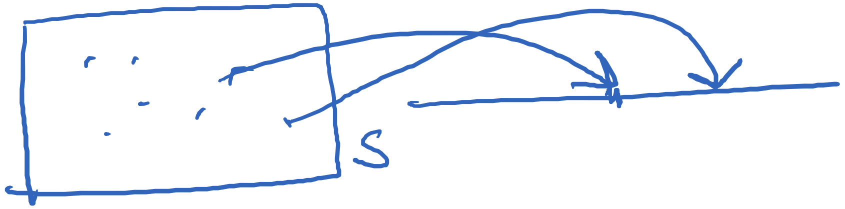
QUANTITATIVE TECHNIQUES FOR MANAGERIAL DECISION - 1 (QTMD1G21-1)

RANDOM EXPERIMENT: TOSSING A COIN TWICE

$$S = \{\overset{2}{HH}, \overset{1}{HT}, \overset{1}{TH}, \overset{0}{TT}\}$$



X = # of heads

$x = 0, 1, 2 \longrightarrow$ Probabilities ??

$$\boxed{P(X=0) = P(\text{No heads}) = \frac{1}{4}}$$

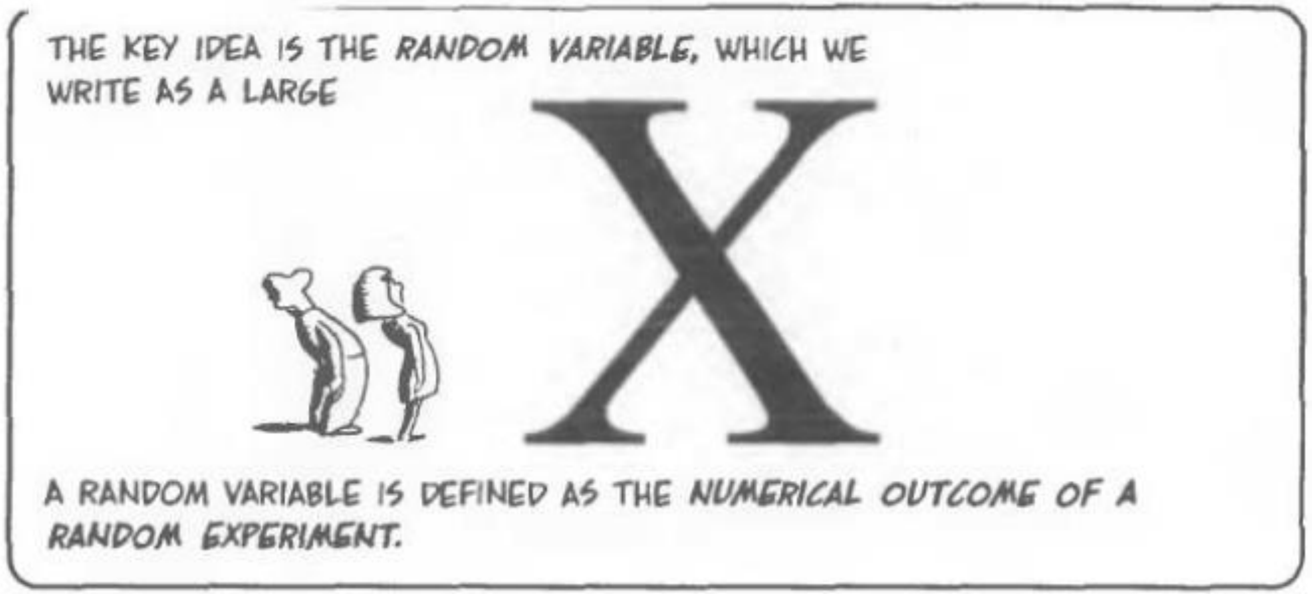
X	P(X=x)
$x \to 0$	1/4
$x \to 1$	1/2
$x \to 2$	1/4

Random variable is a variable quantity whose value is determined by what happens in a chance experiment.

RANDOM VARIABLES

Random variable: Response of random experiments taking different numerical values with certain probabilities.

The probability models describe the random variables.



EXAMPLES

- Number of heads in 1000 coin tosses.
- Number of days it will rain at Jamshedpur in the next month
- Number of errors in each page of a text book.
- The temperature (in Centigrade) of a cup of coffee served at Dhaba
- The time (in seconds) that a customer in a store must wait to receive a credit card authorization
- The quantity of juice (in ml) that is poured in the cups by the fruit seller at the fruit shop

DISCRETE RANDOM VARIABLES

DISCRETE RANDOM VARIABLES

Discrete Random Variable: A random variable whose possible values are either a finite set or a countably infinite set.

Examples:

- Number of days it will rain at Jamshedpur in the next month.
- Number of errors in each page of a text book.
- Number of earthquakes taking place in Jharkhand in the last 100 years.
- Reviewing a single mortgage application and deciding whether the application gets approved(A) or denied(D).
- Number of tails appearing until a head appears in a coin toss

$X = \#$ of errors in 1 pg

$X = 0, 1, 2, 3, \ldots$

$S = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$

$X = \#$ of tails appearing ....

$= 0, 1, 2, \ldots$

PROBABILITY MASS FUNCTION(PMF)

X : Random Variable

Takes values $x_1, x_2, \ldots, x_n$

X	P(X=x)
0	1/4
1	1/2
2	1/4

Probability mass function (p.m.f.):

$$p(x_i) = P(X = x_i)$$

Properties:

$$0 \le P(X = x_i) \le 1$$

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

Value X	Probability $P(X=x_i)$
x_1	$p(x_1)$
x_2	$p(x_2)$
$\vdots$	$\vdots$
x_n	$p(x_n)$
Total	$\sum_{i=1}^{n} p(x_i) = 1$

PROBLEM

An experiment consists of tossing a fair coin three times. Let X be the number of heads following the first tail.

→ r.v.

What is the distribution/p.m.f. of the random variable X.

X = # of heads following the first tail .

Sample space	X
HHH	0
HHT	0
HTH	1
THH	2
TTH	0
THT	1
HTT	0
TTT	0

X = 0, 1, 2 .

Distribution/p.m.f of X ←

X	$P(X=x_i)$
0	5/8
1	2/8
2	1/8

CUMULATIVE DISTRIBUTION FUNCTION

- A random variable X can be characterised by its Cumulative Distribution Function (c.d.f.)

$$F(X) = P(X \leq x)$$

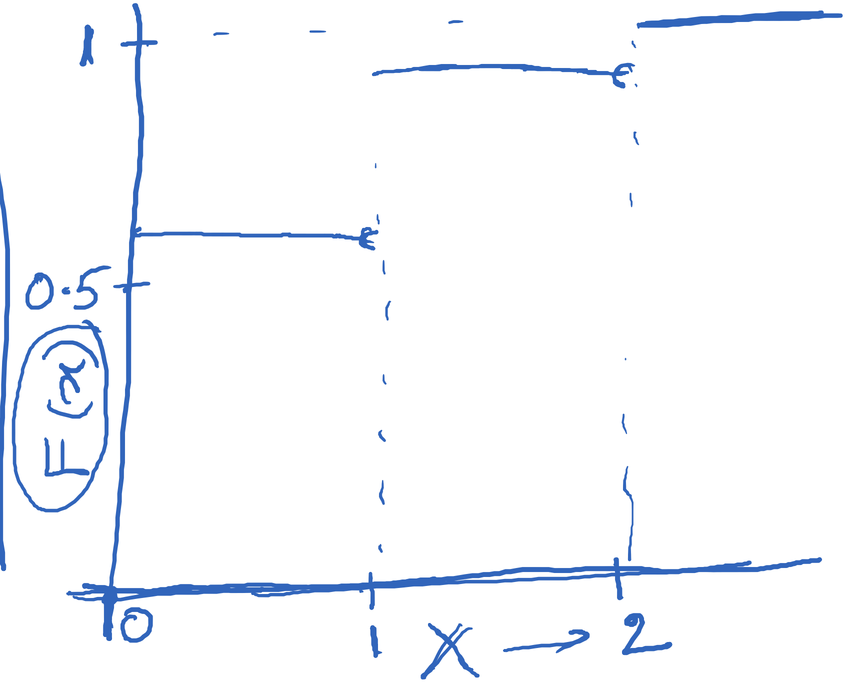
- Distribution Functions tells
  - Number of accidents are below a certain margin.
  - Rainfall is below a certain amount.
  - Salary is above a certain threshold.
  - The patient survives more than some stipulated time limit. (Survival Function = $1 - F(X)$)

PROBLEM

An experiment consists of tossing a fair coin three times. Let X be the number of heads following the first tail.

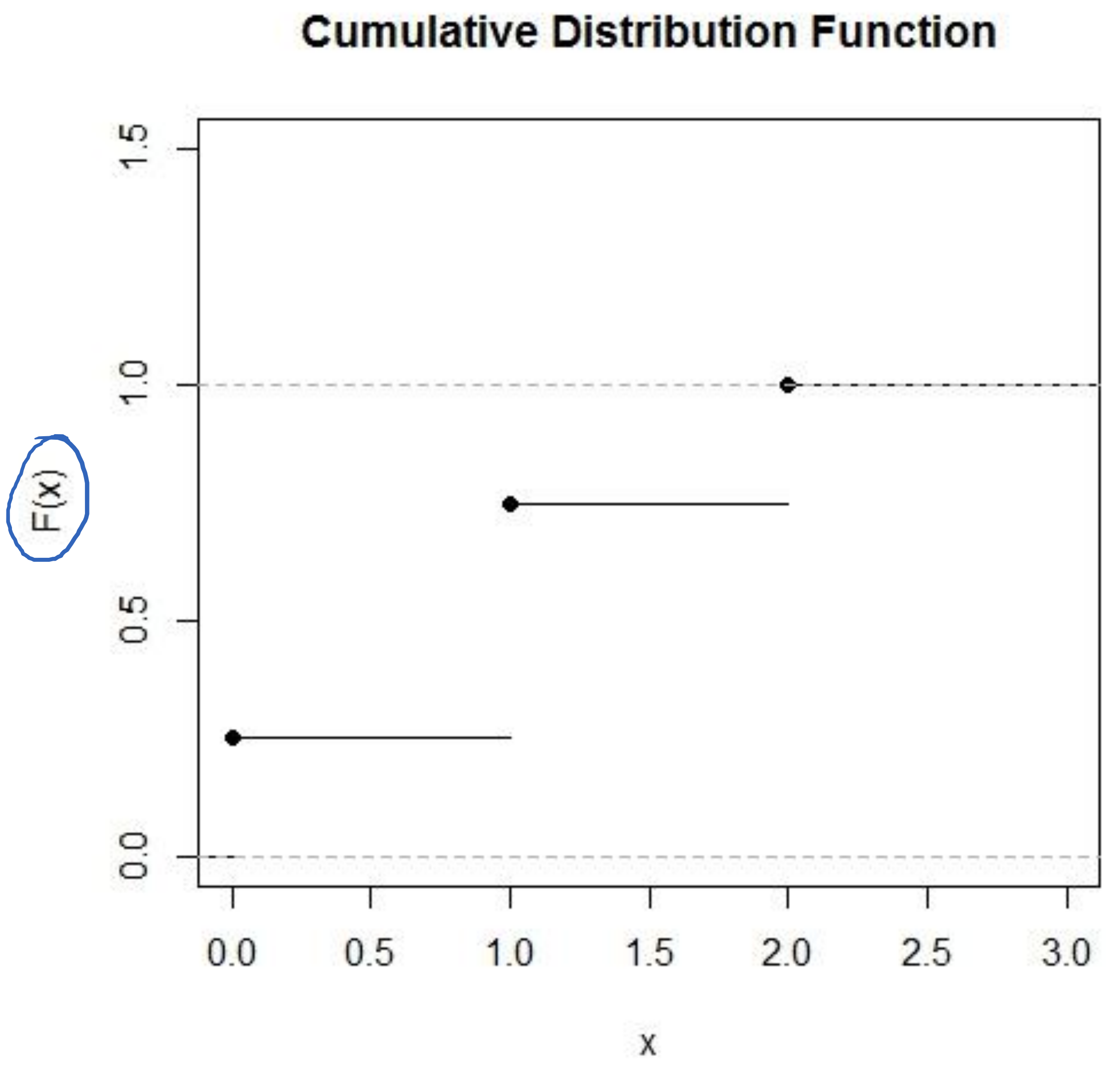
Find the (cumulative) distribution function.

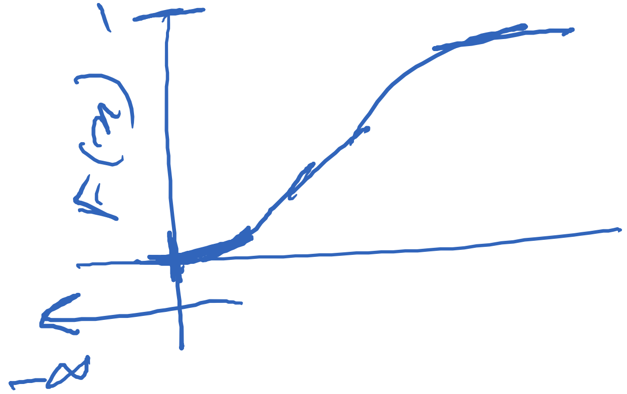
X	$P(X=x_i)$	$F(x)=P(X\le x)$
0	$5/8$	$F(0)=P(X\le 0)=5/8$
1	$2/8$	$F(1)=P(X\le 1)=P(X=0)+P(X=1)=\frac{7}{8}$
2	$1/8$	$F(2)=P(X\le 2)=1$



HOW DOES A CDF LOOK LIKE?

For a Discrete Random Variable, the CDF looks like a step function



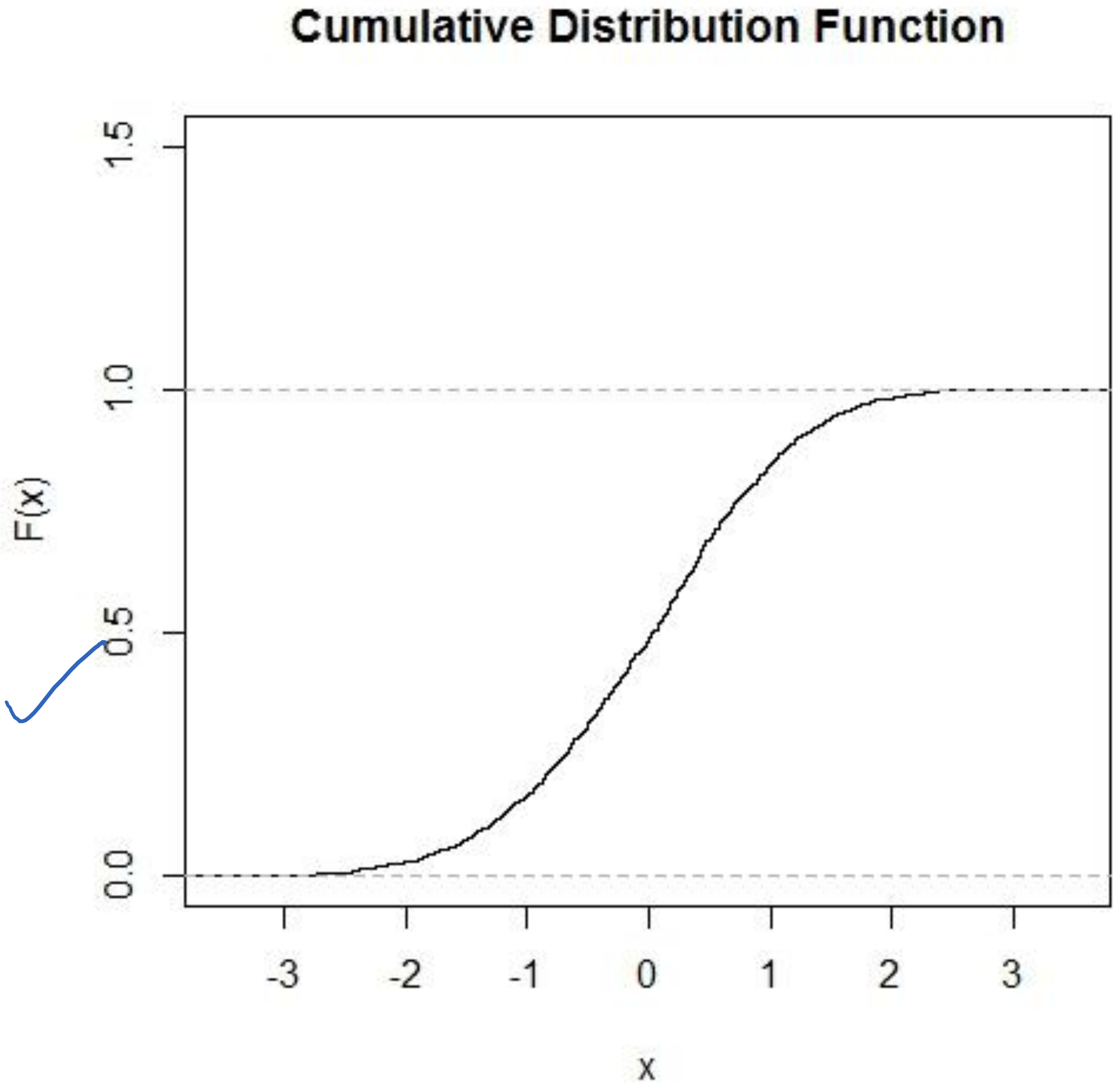


PROPERTIES OF CDF

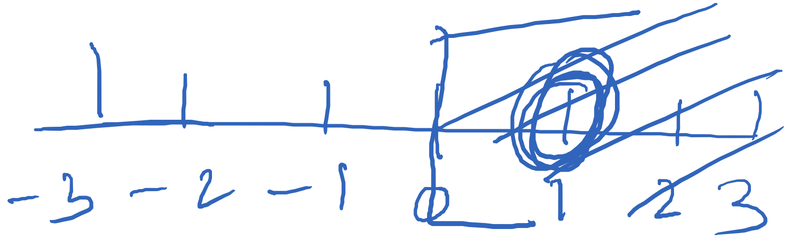
1. $0 \leq F(x) \leq 1$
2. F(x) is non-decreasing
3. F(x) is right continuous
4. $\lim_{x \to -\infty} F(x) = 0, \lim_{x \to \infty} F(x) = 1,$

HOW DOES A CDF LOOK LIKE?

CDF for a Continuous Random Variable



PROBLEM TEXTBOOK 4.10 PG 163



R.V. X = winnings of a gambler

Let X be the winnings of a gambler. Let p(i) = P(X=i) and suppose that,

$p(0) = {}^{1}/_{3}\,,\; p(1) = p(-1) = {}^{13}/_{55}\,,\; p(2) = p(-2) = {}^{1}/_{11}\,,$

$p(3) = p(-3) = {}^{1}/_{165}$

Compute the conditional probability that the gambler wins i, i =1, 2, 3, given that he wins a positive amount.

X	0	1	-1	2	-2	3	-3
$P(X=x)$	$1/3$	$13/55$	$13/55$	$1/11$	$1/11$	$1/165$	$1/165$

$$P(X>0) = P(X=1) + P(X=2) + P(X=3) = \frac{13}{55} + \frac{1}{11} + \frac{1}{165} = \frac{55}{165}$$

$$P(X=1 \mid X>0) = \frac{P(X=1 \cap X>0)}{P(X>0)} = P(X=1)/P(X>0) = \frac{13/55}{55/165}$$

EXPECTATION

Expectation of a random variable is the long-run average, when computed over a very large number of trials.

Expectation can theoretically be measured through the weighted average of the values of the random variable, weighted by the corresponding probabilities.

Consider a discrete random variable X taking values $\mathbf{x_1, x_2, \ldots, x_n}$ with probabilities $\mathbf{p(x_1), p(x_2), \ldots , p(x_n)}$ respectively.

The expectation (or mean/average) of X is $E(X) = \sum_{i=1}^{n} x_i p(x_i)$.

X	$P(X=x_i)$
$x_1 \to$	$p(x_1)$
$x_2 \to$	$p(x_2)$
$\vdots$	$\vdots$
x_n	$p(x_n)$

PROBLEM

Let X be a discrete random variable that takes on the values 0, 1, 2 with probabilities ½, 3/8, 1/8 respectively.

Find E(X).

$$E(X) = 0 \times \frac{1}{2} + 1 \times \frac{3}{8} + 2 \times \frac{1}{8} = \frac{5}{8}$$

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

EXPECTATION OF A FUNCTION OF X

For any function $g(x)$, $E(g(X)) = \sum_{i=1}^{n} g(x_i)p(x_i)$.

$E(g(X))$ exists only if $\boxed{E(|g(X)|)}$ exists.

PROBLEM

p(1) = 0.1, p(2) = 0.2 and p(5) = 0.7. Find $E(X^2)$, $E(X^3)$. $E(x)$

x	$P(X=x)$
1	0·1
2	0·2
5	0·7

$E(x) = 1 \times 0\cdot1 + 2 \times 0\cdot2 + 5 \times 0\cdot7 = 4$

$$E(x^2) = \sum_{i=1}^{3} x_i^2\, p(x_i) = x_1^2 p(x_1) + x_2^2 p(x_2) + x_3^2 p(x_3)$$

$$= 1^2 \times 0\cdot1 + 2^2 \times 0\cdot2 + 5^2 \times 0\cdot7$$

$$= 18\cdot4$$

$$E(x^3) = 1^3 \times 0\cdot1 + 2^3 \times 0\cdot2 + 5^3 \times 0\cdot7$$

$$= 89\cdot2$$

PROPERTIES OF EXPECTATION (TRUE FOR ALL VARIABLES)

- If X is a random variable, and a and b are two constants, then

$$E(a+bX) = a+bE(X).$$

- $E(g(X)) \neq g(E(X))$ in general. For example, $E(X^2) \neq (E(X))^2$.
- $E(b\times g(X)) = b\times E(g(X))$.
- If X and Y are any two random variables such that E(X) and E(Y) exists, then

$$E(X+Y) = E(X) + E(Y).$$

VARIANCE

The variance of a random variable measures its spread.

For a discrete random variable **X** taking values **x_1, x_2, … , x_n** with probabilities **$p(x_1)$, $p(x_2)$, … , $p(x_n)$** respectively, Variance is given by

$$V(X) = \underline{\sum_{i=1}^{n} (x_i - E(X))^2 p(x_i)} = \sum_{i=1}^{n} x_i^2\, p(x_i) - E(X)^2 = \underline{E(X^2) - E(X)^2}$$

SD is the square root of variance.

PROBLEM

Let X be a discrete random variable taking values 0, 1, 2 with probabilities ½, 3/8, 1/8 respectively. Find V(X).

$$V(x) = E(x^2) - (E(x))^2$$

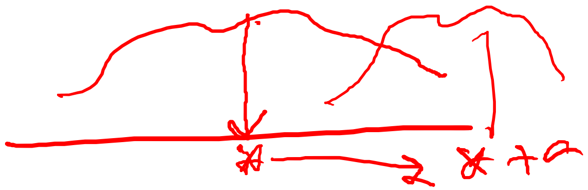
$$E(x) = \frac{5}{8}$$

x	$P(x=x_i)$
0	$1/2$
1	$3/8$
2	$1/8$

$$= 0^2 \times \frac{1}{2} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{1}{8} - \left(\frac{5}{8}\right)^2$$

$$= \frac{7}{8} - \frac{25}{64} = \frac{31}{64}$$

PROPERTIES OF VARIANCE



1. True for all variables: If X is a random variable, and a and b are two constants, then

   $V(bX) = b^2V(X)$, $V(a+X) = V(X)$.

   Together: $V(a+bX) = V(bX) = b^2V(X)$

   Note: $sd(a+bX) = |b|sd(X)$

2. If X and Y are two independent random variables: then

   $V(X+Y) = V(X) + V(Y)$.

   Otherwise, there will be a cross-term called covariance.

RISK AVERSE, RISK NEUTRAL AND RISK LOVING CONSUMER

A risk averse consumer demands a positive expected gain as compensation for taking risk. This compensation increases with the level of risk taken and the degree of risk aversion.

A risk neutral consumer completely ignores risk and always accepts a prospect that offers a positive gain

A risk loving consumer may accept a risky prospect even if the expected gain is negative.

PROBLEM

You are considering two mutual funds for your investment. The possible returns for the funds are dependent on the state of the economy and are given in the table:

You believe that the likelihood is 20% that the economy will be good, 50% that it will be fair, and 30% that it will be poor.

Which fund will you pick if you are risk averse?

STATE OF ECONOMY	FUND 1	FUND 2
Good	20%	40%
Fair	10%	20%
Poor	-10%	-40%

R.V. X = Return from fund 1

R.V. Y = Return from fund 2

X	20	10	-10
P(X=x)	0.2	0.5	0.3

Var (X) = 124

SD (X) = 11.135

$E(X) = 6\%$

$E(Y) = 6\%$

Y	40	20	-40
P(Y=y)	0.2	0.5	0.3

Var(Y) = 964

SD(Y) = 31.048

SPECIAL RANDOM VARIABLES

SPECIAL RANDOM VARIABLES

We now discuss some special random variable needed for specific modelling situations.

Examples:

Probability distribution of the number of rainy days in Jamshedpur this year.

Probability distribution of the waiting time till the first day of rain in terms of number of days after the official onset of monsoon.

Probability distribution of the number of children out of 100 randomly chosen kids of age 10 who have dropped out of school.

Probability distribution of a “rare event”, say defects in a page of our textbook.

LOOK AT THE FOLLOWING RANDOM EXPERIMENTS

- Tossing a coin
- A school student applies or does not apply to college
- A drug is effective or not effective
- An employee travels or does not travel by public transport

-- These experiments have only 2 outcomes:

Success or Failure!

BERNOULLI PROCESS

- The result of each trial may be either a success or a failure

We get to define "success"!

- The probability of success is the same in every trial
- The trials are independent

BERNOULLI RANDOM VARIABLE

X	P(X=x)
0	1-p
✓ 1	✓ p

X: Bernoulli Random Variable

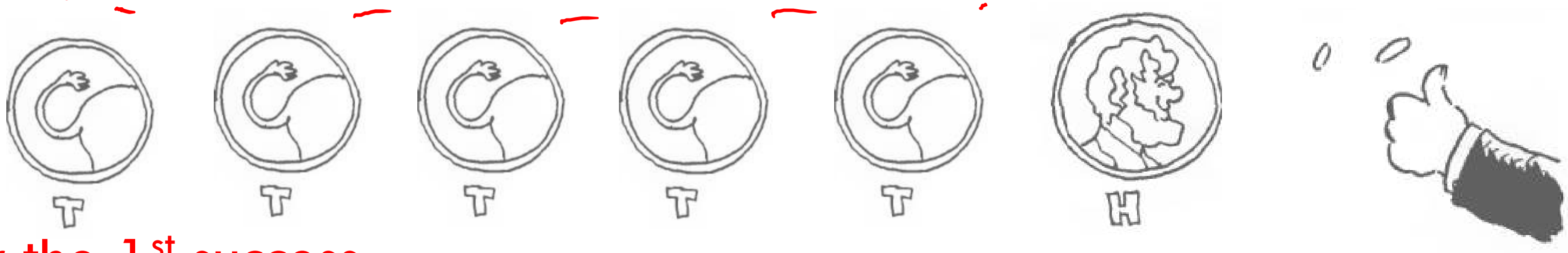
E(X) = p

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

ANOTHER COIN TOSS EXPERIMENT

Suppose we have independent Bernoulli trials with probability of success p

S



Waiting for the 1st success

X= number of trials needed to get the first success

$P(X = k) = (1-p)^{k-1} \cdot p \quad , \quad k = 1, 2, 3, \ldots$

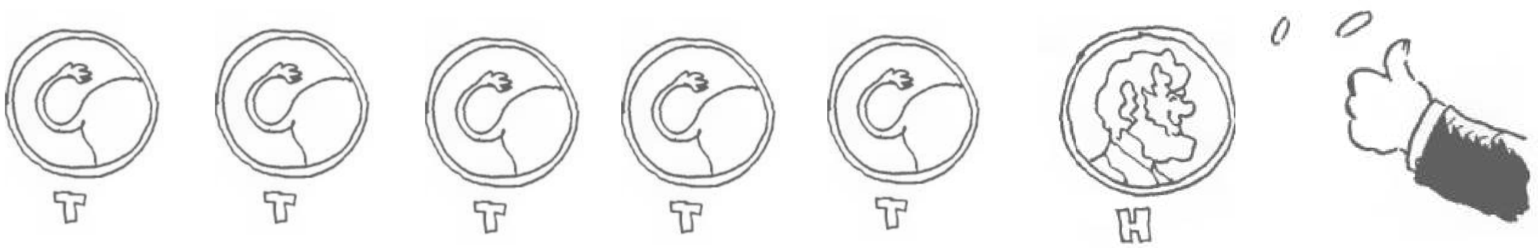
GEOMETRIC RANDOM VARIABLE

$X \sim Geom(p)$

A geometric random variable is the number of Bernoulli trials needed for the first success.

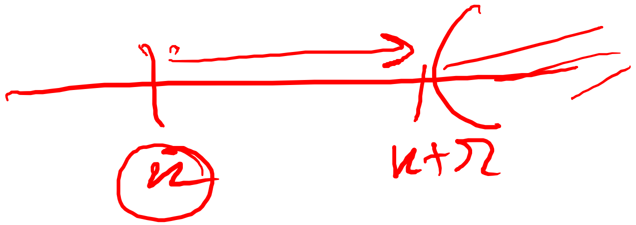
Suppose we have independent Bernoulli trials with probability of success p

Geometric Distribution: Waiting for the 1st success



X= number of trials needed to get the first success

p.m.f. : $P(X = k) = (1 - p)^{k-1}p, k = 1, 2, 3, \ldots$



GEOMETRIC DISTRIBUTION

- $E(X) = 1/p$
- $V(X) = (1-p)/p^2$
- $P(X>k) =$ $\sum_{i=k+1}^{\infty}(1-p)^{i-1}p \quad \left(= 1-\sum_{i=1}^{k}(1-p)^{i-1}p\right)$

$$= (1-p)^k$$

- $P(X>k+r \mid X>k) = P(X>r)$

$$P(X>k+r \mid X>k) = \frac{P(X>k+r \cap X>k)}{P(X>k)} = \frac{P(X>k+r)}{P(X>k)} = \frac{\sum_{i=k+r+1}^{\infty}(1-p)^{i-1}p}{\sum_{i=k}^{\infty}(1-p)^{i-1}p}$$

Memoryless property.

Geometric is the only discrete distribution with memoryless property.

$$= \frac{(1-p)^{n+k}\, p \sum_{j=0}^{\infty} (1-p)^j}{(1-p)^k\, p \sum_{j=0}^{\infty} (1-p)^j}$$

$$= (1-p)^n$$

$$= P(X > n)$$

Hint: $(1-p)^{n+k}p$

$+ (1-p)^{n+k+1} p +$

$(1-p)^{n+k+2} p + \cdots$

$= (1-p)^{n+k} \cdot p \left[1 + (1-p) + (1-p)^2 + \cdots \right]$

X	
0	1/2
1	3/8
2	1/8

→ dgeom

NOTE: ALTERNATIVE FORMULATION

Sometimes a geometric random variable counts number of failures before the first success, instead of number of trials.

Notice that number of failures = number of trials – 1

R uses this definition. #f =

R:

- p.m.f. of no. of failures: `dgeom(x, p)` → # of f before the first success occurs
- c.d.f. of no. of failures: `pgeom(x, p)`

→ Prob of success

$\#f = x - 1$

$\Rightarrow x = \#f + 1$

$P(\#f \le x)$

$= P(X \le x+1)$