Theory of statistics and probability

MinSeok Song

1 Statistics

The dichotomy of Bayesian and Frequentist approach is mostly a matter of division based on foundational philosophical differences, and in some situation is not necessarily natural; each instance can be formally simply seen as a choice of analysis. The dichotomy has mainly developed based on two philosophical school of thoughts: subjective belief (via prior distribution) vs. long-run frequency (via sampling distribution). Frequentist

- They avoid making probabilistic claims about model parameters and hypothesss. Instead, they describe the behavior of statistics and procedures over many hypothetical repeated samples.
- They use Statistics derived from data, and use deterministic approach for inference, with methods like hypothesis testing and confidence intervals to draw conclusions about population parameters based on sample data.
- examples: T-test, linear regression, etc

Bayesian

- They assign prior beliefs (distribution function) on parameters of model.
- They use probabilistic argument on specific hypothesis or parameter values.
- examples: MCMC, Bayesian hierarchical modeling, etc

2 Probability Theory

Definition 1. ***Distribution Function (F)****: A function $F : \mathbb{R} \to [0, 1]$ that:*

1. *Is non-decreasing.*
2. *Is right-continuous.*
3. *Limits to 0 as $x \to -\infty$ and 1 as $x \to \infty$.*

Random Variable (X)*: A measurable function $X : \Omega \to \mathbb{R}$ where:*

1. Ω *is a sample space equipped with a probability measure.*

2. $\mathbb{R}$ *is the real line equipped with the Lebesgue measure.*

Probability Measure (P)*: A measure defined on a sample space* Ω *whose total measure is 1.*
Distribution (D)*: Given a random variable* X*, its distribution is the measure induced by* X *on* $\mathbb{R}$*.*
Density Function (f)*: For a random variable* X*, its density function* f *is any measurable function that satisfies:*

$$\Pr(X \in A) = \int_A f \, d\mu$$

for every measurable set $A \subset \mathbb{R}$*, where* μ *is the Lebesgue measure.*

There are several theorems/properties that look trivial but are not really trivial.

- The Skorokhod Representation Theorem states that for every CDF, there exists a canonical probability space and a random variable on that space with the given CDF.

- Kolmogorov's Existence threorem states that given finite dimensional sets of distribution, under some conditions, there exists a canonical probability space with random variables whose corresponding distributions coincide with our original distributions.

- The existence of distribution function does not guarantee the existence of density function. By Radon-Nikodym, there is a nice characterization when this happens: when distribution (which is a measure defined on $\mathbb{R}$) corresponding to distribution function is absolutely continuous with respect to Lebesgue measure.

- The existence of distribution function guarantees the existence of distribution, and vice versa.

- Thus, Kolmogorov's existence theorem is more general than Skorokhod Representation theorem.

Throughout the theory of probability, we usually assume that the probability measure is complete, because of the following proposition.

Proposition 1. *Let us assume that probability measure is complete. Let* $X_t = Y_t$ *almost everywhere for every* $t \geqslant 0$*. Then there exists* $\tilde{Y}_t = Y_t$ *almost everywhere for every* t *such that* $X_t(w) = \tilde{Y}_t(w)$ *for every* $w \in P$ *for every* t *for measurable set* P *whose measure is 1.*