Axioms of Probability Lecture 2

The Axiomatic Approach

Consider an experiment in which the actual outcome is uncertain but the set of possible outcomes is known.

Def: Sample Space S is the set of all possible outcomes.

Ex. If you roll a die, then

$$S = \{1, 2, 3, 4, 5, 6\}$$

If you toss two coins.

$$S = \{(H,H), (H,T), (T,H), (T,T)\}$$

If you're trying to determine the lifetime of a flashlight.

$$S = \{t : t \geq 0\} = [0, \infty)$$

Remark: Think of S as a collection of all possible universes. A point $s \in S$ contains information about every minute observable detail.

Any subset of a sample space is known as an event.

Ex. Let S be a sample space for the lifetime of a flashlight. Let E be the event that the flashlight works longer than 5 hrs. Then $E = (5, \infty)$.

The union and intersection of events is another event.

Ex. Let $S = \{1,2,3,4,5,6\}$ be the outcomes from rolling a die. If E is the event that the number rolled is even, then $E = \{2,4,6\}$. If F is the event that the number is divisible by 3, then

$F = \{3, 6\}$

$H = E \cap F = \{6\}$, $G = E \cup F = \{2,3,4,6\}$, $\emptyset$, and S are also events in S.

The complement of an event E, denoted by E^c is the set $E^c = \{x : x \in S, x \notin E\} = S - E$.

We have the following properties:

① $$\left(\bigcup_{k=1}^{n} E_k\right)^c = \bigcap_{k=1}^{n} E_k^c$$

② $$\left(\bigcap_{k=1}^{n} E_k\right)^c = \bigcup_{k=1}^{n} E_k^c$$

In words ① means that if an outcome happens that is not in any E_k, it must be not in E_1 and not in E_2, etc. ② is similar.

Axions of Probability

Let S be a sample space and $P: \mathcal{P}(S) \longrightarrow \mathbb{R}$ be a function from the power set of S, $\mathcal{P}(S)$, to $\mathbb{R}$ such that

Axiom I: $0 \leq P(E) \leq 1$ for $E \subset S$

Axiom II: $P(S) = 1$

Axiom III: If $E_1, E_2, E_3 \ldots$ is a sequence of events in S such that $E_i \cap E_j = E_i E_j = \emptyset$ for all $i \neq j$ then

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

<u>Remarks:</u> (a) Axiom III is the chief reason why infinite series play a big role in probability.

(b) If S is uncountably infinite, it is impossible to define a function that satisfies Axioms I–III for all elements of $\mathcal{P}(S) =$
$= \{E : E \subset S\}$.

<u>Ex.</u> What is $P(\emptyset)$? Justify your answer using the axioms of probability.

<u>Solution:</u> Let $E_1 = S$, $E_2 = \emptyset$, $E_3 = \emptyset, \ldots$

Then $S = \bigcup_{k=1}^{\infty} E_k$. By axiom II $P(S) = 1$.

Since $E_i E_j = \emptyset$ for all $i \neq j$

by axiom III we have

$$1 = P(S) = P\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} P(E_k)$$

Thus $1 = 1 + \sum_{k=2}^{\infty} P(\emptyset) \implies 0 = \sum_{k=2}^{\infty} P(\emptyset)$.

By axiom I, $P(\emptyset) \geq 0$. Thus we must have

$P(\emptyset)$.

Alternatively, observe that by axiom III

$$P(\emptyset) = \sum_{k=1}^{\infty} P(\emptyset)$$

By axiom I this series converges. Hence by the divergence theorem for series, setting $a_k = P(\emptyset)$ we must have $\lim_{k \to \infty} a_k = 0$. But a_k is constant so $a_k = 0$ for all k. Hence $P(\emptyset) = 0$.

<u>Ex.</u> If $E_1, E_2, \ldots, E_n$ is a finite sequence of pairwise disjoint events, what is $P\left(\bigcup_{k=1}^{n} E_k\right)$ in terms of $P(E_1), P(E_2), \ldots, P(E_n)$?

Solution: Set $E_{n+1}, E_{n+2}, \ldots = \emptyset$. Then

$$P\left(\bigcup_{k=1}^{n} E_k\right) = P\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} P(E_k) =$$

$$= \sum_{k=1}^{n} P(E_k) + \sum_{k=n+1}^{\infty} P(\emptyset) = \sum_{k=1}^{n} P(E_k)$$

Proposition: $P(E^c) = 1 - P(E)$ for any $E \subset S$.

Proof: $S = E \cup E^c$ where $EE^c = \emptyset$ Hence

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

or $\quad 1 - P(E) = P(E^c)$

Proposition: If $E \subset F$ then $P(E) \leq P(F)$

Proof: $S = E \cup E^c$ so $F = SF = (E \cup E^c)F =$

$= EF \cup E^cF$. Since EF and E^cF are pairwise disjoint, $P(F) = P(EF) + P(E^cF)$

$= P(E) + P(E^cF) \geq P(E)$.

<u>Proposition:</u> $P(E \cup F) = P(E) + P(F) - P(EF)$.

<u>Proof:</u> $E \cup F = E \cup E^c F$ where E and $E^c F$ are disjoint. Hence $P(E \cup F) = P(E \cup E^c F) =$

$= P(E) + P(E^c F)$

Recall that $P(F) = P(EF) + P(E^c F)$

so $P(E^c F) = P(F) - P(EF)$.

In particular,

$$P(E \cup F) = P(E) + \underline{P(E^c F)} = P(E) + \underline{P(F) - P(EF)}.$$

This proposition can be extended to the union of 3 sets.

$$P(E \cup F \cup G) = P([E \cup F] \cup G) = \underline{P(E \cup F)} + P(G)$$

$$- P([E \cup F] G) = \underline{P(E) + P(F) - P(EF)} + P(G)$$

$$- P(EG \cup FG) = P(E) + P(F) + P(G) - P(EF)$$

$$- P(EG) - P(FG) + P(EFG)$$

You should convince yourself that the following general result holds.

<u>Proposition:</u> $P(E_1 \cup E_2 \cup \ldots \cup E_n) =$

$$= \sum_{r=1}^{n} (-1)^{r+1} \sum_{i_1 < i_2 < \ldots < i_r} P(E_{i_1} E_{i_2} \ldots E_{i_r})$$ where

$\sum_{i_1 < i_2 < \ldots < i_r} P(E_{i_1} E_{i_2} \ldots E_{i_r})$ is taken over all of the $\binom{n}{r}$ possible subsets of size r of the set $\{1, 2, \ldots, n\}$

<u>Ex.</u>

$$P(E_1 \cup E_2 \cup E_3 \cup E_4) = \underbrace{P(E_1) + P(E_2) + P(E_3) + P(E_4)}_{\text{All possible single sets}}$$

$$- \underbrace{P(E_1E_2) - P(E_1E_3) - P(E_1E_4) - P(E_2E_3) - P(E_2E_4) - P(E_3E_4)}_{\text{All possible pairs.}}$$

$$+ \underbrace{P(E_1E_2E_3) + P(E_1E_2E_4) + P(E_1E_3E_4) + P(E_2E_3E_4)}_{\text{All possible triples.}}$$

$$- \underbrace{P(E_1E_2E_3E_4)}_{\text{All possible quadruples.}}$$

<u>Ex.</u> I have two cats. Cat A irritates me with probability 0.95 and cat B irritates me with probability 0.01. What is the probability that I will be irritated by neither cat? Assume that the probability that both cats bother me is 0.005.

<u>Solution:</u>

The probability that I am bothered by at least one of the cats is

$$P(A \cup B) = P(A) + P(B) - P(AB) =$$

$$= 0.95 + 0.01 - 0.005 = 0.955$$

Thus the probability that I am bothered by neither cat is

$$P([A \cup B]^c) = 1 - P(A \cup B) = 1 - 0.955 = 0.045$$