Math 4600 - Probability

We start with a definition

Defn. A and B are disjoint events if $A \cap B = \emptyset$ (they have no outcomes in common).

Notation. We use $P(A)$ to denote the probability of event A.

Axioms. We assume that

1. For all A, $0 \le P(A) \le 1$.
2. For the sample space S, $P(S) = 1$.
3. If $A_1, A_2, \ldots$ is a collection of disjoint events, $P(\bigcup_j A_j) = \sum_j P(A_j)$.

Keep in mind that the axioms are <u>assumptions</u>, not theorems. Their job is to give us a foothold to start building a theory - we won't worry about what random <u>means</u> for now.

~~Theorem.~~ If S has n <u>equally likely</u> outcomes the probability of each is $1/n$.

Proof.
$$
\begin{aligned}
1 &= P(S) \quad \swarrow \text{axiom 2} \\
&= P(\{x_1, \ldots, x_n\}) \\
&= P(\{x_1\} \cup \ldots \cup \{x_n\}) \\
&= P(\{x_1\}) + \ldots + P(\{x_n\}) \quad \swarrow \text{axiom 3} \\
&= n\, P(\{x_1\}) \quad \swarrow \text{hypothesis}
\end{aligned}
$$

So $P(\{x_1\}) = 1/n$.

<u>Corollary</u> If S has n equally likely outcomes and A has j outcomes, $P(A) = j/n$.

We'll use this kind of reasoning so often we'll want notation:

<u>Definition</u>. The number of outcomes in A is called the size of A and denoted $|A|$ (or $\#A$).

Example. Suppose that a roll of a 6-sided die has equally likely outcomes.

1) What is $P(\{4\})$?

2) What is $P(\text{an even number})$?

Example. Suppose a 6-sided die has equally likely outcomes, and you roll it twice.

1) What is the new sample space?
2) What is P(the numbers sum to 7)?
3) What is P(the numbers sum to 4)?

Definition. If B_j a collection of nonempty disjoint events and $\bigcup_j B_j = S$, we say B_j is a partition of S.

Example. If $S = \{1, 2, \ldots, 6\}$, then $B_1 = \{1,3,5\}$, $B_2 = \{2,4,6\}$ is a partition.

Note. $P(\bigcup_j B_j) = \sum_j P(B_j)$ ← disjoint (axiom 3) and

$$P(\bigcup_j B_j) = P(S) = 1 \leftarrow \text{axiom 2}$$

so $\sum_j P(B_j) = \underline{1}$.

The particularly useful partition is A, A^c.

Theorem. $P(A^c) = 1 - P(A)$.

Now we prove

Theorem. If $A \subset B$, $P(A) \leq P(B)$.

Proof. Recall $B \setminus A$ means "all outcomes in B but not in A".
We have

$$B = A \cup B \setminus A$$

and this is disjoint, so

$$P(B) = P(A) + P(B \setminus A)$$

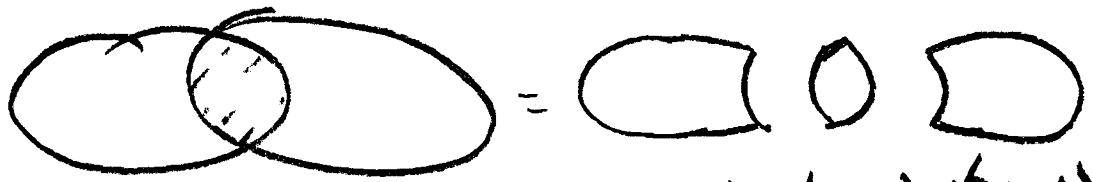
Now $P(B \setminus A) \geq 0$ (axiom 1), so this implies

$$P(B) \geq P(A). \quad \square$$

We now prove something really neat!

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

Proof.

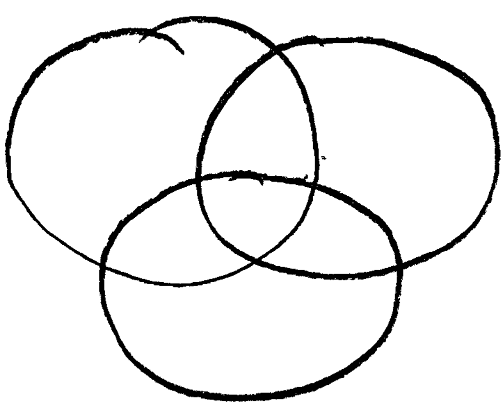


$P(A \cup B) = P(A \setminus B) + P(A \cap B) + P(B \setminus A)$

$= P(A \setminus B) + P(A \cap B) + P(B \setminus A) + P(A \cap B) - P(A \cap B)$

$= P(A) + P(B) - P(A \cap B)$. $\square$

We can make a similar argument with three sets.



Theorem.

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

We should be seeing a pattern...

Theorem (inclusion-exclusion principle)

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{\text{collections } C \\ \text{of } k \text{ events}}} P\left(\bigcap_{j \in C} A_j\right).$$