Ch8: Bivariate Distributions

8.1 Joint distributions of two random variables

- We now consider two or or more random variables that are defined simultaneously on the same sample space.
  - Functions
  - Functions

Joint Probability Mass Functions

- **Definition**
  - *Let X and Y be two discrete random variables defined on the same sample space.*
  - *Let the sets of possible values of X and Y be A and B, respectively. The function*

    $p(x, y) =$

    is called the ***joint probability mass function*** *of X and Y .*
- Note that $p(x, y) \geq 0$. If $x \notin A$ or $y \notin B$, then Also,

Definition

- *Let X and Y have joint probability mass function p(x, y).*
- *Let A be the set of possible values of X and B be the set of possible values of Y .*
- *Then the functions* $p_X(x) =$ *and* $p_Y(y) =$ *are called, respectively, the* ***functions*** *of X and Y.*

Example 8.1

- A small college has 90 male and 30 female professors. An ad hoc committee of five is selected at random to write the vision and mission of the college.
- Let X and Y be the number of men and women on this committee, respectively.

  (a) Find the joint probability mass function of X and Y.

  (b) Find p_X and p_Y, the marginal probability mass functions of X and Y.

Example 8.1

(a)

1. The set of possible values for both X and Y is {0, 1, 2, 3, 4, 5}.
2.

$$p(x,y)=\begin{cases} & \text{if } x, y \in \{0,1,2,3,4,5\},\ x+y=5 \\ 0 & \text{otherwise} \end{cases}$$

(b)

1.

Since $p(x, y) = 0$ if $x + y \neq 5$, $\sum_{y=0}^{5} p(x,y) =$

and $\sum_{x=0}^{5} p(x,y) =$

$p_X(x) =$, $p_Y(y) =$ $x, y \in \{0,1,2,3,4,5\}$

- Let X and Y be discrete random variables with joint probability mass function $p(x, y)$.
- Let the sets of possible values of X and Y be A and B, respectively.
- To find $E(X)$ and $E(Y)$, first we calculate p_X and p_Y, the marginal probability mass functions of X and Y, respectively. Then we will use the following formulas.

$$E(X) = \qquad ; \quad E(Y) =$$

Example 8.3

- Let the joint probability mass function of random variables *X* and *Y* be given by

$$p(x,y)=\begin{cases}\frac{1}{70}x(x+y) & \text{if } x=1,2,3, \quad y=3,4\\ 0 & \text{elsewhere}\end{cases}$$

  Find *E(X)* and *E(Y)*.

1. First we need to calculate
2.

$$px(x)= \qquad =\frac{1}{70}x(x+3)+\frac{1}{70}x(x+4)=\frac{1}{35}x^2+\frac{1}{10}x$$

$$py(y)= \qquad =\frac{1}{5}+\frac{3}{35}y$$

Example 8.3

3.

$$E(X) = \qquad = \sum_{x=1}^{3} x\left(\frac{1}{35}x^2 + \frac{1}{10}x\right) = \frac{17}{7} \approx 2.43$$

$$E(Y) = \qquad = \sum_{y=3}^{4} y\left(\frac{1}{5} + \frac{3}{35}y\right) = \frac{124}{35} \approx 3.54$$

Theorem 8.1

- *Let p(x, y) be the joint probability mass function of discrete random variables X and Y.*
- *Let A and B be the set of possible values of X and Y, respectively.*
- *If h is a function of two variables from* $\mathbf{R}^2$ *to* $\mathbf{R}$*, then h(X, Y) is a discrete random variable with the expected value given by*

$$E[h(X,Y)] =$$

provided that the sum is absolutely convergent.

(generalization of Theorem 4.2)

Corollary

- *For discrete random variables X and Y ,* ***$E(X+Y)=E(X)+E(Y)$.***

Proof: In Theorem 8.1 let h(x,y)=x+y. Then

$$E(X+Y)=$$

$$=$$

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

Joint Probability Density Functions

- **Definition**

  Two random variables X and Y, defined on the same sample space, have a continuous joint distribution if there exists a nonnegative function of two variables, f (x, y) on $\mathbf{R} \times \mathbf{R}$*, such that for any region R in the xy-plane that can be formed from rectangles by a countable number of set operations,*

$$P((X, Y) \in R) =$$

The function f (x, y) is called the ***joint probability density function*** *of X and Y.*

- Let $R = \{(x, y) : x \in A, y \in B\}$, where A and B are *any* subsets of real numbers that can be constructed from intervals by a countable number of set operations. Then (8.2) gives

$$P(X \in A, Y \in B) = \int_B \int_A f(x, y)\,dx\,dy$$

Letting $A = (-\infty,\infty)$, $B = (-\infty,\infty)$, (8.3) implies the relation

Definition

- *Let X and Y have joint probability density function f(x, y); then the functions*

$$f_X(x) =$$

$$f_Y(y) =$$

*are called, respectively, the **marginal probability density functions** of X and Y .*

- Let X and Y be two random variables (discrete, continuous, or mixed).
- The **joint probability distribution function**, or *joint cumulative probability distribution function*, or simply the *joint distribution of X and Y*, is defined by

$$F(t,u) =$$

for all $-\infty < t, u < \infty$.

- The **marginal probability distribution function of *X***, F_X, can be found from F as follows:

$$F_X(t) =$$

$$\equiv F(t, \infty).$$

- Similarly, F_Y, the **marginal probability distribution function of *Y***, is

$$F_Y(u) = P(Y \leq u) \equiv$$

Suppose that the joint density function of X and Y is $f(x, y)$

$F(x, y) = P(X \le x, Y \le y) =$

Assuming that the partial derivatives of F exist, we get

$$f(x, y) =$$

Moreover, $F_X(x) =$

$$= \int_{-\infty}^{x} \left(\int_{-\infty}^{\infty} f(t, u)\, du \right) dt = \int_{-\infty}^{x} f_X(t)\, dt,$$

and similarly, $F_Y(y) =$

And $F_X'(x) = f_X(x), \quad F_Y'(y) = f_Y(y)$

$$E(X) =$$

$$E(Y) =$$

Example 8.4

- The joint probability density function of random variables *X* and *Y* is given by

$$f(x,y)=\begin{cases}\lambda xy^2 & 0\le x\le y\le 1\\ 0 & \text{otherwise.}\end{cases}$$

(a) Determine the value of λ.

(b) Find the marginal probability density functions of *X* and *Y*.

(c) Calculate *E(X)* and *E(Y)*.

Example 8.4

(a) To find λ, note that

Therefore,

$$1 = \int_0^1 \left(\int_x^1 y^2 dy\right)\lambda x dx = \int_0^1 [\frac{1}{3}y^3]_x^1 \lambda x dx = \lambda \int_0^1 (\frac{1}{3} - \frac{1}{3}x^3) x dx$$

$$= \frac{\lambda}{3}\int_0^1 (1 - x^3) x dx = \frac{\lambda}{3}\int_0^1 (x - x^4) dx = \frac{\lambda}{3}[\frac{1}{2}x^2 - \frac{1}{5}x^5]_0^1 = \frac{\lambda}{10}$$

hence $\lambda = 10$.

(b) fx(x) = $= \left[\frac{10}{3}xy^3\right]_x^1 = \frac{10}{3}x(1 - x^3)$

fy(y) = $= \left[5x^2y^2\right]_0^y = 5y^4$

(c) E(X) = $= \frac{5}{9}$; E(Y) $= \int_0^1 y \cdot 5y^4 dy = \frac{5}{6}$

Example 8.5

- For $\lambda > 0$, let

$$F(x,y) = \begin{cases} 1 - \lambda e^{-\lambda(x+y)} & \text{if } x > 0,\ y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Determine if F is the joint probability distribution function of two random variables X and Y.

If F is the joint density of X and Y then

is the joint density function of X and Y. But

$$\frac{\partial^2}{\partial x \partial y} F(x,y) = \begin{cases} & \text{if } x>0,\ y>0 \\ 0 & \text{otherwise.} \end{cases}$$

Since , it cannot be a joint density function.

Therefore, F is not a joint distribution function.

Definition

- *Let S be a subset of the plane with area A(S). A point is said to be **randomly selected** from S if for any subset R of S with area A(R), the probability that R contains the point is*
- This definition is essential in the field of **probability**. By the following examples, we will show how it can help to solve problems readily.

Example 8.7

- A man invites his fiancee to a fine hotel for a Sunday brunch. They decide to meet in the lobby of the hotel between 11:30 A.M. and 12 noon.
- If they arrive at random times during this period, what is the probability that they will meet within 10 minutes?

1. Let *X and Y be the minutes past 11:30 A.M. that the man and his fiancée arrive* at the lobby, respectively.

Let *S = {(x, y) :* *}, and*

R = {(x, y) ∈ S : *}.*

2. $$P(|X-Y|\leq 10) = \quad = \frac{area\,of\;R}{30\times 30} = \frac{area(R)}{900}$$

3. From fig 8.2, R is the shaded region

area(R)= = 500

$$P(|X-Y|\leq 10) = \frac{500}{900} = \frac{5}{9}$$

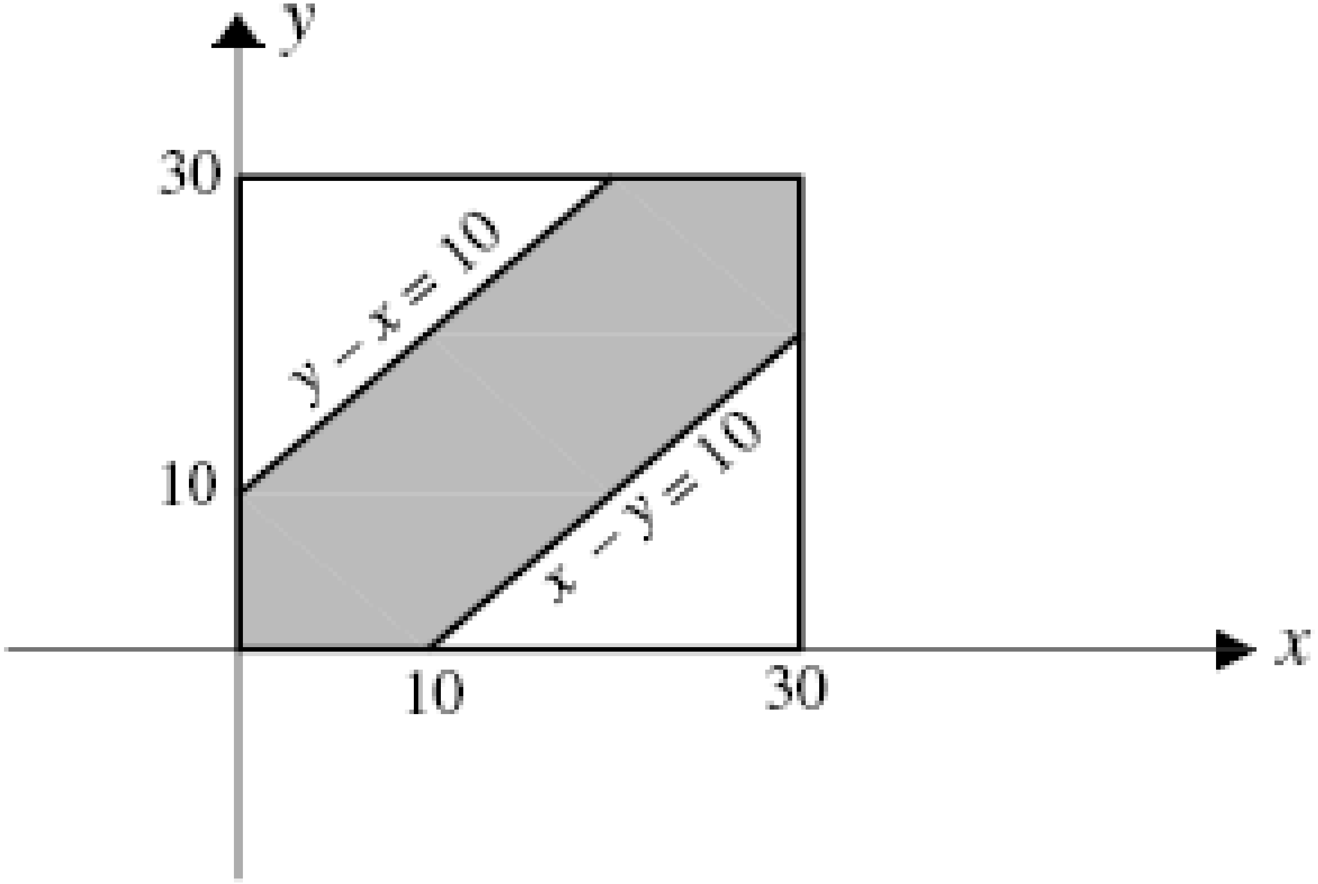

Figure 8.2 Geometric model of Example 8.7.

Theorem 8.2

- *Let f (x, y) be the joint probability density function of random variables X and Y.*
- *If h is a function of two variables from* $\mathbf{R}^2$ *to* $\mathbf{R}$*, then h(X,Y) is a random variable with the expected value given by*

$$E[h(X,Y)] =$$

provided that the integral is absolutely convergent.

Corollary

- *For random variables X and Y,*
  $\boldsymbol{E(X + Y) =}$

Example 8.9

- Let X and Y have joint probability density function

$$f(x,y) = \begin{cases} \frac{3}{2}(x^2 + y^2) & \text{if } 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Find $E(X^2 + Y^2)$.

$$E(X^2+Y^2)=$$

$$=\int_0^1\int_0^1 \frac{3}{2}(x^2+y^2)^2\,dxdy$$

$$=\frac{3}{2}\int_0^1\int_0^1 (x^4+2x^2y^2+y^4)\,dxdy=\frac{14}{15}.$$

8.2 Independent random variables

- Two random variables X and Y are called **independent** if, for arbitrary subsets A and B of real numbers, the events $\{X \in A\}$ and $\{Y \in B\}$ are independent, that is, if

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

- This implies that for any two real numbers a and b,

  $$P(X \leq a, Y \leq b) =$$

Theorem 8.3

- *Let X and Y be two random variables defined on the same sample space.*
- *If F is the joint probability distribution function of X and Y, then X and Y are independent if and only if for all real numbers t and u,*

$$F(t,u) =$$

Independence of Discrete Random Variables

- **Theorem 8.4**
  - *Let X and Y be two discrete random variables defined on the same sample space.*
  - *If $p(x, y)$ is the joint probability mass function of X and Y, then X and Y are independent if and only if for all real numbers x and y,*

$$p(x, y) = \qquad (8.11)$$

- Let X and Y be discrete *independent* random variables with sets of possible values A and B, respectively. Then (8.11) implies that for all $x \in A$ and $y \in B$,

  $$P(X = x \mid Y = y) =$$

  and

  $$P(Y = y \mid X = x) =$$

Example 8.10

- Suppose that 4% of the bicycle fenders, produced by a stamping machine from the strips of steel, need smoothing.
- What is the probability that, of the next 13 bicycle fenders stamped by this machine, two need smoothing and, of the next 20, three need smoothing?

1. Let *X be the number of bicycle fenders among the first 13 that need smoothing.*

   Let *Y be the number of those among the next 7 that need smoothing.*

2. *We want to* calculate *P(X = 2, Y = 1). Since X and Y are independent* binomial random variables with parameters *(13, 0.04) and (7, 0.04), respectively,*

$$P(X=2, Y=1) =$$

$$= \binom{13}{2}(0.04)^2(0.96)^{11}\binom{7}{1}(0.04)^1(0.96)^6 \approx 0.0175$$

Independence of Continuous Random Variables

- **Theorem 8.7**
  - *Let X and Y be jointly continuous random variables with joint probability density function f(x, y).*
  - *Then X and Y are   if and only if f (x, y) is the product of their marginal densities $f_X(x)$ and $f_Y(y)$.*
- *By differentiating F(x,y)=   , you can prove this theorem!*

Example 8.12

- Stores *A* and *B*, which belong to the same owner, are located in two different towns.
- If the probability density function of the weekly profit of each store, in thousands of dollars, is given by

$$f(x) = \begin{cases} x/4 & if\ 1 < x < 3 \\ 0 & otherwise \end{cases}$$

and the profit of one store is independent of the other, what is the probability that next week one store makes at least $500 more than the other store?

1. Let *X and Y denote next week's profits of A and B, respectively. The desired* probability is

Since *X and Y have the same* probability density function, by symmetry, this sum equals 2*P(X > Y + 1/2).*

2. $\because X$ and Y are independent

$\therefore$

$$\because f_X(x) = \begin{cases} x/4 & \text{if } 1 < x < 3 \\ 0 & \text{otherwise} \end{cases} \qquad f_y(y) = \begin{cases} y/4 & \text{if } 1 < y < 3 \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore f(x,y) = \begin{cases} & \\ 0 & \text{otherwise} \end{cases}$$

$$2P(X > Y + \frac{1}{2}) =$$

$$= \qquad = \frac{1}{8}\int_{3/2}^{3} [\frac{xy^2}{2}]_1^{x-1/2} dx$$

$$= \frac{1}{16}\int_{3/2}^{3} x[(x-\frac{1}{2})^2 - 1]dx$$

$$= \frac{1}{16}\int_{3/2}^{3} (x^3 - x^2 - \frac{3}{4}x)dx$$

$$= \frac{1}{16}[\frac{1}{4}x^4 - \frac{1}{3}x^3 - \frac{3}{8}x^2]_{3/2}^{3} = \frac{549}{1024} \approx 0.54$$

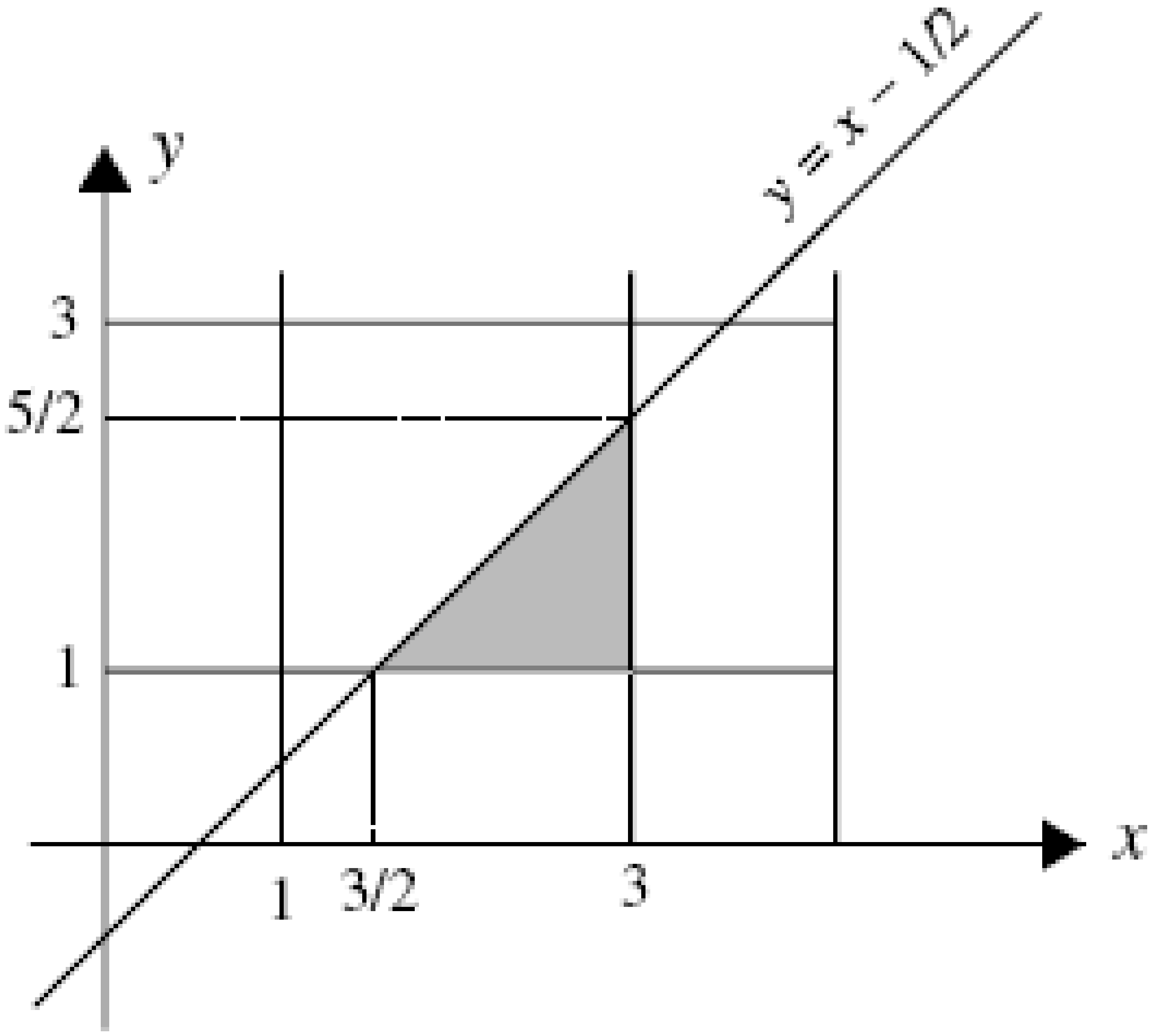

Figure 8.4 Figure of Example 8.12.

Example 8.15

- Prove that two random variables *X* and *Y* with the following joint probability density function are not independent.

$$f(x,y)=\begin{cases}8xy & 0\le x\le y\le 1\\ 0 & otherwise\end{cases}$$

1. $f_X(x)= \qquad =4x(1-x^2), \qquad 0\le x\le 1$

   $f_Y(y)= \qquad =4y^3, \qquad 0\le y\le 1$

2. $\because$

   $\therefore$ X and Y are dependent

8.3 Conditional distributions

- Conditional Distributions: Discrete Case
- Conditional Distributions: Continuous Case

Conditional Distributions: Discrete Case

- Let X be a discrete random variable with set of possible values A, and let Y be a discrete random variable with set of possible values B.
- Let be the joint probability mass function of X and Y, and let be the marginal probability mass functions of X and Y.
- If the value of Y is known, the **conditional probability mass function of X given that $Y = y$** which is denoted by $p_{X|Y}(x|y)$ is defined as follows:

$$p_{X|Y}(x \mid y) =$$

where $x \in A$, $y \in B$, and $p_Y(y) > 0$.

- Note that

$$\sum_{x \in A} p_{X|Y}(x \mid y) = \qquad = \frac{1}{p_Y(y)} p_Y(y) = 1$$

- Hence for any fixed $y \in B$, $p_{X|Y}(x|y)$ is itself a probability mass function with the set of possible values *A*.
- If *X* and *Y* are independent,

$$p_{X|Y}(x \mid y) =$$

$$= P(X = x) = p_X(x)$$

- Similar to $p_{X|Y}(x|y)$, the **conditional distribution function of *X*, given that *Y* = *y*** is defined as follows:

$$F_{X|Y}(x|y) =$$

$$= \qquad = \sum_{t \le x} p_{X|Y}(t|y)$$

Example 8.16

- Let the joint probability mass function of X and Y be given by

$$p(x,y)=\begin{cases}\frac{1}{15}(x+y) & if\ x=0,1,2,\ y=1,2\\ 0 & otherwise\end{cases}$$

Find $p_{X|Y}(x|y)$ and $P(X = 0 \mid Y = 2)$.

1. $p_y(y)=$
2. $p_{X|Y}(x\mid y)=\qquad\qquad, \quad x=0,1,2\ when\ y=1 or 2$

   $P(X=0\mid Y=2)=$

- Let X and Y be discrete random variables, and let the set of possible values of X be A.
- The **conditional expectation of the random variable X given that $Y = y$** is as follows:

$$E(X \mid Y = y) =$$

where $p_Y(y) > 0$.

- If *h* is a function from **R** to **R**, then for the discrete random variables *X* and *Y* with set of possible values *A* for *X*, the expected value of *h(X)* is obtained from

$$E[h(X)\,|\,Y=y]=$$

Example 8.18

- Calculate the expected number of aces in a randomly selected poker hand that is found to have exactly two jacks.

1. Let *X and Y be the number of aces and jacks in a random poker hand,* respectively.
2.

$$E(X \mid Y=2) = \sum_{x=0}^{3} x \frac{\dfrac{\binom{4}{2}\binom{4}{x}\binom{44}{3-x}}{\binom{52}{5}}}{\dfrac{\binom{4}{2}\binom{48}{3}}{\binom{52}{5}}} = \sum_{x=0}^{3} x \frac{\binom{4}{x}\binom{44}{3-x}}{\binom{48}{3}} \approx 0.25$$

Conditional Distributions: Continuous Case

- Let X and Y be two continuous random variables with the joint probability density function $f(x, y)$.
- When the value of Y is known, to find the probability of events concerning X, $f_{X|Y}(x|y)$, the **conditional probability density function of X given that $Y = y$** is used.
- $f_{X|Y}(x|y)$ is defined as follows:

$$f_{X|Y}(x, y) =$$

provided that $f_Y(y) > 0$.

- Note that

$$\int_{-\infty}^{\infty} f_{X|Y}(x \mid y)dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{f_Y(y)} dx = \qquad = \frac{1}{f_Y(y)} f_Y(y) = 1$$

showing that for a fixed *y*, $f_{X|Y}(x|y)$ is itself a probability density function.

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

$$f_{X|Y}(x \mid y) = \qquad = f_X(x)$$

- Also, as we expect, $F_{X|Y}(x|y)$, the **conditional probability distribution function of *X* given that *Y* = *y*** is defined as follows:

  $$F_{X|Y}(x\mid y) =$$

- Therefore,

  $$\frac{d}{dx}F_{X|Y}(x\mid y) =$$

Example 8.21

- Let X and Y be continuous random variables with joint probability density function

$$f(x,y)=\begin{cases}\frac{3}{2}(x^2+y^2) & \text{if } 0<x<1, 0<y<1\\ 0 & \text{otherwise}\end{cases}$$

Find $f_{X|Y}(x|y)$.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \qquad = \frac{3}{2}y^2 + \frac{1}{2}$$

$$f_{X|Y}(x|y) = \qquad = \frac{3(x^2+y^2)}{3y^2+1},\ 0 < x < 1 \ and \ 0 < y < 1$$

Example 8.22

- First, a point *Y* is selected at random from the interval *(0, 1)*. Then another point *X* is chosen at random from the interval *(0, Y)*.
- Find the probability density function of *X*.

1. Let *f (x, y) be the joint probability density function of X and Y*

$$f_X(x) =$$

2. $f_{X|Y}(x|y) = \dfrac{f(x,y)}{f_Y(y)} \Leftrightarrow f(x,y) =$

3. *Y is uniformly distributed over (0, 1),*

$$f_Y(y) = \begin{cases} & \mathit{if}\ 0 < y < 1 \\ 0 & \mathit{elsewhere} \end{cases}$$

4. given *Y = y, X is uniformly distributed over (0, y)*

$$f_{X|Y}(x|y) = \begin{cases} & \mathit{if}\ 0 < y < 1,\ \ 0 < x < y \\ 0 & \mathit{elsewhere} \end{cases}$$

5. $f_X(x) = \qquad\qquad = \ln 1 - \ln x = -\ln x$

$$\therefore\ f_X(x) = \begin{cases} -\ln x & \mathit{if}\ 0 < x < 1 \\ 0 & \mathit{elsewhere} \end{cases}$$

Example 8.23

- Let the conditional probability density function of *X*, given that *Y* = *y*, be

$$f_{X|Y}(x|y)=\frac{x+y}{1+y}e^{-x},\quad 0<x<\infty,\quad 0<y<\infty$$

Find $P(X<1 \mid Y=2)$.

1.

$$f_{X|Y}(x|2)= \qquad\qquad ,\quad 0<x<\infty$$

$$P(X<1|Y=2)=$$

$$=\frac{1}{3}[-xe^{-x}-e^{-x}]_0^1-\frac{2}{3}[e^{-x}]_0^1=1-\frac{4}{3}e^{-1}\approx 0.509$$

- Similar to the case where X and Y are discrete, for continuous random variables X and Y with joint probability density function $f(x, y)$, the **conditional expectation of X given that $Y = y$** is as follows:

$$E(X \mid Y = y) =$$

where $f_Y(y) > 0$.

- If h is a function from **R** to **R**, then, for continuous random variables X and Y, with joint probability density function $f(x, y)$,

$$E[h(X) \mid Y = y] =$$

- In particular, this implies that the conditional variance of X given that $Y =$ y is given by

$$\sigma^2_{X|Y=y} =$$

Example 8.24

- Let X and Y be continuous random variables with joint probability density function

$$f(x,y)=\begin{cases} e^{-y} & if\ \ y>0,\ 0<x<1 \\ 0 & elsewhere \end{cases}$$

Find $E(X \mid Y=2)$.

1. $$E[X \mid Y=2] = \qquad = \int_0^1 x \frac{f(x,2)}{f_Y(2)}\,dx = \int_0^1 x \frac{e^{-2}}{f_Y(2)}\,dx$$

$$\because\ f_Y(2) = \qquad = \int_0^1 e^{-2}\,dx = e^{-2}$$

$$\therefore E(X \mid Y=2) = \qquad = \frac{1}{2}$$

Example 8.25

- The lifetimes of batteries manufactured by a certain company are identically distributed with probability distribution and probability density functions F and f, respectively.
- In terms of F, f, and s, find the expected value of the lifetime of an s-hour-old battery.

1. Let *X be the lifetime of the s-hour-old battery. We want to calculate E(X | X > s). Let* $F_{X|X>s}(t) = P(X \le t \mid X > s)$

2. $E(X \mid X > s) =$

$$\because F_{X|X>s}(t) = \qquad = \begin{cases} 0 & \text{if } t \le s \\ \dfrac{P(s < X \le t)}{P(X > s)} & \text{if } t > s \end{cases}$$

$$= \begin{cases} 0 & \text{if } t \le s \\ \dfrac{F(t) - F(s)}{1 - F(s)} & \text{if } t > s \end{cases}$$

$$\therefore f_{X|X>s}(t) = \frac{d}{dt} F_{X|X>s}(t) = \begin{cases} 0 & \text{if } t \le s \\ & \text{if } t > s \end{cases}$$

$$\Rightarrow E(X \mid X > s) = \qquad = \frac{1}{1 - F(s)} \int_s^{\infty} t f(t) dt$$