Outcomes, Sample Space

- An outcome is a result of an experiment. I An experiment means any action that can have a number of possible results, but which result will actually occur cannot be predicted with certainty prior to the experiment. *e.g.* Tossing of a coin.
- The set of all possible outcomes of an experiment is the sample space or the outcome space.
- A set of outcomes or a subset of the sample space is an event.

Outcomes, Sample Space

- Toss of a coin. Sample space is $S = \{H, T\}$
  The model for probabiliteis is $P(H) = 0.5, P(T) = 0.5$.
- Roll a dice. Sample space is $S = \{1, 2, 3, 4, 5, 6\}$
  with $P(i) = \frac{1}{6}$ for $i = 1, 2, \ldots 6$.
- Toss a quarter, a dime and a nickle together. Sample space is $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$
  Reasonable to model that the outcomes are equally likely. So each outcome carries probability $\frac{1}{8}$.

Outcomes, Sample Space

Again toss a quarter, a dime and a nickle together, and concentrate only on the number of heads. The sample space now is $S = \{0, 1, 2, 3\}$.

Do we have reason to model these as equally likely?

A little thought would convince us otherwise. An experiment would also reveal the same: if we toss the three coins 100 times we would observe that 1, 2 occur far more than 0, 3. So the events are not equally likely.
How do we assign the probabilities?

0 heads if TTT occurs
1 head if HTT, THT or TTH occurs
2 heads if HHT, HTH or THH occurs
3 heads if HHH occurs

$P(0) = P(3) = \frac{1}{8},\ P(1) = P(2) = \frac{3}{8}$

Outcomes, Sample Space, Events

When an experiment **E** results in m equally likely outcomes $\{e_1, e_2, \ldots, e_m\}$, probability of any event A is simply

$$P(A) = \frac{\#A}{m}$$

which is often read as ratio of number of favorable outcomes and the total number of outcomes.

The basic principle of counting

Two state Experiment, Stage A and Stage B
Stage A can result in any one of m possible outcomes.
For each outcome of A there are n possible outcomes for stage B.
Then together there are mn possible outcomes.

$$\begin{array}{cccc} (1,1), & (1,2), & \ldots, & (1,n) \\ (2,1), & (2,2), & \ldots, & (2,n) \\ \vdots & & & \\ (m,1), & (m,2), & \ldots, & (m,n) \end{array}$$

Events

Event: A subset of the sample space (technical restrictions unimportant for now)

Often use early-alphabet capital letters (*e.g.*, A, B, C) for events.

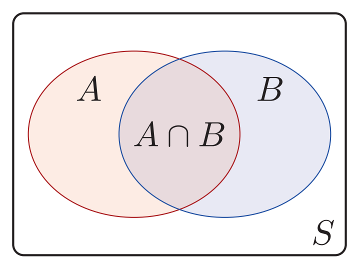
What is the difference between an outcome and an event?

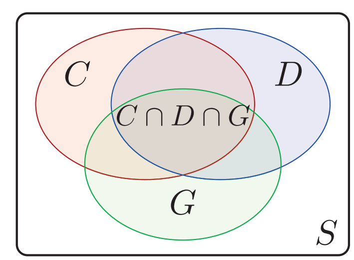
Sets, subsets

- Null event and full sample space: $\emptyset$ and S
- Subset notation: $A \subset B$ or, alternatively, $A \supset B$ or A implies B
- Union: $A \cup B$
- Intersection: $A \cap B$ or simply AB
- Complement: A' or, alternatively, A^c

- Mutually exclusive events: Have empty pairwise intersection
- Exhaustive events: Have union equal to S
- Partition: A group of mutually exclusive and exhaustive events

Venn Diagrams





Algebra of sets

Commutative property: $A \cup B = B \cup A, \quad A \cap B = B \cap A$

Associative property:

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

Distributive property:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

DeMorgan's Laws:

$$(A \cup B)^c = A^c \cap B^c$$
$$(A \cap B)^c = A^c \cup B^c$$

More generally,

$$\left(\bigcup_{i=1}^{n} A_i\right)^c = \bigcap_{i=1}^{n} A_i^c, \text{ and } \left(\bigcap_{i=1}^{n} A_i\right)^c = \bigcup_{i=1}^{n} A_i^c$$

Example

Roll two dice, so the sample space S has 36 elements or outcomes.

$$\begin{matrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{matrix}$$

$A = \{s \in S$: sum of the results of the two dice in s is even$\}$
$B = \{$first die is even$\}$ (using abbreviated notation)
$C = \{$second die is $> 5\}$
$D = \{$sum is odd$\}$
What is $A \cap D$?
$(B \cup C^c \cup A)^c$?
$(A^c \cap B) \cup (A \cap B)$?

Axioms of Probability

Axiom 1: For all $A \subset S, \quad P(A) \geq 0$ (Nonnegativity)

Axiom 2: $P(S) = 1$

Axiom 3: Whenever $A_1, A_2, \ldots$ are mutually exclusive,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad \text{(Countable additivity)}$$

Simple propositions

P1: $P(\emptyset) = 0$
Proof: If $A_i = \emptyset$ for all i, then A_i are mutually exclusive. So by Axiom 3, $P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$. This is meaningless unless $P(\emptyset) = 0$.

P2: If $A_1, A_2, \ldots, A_n$ are mutually exclusive, then by **P1**,
$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$ (Finite addititivity)
Proof: Follows from Axion 3, by taking $A_i = \emptyset$ for all $i > n$.

P3: $P(A) = 1 - P(A^c)$
Proof: As A and A^c are mutually exclusive,
$1 = P(S) = P(A) + P(A^c)$ by **P2** and Axiom 2.

P4: If $A \subset B$, then $P(A) \leq P(B)$
Proof: $B = A \cup (B \cap A^c)$, and A and $B \cap A^c$ are mutually exclusive. So $P(B) = P(A) + P(B \cap A^c) \geq P(A)$ by **P2** and Axiom 1.

Simple propositions

P5: For any A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

P5$*$: For any A, B and C,

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) \\ - P(A \cap B) - P(B \cap C) - P(A \cap C) \\ + P(A \cap B \cap C) \end{aligned}$$

P5$**$: Inclusion-exclusion identity. For any $E_1, E_2, \ldots, E_n$,

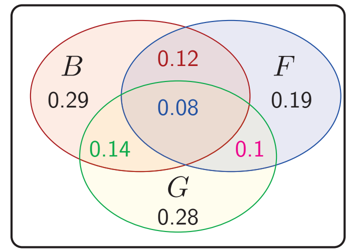
$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) - \sum\sum_{i<j} P(E_i \cap E_j) + \cdots + (-1)^{n+1} P(E_1 \cap \cdots \cap E_n)$$

Examples

A survey of a group's viewing habits of gymnastics (G), baseball (B), and football (F) revealed that:
28% watched G, 29% watched B, 19% watched F,
14% watched G and B,
12% watched B and F,
10% watched G and F,
and 8% watched all three sports.

What percentage of the group watched none of the three sports? The event that none of the three sports is watched $(B \cup G \cup F)^c$.

$$P((B \cup G \cup F)^c) = 1 - P(B \cup G \cup F)$$



Examples

Poker: Without all of the betting, poker boils down to selecting 5 cards from 52-card deck.

How many outcomes are there? (2,598.960)

How many outcomes give you a full house? (This occurs when the cards have denominations a, a, a, b, b, where a and b are distinct, *i.e.* three of a kind and a pair.)

What is P(full house)?

How about P(two pairs)? (This occurs when the cards have denominations a, a, b, b, c, where a, b, c are all distinct.)

Matching Problem: E_i is the event that i-th match occurs, $i = 1, \ldots n$.

$$P(E_i) = \frac{(n-1)!}{n!}, \quad P(E_i \cap E_j) = \frac{(n-2)!}{n!},$$

$$P(E_i \cap E_j \cap E_k) = \frac{(n-3)!}{n!}$$

n probabilities $P(E_i)$, $\binom{n}{2}$ pairs $P(E_i \cap E_j)$ etc.

$$P(\text{atleast one match}) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + \frac{(-1)^{n+1}}{n!} \approx 1 - \frac{1}{e}$$

Conditional Probability

Definition: For any events A and B with $P(B) > 0$,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Left-hand side: "Conditional probability of A given that B has occurred" or simply "probability of A given B."

Conditional Probability

Roll two dice, so the sample space S has 36 elements or outcomes.

$$\begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{array}$$

$A = \{$sum of the results of the two dice in s is even$\}$
$B = \{$first die is even$\}$
$C = \{$second die is $> 5\}$
$D = \{$sum is prime$\}$
$P(A \cap D|B)$? $P(B|C)$? $P(D|C)$?

Conditional Probability

A public health researcher examines the medical records of a group of 937 men who died in 1999 and discovers that 210 of the men died from causes related to heart disease (H). Moreover, 312 of the 937 men had at least one parent who suffered from heart disease (D), and, of these 312 men, 102 died from causes related to heart disease.

$$P(H) = \frac{210}{937} \quad P(D) = \frac{312}{937} \quad P(H|D) = \frac{102}{312}$$

Determine the probability that a man randomly selected from this group died of causes related to heart disease, given that neither of his parents suffered from heart disease.
$P(H|D^c)$?
Venn diagrams can help with conditional probability.

The multiplication rule (of conditioning)

From

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

we get $P(A \cap B) = P(A|B)P(B)$.
Similarly, $P(A \cap B) = P(B|A)P(A)$.
. . . assuming, of course, that $P(B) \neq 0$ and $P(A) \neq 0$.

Example: The classic birthday problem

If there are n people in a room, what is the probability that at least one pair has the same birthday? (Assume nobody related, all 365 days equally likely, ignore Feb 29.)
Easiest attack: find 1 P(no match)
One method for P(no match): Use counting as in chapter 1
Well use the multiplication rule instead. . .

Goal: Find P(no match)

Imagine we draw birthdays one at a time at random. Let A_i be the event 'no match is created at the ith draw'. Then 'no match' is just $A_1 \cap A_2 \cap \cdots \cap A_n$.

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)\ldots \\ &= \frac{365}{365} \times \frac{364}{365} \times \frac{363}{365} \times \ldots \end{aligned}$$

For $n = 23, P(match) = 0.507$; for $n = 50, P(match) = 0.97$; for $n = 57, P(match) = 0.99$; for $n = 60, P(match) = 0.9941$.

Independnet Events

What does it mean for two events A and B to be independent?

Intuition: Knowing that B occurs has no influence on P(A).

In symbols: $P(A|B)$ is the same as P(A).

By the multiplication rule: $P(A \cap B) = P(A)P(B)$.

This is taken as the mathematical definition of independence.

Independnet Events

For three events A, B, C to be independent, it is not enough that

$$P(A \cap B \cap C) = P(A)P(B)P(C). \qquad (1)$$

Example: $S = \{1, 2, 3, 4, 5, 6, 7, 8\}$; sample space with 8 equally likely outcomes.
$A = \{1, 2, 3, 4\}$, $B = \{1, 5, 6, 7\}$, $C = \{1, 5, 6, 8\}$
None of the pairs of events are independent but (1) holds.
So we also need

$$\begin{aligned} P(A \cap B) &= P(A)P(B) \\ P(A \cap C) &= P(A)P(C) \\ P(B \cap C) &= P(B)P(C) \end{aligned}$$

Example: $D = \{1, 2, 5, 6\}$, $E = \{1, 2, 7, 8\}$.
A, D are independent, A, E are independent, E, D are independent, but $P(A \cap D \cap E) = \frac{1}{4} \neq P(A)P(D)P(E)$.
Generalize this idea for more than 3 events.

Independnet Events

In problems involving independence. . .
Very common: Independence is given or may be assumed; you then use the formula(s) in the definition.
Quite rare: You use the definition to check for independence.

Examples:

Roll two dice. What is P(odd number on at least one die)?

Flip seven coins. What is P(exactly three Heads)?

Two warnings:
Independence is not the same as mutual exclusivity (!!)
Venn diagrams are no good for depicting independence.

Law of Total Probability

Suppose that $B_1, B_2, \ldots, B_n$ is a partition of S. We know that for any A:

$$A = \bigcup_{i=1}^{n} (A \cap B_i)$$

Furthermore:

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

The latter is sometimes called the Law of Total Probability.
By the multiplication rule:

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

Bayes' Theorem

To find $P(B_k|A)$ for any of the B_k in a partition,
use the Law of Total Probability in the denominator of $P(B_k|A)$:

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

This is called Bayes Theorem. Note how Bayes Theorem 'switches the order of conditioning.'

Examples

Example (very common):
A blood test indicates the presence of a particular disease 95% of the time when the disease is actually present. The same test indicates the presence of the disease 0.5% of the time when the disease is not present. One percent of the population actually has the disease.

Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

Example:
During a criminal trial, suppose the jury believes that there is a 50% probability that the defendant is guilty. Then, an expert witness testifies that a new piece of evidence has surfaced that proves that the true criminal is left-handed, just like the defendant. If we suppose that 10% of the general population is left-handed, how certain should the jury now be of the defendants guilt in light of this new information?