Statistics Inference I

Probability Theory, Lecture 1

Haoming Wang

23 June 2020

This is the Lecture note for the *Mathematical Statistics*. The reference materials is *Statistics Inference second edition, George Casella, Roger L. Berger*. The course covers first 5 chapters of the book: Probability Theory, Transformations and Expectations, Common Families of Distributions, Multiple Random Variables, Properties of a Random Sample.

Content:

Set

Definition 1 (Countable Set)**.** Given a set A, if $\exists$ a bijection $A \xrightarrow{f} \mathbb{N}$, then we say A is a countable set.

Note 1. That is for $\forall n \in \mathbb{N}, \exists! a \in A$, s.t. $f(a) = n$. If A is a countable set, then the cardinality of A is equal to the cardinality of $\mathbb{N}$.

Proposition 1. *Given countable sets $A_1, A_2, \cdots, A_n$, then $\cup_{i=1}^n A_i$ is a countable set.*

$$A_1 = \{\overset{1}{a_{11}}, \overset{6}{a_{12}}, \overset{11}{a_{13}}, \cdots\}$$
$$A_2 = \{\overset{2}{a_{21}}, \overset{7}{a_{22}}, \overset{11}{a_{23}}, \cdots\}$$
$$A_3 = \{\overset{3}{a_{31}}, \overset{8}{a_{32}}, \overset{12}{a_{33}}, \cdots\}$$
$$A_4 = \{\overset{4}{a_{41}}, \overset{9}{a_{42}}, \overset{13}{a_{43}}, \cdots\}$$
$$A_5 = \{\overset{5}{a_{51}}, \overset{10}{a_{52}}, \overset{14}{a_{53}}, \cdots\}$$

Proof. Just give an intuitive proof in the margin figure, note to skip all elements that have been numbered. ☐

Proposition 2. *Rational number set $\mathbb{Q}$ is a countable set. Real number set $\mathbb{R}$ is not a countable set.*

Proof. The proofs have been given in *Introduction to Topology/Naive Set Theory 2.pdf*. ☐

Proposition 3. *The intersection of finite open sets is still an open set. The union of any open sets is still an open set.*

Proof. Assume that $A_i (i = 1, \cdots, n)$ is open, If $\cap_{i=1}^n A_i = \emptyset$, then it is open. If $\cap_{i=1}^n A_i \neq \emptyset$, then for any $a \in \cap_{i=1}^n A_i$, $a \in A_i$ for $i = 1, \cdots, n$, and $\exists r_i > 0$, s.t. $B_{r_i}(a) \subseteq A_i$. Let $r = \min\{r_1, \cdots, r_n\}$, then $B_r(a) \subseteq A_i (i = 1, \cdots, n)$, and $B_r(a) \subseteq \cap_{i=1}^n A_i$. Thus $\cap_{i=1}^n A_i$ is an open set. The case for union is trivial. ☐

Example 1. The intersection of countable open sets could not be an open set, for example, let $A_n = (0 - 1/n, 1 + 1/n)$, then $\cap_{n=1}^\infty A_i = [0, 1]$. It can be prove that $\mathbb{R} \backslash \cap_{n=1}^\infty A_i$ is open.

Sample space, Event and Sample

Define a random trial, the set of all outcomes is called sample space, denote as S (or Ω). The element of the sample space S, that is the outcome, is called sample point, any measurable subset of sample space S is called event.

Example 2. Given a sample space S:
1. $\emptyset$ is an event, called empty-event;
2. S is also an event, called total-event.

Given a random trial with sample space S and an event $A \subseteq S$. If the outcome of the random trial belongs to A, we say event A occurs, otherwise we say A does not occur.

Definition 2 (Disjoin)**.** Given two events A, B of the sample space S, if $A \cap B = \emptyset$, we say A and B are disjoin. For (countable or finite) events $A_i (i \in I)$, if $A_i \cap A_j = \emptyset$ for any $i, j \in I$ and $i \neq j$, we say $A_i (i \in I)$ are pairwise disjoin.

Definition 3 (Partition)**.** Given (countable or finite) event of sample space S, $A_i (i \in I)$, if:
1. $A_i (i \in I)$ are pairwise disjoin;
2. $\cup_{i \in I} A_i = S$,

we say $A_i (i \in I)$ is a partition of S.

Example 3. Define $A_i = [i, i+1), i \in \mathbb{Z}$, then $A_i (i \in \mathbb{Z})$ is a partition of $\mathbb{R}$.

Probability

Intuitively, probability is a measure that maps the possibility of the occurrence of an uncertainty event, that is the outcome of the random trial belongs to the event, to $[0, 1]$.

Classical probability

Each outcome of the random trial has the **same possibility** to occur. Then the probability of an event A is $\mathbb{P}(A) = \frac{n(A)}{n(S)}$. ($n(\cdot)$ represents the amount of sample points in a set)

Example 4. Roll two fair dice at once, the probability of the event where the sum of the points is 7 is $\frac{n(A)}{n(S)} = \frac{6}{36} = \frac{1}{6}$. Here $A = \{(1,6), (2,5), (3,4), (6,1), (5,2), (4,3)\}$.

Experience probability

Repeat a random trial n times. Denote the number of occurrences of event A as $n(A)$, then the probability of A is $\mathbb{P}(A) = \lim_{n \to \infty} \frac{n(A)}{A}$.

Axiomatic definition

Let S be a sample space, the set $\mathcal{P}(A) = \{A | A \subseteq S\}$ is called the power set of S.

Note 2. If S is a finite set, $n(S) = n$, then the number of elements in $\mathcal{P}(S)$ is 2^n, thus $\mathcal{P}(S)$ sometimes also be denoted as 2^S.

Definition 4 (σ algebra)**.** Let S be a set, $\mathcal{F}$ is constituted by some subsets of S. If

1. $\emptyset \in \mathcal{F}$;
2. if $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;
3. if countable subsets of S, $A_i (i \in I) \in \mathcal{F} \Rightarrow \cup_{i \in I} A_i \in \mathcal{F}$,

then we say $\mathcal{F}$ is a σ algebra on S.

Example 5. Let S is an nonempty set, then $\mathcal{F}_1 = \{\emptyset, S\}$ is the smallest σ algebra on S; the power set of S, 2^S, is the largest σ algebra on S.

Example 6. If $S = \{a\}$, then S has only one σ algebra $\mathcal{F} = \{\emptyset, \{a\}\}$; If $S = \{a, b\}$, then S has two σ algebras: $\mathcal{F}_1 = \{\emptyset, \{a, b\}\}$, $\mathcal{F}_2 = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$; If $S = \{a, b, c\}$, its all σ algebras are

$$\begin{aligned}
\mathcal{F}_1 &= \{\emptyset, \{a, b, c\}\} \\
\mathcal{F}_2 &= \{\emptyset, \{a, b, c\}, \{a\}, \{b, c\}\} \\
\mathcal{F}_3 &= \{\emptyset, \{a, b, c\}, \{b\}, \{a, c\}\} \\
\mathcal{F}_4 &= \{\emptyset, \{a, b, c\}, \{c\}, \{a, b\}\} \\
\mathcal{F}_5 &= \{\emptyset, \{a, b, c\}, \{a\}, \{b\}, \{a, b\}, \{c\}, \{a, c\}, \{b, c\}\}.
\end{aligned}$$

Note 3. In $\mathcal{F}_5$, $\{a\} \cup \{b\} = \{a, b\}$, and then $\{a, b\}^c = \{c\}$, and then union out $\{a, c\}, \{b, c\}$. It is the power set of S.

Definition 5 (Generate)**.** Let S be a set, $A \subseteq S$. The smallest σ algebra that contains A is called the σ algebra generated by A, denote as $\sigma(A)$.

Example 7. Let $S = \{a, b, c\}$, then

$$\begin{aligned}
&\sigma(\{a\}) = \{\emptyset, \{a, b, c\}, \{a\}, \{b, c\}\}, \\
&f\sigma(\{a\}, \{b\}) = 2^S
\end{aligned}$$

Definition 6 (Axiomatic definition for probability)**.** Assume that S is the sample space of a random trial, $\mathcal{F}$ is a σ algebra on S, map $\mathbb{P} : \mathcal{F} \to [0, 1]$ satisfies:

1. $\mathbb{P}(A) \geq 0$, for $\forall A \in \mathcal{F}$;
2. $\mathbb{P}(S) = 1$;
3. If (countable or finite) events $A_i (i \in I)$ are pairwise disjoin, then $\mathbb{P}(\cup_{i \in I} A_i) = \sum_{i \in I} \mathbb{P}(A_i)$.

We say $\mathbb{P}$ is a probability on $\mathcal{F}$, and $(S, \mathcal{F}, \mathbb{P})$ is called probability space.