5-1 Two or more random variables

- **5-1.1 Joint probability distributions**
  - Ex. 5-1. Calls are made to check the airline schedule at your departure city. You monitor the number of bars of signal strength on your cell phone and the number of times you have to state the name of your departure city before the voice system recognizes the name.
    - ☞ Let X denote the number of bars of signal strength on your cell phone and Y denote the number of times you need to state your departure city.
  - If X and Y are discrete random variables, the joint probability distribution of X and Y is a description of the set of points (x, y) in the range of (X, Y) along with the probability of each point.

☞The joint probability distribution of two random variables is sometimes referred to as the **bivariate probability distribution** or **bivariate distribution** of the random variables.

☞ $P(X=x \text{ and } Y=y)=P(X=x,Y=y)$

Ex. $P(X=1,Y=4)=0.15$

	x = number of bars of signal strength		
y = number of times city name is stated	**1**	**2**	**3**
4	0.15	0.1	0.05
3	0.02	0.1	0.05
2	0.02	0.03	0.2
1	0.01	0.02	0.25

- Joint probability distribution of two discrete random variables
  - $P(X=x,Y=y)$
  - Def. The **joint probability mass function** of the discrete random variables X and Y, denoted as $f_{XY}(x, y)$, satisfies
    - ☞ $f_{XY}(x, y) \geq 0$
    - ☞ $\sum_x \sum_y f_{XY}(x, y) = 1$
    - ☞ $f_{XY}(x, y) = P(X=x, Y=y)$
  - Ex. 5-1
    - ☞ $f_{XY}(1,4) = P(X=1, Y=4) = 0.15$

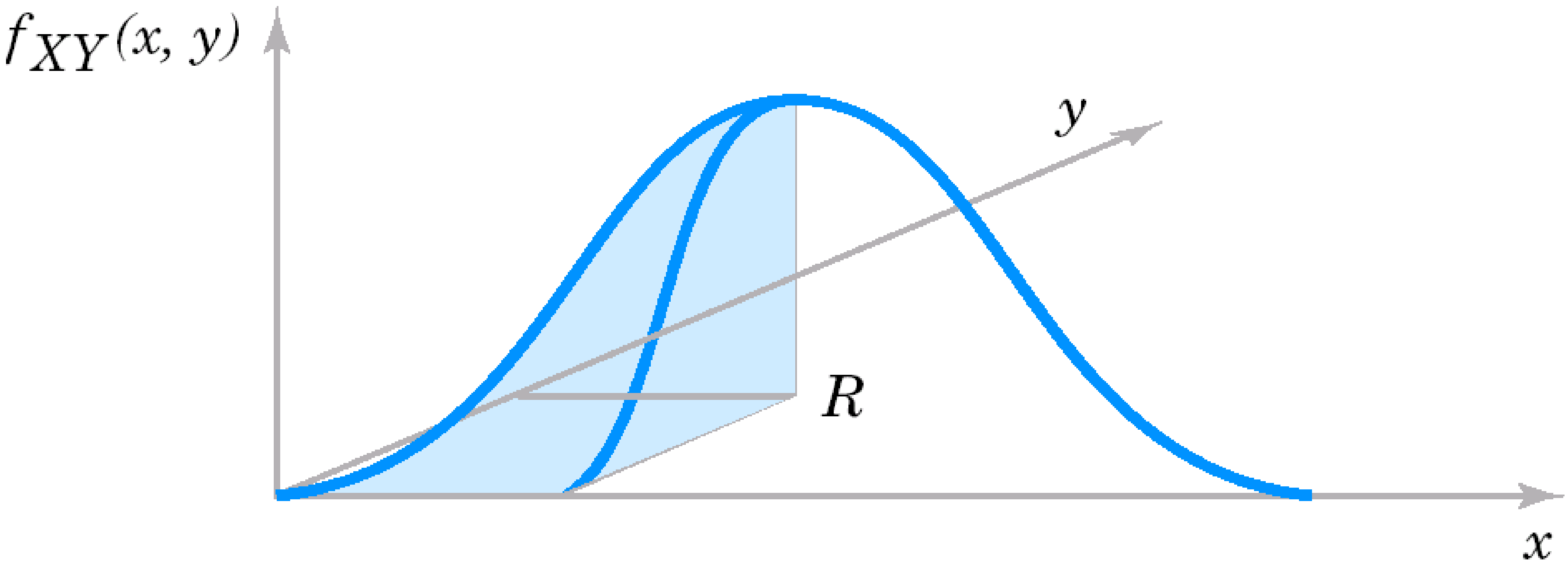
- Def. A **joint probability density function** for the continuous random variables X and Y, denoted as $f_{XY}(x,y)$, satisfies the following properties:

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

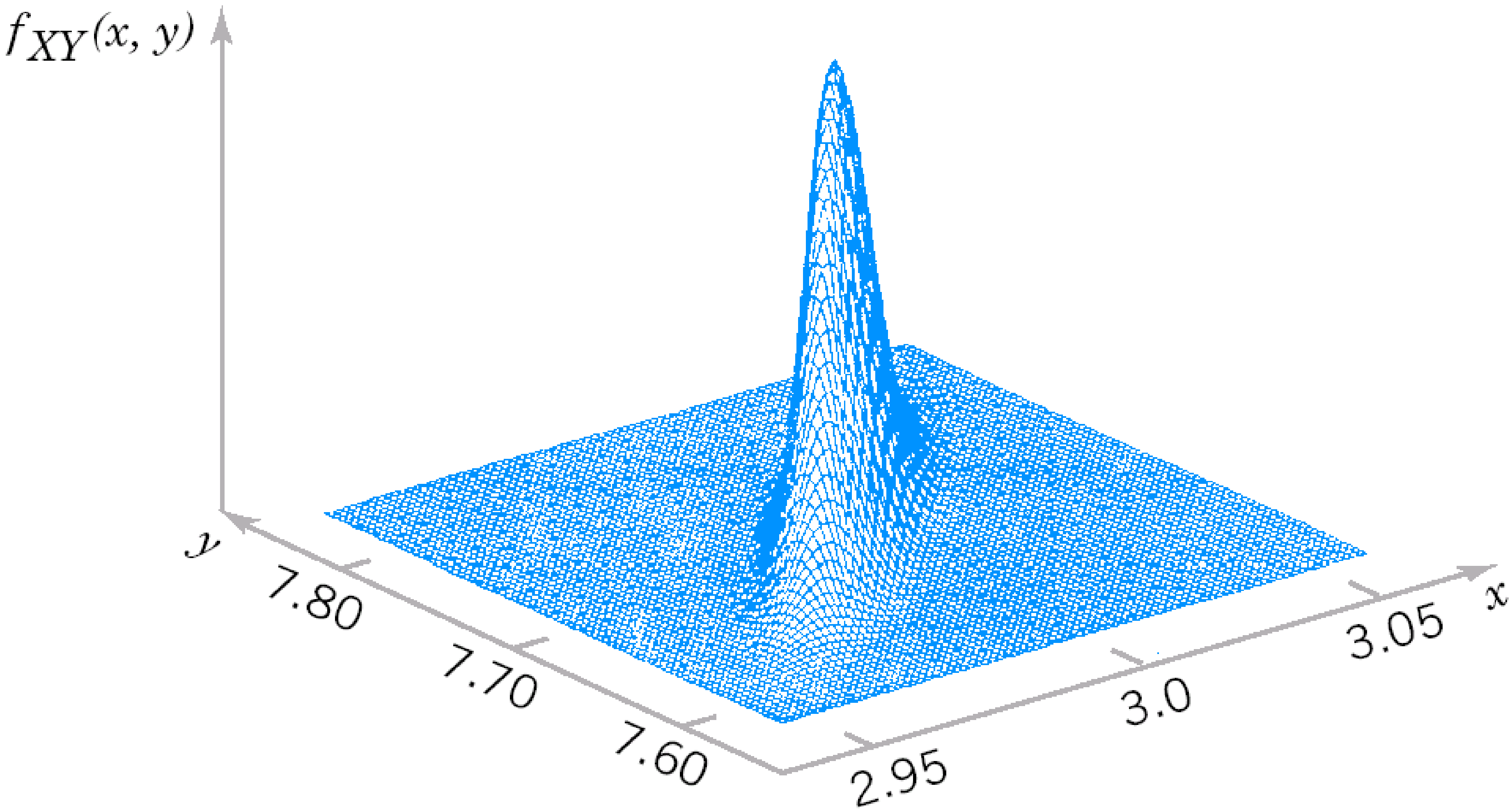
(2) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1$$

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

$$P([X, Y] \in R) = \iint_R f_{XY}(x, y)\, dx\, dy$$



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



$f_{XY}(x, y)$
y
7.80
7.70
7.60
x
2.95
3.0
3.05

- Ex. 5-2 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 $f_{XY}(x,y)=6\times10^{-6}\exp(-0.001x-0.002y)$ for $x<y$

☞ $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dydx$

$$=\int_0^{\infty}(\int_x^{\infty} 6\times10^{-6}e^{-0.001x-0.002y}dy)dx=1$$

☞ $P(X\leq1000, Y\leq2000)=?$

$$P(X\leq1000, Y\leq2000)=\int_0^{1000}\int_x^{2000} f_{XY}(x,y)dydx=0.915$$

- **5-1.2 Marginal probability distributions**
  - Ex. 5-3. in Fig. 5-6
    - ☞ $P(X=3)=P(X=3,Y=1)+P(X=3,Y=2)+P(X=3,Y=3)+P(X=3,Y=4)$

      $=0.25+0.2+0.05+0.05=0.55$

	x = number of bars of signal strength			
y = number of times city name is stated	**1**	**2**	**3**	**Marginal probability distribution of Y**
4	0.15	0.1	0.05	0.3
3	0.02	0.1	0.05	0.17
2	0.02	0.03	0.2	0.25
1	0.01	0.02	0.25	0.28
	0.2	0.25	0.55	
	Marginal probability distribution of X			

- Def. If X and Y are discrete random variables with joint probability mass function $f_{XY}(x, y)$, then the **marginal probability mass functions** of X and Y are

$$f_X(x) = P(X = x) = \sum_y f_{XY}(x, y) \text{ and } f_Y(y) = P(Y = y) = \sum_x f_{XY}(x, y)$$

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

☞ $E(X)$ and $V(X)$ can be obtained by first calculating the marginal probability distribution of X then determining $E(X)$ and $V(X)$ by the usual method.

☞ In Fig. 5-6, $E(X)=1(0.2)+2(0.25)+3(0.55)=2.35$

- Def. 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

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

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

- Ex. 5-13. $f_Y(y)$=? $\quad f_Y(y) = \int_0^y 6\times 10^{-6} e^{-0.001x-0.002y} dx$

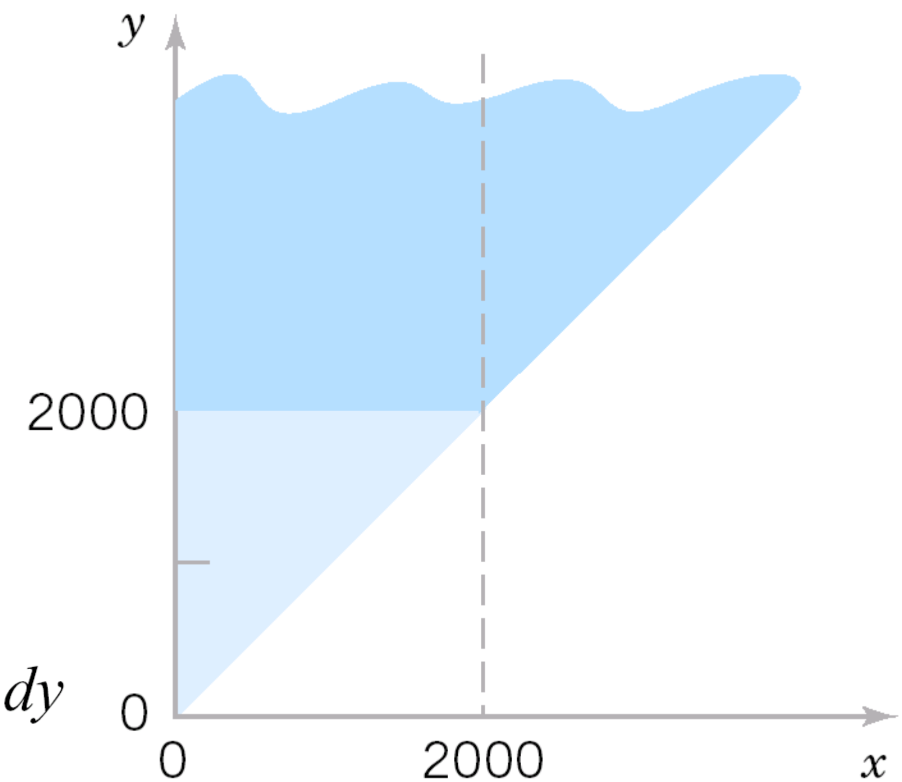
$$= 6\times 10^{-3} e^{-0.002y}(1-e^{-0.001y}) \text{ for } y>0$$

☞ $P(a<Y<b) = \int_a^b f_Y(y)dy$

➢Ex. 5-4 $P(Y>2000)$

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

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



☞ $P(Y > 2000)$

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

= 0.05

- Mean and variance from joint distribution

$$E(X) = \mu_X = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_{-\infty}^{\infty} x \left[\int_y f_{XY}(x, y)\, dy \right] dx$$

$$= \iint_R x f_{XY}(x, y)\, dx\, dy$$

$$V(X) = \sigma_x^2 \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\, dx = \int_{-\infty}^{\infty} (x - \mu_X)^2 \left[\int_y f_{XY}(x, y)\, dy \right] dx$$

$$= \iint_R (x - \mu_X)^2 f_{XY}(x, y)\, dx\, dy$$