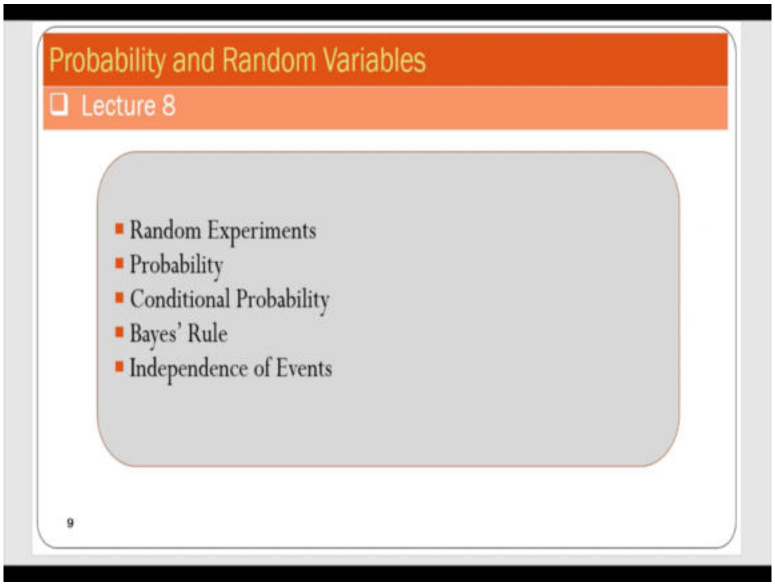
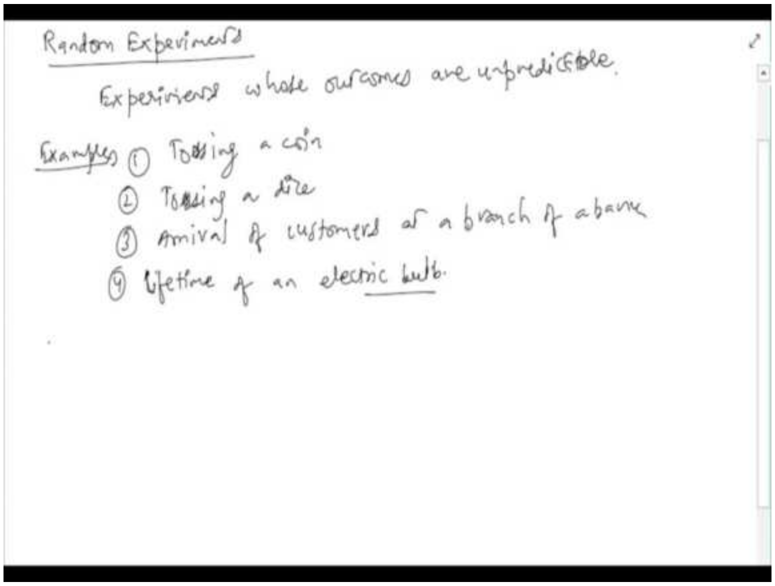
Mathematical Aspects of Biomedical Electronic System Design
Professor Chandramani Singh
Indian Institute of Science, Bangalore
Lecture - 22
Introduction to Probability

(Refer Slide Time: 00:38)



Hello, everyone. Welcome to today's lecture of the course mathematical aspects of biomedical electronic system design. In this lecture, we will begin a new module probability and random variables. In the first lecture of this module, we will learn about random experiments, probability, conditional probability, Baye's rule, independence of events et-cetera. So, let us begin today's lecture.

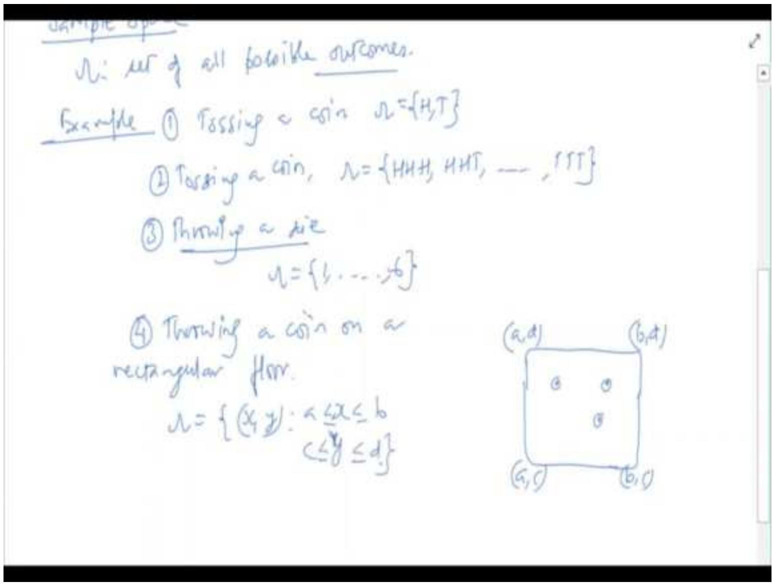
(Refer Slide Time: 00:58)



Let me start with introducing random experiments. Any experiment whose outcome is unpredictable is referred to as random experiment; experiments whose outcomes are unpredictable. For instance, we can think of so tossing a coin in which case outcome could be either head or tail; tossing a dice, the outcome could be any number from 1 to 6.

Arrival of customers say at the branch of a bank; so, number of customers that arrive in a day who is not predictable. Or say, lifetime of an electric bulb, it is unpredictable electric bulb. We study this notion of probability in the context of random experiment. Before I introduce probability, I will introduce some other nomenclature.

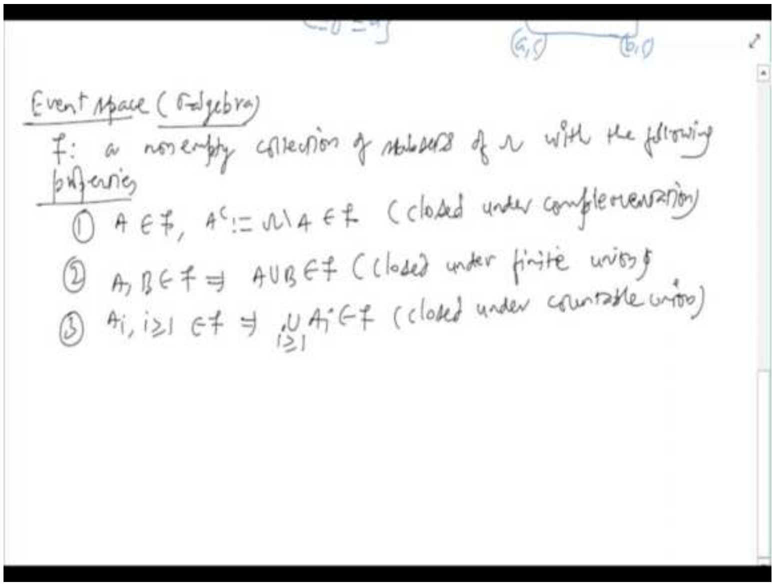
(Refer Slide Time: 3:00)



For instance, let me tell what it means by sample space. Sample space usually denoted by the symbol Ω is set of all possible outcomes. For example, when I am tossing a coin, sample space is this set containing head and tail. Tossing a coin thrice, same coin if I toss thrice; now my sample space is set of tuples of the type HHH, HHT, … et-cetera, TTT. So, there will be eight such combinations, say throwing a die, now my sample space will be 1 to 6; this is the set of all possible outcomes.

Next, consider that we have a rectangular floor; let us say here is a floor and its coordinates are (a,c),σ (b,c), (a,d) and (b,d). I throw a coin on this floor, the coin lands somewhere; and the point where the coin lands is the outcome of my experiment. In this so the experiment is throwing a coin on a rectangular floor. In this case, my sample space is collection of all coins the form x, y, where x is varying between a and b; and y is varying between c and d.

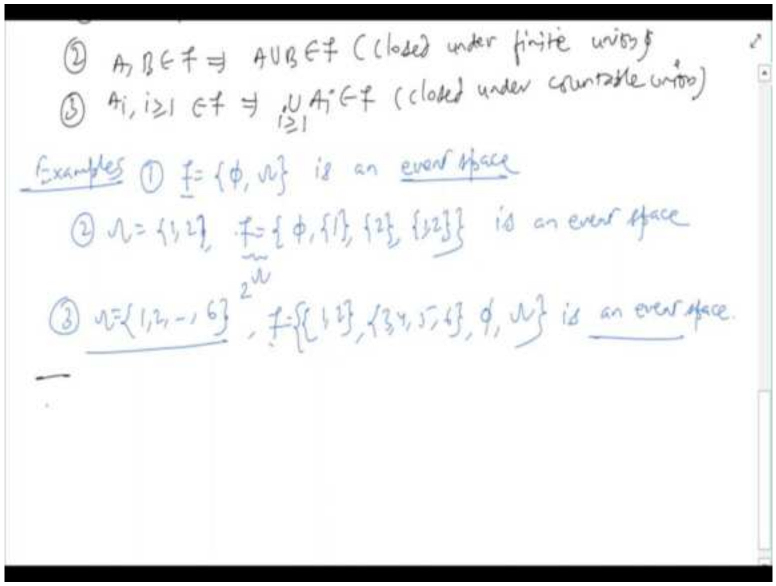
(Refer Slide Time: 05:30)



The next notion is that of event space often called σ algebra. An event space or σ algebra usually denoted by the symbol script F is a non-empty collection of subsets of Ω; collection of subsets of Ω with the following properties. First property is, if A belongs to this set, then A complement that is Ω - A should also belong to F; this property is called closeness under complementation, closed under complementation.

Second property is if A, B belong to F, then their union also belongs to F; this is called closeness under finite union, closed under finite union. The third property is if collection of sets A_i, $i \geq 1 \in F$; then their union also is in F, this is called closeness under countable union. Let us see a few examples of event spaces.

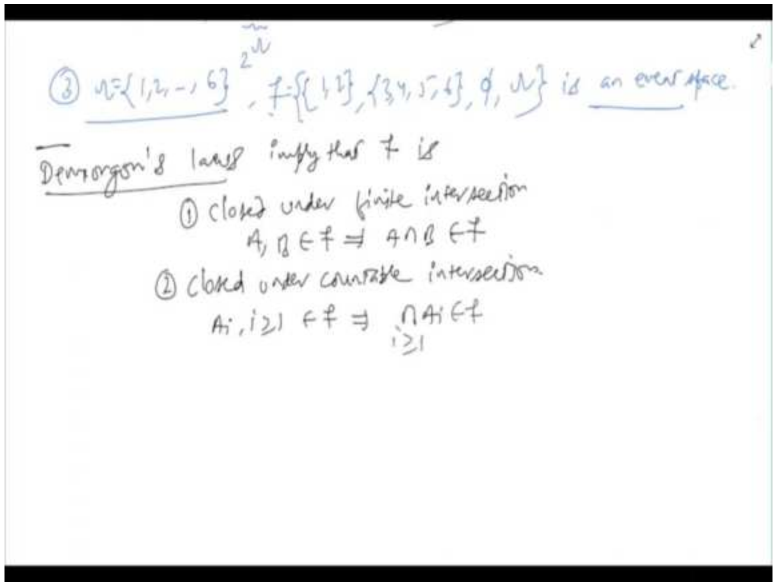
(Refer Slide Time: 07:40)



For any sample space Ω, the collection of empty set and Ω is an event space. It can be checked that this collection satisfies all these three constraints. If Ω = {1, 2} then a collection of subsets {Φ, {1}, {2}, {1,2}} this is an event space. Notice that this collection is also what we call power set of Ω. Similarly, if Ω = {1, 2, 6}, the collection {{1,2}, {3,4,5,6}, Φ, Ω}, this is an event space. Again, it can be checked that, this script F satisfies all the three properties mentioned above.

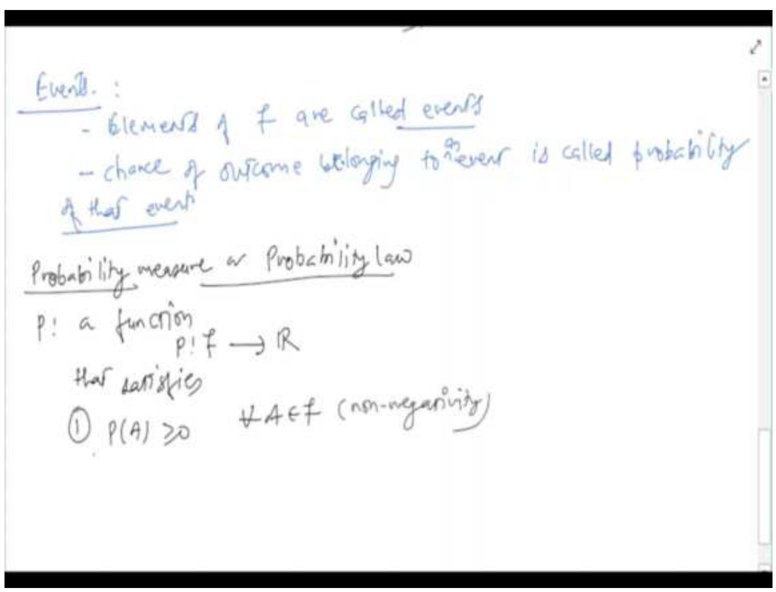
At this point, let us recall DeMorgan's laws set theory. According to these laws if A and B are two subsets of the sample space Ω; then complement of union of A and B is a complement intersection B complement. Similarly, the complement of intersection of A and B is A complement union B complement. Before we move further let us observe a few more properties of event spaces that follow DeMorgan's laws.

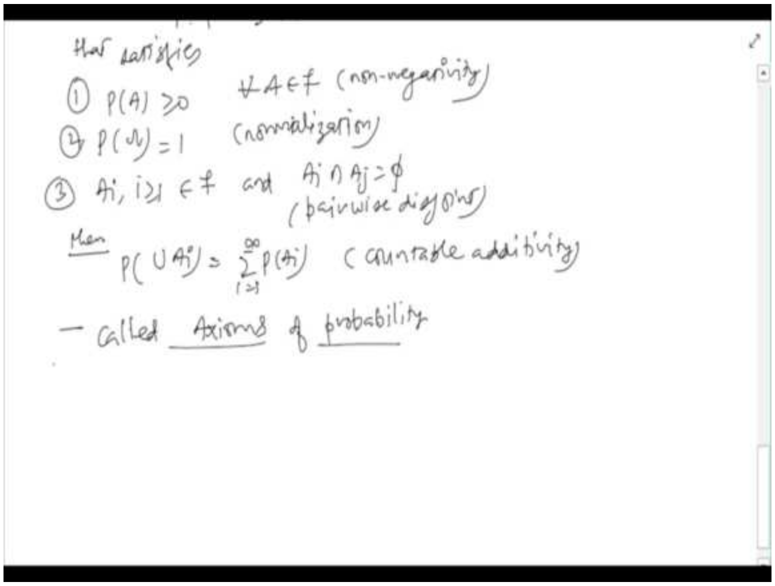
(Refer Slide Time: 09:51)



DeMorgan's laws imply that F is closed under finite intersection; that is if A, B ∈ F, then intersection of A ∩ B ∈ F. F is also closed under countable intersection that is if A_i, i ≥1 ∈ F; then their intersection will also belong to F. the next notion is that of events.

(Refer Slide Time: 10:56)



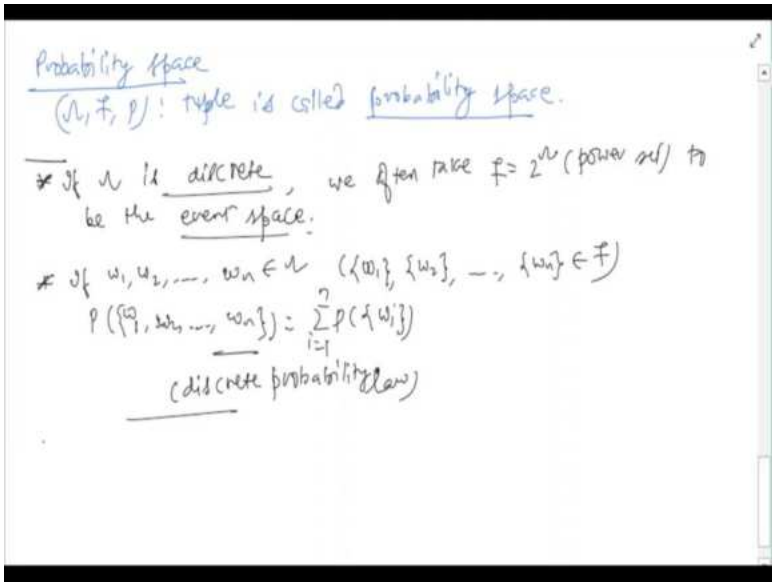


Elements of event space are called events; elements of F are called events. In probability theory we are interested in chances of outcome belonging to an event. In fact, chance of outcome belonging to an event is called probability of that event. If you are interested in probability of an event A, then the outcomes belonging to A are referred to as favorable outcomes.

Let us formalize this notion of probability; we do it via introducing the notion of probability measure; or often called probability law. It is denoted by symbol P which is nothing but a function; P is a function from F to R that satisfies the following properties. The first one is, $P(A) \geq 0$ $\forall A \in F$; this property is called non-negativity. The second property says P of the whole sample space $P(\Omega) = 1$; this is normalization.

And the third one is if A_i, $i \geq 1 \in F$, and they are disjoint; that is $A_i \cap A_j = \Phi$, they are pairwise disjoint. Then, $P(\cup A_i) = \sum_{i=1}^{\infty} P(A_i)$; this property is called countable additivity. The above three properties together are also referred to as axioms of probability, having defined sample space, event space and probability law; we can now define so called probability space.

(Refer Slide Time: 14:37)

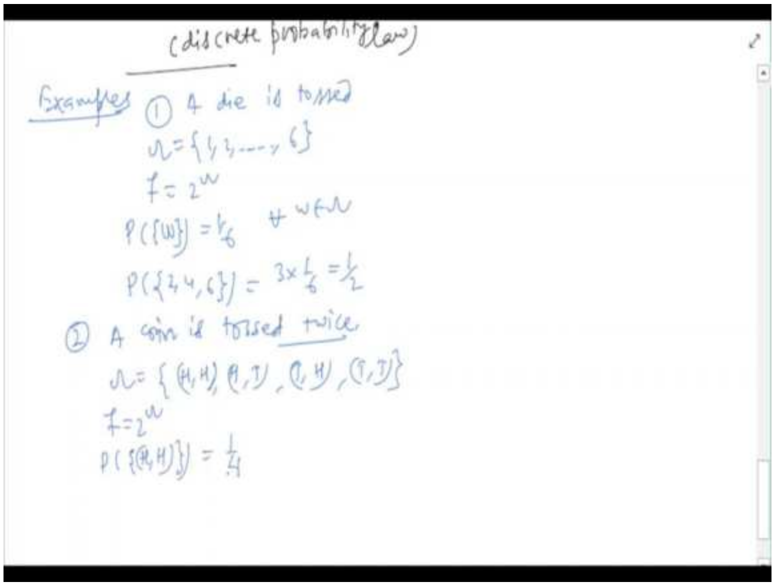


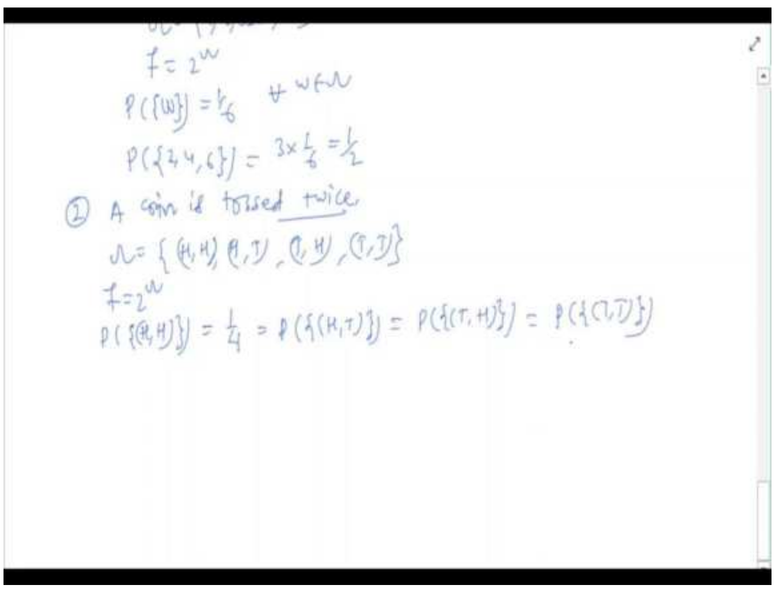
The tuple (Ω, F, P) is called probability space. Before we see examples of probability laws or probability spaces; let me make a couple of statements about the scenarios where Ω is discrete set, or in particular is finite set. If Ω is discrete, we often take 2^{Ω} to be the event space. So, though it could be any collection of subsets of Ω; that satisfies the three properties mentioned above, we often take it to be the power set.

Moreover, if ω_1, ω_2 … ω_n are elements of Ω; then the singleton sets $\{\{\omega_1\}, \{\omega_2\} \dots \{\omega_n\} \in F\}$. And in this case the probability of the set $P(\{\omega_1, \omega_2 \dots \omega_n\}) = \sum_{i=1}^{n} P(\{\omega_i\})$ is simply sum of the probability of these singletons.

This equality is often referred to as discrete probability law. If the sample space is finite, say it has n total outcomes, and all the outcomes are equally likely; in the sense that all singleton events have same probability. Then probability of any event A will be number of favourable outcomes; that is number of elements in A divided by n. Let us see a few examples now.

(Refer Slide Time: 17:25)

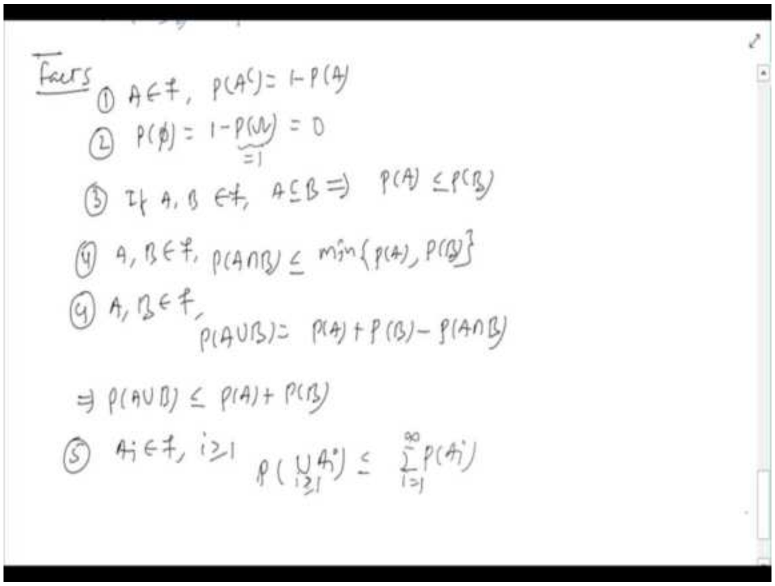


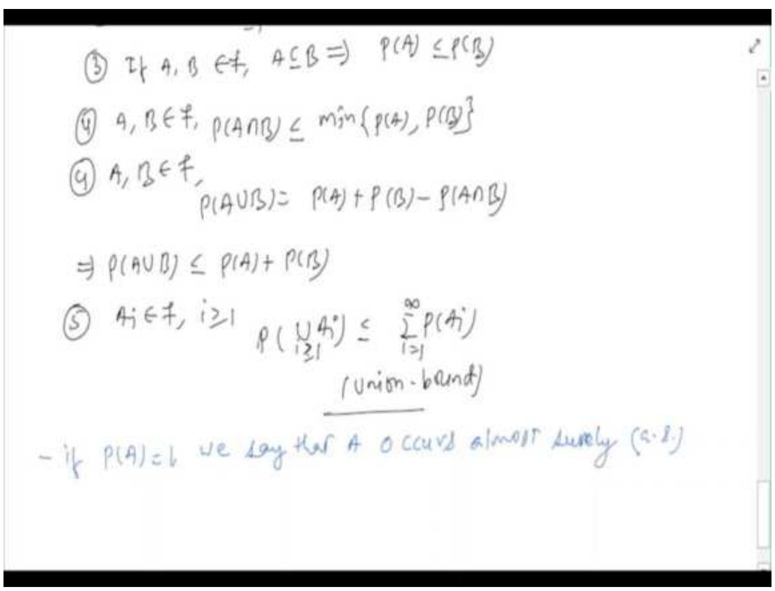


First consider the experiment where a die is tossed. In this case, sample space will be as we have already seen Ω= {1, 2,…. ,6}; event space F=2^{Ω} that is collection of all subsets of Ω. Probability of the singleton sets P(Ω) = $\frac{1}{6}$, $\forall \omega \in \Omega$ that is probability of outcome being 1, being 2, being 3 etcetera; all these probabilities will be $\frac{1}{6}$. On the other hand, if you are interested in probability of say {2, 4, 6}; then following discrete probability law it will be $3 \times \frac{1}{6} = \frac{1}{2}$.

Let us see another example. Now, consider the experiment where a coin is tossed twice. Now, the sample space is Ω ={{HH}, {HT}, {TH}, {TT}}; as before the event space will be power set of the sample space. Probability of an event which is a singleton set, say P({(H,H)}) = $\frac{1}{4}$ = P({(H,T)}) = P({(T,H)}) = P({(T,T)}). Now, we will learn a few facts about probability or probability laws.

(Refer Slide Time: 19:41)





This could be easily verified. The first one says that if A ∈ F, then $P(A^C)$ =1- P(A). The second one is a special case of first. Suppose we take A =Ω, then A^C = Φ; so, P(Φ) =1- P(Ω). And this we know is 1 by the axioms of probability; so, P(Φ) = 0.

The third property says that if A and B are events such that A ⊆ B; then P(A) ≤ P(B). Fourth property says that if A, B are events, then P(A∩B) ≤ min{P(A), P(B)}.

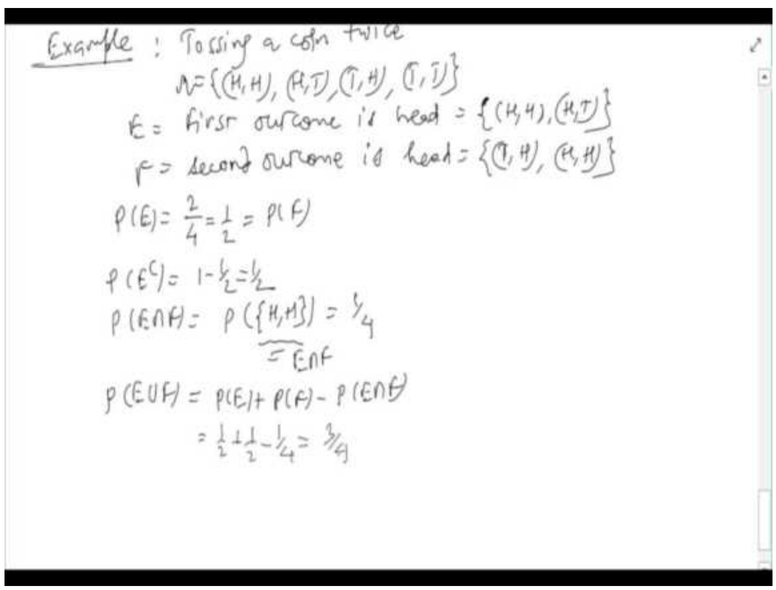
The next property says that as before if A and B are events, then P(A∪B) = P(A) + P(B) – P(A∩B). All these properties can be easily verified using set theory. In particular the fourth property implies that P(A∪B) ≤ P(A) + P(B).

The next property says $A_i \in F$, $i \geq 1$; then

$$P(\cup_{i\geq 1} A_i) \leq \sum_{i=1}^{\infty} P(\{A_i\}) .$$

These properties referred to as union bound. A quick point before we move ahead, if for an event A P(A) = 1; we often say that A happens or A occurs almost surely. In short it is written as this; let us see an example.

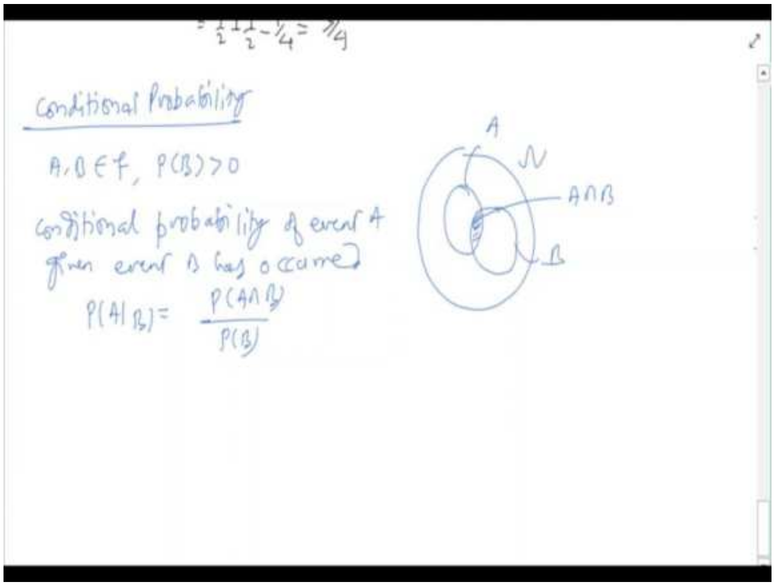
(Refer Slide Time: 22:23)



Consider the random experiment of tossing a coin twice. Recall that in this case Ω ={{HH}, {HT}, {TH}, {TT}}. Define event E as the first outcome being a head, first outcome is head; in other words, this is the event ={{HH}, {HT}}. Similarly, let us define F to be the event that second outcome is head; that is F ={{TH}, {HH}}. In this case P(E) $= \frac{2}{4} = \frac{1}{2}$, so is P(F). Then, following the above properties, $P(E^C) = 1 - \frac{1}{2} = \frac{1}{2}$.

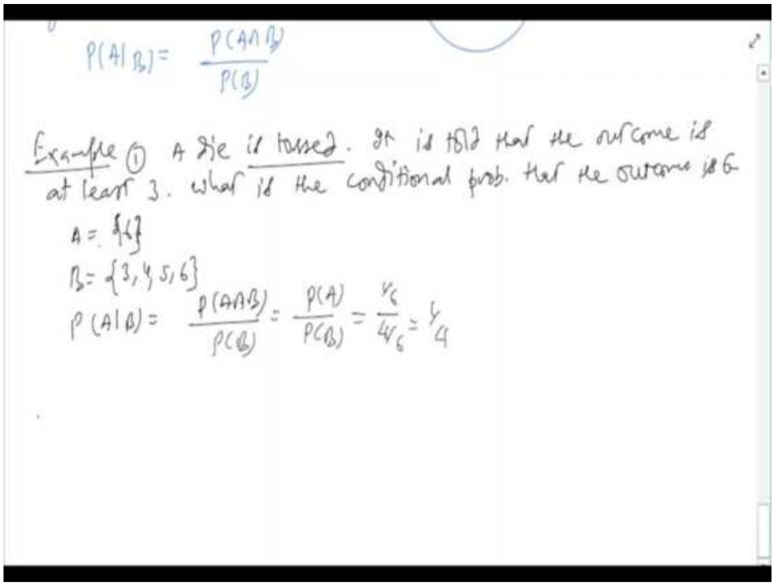
P(E∩F) = P({H,H}); notice that this is the event E intersection F. So, P(E∩F) = $\frac{1}{4}$. And so, P(E∪F) again following one of the properties above = P(E) + P(F) – P(E∩F); which will be $= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}$. The next notion that we will see is of conditional probability.

(Refer Slide Time: 24:29)



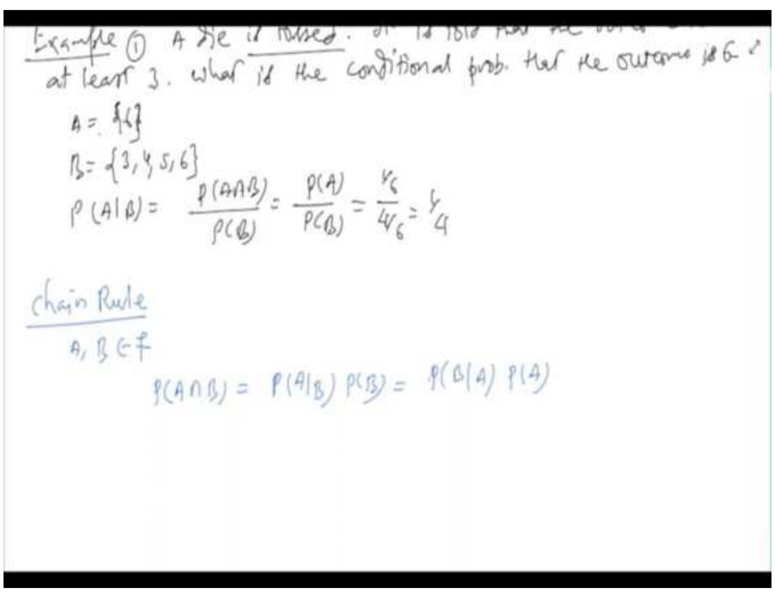
Suppose we have a sample space Ω; there are two events A and B. So, A and B are in event space, and B is such that P(B) > 0; notice that this is the event A∩B. In this case we are interested in conditional probability of event A given event B; conditional probability of event A given event B has occurred. It is denoted as P(A|B), and it turns out to be $= \frac{P(A \cap B)}{P(B)}$. Here we see why we assumed P(B) > 0; it ensures that the division is well defined. Let us see an example.

(Refer Slide Time: 25:59)



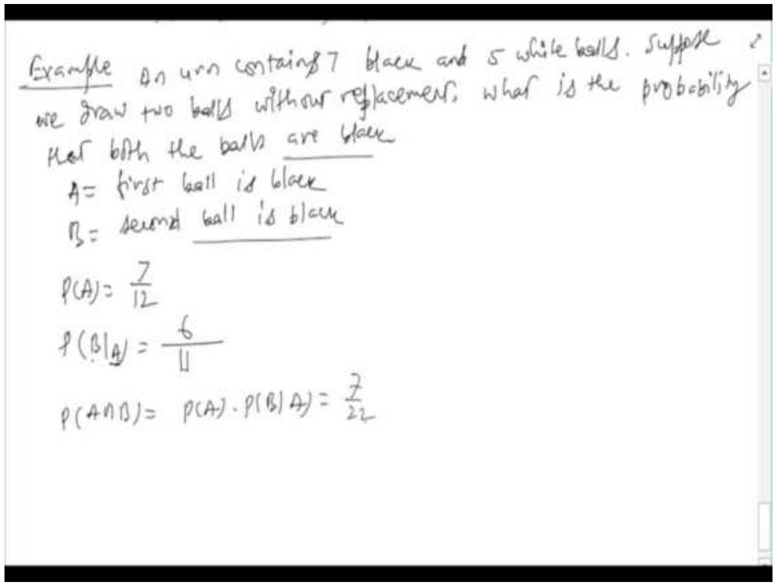
Suppose A die is tossed and it is told that the outcome is at least 3; it is told that the outcome is at least 3. What is the conditional probability that the outcome is 6, probability that the outcome is 6? Here if I define event A to be the event that outcome is 6; and B to be the event that outcome is at least 3. Then $P(A|B) = \frac{P(A \cap B)}{P(B)}$. $A \cap B$ is same as A, so this becomes $= \frac{P(A)}{P(B)}$; this becomes $= \frac{1/6}{4/6} = \frac{1}{4}$. The next rule that we will see is called chain rule.

(Refer Slide Time: 27:28)



As before let us assume that A, B are two events; then we can rearrange the terms in the conditional probability formula to deduce that $P(A \cap B) = \mathrm{P(A|B)P(B)}$. In fact, it also turns out to be $=$ $\mathrm{P(B|A)P(A)}$; this equality is called chain rule. Let us again see an example.

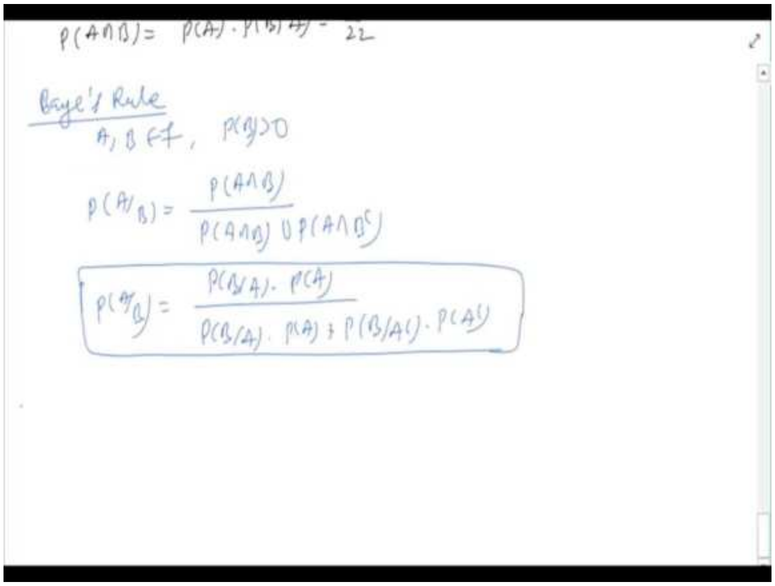
(Refer Slide Time: 28:04)



Assume that an urn contains 7 black and 5 white balls. Suppose we draw two balls from the urn without replacement; what is the probability that both the balls are black? We are interested in probability that both the balls are black. Let us define events A and B as follows. A is the event that first ball is black and B is the event that second ball is black. Then probability of A is see there are 7 black balls and 5 white balls; so, there are total 12 balls probability that I will draw a black ball is $\frac{7}{12}$.

Similarly, P(B|A) can be computed as follows; so, given that the first ball was black ball. Now, there are only 6 black balls; so, there are total 6 + 5 =11 balls, out of which 6 are black. So, given that the first ball was black; the probability that second one will also be black will be $\frac{6}{11}$. Now, I can use chain rule to compute $P(A \cap B)$= P(A).P(B|A); it turns out to be $= \frac{7}{22}$. Next, we will see a rule, a very important one refer to as Baye's rule.

(Refer Slide Time: 30:38)

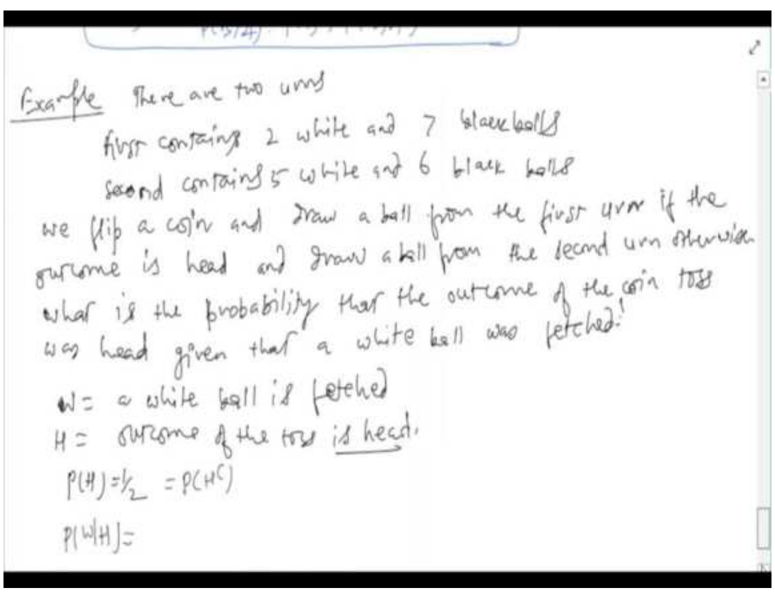


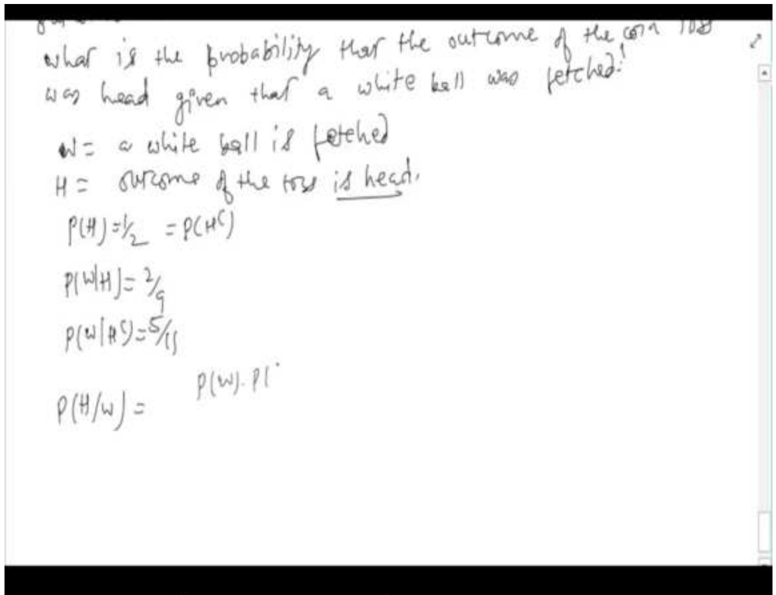
As before, assume that A and B are two events such that probability of B is strictly positive. Then

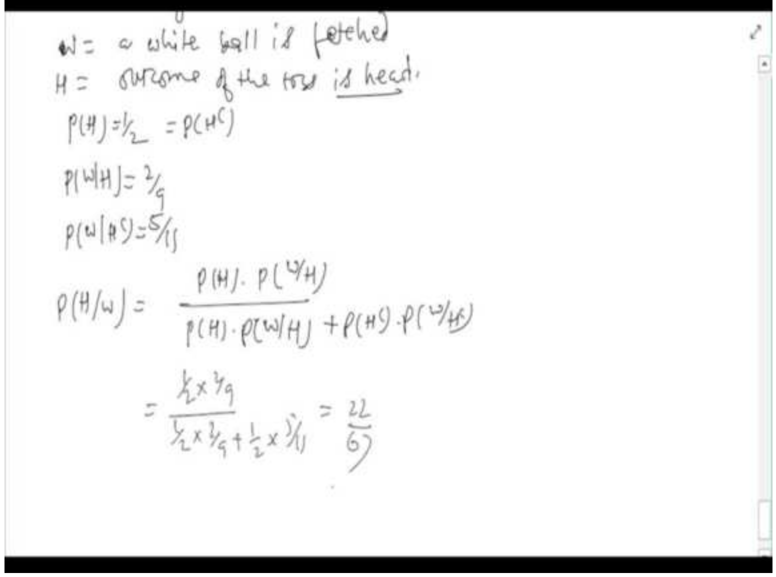
$$\mathrm{P(A|B)} = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(A \cap B) \cup P(A \cap B^C)} = \frac{P(B|A).P(A)}{P(B|A).P(A) + P(B|A^C).P(A^C)}.$$

This equality is referred to as Baye's rule; let us see an example.

(Refer Slide Time: 31:55)







Suppose there are two urns, first contains 2 white and 7 black balls; whereas the second one contains 5 white and 6 black balls. We also have a coin, so we flip a coin and draw a ball from the first urn; if the outcome of coin toss is head and draw a ball from the second coin otherwise, second urn otherwise.

We are interested in the probability that outcome of the toss was head, given that a white ball was fetched. What is the probability that the outcome of the coin toss was head, given that a white ball was fetched? To answer this question let us again define events W and H as follows. W is event that a white ball is fetched; and H is the event that outcome of the toss is head.

Clearly P(H) = $\frac{1}{2}$ =P(H^C). P(W|H) that is given that the outcome of coin toss was head. What is the probability that a white ball would be fetched will be $\frac{2}{9}$? Similarly, P(W|H^C) would be, there are 5 white balls and 6 black balls; so, = $\frac{5}{11}$.

Now, we have to compute probability that the outcome of coin toss was head, given that a white ball was fetched. So, P(H|W), this will be probability of W times probability of. This will be

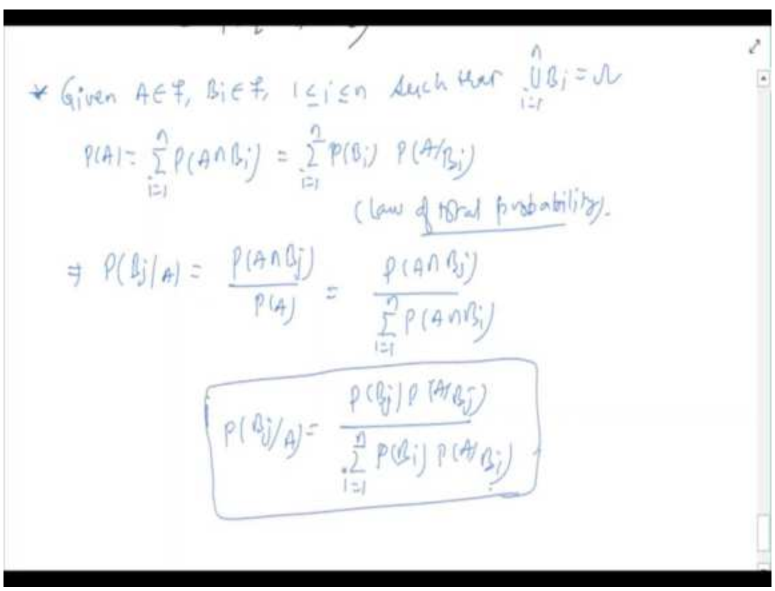
$$\mathrm{P(H|W)} = \frac{P(H).P(W|H)}{P(H).P(W|H) + P(H^C).P(W|H^C)}.$$

This will be

$$= \frac{\frac{1}{2} \times \frac{2}{9}}{\frac{1}{2} \times \frac{2}{9} + \frac{1}{2} \times \frac{5}{11}} = \frac{22}{67}.$$

So, this is an illustration of Bayes rule.

(Refer Slide Time: 36:16)



Let us now generalize Bayes rule to more than two events. Given an event A and events B_1 to B_n, so B_i, i varying from 1 to n, such that $\bigcup_{i=1}^{n} B_i = \Omega$. Probability of A can be written as

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(B_i)P(A|B_i).$$

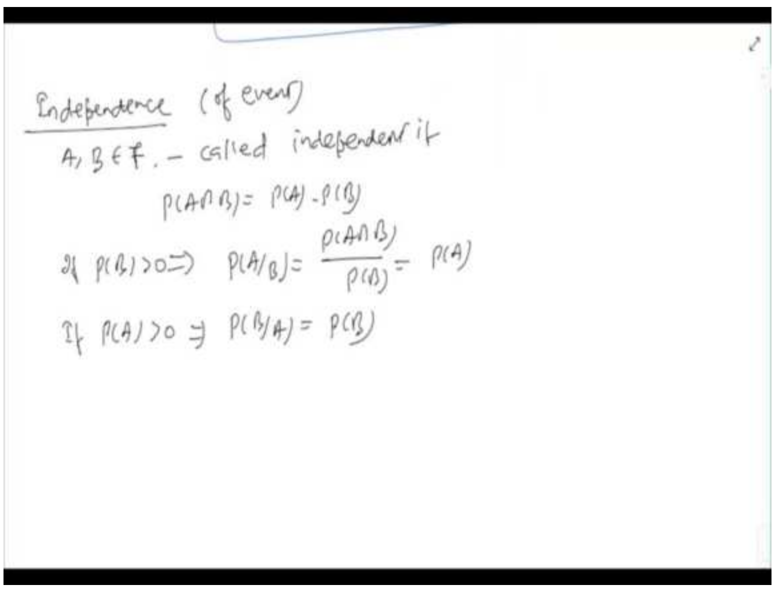
This equality is called law of total probability.

Next, I can write

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A \cap B_j)}{\sum_{i=1}^{n} P(A \cap B_i)} = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}.$$

This equality is generalization of Baye's rule to more than two events.

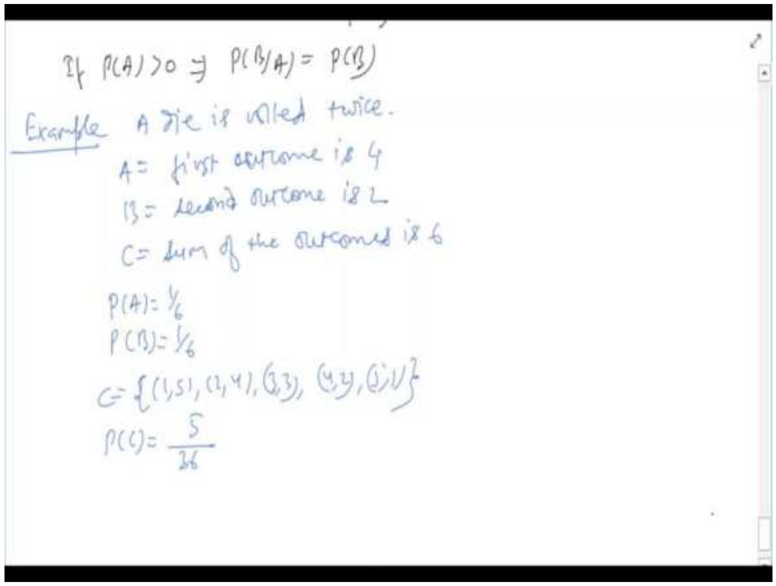
(Refer Slide Time: 38:23)



Let us now move to the next notion, that is of independence. In the forthcoming lectures we will study independence of random variables; so just to distinguish this notion, we explicitly say here independence of events. Let us define what we mean by independence of events. Say A and B are two events, these are called independent, if $P(A \cap B) = P(A).P(B)$.

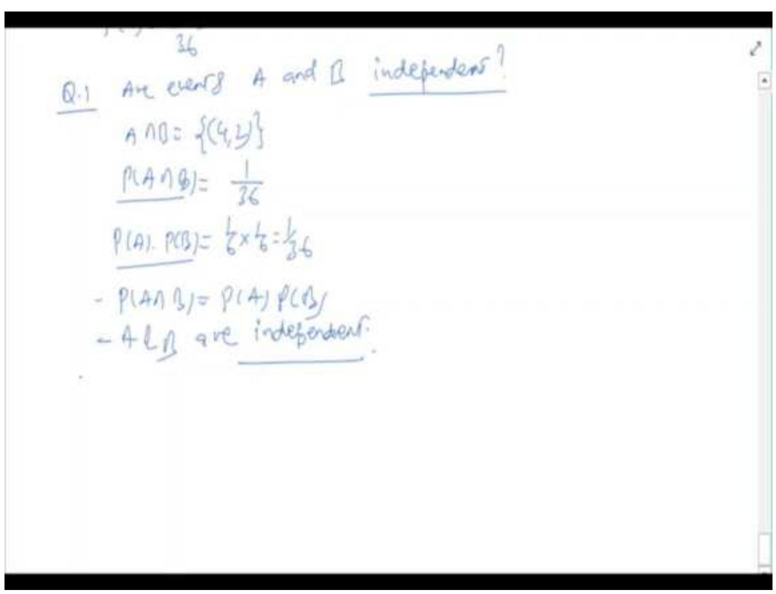
Notice that if P(B) is strictly positive, then A and B being independent implies that P(A|B), which is $\frac{P(A \cap B)}{P(B)} = P(A)$. Similarly, if P(A) > 0, then A and B being independent implies that P(B|A) = P(B). Let us see an example.

(Refer Slide Time: 39:43)



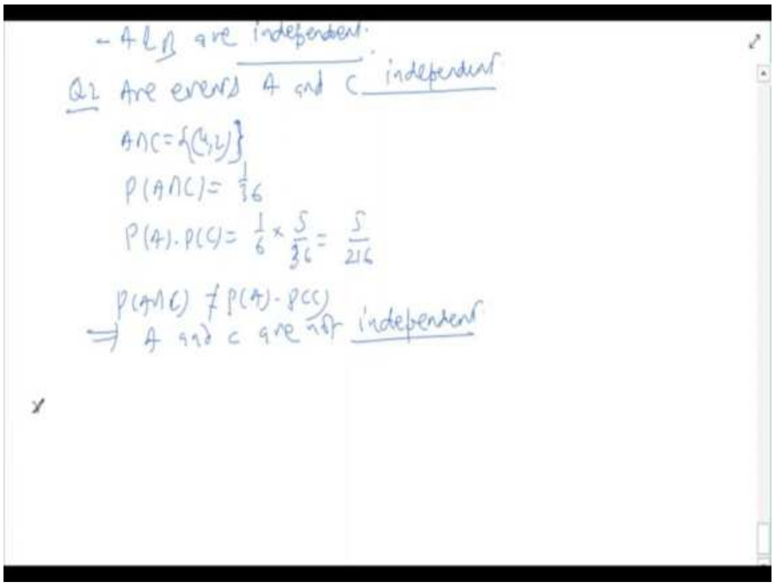
Consider the random experiment of A die being rolled twice, rolled twice. Let us define the following events, A is the event that first outcome is 4; B is the event that second outcome is 2; and C is the event that sum of the outcomes is 6. Then, clearly P(A) = $\frac{1}{6}$, P(B) = $\frac{1}{6}$, C ={(1,5),(2,4),(3,3),(4,2),(5,1)}. So, P(C) is the number of favorable outcomes is 5, and total number of outcomes in tossing a die twice is 36; so, P(C) = $\frac{5}{36}$.

(Refer Slide Time: 41:14)



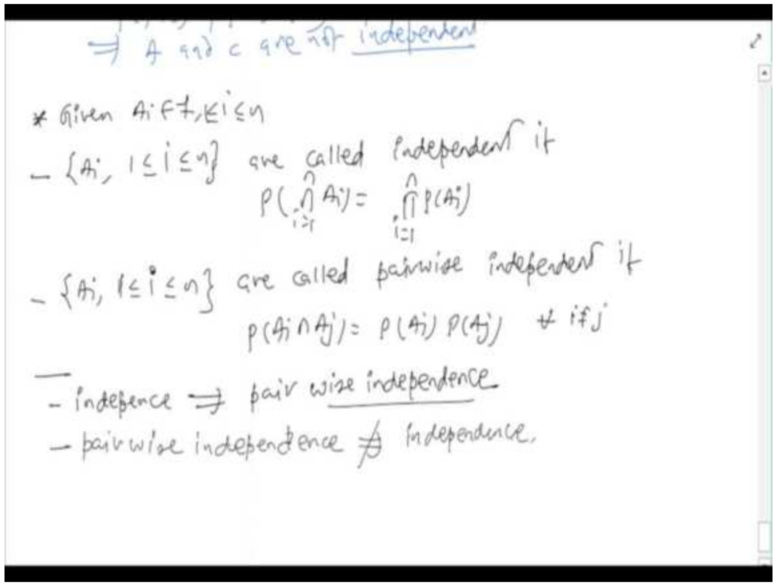
Let us ask the following questions. Are events A and B independent? To answer this question let us observe $A \cap B$. $A \cap B$ is the event in which the first outcome is 4 and second one is 2; so, this is $A \cap B$. P($A \cap B$) = $\frac{1}{36}$; only one of the 36 possible outcomes is favorable. On the other hand, P(A).P(B) $= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$. So, we see that P($A \cap B$) = P(A).P(B). So, A and B are indeed independent.

(Refer Slide Time: 42:24)



Now, let us ask another question, are events A and C independent? Again, if we see $A \cap C$, it turns out to be same singleton set {(4, 2)}; P($A \cap C$) = $\frac{1}{36}$. However, now P(A).P(C); it becomes $\frac{1}{6} \times \frac{5}{36} = \frac{5}{216}$. Clearly, P($A \cap C$) ≠ P(A).P(C); so, A and C are not independent. Similar calculation reveals that B and C are also not independent. Let us now extend this notion of independence to more than two events.

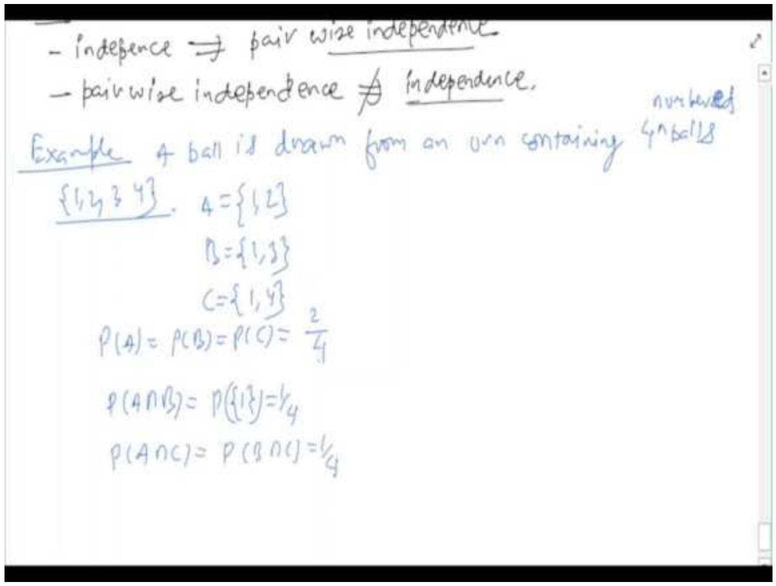
(Refer Slide Time: 43:34)

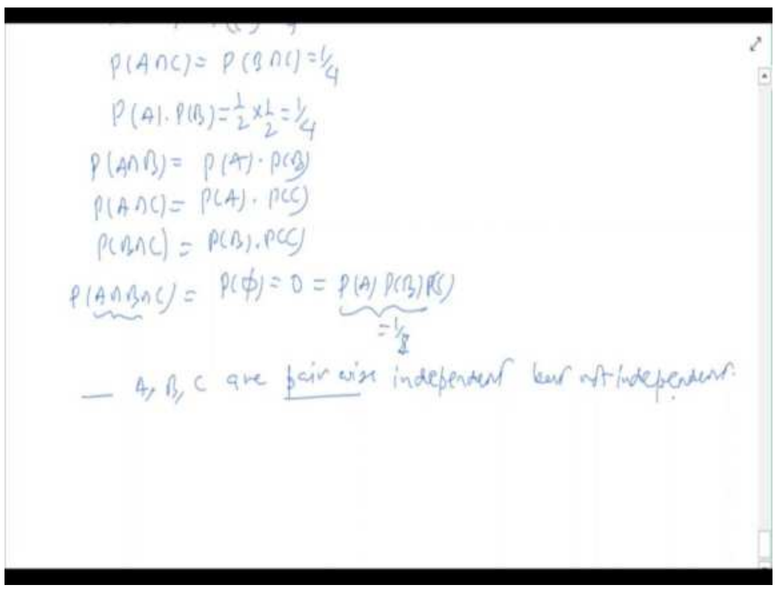


So, given events A_i, so i is changing again from 1 to n. This events A_i are called independent, if $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$. There is another notion here and that is a pairwise independence. So, A_i, i ranging from 1 to n are called pairwise independent, if $P(A_i \cap A_j) = P(A_i)P(A_j)$, for all pairs, this is for all $i \neq j$.

It can be seen that independence is a stronger notion that than pairwise independence. What we mean that independence implies pair wise independence. On the other hand, if a collection of events is pairwise independent; that does not mean that those events will be independent as well. Pairwise independence does not imply independence. Let us see an example to illustrate this point.

(Refer Slide Time: 45:43)



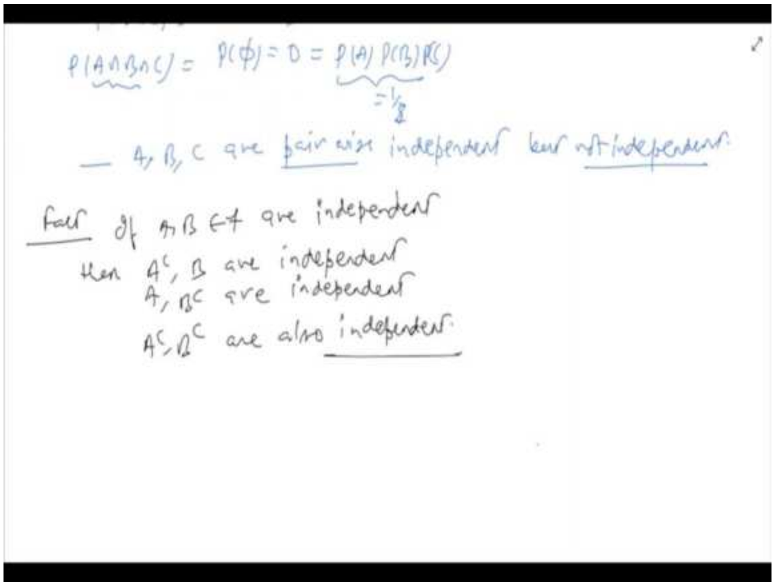


Suppose a ball is drawn from an urn containing 4 balls, and say these balls are numbered 1, 2, 3, 4. So, four numbered balls and define events A, B and C as follows. A is outcome is either 1 or 2, B is 1, 3, and C is 1, 4. In this case P(A)=P(B)=P(C)=$\frac{2}{4}$. Clearly, in each case out of four outcomes, there are two favorable outcomes.

Next P($A \cap B$), this is P({1})=$\frac{1}{4}$; and so are P($A \cap C$) and P($B \cap C$). Further, P(A)P(B) = $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. So, we see that P($A \cap B$) = P(A)P(B); and similarly, P($A \cap C$) = P(A)P(C); P($B \cap C$) = P(B)P(C).

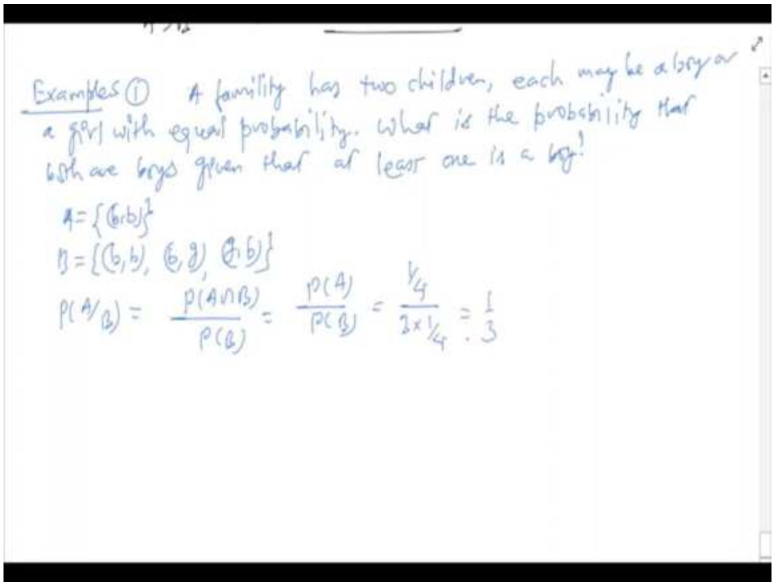
Finally, if we see P($A \cap B \cap C$) = P(Φ) = 0 $\neq$ P(A)P(B)P(C). Notice that this is $\frac{1}{8}$. So, clearly in this example A, B, C are pairwise independent, but not independent. We observe the following important fact about independence of events.

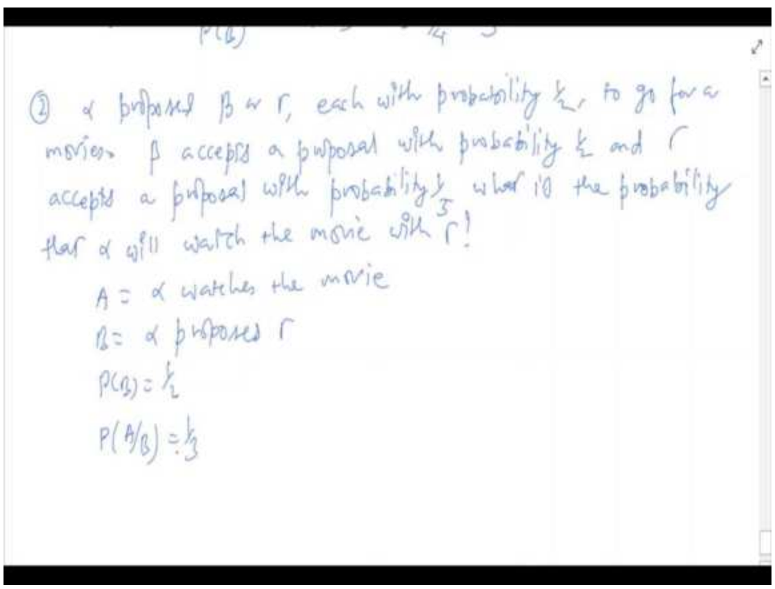
(Refer Slide Time: 48:39)

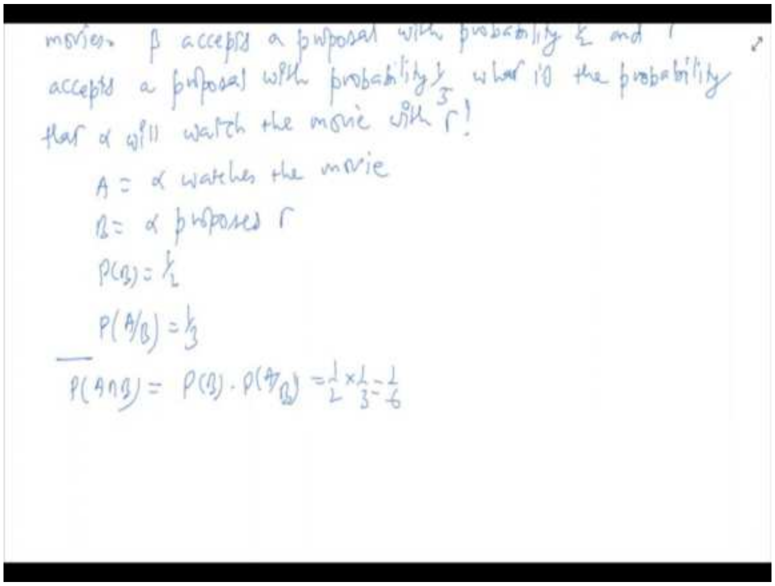


So, fact: If A and B events A and B are independent, then A^C to B are independent; A, B^C are independent, and so are A^C, B^C are also independent. Let us now consider a few more examples.

(Refer Slide Time: 49:27)







First example is suppose A family has two children, each maybe a boy or a girl with equal probability. We are interested in the conditional probability that both are boys given that at least one of them is a boy. What is the probability that both are boys given that at least one is a boy? We can again define events A and B as follows, A is the event that both are boys, and B is the event that at least one of them is boy; so, it is {(bb), (bg), (gb)}.

Then we are interested in P(A|B). We can use the conditional probability formula that says that

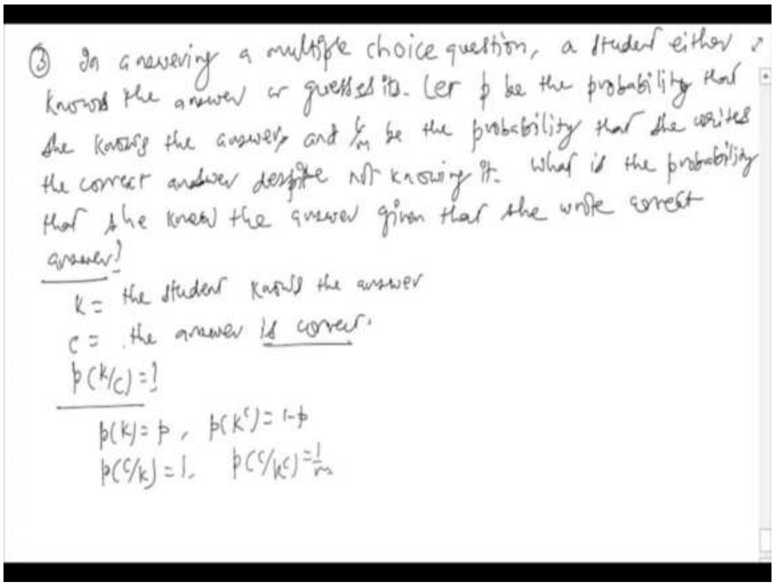
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{1/4}{3 \times 1/4} = \frac{1}{3}.$$

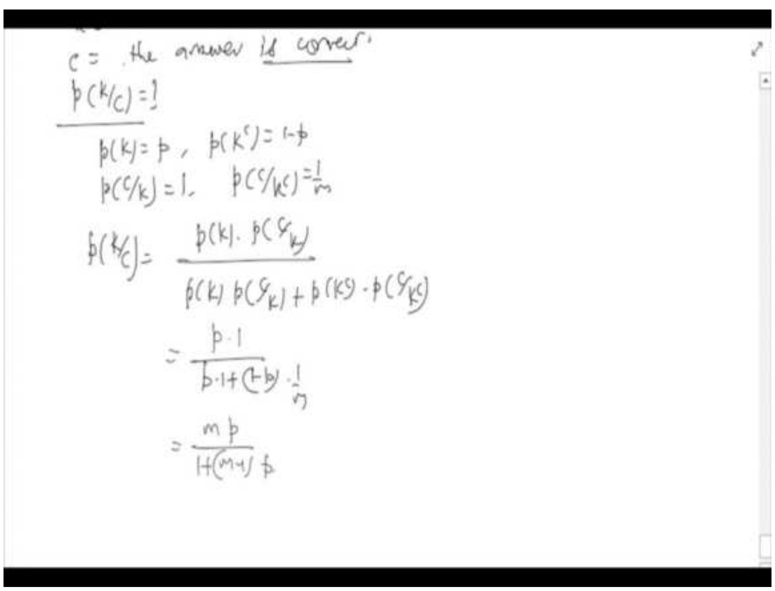
$P(A) = 1/4$; since each one maybe boy or girl with equal probability. Probability that both are boys is $1/2 \times 1/2$; that is $1/4$. And $P(B) = 1/4 + 1/4 + 1/4 = 3 \times 1/4$.

Let us see another example. Suppose a person α proposes a person's β or γ each with equal probability; so, each with probability half to go for a movie. Person β accepts a proposal with probability half, and γ accepts a proposal with probability $\frac{1}{3}$. We are interested in probability that α will watch the movie with γ. What is the probability that α will watch the movie with γ?

Again, we can define two events. Event A is α watches the movie and event B is α proposes γ; then probability of B is clearly is half. Given that α proposes γ, the conditional probability that A will watch movie with γ is $\frac{1}{3}$; because this is the probability with which γ accepts the proposal. We are interested in $P(A \cap B)$ that is α proposes γ and γ accepts the proposal. So, following chain rule, $P(A \cap B) = P(B)P(A|B)$. So, it is simply $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$.

(Refer Slide Time: 54:28)





Let us see one more example. In answering a multiple-choice question, a student either knows the answer, answer or guesses it; or if he does not know, if he or she does not know, then he or she guesses it. Let P be the probability that the student knows the answer, that she knows the answer;

and $\frac{1}{n}$ be the probability that the answer is correct. She writes the correct answer despite not knowing it, despite not knowing it; that is, she guesses correctly.

We are interested in the probability that student knew the answer, given that the answer is correct. What is the probability that she knew the answer given that she wrote correct answer? We can again define two events. K is the event that the student knows the answer, and C is the event that the answer is correct.

We are interested in P(K|C); towards this notice that P(K) = P, $P(K^C)$ = 1-P, P(C|K); that is the student answers correctly, given that she knows the answer is 1. And $P(C|K^C) = \frac{1}{m}$.

Now, we can use Baye's rule to compute

$$P(K|C) = \frac{P(K)\mathrm{P(C|K)}}{\mathrm{P(K)P(C|K) + P(K^{\mathit{C}})P(C|K^{\mathit{C}})}}.$$

We know all these probabilities. Let us substitute these

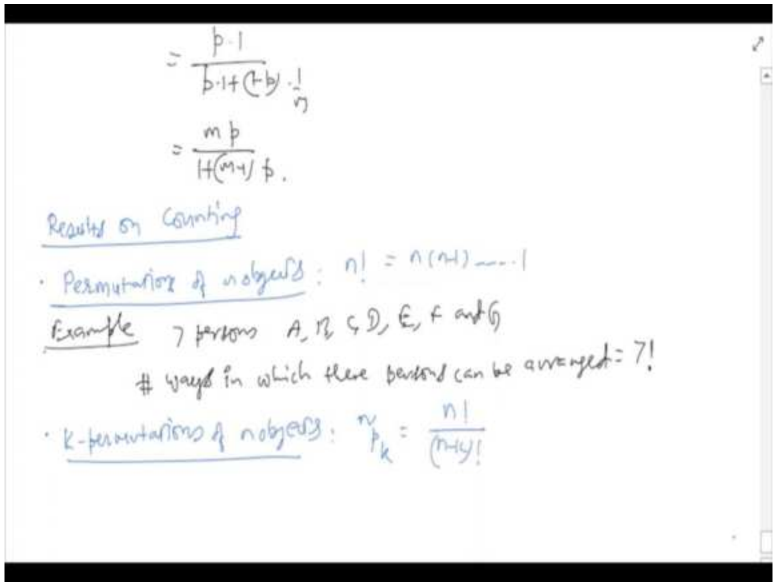
$$= \frac{P.1}{\mathrm{P.1 + (1 - P)}(\frac{1}{m})};$$

it turns out to be

$$= \frac{mP}{\mathrm{1 + (m - 1)P}}.$$

Before we end this lecture let us see a few results on counting.

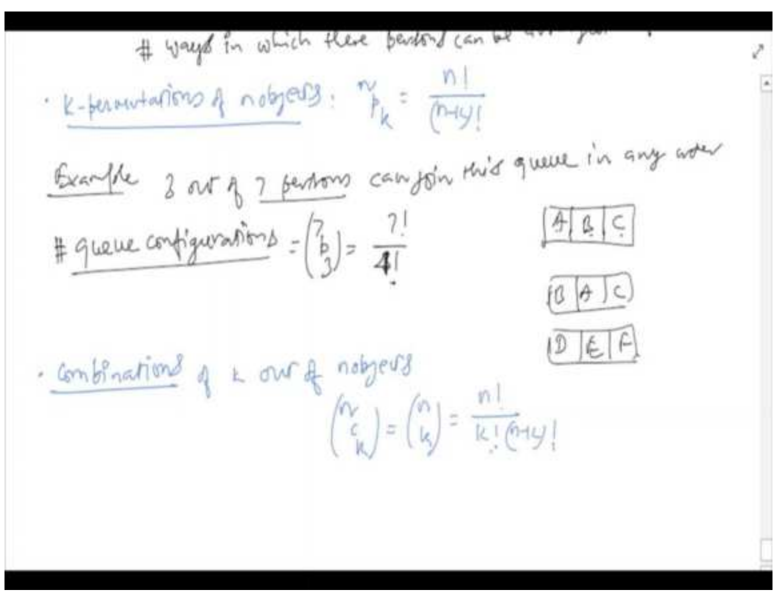
(Refer Slide Time: 58:43)



These results will be quite useful in computing probabilities; the first one is on permutations, permutations of n objects. Number of permutations of an object turns out to be $n! = n(n-1) \ldots 1$. For instance, suppose we have 7 persons A, B, C, D, E, F and G; and we want to arrange these people in a row. In how many ways we can arrange these people in a row?

That is exactly the number of which we can permute these letters A to G; number of ways in which these persons can be arranged, and so this is factorial 7. The next result is on K-permutations, K permutation of n objects; these are denoted by symbol $^{n}P_{k} = \frac{n!}{(n-k)!}$.

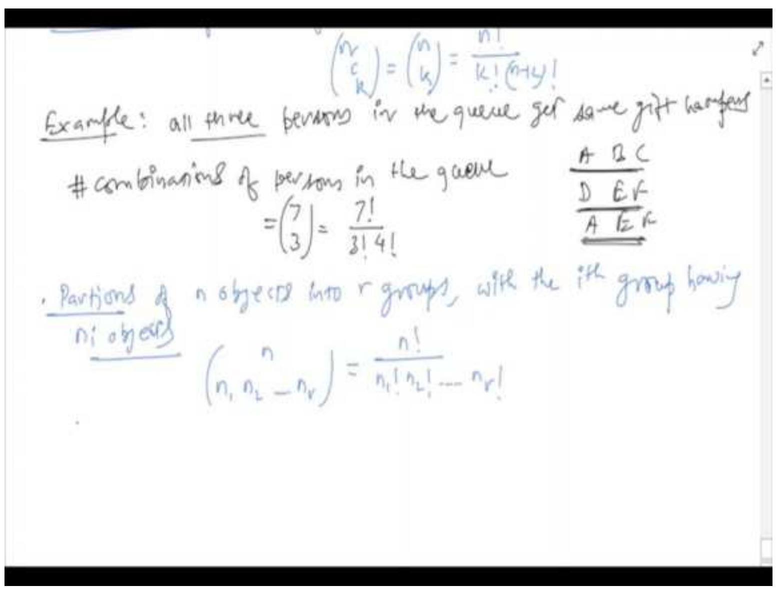
(Refer Slide Time: 1:00:42)



To see an example, suppose the seven persons in the above example have to join a queue; the queue has three positions. These seven persons can come and join in any order, 3 out of 7 persons can join this queue in any order. So, this queue may look like ABC or BAC or DEF. So, each of this we call a queue configuration. So, number of queue configurations will be equal to $^7P_3$, which will be $\frac{7!}{4!}$.

Notice that here we are also concerned about the order in which these persons appear in the queue. So, ABC is a different configuration than BAC. There are situations where we do not worry about the ordering; we only see the entities that are in the collection. And this notion is captured by combinations; we now look at combinations of K out of n objects. This is denoted by n choose k or sometimes by simply $^nk = \frac{n!}{k!(n-k)!}$.

(Refer Slide Time: 1:02:53)

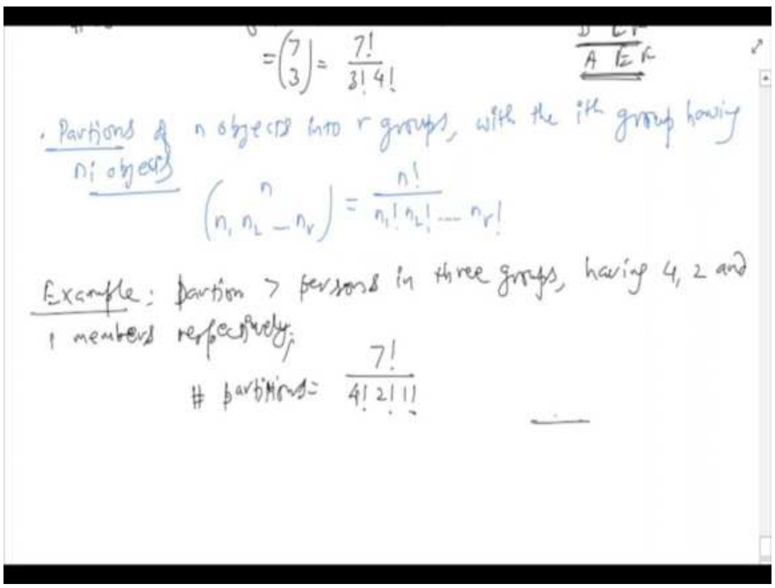


For instance, in the above example if all three persons in the queue get the same gift hampers; all three persons in the queue get same gift, similar gift hampers. So, either A, B, C may get gift hampers or D, E F or A, E, F. So, now we are interested in combinations of persons that get gift hampers; the ordering in the queue does not matter. So, A, B, C and B, A, C will not be counted separately.

So, what is the number of combinations of persons in the queue? This will be $= \binom{7}{3} = \frac{7!}{3!4!}$. The above example can be seen as if we are dividing the 7 people into two groups, those that are in queue and those that are not in queue. Basically, we want to see in how many ways we can partition with seven people in groups of 3 and 4 people. This idea could be generalized to more than two groups.

So, now we look at partitions to more than two groups, so partitions of n objects into r groups; and say i^{th} group can have n_i objects, with the i^{th} group having n_i objects. The number of possible partitions is denoted by following symbol $\binom{n}{n_1 n_2 \ldots . n_r} = \frac{n!}{n_1! n_2! \ldots . n_r!}$.

(Refer Slide Time: 1:05:28)



For instance, in the above example if we have to divide 7 persons in three groups; partition 7 persons in three groups, with these groups having say 4, 2 and 1 members respectively. Then, number of partitions $= \frac{7!}{4!2!1!}$. These were a few results about counting. This brings us to the end of this lecture. Thank you.