5-2 Two Continuous Random Variables

5-2.1 Joint Probability Distribution

Definition

A **joint probability density function** for the continuous random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies the following properties:

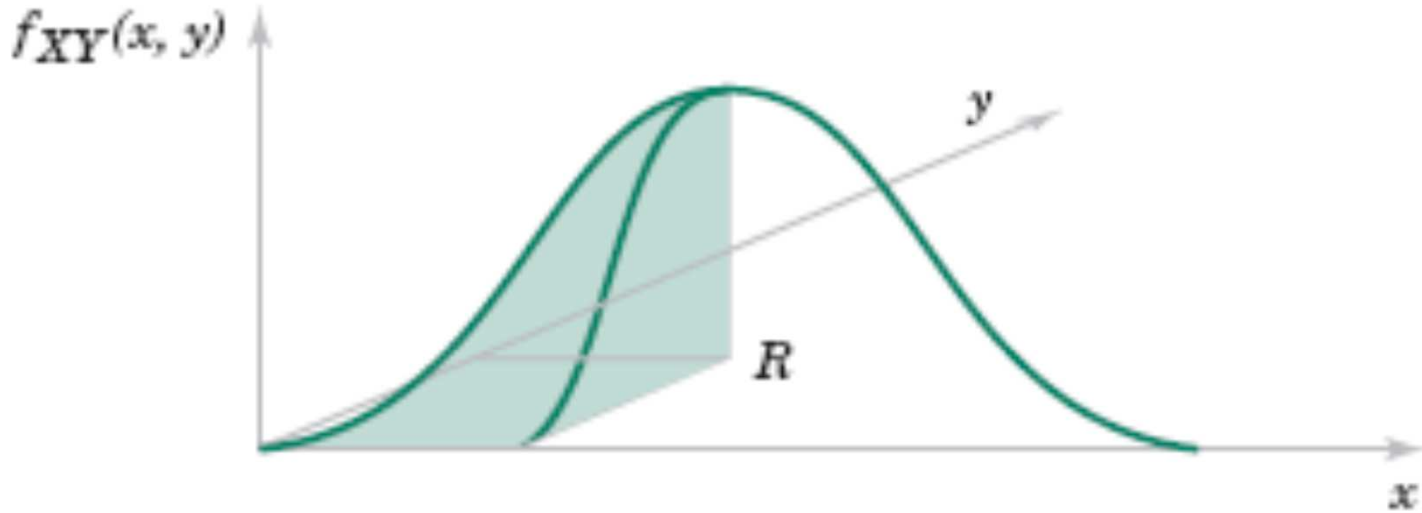
(1) $f_{XY}(x, y) \geq 0$ for all x, y

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

(3) For any region R of two-dimensional space

$$P((X, Y) \in R) = \iint_R f_{XY}(x, y)\, dx\, dy \qquad (5\text{-}14)$$

5-2 Two Continuous Random Variables



Probability that (X, Y) is in the region R is determined by the volume of $f_{XY}(x, y)$ over the region R.

Figure 5-6 Joint probability density function for random variables *X* and *Y*.

5-2 Two Continuous Random Variables

Example 5-12

Let the random variable X denote the time until a computer server connects to your machine (in milliseconds), and let Y denote the time until the server authorizes you as a valid user (in milliseconds). Each of these random variables measures the wait from a common starting time and $X < Y$. Assume that the joint probability density function for X and Y is

$$f_{XY}(x, y) = 6 \times 10^{-6} \exp(-0.001x - 0.002y) \quad \text{for } x < y$$

Reasonable assumptions can be used to develop such a distribution, but for now, our focus is only on the joint probability density function.

5-2 Two Continuous Random Variables

Example 5-12

The region with nonzero probability is shaded in Fig. 5-8. The property that this joint probability density function integrates to 1 can be verified by the integral of $f_{XY}(x, y)$ over this region as follows:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dy\, dx = \int_{0}^{\infty}\left(\int_{x}^{\infty} 6 \times 10^{-6} e^{-0.001x-0.002y}\, dy\right) dx$$

$$= 6 \times 10^{-6} \int_{0}^{\infty}\left(\int_{x}^{\infty} e^{-0.002y}\, dy\right) e^{-0.001x}\, dx$$

$$= 6 \times 10^{-6} \int_{0}^{\infty}\left(\frac{e^{-0.002x}}{0.002}\right) e^{-0.001x}\, dx$$

$$= 0.003\left(\int_{0}^{\infty} e^{-0.003x}\, dx\right) = 0.003\left(\frac{1}{0.003}\right) = 1$$

5-2 Two Continuous Random Variables

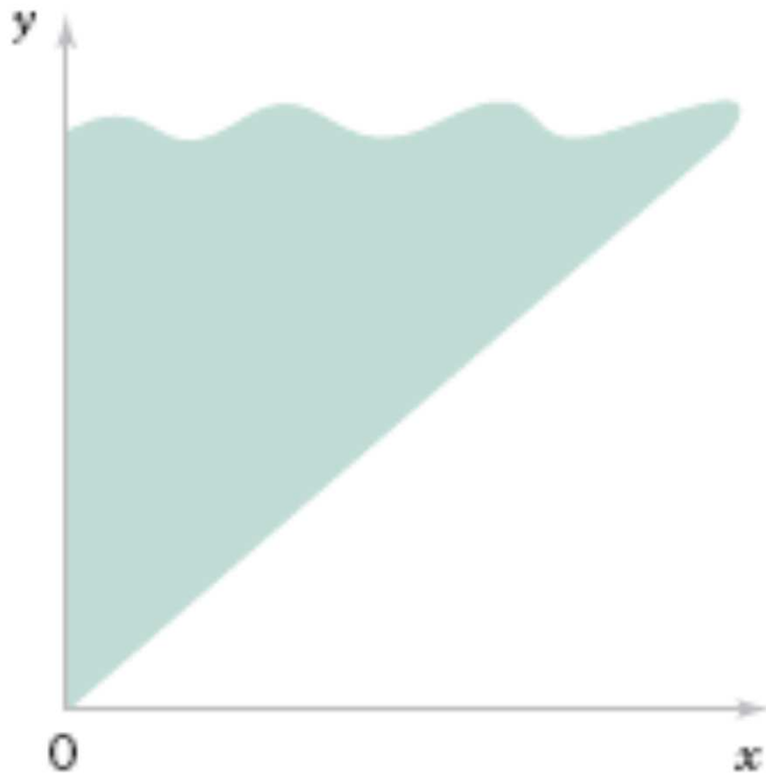


Figure 5-8 The joint probability density function of X and Y is nonzero over the shaded region.

5-2 Two Continuous Random Variables

Example 5-12

The probability that $X < 1000$ and $Y < 2000$ is determined as the integral over the darkly shaded region in Fig. 5-9.

$$P(X \leq 1000, Y \leq 2000) = \int_0^{1000} \int_x^{2000} f_{XY}(x, y)\, dy\, dx$$

$$= 6 \times 10^{-6} \int_0^{1000} \left(\int_x^{2000} e^{-0.002y}\, dy \right) e^{-0.001x}\, dx$$

$$= 6 \times 10^{-6} \int_0^{1000} \left(\frac{e^{-0.002x} - e^{-4}}{0.002} \right) e^{-0.001x}\, dx$$

$$= 0.003 \int_0^{1000} e^{-0.003x} - e^{-4} e^{-0.001x}\, dx$$

$$= 0.003 \left[\left(\frac{1 - e^{-3}}{0.003} \right) - e^{-4} \left(\frac{1 - e^{-1}}{0.001} \right) \right]$$

$$= 0.003(316.738 - 11.578) = 0.915$$

5-2 Two Continuous Random Variables

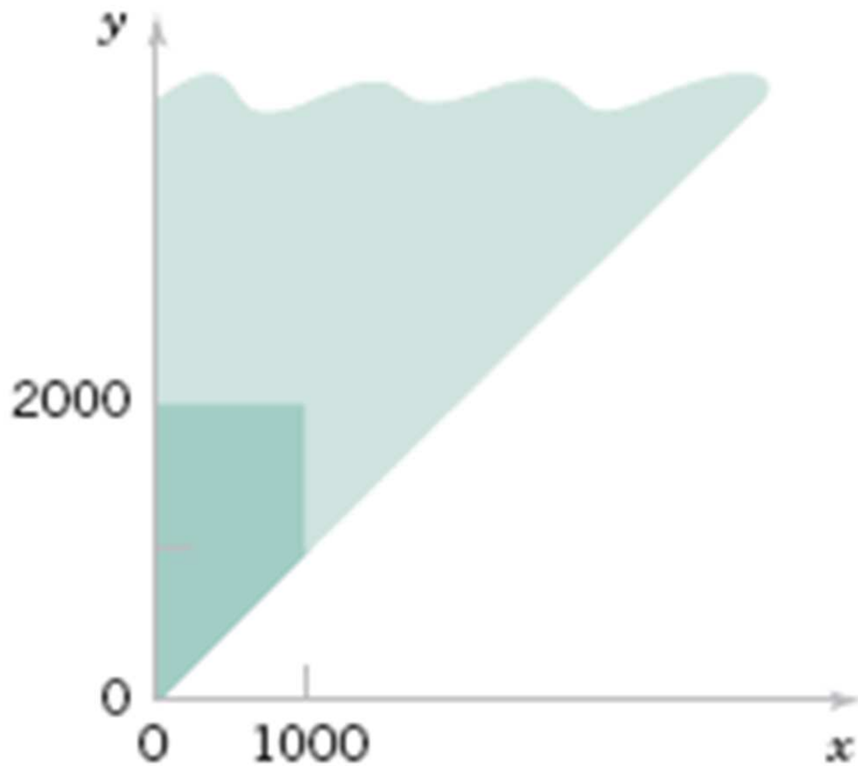


Figure 5-9 Region of integration for the probability that $X < 1000$ and $Y < 2000$ is darkly shaded.

5-2 Two Continuous Random Variables

5-2.2 Marginal Probability Distributions

Definition

If the joint probability density function of continuous random variables X and Y is $f_{XY}(x, y)$, the **marginal probability density functions** of X and Y are

$$f_X(x) = \int_y f_{XY}(x, y)\, dy \quad \text{and} \quad f_Y(y) = \int_x f_{XY}(x, y)\, dx \qquad (5\text{-}15)$$

where the first integral is over all points in the range of (X, Y) for which $X = x$ and the second integral is over all points in the range of (X, Y) for which $Y = y$

5-2 Two Continuous Random Variables

Example 5-13

For the random variables that denote times in Example 5-12, calculate the probability that Y exceeds 2000 milliseconds.

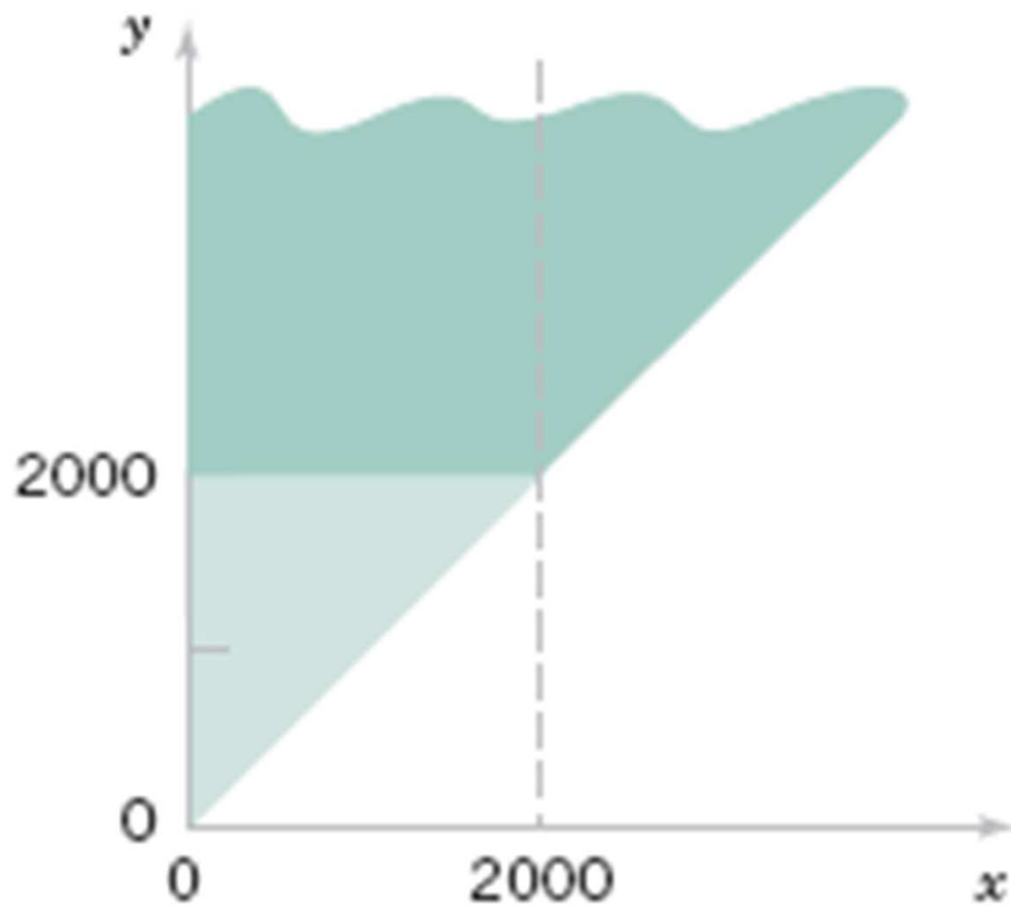
This probability is determined as the integral of $f_{XY}(x, y)$ over the darkly shaded region in Fig. 5-10. The region is partitioned into two parts and different limits of integration are determined for each part.

$$P(Y > 2000) = \int_0^{2000} \left(\int_{2000}^{\infty} 6 \times 10^{-6} e^{-0.001x - 0.002y} \, dy \right) dx$$

$$+ \int_{2000}^{\infty} \left(\int_{x}^{\infty} 6 \times 10^{-6} e^{-0.001x - 0.002y} \, dy \right) dx$$

5-2 Two Continuous Random Variables

Figure 5-10 Region of integration for the probability that $Y < 2000$ is darkly shaded and it is partitioned into two regions with $x < 2000$ and and $x > 2000$.



5-2 Two Continuous Random Variables

Example 5-13

The first integral is

$$6 \times 10^{-6} \int_{0}^{2000} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{2000}^{\infty} \right) e^{-0.001x}\, dx = \frac{6 \times 10^{-6}}{0.002} e^{-4} \int_{0}^{2000} e^{-0.001x}\, dx$$

$$= \frac{6 \times 10^{-6}}{0.002} e^{-4} \left(\frac{1 - e^{-2}}{0.001} \right) = 0.0475$$

The second integral is

$$6 \times 10^{-6} \int_{2000}^{\infty} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{x}^{\infty} \right) e^{-0.001x}\, dx = \frac{6 \times 10^{-6}}{0.002} \int_{2000}^{\infty} e^{-0.003x}\, dx$$

$$= \frac{6 \times 10^{-6}}{0.002} \left(\frac{e^{-6}}{0.003} \right) = 0.0025$$

5-2 Two Continuous Random Variables

Example 5-13

Therefore,

$$P(Y > 2000) = 0.0475 + 0.0025 = 0.05.$$

Alternatively, the probability can be calculated from the marginal probability distribution of Y as follows. For $y > 0$

$$f_Y(y) = \int_0^y 6 \times 10^{-6} e^{-0.001x - 0.002y}\, dx = 6 \times 10^{-6} e^{-0.002y} \int_0^y e^{-0.001x}\, dx$$

$$= 6 \times 10^{-6} e^{-0.002y} \left(\frac{e^{-0.001x}}{-0.001} \Big|_0^y \right) = 6 \times 10^{-6} e^{-0.002y} \left(\frac{1 - e^{-0.001y}}{0.001} \right)$$

$$= 6 \times 10^{-3} e^{-0.002y} (1 - e^{-0.001y}) \quad \text{for } y > 0$$

5-2 Two Continuous Random Variables

Example 5-13

We have obtained the marginal probability density function of Y. Now,

$$P(Y > 2000) = 6 \times 10^{-3} \int_{2000}^{\infty} e^{-0.002y}(1 - e^{-0.001y})\, dy$$

$$= 6 \times 10^{-3} \left[\left(\frac{e^{-0.002y}}{-0.002} \Bigg|_{2000}^{\infty} \right) - \left(\frac{e^{-0.003y}}{-0.003} \Bigg|_{2000}^{\infty} \right) \right]$$

$$= 6 \times 10^{-3} \left[\frac{e^{-4}}{0.002} - \frac{e^{-6}}{0.003} \right] = 0.05$$

5-2 Two Continuous Random Variables

5-2.3 Conditional Probability Distributions

Definition

Given continuous random variables X and Y with joint probability density function $f_{XY}(x, y)$, the **conditional probability density function** of Y given $X = x$ is

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)} \quad \text{for} \quad f_X(x) > 0 \qquad (5\text{-}16)$$

5-2 Two Continuous Random Variables

5-2.3 Conditional Probability Distributions

Because the conditional probability density function $f_{Y|x}(y)$ is a probability density function for all y in R_x, the following properties are satisfied:

(1) $f_{Y|x}(y) \geq 0$

(2) $\int_{R_x} f_{Y|x}(y)\, dy = 1$

(3) $P(Y \in B | X = x) = \int_B f_{Y|x}(y)\, dy$ for any set B in the range of Y

(5-17)

5-2 Two Continuous Random Variables

Example 5-14

For the random variables that denote times in Example 5-12, determine the conditional probability density function for Y given that $X = x$.

First the marginal density function of x is determined. For $x > 0$

$$f_X(x) = \int_x^\infty 6 \times 10^{-6} e^{-0.001x - 0.002y} \, dy = 6 \times 10^{-6} e^{-0.001x} \left(\frac{e^{-0.002y}}{-0.002} \Bigg|_x^\infty \right)$$

$$= 6 \times 10^{-6} e^{-0.001x} \left(\frac{e^{-0.002x}}{0.002} \right) = 0.003 e^{-0.003x} \quad \text{for} \quad x > 0$$

This is an exponential distribution with $\lambda = 0.003$. Now, for $0 < x$ and $x < y$ the conditional probability density function is

$$f_{Y|x}(y) = f_{XY}(x, y)/f_x(x) = \frac{6 \times 10^{-6} e^{-0.001x - 0.002y}}{0.003 e^{-0.003x}}$$

$$= 0.002 e^{0.002x - 0.002y} \quad \text{for } 0 < x \quad \text{and} \quad x < y$$

5-2 Two Continuous Random Variables

Example 5-14

The conditional probability density function of Y, given that $x = 1500$, is nonzero on the solid line in Fig. 5-11.

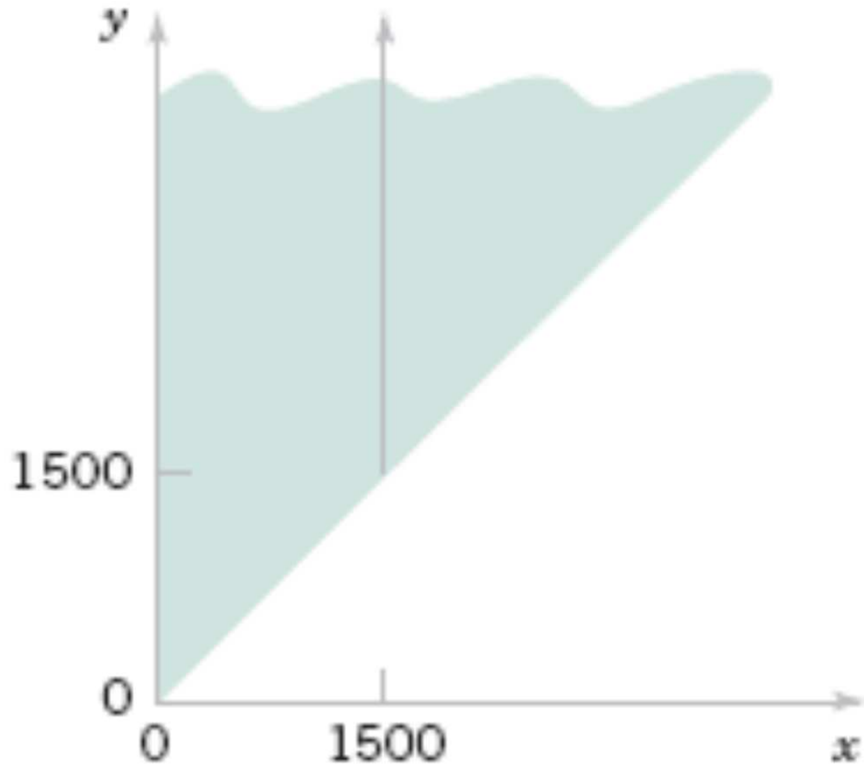


Figure 5-11 The conditional probability density function for Y, given that x = 1500, is nonzero over the solid line.

5-2 Two Continuous Random Variables

Definition: Conditional Mean and Variance

The **conditional mean** of Y given $X = x$, denoted as $E(Y|x)$ or $\mu_{Y|x}$, is

$$E(Y|x) = \int y f_{Y|x}(y)\, dy$$

and the **conditional variance** of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$, is

$$V(Y|x) = \int_{-\infty} (y - \mu_{Y|x})^2 f_{Y|x}(y)\, dy = \int y^2 f_{Y|x}(y)\, dy - \mu^2_{Y|x} \qquad (5\text{-}18)$$

5-2 Two Continuous Random Variables

5-2.4 Independence

Definition

For continuous random variables X and Y, if any one of the following properties is true, the others are also true, and X and Y are said to be **independent.**

(1) $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y

(2) $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$

(4) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for any sets A and B in the range of X and Y, respectively. (5-19)

5-2 Two Continuous Random Variables

Example 5-16

For the joint distribution of times in Example 5-12, the

- Marginal distribution of Y was determined in Example 5-13.
- Conditional distribution of Y given $X = x$ was determined in Example 5-14.

Because the marginal and conditional probability densities are not the same for all values of x, property (2) of Equation 5-18 implies that the random variables are not independent. The fact that these variables are not independent can be determined quickly by noticing that the range of (X, Y), shown in Fig. 5-8, is not rectangular. Consequently, knowledge of X changes the interval of values for Y that receives nonzero probability.

5-2 Two Continuous Random Variables

Example 5-18

Let the random variables X and Y denote the lengths of two dimensions of a machined part, respectively. Assume that X and Y are independent random variables, and further assume that the distribution of X is normal with mean 10.5 millimeters and variance 0.0025 (millimeter)^2 and that the distribution of Y is normal with mean 3.2 millimeters and variance 0.0036 (millimeter)^2. Determine the probability that $10.4 < X < 10.6$ and $3.15 < Y < 3.25$.

Because X and Y are independent,

$$P(10.4 < X < 10.6, 3.15 < Y < 3.25) = P(10.4 < X < 10.6)P(3.15 < Y < 3.25)$$

$$= P\left(\frac{10.4 - 10.5}{0.05} < Z < \frac{10.6 - 10.5}{0.05}\right) P\left(\frac{3.15 - 3.2}{0.06} < Z < \frac{3.25 - 3.2}{0.06}\right)$$

$$= P(-2 < Z < 2)P(-0.833 < Z < 0.833) = 0.566$$

where Z denotes a standard normal random variable.

5-2 Two Continuous Random Variables

Example 5-20

In an electronic assembly, let the random variables X_1, X_2, X_3, X_4 denote the lifetimes of four components in hours. Suppose that the joint probability density function of these variables is

$$f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4) = 9 \times 10^{-2} e^{-0.001x_1 - 0.002x_2 - 0.0015x_3 - 0.003x_4}$$
$$\text{for } x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$$

What is the probability that the device operates for more than 1000 hours without any failures?

The requested probability is $P(X_1 > 1000, X_2 > 1000, X_3 > 1000, X_4 > 1000)$, which equals the multiple integral of $f_{X_1X_2X_3X_4}(x_1, x_2, x_3, x_4)$ over the region $x_1 > 1000, x_2 > 1000, x_3 > 1000, x_4 > 1000$. The joint probability density function can be written as a product of exponential functions, and each integral is the simple integral of an exponential function. Therefore,

$$P(X_1 > 1000, X_2 > 1000, X_3 > 1000, X_4 > 1000) = e^{-1-2-1.5-3} = 0.00055$$

5-2 Two Continuous Random Variables

Definition: Marginal Probability Density Function

If the joint probability density function of continuous random variables $X_1, X_2, \ldots, X_p$ is $f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$, the **marginal probability density function** of X_i is

$$f_{X_i}(x_i) = \iint\limits_{R_{x_i}} \cdots \int f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_{i-1}\, dx_{i+1} \ldots dx_p \qquad (5\text{-}21)$$

where the integral is over all points in the range of $X_1, X_2, \ldots, X_p$ for which $X_i = x_i$.

5-2 Two Continuous Random Variables

Mean and Variance from Joint Distribution

$$E(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} x_i f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p$$

and (5-22)

$$V(X_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (x_i - \mu_{X_i})^2 f_{X_1X_2\ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_1\, dx_2 \ldots dx_p$$

5-2 Two Continuous Random Variables

Distribution of a Subset of Random Variables

If the joint probability density function of continuous random variables $X_1, X_2, \ldots, X_p$ is $f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)$, the **probability density function** of $X_1, X_2, \ldots, X_k$, $k < p$, is

$$f_{X_1 X_2 \ldots X_k}(x_1, x_2, \ldots, x_k) = \int_{R_{x_1 x_2 \ldots x_k}} \int \cdots \int f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots, x_p)\, dx_{k+1}\, dx_{k+2} \ldots dx_p \qquad (5\text{-}23)$$

where the integral is over all points in the range of $X_1, X_2, \ldots, X_k$ for which $X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k$.

5-2 Two Continuous Random Variables

Conditional Probability Distribution

Definition

Continuous random variables $X_1, X_2, \ldots, X_p$ are **independent** if and only if

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2 \ldots, x_p) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_p}(x_p) \quad \text{for } all\ x_1, x_2, \ldots, x_p \qquad (5\text{-}24)$$

5-2 Two Continuous Random Variables

Example 5-23

Suppose X_1, X_2, and X_3 represent the thickness in micrometers of a substrate, an active layer, and a coating layer of a chemical product. Assume that X_1, X_2, and X_3 are independent and normally distributed with $\mu_1 = 10000, \mu_2 = 1000, \mu_3 = 80, \sigma_1 = 250, \sigma_2 = 20$, and $\sigma_3 = 4$, respectively.

The specifications for the thickness of the substrate, active layer, and coating layer are $9200 < x_1 < 10800$, $950 < x_2 < 1050$, and $75 < x_3 < 85$, respectively. What proportion of chemical products meets all thickness specifications? Which one of the three thicknesses has the least probability of meeting specifications?

The requested probability is $P(9200 < X_1 < 10800, 950 < X_2 < 1050, 75 < X_3 < 85$. Because the random variables are independent,

$$P(9200 < X_1 < 10800, 950 < X_2 < 1050, 75 < X_3 < 85)$$
$$= P(9200 < X_1 < 10800)P(950 < X_2 < 1050)P(75 < X_3 < 85)$$

5-2 Two Continuous Random Variables

Example 5-23

After standardizing, the above equals

$$P(-3.2 < Z < 3.2)P(-2.5 < Z < 2.5)P(-1.25 < Z < 1.25)$$

where Z is a standard normal random variable. From the table of the standard normal distribution, the above equals

$$(0.99862)(0.98758)(0.78870) = 0.7778$$

The thickness of the coating layer has the least probability of meeting specifications. Consequently, a priority should be to reduce variability in this part of the process.