Questions

1. What's the difference between **probability** and **statistics**?

2. What's the difference between them and **math**?

Prof. Chaing: Mathematics is in a deterministic environment. With math, we can describe a dynamic system in a deterministic way. Probability involves in a probabilistic system, where the events follow certain distribution and act in like flipping a coin. As to statistics, what we face is not beautiful mathematical objects. Instead, we have to deal with data. We assume there is a distribution or general function behind the data. This function might be deterministic or probabilistic. What a statistician aim to do is to find out (formally speaking, estimate) such underlying distribution.

1 Probabilities, Random variables, Distribution

1.1 Set Theory

In the world of probability and statistics, we use **set** to intuitively model the object we concerned. As a result, we need to define some basic elements and operations. Furthermore, we will derive some basic properties.

Sample space   Here we use Ω to represent sample space, which is composed of all the possible **outcomes**. The element in sample space is often denote as ω.

Event   What we really care about is the subset of sample space, which is intuitively being an event. With this notion, we can interact with various outcomes and play with events that share the common outcomes etc.

Set operations   Here, we use $A, B, \{A_n\}$ to denote events and a sequence of events.

- Union: $A \cup B := \{w : w \in A \ or \ w \in B\}$

- Intersection: $A \cap B := \{w : w \in A \ and \ w \in B\}$ Here, we can drop out the geometric notion of intersection and think of it as **the outcomes that the two events share**.

- Complementation: $A^C := \{w : w \in \Omega \ and \ w \notin A\}$ The outcomes that haven't happened?

- Set difference: $B \backslash A := B \cap A^C$

limsup Intuitively, it's analogous to the smallest upper bound for an infinite sequence.

$$\limsup A_n := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

If we let $B_n := \bigcup_{k=n}^{\infty} A_k$, we can see that $B_1 \supseteq B_2 \supseteq \dots$. Formally, we also have

$$\limsup A_n = \{w : w \in A_n, for\ infinitely\ many\ n\}$$

Intuitively, B_k is the outcomes that share by all the events with index greater than k. Or, for any element ω and barrier m, $\exists n \geq m$ such that $\omega \in A_n$.

liminf It's the analogy to the largest lower bound for an infinite sequence.

$$\liminf A_n := \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

If we let $C_n := \bigcap_{k=n}^{\infty} A_k$, we can see that $C_1 \subseteq C_2 \subseteq \dots$. Formally, we also have

$$\liminf A_n = \{w : w \in A_n\ for\ all\ but\ finite\ many\ n\}$$

Or, for any element ω in $\limsup A_n$, $\exists m$ such that $\forall n \geq m$, $\omega \in A_n$. That is, after some barrier, ω will appear in all A_n afterwards.

Limit of sequence of events We say a sequence of events $\{A_n\}$ converges to event A if

$$\limsup A_n = \liminf A_n = A$$

Property 1 *Let A, B, C be events in sample space Ω, then the following results hold:*

1. *Community: $A \cap B = B \cap A$, $A \cup B = B \cup A$*

2. *Associativity: $A \cap (B \cap C) = (A \cap B) \cap C$, $A \cup (B \cup C) = (A \cup B) \cup C$*

3. *Distributed law: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$*

4. *De Morgan's law: $(A \cap B)^C = A^C \cup B^C$, $(A \cup B)^C = A^C \cap B^C$*

2 Small talk

Prof. Chiang: Can you give an example about statistics?
Me: SVM
Prof. Chiang: What is SVM?
Me: blablabla...
Prof. Chiang: For me, SVM and others techniques are just a chance mechanism. They provides a model or platform for us to analysis the data in some way. But deep in to the heart of data analysis, what we deal with is a general function $G(y, x_1, ..., x_k)$. And all techniques are to model the G with certain assumption and structures.