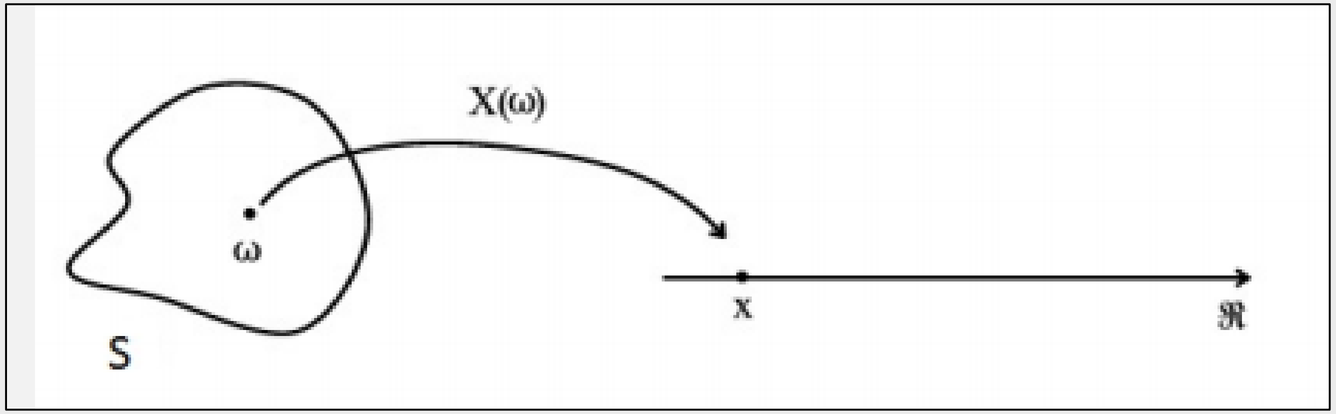
5. RANDOM VARIABLES & PROBABILITY DISTRIBUTIONS [IT2110]

By SLIIT Mathematics Unit

Faculty of Humanities and Sciences

RANDOM VARIABLES

- A random variable (r.v.) X is a function defined on a sample space (S), that associates a real number, $X(\omega) = x$, with each outcome ω in S.
- ***Simple Definition :*** A random variable is a mapping between all the outcomes in the sample space with a set of real numbers.



- Random variables are denoted by using capital letters

Random Variables

- Discrete random variables
- Continuous random variables

Discrete Random Variables

- A random variable is said to be discrete, if it can assume only ***distinct*** values.

OR

- In other words, it can assume only countable number of values.

Examples

- Toss a coin 5 times. Let X be the number of heads appeared. Then,

  X - 0,1,2,3,4,5

- Roll a die twice. Let X be the number of times 4 comes up.

  X - 0,1,2

- Suppose we toss two coins. Assume that all the outcomes are equally likely (fair coins). Let Y be the number of heads appeared. Then,

  Y - 0,1,2

PROBABILITY DISTRIBUTIONS

- The set of all ordered pairs (x, P r(X = x)) of a discrete r.v. (X), is known as the probability distribution
- This is also known as the ***probability mass function*** (p.m.f.) and is denoted by $P_X(x)$.
- ***Simple Definition:*** All the possible values of a discrete random variable with their corresponding probability values is known as probability distribution.

Properties

- $P_X(x)$ refers to P(X=x).
- The probability distribution function is always non-negative.
- $\sum_{all\ x} P_X(x) = 1$
- The cumulative distribution function (c.d.f.) F of the random variable X is defined by

$$F_X(x) = \Pr(X \leq x)$$

Example

- Suppose we toss two coins. Assume that all the outcomes are equally likely (fair coins). Let Y be the number of heads appeared. Then,

Y	**0**	**1**	**2**
Pr(Y=y)	1/4	2/4	1/4
$F_Y(y)$ = Pr(Y≤y)	1/4	3/4	1

Expected Value & Variance

- This is same as mean of the random variable.
- Let X be a discrete random variable with p.m.f. P_X(x). Then the expected value of X, denoted by E(X), is defined by

$$\boldsymbol{E(X) = \sum_{all\, x} x * Pr(X = x)}$$

- The variance of a random variable X is defined by

$$\boldsymbol{V(X) = E(X - E(X))^2 = E\left(X^2\right) - [E(X)]^2}$$

Properties [E(X)]

Let X & Y be two random variables. Then,

- $E(c) = c$
- $E[g(X)] = \sum_{all\ x} g(x) * Pr(X = x)$
- $E[g(X)+c] = E[g(X)] + c$
- $E[c*g(X)] = c*E[g(X)]$
- $E[X+Y] = E[X] + E[Y]$

Properties [V(X)]

Let X & Y be two random variables. Then,

- $V(c) = 0$
- $V[g(X)+c] = V[g(X)]$
- $V[c*g(X)] = c^2 *V[g(X)]$
- $V[X+Y] = V[X] + V[Y] + 2Cov(X,Y)$
- $V[X-Y] = V[X] + V[Y] - 2Cov(X,Y)$
- *If X & Y are independent then,* $Cov(X,Y) = 0$

Covariance

- Covariance is a measure of how the changes in one variable are associated with the changes in second variable.

$$\boldsymbol{Cov(X,Y) = \sum_{i=1}^{N} [X_i - E(X)][Y_i - E(Y)]P(X_iY_i)}$$

Examples

1. Suppose we toss two coins. Assume that all the outcomes are equally likely (fair coins). Let Y be the number of heads appeared. Find E(Y) and Var(Y).

Y	**0**	**1**	**2**
Pr(Y=y)	0.25	0.5	0.25

E(Y) = $\sum_{all\ y} \boldsymbol{y * Pr(Y = y)}$ = (0*0.25) + (1*0.5) + (2*0.25) = 1

E(Y^2) = $\sum_{all\ y} \boldsymbol{y^2 * Pr(Y = y)}$ = (0*0.25) + (1*0.5) + (4*0.25) = 1.5

Var(Y) = E(Y^2) - $[E(Y)]^2$ = 1.5 - 1^2 = 0.5

Examples

2. To find out the prevalence of smallpox vaccine use, a researcher inquired into the number of times a randomly selected 200 people aged 16 and over in an African village had been vaccinated. He obtained the following figures: never, 16 people; once, 30; twice, 58; three times, 51; four times, 38; five times, 7. Assuming these proportions continue to hold exhaustively for the population of that village, what is the expected number of times those people in the village had been vaccinated, and what is the standard deviation?

3. Let X and Y be two independent random variables. Suppose that we know $Var(2X-Y)=6$ and $Var(X+2Y)=9$. Find $Var(X)$ and $Var(Y)$.

Discrete Probability Distributions

Discrete Probability Distributions

- *Bernoulli Distribution*
- *Binomial Distribution*
- *Poisson Distribution*

Conditions for Discrete Distributions

Bernoulli Distribution	**Binomial Distribution**	**Poisson Distribution**
Only two possible outcomes (Success & Failure)	For each trial, only two possible outcomes (Success & Failure)	For each trial, only two possible outcomes (Success & Failure)
Only one trial	No of trials (n) are fixed	Trials (n) is large
	Probability of success (p) is constant for each & every trial	The occurrences are independent of each other
	Trials are independent	***[Assume that the numbers of occurrences in disjoint Intervals]***
Eg:- Tossing a coin once	*Eg:- Tossing a coin 10 (n) times*	*Eg:- Number of defects in a lot*

Discrete Distributions

Bernoulli Distribution	Binomial Distribution	Poisson Distribution
X – Getting the success	***X*** - The number of successes in n number of trials.	***X*** - The number of occurrences for a given rate of occurrence (λ)
X ~ Bernoulli(p)	***X ~ Bin(n,p)***	***X ~ Poisson(λ)***
$\boldsymbol{P_X(x) = p^x(1-p)^{1-x}}$ ***[p.m.f.]***	$\boldsymbol{P_X(x) = \binom{n}{x} p^x(1-p)^{n-x}}$;x = 0,1,2,…,n ***[p.m.f.]***	$\boldsymbol{P_X(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}}$;x = 0,1,2,… ***[p.m.f.]***
E(X) = p	***E(X) = np***	***E(X) = λ***
V(X) = p(1-p)	***V(X) = np(1-p)***	***V(X) = λ***

Binomial Distributions

- An expansion of the Bernoulli distribution.
- Each trial has a Bernoulli distribution.

Examples

1) It is known that screws produced by a certain machine will be defective with probability 0.01 independently of each other. If we randomly pick 10 screws produced by this machine, what is the probability that

   a) exactly six screws will be defective?

   b) at most 3 screws will be defective?

   c) at least 2 screws will be defective?

   d) What is the expected number of defectives?

   e) What is the variance of defectives?

Examples

2) Fifty seeds were planted and it is known that the probability of any seed germinating is 0.4. Assuming that the seeds have no other factors in germinating, find the following probabilities.

   a) More than 12 seeds germinate.

   b) More than 15 but fewer than 30 seeds germinate.

Poisson Distribution - Example

1) Suppose that, on average, in every two pages of a book there is one typographical error. What is the probability of at least one error on a certain page of the book?

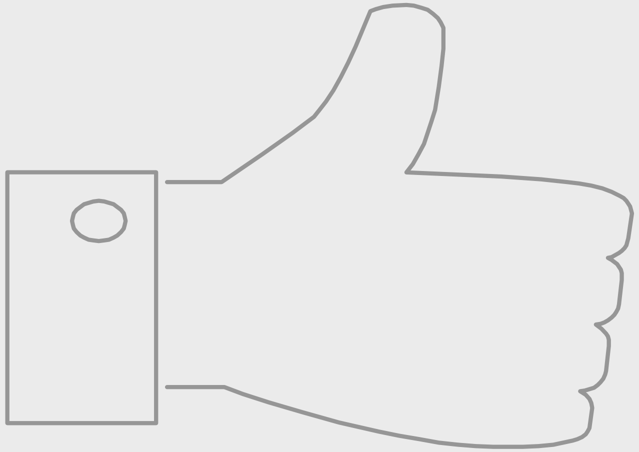
Poisson Approximation

- If $X \sim Bin(n, p)$, then X can be approximated with a Poisson distribution with the rate parameter (λ) being equal to np if p is quite small and n is large.

- Usually this approximation can be used if ***$p < 0.1$ and $n > 50$.***

Example

1) If the probability that an individual suffers an adverse reaction from a particular drug is known to be 0.001, determine the probability that out of 2000 individuals, (a) exactly three and (b) more than two individuals will suffer an adverse reaction.



THANKS!

Any questions?