Higher Theory of Statistics
Math2901 UNSW

Jeremy Le
(Based of Hussain Nawaz's Notes)

2023T2

Contents

1 Probability

1.1 Experiment, Sample Space, Event

Experiment An experiment is any process leading to recorded observations.

Outcome An outcome is a possible result of an experiment.

Sample Space The set Ω of all possible outcomes is the sample space of an experiment. Ω is discrete if it contains a countable (finite or countably infinite) number of outcomes.

Events An event is a set of outcomes, i.e. a subset of Ω. An event occurs if the result of the experiment is one of the outcomes in that event.

Mutual Exclusion Events are mutually exclusive (disjoint) if they have no outcomes in common.

Set Operations If you have trouble remembering the above rules, then one can essentially replace $\cup$ by multiplication and $\cap$ by addition.
(The associative law) If A, B, C are sets then

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

(Distributive Law) If A, B, C are sets then

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

1.2 Sigma-algebra

The $\sigma-$algebra must be defined for rigorously working with probability. The $\sigma-$algebra can be thought of as the family of all possible events in a sample space. Analogously, this may be conceptualised as the power set of the sample space.

Probability A probability is a set function, which is usually denoted by $\mathbb{P}$, that maps events from the $\sigma-$algebra to $[0, 1]$ and satisfies certain properties.

Probability Space The triplet $(\Omega, \mathcal{A}, \mathbb{P})$ is the probability/sample space where

- Ω is the sample space,
- $\mathcal{A}$ is the σ-algebra,
- $\mathbb{P}$ is the probability function.

Properties of Probability Given the probability/sample space $(\Omega, \mathcal{A}, \mathbb{P})$, the probability function $\mathbb{P}$ must satisfy

- For every set $A \in \mathcal{A}$, $\mathbb{P}(A) \geq 0$
- $\mathbb{P}(\Omega) = 1$
- (Countably additive) Suppose the family of sets $(A_i)_{i \in \mathbb{N}}$ are mutually exclusive, then

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

Probability Lemmas

- Given a family of disjoint sets $(A_i)_{i=1,\ldots,k}$

$$\mathbb{P}\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \mathbb{P}(A_i)$$

- $\mathbb{P}(\phi) = 0$
- For any $A \in \mathcal{A}, \mathbb{P}(A) \leq 1$ and $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- Suppose $B, A \in \mathcal{A}$ and $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

Continuity from Below Given an increasing sequence of events $A_1 \subset A_2 \subset \ldots$ then,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

Continuity from Above Given a decreasing sequence of events $A_1 \supset A_2 \supset \ldots$ then,

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mathbb{P}(A_n)$$

1.3 Conditional Probability and Independence

Conditional Probability The conditional probability that an event A occurs given that an event B has occurred is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cup B)}{\mathbb{P}(B)}, \quad \mathbb{P}(B) > 0$$

Independence Events A and B are independent if $\mathbb{P}(A \cup B) = \mathbb{P}(A)\mathbb{P}(B)$.

Lemma - Given two events A and B then $\mathbb{P}(A|B) = \mathbb{P}(A)$ if and only if $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Independence of Sequences

- A countable sequence of event $(A_i)_{i=\mathbb{N}}$ is pairwise independent if $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for all $i \neq j$.
- A countable sequence of events $(A_i)_{i=\mathbb{N}}$ are independent if for any sub-collection $A_{i_1}, \ldots, A_{i_n}$ we have
$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_n}) = \prod_{j=1}^{n} \mathbb{P}(A_{i_j})$$

Independence implies pairwise independence, but pairwise independence does not imply independence.

Multiplicative Law Given events A and B then

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B),$$

and similarly, if you have events A, B, C then

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(A_3|A_2 \cap A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$$

Additive Law Let A and B be events then

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Law of Total Probability Suppose $(A_i)_{i=1,\ldots,k}$ are mutually exclusive and exhaustive of Ω, that is $\bigcup_{i=1}^{k} A_i = \Omega$, then for any event B, we have

$$\mathbb{P}(B) = \sum_{i=1}^{k} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Bayes Formula Given sets B, A and a family of disjoint and exhaustive sets $(A_i)_{i=1,\ldots,k}$ then
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^{k} \mathbb{P}(B|A_i)\mathbb{P}(A_i)}$$

1.4 Descriptive Statistics and R

Categorical Data can be sorted into a finite set of (unordered) categories. e.g. Gender

Quantitative Responses are measured on some sort of scale. e.g. Weight.

Numerical Summaries of the Quantitative Data Given observations $x = (x_1, \ldots, x_n)$. The sample mean (estimated mean) or average is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Sample variance (estimated variance)

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$