Probability & Random Processes (Math3211)

Chapter 3: Probability Distributions

Tatek Getachew (PhD)

November 5, 2025

Outline

Random Variable and Probability Distribution

Definition: Random variable

is a function that assign areal number for each outcome of a random experiment.
:- a real numerical valued function defined over a sample space.

eg: Toss a coin twice. Let H denote a head and T denote a tail that will show up when the experiment is performed

$$S=\{HH, HT, TH, TT\}$$

R.V x be the number of heads that will show up in tossing the coin twice
$x=\{0, 1, 2\}$

$$x(HH)=2$$
$$x(HT)=1$$
$$x(TH)=1$$
$$x(TT)=0$$

Defn The probability distribution function of a random variable x is a function that assign probabilities for each value of the random variable

Eg Suppose a fair coin is tossed tree times. Let the r.v x represent the number of heads that will show up when the experiment is performed. Find the probability distribution of x.

Soln The sample space

S={HHH, HHT, HTH, HTT, THH, TTH, THT, TTT}

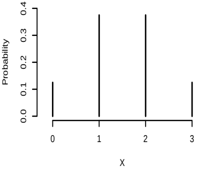
$\Longrightarrow$ The possible values of x are 0, 1, 2, 3

$$\begin{aligned} P(x=0) &= P(TTT) = 1/8 \\ P(x=1) &= P(HTT, THT, TTH) = 3/8 \\ P(x=2) &= P(HHT, THH, HTH) = 3/8 \\ P(x=3) &= P(HHH) = 1/8 \end{aligned}$$

$\therefore$ the probability distribution of x is

x	P(X=x)
0	1/8
1	3/8
2	3/8
3	1/8

$$P(X=x) = \begin{cases} \frac{1}{8} & \text{if } x = 0, 3 \\ \frac{3}{8} & \text{if } x = 1, 2 \end{cases}$$



A random variable can be

- Discrete r.v:- if it takes finite or countably infinite values
  Eg -the number of car accidents occurring at a certain place
  - the number of children in a family
  - the number of spots on rolling a die
- Continuous r.v:- if it takes infinite values in a specified interval
  Eg height, weight, time etc

Based on the random variable we have probability distribution can be classified in to two

1. Discrete Probability Distribution:- is a probability distribution of a discrete random variable

Defn

Let x be a discrete r.v with possible values $x_1, \; x_2, \; \cdots x_n$. Then the function P() with domain **R** and counter domain [0, 1] denoted by

$$P(X = x) = \begin{cases} P(X = x) & \text{if x= } \; x_1, \; x_2, \; \cdots x_n \\ 0 & \text{if it is not} \end{cases}$$

is called a discrete probability, function iff

i) P(x)$\geq$ 0 all x

ii) $\sum P(x_i)$=1

Eg. consider the above example

$$P(X=x)=\begin{cases} 1/8 & \text{if} x=0,3 \\ 3/8 & \text{if} x=1,2 \\ 0 & \text{otherwise} \end{cases}$$

we can show that P is a discrete probability function

- P(x) $\geq$ 0 for all x
- $\sum P(x_i) = P(x_1) + P(x_2) + P(x_3) + P(x_4)$=1/8+3/8+3/8+1/8=1

P is discrete proba. function

Defn: Continuous Probability Distribution

The probability distribution of a continuous r.v called continuous probability distribution. The function that describe the probability distribution of a continuous r.v is probability density function (pdf) of x and has the following properties

i f(x)$\geq$ 0 all x

ii $\int_{-\infty}^{\infty}$ f(x)dx=1

iii $P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b) = \int_a^b f(x)dx$

Eg.

$$f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Show that f(x) is Pdf for x?

- f(x)$\geq$ x
- $\int_{-\infty}^{\infty} f(x)dx$=1, $\implies \int_{-\infty}^{0} f(x)dx + \int_0^1 f(x)dx + \int_1^{\infty} f(x)dx$
  $= 0+ \int_0^1 2xdx+0=x^2|_0^1=1$

f(x) is a Pdf

Binomial Probability Distribution

Binomial Experiment

For a random experiment to be classified as a binomial experiment, it must satisfy the following properties

- The experiment consists of a sequence of n trials
- For each trial there are only two possible outcomes or the outcomes can be partitioned in to two. We refer one as a success and the other as a failure.
- The probability of success denoted by p is the same for each trial
- the trials are independent

Eg. Tossing a coin 20 times

Rolling a die 5 times, S={1, 2, 3, 4, 5, 6}, S={even, odd}

In a binomial experiment our interest is the number of success occurring in n trial. If we let x denote the number of success occurring in the n trials. We see that x can assume values 0, 1, 2,3, $\cdots$,n. Since the values are finite x is a discrete r.v.

The prob distribution associated with this r.v x is called the Binomial Probability distribution

Defn

If P is the probability of success in a single trial, the probability of getting x success in n trials of the experiment given by

$$P(X = x) = \binom{n}{x} \cdot p^x q^{n-x}, \quad \text{x} = 0,\ 1, \cdots,\ \text{n}$$

Eg. If a fair coin is tossed 5 times, what is the probability that

a exactly two head will show up?

b at least two heads will show up?

c No head will show up

Soln Let x be the number of heads will show up

Given P=1/2, and n=5

$$a) \quad P(x = 2) = \binom{5}{2} \cdot (1/2)^2 (1/2)^{5-2} = 5/16$$

$$b) \quad P(x \geq 2) = 1 - P(x < 2) = 1 - (P(x = 0) - P(x = 1))$$

$$c)P(x=0)=\binom{5}{0}\cdot(1/2)^0(1/2)^5=1/32$$

Defn

A r.v x is said to have a binomial probability distribution with parameters n and p, denoted by x$\sim$ Bin(n,p) if it is prob. function is given by

$$P(X=x)=\binom{n}{x}\cdot p^x q^{n-x}$$

Eg. A rail way company claimed that 95% of its trains arrive a station on time. If three of these trains are scheduled for a day, what is the probability that

a all arrive on time?

b one of the three is late?

Given p=0.95, and n=3, x$\sim$ Bin(n,p)

Theorem If r.v x$\sim$ Bin(n,p) then E[x]=np, Var(x)=np(1-p)

Eg. Find the mean and the variance of x for the above example?

Ex. In a large community 40% of the population are video owners. If a random sample of 10 persons is taken. What is the probability that at least 2 will be video owners?

The Poisson Probability Distribution

A probability distribution that is often useful when dealing with the occurrence of an event over specified interval of time, space, distance and so on is the Poisson prob. distribution

Eg. the number of telephone call per hour
the number of bacteria in a given culture
the number of typing error on a page

Defn

A r.v x is defined to have a Poisson probability distn with parameter λ, if its probability function is given by

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, \qquad x=0,\ 1,\ 2, \cdots$$

where λ is the average number of outcomes per unit time, space, etc
Eg. The number of telephone calls arriving a certain switch board had a Poisson prob. distn with 5 calls per minute on the average. What is the probab that

- i two calls will arrive on the next minute
- ii at least 4 calls will arrive on the next 3 minute

Theorem

If a r.v x$\sim$Poiss(λ) E[x]=λ, Var(x)=λ

Poisson Approximation to The Binomial

In a binomial distribution if p is very close to zero and n is very large, the binomial probabilities can be approximated by the Poisson distribution with λt= np

Eg. If approximately 2% of the people in a room of 200 people are left-handed, find the probability that exactly 5 people there are left-handed.

Soln

Given n=200 and p=0.02, $\implies \lambda = np = 200 \times 0.02 = 4$

What we require is P(X=5)

Using Poison approximation to the Binomial

$$P(X=5) = \frac{e^{-\lambda}.\lambda^x}{x!} = \frac{(2.718)^4 \cdot 4^5}{5!} = 0.1563$$

Common Continuous Probability Distributions

Normal Distribution (Gaussian PDF)

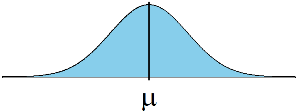
A continuous r.v x with mean μ and variance σ^2 said to have normal distribution with parameters μ and σ^2 if its pdf is given by

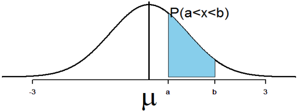
$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad -\infty < \mu < \infty, \quad \sigma > 0$$

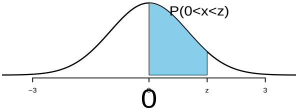
Properties of a normal distribution

1. The area under the curve is 1. $\implies \int_{-\infty}^{\infty} f(x)dx = 1$
2. The curve is symmetrical about the mean μ $(\bar{x} = \tilde{x} = \hat{x})$, the curve is Mesokurtic
3. In general P$(a < x < b)$is the area under the curve between a and b, i.e
4. Define Z$=\frac{x-\mu}{\sigma}$ where x$\sim$N$(\mu,\ \sigma^2)$, then z$\sim$N(0,1) and it is called standard normal variable and its pdf is given by

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}$$





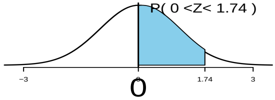


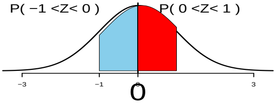
Reading Standard Normal Table

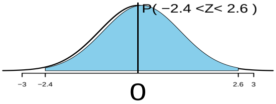
- The table gives P(0$<$ Z $<$z)=$\int_0^1 f(z)dz$ that is
- The integral part with the first decimal part of 2 is presented in the first column and the rest decimal part is presented in the first row of the table. The point joining the row and the the column gives P(0$<$ Z $<$z)

Example

1. $P(0 < Z < 1.74)$=0.4591
2. $P(-1.0 < Z < 0)$=$P(0 < Z < 1)$=0.3413
3. $P(-2.4 < Z < 2.6)$=$P(-2.4 < Z < 0)$+$P(0 < Z < 2.6)$=$P(0 < Z < 2.4)$+$P(0 < Z < 2.6)$=0.4918+0.4953=0.9871







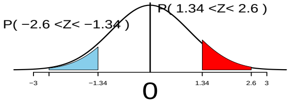
4 $P(-2.6 < Z < -1.34)$=$P(1.34 < Z < 2.6)$=$P(0 < Z < 2.6)$-$P(0 < Z < 1.34)$=0.4953-0.4099=

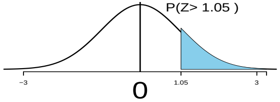
5 $P(Z > 1.05)$=0.5-$P(0 < Z < 1.05)$=0.5-0.3531=0.1469

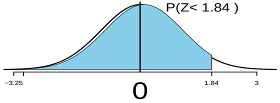
6 $P(Z < 1.84)$=0.5+$P(0 < Z < 1.84)$=0.9671

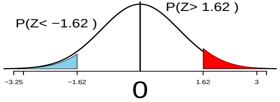
7 $P(Z < -1.62)$=$P(Z > 1.62)$=0.5-$P(0 < Z < 1.62)$=0.0529

8 $P(Z > -2.88)$









Eg. The class with 475 students the $\bar{x}$ score on the final is 80 and the standard deviation 12. Assuming that the score for normal distribution. What is the probability that the score of a randomly selected 5 students with scores with

a between 69 and 75

b above 90

c below 82

How many students recorded scores between

d 69 and 75

e above 82

f below 82

Soln let x be the score of a randomly selected student then x$\sim$N($\mu = 80, \sigma^2 = 144$)
$\implies$ $Z = \frac{x-\mu}{\sigma}$ $\sim$N(0,1)

a P($69 < Z < 75$)=P($\frac{69-80}{12} < Z < \frac{75-80}{12}$)=P(-0.92< Z <-0.42)=P(0.42< Z <0.92)

b P($x > 90$)=P($Z > \frac{90-80}{12}$)=P(Z<0.83)=0.5-P(0< Z <0.83)

c P(x<82)=P(Z< $\frac{82-80}{12}$)=P(Z<0.17)=0.5+P(0< Z <0.17)

d n.P(69< x <75)=475$\times$0.1584

e n.P(x<90)=475$\times$0.2033=

f n.P(x<82)=475$\times$0.5675

Ex. The weight of students in a class has a normal distribution with mean 60kg and s.d of 5kg. Find the value below which 20% of the students have their weight and the value above which which 15% of the students have their weight.

Soln Let x be the weight of a randomly selected student

$x_1 =$ the value below which 20% of the students have their weight

$x_2 =$ the value above which 15% of the students have their weight

Given x$\sim$N$(\mu = 60,\ \sigma^2 = 25) \implies z = \frac{x-60}{\sigma}$ $\sim$N(0,1)

P$(x < x_1)$=0.2=P$(z < \frac{x_1-60}{5})$=0.2

P$(x < x_1)$=P$(Z < z_1)$=0.2

z_1 is negative number

Ex. The amount of time it takes for a person to learn how to operate a certain machine is a random variable having normal distribution. If past records show that 15.8% of such persons spent less than 4.4 hours and 6.68% of them spent more than 7.4 hours. What was the average time and the s.d taken to learn how to operate the machine?

Cumulative Distribution Function (CDF)

Defn

Let X be a random variable, discrete or continuous. We define F to be the cumulative distribution function of the rv.X (abbreviated as c.d.f) where F(X)=P(X<x).

- Properties of cumulative distribution functions are summarized as follows:
  1) $F_X(-\infty) = 0, F_X(\infty) = 1,$
  2) $0 \leq F_X(x) \leq 1,$
  3) For $x_1 < x_2, F_X(x_1) \leq F_X(x_2),$
  4) For $x_1 < x_2, Pr(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1).$
- EXAMPLE 3.1: Which of the following mathematical functions could be the CDF of some random variable?
  a) $F_X(x) = \frac{1}{2} + \frac{1}{\pi} tan^{-1}(x),$
  b) $F_X(x) = [1 - e^{-x}]u(x),$ ($u(x)$ is the unit step function),
  c) $F_X(x) = e^{-x^2},$
  d) $F_X(x) = x^2 u(x).$

- we need to check that the function starts at 0 when $x = -\infty$, ends at 1 when $x = \infty$, and is monotonic increasing in between.
- The first two functions satisfy these properties and thus are valid CDFs, while the last two do not.

Theorem

If X is a discrete rv, F(X)=$\sum_j P(X = x_j)$, where the sum is taken over all j satisfying $x_j \leq x$.
If X takes on only a finite number of values $x_1,\ x_2, \cdots,\ x_n$ then the c.d.f is given by

$$F(x) = \begin{cases} 0, & -\infty < x < x_1 \\ P(x_1) & x_1 \leq x < x_2 \\ P(x_1) + P(x_2) & x_2 \leq x < x_3 \\ & \\ P(x_1) + P(x_2) + \cdots + P(x_n), & x_n \leq x < \infty \end{cases}$$

Eg.

a) Find the distribution function for the r.v.X representing the number of heads in the two tosses of a fair coin.

b) Obtain its graph

Defn

The cumulative distribution function F(x) of a continuous rv.X with density function f(x) is given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

As an immediate consequence of the above definition, one can write the two results:

i P($a \leq X \leq b$)=P($a < X < b$)=F(b)−F(a)

ii $f(x) = \frac{dF(x)}{dx}$

the PDF of a random variable is the derivative of its CDF, Conversely, the CDF of a random variable can be expressed as the integral of its PDF.

- EXAMPLE 3.3: Suppose a random variable has a CDF given by $F_X(x) = (1 - e^{-x})u(x)$. Find the following quantities:
  a) $Pr(X > 5)$,
  b) $Pr(X < 5)$,
  c) $Pr(3 < X < 7)$,
  d) $Pr(X > 5|X < 7)$

Soln

a) $Pr(X > 5) = 1 - Pr(X \leq 5) = 1 - F_X(5) = e^{-5}$
b) $F_X(5) = Pr(X < 5 \cup X = 5) = Pr(X < 5) + Pr(X = 5)$.
   In this case, since X is a continuous random variable, $Pr(X = 5) = 0$
   $Pr(X < 5) = FX(5) = 1 - exp(-5)$.
c) $F_X(7) - FX(3) = Pr(3 < X7) = e^{-3}e^{-7}$.
d) we invoke the definition of conditional probability to write the required quantity in terms of the CDF of X:

$$Pr(X > 5|X < 7) = \frac{Pr(\{X > 5\} \cap \{X < 7\})}{Pr(X < 7)} = \frac{Pr(5 < X < 7)}{Pr(X < 7)}$$

$$= \frac{F_X(7)F_X(5)}{F_X(7)} = \frac{e^{-5} - e^{-7}}{1 - e^{-7}}$$

Some properties of PDFs are

1) $f_X(x) \geq 0$;
2) $f_X(x) = \frac{dF_X(x)}{dx}$;
3) $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$;
4) $\int_{-\infty}^{\infty} f_X(x)dx = 1$;
5) $Pr(a < X \leq b) = \int_a^b f_X(x)$.

EXAMPLE 3.4: Which of the following are valid probability density functions?

a) $f_X(x) = e^{-x}u(x)$;
b) $f_X(x) = e^{-|x|}$;
c)
$$f_X(x) = \begin{cases} \frac{3}{4}(x^2 - 1), & |x| < 2 \\ 0 & Otherwise; \end{cases}$$
d)
$$f_X(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0 & Otherwise; \end{cases}$$
e) $f_X(x) = 2xe^{-x^2}u(x)$.

- The function in part (c) takes on negative values while the function in part (b) is not properly normalized, and hence these are not valid PDFs. The other three functions are valid PDFs.
- EXAMPLE 3.5: A random variable has a CDF given by $F_X(x) = (1 - e^{-\lambda x})u(x)$. Its PDF is then given by

$$f_X(x) = \frac{dF_X(x)}{dx} = \lambda e^{-\lambda x} u(x).$$

- Likewise, if a random variable has a PDF given by $f_X(x) = 2xe^{-x^2} u(x)$, then its CDF is given by

$$\begin{aligned} F_X(x) = \int_{-\infty}^{x} f_X(y)dy &= \int_{-\infty}^{x} 2ye^{-y^2} u(y)dy \\ &= \int_{-\infty}^{x} 2ye^{-y^2} dyu(x) = (1 - e^{-x^2})u(x). \end{aligned}$$

- Whether the Gaussian CDF is to be evaluated by using a table or a program, the required CDF must be converted into one of a few commonly used standard forms.

DEFn: A Gaussian random variable

is one whose probability density function can be written in the general form

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

- A few of these common forms are
- error function integral, $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x exp\left(-t^2\right) dt$,
- complementary error function integral, $erfc(x) = 1 - erf(x) = \int_x^\infty exp\left(-t^2\right) dt$;
- Φ-function, $\Phi = Pr(X < x) = \frac{1}{\sqrt{2pi}} \int_{-\infty}^x exp\left(\frac{-t^2}{2}\right) dt$
- Q-function, $Q(x) = Pr(X > x) = \frac{1}{\sqrt{2pi}} \int_x^\infty exp\left(\frac{-t^2}{2}\right) dt$.

- it is a more straightforward thing to express the Gaussian CDF in terms of a Φ-function or a Q-function.
- Also, the Q-function seems to be enjoying the most common usage in the engineering literature in recent years.
- Perhaps the advantage is clearer if we note that the $\Phi-$function is simply the CDF of a standard normal random variable.
- For general Gaussian random variables that are not in the normalized form, the CDF can be expressed in terms of a $\Phi-$function using a simple transformation. Starting with the Gaussian CDF in Equation 3.13, make the transformation $t = (y-m)/\sigma$, resulting in

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2pi}} exp\left(\frac{-(y-m)^2}{2\sigma^2}\right) dy = \int_{-\infty}^{\frac{x-m}{\sigma}} \frac{1}{\sqrt{2pi}} exp\left(\frac{-t^2}{2}\right) dt$$
$$= \Phi\left(\frac{x-m}{\sigma}\right)$$

The Q-function is more natural for evaluating probabilities

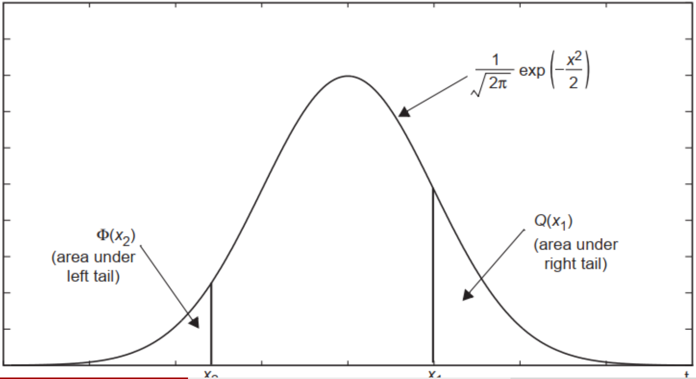
$$Pr(X > x) = \int_{\frac{x-m}{\sigma}}^{\infty} \frac{1}{\sqrt{2pi}} exp\left(\frac{-t^2}{2}\right) dt = Q\left(\frac{x-m}{\sigma}\right)$$

it is apparent that the relationship between the Φ-function and the Q-function is

$$Q(x) = 1 - \Phi(x).$$

This and other symmetry relationships can be visualized using the graphical definitions of the 8-function (phi function) and the Q-function

$$F_X(x) = 1 - Q\left(\frac{x-m}{\sigma}\right)$$



- EXAMPLE 3.7: A random variable has a PDF given by

$$f_X(x) = \frac{1}{\sqrt{8\pi}} exp\left(-\frac{(x+3)^2}{8}\right).$$

- Find each of the following probabilities and express the answers in terms of Q-functions.
  a) $Pr(X \leq 0)$,
  b) $Pr(X > 4)$,
  c) $Pr(|X+3| < 2)$,
  d) $Pr(|X-2| > 1)$.
- For the given Gaussian pdf, m = -3 and = 2. For part (a), $Pr(X \leq 0) = \Phi((0-(-3))/2) = \Phi(1.5)$.
- This can be rewritten in terms of a Q-function as $Pr(X \leq 0) = 1 - Q(1.5)$.
- (b) is easier to express directly in terms of a Q-function. $Pr(X > 4) = Q((4-(-3))/2) = Q(3.5)$.
- (c), the probability of the random variable X falling in an interval is required. This event can be rewritten as $Pr(|X+3| < 2) = Pr(-5 < X < -1) = Pr(X > -5) - Pr(X > -1)$ $= Q(-1) - Q(1) = 1 - 2Q(1)$

Other Important Random Variables

- Uniform Random Variable $f_X(x) = \frac{1}{b-a}, \quad a < x < b;$
- Exponential Random Variable $f_X(x) = \frac{1}{b} exp\left(\frac{-x}{b}\right) u(x),$
- Laplace Random Variable
- Gamma Random Variable
- Erlang Random Variable
- Rayleigh Random Variable
- Rician Random Variable
- Cauchy Random Variable

Pairs of Random Variables

- A pair of random variables can be used to characterize this relationship; one for the input and another for the output.
- Another class of examples involving random variables is one involving spatial coordinates in two dimensions.
- Examples of two random quantities that may or may not be related to one another, for example,
  - the height and weight of a student,
  - the grade point average and GRE scores of a student,
  - the temperature and relative humidity at a certain place and time.
- A two-dimensional random variable in the sample space S to ordered pairs x, y viewed as a combination of two simpler sample spaces.
- The range of student heights could fall sample space S_1, while the range of student weights could fall within the space S_2. The overall sample space of the experiment could then be viewed as $S = S_1 \times S_2$.
- For any outcome $s \in S$ of this experiment, the pair of random variables (X, Y) is merely a mapping of the outcome s to a pair of numerical values (x(s), y(s))

Joint Cumulative Distribution Functions

Joint cumulative distribution function

of a pair of random variables, {X, Y}, is $F_{X,Y}(x, y) = Pr(X \leq x, Y \leq y)$. That is, the joint CDF is the joint probability of the two events $\{X \leq x\}$ and $\{Y \leq y\}$.

- Hence, $F_{X,Y}(x, y)$ evaluated at either x = -∞ or y = -∞ (or both) must be zero and
- $F_{X,Y}(\infty, \infty)$ must be one.
- for $x_1 \leq x_2$ and $y_1 \leq y_2$, $\{X \leq x_1\} \cap \{Y \leq y_1\}$ is a subset of $\{X \leq x_2\} \cap \{Y \leq y_2\}$ so that $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$.
- The CDF is a monotonic, nondecreasing function of both x and y.
- A joint CDF to evaluate the probability that the pair of random variables (X, Y) falls into a rectangular region bounded by the points (x_1, y_1), (x_2, y_1), (x_1, y_2), and (x_2, y_2).

- properties of joint CDFs are
  1) $F_{X,Y}(-\infty, -\infty) = F_{X,Y}(-\infty, y) = F_{X,Y}(x, -\infty) = 0;$
  2) $F_{X,Y}(\infty, \infty) = 1;$
  3) $0 \leq F_{X,Y}(x, y) \leq 1;$
  4) $F_{X,Y}(x, \infty) = F_X(x), F_{X,Y}(\infty, y) = F_Y(y);$
  5) $Pr(x_1 < X_1 \leq x_2, y_1 < Y_1 \leq y_2) =$
     $F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1) \geq 0.$
- we are interested in also calculating the probability of the pair of random variables falling in a region that is not rectangular (e.g., a circle or triangle).
- This can be done by forming the required region using many infinitesimal rectangles and then repeatedly applying property

Marginal Distribution

- In the case of two or more random variables, the statistics of each individual variable are called marginal Distribution/statistics.

i Marginal cdf of X and Y

$$F_X(x) = \lim_{y\to\infty} F_{XY}(X, Y) = F_{XY}(X, \infty)$$

$$F_Y(y) = \lim_{x\to\infty} F_{XY}(X, Y) = F_{XY}(\infty, Y)$$

ii Marginal pdf of X and Y

$$f_x(x) = \int_{-\infty}^{\infty} f_{xy}(x, y)dy$$

$$f_y(y) = \int_{-\infty}^{\infty} f_{xy}(x, y)dx$$

- One of the simplest example of a pair of random variables is one that is uniformly distributed over the unit square (i.e., $0 < x < 1, 0 < y < 1$). The CDF of such a random variable is

$$F_{X,Y}(x,y) = \begin{cases} 0 & x < 0 \text{ or } y < 0 \\ x & 0 \le x \le 1,\ y > 1 \\ y & x > 1,\ 0 \le y \le 1 \\ xy & 0 \le x \le 1,\ 0 \le y \le 1 \\ 1 & x > 1,\ y > 1 \end{cases}$$

- this function does indeed satisfy all the properties of a joint CDF. From this joint CDF, the marginal CDF of X can be found to be

$$F_X(x) = F_{X,Y}(x,\infty) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

- the marginal CDF of X is also a uniform distribution. The same statement holds for Y as well.

THEOREM

The joint PDF $f_{X,Y}(x,y)$ can be obtained from the joint CDF $F_{X,Y}(x,y)$ by taking a partial derivative with respect to each variable. That is

$$f_{X,Y}(x,y) = \frac{\delta^2}{\delta x \delta y} F_{X,Y}(x,y)$$

- This theorem shows that we can obtain a joint PDF from a joint CDF by differentiating with respect to each variable.
- The converse of this statement would be that we could obtain a joint CDF from a joint PDF by integrating with respect to each variable. Specifically,

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)dudv$$

- From the joint CDF given in above Example 5.1, it is easily found the joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 1 & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

- how much simpler the joint PDF is to specify than is the joint CDF.

several properties of joint PDFs are summarized as follows:

1) $f_{X,Y}(x,y) \geq 0$;
2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = 1$
3) $F_{X,Y}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y}(u,v)dudv$
4) $f_{X,Y}(x,y) = \frac{\delta^2}{\delta x \delta y} F_{X,Y}(x,y)$
5) $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy, \;\; f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx,$
6) $Pr(x_1 < X_1 \leq x_2, y_1 < Y_1 \leq y_2) = \int_{y_1}^{y_2}\int_{x_1}^{x_2} f_{X,Y}(x,y)dxdy$

Eg: Suppose a pair of random variables is jointly uniformly distributed over the unit circle. That is, the joint PDF $f_{X,Y}(x,y)$ is constant anywhere such that $x^2 + y^2 < 1$:

$$f_{X,Y}(x,y) = \begin{cases} c & x^2 + y^2 < 1 \\ 0 & \text{otherwise} \end{cases}.$$

The constant c can be determined using the normalization integral for joint PDFs:

$$\int\int_{x^2+y^2<1} c\,dx\,dy = 1 \implies c = \frac{1}{\pi}$$

- The marginal PDF of X is found by integrating y out of the joint PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}dy = \frac{2}{\pi}\sqrt{1-x^2}, for -1 \leq x \leq 1$$

- By symmetry, the marginal PDF of Y would have the same functional form:

$$f_Y(y) = \frac{2}{\pi}\sqrt{1-y^2}, for \quad -1 \leq x \leq 1.$$

- Although X and Y were jointly uniformly distributed, the marginal distributions are not uniform.
- One may be able to form many joint PDFs that produce the same marginal PDFs. For example, suppose we form

$$f_{X,Y}(x,y) = \begin{cases} \frac{4}{\pi^2}\sqrt{(1-x^2)(1-y^2)} & -1 \leq x \leq 1,\ -1 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

- It is easy to verify that this is a valid joint PDF and leads to the same marginal PDFs.
- Yet, this is clearly a completely different joint PDF than the uniform distribution with which we started.
- This reemphasizes the need to specify the joint distributions of random variables and not just their marginal distributions.

Eg: Suppose a pair of random variables has the joint PDF given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi} exp\left(-\frac{x^2+y^2}{2}\right)$$

- The probability that the point (X, Y) falls inside the unit circle is given by

$$Pr(X^2+Y^2<1) = \int\int_{x^2+y^2<1} \frac{1}{2\pi} exp\left(-\frac{x^2+y^2}{2}\right) dxdy$$

- Converting this integral to polar coordinates results in

$$Pr(X^2+Y^2<1) = \int_0^{2\pi}\int_0^1 \frac{r}{2\pi} exp\left(-\frac{r^2}{2}\right) drd\theta = \int_0^r rexp\left(-\frac{r^2}{2}\right) dr$$

$$= -exp\left(-\frac{r^2}{2}\right)|_0^1 = 1 - exp\left(-\frac{1}{2}\right)$$

Joint Probability Mass Functions

- When the random variables are discrete rather than continuous, it is often more convenient to work with probability mass functions rather than PDFs or CDFs

The joint probability mass function

for a pair of discrete random variables X and Y is given by
$P_{X,Y}(x, y) = Pr(\{X = x\} \cap \{Y = y\})$.

- In particular, suppose the random variable X takes on values from the set $\{x_1, x_2, \ldots, x_M\}$ and the random variable Y takes on values from the set $\{y_1, y_2, \ldots, y_N\}$.
- Here, either M and/or N could be potentially infinite, or both could be finite.

Several properties of the joint probability mass function

1) $0 \leq P_{X,Y}(x_m, y_n) \leq 1$;
2) $\sum_{m=1}^{M} \sum_{n=1}^{N} P_{X,Y}(x_m, y_n) = 1$;
3) $P_X(x_m) = \sum_{n=1}^{N} P_{X,Y}(x_m, y_n), \qquad P_Y(y_n) = \sum_{n=1}^{N} P_{X,Y}(x_m, y_n)$
4) $Pr((X, Y) \in A) = \sum \sum_{(X,Y) \in A} P_{X,Y}(x_m, y_n)$.

- EXAMPLE 5.7: A pair of discrete random variables N and M have a joint PMF given by

$$P_{N,M}(n,m) = \frac{(n+m)!}{n!m!}\frac{a^n b^m}{(a+b+1)^{n+m+1}}, m = 0,1,2,3,\ldots, n = 0,1,2,$$

- The marginal PMF of N can be found by summing over m in the joint PMF:

$$P_N(n) = \sum_{m=0}^{\infty} P_{N,M}(n,m) = \sum_{m=0}^{\infty} \frac{(n+m)!}{n!m!}\frac{a^n b^m}{(a+b+1)^{n+m+1}}$$

- To evaluate this series, the following identity is used:

$$\sum_{m=0}^{\infty} \frac{(n+m)!}{n!m!} x^m = \left(\frac{1}{1-x}\right)^{n+1}$$

- The marginal PMF then reduces to

$$P_N(n) = \frac{a^n}{(a+b+1)^{n+1}} \sum_{m=0}^{\infty} \frac{(n+m)!}{n!m!} \frac{b^m}{(a+b+1)^m}$$

$$\frac{a^n}{(a+b+1)^{n+1}} \left(\frac{1}{1-\frac{b}{a+b+1}} \right)^{n+1} = \frac{a^n}{(1+a)^{n+1}}.$$

- Likewise, by symmetry, the marginal PMF of M is

$$P_M(m) = \frac{b^m}{(1+b)^{m+1}}.$$

- Hence, the random variables M and N both follow a geometric distribution.

Conditional Distribution, Density, and Mass Functions

- The notion of conditional distribution functions and conditional density functions where the conditioning event is related to another random variable.
- For example, the score a student achieves on a test given the number of hours the student studied for the test.
- the outside temperature given that the humidity .
- A pair of discrete random variables X and Y with a PMF, $P_{X,Y}(x,y)$, the PMF of X given that the value of Y has been observed.

$$Pr(X = x|Y = y) = \frac{Pr(X = x, Y = y)}{Pr(Y = y)} = \frac{P_{X,Y}(x,y)}{PY(y)} \quad (1)$$

- We refer to this as the conditional PMF of X given Y. By way of notation we write

$$P_{X|Y}(x|y) = \frac{P_{X,Y}(x,y)}{PY(y)}$$

.

- EXAMPLE 5.8: Using the joint PMF given in Example 5.7, along with the marginal PMF found in that example, it is found that

$$P_{N|M}(n|m) = \frac{P_{M,N}(m,n)}{PM(m)} = \frac{(n+m)!}{n!m!}\frac{a^n b^m}{(a+b+1)^{n+m+1}}\frac{(1+b)^{m+1}}{b^m}$$

$$= \frac{n+m)!}{n!m!}\frac{a^n(1+b)^{m+1}}{(a+b+1)^{n+m+1}}$$

- Note that the conditional PMF of N given M is quite different than the marginal PMF of N. That is, knowing M changes the distribution of N.
- Equation (1) can be extended to the case of continuous random variables and PDFs.

Theorem

The conditional PDF of a random variable X given that $Y = y$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

- EXAMPLE 5.9: A certain pair of random variables has a joint PDF given by

$$f_{X,Y}(x,y) = \frac{2abc}{(ax+by+c)^3}u(x)u(y)$$

- for some positive constants a, b, and c. The marginal PDFs are easily found to be

$$f_X(x) = \int_0^\infty f_{X,Y}(x,y)dy = \frac{ac}{(ax+c)^2}u(x)$$

$$f_Y(y) = \int_0^\infty f_{X,Y}(x,y)dx = \frac{bc}{(by+c)^2}u(y)$$

- The conditional PDF of X given Y then works out to be

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{2a(by+c)^2}{(ax+by+c)^3}u(x)$$

.

- The conditional PDF of Y given X could also be determined in a

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{2b(ax+c)^2}{(ax+by+c)^3}u(y)$$

Independent Random Variables

- The two events A and B are statistically independent if $Pr(A, B) = Pr(A)Pr(B)$.
- Restated in terms of the random variables, this condition becomes

$$Pr(X \leq x, Y \leq y) = Pr(X \leq x)Pr(Y \leq y) \implies F_{X,Y}(x, y) = F_X(x)F_Y$$

- Hence, two random variables are statistically independent if their joint CDF factors into a product of the marginal CDFs.
- for statistically independent random variables, the joint PDF factors into a product of the marginal PDFs:

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

- The conditional PDFs.

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

- if X and Y are independent, knowing the value of the random variable X should not change the distribution of Y and vice versa.

- EXAMPLE 5.13: Returning once again to the joint PDF of Example 5.10, we saw in that example that the marginal PDF of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right),$$

- while the conditional PDF of X given Y is

$$f_{X|Y}(x|y) = \sqrt{\frac{2}{3\pi}} exp\left(-\frac{2}{3}(x - \frac{y}{2})^2\right).$$

- Clearly, these two random variables are not independent.

The joint pdf of two continuous random variables X and Y is given by:

$$f_{x,y}(x,y) = \begin{cases} kxy, & 0 < x < 1, 0 < y < 1 \\ 0 & \text{Otherwise;} \end{cases}$$

where k is constant

a) Find the value of k
b) Find the marginal pdf of X and Y
c) Are X and Y independent?
d) Find the P(X+Y<1)
e) Find the conditional PDF of X and Y X|Y, Y|X

$$f_{x,y}(x,y) = \begin{cases} k, & 0 < y <\leq x1 \\ 0 & \text{Otherwise;} \end{cases}$$

where k is constant

a) Find the value of k
b) Find the marginal pdf of X and Y
c) Are X and Y independent?
d) Find the P(0<X< $\frac{1}{2}$)
e) Find the conditional PDF of X and Y X|Y, Y|X

The joint pmf of two discrete random variables X and Y is given by:

$$P_{x,y}(x,y) = \begin{cases} k(2x_i + y_i), & x_i = 1,\ 2; \quad y_i = 1,\ 2; \\ 0 & \text{Otherwise;} \end{cases}$$

a) Find the value of k
b) Find the marginal pdf of X and Y
c) Are X and Y independent?
d) Find the P($x_i = 2,\ y_i = 1$)
e) Find the conditional PDF of X and Y X|Y, Y|X