The axioms of Probability

Given a random experiment with sample space S, we say an event $E \subseteq S$ <u>occurs</u> if some element of E is the outcome of the experiment.

<u>Probability</u> is a way to measure the likelihood of an event.

One way to define probability is in terms of relative frequency. Given an event E, let $n(E)$ be the number of times E occurs when the experiment is run n times. Then we may define the probability of E as

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}.$$

However, we have no reason to believe that this will necessarily converge, and so we take a more axiomatic approach to probability.

<u>First Example:</u> <u>Equally Likely outcomes</u>

Given a sample space with n-outcomes, each of which is equally likely, each outcome is assigned a probability of $1/n$.

Ex: A fair coin: $S = \{H, T\}$, $P(\{H\}) = 1/2$, $P(\{T\}) = 1/2$.

Ex: A fair die: $P(\{1\}) = 1/6$. etc.

Axioms of Probability.

Given a sample space S, we define probability to be a function from the set of events to the interval $[0, 1]$. P satisfies the following axioms:

1) $P(S) = 1$.
2) $0 \leq P(E) \leq 1$ for any $E \subseteq S$.
3) For any sequence $E_1, E_2, \dots$ of mutually exclusive events (i.e. $E_i \cap E_j = \emptyset$ if $i \neq j$) then

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

As a particular consequence of axiom 3), if S is a discrete sample space then for any $E \subseteq S$, $P(E) = \sum_{x \in E} P(\{x\})$.

For example, for a die,

$$P(\{\text{even roll}\}) = P(\{2\}) + P(\{4\}) + P(\{6\})$$
$$= 1/6 + 1/6 + 1/6 = 3/6 = 1/2.$$

Prop: $P(E^c) = 1 - P(E)$.

Pf: $E \cap E^c = \emptyset$ and $E \cup E^c = S$,

so by axioms 2 and 3,

$$1 = P(S) = P(E \cup E^c) = P(E) + P(E^c)$$

so $P(E^c) = 1 - P(E)$. $\square$

In particular, $P(\emptyset) = 1 - P(\emptyset^c)$
$= 1 - P(S) = 1 - 1 = 0$.

In general, for events $E \subseteq F$, we have $P(E) \leq P(F)$. This is because F can be written as

$$F = E \cup (F \cap E^c)$$

and since $E \cap (F \cap E^c) = (E \cap F) \cap E^c = E \cap E^c = \emptyset$, we have that

$$P(F) = P(E) + \underbrace{P(F \cap E^c)}_{\geq 0} \geq P(E).$$

For non-mutually exclusive events we need to subtract the probability of double counted events:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F).$$

For example: Rolling a die:

$$P(\{1,2,3,4\}) = 4/6 = 2/3$$

$$P(\{3,4,5,6\}) = 4/6 = 2/3.$$

$$\begin{aligned} 1 = P(\{1,2,3,4,5,6\}) &= P(\{1,2,3,4\}) + P(\{3,4,5,6\}) \\ &\quad - P(\{3,4\}) \\ &= 2/3 + 2/3 - 2/6 \\ &= 4/3 - 1/3 = 3/3 = 1 \end{aligned}$$

More general:

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

note.

Ex: In a game of 5 card stud, the probability of getting a royal flush, 10, J, Q, K, A is $\frac{4}{\binom{52}{5}} \approx 0.000002\%$

Conditional Probability.

Suppose you have a deck of 52 cards. The probability that you draw an ace from a shuffled deck is $4/52$. Suppose one random card is missing (you don't know which; say your friend draws a random card before you). What is the probability of drawing an ace now? Now things are somewhat more complicated: if the missing card was an ace, then the probability is $3/51$, but if it wasn't an ace, then the probability is $4/51$ (51 since one card is already missing).

That is, the probability depends / is "conditional" on the outcome of the first draw.

Given events A and B, we write

$$P(B|A)$$

to mean "the probability of B given an occurence of A" i.e. $P(B|A)$ is the answer to the question "what is the probability of B if we already know that A has occured?"

To Be Continued...