May 6 Lec—Independence, Bayes' Rules, Multivariable Coefficients, Discrete Random Variable

Independence

Definition: Independent events

Given a sample space S and a probability P, we say that two events A, B are **independent** $P(A \cap B) = P(A) \cdot P(B)$, or equivalently $P(A|B) = P(A)$ or $P(B|A) = P(B)$.

The equivalent definition makes sense because $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and if $P(A \cap B) = P(A) \cdot P(B)$ then $P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$.
Independence means that the chances of A occurring does not depend on the chances of B occurring.

Disjointness (mutual exclusion) is not independence! Two disjoint sets A, B have $P(A \cap B) = 0$. Since $P(A \cap B) = P(A) \cdot P(B)$, we see that one of $P(A)$ and $P(B)$ must be 0, i.e. either A or B is null. So two non-null, nontrivial events cannot be both independent and mutually exclusive.
Do not try to draw pictoral representations of independence—you can't!

Examples

1. Roll a fair die. Let event A be "Result is even," let event B be "Result is either 1 or 2 or 5," and let event C be "Result is either 1 or 2." (We would want to intuitively guess the events' independence or dependence, but don't.)

$$P(A) = P(B) = \frac{1}{2}, P(C) = \frac{1}{3}$$

$$P(A \cap B) = \frac{1}{6} \neq P(A) \cdot P(B)$$

so A, B are not independent.

On the other hand,

$$P(A \cap C) = \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = P(A) \cdot P(C)$$

So A, C are indeed independent (contrary to intuition).
Also,

$$P(B \cap C) = P(C)$$

and equivalently

$$P(B|C) = 1 \neq P(B)$$

so B, C are not independent.

2. A coin is such that $P(\{H\}) = \frac{2}{3}$ (implying $P(\{T\}) = \frac{1}{2}$. Flip the coin until H appears. Our sample space is countable:
   $S = \{w_1, w_2, \ldots, w_n, \ldots\}$ where $w_n = \underbrace{T \cdot \cdots \cdot T}_{n-1} \cdot H$. Intuitively, the different trials are independent, so let's assume that for this example.

$$P(\{w_1\}) = P(\{H\}) = \frac{2}{3}$$

$$P(\{w_2\}) = P(\{TH\}) = P(\text{tail on 1st flip} \cap \text{head on 2nd flip}) = \frac{1}{3} \cdot \frac{2}{3}$$

More generally, $w_n = \underbrace{T \cdot \cdots \cdot T}_{n-1} \cdot H$, so applying induction

$$P(\{w_n\}) = \left(\frac{1}{3}\right)^{n-1} \cdot \frac{2}{3} = 2 \cdot (\frac{1}{3})^n$$

which is the model we saw in an earlier lecture.

Properties of independence

1. $\forall A \subseteq S$, A and S are independent; so are A and $\emptyset$.
2. If A and B are independent, then so are A^c and B (analogously A and B^c). Therefore A^c and B^c are also independent.

Proof

1. Trivial.

2.

$$
\begin{aligned}
P(A^c \cap B) &= P(B \setminus A) \\
&= P(B) - P(A \cap B) \\
&= P(B) - P(A)P(B) \\
&= P(B) \cdot (1 - P(A)) \\
&= P(B) \cdot P(A^c)
\end{aligned}
$$

Example

Suppose that a population is subject to two diseases A and B. The prevalence (proportion of population carrying the disease) of A is 50% and the prevalence of B is 60%. Any member of the population can have disease A indpeendently of their status regarding disease B (and vice versa). What is the probability that a selected member of the population does not have either disease?
The answer is

$$P(A^c \cdot B^c) = P(A^c) \cdot P(B^c) = (1 - P(A)) \cdot (1 - P(B)) = 0.5 \cdot 0.4 = 0.2$$

Another way of doing this (not preferable because it's longer) is

$$1 - P(A \cup B) = 1 - (P(A) + P(B) - P(A \cap B))$$

Bayes' Rules

Definition: Partition

Let S be a sample space. We say that $A_1, A_2, \ldots, A_n, \ldots$ form a **partition** of S if the A_i's are pairwise disjoint and

$$S = \bigcup_n A_n$$

Theorem

Let S be a sample space and $E_1, E_2, \ldots, E_N$ be a finite partition of S. Then for any event A we have:

1. $$P(A) = \sum_{k=1}^{N} P(A|E_k) \cdot P(E_k)$$

2. For any fixed $n \in \{1, 2, \ldots, N\}$ we have
$$P(E_n|A) = \frac{P(A|E_n) \cdot P(E_n)}{\sum_{k=1}^{N} P(A|E_k) \cdot P(E_k)}$$

Proof

2 follows from 1 by $P(E_n|A) = \frac{P(A \cap E_n)}{P(A)}$.
We prove 1:

$$A = A \cap S = A \cap (\bigcup_{k=1}^{N} E_k)$$

$$= \bigsqcup_{k=1}^{N} (A \cap E_k)$$

We have that the $A \cap E_k$ are pirwise disjoint. So

$$P(A) = \sum_{k=1}^{N} P(A \cap E_k) = \sum_{k=1}^{N} P(A|E_k) P(E_k)$$

Examples

Urn A contains 3 red balls and 4 blue balls. Urn B contains 4 red balls and 3 blue balls. Draw a ball from A, put it in B, and a ball from B. Let $A := \text{ball drawn from B is blue}$. What is $P(A)$?
We partition the sample space with events

$$E_R := \text{ball drawn from urn A is red}$$

$$E_B := \text{ball drawn from urn A is blue}$$

We calculate:

$$P(E_R) = \frac{3}{7}, P(E_B) = \frac{4}{7}$$

$$P(A|E_R) = \frac{3}{8}, P(A|E_B) = \frac{4}{8}$$

so $P(A) = P(A|E_R) \cdot P(E_R) + P(A|E_B) \cdot P(E_B)$

$$= \frac{3}{8} \cdot \frac{3}{7} + \frac{4}{8} \cdot \frac{4}{7} = \frac{25}{56}$$

Part 2 of the same question:
What is the probability that the ball originally drawn from urn A is red, given that the ball later drawn from urn B is blue?

$$P(E_R|A) = \frac{P(A|E_R) \cdot P(E_R)}{P(A)} = \frac{9}{25}$$

Example 2:
Refer to 2.132 from the textbook.
A missing plane has equal probability of going tdown in 3 regions R_1, R_2, R_3 (i.e. the probability of the plane going down in any one of the regions is $\frac{1}{3}$). Let $1 - \alpha_i$ be the probability of finding the plane if it went down in R_i. Find the probability that the plane is in R_2 given that the search in R_1 is unsuccessful.
Let A be the event that the search in R_1 is unsuccessful. We want

$$P(R_2|A) = \frac{P(A|R_2) \cdot P(R_2)}{P(A)}$$

and

$$P(A) = P(A|R_1) \cdot P(R_1) + P(A|R_2) \cdot P(R_2) + P(A|R_3) \cdot P(R_3)$$

Using some common sense:
$P(A|R_1)$ is the probability that the search in R_1 is unsuccessful given that the plane is in R_1, i.e. $1 - (1 - \alpha_1) = \alpha_1$.
$P(A|R_2)$ is the probability that the search in R_1 is unsuccessful given that the plane actually went down in R_2, and this probability is (what

are you thinking otherwise???) 1.
By the same reasoning, $P(A|R_3)$ is also equal to 1.
So

$$P(A) = \alpha_1 \cdot \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{1}{3}(a_1 + 2)$$

$$P(A|R_2) = 1$$

$$P(R_2) = \frac{1}{3}$$

$$\implies P(R_2|A) = \frac{P(A|R_2) \cdot P(R_2)}{P(A)} = \frac{1}{a_1 + 2}$$

Similarly (try it out yourself) the probability that the plane is in R_1 given that the search in R_1 is unsuccessful is $\frac{\alpha_1}{2+\alpha_1}$.

Multinomial Coefficients

Let S be a set such that $card(S) = N \in \mathbb{N}$.

Question 1 (review of the binomial coefficient): In how many ways can we split S into 2 subsets A_1, A_2 such that $card(A_1) = n_1, card(A_2) = n_2 := N - n_1$?
We know that (with labeled subsets) A_1, A_2 the answer is
$C_{n_1}^N = \frac{N!}{n_1!(N-n_1)!} = \frac{N!}{n_1!n_2!}$. (Keep in mind that if the order of A_1, A_2 doesn't matter we would get fewer choices than this.)

Question 2 (our main question): Same as above, but split S into 3 subsets of size n_1, n_2, n_3 that sum up to N. The answer is $C_{n_1}^N \cdot C_{n_2}^{N-n_1}$ (the third subset is uniquely determined at this point, no need to include another factor $C_{n_3}^{N-n_1-n_2}$).
Simplifying, we get

$$C_{n_1}^N \cdot C_{n_2}^{N-n_1} = \frac{N!}{n_1!(N-n_1)!} \cdot \frac{(N-n_1)!}{n_2!(N-n_1-n_2)!}$$

$$= \frac{N!}{n_1!n_2!n_3!}$$

We generalize this result:

Formula of the multinomial coefficient

The number of ways in which we can split a set S with N elements into k subsets with $n_1, n_2, \dots, n_k$ elements is denoted

$$\begin{bmatrix} N \\ n_1, n_2, \dots, n_k \end{bmatrix} = \frac{N!}{n_1! n_2! \dots n_k!}$$

(digression into the Multinomial Theorem which we won't use)

Example

A manager wants to divide a group of 10 employees into 3 teams of 3, 3, and 4 employees. Paul and John are two employees that hate each other to the bones. If the splitting is done at random, what is the probability that Paul and John are not on the same team?
Let A be the probability that these two are on the same team.

$$P(A^c) = \frac{\begin{bmatrix} 8 \\ 1,3,4 \end{bmatrix} + \begin{bmatrix} 8 \\ 3,1,4 \end{bmatrix} + \begin{bmatrix} 8 \\ 3,3,2 \end{bmatrix}}{\begin{bmatrix} 10 \\ 3,3,4 \end{bmatrix}}$$

End of Chapter 2!

Discrete Random Variable

Definition

Let S be a sample space. A random variable is a function $X : S \to \mathbb{R}$. The range of X, denoted $X(S)$, is called the support of the distribution of X.

Examples

1. Roll a die twice, and let X be the sum of the numbers obtained. $S = S_1 \times S_2$ where $S_1 = S_2 = \{w_1, w_2, w_3, w_4, w_5, w_6\}$, and $X(w_i, w_j) = i + j$

.

$X(S) = \{2, \ldots, 12\}$

2. Flip a coin until H appears. Let X be the number of trials needed. $S = \{w_1, w_2, \ldots, w_n, \ldots\}$ (a countably infinite set). So $X(w_1) = 1, X(w_2) = 2, \ldots, X(w_n) = n, \ldots$ So $X(S) = \mathbb{N}^* = \mathbb{N} \setminus \{0\}$.

Definition: Discrete random variable

A random variable X is said to be discrete if $X(S)$ is a discrete (i.e. finite, or countably infinite) subset of $\mathbb{R}$.

Definition: Probability function

Let S be a sample space, P a probability on $\mathcal{P}(S)$ and X be a discrete random variable. The probability function of X, denoted p_X, is a function defined on $\mathbb{R}$ as follows: $\forall x \in \mathbb{R}$, define the event $\{X = x\} = \{\omega \in S | X(\omega) = x\}$, and we have $p_X(x) = P(\{X = x\})$.

Remark

If $x \notin X(S)$, then $\{X = x\} = \emptyset$ and $p_X(x) = 0$. So $p_X(x) = 0$ whenever $x \notin X(S)$.

Examples

Roll a *fair* die twice. Let X be the sum of the numbers rolled.
$X(S) = \{2, \ldots, 12\}$.
$\forall x \in \mathbb{R}, x \notin X(S)$, we have $p_X(x) = 0$.
Let us construct a table for $x \in S$.

x	$p_X(x)$
2	$\frac{1}{36}$
3	$\frac{2}{36}$

x	$p_X(x)$
4	$\frac{3}{36}$
5	$\frac{4}{36}$
6	$\frac{5}{36}$
7	$\frac{6}{36}$
8	$\frac{5}{36}$
9	$\frac{4}{36}$
10	$\frac{3}{36}$
11	$\frac{2}{36}$
12	$\frac{1}{36}$