1. Probability and counting

1.2.Sample Spaces and Pebble World

The **sample space** S of an experiment = set of all possible outcomes of the experiment. Can be finite or infinite.

An **event** A is a subset of the S, and we say that A **occurred** if the actual outcome is in A.

not A = A^c

Example:

$A, B \subseteq S$

$A \cup B$ if and only if at least one of A, B occurs.

$A \cap B$ if and only if both of A, B.

De Morgan's laws:

$(A \cup B)^c = A^c \cap B^c$ | A OR B not occur = A not occur AND B not occur
$(A \cap B)^c = A^c \cup B^c$ | A AND B not occur = A not occur OR B not occur

1.3 Naive definition of probability

$$P_{naive}(A) = \frac{|A|}{|S|}$$

Requires S to be finite, with equal mass for each outcome.

1.4 How to count

Multiplication rule:

Consider experiment C = 2 sub-experiments A, B. A has a outcomes, B has b outcomes. Then C has ab outcomes.

Sampling without replacement:

n objects and making k choices from them, one at a time without replacement. Number of outcomes: $n(n-1)\dots(n-k+1)$ for $n \leq k$.

A **permutation** of $1, 2, \dots, n$ is an arrangement of them in some order, e.g., 3,5,1,2,4 is a permutation of 1,2,3,4,5.

There are $n!$ permutations of $1, 2, \dots, n$

Binomial coefficient: $\binom{n}{k}$ = "n choose k" number of subsets of size k for a set of size n. For example, $\binom{4}{2} = 6$

For $n \leq k$:

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n!}{(n-k)!k!}$$

For $n > k$:

$$\binom{n}{k} = 0$$

1.5 Story Proofs

Example: **Vandermonde's identity**

$$\binom{n+m}{k} = \sum_{j=0}^{k} \binom{m}{j} \binom{n}{k-j}$$

Story proof:

Consider a group of m peacocks and n toucans, from which a set of size k birds will be chosen. There are $\binom{n+m}{k}$ possibilities for this set of birds. If there are j peacocks in the set, then there must be $k-j$ toucans in the set. The right-hand side of Vandermonde's identity sums up the cases for j.

1.6 General definition of probability

A **probability space** = *sample space* S and *probability function* P. Event $A \subseteq S$. $P(A) = [0, 1]$.

Axioms for *probability function* P:

1. $P(\emptyset) = 0, P(1) = 1$
2. If $A_1, A_2, \ldots$ are disjoint:

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

("A_i, A_j are disjoint" means that $A_i \cap A_j = \emptyset$ if $i \neq j$)

The **frequentist view** of probability: if we say a coin has probability 1/2 of Heads, that means the coin would land Heads 50% of the time if we tossed it over and over and over (long-run frequency).

The **Bayesian view** of probability: represents a degree of belief about the event in question. We can assign probabilities to hypotheses even if it isn't possible to repeat the experiment over and over again.

Properties of probability:

$\forall A, B$,

1. $P(A^C) = 1 - P(A)$
2. If $A \subseteq B$, then $P(A) \leq P(B)$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ *inclusion-exclusion*

Inclusion-exclusion for n events:

$\forall A_1, .. A_n$,

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

$$+(-1)^{n+1} P(A_i \cap \ldots \cap A_n)$$

Can be simply used if A_i are symmetric else it is very tedious.