Combinatorics

Definition 2

The number of ways of choosing r items from n distinct items (in any order) is

$$^nC_r = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

Combinatorics

Definition 2

The number of ways of choosing r items from n distinct items (in any order) is

$$^{n}C_r = \frac{n!}{r!(n-r)!} = \binom{n}{r}$$

Definition 3 (combination)

*We called this type of selections, i.e. one that does not considering the order or the arrangement important, **combinations**.*

How many ways can a committee of 6 people be chosen from 10 people?

How many different poker hands are there? i.e. how many different choices are there of drawing 5 cards from a stand deck of 52 playing cards?

Practice

How many ways can a group of 4 people be chosen from 6 men and 4 women if there must be 2 people of each gender in the group? [Hint: use Theorem 1]

Practice

As you will find throughout this course, there will often be multiple ways of arriving to the same conclusion

Practice

As you will find throughout this course, there will often be multiple ways of arriving to the same conclusion

In how many different ways can six tosses of a coin yield two heads and four tail?

Direct enumeration

Using Definition 2

Using Theorem 4

Probability & Counting

In a lottery, 6 numbers are drawn from a drum with 49 numbers without replacement. How many different winning lines are possible?

Bayes Theorem

Revisit Conditional Probability

- We have seen the **AND Rule**: if E and F are two events then

$$P(E \cap F) = P(E)P(F \mid E),$$

where $F \mid E$ means the occurrence of an event F given that an event E has already occurred.

Revisit Conditional Probability

- We have seen the **AND Rule**: if E and F are two events then

$$P(E \cap F) = P(E)P(F \mid E),$$

  where $F \mid E$ means the occurrence of an event F given that an event E has already occurred.
- Dividing both sides of the above equation by $P(E)$ gives us the definition of a conditional probability.

Revisit Conditional Probability

- We have seen the **AND Rule**: if E and F are two events then

$$P(E \cap F) = P(E)P(F \mid E),$$

  where $F \mid E$ means the occurrence of an event F given that an event E has already occurred.

- Dividing both sides of the above equation by $P(E)$ gives us the definition of a conditional probability.

- The probability that an event F occurs given that an event E has already occurred is given by

$$P(F \mid E) = \frac{P(E \cap F)}{P(E)}.$$

Conditional Probability II

- Using the fact that $P(E \cap F) = P(F \cap E)$ and noting that $P(F \cap E) = P(E \mid F)P(F)$, we can write the expression for $P(F \mid E)$ as follows.

$$P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} \tag{1}$$

Conditional Probability II

- Using the fact that $P(E \cap F) = P(F \cap E)$ and noting that $P(F \cap E) = P(E \mid F)P(F)$, we can write the expression for $P(F \mid E)$ as follows.

$$P(F \mid E) = \frac{P(E \mid F)P(F)}{P(E)} \tag{1}$$

- Now, we can also rewrite the term $P(E)$ in this equation. However, first we will look at an example to help motivate the situation.

Conditional Probability Example

Suppose two cards are dealt from a well-shuffled deck of 52 cards.

What is the probability that the second card is black?

The Partition Theorem

What we are actually used in the last example is The Partition Theorem. This approach can be generalized to get a formula for $P(E)$ in terms of conditional probabilities.

The Partition Theorem

What we are actually used in the last example is The Partition Theorem. This approach can be generalized to get a formula for $P(E)$ in terms of conditional probabilities.

Theorem 5 (AKA Rule of total probability)

The Partition Theorem*: If $F_1, F_2, \ldots, F_n$ are mutually exclusive events of which one must occur, then*

$$P(E) = \sum_{i=1}^{n} P(E \mid F_i) P(F_i)$$

The Partition Theorem

What we are actually used in the last example is The Partition Theorem. This approach can be generalized to get a formula for $P(E)$ in terms of conditional probabilities.

Theorem 5 (AKA Rule of total probability)

The Partition Theorem: *If $F_1, F_2, \ldots, F_n$ are mutually exclusive events of which one must occur, then*

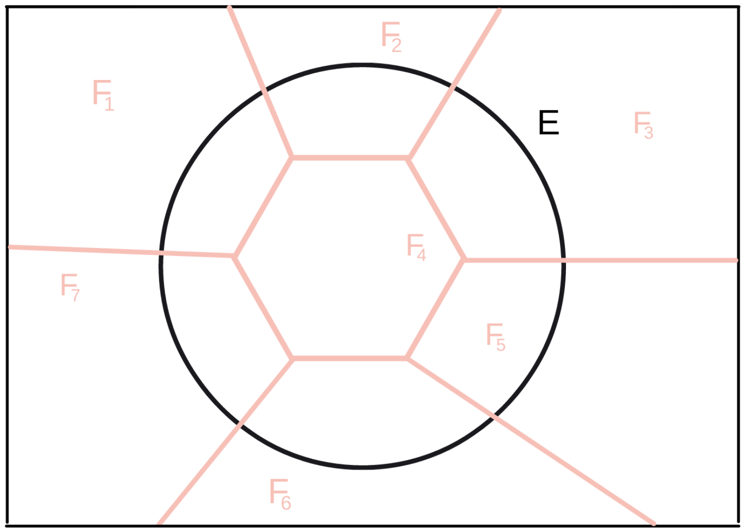
$$P(E) = \sum_{i=1}^{n} P(E \mid F_i)P(F_i)$$

Partition

- The events $F_1, F_2, \ldots, F_n$ make up a partition of the sample space of an experiment.
- Another way to say this is that $F_1, F_2, \ldots, F_n$ mutually exclusive and exhaustive events.
- Recall that $F_1, F_2, \ldots, F_n$ are mutually exclusive if no two have any common outcomes.
- The events are exhaustive if one F_i must occur. That is to say,

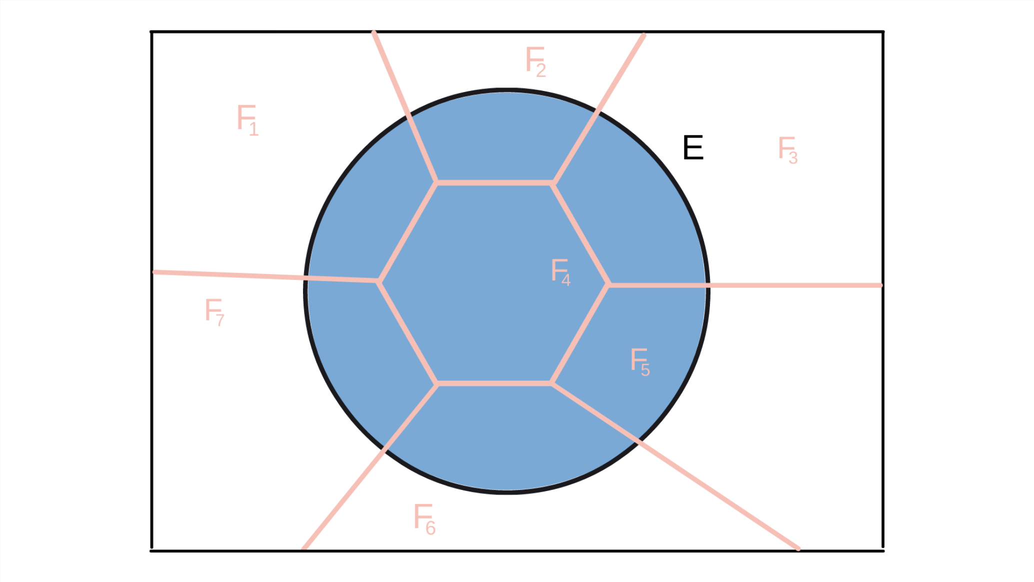
$$S = F_1 \cup F_2 \cup \cdots \cup F_n$$

Visualization of the Partition Theorem



$$
\begin{aligned}
P(E) &= P(E \mid F_1) \cdot P(F_1) + \cdots + P(E \mid F_7) \cdot P(F_7) \\
&= \sum_{i=1}^{7} P(E \mid F_i) \cdot P(F_i)
\end{aligned}
$$

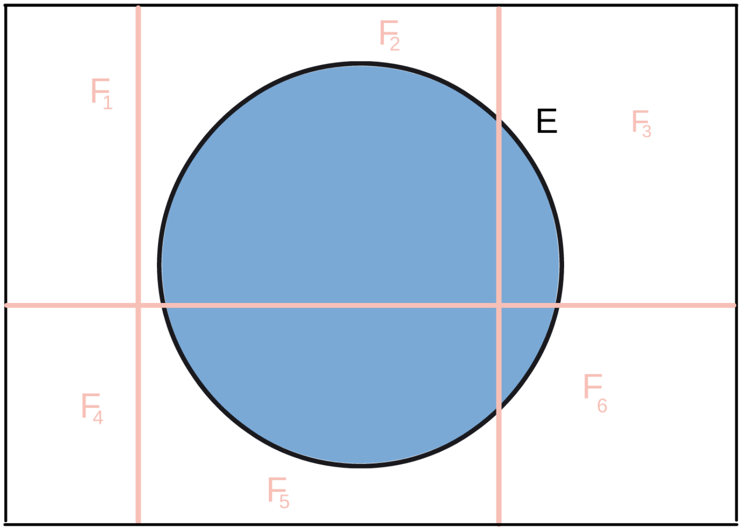
Visualization of the Partition Theorem



$$P(E) = P(E \mid F_1) \cdot P(F_1) + \cdots + P(E \mid F_7) \cdot P(F_7)$$

$$= \sum_{i=1}^{7} P(E \mid F_i) \cdot P(F_i)$$

Visualization of the Partition Theorem



$$P(E) = P(E \mid F_1) \cdot P(F_1) + \cdots + P(E \mid F_7) \cdot P(F_6)$$
$$= \sum_{i=1}^{6} P(E \mid F_i) \cdot P(F_i)$$

Example 4

In a region, 31% of people are smokers, 19% of smokers develop lung cancer and 2% of non-smokers develop lung cancer. What is the probability that a person chosen at random will develop lung cancer; i.e. what percentage of this region will suffer from lung cancer? *Hint: It may be useful to represent this information in a tree diagram*

Example 4

In a region, 31% of people are smokers, 19% of smokers develop lung cancer and 2% of non-smokers develop lung cancer. What is the probability that a person chosen at random will develop lung cancer; i.e. what percentage of this region will suffer from lung cancer? Hint: It may be useful to represent this information in a tree diagram

Given, $P(S) = 0.31$

Example 4

In a region, 31% of people are smokers, 19% of smokers develop lung cancer and 2% of non-smokers develop lung cancer. What is the probability that a person chosen at random will develop lung cancer; i.e. what percentage of this region will suffer from lung cancer? Hint: It may be useful to represent this information in a tree diagram

Given, $P(S) = 0.31 \rightarrow P(\overline{S}) = 1 - 0.31 = 0.69$ (compliment rule)

Example 4

In a region, 31% of people are smokers, 19% of smokers develop lung cancer and 2% of non-smokers develop lung cancer. What is the probability that a person chosen at random will develop lung cancer; i.e. what percentage of this region will suffer from lung cancer? Hint: It may be useful to represent this information in a tree diagram

Given, $P(S) = 0.31 \rightarrow P(\overline{S}) = 1 - 0.31 = 0.69$ (compliment rule)

Given, $P(B \mid A) = 0.19$ and $P(B \mid \overline{A}) = 0.02$

Example 4

In a region, 31% of people are smokers, 19% of smokers develop lung cancer and 2% of non-smokers develop lung cancer. What is the probability that a person chosen at random will develop lung cancer; i.e. what percentage of this region will suffer from lung cancer? *Hint: It may be useful to represent this information in a tree diagram*

Given, $P(S) = 0.31 \rightarrow P(\overline{S}) = 1 - 0.31 = 0.69$ (compliment rule)

Given, $P(B \mid A) = 0.19$ and $P(B \mid \overline{A}) = 0.02$

$$
\begin{aligned}
P(B) &= P(B \cap A) + P(B \cap \overline{A}) && \text{(partition rule)} \\
&= (0.19)(0.31) + (0.02)(0.69) \\
&= 0.0589 + 0.0138 = 0.0727
\end{aligned}
$$

Why?

- Why is it the case that $P(B) = P(B \cap A) + P(B \cap \overline{A})$?
- There are two ways of reasoning this.
- There are two groups that suffer from cancer – smokers and non-smokers – so we add the probabilities accordingly (this can be seen best using a tree diagram).
- Or, in more technical language, the set the people in the study can be **partitioned** by A $(=F_1)$ and $\overline{A}$ $(= F_2)$, allowing us to use Theorem 5.

Bayes' Theorem

Theorem 6 (Bayes' Theorem)

If $F_1, F_2, \ldots, F_n$ are mutually exclusive events of which one must occur, and $P(F_i) \neq 0$ for $i = 1, 2, ..., n$ then for any event E for which $P(E) > 0$

$$P(F_j \mid E) = \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i)P(F_i)}, \quad j = 1, 2, \ldots, n.$$

$$P(F_j \mid E) \overset{\text{cond prob}}{=} \frac{P(E \cap F_j)}{P(E)} \overset{\text{mult. rule}}{=} \frac{P(E \mid F_j)P(F_j)}{P(E)}$$

$$\overset{\text{Thrm 5}}{=} \frac{P(E \mid F_j)P(F_j)}{\sum_{i=1}^{n} P(E \mid F_i) \cdot P(F_i)}$$

Bayes' Theorem Example

Returning to the lung cancer example, work out the probability a person with lung cancer also smokes

Clicker 3

How do we express this probability **in symbols**. *Assume A is the event that the person smokes and B is the event that the person develops lung cancer?*

1. $P(A \mid B)$
2. $P(B \mid A)$
3. $P(A \cup B)$
4. $P(A \cap B)$
5. none of the above

Bayes' Theorem Example

Returning to the lung cancer example, work out the probability a person with lung cancer also smokes

Clicker 3

How do we express this probability **in symbols**. *Assume A is the event that the person smokes and B is the event that the person develops lung cancer?*

1. $P(A \mid B)$
2. $P(B \mid A)$
3. $P(A \cup B)$
4. $P(A \cap B)$
5. none of the above

$$
\begin{aligned}
P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid \overline{A})P(\overline{A})} \\
&= \frac{(0.19)(0.31)}{(0.19)(0.31) + (0.02)(0.69)} \\
&= 0.8102
\end{aligned}
$$

Bayes' Theorem Example

Example 5

The members of a firm rent cars from three rental agencies: 60% from agency 1, 30% from agency 2, and 10% from agency 3. If 9% of cars from agency 1 need a tune-up, 20% of the cars from agency 2 need a tune-up, and 6% of cars from agency 3 need a tune-up, what is the probability that a rental car delivered to the firm needs a tune-up?

Bayes' Theorem Example

Example 6

Referring back to Ex 5, if we discover that the rental car does in fact need a tune-up. What is the probability the delivered car is from agency 2.

Try on your own

Exercise 1 Personal computers are assembled on two production lines, 60% are assembled on Line 1 and 40% on Line 2. QC records show that both lines are not equally reliable: 95% of units assembled by Line 1 require no rework, while the figure for Line 2 is 88%.

1. What percentage of all computers require rework?
2. If a computer is found to require rework, what is the probability that it came from Line 1?