Introduction Into Probability Theory

MTH 231
Lecture 3
Chapter III

Random Variable and Probability Distributions

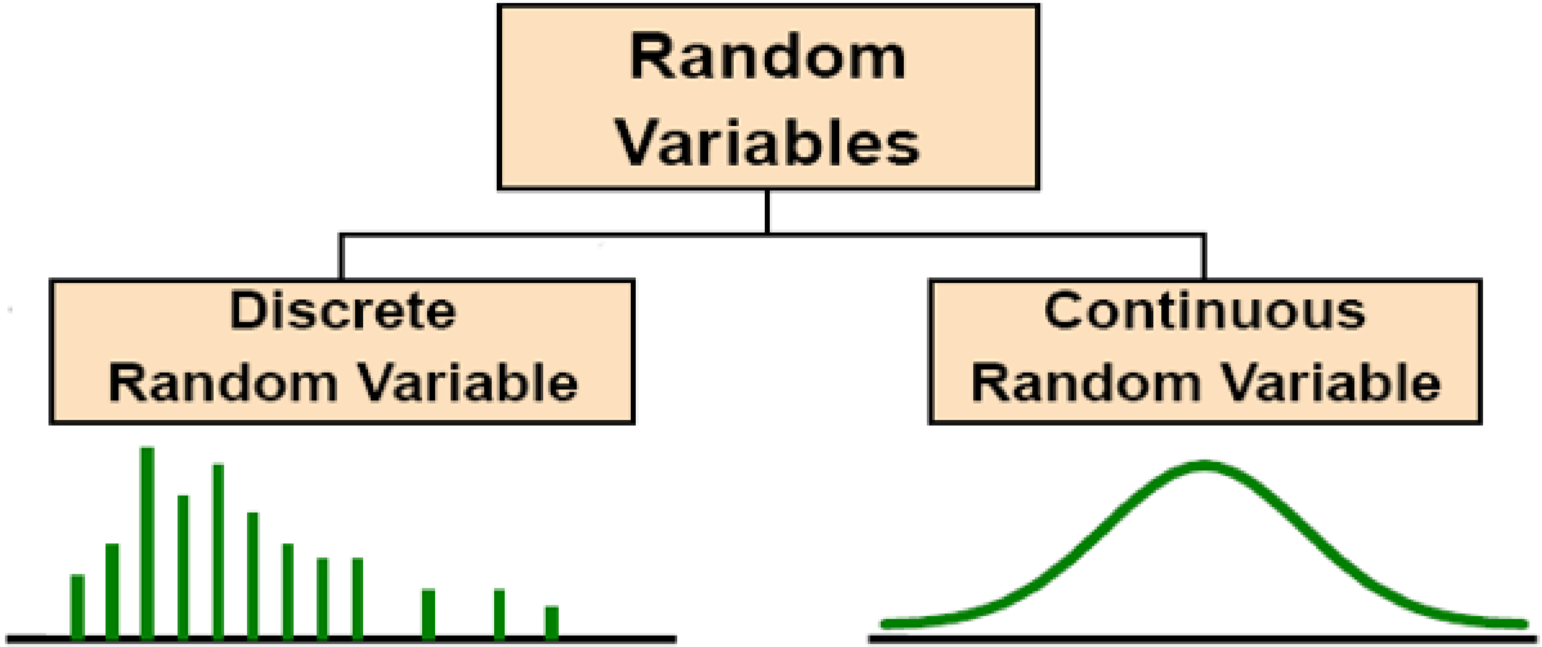
Today's lecture

- Discrete & continuous random variables
- Probability Distributions
- Calculate the 'expected value' and the variance

Discrete or Continuous?

Random Variable

- **Random Variable**
  - Represents a possible numerical value from an uncertain event
  - It is a function which associates a real number with each element in the sample space. It may be discrete or continuous.



Discrete and Continuous Data

Discrete data can only take on certain individual values.

Continuous data can take on any value in a certain range.

Example 1

Number of pages in a book is a **discrete variable.**

Example 2

Length of a film is a **continuous variable.**

Example 3

Shoe size is a **Discrete variable.** E.g. $5, 5\frac{1}{2}, 6, 6\frac{1}{2}$ etc. Not in between.

Example 4

Temperature is a **continuous variable.**

Example 5

Number of people in a race is a **discrete variable.**

Example 6

Time taken to run a race is a **continuous variable.**

Discrete probability distribution

If a variable X can assume a discrete set of values $x_1, x_2, \ldots, x_n$ with respective probabilities $P(x_1), P(x_2), \ldots, P(x_n)$, where

$$\sum_{i=1}^{n} P(x_i) = 1$$

we say that a discrete probability distribution for X has been defined .

The function $P(x_i) = P(X = x_i)$ is called the **probability function**, **probability mass function** or **probability distribution**:

$$P(x_i) \geq 0, \qquad i = 1, 2, \ldots, n$$

Properties of probability distribution

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. $P(X = x) = f(x)$.

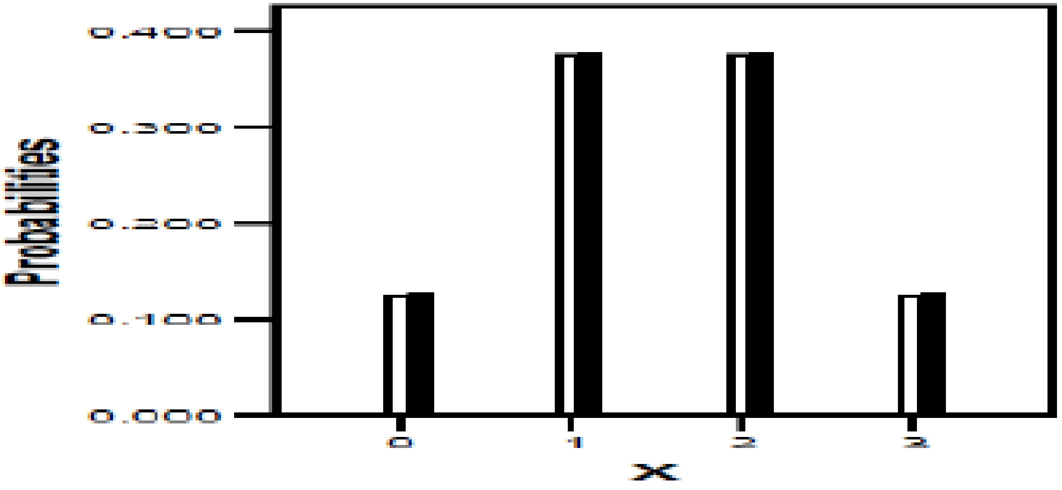
➢**Example (1)**: If X is the number of heads in three tosses of a coin. Find the probability function of the random variable.

Solution: Let H denotes a head , and T a tail of coin,
X = the number of heads obtained, then **X = 0 , 1 , 2 , 3** .
Now, for a random sample of three tosses, the following mutually exclusive events can occur with probabilities.

P(X = 0) = P(0) = P{(T, T, T)} = 1/8
P(X = 1) = P(1) = P{ (H , T , T), (T , H , T),(T , T , H)} = 3/8
P(X = 2) = P(2) = P{ (H , H , T),(H , T , H),(T , H , H)} = 3/8
P(X = 3) = P(3) = P{(H, H, H)} = 1/8.

Therefore, the probability function is

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



Graph of P(x)

Example (2): Suppose that X, the score on the uppermost face of a loaded die has a probability function

$$P(x) = k\,x\,; \qquad x = 1, 2, 3, 4, 5, 6.$$

(a) Find k,

(b) write P(x) in tabular form, and find the probability that $X \geq 3$.

Solution:

(a) Since $\sum_{x=1}^{6} P(x) = 1$ then, $k + 2k + 3k + \ldots + 6k = 1$ so, $21k = 1$

then, $k = 1/21$, $x = 1, 2, 3,4,5,6.$

Then,

$$P(x) = \frac{x}{21} \;; \qquad x = 1, 2, 3, 4, 5, 6$$

X	1	2	3	4	5	6
P(x) = P(X = x)	1/21	2/21	3/21	4/21	5/21	6/21

(b) $$P(X \geq 3) = \frac{3}{21} + \frac{4}{21} + \frac{5}{21} + \frac{6}{21} = \frac{6}{7}$$

The Cumulative Distribution of the Discrete Random Variable

- The cumulative distribution F(x) of the discrete random variable X with probability distribution P(x) is

$$F(x) = P(X \le x) = \sum_{t \le x} P(t), -\infty < x < \infty$$

- Properties of F(x)

  - $F(-\infty) = 0,$
  - $F(\infty) = 1$

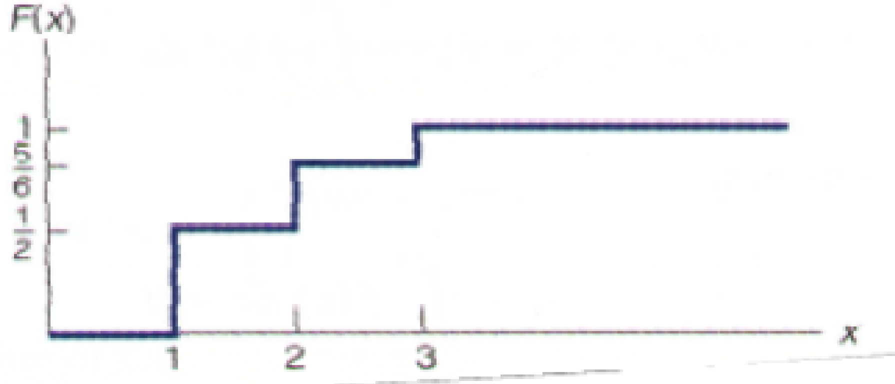
The Cumulative Distribution of the Discrete Random Variable

Note that: If X is a discrete random variable whose set of possible values are $x_1, x_2, x_3, \ldots$, where $x_1 < x_2 < x_3 < \ldots$ then, its distribution function F is a *step function* . That is, the value of F is a constant in the interval $[x_{i-1}, x_i)$ and then takes a step (or jump) of size $P(x_i)$ at x_i .
Then the cumulative distribution function F of X is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{2} & 1 \le x < 2 \\ \frac{5}{6} & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

X	1	2	3
P(X = x)=**P(x)**	1/2	1/3	1/6
F(x)	$\frac{1}{2}$	$\frac{1}{2}+\frac{1}{3}=\frac{5}{6}$	$\frac{1}{2}+\frac{1}{3}+\frac{1}{6}=1$

This is graphically presented in the following figure .



Graph of F(x)

Example

Rolling a Die

Construct a probability distribution for rolling a single die.

Solution

Since the sample space is 1, 2, 3, 4, 5, 6 and each outcome has a probability of $\frac{1}{6}$, the distribution is as shown.

Outcome X	1	2	3	4	5	6
Probability $P(X)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Example

- Let X be the number of heads when 3 coins are tossed. Find the probability distribution function and the following probabilities

  i) The probability of obtain at least 2 heads .

  ii) The probability of obtain no heads.

  iii) The cumulative distribution function .

- Solution

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

i) $P(X \geq 2) = P(x = 2) + P(x = 3)$

=3/8 +1/8 = 4/8 .

ii) $P(X = 0) = 1/8$.

iii) F(X)=P(X$\leq x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{4}{8}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & 3 \leq x \end{cases}$

Continuous Random Variables

It is a random variable that can assume any value in some interval of real numbers. A continuous random variable has probability of zero of assuming exactly any of its values. It follows that

$$P\,(a \le X \le b) = P(X = a) + P\,(a < X < b) + P(X = b) = P\,(a < X < b)$$

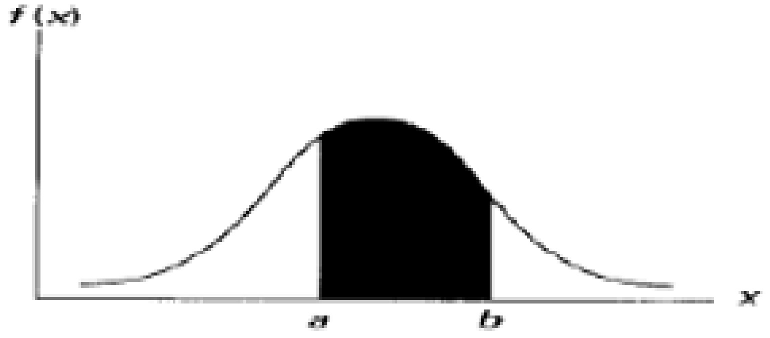
since,

$$P(X = a) = P(X = b) = 0.$$

With continuous variables, **f(x)** is usually called the **density function**.
A probability density function is nonnegative and constructed so that the area under its curve bounded by the x axis is equal to 1 when computed over the range of X for which f(x) is defined.

In the following figure, the probability that X assumes a value between a and b is equal to the shaded area under the curve of the density function between x = a and x = b, and from integral calculus is given by

$$P(a < X < b) = \int_a^b f(x)\,dx$$



P (a < X < b)

Definition (1): The function f(x) is a probability density function (p.d.f.) for the continuous random variable X, defined over the set of real numbers R, if:

$$(1)\ f(x) \geq 0\text{, for all } x$$

$$(2)\ \int_{-\infty}^{\infty} f(x)\,dx = 1$$

Probability Density Function Properties

- Definition:

For a continuous random variable X, a **probability density function** is a function such that

(1) $f(x) \geq 0$

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

(3) $P(a \leq X \leq b) = \int_{a}^{b} f(x)\,dx =$ area under $f(x)$ from a to b

for any a and b (4-1)

➢**Example (1)**: Suppose that X is a continuous random variable whose density function is given by

$$f(x)=\begin{cases} C(4x-2x^2) & 0<x<2 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is the value of C

(b) Find P(X > 1)

Solution:

(a) Since, f(x) is a p.d.f., we must have that $\int_{-\infty}^{\infty} f(x)dx = 1$

$$C\int_0^2 (4x-2x^2)dx = 1 \Longrightarrow C\left[2x^2-\frac{2x^3}{3}\right]_{x=0}^{x=2} = 1 \Longrightarrow C=\frac{3}{8}$$

(b) Find $P(X>1) = \int_1^{\infty} f(x)dx = \frac{3}{8}\int_1^2 (4x-2x^2)dx = \frac{1}{2}$

➢**Example (2)**: The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that:

(a) a computer will function between 50 and 100 hours before breaking down;

(b) it will function less than 100 hours?

Solution:

(a) Since

$$1 = \int_{-\infty}^{\infty} f(x)dx = \lambda \int_{0}^{\infty} e^{-x/100} dx \Rightarrow 1 = \lambda(-100)e^{-x/100}\Big|_0^{\infty} = -100\lambda(0-1) = 100\lambda$$

$$\Rightarrow \quad \lambda = \frac{1}{100}$$

Hence , the probability that a computer will function between 50 and 150 hours before breaking down is given by

$$P(50 < X < 150) = \int_{50}^{100} \frac{1}{100} e^{-x/100} dx \quad = -e^{-x/100}\Big|_{50}^{100} \quad = e^{-1/2} - e^{-3/2}$$

$$\approx 0.384$$

(b) Similarly

$$P(X < 100) = \int_{0}^{100} \frac{1}{100} e^{-x/100} dx = -e^{-x/100}\Big|_{0}^{100} = 1 - e^{-1} \approx 0.633$$

In other words, approximately 63.3 percent of the time a computer will fail before registering 100 hours of use.

The Cumulative Distribution of the Continuous Random Variable

- The cumulative distribution F(x) of the continuous random variable X with probability distribution P(x) is

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

- Properties of F(x)
  - $F(-\infty) = 0,$
  - $F(\infty) = 1,$
  - $f(x) = \frac{dF(x)}{dx}, when\ F(x) given\ ,$
  - $P(a < X < b) = F(b) - F(a).$

Example

➢ Consider the density function

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

I) Evaluate K.

II) Find F(x) and use it to evaluate $P(0.3 < X < 0.6)$.

➢ **Solution**

i) $\int_{-\infty}^{\infty} f(x)dx = 1, \quad \int_0^1 k\sqrt{x}dx = 1, \; then \quad k = \frac{3}{2}.$

ii) $F(x) = \int_{-\infty}^{x} f(t)dt = \int_0^x f(t)dt = x^{3/2},$

$P(0.3 < X < 0.6) = F(0.6) - F(0.3) = 0.6^{3/2} - 0.3^{3/2}.$

Example (4):

Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

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

.

(a) Verify that $f(x)$ is a density function.

(b) Find $P(0 < X \leq 1)$.

(c) find $F(x)$, and use it to evaluate $P(0 < X \leq 1)$.

Solution

(a) Obviously, $f(x) \geq 0$. To verify condition 2 in Definition 3.6, we have

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9}\Big|_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1.$$

(b)

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \frac{x^3}{9}\bigg|_0^1 = \frac{1}{9}.$$

For $-1 < x < 2$,

(c)

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-1}^{x} \frac{t^2}{3} dt = \frac{t^3}{9}\bigg|_{-1}^{x} = \frac{x^3 + 1}{9}.$$

Therefore,

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{x^3+1}{9}, & -1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9},$$

Questions!

We must become more comfortable with probability and uncertainty