Introduction to Mathematical Statistics

X. Joan Hu

Department of Statistics and Actuarial Science
Simon Fraser University

Fall 2020

What to do today (Sept 14, 2020)?

1. Introduction

2. Probability and Distribution (Chp1-3)
2.1 Probability (Chp1.1-4)

2.1.1 Introduction

2.1.2 Set Theory

2.1.3 Preliminaries on Probability

2.1.4 Conditional Probability and Independence

2.1.1 Introduction: Why to study probability?

In order to understand a stochastic (random) system, ...

- We call a phenomenon *random* if individual outcomes are uncertain but there is nonetheless a regular distribution of outcomes in a large number of repetitions.
- *"limited information* $\Rightarrow$ *uncertainty"*: $\Rightarrow$ probability

2.1.1 Introduction: What is "probability"?

Probability is a measure of the expectation

that an event will occur or a statement is true. (*Wikipedia, the free encyclopedia*)

- **Frequency definition.** In hypothetical identical trials of an experiment, the **probability** of event A is the long term relative frequency of A.
- **Classical definition.** (Symmetry definition) The extent to which an event is likely to occur is measured by the ratio of favourable cases to the whole number of cases possible:

$$P(A) = \frac{\text{number of outcomes leading to A}}{\text{number of all outcomes in the experiment}} = \frac{\#A}{\#S}$$

- **Axiom Definition.** (Kolmogorov, 1933) A probability measure P satisfies three axioms (*basic probability rules*): for a sample space S and a collection of its subsets (i.e. events) $\mathcal{F}$,

  (1). for any event A, $P(A) \geq 0$;
  (2). $P(S) = 1$ for the sample space S;
  (3). if $A_1, A_2, \ldots$ are disjoint, $P(\bigcup A_i) = \sum P(A_i)$.

2.1.1 Introduction: What is "probability"?

Further discussion.

- the Classical Defn: very ideal, not practical
- the Frequency Defn: good interpretation, hypothetical setting, frequentist
- Kolmogorov's Axiom Defn: mathematical, a measure with $P(S) = 1$
- the Subjective (Personal) Defn – *A personal probability of an event is a number between 0 and 1 that expresses an individual's judgment of how likely the outcome is*: a measure for degree of belief

2.1.2 Set Theory: Basic Concepts

- **set, element/point, countable**

$$C_1 = \{x : x = 1, 2, 3\} = \{1, 2, 3\}; C_2 = \{(x, y) : x^2 + y^2 \leq 1\}$$

- **subset, null set/empty set, space**

$$C_{1b} = \{1\} \subset C_1; \quad C_{2b} = \{(x, y) : x^2 + y^2 \leq 2\} \supset C_2$$

More examples ... ...

$$D = \emptyset$$

$$\mathcal{D}_1 = \{1, 2, \ldots\} \text{ all the natural integers;}$$

$$\mathcal{D}_2 = \{(x, y) : -\infty < x, y < \infty\} = \mathcal{R}^2$$

2.1.2 Set Theory: Operation

"A union B". "A or B"; $A \bigcup B$

- Commutativity: $A \bigcup B = B \bigcup A$
- Associativity: $A \bigcup B \bigcup C = A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C$
- Countable union: $C_1 \bigcup C_2 \bigcup \ldots = \bigcup_{j=1}^{\infty} C_j$

Using *Venn Diagram* to present ... ...

2.1.2 Set Theory: Operation

"A intersect B". "A and B"; $A \bigcap B \equiv AB$

- Commutativity: $A \bigcap B = B \bigcap A$
- Associativity: $A \bigcap B \bigcap C = A \bigcap (B \bigcap C) = (A \bigcup B) \bigcup C$
- Countable intersection: $C_1 \bigcap C_2 \bigcap \ldots = \bigcap_{j=1}^{\infty} C_j$
- Distributive laws:

$$C_1 \bigcap (C_2 \bigcup C_3) = (C_1 C_2) \bigcup (C_1 C_3);$$

$$C_1 \bigcup (C_2 \bigcap C_3) = (C_1 \bigcup C_2) \bigcap (C_1 \bigcup C_3)$$

2.1.2 Set Theory: Operation

"A complement". "not A"; $\bar{A} \equiv A' \equiv A^c$

$$\bar{\mathcal{S}} = \emptyset; \quad \bar{\emptyset} = \mathcal{S}; \quad \bar{\bar{C}} = C$$

- Disjoint: $A \bigcap B = \emptyset$ (A and B are *mutually exlcusive*.)
- DeMorgan's Laws:

$$\overline{A \bigcup B} = \bar{A} \bigcap \bar{B}; \quad \overline{A \bigcap B} = \bar{A} \bigcup \bar{B}$$

2.1.2 Set Theory: Set Function

set function.

a mapping Q : a collection of sets $\longrightarrow \mathcal{R} = (-\infty, \infty)$

- e.g. C is a set of integers, and $Q(C)$ is the number of points in C:
  - $C = \{1, 2, 3\}$; $C = \{-1, 0\}$; $C = \emptyset$; $C = \{1, 2, \ldots\}$
- e.g. C is a set in $\mathcal{R}^2$ (the 2-dim real space), and $Q(C)$ is the area of C (ie, $\int_C 1dxdy$):
  - $C = \{(x, y) : x^2 + y^2 \leq 1\}$;
  - $C = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$

2.1.2 Set Theory: Examples

Example 2.1 $C_1 = \{0, 1, 2, 3\}$, $C_2 = \{2, 3, 4\}$, $C_3 = \{-1, 0\}$

- $C_1 \bigcup C_2$, $C_1 \bigcup C_2 \bigcup C_3$
- $C_1 \bigcap C_2$, $C_1 \bigcap C_2 \bigcap C_3$
- $C_1 \bigcup (C_2 \bigcap C_3) = (C_1 \bigcup C_2) \bigcap (C_1 \bigcup C_3)$
- If $Q(C)$ is the number of points in C, $Q(C_1) =?$ and $Q(C_2) =?$

2.1.2 Set Theory: Examples

Example 2.2 a sequence of sets
$C_k = \{(x, y) : 1/k \leq x^2 + y^2 \leq 3 - 1/k\}$ for $k = 1, 2, \ldots$

- $C_1 = \{(x, y) : 1 \leq x^2 + y^2 \leq 2\}$,
  $C_2 = \{(x, y) : 1/2 \leq x^2 + y^2 \leq 5/2\}$,
  $C_3 = \{(x, y) : 1/3 \leq x^2 + y^2 \leq 8/3\}$
- $C_1 \bigcup C_2$, $C_1 \bigcap C_2$
- $C_1 \bigcap (C_2 \bigcup C_3)$
- If $Q(C)$ is the area of C, $Q(C_1) =?$ and $Q(C_2) =?$
- How about $\bigcup_{k=}^{\infty} C_k$ and $\bigcap_{k=}^{\infty} C_k$?

2.1.2 Set Theory: Remarks

- σ**-field/**σ**-algebra.** A σ-field on a set Ω is a collection $\mathcal{F}$ of subsets of Ω that includes "empty set", and is closed under "complement", "countable unions" and "countable intersections".

e.g. $\Omega = \{1, 2, 3\}$: $\mathcal{F}_1 = \{\emptyset, \Omega, \{1\}, \{2, 3\}\}$
$\mathcal{F}_2 = \{\emptyset, \Omega, \{1\}, \{2, 3\}, \{2\}, \{1, 3\}\}$

- **measure** μ**.** Let $\mathcal{F}$ be a σ-field on Ω. A set function μ: $\mathcal{F} \longrightarrow \mathcal{R}$ is called a measure if it satisfies the following.
  - Non-negativity: $\mu(E) \geq 0$ for all $E \in \mathcal{F}$.
  - Null empty set: $\mu(\emptyset) = 0$.
  - Countable additivity: $\mu(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$ if $E_j \in \mathcal{F}$ and are disjoint.

e.g. $\Omega = \{1, 2, 3\}$, $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2, 3\}, \{2\}, \{1, 3\}\}$, $Q(C)$: the number of elements in C

2.1.3 Preliminaries on Probability: Definition

Consider an experiment.

- Its *sample space* S is the set of all its possible outcome.

- Its associated *events* are the subsets of S from a σ-field $\mathcal{F}$ definied on S.

- **Axiom Definition of Probability:** (Kolmogorov, 1933) A probability measure P satisfies three axioms (*basic probability rules*): for a sample space S and a collection of its subsets (i.e. events) $\mathcal{F}$,

  (1). For any event $A \in \mathcal{F}$, $P(A) \geq 0$.
  (2). $P(S) = 1$ for the sample space S.
  (3). If $A_1, A_2, \ldots$ are disjoint (mutually exclusive),
  $P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$.

A probability measure P is a measure defined on $\mathcal{F}$ with $P(S) = 1$.

2.1.3 Preliminaries on Probability: Definition

Probability Space.

A probability space consists of three components $(\Omega, \mathcal{F}, P)$:

- a sample space Ω, the set of all possible outcomes.
- a σ-field $\mathcal{F}$, a set of events containing zero or more outcomes.
- a set function P (probability measure), the assignment of probabilities to the events

e.g. A fair coin is flipped three times. What is the probability of getting at least two heads?

2.1.3 Preliminaries on Probability: Basic Properties

Useful derivations from the Kolmogorov defn (*induced probability rules*):

- $P(\bar{A}) = 1 - P(A) \Longrightarrow P(\emptyset) = 0$
- If $A \subseteq B$, $P(A) \leq P(B)$
- Inclusion Exclusion Formula.
  - $P(A \bigcup B) = P(A) + P(B) - P(AB)$
  - $P(A \bigcup B \bigcup C) =$ $\left[P(A)+P(B)+P(C)\right]-\left[P(AB)+P(BC)+P(AC)\right]+P(ABC)$
  - How about $P(\bigcup_{k=1}^{K} A_k) =?$

2.1.3 Preliminaries on Probability: Example

Example 2.3 A tour guide collects the passports from a group of K tourists for registration purpose. The passports are randomly given back to the tourists later. What is the probability that one or more tourists get their own passports?
Let A_k be the kth tourist having his/her passport. $P(A_k) = 1/K$.

- when $K = 1$?
- when $K = 3$? $[2/3 = .67]$
- when K is a given integer?
  $[1 - 1/2! + 1/3! - 1/4! + ... \approx 1 - e^{-1} = .632]$ Note that

$$P\left(\bigcup_{i=1}^{K} A_i\right) = \sum_{i=1}^{K} P(A_i) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k)$$

$$- \sum_{i<j<k<l} P(A_iA_jA_kA_l) + \ldots$$

2.1.3 Preliminaries on Probability: Counting Rules

For $0 \leq r \leq n$,

- **permutation.** The number of *permutations* of r objects chosen from n distinct objects

$$n^{(r)} = n(n-1)(n-2)\dots(n-r+1) = \frac{n!}{(n-r)!}$$

  Special Case: $r = n$

- **combination.** The number of *combinations* of r objects chosen from n distinct objects

$$\binom{n}{r} = \frac{n^{(r)}}{r!} = \frac{n!}{r!(n-r)!}$$

- $\binom{n}{r} = \binom{n}{n-r}$.

- **binomial coefficients.** $\binom{n}{k}$:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

2.1.3 Preliminaries on Probability: Further Properties.

Let $\{C_n\}$ be a sequence of events from $(\Omega, \mathcal{F}, P)$.

- If $C_n \subseteq C_{n+1}$ ($\{C_n\}$ is nondecreasing), denote $\lim_{n\to\infty} C_n = \bigcup_{n=1}^{\infty} C_n$, and then

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P(\bigcup_{n=1}^{\infty} C_n).$$

- If $C_n \supseteq C_{n+1}$ ($\{C_n\}$ is nonincreasing), denote $\lim_{n\to\infty} C_n = \bigcap_{n=1}^{\infty} C_n$, and then

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P(\bigcap_{n=1}^{\infty} C_n).$$

- Boole's Inequality. $P(\bigcup_{n=1}^{\infty} C_n) \leq \sum_{n=1}^{\infty} P(C_n)$

What will we study in the next class?

1. *Introduction*
2. **Probability and Distributions**
   - **2.1 Probability (Chp1.1-4)**
   - **2.2 Random Variables and Distributions (Chp1.5-10)**
   - *2.3 Multivariate Distributions (Chp2)*
   - *2.4 Some Important Distributions (Chp3)*
3. *Essential Topics in Mathematical Statistics (Chp 4-6)*
4. *Further Topics, Selected from Chp 7-11*