CS5805 : Machine Learning I
Lecture # 6

Reza Jafari, Ph.D

Collegiate Associate Professor
rjafari@vt.edu

COLLEGE OF ENGINEERING
COMPUTER SCIENCE
VIRGINIA TECH.

Introduction to Probability

- What is a probability?

Introduction to Probability

- What is a probability?
- One of the definition is based on the relative frequency :

$$P(E) = \frac{\#times\ an\ E\ occurs}{\#Trials}$$
$$as \#Trials \to \infty$$

- What is a probability?
- One of the definition is based on the relative frequency :

$$P(E) = \frac{\#times\ an\ E\ occurs}{\#Trials}$$
$$as\#Trials \to \infty$$

- The problem with this definition is that # Trials$\to \infty$ is impossible.

Other definition of Probability

- **Axiomatic Probability**:
  1. P(E)$\geq$ 0, non-negative
  2. P(S) $=$ 1, S: is the sample space
  3. If $A_1, A_2, ..., An$ is mutually exclusive (disjoint) then

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

Other definition of Probability

- **Axiomatic Probability**:
  1. P(E)$\geq$ 0, non-negative
  2. P(S) $=$ 1, S: is the sample space
  3. If $A_1, A_2, ..., An$ is mutually exclusive (disjoint) then

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

- Mutually exclusive events: are the events that can not occur simultaneously.

Other definition of Probability

- **Axiomatic Probability**:
  1. P(E)$\geq$ 0, non-negative
  2. P(S) $=$ 1, S: is the sample space
  3. If $A_1, A_2, ..., An$ is mutually exclusive (disjoint) then

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

- Mutually exclusive events: are the events that can not occur simultaneously.
- In other words, if one event has already occurred, the other event can not occur.

Other definition of Probability

- **Axiomatic Probability**:
  1. P(E)$\geq$ 0, non-negative
  2. P(S) $=$ 1, S: is the sample space
  3. If $A_1, A_2, ..., An$ is mutually exclusive (disjoint) then

$$P(A_1 \cup A_2 \cup ... \cup A_n) = P(A_1) + P(A_2) + ... + P(A_n)$$

- Mutually exclusive events: are the events that can not occur simultaneously.
- In other words, if one event has already occurred, the other event can not occur.
- Flipping a coin: Once you get H, there is no way to get T.

Disjoint Events

Independence

- Two events are independent, if the probability of the outcome of one event does not influence the probability of outcome of another event.

Independence

- Two events are independent, if the probability of the outcome of one event does not influence the probability of outcome of another event.
- Let event $A_1, A_2, ..., A_n$ are independent events (meaning occurness one does not depend on the other) then:

$$P(A_1 \cap A_2 ... \cap A_n) = P(A_1) * P(A_2) ... * P(A_n)$$

Independence

- Two events are independent, if the probability of the outcome of one event does not influence the probability of outcome of another event.
- Let event $A_1, A_2, ..., A_n$ are independent events (meaning occurness one does not depend on the other) then:

$$P(A_1 \cap A_2 ... \cap A_n) = P(A_1) * P(A_2) ... * P(A_n)$$

- Hence, in order to prove that an event A is independent of event B, it requires to show that

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

Independence

- Two events are independent, if the probability of the outcome of one event does not influence the probability of outcome of another event.
- Let event $A_1, A_2, ..., A_n$ are independent events (meaning occurness one does not depend on the other) then:

$$P(A_1 \cap A_2 ... \cap A_n) = P(A_1) * P(A_2) ... * P(A_n)$$

- Hence, in order to prove that an event A is independent of event B, it requires to show that

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

- Does independence imply disjoint?

Independence

- Two events are independent, if the probability of the outcome of one event does not influence the probability of outcome of another event.
- Let event $A_1, A_2, ..., A_n$ are independent events (meaning occurness one does not depend on the other) then:

$$P(A_1 \cap A_2 ... \cap A_n) = P(A_1) * P(A_2) ... * P(A_n)$$

- Hence, in order to prove that an event A is independent of event B, it requires to show that

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

- Does independence imply disjoint?
- Does disjoint imply independence?

Conditional Probability

- The conditional probability of an event B is the probability that the event will occur given the knowledge that an event A has already occurred.

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

Conditional Probability

- The conditional probability of an event B is the probability that the event will occur given the knowledge that an event A has already occurred.

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

- Or the joint probability between A and B can be calculated as

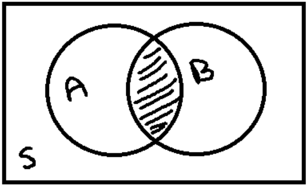
$$P(A \cap B) = P(B|A) * P(A)$$

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

Example

- A and B are not mutually exclusive in below venn diagram, hence:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Example

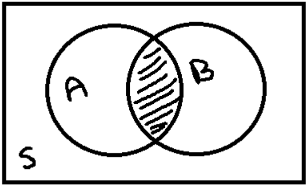
- A and B are not mutually exclusive in below venn diagram, hence:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- If A and B are independent then $P(B|A) = P(B)$ because:

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

- Similarly $P(A|B) = P(A)$



Example

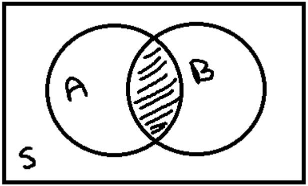
- A and B are not mutually exclusive in below venn diagram, hence:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- If A and B are independent then $P(B|A) = P(B)$ because:

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

- Similarly $P(A|B) = P(A)$
- Conditional Probability is extremely important in parameter estimation and forecasting of time-series model.



Theorem of Total Probability

- If $B_1, B_2, B_3, ...$ is a partition of the sample space S, then for any event A we have:

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$$

Theorem of Total Probability

- If $B_1, B_2, B_3, ...$ is a partition of the sample space S, then for any event A we have:

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i)$$

- **Example**: There are three bags that each contain 100 marbles:
  - Bag 1 has 75 red and 25 blue marbles;
  - Bag 2 has 60 red and 40 blue marbles;
  - Bag 3 has 45 red and 55 blue marbles;

  I choose one of the bags at random and then pick a marble from the chosen bag, also at random. What is the probability that the chosen marble is red?

Total Probability and Decision tree

- The decision tree is a simple and convenient method of visualizing problem with the total probability rule.

Total Probability and Decision tree

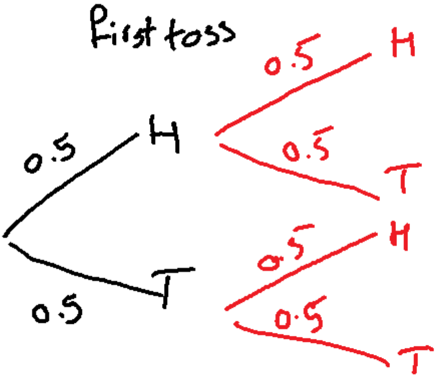
- The decision tree is a simple and convenient method of visualizing problem with the total probability rule.
- The decision tree depicts all possible events in the sequence and quickly identify the relationships between events and calculate the conational probability.

Total Probability and Decision tree

- The decision tree is a simple and convenient method of visualizing problem with the total probability rule.
- The decision tree depicts all possible events in the sequence and quickly identify the relationships between events and calculate the conational probability.
- Let toss a fair coin twice. What is the probability of getting TT?

Total Probability and Decision tree

- The decision tree is a simple and convenient method of visualizing problem with the total probability rule.
- The decision tree depicts all possible events in the sequence and quickly identify the relationships between events and calculate the conational probability.
- Let toss a fair coin twice. What is the probability of getting TT?



Example

- You are a stock analyst for a company. You discovered that the company is planning to launch a new project that is likely to affect the company stock price. You identified the following probabilities:
  1. There is a 60% probability of launching a new project.
  2. If a company launches the project, there is a 75% chance than the company stock will increase.
  3. If a company does not launch the project, there is a 30% chance probability that company stock price will increase.

Example

- You are a stock analyst for a company. You discovered that the company is planning to launch a new project that is likely to affect the company stock price. You identified the following probabilities:
  1. There is a 60% probability of launching a new project.
  2. If a company launches the project, there is a 75% chance than the company stock will increase.
  3. If a company does not launch the project, there is a 30% chance probability that company stock price will increase.
- Find the probability that company's stock will increase?

Example

- You are a stock analyst for a company. You discovered that the company is planning to launch a new project that is likely to affect the company stock price. You identified the following probabilities:
    1. There is a 60% probability of launching a new project.
    2. If a company launches the project, there is a 75% chance than the company stock will increase.
    3. If a company does not launch the project, there is a 30% chance probability that company stock price will increase.
- Find the probability that company's stock will increase?
- Find the probability that company's stock will increase given that new project is launched?

Example

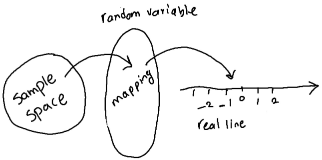
- You are a stock analyst for a company. You discovered that the company is planning to launch a new project that is likely to affect the company stock price. You identified the following probabilities:
  1. There is a 60% probability of launching a new project.
  2. If a company launches the project, there is a 75% chance than the company stock will increase.
  3. If a company does not launch the project, there is a 30% chance probability that company stock price will increase.
- Find the probability that company's stock will increase?
- Find the probability that company's stock will increase given that new project is launched?
- Find the probability that company's stock will increase given that new project is not launched?

Example

- You are a stock analyst for a company. You discovered that the company is planning to launch a new project that is likely to affect the company stock price. You identified the following probabilities:
    1. There is a 60% probability of launching a new project.
    2. If a company launches the project, there is a 75% chance than the company stock will increase.
    3. If a company does not launch the project, there is a 30% chance probability that company stock price will increase.
- Find the probability that company's stock will increase?
- Find the probability that company's stock will increase given that new project is launched?
- Find the probability that company's stock will increase given that new project is not launched?
- Should the company launch the project or not?

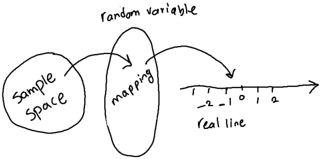
Random Variable

- Given an experiment with events and their associated probabilities. A random variable is a mapping from a sample space to the real line.



Random Variable

- Given an experiment with events and their associated probabilities. A random variable is a mapping from a sample space to the real line.



- Random variables map events to numbers where functions maps numbers to numbers. For example:

$$y = \sin(\frac{\pi}{2}) \tag{1}$$

Probability Distribution Function

- What is the purpose of this mapping?

$$F_x(\lambda) = P(\{S : X(s) \leq \lambda\}) \tag{2}$$

Probability Distribution Function

- What is the purpose of this mapping?
  - It shows probability distribution function.

$$F_x(\lambda) = P(\{S : X(s) \leq \lambda\}) \tag{2}$$

Probability Distribution Function

- What is the purpose of this mapping?
  - It shows probability distribution function.

$$F_x(\lambda) = P(\{S : X(s) \leq \lambda\}) \tag{2}$$

- λ is the dummy variable and X is the random variable.

Probability Distribution Function

- What is the purpose of this mapping?
  - It shows probability distribution function.

$$F_x(\lambda) = P(\{S : X(s) \leq \lambda\}) \tag{2}$$

- λ is the dummy variable and X is the random variable.
- It is very important to use different letters for random variable (X) and dummy variable (λ)

Probability Distribution Function

- What is the purpose of this mapping?
  - It shows probability distribution function.

$$F_x(\lambda) = P(\{S : X(s) \leq \lambda\}) \tag{2}$$

- λ is the dummy variable and X is the random variable.
- It is very important to use different letters for random variable (X) and dummy variable (λ)
- Random variable is a mapping and the dummy variable is the place holder.

Probability Density Function

- The derivative of probability distribution function is defined as probability density function:

$$f_x(\lambda) = \frac{d}{d\lambda} F_x(\lambda) \tag{3}$$

or

$$F_x(\lambda) = \int_{-\infty}^{\lambda} f_x(\mu) d\mu \tag{4}$$

Probability Density Function

- The derivative of probability distribution function is defined as probability density function:

$$f_x(\lambda) = \frac{d}{d\lambda} F_x(\lambda) \tag{3}$$

or

$$F_x(\lambda) = \int_{-\infty}^{\lambda} f_x(\mu) d\mu \tag{4}$$

- In other word the area under the density function shows probability of the event occur. Hence, the area under the density function from $-\infty$ to ∞ must be equal to 1.

Probability Density Function

- The derivative of probability distribution function is defined as probability density function:

$$f_x(\lambda) = \frac{d}{d\lambda} F_x(\lambda) \tag{3}$$

or

$$F_x(\lambda) = \int_{-\infty}^{\lambda} f_x(\mu) d\mu \tag{4}$$

- In other word the area under the density function shows probability of the event occur. Hence, the area under the density function from $-\infty$ to ∞ must be equal to 1.
- In order to define a random variable, the density function is needed.

Probability Density Function

- The derivative of probability distribution function is defined as probability density function:

$$f_x(\lambda) = \frac{d}{d\lambda} F_x(\lambda) \tag{3}$$

or

$$F_x(\lambda) = \int_{-\infty}^{\lambda} f_x(\mu) d\mu \tag{4}$$

- In other word the area under the density function shows probability of the event occur. Hence, the area under the density function from $-\infty$ to ∞ must be equal to 1.
- In order to define a random variable, the density function is needed.
- Knowing density function, reveal all information about the random variable.

Summery

Consider a continuous random variable X with probability density function $f_X(\lambda)$. We have:

1. $f_X(\lambda) \geq 0$ for all λ in R
2. $\int_{-\infty}^{\infty} f_X(\lambda) d\lambda = 1$
3. $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(\lambda) d\lambda$

Example

Let consider a uniform density function for a random variable X. This random variable is said to have *Uniform(a,b)* distribution. Or it can mathematically to be written as:

$$f_X(\lambda) = \begin{cases} c, & 0 < a < \lambda < b \\ 0, & \text{Else} \end{cases}$$

Graph the above density function and find c in terms if a and b.

Example

Let X be a continuous random variable with the following pdf:

$$f_X(\lambda) = \begin{cases} ce^{-\lambda}, & \lambda \geq 0 \\ 0, & \text{Else} \end{cases}$$

1. Find c.
2. Find the probability distribution function $F_X(\lambda)$
3. Find $P(1 < X \leq 3)$
4. Find $P(X = 2)$
5. Find $P(X \in [0,1] \cup [3,4])$

Example

Let X be a continuous random variable that is equally likely to be any valye between 80 and 100.

1. Graph the corresponding probability density function.
2. Find $P(90 < X \leq 95)$

Joint Random Variables

- Two random variables X and Y are jointly continuous if they have a joint probability distribution function defined as:

$$F_{X,Y}(\lambda_1, \lambda_2) = P(\{s : X(s) \leq \lambda_1\} \bigcap \{t : Y(t) \leq \lambda_2\})$$

Joint Random Variables

- Two random variables X and Y are jointly continuous if they have a joint probability distribution function defined as:

$$F_{X,Y}(\lambda_1, \lambda_2) = P(\{s : X(s) \leq \lambda_1\} \bigcap \{t : Y(t) \leq \lambda_2\})$$

- where the joint density function is defined as:

$$f_{X,Y}(\lambda_1, \lambda_2) = \frac{\partial^2 F_{X,Y}(\lambda_1, \lambda_2)}{\partial \lambda_1, \partial \lambda_2}$$

Joint Random Variables

- Two random variables X and Y are jointly continuous if they have a joint probability distribution function defined as:

$$F_{X,Y}(\lambda_1, \lambda_2) = P(\{s : X(s) \leq \lambda_1\} \bigcap \{t : Y(t) \leq \lambda_2\})$$

- where the joint density function is defined as:

$$f_{X,Y}(\lambda_1, \lambda_2) = \frac{\partial^2 F_{X,Y}(\lambda_1, \lambda_2)}{\partial \lambda_1, \partial \lambda_2}$$

- properties of Joint density function is

1. $$f_{X,Y}(\lambda_1, \lambda_2) \geq 0 \quad \forall \quad \lambda_1 and \quad \lambda_2$$

2. $$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 = 1$$

Example

Let X and Y be two jointly continuous random variables (uniformly distributed) with joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} \lambda_1 + c\lambda_2^2, & 0 \leq \lambda_1 \leq 1, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$

1. Find c.
2. Find $P(0 < X \leq \frac{1}{2}, 0 < Y \leq \frac{1}{2})$

Example

Let X and Y be two jointly continuous random variables with the joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c\lambda_1^2 + \frac{\lambda_1 \lambda_2}{3}, & 0 \leq \lambda_1 \leq 1, 0 \leq \lambda_2 \leq 2 \\ 0, & \text{Else} \end{cases}$$

1. Find c.
2. Find $P(X + Y \geq 1)$

Marginal Probability Density Function

- We can find the marginal probability density functions of random variable X and Y from their joint density function.

$$f_X(\lambda_1) = \int_{-\infty}^{\infty} f_{X,Y}(\lambda_1, \lambda_2) d\lambda_2$$

$$f_Y(\lambda_2) = \int_{-\infty}^{\infty} f_{X,Y}(\lambda_1, \lambda_2) d\lambda_1$$

- In general, we say a set of random variables $X_1(\lambda_1) \dots X_n(\lambda_n)$ is (mutually) **independent** if

$$f_{X_1, \dots X_n}(\lambda_1 \dots \lambda_n) = \prod_{i=1}^{n} f_{Xi}(\lambda_i)$$

Marginal Probability Density Function

- We can find the marginal probability density functions of random variable X and Y from their joint density function.

$$f_X(\lambda_1) = \int_{-\infty}^{\infty} f_{X,Y}(\lambda_1, \lambda_2) d\lambda_2$$

$$f_Y(\lambda_2) = \int_{-\infty}^{\infty} f_{X,Y}(\lambda_1, \lambda_2) d\lambda_1$$

- To check if random variable X and Y are independent, it requires to show:

$$f_{XY}(\lambda_1, \lambda_2) \stackrel{?}{=} f_X(\lambda_1) * f_Y(\lambda_2)$$

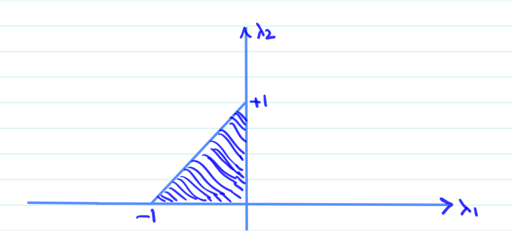
- In general, we say a set of random variables $X_1(\lambda_1) \dots X_n(\lambda_n)$ is (mutually) **independent** if

$$f_{X_1,\dots X_n}(\lambda_1 \dots \lambda_n) = \prod_{i=1}^{n} f_{Xi}(\lambda_i)$$

Example

- Let consider two random variables X and Y with the following joint probability density function:

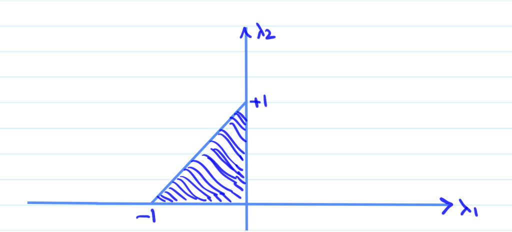
$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c, & -1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$



Example

- Let consider two random variables X and Y with the following joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c, & -1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$

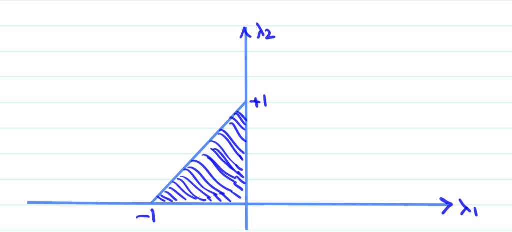


- Find the constant c.

Example

- Let consider two random variables X and Y with the following joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c, & -1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$

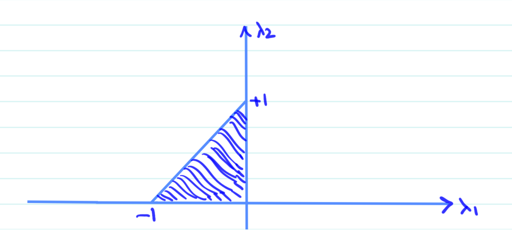


- Find the constant c.
- Find the marginal density $f_X(\lambda_1)$.

Example

- Let consider two random variables X and Y with the following joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c, & -1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$

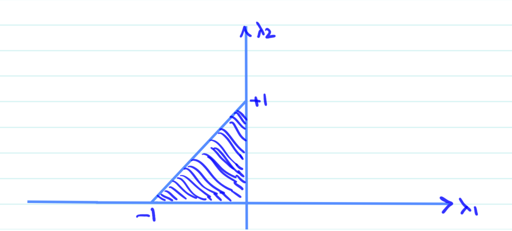


- Find the constant c.
- Find the marginal density $f_X(\lambda_1)$.
- Find the marginal density $f_Y(\lambda_2)$.

Example

- Let consider two random variables X and Y with the following joint probability density function:

$$f_{XY}(\lambda_1, \lambda_2) = \begin{cases} c, & -1 \leq \lambda_1 \leq 0, 0 \leq \lambda_2 \leq 1 \\ 0, & \text{Else} \end{cases}$$



- Find the constant c.
- Find the marginal density $f_X(\lambda_1)$.
- Find the marginal density $f_Y(\lambda_2)$.
- Are X and Y independent?

Conditional Density Function

- Let X and Y be two continuous random variables. The conditional probability density function of X given Y is defined as :

$$f_{X|Y}(\lambda_1|\lambda_2) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_Y(\lambda_2)}$$

$$P(a \leq Y \leq b|X = \lambda_1) = \int_a^b f_{Y|X}(\lambda_2|\lambda_1)d\lambda_2$$

$$f_{Y|X}(\lambda_2|\lambda_1) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_X(\lambda_1)}$$

$$P(a \leq X \leq b|Y = \lambda_2) = \int_a^b f_{X|Y}(\lambda_1|\lambda_2)d\lambda_1$$

Conditional Density Function

- Let X and Y be two continuous random variables. The conditional probability density function of X given Y is defined as :

$$f_{X|Y}(\lambda_1|\lambda_2) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_Y(\lambda_2)}$$

$$P(a \leq Y \leq b|X = \lambda_1) = \int_a^b f_{Y|X}(\lambda_2|\lambda_1)d\lambda_2$$

$$f_{Y|X}(\lambda_2|\lambda_1) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_X(\lambda_1)}$$

$$P(a \leq X \leq b|Y = \lambda_2) = \int_a^b f_{X|Y}(\lambda_1|\lambda_2)d\lambda_1$$

Conditional Density Function

- Let X and Y be two continuous random variables. The conditional probability density function of X given Y is defined as :

$$f_{X|Y}(\lambda_1|\lambda_2) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_Y(\lambda_2)}$$

$$P(a \leq Y \leq b|X = \lambda_1) = \int_a^b f_{Y|X}(\lambda_2|\lambda_1)d\lambda_2$$

$$f_{Y|X}(\lambda_2|\lambda_1) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_X(\lambda_1)}$$

$$P(a \leq X \leq b|Y = \lambda_2) = \int_a^b f_{X|Y}(\lambda_1|\lambda_2)d\lambda_1$$

Conditional Density Function

- Let X and Y be two continuous random variables. The conditional probability density function of X given Y is defined as :

$$f_{X|Y}(\lambda_1|\lambda_2) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_Y(\lambda_2)}$$

$$P(a \leq Y \leq b|X = \lambda_1) = \int_a^b f_{Y|X}(\lambda_2|\lambda_1) d\lambda_2$$

$$f_{Y|X}(\lambda_2|\lambda_1) = \frac{f_{X,Y}(\lambda_1, \lambda_2)}{f_X(\lambda_1)}$$

$$P(a \leq X \leq b|Y = \lambda_2) = \int_a^b f_{X|Y}(\lambda_1|\lambda_2) d\lambda_1$$

- For example want to know the probability density function of the outside temperature given that the humidity is known to be below 50%

Example

- Let consider a random variable X and Y with the following joint density function as follow and the plot given in the previous example. Find the conditional density function $f_{X|Y}(\lambda_1|\lambda_2)$ and $f_{Y|X}(\lambda_2|\lambda_1)$

$$f_{X,Y}(\lambda_1, \lambda_2) = \begin{cases} 2: & -1 \leq \lambda_1 \leq 0 \text{ and } 0 \leq \lambda_2 \leq 1 \\ 0: & \text{Otherwise} \end{cases}$$

Example

- R is a random variable that is equally likely to be any value between 80 and 100.
  - Find $P(90 \leq R \leq 95)$
  - Find $P(90 \leq R \leq 95 | 85 \leq R \leq 95)$

Expectation

- Expected value of a random variable X reduces density function to one number which is the mean value of

$$\mu_x = E[x] = \int_{-\infty}^{\infty} \lambda f_X(\lambda) d\lambda$$

Expectation

- Expected value of a random variable X reduces density function to one number which is the mean value of

$$\mu_x = E[x] = \int_{-\infty}^{\infty} \lambda f_X(\lambda) d\lambda$$

- Similarly the variance of random variable X is defined as :

$$\sigma_x^2 = E[(x-\mu_x)^2] = \int_{-\infty}^{\infty} (\lambda-\mu_x)^2 f_X(\lambda) d\lambda = E[x^2] - \mu_x^2$$

Expectation

- Expected value of a random variable X reduces density function to one number which is the mean value of

$$\mu_x = E[x] = \int_{-\infty}^{\infty} \lambda f_X(\lambda) d\lambda$$

- Similarly the variance of random variable X is defined as :

$$\sigma_x^2 = E[(x - \mu_x)^2] = \int_{-\infty}^{\infty} (\lambda - \mu_x)^2 f_X(\lambda) d\lambda = E[x^2] - \mu_x^2$$

- Covariance between two random variable X and Y is defined as :

$$\sigma_{xy} = E[(x - \mu_x)(y - \mu_y)]$$

Expectation

- Expected value of a random variable X reduces density function to one number which is the mean value of

$$\mu_x = E[x] = \int_{-\infty}^{\infty} \lambda f_X(\lambda) d\lambda$$

- Similarly the variance of random variable X is defined as :

$$\sigma_x^2 = E[(x-\mu_x)^2] = \int_{-\infty}^{\infty} (\lambda-\mu_x)^2 f_X(\lambda) d\lambda = E[x^2] - \mu_x^2$$

- Covariance between two random variable X and Y is defined as :

$$\sigma_{xy} = E[(x-\mu_x)(y-\mu_y)]$$

- For uncorrelated random variables X and Y, the covariance is zero hence the correlation coefficient is zero.

Example

- Let consider a continuous random variable X to be defined uniformly between 0 and 10. What is the mean of X and what is the variance?

Example

- Let consider a continuous random variable X to be defined uniformly between 0 and 10. What is the mean of X and what is the variance?
- **Answer** : $\mu_x = 5$ and $\sigma_x^2 = 8.3$