Probability and Statistics notes

Aditya Mehrotra

April 2021

Contents

1 Probability

1.1 Basic mathematical definitions

Compliment: Given a set of of elements A where $A \subseteq S$ for some set of elements S, we define the compliment of A in S as the set of elements of S that are not in A. Mathematically, this is expressed as A^c or $S - A$.

Difference: The difference between sets A and B is all elements that are in A but not in B. Mathematically, this is expressed as $A - B$

Disjoint: A and B are disjoint if they have no common elements. Mathematically, this is when $A \cap B = \emptyset$

Product of sets: The cartesian product of two sets S and T, denoted by $S \times T$:

$$S \times T = \{(s,t) \mid s \in S, t \in T\}$$

Cardinality of a set: The number of elements in a set S, mathematically denoted by $|S|$

Inclusion exclusion principle:

$$|A \cup B| = |A| + |B| - |A \cap B|$$

1.2 Basic probability terminology

Experiment: A repeatable procedure with well-defined possible outcomes

Sample space: The set of all possible outcomes. This is denoted by Ω and sometimes by S.

Event: A subset of the sample space

Probability function: A function assigning a probability for each outcome. More specifically, for a discrete sample space S, a probability function P assigns each outcome $\omega \in S$ a number $P(\omega)$. This number is called the probability of ω.

P must satisfy two rules:

Rule 1: $0 \leq P(\omega) \leq 1$ (probabilities are between 0 and 1)

Rule 2: Given $S = \{\omega_1, \omega_2, \omega_3, ..., \omega_n\}$, $\sum_{j=1}^{n} P(\omega_j) = 1$.

(The sum of probabilities of all possible outcomes add to 1)

Additionally, the probability of an event $E \subseteq S$ is the sum of the probabilities of all outcomes in E. More specifically,

$$P(E) = \sum_{\omega \in E} P(\omega)$$

Discrete sample space: A listable sample space, can be finite or infinite.

1.3 Basic probability rules

For some events A, L and R contained in a sample space Ω:

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

If L and R are disjoint, then $P(L \cup R) = P(L) + P(R)$

If L and R are not disjoint, we have the inclusion-exclusion principle:

$$P(L \cup R) = P(L) + P(R) - P(L \cap R)$$

1.4 Conditional Probability

Conditional probability: The probability of A given B. Mathematically, this is expressed as $P(A|B)$. More rigorously, given events A and B, the conditional probability of A given B is,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \textit{ provided that } P(B) \neq 0$$

This gives us the multiplication rule,

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Law of Total Probability: Suppose the sample space Ω is divided into n disjoint events, $B_1, B_2, ..., B_n$. Then for any event A:

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_n)$$

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_n)P(B_n)$$

1.5 Independence