CH 1.2 Sample Spaces

Definition: (1.2.1) Sample space and event

The sample space S of an experiment is the set of all possible outcomes of the experiment. An event A is a subset of the sample space S, and say that A occurred if the actual outcome is in A.

Ex 1.2.2

A coin is flipped 10 times. Let the sample space of all possible combinations be denoted as

$$A = \{(s_1, s_2, \dots, s_{10}) : s_j \in \{0,1\} \text{ for } 1 \le j \le 10\}$$

where $H=1$, $T=0$. Let A_j be the event that the jth flip is heads.

Let B be the event that at least one flip was heads. Then $B = \bigcup_{j=1}^{10} A_j$.

Let C be the event that all flips were heads. Then $C = \bigcap_{j=1}^{10} A_j$.

Let D be the event that there were at least two consecutive heads. Then $D = \bigcup_{j=1}^{9} (A_j \cap A_{j+1})$.

Remark:

English	Sets
s is a possible outcome	$s \in S$
A is an event	$A \subseteq S$
A occurred	$s_{actual} \in A$
something must happen	$s_{actual} \in S$
A or B, but not both	$(A \cap B^c) \cup (A^c \cap B)$
A implies B	$A \subseteq B$
$A_1, \dots, A_n$ are a partition of S	$A_1 \cup \dots \cup A_n = S$, $A_i \cap A_j = \emptyset$ for $i \neq j$

CH 1.6 Non-naive Def. of Probability

Definition: (1.6.1) General definition of probability

A probability space consists of a sample space S and a probability function P which takes an event $A \subseteq S$ as input and returns $P(A)$, a real number between 0 and 1, as output. The function P must satisfy the following Axioms:

(i) $P(\emptyset) = 0$, $P(S) = 1$

(ii) If $A_1, A_2, \ldots$ are disjoint events, then

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

In the Pebble World analogy, probability is like mass. The mass of an empty pile of pebbles is 0, the total mass of all pebbles is 1, and if we have non-overlapping piles of pebbles, we can get their combined mass by adding the masses of the individual masses. We can also have a countably infinite number of pebbles as long as their mass is 1

We can even have uncountable sample spaces, such as having S be an area in a plane.

Any function P that satisfies the two axioms is considered a valid probability function.

The frequentist view of probability is that it represents a long-run frequency over a large number of repetitions of an experiment.

The Bayesian view of probability is that it represents a degree of belief about the event in question.

Theorem: (1.6.2) Properties of Probability

(i) $P(A^c) = 1 - P(A)$

(ii) If $A \subseteq B$, then $P(A) \leq P(B)$

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

Proof: (i)

Since A and A^c are disjoint and their union is S, the second axiom gives

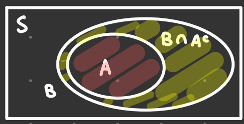
$$P(S) = P(A \cup A^c) = P(A) + P(A^c)$$

$P(S) = 1 \Rightarrow P(A) + P(A^c) = 1$

$\blacksquare$

Proof: (ii)

If $A \subseteq B$, then we can write B as the union of A and $B \cap A^c$, where $B \cap A^c$ is the part of B not also in A.

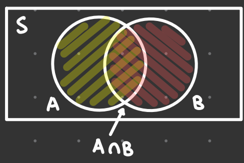


Since A and $B \cap A^c$ are disjoint, we can apply the second axiom:

$$P(B) = P(A \cup (B \cap A^c)) = P(A) + P(B \cap A^c)$$
$$\geq P(A) \qquad \text{since } P(B \cap A^c) \geq 0$$

$\square$

Proof: (iii)



The shaded regions combined represents $A \cup B$, but $P(A \cup B) \neq P(A) + P(B)$.

We can write $A \cup B$ as the union of the disjoint events A and $B \cap A^c$. Thus by the second axiom,

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

So it suffices to show that $P(B \cap A^c) = P(B) - P(A \cap B)$. Since $A \cap B$ and $B \cap A^c$ are disjoint and their union is B, another application of the second axiom gives

$$P(A \cap B) + P(B \cap A^c) = P(B)$$

Theorem: (1.6.3) Inclusion-exclusion

For any events $A_1, \dots, A_n$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P(A_1 \cap \dots \cap A_n)$$