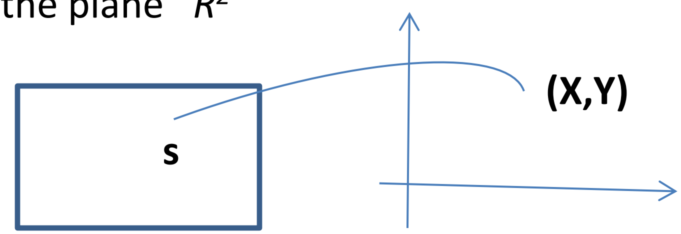
5.1 Jointly Distributed Random Variables

One random variable X: a function from the sample space *S* into the set of real numbers *R.*

Two random variables defined on the same sample space (X,Y): a function from the sample space *S* into the plane R^2



If the set of possible values of (X,Y) is countable (in particular if each X and Y are discrete), then the joint distribution of (X,Y) is called ***discrete*.**

If the set of possible values of (X,Y) is R^2 or a region in R^2 and (X,Y) have a joint probability density function, then the joint distribution of (X,Y) is called ***continuous*.**

Two Discrete Random Variables

DEFINITION

Let X and Y be two discrete rv's defined on the sample space $\mathcal{S}$ of an experiment. The **joint probability mass function** $p(x, y)$ is defined for each pair of numbers (x, y) by

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

It must be the case that $p(x, y) \geq 0$ and $\sum_x \sum_y p(x, y) = 1$.

Now let A be any set consisting of pairs of (x, y) values (e.g., $A = \{(x, y): x + y = 5\}$ or $\{(x, y): \max(x, y) \leq 3\}$). Then the probability $P[(X, Y) \in A]$ is obtained by summing the joint pmf over pairs in A:

$$P[(X, Y) \in A] = \sum_{(x, y) \in A} \sum p(x, y)$$

Example 1. Joint probability table

$p(x, y)$		y = 0	100	200
x	100	.20	.10	.20
	250	.05	.15	.30

Here P(X = 100 and Y = 100) = 0.10 and P(X+Y ≥ 300) = P((X,Y) ∈ A) = 0.65.

DEFINITION

The **marginal probability mass functions** of X and of Y, denoted by $p_X(x)$ and $p_Y(y)$, respectively, are given by

$$p_X(x) = \sum_y p(x, y) \qquad p_Y(y) = \sum_x p(x, y)$$

$p_X(x) = P(X = x)$ is the pmf of *X*, *$p_Y(y) = P(Y = y)$* is the pmf of *Y*.

Example 1 cont. Marginal distributions: Find the distribution of Y

$p(x, y)$		y = 0	100	200	$p_X(x)$
x	100	.20	.10	.20	**0.50**
	250	.05	.15	.30	**0.50**
	$p_Y(y)$	**.25**	**.25**	**.50**	

Conditional distributions are defined as

$$p_{Y|X}(y|x) = P(Y = y|\, X = x) = \frac{P(Y = y \text{ and } X = x)}{P(X = x)}$$

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{P(X = x \text{ and } Y = y)}{P(Y = y)}.$$

Example 1 cont. Find the conditional distribution of Y, given that X = 250, and compute E(Y|X=250)

y	0	100	200
$p_{Y\|X}(y\|250)$	0.1	0.3	0.6

$E(Y|X=250) = 0 \times 0.1 + 100 \times 0.3 + 200 \times 0.6 = 150$

Two Continuous Random Variables

DEFINITION

Let X and Y be continuous rv's. A **joint probability density function** $f(x, y)$ for these two variables is a function satisfying $f(x, y) \geq 0$ and $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$. Then for any two-dimensional set A

$$P[(X, Y) \in A] = \iint_A f(x, y)\, dx\, dy$$

In particular, if A is the two-dimensional rectangle $\{(x, y): a \leq x \leq b, c \leq y \leq d\}$, then

$$P[(X, Y) \in A] = P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x, y)\, dy\, dx$$

We can think of $f(x, y)$ as specifying a surface at height $f(x, y)$ above the point (x, y) in a three-dimensional coordinate system. Then $P[(X, Y) \in A]$ is the volume underneath this surface and above the region A, analogous to the area under a curve in the one-dimensional case. This is illustrated in Figure 5.1.

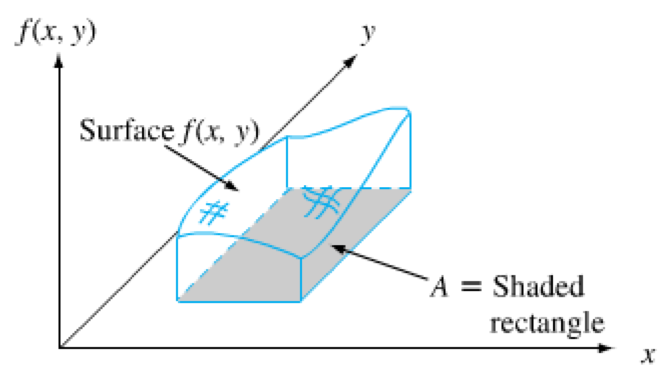


Figure 5.1 $P[(X, Y) \in A$ = volume under density surface above A

Properties of joint pdf:

1. $f(x, y)$ is defined for every real x,y
2. $f(x, y) \geq 0$
3. The set $D = \{ (x, y) \in R^2\colon\ f(x, y) > 0 \}$ is the set of all possible values of *(X,Y)*
4. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx\, dy = 1$

DEFINITION

The **marginal probability density functions** of X and Y, denoted by $f_X(x)$ and $f_Y(y)$, respectively, are given by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \qquad \text{for } -\infty < x < \infty$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \qquad \text{for } -\infty < y < \infty$$

$f_X(x)$ is the pdf of X, $f_Y(y)$ is the pdf of Y.

DEFINITION

Let X and Y be two continuous rv's with joint pdf $f(x, y)$ and marginal X pdf $f_X(x)$. Then for any X value x for which $f_X(x) > 0$, the **conditional probability density function of Y given that $X = x$** is

$$f_{Y|X}(y \mid x) = \frac{f(x, y)}{f_X(x)} \qquad -\infty < y < \infty$$

If X and Y are discrete, replacing pdf's by pmf's in this definition gives the **conditional probability mass function of Y when $X = x$.**

DEFINITION

Two random variables X and Y are said to be **independent** if for every pair of x and y values,

$$p(x, y) = p_X(x) \cdot p_Y(y) \qquad \text{when } X \text{ and } Y \text{ are discrete}$$

or (5.1)

$$f(x, y) = f_X(x) \cdot f_Y(y) \qquad \text{when } X \text{ and } Y \text{ are continuous}$$

If (5.1) is not satisfied for all (x, y), then X and Y are said to be **dependent.**

Example (example 5.5, p. 197) .

Let
$$f(x, y) = \begin{cases} 24xy & 0 \le x \le 1, 0 \le y \le 1, x + y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

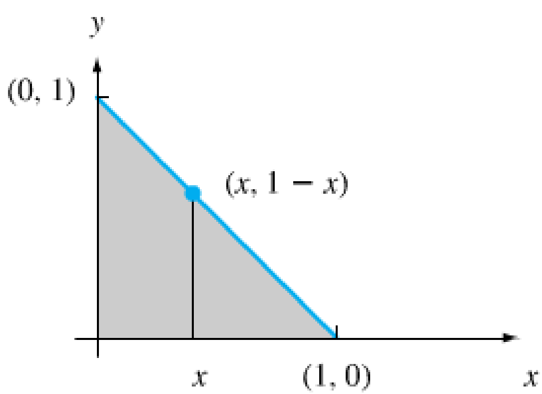


Figure 5.2 Region of positive density for Example 5.5

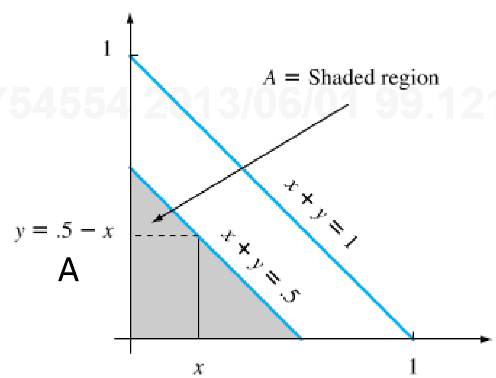


Figure 5.3 Computing $P[(X, Y) \in A]$ for Example 5.5

1. Find the probability P(X+Y ≤ 0.5) = P((X,Y) ∈ A)

P(X+Y ≤ 0.5) = $P((X, Y) \in A) = \int_A \int f(x, y)\, dx\, dy = \int_0^{.5} \int_0^{.5-x} 24xy\, dy\, dx = .0625$

2. Find the pdf of X

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \begin{cases} \int_0^{1-x} 24xy\, dy = 12x(1-x)^2 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

3. Find the pdf of Y $f_Y(y) = 12\, y\, (1-y)^2$, $0 \le y \le 1$
4. Are X and Y independent? No, $f(x,y) \neq f_X(x) f_Y(y)$
5. Compute P(Y > 0.5| X = 0.25).

$$f_{Y|X}(y|0.25) = \frac{f(0.25, y)}{f_X(0.25)} = \begin{cases} \frac{32}{9} y & if \quad 0 \le y \le .75 \\ 0 & otherwise \end{cases}$$

Hence

$$P(Y > 0.5|X = 0.25) = \int_{0.5}^{\infty} f_{Y|X}(y|0.25)\, dy = \int_{0.5}^{0.75} \frac{32}{9} y\, dy = \frac{5}{9}$$

6. Compute E(Y| X = 0.25)

$$E(Y|X = 0.25) = \int_{-\infty}^{\infty} y\, f_{Y|X}(y|0.25)\, dy = \int_{0}^{0.75} \frac{32}{9} y^2\, dy = \frac{1}{2}$$

5.1 EXERCISES

A restaurant serves three fixed-price dinners costing \$12, \$15, and \$20. For a randomly selected couple dining at this restaurant, let X = the cost of the man's dinner and Y = the cost of the woman's dinner. The joint pmf of X and Y is given in the following table:

$p(x, y)$		y = 12	15	20
x	12	.05	.05	.10
	15	.05	.10	.35
	20	0	.20	.10

a. Compute the marginal pmf's of X and Y.
b. What is the probability that the man's and the woman's dinner cost at most \$15 each?
c. Are X and Y independent? Justify your answer.
d. What is the expected total cost of the dinner for the two people?

e. Compute $P(X \geq 15 \mid Y = 20)$
f. Determine the conditional pmf of Y, given that X=15.
g. Compute E(Y|X=15)

ANSWERS:
a. Compute row and column sums on margins of the table
b. 0.25, c.NO, d. 36.05, e. 9/11, f. ,1, .2, .7

12. Two components of a minicomputer have the following joint pdf for their useful lifetimes X and Y:

$$f(x, y) = \begin{cases} xe^{-x(1+y)} & x \geq 0 \text{ and } y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

a. What is the probability that the lifetime X of the first component exceeds 3?
b. What are the marginal pdf's of X and Y? Are the two lifetimes independent? Explain.
c. What is the probability that the lifetime of at least one component exceeds 3?

ANSWERS:
a. $f_X(x) = e^{-x}$, $x>0$; $f_Y(y) = 1/(1+y)^2$, $y>0$;
b.NO
c.0.2998

More than Two Random variables

DEFINITION

If $X_1, X_2, \ldots, X_n$ are all discrete random variables, the joint pmf of the variables is the function

$$p(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

If the variables are continuous, the joint pdf of $X_1, \ldots, X_n$ is the function $f(x_1, x_2, \ldots, x_n)$ such that for any n intervals $[a_1, b_1], \ldots, [a_n, b_n]$,

$$P(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

DEFINITION

The random variables $X_1, X_2, \ldots, X_n$ are said to be **independent** if for *every* subset $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ of the variables (each pair, each triple, and so on), the joint pmf or pdf of the subset is equal to the product of the marginal pmf's or pdf's.