Conditional Distribution and Density Functions, Vector Random Variables, Joint and Marginal Distribution and densities, Statistical Independence, Sum of Random Variables, Central Limit Theorem

Dr. G. Omprakash

Assistant Professor, ECE, KLEF

Table of Contents

Conditional Distribution

The **conditional probability** of event A given event B is

$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)}$$

Conditional Distribution: Let A be an event $[X \le x]$ for the random variable X. Then the conditional distribution of X is given by

$$F_X(x \setminus B) = P[X \le x \setminus B] = \frac{P[X \le x \cap B]}{P(B)}$$

Properties of Conditional Distribution

- $F_X(-\infty \setminus B) = 0$
- $F_X(\infty \setminus B) = 1$
- $0 \le F_X(x \setminus B) \le 1$
- $F_X(x_1 \setminus B) \le F_X(x_2 \setminus B)$ if $x_1 < x_2$
  - $F_X(x \setminus B)$ is a nondecreasing function of x
- $P\{x_1 < X \le x_2 \setminus B\} = F_X(x_2 \setminus B) - F_X(x_1 \setminus B)$
- $F_X(x^+ \setminus B) = F_X(x \setminus B)$
  - $F_X(x \setminus B)$ is a function continuous from the right

Conditional Density function

Conditional density function of the random variable X is defined as the derivative of the distribution function

$$f_X(x \setminus B) = \frac{dF_X(x \setminus B)}{dx}$$

Properties

1. $f_X(x \setminus B) \geq 0$ for all x
2. $\int_{-\infty}^{\infty} f_X(x \setminus B) dx = 1$
3. $F_X(x \setminus B) = \int_{-\infty}^{x} f_X(\xi \setminus B) d\xi$
4. $P\{x_1 < X \leq x_2 \setminus B\} = \int_{x_1}^{x_2} f_X(x \setminus B)$

Example

Two boxes have red, green and blue balls in them. Number of balls are given in the table. The experiment is to select a box randomly and then pick a ball from selected box. Box2 is larger than Box1, so it is selected frequently. $P(B_1) = 0.2$, $P(B_2) = 0.8$. Draw $F_X(x \setminus B_1), f_X(x \setminus B_1), F_X(x), f_X(x)$

		Box		
x_i	**Ball color**	1	2	**Total**
1	Red	5	80	85
2	Green	35	60	95
3	Blue	60	10	70
	Total	100	150	250

Table: Numbers of colored balls in two boxes

Example

$P(X = 1 \setminus B = B_1) = \frac{5}{100}; P(X = 1 \setminus B = B_2) = \frac{80}{150};$
$P(X = 2 \setminus B = B_1) = \frac{35}{100}; P(X = 2 \setminus B = B_2) = \frac{60}{150};$
$P(X = 3 \setminus B = B_1) = \frac{60}{100}; P(X = 3 \setminus B = B_2) = \frac{10}{150};$

Example

$P(X = 1 \setminus B = B_1) = \frac{5}{100}; P(X = 1 \setminus B = B_2) = \frac{80}{150};$
$P(X = 2 \setminus B = B_1) = \frac{35}{100}; P(X = 2 \setminus B = B_2) = \frac{60}{150};$
$P(X = 3 \setminus B = B_1) = \frac{60}{100}; P(X = 3 \setminus B = B_2) = \frac{10}{150};$
The conditional probability density $f_X(x \setminus B_1)$ is given by:

$$f_X(x \setminus B_1) = \frac{5}{100}\delta(x - 1) + \frac{35}{100}\delta(x - 2) + \frac{60}{100}\delta(x - 3)$$

By direct integration

$$F_X(x \setminus B_1) = \frac{5}{100}u(x - 1) + \frac{35}{100}u(x - 2) + \frac{60}{100}u(x - 3)$$

Finding $F_X(x)$ and $f_X(x)$

Using Total probability

$$P(X = 1) = P(X = 1 \setminus B = B_1)P(B_1) + P(X = 1 \setminus B = B_2)P(B_2)$$
$$= \frac{5}{100}\frac{2}{10} + \frac{80}{150}\frac{8}{10} = 0.437$$
$$P(X = 2) = \frac{35}{100}\frac{2}{10} + \frac{60}{150}\frac{8}{10} = 0.390$$
$$P(X = 3) = \frac{60}{100}\frac{2}{10} + \frac{10}{150}\frac{8}{10} = 0.173$$

$$f_X(x) = 0.437\delta(x - 1) + 0.390\delta(x - 2) + 0.173\delta(x - 3)$$

By direct integration

$$F_X(x) = 0.437u(x - 1) + 0.390u(x - 2) + 0.173u(x - 3)$$

Conditional Event (Second type)

In the previous example $F_X(x \setminus B)$, B is an event. Let us say B depends on the random variable X such that $B = \{X \leq b\}$

$$F_X(x \setminus X \leq b) = P\{X \leq x \setminus X \leq b\} = \frac{P(X \leq x \cap X \leq b)}{P(X \leq b)}$$

For $x \geq b$ above equation is

$$F_X(x \setminus X \leq b) = \frac{P(X \leq x \cap X \leq b)}{P(X \leq b)}$$
$$= \frac{P(X \leq b)}{P(X \leq b)} = 1$$

for $x < b$ above equation is

$$F_X(x \setminus X \leq b) = \frac{P(X \leq x \cap X \leq b)}{P(X \leq b)}$$

$$= \frac{P(X \leq x)}{P(X \leq b)} = \frac{F_X(x)}{F_X(b)}$$

$$F_X(x \setminus X \leq b) = \begin{cases} \frac{F_X(x)}{F_X(b)} & x < b \\ 1 & x \geq b \end{cases}$$

The conditional density function is given by

$$f_X(x \setminus X \leq b) = \begin{cases} \frac{f_X(x)}{F_X(b)} & x < b \\ 0 & x \geq b \end{cases}$$

where $F_X(b) = \int\limits_{-\infty}^{b} f_X(x) dx$

Problems

The radial "miss-distance" of landings from parachuting sky divers, as measured from a target's center, is a Rayleigh random variable with $b = 800m^2$ and $a = 0$. The target is a circle of 50-m radius with a bull's eye of 10-m radius. Find the probability of a parachuter hitting the bull's eye given that the landing is on the target.

$$F_X(x) = [1 - e^{-(x-a)^2/b}]u(x) = [1 - e^{-x^2/800}]u(x)$$

$$P(\text{Hitting bull's eye} \setminus \text{Landing is on target}) = \frac{P(X \leq 10)}{P(X \leq 50)} = \frac{F_X(10)}{F_X(50)}$$
$$= \frac{[1 - e^{-100/800}]}{[1 - e^{-2500/800}]} = 0.1229$$

Vector Random variables

Why Do We Need Vector Random Variables?

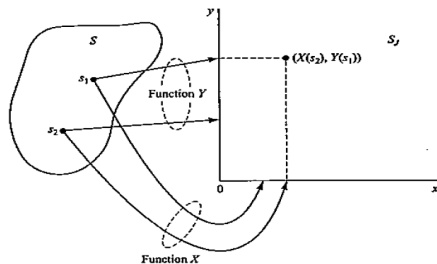Real-world systems almost always involve multiple random quantities interacting with each other.

- A patient's systolic blood pressure and heart rate are both random and often correlated.
    - $X = [BloodPressure, HeartRate]$;
- We can track a portfolio of 5–10 stocks
    - $R = [R_1, R_2, R_3, R_4, R_5]$
- Speed and direction of wind
- Temperature and humidity
- X and Y position of a robot

Vector Random Variables

Consider two random variables X and Y defined on the sample space S

- Ordered pair (x, y) is considered to be a *random point* in the xy plane
- This order pair of number (x, y) is known as specific value of *vector random variable*
- X and Y denote the two dimensional vector random variable
- The plane of all points (x, y) : *joint sample space* or *range sample space*

Source: Peebles

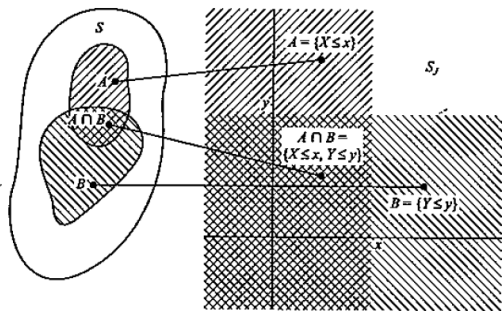Figure: Mapping from sample space S to joint sample space S_J

Vector Random Variables

Source: Peebles

- Two events A and B refer to sample space S
- Event A is defined as $A = \{X \leq x\}$
- Event B is defined as $B = \{Y \leq y\}$

Figure: Comparison of events in S and S_J

In a more general case for N random variables $X_1, X_2, X_3, ..., X_N \implies N$ dimensional random vector

Joint Probability Distribution

Consider two events $A = \{X \leq x\}$ and $B = \{Y \leq y\}$ with distribution functions defined as

$$F_X(x) = P\{X \leq x\}$$
$$F_Y(y) = P\{Y \leq y\}$$

The probability of the joint event $\{X \leq x, Y \leq y\}$ is defined as the *joint probability distribution function*. It is denoted by $F_{X,Y}(x, y)$

$$F_{X,Y}(x, y) = P\{X \leq x, Y \leq y\}$$

Note: $P\{X \leq x, Y \leq y\} = P(A \cap B)$.
HA: Solved examples 4.2 and 4.3 in Ramesh Babu book

Properties of Joint Distribution

- $F_{X,Y}(-\infty, -\infty) = 0;$ $\qquad F_{X,Y}(-\infty, y) = 0;$ $\qquad F_{X,Y}(x, -\infty) = 0$
- $F_{X,Y}(\infty, \infty) = 1$
- $0 \leq F_{X,Y}(x, y) \leq 1$
- $F_{X,Y}(x, y)$ is a nondecreasing function of both x and y
- $F_{X,Y}(x_2, y_2) + F_{X,Y}(x_1, y_1) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) = P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} \geq 0$
- $F_{X,Y}(x, \infty) = F_X(x)$ $\qquad F_{X,Y}(\infty, y) = F_Y(y)$

Joint Density Function

The *joint probability density function*, denoted by $f_{X,Y}(x, y)$, is defined as the second derivative of the joint distribution function

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

if X and Y are discrete

$$f_{X,Y}(x, y) = \sum_{n=1}^{N} \sum_{m=1}^{M} P(x_n, y_m)\delta(x - x_n)\delta(y - y_n)$$

Properties of Joint Density

- $f_{X,Y}(x, y) \geq 0$

- $\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$

- $F_{X,Y}(x, y) = \int\limits_{-\infty}^{y} \int\limits_{-\infty}^{x} f_{X,Y}(u, v)\, du\, dv$

- $F_X(x) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{\infty} f_{X,Y}(u, v)\, du\, dv;$ $\qquad F_Y(y) = \int\limits_{-\infty}^{y} \int\limits_{-\infty}^{\infty} f_{X,Y}(u, v)\, du\, dv$

- $P\{x_1 < X \leq x_2, y_1 < Y \leq y_2\} = \int\limits_{y_1}^{y_2} \int\limits_{x_1}^{x_2} f_{X,Y}(x, y)\, dx\, dy$

- $f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy$ $\qquad f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx$

Marginal Distributions

$F_{X,Y}(x, \infty) = F_X(x); \qquad F_{X,Y}(\infty, y) = F_Y(y).$

The functions $F_X(x)$ and $F_Y(y)$ obtained in this manner are called *marginal distribution functions*.

- We know $P\{X \leq x, Y \leq y\} = P(A \cap B)$.
- set $y = \infty \implies B = \{Y \leq \infty\} = S$
  - $B = \{Y \leq \infty\}$ is equivalent to making B the certain event
- $P(A \cap B) = P(A \cap S) = P(A) = P\{X \leq x\} = F_X(x)$

Marginal Density functions

The marginal density functions are defined as the derivatives of the marginal distribution functions

$$f_X(x) = \frac{dF_X(x)}{dx}$$
$$f_Y(y) = \frac{dF_Y(y)}{dy}$$

If $f_{X,Y}(x, y)$ is given then

$$f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)dy \qquad f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y)dx$$

Problem

Let X, Y and Z be three jointly continuous random variables with joint PDF. Find c and marginal PDF of X (Ex 24)

$$f_{X,Y,Z}(x,y,z) = \begin{cases} c(x + 2y + 3z) & 0 \le x, y, z \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Problem

Let X, Y and Z be three jointly continuous random variables with joint PDF. Find c and marginal PDF of X (Ex 24)

$$f_{X,Y,Z}(x,y,z) = \begin{cases} c(x + 2y + 3z) & 0 \leq x, y, z \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

c=1/3,

$$f_X(x) = \begin{cases} \frac{1}{3}(x + \frac{5}{2}) & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Problem

If $g(x, y)$ is a valid PDF. find the value of b

$$g_{X,Y}(x, y) = \begin{cases} be^{-x}cos(y) & 0 \le x \le 2 \text{ and } 0 \le y \le \pi/2 \\ 0 & \text{all other } x \text{ and } y \end{cases}$$

Problem

If $g(x, y)$ is a valid PDF. find the value of b

$$g_{X,Y}(x, y) = \begin{cases} be^{-x}\cos(y) & 0 \le x \le 2 \text{ and } 0 \le y \le \pi/2 \\ 0 & \text{all other } x \text{ and } y \end{cases}$$

$$b \int\limits_0^2 e^{-x}dx \int\limits_0^{\pi/2} \cos(y)dy = 1$$

Problem

If $g(x, y)$ is a valid PDF. find the value of b

$$g_{X,Y}(x, y) = \begin{cases} be^{-x}cos(y) & 0 \le x \le 2 \text{ and } 0 \le y \le \pi/2 \\ 0 & \text{all other } x \text{ and } y \end{cases}$$

$$b \int\limits_0^2 e^{-x}dx \int\limits_0^{\pi/2} cos(y)dy = 1$$

$b = \frac{1}{1-e^{-2}}$

Home Assignment

Given the joint density function

$$f_{X,Y}(x,y) = \begin{cases} \frac{x(1+3y^2)}{4} & 0 \le x \le 2 \text{ and } 0 \le y \le 1 \\ 0 & \text{all other } x \text{ and } y \end{cases}$$

Find $f_X(x), f_Y(y), f_{X,Y}(x/y)$. (Example 25)

Problem

For the bivariate probability distribution given below, find $P(X \leq 1)$, $P(Y \leq 3)$, $P(X \leq 1, Y \leq 3)$, $P(X \leq 1 \setminus Y \leq 3)$

$X \backslash Y$		1	2	3	4	5	6
	0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
X	1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
	2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

Table: Joint Probability Table of X and Y

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$
$$= \sum_{j=1}^{6} P(X = 0, Y = j) + \sum_{j=1}^{6} P(X = 1, Y = j) = \frac{7}{8}$$

Problem

$$P(Y \le 3) = P(Y = 1) + P(Y = 2)P(Y = 3) = \frac{23}{64}$$

$$P(X \le 1, Y \le 3) = \frac{9}{32}$$

$$P(X \le 1 \setminus Y \le 3) = \frac{18}{23}$$

For Discrete Random Variable X

$$F_X(x) = \sum_i p_i u(x - x_i)$$

$$f_X(x) = p(x) = \sum_i p_i \delta(x - x_i)$$

For Discrete Random Variables X and Y

$$F_{X,Y}(x, y) = \sum_i \sum_j p(x_i, y_j) u(x - x_i) u(y - y_j)$$

$$f_{X,Y}(x, y) = p(x_i, y_j) = \sum_i \sum_j p(x_i, y_j) \delta(x - x_i) \delta(y - y_j)$$

$p(x_i, y_j)$ is called the *mass function*

Conditional Distribution and Density

The conditional distribution of random variable X, given some event B is defined as

$$F_X(x \setminus B) = P\{X \le x \setminus B\} = \frac{P\{X \le x \cap B\}}{P(B)} \quad (1)$$

Corresponding density function is defined by taking the derivative

$$f_X(x \setminus B) = \frac{dF_X(x \setminus B)}{dx}$$

Point Conditioning

In practical problems, we are interested in distribution function of one random variable X conditioned by the second random variable Y has some specific value.

$$B = \{y - \Delta y < Y \leq y + \Delta y\} \qquad \text{where} \quad \Delta y \to 0$$

Equation 1 can be written as

$$F_X(x \setminus y - \Delta y < Y \leq y + \Delta y) = \frac{P\{X \leq x \setminus y - \Delta y < Y \leq y + \Delta y\}}{P\{y - \Delta y < Y \leq y + \Delta y\}}$$

$$= \frac{\int\limits_{y-\Delta y}^{y+\Delta y} \int\limits_{-\infty}^{x} f_{X,Y}(\xi_1, \xi_2) d\xi_1 d\xi_2}{\int\limits_{y-\Delta y}^{y+\Delta y} f_Y(\xi) d\xi}$$

Point Conditioning

Since Δy is very small quantity the above expression is written as

$$F_X(x \setminus y - \Delta y < Y \leq y + \Delta y) = \frac{\int\limits_{-\infty}^{x} f_{X,Y}(\xi_1, \xi_2) d\xi_1 2\Delta y}{f_Y(\xi) 2\Delta y}$$

$$F_X(x \setminus B) = \frac{\int\limits_{-\infty}^{x} f_{X,Y}(\xi_1, \xi_2) d\xi_1}{f_Y(\xi)}$$

Differentiate both sides w.r.to x to get conditional density

$$\frac{d}{dx}\{F_X(x \setminus B)\} = \frac{d}{dx}\left(\frac{\int\limits_{-\infty}^{x} f_{X,Y}(\xi_1, \xi_2) d\xi_1}{f_Y(\xi)}\right) \implies f_X(x \setminus B) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

Point Conditioning

For discrete random variables with values x_i and y_j, as $\Delta y \rightarrow 0$, the conditional distribution is given by

$$F_X(x \setminus Y = y_k) = \sum_{l=1}^{N} \frac{P(x_l, y_k)}{P(y_k)} u(x - x_l)$$

by differentiating the above equation, the conditional density is given by

$$f_X(x \setminus Y = y_k) = \sum_{l=1}^{N} \frac{P(x_l, y_k)}{P(y_k)} \delta(x - x_l)$$

Interval Conditioning

$$B = \{y_a < Y \le y_b\}$$

$$F_X(x \setminus y_a < Y \le y_b) = \frac{F_{X,Y}(x, y_b) - F_{X,Y}(x, y_a)}{F_Y(y_b) - F_Y(y_b)}$$

Consider the denominator term

$$F_Y(y_b) - F_Y(y_a) = \int_{y_a}^{y_b} f_Y(y)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

$$F_Y(y_b) - F_Y(y_a) = \int_{y_a}^{y_b} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy$$

Interval conditioning

$$F_X(x \setminus y_a < Y \le y_b) = \frac{F_{X,Y}(x, y_b) - F_{X,Y}(x, y_a)}{F_Y(y_b) - F_Y(y_b)}$$

$$= \frac{\int\limits_{y_a}^{y_b} \int\limits_{-\infty}^{x} f_{X,Y}(\xi, y) d\xi dy}{\int\limits_{y_a}^{y_b} \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy}$$

$$f_X(x \setminus y_a < Y \le y_b) = \frac{\int\limits_{y_a}^{y_b} f_{X,Y}(x, y) dy}{\int\limits_{y_a}^{y_b} \int\limits_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy}$$

Independence-Recap

Two events A and B are *statistically independent* if the probability of occurrence of one event is not affected by the occurrence of the other event.

$$P(A \setminus B) = P(A) \qquad or \qquad P(B \setminus A) = P(B)$$

From conditional probability

$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)}$$

Substitute eqn(34) in the above equation. We get

$$P(A) = \frac{P(A \cap B)}{P(B)} \implies P(A \cap B) = P(A)P(B);$$

Two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

Statistical Independence

- The concept of statistical independence is used when two (or more) random variables **do not influence each other**
- When variables are independent, **probability calculations become much simpler** because we can replace complicated joint probability terms with products of marginals.
- **Reliability Engineering**
    - If components in a system fail independently:
        - P(system works)=P(component 1 works)× P(component 2 works)
- **Finance / Risk Modeling**
    - If two asset returns are independent:
    - Easier to compute risk of combined investments.
- **Naïve Bayes** assumes conditional independence of features given the class.
    - $P(X_1, X_2, ..., X_n \setminus Y) = \prod_{i=1}^{n} P(X_i \setminus Y)$

Statistical Independence

Two random variables X and Y defining the events $A = \{X \leq x\}$ and $B = \{Y \leq y\}$, where x and y are real numbers. X and Y are said to be *statistically independent random variables* if and only if

$$P\{X \leq x, Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$$

From the definition of distribution function

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

From the definition of density function

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

Statistical Independence

The conditional distribution for independent events is

$$F_X(x \setminus Y \leq y) = \frac{P\{X \leq x, Y \leq y\}}{P\{Y \leq y\}} = \frac{F_{X,Y}(x,y)}{F_Y(y)} = \frac{F_X(x)F_Y(y)}{F_Y(y)} = F_X(x)$$

$$F_X(x \setminus Y \leq y) = F_X(x)$$
$$F_Y(y \setminus X \leq x) = F_Y(y)$$

Similarly the conditional density becomes

$$f(x \setminus Y \leq y) = f_X(x)$$
$$f(y \setminus X \leq x) = f_Y(y)$$

Problems

The joint density of two random variables X and Y is

$$f_{X,Y}(x,y) = \frac{1}{12} u(x) u(y) e^{-(x/4)-(y/3)}$$

Are X and Y statistically independent?

Problems

The joint density of two random variables X and Y is

$$f_{X,Y}(x,y) = \frac{1}{12} u(x)u(y)e^{-(x/4)-(y/3)}$$

Are X and Y statistically independent?

$$f_X(x) = (1/4)u(x)e^{-(x/4)}$$
$$f_Y(y) = (1/3)u(y)e^{-(y/3)}$$

Since $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, then X and Y are independent

Problem

The joint PDF of RV (X, Y) is given by $f(x, y) = kxye^{-(x^2+y^2)}$, $x > 0$ and $y > 0$. Find the value of k and prove X and Y are independent.

Problem

The joint PDF of RV (X, Y) is given by $f(x, y) = kxye^{-(x^2+y^2)}$, $x > 0$ and $y > 0$. Find the value of k and prove X and Y are independent.

$k = 4, f_X(x) = 2xe^{-x^2}, x > 0$
$f_Y(y) = 2ye^{-y^2}, y > 0$

Problem

(X, Y) is a bivariate random variable in which X and Y are independent. If X is a uniform random variable over $(0, 0.5)$ and Y is an exponential random variable with parameter $\lambda = 4$, find the joint pdf of (X, Y) (P4.50 in Ramesh Babu).

Sum of Random Variables

Sum of random variables shows up in many real-world scenarios whenever multiple uncertain effects combine to produce a total outcome.

- Total Measurement Error. In physics experiments
    - Total Error=Thermal Noise+Quantization Noise+Interference Noise
- Total Waiting Time in Queues
    - Wait time=check-in + security + boarding at an airport
- Total Signal from Multiple Sources
    - Noise + interference + intended signal are all random variables whose sum forms the received signal.
- Central Limit Theorem in Action
    - height of a plant = genetics + soil + rainfall + random variation

Sum of Two Random Variables

Let W be a random variable equal to sum of two independent random variables X and Y

$$W = X + Y$$

Eg: Signal voltage plus noise at some receiver

$$F_W(w) = P(W \leq w) = P(X + Y \leq w) = \iint\limits_{X+Y \leq Z} f_{X,Y}(x,y)dxdy$$

Since X and Y are independent

$$F_W(w) = \iint\limits_{X+Y \leq Z} f_X(x)f_Y(y))dxdy$$

y limits are from $-\infty$ to ∞; x limits are from $-\infty$ to $x = w - y$

Sum of Random Variables

$$F_W(w) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{w-y} f_X(x) f_Y(y)) \, dx \, dy = \int\limits_{-\infty}^{\infty} f_Y(y) \, dy \int\limits_{-\infty}^{w-y} f_X(x) \, dx$$

To get the density function, differentiate above equation using Leibniz's rule

$$f_W(w) = \int\limits_{-\infty}^{\infty} f_Y(y) f_X(w-y) \, dy$$

The density function of the sum of two statistically independent random variables is the convolution of their individual density functions

Problem

Two independent random variables X and Y have densities
$f_X(x) = 5e^{-5x}u(x)$ and $f_Y(y) = 2e^{-2y}u(y)$. Find the density of the sum
$Z = X + Y$. (P4.60 in Ramesh Babu)

$$f_Z(z) = \int\limits_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx$$

$$f_Z(z) = \frac{10}{3}(e^{-2z} - e^{-5z})u(z)$$

Sum of several Random Variables

Let Y be the sum of N independent random variables $Y = X_1 + X_2 + \cdots + X_N$

$$Y_1 = X_1 + X_2 \implies f_{Y_1}(y_1) = f_{X_1}(x_1) * f_{X_2}(x_2)$$

$$Y_2 = X_1 + X_2 + X_3 = Y_1 + X_3 \implies f_{Y_2}(y_2) = f_{Y_1}(y_1) * f_{X_3}(x_3)$$
$$= f_{X_1}(x_1) * f_{X_2}(x_2) * f_{X_3}(x_3)$$

The density function of Y is given by

$$f_Y(y) = f_{X_1}(x_1) * f_{X_2}(x_2) * \cdots * f_{X_N}(x_N)$$

Central Limit Theorem

Central Limit Theorem *says that the probability function of the sum of a large number of random variables approaches a Gaussian distribution.*

Unequal distributions Let $\tilde{X}_i$ and $\sigma_{X_i}^2$ be the means and variances of N random variables X_i, $i = 1, 2, ..., N$, which may have arbitrary probability densities. The central limit theorem states that the sum $Y_N = X_1 + X_2 + \cdots X_N$, which has mean $\tilde{Y}_N = \tilde{X}_1 + \tilde{X}_2 + \cdots + \tilde{X}_N$ and variance $\sigma_{Y_N}^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \cdots \sigma_{X_N}^2$, has a probability distribution that asymptotically approaches gaussian as $N \to \infty$.

$$\sigma_{X_i}^2 > B_1 > 0 \qquad i = 1, 2, \cdots, N$$
$$E[|X_i - \tilde{X}_i|^3] < B_2 \qquad i = 1, 2, \cdots, N$$

These conditions guarantee that no one random variable in the sum dominates

Central Limit Theorem

Equal distributions: Let $\tilde{X}$ and σ_X^2 be the mean and variance of N statistically independent continuous random variables X_i, $i = 1, 2, ..., N$, with same probability densities. The central limit theorem states that the sum $Y_N = X_1 + X_2 + \cdots X_N$ has a probability distribution that asymptotically approaches gaussian as $N \rightarrow \infty$.

Appendix

Central Limit Theorem for equal distributions

To prove: the characteristic function of W_N is that of zero mean and unit-variance gaussian random variable

$$\Phi_{W_N}(\omega) = e^{-\omega^2/2}$$

and

$$W_N = (Y_N - \tilde{Y}_N)/\sigma_{Y_N} = \frac{\sum\limits_{i=1}^{N}(X_i - \tilde{X}_i)}{\left[\sum\limits_{i=1}^{N}\sigma_{X_i}^2\right]^{1/2}}$$

$$= \frac{1}{\sqrt{N}\sigma_X}\sum_{i=1}^{N}(X_i - \tilde{X}_i)$$

Acknowledge various sources for the images.
Thankyou