Take an example. Say, $\mathcal{U}$ contains n events $\{A_1, A_2, \dots, A_n\}$. Then mutual independence implies

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j) \quad \text{for } 1 \le i \ne j \le n.$$

(hence, $\mathcal{U}$ is pairwise independent).

$$\mathbb{P}(A_i \cap A_j \cap A_k) = \mathbb{P}(A_i)\,\mathbb{P}(A_j)\,\mathbb{P}(A_k) \quad \text{for } 1 \le i \ne j \ne k \le n.$$

$$\vdots$$

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

For examples and exercises, see Rohatgi & Saleh.

Lecture - IV. Random variables

Definition.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A map $X: \Omega \to \mathbb{R}$ is called random variable if $\{\omega: X(\omega) \in A\} \in \mathcal{F}$ for every $A \in \mathcal{B}(\mathbb{R})$.

Sometimes, it is said that a random variable is a "measurable map." A "measurable" map means a map which preserves the "measurable sets". "Measurable sets" in our context means 'events". Remember that we can compute probability only for the events (the members of the σ-field). We have seen that the Borel sets are the events if $\Omega = \mathbb{R}$. If X is a map from Ω to $\mathbb{R}$ and we want to compute the probability $\mathbb{P}(X \in A)$ for any $A \in \mathcal{B}(\mathbb{R})$, then we must have

$$\{\omega: X(\omega) \in A\} \in \mathcal{F}.$$

as we can compute probability $\mathbb{P}$ only ~~of~~ for the members/elements of $\mathcal{F}$. So the requirement ~~is very natural~~ of preserving ~~...~~ "events/measurable sets" by X is natural.

Recall that $\mathcal{B}(\mathbb{R})$ is Borel σ-field "generated" by the intervals $\{(-\infty, a) : a \in \mathbb{R}\}$ or $\{(b, \infty) : b \in \mathbb{R}\}$ that is, $\mathcal{B}(\mathbb{R})$ is the smallest σ-field containing

$\{(-\infty, a) : a \in \mathbb{R}\}$ and $\{(b, \infty) : b \in \mathbb{R}\}$.

Theorem (Sufficient condition to be a random variable).

Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a real valued map. ~~then we say~~ If $\{\omega : X(\omega) < a\} \in \mathcal{F}$ for every $a \in \mathbb{R}$, then X is a random variable.

Exercise-1. As a consequence of the theorem, show that if $\{\omega : X(\omega) \leq b\} \in \mathcal{F}$ for every $b \in \mathbb{R}$, then a map $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is a random variable.

Note: The proof of this theorem is not included in your syllabus and will be touched in a more advanced course. Intuitively, the idea is that $\{(-\infty, a) : a \in \mathbb{R}\}$ is "generator" of Borel σ-field and it is enough to check on this smaller subcollection of Borel sets.

Some important facts:

~~Here~~ In the definition of ~~B(R)~~ random variable, $\mathbb{R}$ is used as the range of the map X. But we say $X : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (R, \mathcal{G})$ is a random variable ~~with~~ when $R \subset \mathbb{R}$ for example $R = \mathbb{N}, \mathbb{Z}, \mathbb{N}_0, \mathbb{Q}, [a,b]$ etc. By this I mean, the range of any random variable must be contained in $\mathbb{R}$. Althoug, we are not going to verify these

* The σ-field associated to $\mathbb{N}$ is $2^{\mathbb{N}}$. It can be shown that

$$2^{\mathbb{N}} = \{ B \cap \mathbb{N} : B \in \mathcal{B}(\mathbb{R}) \}.$$

* The σ-field associated to $\{1, 2, 3, \dots n\}$ is $2^{\{1,2,3,\dots n\}}$. It can be shown that

$$2^{\{1,2,3,\dots n\}} = \{ (1,2,3,\dots n) \cap B : B \in \mathcal{B}(\mathbb{R}) \}.$$

* The σ-field associated to $\mathbb{Z}$ is $2^{\mathbb{Z}}$ and

$$2^{\mathbb{Z}} = \{ \mathbb{Z} \cap B : B \in \mathcal{B}(\mathbb{R}) \}.$$

* The σ-field associated to $[a,b]$ is defined to be

$$\mathcal{B}([a,b]) = \{ [a,b] \cap B : B \in \mathcal{B}(\mathbb{R}) \}.$$

Therefore, Borel σ-field $\mathcal{B}(\mathbb{R})$ generalizes and contains all the members of the σ-fields associated to any of its subset.

This is why we use $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ as the range of a random variable rather than mentioning each of interesting subsets of $\mathbb{R}$.

An algorithmic construction of random variables

* Based on the problem/purpose, choose a subset $R \subset \mathbb{R}$. Check that $R \in \mathcal{B}(\mathbb{R})$. (usually we don't check it as the subset of $\mathbb{R}$ which is not a member of $\mathcal{B}(\mathbb{R})$ is very complicated and never occurs)

* Declare $\Omega = R$ and $\mathcal{F} = \mathcal{B}(R) = \{B \cap R : B \in \mathcal{B}(\mathbb{R})\}$.

* Define or write down the map X explicitly that is, specify the value of $X(\omega)$ for each $\omega \in R$.

(Usually, it turns all the maps that we can write down are random variables. There are pathological cases but they ~~occur~~ ~~never~~ seldom occur}. So we can ignore such a situation)

Advantage.

* If we can write down or describe random objects or random experiments in terms of the real ~~numb~~ real numbers then we can immediately use the

arithmetic and analytic properties of real line

to analyse the random experiment or random objects.

We have learned plethora of tools (algebra, calculus) to use these properties.

Example. (Coin tossing)

Suppose that we are tossing n ~~coin~~ coins together such that the tossings are independent. We are interested in number of heads. We further assume that probability of head from each coin is p.

It is very natural that the number of heads is a non-negative integer, and less than or equal to n. So

$$\Omega = \{0, 1, 2, \ldots n\}, \ \mathcal{F} = 2^{\{0,1,2,\ldots n\}}.$$

$$\mathbb{P}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k}, \ k = 0, 1, 2, \ldots n.$$

We can derive the formula using combinatorial technique. Our favourable outcome must contain k heads and $(n-k)$ tails. Each of such arrangement has probability

$$p^k (1-p)^{n-k}.$$ (as the outcome of one experiment is independent of the other.)

So we are now remained with counting the number of such arrangements of heads and tails. The important fact is that the order in which the heads appear does not influence the probability and number of heads in the arrangement. So, total number of such arrangements is

$$\binom{n}{k}.$$

Combining these two facts we have our desired answer.

* We have seen that if Ω is finite and we specify probability of each of the sample point, then we have specified the probability on the entire space. $\Omega = \{0, 1, 2, \ldots n\}$.

<u>An alternative way to see this.</u>

Consider a coin tossing experiment where n coins are tossed independently. We are interested in the number of heads in a particular arrangement.

(PROBLEM IS SAME).

Define $\tilde{\Omega} = \{H, T\}^n = \{(a_1, a_2, \dots a_n) : a_i \in \{H, T\}$ for all $i = 1, 2, \dots n\}$

$\tilde{\mathcal{F}} = 2^{\tilde{\Omega}}$.

$$\tilde{\mathbb{P}}(\{a_1, a_2, \dots a_n\}) = p^{\#\{i : a_i = H\}} (1-p)^{\#\{i ; a_i = T\}}$$

$$= p^{|\{i : a_i = H\}|} (1-p)^{|\{i : a_i = T\}|}.$$

We can check that $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ is a probability space. (I leave it for you to check it as an exercise).

Define a map $X : (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}) \to (\Omega, \mathcal{F}, \mathbb{P})$

where $X(\{w\}) = X(\{a_1 \dots a_n\}) = |\{i : a_i = H\}|$.

for every $w = (a_1, \dots a_n) \in \tilde{\Omega}$. Here

$\Omega = \{0, 1, 2, \dots, n\}$, $\mathcal{F} = 2^{\Omega} = 2^{\{0, 1, 2, \dots, n\}}$

<u>Claim - I.</u> X is a random variable.

We know Ω is finite. ~~then~~ and therefore $\mathcal{F}$ is generated by the singletones $(\{i\} : i = 1, 2, \dots n)$

So it is enough to check that

$X^{-1}(\{k\}) \in \tilde{\mathcal{F}}$ for every $k \in \Omega$.

$$X^{-1}(\{k\}) = \{(a_1, a_2, \ldots, a_n) : |\{i : a_i = H\}| = k\} \subseteq \tilde{\Omega}$$

and so $X^{-1}(\{k\}) \in \tilde{\mathcal{F}} = 2^{\tilde{\Omega}}$. So we checked that $\cancel{X^{-1}(\{k\})}$ X is a random variable.

We usually write

$$\tilde{\mathbb{P}}(\{\omega : X(\omega) = k\}) = \tilde{\mathbb{P}}(\{(a_1, \ldots, a_n) \in \tilde{\Omega} : X(a_1, \ldots, a_n) = k\})$$

$$=: \mathbb{P}(X = k) \longrightarrow (**_1)$$

$\mathbb{P}$ is a set function on $(\Omega, \mathcal{F})$ which is induced by the random variable X from $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. We specify the ~~proba~~ set function $\mathbb{P}$ in terms of $\tilde{\mathbb{P}}$. It can be checked that $\tilde{\mathbb{P}}$ is a probability implies $\mathbb{P}$ is a probability. We can compute

$$\tilde{\mathbb{P}}(\{(a_1, a_2, \ldots, a_n) \in \tilde{\Omega} : X(a_1, a_2, \ldots, a_n) = k\})$$

$$= \binom{n}{k} p^k (1-p)^k \quad \text{for every } k = 0, 1, 2, \ldots n$$

using the same argument as before.

<u>Summary</u>. We started with probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and the we introduced a random

variable $X : (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}) \longrightarrow (\Omega, \mathcal{F}, \mathbb{P})$. So we constructed a new probability space from a given one using the random variable X.

Note that the elements in $\tilde{\Omega}$ has more information ~~that~~ than Ω as a typical element $(a_1, a_2, \ldots, a_n) \in \tilde{\Omega}$ also has the information on the order in which the heads and tails appear.

There is nothing special about finite sample space or coin ~~toss~~ tossing.

Exercise (Conceptually difficult).
Let us throw a dice for infinitely many times and ~~independently~~ independently (for the moment, consider this hypothetical situation but this will be formalized later). ~~Clear~~ Clearly $\Omega = \{1, 2, 3, 4, 5, 6\}^{\mathbb{N}}$ that is each outcome is a sequence whose elements are one of $\{1, 2, 3, 4, 5, 6\}$. We are interested in the first time 6 occurs in this sequence. Note that Ω is countable and hence we can consider the σ-field $\mathcal{F} = 2^{\Omega}$. Suppose that the die is fair. Then we have a probability on $(\Omega, \mathcal{F})$