WEEK 1 : TUTORIAL 1

Probability Basics (2)

Random variable

A random variable is a function $X : \Omega \to R$ i.e., it is a function from the sample space to the real numbers.

Example:

- Sum of outcomes on rolling 3 dice
- No. of heads observed when tossing a fair coin 3 times

Induced Probability Function

Consider the previous example of tossing a fair coin 3 times. Let X be the no. of heads obtained in the three tosses. Enumerating the elementary outcomes, we observe the value of X as

ω	HHH	HHT	HTH	THH	TTH	THT	HTT	TTT
$X(\omega)$	3	2	2	2	1	1	1	0

Instead of using the probability measure defined on the elementary outcomes or events, we should ideally like to measure the probability of the random variable taking on values in its range

x	0	1	2	3
$P_X(X=x)$	1/8	3/8	3/8	1/8

Let $\Omega = \{\omega_1, \omega_2, \dots\}$ be a sample space and P be a probability measure (function)

Let X be a random variable with range $\mathcal{X} = \{x_1, x_2 \cdots x_m\}$

We define the induced probability function P_X on $\mathcal{X}$ as

$$P_X(X = x_i) = P(\{\omega_j \in \Omega : X(\omega_j) = x_i\})$$

Cumulative Distribution Function

The cdf of a random variable X, denoted by $F_X(x)$ is defined by

$$F_X(x) = P_X(X \leq x), \text{ for all } x$$

Example

x	$(-\infty, 0]$	$(-\infty, 1]$	$(-\infty, 2]$	$(-\infty, 3]$	$(-\infty, \infty)$
$F_X(x)$	$1/8$	$1/2$	$7/8$	1	1

Properties of cdf

A function $F_X(x)$ is a cdf iff the following three conditions hold:

- (Monotonicity) If $x \leq y$, then $F_X(x) \leq F_X(y)$
- (Limiting Values) $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$
- (Right-continuity) For every x, we have $\lim_{y \downarrow x} F_X(y) = F_X(x)$

Continuous and Discrete Random Variable

Random variable X is continuous if $F_X(x)$ is a continuous function of x

Random variable X is discrete if $F_X(x)$ is a step function of x

Probability Mass Function (pmf)

The pmf of a discrete random variable X is given by

$$f_X(x) = P(X=x), \text{ for all } x$$

Example:

For a geometric random variable X with parameter p,

$$f_X(x) = \begin{cases} (1-p)^{x-1}p & \text{for } x = 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Properties:

- $f_X(x) \geq 0$, for all x
- $\sum_x f_X(x) = 1$

Probability Density Function (pdf)

The pdf of a continuous random variable is the function $f_X(x)$ which satisfies

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt, \text{ for all } x$$

Properties:

- $f_X(x) \geq 0$, for all x
- $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

Expectation

The expected value or mean of a r.v. X, denoted by $E[X]$, is given by

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{(continuous RV)}$$

$$E[X] = \sum_{x:P(x)>0} x f_X(x)$$

$$= \sum_{x:P(x)>0} x\, P(X=x) \quad \text{(discrete RV)}$$

Properties of Expectation

Let X be a r.v. and let a, b, c be constants. Then, for functions $g_1(X)$ and $g_2(X)$ whose expectations exists

- $E(a g_1(X) + b g_2(X) + c) = a E g_1(X) + b E g_2(X) + c$
- If $g_1(X) \geq 0$ for all X, then $E g_1(X) \geq 0$
- If $g_1(X) \geq g_2(X)$ for all X, then $E g_1(X) \geq E g_2(X)$
- If $a \leq g_1(X) \leq b$ for all X, then $a \leq E g_1(X) \leq b$

Moments

For each integer n, the n^{th} moment of X is

$$\mu'_n = E X^n$$

The n^{th} central moment of X is

$$\mu_n = E(X-\mu)^n$$

Variance

The variance of a r.v. X is its second central moment

$$\text{Var}\, X = E(X-\mu)^2 = E(X-EX)^2 = EX^2 - (EX)^2$$

The +ve square root of $\text{Var}\, X$ is the standard deviation of X

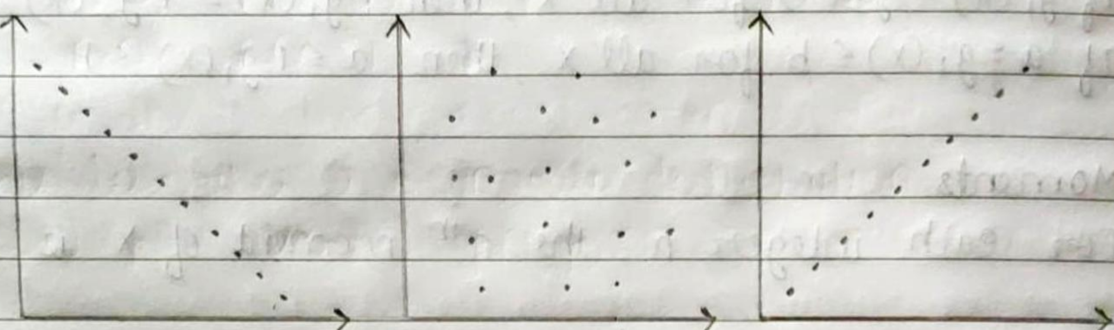
Note: $Var(aX+b) = a^2 \, Var X$
where a, b are constants

Covariance

The covariance of two r.v. X and Y is

$$cov(X,Y) = E[(X-EX)(Y-EY)]$$

It is a measure of how much two r.v. change together.



Large −ve c.v. Neare zero c.v. Large +ve c.v.

Correlation

The correlation of two r.v. X and Y is

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)\,var(Y)}}$$

Note:

- For correlation to be defined, indivisual variances must be non-zero and finite.
- $\rho(X,Y)$ lies b/w -1 and $+1$

Probability Distributions

Consider two variables X and Y, and suppose we know the corresponding pmf f_X and f_Y

Can we answer the following question:

$$P(X = x \text{ and } Y = y) = ?$$

Joint Distributions

To capture the properties of two r.v. X and Y, we use the joint PMF

$$f_{X,Y} : R^2 \to [0, 1], \text{ defined by}$$
$$f_{X,Y}(x, y) = P(X = x, Y = y)$$

Marginal Distributions

Suppose we have the joint PMF

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

From this joint PMF, we can obtain the PMF's of the two r.v.

$f_X = \Sigma_y f_{X,Y}(x, y)$ (marginal PMF of R.V. X)
$f_Y = \Sigma_x f_{X,Y}(x, y)$ (marginal PMF of R.V. Y)

Conditional Distributions

Like joint distributions, we can also consider conditional distributions

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

Using conditional probability definition, we have

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

Note that the above conditional probability is undefined if $f_Y(y) = 0$

Bernoulli Distribution

Consider a R.V. X taking one of two possible values (either 0 or 1). Let the PMF ~~x~~ of X be given by

$$f_X(0) = P(X=0) = 1-p \qquad (0 \le p \le 1)$$
$$f_X(1) = P(X=1) = p$$

This describes a Bernoulli distribution

$E[X] = p \qquad Var[X] = p(1-p)$

Binomial Distribution

Consider the situation where we perform n independent Bernoulli trials where

- probability of success (for each trial) = p
- probability of failure = 1-p

Let X be the no. of success in the n trials, then we have

$$P(X = x|n, p) = \binom{n}{x} p^x (1-p)^{n-x} \quad ; 0 \le x \le n$$

$E[X] = np$

$Var(X) = np(1-p)$

Geometric Distribution

Suppose we perform a series of independent Bernoulli trials, each with a probability p of success. Let X represent the no. of trials before the first success, then we have

$$P(X = x|p) = (1-p)^{x-1} p \qquad x = 1, 2, 3 \ldots$$

$$E[X] = \frac{1}{p} \qquad Var(X) = \frac{(1-p)}{p^2}$$

Uniform Distribution

A continuous R.V. X is said to be uniformly distributed on an interval [a, b] if its PDF is given by

$$f_X(x|a,b) = \begin{cases} \frac{1}{b-a} & \text{if } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \frac{(a+b)}{2} \qquad Var(X) = \frac{(b-a)^2}{12}$$

Normal Distribution

A continuous RV X is said to be normally distributed with parameters μ and σ^2 if the density of X is

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

Importance of Normal Distribution

Roughly, the central limit theorem states that the distribution of the sum (or average) of a large no of independent, identically distributed variables will be approximately normal, regardless of the underlying distribution.

Multivariate Normal Distribution

$$N(x|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^D |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

where,

- μ is the D-dimensional mean vector
- Σ is the $D \times D$ covariance matrix
- $|\Sigma|$ is the determinant of the covariance matrix.

Beta Distribution

The pdf of the beta distribution in the range $[0, 1]$, with shape parameters α, β is given by

$$f(x|\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

where the gamma function is an extension of the factorial function

$$E[X] = \frac{\alpha}{(\alpha+\beta)} \qquad Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$