Recall :-

– Random variables :- many different distributions out of Bernoulli trials; creating new sample spaces for each distribution, same questions on different frameworks.

Work on – one sample space – define functions on them; using these functions answer all the relevant questions. – Random variables.

- $X : S \to \mathbb{R}$ – (temporary definition) –↗
- S, – sample space $\mathbb{P}$ – Probability on it
  $X : S \to \mathbb{R}$ & $Q(B) = \mathbb{P}(X^{-1}(B))$
  then Q defines a Probability on the range of X.
- $X : S \to T$ S – countable (finite) equipped with Probability $\mathbb{P}$ & T is countable subset of $\mathbb{R}$ then X is called as a Discrete Random variable.

$f_X(t) = \mathbb{P}(X = t)$ – Probability mass function of X.

ie $\Rightarrow \mathbb{P}(X \in A) = \sum_{t \in A} f_X(t)$.

- $X : S \to T$, $Y : S \to T$ be discrete random variables. X and Y have equal distribution if

$$\underbrace{\mathbb{P}(X = t) = \mathbb{P}(Y = t)}_{\text{Same Probability mass function}} \quad \forall t \in T$$

Independent Random Variables :-

Recall :- Two events A and B are independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B) \quad (\equiv 2^2 \text{ equations}) \; (**)$$

$$[\because \iff \mathbb{P}(A|B) = \mathbb{P}(A) = \mathbb{P}(A|B^c)]$$

$$\mathbb{P}(B|A) = \mathbb{P}(B) = \mathbb{P}(B|A^c)$$

$A_1, \dots, A_n$, $n \geq 2$ are independent if

$$P(A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \dots A_n^{\varepsilon_n}) = \prod_{i=1}^{n} P(A_i^{\varepsilon_i}) \quad -(2^n \text{ equations})$$

$\varepsilon_i = 0, 1$, $\quad A_i^{\varepsilon_i} = \begin{cases} A_i & \text{if } \varepsilon_i = 1, \\ A_i^c & \text{if } \varepsilon_i = 0 \end{cases}$

The above notion can also be considered for random variables.
– understand relationship between random variables.

Definition 3.2.1 :- (Independence of two random variables) Two random variables X and Y are independent if $\{X \in A\}$ and $\{Y \in B\}$ are independent for every event B in the range of X and every event A in the range of Y.

i.e. $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\,\mathbb{P}(Y \in B)$ for all events A – in the range of X and events B in the range of Y.

Ex :- X and Y are discrete random variables. Then X and Y are independent iff

$$\mathbb{P}(X = t, Y = s) = \mathbb{P}(X = t)\,\mathbb{P}(Y = s)$$

$\forall$ $t \in \text{Range}(X)$ and $s \in \text{Range}(Y)$.

Definition 3.2.3 (Mutual independence of Random Variables) A finite collection of random variables $X_1, \dots, X_n$ are mutually independent if the sets $(X_j \in A_j)$ are mutually independent for all events A_j in the ranges of the corresponding X_j.

An arbitrary collection of random variables X_t where $t \in I$, for some index set I, is mutually independent if every finite sub-collection is mutually independent.

Example 3.2.2 – Roll a dice two times
– 36 outcomes and all equally likely outcomes

Each Rolls outcome can be viewed as a Random variable.

X – outcome of 1^{st} roll $\qquad X \sim \text{uniform}\{1,2,3,4,5,6\}$

Y – outcome of 2^{nd} roll $\qquad Y \sim \text{uniform}\{1,2,3,4,5,6\}$

$$\mathbb{P}(X=x) = 1/6 \quad \forall x \in \{1,2,..6\}$$

$$\mathbb{P}(Y=y) = 1/6 \quad \forall y \in \{1,2,..6\}$$

(Independent) $\qquad \mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y) \quad \forall x \in \text{Range}(X),\ y \in \text{Range}(Y)$

– Suppose we repeat an experiment n times for $n \geq 1$.
  – trials are independent
  – X_i – denote outcome of the ith trial. $i=1,..,n$

– for any $i \in \{1,..,n\}$ all X_i have the same distribution and ($i \neq j$ X_i and X_j are independent) $(X_1, .. X_n)$ are mutually independent.

in such case one says $X_1,..X_n$ are i.i.d. random variables

i.i.d. → independent, identically distributed.

Example 3.2.4: let $X_1,..,X_n$ be a sequence of i.i.d. random variable, such that $X_1 \sim \text{Geometric}(p)$, $0<p<1$.

(i.e. X_i - outcome of experiment i ; Experiment is :- observe the time taken for the first success)

What is the Probability all random variables are greater than or equal to j ? (for some $j \geq 1$)

i.e. Find $\mathbb{P}(X_1 \geq j, X_2 \geq j, \dots, X_n \geq j) = ?$

Step 1 : Find $\mathbb{P}(X_1 \geq j)$

$X_1 \sim \text{Geometric}(p)$; $\mathbb{P}(X_1 = k) = p(1-p)^{k-1} \quad \forall k \geq 1.$

$$\therefore \quad \mathbb{P}(X_1 \geq j) = \sum_{k=j}^{\infty} \mathbb{P}(X_1 = k)$$

Series & computation can be made precise

$$= \sum_{k=j}^{\infty} p(1-p)^{k-1}$$

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

$$= (1-p)^{j-1}$$

Step 2 : (Use independence) $(X_i \geq j) \ \forall i = 1, \dots, n$ are events concerning X_i

$$\mathbb{P}(X_1 \geq j, \dots, X_n \geq j) = \prod_{i=1}^{n} \mathbb{P}(X_i \geq j)$$

(identically distributed) $$= \prod_{i=1}^{n} (1-p)^{j-1}$$

$$= \left[(1-p)^{j-1}\right]^n \qquad \triangle$$

3.2.2 Conditional, Joint, and Marginal distribution.

Example: (1) Suppose we choose a number randomly from the set $\{1,2\}$ ≡ let X denote the outcome and note down

(2) – We toss a coin X times. ≡ let Y denote the outcome the number of Heads of (1).

– clearly (2) depends on the outcome

$$\text{range}(X) = \{1,2\}$$

$$\text{range}(Y) = \{0,1,2\}$$

– X and Y are "not" independent

intuitively clear

– How to quantify this?

ⓐ $\mathbb{P}(X=1) = \frac{1}{2} = \mathbb{P}(X=2)$ as $X \sim \text{uniform}\{1,2\}$

ⓑ $\mathbb{P}(Y=0 \mid X=1) = 1-p$ ... (ie toss a coin once and outcome is a Tail)

(Conditional Probability)

$\mathbb{P}(Y=1 \mid X=1) = p$ ... (ie toss a coin once and outcome is a Head)

$\mathbb{P}(Y=0 \mid X=2) = (1-p)^2$ ... (ie toss a coin 2 times and no head appears)

$\mathbb{P}(Y=1 \mid X=2) = \binom{2}{1} p^1 (1-p)^{2-1}$ ... (ie toss a coin 2 times and one head appears)

$= 2\,p(1-p)$

$\mathbb{P}(Y=2 \mid X=2) = p^2$ ... (ie toss a coin 2 times and 2 heads appear)

(c) $\text{range}(Y) = \{0, 1, 2\}$

$$\mathbb{P}(Y=0) = \mathbb{P}((Y=0, X=1) \cup (Y=0, X=2))$$

mutually exclusive event

$$= \mathbb{P}(Y=0, X=1) + \mathbb{P}(Y=0, X=2)$$

& (a), (b)

$$\doteq \mathbb{P}(Y=0 \mid X=1)\,\mathbb{P}(X=1) + \mathbb{P}(Y=0 \mid X=2)\,\mathbb{P}(X=2)$$

$$= \boxed{\overset{(b)}{(1-p)}\ \overset{(a)}{\tfrac{1}{2}}} + \boxed{\overset{(b)}{(1-p)^2}\ \overset{(a)}{\tfrac{1}{2}}}$$

$$= \frac{1}{2}\left[(1-p) + (1-p)^2\right]$$

One can compute $\mathbb{P}(Y=1)$ & $\mathbb{P}(Y=2)$ similarly.

Definition 3.2.5 :- Let X be a random variable on a sample space S. Let $A \subseteq S$ be an event such that $\mathbb{P}(A) > 0$. Then the probability φ described by

$$\varphi(B) = \mathbb{P}(X \in B \mid A) := \frac{\mathbb{P}((X \in B) \cap A)}{\mathbb{P}(A)}$$

is called the "conditional distribution" of X given the event A.

In previous example :- $A = \{X = 1\}$ and we computed $\mathbb{P}(Y=1 \mid A) = p$ and $\mathbb{P}(Y=0 \mid A) = 1-p)$

$\therefore$ Conditional distribution $Y \mid A \sim$ Bernoulli (p)

we also showed if $C = \{X = 2\}$ then the conditional distribution of $Y \mid C \sim$ Binomial $(2, p)$.

- In example :- Dependence between Y and X was clear and specified by the conditional distribution of Y given an outcome of X.

- there is another way - specify the "joint distribution". $\boxed{?}$

Definition 3.2.7. :- If X and Y are discrete random variables, the "joint distribution" of X and Y is the probability φ on the pairs of values in the range of X and Y defined by

$$\varphi((a,b)) = \mathbb{P}(X=a, Y=b).$$

The definition may be expanded to a finite collection of random variables $(X_1, \ldots X_n)$ for which the joint distribution of all n variables is the probability φ defined by

$$\varphi(\{a_1, a_2, \ldots, a_n\}) = \mathbb{P}(X_1=a_1, \ldots X_n=a_n).$$