Chap-3

Random Variables and Probability Distributions

A random variable is a function that associates a. real number with each element in the sample space.

- A random variable is called discrete random variable if its set of possible outcome is countable.
- otherwise it’s a continuous random variable i.e

  sample contain infinite number of possibilities

Discrete Probability Distributions

The set of ordered pairs $(x, f(x))$ is a **probability function, probability mass function, or probability distribution** of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

Example

Class Activity:

3.3 Let W be a random variable giving the number of heads minus the number of tails in three tosses of a coin. List the elements of the sample space S for the three tosses of the coin and to each sample point assign a value w of W.

3.8: A shipment of 8 similar microcomputers to a. retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Practice:

3.5 Determine the value c so that each of the following functions can serve as a probability distribution of the discrete random variable X:

(a) $f(x) = c(x^2 + 4)$, for a: $= 0, 1, 2, 3$;

(b) $f(x) = c\binom{2}{x}\binom{3}{3-x}$, for $x = 0, 1, 2$.

3.10 Find a formula for the probability distribution of the random variable X representing the outcome when a single die is rolled once.

3.11 A shipment of 7 television sets contains 2 defective sets. A hotel makes a random purchase of 3 of the sets. If x is the number of defective sets purchased by the hotel, find the probability distribution of X. Express the results graphically as a probability histogram.

Class Activity (Quiz):

3.22 Three cards are drawn in succession from a deck without replacement. Find the probability distribution for the number of spades.

3.26 From a box containing 4 black balls and 2 green balls, 3 balls are drawn in succession, each ball being replaced in the box before the next draw is made. Find the probability **distribution** for the number of green balls.

Solution:

3.22 Denote by X the number of spades int he three draws. Let S and N stand for a spade and not a spade, respectively. Then
$P(X=0) = P(NNN) = (39/52)(38/51)(37/50) = 703/1700$,
$P(X=1) = P(SNN) + P(NSN) + P(NNS) = 3(13/52)(39/51)(38/50) = 741/1700$,
$P(X=3) = P(SSS) = (13/52)(12/51)(11/50) = 11/850$, and
$P(X=2) = 1 - 703/1700 - 741/1700 - 11/850 = 117/850$.
The probability mass function for X is then

x	0	1	2	3
$f(x)$	703/1700	741/1700	117/850	11/850

Solution:

3.26 Denote by X the number of green balls in the three draws. Let G and B stand for the colors of green and black, respectively.

Simple Event	x	$P(X = x)$
BBB	0	$(2/3)^3 = 8/27$
GBB	1	$(1/3)(2/3)^2 = 4/27$
BGB	1	$(1/3)(2/3)^2 = 4/27$
BBG	1	$(1/3)(2/3)^2 = 4/27$
BGG	2	$(1/3)^2(2/3) = 2/27$
GBG	2	$(1/3)^2(2/3) = 2/27$
GGB	2	$(1/3)^2(2/3) = 2/27$
GGG	3	$(1/3)^3 = 1/27$

The probability mass function for X is then

x	0	1	2	3
$P(X = x)$	$8/27$	$4/9$	$2/9$	$1/27$

Continuous Probability Distribution

The function $f(x)$ is a **probability density function** for the continuous random variable A", defined over the set of real numbers R, if

1. /(a?) ≥ 0, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\, dx = = 1$.
3. $P(a < X < b) = \int_a^b f(x)\, dx$.

Practice:

3.6 The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20{,}000}{(x+100)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that a bottle of this medicine will have a shell life of

(a) at least 200 days;

(b) anywhere from 80 to 120 days.

Class Activity

3.7 The total number of hours, measured in units of 100 hours, that a family runs a vacuum cleaner over a period of one year is a continuous random variable X that has the density function

$$f(x) = \begin{cases} x, & 0 < x < 1, \\ 2 - x, & 1 \le x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the probability that over a period of one year, a family runs their vacuum cleaner

(a) less than 120 hours;

(b) between 50 and 100 hours.

Solution:

(a) $P(X < 1.2) = \int_0^1 x\, dx + \int_1^{1.2} (2 - x)\, dx = \frac{x^2}{2}\Big|_0^1 + \left(2x - \frac{x^2}{2}\right)\Big|_1^{1.2} = 0.68.$

(b) $P(0.5 < X < 1) = \int_{0.5}^1 x\, dx = \frac{x^2}{2}\Big|_{0.5}^1 = 0.375.$

Practice:

3.13 The probability distribution of A, the number of imperfections per 10 meters of a synthetic fabric in continuous rolls of uniform width, is given by

x	0	1	2	3	4
$f(x)$	0.41	0.37	0.16	0.05	0.01

Construct the cumulative distribution function of X.

3.14 The waiting time, in hours, between successive speeders spotted by a radar unit is a continuous random variable with **cumulative** distribution function

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-8x}, & x \geq 0. \end{cases}$$

Find the probability of waiting less than 12 **minutes** between successive speeders

(a) using the cumulative distribution function of X;

(b) using the probability density function of X.

Practice:

3.17 A continuous random variable X that can assume values between $x = 1$ and $x = 3$ has a density function given by $f(x) = 1/2$.

(a) Show that the area under the curve is equal to **1**.

(b) Find $P(2 < X < 2.5)$.

(c) Find $P(X \leq 1.6)$.

3.18 A continuous random variable X that can assume values between $x = 2$ and $x = 5$ has a density function given by $f(x) = 2(1 + x)/27$. Find

(a) $P(X < 4)$;

(b) $P(3 \leq X < 4)$.

Examples:

3.11 A shipment of 7 television sets contains 2 defective sets. A hotel makes a random purchase of 3 of the sets. If x is the number of defective sets purchased by the hotel, find the probability distribution of X. Express the results graphically as a probability histogram.

3.15 Find the cumulative distribution function of the random variable X representing the number of defectives in Exercise 3.11. Then using $F(x)$, find

(a) $P(X = 1)$;

(b) $P(0 < X \leq 2)$.

Solution:

3.11

$$f(x) = \frac{\binom{2}{x}\binom{5}{3-x}}{\binom{7}{3}}, \qquad x = 0, 1, 2.$$

x	0	1	2
$f(x)$	2/7	4/7	1/7

3.15 The c.d.f. of X is

$$F(x) = \begin{cases} 0, & \text{for } x < 0, \\ 2/7, & \text{for } 0 \leq x < 1, \\ 6/7, & \text{for } 1 \leq x < 2, \\ 1, & \text{for } x \geq 2. \end{cases}$$

(a) $P(X = 1) = P(X \leq 1) - P(X \leq 0) = 6/7 - 2/7 = 4/7$;

(b) $P(0 < X \leq 2) = P(X \leq 2) - P(X \leq 0) = 1 - 2/7 = 5/7$.

Joint Probability Distributions

The function $f(x, y)$ is a **joint probability distribution** or **probability mass function** of the discrete random variables X and Y if

1. $f(x, y) \geq 0$ for all (x, y),
2. $\sum_x \sum_y f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$.

For any region A in the xy plane, $P[(X, Y) \in A] = \sum\sum_A f(x, y)$.

Example:

Two ballpoint pens are selected at random from a box that contains 3 blue pens, 2 red pens, and 3 green pens. If X is the number of blue pens selected and Y is the number of red pens selected, find

(a) the joint probability function $f(x, y)$,

(b) $P[(X, Y) \in A]$, where A is the region $\{(x, y) | x + y \leq 1\}$.

possible pairs of values (x, y) are $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$, $(0, 2)$, and $(2, 0)$

$\binom{8}{2} = 28.$ $\qquad \binom{2}{1}\binom{3}{1} = 6.$ $\qquad f(0, 1) = 6/28$

Solution:

a) $$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}},$$

$x = 0, 1, 2;\ y = 0, 1, 2;$ and $0 \le x + y \le 2.$

$f(x,y)$		x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

b) $$P[(X,Y) \in A] = P(X + Y \le 1) = f(0,0) + f(0,1) + f(1,0)$$
$$= \frac{3}{28} + \frac{3}{14} + \frac{9}{28} = \frac{9}{14}.$$

The function $f(x, y)$ is a **joint density function** of the continuous random variables X and Y if

1. $f(x, y) \geq 0$, for all (x, y),
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$,
3. $P[(X, Y) \in A] = \int \int_A f(x, y)\, dx\, dy$, for any region A in the xy plane.

Example:

A privately owned business operates both a drive-in facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-in and the walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y), & 0 \leq x \leq 1, 0 \leq y \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify condition 2 of Definition 3.9.

(b) Find $P[(X, Y) \in A]$, where $A = \{(x, y) \mid 0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$

Solution:

a) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\ dx\ dy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)\ dx\ dy$$

b) $$P[(X,Y)\in A] = P\left(0<X<\frac{1}{2}, \frac{1}{4}<Y<\frac{1}{2}\right)$$
$$= \int_{1/4}^{1/2}\int_0^{1/2} \frac{2}{5}(2x+3y)\ dx\ dy \qquad = \frac{13}{160}.$$

Marginal Distribution:

The **marginal distributions** of X alone and of Y alone are

$$g(x) = \sum_y f(x, y) \quad \text{and} \quad h(y) = \sum_x f(x, y)$$

for the discrete case, and

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

for the continuous case.

Example:

Show that the column and row totals of Table 3.1 give the marginal distribution of X alone and of Y alone.

	$f(x, y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Solution:

$$g(0) = f(0,0) + f(0,1) + f(0,2) = \frac{3}{28} + \frac{3}{14} + \frac{1}{28} = \frac{5}{14},$$

$$g(1) = f(1,0) + f(1,1) + f(1,2) = \frac{9}{28} + \frac{3}{14} + 0 = \frac{15}{28},$$

$$g(2) = f(2,0) + f(2,1) + f(2,2) = \frac{3}{28} + 0 + 0 = \frac{3}{28},$$

x	0	1	2
$g(x)$	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$

y	0	1	2
$h(y)$	$\frac{15}{28}$	$\frac{3}{7}$	$\frac{1}{28}$

Example:

Find $g(x)$ and $h(y)$ for the joint density function of Example 3.15.

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1, 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Solution

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{2}{5}(2x+3y)\,dy$$

$$= \left(\frac{4xy}{5} + \frac{6y^2}{10}\right)\Bigg|_{y=0}^{y=1} = \frac{4x+3}{5},$$

for $0 \le x \le 1$, and $g(x) = 0$ elsewhere.

Similarly,

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{2}{5}(2x+3y)\,dx = \frac{2(1+3y)}{5},$$

for $0 \le y \le 1$, and $h(y) = 0$ elsewhere.

The Conditional Probability Distribution

Let X and Y be two random variables, discrete or continuous. The **conditional distribution** of the random variable Y given that $X = x$ is

$$f(y|x) = \frac{f(x, y)}{g(x)}, \text{ provided } g(x) > 0.$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x|y) = \frac{f(x, y)}{h(y)}, \text{ provided } h(y) > 0.$$

$$P(a < X < b \mid Y = y) = \sum_{a<x<b} f(x|y),$$

$$P(a < X < b \mid Y = y) = \int_a^b f(x|y)\, dx.$$

Example:

The joint density for the random variables (X, Y), where X is the unit temperature change and Y is the proportion of spectrum shift that a certain atomic particle produces, is

$$f(x, y) = \begin{cases} 10xy^2, & 0 < x < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal densities $g(x)$, $h(y)$, and the conditional density $f(y|x)$.

(b) Find the probability that the spectrum shifts more than half of the total observations, given that the temperature is increased by 0.25 unit.

(a) By definition,

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_{x}^{1} 10xy^2\, dy$$

$$= \frac{10}{3} xy^3 \Big|_{y=x}^{y=1} = \frac{10}{3} x(1 - x^3),\ 0 < x < 1,$$

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_{0}^{y} 10xy^2\, dx = 5x^2y^2 \Big|_{x=0}^{x=y} = 5y^4,\ 0 < y < 1.$$

$$f(y|x) = \frac{f(x, y)}{g(x)} = \frac{10xy^2}{\frac{10}{3}x(1 - x^3)} = \frac{3y^2}{1 - x^3},\ 0 < x < y < 1.$$

(b) Therefore,

$$P\left(Y > \frac{1}{2} \,\middle|\, X = 0.25\right) = \int_{1/2}^{1} f(y \mid x = 0.25)\, dy$$

$$= \int_{1/2}^{1} \frac{3y^2}{1 - 0.25^3}\, dy = \frac{8}{9}.$$

Class Activity (Quiz)

Given the joint density function

$$f(x, y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1, \\ 0, & \text{elsewhere,} \end{cases}$$

find $g(x)$, $h(y)$, $f(x|y)$, and evaluate $P(\frac{1}{4} < X < \frac{1}{2} \mid Y = \frac{1}{3})$.

Solution: By definition of the marginal density. for $0 < x < 2$,

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy = \int_0^1 \frac{x(1+3y^2)}{4} dy$$

$$= \left(\frac{xy}{4} + \frac{xy^3}{4}\right)\bigg|_{y=0}^{y=1} = \frac{x}{2},$$

and for $0 < y < 1$,

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^2 \frac{x(1+3y^2)}{4} dx$$

$$= \left(\frac{x^2}{8} + \frac{3x^2y^2}{8}\right)\bigg|_{x=0}^{x=2} = \frac{1+3y^2}{2}.$$

using the conditional density definition, for $0 < x < 2$,

$$f(x|y) = \frac{f(x, y)}{h(y)} = \frac{x(1+3y^2)/4}{(1+3y^2)/2} = \frac{x}{2},$$

$$P\left(\frac{1}{4} < X < \frac{1}{2} \,\middle|\, Y = \frac{1}{3}\right) = \int_{1/4}^{1/2} \frac{x}{2} dx = \frac{3}{64}.$$

Practice:

3.38 If the joint probability distribution of X and Y is given by

$$f(x, y) = \frac{x+y}{30}, \quad \text{for } x = 0, 1, 2, 3; \; y = 0, 1, 2,$$

find

(a) $P(X \leq 2, Y = 1)$;
(b) $P(X > 2, Y \leq 1)$;
(c) $P(X > Y)$;
(d) $P(X + Y = 4)$.

3.46 Referring to Exercise 3.38, find

(a) the marginal distribution of X;
(b) the marginal distribution of Y.

Practice:

3.42 Let X and Y denote the lengths of life, in years, of two components in an electronic system. If the joint density function of these variables is

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

find $P(0 < X < 1 \mid Y = 2)$

3.43 Let X denote the reaction time, in seconds, to a certain stimulus and Y denote the temperature (°F) at which a certain reaction starts to take place. Suppose that two random variables X and Y have the joint density

$$f(x,y) = \begin{cases} 4xy, & 0 < x < 1,\ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find

(a) $P(0 \le X \le \frac{1}{2}$ and $\frac{1}{4} \le Y \le \frac{1}{2})$;

(b) $P(X < Y)$.

Practice:

3.40 A fast-food restaurant operates both a drive-through facility and a walk-in facility. On a randomly selected day, let X and Y, respectively, be the proportions of the time that the drive-through and walk-in facilities are in use, and suppose that the joint density function of these random variables is

$$f(x,y) = \begin{cases} \frac{2}{3}(x+2y), & 0 \le x \le 1,\ 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find the marginal density of X.

(b) Find the marginal density of Y.

(c) Find the probability that the drive-through facility is busy less than one-half of the time.