} \]

• Since \( x + y \leq z \) if and only if \( y \leq -x + z \), it is easy to see that \( A_z \) is the shaded region in Figure 1.
Figure 1: The shaded region is $A_z = \{(x, y) : x + y \leq z\}$. 
Example 7.3

• Express the cdf of $U := \max(X, Y)$ in the form $P((X, Y) \in A_u)$ for some set $A_u$. 
Solution

- To find the cdf of \( U \), begin with

\[
F_U(u) = P(U \leq u) = P(\max(X, Y) \leq u)
\]

- Since the large of \( X \) and \( Y \) is less than or equal to \( u \) if and only if \( X \leq u \) and \( Y \leq u \),

\[
P(\max(X, Y) \leq u) = P(X \leq u, Y \leq u) = P((X, Y) \in A_u),
\]

where \( A_u = \{(x, y) : x \leq u \text{ and } y \leq u\} \) is the shaded “southwest” region shown in Figure 2.
Figure 2: The shaded region is \( \{(x, y) : x \leq u \text{ and } y \leq u\} \).
Example 7.4

- Express the cdf of $V := \min(X, Y)$ in the form $P((X, Y) \in A_v)$ for some set $A_v$. 
Solution

- To find the cdf of $V$, begin with

$$F_V(v) = P(V \leq v) = P(\min(X, Y) \leq v).$$

- Since the smaller of $X$ and $Y$ is less than or equal to $v$ if and only if either $X \leq v$ or $Y \leq v$,

$$P(\min(X, Y) \leq v) = P(X \leq v \text{ or } Y \leq v) = P((X, Y) \in A_v),$$

where $A_v = \{(x, y) : x \leq v \text{ or } y \leq v\}$ is the shaded region shown in Figure 3.
Figure 3: The shaded region is \( \{(x, y) : x \leq v \text{ or } y \leq v\} \).
Product sets and marginal probabilities

• The **Cartesian product** of two univariate sets $B$ and $C$ is defined by

$$B \times C := \{(x, y) : x \in B \text{ and } y \in C\}.$$ 

• In other words,

$$(x, y) \in B \times C \iff x \in B \text{ and } y \in C.$$ 

• For example, if $B = [1, 3]$ and $C = [0.5, 3.5]$, then $B \times C$ is the rectangle in Figure 4.
Figure 4: The Cartesian product $[1, 3] \times [0.5, 3.5]$. 
Joint cumulative distribution functions

- The joint cumulative distribution function of $X$ and $Y$ is defined by
  \[
  F_{XY}(x, y) = P(X \leq x, Y \leq y). \tag{1}
  \]
- We can also write this using a Cartesian product set as
  \[
  F_{XY}(x, y) = P((X, Y) \in (-\infty, x] \times (-\infty, y]).
  \]
- In other words, $F_{XY}(x, y)$ is the probability that $(X, Y)$ lies in the southwest region shown in Figure 5.
Figure 5: The Cartesian product \((-\infty, x] \times (-\infty, y]\).
Rectangle formula

- The joint cdf is important because it can be used to compute $P((X, Y) \in A)$.

- For example, $P(a < X \leq b, c < Y \leq d)$, which is the probability that $(X, Y)$ belongs to the rectangle $(a, b] \times (c, d]$ as shown in Figure 6, is given by the rectangle formula

\[ F_{XY}(b, d) - F_{XY}(a, d) - F_{XY}(b, c) + F_{XY}(a, c). \]
Figure 6: The rectangle \((a, b] \times (c, d]\).
Example 7.5

- If $X$ and $Y$ have joint cdf $F_{XY}$, find the joint cdf of $U := \max(X, Y)$ and $V := \min(X, Y)$. 
Solution (1/2)

- Begin with
  \[ F_{UV}(u, v) = P(U \leq u, V \leq v). \]

- From Example 7.3, we know that \( U := \max(X, Y) \leq u \) if and only if \((X, Y)\) lies in the southwest region shown in Figure 2.

- From Example 7.4, we know that \( V := \min(X, Y) \leq v \) if and only if \((X, Y)\) lies in the region shown in Figure 3.

- Hence, \( U \leq u \) and \( V \leq v \) if and only if \((X, Y)\) lies in the intersection of these two regions.

- The form of this intersection depends on whether \( u > v \) or \( u \leq v \).
Solution (2/2)

- If $u \leq v$, then the southwest region region in Figure 2 is a subset of the region in Figure 3.

- Their intersection is the smaller set, and so

\[
P(U \leq u, V \leq v) = P(U \leq u) = F_U(u) = F_{XY}(u, u), \quad u \leq v.
\]

- If $u > v$, the intersection is shown in Figure 7.

\[
P(U \leq u, V \leq v)
= F_{XY}(u, u) - P(v < X \leq u, v < Y \leq u)
= F_{XY}(u, u) - (F_{XY}(u, u) - F_{XY}(v, u) - F_{XY}(u, v) + F_{XY}(v, v))
= F_{XY}(v, u) + F_{XY}(u, v) - F_{XY}(v, v), \quad u > v.
\]
Figure 7: The intersection of the shaded regions in Figures 2 and 3.
Marginal cumulative distribution functions

- It is possible to obtain the **marginal cumulative distributions** \( F_X \) and \( F_Y \) directly from \( F_{XY} \).

- More precisely, it can be shown that

\[
F_X(x) = \lim_{y \to \infty} F_{XY}(x, y) =: F_{XY}(x, \infty),
\]

and

\[
F_Y(y) = \lim_{x \to \infty} F_{XY}(x, y) =: F_{XY}(\infty, y).
\]
Example 7.7

• If

\[ F_{XY}(x, y) = \begin{cases} \frac{y + e^{-x(y+1)}}{y+1} - e^{-x}, & x, y > 0, \\ 0, & \text{otherwise}. \end{cases} \]

• Find both of the marginal cumulative distribution functions, \( F_X(x) \) and \( F_Y(y) \).
Solution

- The marginal cdf of $X$ is

$$F_X(x) = \begin{cases} 
1 - e^{-x}, & x > 0, \\
0, & x \leq 0.
\end{cases}$$

- The marginal cdf of $Y$ is

$$F_Y(y) = \begin{cases} 
\frac{y}{y+1}, & y > 0, \\
0, & y \leq 0.
\end{cases}$$
Independence

- We record here that jointly continuous random variable $X$ and $Y$ are **independent** if and only if their joint cdf factors into the product of their marginal cdfs.

\[
F_{XY}(x, y) = F_X(x)F_Y(y)
\]
Homework

- Problems 1, 2, 6.
• 7.1 Joint and marginal probabilities

• **7.2 Jointly continuous random variables**

• 7.3 Conditional probability and expectation

• 7.4 The bivariate normal

• 7.5 Extension to three or more random variables
• In analogy with the univariate case, we say that two random variables $X$ and $Y$ are **jointly continuous** with **joint density** $f_{XY}(x, y)$ if

$$P((X, Y) \in A) = \int_A \int f_{XY}(x, y) \, dx \, dy$$

for some nonnegative function $f_{XY}$ that integrates to one; i.e.,

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx \, dy = 1.$$
Example 7.10

- Suppose that a random, continuous-valued signal $X$ is transmitted over a channel subject to additive, continuous-valued noise $Y$.

- The received signal is $Z = X + Y$.

- Find the cdf and density of $Z$ if $X$ and $Y$ are jointly continuous random variables with joint density $f_{XY}$. 
Solution (1/2)

- Write

\[ F_Z(z) = P(Z \leq z) = P(X + Y \leq z) = P((X, Y) \in A_z), \]

where \( A_z := \{(x, y) : x + y \leq z\} \) was sketched in Figure 1.

- With the figure in mind, the double integral \( P(X + Y \leq z) \) can be computed using

\[ F_Z(z) = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{z-x} f_{XY}(x, y) dy \right] dx. \]
Solution (2/2)

- Now carefully differentiate with respect to $z$.

\[ f_z(z) = \frac{\partial}{\partial z} \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{z-x} f_{XY}(x, y)dy \right] dx \]
\[ = \int_{-\infty}^{\infty} \frac{\partial}{\partial z} \left[ \int_{-\infty}^{z-x} f_{XY}(x, y)dy \right] dx \]
\[ = \int_{-\infty}^{\infty} f_{XY}(x, z-x)dx. \]

- Recall that

\[ \frac{\partial}{\partial z} \int_{-\infty}^{g(z)} h(y)dy = h(g(z))g'(z). \]
• The **marginal densities** \( f_X(x) \) and \( f_Y(y) \) can be obtained from the joint density \( f_{XY} \).

\[
f_X(x) = \int_{-\infty}^{-\infty} f_{XY}(x, y) \, dy. \tag{4}
\]

\[
f_Y(y) = \int_{-\infty}^{-\infty} f_{XY}(x, y) \, dx. \tag{5}
\]

• Thus, to obtain the marginal densities, integrate out the unwanted variable.
Independence

- We record here that jointly continuous random variable $X$ and $Y$ are independent if and only if their joint density factors into the product of their marginal densities.

\[ f_{XY}(x, y) = f_X(x) f_Y(y) \]
Expectation

- If $X$ and $Y$ are jointly continuous with joint density $f_{XY}$, then the expectation of $g(X, Y)$ is given by

$$
E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{XY}(x, y) \, dx \, dy.
$$

(6)
Homework

- Problems 9.
• 7.1 Joint and marginal probabilities
• 7.2 Jointly continuous random variables
• **7.3 Conditional probability and expectation**
• 7.4 The bivariate normal
• 7.5 Extension to three or more random variables
• We define the **conditional density** of $Y$ given $X$ by

$$f_{Y|X}(y|x) := \frac{f_{XY}(x,y)}{f_X(x)}, \quad \text{for } x \text{ with } f_X(x) > 0. \quad (7)$$

• The **conditional cdf** is

$$F_{Y|X}(y|x) := P(Y \leq y|X = x) = \int_{-\infty}^{y} f_{Y|X}(t|x)dt. \quad (8)$$

• Note also that if $X$ and $Y$ are independent, the joint density factors, and so $f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)$.

• It then follows that $F_{Y|X}(y|x) = F_Y(y)$.

• In other words, we can “drop the conditioning”.

---

38
• Our definition of conditional probability satisfies the following law of total probability.

\[ P((X, Y) \in A) = \int_{-\infty}^{\infty} P((X, Y) \in A \mid X = x) f_X(x) \, dx. \]  

(9)

• We also have the substitution law,

\[ P((X, Y) \in A \mid X = x) = P((x, Y) \in A \mid X = x) \]  

(10)
Example 7.14

- Suppose that a random, continuous-valued signal $X$ is transmitted over a channel subject to additive, continuous-valued noise $Y$.
- The received signal is $Z = X + Y$.
- Find the cdf and density of $Z$ if $X$ and $Y$ are jointly have joint density $f_{XY}$. 
Solution (1/2)

- We use the laws of total probability and substitution.

\[
F_Z(z) = P(Z \leq z) = \int_{-\infty}^{\infty} P(Z \leq z | Y = y) f_Y(y) dy \\
= \int_{-\infty}^{\infty} P(X + Y \leq z | Y = y) f_Y(y) dy \\
= \int_{-\infty}^{\infty} P(X + y \leq z | Y = y) f_Y(y) dy \\
= \int_{-\infty}^{\infty} P(X \leq z - y | Y = y) f_Y(y) dy \\
= \int_{-\infty}^{\infty} F_{X|Y}(z - y | y) f_Y(y) dy.
\]
Solution (2/2)

- By differentiating with respect to $z$,

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X|Y}(z - y | y) f_Y(y) \, dy = \int_{-\infty}^{\infty} f_{XY}(z - y, y) \, dy.$$  

- If $X$ and $Y$ are independent, we can drop the conditioning and obtain

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) \, dy.$$  

42
Conditional expectation

- Law of total probability

\[ E[g(X, Y)] = \int_{-\infty}^{\infty} E[g(X, Y) | X = x] f_X(x) dx \]  \hspace{1cm} (11)

- Substitution law

\[ E[g(X, Y) | X = x] = E[g(x, Y) | X = x] \]  \hspace{1cm} (12)
Example 7.18

- Let $X \sim \text{exp}(1)$, and suppose that given $X = x$, $Y$ is conditionally normal with $f_{Y|X}(y|x) \sim N(0, x^2)$.

- Evaluate $E[Y^2]$ and $E[Y^2 X^3]$. 
Solution (1/2)

- We use the law of total probability.

\[
E[Y^2] = \int_{-\infty}^{\infty} E[Y^2|X = x] f_X(x) dx
\]

\[
= \int_{-\infty}^{\infty} x^2 f_X(x) dx
\]

\[
= E[X^2] = 2.
\]
Solution (2/2)

- We use the laws of total probability and substitution.

\[ \mathbb{E}[Y^2X^3] = \int_{-\infty}^{\infty} \mathbb{E}[Y^2X^3|X = x] f_X(x) dx \]

\[ = \int_{-\infty}^{\infty} \mathbb{E}[Y^2x^3|X = x] f_X(x) dx \]

\[ = \int_{-\infty}^{\infty} x^3 \mathbb{E}[Y^2|X = x] f_X(x) dx \]

\[ = \int_{-\infty}^{\infty} x^5 f_X(x) dx \]

\[ = \mathbb{E}[X^5] = 5! \]
Homework

- Problems 26, 30, 31, 32, 34, 36, 37, 39(c).
• 7.1 Joint and marginal probabilities
• 7.2 Jointly continuous random variables
• 7.3 Conditional probability and expectation
• 7.4 The bivariate normal
• 7.5 Extension to three or more random variables
• The **bivariate Gaussian** or **bivariate normal** density is a generalization of the univariate \( N(m, \sigma^2) \) density.

• Recall that the standard \( N(0, 1) \) density is given by

\[
\psi(x) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).
\]

• The general \( N(m, \sigma^2) \) density can be written in terms of \( \psi \) as

\[
\frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2} \left(\frac{x - m}{\sigma}\right)^2\right] = \frac{1}{\sigma} \cdot \psi \left(\frac{x - m}{\sigma}\right)
\]
• In order to define the general bivariate Gaussian density, it is convenient to define a standard bivariate density first.

• So, for $|\rho| < 1$, put

$$\psi_\rho(u, v) := \frac{\exp \left( \frac{-1}{2(1-\rho^2)} [u^2 - 2\rho uv + v^2] \right)}{2\pi \sqrt{1 - \rho^2}}.$$  \hspace{1cm} (13)

• For fixed $\rho$, this function of two variables $u$ and $v$ defines a surface.

• The surface corresponding to $\rho = 0$ is shown in Figure 8.
• From the figure and from the formula (13), we see that $\psi_0$ is circularly symmetric.

• For $u^2 + v^2 = r^2$, $\psi_0(u, v) = \frac{e^{-r^2/2}}{2\pi}$ does not depend on the particular values of $u$ and $v$, but only on the radius of the circle on which they lie.

• Some of these circles are shown in Figure 9.
Figure 8: The Gaussian surface $\psi_\rho(u, v)$ with $\rho = 0$. 
Figure 9: The level curves of $\psi_\rho(u, v)$ with $\rho = 0$. 
• We also point out that for $\rho = 0$, the formula (13) factors into the product of two univariate $N(0, 1)$ densities, i.e.,

$$\psi_0(u, v) = \psi(u)\psi(v).$$

• For $\rho \neq 0$, $\psi_\rho$ does not factor.

• In other words, $U$ and $V$ are independent if and only if $\rho = 0$.

• A plot of $\psi_\rho$ for $\rho = -0.85$ is shown in Figure 10.

• It turns out that now $\psi_\rho$ is constant on ellipse instead of circles.

• The axes of the ellipses are not parallel to the coordinate axes, as shown in Figure 11.
Figure 10: The Gaussian surface $\psi_\rho(u, v)$ with $\rho = -0.85$.
Figure 11: The level curves of $\psi_\rho(u, v)$ with $\rho = -0.85$. 
We can now define the general bivariate Gaussian density by

\[
\exp \left( \frac{-1}{2(1-\rho^2)} \left[ \left( \frac{x-m_X}{\sigma_X} \right)^2 - 2\rho \left( \frac{x-m_X}{\sigma_X} \right) \left( \frac{y-m_Y}{\sigma_Y} \right) + \left( \frac{y-m_Y}{\sigma_Y} \right)^2 \right] \right)
\]

\[
\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}.
\]

It can be shown that the marginals are \( f_X \sim N(m_X, \sigma_X^2) \), \( f_Y \sim N(m_Y, \sigma_Y^2) \) and that \( \mathbb{E} \left[ \left( \frac{X - m_X}{\sigma_X} \right) \left( \frac{Y - m_Y}{\sigma_Y} \right) \right] = \rho \).

Hence, \( \rho \) is the correlation coefficient between \( X \) and \( Y \).

A plot of \( f_{XY} \) with \( m_X = m_Y = 0, \sigma_X = 1.5, \sigma_Y = 0.6, \) and \( \rho = 0 \) is shown in Figure 12.

The corresponding elliptical level curves are shown in Figure 13.
Figure 12: The bivariate Gaussian density $f_{XY}(x, y)$ with $m_X = m_Y = 0$, $\sigma_X = 1.5$, $\sigma_Y = 0.6$, and $\rho = 0$. 
Figure 13: The level curves of the bivariate Gaussian density in Figure 12.
Homework

• Problems 47, 48, 49.
7.1 Joint and marginal probabilities
7.2 Jointly continuous random variables
7.3 Conditional probability and expectation
7.4 The bivariate normal
7.5 Extension to three or more random variables
For expectations, we have

\[
E[g(X, Y, Z)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y, z) f_{XYZ}(x, y, z) \, dx \, dy \, dz
\]

(15)

(by the law of total probability)

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[g(X, Y, Z)|Y = y, Z = z] f_{YZ}(y, z) \, dy \, dz
\]

(16)

(by the substitution law)
Homework

• Problems 57, 58.