機率與統計

黃義雄 著
Joey Yu Hsu, MD 編

May 31, 2024

since 2013/02/02

added content from 洪弘 since 2024/05/31

Contents

List of Figures

List of Tables

Chapter 1

recolor

```
\usepackage{expl3,xparse}
\usepackage{xcolor}

\ExplSyntaxOn
\NewDocumentCommand{\recolor}{m}
{
    \tl_set:Nn \l_tmpa_tl { #1 }
    \regex_replace_all:nnN { 2 } { \c{ensuremath}{\c{color}{red}{2}} } \l_tmpa_tl
    \tl_use:N \l_tmpa_tl
}
\ExplSyntaxOff
```

$$c^{\color{red}2} = a^{\color{red}2} + b^{\color{red}2}$$

Chapter 2

機率

2.1 前言

2.1.1 統計學

- statistics
- 對研究主體進行資料蒐集、整理、分析、陳示與推論，以便在不確定之情況下做決策的的學科
- 近代統計學是建立在機率論的基礎上

2.2 機率之定義及基本定理

2.2.1 隨機實驗

Definition 2.2.1. (simple)outcome = sample point

$$\omega$$

Definition 2.2.2. random experiment = statistical experiment = experiment

- every possible outcome can be described before the experiment
- outcomes are not predictable before the experiment
- (an experiment can be carried out repeatedly)

Definition 2.2.3. sample space

$$\Omega = S = U$$

$$\Omega = \{\omega_1, \omega_2, \dots\}$$
$$\Omega_1 = \{\omega_{11}, \omega_{12}, \dots\}, \Omega_2 = \{\omega_{21}, \omega_{22}, \dots\}, \dots$$

$$\begin{aligned} \Omega = \Omega_1 \times \Omega_2 &= \{(\omega_{11}, \omega_{21}), (\omega_{11}, \omega_{22}), \dots, (\omega_{12}, \omega_{21}), (\omega_{12}, \omega_{22}), \dots, \dots\} \\ &= \{\omega_1, \omega_2, \dots\} \Leftrightarrow \begin{cases} \omega_1 = (\omega_{11}, \omega_{21}) \\ \omega_2 = (\omega_{11}, \omega_{22}) \\ \vdots \end{cases} \end{aligned}$$

$$\Omega = \prod_{j \in J \subset \mathbb{N}} \Omega_j = \mathop{\times}_{j \in J \subset \mathbb{N}} \Omega_j$$

Definition 2.2.4. sample point

$$\forall \omega \in \Omega \, (\omega \text{ is a sample point})$$

Definition 2.2.5. event

$$\forall E \subseteq \Omega \, (E \text{ is an event})$$

$$E \subseteq \Omega \Leftrightarrow E \in 2^{\Omega}$$

2.2.2 基本事件

$$E_1, E_2 \subseteq \Omega$$

Definition 2.2.6. additive event

$$E_1 \cup E_2$$

Definition 2.2.7. product event

$$E_1 \cap E_2 = E_1 E_2$$

Definition 2.2.8. complementary event = inverse event

$$\bar{E} = E^{\mathsf{C}}$$

Definition 2.2.9. differential event

$$E_1 - E_2 = E_1 \backslash E_2$$

Definition 2.2.10. impossible event = empty event

$$\emptyset$$

Definition 2.2.11. mutually exclusive event

$$E_1 \cap E_2 = \emptyset \Leftrightarrow E_1, E_2 \text{ are mutually exclusive}$$

Definition 2.2.12. elementary event = simple event

$$\forall E \subseteq \Omega \left(|E| = 1 \Leftrightarrow E \text{ is an elementary event} \right)$$

Definition 2.2.13. sample space of flipping a coin
flip = toss = throw

$$\{H, T\} = \{\text{head}, \text{tail}\} = \{\text{正面}, \text{反面}\}$$

$$\{H, T\}^2 = \{H, T\} \times \{H, T\} = \{(H, H), (H, T), (T, H), (T, T)\} = \{HH, HT, TH, TT\}$$

Definition 2.2.14. game tree

$$\begin{tikzcd}
 & & & w^3 \arrow[dr] & \\
 & & w^2 \arrow[ur] \arrow[dr] & & w^3 l \\
 & w \arrow[ur] \arrow[dr] & & w^2 l \arrow[ur] \arrow[dr] & \\
o \arrow[ur] \arrow[dr] & & wl \arrow[ur] \arrow[dr] & & w^2 l^2 \\
 & l \arrow[ur] \arrow[dr] & & wl^2 \arrow[ur] \arrow[dr] & \\
 & & l^2 \arrow[ur] \arrow[dr] & & wl^3 \\
 & & & l^3 \arrow[ur] &
\end{tikzcd}$$

2.2.3 機率之定義

Definition 2.2.15. probability

- classic probability

$$\mathsf{P}(E) = \frac{|E|}{|\Omega|} = \frac{\#(E)}{\#(\Omega)} = \frac{\mathsf{n}(E)}{\mathsf{n}(\Omega)}$$

- experimental probability = probability estimated by relative frequency in long-term

$$\mathsf{P}(E) = \lim_{\#(\text{experiments})\to\infty} \frac{\#(\text{outcomes in } E)}{\#(\text{experiments})}$$

  - $\because$ 無論$\#(\text{experiments})$多大，$\left|\frac{\#(\text{outcomes in } E)}{\#(\text{experiments})} - \mathsf{P}(E)\right| \geq \epsilon$仍可能發生
  - $\therefore \not\Leftrightarrow \forall\epsilon > 0 \exists N \in \mathbb{N} \forall \#(\text{experiments}) \in \mathbb{N} \left(\#(\text{experiments}) > N \to \left|\frac{\#(\text{outcomes in } E)}{\#(\text{experiments})} - \mathsf{P}(E)\right| < \epsilon\right)$

- 學派
  - classic
  - frequentist
  - 主觀機率學派: e.g. L.J. Savage, Raiffa
    - ⋆ market preference: risk neutral probability
    - ⋆ personal belief
  - neurological

- 雖機率的哲學看法不一，但其代數性質及運算皆一致

- concept of probability distribution

$$p : X(\Omega) \to \left(\{0\} \cup \mathbb{R}^+\right) \Leftrightarrow p : X(\Omega) \to \left(\mathbb{R} - \mathbb{R}^-\right)$$

$$\mathsf{P}(E) = \sum_{\omega \in E} p(X(\omega))$$

2.2.4 機率的公理體系及有關之運算定理

Definition 2.2.16. σ-algebra = σ-field

$$\mathcal{S} \text{ is a } \sigma\text{-algebra over } S \Leftrightarrow \begin{cases} \emptyset \neq S \in \mathcal{S} \subseteq 2^S \\ \forall A \in \mathcal{S}\left(S \backslash A = \bar{A} \in \mathcal{S}\right) \\ \forall i \in I \subseteq \mathbb{N}\left(A_i \in \mathcal{S}\right) \Rightarrow \left(\bigcup_{i \in I} A_i\right) \in \mathcal{S} \end{cases}$$

Definition 2.2.17. measure

$$\mu \text{ is a measure from } \mathcal{S} \Leftrightarrow \begin{cases} \mathcal{S} \text{ is a } \sigma\text{-algebra over } S \\ \mu : \mathcal{S} \to \mathbb{R} \\ \forall A \in \mathcal{S}\left(\mu(A) \geq 0\right) \\ \mu(\emptyset) = 0 \\ \begin{cases} \forall i \in I\left(A_i \in \mathcal{S}\right) \\ \forall i_1, i_2 \in I \subseteq \mathbb{N}\left(A_{i_1} \cap A_{i_2} = \emptyset\right) \end{cases} \Rightarrow \mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i) \end{cases} \tag{2.2.1}$$

- measurable set

$$\forall A \in \mathcal{S}\left(A \text{ is a measurable set}\right)$$

- measurable space

$$(S, \mathcal{S}) \text{ is a measurable space}$$

- measure space

$$(S, \mathcal{S}, \mu) \text{ is a measure space}$$

Definition 2.2.18. probability space

$$(\Omega, \mathcal{F}, \mathsf{P}) \text{ is a probability space} \Leftrightarrow \begin{cases} \mathcal{F} \text{ is a } \sigma\text{-algebra over } \Omega \\ (\Omega, \mathcal{F}, \mathsf{P}) \text{ is a measure space} \\ \mathsf{P} \text{ is a probability measure} \end{cases}$$

$$\mathsf{P} \text{ is a probability measure} \Leftrightarrow \begin{cases} \mathsf{P} \text{ is a measure from } \mathcal{F} \\ 2.2.2 \\ 2.2.3 \end{cases}$$

Axiom 1.

$$\forall E \in \mathcal{F} \left(\mathsf{P}\left(E\right) \geq 0\right) \tag{2.2.2}$$

Axiom 2.

$$\mathsf{P}\left(\Omega\right) = 1 \tag{2.2.3}$$

Axiom 3. additivity

- countable additivity: 隱含在

$$\mathsf{P} \text{ is a measure from } \mathcal{F}$$

 中

$$\begin{aligned} &\begin{cases} \forall i \in I \left(E_i \in \mathcal{F}\right) \\ \forall i_1, i_2 \in I \subseteq \mathbb{N} \left(E_{i_1} \cap E_{i_2} = \emptyset\right) \end{cases} \Rightarrow \mathsf{P}\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} \mathsf{P}\left(E_i\right) \text{ 2.2.1} \\ \Leftrightarrow &\begin{cases} \forall n \in \mathbb{N} \left(B_n \in \mathcal{F}\right) \\ \forall n \in \mathbb{N} \left(B_{n+1} \subseteq B_n\right) \\ \bigcap_{n \in \mathbb{N}} B_n = B \end{cases} \Rightarrow \mathsf{P}\left(B\right) = \lim_{n \to \infty} \mathsf{P}\left(B_n\right) \end{aligned}$$

  - finite additivity: 較弱的型式

Theorem 2.2.1. empty event

$$\mathsf{P}\left(\emptyset\right) = 0$$

為何我覺得這已隱含在measure的定義裡？

Theorem 2.2.2. inverse probability

$$\mathsf{P}\left(\bar{E}\right) = 1 - \mathsf{P}\left(E\right)$$

Proof.

$$\begin{aligned} & 1 \overset{2.2.3}{=} \mathsf{P}\left(\Omega\right) = \mathsf{P}\left(E \cup \bar{E}\right) \overset{3}{=} \mathsf{P}\left(E\right) + \mathsf{P}\left(\bar{E}\right) \\ \Rightarrow \quad & \mathsf{P}\left(\bar{E}\right) = 1 - \mathsf{P}\left(E\right) \end{aligned}$$

- $\mathsf{P}\left(E_1 \cup E_2\right) = 1 - \mathsf{P}\left(\bar{E}_1 \cap \bar{E}_2\right)$
- $\mathsf{P}\left(E_1 \cup E_2 \cup E_3\right) = \mathsf{P}\left(E_1 \cup \left(E_2 \cup E_3\right)\right) = 1 - \mathsf{P}\left(\bar{E}_1 \cap \overline{\left(E_2 \cup E_3\right)}\right) = 1 - \mathsf{P}\left(\bar{E}_1 \cap \left(\bar{E}_2 \cap \bar{E}_3\right)\right) = 1 - \mathsf{P}\left(\bar{E}_1 \cap \bar{E}_2 \cap \bar{E}_3\right)$
- $\mathsf{P}\left(\bigcup_{i \in I} E_i\right) = 1 - \mathsf{P}\left(\bigcap_{i \in I} \bar{E}_i\right)$

□

Theorem 2.2.3. $\mathsf{P}\left(E\right) \in [0, 1]$

Proof.

$$\begin{cases} \mathsf{P}(E) \overset{2.2.2}{\geq} 0 \\ 0 \overset{2.2.2}{\leq} \mathsf{P}(\bar{E}) \overset{\text{thm:}2.2.2}{=} 1 - \mathsf{P}(E) \end{cases}$$
$$\Rightarrow \begin{cases} 0 \leq \mathsf{P}(E) \\ \mathsf{P}(E) \leq 1 \end{cases} \Rightarrow 0 \leq \mathsf{P}(E) \leq 1$$

□

Theorem 2.2.4. inclusion-exclusion principle

- $\mathsf{P}(E_1 \cup E_2) = \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2)$
- $\mathsf{P}(E_1 \cup E_2 \cup E_3) = \mathsf{P}(E_1) + \mathsf{P}(E_2) + \mathsf{P}(E_3) - \mathsf{P}(E_1 \cap E_2) - \mathsf{P}(E_2 \cap E_3) - \mathsf{P}(E_3 \cap E_1) + \mathsf{P}(E_1 \cap E_2 \cap E_3)$
- 通式
$$\mathsf{P}\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n}\left((-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} \mathsf{P}\left(\bigcap_{i \in \{i_1, \ldots, i_k\}} E_i\right)\right)$$

Example 2.2.1. 例4

$$E_1 \cap E_2 = \emptyset \Rightarrow \mathsf{P}(E_1) \leq \mathsf{P}(\bar{E}_2)$$

Proof.

$$\begin{aligned} & 1 \overset{\text{thm:}2.2.3}{\geq} \mathsf{P}(E_1 \cup E_2) \overset{\text{thm:}2.2.4}{=} \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2) \overset{E_1 \cap E_2 = \emptyset}{=} \mathsf{P}(E_1) + \mathsf{P}(E_1) \\ \Rightarrow \quad & 1 - \mathsf{P}(E_2) \geq \mathsf{P}(E_1) \\ \overset{\text{thm:}2.2.2}{\Rightarrow} \quad & \mathsf{P}(\bar{E}_2) \geq \mathsf{P}(E_1) \end{aligned}$$

□

Example 2.2.2. 例5

$$E_1 \cap E_2 \subseteq E \Rightarrow \mathsf{P}(\bar{E}_1) + \mathsf{P}(\bar{E}_2) \geq \mathsf{P}(\bar{E})$$

Proof.

$$\begin{aligned} & E_1 \cap E_2 \subseteq E \\ \Rightarrow \quad & \mathsf{P}(E_1 \cap E_2) \leq \mathsf{P}(E) \\ \Rightarrow \quad & -\mathsf{P}(E_1 \cap E_2) \geq -\mathsf{P}(E) \\ \Rightarrow \quad & 1 - \mathsf{P}(E_1 \cap E_2) \geq 1 - \mathsf{P}(E) \\ \overset{\text{thm:}2.2.2}{\Rightarrow} \quad & \mathsf{P}\left(\overline{E_1 \cap E_2}\right) \geq \mathsf{P}(\bar{E}) \\ \Rightarrow \quad & \mathsf{P}(\bar{E}_1 \cup \bar{E}_2) \geq \mathsf{P}(\bar{E}) \\ \Rightarrow \quad & \mathsf{P}(\bar{E}) \leq \mathsf{P}(\bar{E}_1 \cup \bar{E}_2) = \mathsf{P}(\bar{E}_1) + \mathsf{P}(\bar{E}_2) - \mathsf{P}(\bar{E}_1 \cap \bar{E}_2) \leq \mathsf{P}(\bar{E}_1) + \mathsf{P}(\bar{E}_2) \end{aligned}$$

□

2.2.5 組合分析在機率之應用

- combinatorial analysis

Example 2.2.3. 例6: birthday

$$\begin{cases} |\mathsf{date}| = 365 \\ \Omega_n^{365} = \overset{n}{\underset{i=1}{\times}} \mathsf{date}_i \end{cases}$$

$$
\begin{aligned}
&\mathsf{P}\left(n \text{ people w/o common birthday}\right)\\
=\ & \frac{\mathsf{P}_n^{|\text{date}|}}{|\Omega|} = \frac{\frac{365!}{(365-n)!}}{365^n}\\
=\ & \frac{365 \cdot 364 \cdot 363 \cdot \cdots \cdot (365-(n-1))}{365^n}\\
=\ & \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \cdots \cdot \frac{365-(n-1)}{365}\\
=\ & \left(1-\frac{0}{365}\right)\left(1-\frac{1}{365}\right)\left(1-\frac{2}{365}\right)\cdots\left(1-\frac{n-1}{365}\right) = \prod_{i=1}^{n} \frac{366-i}{365}
\end{aligned}
$$

Figure 2.2.1: birthday

`birthday-1.png` `birthday-2.png`

$$
f(n) = \prod_{i=1}^{n} \frac{366-i}{365}
$$

Example 2.2.4. 例7: occupancy problem

$$
\begin{cases}
|\text{box}| = n \geq m = |\text{ball}|\\
\Omega_m^n = \mathop{\times}\limits_{i=1}^{m} \text{box}_i
\end{cases}
$$

$$
m \geq x \in \mathbb{N}
$$

$$
\begin{aligned}
&\mathsf{P}\left(x \text{ different balls occupy one specific box}\right)\\
=\ & \frac{\mathsf{C}_{x\text{ balls}}^{m\text{ balls}} \left|\Omega_{m-x}^{n-1}\right|}{\left|\Omega_m^n\right|} = \frac{\mathsf{C}_x^m\,(n-1)^{m-x}}{n^m}\\
=\ & \frac{\mathsf{C}_x^m\,(n-1)^{m-x}}{n^{m-x}\cdot n^x} = \mathsf{C}_x^m \left(\frac{n-1}{n}\right)^{m-x} \frac{1}{n^x} = \mathsf{C}_x^m \left(1-\frac{1}{n}\right)^{m-x} \left(\frac{1}{n}\right)^x
\end{aligned}
$$

Example 2.2.5. 例8: Bose-Einstein distribution

$$
\begin{cases}
m \text{ same balls}\\
n \text{ boxes}
\end{cases}
\Rightarrow \mathsf{P}\left(\exists!\text{specific box}\left(x \text{ balls in the specific box}\right)\right) = \frac{\mathsf{C}_{m-x}^{m+n-x-2}}{\mathsf{C}_m^{m+n-1}}
$$

Proof.

$$
\sum_{i=1}^{n} x_i = m\text{的非負整數解集合} = \Omega_m^n
$$

$$
\left|\Omega_m^n\right| = \frac{(m \text{ balls} + n-1 \text{ between-box edges})!}{(m \text{ balls})!\,(n-1 \text{ between-box edges})!} = \mathsf{C}_m^{m+n-1}
$$

$$
\sum_{i=1}^{n-1} x_i = m-x\text{的非負整數解集合} = \Omega_{m-x}^{n-1}
$$

$$
\begin{aligned}
&\mathsf{P}\left(\exists!\text{specific box}\left(x \text{ balls in the specific box}\right)\right)\\
=\ & \frac{\left|\Omega_{m-x}^{n-1}\right|}{\left|\Omega_m^n\right|} = \frac{\mathsf{C}_{(m-x)}^{(m-x)+(n-1)-1}}{\mathsf{C}_m^{m+n-1}} = \frac{\mathsf{C}_{m-x}^{m+n-x-2}}{\mathsf{C}_m^{m+n-1}}
\end{aligned}
$$

□

2.2.6 問題1-2

2.3 條件機率、機率獨立與貝氏定理

Definition 2.3.1. conditional probability

$$\mathsf{P}(E_1) \neq 0 \Rightarrow \mathsf{P}(E_2 \mid E_1) = \frac{|E_1 \cap E_2|}{|E_1|} = \frac{\frac{|E_1 \cap E_2|}{|\Omega|}}{\frac{|E_1|}{|\Omega|}} = \frac{\mathsf{P}(E_1 \cap E_2)}{\mathsf{P}(E_1)} \Leftrightarrow \mathsf{P}(E_1 \cap E_2) = \mathsf{P}(E_2 \mid E_1)\,\mathsf{P}(E_1)$$

Example 2.3.1. 例2

$$\begin{cases} \mathsf{P}(E_1) = p_1 \neq 0 \\ \mathsf{P}(E_2) = p_2 \end{cases} \Rightarrow \mathsf{P}(E_2 \mid E_1) \geq \frac{p_2 + p_1 - 1}{p_1}$$

Proof.

$$\begin{aligned} & 1 \overset{\text{thm:2.2.3}}{\geq} \mathsf{P}(E_1 \cup E_2) \overset{\text{thm:2.2.4}}{=} \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2) \\ \Rightarrow\ & \mathsf{P}(E_1 \cap E_2) \geq \mathsf{P}(E_1) + \mathsf{P}(E_2) - 1 = p_1 + p_2 - 1 \\ \Rightarrow\ & \frac{p_1 + p_2 - 1}{\mathsf{P}(E_1)} \leq \frac{\mathsf{P}(E_1 \cap E_2)}{\mathsf{P}(E_1)} = \mathsf{P}(E_2 \mid E_1) \Rightarrow \mathsf{P}(E_2 \mid E_1) \geq \frac{p_1 + p_2 - 1}{\mathsf{P}(E_1)} = \frac{p_1 + p_2 - 1}{p_1} \end{aligned}$$

□

Theorem 2.3.1. total probability theorem

- $\mathsf{P}(E_1) = \mathsf{P}(E_1 \mid E_2)\,\mathsf{P}(E_2) + \mathsf{P}(E_1 \mid \bar{E}_2)\,\mathsf{P}(\bar{E}_2)$

  Proof.

$$\begin{aligned} (E_1 \cap E_2) \cup (E_1 \cap \bar{E}_2) &= ((E_1 \cap E_2) \cup E_1) \cap ((E_1 \cap E_2) \cup \bar{E}_2) \\ &= E_1 \cap ((E_1 \cup \bar{E}_2) \cap (E_2 \cup \bar{E}_2)) \\ &= E_1 \cap ((E_1 \cup \bar{E}_2) \cap \Omega) \\ &= E_1 \cap (E_1 \cup \bar{E}_2) \\ &= E_1 \end{aligned} \tag{2.3.1}$$

$$\begin{aligned} (E_1 \cap E_2) \cap (E_1 \cap \bar{E}_2) &= ((E_1 \cap E_2) \cap E_1) \cap ((E_1 \cap E_2) \cap \bar{E}_2) \\ &= (E_1 \cap E_2) \cap (E_1 \cap (E_2 \cap \bar{E}_2)) \\ &= (E_1 \cap E_2) \cap (E_1 \cap \emptyset) \\ &= (E_1 \cap E_2) \cap \emptyset \\ &= \emptyset \end{aligned} \tag{2.3.2}$$

$$\begin{aligned} \mathsf{P}(E_1) &\overset{2.3.1}{=} \mathsf{P}((E_1 \cap E_2) \cup (E_1 \cap \bar{E}_2)) \\ &\overset{\text{thm:2.2.4}}{=} \mathsf{P}(E_1 \cap E_2) + \mathsf{P}(E_1 \cap \bar{E}_2) + \mathsf{P}((E_1 \cap E_2) \cap (E_1 \cap \bar{E}_2)) \\ &\overset{2.3.2}{=} \mathsf{P}(E_1 \cap E_2) + \mathsf{P}(E_1 \cap \bar{E}_2) + \mathsf{P}(\emptyset) \\ &\overset{\text{thm:2.2.1}}{=} \mathsf{P}(E_1 \cap E_2) + \mathsf{P}(E_1 \cap \bar{E}_2) + 0 = \mathsf{P}(E_1 \cap E_2) + \mathsf{P}(E_1 \cap \bar{E}_2) \\ &\overset{\text{dfn:2.3.1}}{=} \mathsf{P}(E_1 \mid E_2)\,\mathsf{P}(E_2) + \mathsf{P}(E_1 \mid \bar{E}_2)\,\mathsf{P}(\bar{E}_2) \end{aligned}$$

□

- $\begin{cases} \bigcup_{j \in J} (E_{2j} \cap E_1) = E_1 \\ \forall j_1, j_2 \in J\, (E_{2j_1} \cap E_{2j_2} = \emptyset) \end{cases} \Rightarrow \mathsf{P}(E_1) = \sum_{j \in J} \mathsf{P}(E_1 \mid E_{2j})\,\mathsf{P}(E_{2j})$

2.3.1 差分方程式之應用

$$\begin{aligned} p_n = \mathsf{P}(E_n) &= \mathsf{P}(E_n \mid E_{n-1})\,\mathsf{P}(E_{n-1}) + \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right)\mathsf{P}\left(\overline{E_{n-1}}\right) \\ &= \mathsf{P}(E_n \mid E_{n-1})\,p_{n-1} + \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right)(1-p_{n-1}) \\ &= \left(\mathsf{P}(E_n \mid E_{n-1}) - \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right)\right)p_{n-1} + \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right) \\ &= a\cdot p_{n-1} + b \wedge \begin{cases} a = \mathsf{P}(E_n \mid E_{n-1}) - \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right) \\ b = \mathsf{P}\left(E_n \mid \overline{E_{n-1}}\right) \end{cases} \end{aligned}$$

$$\begin{aligned} p_n &= a\cdot p_{n-1} + b \\ &\Downarrow\ a \neq 1 \\ p_n - \frac{b}{1-a} &= a\cdot p_{n-1} + b - \frac{b}{1-a} = a\cdot p_{n-1} - \frac{a\cdot b}{1-a} \\ &= a\left(p_{n-1} - \frac{b}{1-a}\right) \\ &= a^n\left(p_0 - \frac{b}{1-a}\right) \\ &\Downarrow \\ p_n &= a^n\left(p_0 - \frac{b}{1-a}\right) + \frac{b}{1-a} \\ &= a^n\cdot p_0 + \frac{b}{1-a}(1-a^n) \end{aligned}$$

$$a = 1 \Rightarrow p_n = p_0 + n\cdot b$$

為何等號兩邊要減$\frac{b}{1-a}$呢？是為了形成等比數列

$$\begin{aligned} p_n &= ap_{n-1} + b \\ p_n - c &= ap_{n-1} + b - c \\ &\overset{a\neq 0}{=} a\left(p_{n-1} + \frac{b-c}{a}\right) \\ \text{let } -c &= \frac{b-c}{a} \\ (1-a)\,c &= b \end{aligned}$$

Example 2.3.2. 例8

$$p_n = (1-p)\,p_{n-1} + p\,(1-p_{n-1}) = (1-2p)\,p_{n-1} + p$$

Condition 1. $p = 1$或$p = 0$原式的意義都不對勁

Condition 2. $\begin{cases} 1-2p \neq 1 \Leftrightarrow p \neq 1 \\ p \neq 0 \end{cases}$

$$p_n = (1-2p)^n\left(p_0 - \frac{p}{1-(1-2p)}\right) + \frac{p}{1-(1-2p)} = (1-2p)^n\left(p_0 - \frac{1}{2}\right) + \frac{1}{2}$$

$$0 \text{ is even} \Rightarrow p_0 = 1$$

$$\begin{cases} p_n = (1-2p)^n\left(p_0 - \frac{1}{2}\right) + \frac{1}{2} \\ p_0 = 1 \end{cases}$$

$$\begin{aligned} \Rightarrow\quad p_n &= (1-2p)^n\left(1 - \frac{1}{2}\right) + \frac{1}{2} \\ &= (1-2p)^n\left(\frac{1}{2}\right) + \frac{1}{2} \\ &= \frac{1}{2}\left((1-2p)^n + 1\right) \end{aligned}$$

$$\lim_{n\to\infty} p_n = \begin{cases} \frac{1}{2}\left(0^n+1\right) & p=\frac{1}{2} \\ \frac{1}{2}\left(\left(\lim_{n\to\infty}\left(1-2p\right)^n\right)+1\right) & p\in(0,1)-\left\{\frac{1}{2}\right\} \end{cases}$$
$$= \begin{cases} \frac{1}{2}\left(0+1\right) & p=\frac{1}{2} \\ \frac{1}{2}\left(0+1\right) & p\in(0,1)-\left\{\frac{1}{2}\right\} \end{cases} = \frac{1}{2}$$

Example 2.3.3. 例9: gambler’s ruin

$$\begin{cases} x_A(n)=x \\ x_B(n)=a+b-x \end{cases}$$

$$\begin{cases} \begin{cases} x_A(n)=x_A(n-1)+1 \\ x_B(n)=x_B(n-1)-1 \end{cases} & C_n=H \\ \begin{cases} x_A(n)=x_A(n-1)-1 \\ x_B(n)=x_B(n-1)+1 \end{cases} & C_n=T \end{cases}$$

$$\begin{cases} \begin{cases} \mathsf{P}(C_n=H)=p \\ \mathsf{P}(C_n=T)=1-\mathsf{P}(C_n=H)=1-p=q \end{cases} \\ p>0\wedge pq>0 \end{cases}$$

$$A \text{ lost} \Leftrightarrow A \text{ lost finally} \Leftrightarrow x_A=0$$

$$\begin{aligned} p_x &= \mathsf{P}\left(A \text{ lost with initial } x \text{ after } C_{n-1} \text{ before } C_n\right) \\ &= \mathsf{P}\left(C_n=H\wedge A \text{ lost with initial } x+1\right)+\mathsf{P}\left(C_n=T\wedge A \text{ lost with initial } x-1\right) \\ &= \mathsf{P}\left(A \text{ lost with initial } x+1 \mid C_n=H\right)\mathsf{P}\left(C_n=H\right) \\ &\quad +\mathsf{P}\left(A \text{ lost with initial } x-1 \mid C_n=T\right)\mathsf{P}\left(C_n=T\right) \\ &= \mathsf{P}\left(A \text{ lost with initial } x+1\right)\mathsf{P}\left(C_n=H\right)+\mathsf{P}\left(A \text{ lost with initial } x-1\right)\mathsf{P}\left(C_n=T\right) \\ &= p_{x+1}\cdot p+p_{x-1}\cdot q=pp_{x+1}+qp_{x-1} \end{aligned}$$

上式用到了

$$\mathsf{P}\left(A \text{ lost with initial } x+1 \mid C_n=H\right)=\mathsf{P}\left(A \text{ lost with initial } x+1\right)$$

的假設，假設

$$A \text{ lost with initial } x+1, C_n=H$$

兩件事彼此獨立，然而難道任何事都像這樣
「一次的勝負無關最終的輸贏」嗎？

$$\begin{aligned} pp_{x+1}+qp_{x-1} &= p_x \overset{p+q=1}{=} (p+q)\,p_x \\ p\left(p_{x+1}-p_x\right) &= q\left(p_x-p_{x-1}\right) \\ p_{x+1}-p_x &= \frac{q}{p}\left(p_x-p_{x-1}\right) \text{ 化成如此很對稱} \\ &= \begin{cases} p_x-p_{x-1} & p=q \\ \frac{q}{p}\left(p_x-p_{x-1}\right) & p\neq q \end{cases} \\ &= \begin{cases} p_1-p_0 & p=q \\ \left(\frac{q}{p}\right)^x\left(p_1-p_0\right) & p\neq q \end{cases} \end{aligned}$$

$$\begin{cases} p_0=1 \\ p_{a+b}=0 \end{cases}$$

$$\begin{aligned}
p_{x+1}-p_x &= \begin{cases} p_1-1 & p=q \\ \left(\frac{q}{p}\right)^x (p_1-1) & p\neq q \end{cases} \overset{p_1-1=c}{=} \begin{cases} c & p=q \\ \left(\frac{q}{p}\right)^x c & p\neq q \end{cases} \\
p_{x+1} &= p_x + \begin{cases} c & p=q \\ \left(\frac{q}{p}\right)^x c & p\neq q \end{cases} \\
&= p_0 + \begin{cases} \sum_{k=0}^{x} c & p=q \\ \sum_{k=0}^{x}\left(\frac{q}{p}\right)^k c = c\sum_{k=0}^{x}\left(\frac{q}{p}\right)^k & p\neq q \end{cases} \\
&= p_0 + \begin{cases} (x+1)\,c & p=q \\ c\cdot\frac{1\cdot\left(\left(\frac{q}{p}\right)^{x+1}-1\right)}{\frac{q}{p}-1} & p\neq q \end{cases} \\
p_x &= p_0 + c\cdot\begin{cases} x & p=q \\ \frac{\left(\frac{q}{p}\right)^x-1}{\frac{q}{p}-1} & p\neq q \end{cases}
\end{aligned}$$

$$\begin{aligned}
& p=q \\
\Rightarrow\ & 0=p_{a+b}=p_0+c\,(a+b)=1+(a+b)\,c \\
\Rightarrow\ & c=\frac{-1}{a+b} \\
\Rightarrow\ & p_x=p_0+c\cdot x=1-\frac{x}{a+b}=\frac{a+b-x}{a+b}
\end{aligned}$$

$$\begin{aligned}
& p\neq q \\
\Rightarrow\ & 0=p_{a+b}=p_0+c\cdot\frac{\left(\frac{q}{p}\right)^{a+b}-1}{\frac{q}{p}-1}=1+\frac{\left(\frac{q}{p}\right)^{a+b}-1}{\frac{q}{p}-1}\cdot c \\
\Rightarrow\ & c=\frac{-\left(\frac{q}{p}-1\right)}{\left(\frac{q}{p}\right)^{a+b}-1} \\
\Rightarrow\ & p_x=p_0+c\cdot x=1-\frac{\frac{q}{p}-1}{\left(\frac{q}{p}\right)^{a+b}-1}\cdot\frac{\left(\frac{q}{p}\right)^{x}-1}{\frac{q}{p}-1}=1-\frac{\left(\frac{q}{p}\right)^{x}-1}{\left(\frac{q}{p}\right)^{a+b}-1}=\frac{\left(\frac{q}{p}\right)^{a+b}-\left(\frac{q}{p}\right)^{x}}{\left(\frac{q}{p}\right)^{a+b}-1}
\end{aligned}$$

2.3.2 優勝比

Figure 2.3.1: odds ratio

odds ratio.png

$$\begin{aligned}
& \frac{a+(1-a-b)\,(a+(1-a-b)\,(a+(1-a-b)\cdots))}{a+b+(1-a-b)\,(a+b+(1-a-b)\,(a+b+(1-a-b)\cdots))} \\
=\ & \frac{a+a\,(1-a-b)+a\,(1-a-b)^2+\cdots}{(a+b)+(a+b)\,(1-a-b)+(a+b)\,(1-a-b)^2+\cdots} \\
=\ & \frac{\frac{a}{1-(1-a-b)}}{\frac{a+b}{1-(1-a-b)}}=\frac{a}{a+b}
\end{aligned}$$

Theorem 2.3.2. odds ratio

$$\begin{aligned}&\begin{cases}\mathsf{P}(E_1)=p_1\\\mathsf{P}(E_2)=p_2\\E_1\cap E_2=\emptyset\\p_1+p_2\neq 0\end{cases}\\\Rightarrow\quad&\mathsf{P}(\text{trials until } E_1 \text{ or } E_2 \text{ happens} \Rightarrow E_1 \text{ happens earlier than } E_2)\\&=\mathsf{P}(E_1\mid E_1\cap E_2)=\frac{p_1}{p_1+p_2}\end{aligned}$$

Theorem 2.3.3. multiplication priciple

- $\mathsf{P}(E_1\cap E_2)\overset{2.3.1}{=}\mathsf{P}(E_1)\,\mathsf{P}(E_2\mid E_1)$
- $\mathsf{P}(E_1\cap E_2\cap E_3)=\mathsf{P}(E_1)\,\mathsf{P}(E_2\mid E_1)\,\mathsf{P}(E_3\mid E_1\cap E_2)$
- $\mathsf{P}(\bigcap_{i=0}^{n}E_i)=\prod_{i=0}^{n}\mathsf{P}\left(E_{i+1}\mid\bigcap_{j=0}^{i}E_j\right)\wedge\bigwedge_{i=0}^{n}(E_i\subseteq E_0=\Omega)$

2.3.3 機率獨立

Definition 2.3.2. independence

$$\begin{aligned}&E_1,E_2 \text{ are independent}\\\Leftrightarrow\quad&\mathsf{P}(E_1\cap E_2)=\mathsf{P}(E_1)\,\mathsf{P}(E_2)\\\Leftrightarrow\quad&\begin{cases}\mathsf{P}(E_2\mid E_1)=\mathsf{P}(E_2)&\mathsf{P}(E_1)\neq 0\\\mathsf{P}(E_1\mid E_2)=\mathsf{P}(E_1)&\mathsf{P}(E_2)\neq 0\end{cases}\end{aligned}$$

Theorem 2.3.4. independence triad

- $\begin{cases}\mathsf{P}(E_1\cap E_2)=0\\\mathsf{P}(E_1)\,\mathsf{P}(E_2)=0\end{cases}\Rightarrow\mathsf{P}(E_1\cap E_2)=\mathsf{P}(E_1)\,\mathsf{P}(E_2)$
- $\begin{cases}\mathsf{P}(E_1\cap E_2)=0\\\mathsf{P}(E_1\cap E_2)=\mathsf{P}(E_1)\,\mathsf{P}(E_2)\end{cases}\Rightarrow\mathsf{P}(E_1)\,\mathsf{P}(E_2)=0$
- $\begin{cases}\mathsf{P}(E_1)\,\mathsf{P}(E_2)=0\\\mathsf{P}(E_1\cap E_2)=\mathsf{P}(E_1)\,\mathsf{P}(E_2)\end{cases}\Rightarrow\mathsf{P}(E_1\cap E_2)=0$

Theorem 2.3.5. inverse independence

$$\begin{aligned}[0]\qquad&\mathsf{P}(E_1\cap E_2)=\mathsf{P}(E_1)\,\mathsf{P}(E_2)\\\Rightarrow\quad&\begin{cases}\mathsf{P}\left(E_1\cap\overline{E_2}\right)=\mathsf{P}(E_1)\,\mathsf{P}\left(\overline{E_2}\right)&[1]\\\mathsf{P}\left(\overline{E_1}\cap E_2\right)=\mathsf{P}\left(\overline{E_1}\right)\mathsf{P}(E_2)&[2]\\\mathsf{P}\left(\overline{E_1}\cap\overline{E_2}\right)=\mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right)&[3]\end{cases}\end{aligned}$$

Proof. $[1]$

$$\begin{aligned}\mathsf{P}\left(E_1\cap\overline{E_2}\right)&=\mathsf{P}(E_1-(E_1\cap E_2))\\&=\mathsf{P}(E_1)-\mathsf{P}(E_1\cap E_2)\\&\overset{[0]}{=}\mathsf{P}(E_1)-\mathsf{P}(E_1)\,\mathsf{P}(E_2)\\&=\mathsf{P}(E_1)\,(1-\mathsf{P}(E_2))\\&\overset{\text{thm:}2.2.2}{=}\mathsf{P}(E_1)\,\mathsf{P}\left(\overline{E_2}\right)\end{aligned}$$

$[2]$ similar to $[1]$

[3]

$$\begin{aligned}
\mathsf{P}\left(\overline{E_1}\cap\overline{E_2}\right) &\overset{2.2.2}{=} 1-\mathsf{P}\left(E_1\cup E_2\right)\\
&\overset{\text{thm:}2.2.4}{=} 1-\left(\mathsf{P}\left(E_1\right)+\mathsf{P}\left(E_2\right)-\mathsf{P}\left(E_1\cap E_2\right)\right)\\
&= 1-\mathsf{P}\left(E_1\right)-\left(\mathsf{P}\left(E_2\right)-\mathsf{P}\left(E_1\cap E_2\right)\right)\\
&\overset{\text{thm:}2.2.2}{=} \mathsf{P}\left(\overline{E_1}\right)-\left(\mathsf{P}\left(E_2\right)-\mathsf{P}\left(E_1\cap E_2\right)\right)\\
&\overset{[0]}{=} \mathsf{P}\left(\overline{E_1}\right)-\left(\mathsf{P}\left(E_2\right)-\mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)\right)\\
&= \mathsf{P}\left(\overline{E_1}\right)-\left(1-\mathsf{P}\left(E_1\right)\right)\mathsf{P}\left(E_2\right)\\
&\overset{\text{thm:}2.2.2}{=} \mathsf{P}\left(\overline{E_1}\right)-\mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(E_2\right)\\
&= \mathsf{P}\left(\overline{E_1}\right)\left(1-\mathsf{P}\left(E_2\right)\right)\\
&\overset{\text{thm:}2.2.2}{=} \mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right)
\end{aligned}$$

□

Theorem 2.3.6.

$$\begin{cases}\mathsf{P}\left(E_1\cap E_2\right)=\mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)\\ \mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)\neq 0\end{cases}\Rightarrow \mathsf{P}\left(E_1\cap E_2\right)\neq 0$$

Proof. according to 2.3.4,

$$\begin{aligned}
&\begin{cases}\mathsf{P}\left(E_1\cap E_2\right)=0 & [1]\\ \mathsf{P}\left(E_1\cap E_2\right)=\mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right) & [2]\end{cases}\Rightarrow \mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)=0\ [3]\\
\Rightarrow &\begin{cases}[2]\equiv\mathsf{T}\\ [3]\equiv\mathsf{F}\end{cases}\Rightarrow [1]\equiv\mathsf{F}
\end{aligned}$$

□

Definition 2.3.3.

- E_1, E_2, E_3 are independent $\Leftrightarrow \begin{cases}\text{pairwise independence}\Leftrightarrow\begin{cases}\mathsf{P}\left(E_1\cap E_2\right)=\mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)\\ \mathsf{P}\left(E_2\cap E_3\right)=\mathsf{P}\left(E_2\right)\mathsf{P}\left(E_3\right)\\ \mathsf{P}\left(E_1\cap E_1\right)=\mathsf{P}\left(E_3\right)\mathsf{P}\left(E_1\right)\end{cases}\\ \mathsf{P}\left(E_1\cap E_2\cap E_3\right)=\mathsf{P}\left(E_1\right)\mathsf{P}\left(E_2\right)\mathsf{P}\left(E_3\right)\end{cases}$

- 通式

$$\begin{aligned}
&E_1, E_2, \ldots, E_n \text{ are independent}\\
\Leftrightarrow &\bigwedge_{i=2}^{n}\left(\bigwedge_{1\leq j_1<j_2<\cdots<j_i\leq n}\left(\mathsf{P}\left(\bigcap_{j\in\{j_1,j_2,\ldots,j_i\}}E_j\right)=\prod_{j\in\{j_1,j_2,\ldots,j_i\}}\mathsf{P}\left(E_j\right)\right)\right)
\end{aligned}$$

  –

$$\begin{aligned}
\#\text{ of equations} &= \sum_{i=2}^{n}\mathsf{C}_i^n\\
&= \left(\sum_{i=2}^{n}\left(\mathsf{C}_i^n\cdot 1^{n-i}\cdot 1^i\right)\right)+\left(\sum_{i=0}^{1}\left(\mathsf{C}_i^n\cdot 1^{n-i}\cdot 1^i\right)\right)-\left(\sum_{i=0}^{1}\left(\mathsf{C}_i^n\cdot 1^{n-i}\cdot 1^i\right)\right)\\
&= \left(\sum_{i=0}^{n}\left(\mathsf{C}_i^n\cdot 1^{n-i}\cdot 1^i\right)\right)-\left(\sum_{i=0}^{1}\left(\mathsf{C}_i^n\cdot 1^{n-i}\cdot 1^i\right)\right)\\
&= (1+1)^n-(1+n)=2^n-(n+1)=2^n-n-1
\end{aligned}$$

Fact 2.3.1. $\begin{cases}\text{pairwise independence}\nRightarrow\text{independence}\\ \text{independence}\Rightarrow\text{pairwise independence}\end{cases}$

Theorem 2.3.7. E_1, E_2, E_3 are independent $\Rightarrow \overline{E_1}, \overline{E_2}, \overline{E_3}$ are independent

Proof.

$$
\begin{aligned}
E_1, E_2, E_3 \text{ are independent} \quad & \Rightarrow \quad \begin{cases} \mathsf{P}(E_1 \cap E_2) = \mathsf{P}(E_1)\,\mathsf{P}(E_2) \\ \mathsf{P}(E_2 \cap E_3) = \mathsf{P}(E_2)\,\mathsf{P}(E_3) \\ \mathsf{P}(E_1 \cap E_1) = \mathsf{P}(E_3)\,\mathsf{P}(E_1) \end{cases} \\
& \overset{\text{thm:2.3.5}}{\Rightarrow} \quad \begin{cases} \mathsf{P}\left(\overline{E_1} \cap \overline{E_2}\right) = \mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right) \\ \mathsf{P}\left(\overline{E_2} \cap \overline{E_3}\right) = \mathsf{P}\left(\overline{E_2}\right)\mathsf{P}\left(\overline{E_3}\right) \\ \mathsf{P}\left(\overline{E_3} \cap \overline{E_1}\right) = \mathsf{P}\left(\overline{E_3}\right)\mathsf{P}\left(\overline{E_1}\right) \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{P}\left(\overline{E_1} \cap \overline{E_2} \cap \overline{E_3}\right) & \overset{2.2.2}{=} 1 - \mathsf{P}(E_1 \cup E_2 \cup E_3) \\
& \overset{\text{thm:2.2.4}}{=} 1 - (\mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2))\, 1 - \mathsf{P}(E_1) - (\mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2)) \\
& = \text{similar to thm:2.3.5} \\
& = \mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right)\mathsf{P}\left(\overline{E_3}\right)
\end{aligned}
$$

$$
\begin{cases} \begin{cases} \mathsf{P}\left(\overline{E_1} \cap \overline{E_2}\right) = \mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right) \\ \mathsf{P}\left(\overline{E_2} \cap \overline{E_3}\right) = \mathsf{P}\left(\overline{E_2}\right)\mathsf{P}\left(\overline{E_3}\right) \\ \mathsf{P}\left(\overline{E_3} \cap \overline{E_1}\right) = \mathsf{P}\left(\overline{E_3}\right)\mathsf{P}\left(\overline{E_1}\right) \end{cases} \\ \mathsf{P}\left(\overline{E_1} \cap \overline{E_2} \cap \overline{E_3}\right) = \mathsf{P}\left(\overline{E_1}\right)\mathsf{P}\left(\overline{E_2}\right)\mathsf{P}\left(\overline{E_3}\right) \end{cases} \Leftrightarrow \overline{E_1}, \overline{E_2}, \overline{E_3} \text{ are independent}
$$

□

Example 2.3.4. 例16

$$\mathsf{P}(A \mid B) = \mathsf{P}(A \mid B \cap C)\,\mathsf{P}(C) + \mathsf{P}\left(A \mid B \cap \bar{C}\right)\mathsf{P}\left(\bar{C}\right)$$

2.3.4 機率在電流流通模型之應用

$$
\begin{tikzcd}
 & s_1 \arrow[r] & e_1 \arrow[r] & s_2 \arrow[r] & e_2 \arrow[dr] & \\
s \arrow[ur] \arrow[drr] & & & & & e \\
 & & s_3 \arrow[r] & e_3 \arrow[urr] & &
\end{tikzcd}
$$

$$
\begin{cases} \mathsf{P}(s \to s_i) = 1 \Leftrightarrow s \to s_i = \mathsf{T} \\ \mathsf{P}(e_i \to e) = 1 \Leftrightarrow e_i \to e = \mathsf{T} \\ \mathsf{P}(e_i \to s_j) = 1 \Leftrightarrow e_i \to e_j = \mathsf{T} \quad i \neq j \end{cases}
$$

$$
\begin{aligned}
\mathsf{P}(E) = \mathsf{P}(s \to e) & = \mathsf{P}((s \to s_1 \to e_2 \to e) \vee (s \to s_3 \to e_3 \to e)) \\
& = \mathsf{P}(((s \to s_1) \wedge (s_1 \to e_2) \wedge (e_2 \to e)) \vee (s \to s_3 \to e_3 \to e)) \\
& = \mathsf{P}((\mathsf{T} \wedge (s_1 \to e_2) \wedge \mathsf{T}) \vee (s \to s_3 \to e_3 \to e)) \\
& = \mathsf{P}((s_1 \to e_2) \vee (s \to s_3 \to e_3 \to e)) \\
& = \mathsf{P}((s_1 \to e_2) \vee (s_3 \to e_3)) \\
& = \mathsf{P}((s_1 \to e_1 \to s_2 \to e_2) \vee (s_3 \to e_3)) \\
& = \mathsf{P}(((s_1 \to e_1) \wedge (e_1 \to s_2) \wedge (s_2 \to e_2)) \vee (s_3 \to e_3)) \\
& = \mathsf{P}(((s_1 \to e_1) \wedge \mathsf{T} \wedge (s_2 \to e_2)) \vee (s_3 \to e_3)) \\
& = \mathsf{P}(((s_1 \to e_1) \wedge (s_2 \to e_2)) \vee (s_3 \to e_3)) \\
& = \mathsf{P}((E_1 \cap E_2) \cup E_3)
\end{aligned}
$$

2.3.5 貝氏定理

Theorem 2.3.8. Bayes theorem

- $\mathsf{P}(B \mid A) = \frac{\mathsf{P}(A|B)\mathsf{P}(B)}{\mathsf{P}(A)} = \frac{\mathsf{P}(A|B)\mathsf{P}(B)}{\mathsf{P}(A|B)\mathsf{P}(B)+\mathsf{P}\left(A|\bar{B}\right)\mathsf{P}\left(\bar{B}\right)} = \frac{1}{1+\frac{\mathsf{P}\left(A|\bar{B}\right)\mathsf{P}\left(\bar{B}\right)}{\mathsf{P}(A|B)\mathsf{P}(B)}} = \frac{1}{1+\frac{\mathsf{P}\left(A|\bar{B}\right)(1-\mathsf{P}(B))}{\mathsf{P}(A|B)\mathsf{P}(B)}}$
- $\begin{cases} \bigcup_{j\in J}(B_j \cap A) = A \\ \forall j_1, j_2 \in J\,(B_{j_1} \cap B_{j_2} = \emptyset) \end{cases} \Rightarrow \forall i \in J \left(\mathsf{P}(B_i \mid A) = \frac{\mathsf{P}(A|B_i)\mathsf{P}(B_i)}{\sum_{j\in J}\mathsf{P}(A|B_j)\mathsf{P}(B_j)} = \frac{\mathsf{P}(A|B_i)\mathsf{P}(B_i)}{\mathsf{P}(A)}\right)$

Proof. according to dfn:2.3.1 and thm:2.3.1 □

Example 2.3.5. 例18

$$\begin{cases} \mathsf{P}(+ \mid D) = 98\% \\ \mathsf{P}\left(+ \mid \bar{D}\right) = 3\% \\ \mathsf{P}(D) = 5\% \end{cases}$$

$$\begin{aligned} \mathsf{P}\left(\bar{D} \mid +\right) &= \frac{\mathsf{P}\left(+ \mid \bar{D}\right)\mathsf{P}\left(\bar{D}\right)}{\mathsf{P}\left(+ \mid \bar{D}\right)\mathsf{P}\left(\bar{D}\right) + \mathsf{P}(+ \mid D)\mathsf{P}(D)} \\ &= \frac{1}{1+\frac{\mathsf{P}(+|D)\mathsf{P}(D)}{\mathsf{P}\left(+|\bar{D}\right)\mathsf{P}\left(\bar{D}\right)}} = \frac{1}{1+\frac{\mathsf{P}(+|D)\mathsf{P}(D)}{\mathsf{P}\left(+|\bar{D}\right)(1-\mathsf{P}(D))}} \\ &= \frac{1}{1+\frac{98\cdot 5}{3\cdot 95}} = \frac{1}{1+\frac{98}{3\cdot 19}} = \frac{57}{57+98} = \frac{57}{155} \end{aligned}$$

Example 2.3.6. 例18

$$\begin{cases} \mathsf{P}(K) = p \\ \mathsf{P}(C \mid K) = 1 \\ \mathsf{P}\left(C \mid \bar{K}\right) = \frac{1}{m} \end{cases}$$

$$\begin{aligned} \mathsf{P}(K \mid C) &= \frac{1}{1+\frac{\mathsf{P}\left(C|\bar{K}\right)\mathsf{P}\left(\bar{K}\right)}{\mathsf{P}(C|K)\mathsf{P}(K)}} = \frac{1}{1+\frac{\mathsf{P}\left(C|\bar{K}\right)(1-\mathsf{P}(K))}{\mathsf{P}(C|K)\mathsf{P}(K)}} \\ &= \frac{1}{1+\frac{\frac{1}{m}(1-p)}{1\cdot p}} = \frac{1}{1+\frac{1}{m}\left(\frac{1}{p}-1\right)} = \frac{m}{m-1+\frac{1}{p}} \propto p \end{aligned}$$

2.3.6 問題1-3

Chapter 3

隨機變數之分配

3.1 隨機變數之概念

Definition 3.1.1. random variable = RV

$$\begin{aligned}&\begin{cases}\Omega \text{ is a sample space}\\ \omega \in \Omega\\ X : \Omega \to \mathbb{R}\end{cases}\\ \Leftrightarrow \quad & X(\omega) \text{ is a random variable}\\ \Leftrightarrow \quad & X \text{ is a random variable}\end{aligned}$$

- dicrete RV

$$X(\Omega) \text{ is countable} \Leftrightarrow X \text{ is a discrete RV}$$

- continuous RV

$$X(\Omega) \text{ is uncountable} \Leftrightarrow X \text{ is a continuous RV}$$

3.2 機率密度函數

3.2.1 機率密度函數之定義

Definition 3.2.1. probability function = F

- probability mass function = PMF

$$\begin{cases}X \text{ is a discrete RV}\\ f : X(\Omega) \to (\mathbb{R}^+ \cup \{0\})\\ \sum_{x \in X(\Omega)} f(x) = 1\end{cases} \quad \Leftrightarrow f \text{ is a PMF}$$

- probability density function = PDF

$$\begin{cases}X \text{ is a continuous RV}\\ f : X(\Omega) \to (\mathbb{R}^+ \cup \{0\})\\ \int_{x \in X(\Omega)} f(x)\ \mathrm{d}x = 1\end{cases} \quad \Leftrightarrow f \text{ is a PDF}$$

Definition 3.2.2. event probability

$$\mathsf{P}(E) = \begin{cases}\sum_{x \in X(E)} f(x) & X \text{ is a discrete RV}\\ \int_{x \in X(E)} f(x)\ \mathrm{d}x & X \text{ is a continuous RV}\end{cases} = \begin{cases}\sum_{\omega \in E} f(X(\omega)) & X \text{ is a discrete RV}\\ \int_{\omega \in E} f(X(\omega))\ \mathrm{d}x & X \text{ is a continuous RV}\end{cases}$$

Fact 3.2.1. X is a continuous RV $\Rightarrow \mathsf{P}(a \leq X \leq b) = \mathsf{P}(a < X \leq b) = \mathsf{P}(a \leq X < b) = \mathsf{P}(a < X < b) = \int_a^b f(x)\ \mathrm{d}x$

Proof.

$$\begin{aligned}\mathsf{P}(a \leq X \leq b) &= \mathsf{P}(X = a) + \mathsf{P}(a < X \leq b)\\ &= \int_a^a f(x)\,\mathsf{d}x + \mathsf{P}(a < X \leq b)\\ &= 0 + \mathsf{P}(a < X \leq b)\\ &= \mathsf{P}(a < X \leq b)\end{aligned}$$

□

Example 3.2.1. 例6

$$\Omega_1 = \Omega_2 = \{1, 2, \ldots, 9\}$$

$$\Omega = \Omega_1 \times \Omega_2$$

$$\begin{cases} X_1(\Omega_1) = \Omega_1 \\ X_2(\Omega_2) = \Omega_2 \\ X_1, X_2 \text{ are identity functions} \\ X = \max(X_1, X_2) \end{cases}$$

$$\mathsf{P}(X_1 = x_1) = \frac{|\{x_1\}|}{|\Omega_1|} = \frac{1}{9}$$

$$\mathsf{P}(X_1 \leq x) = \frac{|\{1, 2, \ldots, x\}|}{|\Omega_1|} = \frac{x}{9}$$

$$\begin{cases} E_1 = \{(x_1, x_2) \mid x = x_1 \geq x_2\} \\ E_2 = \{(x_1, x_2) \mid x_1 \leq x_2 = x\} \\ E_1, E_2 \subseteq \Omega \end{cases}$$

$$\begin{aligned}\mathsf{P}(X = x) &= \mathsf{P}(E_1 \cup E_2)\\ &= \mathsf{P}(E_1) + \mathsf{P}(E_2) - \mathsf{P}(E_1 \cap E_2)\\ &= \mathsf{P}(x = X_1 \geq X_2) + \mathsf{P}(X_1 \leq X_2 = x) - \mathsf{P}((x = X_1 \geq X_2) \wedge (X_1 \leq X_2 = x))\\ &= \mathsf{P}(x = X_1 \wedge x \geq X_2) + \mathsf{P}(X_1 \leq x \wedge X_2 = x) - \mathsf{P}(X_1 = x = X_2)\\ &= \mathsf{P}(x = X_1)\,\mathsf{P}(x \geq X_2) + \mathsf{P}(X_1 \leq x)\,\mathsf{P}(X_2 = x) - \mathsf{P}(X_1 = x \wedge x = X_2)\\ &= \mathsf{P}(X_1 = x)\,\mathsf{P}(X_2 \leq x) + \mathsf{P}(X_1 \leq x)\,\mathsf{P}(X_2 = x) - \mathsf{P}(X_1 = x)\,\mathsf{P}(X_2 = x)\\ &= \frac{1}{9} \cdot \frac{x}{9} + \frac{x}{9} \cdot \frac{1}{9} - \frac{1}{9} \cdot \frac{1}{9} = \frac{2x - 1}{81}\end{aligned}$$

$$\begin{aligned}\mathsf{P}(X = x) &= \mathsf{P}(X \leq x) - \mathsf{P}(X \leq x - 1)\\ &= \mathsf{P}(X_1 \leq x \wedge X_2 \leq x) - \mathsf{P}(X_1 \leq x - 1 \wedge X_2 \leq x - 1)\\ &= \mathsf{P}(X_1 \leq x)\,\mathsf{P}(X_2 \leq x) - \mathsf{P}(X_1 \leq x - 1)\,\mathsf{P}(X_2 \leq x - 1)\\ &= \frac{x}{9} \cdot \frac{x}{9} - \frac{x - 1}{9} \cdot \frac{x - 1}{9} = \frac{x^2 - (x - 1)^2}{81} = \frac{2x - 1}{81}\end{aligned}$$

3.2.2 分配函數

Definition 3.2.3. (cumulative)distribution function = CDF = DF

$$F(x) = \mathsf{P}(X \leq x)$$

Fact 3.2.2. $F(x) \in [0, 1]$

Fact 3.2.3. $(x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)) \Leftrightarrow F(x)$ is monotonically increasing

Proof.

$$\begin{aligned}
F(x_2) &= \mathsf{P}(X \le x_2) \\
&\overset{x_1 \le x_2}{=} \mathsf{P}(X \le x_1 \vee x_1 < X \le x_2) \\
&= \mathsf{P}(X \le x_1) + \mathsf{P}(x_1 < X \le x_2) - \mathsf{P}(X \le x_1 \wedge x_1 < X \le x_2) \\
&\overset{x_1 \le x_2}{=} \mathsf{P}(X \le x_1) + \mathsf{P}(x_1 < X \le x_2) - \mathsf{P}(\emptyset) \\
&= \mathsf{P}(X \le x_1) + \mathsf{P}(x_1 < X \le x_2) \\
&\ge \mathsf{P}(X \le x_1) = F(x_1)
\end{aligned}$$

□

Fact 3.2.4. $\begin{cases} F(\infty) = \lim_{x\to\infty} F(x) = 1 \\ F(-\infty) = \lim_{x\to-\infty} F(x) = 0 \end{cases}$

Fact 3.2.5. $\mathsf{P}(a < X \le b) = F(b) - F(a)$

Proof.

$$F(b) = \mathsf{P}(X \le b) = \mathsf{P}(X \le a) + \mathsf{P}(a < X \le b) = F(a) + \mathsf{P}(a < X \le b)$$
$$F(b) - F(a) = \mathsf{P}(a < X \le b)$$

□

Fact 3.2.6. X is a continuous RV $\Rightarrow F(x)$ is continuous

- X is a continuous RV $\Rightarrow F(x)$ is right-continuous

  Proof.

$$\begin{aligned}
F(a^+) - F(a) &= \left(\lim_{x\to a^+} F(x)\right) - F(a) \\
&= \lim_{x\to a^+} (F(x) - F(a)) \\
&= \lim_{x\to a^+} \mathsf{P}(a \le X \le x) \\
&= \lim_{h\to 0} \mathsf{P}(a \le X \le a+h) \\
&= \mathsf{P}\left(\lim_{h\to 0} a \le X \le \lim_{h\to 0}(a+h)\right) \\
&\overset{\text{squeeze thm}}{=} \mathsf{P}(X = a) = \int_a^a f(x)\,\mathrm{d}x = 0
\end{aligned}$$
$$F(a^+) = F(a)$$

□

- X is a continuous RV $\Rightarrow F(x)$ is left-continuous

$$F(b^-) = F(b)$$

Fact 3.2.7. $F(x) = \mathsf{P}(X \le x) = \begin{cases} \sum_{s\le x} f(s) & X \text{ is a discrete RV} \\ \int_{-\infty}^{x} f(s)\,\mathrm{d}s & X \text{ is a continuous RV} \end{cases}$

Claim 3.2.1. X is a continuous RV $\Rightarrow \mathrm{d}\mathsf{P}(X = x) = \mathrm{d}p(x) = f(x)\,\mathrm{d}x$

Fact 3.2.8. 3.2.7 $\overset{\text{fundamental thm of calculus}}{\Rightarrow} f(x) = \begin{cases} F(x) - F(x^-) & X \text{ is a discrete RV} \\ \frac{\mathrm{d}}{\mathrm{d}x} F(x) & X \text{ is a continuous RV} \end{cases}$

Fact 3.2.9. X is a discrete RV $\Rightarrow F(x)$ is stair-case

Definition 3.2.4. symmetric PF

$$\exists x_0 \in X(\Omega)\, \forall x \in X(\Omega)\, (f(-x+x_0) = f(x_0+x)) \Leftrightarrow f \text{ is symmetric about } x_0$$

Fact 3.2.10. $\begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow F(0) = \frac{1}{2}$

Proof.

$$
\begin{aligned}
1 &= \int_{-\infty}^{\infty} f(x)\,\mathsf{d}x \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{0}^{\infty} f(x)\,\mathsf{d}x \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{x=0}^{x=\infty} f(x)\,\mathsf{d}x \\
&\overset{x=-x'}{=} \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{-x'=0}^{-x'=\infty} f(-x')\,\mathsf{d}(-x') \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x - \int_{-x'=\infty}^{-x'=0} f(-x')\,\mathsf{d}(-x') \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x - \int_{x'=-\infty}^{x'=0} f(-x')\,\mathsf{d}(-x') \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x - \left(-\int_{x'=-\infty}^{x'=0} f(-x')\,\mathsf{d}x'\right) \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{x'=-\infty}^{x'=0} f(-x')\,\mathsf{d}x' \\
&\overset{[0]}{=} \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{x'=-\infty}^{x'=0} f(-x')\,\mathsf{d}x' \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{x'=-\infty}^{x'=0} f(x')\,\mathsf{d}x' \\
&= \int_{-\infty}^{0} f(x)\,\mathsf{d}x + \int_{-\infty}^{0} f(x)\,\mathsf{d}x \\
&= 2\int_{-\infty}^{0} f(x)\,\mathsf{d}x
\end{aligned}
$$

$$
\frac{1}{2} = \int_{-\infty}^{0} f(x)\,\mathsf{d}x = F(0)
$$

□

Fact 3.2.11. $\mathsf{P}(X \geq x) = \begin{cases} \mathsf{P}(X \geq 0) - \mathsf{P}(x > X \geq 0) & x \geq 0 \\ \mathsf{P}(X \geq 0) + \mathsf{P}(0 > X \geq x) & x < 0 \end{cases}$

$$
\wedge \begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow \mathsf{P}(X \geq 0) = 1 - \mathsf{P}(X < 0) = 1 - \mathsf{P}(X \leq 0) = 1 - \frac{1}{2} = \frac{1}{2}
$$

Fact 3.2.12. $\begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow F(x) = 1 - F(-x) \Leftrightarrow F(-x) = 1 - F(x) \Leftrightarrow F(-x) + F(x) = 1$

Proof.

$$
\begin{aligned}
F(-x)+F(x) &= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s+\int_{-\infty}^{x} f(s)\,\mathrm{d}s \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s-\int_{x}^{-\infty} f(s)\,\mathrm{d}s \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s-\int_{s=x}^{s=-\infty} f(s)\,\mathrm{d}s \\
&\overset{s=-s'}{=} \int_{-\infty}^{-x} f(s)\,\mathrm{d}s-\int_{-s'=x}^{-s'=-\infty} f(-s')\,\mathrm{d}(-s') \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s-\int_{s'=-x}^{s'=\infty} f(-s')\,\mathrm{d}(-s') \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s-\left(-\int_{s'=-x}^{s'=\infty} f(-s')\,\mathrm{d}s'\right) \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s+\int_{s'=-x}^{s'=\infty} f(-s')\,\mathrm{d}s' \\
&\overset{[0]}{=} \int_{-\infty}^{-x} f(s)\,\mathrm{d}s+\int_{s'=-x}^{s'=\infty} f(s')\,\mathrm{d}s' \\
&= \int_{-\infty}^{-x} f(s)\,\mathrm{d}s+\int_{-x}^{\infty} f(s)\,\mathrm{d}s \\
&= \int_{-\infty}^{\infty} f(s)\,\mathrm{d}s=1
\end{aligned}
$$

□

Fact 3.2.13. $\begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow \mathsf{P}(|X|\geq x)=\begin{cases} 2(1-F(x)) & x\geq 0 \\ 0 & x<0 \end{cases}$

Proof.

$$
\begin{aligned}
\mathsf{P}(|X|\geq x) &= \mathsf{P}(X\geq x \vee X\leq -x) \\
&= \mathsf{P}(X\geq x)+\mathsf{P}(X\leq -x)-\mathsf{P}(X\geq x\wedge X\leq -x) \\
&= \mathsf{P}(X\geq x)+\mathsf{P}(X\leq -x)-\mathsf{P}(\emptyset) \\
&= 1-\mathsf{P}(X<x)+\mathsf{P}(X\leq -x)-0 \\
&\overset{\begin{cases}\mathsf{P}(X=x)=0 \\ \mathsf{P}(X<x\wedge X=x)=\mathsf{P}(\emptyset)\end{cases}}{=} 1-\Big(\mathsf{P}(X<x)+\mathsf{P}(X=x) \\
&\qquad -\mathsf{P}(X<x\wedge X=x)\Big)+\mathsf{P}(X\leq -x) \\
&= 1-\mathsf{P}(X\leq x)+\mathsf{P}(X\leq -x) \\
&= 1-F(x)+F(-x) \\
&\overset{3.2.12}{=} F(-x)+F(-x)=2F(-x) \\
&\overset{3.2.12}{=} 2(1-F(x))
\end{aligned}
$$

□

- $\begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow \mathsf{P}(|X|\geq |x|)=\mathsf{P}(|X|>|x|)=2(1-F(|x|))=2F(-|x|)$

Fact 3.2.14. $\begin{cases} X \text{ is a continuous RV} \\ f \text{ is symmetric about } 0 \quad [0] \end{cases} \Rightarrow \mathsf{P}(|X|\leq |x|)=\mathsf{P}(|X|<|x|)=2F(|x|)-1=1-2F(-|x|)$

Proof.

$$\begin{aligned}
\mathsf{P}(|X| \le |x|) &= \mathsf{P}(|X| < |x|) + \mathsf{P}(|X| = |x|) - \mathsf{P}(\emptyset) \\
&= \mathsf{P}(|X| < |x|) + 0 - 0 \\
&= 1 - \mathsf{P}(|X| \ge |x|) \\
&\overset{3.2.13}{=} \begin{cases} 1 - 2(1 - F(|x|)) = 2F(|x|) - 1 \\ 1 - 2F(-|x|) \end{cases}
\end{aligned}$$

□

3.2.3 p分位數

Definition 3.2.5. $100p^{\text{th}}$ percentile

$$F(x) = \mathsf{P}(X \le x) \ge p \in [0,1] \Leftrightarrow x \text{ is the } 100p^{\text{th}} \text{ percentile}$$

- median

$$F(x) = \frac{1}{2} \Leftrightarrow x \text{ is the } 50^{\text{th}} \text{ percentile} \Leftrightarrow x \text{ is the median}$$

3.2.4 截斷機率函數

Definition 3.2.6. truncated PDF = TPDF

- $\mathsf{P}(x \le a) \ne 0 \Rightarrow f(x \mid X \le a) = \begin{cases} 0 & x \ge a \\ \frac{f(x)}{F(a)} & x < a \end{cases}$

$$\begin{aligned}
F(x \mid X \le a) &= \mathsf{P}(X \le x \mid X \le a) \\
&= \frac{\mathsf{P}(X \le x \wedge X \le a)}{\mathsf{P}(X \le a)} \\
&= \frac{\mathsf{P}(X \le x \wedge X \le a)}{F(a)} \\
&= \begin{cases} \frac{\mathsf{P}(X \le a \le x)}{F(a)} & x \ge a \\ \frac{\mathsf{P}(X \le x < a)}{F(a)} & x < a \end{cases} \\
&= \begin{cases} \frac{\mathsf{P}(X \le a)}{F(a)} & x \ge a \\ \frac{\mathsf{P}(X \le x)}{F(a)} & x < a \end{cases} \\
&= \begin{cases} \frac{F(a)}{F(a)} = 1 & x \ge a \\ \frac{F(x)}{F(a)} & x < a \end{cases}
\end{aligned}$$

$$\begin{aligned}
f(x \mid X \le a) &\overset{3.2.8}{=} \frac{\mathsf{d}}{\mathsf{d}x} F(x \mid X \le a) \\
&= \begin{cases} \frac{\mathsf{d}}{\mathsf{d}x} 1 & x \ge a \\ \frac{\mathsf{d}}{\mathsf{d}x} \frac{F(x)}{F(a)} = \frac{1}{F(a)} \frac{\mathsf{d}}{\mathsf{d}x} F(x) & x < a \end{cases} \\
&= \begin{cases} 0 & x \ge a \\ \frac{1}{F(a)} f(x) = \frac{f(x)}{F(a)} & x < a \end{cases}
\end{aligned}$$

- $\mathsf{P}(a < x \le b) \ne 0 \Rightarrow f(x \mid a < X \le b) = \begin{cases} 0 & x \notin (a,b] \\ \frac{f(x)}{F(b) - F(a)} & x \in (a,b] \end{cases}$

3.2.5 習題2-2

3.3 衍生性PDF

3.3.1 離散型隨機變數

3.3.2 連續型隨機變數

Definition 3.3.1. derivative PF = DPF

- derivative PMF = DPMF

$$Y = t(X) \Leftrightarrow y = t(x) \Rightarrow f_Y(y) = \sum_{x \in \{x \mid t(x) = y\}} f_X(x)$$

- derivative PDF = DPDF

$$\begin{cases} \bigvee_{i \in I} \left(x \in (l_i, u_i] \wedge \begin{cases} y = t_i(x) \\ x = t_i^{-1}(y) \end{cases} \right) \\ \bigcup_{i \in I} (l_i, u_i] = X(\Omega) \end{cases} \Rightarrow f_Y(y) = \sum_{x \in \bigcup_{i \in I} \{x \mid t_i(x) = y\}} \left(f_X(x) \cdot \left| \frac{\mathrm{d}x}{\mathrm{d}y} \right| \right)$$

$$F_X(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f_X(s)\, \mathrm{d}s$$

$$F_X(x \mid l_i < X \leq u_i) = \mathrm{P}(X \leq x \mid l_i < X \leq u_i) = \int_{l_i}^{x} \frac{f_X(s)}{F(u_i) - F(l_i)}\, \mathrm{d}s$$

$$\begin{aligned} F_X(x \mid l_i < X \leq u_i) &= F_Y\left(t_i^{-1}(y) \mid l_i < t_i^{-1}(Y) \leq u_i\right) \\ &= \begin{cases} F_Y(y \mid t_i(l_i) < Y \leq t_i(u_i)) & t_i(l_i) < t_i(u_i) \\ F_Y(y \mid t_i(u_i) \leq Y < t_i(l_i)) & t_i(u_i) < t_i(l_i) \end{cases} \end{aligned}$$

$$\begin{aligned} f_X(x \mid l_i < X \leq u_i) &= \frac{\mathrm{d}}{\mathrm{d}x} F_X(x \mid l_i < X \leq u_i) \\ &= \frac{\mathrm{d}}{\mathrm{d}x} \begin{cases} F_Y(y \mid t_i(l_i) < Y \leq t_i(u_i)) & t_i(l_i) < t_i(u_i) \\ F_Y(y \mid t_i(u_i) \leq Y < t_i(l_i)) & t_i(u_i) < t_i(l_i) \end{cases} \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}y} \begin{cases} F_Y(y \mid t_i(l_i) < Y \leq t_i(u_i)) & t_i(l_i) < t_i(u_i) \\ F_Y(y \mid t_i(u_i) \leq Y < t_i(l_i)) & t_i(u_i) < t_i(l_i) \end{cases} \right) \cdot \frac{\mathrm{d}y}{\mathrm{d}x} \\ &= \left(\begin{cases} f_Y(y \mid t_i(l_i) < Y \leq t_i(u_i)) & t_i(l_i) < t_i(u_i) \\ f_Y(y \mid t_i(u_i) \leq Y < t_i(l_i)) & t_i(u_i) < t_i(l_i) \end{cases} \right) \cdot \frac{\mathrm{d}y}{\mathrm{d}x} \end{aligned}$$

3.3.3 $Y = F(X) \sim U(0,1)$與模擬

Theorem 3.3.1. fundamental theorem of simulation

$$\begin{cases} F_X(x) = \mathrm{P}(X \leq x) \\ Y = F_X(X) \end{cases} \Rightarrow f_Y(y) = \begin{cases} 1 & y \in [0,1] \\ 0 & y \notin [0,1] \end{cases} \Leftrightarrow Y \sim U(0,1)$$

Proof.

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \leq y) \\ &= \mathrm{P}(F_X(X) \leq y) \\ &= \mathrm{P}\left(X \leq F_X^{-1}(y)\right) \\ &= F_X\left(F_X^{-1}(y)\right) \\ &= y \end{aligned}$$

$$\begin{cases} F_Y(y) = \mathrm{P}(Y \leq y) \in [0,1] \\ F_Y(y) = y \end{cases} \Rightarrow y \in [0,1]$$

$$F_Y(y) = \begin{cases} 1 & y > 1 \\ y & 0 \le y \le 1 \\ 0 & y < 0 \end{cases}$$

$$f_Y(y) = \frac{\mathsf{d}}{\mathsf{d}y} F_Y(y) = \frac{\mathsf{d}}{\mathsf{d}y} \begin{cases} 1 & y > 1 \\ y & 0 \le y \le 1 \\ 0 & y < 0 \end{cases} = \begin{cases} 0 & y > 1 \\ 1 & 0 \le y \le 1 \\ 0 & y < 0 \end{cases} = \begin{cases} 1 & y \in [0,1] \\ 0 & y \notin [0,1] \end{cases}$$

□

- $\begin{cases} X, Y \text{ are continuous RV} \\ Y = F_X(X) \end{cases} \quad \Leftrightarrow y = F_X(x) = \int_{-\infty}^{x} f_X(s)\ \mathsf{d}s$

3.3.4 習題2-3

3.4 隨機變數之期望值與變異數

Definition 3.4.1. general definition of expectation

$$\mathsf{E}(t(X)) = \begin{cases} \sum_{x \in X(\Omega)} (t(x) \cdot f(x)) & X \text{ is a discrete RV} \\ \int_{x \in X(\Omega)} (t(x) \cdot f(x))\ \mathsf{d}x & X \text{ is a continuous RV} \end{cases}$$

- expectation = expected value = 期望值

$$\begin{aligned} t(X) = X \quad &\Rightarrow \quad \mathsf{E}(t(X)) = \mathsf{E}(X) = \begin{cases} \sum_{x \in X(\Omega)} (x \cdot f(x)) & X \text{ is a discrete RV} \\ \int_{x \in X(\Omega)} (x \cdot f(x))\ \mathsf{d}x & X \text{ is a continuous RV} \end{cases} \\ &\Leftrightarrow \quad \mathsf{E}(X) \text{ is the expectation of } X \end{aligned}$$

- variance = 變異數

$$\begin{aligned} t(X) = (X - \mathsf{E}(X))^2 \quad &\Rightarrow \quad \mathsf{E}(t(X)) = \mathsf{E}\left((X - \mathsf{E}(X))^2\right) = \begin{cases} \sum_{x \in X(\Omega)} \left((x - \mathsf{E}(X))^2 \cdot f(x)\right) \\ \int_{x \in X(\Omega)} \left((x - \mathsf{E}(X))^2 \cdot f(x)\right)\ \mathsf{d}x \end{cases} \\ &\Leftrightarrow \quad \mathsf{V}(X) = \mathsf{E}\left((X - \mathsf{E}(X))^2\right) \text{ is the variance of } X \end{aligned}$$

- moment generating function = 動差母函數

$$\begin{aligned} t(X) = \mathsf{e}^{\xi X} \quad &\Rightarrow \quad \mathsf{E}(t(X)) = \mathsf{E}\left(\mathsf{e}^{\xi X}\right) = \begin{cases} \sum_{x \in X(\Omega)} \left(\mathsf{e}^{\xi X} \cdot f(x)\right) & X \text{ is a discrete RV} \\ \int_{x \in X(\Omega)} \left(\mathsf{e}^{\xi X} \cdot f(x)\right)\ \mathsf{d}x & X \text{ is a continuous RV} \end{cases} \\ &\Leftrightarrow \quad \mathsf{M}(\xi) = \mathsf{E}\left(\mathsf{e}^{\xi X}\right) \text{ is the expectation of } X \end{aligned}$$

- linearity of expectation

$$\begin{aligned} & \mathsf{E}\left(\sum_{j \in J} (a_j \cdot t_j(X) + b_j)\right) \\ = & \sum_{j \in J} (a_j \cdot \mathsf{E}(t_j(X)) + \mathsf{E}(b_j)) \\ = & \sum_{j \in J} (a_j \cdot \mathsf{E}(t_j(X)) + b_j) \\ = & \left(\sum_{j \in J} (a_j \cdot \mathsf{E}(t_j(X)))\right) + \sum_{j \in J} b_j \end{aligned}$$

Proof. according to the definition of expecation and linearity of integration, apparently □

Definition 3.4.2. standard deviation

$$\sqrt{\mathsf{V}(X)}$$

Theorem 3.4.1. $\mathsf{V}(X) = \mathsf{E}\left(X^2\right) - \left(\mathsf{E}(X)\right)^2$

Proof.

$$\begin{aligned}
\mathsf{V}(X) &= \mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2\right) = \mathsf{E}\left(X^2 - 2X \cdot \mathsf{E}(X) + \left(\mathsf{E}(X)\right)^2\right) \\
&\overset{3.4.1}{=} \mathsf{E}\left(X^2\right) - 2\mathsf{E}(X) \cdot \mathsf{E}(X) + \mathsf{E}\left(\left(\mathsf{E}(X)\right)^2\right) \\
&= \mathsf{E}\left(X^2\right) - 2\mathsf{E}(X) \cdot \mathsf{E}(X) + \mu^2 \\
&= \mathsf{E}\left(X^2\right) - 2\left(\mathsf{E}(X)\right)^2 + \left(\mathsf{E}(X)\right)^2 \\
&= \mathsf{E}\left(X^2\right) - \left(\mathsf{E}(X)\right)^2
\end{aligned}$$

□

Theorem 3.4.2. $\mathsf{V}(aX + b) = a^2\mathsf{V}(X)$

Proof.

$$\begin{aligned}
\mathsf{V}(aX + b) &= \mathsf{E}\left(\left(aX + b - \mathsf{E}(aX + b)\right)^2\right) \\
&\overset{3.4.1}{=} \mathsf{E}\left(\left(aX + b - (a\mathsf{E}(X) + b)\right)^2\right) \\
&= \mathsf{E}\left(\left(a\left(X - \mathsf{E}(X)\right)\right)^2\right) \\
&= \mathsf{E}\left(a^2\left(X - \mathsf{E}(X)\right)^2\right) \\
&\overset{3.4.1}{=} a^2\mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2\right) \\
&\overset{3.4.1}{=} a^2\mathsf{V}(X)
\end{aligned}$$

□

Theorem 3.4.3. $\mathsf{E}(X - \mathsf{E}(X)) = 0$

Proof.

$$\begin{aligned}
\mathsf{E}(X - \mathsf{E}(X)) &\overset{3.4.1}{=} \mathsf{E}(X) - \mathsf{E}(\mathsf{E}(X)) \\
&= \mathsf{E}(X) - \mathsf{E}(X) \\
&= 0
\end{aligned}$$

□

Theorem 3.4.4. $\mathsf{V}(X) \leq \mathsf{E}\left((X - c)^2\right)$

Proof.

$$\begin{aligned}
\mathsf{E}\left((X - c)^2\right) &= \mathsf{E}\left(\left(\left(X - \mathsf{E}(X)\right) + \left(\mathsf{E}(X) - c\right)\right)^2\right) \\
&= \mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2 + 2\left(X - \mathsf{E}(X)\right)\left(\mathsf{E}(X) - c\right) + \left(\mathsf{E}(X) - c\right)^2\right) \\
&\overset{3.4.1}{=} \mathsf{E}\left(\left(X - \mathsf{E}(X)\right)^2\right) + 2\left(\mathsf{E}(X) - c\right)\mathsf{E}\left(X - \mathsf{E}(X)\right) + \mathsf{E}\left(\left(\mathsf{E}(X) - c\right)^2\right) \\
&\overset{3.4.3}{=} \mathsf{V}(X) + \left(\mathsf{E}(X) - c\right)^2 \\
&\geq \mathsf{V}(X)
\end{aligned}$$

$$c = \mathsf{E}(X) \Rightarrow \mathsf{V}(X) = \mathsf{E}\left((X - c)^2\right)$$

□

Theorem 3.4.5. $\begin{cases} f \text{ is the PDF of } X \\ f \text{ is symmetric about } x_0 & [1] \\ \exists \mathsf{E}(X) \in (-\infty, \infty) \left(\mathsf{E}(X) = \int_{-\infty}^{\infty} (x \cdot f(x)) \, \mathsf{d}x \right) & [2] \end{cases} \Rightarrow \mathsf{E}(X) = x_0$

Proof.

$$\begin{aligned}
\mathsf{E}(X) - x_0 &\overset{[2]}{=} \left(\int_{-\infty}^{\infty} (x \cdot f(x)) \, \mathsf{d}x \right) - x_0 \\
&= \int_{-\infty}^{\infty} (x \cdot f(x)) \, \mathsf{d}x - \int_{-\infty}^{\infty} (x_0 \cdot f(x)) \, \mathsf{d}x \\
&= \int_{-\infty}^{\infty} ((x - x_0) \cdot f(x)) \, \mathsf{d}x \\
&= \frac{1}{2} \left(\int_{-\infty}^{\infty} ((x - x_0) \cdot f(x)) \, \mathsf{d}x + \int_{-\infty}^{\infty} ((x - x_0) \cdot f(x)) \, \mathsf{d}x \right) \\
&= \frac{1}{2} \left(\int_{x_1=-\infty}^{x_1=\infty} ((x_1 - x_0) \cdot f(x_1)) \, \mathsf{d}x_1 + \int_{x_2=-\infty}^{x_2=\infty} ((x_2 - x_0) \cdot f(x_2)) \, \mathsf{d}x_2 \right) \\
\begin{cases} x_1 = x_0 + x_1' \\ x_2 = -x_2' + x_0 \end{cases} &= \frac{1}{2} \Bigg(\int_{x_0+x_1'=-\infty}^{x_0+x_1'=\infty} ((x_0 + x_1' - x_0) \cdot f(x_0 + x_1')) \, \mathsf{d}(x_0 + x_1') \\
&\quad + \int_{-x_2'+x_0=-\infty}^{-x_2'+x_0=\infty} ((-x_2' + x_0 - x_0) \cdot f(-x_2' + x_0)) \, \mathsf{d}(-x_2' + x_0) \Bigg) \\
&= \frac{1}{2} \Bigg(\int_{x_1'=-\infty}^{x_1'=\infty} (x_1' \cdot f(x_0 + x_1')) \, \mathsf{d}x_1' \\
&\quad - \int_{x_2'=\infty}^{x_2'=-\infty} ((-x_2') \cdot f(-x_2' + x_0)) \, \mathsf{d}x_2' \Bigg) \\
&= \frac{1}{2} \left(\int_{-\infty}^{\infty} (x \cdot f(x_0 + x)) \, \mathsf{d}x - \int_{\infty}^{-\infty} ((-x) \cdot f(-x + x_0)) \, \mathsf{d}x \right) \\
&= \frac{1}{2} \left(\int_{-\infty}^{\infty} (x \cdot f(x_0 + x)) \, \mathsf{d}x + \int_{-\infty}^{\infty} ((-x) \cdot f(-x + x_0)) \, \mathsf{d}x \right) \\
&= \frac{1}{2} \int_{-\infty}^{\infty} (x \cdot f(x_0 + x) + (-x) \cdot f(-x + x_0)) \, \mathsf{d}x
\end{aligned}$$

$$\begin{aligned}
& x \cdot f(x_0 + x) + (-x) \cdot f(-x + x_0) \\
=\ & x \cdot (f(x_0 + x) - f(-x + x_0)) \\
\overset{[1]}{=}\ & x \cdot 0 = 0
\end{aligned}$$

$$\begin{aligned}
& \mathsf{E}(X) - x_0 = \frac{1}{2} \int_{-\infty}^{\infty} (x \cdot f(x_0 + x) + (-x) \cdot f(-x + x_0)) \, \mathsf{d}x \\
& = \frac{1}{2} \int_{-\infty}^{\infty} 0 \, \mathsf{d}x = 0 \\
\Rightarrow\ & \mathsf{E}(X) = x_0
\end{aligned}$$

□

3.4.1 $\mathsf{E}(X \mid A)$

Definition 3.4.3. conditional expectation

$$\mathsf{E}(X \mid A) = \begin{cases} \sum_{x \in X(\Omega)} (x \cdot f(x \mid A)) & X \text{ is a discrete RV} \\ \int_{x \in X(\Omega)} (x \cdot f(x \mid A)) \, \mathsf{d}x & X \text{ is a continuous RV} \end{cases}$$

Example 3.4.1. 例4

$$X \text{ is a continuous RV} \Rightarrow \mathsf{E}(X \mid X \geq a) = \begin{cases} \frac{\int_{-\infty}^{\infty}(x \cdot f(x))\,\mathrm{d}x}{1-F(a)} \\ \frac{\int_{-\infty}^{\infty}(x \cdot f(x))\,\mathrm{d}x}{\int_a^{\infty} f(x)\,\mathrm{d}x} \end{cases}$$

Proof.

$$f(x \mid X \geq a) \overset{3.2.6}{=} \frac{f(x)}{F(\infty) - F(a)} = \begin{cases} \frac{f(x)}{1-F(a)} \\ \frac{f(x)}{\int_a^{\infty} f(x)\,\mathrm{d}x} \end{cases}$$

$$\begin{aligned} \mathsf{E}(X \mid X \geq a) &= \int_{-\infty}^{\infty} (x \cdot f(x \mid X \geq a))\,\mathrm{d}x \\ &= \begin{cases} \int_{-\infty}^{\infty}\left(x \cdot \frac{f(x)}{1-F(a)}\right)\mathrm{d}x \\ \int_{-\infty}^{\infty}\left(x \cdot \frac{f(x)}{\int_a^{\infty} f(x)\,\mathrm{d}x}\right)\mathrm{d}x \end{cases} \\ &= \begin{cases} \frac{\int_{-\infty}^{\infty}(x \cdot f(x))\,\mathrm{d}x}{1-F(a)} \\ \frac{\int_{-\infty}^{\infty}(x \cdot f(x))\,\mathrm{d}x}{\int_a^{\infty} f(x)\,\mathrm{d}x} \end{cases} \end{aligned}$$

□

Example 3.4.2. 例6

$$\begin{cases} X \in [a,b] \\ \mathsf{E}(X) = \frac{a+b}{2} \end{cases} \quad \Rightarrow \mathsf{V}(X) \leq \frac{(b-a)^2}{4}$$

Proof.

$$\begin{aligned} & a \leq X \leq b \\ \Rightarrow\ & a - \frac{a+b}{2} \leq X - \frac{a+b}{2} \leq b - \frac{a+b}{2} \\ \Rightarrow\ & -\frac{b-a}{2} \leq X - \frac{a+b}{2} \leq \frac{b-a}{2} \\ \Rightarrow\ & \left| X - \frac{a+b}{2} \right| \leq \frac{b-a}{2} \\ \Rightarrow\ & \left(X - \frac{a+b}{2}\right)^2 \leq \left(\frac{b-a}{2}\right)^2 \\ \Rightarrow\ & \mathsf{E}\left(\left(X - \frac{a+b}{2}\right)^2\right) \leq \mathsf{E}\left(\left(\frac{b-a}{2}\right)^2\right) = \left(\frac{b-a}{2}\right)^2 \end{aligned}$$

$$\mathsf{E}(X) = \frac{a+b}{2} \Rightarrow \mathsf{V}(X) = \mathsf{E}\left((X - \mathsf{E}(X))^2\right) = \mathsf{E}\left(\left(X - \frac{a+b}{2}\right)^2\right) \leq \left(\frac{b-a}{2}\right)^2 = \frac{(b-a)^2}{4}$$

□

Theorem 3.4.6. Schwarz inequality

$$(\mathsf{E}(XY))^2 \leq \mathsf{E}(X^2)\,\mathsf{E}(Y^2)$$

Proof.

$$\begin{aligned} \mathsf{E}\left((X - \lambda Y)^2\right) &= \mathsf{E}\left(X^2 - 2\lambda XY + \lambda^2 Y^2\right) \\ &= \mathsf{E}(X^2) - 2\lambda \mathsf{E}(XY) + \lambda^2 \mathsf{E}(Y^2) \\ &= \mathsf{E}(Y^2) \cdot \lambda^2 - 2\mathsf{E}(XY) \cdot \lambda + \mathsf{E}(X^2) \end{aligned}$$

$$\begin{aligned} & \forall \lambda \in \mathbb{R}\left(\mathsf{E}\left((X - \lambda Y)^2\right) \geq 0\right) \\ \Leftrightarrow\ & (-2\mathsf{E}(XY))^2 - 4\mathsf{E}(Y^2) \cdot \mathsf{E}(X^2) \leq 0 \\ \Leftrightarrow\ & 4(\mathsf{E}(XY))^2 - 4\mathsf{E}(X^2) \cdot \mathsf{E}(Y^2) \leq 0 \\ \Leftrightarrow\ & (\mathsf{E}(XY))^2 - \mathsf{E}(X^2) \cdot \mathsf{E}(Y^2) \leq 0 \\ \Leftrightarrow\ & (\mathsf{E}(XY))^2 \leq \mathsf{E}(X^2) \cdot \mathsf{E}(Y^2) \end{aligned}$$

□

3.4.2 期望值與變異數之近似式

Theorem 3.4.7. expectation and variance transformation approximation

$$Y = t(X) \Rightarrow \begin{cases} \mathsf{E}(Y) \approx t(\mathsf{E}(X)) + \frac{\ddot{t}(\mathsf{E}(X))}{2}\mathsf{V}(X) \\ \mathsf{V}(Y) \approx \left(\dot{t}(\mathsf{E}(X))\right)^2 \mathsf{V}(X) \end{cases}$$

Proof.

$$\begin{aligned}
\mathsf{E}(Y) = \mathsf{E}(t(X)) &\overset{\text{Taylor thm}}{\approx} \mathsf{E}\left(\frac{t(\mathsf{E}(X))}{0!}(X-\mathsf{E}(X))^0 + \frac{\dot{t}(\mathsf{E}(X))}{1!}(X-\mathsf{E}(X))^1 + \frac{\ddot{t}(\mathsf{E}(X))}{2!}(X-\mathsf{E}(X))^2\right) \\
&= \mathsf{E}\left(t(\mathsf{E}(X)) + \dot{t}(\mathsf{E}(X))(X-\mathsf{E}(X)) + \frac{\ddot{t}(\mathsf{E}(X))}{2}(X-\mathsf{E}(X))^2\right) \\
&\overset{3.4.1}{=} \mathsf{E}(t(\mathsf{E}(X))) + \dot{t}(\mathsf{E}(X))\,\mathsf{E}(X-\mathsf{E}(X)) + \frac{\ddot{t}(\mathsf{E}(X))}{2}\mathsf{E}\left((X-\mathsf{E}(X))^2\right) \\
&\overset{3.4.3}{=} t(\mathsf{E}(X)) + \dot{t}(\mathsf{E}(X))\cdot 0 + \frac{\ddot{t}(\mathsf{E}(X))}{2}\mathsf{V}(X) = t(\mathsf{E}(X)) + \frac{\ddot{t}(\mathsf{E}(X))}{2}\mathsf{V}(X)
\end{aligned}$$

$$\begin{aligned}
\mathsf{V}(Y) = \mathsf{V}(t(X)) &\overset{\text{Taylor thm}}{\approx} \mathsf{V}\left(\frac{t(\mathsf{E}(X))}{0!}(X-\mathsf{E}(X))^0 + \frac{\dot{t}(\mathsf{E}(X))}{1!}(X-\mathsf{E}(X))^1\right) \\
&= \mathsf{V}\left(t(\mathsf{E}(X)) + \dot{t}(\mathsf{E}(X))(X-\mathsf{E}(X))\right) \\
&= \mathsf{V}\left(\dot{t}(\mathsf{E}(X))\cdot X + \left(t(\mathsf{E}(X)) - \dot{t}(\mathsf{E}(X))\cdot\mathsf{E}(X)\right)\right) \\
&\overset{3.4.2}{=} \left(\dot{t}(\mathsf{E}(X))\right)^2 \mathsf{V}(X)
\end{aligned}$$

□

3.4.3 期望值與分配函數之關係

Theorem 3.4.8. $\mathsf{E}(X) = \int_0^\infty (1-F(x))\,\mathsf{d}x - \int_{-\infty}^0 F(x)\,\mathsf{d}x = \int_0^\infty (F(\infty)-F(x))\,\mathsf{d}x - \int_{-\infty}^0 (F(x)-F(-\infty))\,\mathsf{d}x$

Proof.

$$\int_0^\infty (1-F(x))\,\mathsf{d}x = \int_0^\infty \int_x^\infty f(s)\,\mathsf{d}s\,\mathsf{d}x$$

$$\int_{-\infty}^0 F(x)\,\mathsf{d}x = \int_{-\infty}^0 \int_{-\infty}^x f(s)\,\mathsf{d}s\,\mathsf{d}x$$

□

3.4.4 動差母函數

Definition 3.4.4. moment generating function = MGF

$$\begin{aligned}
\Rightarrow\quad & \forall n \in \mathbb{N}\left(\mathsf{E}(X^n) \in (-\infty,\infty)\right) \\
& \mathsf{M}(\xi) = \mathsf{E}\left(\mathsf{e}^{\xi X}\right) \\
\overset{\text{Maclaurine thm}}{=} & \mathsf{E}\left(\sum_{n=0}^\infty \frac{(\xi X)^n}{n!}\right) \\
\overset{3.4.1}{=} & \sum_{n=0}^\infty \frac{\mathsf{E}(X^n)}{n!}\xi^n
\end{aligned}$$

$$
\begin{aligned}
& \forall n \in \mathbb{N}\left(\mathsf{E}\left(X^n\right) \in (-\infty, \infty)\right) \\
\Rightarrow \quad & \mathsf{E}\left(\frac{1}{1-\xi X}\right) \\
\overset{\text{Maclaurine thm}}{=} \quad & \mathsf{E}\left(\sum_{n=0}^{\infty} (\xi X)^n\right) \\
\overset{3.4.1}{=} \quad & \sum_{n=0}^{\infty} \left(\mathsf{E}\left(X^n\right) \xi^n\right)
\end{aligned}
$$

Theorem 3.4.9. $\mathsf{M}(\xi) = \mathsf{E}\left(\mathsf{e}^{\xi X}\right) \in (-\infty, \infty) \Rightarrow \exists! f\left(f : \mathsf{M}(\xi) \leftrightarrow \mathsf{P}(X = x)\right)$

Fact 3.4.1. $\mathsf{M}(0) = 1$

Proof.

$$\mathsf{M}(0) = \mathsf{E}\left(\mathsf{e}^{0 \cdot X}\right) = \mathsf{E}\left(\mathsf{e}^0\right) = \mathsf{E}(1) = 1$$

□

Fact 3.4.2. $\dot{\mathsf{M}}(0) = \mathsf{E}(X)$

Proof.

$$
\begin{aligned}
\dot{\mathsf{M}}(\xi) &= \frac{\mathsf{d}}{\mathsf{d}\xi} \mathsf{E}\left(\mathsf{e}^{\xi X}\right) \\
&= \frac{\mathsf{d}}{\mathsf{d}\xi} \sum_{n=0}^{\infty} \frac{\mathsf{E}\left(X^n\right)}{n!} \xi^n = \frac{\mathsf{d}}{\mathsf{d}\xi}\left(1 + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^n\right)}{n!} \xi^n\right) \\
&= \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^n\right)}{n!} \frac{\mathsf{d}}{\mathsf{d}\xi} \xi^n = \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^n\right)}{(n-1)!} \xi^{n-1} \\
&= \sum_{n=0}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{n!} \xi^n = \mathsf{E}(X) + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{n!} \xi^n \\
\dot{\mathsf{M}}(0) &= \mathsf{E}(X) + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{n!} 0^n = \mathsf{E}(X) + 0 = \mathsf{E}(X)
\end{aligned}
$$

□

Fact 3.4.3. $\ddot{\mathsf{M}}(0) - \left(\dot{\mathsf{M}}(0)\right)^2 = \mathsf{V}(X)$

Proof.

$$
\begin{aligned}
\ddot{\mathsf{M}}(\xi) &= \frac{\mathsf{d}}{\mathsf{d}\xi} \dot{\mathsf{M}}(\xi) \\
&= \frac{\mathsf{d}}{\mathsf{d}\xi}\left(\mathsf{E}(X) + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{n!} \xi^n\right) \\
&= \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{n!} \frac{\mathsf{d}}{\mathsf{d}\xi} \xi^n = \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+1}\right)}{(n-1)!} \xi^{n-1} \\
&= \sum_{n=0}^{\infty} \frac{\mathsf{E}\left(X^{n+2}\right)}{n!} \xi^n = \mathsf{E}\left(X^2\right) + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+2}\right)}{n!} \xi^n \\
\ddot{\mathsf{M}}(0) &= \mathsf{E}\left(X^2\right) + \sum_{n=1}^{\infty} \frac{\mathsf{E}\left(X^{n+2}\right)}{n!} 0^n = \mathsf{E}\left(X^2\right) + 0 = \mathsf{E}\left(X^2\right)
\end{aligned}
$$

$$\mathsf{V}(X) = \mathsf{E}\left(X^2\right) - \left(\mathsf{E}(X)\right)^2 = \ddot{\mathsf{M}}(0) - \left(\dot{\mathsf{M}}(0)\right)^2$$

□

- indeterminate form

$$\begin{cases} \mathsf{E}(X) = \lim_{\xi\to 0} \dot{\mathsf{M}}(\xi) \\ \mathsf{V}(X) = \lim_{\xi\to 0} \left(\ddot{\mathsf{M}}(\xi) - \left(\dot{\mathsf{M}}(\xi)\right)^2\right) \end{cases}$$

  - L' Hôpital rule

Definition 3.4.5. cumulant

$$\mathsf{C}(\xi) = \ln(\mathsf{M}(\xi)) = \ln \mathsf{M}(\xi)$$

Fact 3.4.4. $\dot{\mathsf{C}}(0) = \mathsf{E}(X)$

Proof.

$$\begin{aligned} \dot{\mathsf{C}}(\xi) &= \frac{\mathsf{d}}{\mathsf{d}\xi} \ln \mathsf{M}(\xi) \\ &= \frac{\dot{\mathsf{M}}(\xi)}{\mathsf{M}(\xi)} \\ \dot{\mathsf{C}}(0) &= \frac{\dot{\mathsf{M}}(0)}{\mathsf{M}(0)} = \frac{\mathsf{E}(X)}{1} = \mathsf{E}(X) \end{aligned}$$

□

Fact 3.4.5. $\ddot{\mathsf{C}}(0) = \mathsf{V}(X)$

Proof.

$$\begin{aligned} \ddot{\mathsf{C}}(\xi) &= \frac{\mathsf{d}}{\mathsf{d}\xi}\dot{\mathsf{C}}(\xi) = \frac{\mathsf{d}}{\mathsf{d}\xi}\frac{\dot{\mathsf{M}}(\xi)}{\mathsf{M}(\xi)} \\ &= \frac{\ddot{\mathsf{M}}(\xi)\,\mathsf{M}(\xi) - \left(\dot{\mathsf{M}}(\xi)\right)^2}{(\mathsf{M}(\xi))^2} \\ \ddot{\mathsf{C}}(0) &= \frac{\ddot{\mathsf{M}}(0)\,\mathsf{M}(0) - \left(\dot{\mathsf{M}}(0)\right)^2}{(\mathsf{M}(0))^2} \\ &= \frac{\ddot{\mathsf{M}}(0)\cdot 1 - \left(\dot{\mathsf{M}}(0)\right)^2}{1^2} = \ddot{\mathsf{M}}(0) - \left(\dot{\mathsf{M}}(0)\right)^2 \\ &= \mathsf{V}(X) \end{aligned}$$

□

Definition 3.4.6. characteristic function

$$\Psi(\xi) = \mathsf{E}\left(\mathsf{e}^{\mathsf{i}\xi X}\right)$$

3.4.5 幾個基本的機率不等式

Theorem 3.4.10. Chebyshev inequality

$$k^2\mathsf{V}(X) > 0 \Rightarrow \mathsf{P}\left(|X - \mathsf{E}(X)| > |k|\sqrt{\mathsf{V}(X)}\right) \le \frac{1}{k^2}$$

Proof.

$$E = \left\{x \mid |x - \mathsf{E}(X)| > |k|\sqrt{\mathsf{V}(X)}\right\}$$

$$
\begin{aligned}
\mathsf{V}(X) &= \int_{x\in X(\Omega)} \left((x-\mathsf{E}(X))^2\cdot f(x)\right)\,\mathsf{d}x \\
&= \int_{x\in E\cup\bar{E}} \left((x-\mathsf{E}(X))^2\cdot f(x)\right)\,\mathsf{d}x \\
&= \int_{x\in E} \left((x-\mathsf{E}(X))^2\cdot f(x)\right)\,\mathsf{d}x + \int_{x\in \bar{E}} \left((x-\mathsf{E}(X))^2\cdot f(x)\right)\,\mathsf{d}x \\
&\geq \int_{x\in E} \left((x-\mathsf{E}(X))^2\cdot f(x)\right)\,\mathsf{d}x \\
&\geq \int_{x\in E} \left(k^2\mathsf{V}(X)\cdot f(x)\right)\,\mathsf{d}x = k^2\mathsf{V}(X)\int_{x\in E} f(x)\,\mathsf{d}x \\
&\Downarrow\ k^2\mathsf{V}(X)>0 \\
\frac{1}{k^2} &\geq \int_{x\in E} f(x)\,\mathsf{d}x = \mathsf{P}\left(|X-\mathsf{E}(X)|>|k|\sqrt{\mathsf{V}(X)}\right)
\end{aligned}
$$

□

Theorem 3.4.11. Markov inequality

$$
\begin{cases}
\mathsf{P}(X<0)=0 & [1] \\
\mathsf{E}(t(X))\in(-\infty,\infty) & \\
a\neq 0 &
\end{cases}
\Rightarrow \mathsf{P}(t(X)\geq a)\begin{cases}\leq \frac{\mathsf{E}(t(X))}{a} & a>0 \\ \geq \frac{\mathsf{E}(t(X))}{a} & a<0\end{cases}
$$

Proof.

$$E=\{x\mid t(x)\geq a\}\cap[0,\infty)$$

$$
\begin{aligned}
\mathsf{E}(t(X)) &= \int_{-\infty}^{\infty} (t(x)\cdot f(x))\,\mathsf{d}x \\
&\stackrel{[1]}{=} \int_{0}^{\infty} (t(x)\cdot f(x))\,\mathsf{d}x \\
&= \int_{x\in E} (t(x)\cdot f(x))\,\mathsf{d}x + \int_{x\in \bar{E}} (t(x)\cdot f(x))\,\mathsf{d}x \\
&\geq \int_{x\in E} (t(x)\cdot f(x))\,\mathsf{d}x \\
&\geq \int_{x\in E} (a\cdot f(x))\,\mathsf{d}x = a\int_{x\in E} f(x)\,\mathsf{d}x \\
\Rightarrow &\quad \mathsf{P}(t(X)\geq a) = \int_{x\in E} f(x)\,\mathsf{d}x \begin{cases}\leq \frac{\mathsf{E}(t(X))}{a} & a>0 \\ \geq \frac{\mathsf{E}(t(X))}{a} & a<0\end{cases}
\end{aligned}
$$

□

Theorem 3.4.12. Jensen inequality

$$f \text{ is convex} \Leftrightarrow \left(x_1\neq x_2 \Rightarrow \frac{f(x_1)+f(x_2)}{2} > f\left(\frac{x_1+x_2}{2}\right)\right)$$

$$t \text{ is convex} \Rightarrow \mathsf{E}(t(X)) > t(\mathsf{E}(X))$$

Proof.

$$t \text{ is convex} \Rightarrow \ddot{t}>0\,[1]$$

$$
\begin{aligned}
\mathsf{E}\left(t\left(X\right)\right) & \overset{\text{Taylor thm}}{\approx} \mathsf{E}\left(\frac{t\left(\mathsf{E}\left(X\right)\right)}{0!}\left(X-\mathsf{E}\left(X\right)\right)^{0}+\frac{\dot{t}\left(\mathsf{E}\left(X\right)\right)}{1!}\left(X-\mathsf{E}\left(X\right)\right)^{1}+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2!}\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & = \mathsf{E}\left(t\left(\mathsf{E}\left(X\right)\right)+\dot{t}\left(\mathsf{E}\left(X\right)\right)\left(X-\mathsf{E}\left(X\right)\right)+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2}\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & \overset{3.4.1}{=} \mathsf{E}\left(t\left(\mathsf{E}\left(X\right)\right)\right)+\dot{t}\left(\mathsf{E}\left(X\right)\right)\mathsf{E}\left(X-\mathsf{E}\left(X\right)\right)+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2}\mathsf{E}\left(\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & = t\left(\mathsf{E}\left(X\right)\right)+\dot{t}\left(\mathsf{E}\left(X\right)\right)\mathsf{E}\left(X-\mathsf{E}\left(X\right)\right)+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2}\mathsf{E}\left(\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & = t\left(\mathsf{E}\left(X\right)\right)+\dot{t}\left(\mathsf{E}\left(X\right)\right)\cdot 0+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2}\mathsf{E}\left(\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & = t\left(\mathsf{E}\left(X\right)\right)+\frac{\ddot{t}\left(\mathsf{E}\left(X\right)\right)}{2}\mathsf{E}\left(\left(X-\mathsf{E}\left(X\right)\right)^{2}\right)\\
 & \overset{[1]}{>} t\left(\mathsf{E}\left(X\right)\right)
\end{aligned}
$$

□

3.4.6 習題2-4

Chapter 4

多變量隨機變數

4.1 結合機率密度函數及結合分配函數

4.1.1 前言—一個引例

4.1.2 結合機率密度函數

$$X_1, X_2 \text{ are discrete} \Rightarrow \mathsf{P}\left(X_1 = x_1 \wedge X_2 = x_2\right) = f_{X_1,X_2}\left(x_1, x_2\right) = f\left(x_1, x_2\right)$$
$$X_1, X_2 \text{ are continuous} \Downarrow$$
$$\mathsf{dP}\left(X_1 = x_1 \wedge X_2 = x_2\right) = f_{X_1,X_2}\left(x_1, x_2\right)\, \mathsf{d}x_1\, \mathsf{d}x_2 = f\left(x_1, x_2\right)\, \mathsf{d}x_1\, \mathsf{d}x_2$$

Definition 4.1.1. joint probability function = JPF

- probability mass function = JPMF

$$\begin{cases} X, Y \text{ are discrete RV} \\ f : X(\Omega) \times Y(\Omega) \to \left(\mathbb{R}^+ \cup \{0\}\right) \\ \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} f(x, y) = 1 \end{cases} \Leftrightarrow f \text{ is a JPMF}$$

- probability density function = JPDF

$$\begin{cases} X, Y \text{ are continuous RV} \\ f : X(\Omega) \times Y(\Omega) \to \left(\mathbb{R}^+ \cup \{0\}\right) \\ \int\limits_{y \in Y(\Omega)} \int\limits_{x \in X(\Omega)} f(x, y)\, \mathsf{d}x\, \mathsf{d}y = 1 \end{cases} \Leftrightarrow f \text{ is a JPDF}$$

- mixed

$$\int\limits_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} f(x, y)\, \mathsf{d}y = 1$$

or

$$\sum_{y \in Y(\Omega)} \int\limits_{x \in X(\Omega)} f(x, y)\, \mathsf{d}x = 1$$

- $f(x, y) = f_{X,Y}(x, y)$

Definition 4.1.2. joint (cumulative)distribution function = JCDF = JDF

$$F(x, y) = \mathsf{P}(X \le x \wedge Y \le y) = \begin{cases} \sum_{v \le y} \sum_{u \le x} f(u, v) & X, Y \text{ are discrete RV} \\ \int\limits_{-\infty}^{y} \int\limits_{-\infty}^{x} f(u, v)\, \mathsf{d}u\, \mathsf{d}v & X, Y \text{ are continuous RV} \end{cases}$$

Fact 4.1.1. $f(x, y) = \begin{cases} \left(F(x, y) - F\left(x, y^-\right)\right) - \left(F\left(x^-, y\right) - F\left(x^-, y^-\right)\right) & X, Y \text{ are discrete RV} \\ \frac{\partial}{\partial x} \frac{\partial}{\partial y} F(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y) & X, Y \text{ are continuous RV} \end{cases}$

Definition 4.1.3. trunted JPF

$$f(x, y \mid A)$$

4.2 邊際密度函數、條件密度函數與機率獨立

$$X_1, X_2 \text{ are discrete} \Rightarrow \mathsf{P}(X_1 = x_1 \wedge X_2 = x_2) = f_{X_1,X_2}(x_1, x_2) = f(x_1, x_2)$$

$$\begin{aligned}
\mathsf{P}(X_1 = x_1) &= \mathsf{P}\left(\bigvee_{x_2 \in X_2(\Omega)} (X_1 = x_1 \wedge X_2 = x_2)\right) \\
&= \sum_{x_2 \in X_2(\Omega)} \mathsf{P}(X_1 = x_1 \wedge X_2 = x_2) \\
&= \sum_{x_2 \in X_2(\Omega)} f_{X_1,X_2}(x_1, x_2) = \sum_{x_2 \in X_2(\Omega)} f(x_1, x_2)
\end{aligned}$$

$$X_1, X_2 \text{ are continuous} \Downarrow$$

$$\mathsf{dP}(X_1 = x_1 \wedge X_2 = x_2) = f_{X_1,X_2}(x_1, x_2)\, \mathsf{d}x_1\, \mathsf{d}x_2 = f(x_1, x_2)\, \mathsf{d}x_1\, \mathsf{d}x_2$$

$$\begin{aligned}
\mathsf{dP}(X_1 = x_1) &= \mathsf{dP}\left(\bigvee_{x_2 \in X_2(\Omega)} (X_1 = x_1 \wedge X_2 = x_2)\right) \\
&= \int_{x_2 \in X_2(\Omega)} \mathsf{dP}(X_1 = x_1 \wedge X_2 = x_2) \\
&= \int_{x_2 \in X_2(\Omega)} f_{X_1,X_2}(x_1, x_2)\, \mathsf{d}x_1\, \mathsf{d}x_2 = \int_{x_2 \in X_2(\Omega)} f(x_1, x_2)\, \mathsf{d}x_1\, \mathsf{d}x_2 \\
&= \left(\int_{x_2 \in X_2(\Omega)} f_{X_1,X_2}(x_1, x_2)\, \mathsf{d}x_2\right) \mathsf{d}x_1 = \left(\int_{x_2 \in X_2(\Omega)} f(x_1, x_2)\, \mathsf{d}x_2\right) \mathsf{d}x_1
\end{aligned}$$

Definition 4.2.1. marginal probability function = MPF

$$\begin{aligned}
& f \text{ is a JPF of } X_1, X_2 \\
\Rightarrow & \begin{cases}
f_1(x) = f_1(x_1) = f_{X_1}(x_1) = f(x_1) = \begin{cases} \sum_{x_2 \in X_2(\Omega)} f(x_1, x_2) & X_2 \text{ is a discrete RV} \\ \int_{x_2 \in X_2(\Omega)} f(x_1, x_2)\, \mathsf{d}x_2 & X_2 \text{ is a continuous RV} \end{cases} \\
f_2(x) = f_2(x_2) = f_{X_2}(x_2) = f(x_2) = \begin{cases} \sum_{x_1 \in X_1(\Omega)} f(x_1, x_2) & X_1 \text{ is a discrete RV} \\ \int_{x_1 \in X_1(\Omega)} f(x_1, x_2)\, \mathsf{d}x_1 & X_1 \text{ is a continuous RV} \end{cases}
\end{cases}
\end{aligned}$$

Definition 4.2.2. conditional probability function = CPF

$$f_2(x_2) \neq 0 \Rightarrow f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

$$f_1(x_1) \neq 0 \Rightarrow f(x_2 \mid x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

Definition 4.2.3. independence

$$\begin{aligned}
& \begin{cases}
f \text{ is a JPF of } X_1, X_2, \ldots, X_n \\
\forall i \in \mathbb{N} \cap [1, n]\ (f_i \text{ is the MPF of } X_i) \\
f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i)
\end{cases} \\
\Leftrightarrow\ & X_1, X_2, \ldots, X_n \text{ are independent}
\end{aligned}$$

Theorem 4.2.1. $\begin{cases} X_1, X_2 \text{ are independent} \\ f_2(x_2) \neq 0 \end{cases} \Rightarrow f(x_1 \mid x_2) \overset{4.2.2}{=} \frac{f(x,y)}{f_2(x_2)} \overset{4.2.3}{=} \frac{f_1(x_1) f_2(x_2)}{f_2(x_2)} = f_1(x_1)$

Definition 4.2.4. 卡氏分割

Theorem 4.2.2. $\begin{cases} f \text{ is a JPF of } X_1, X_2, \ldots, X_n \\ \forall i \in \mathbb{N} \cap [1, n]\ (t_i : 2^{\mathbb{R}} \to \mathbb{R}) \\ f(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} t_i(x_i) \end{cases} \Rightarrow X_1, X_2, \ldots, X_n \text{ are independent}$

4.2.1 雜例

4.3 多變量隨機變數之期望值

4.3.1 隨機變數函數之期望值

Definition 4.3.1. expectation

$$\mathsf{E}\left(t\left(X_1, X_2, \ldots, X_n\right)\right) =$$

$$\sum_{x_1\in X_1(\Omega)} \sum_{x_2\in X_2(\Omega)} \cdots \sum_{x_n\in X_n(\Omega)} \left(t\left(x_1, x_2, \ldots, x_n\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \quad X_1, X_2, \ldots, X_n \text{ are discrete}$$

$$\int_{x_1\in X_1(\Omega)} \int_{x_2\in X_2(\Omega)} \cdots \int_{x_n\in X_n(\Omega)} \left(t\left(x_1, x_2, \ldots, x_n\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \mathsf{d}x_1 \, \mathsf{d}x_2 \cdots \mathsf{d}x_n$$

$$X_1, X_2, \ldots, X_n \text{ are continuous}$$

$$= \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \left(t\left(x_1, x_2, \ldots, x_n\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \left(t\left(x_1, x_2, \ldots, x_n\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases}$$

$$\begin{cases} \sum_{x\in X(\Omega)}^{(n)} = \sum_{x_1\in X_1(\Omega)} \sum_{x_2\in X_2(\Omega)} \cdots \sum_{x_n\in X_n(\Omega)} = \sum_{\left(\overset{n}{\underset{i=1}{,}} x_i\right) \in \overset{n}{\underset{i=1}{\times}} X_i(\Omega)} \\ \int_{x\in X(\Omega)}^{(n)} = \int_{x_1\in X_1(\Omega)} \int_{x_2\in X_2(\Omega)} \cdots \int_{x_n\in X_n(\Omega)} = \int \cdots \int_{\left(\overset{n}{\underset{i=1}{,}} x_i\right) \in \overset{n}{\underset{i=1}{\times}} X_i(\Omega)} \\ \mathsf{d}x^n = \prod_{i=1}^{n} \mathsf{d}x_i \end{cases}$$

- $\mathsf{E}\left(X_i\right) = \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \left(x_i \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \left(x_i \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases}$
- $\mathsf{V}\left(X_i\right) = \mathsf{E}\left(\left(X_i - \mathsf{E}\left(X_i\right)\right)^2\right) =$

$$\begin{cases} \sum_{x\in X(\Omega)}^{(n)} \left(\left(x_i - \mathsf{E}\left(X_i\right)\right)^2 \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \left(\left(x_i - \mathsf{E}\left(X_i\right)\right)^2 \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases}$$

Theorem 4.3.1. independence in expectation

$$\begin{cases} X_1, X_2, \ldots, X_n \text{ are continuous RV} \\ X_1, X_2, \ldots, X_n \text{ are independent} \end{cases} \Rightarrow \mathsf{E}\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) = \prod_{i=1}^{n} \mathsf{E}\left(t_i\left(x_i\right)\right)$$

Proof.

$$\begin{aligned}
& \mathsf{E}\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) \\
= & \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \left(\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \left(\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) \cdot f\left(x_1, x_2, \ldots, x_n\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases} \\
\overset{4.2.3}{=} & \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \left(\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) \cdot \prod_{i=1}^{n} f_i\left(x_i\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \left(\left(\prod_{i=1}^{n} t_i\left(x_i\right)\right) \cdot \prod_{i=1}^{n} f_i\left(x_i\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases} \\
= & \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \prod_{i=1}^{n} \left(t_i\left(x_i\right) \cdot f_i\left(x_i\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \int_{x\in X(\Omega)}^{(n)} \prod_{i=1}^{n} \left(t_i\left(x_i\right) \cdot f_i\left(x_i\right)\right) \mathsf{d}x^n & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases} \\
\overset{\text{integration order}}{=} & \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \prod_{i=1}^{n} \left(t_i\left(x_i\right) \cdot f_i\left(x_i\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \prod_{i=1}^{n} \int_{x_i\in X_i(\Omega)} \left(t_i\left(x_i\right) \cdot f_i\left(x_i\right)\right) \mathsf{d}x_i & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases} \\
= & \begin{cases} \sum_{x\in X(\Omega)}^{(n)} \prod_{i=1}^{n} \left(t_i\left(x_i\right) \cdot f_i\left(x_i\right)\right) & X_1, X_2, \ldots, X_n \text{ are discrete RV} \\ \prod_{i=1}^{n} \mathsf{E}\left(t_i\left(x_i\right)\right) & X_1, X_2, \ldots, X_n \text{ are continuous RV} \end{cases}
\end{aligned}$$

□

4.3.2 多維隨機變數之動差母函數

Definition 4.3.2. MGF in n dimension

$$
\begin{aligned}
& \forall m_i \in \mathbb{N}\left(\mathsf{E}\left(\prod_{i=1}^{n}(X_i)^{m_i}\right) \in (-\infty,\infty)\right) \\
\Rightarrow \quad & \mathsf{M}(\xi_1,\xi_2,\ldots,\xi_n) = \mathsf{E}\left(\mathsf{e}^{\sum_{i=1}^{n}(\xi_i X_i)}\right) \\
= \quad & \mathsf{E}\left(\sum_{m=0}^{\infty}\frac{\left(\sum_{i=1}^{n}(\xi_i X_i)\right)^m}{m!}\right) \\
= \quad & \sum_{m=0}^{\infty}\mathsf{E}\left(\frac{\left(\sum_{i=1}^{n}(\xi_i X_i)\right)^m}{m!}\right) \\
= \quad & \sum_{m=0}^{\infty}\frac{\mathsf{E}\left(\left(\sum_{i=1}^{n}(\xi_i X_i)\right)^m\right)}{m!} \\
= \quad & \sum_{m=0}^{\infty}\frac{\mathsf{E}\left(\sum_{\begin{cases} m_i \in \mathbb{N}\cup\{0\} \\ \sum_{i=1}^{n} m_i = m\end{cases}}\left(\binom{m}{m_1,m_2,\ldots,m_n}\prod_{i=1}^{n}(\xi_i X_i)^{m_i}\right)\right)}{m!} \\
= \quad & \sum_{m=0}^{\infty}\frac{\sum\left(\mathsf{C}\cdot\mathsf{E}\left(\prod(\xi_i X_i)^{m_i}\right)\right)}{m!} \\
= \quad & \sum_{m=0}^{\infty}\frac{\sum\left(\mathsf{C}\cdot\mathsf{E}\left(\prod(X_i)^{m_i}\right)\cdot\prod(\xi_i)^{m_i}\right)}{m!}
\end{aligned}
$$

$$
\sum = \sum_{\begin{cases} m_i \in \mathbb{N}\cup\{0\} \\ \sum_{i=1}^{n} m_i = m\end{cases}}
$$

$$
\mathsf{C} = \binom{m}{m_1,m_2,\ldots,m_n}
$$

$$
\prod = \prod_{i=1}^{n}
$$

- independence condition

$$
\begin{aligned}
\mathsf{M}(\xi_1,\xi_2,\ldots,\xi_n) &= \sum_{m=0}^{\infty}\frac{\sum\left(\mathsf{C}\cdot\mathsf{E}\left(\prod(X_i)^{m_i}\right)\cdot\prod(\xi_i)^{m_i}\right)}{m!} \\
&\overset{4.3.1}{=} \sum_{m=0}^{\infty}\frac{\sum\left(\mathsf{C}\cdot\prod\mathsf{E}\left((X_i)^{m_i}\right)\cdot\prod(\xi_i)^{m_i}\right)}{m!}
\end{aligned}
$$

Fact 4.3.1. $\mathsf{E}(X_k) = \partial_k\mathsf{M}(0,0,\ldots,0)$

Proof.

$$
\frac{\partial}{\partial\xi_k} = \mathsf{D}_k = \partial_k
$$

$$
\begin{aligned}
\partial_k \mathsf{M}(\xi_1,\xi_2,\ldots,\xi_n) &= \partial_k \sum_{m=0}^{\infty} \frac{\sum \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot \prod (\xi_i)^{m_i}\right)}{m!} \\
&= \sum_{m=0}^{\infty} \frac{\sum \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot \partial_k \prod (\xi_i)^{m_i}\right)}{m!} \\
&= \sum_{m=0}^{\infty} \frac{\sum \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)}{m!} \\
&= 0 + ((n-1)\cdot 0 + \mathsf{E}(X_k)) \\
&\quad + \sum_{m=2}^{\infty} \frac{\sum \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)}{m!} \\
&= \mathsf{E}(X_k) \\
&\quad + \sum_{m=2}^{\infty} \frac{\sum_{m_k=0}^{m} \sum_{\sum_{i\neq k} m_i = m-m_k} \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)}{m!}
\end{aligned}
$$

$$
\begin{aligned}
&\sum_{m=2}^{\infty} \frac{\sum \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)}{m!} \\
=\;& \sum_{m=2}^{\infty} \frac{\sum_{m_k=0}^{m} \sum_{\sum_{i\neq k} m_i = m-m_k} \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)}{m!} \\
=\;& \sum_{m=2}^{\infty} \frac{1}{m!}\Bigg(0 + \sum_{\sum_{i\neq k} m_i = m-1} \left(\mathsf{C}\cdot\mathsf{E}\left(X_k \prod_{i\neq k} (X_i)^{m_i}\right)\cdot \prod_{i\neq k} (\xi_i)^{m_i}\right) \\
&+ \sum_{m_k=2}^{m} \sum_{\sum_{i\neq k} m_i = m-m_k} \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)\Bigg) \\
=\;& \sum_{m=2}^{\infty} \frac{1}{m!}\Bigg(\sum_{\sum_{i\neq k} m_i = m-1} \left(\mathsf{C}\cdot\mathsf{E}\left(X_k \prod_{i\neq k} (X_i)^{m_i}\right)\cdot \prod_{i\neq k} (\xi_i)^{m_i}\right) \\
&+ \sum_{m_k=2}^{m} \sum_{\sum_{i\neq k} m_i = m-m_k} \left(\mathsf{C}\cdot\mathsf{E}\left(\prod (X_i)^{m_i}\right)\cdot m_k (\xi_k)^{m_k-1} \prod_{i\neq k} (\xi_i)^{m_i}\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\partial_k \mathsf{M}(0,0,\ldots,0) &= \mathsf{E}(X_k) + \sum_{m=2}^{\infty} \frac{0+0}{m!} \\
&= \mathsf{E}(X_k)
\end{aligned}
$$

- $\mathsf{E}(X_{k_1}X_{k_2}) = \partial_{k_2}\partial_{k_1}\mathsf{M}(0,0,\ldots,0) = \partial_{k_1}\partial_{k_2}\mathsf{M}(0,0,\ldots,0) = \partial^2_{k_1,k_2}\mathsf{M}(0,0,\ldots,0)$
  - $\mathsf{E}\left((X_k)^2\right) = \partial_k\partial_k\mathsf{M}(0,0,\ldots,0) = \partial^2_{k,k}\mathsf{M}(0,0,\ldots,0) == \partial^2_k\mathsf{M}(0,0,\ldots,0)$
- $\mathsf{E}\left(\prod_{k=1}^{n} (X_k)^{p_k}\right) = \partial^{\sum_{k=1}^{n} p_k}_{\overset{n}{\underset{k=1}{,}}(k)^{p_k}}\mathsf{M}(0,0,\ldots,0)$, e.g.

$$
\begin{aligned}
\mathsf{E}\left((X_1)^2 (X_3)^4 (X_6)^5\right) &= \partial^{2+4+5}_{(1)^2,(3)^4,(6)^5}\mathsf{M}(0,0,\ldots,0) \\
&= \partial^{11}_{(1)^2,(3)^4,(6)^5}\mathsf{M}(0,0,\ldots,0)
\end{aligned}
$$

□

Theorem 4.3.2. $\mathsf{C}(\xi_x,\xi_y)=\ln\mathsf{M}(\xi_x,\xi_y)\Rightarrow\begin{cases}\mathsf{E}(X)=\partial_x\mathsf{C}(0,0)\\ \mathsf{V}(X)=\partial_x^2\mathsf{C}(0,0)\\ \mathsf{E}(Y)=\partial_y\mathsf{C}(0,0)\\ \mathsf{V}(Y)=\partial_y^2\mathsf{C}(0,0)\end{cases}$

$$\begin{aligned}\partial_x\mathsf{C}(\xi_x,\xi_y) &= \partial_x\ln\mathsf{M}(\xi_x,\xi_y)=\frac{\partial_x\mathsf{M}(\xi_x,\xi_y)}{\mathsf{M}(\xi_x,\xi_y)}\\ \partial_x\mathsf{C}(0,0) &= \frac{\partial_x\mathsf{M}(0,0)}{\mathsf{M}(0,0)}=\frac{\mathsf{E}(X)}{1}=\mathsf{E}(X)\end{aligned}$$

$$\begin{aligned}\partial_x^2\mathsf{C}(\xi_x,\xi_y) &= \partial_x\partial_x\mathsf{C}(\xi_x,\xi_y)\\ &= \partial_x\frac{\partial_x\mathsf{M}(\xi_x,\xi_y)}{\mathsf{M}(\xi_x,\xi_y)}\\ &= \frac{\left(\partial_x^2\mathsf{M}(\xi_x,\xi_y)\right)\mathsf{M}(\xi_x,\xi_y)-\left(\partial_x\mathsf{M}(\xi_x,\xi_y)\right)^2}{\left(\mathsf{M}(\xi_x,\xi_y)\right)^2}\\ \partial_x^2\mathsf{C}(0,0) &= \frac{\left(\partial_x^2\mathsf{M}(0,0)\right)\mathsf{M}(0,0)-\left(\partial_x\mathsf{M}(0,0)\right)^2}{\left(\mathsf{M}(0,0)\right)^2}\\ &= \frac{\left(\mathsf{E}\left(X^2\right)\right)\cdot 1-\left(\mathsf{E}(X)\right)^2}{1^2}=\mathsf{E}\left(X^2\right)-\left(\mathsf{E}(X)\right)^2\\ &= \mathsf{V}(X)\end{aligned}$$

Fact 4.3.2.

$$\begin{aligned}&\begin{cases}\mathsf{M}(\xi_1,\xi_2,\ldots,\xi_n)=\mathsf{E}\left(\mathrm{e}^{\sum_{i=1}^n(\xi_iX_i)}\right)\\ \mathsf{C}(\xi_1,\xi_2,\ldots,\xi_n)=\ln\mathsf{M}(\xi_1,\xi_2,\ldots,\xi_n)\end{cases}\\ \Rightarrow\quad&\forall k\in\mathbb{N}\cap[1,n]\left(\begin{cases}\mathsf{E}(X_k)=\partial_k\mathsf{C}(0,0,\ldots,0)=\partial_k\mathsf{M}(0,0,\ldots,0)\\ \mathsf{V}(X_k)=\partial_k^2\mathsf{C}(0,0,\ldots,0)=\partial_k^2\mathsf{M}(0,0,\ldots,0)-\left(\mathsf{E}(X_k)\right)^2\end{cases}\right)\end{aligned}$$

4.4 條件期望值

4.4.1 定義

Definition 4.4.1. conditional expectation

$$\mathsf{E}(t(X_1,X_2)\mid X_1=x_1)\overset{f_1(x_1)\neq 0}{=}\begin{cases}\sum_{x_2\in X_2(\Omega)}\left(t(x_1,x_2)\cdot f(x_2\mid x_1)\right) & X_1\text{ is dicrete}\\ \int_{x_2\in X_2(\Omega)}\left(t(x_1,x_2)\cdot f(x_2\mid x_1)\right)\mathrm{d}x_2 & X_1\text{ is continuous}\end{cases}$$

$$\mathsf{E}(t(X_1,X_2)\mid X_2=x_2)\overset{f_2(x_2)\neq 0}{=}\begin{cases}\sum_{x_1\in X_1(\Omega)}\left(t(x_1,x_2)\cdot f(x_1\mid x_2)\right) & X_2\text{ is dicrete}\\ \int_{x_1\in X_1(\Omega)}\left(t(x_1,x_2)\cdot f(x_1\mid x_2)\right)\mathrm{d}x_1 & X_2\text{ is continuous}\end{cases}$$

- conditional mean of Y, given $X=x$

$$\mathsf{E}(Y\mid x)=\mathsf{E}(Y\mid X=x)\overset{f_X(x)\neq 0}{=}\begin{cases}\sum_{y\in Y(\Omega)}\left(y\cdot f(y\mid x)\right) & Y\text{ is dicrete}\\ \int_{y\in Y(\Omega)}\left(y\cdot f(y\mid x)\right)\mathrm{d}y & Y\text{ is continuous}\end{cases}$$

- conditional mean of X, given $Y=y$

$$\mathsf{E}(X\mid y)=\mathsf{E}(X\mid Y=y)\overset{f_Y(y)\neq 0}{=}\begin{cases}\sum_{x\in X(\Omega)}\left(x\cdot f(x\mid y)\right) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(x\cdot f(x\mid y)\right)\mathrm{d}x & X\text{ is continuous}\end{cases}$$

- conditional variance of Y, given $X=x$

$$\mathsf{V}(Y\mid x)=\mathsf{V}(Y\mid X=x)\overset{f_X(x)\neq 0}{=}\begin{cases}\sum_{y\in Y(\Omega)}\left(\left(y-\mathsf{E}(Y\mid x)\right)^2\cdot f(y\mid x)\right) & Y\text{ is dicrete}\\ \int_{y\in Y(\Omega)}\left(\left(y-\mathsf{E}(Y\mid x)\right)^2\cdot f(y\mid x)\right)\mathrm{d}y & Y\text{ is continuous}\end{cases}$$

- conditional variance of X, given $Y=y$

$$\mathsf{V}(X\mid y)=\mathsf{V}(X\mid Y=y)\overset{f_Y(y)\neq 0}{=}\begin{cases}\sum_{x\in X(\Omega)}\left((x-\mathsf{E}(X\mid y))^2\cdot f(x\mid y)\right) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\left((x-\mathsf{E}(X\mid y))^2\cdot f(x\mid y)\right)\mathrm{d}x & X\text{ is continuous}\end{cases}$$

- $\mathsf{E}(Y\mid x)=\mathsf{E}(Y\mid X=x)$, $\mathsf{E}(X\mid y)=\mathsf{E}(X\mid Y=y)$ are **not** RV, but $\mathsf{E}(Y\mid X)$, $\mathsf{E}(X\mid Y)$ are RV
  - $\mathsf{E}(Y\mid x)=\mathsf{E}(Y\mid X=x)$ is a function of x, and $\mathsf{E}(X\mid y)=\mathsf{E}(X\mid Y=y)$ is a function of y
  - $\mathsf{E}(Y\mid X)$ is a RV with X, and $\mathsf{E}(X\mid Y)$ is a RV with Y

- $\mathsf{E}(Y\mid x)=a+bx$ is the heuristic equation into the field of regression research

Example 4.4.1. 例1

Fact 4.4.1. $\mathsf{E}(Y+c\mid x)=\mathsf{E}(Y\mid x)+c$

Fact 4.4.2. $\mathsf{E}(cY\mid x)=c\cdot\mathsf{E}(Y\mid x)$

4.4.2 $\mathsf{E}(\mathsf{E}(Y\mid X))=\mathsf{E}(Y)$之應用

Theorem 4.4.1. $\begin{cases}\mathsf{E}(Y\mid X)\in(-\infty,\infty)\Rightarrow\mathsf{E}(\mathsf{E}(Y\mid X))=\mathsf{E}(Y)\\ \mathsf{E}(X\mid Y)\in(-\infty,\infty)\Rightarrow\mathsf{E}(\mathsf{E}(X\mid Y))=\mathsf{E}(X)\end{cases}$

Proof.

$$\begin{aligned}\mathsf{E}(Y\mid X)&\overset{f_X(X)\neq 0}{=}\begin{cases}\sum_{y\in Y(\Omega)}(y\cdot f(y\mid X)) & Y\text{ is dicrete}\\ \int_{y\in Y(\Omega)}(y\cdot f(y\mid X))\,\mathrm{d}y & Y\text{ is continuous}\end{cases}\\
\mathsf{E}(\mathsf{E}(Y\mid X))&=\begin{cases}\begin{cases}\sum_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)}(y\cdot f(y\mid x))\right)f_X(x) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)}(y\cdot f(y\mid x))\right)f_X(x)\,\mathrm{d}x & X\text{ is continuous}\end{cases} & Y\text{ is dicrete}\\ \begin{cases}\sum_{x\in X(\Omega)}\left(\int_{y\in Y(\Omega)}(y\cdot f(y\mid x))\,\mathrm{d}y\right)f_X(x) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(\int_{y\in Y(\Omega)}(y\cdot f(y\mid x))\,\mathrm{d}y\right)f_X(x)\,\mathrm{d}x & X\text{ is continuous}\end{cases} & Y\text{ is continuous}\end{cases}\\
&\overset{4.2.2}{=}\begin{cases}\begin{cases}\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}(y\cdot f(x,y)) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)}(y\cdot f(x,y))\right)\mathrm{d}x & X\text{ is continuous}\end{cases} & Y\text{ is dicrete}\\ \begin{cases}\sum_{x\in X(\Omega)}\left(\int_{y\in Y(\Omega)}(y\cdot f(x,y))\,\mathrm{d}y\right) & X\text{ is dicrete}\\ \int_{x\in X(\Omega)}\int_{y\in Y(\Omega)}(y\cdot f(x,y))\,\mathrm{d}y\,\mathrm{d}x & X\text{ is continuous}\end{cases} & Y\text{ is continuous}\end{cases}\\
&=\mathsf{E}(Y)\end{aligned}$$

for the same reason,

$$\mathsf{E}(\mathsf{E}(X\mid Y))=\mathsf{E}(X)$$

□

Theorem 4.4.2. $\begin{cases}\mathsf{E}(t(Y)\mid X)\in(-\infty,\infty)\Rightarrow\mathsf{E}(\mathsf{E}(t(Y)\mid X))=\mathsf{E}(t(Y))\\ \mathsf{E}(t(X)\mid Y)\in(-\infty,\infty)\Rightarrow\mathsf{E}(\mathsf{E}(t(X)\mid Y))=\mathsf{E}(t(X))\end{cases}$

Theorem 4.4.3. $\begin{cases}\mathsf{E}(X\mid Y),\mathsf{E}\left(X^2\mid Y\right)\in(-\infty,\infty)\Rightarrow\mathsf{V}(X)=\mathsf{V}(\mathsf{E}(X\mid Y))+\mathsf{E}(\mathsf{V}(X\mid Y))\\ \mathsf{E}(Y\mid X),\mathsf{E}\left(Y^2\mid X\right)\in(-\infty,\infty)\Rightarrow\mathsf{V}(Y)=\mathsf{V}(\mathsf{E}(Y\mid X))+\mathsf{E}(\mathsf{V}(Y\mid X))\end{cases}$

Proof.

$$
\begin{aligned}
& \mathsf{V}(\mathsf{E}(X \mid Y)) + \mathsf{E}(\mathsf{V}(X \mid Y)) \\
= & \left(\mathsf{E}\left((\mathsf{E}(X \mid Y))^2\right) - (\mathsf{E}(\mathsf{E}(X \mid Y)))^2\right) \\
& + \left(\mathsf{E}\left(\mathsf{E}(X^2 \mid Y) - (\mathsf{E}(X \mid Y))^2\right)\right) \\
= & \underline{\mathsf{E}\left((\mathsf{E}(X \mid Y))^2\right)} - (\mathsf{E}(\mathsf{E}(X \mid Y)))^2 \\
& + \mathsf{E}(\mathsf{E}(X^2 \mid Y)) - \underline{\mathsf{E}\left((\mathsf{E}(X \mid Y))^2\right)} \\
= & \mathsf{E}(\mathsf{E}(X^2 \mid Y)) - (\mathsf{E}(\mathsf{E}(X \mid Y)))^2 \\
\overset{4.4.2}{=} & \mathsf{E}(X^2) - (\mathsf{E}(X))^2 = \mathsf{V}(X)
\end{aligned}
$$

for the same reason,

$$\mathsf{V}(\mathsf{E}(Y \mid X)) + \mathsf{E}(\mathsf{V}(Y \mid X)) = \mathsf{V}(Y)$$

□

Corollary 4.4.1.

$$
\begin{cases}
\mathsf{E}(X \mid Y), \mathsf{E}(X^2 \mid Y) \in (-\infty, \infty) \Rightarrow \mathsf{V}(X) = \mathsf{V}(\mathsf{E}(X \mid Y)) + \mathsf{E}(\mathsf{V}(X \mid Y)) \overset{\mathsf{E}(\mathsf{V}(X|Y))\geq 0}{\geq} \mathsf{V}(\mathsf{E}(X \mid Y)) \\
\mathsf{E}(Y \mid X), \mathsf{E}(Y^2 \mid X) \in (-\infty, \infty) \Rightarrow \mathsf{V}(Y) = \mathsf{V}(\mathsf{E}(Y \mid X)) + \mathsf{E}(\mathsf{V}(Y \mid X)) \overset{\mathsf{E}(\mathsf{V}(Y|X))\geq 0}{\geq} \mathsf{V}(\mathsf{E}(Y \mid X))
\end{cases}
$$

4.4.3 $\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}(N)\,\mathsf{E}(\bar{X})$

Theorem 4.4.4. $\begin{cases} \left(\overset{n}{\underset{i=1}{;}} X_i\right), N \text{ are independent} \\ X_i \sim \mathsf{d}(\mu, \sigma) = \text{ambiguis distribution with } \mu, \sigma \end{cases} [1] \Rightarrow \mathsf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}(N)\,\mathsf{E}(\bar{X})$

Proof.

$$
\begin{aligned}
\mathsf{E}\left(\sum_{i=1}^{N} X_i \mid n\right) = \mathsf{E}\left(\sum_{i=1}^{N} X_i \mid N = n\right) & \overset{[1]}{=} \mathsf{E}\left(\sum_{i=1}^{n} X_i\right) = \mathsf{E}(n\bar{X}) = n \cdot \mathsf{E}(\bar{X}) \\
& \Downarrow \\
\mathsf{E}\left(\sum_{i=1}^{N} X_i \mid N\right) & = N \cdot \mathsf{E}(\bar{X}) \\
& \Downarrow \ 4.4.2 \\
\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) = \mathsf{E}\left(\mathsf{E}\left(\sum_{i=1}^{N} X_i \mid N\right)\right) & = \mathsf{E}(N \cdot \mathsf{E}(\bar{X})) \\
& = \mathsf{E}(N) \cdot \mathsf{E}(\bar{X})
\end{aligned}
$$

□

Definition 4.4.2. generalization of conditional expectation

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) \neq 0$$

$$
\begin{aligned}
\mathsf{E}(X_0 \mid x_1, x_2, \ldots, x_n) & = \mathsf{E}(X_0 \mid X_1 = x_1 \wedge X_2 = x_2, \ldots, X_n = x_n) \\
& = \begin{cases} \sum_{x_0 \in X_0(\Omega)} (x_0 \cdot f(x_0 \mid x_1, x_2, \ldots, x_n)) & X_0 \text{ is dicrete} \\ \int_{x_0 \in X_0(\Omega)} (x_0 \cdot f(x_0 \mid x_1, x_2, \ldots, x_n))\, \mathsf{d}x_0 & X_0 \text{ is continuous} \end{cases}
\end{aligned}
$$

4.5 相關係數

4.5.1 共變數

Definition 4.5.1. covariance = cov.

$$\begin{aligned}
\mathrm{cov}(X,Y)=\sigma_{XY}=\mathsf{V}(X,Y) &= \mathsf{E}((X-\mathsf{E}(X))(Y-\mathsf{E}(Y))) \\
&= \mathsf{E}(XY-Y\cdot\mathsf{E}(X)-X\cdot\mathsf{E}(Y)+\mathsf{E}(X)\cdot\mathsf{E}(Y)) \\
&= \mathsf{E}(XY)-\mathsf{E}(Y\cdot\mathsf{E}(X))-\mathsf{E}(X\cdot\mathsf{E}(Y))+\mathsf{E}(\mathsf{E}(X)\cdot\mathsf{E}(Y)) \\
&= \mathsf{E}(XY)-\mathsf{E}(Y)\cdot\mathsf{E}(X)-\mathsf{E}(X)\cdot\mathsf{E}(Y)+\mathsf{E}(X)\cdot\mathsf{E}(Y) \\
&= \mathsf{E}(XY)-\mathsf{E}(X)\cdot\mathsf{E}(Y)=\mathsf{E}(XY)-\mathsf{E}(X)\,\mathsf{E}(Y)
\end{aligned}$$

Theorem 4.5.1. X,Y are independent $\Rightarrow \mathsf{V}(X,Y)=0$

Proof.

$$\begin{aligned}
\mathsf{V}(X,Y) &= \mathsf{E}(XY)-\mathsf{E}(X)\,\mathsf{E}(Y) \\
&\overset{X,Y\text{ are independent}}{=} \mathsf{E}(X)\,\mathsf{E}(Y)-\mathsf{E}(X)\,\mathsf{E}(Y)=0
\end{aligned}$$

□

Theorem 4.5.2. $\mathsf{V}\left(\sum_{j=1}^{n_1} b_{1j}X_{1j},\sum_{j=1}^{n_2} b_{2j}X_{2j}\right)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} b_{1i}b_{2j}\mathsf{V}(X_{1i},X_{2j})$

Proof.

$$\begin{aligned}
\mathsf{V}\left(\sum_{i=1}^{n_1} b_{1i}X_{1i},\sum_{j=1}^{n_2} b_{2j}X_{2j}\right) &= \mathsf{E}\left(\left(\sum_{i=1}^{n_1} b_{1i}X_{1i}\right)\left(\sum_{j=1}^{n_2} b_{2j}X_{2j}\right)\right)-\mathsf{E}\left(\sum_{i=1}^{n_1} b_{1i}X_{1i}\right)\mathsf{E}\left(\sum_{j=1}^{n_2} b_{2j}X_{2j}\right) \\
&= \mathsf{E}\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} b_{1i}b_{2j}X_{1i}X_{2j}\right)-\left(\sum_{i=1}^{n_1}\mathsf{E}(b_{1i}X_{1i})\right)\left(\sum_{j=1}^{n_2}\mathsf{E}(b_{2j}X_{2j})\right) \\
&= \left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} b_{1i}b_{2j}\mathsf{E}(X_{1i}X_{2j})\right)-\left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\mathsf{E}(b_{1i}X_{1i})\,\mathsf{E}(b_{2j}X_{2j})\right) \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} b_{1i}b_{2j}\left(\mathsf{E}(X_{1i}X_{2j})-\mathsf{E}(X_{1i})\,\mathsf{E}(X_{2j})\right) \\
&= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} b_{1i}b_{2j}\mathsf{V}(X_{1i},X_{2j})
\end{aligned}$$

□

Theorem 4.5.3. $\mathsf{V}(X\pm Y)=\mathsf{V}(X)+\mathsf{V}(Y)\pm 2\cdot\mathsf{V}(X,Y)$

Proof.

$$\begin{aligned}
\mathsf{V}(X\pm Y) &= \mathsf{E}\left((X\pm Y)^2\right)-(\mathsf{E}(X\pm Y))^2 \\
&= \mathsf{E}\left(X^2\pm 2XY+Y^2\right) \\
&\quad -(\mathsf{E}(X)\pm\mathsf{E}(Y))^2 \\
&= \mathsf{E}\left(X^2\right)\pm 2\mathsf{E}(XY)+\mathsf{E}\left(Y^2\right) \\
&\quad -\left((\mathsf{E}(X))^2\pm 2\mathsf{E}(X)\,\mathsf{E}(Y)+(\mathsf{E}(Y))^2\right) \\
&= \left(\mathsf{E}\left(X^2\right)-(\mathsf{E}(X))^2\right)+\left(\mathsf{E}\left(Y^2\right)-(\mathsf{E}(Y))^2\right) \\
&\quad \pm 2\left(\mathsf{E}(XY)-\mathsf{E}(X)\,\mathsf{E}(Y)\right) \\
&= \mathsf{V}(X)+\mathsf{V}(Y)\pm 2\cdot\mathsf{V}(X,Y)
\end{aligned}$$

□

Theorem 4.5.4. $\mathsf{V}\left(\sum_{i=1}^{n} X_i\right) = \left(\sum_{i=1}^{n} \mathsf{V}(X_i)\right) + 2\sum_{1\le i<j\le n} \mathsf{V}(X_i, X_j)$

Proof.

$$
\begin{aligned}
\mathsf{V}\left(\sum_{i=1}^{n} X_i\right) &= \mathsf{E}\left(\left(\sum_{i=1}^{n} X_i\right)^2\right) - \left(\mathsf{E}\left(\sum_{i=1}^{n} X_i\right)\right)^2 \\
&= \mathsf{E}\left(\left(\sum_{i=1}^{n} (X_i)^2\right) + 2\sum_{1\le i<j\le n} X_i X_j\right) - \left(\sum_{i=1}^{n} \mathsf{E}(X_i)\right)^2 \\
&= \left(\sum_{i=1}^{n} \mathsf{E}\left((X_i)^2\right)\right) + 2\sum_{1\le i<j\le n} \mathsf{E}(X_i X_j) \\
&\quad - \left(\sum_{i=1}^{n} (\mathsf{E}(X_i))^2 + 2\sum_{1\le i<j\le n} \mathsf{E}(X_i)\,\mathsf{E}(X_j)\right) \\
&= \left(\sum_{i=1}^{n} \left(\mathsf{E}\left((X_i)^2\right) - (\mathsf{E}(X_i))^2\right)\right) + 2\sum_{1\le i<j\le n} (\mathsf{E}(X_i X_j) - \mathsf{E}(X_i)\,\mathsf{E}(X_j)) \\
&= \left(\sum_{i=1}^{n} \mathsf{V}(X_i)\right) + 2\sum_{1\le i<j\le n} \mathsf{V}(X_i, X_j)
\end{aligned}
$$

□

4.5.2 相關係數

Definition 4.5.2. correlation coefficient = CC

$$
\begin{aligned}
\mathsf{CC}(X,Y) = \rho_{XY} = \mathsf{R}(X,Y) &= \mathsf{V}\left(\frac{X - \mathsf{E}(X)}{\sqrt{\mathsf{V}(X)}}, \frac{Y - \mathsf{E}(Y)}{\sqrt{\mathsf{V}(Y)}}\right) \\
&= \mathsf{E}\left(\left(\frac{X - \mathsf{E}(X)}{\sqrt{\mathsf{V}(X)}}\right)\left(\frac{Y - \mathsf{E}(Y)}{\sqrt{\mathsf{V}(Y)}}\right)\right) = \frac{\mathsf{E}((X - \mathsf{E}(X))(Y - \mathsf{E}(Y)))}{\sqrt{\mathsf{V}(X)\,\mathsf{V}(Y)}} \\
&= \frac{\mathsf{V}(X,Y)}{\sqrt{\mathsf{V}(X)\,\mathsf{V}(Y)}} \\
&= \frac{\mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y)}{\sqrt{\mathsf{V}(X)\,\mathsf{V}(Y)}} = \frac{\mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y)}{\sqrt{\left(\mathsf{E}(X^2) - (\mathsf{E}(X))^2\right)\left(\mathsf{E}(Y^2) - (\mathsf{E}(Y))^2\right)}}
\end{aligned}
$$

$$
\begin{cases}
\frac{\mathsf{V}(X,Y)}{\mathsf{V}(X)} = \mathsf{R}(X,Y)\sqrt{\frac{\mathsf{V}(Y)}{\mathsf{V}(X)}} = \frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y) \\
\frac{\mathsf{V}(X,Y)}{\mathsf{V}(Y)} = \mathsf{R}(X,Y)\sqrt{\frac{\mathsf{V}(X)}{\mathsf{V}(Y)}} = \frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)
\end{cases}
$$

Theorem 4.5.5. $\mathsf{V}(b_1X_1 \pm b_2X_2) = (b_1)^2\,\mathsf{V}(X_1) + (b_2)^2\,\mathsf{V}(X_2) \pm 2b_1b_2 \cdot \mathsf{V}(X_1, X_2)$

$$
\mathsf{V}(b_1X_1 \pm b_2X_2) = \left(b_1\sqrt{\mathsf{V}(X_1)}\right)^2 \pm 2\mathsf{R}(X_1,X_2)\,b_1b_2\sqrt{\mathsf{V}(X_1)}\sqrt{\mathsf{V}(X_2)} + \left(b_2\sqrt{\mathsf{V}(X_2)}\right)^2 \tag{4.5.1}
$$

Proof.

$$
\begin{aligned}
\mathsf{V}(b_1X_1 \pm b_2X_2) &\overset{4.5.3}{=} \mathsf{V}(b_1X_1) + \mathsf{V}(b_2X_2) \pm 2 \cdot \mathsf{V}(b_1X_1, b_2X_2) \\
&= (b_1)^2\,\mathsf{V}(X_1) + (b_2)^2\,\mathsf{V}(X_2) \pm 2b_1b_2 \cdot \mathsf{V}(X_1, X_2) \\
&= \left(b_1\sqrt{\mathsf{V}(X_1)}\right)^2 + \left(b_2\sqrt{\mathsf{V}(X_2)}\right)^2 \pm 2b_1b_2\sqrt{\mathsf{V}(X_1)}\sqrt{\mathsf{V}(X_2)}\mathsf{R}(X_1,X_2) \\
&= \left(b_1\sqrt{\mathsf{V}(X_1)}\right)^2 \pm 2\mathsf{R}(X_1,X_2)\,b_1b_2\sqrt{\mathsf{V}(X_1)}\sqrt{\mathsf{V}(X_2)} + \left(b_2\sqrt{\mathsf{V}(X_2)}\right)^2
\end{aligned}
$$

□

4.5.2.1 相關係數重要性質

Fact 4.5.1. $\mathsf{R}(X,Y)\in[-1,1]$

Proof.

$$\begin{aligned}
0 &\leq \mathsf{V}\left(\frac{X}{\sqrt{\mathsf{V}(X)}}+\frac{Y}{\sqrt{\mathsf{V}(Y)}}\right)\\
&\overset{4.5.1}{=} \left(\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(X)}}\right)^2+2\mathsf{R}(X,Y)\frac{\sqrt{\mathsf{V}(X)}\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}\sqrt{\mathsf{V}(Y)}}+\left(\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(Y)}}\right)^2\\
&= 1+2\mathsf{R}(X,Y)+1\\
&= 2+2\mathsf{R}(X,Y)
\end{aligned}$$

$$\mathsf{R}(X,Y)\geq -1$$

$$\begin{aligned}
0 &\leq \mathsf{V}\left(\frac{X}{\sqrt{\mathsf{V}(X)}}-\frac{Y}{\sqrt{\mathsf{V}(Y)}}\right)\\
&\overset{4.5.1}{=} \left(\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(X)}}\right)^2-2\mathsf{R}(X,Y)\frac{\sqrt{\mathsf{V}(X)}\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}\sqrt{\mathsf{V}(Y)}}+\left(\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(Y)}}\right)^2\\
&= 1-2\mathsf{R}(X,Y)+1\\
&= 2-2\mathsf{R}(X,Y)
\end{aligned}$$

$$\mathsf{R}(X,Y)\leq 1$$

$$\begin{cases}\mathsf{R}(X,Y)\leq 1\\ \mathsf{R}(X,Y)\geq -1\end{cases}\Rightarrow -1\leq \mathsf{R}(X,Y)\leq 1$$

□

Theorem 4.5.6. $\begin{cases}X,Y \text{ are independent}\Rightarrow \mathsf{R}(X,Y)=0\\ \mathsf{R}(X,Y)=0\not\Rightarrow X,Y \text{ are independent}\end{cases}$

Proof.

$$=X,Y \text{ are independent}\Rightarrow \mathsf{V}(X,Y)=0\Rightarrow \mathsf{R}(X,Y)=\frac{\mathsf{V}(X,Y)}{\sqrt{\mathsf{V}(X)\mathsf{V}(Y)}}=0$$

□

4.5.3 迴歸方程式與相關係數

Definition 4.5.3. regression line

- regression line of Y on X

$$\mathsf{E}(Y\mid x)=a+bx$$

- regression line of X on Y

$$\mathsf{E}(X\mid y)=c+dx$$

Theorem 4.5.7. $\mathsf{E}(Y\mid x)=a+bx\Rightarrow\begin{cases}a=\mathsf{E}(Y)-\mathsf{E}(X)\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\\ b=\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\end{cases}$

$$\begin{aligned}
&\Leftrightarrow\ \mathsf{E}(Y\mid x)=\mathsf{E}(Y)+\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)(x-\mathsf{E}(X))\\
&\Leftrightarrow\ \mathsf{E}(Y\mid x)-\mathsf{E}(Y)=\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)(x-\mathsf{E}(X))
\end{aligned}$$

Proof.

$$\begin{aligned}
a+bx=\mathsf{E}(Y\mid x) &= \begin{cases}\sum_{y\in Y(\Omega)}\left(y\cdot f(y\mid x)\right) & Y \text{ is dicrete}\\ \int_{y\in Y(\Omega)}\left(y\cdot f(y\mid x)\right)\,\mathrm{d}y & Y \text{ is continuous}\end{cases}\\
&= \begin{cases}\sum_{y\in Y(\Omega)}\left(y\cdot \frac{f(x,y)}{f_X(x)}\right) & Y \text{ is dicrete}\\ \int_{y\in Y(\Omega)}\left(y\cdot \frac{f(x,y)}{f_X(x)}\right)\,\mathrm{d}y & Y \text{ is continuous}\end{cases}\\
(a+bx)\,f_X(x) &= \begin{cases}\sum_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right) & Y \text{ is dicrete}\\ \int_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right)\,\mathrm{d}y & Y \text{ is continuous}\end{cases}
\end{aligned}$$

$$\begin{aligned}
\mathsf{E}(Y) &= \begin{cases}\begin{cases}\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right)\right)\mathrm{d}x & X \text{ is continuous}\end{cases} & Y \text{ is dicrete}\\ \begin{cases}\sum_{x\in X(\Omega)}\left(\int_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right)\mathrm{d}y\right) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}\int_{y\in Y(\Omega)}\left(y\cdot f(x,y)\right)\mathrm{d}y\,\mathrm{d}x & X \text{ is continuous}\end{cases} & Y \text{ is continuous}\end{cases}\\
&= \begin{cases}\sum_{x\in X(\Omega)}(a+bx)\,f_X(x) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}(a+bx)\,f_X(x)\,\mathrm{d}x & X \text{ is continuous}\end{cases}\\
&= a+b\cdot\mathsf{E}(X)
\end{aligned}$$

$$\begin{aligned}
\mathsf{E}(XY) &= \begin{cases}\begin{cases}\sum_{x\in X(\Omega)}\sum_{y\in Y(\Omega)}\left(xy\cdot f(x,y)\right) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}\left(\sum_{y\in Y(\Omega)}\left(xy\cdot f(x,y)\right)\right)\mathrm{d}x & X \text{ is continuous}\end{cases} & Y \text{ is dicrete}\\ \begin{cases}\sum_{x\in X(\Omega)}\left(\int_{y\in Y(\Omega)}\left(xy\cdot f(x,y)\right)\mathrm{d}y\right) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}\int_{y\in Y(\Omega)}\left(xy\cdot f(x,y)\right)\mathrm{d}y\,\mathrm{d}x & X \text{ is continuous}\end{cases} & Y \text{ is continuous}\end{cases}\\
&= \begin{cases}\sum_{x\in X(\Omega)}x\,(a+bx)\,f_X(x) & X \text{ is dicrete}\\ \int_{x\in X(\Omega)}x\,(a+bx)\,f_X(x)\,\mathrm{d}x & X \text{ is continuous}\end{cases}\\
&= a\cdot\mathsf{E}(X)+b\left(\mathsf{V}(X)+(\mathsf{E}(X))^2\right)
\end{aligned}$$

$$\begin{aligned}
&\begin{cases}a+\mathsf{E}(X)\,b=\mathsf{E}(Y)\\ \mathsf{E}(X)\,a+\left(\mathsf{V}(X)+(\mathsf{E}(X))^2\right)b=\mathsf{E}(XY)\end{cases}\\
\Rightarrow &\begin{cases}a=\frac{\mathsf{E}(Y)\left(\mathsf{V}(X)+(\mathsf{E}(X))^2\right)-\mathsf{E}(XY)\mathsf{E}(X)}{\left(\mathsf{V}(X)+(\mathsf{E}(X))^2\right)-(\mathsf{E}(X))^2}=\frac{\mathsf{E}(Y)\mathsf{V}(X)-\mathsf{E}(X)(\mathsf{E}(XY)-\mathsf{E}(X)\mathsf{E}(Y))}{\mathsf{V}(X)}\\ b=\frac{\mathsf{E}(XY)-\mathsf{E}(X)\mathsf{E}(Y)}{\mathsf{V}(X)}\end{cases}\\
\Rightarrow &\begin{cases}a=\mathsf{E}(Y)-\mathsf{E}(X)\frac{\mathsf{V}(X,Y)}{\mathsf{V}(X)}=\mathsf{E}(Y)-\mathsf{E}(X)\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\\ b=\frac{\mathsf{V}(X,Y)}{\mathsf{V}(X)}=\frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\end{cases}
\end{aligned}$$

□

Theorem 4.5.8. $\mathsf{E}(X\mid y)=c+dy\Rightarrow\begin{cases}c=\mathsf{E}(X)-\mathsf{E}(Y)\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)\\ d=\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)\end{cases}$

$$\Leftrightarrow\quad \mathsf{E}(X\mid y)=\mathsf{E}(X)+\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)\,(y-\mathsf{E}(Y))$$

$$\Leftrightarrow\quad \mathsf{E}(X\mid y)-\mathsf{E}(X)=\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)\,(y-\mathsf{E}(Y))$$

Corollary 4.5.1. $\begin{cases}\mathsf{E}(Y\mid x)=a+bx\\ \mathsf{E}(X\mid y)=c+dy\end{cases}\Rightarrow\begin{cases}(\mathsf{R}(X,Y))^2=bd\\ \mathsf{R}(X,Y)=\operatorname{sgn}(b)\sqrt{bd}=\operatorname{sgn}(d)\sqrt{bd}\end{cases}$

Proof.

$$
\begin{aligned}
bd &= \frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\cdot\frac{\sqrt{\mathsf{V}(X)}}{\sqrt{\mathsf{V}(Y)}}\mathsf{R}(X,Y)\\
&= (\mathsf{R}(X,Y))^2 \geq 0
\end{aligned}
$$

$$
\begin{aligned}
b\cdot\mathsf{R}(X,Y) &= \frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}\mathsf{R}(X,Y)\cdot\mathsf{R}(X,Y)\\
&= \frac{\sqrt{\mathsf{V}(Y)}}{\sqrt{\mathsf{V}(X)}}(\mathsf{R}(X,Y))^2 \geq 0
\end{aligned}
$$

for the same reason,

$$d\cdot\mathsf{R}(X,Y)\geq 0$$

$$\begin{cases} b\cdot\mathsf{R}(X,Y)\geq 0\\ d\cdot\mathsf{R}(X,Y)\geq 0\end{cases} \Rightarrow \mathsf{sgn}(\mathsf{R}(X,Y)) = \mathsf{sgn}(b) = \mathsf{sgn}(d)$$

□

- more discussion about regression into 10

4.6 二元隨機變數之函數

$$\begin{cases} x' = x'(x,y)\\ y' = y'(x,y)\end{cases} \Leftrightarrow \begin{cases} x = x(x',y')\\ y = y(x',y')\end{cases}$$

$$\mathsf{P}((X,Y)\in S) = \mathsf{P}((X',Y')\in S')$$

$$\begin{cases} X,Y \text{ are continuous}\\ X',Y' \text{ are continuous}\end{cases}$$

$$J = \det\left[\frac{\partial(x,y)}{\partial(x',y')}\right] = \det\begin{bmatrix} \frac{\partial x}{\partial x'} & \frac{\partial y}{\partial x'}\\ \frac{\partial x}{\partial y'} & \frac{\partial y}{\partial y'}\end{bmatrix} = \det\begin{bmatrix} \frac{\partial x}{\partial x'} & \frac{\partial x}{\partial y'}\\ \frac{\partial y}{\partial x'} & \frac{\partial y}{\partial y'}\end{bmatrix}$$

$$
\begin{aligned}
\mathsf{P}((X,Y)\in S) &= \iint_S f_{X,Y}(x,y)\,\mathsf{d}x\,\mathsf{d}y\\
&\overset{\text{go study calculus}}{=} \iint_{S'} f_{X,Y}(x(x',y'),y(x',y'))\,J\,\mathsf{d}x'\,\mathsf{d}y'\\
&= \iint_{S'} f_{X',Y'}(x',y')\,\mathsf{d}x'\,\mathsf{d}y' = \mathsf{P}((X',Y')\in S')
\end{aligned}
$$

$$
\begin{cases} f_{X',Y'}(x',y') = \sum_{\substack{x'=x'(x,y)\\ y'=y'(x,y)}} f_{X,Y}(x(x',y'),y(x',y'))\,J\\ f_{X',Y'}(x',y')\geq 0\end{cases}
$$

$$\Rightarrow\quad f_{X',Y'}(x',y') = \sum_{\substack{x'=x'(x,y)\\ y'=y'(x,y)}} f_{X,Y}(x(x',y'),y(x',y'))\,|J|$$

Theorem 4.6.1. transformation from two to one

$$\begin{cases} x = x(x_1,x_2)\\ x_2 = x_2\end{cases} \Leftrightarrow \begin{cases} x_1 = x_1(x,x_2)\\ x_2 = x_2\end{cases}$$

$$\begin{cases} X_1,X_2 \text{ are continuous}\\ X,X_2 \text{ are continuous}\end{cases}$$

$$
\begin{aligned}
f_{X,X_2}(x,x_2) &= \sum_{x=x(x_1,x_2)} f_{X_1,X_2}(x_1(x,x_2),x_2)\left|\text{det}\begin{bmatrix} \frac{\partial x_1}{\partial x} & \frac{\partial x_2}{\partial x} \\ \frac{\partial x_1}{\partial x_2} & \frac{\partial x_2}{\partial x_2} \end{bmatrix}\right| \\
&= \sum_{x=x(x_1,x_2)} f_{X_1,X_2}(x_1(x,x_2),x_2)\left|\text{det}\begin{bmatrix} \frac{\partial x_1}{\partial x} & 0 \\ \frac{\partial x_1}{\partial x_2} & 1 \end{bmatrix}\right| \\
&= \sum_{x=x(x_1,x_2)} f_{X_1,X_2}(x_1(x,x_2),x_2)\left|\frac{\partial x_1}{\partial x}\right|
\end{aligned}
$$

$$
\begin{aligned}
f_X(x) &= \int\limits_{x_2\in X_2(\Omega)} f_{X,X_2}(x,x_2)\,\mathrm{d}x_2 \\
&= \int\limits_{x_2\in X_2(\Omega)} \sum_{x=x(x_1,x_2)} f_{X_1,X_2}(x_1(x,x_2),x_2)\left|\frac{\partial x_1}{\partial x}\right|\,\mathrm{d}x_2
\end{aligned}
$$

Theorem 4.6.2. transformation by the concept of probability summation

X, X_2 are continuous

$$
\begin{aligned}
f_X(x)\,\mathrm{d}x &= \mathrm{dP}(X=x) \\
&= \mathrm{dP}\left(\bigvee_{x=x\left(\mathop{;}_{i=1}^{n} x_i\right)}\left(\bigwedge_{i=1}^{n} X_i=x_i\right)\right) \\
&= \sum_{x=x\left(\mathop{;}_{i=1}^{n} x_i\right)} \mathrm{dP}\left(\bigwedge_{i=1}^{n} X_i=x_i\right) \\
&= \sum_{x=x\left(\mathop{;}_{i=1}^{n} x_i\right)} f_{\mathop{;}_{i=1}^{n} X_i}\left(\mathop{;}_{i=1}^{n} x_i\right)\prod_{i=1}^{n}\mathrm{d}x_i
\end{aligned}
$$

Chapter 5

重要機率分配

5.1 超幾何分配

Definition 5.1.1. hypergeometric distribution

- case for two

$$f_X(x) = \frac{\binom{N_2}{n-x}\binom{N_1}{x}}{\binom{N_1+N_2}{n}} = \frac{\binom{N-N_1}{n-x}\binom{N_1}{x}}{\binom{N}{n}}$$

- case for three

$$f_{X_1,X_2}(x_1,x_2) = \frac{\binom{N_3}{n-(x_1+x_2)}\binom{N_2}{x_2}\binom{N_1}{x_1}}{\binom{N_1+N_2+N_3}{n}} = \frac{\binom{N-(N_1+N_2)}{n-(x_1+x_2)}\binom{N_2}{x_2}\binom{N_1}{x_1}}{\binom{N}{n}}$$

- case for $m \in \mathbb{N}$

$$f_{\overset{m}{\underset{i=1}{,}} X_i}\left(\overset{m}{\underset{i=1}{,}} x_i\right) = \frac{\binom{N_m}{n-\sum_{i=1}^{m-1} x_i}\prod_{i=1}^{m}\binom{N_i}{x_i}}{\binom{\sum_{i=1}^{m} N_i}{n}} = \frac{\binom{N-\sum_{i=1}^{m-1} N_i}{n-\sum_{i=1}^{m-1} x_i}\prod_{i=1}^{m}\binom{N_i}{x_i}}{\binom{N}{n}}$$

- draw **without** replacement **vs.** draw **with** replacement

5.2 Bernoulli試行及其有關之機率分配

5.3 卜瓦松分配、指數分配與gamma分配

5.4 常態分配

Definition 5.4.1. normal distribution

$$X \sim \mathsf{n}(\mu,\sigma) \Leftrightarrow f_X(x) = \frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Fact 5.4.1. $X \sim \mathsf{n}(\mu,\sigma) \Rightarrow \begin{cases} \mathsf{M}_X(\xi) = \mathsf{e}^{\left(\mu\xi + \frac{\sigma^2}{2}\xi^2\right)} \\ \mathsf{C}_X(\xi) = \mu\xi + \frac{\sigma^2}{2}\xi^2 \\ \mathsf{E}(X) = \mu \\ \mathsf{V}(X) = \sigma^2 \end{cases}$

$$\begin{aligned} \mathsf{M}_X(\xi) &= \mathsf{E}\left(\mathsf{e}^{\xi X}\right) \\ &= \int_{-\infty}^{\infty} \mathsf{e}^{\xi x} f_X(x)\, \mathsf{d}x \\ &= \int_{-\infty}^{\infty} \mathsf{e}^{\xi x} \cdot \frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\, \mathsf{d}x \\ &= \int_{-\infty}^{\infty} \frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + \xi x}}{\sigma\sqrt{2\pi}}\, \mathsf{d}x \end{aligned}$$

$$\begin{aligned}
&\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+\xi x\\
=\;&\frac{-1}{2\sigma^2}\left(x^2-2\mu x+\mu^2-2\sigma^2\xi x\right)\\
=\;&\frac{-1}{2\sigma^2}\left(x^2-2\left(\mu+\sigma^2\xi\right)x+\mu^2\right)
\end{aligned}$$

$$\begin{aligned}
&x^2-2\left(\mu+\sigma^2\xi\right)x+\mu^2\\
=\;&x^2-2\left(\mu+\sigma^2\xi\right)x+\left(\mu+\sigma^2\xi\right)^2-\left(\mu+\sigma^2\xi\right)^2+\mu^2\\
=\;&\left(x-\left(\mu+\sigma^2\xi\right)\right)^2+\mu^2-\left(\mu+\sigma^2\xi\right)^2\\
=\;&\left(x-\left(\mu+\sigma^2\xi\right)\right)^2+\left(\mu+\left(\mu+\sigma^2\xi\right)\right)\left(\mu-\left(\mu+\sigma^2\xi\right)\right)\\
=\;&\left(x-\left(\mu+\sigma^2\xi\right)\right)^2+\left(2\mu+\sigma^2\xi\right)\left(-\sigma^2\xi\right)
\end{aligned}$$

$$\begin{aligned}
&\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+\xi x\\
=\;&\frac{-1}{2\sigma^2}\left(x^2-2\left(\mu+\sigma^2\xi\right)x+\mu^2\right)\\
=\;&\frac{-1}{2\sigma^2}\left(\left(x-\left(\mu+\sigma^2\xi\right)\right)^2+\left(2\mu+\sigma^2\xi\right)\left(-\sigma^2\xi\right)\right)\\
=\;&\frac{-\left(x-\left(\mu+\sigma^2\xi\right)\right)^2}{2\sigma^2}+\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)
\end{aligned}$$

$$\begin{aligned}
\mathsf{M}_X(\xi)\;&=\;\int_{-\infty}^{\infty}\frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+\xi x}}{\sigma\sqrt{2\pi}}\,\mathsf{d}x\\
&=\;\int_{-\infty}^{\infty}\frac{\mathsf{e}^{\frac{-\left(x-\left(\mu+\sigma^2\xi\right)\right)^2}{2\sigma^2}+\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}}{\sigma\sqrt{2\pi}}\,\mathsf{d}x\\
&=\;\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}\int_{-\infty}^{\infty}\frac{\mathsf{e}^{\frac{-\left(x-\left(\mu+\sigma^2\xi\right)\right)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}\,\mathsf{d}x\\
&\overset{x'=x-\sigma^2\xi}{=}\;\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}\int_{-\infty}^{\infty}\frac{\mathsf{e}^{\frac{-\left(x'-\mu\right)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}\,\mathsf{d}\left(x'+\sigma^2\xi\right)\\
&=\;\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}\int_{-\infty}^{\infty}\frac{\mathsf{e}^{\frac{-\left(x'-\mu\right)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}\,\mathsf{d}x'\\
&=\;\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}\int_{-\infty}^{\infty}f_{X'}\left(x'\right)\mathsf{d}x'\wedge X'\sim\mathsf{n}\left(\mu,\sigma\right)\\
&=\;\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}\cdot 1=\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}
\end{aligned}$$

$$\mathsf{C}_X(\xi)=\ln\mathsf{M}_X(\xi)=\ln\mathsf{e}^{\left(\mu\xi+\frac{\sigma^2}{2}\xi^2\right)}=\mu\xi+\frac{\sigma^2}{2}\xi^2$$

$$\mathsf{E}(X)=\dot{\mathsf{C}}_X(\xi)\mid_{\xi=0}=\left[\mu+\sigma^2\xi\right]_{\xi=0}=\mu$$

$$\mathsf{V}(X)=\ddot{\mathsf{C}}_X(\xi)\mid_{\xi=0}=\left[\sigma^2\right]_{\xi=0}=\sigma^2$$

Definition 5.4.2. standardized normal distribution = standarized normal distribution = standard normal distribution

$$Z\sim\mathsf{n}(0,1)=\mathsf{n}(\mu,\sigma)\mid_{\begin{cases}\mu=0\\\sigma=1\end{cases}}\Leftrightarrow f_Z(z)=\frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\mid_{\begin{cases}\mu=0\\\sigma=1\end{cases}}=\frac{\mathsf{e}^{\frac{-z^2}{2}}}{\sqrt{2\pi}}$$

Theorem 5.4.1. standardization = standarization

$$\begin{cases} X \sim \mathsf{n}(\mu, \sigma) \\ Z = \frac{X-\mu}{\sigma} \end{cases} \Rightarrow Z \sim \mathsf{n}(0,1)$$

Proof.

$$\begin{aligned} f_Z(z) &= f_X(x(z)) \left| \frac{\mathsf{d}x}{\mathsf{d}z} \right| \\ &= \frac{\mathsf{e}^{\frac{-z^2}{2}}}{\sigma\sqrt{2\pi}} \cdot \left| \frac{\mathsf{d}}{\mathsf{d}z}(\sigma z + \mu) \right| \\ &= \frac{\mathsf{e}^{\frac{-z^2}{2}}}{\sigma\sqrt{2\pi}} \cdot |\sigma| = \frac{\mathsf{e}^{\frac{-z^2}{2}}}{\sigma\sqrt{2\pi}} \cdot \sigma \\ &= \frac{\mathsf{e}^{\frac{-z^2}{2}}}{\sqrt{2\pi}} \end{aligned}$$

□

Fact 5.4.2. $Z \sim \mathsf{n}(0,1) \Rightarrow \begin{cases} \mathsf{M}_Z(\xi) = \mathsf{e}^{\left(0\cdot\xi + \frac{1}{2}\xi^2\right)} = \mathsf{e}^{\frac{\xi^2}{2}} \\ \mathsf{C}_Z(\xi) = 0 \cdot \xi + \frac{1}{2}\xi^2 = \frac{\xi^2}{2} \\ \mathsf{E}(Z) = 0 \\ \mathsf{V}(Z) = 1 \end{cases}$

Definition 5.4.3. lognormal distribution

$$X \sim \mathsf{n}(\mu, \sigma) \Rightarrow \mathsf{e}^X \sim \mathsf{logn}(\mu, \sigma)$$

5.5 一致分配

5.6 二元常態分配

Chapter 6

抽樣分配

6.1 樣本分配與抽樣分配

Definition 6.1.1. sampl? distribution

- sample distribution = SmplD = $f_{\mathop{,}\limits_{i=1}^{n} X_i}\left(\mathop{,}\limits_{i=1}^{n} x_i\right)$

$$\begin{cases} \mathop{,}\limits_{i=1}^{n} X_i \text{ are independent} \\ X_i \sim \mathrm{d}(\mu,\sigma) = \text{arbitrary distribution with } \mu,\sigma \end{cases}$$

$$\forall i \in \mathbb{N} \cap [1,n] \left(f(x_i) = f_{X_i}(x_i) = \mathrm{d}(x_i;\mu,\sigma)\right)$$

$$f\left(\mathop{,}\limits_{i=1}^{n} x_i\right) = f_{\mathop{,}\limits_{k=1}^{n} X_i}\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \overset{\text{independence}}{=} \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} f(x_i)$$

  - if no independence, $\neg\forall i \in \mathbb{N} \cap [1,n] \left(f(x_i) = f_{X_i}(x_i) = \mathrm{d}(x_i;\mu,\sigma)\right)$

- sampling distribution = SmplingD = $f_G\left(g\left(\mathop{,}\limits_{i=1}^{n} x_i\right)\right)$

$$g\left(\mathop{,}\limits_{i=1}^{n} x_i\right)$$

$$f(g) = f_G(g) = f_n(g) = f_{G,n}(g) = f_G\left(g\left(\mathop{,}\limits_{i=1}^{n} x_i\right)\right) = f\left(g\left(\mathop{,}\limits_{i=1}^{n} x_i\right)\right)$$

  - SmplingD can be derived from SmplD by 4.6.1

Definition 6.1.2. random sample; independent and identically distributed, IID

$$\begin{aligned} & \begin{cases} \mathop{,}\limits_{i=1}^{n} X_i \text{ are independent} \\ \exists!\mathrm{d} = \text{arbitrary distribution} \forall i \in \mathbb{N} \cap [1,n] (X_i \sim \mathrm{d}) \end{cases} \\ \Leftrightarrow \quad & \begin{cases} \mathop{,}\limits_{i=1}^{n} X_i \text{ is a random sample of size } n \\ \mathop{,}\limits_{i=1}^{n} X_i \text{ is a random sample} \sim \mathrm{d} \end{cases} \\ \Leftrightarrow \quad & \mathop{,}\limits_{i=1}^{n} X_i \overset{\text{IID}}{\sim} \mathrm{d} \end{aligned}$$

6.1.1 樣本統計量之抽樣分配

Definition 6.1.3. parameter

$$\begin{cases} g_{\pi_k} : X(\Omega) \to \mathbb{R} \\ \bigcirc : \mathbb{R}^\infty \to \mathbb{R} \\ \pi_k = \mathop{\bigcirc}\limits_{x \in X(\Omega)} g_{\pi_k} \end{cases} \Leftrightarrow \mathrm{d}\left(\mathop{,}\limits_{k=1}^{m} \pi_k\right) \Rightarrow \pi_k \text{ is a parameter of the distribution d}$$

$$\begin{cases} g_{\theta_k} : X(\Omega) \to \mathbb{R} \\ \bigcirc : \mathbb{R}^\infty \to \mathbb{R} \\ \theta_k = \mathop{\bigcirc}\limits_{x \in X(\Omega)} g_{\theta_k} \end{cases} \Leftrightarrow \mathsf{d}\left(\mathop{,}_{k=1}^{m} \theta_k\right) \Rightarrow \theta_k \text{ is a parameter of the distribution d}$$

- e.g.

$$\mu, \sigma \text{ are parameters of n}(\mu, \sigma)$$

- parameter classification
  - location parameter
  - dispersion parameter = scale parameter
  - shape parameter
- a parameter is about the population

Definition 6.1.4. statistic

$$g : (X(\Omega) / \{\text{unknown parameters}\}) \to \mathbb{R} \Leftrightarrow g \text{ is a statistic}$$

- e.g.

$$g = g\left(\mathop{,}_{i=1}^{n} X_i\right) \Rightarrow g \text{ is a statistic}$$

$$\begin{cases} g = g\left(\mathop{,}_{i=1}^{n} X_i, \mathop{,}_{k=1}^{m} \pi_k\right) \\ \mathop{,}_{k=1}^{m} \pi_k \text{ are } \textbf{unknown} \end{cases} \Rightarrow g \text{ is } \textbf{not} \text{ a statistic}$$

$$\begin{cases} g = g\left(\mathop{,}_{i=1}^{n} X_i, \mathop{,}_{k=1}^{m} \pi_k\right) \\ \mathop{,}_{k=1}^{m} \pi_k \text{ are } \textbf{known} \end{cases} \Rightarrow g \text{ is a statistic}$$

- a statistic is about the sample

Fact 6.1.1. $X_i \sim \mathsf{d}(\mu, \sigma) = \text{arbitrary distribution with } \mu, \sigma \Rightarrow \begin{cases} \mathsf{E}(X_i) = \mu \\ \mathsf{V}(X_i) = \sigma^2 \end{cases}$

Theorem 6.1.1. expectation and variance of sample mean

- simple random sampling with replacement = SRSWR = SRSw/R

$$\begin{aligned} &\begin{cases} \mathop{,}_{i=1}^{n} X_i \overset{\text{IID}}{\sim} \mathsf{d}(\mu, \sigma) = \text{arbitrary distribution with } \mu, \sigma \\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \end{cases} \\ \Rightarrow &\begin{cases} \mathsf{E}(\bar{X}) = \mu = \mathsf{E}(X_i) \\ \mathsf{V}(\bar{X}) = \frac{\sigma^2}{n} = \frac{\mathsf{V}(X_i)}{n} \end{cases} \end{aligned}$$

Proof.

$$\begin{aligned} \mathsf{E}(\bar{X}) &= \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\mathsf{E}\left(\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathsf{E}(X_i) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n\mu = \mu \end{aligned}$$

$$
\begin{aligned}
\mathsf{V}\left(\bar{X}\right) &= \mathsf{V}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) \\
&= \frac{1}{n^2}\mathsf{V}\left(\sum_{i=1}^{n}X_i\right) \\
&= \frac{1}{n^2}\left(\mathsf{E}\left(\left(\sum_{i=1}^{n}X_i\right)^2\right)-\left(\mathsf{E}\left(\sum_{i=1}^{n}X_i\right)\right)^2\right) \\
&= \frac{1}{n^2}\left(\mathsf{E}\left(\sum_{i=1}^{n}\left(X_i\right)^2+2\sum_{1\le i_1<i_2\le n}X_{i_1}X_{i_2}\right)-\left(\sum_{i=1}^{n}\mathsf{E}\left(X_i\right)\right)^2\right) \\
&= \frac{1}{n^2}\Bigg(\mathsf{E}\left(\sum_{i=1}^{n}\left(X_i\right)^2\right)+\mathsf{E}\left(2\sum_{1\le i_1<i_2\le n}X_{i_1}X_{i_2}\right) \\
&\quad -\left(\sum_{i=1}^{n}\left(\mathsf{E}\left(X_i\right)\right)^2+2\sum_{1\le i_1<i_2\le n}\mathsf{E}\left(X_{i_1}\right)\mathsf{E}\left(X_{i_2}\right)\right)\Bigg) \\
&= \frac{1}{n^2}\Bigg(\sum_{i=1}^{n}\mathsf{E}\left(\left(X_i\right)^2\right)+2\mathsf{E}\left(\sum_{1\le i_1<i_2\le n}X_{i_1}X_{i_2}\right) \\
&\quad -\sum_{i=1}^{n}\left(\mathsf{E}\left(X_i\right)\right)^2-2\sum_{1\le i_1<i_2\le n}\mathsf{E}\left(X_{i_1}\right)\mathsf{E}\left(X_{i_2}\right)\Bigg) \\
&= \frac{1}{n^2}\Bigg(\sum_{i=1}^{n}\left(\mathsf{E}\left(\left(X_i\right)^2\right)-\left(\mathsf{E}\left(X_i\right)\right)^2\right) \\
&\quad +2\sum_{1\le i_1<i_2\le n}\left(\mathsf{E}\left(X_{i_1}X_{i_2}\right)-\mathsf{E}\left(X_{i_1}\right)\mathsf{E}\left(X_{i_2}\right)\right)\Bigg) \\
&= \frac{1}{n^2}\left(\left(\sum_{i=1}^{n}\mathsf{V}\left(X_i\right)\right)+2\left(\sum_{1\le i_1<i_2\le n}\mathsf{V}\left(X_{i_1},X_{i_2}\right)\right)\right) \\
&= \frac{1}{n^2}\left(\left(\sum_{i=1}^{n}\sigma^2\right)+2\left(\sum_{1\le i_1<i_2\le n}\mathsf{V}\left(X_{i_1},X_{i_2}\right)\right)\right) \\
&= \frac{1}{n^2}\left(\left(n\sigma^2\right)+2\left(\sum_{1\le i_1<i_2\le n}\mathsf{V}\left(X_{i_1},X_{i_2}\right)\right)\right) \\
&= \frac{\sigma^2}{n}+\frac{2}{n^2}\left(\sum_{1\le i_1<i_2\le n}\mathsf{V}\left(X_{i_1},X_{i_2}\right)\right)
\end{aligned}
$$

$\overset{n}{\underset{i=1}{;}} X_i$ are independent $\Rightarrow \mathsf{V}\left(X_{i_1},X_{i_2}\right)=\begin{cases}\mathsf{V}\left(X_{i_1}\right)=\mathsf{V}\left(X_{i_2}\right) & i_1=i_2\\ 0 & i_1\neq i_2\end{cases}$

$$
\begin{aligned}
\mathsf{V}\left(\bar{X}\right) &= \frac{\sigma^2}{n}+\frac{2}{n^2}\left(\sum_{1\le i_1<i_2\le n}\mathsf{V}\left(X_{i_1},X_{i_2}\right)\right) \\
&= \frac{\sigma^2}{n}+\frac{2}{n^2}\left(\binom{n}{2}\cdot 0\right)=\frac{\sigma^2}{n}
\end{aligned}
$$

□

- simple random sampling without replacement = SRSWOR = SRSw/oR

$$
\begin{aligned}
& C_i \text{ is the selection indicator} \\
\Leftrightarrow & \begin{cases} C_i \text{ is a RV} \\ C_i \in \{0,1\} & \text{indicator property} \\ \sum_{i=1}^{N} C_i = n & \text{selection or sampling property} \end{cases} \\
\Rightarrow & \begin{cases} \mathsf{E}(C_i) = \frac{1}{N}\sum_{i=1}^{N} C_i = \frac{n}{N} \\ \mathsf{V}(C_i) = \frac{1}{N}\sum_{i=1}^{N}(C_i - \mathsf{E}(C_i))^2 \\ = \frac{1}{N}\left((N-n)\left(\frac{n}{N}\right)^2 + n\left(1-\frac{n}{N}\right)^2\right) \\ = \frac{n(N-n)}{N^2}\left(\frac{n}{N}+\frac{N-n}{N}\right) = \frac{n}{N}\frac{N-n}{N} = \frac{n\left(1-\frac{n}{N}\right)}{N} \end{cases}
\end{aligned}
$$

$$
\begin{cases} \{x_i\}_{i=1}^{N} \text{ is the population} \wedge \begin{cases} \mu = \frac{1}{N}\sum_{i=1}^{N} x_i \\ \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2} \end{cases} \\ X_i = C_i x_i \wedge C_i \text{ is the selection indicator} \\ \bar{X} = \frac{\sum_{i=1}^{N} X_i}{\sum_{i=1}^{N} C_i} = \frac{1}{n}\sum_{i=1}^{N} X_i \end{cases}
$$

$$
\Rightarrow \begin{cases} \mathsf{E}(\bar{X}) = \mu = \mathsf{E}(X_i) \\ \mathsf{V}(\bar{X}) = \frac{N-n}{N-1}\frac{\sigma^2}{n} = \frac{N-n}{N-1}\frac{\mathsf{V}(X_i)}{n} = \frac{1-\frac{n}{N}}{1-\frac{1}{N}}\frac{\sigma^2}{n} \overset{N\to\infty}{\to} \frac{\sigma^2}{n} \end{cases}
$$

Proof.

$$
\begin{aligned}
\mathsf{E}(\bar{X}) &= \mathsf{E}\left(\frac{1}{n}\sum_{i=1}^{N} X_i\right) = \frac{1}{n}\mathsf{E}\left(\sum_{i=1}^{N} X_i\right) \\
&= \frac{1}{n}\sum_{i=1}^{N}\mathsf{E}(X_i) = \frac{1}{n}\sum_{i=1}^{N}\mathsf{E}(C_i x_i) \\
&= \frac{1}{n}\sum_{i=1}^{N} x_i \mathsf{E}(C_i) = \frac{1}{n}\sum_{i=1}^{N} x_i \cdot \frac{n}{N} \\
&= \frac{n}{n}\cdot\frac{\sum_{i=1}^{N} x_i}{N} = 1\cdot\mu = \mu
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}(\bar{X}) &= \mathsf{V}\left(\frac{1}{n}\sum_{i=1}^{N} X_i\right) = \frac{1}{n^2}\mathsf{V}\left(\sum_{i=1}^{N} X_i\right) \\
&\overset{4.5.4}{=} \frac{1}{n^2}\left(\sum_{i=1}^{N}\mathsf{V}(X_i) + 2\sum_{1\le i_1<i_2\le N}\mathsf{V}(X_{i_1}, X_{i_2})\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{N}\mathsf{V}(C_i x_i) + 2\sum_{1\le i_1<i_2\le N}\mathsf{V}(C_{i_1}x_{i_1}, C_{i_2}x_{i_2})\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{N} x_i^2\mathsf{V}(C_i) + 2\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\mathsf{V}(C_{i_1}, C_{i_2})\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}(C_{i_1}, C_{i_2}) &= \mathsf{E}(C_{i_1}C_{i_2}) - \mathsf{E}(C_{i_1})\mathsf{E}(C_{i_2}) \\
&= \left(1\cdot\frac{n}{N}\frac{n-1}{N-1} + 0\cdot\ldots\right) - \frac{n}{N}\cdot\frac{n}{N} \\
&= \frac{n}{N}\left(\frac{n-1}{N-1} - \frac{n}{N}\right) = \frac{n}{N}\frac{nN - N - nN + n}{N(N-1)} \\
&= \frac{n}{N}\frac{n-N}{N(N-1)} = \frac{-n}{N}\frac{N-n}{N(N-1)} = \frac{-n\left(1-\frac{n}{N}\right)}{N(N-1)}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}\left(\bar{X}\right) &= \frac{1}{n^2}\left(\sum_{i=1}^{N} x_i^2 \mathsf{V}\left(C_i\right) + 2\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\mathsf{V}\left(C_{i_1},C_{i_2}\right)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^{N} x_i^2 \frac{n\left(1-\frac{n}{N}\right)}{N} + 2\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\frac{-n\left(1-\frac{n}{N}\right)}{N\left(N-1\right)}\right) \\
&= \frac{1}{n^2}\frac{n\left(1-\frac{n}{N}\right)}{N}\left(\sum_{i=1}^{N} x_i^2 - \frac{2}{N-1}\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\right) \\
&= \frac{1-\frac{n}{N}}{n\cdot N}\left(\sum_{i=1}^{N} x_i^2 - \frac{2}{N-1}\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\right)
\end{aligned}
$$

$$
\begin{aligned}
& 2\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2} \\
=& \left(\sum_{i=1}^{N} x_i\right)^2 - \sum_{i=1}^{N} x_i^2 \\
=& \left(N\mu\right)^2 - \sum_{i=1}^{N} x_i^2
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}\left(\bar{X}\right) &= \frac{1-\frac{n}{N}}{n\cdot N}\left(\sum_{i=1}^{N} x_i^2 - \frac{2}{N-1}\sum_{1\le i_1<i_2\le N} x_{i_1}x_{i_2}\right) \\
&= \frac{1-\frac{n}{N}}{n\cdot N}\left(\frac{\left(N-1\right)\sum_{i=1}^{N} x_i^2}{N-1} - \frac{\left(N\mu\right)^2 - \sum_{i=1}^{N} x_i^2}{N-1}\right) \\
&= \frac{1-\frac{n}{N}}{n\cdot N}\left(\frac{N\sum_{i=1}^{N} x_i^2 - \left(N\mu\right)^2}{N-1}\right) = \frac{1-\frac{n}{N}}{n\cdot N}\cdot N\left(\frac{\sum_{i=1}^{N} x_i^2 - N\mu^2}{N-1}\right) \\
&= \frac{1-\frac{n}{N}}{n}\cdot\frac{\sum_{i=1}^{N}\left(x_i-\mu\right)^2}{N-1} = \frac{1-\frac{n}{N}}{n}\cdot\frac{N\sigma^2}{N-1} = \frac{N-n}{N-1}\frac{\sigma^2}{n} = \frac{1-\frac{n}{N}}{1-\frac{1}{N}}\frac{\sigma^2}{n} \overset{N\to\infty}{\to} \frac{\sigma^2}{n}
\end{aligned}
$$

$\square$

Theorem 6.1.2. $\begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ are independent} & [1] \\ X_i \sim \mathsf{d}\left(\mu,\sigma\right) = \text{arbitrary distribution with } \mu,\sigma \\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \text{sampling with replacement} \\ \mathsf{E}\left(S^2\right) = \sigma^2 \end{cases} \Rightarrow S^2 = \frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}$

Proof.

$$
\begin{aligned}
\mathsf{E}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right) &= \sum_{i=1}^{n}\mathsf{E}\left(\left(X_i-\bar{X}\right)^2\right)\\
&= \sum_{i=1}^{n}\mathsf{E}\left(\left(X_i-\bar{X}\right)^2\right)\\
&= \sum_{i=1}^{n}\mathsf{E}\left(\left((X_i-\mu)-\left(\bar{X}-\mu\right)\right)^2\right)\\
&= \sum_{i=1}^{n}\mathsf{E}\left((X_i-\mu)^2-2\left(X_i-\mu\right)\left(\bar{X}-\mu\right)+\left(\bar{X}-\mu\right)^2\right)\\
&= \sum_{i=1}^{n}\left(\mathsf{E}\left((X_i-\mu)^2\right)-2\cdot\mathsf{E}\left((X_i-\mu)\left(\bar{X}-\mu\right)\right)+\mathsf{E}\left(\left(\bar{X}-\mu\right)^2\right)\right)\\
&= \sum_{i=1}^{n}\left(\mathsf{V}\left(X_i\right)-2\cdot\mathsf{V}\left(X_i,\bar{X}\right)+\mathsf{V}\left(\bar{X}\right)\right)\\
&= \sum_{i=1}^{n}\left(\sigma^2-2\cdot\mathsf{V}\left(X_i,\bar{X}\right)+\frac{\sigma^2}{n}\right)=\sum_{i=1}^{n}\left(\frac{n+1}{n}\sigma^2-2\cdot\mathsf{V}\left(X_i,\bar{X}\right)\right)\\
&= \left(\sum_{i=1}^{n}\left(\frac{n+1}{n}\sigma^2\right)\right)-2\sum_{i=1}^{n}\mathsf{V}\left(X_i,\bar{X}\right)\\
&= n\left(\frac{n+1}{n}\sigma^2\right)-2\sum_{i=1}^{n}\mathsf{V}\left(X_i,\bar{X}\right)=(n+1)\,\sigma^2-2\sum_{i=1}^{n}\mathsf{V}\left(X_i,\bar{X}\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}\left(X_i,\bar{X}\right) &= \mathsf{V}\left(X_i,\frac{1}{n}\sum_{j=1}^{n}X_j\right)\\
&= \mathsf{V}\left(X_i,\sum_{j=1}^{n}\frac{X_j}{n}\right)\\
&\overset{4.5.2}{=} \sum_{j=1}^{n}\frac{\mathsf{V}\left(X_i,X_j\right)}{n}\\
&= \frac{1}{n}\sum_{j=1}^{n}\mathsf{V}\left(X_i,X_j\right)\\
&= \frac{1}{n}\left(\mathsf{V}\left(X_i,X_i\right)+\sum_{j\neq i}\mathsf{V}\left(X_i,X_j\right)\right)\\
&\overset{[1]}{=} \frac{1}{n}\left(\mathsf{V}\left(X_i\right)+(n-1)\cdot 0\right)\\
&= \frac{1}{n}\left(\sigma^2+0\right)=\frac{\sigma^2}{n}
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{E}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2\right) &= (n+1)\,\sigma^2-2\sum_{i=1}^{n}\mathsf{V}\left(X_i,\bar{X}\right)\\
&= (n+1)\,\sigma^2-2\sum_{i=1}^{n}\frac{\sigma^2}{n}\\
&= (n+1)\,\sigma^2-2\cdot n\cdot\frac{\sigma^2}{n}\\
&= (n+1)\,\sigma^2-2\sigma^2=(n-1)\,\sigma^2\\
\mathsf{E}\left(\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}\right) &= \sigma^2
\end{aligned}
$$

$$\mathsf{E}\left(S^2\right)=\sigma^2 \Rightarrow S^2=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}$$

□

Theorem 6.1.3. Cochran theorem; Basu theorem

Proof. □

6.1.2 $\bar{X}$之抽樣分配

Theorem 6.1.4. two population

$$\begin{cases} \overset{n_1}{\underset{j=1}{,}} X_{1j}, \overset{n_2}{\underset{j=1}{,}} X_{2j} \text{ are independent} & [1] \\ \begin{cases} X_{1j}\sim \mathsf{d}\left(\mu_1,\sigma_1\right)=\text{arbitrary distribution with } \mu_1,\sigma_1 \\ \bar{X}_1=\frac{1}{n_1}\sum_{j=1}^{n_1}X_{1j} \end{cases} \\ \begin{cases} X_{2j}\sim \mathsf{d}\left(\mu_2,\sigma_2\right)=\text{arbitrary distribution with } \mu_2,\sigma_2 \\ \bar{X}_2=\frac{1}{n_2}\sum_{j=1}^{n_2}X_{2j} \end{cases} \\ \text{sampling with replacement} \end{cases}$$

$$\Rightarrow \begin{cases} \mathsf{E}\left(\bar{X}_1\pm\bar{X}_2\right)=\mu_1\pm\mu_2 \\ \mathsf{V}\left(\bar{X}_1\pm\bar{X}_2\right)=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2} \end{cases}$$

Proof.

$$\mathsf{E}\left(\bar{X}_1\pm\bar{X}_2\right)\overset{6.1.1}{=}\mathsf{E}\left(\bar{X}_1\right)\pm\mathsf{E}\left(\bar{X}_2\right)=\mu_1\pm\mu_2$$

$$\begin{aligned}
\mathsf{V}\left(\bar{X}_1\pm\bar{X}_2\right) &\overset{4.5.3}{=} \mathsf{V}\left(\bar{X}_1\right)+\mathsf{V}\left(\bar{X}_2\right)\pm 2\cdot\mathsf{V}\left(\bar{X}_1,\bar{X}_2\right) \\
&\overset{6.1.1}{=} \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\pm 2\cdot\mathsf{V}\left(\frac{1}{n_1}\sum_{j_1=1}^{n_1}X_{1j_1},\frac{1}{n_2}\sum_{j_2=1}^{n_2}X_{2j_2}\right) \\
&= \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\pm 2\cdot\mathsf{V}\left(\sum_{j_1=1}^{n_1}\frac{X_{1j_1}}{n_1},\sum_{j_2=1}^{n_2}\frac{X_{2j_2}}{n_2}\right) \\
&\overset{4.5.2}{=} \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\pm 2\cdot\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\frac{\mathsf{V}\left(X_{1j_1},X_{2j_2}\right)}{n_1n_2} \\
&\overset{[1]\Rightarrow\mathsf{V}\left(X_{1j_1},X_{2j_2}\right)=0}{=} \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\pm 2\cdot\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2}\frac{0}{n_1n_2} \\
&= \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}
\end{aligned}$$

□

Theorem 6.1.5. two normally-distributed population

$$\begin{cases} \overset{n_1}{\underset{j=1}{,}} X_{1j}, \overset{n_2}{\underset{j=1}{,}} X_{2j} \text{ are independent} \\ \begin{cases} X_{1j}\sim \mathsf{n}\left(\mu_1,\sigma_1\right) \\ \bar{X}_1=\frac{1}{n_1}\sum_{j=1}^{n_1}X_{1j} \end{cases} \\ \begin{cases} X_{2j}\sim \mathsf{n}\left(\mu_2,\sigma_2\right) \\ \bar{X}_2=\frac{1}{n_2}\sum_{j=1}^{n_2}X_{2j} \end{cases} \\ \text{sampling with replacement} \end{cases} \Rightarrow \frac{\left(\bar{X}_1\pm\bar{X}_2\right)-\left(\mu_1\pm\mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim\mathsf{n}\left(0,1\right)$$

Proof.

$$6.1.1\Rightarrow\begin{cases} \bar{X}_1\sim\mathsf{n}\left(\mu_1,\frac{\sigma_1}{\sqrt{n_1}}\right) \\ \bar{X}_2\sim\mathsf{n}\left(\mu_2,\frac{\sigma_2}{\sqrt{n_2}}\right) \end{cases}$$

$$\bar{X}_1 \pm \bar{X}_2 \overset{6.1.4}{\sim} \mathsf{n}\left(\mu_1 \pm \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

$$\frac{\left(\bar{X}_1 \pm \bar{X}_2\right) - (\mu_1 \pm \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim \mathsf{n}(0,1)$$

□

6.2 順序統計量

6.3 中央極限定理

Theorem 6.3.1. central limit theorem (weak form) = CLT-W

$$\begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ are independent} & [1] \\ X_i \sim \mathsf{d}(\mu,\sigma) = \text{arbitrary distribution with } \mu,\sigma & [2] \\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \text{sampling with replacement} \\ \mathsf{M}_{X_i}(\xi_i) \in (-\infty,\infty) & \text{weak point} \\ Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)-\mu}{\sigma/\sqrt{n}} = \frac{\left(\sum_{i=1}^{n} X_i\right)-n\mu}{\sigma\sqrt{n}} \end{cases} \Rightarrow \left(\lim_{n\to\infty} Z\right) \sim \mathsf{n}(0,1)$$

Proof.

$$\mathsf{M}_{X_i}(\xi_i) \in (-\infty,\infty) \Rightarrow \mathsf{M}_{X_i-\mu}(\xi_i) \in (-\infty,\infty)$$

$$\begin{aligned} \mathsf{M}_{X_i-\mu}(\xi) &\overset{\text{Maclaurin thm}}{=} \mathsf{M}_{X_i-\mu}(0) + \dot{\mathsf{M}}_{X_i-\mu}(0)\,\xi \\ &\quad + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)}{2}\xi^2 \wedge \epsilon \in [0,\xi] \\ &= \mathsf{E}\left(\mathsf{e}^{0\cdot(X_i-\mu)}\right) + \mathsf{E}\left((X_i-\mu)\,\mathsf{e}^{0\cdot(X_i-\mu)}\right)\xi \\ &\quad + \frac{\mathsf{E}\left((X_i-\mu)^2\,\mathsf{e}^{\epsilon\cdot(X_i-\mu)}\right)}{2}\xi^2 \wedge \epsilon \in [0,\xi] \\ &= \mathsf{E}(1) + \mathsf{E}((X-\mu))\,\xi \\ &\quad + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)}{2}\xi^2 \wedge \epsilon \in [0,\xi] \\ &= 1 + 0\cdot\xi + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)}{2}\xi^2 \wedge \epsilon \in [0,\xi] \\ &= 1 + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)}{2}\xi^2 \wedge \epsilon \in [0,\xi] \\ &= 1 + \frac{\sigma^2}{2}\xi^2 + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2}\xi^2 \wedge \epsilon \in [0,\xi] \end{aligned} \tag{6.3.1}$$

$$
\begin{aligned}
\mathsf{M}_Z(\xi) &= \mathsf{E}\left(\mathsf{e}^{\xi Z}\right) \\
&= \mathsf{E}\left(\mathsf{e}^{\xi \frac{\left(\sum_{i=1}^n X_i\right)-n\mu}{\sigma\sqrt{n}}}\right) \\
&= \mathsf{E}\left(\mathsf{e}^{\xi \frac{\sum_{i=1}^n (X_i-\mu)}{\sigma\sqrt{n}}}\right) \\
&= \mathsf{E}\left(\prod_{i=1}^n \mathsf{e}^{\xi \frac{X_i-\mu}{\sigma\sqrt{n}}}\right) \\
&\overset{[1]}{=} \prod_{i=1}^n \mathsf{E}\left(\mathsf{e}^{\xi \frac{X_i-\mu}{\sigma\sqrt{n}}}\right) = \prod_{i=1}^n \mathsf{E}\left(\mathsf{e}^{\frac{\xi}{\sigma\sqrt{n}}(X_i-\mu)}\right) \\
&= \prod_{i=1}^n \mathsf{M}_{X_i-\mu}\left(\frac{\xi}{\sigma\sqrt{n}}\right) \\
&\overset{[2]}{=} \left(\mathsf{M}_{X_i-\mu}\left(\frac{\xi}{\sigma\sqrt{n}}\right)\right)^n \\
&= \left(1+\frac{\sigma^2}{2}\left(\frac{\xi}{\sigma\sqrt{n}}\right)^2 + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2}\left(\frac{\xi}{\sigma\sqrt{n}}\right)^2\right)^n \wedge \epsilon \in \left[0, \frac{\xi}{\sigma\sqrt{n}}\right] \\
&= \left(1+\frac{\xi^2}{2n} + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2n\sigma^2}\xi^2\right)^n \wedge \epsilon \in \left[0, \frac{\xi}{\sigma\sqrt{n}}\right]
\end{aligned}
$$

$$
\begin{aligned}
\lim_{n\to\infty} \mathsf{M}_Z(\xi) &= \lim_{n\to\infty}\left(\left(1+\frac{\xi^2}{2n} + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2n\sigma^2}\xi^2\right)^n \wedge \epsilon \in \left[0, \frac{\xi}{\sigma\sqrt{n}}\right]\right) \\
&= \lim_{n\to\infty}\left(1+\frac{\xi^2}{2n} + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2n\sigma^2}\xi^2\right)^n \wedge \lim_{n\to\infty}\epsilon = 0
\end{aligned}
$$

$$
\begin{aligned}
\lim_{n\to\infty}\epsilon = 0 \quad \Rightarrow \quad & \lim_{n\to\infty}\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon) \\
& = \ddot{\mathsf{M}}_{X_i-\mu}(0) = \mathsf{E}\left((X_i-\mu)^2\,\mathsf{e}^{0\cdot(X_i-\mu)}\right) \\
& = \mathsf{E}\left((X_i-\mu)^2\right) = \mathsf{V}(X_i) = \sigma^2 \\
\Rightarrow \quad & \lim_{n\to\infty}\left(\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2\right) = 0
\end{aligned}
$$

$$
\begin{aligned}
\lim_{n\to\infty} \mathsf{M}_Z(\xi) &= \lim_{n\to\infty}\left(1+\frac{\xi^2}{2n} + \frac{\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2}{2n\sigma^2}\xi^2\right)^n \\
&\quad \wedge \lim_{n\to\infty}\left(\ddot{\mathsf{M}}_{X_i-\mu}(\epsilon)-\sigma^2\right) = 0 \\
&= \lim_{n\to\infty}\left(1+\frac{\xi^2}{2n} + \frac{0}{2n\sigma^2}\xi^2\right)^n \\
&= \lim_{n\to\infty}\left(1+\frac{\xi^2}{2n}\right)^n = \lim_{n\to\infty}\left(1+\frac{\frac{\xi^2}{2}}{n}\right)^n \\
&= \mathsf{e}^{\frac{\xi^2}{2}}
\end{aligned}
$$

$$
\lim_{n\to\infty} \mathsf{M}_Z(\xi) = \mathsf{e}^{\frac{\xi^2}{2}} \Rightarrow \left(\lim_{n\to\infty}\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) = \left(\lim_{n\to\infty} Z\right) \sim \mathsf{n}(0,1)
$$

□

6.4 基本抽樣分配

Definition 6.4.1. degree of freedom = d_f = df = DF = DoF

Definition 6.4.2. factorial

$$n! = \begin{cases} \prod_{i=1}^{n} i = \prod_{i=0}^{n-1}(n-i) = \prod_{i=1}^{n}(n-(i-1)) = \prod_{i=1}^{n}(n+1-i) & n \in \mathbb{N} \\ 1 & n = 0 \end{cases}$$

Definition 6.4.3. double factorial

$$\begin{aligned} n!! &= \prod_{i=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} (n-2i) \\ &= \begin{cases} \prod_{i=0}^{\left\lfloor \frac{(2m-1)-1}{2} \right\rfloor} ((2m-1)-2i) & n = 2m-1 \\ \prod_{i=0}^{\left\lfloor \frac{2m-1}{2} \right\rfloor} (2m-2i) & n = 2m \end{cases} \wedge m \in \mathbb{N} \\ &= \begin{cases} \prod_{i=0}^{m-1} (2(m-i)-1) & n = 2m-1 \\ \prod_{i=0}^{\left\lfloor m-\frac{1}{2} \right\rfloor = m-1} 2(m-i) = 2^m \prod_{i=0}^{m-1}(m-i) = m!2^m & n = 2m \end{cases} \wedge m \in \mathbb{N} \end{aligned}$$

$$m \in \mathbb{N} \Rightarrow (2m)! = (2m)!!\,(2m-1)!! \Rightarrow (2m-1)!! = \frac{(2m)!}{(2m)!!} \overset{(2m)!!=m!2^m}{=} \frac{(2m)!}{m!2^m}$$

Definition 6.4.4. gamma function

$$\Gamma(z) = \int_0^\infty s^{z-1} \mathrm{e}^{-s}\, \mathrm{d}s$$

- $\int_0^\infty s^{z-1}\mathrm{e}^{-a\cdot s}\,\mathrm{d}s \overset{a\neq 0}{=} \frac{\Gamma(z)}{a^z}$

$$\begin{aligned} \int_0^\infty s^{z-1}\mathrm{e}^{-a\cdot s}\,\mathrm{d}s &\overset{\begin{cases} s' = a \cdot s \\ a \neq 0 \end{cases}}{=} \int_0^\infty \left(\frac{s'}{a}\right)^{z-1} \mathrm{e}^{-s'}\,\mathrm{d}\frac{s'}{a} \\ &= \frac{\int_0^\infty (s')^{z-1}\mathrm{e}^{-s'}\,\mathrm{d}s'}{a^z} \\ &= \frac{\Gamma(z)}{a^z} \end{aligned} \tag{6.4.1}$$

- $\Gamma(z) = 2\int_0^\infty s^{2z-1}\mathrm{e}^{-s^2}\,\mathrm{d}s$

$$\begin{aligned} \Gamma(z) &= \int_0^\infty s^{z-1}\mathrm{e}^{-s}\,\mathrm{d}s \\ &\overset{s'=\sqrt{s}}{=} \int_0^\infty \left((s')^2\right)^{z-1} \mathrm{e}^{-(s')^2}\,\mathrm{d}(s')^2 \\ &= \int_0^\infty (s')^{2(z-1)}\mathrm{e}^{-(s')^2} \cdot 2s'\,\mathrm{d}s' \\ &= 2\int_0^\infty (s')^{2z-1}\mathrm{e}^{-(s')^2}\,\mathrm{d}s' \\ &= 2\int_0^\infty s^{2z-1}\mathrm{e}^{-s^2}\,\mathrm{d}s \end{aligned}$$

- $\forall z \in \mathbb{N}\,(\Gamma(z+1) = z \cdot \Gamma(z))$

$$\begin{aligned} \Gamma(z+1) &= \int_0^\infty s^{(z+1)-1}\mathrm{e}^{-s}\,\mathrm{d}s \\ &= -\int_0^\infty s^z\,\mathrm{d}\left(\mathrm{e}^{-s}\right) \\ &\overset{\text{by parts}}{=} -\left(\left[s^z\mathrm{e}^{-s}\right]_{s=0}^{\infty} - \int_0^\infty \mathrm{e}^{-s}\,\mathrm{d}(s^z)\right) \\ &\overset{z\in\mathbb{N}}{=} -\left(0 - \int_0^\infty z \cdot s^{z-1}\mathrm{e}^{-s}\,\mathrm{d}s\right) \\ &= z \cdot \int_0^\infty s^{z-1}\mathrm{e}^{-s}\,\mathrm{d}s = z \cdot \Gamma(z) \end{aligned}$$

- $\Gamma(0)$

$$\begin{aligned}
\Gamma(0) &= \int_0^\infty s^{0-1}\mathsf{e}^{-s}\,\mathsf{d}s \\
&= \int_0^\infty s^{-1}\mathsf{e}^{-s}\,\mathsf{d}s \\
&= -\int_0^\infty s^{-1}\,\mathsf{d}\mathsf{e}^{-s} \\
&= -\left(\left[s^{-1}\mathsf{e}^{-s}\right]_{s=0}^{\infty} - \int_0^\infty \mathsf{e}^{-s}\,\mathsf{d}s^{-1}\right) \\
&= -\left(\left[s^{-1}\mathsf{e}^{-s}\right]_{s=0}^{\infty} + \int_0^\infty s^{-2}\mathsf{e}^{-s}\,\mathsf{d}s\right) \\
&= -\left([0-\infty] + \Gamma(-1)\right) = \infty - \Gamma(-1)
\end{aligned}$$

- $\Gamma(1) = 1$

$$\begin{aligned}
\Gamma(1) &= \int_0^\infty s^{1-1}\mathsf{e}^{-s}\,\mathsf{d}s \\
&= \int_0^\infty s^{0}\mathsf{e}^{-s}\,\mathsf{d}s = \int_0^\infty \mathsf{e}^{-s}\,\mathsf{d}s \\
&= \left[-\mathsf{e}^{-s}\right]_{s=0}^{\infty} = [-0-(-1)] = 1
\end{aligned}$$

- $\forall n \in \mathbb{N}\left(\Gamma(n) = (n-1)!\right)$

$$\begin{cases}\Gamma(1) = 1 \\ \forall z \in \mathbb{N}\left(\Gamma(z+1) = z\cdot\Gamma(z)\right)\end{cases}$$

$$\Gamma(2) = \Gamma(1+1) = 1\cdot\Gamma(1) = 1\cdot 1 = 1$$

$$\Gamma(3) = \Gamma(2+1) = 2\cdot\Gamma(2) = 2\cdot 1 = 2$$

$$\begin{aligned}
\Gamma(n) &= \Gamma((n-1)+1) \\
&= (n-1)\,\Gamma(n-1) \\
&= (n-1)(n-2)\,\Gamma(n-2) \\
&= (n-1)(n-2)\cdots 2\cdot\Gamma(2) \\
&= (n-1)(n-2)\cdots 2\cdot 1 = (n-1)!
\end{aligned}$$

- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

$$
\begin{aligned}
\Gamma\left(\frac{1}{2}\right) &= \int_0^\infty s^{\frac{1}{2}-1}\mathsf{e}^{-s}\,\mathsf{d}s = \int_0^\infty s^{\frac{-1}{2}}\mathsf{e}^{-s}\,\mathsf{d}s \\
&\overset{s=s_0^2}{=} \int_0^\infty \left(s_0^2\right)^{\frac{-1}{2}}\mathsf{e}^{-s_0^2}\,\mathsf{d}s_0^2 \\
&= \int_0^\infty s_0^{-1}\cdot\mathsf{e}^{-s_0^2}\cdot 2s_0\,\mathsf{d}s_0 \\
&= 2\int_0^\infty \mathsf{e}^{-s_0^2}\,\mathsf{d}s_0 = 2\sqrt{\left(\int_0^\infty \mathsf{e}^{-s_0^2}\,\mathsf{d}s_0\right)^2} \\
&= 2\sqrt{\left(\int_0^\infty \mathsf{e}^{-s_1^2}\,\mathsf{d}s_1\right)\left(\int_0^\infty \mathsf{e}^{-s_2^2}\,\mathsf{d}s_2\right)} \\
&\overset{\text{integration order}}{=} 2\sqrt{\left(\int_0^\infty\int_0^\infty \mathsf{e}^{-s_1^2}\mathsf{e}^{-s_2^2}\,\mathsf{d}s_1\,\mathsf{d}s_2\right)} \\
&= 2\sqrt{\left(\int_0^\infty\int_0^\infty \mathsf{e}^{-\left(s_1^2+s_2^2\right)}\,\mathsf{d}s_1\,\mathsf{d}s_2\right)} \\
&= 2\sqrt{\left(\int_0^{\frac{\pi}{2}}\int_0^\infty \mathsf{e}^{-r^2}\,r\mathsf{d}r\,\mathsf{d}\theta\right)} = 2\sqrt{\left(\int_0^{\frac{\pi}{2}}\left(\frac{1}{2}\int_0^\infty \mathsf{e}^{-r^2}\,\mathsf{d}r^2\right)\mathsf{d}\theta\right)} \\
&= 2\sqrt{\left(\int_0^{\frac{\pi}{2}}\frac{\left[-\mathsf{e}^{-r^2}\right]_{r=0}^\infty}{2}\,\mathsf{d}\theta\right)} = 2\sqrt{\left(\int_0^{\frac{\pi}{2}}\frac{-0-(-1)}{2}\,\mathsf{d}\theta\right)} \\
&= 2\sqrt{\left(\int_0^{\frac{\pi}{2}}\frac{1}{2}\,\mathsf{d}\theta\right)} = 2\sqrt{\left[\frac{\theta}{2}\right]_{\theta=0}^{\frac{\pi}{2}}} = 2\sqrt{\frac{\pi}{4}} = \sqrt{\pi}
\end{aligned}
$$

- $n\in\mathbb{N}\Rightarrow\Gamma\left(\frac{1}{2}+n\right)=\frac{(2n-1)!!}{2^n}\sqrt{\pi}=\frac{(2n)!}{n!4^n}\sqrt{\pi}$

$$
\begin{cases}\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}\\ \Gamma(z+1)=z\cdot\Gamma(z)\end{cases}\quad\Downarrow
$$

$$
\begin{aligned}
\Gamma\left(\frac{1}{2}+1\right)=\Gamma\left(\frac{3}{2}\right)=\Gamma\left(\frac{1}{2}+1\right) &= \frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)=\frac{1}{2}\cdot\sqrt{\pi} \\
\Gamma\left(\frac{1}{2}+2\right)=\Gamma\left(\frac{5}{2}\right)=\Gamma\left(\frac{3}{2}+1\right) &= \frac{3}{2}\cdot\Gamma\left(\frac{3}{2}\right)=\frac{3}{2}\cdot\frac{1}{2}\cdot\sqrt{\pi} \\
&\vdots \\
\Gamma\left(\frac{1}{2}+n\right)=\Gamma\left(\frac{1+2n}{2}\right)=\Gamma\left(\frac{2n-1}{2}+1\right) &= \frac{2n-1}{2}\cdot\Gamma\left(\frac{2n-1}{2}\right) \\
&= \frac{2n-1}{2}\cdot\frac{2n-3}{2}\cdot\cdots\cdot\frac{1}{2}\cdot\sqrt{\pi} \\
&= \frac{(2n-1)!!}{2^n}\sqrt{\pi} \\
&\overset{(2n-1)!!=\frac{(2n)!}{n!2^n}}{=} \frac{\frac{(2n)!}{n!2^n}}{2^n}\sqrt{\pi} \\
&= \frac{(2n)!}{n!4^n}\sqrt{\pi}
\end{aligned}
$$

- $n\in\mathbb{N}\Rightarrow\Gamma\left(\frac{1}{2}-n\right)=\frac{(-2)^n}{(2n-1)!!}\sqrt{\pi}=\frac{n!(-4)^n}{(2n)!}\sqrt{\pi}$

$$
\begin{cases}\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}\\ \Gamma(z+1)=z\cdot\Gamma(z)\overset{z\neq 0}{\Rightarrow}\Gamma(z)=\frac{\Gamma(z+1)}{z}\Rightarrow\Gamma(z-1)=\frac{\Gamma(z)}{z-1}\wedge z\neq 1\end{cases}
$$

$$
\Gamma\left(\frac{1}{2}-1\right)=\Gamma\left(\frac{-1}{2}\right)=\Gamma\left(\frac{1}{2}-1\right)=\frac{\Gamma\left(\frac{1}{2}\right)}{\frac{-1}{2}}=\frac{\sqrt{\pi}}{\frac{-1}{2}}
$$

$$\Gamma\left(\frac{1}{2}-2\right)=\Gamma\left(\frac{-3}{2}\right)=\Gamma\left(\frac{-1}{2}-1\right)=\frac{\Gamma\left(\frac{-1}{2}\right)}{\frac{-3}{2}}=\frac{\sqrt{\pi}}{\frac{-1}{2}\cdot\frac{-3}{2}}$$

$$\begin{aligned}
\Gamma\left(\frac{1}{2}-n\right) &= \Gamma\left(\frac{1-2n}{2}\right)=\frac{\Gamma\left(\frac{1}{2}-(n-1)\right)}{\frac{1-2n}{2}}=\frac{\sqrt{\pi}}{\frac{-1}{2}\cdot\frac{-3}{2}\cdot\ldots\cdot\frac{1-2n}{2}}\\
&= \frac{\sqrt{\pi}}{\frac{1}{-2}\cdot\frac{3}{-2}\cdot\ldots\cdot\frac{2n-1}{-2}}=\frac{\sqrt{\pi}}{\frac{(2n-1)!!}{(-2)^n}}=\frac{(-2)^n}{(2n-1)!!}\sqrt{\pi}\\
&\overset{(2n-1)!!=\frac{(2n)!}{n!2^n}}{=} \frac{(-2)^n}{\frac{(2n)!}{n!2^n}}\sqrt{\pi}=\frac{n!\,(-4)^n}{(2n)!}\sqrt{\pi}
\end{aligned}$$

- $\forall n\in\mathbb{N}\left(\Gamma\left(\frac{n}{2}\right)=\left(\frac{1}{\pi}\right)^k\prod_{i=1}^{\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor-1}4^k\left(\frac{n}{2}-4^k i\right)\wedge k=\frac{n}{2}-\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor\right)$

$$\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \begin{cases}\Gamma\left(\frac{1}{2}+(m-1)\right) & n=2m-1\\ \Gamma(m) & n=2m\end{cases}\wedge m\in\mathbb{N}\\
&= \begin{cases}\frac{(2(m-1))!}{(m-1)!4^{m-1}}\sqrt{\pi} & n=2m-1\\ (m-1)! & n=2m\end{cases}\wedge m\in\mathbb{N}\\
&= \begin{cases}\frac{\prod_{i=1}^{2(m-1)}i}{4^{m-1}\prod_{i=1}^{m-1}i}\sqrt{\pi} & n=2m-1\\ \prod_{i=1}^{m-1}i=\prod_{i=1}^{m-1}(m-i)=\prod_{i=1}^{m-1}\left(\frac{2m}{2}-i\right) & n=2m\end{cases}\wedge m\in\mathbb{N}
\end{aligned}$$

$$\begin{aligned}
\frac{\prod_{i=1}^{2(m-1)}i}{4^{m-1}\prod_{i=1}^{m-1}i} &= \frac{\left(\prod_{i=1}^{m-1}i\right)\left(\prod_{i=m}^{2(m-1)}i\right)}{4^{m-1}\prod_{i=1}^{m-1}i}\\
&= \frac{\prod_{i=m}^{2(m-1)}i}{4^{m-1}}=\frac{\prod_{i=1}^{m-1}(m-1+i)}{4^{m-1}}\\
&= \prod_{i=1}^{m-1}\frac{m-1+i}{4}=\prod_{i=1}^{m-1}\frac{m-1+(m-i)}{4}\\
&= \prod_{i=1}^{m-1}\frac{2m-1-i}{4}=\prod_{i=1}^{m-1}\left(\frac{2m-1}{4}-\frac{i}{4}\right)\\
&= \prod_{i=1}^{m-1}\frac{1}{2}\left(\frac{2m-1}{2}-\frac{i}{2}\right)
\end{aligned}$$

$$\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \begin{cases}\frac{\prod_{i=1}^{2(m-1)}i}{4^{m-1}\prod_{i=1}^{m-1}i}\sqrt{\pi} & n=2m-1\\ \prod_{i=1}^{m-1}\left(\frac{2m}{2}-i\right) & n=2m\end{cases}\wedge m\in\mathbb{N}\\
&= \begin{cases}\sqrt{\pi}\prod_{i=1}^{m-1}\frac{1}{2}\left(\frac{2m-1}{2}-\frac{i}{2}\right) & n=2m-1\\ \prod_{i=1}^{m-1}\left(\frac{2m}{2}-i\right) & n=2m\end{cases}\wedge m\in\mathbb{N}\\
&= \begin{cases}\sqrt{\pi}\prod_{i=1}^{\left\lfloor\frac{n+1}{2}\right\rfloor-1}\frac{1}{2}\left(\frac{n}{2}-\frac{i}{2}\right) & n=2m-1\\ \prod_{i=1}^{\left\lfloor\frac{n+1}{2}\right\rfloor-1}\left(\frac{n}{2}-i\right) & n=2m\end{cases}\wedge m\in\mathbb{N}\\
&= \left(\frac{1}{\pi}\right)^k\prod_{i=1}^{\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor-1}4^k\left(\frac{n}{2}-4^k i\right)\wedge k=\frac{n}{2}-\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor\\
&= \left(\sqrt{\pi}\right)^k\prod_{i=1}^{\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor-1}\left(\frac{1}{2}\right)^k\left(\frac{n}{2}-\left(\frac{1}{2}\right)^k i\right)\wedge k=2\left(\left\lfloor\frac{n}{2}+\frac{1}{2}\right\rfloor-\frac{n}{2}\right)
\end{aligned}$$

- $\forall n \in \mathbb{N}\left(2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) = (2\pi)^{-\frac{n}{2}}\left(2^{n-2(k-1)}\pi\right)^k \prod_{i=1}^{k-1}\left(\frac{n}{2} - 2^{n-2k}i\right) \wedge k = \left\lfloor \frac{n}{2} + \frac{1}{2}\right\rfloor\right)$

$$\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \left(\sqrt{\pi}\right)^k \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{1}{2}\right)^k \left(\frac{n}{2} - \left(\frac{1}{2}\right)^k i\right) \wedge k = 2\left(\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor - \frac{n}{2}\right) \\
&= \left(\sqrt{\pi}\right)^{2\left(\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-\frac{n}{2}\right)} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{1}{2}\right)^{2\left(\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-\frac{n}{2}\right)} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left(\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-\frac{n}{2}\right)} i\right) \\
&= (\pi)^{-\frac{n}{2}+\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (\pi)^{-\frac{n}{2}+\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} 2^{n-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (\pi)^{-\frac{n}{2}+\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} 2^{-n+(n+2)\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor^2} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \pi^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} 2^{-\frac{n}{2}+(n+2)\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor^2} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right)
\end{aligned}$$

$$\begin{aligned}
2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right) &= (2\pi)^{-\frac{n}{2}} \pi^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} 2^{(n+2)\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor^2} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \pi^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} 2^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor\left((n+2)-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor\right)} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \left(2^{n+2\left(1-\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor\right)}\pi\right)^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - \left(\frac{1}{2}\right)^{2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-n} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \left(2^{n+2\left(1-\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor\right)}\pi\right)^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - 2^{n-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \left(2^{n-2\left(\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1\right)}\pi\right)^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} \prod_{i=1}^{\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor-1} \left(\frac{n}{2} - 2^{n-2\left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor} i\right) \\
&= (2\pi)^{-\frac{n}{2}} \left(2^{n-2(k-1)}\pi\right)^k \prod_{i=1}^{k-1} \left(\frac{n}{2} - 2^{n-2k} i\right) \wedge k = \left\lfloor \frac{n}{2}+\frac{1}{2}\right\rfloor
\end{aligned}$$

6.4.1 卡方分配

Definition 6.4.5. χ^2 distribution = chi-square distribution

$$\begin{aligned}
X \sim \chi^2(n) \quad &\Leftrightarrow \quad f_X\left(\chi^2; n\right) = f_{X,n}\left(\chi^2\right) = f_n\left(\chi^2\right) = f\left(\chi^2\right) = \frac{\left(\chi^2\right)^{\frac{n}{2}-1} \mathsf{e}^{\frac{-\left(\chi^2\right)}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \\
&\Leftrightarrow \quad \mathsf{dP}\left(X = \chi^2\right) = f_X\left(\chi^2; n\right) \mathsf{d}x \\
&\Leftrightarrow \quad \mathsf{dP}(X = x) = \begin{cases} \dfrac{x^{\frac{n}{2}-1}\mathsf{e}^{\frac{-x}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\mathsf{d}x & x > 0 \\ 0\,\mathsf{d}x & x \le 0 \end{cases}
\end{aligned}$$

Fact 6.4.1. $X \sim \chi^2(n) \Rightarrow \begin{cases} \mathsf{M}(\xi) = (1-2\xi)^{\frac{-n}{2}} \\ \mathsf{E}(X) = n \\ \mathsf{V}(X) = 2n \end{cases}$

Proof.

$$
\begin{aligned}
\mathsf{M}(\xi)=\mathsf{E}\left(\mathsf{e}^{\xi X}\right) &= \int_{-\infty}^{\infty}\mathsf{e}^{\xi x}\,\mathsf{dP}(X=x)\\
&= \int_{-\infty}^{0}\mathsf{e}^{\xi x}\,\mathsf{dP}(X=x)+\int_{0}^{\infty}\mathsf{e}^{\xi x}\,\mathsf{dP}(X=x)\\
&= \int_{-\infty}^{0}\mathsf{e}^{\xi x}\cdot 0\,\mathsf{d}x+\int_{0}^{\infty}\mathsf{e}^{\xi x}\cdot\frac{x^{\frac{n}{2}-1}\mathsf{e}^{\frac{-x}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\,\mathsf{d}x\\
&= 0+\frac{\int_0^{\infty}x^{\frac{n}{2}-1}\mathsf{e}^{-\left(\frac{1}{2}-\xi\right)x}\,\mathsf{d}x}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\\
&\overset{6.4.1}{=} \frac{\frac{\Gamma\left(\frac{n}{2}\right)}{\left(\frac{1}{2}-\xi\right)^{\frac{n}{2}}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}=(1-2\xi)^{\frac{-n}{2}}
\end{aligned}
$$

$$
\mathsf{C}(\xi)=\ln\mathsf{M}(\xi)=\ln(1-2\xi)^{\frac{-n}{2}}=\frac{-n}{2}\ln(1-2\xi)
$$

$$
\begin{aligned}
\mathsf{E}(X)=\dot{\mathsf{C}}(0) &= \frac{\mathsf{d}}{\mathsf{d}\xi}\left(\frac{-n}{2}\ln(1-2\xi)\right)|_{\xi=0}\\
&= \frac{-n}{2}\frac{-2}{1-2\xi}|_{\xi=0}\\
&= \frac{n}{1-2\xi}|_{\xi=0}=n
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{V}(X)=\ddot{\mathsf{C}}(0) &= \frac{\mathsf{d}}{\mathsf{d}\xi}\frac{n}{1-2\xi}|_{\xi=0}\\
&= -n(1-2\xi)^{-2}\cdot(-2)|_{\xi=0}\\
&= 2n(1-2\xi)^{-2}|_{\xi=0}\\
&= 2n
\end{aligned}
$$

□

Definition 6.4.6. χ^2 distribution (alternate definition)

$$
\begin{cases}\overset{n}{\underset{i=1}{,}}Z_i \text{ are independent}\\ Z_i\sim\mathsf{n}(0,1)\end{cases}\Leftrightarrow\left(\sum_{i=1}^{n}(Z_i)^2\right)\sim\chi^2(n)
$$

Theorem 6.4.1.

$$
\begin{aligned}
\left(\sum_{i=1}^{n}(Z_i)^2\right)\sim\chi^2(n)\quad\Rightarrow\quad & f_{\sum_{i=1}^{n}(Z_i)^2}\left(\chi^2=\sum_{i=1}^{n}(z_i)^2\right)\\
&=\begin{cases}\dfrac{\left(\chi^2\right)^{\frac{n}{2}-1}\mathsf{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} & \chi^2>0\\ 0 & \chi^2\le 0\end{cases}
\end{aligned}
$$

Proof.

$$
\chi^2_{[n]}=\sum_{i=1}^{n}(z_i)^2
$$

$$
\chi^2_{[1]}=z_1^2,\chi^2_{[2]}=z_1^2+z_2^2,\ldots
$$

$$
\begin{aligned}
&\begin{cases}\chi^2_{[2]}=\chi^2_{[2]}(z_1,z_2)=z_1^2+z_2^2\\ z_2^2=z_2^2\end{cases}\\
\Leftrightarrow&\begin{cases}z_1^2=z_1^2\left(\chi^2_{[2]},z_2^2\right)=\chi^2_{[2]}-z_2^2\\ z_2^2=z_2^2\end{cases}
\end{aligned}
$$

$$Z_i \sim \mathsf{n}(0,1) \Leftrightarrow f_{Z_i}(z_i) = \frac{\mathsf{e}^{\frac{-z_i^2}{2}}}{\sqrt{2\pi}}$$

$$Z_i^2 = (Z_i)^2 \Leftrightarrow Z_i = \pm\sqrt{Z_i^2}$$

$$\begin{aligned} f_{Z_i^2}\left(z_i^2\right) &= \sum f_{Z_i}\left(z_i\left(z_i^2\right)\right)\left|\frac{\mathsf{d}z_i}{\mathsf{d}z_i^2}\right| \\ &\overset{(z_i)^2=(-z_i)^2}{=} 2\cdot\frac{\mathsf{e}^{\frac{-z_i^2}{2}}}{\sqrt{2\pi}}\cdot\left|\frac{1}{2\sqrt{z_i^2}}\right| \\ &= \frac{\left(z_i^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z_i^2}{2}}}{\sqrt{2\pi}} \end{aligned}$$

$$f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right) = f_{Z_1^2}\left(z_1^2\right) = \frac{\left(z_1^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z_1^2}{2}}}{\sqrt{2\pi}} = \frac{\left(\chi^2_{[1]}\right)^{\frac{1}{2}-1}\mathsf{e}^{\frac{-\chi^2_{[1]}}{2}}}{2^{\frac{1}{2}}\Gamma\left(\frac{1}{2}\right)}$$

$$\begin{aligned} f_{Z_1^2,Z_2^2}\left(z_1^2,z_2^2\right) &\overset{\text{independence}}{=} f_{Z_1^2}\left(z_1^2\right)f_{Z_2^2}\left(z_2^2\right) \\ &= \frac{\left(z_1^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z_1^2}{2}}}{\sqrt{2\pi}}\cdot\frac{\left(z_2^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z_2^2}{2}}}{\sqrt{2\pi}} \\ &= \frac{\left(z_1^2z_2^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\left(z_1^2+z_2^2\right)}{2}}}{\left(\sqrt{2\pi}\right)^2} \end{aligned}$$

$$\begin{cases} \chi^2_{[2]} = \chi^2_{[2]}(z_1,z_2) = z_1^2+z_2^2 \\ z_2^2 = z_2^2 \end{cases}$$

$$\Leftrightarrow \begin{cases} z_1^2 = z_1^2\left(\chi^2_{[2]},z_2^2\right) = \chi^2_{[2]} - z_2^2 \\ z_2^2 = z_2^2 \end{cases}$$

$$\begin{aligned} f_{\chi^2_{[2]},Z_2^2}\left(\chi^2_{[2]},z_2^2\right) &= f_{Z_1^2,Z_2^2}\left(z_1^2\left(\chi^2_{[2]},z_2^2\right),z_2^2\right)\left|\frac{\partial z_1^2}{\partial\chi^2_{[2]}}\right| \\ &= \frac{\left(\left(\chi^2_{[2]}-z_2^2\right)z_2^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\left(\left(\chi^2_{[2]}-z_2^2\right)+z_2^2\right)}{2}}}{\left(\sqrt{2\pi}\right)^2}\cdot 1 \\ &= \frac{\left(\left(\chi^2_{[2]}-z_2^2\right)z_2^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \end{aligned}$$

$$\begin{aligned}
f_{\chi^2_{[2]}}\left(\chi^2_{[2]}\right) &= \int_0^{\chi^2_{[2]}} f_{\chi^2_{[2]},Z_2^2}\left(\chi^2_{[2]}, z_2^2\right) \mathrm{d}z_2^2 \\
&= \int_0^{\chi^2_{[2]}} \frac{\left(\left(\chi^2_{[2]} - z_2^2\right) z_2^2\right)^{\frac{-1}{2}} \mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \mathrm{d}z_2^2 \\
&= \frac{\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \int_0^{\chi^2_{[2]}} \frac{\mathrm{d}z_2^2}{\sqrt{z_2^2}\sqrt{\chi^2_{[2]} - z_2^2}} \\
&= \frac{\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \int_0^{\sqrt{\chi^2_{[2]}}} \frac{2\,\mathrm{d}\sqrt{z_2^2}}{\sqrt{\chi^2_{[2]} - z_2^2}} \\
&\overset{\chi^2_{[2]}>0}{=} \frac{2\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \int_0^1 \frac{\mathrm{d}\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}}{\sqrt{1 - \left(\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}\right)^2}}
\end{aligned}$$

$$0 \le \frac{z_2^2}{\chi^2_{[2]}} = 1 - \frac{z_1^2}{z_1^2 + z_2^2} \le 1$$

$$\begin{aligned}
\int_0^1 \frac{\mathrm{d}\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}}{\sqrt{1 - \left(\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}\right)^2}} &\overset{\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}=\sin\theta}{=} \int_0^1 \frac{\mathrm{d}\sin\theta}{\sqrt{1 - (\sin\theta)^2}} \\
&= \int_0^{\frac{\pi}{2}} \frac{\cos\theta\,\mathrm{d}\theta}{\sqrt{(\cos\theta)^2}} = \int_0^{\frac{\pi}{2}} \frac{\cos\theta\,\mathrm{d}\theta}{|\cos\theta|} \\
&= \int_0^{\frac{\pi}{2}} \frac{\cos\theta\,\mathrm{d}\theta}{\cos\theta} \because \forall\theta \in \left[0, \frac{\pi}{2}\right) (0 < \cos\theta \le 1) \\
&= \int_0^{\frac{\pi}{2}} \mathrm{d}\theta = [\theta]_{\theta=0}^{\frac{\pi}{2}} = \frac{\pi}{2}
\end{aligned}$$

$$\begin{aligned}
f_{\chi^2_{[2]}}\left(\chi^2_{[2]}\right) &\overset{\chi^2_{[2]}>0}{=} \frac{2\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \int_0^1 \frac{\mathrm{d}\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}}{\sqrt{1 - \left(\sqrt{\frac{z_2^2}{\chi^2_{[2]}}}\right)^2}} \\
&= \frac{2\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{\left(\sqrt{2\pi}\right)^2} \cdot \frac{\pi}{2} = \frac{\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{2} \wedge \chi^2_{[2]} > 0 \\
&= \frac{\left(\chi^2_{[2]}\right)^{\frac{2}{2}-1} \mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{2^{\frac{2}{2}}\Gamma\left(\frac{2}{2}\right)} \wedge \chi^2_{[2]} > 0
\end{aligned}$$

$$\begin{aligned}
f_{\chi^2_{[2]},Z_3^2}\left(\chi^2_{[2]}, z_3^2\right) &\overset{\text{independence}}{=} f_{\chi^2_{[2]}}\left(\chi^2_{[2]}\right) f_{Z_3^2}\left(z_3^2\right) \\
&= \frac{\mathrm{e}^{\frac{-\chi^2_{[2]}}{2}}}{2} \cdot \frac{\left(z_3^2\right)^{\frac{-1}{2}} \mathrm{e}^{\frac{-z_3^2}{2}}}{\sqrt{2\pi}} \\
&= \frac{\left(z_3^2\right)^{\frac{-1}{2}} \mathrm{e}^{\frac{-\left(\chi^2_{[2]}+z_3^2\right)}{2}}}{2\sqrt{2\pi}}
\end{aligned}$$

$$\begin{cases}\chi^2_{[3]} = \chi^2_{[3]}\left(\chi^2_{[2]}, z_3\right) = \chi^2_{[2]} + z_3^2 \\ z_3^2 = z_3^2\end{cases}$$

$$\Leftrightarrow \quad \begin{cases}\chi^2_{[2]} = \chi^2_{[2]}\left(\chi^2_{[3]}, z_3^2\right) = \chi^2_{[3]} - z_3^2 \\ z_3^2 = z_3^2\end{cases}$$

$$\begin{aligned}
f_{\chi^2_{[3]},Z_3^2}\left(\chi^2_{[3]}, z_3^2\right) &= f_{\chi^2_{[2]},Z_3^2}\left(\chi^2_{[2]}\left(\chi^2_{[3]}, z_3^2\right), z_3^2\right)\left|\frac{\partial\chi^2_{[2]}}{\partial\chi^2_{[3]}}\right| \\
&= \frac{\left(z_3^2\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-\left(\left(\chi^2_{[3]}-z_3^2\right)+z_3^2\right)}{2}}}{2\sqrt{2\pi}} \cdot 1 \\
&= \frac{\left(z_3^2\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2\sqrt{2\pi}}
\end{aligned}$$

$$\begin{aligned}
f_{\chi^2_{[3]}}\left(\chi^2_{[3]}\right) &= \int_0^{\chi^2_{[3]}} f_{\chi^2_{[3]},Z_3^2}\left(\chi^2_{[3]}, z_3^2\right) \mathsf{d}z_3^2 \\
&= \int_0^{\chi^2_{[3]}} \frac{\left(z_3^2\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2\sqrt{2\pi}} \mathsf{d}z_3^2 \\
&= \frac{\mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2\sqrt{2\pi}} \int_0^{\chi^2_{[3]}} \frac{\mathsf{d}z_3^2}{\sqrt{z_3^2}} \\
&= \frac{\mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2\sqrt{2\pi}} \int_0^{\sqrt{\chi^2_{[3]}}} 2\,\mathsf{d}\sqrt{z_3^2} \\
&= \frac{\mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2\sqrt{2\pi}} \cdot 2\sqrt{\chi^2_{[3]}} = \frac{\sqrt{\chi^2_{[3]}}\mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{\sqrt{2\pi}} = \frac{\left(\chi^2_{[3]}\right)^{\frac{3}{2}-1} \mathsf{e}^{\frac{-\chi^2_{[3]}}{2}}}{2^{\frac{3}{2}}\Gamma\left(\frac{3}{2}\right)}
\end{aligned}$$

$$\begin{aligned}
&f_{\chi^2_{[n+1]},Z_{n+1}^2}\left(\chi^2_{[n+1]}, z_{n+1}^2\right) = f_{\chi^2_{[n]},Z_{n+1}^2}\left(\chi^2_{[n]}, z_{n+1}^2\right)\left\|\frac{\partial\left(\chi^2_{[n]}, z_{n+1}^2\right)}{\partial\left(\chi^2_{[n+1]}, z_{n+1}^2\right)}\right\| \\
=&f_{\chi^2_{[n]},Z_{n+1}^2}\left(\chi^2_{[n]}, z_{n+1}^2\right)\left|\frac{\partial\chi^2_{[n]}}{\partial\chi^2_{[n+1]}}\right| = f_{\chi^2_{[n]},Z_{n+1}^2}\left(\chi^2_{[n]}, z_{n+1}^2\right)\left|\frac{\partial\left(\chi^2_{[n+1]} - z_{n+1}^2\right)}{\partial\chi^2_{[n+1]}}\right| \\
=&f_{\chi^2_{[n]},Z_{n+1}^2}\left(\chi^2_{[n]}, z_{n+1}^2\right)|1| = f_{\chi^2_{[n]},Z_{n+1}^2}\left(\chi^2_{[n]}, z_{n+1}^2\right) = f_{\chi^2_{[n]}}\left(\chi^2_{[n]}\right) f_{Z_{n+1}^2}\left(z_{n+1}^2\right)
\end{aligned}$$

$$\begin{cases}
f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right) = \frac{\left(\chi^2_{[1]}\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-\chi^2_{[1]}}{2}}}{\sqrt{2\pi}} \\
f_{Z_i^2}\left(z_i^2\right) = \frac{\left(z_i^2\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-z_i^2}{2}}}{\sqrt{2\pi}} \Rightarrow f_{Z_1^2}\left(z_1^2\right) = \frac{\left(z_1^2\right)^{\frac{-1}{2}} \mathsf{e}^{\frac{-z_1^2}{2}}}{\sqrt{2\pi}} \\
\chi^2_{[n+1]} = \chi^2_{[n]} + z_{n+1}^2 \\
f_{\chi^2_{[n+1]},Z_{n+1}^2}\left(\chi^2_{[n+1]}, z_{n+1}^2\right) = f_{\chi^2_{[n]}}\left(\chi^2_{[n]}\right) f_{Z_{n+1}^2}\left(z_{n+1}^2\right) \\
f_{\chi^2_{[n+1]}}\left(\chi^2_{[n+1]}\right) = \int_0^{\chi^2_{[n+1]}} f_{\chi^2_{[n+1]},Z_{n+1}^2}\left(\chi^2_{[n+1]}, z_{n+1}^2\right) \mathsf{d}z_{n+1}^2
\end{cases}$$

$$
\begin{aligned}
f_{\chi^2_{[n]}}\left(\chi^2_{[n]}\right) &= \int_0^{\chi^2_{[n]}} f_{\chi^2_{[n]},Z^2_n}\left(\chi^2_{[n]}, z^2_n\right) \mathsf{d}z^2_n = \int_0^{\chi^2_{[n]}} f_{\chi^2_{[n-1]}}\left(\chi^2_{[n-1]}\right) f_{Z^2_n}\left(z^2_n\right) \mathsf{d}z^2_n \\
&= \int_0^{\chi^2_{[n]}}\left(\int_0^{\chi^2_{[n-1]}} f_{\chi^2_{[n-2]}}\left(\chi^2_{[n-2]}\right) f_{Z^2_{n-1}}\left(z^2_{n-1}\right) \mathsf{d}z^2_{n-1}\right) f_{Z^2_n}\left(z^2_n\right) \mathsf{d}z^2_n \\
&= \int_0^{\chi^2_{[n]}}\left(\int_0^{\chi^2_{[n-1]}} f_{\chi^2_{[n-2]}}\left(\chi^2_{[n-2]}\right) f_{Z^2_{n-1}}\left(z^2_{n-1}\right) f_{Z^2_n}\left(z^2_n\right) \mathsf{d}z^2_{n-1}\right) \mathsf{d}z^2_n \\
&\vdots \\
&= \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right)\cdots f_{Z^2_{n-1}}\left(z^2_{n-1}\right) f_{Z^2_n}\left(z^2_n\right) \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \prod_{i=1}^n f_{Z^2_i}\left(z^2_i\right) \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \prod_{i=1}^n \frac{\left(z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z^2_i}{2}}}{\sqrt{2\pi}} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \frac{\left(\prod_{i=1}^n z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\sum_{i=1}^n z^2_i}{2}}}{\left(\sqrt{2\pi}\right)^n} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= \frac{\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \left(\prod_{i=1}^n z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\sum_{i=1}^n z^2_i}{2}} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n}{\left(\sqrt{2\pi}\right)^n} \\
&= (2\pi)^{\frac{-n}{2}}\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \left(\prod_{i=1}^n z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\sum_{i=1}^n z^2_i}{2}} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= (2\pi)^{\frac{-n}{2}}\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \left(\prod_{i=1}^n z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&= (2\pi)^{\frac{-n}{2}}\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} \left(\prod_{i=1}^n z^2_i\right)^{\frac{-1}{2}} \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n \\
&\text{or} \\
&= \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}} f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right)\prod_{i=2}^n f_{Z^2_i}\left(z^2_i\right) \mathsf{d}z^2_2\cdots \mathsf{d}z^2_{n-1}\,\mathsf{d}z^2_n
\end{aligned}
$$

$$
\begin{aligned}
f_{\chi^2_{[n+1]}}\left(\chi^2_{[n+1]}\right) &= \int_0^{\chi^2_{[n+1]}} f_{\chi^2_{[n+1]},Z^2_{n+1}}\left(\chi^2_{[n+1]}, z^2_{n+1}\right) \mathsf{d}z^2_{n+1} \\
f_{\chi^2_{[0+1]}}\left(\chi^2_{[0+1]}\right) &= \int_0^{\chi^2_{[0+1]}} f_{\chi^2_{[0+1]},Z^2_{0+1}}\left(\chi^2_{[0+1]}, z^2_{0+1}\right) \mathsf{d}z^2_{0+1} \\
f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right) &= \int_0^{\chi^2_{[1]}} f_{\chi^2_{[1]},Z^2_1}\left(\chi^2_{[1]}, z^2_1\right) \mathsf{d}z^2_1
\end{aligned}
$$

weired definition

$$
f_{\chi^2_{[1]},Z^2_1}\left(\chi^2_{[1]}, z^2_1\right)
$$

but we continue calculation

$$
\begin{aligned}
\chi^2_{[n+1]} &= \chi^2_{[n]} + z^2_{n+1} \\
\chi^2_{[0+1]} &= \chi^2_{[0]} + z^2_{0+1} \\
\chi^2_{[1]} &= \chi^2_{[0]} + z^2_1 \\
\chi^2_{[1]} = z^2_1 \Downarrow & \\
\chi^2_{[0]} &= 0
\end{aligned}
$$

$$
\begin{aligned}
f_{\chi^2_{[n+1]},Z^2_{n+1}}\left(\chi^2_{[n+1]},z^2_{n+1}\right) =& f_{\chi^2_{[n]}}\left(\chi^2_{[n]}\right) f_{Z^2_{n+1}}\left(z^2_{n+1}\right)\\
f_{\chi^2_{[0+1]},Z^2_{0+1}}\left(\chi^2_{[0+1]},z^2_{0+1}\right) =& f_{\chi^2_{[0]}}\left(\chi^2_{[0]}\right) f_{Z^2_{0+1}}\left(z^2_{0+1}\right)\\
f_{\chi^2_{[1]},Z^2_{1}}\left(\chi^2_{[1]},z^2_{1}\right) =& f_{\chi^2_{[0]}}\left(\chi^2_{[0]}\right) f_{Z^2_{1}}\left(z^2_{1}\right)
\end{aligned}
$$

only one result for $\chi^2_{[0]}$

$$
\chi^2_{[0]} = 0
$$

$$
\mathsf{P}\left(\chi^2_{[0]} = 0\right) = 1
$$

$$
f_{\chi^2_{[0]}}\left(\chi^2_{[0]}\right) = 1 = \delta\left(\chi^2_{[1]} - z^2_1\right)\Big|_{\chi^2_{[1]}=z^2_1}
$$

$$
f_{\chi^2_{[1]},Z^2_1}\left(\chi^2_{[1]},z^2_1\right) = f_{\chi^2_{[0]}}\left(\chi^2_{[0]}\right) f_{Z^2_1}\left(z^2_1\right) = 1\cdot f_{Z^2_1}\left(z^2_1\right) = f_{Z^2_1}\left(z^2_1\right)
$$

$$
\begin{aligned}
f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right) =& \int_0^{\chi^2_{[1]}} f_{\chi^2_{[1]},Z^2_1}\left(\chi^2_{[1]},z^2_1\right)\mathsf{d}z^2_1\\
=& \int_0^{\chi^2_{[1]}} f_{Z^2_1}\left(z^2_1\right)\mathsf{d}z^2_1
\end{aligned}
$$

$$
\begin{aligned}
f_{\chi^2_{[n]}}\left(\chi^2_{[n]}\right) =& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\left(f_{\chi^2_{[1]}}\left(\chi^2_{[1]}\right)\right)\prod_{i=2}^{n} f_{Z^2_i}\left(z^2_i\right)\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\left(\int_0^{\chi^2_{[1]}} f_{Z^2_1}\left(z^2_1\right)\mathsf{d}z^2_1\right)\prod_{i=2}^{n} f_{Z^2_i}\left(z^2_i\right)\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}} f_{Z^2_1}\left(z^2_1\right)\prod_{i=2}^{n} f_{Z^2_i}\left(z^2_i\right)\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\prod_{i=1}^{n} f_{Z^2_i}\left(z^2_i\right)\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\prod_{i=1}^{n}\frac{\left(z^2_i\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-z^2_i}{2}}}{\sqrt{2\pi}}\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\left(\prod_{i=1}^{n} z^2_i\right)^{\frac{-1}{2}}\frac{\mathsf{e}^{\frac{-\sum_{i=1}^{n} z^2_i}{2}}}{\left(\sqrt{2\pi}\right)^n}\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\left(\prod_{i=1}^{n} z^2_i\right)^{\frac{-1}{2}}\frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\left(\prod_{i=1}^{n} z^2_i\right)^{\frac{-1}{2}}\mathsf{d}z^2_1\mathsf{d}z^2_2\cdots\mathsf{d}z^2_{n-1}\mathsf{d}z^2_n\\
=& \frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}\int_0^{\chi^2_{[n]}}\int_0^{\chi^2_{[n-1]}}\cdots\int_0^{\chi^2_{[2]}}\int_0^{\chi^2_{[1]}}\frac{\mathsf{d}z^2_1}{z_1}\frac{\mathsf{d}z^2_2}{z_2}\cdots\frac{\mathsf{d}z^2_n}{z_n}\frac{\mathsf{d}z^2_n}{z_n}\\
=& \frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}2\int_0^{\chi_{[n]}}2\int_0^{\chi_{[n-1]}}\cdots 2\int_0^{\chi_{[2]}}2\int_0^{\chi_{[1]}}\mathsf{d}z_1\mathsf{d}z_2\cdots\mathsf{d}z_{n-1}\mathsf{d}z_n\\
=& \frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}\int_{-\chi_{[n]}}^{\chi_{[n]}}\int_{-\chi_{[n-1]}}^{\chi_{[n-1]}}\cdots\int_{-\chi_{[2]}}^{\chi_{[2]}}\int_{-\chi_{[1]}}^{\chi_{[1]}}\mathsf{d}z_1\mathsf{d}z_2\cdots\mathsf{d}z_{n-1}\mathsf{d}z_n\\
=& \frac{\mathsf{e}^{\frac{-\chi^2_{[n]}}{2}}}{\left(\sqrt{2\pi}\right)^n}\int_{-\chi_{[n]}}^{\chi_{[n]}}\int_{-\chi_{[n-1]}}^{\chi_{[n-1]}}\cdots\int_{-\chi_{[2]}}^{\chi_{[2]}}\int_{-\chi_{[1]}}^{\chi_{[1]}}\left(\prod_{i=1}^{n}\mathsf{d}z_i\right)
\end{aligned}
$$

$$\int_{-\chi_{[n]}}^{\chi_{[n]}}\int_{-\chi_{[n-1]}}^{\chi_{[n-1]}}\cdots\int_{-\chi_{[2]}}^{\chi_{[2]}}\int_{-\chi_{[1]}}^{\chi_{[1]}}\left(\prod_{i=1}^{n}\mathsf{d}z_i\right)=\sum_{\chi_{[n]}^2=\sum_{i=1}^{n}z_i^2}\left(\prod_{i=1}^{n}\mathsf{d}z_i\right)$$

□

Definition 6.4.7. $(n-1)$ -sphere and n-ball

$$S^{n-1}=S_r^{n-1}=\left\{\left(\overset{n}{\underset{i=1}{,}}x_i\right)\mid\begin{cases}\forall i\in\mathbb{N}\cap[1,n]\,(x_i\in\mathbb{R})\\ \sum_{i=1}^{n}x_i^2=r^2>0\end{cases}\right\}$$
$$\Leftrightarrow\quad S_r^{n-1}\text{ is an }(n-1)\text{-sphere in }n\text{-Euclidean space}(\mathbb{R}^n)\text{ with radius }r$$

$$B^{n}=B_r^{n}=\left\{\left(\overset{n}{\underset{i=1}{,}}x_i\right)\mid\begin{cases}\forall i\in\mathbb{N}\cap[1,n]\,(x_i\in\mathbb{R})\\ \sum_{i=1}^{n}x_i^2\le r^2>0\end{cases}\right\}$$
$$\Leftrightarrow\quad B_r^{n}\text{ is an }n\text{-ball in }n\text{-Euclidean space}(\mathbb{R}^n)\text{ with radius }r$$

Lemma 6.4.1. $\begin{cases}V_n\text{ is the volum of }B_r^n\\ A_n\text{ is the surface area of }B_r^n\end{cases}\Rightarrow\begin{cases}V_n=c_nr^n=\frac{\left(\sqrt{\pi}\right)^n}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)}r^n\\ A_n=nc_nr^{n-1}=\frac{\left(\sqrt{\pi}\right)^n}{\frac{1}{2}\Gamma\left(\frac{n}{2}\right)}r^{n-1}\end{cases}$

Proof.

$$V_n=c_nr^n$$
$$\prod_{i=1}^{n}\mathsf{d}x_i=\mathsf{d}V_n=A_n\,\mathsf{d}r$$
$$A_n=\frac{\mathsf{d}V_n}{\mathsf{d}r}=\frac{\mathsf{d}}{\mathsf{d}r}c_nr^n=c_n\frac{\mathsf{d}}{\mathsf{d}r}r^n=c_n\cdot nr^{n-1}=nc_nr^{n-1}$$
$$\sum_{i=1}^{n}x_i^2=r^2$$

$$\begin{aligned}
\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathsf{e}^{-r^2}\prod_{i=1}^{n}\mathsf{d}x_i=\int_0^\infty\mathsf{e}^{-r^2}\,\mathsf{d}V_n&=\int_0^\infty\mathsf{e}^{-r^2}A_n\,\mathsf{d}r\\
\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mathsf{e}^{-\sum_{i=1}^{n}x_i^2}\prod_{i=1}^{n}\mathsf{d}x_i=&=\int_0^\infty\mathsf{e}^{-r^2}\cdot nc_nr^{n-1}\,\mathsf{d}r\\
\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left(\prod_{i=1}^{n}\mathsf{e}^{-x_i^2}\right)\left(\prod_{i=1}^{n}\mathsf{d}x_i\right)=&=nc_n\int_0^\infty r^{n-1}\mathsf{e}^{-r^2}\,\mathsf{d}r\\
\prod_{i=1}^{n}\left(\int_{-\infty}^{\infty}\mathsf{e}^{-x_i^2}\,\mathsf{d}x_i\right)=&=nc_n\cdot\frac{1}{2}\cdot2\int_0^\infty r^{2\cdot\frac{n}{2}-1}\mathsf{e}^{-r^2}\,\mathsf{d}r\\
\prod_{i=1}^{n}\left(\int_{-\infty}^{\infty}\mathsf{e}^{-s^2}\,\mathsf{d}s\right)=&=\frac{nc_n}{2}\cdot\Gamma\left(\frac{n}{2}\right)\\
\left(\int_{-\infty}^{\infty}\mathsf{e}^{-s^2}\,\mathsf{d}s\right)^n=&\\
\left(\int_0^\infty\mathsf{e}^{-s^2}\,\mathsf{d}s-\int_{-\infty}^{0}\mathsf{e}^{-s^2}\,\mathsf{d}s\right)^n=&\\
\left(2\int_0^\infty s^{2\cdot\frac{1}{2}-1}\mathsf{e}^{-s^2}\,\mathsf{d}s\right)^n=&\\
\left(\Gamma\left(\frac{1}{2}\right)\right)^n=&\\
\left(\sqrt{\pi}\right)^n=&
\end{aligned}$$

$$c_n = \frac{\left(\sqrt{\pi}\right)^n}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)} \Rightarrow \begin{cases} V_n = c_n r^n = \frac{\left(\sqrt{\pi}\right)^n}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)} r^n \\ A_n = n c_n r^{n-1} = \frac{\left(\sqrt{\pi}\right)^n}{\frac{1}{2}\Gamma\left(\frac{n}{2}\right)} r^{n-1} \end{cases}$$

□

Theorem 6.4.2. $\begin{cases} X^2 = \left(\sum_{i=1}^n (Z_i)^2\right) \sim \chi^2(n) \\ \chi^2 > 0 \end{cases} \Rightarrow f_{X^2}\left(\chi^2\right) \mathrm{d}\chi^2 = \frac{\left(\chi^2\right)^{\frac{n}{2}-1} \mathrm{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \mathrm{d}\chi^2$

Proof.

$$\begin{aligned}
& f_{X^2}\left(\chi^2\right) \mathrm{d}\chi^2 \\
\overset{4.6.2}{=} & \sum_{\chi^2 = \sum_{i=1}^n z_i^2} f_{\underset{i=1}{\overset{n}{,}} Z_i}\left(\underset{i=1}{\overset{n}{,}} z_i\right) \prod_{i=1}^n \mathrm{d}z_i \\
\overset{\text{independence}}{=} & \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\prod_{i=1}^n f_{Z_i}(z_i)\, \mathrm{d}z_i\right) \\
\overset{Z_i \sim \mathrm{n}(0,1)}{=} & \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\prod_{i=1}^n \frac{\mathrm{e}^{\frac{-z_i^2}{2}}}{\sqrt{2\pi}} \mathrm{d}z_i\right) \\
= & \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\frac{\mathrm{e}^{\frac{-\sum_{i=1}^n z_i^2}{2}}}{\left(\sqrt{2\pi}\right)^n} \prod_{i=1}^n \mathrm{d}z_i\right) \\
= & \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\frac{\mathrm{e}^{\frac{-\chi^2}{2}}}{\left(\sqrt{2\pi}\right)^n} \prod_{i=1}^n \mathrm{d}z_i\right) \\
= & \frac{\mathrm{e}^{\frac{-\chi^2}{2}}}{\left(\sqrt{2\pi}\right)^n} \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\prod_{i=1}^n \mathrm{d}z_i\right)
\end{aligned}$$

$$\begin{aligned}
& \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\prod_{i=1}^n \mathrm{d}z_i\right) \\
\overset{\begin{cases}\chi^2 > 0 \\ 6.4.1\end{cases}}{=} & \left[A_n\, \mathrm{d}r\right]_{r=\sqrt{\chi^2}} \\
= & \left[\frac{\left(\sqrt{\pi}\right)^n}{\frac{1}{2}\Gamma\left(\frac{n}{2}\right)} r^{n-1}\, \mathrm{d}r\right]_{r=\sqrt{\chi^2}} \\
= & \frac{\left(\sqrt{\pi}\right)^n}{\frac{1}{2}\Gamma\left(\frac{n}{2}\right)} \left(\sqrt{\chi^2}\right)^{n-1} \mathrm{d}\sqrt{\chi^2} \\
= & \frac{2\left(\sqrt{\pi}\right)^n}{\Gamma\left(\frac{n}{2}\right)} \left(\chi^2\right)^{\frac{n-1}{2}} \cdot \frac{1}{2\sqrt{\chi^2}} \mathrm{d}\chi^2 \\
= & \frac{\left(\sqrt{\pi}\right)^n}{\Gamma\left(\frac{n}{2}\right)} \left(\chi^2\right)^{\frac{n}{2}-1} \mathrm{d}\chi^2
\end{aligned}$$

$$\begin{aligned}
f_{X^2}\left(\chi^2\right) \mathrm{d}\chi^2 &= \frac{\mathrm{e}^{\frac{-\chi^2}{2}}}{\left(\sqrt{2\pi}\right)^n} \sum_{\chi^2 = \sum_{i=1}^n z_i^2} \left(\prod_{i=1}^n \mathrm{d}z_i\right) \\
&= \frac{\mathrm{e}^{\frac{-\chi^2}{2}}}{\left(\sqrt{2\pi}\right)^n} \frac{\left(\sqrt{\pi}\right)^n}{\Gamma\left(\frac{n}{2}\right)} \left(\chi^2\right)^{\frac{n}{2}-1} \mathrm{d}\chi^2 \\
&= \frac{\left(\chi^2\right)^{\frac{n}{2}-1} \mathrm{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \mathrm{d}\chi^2
\end{aligned}$$

□

6.4.2 χ^2分配之極限分配

6.4.3 χ^2分配與其它機率分配之關係

Theorem 6.4.3. $X \sim \mathsf{n}(0,1) = \mathsf{n}(\mu,\sigma)\,|_{\begin{cases}\mu=0\\ \sigma=1\end{cases}} \Rightarrow X^2 \sim \chi^2(1) = \chi^2(n)\,|_{n=1}$

Proof.

$$f_X(x) = \mathsf{n}(x;0,1) = \mathsf{n}(x;\mu,\sigma)\,|_{\begin{cases}\mu=0\\ \sigma=1\end{cases}} = \frac{\mathsf{e}^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}\,|_{\begin{cases}\mu=0\\ \sigma=1\end{cases}} = \frac{\mathsf{e}^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$$

$$\begin{aligned}
f_{X^2}(x^2) &= \sum_{x\in\left\{x|x=\sqrt{x^2}\right\}\cup\left\{x|x=-\sqrt{x^2}\right\}} \left(f_X(x)\cdot\left|\frac{\mathsf{d}x}{\mathsf{d}x^2}\right|\right) \\
&\overset{x\neq 0}{=} 2\left(f_X(x)\cdot\frac{\mathsf{d}\sqrt{x^2}}{\mathsf{d}x^2}\right) = 2\cdot f_X(x)\cdot\frac{1}{2}\left(x^2\right)^{\frac{1}{2}-1} \\
&= f_X(x)\cdot\left(x^2\right)^{\frac{-1}{2}} = \frac{\left(x^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-x^2}{2}}}{\sqrt{2\pi}}
\end{aligned}$$

$$f_{X^2}(x^2) = \begin{cases}\frac{\left(x^2\right)^{\frac{-1}{2}}\mathsf{e}^{\frac{-x^2}{2}}}{\sqrt{2\pi}} \overset{\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}}{=} \frac{\left(x^2\right)^{\frac{n}{2}-1}\mathsf{e}^{\frac{-x^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\,|_{n=1} & x^2>0\\ 0 & x^2\leq 0\end{cases}$$

$$f_{X^2}(x^2) = f_{X^2}(\chi^2;n)\,|_{\begin{cases}\chi^2=x^2\\ n=1\end{cases}} \Rightarrow X^2 \sim \chi^2(n)\,|_{n=1} = \chi^2(1)$$

□

Theorem 6.4.4. $\begin{cases}X_1,X_2,\ldots,X_n \text{ are independent} & [1]\\ \forall i\in\mathbb{N}\cap[1,n]\left(X_i\sim\mathsf{n}(0,1)=\mathsf{n}(\mu,\sigma)\,|_{\begin{cases}\mu=0\\ \sigma=1\end{cases}}\right) & [2]\end{cases} \Rightarrow \left(\sum_{i=1}^n (X_i)^2\right)\sim\chi^2(n)$

Proof.

$$\begin{aligned}
[2] &\overset{6.4.3}{\Rightarrow} (X_i)^2 \sim \chi^2(1) \\
&\overset{6.4.1}{\Rightarrow} \mathsf{E}\left(\mathsf{e}^{\xi\cdot(X_i)^2}\right) = (1-2\xi)^{\frac{-n}{2}}\,|_{n=1} = (1-2\xi)^{\frac{-1}{2}}
\end{aligned} \tag{6.4.2}$$

$$\begin{aligned}
\mathsf{M}(\xi) &= \mathsf{E}\left(\mathsf{e}^{\xi\sum_{i=1}^n(X_i)^2}\right) = \mathsf{E}\left(\mathsf{e}^{\sum_{i=1}^n\xi\cdot(X_i)^2}\right) = \mathsf{E}\left(\prod_{i=1}^n\mathsf{e}^{\xi\cdot(X_i)^2}\right) \\
&\overset{[1]}{=} \prod_{i=1}^n\mathsf{E}\left(\mathsf{e}^{\xi\cdot(X_i)^2}\right) \\
&\overset{6.4.2}{=} \prod_{i=1}^n(1-2\xi)^{\frac{-1}{2}} = \left((1-2\xi)^{\frac{-1}{2}}\right)^n = (1-2\xi)^{\frac{-n}{2}}
\end{aligned}$$

$$f_{\sum_{i=1}^n(X_i)^2}\left(\sum_{i=1}^n(x_i)^2\right) = f_{\sum_{i=1}^n(X_i)^2}(\chi^2;n)\,|_{\chi^2=\sum_{i=1}^n(x_i)^2} \Rightarrow \left(\sum_{i=1}^n(X_i)^2\right)\sim\chi^2(n)$$

□

Corollary 6.4.1. $\begin{cases}X_1,X_2,\ldots,X_n \text{ are independent}\\ \forall i\in\mathbb{N}\cap[1,n]\left(X_i\sim\mathsf{n}(\mu,\sigma)\right)\end{cases} \Rightarrow \left(\sum_{i=1}^n\left(\frac{X_i-\mu}{\sigma}\right)^2\right)\sim\chi^2(n)$

6.4.4 χ^2分配之加法性

Theorem 6.4.5. $\begin{cases} X_1, X_2 \text{ are independent} & [1] \\ \begin{cases} X_1 \sim \chi^2(n_1) \\ X_2 \sim \chi^2(n_2) \end{cases} \end{cases} \Rightarrow (X_1 \pm X_2) \sim \chi^2(n_1 \pm n_2)$

Proof.

$$\begin{aligned} \mathrm{M}(\xi) &= \mathrm{E}\left(\mathrm{e}^{\xi(X_1 \pm X_2)}\right) = \mathrm{E}\left(\mathrm{e}^{\xi X_1 \pm \xi X_2}\right) = \mathrm{E}\left(\mathrm{e}^{\xi X_1}\mathrm{e}^{\pm\xi X_2}\right) \\ &\overset{[1]}{=} \mathrm{E}\left(\mathrm{e}^{\xi X_1}\right)\mathrm{E}\left(\mathrm{e}^{\pm\xi X_2}\right) = (1-2\xi)^{\frac{-n_1}{2}}(1-2\xi)^{\frac{\mp n_2}{2}} \\ &= (1-2\xi)^{\frac{-(n_1 \pm n_2)}{2}} \end{aligned}$$

$$f_{X_1 \pm X_2}(x_1 \pm x_2) = f_{X_1 \pm X_2}\left(\chi^2; n\right) \Big|_{\begin{cases} \chi^2 = x_1 \pm x_2 \\ n = n_1 \pm n_2 \end{cases}} \Rightarrow (X_1 \pm X_2) \sim \chi^2(n_1 \pm n_2)$$

□

Corollary 6.4.2. $\begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \chi^2(n_i) \end{cases} \Rightarrow \left(\sum_{i=1}^n X_i\right) \sim \chi^2\left(\sum_{i=1}^n n_i\right)$

Theorem 6.4.6. sampling from normally-distributed population

$$\begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathrm{n}(\mu, \sigma) \end{cases} \Rightarrow \sum_{i=1}^n X_i \sim \mathrm{n}\left(n\mu, \sigma\sqrt{n}\right) \Rightarrow \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \sim \mathrm{n}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Theorem 6.4.7. $\begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathrm{n}(\mu, \sigma) \\ S = \sqrt{\frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}\right)^2} \end{cases} \Rightarrow \begin{cases} \left(\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2\right) = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2} \sim \chi^2(n) & [1] \\ \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{\sigma^2} \sim \chi^2(n-1) \end{cases}$

Proof.

$$6.4.1 \Leftrightarrow [1]$$

$$\begin{aligned} \sum_{i=1}^n \left(X_i - \bar{X}\right)^2 &= \sum_{i=1}^n \left((X_i - \mu) - (\bar{X} - \mu)\right)^2 \\ &= \sum_{i=1}^n \left((X_i - \mu)^2 - 2(X_i - \mu)\left(\bar{X} - \mu\right) + \left(\bar{X} - \mu\right)^2\right) \\ &= \sum_{i=1}^n (X_i - \mu)^2 - 2\left(\bar{X} - \mu\right)\sum_{i=1}^n (X_i - \mu) + \sum_{i=1}^n \left(\bar{X} - \mu\right)^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - 2\left(\bar{X} - \mu\right) \cdot n\left(\bar{X} - \mu\right) + n\left(\bar{X} - \mu\right)^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - n\left(\bar{X} - \mu\right)^2 \end{aligned}$$

$$\sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 - \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 \tag{6.4.3}$$

$$\begin{aligned} \begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathrm{n}(\mu, \sigma) \end{cases} &\Rightarrow \bar{X} \sim \mathrm{n}\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \\ &\Rightarrow \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathrm{n}(0, 1) \\ &\Rightarrow \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 \sim \chi^2(1) \end{aligned}$$

$$\begin{cases} \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n) & [1] \\ \left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 \sim \chi^2(1) & [2] \end{cases}$$

$$\overset{6.4.5}{\underset{6.4.3}{\Rightarrow}} \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 = [1]-[2] \sim \chi^2(n-1)$$

□

Example 6.4.1. 例5

$$\begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathsf{n}(\mu,\sigma) \end{cases} \Rightarrow \begin{cases} \mathsf{E}\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})}{n}\right) = \frac{n-1}{n}\sigma^2 \\ \mathsf{V}\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})}{n}\right) = \frac{2(n-1)}{n^2}\sigma^4 \end{cases}$$

Proof.

$$\begin{aligned} \begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathsf{n}(\mu,\sigma) \end{cases} &\Rightarrow \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1) \\ &\Rightarrow \mathsf{E}\left(\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2\right) = n-1 \\ &\Rightarrow \mathsf{E}\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right) = (n-1)\sigma^2 \\ &\Rightarrow \mathsf{E}\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})}{n}\right) = \frac{n-1}{n}\sigma^2 \end{aligned}$$

$$\begin{aligned} &\Rightarrow \mathsf{V}\left(\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2\right) = 2(n-1) \\ &\Rightarrow \mathsf{V}\left(\frac{\sigma^2}{n}\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2\right) = \left(\frac{\sigma^2}{n}\right)^2 \cdot 2(n-1) \\ &\Rightarrow \mathsf{V}\left(\frac{\sum_{i=1}^{n}(X_i-\bar{X})}{n}\right) = \frac{2(n-1)}{n^2}\sigma^4 \end{aligned}$$

□

6.4.5 t分配

Definition 6.4.8. t distribution

$$\begin{cases} Z,\Sigma^2 \text{ are independent} \\ Z \sim \mathsf{n}(0,1) \\ \Sigma^2 \sim \chi^2(n) \end{cases} \Rightarrow T = \frac{Z}{\sqrt{\frac{\Sigma^2}{n}}} \sim \mathsf{t}(n)$$

Theorem 6.4.8. $T \sim \mathsf{t}(n) \Rightarrow f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{-2}}$

Proof.

$$\begin{aligned} &\begin{cases} t = t(z,\chi^2) = \frac{z}{\sqrt{\frac{\chi^2}{n}}} \\ \chi^2 = \chi^2 \end{cases} \\ \Leftrightarrow &\begin{cases} z = z(t,\chi^2) = t\sqrt{\frac{\chi^2}{n}} \\ \chi^2 = \chi^2 \end{cases} \end{aligned}$$

$$\begin{aligned} f_{Z,\Sigma^2}\left(z,\chi^2\right) &\overset{\text{independence}}{=} f_Z(z)\, f_{\Sigma^2}\left(\chi^2\right) \\ &= \frac{\mathrm{e}^{\frac{-z^2}{2}}}{\sqrt{2\pi}} \cdot \frac{\left(\chi^2\right)^{\frac{n}{2}-1}\mathrm{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \end{aligned}$$

$$\begin{aligned} f_{T,\Sigma^2}\left(t,\chi^2\right) &= \sum f_{Z,\Sigma^2}\left(z\left(t,\chi^2\right),\chi^2\right)\left|\frac{\partial z}{\partial t}\right| \\ &= \frac{\mathrm{e}^{\left(\frac{-1}{2}\left(t\sqrt{\frac{\chi^2}{n}}\right)^2\right)}}{\sqrt{2\pi}} \cdot \frac{\left(\chi^2\right)^{\frac{n}{2}-1}\mathrm{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left|\sqrt{\frac{\chi^2}{n}}\right| \\ &= \frac{\mathrm{e}^{\left(\frac{-t^2}{2}\frac{\chi^2}{n}\right)}}{\sqrt{2\pi}} \cdot \frac{\left(\chi^2\right)^{\frac{n}{2}-1}\mathrm{e}^{\frac{-\chi^2}{2}}}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} \cdot \sqrt{\frac{\chi^2}{n}} \\ &= \frac{\left(\chi^2\right)^{\frac{n-1}{2}}\mathrm{e}^{\frac{1+\frac{t^2}{n}}{-2}\chi^2}}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \end{aligned}$$

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{T,\Sigma^2}\left(t,\chi^2\right)\mathrm{d}\chi^2 \\ &= \int_0^\infty \frac{\left(\chi^2\right)^{\frac{n-1}{2}}\mathrm{e}^{\frac{1+\frac{t^2}{n}}{-2}\chi^2}}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\mathrm{d}\chi^2 \\ &= \frac{\int_0^\infty \left(\chi^2\right)^{\frac{n-1}{2}}\mathrm{e}^{\frac{1+\frac{t^2}{n}}{-2}\chi^2}\mathrm{d}\chi^2}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \\ &= \frac{\int_0^\infty \left(\chi^2\right)^{\frac{n+1}{2}-1}\mathrm{e}^{\frac{1+\frac{t^2}{n}}{-2}\chi^2}\mathrm{d}\chi^2}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \\ &\overset{6.4.1}{=} \frac{\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left(\frac{1+\frac{t^2}{n}}{2}\right)^{\frac{n+1}{2}}}}{2^{\frac{n+1}{2}}\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{-2}} \end{aligned}$$

□

Theorem 6.4.9. $\begin{cases} \overset{n}{\underset{i=1}{,}} X_i \text{ are independent} \\ X_i \sim \mathsf{n}(\mu,\sigma) \\ \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i \\ S = \sqrt{\frac{1}{n-1}\sum_{i=1}^n\left(X_i-\bar{X}\right)^2} \end{cases} \Rightarrow \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \mathsf{t}(n-1)$

Proof.

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathsf{n}(0,1)$$

$$\sigma = \sqrt{\sigma^2}$$

$$\begin{aligned} \sqrt{\frac{S^2}{\sigma^2}} &= \sqrt{\frac{S^2(n-1)}{\sigma^2(n-1)}} \\ &= \sqrt{\frac{\frac{S^2(n-1)}{\sigma^2}}{n-1}} = \sqrt{\frac{\frac{\sum_{i=1}^n\left(X_i-\bar{X}\right)^2}{\sigma^2}}{n-1}} \\ &= \sqrt{\frac{\sum_{i=1}^n\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}{n-1}} \end{aligned}$$

$$\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \overset{6.4.7}{\sim} \chi^2(n-1)$$

$$\begin{aligned}\frac{\bar{X}-\mu}{S/\sqrt{n}} &= \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\frac{\sigma}{S}\\ &= \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{S^2}{\sigma^2}}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}{n-1}}}\end{aligned}$$

$$\begin{cases}\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2}{n-1}}}\\ \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathsf{n}(0,1)\\ \sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)\end{cases} \Rightarrow \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim \mathsf{t}(n-1)$$

□

6.4.6 F分配

Definition 6.4.9. F distribution

$$\begin{cases}\Sigma_1^2, \Sigma_2^2 \text{ are independent}\\ \Sigma_1^2 \sim \chi^2(n_1)\\ \Sigma_2^2 \sim \chi^2(n_2)\end{cases} \Rightarrow F = \frac{\frac{\Sigma_1^2}{n_1}}{\frac{\Sigma_2^2}{n_2}} = \frac{\Sigma_1^2/n_1}{\Sigma_2^2/n_2} \sim \mathsf{F}(n_1, n_2)$$

Theorem 6.4.10. $F \sim \mathsf{F}(n_1,n_2) \Rightarrow f_F(\mathsf{f}) = \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(1+\frac{n_1}{n_2}\mathsf{f}\right)^{\frac{n_1+n_2}{-2}}\mathsf{f}^{\frac{n_1}{2}-1}$

Proof.

$$\begin{aligned}&\begin{cases}\mathsf{f} = \mathsf{f}\left(\chi_1^2,\chi_2^2\right) = \frac{\chi_1^2/n_1}{\chi_2^2/n_2}\\ \chi_2^2 = \chi_2^2\end{cases}\\ \Leftrightarrow &\begin{cases}\chi_1^2 = \chi_1^2\left(\mathsf{f},\chi_2^2\right) = \frac{n_1}{n_2}\chi_2^2\mathsf{f}\\ \chi_2^2 = \chi_2^2\end{cases}\end{aligned}$$

$$\begin{aligned}f_{\Sigma_1^2,\Sigma_2^2}\left(\chi_1^2,\chi_2^2\right) &\overset{\text{independence}}{=} f_{\Sigma_1^2}\left(\chi_1^2\right)f_{\Sigma_2^2}\left(\chi_2^2\right)\\ &= \frac{\left(\chi_1^2\right)^{\frac{n_1}{2}-1}\mathsf{e}^{\frac{-\chi_1^2}{2}}}{2^{\frac{n_1}{2}}\Gamma\left(\frac{n_1}{2}\right)} \cdot \frac{\left(\chi_2^2\right)^{\frac{n_2}{2}-1}\mathsf{e}^{\frac{-\chi_2^2}{2}}}{2^{\frac{n_2}{2}}\Gamma\left(\frac{n_2}{2}\right)}\end{aligned}$$

$$\begin{aligned}f_{F,\Sigma_2^2}\left(\mathsf{f},\chi_2^2\right) &= \sum f_{\Sigma_1^2,\Sigma_2^2}\left(\chi_1^2\left(\mathsf{f},\chi_2^2\right),\chi^2\right)\left|\frac{\partial\chi_1^2}{\partial\mathsf{f}}\right|\\ &= \frac{\left(\chi_1^2\right)^{\frac{n_1}{2}-1}\mathsf{e}^{\frac{-\chi_1^2}{2}}}{2^{\frac{n_1}{2}}\Gamma\left(\frac{n_1}{2}\right)} \cdot \frac{\left(\chi_2^2\right)^{\frac{n_2}{2}-1}\mathsf{e}^{\frac{-\chi_2^2}{2}}}{2^{\frac{n_2}{2}}\Gamma\left(\frac{n_2}{2}\right)}\left|\frac{n_1}{n_2}\chi_2^2\right|\\ &= \frac{\left(\chi_1^2\right)^{\frac{n_1}{2}-1}\left(\chi_2^2\right)^{\frac{n_2}{2}-1}\mathsf{e}^{\frac{\chi_1^2+\chi_2^2}{-2}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)} \cdot \frac{n_1}{n_2}\chi_2^2\\ &= \frac{n_1}{n_2}\frac{\left(\chi_1^2\right)^{\frac{n_1}{2}-1}\left(\chi_2^2\right)^{\frac{n_2}{2}}\mathsf{e}^{\frac{\chi_1^2+\chi_2^2}{-2}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\\ &= \frac{n_1}{n_2}\frac{\left(\frac{n_1}{n_2}\chi_2^2\mathsf{f}\right)^{\frac{n_1}{2}-1}\left(\chi_2^2\right)^{\frac{n_2}{2}}\mathsf{e}^{\frac{\frac{n_1}{n_2}\chi_2^2\mathsf{f}+\chi_2^2}{-2}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\\ &= \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}\frac{(\mathsf{f})^{\frac{n_1}{2}-1}\left(\chi_2^2\right)^{\frac{n_1+n_2}{2}-1}\mathsf{e}^{\frac{\chi_2^2\left(1+\frac{n_1}{n_2}\mathsf{f}\right)}{-2}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\end{aligned}$$

$$\begin{aligned}
f_F(\mathsf{f}) &= \int_0^\infty f_{F,\Sigma_2^2}\left(\mathsf{f},\chi_2^2\right)\,\mathsf{d}\chi_2^2 \\
&= \int_0^\infty \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{(\mathsf{f})^{\frac{n_1}{2}-1}\left(\chi_2^2\right)^{\frac{n_1+n_2}{2}-1}\mathsf{e}^{\frac{\chi_2^2\left(1+\frac{n_1}{n_2}\mathsf{f}\right)}{-2}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\,\mathsf{d}\chi_2^2 \\
&= \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{\int_0^\infty \left(\chi_2^2\right)^{\frac{n_1+n_2}{2}-1}\mathsf{e}^{\frac{\chi_2^2\left(1+\frac{n_1}{n_2}\mathsf{f}\right)}{-2}}\,\mathsf{d}\chi_2^2}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\mathsf{f}^{\frac{n_1}{2}-1} \\
&\overset{6.4.1}{=} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\left(\frac{1}{2}\left(1+\frac{n_1}{n_2}\mathsf{f}\right)\right)^{\frac{n_1+n_2}{2}}}}{2^{\frac{n_1+n_2}{2}}\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\mathsf{f}^{\frac{n_1}{2}-1} \\
&= \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right)\Gamma\left(\frac{n_2}{2}\right)}\left(1+\frac{n_1}{n_2}\mathsf{f}\right)^{\frac{n_1+n_2}{-2}}\mathsf{f}^{\frac{n_1}{2}-1}
\end{aligned}$$

□

Theorem 6.4.11. $F_{1-\alpha}(n_2,n_1) = \frac{1}{F_\alpha(n_1,n_2)}$

$$F \sim \mathsf{F}(n_1,n_2) \Rightarrow \frac{1}{F} \sim \mathsf{F}(n_2,n_1)\,[0]$$

$$\begin{aligned}
\alpha &= \mathsf{P}\left(F > F(\alpha;n_1,n_2)\right) = \mathsf{P}\left(F > F_\alpha(n_1,n_2)\right) \\
&= \mathsf{P}\left(\frac{1}{F} < \frac{1}{F_\alpha(n_1,n_2)}\right) \\
&\Downarrow \\
\mathsf{P}\left(\frac{1}{F} \geq F_{1-\alpha}(n_2,n_1)\right) \overset{[0]}{=} 1-\alpha &= 1-\mathsf{P}\left(\frac{1}{F} < \frac{1}{F_\alpha(n_1,n_2)}\right) \\
&= \mathsf{P}\left(\frac{1}{F} \geq \frac{1}{F_\alpha(n_1,n_2)}\right) \\
&\Downarrow \\
F_{1-\alpha}(n_2,n_1) &= \frac{1}{F_\alpha(n_1,n_2)}
\end{aligned}$$

Theorem 6.4.12. $\begin{cases} \overset{2}{\underset{i=1}{,}}\overset{n_i}{\underset{j=1}{,}} X_{ij} \text{ are independent} \\ X_{ij} \sim \mathsf{n}(\mu_i,\sigma_i) \\ \bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij} \\ S_i = \sqrt{\frac{1}{n_i-1}\sum_{j=1}^{n_i}\left(X_{ij}-\bar{X}_i\right)^2} \end{cases} \Rightarrow \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim \mathsf{F}(n_1-1,n_2-1)$

Proof.

$$\frac{(n_i-1)S_i^2}{\sigma_i^2} \sim \chi^2(n_i-1) \Rightarrow \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} = \frac{\frac{(n_1-1)S_1^2}{\sigma_1^2}\Big/ n_1-1}{\frac{(n_2-1)S_2^2}{\sigma_2^2}\Big/ n_2-1} \sim \mathsf{F}(n_1-1,n_2-1)$$

□

Chapter 7

估計理論

- statistical inference
  - estimation
  - hypothesis testing

7.1 不偏性與最小變異性

7.1.1 評判推定量之準繩

Definition 7.1.1. estimator and estimate

$$
\begin{aligned}
&\mathop{;}_{i=1}^{n} X_i \text{ is a random sample} \sim \mathsf{d}(\theta) \\
\Leftrightarrow\ &\begin{cases} \mathop{;}_{i=1}^{n} X_i \text{ are independent} \\ \exists!\mathsf{d} = \mathsf{d}(\theta) = \text{arbitrary distribution} \forall i \in \mathbb{N} \cap [1, n] \left(X_i \sim \mathsf{d}(\theta)\right) \\ \theta \text{ is a parameter of } \mathsf{d} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
&\begin{cases} \mathop{;}_{i=1}^{n} X_i \text{ is a random sample} \sim \mathsf{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\mathop{;}_{i=1}^{n} X_i\right) \text{ is to estimate } \theta \\ \hat{\theta} = \hat{\theta}\left(\mathop{;}_{i=1}^{n} (X_i = x_i)\right) = \hat{\theta}\left(\mathop{;}_{i=1}^{n} x_i\right) \end{cases} \\
\Leftrightarrow\ &\begin{cases} \hat{\Theta} \text{ is an estimator of } \theta \\ \hat{\theta} \text{ is an estimate of } \theta \end{cases}
\end{aligned}
$$

7.1.2 不偏性

Definition 7.1.2. unbiased estimator

$$\hat{\Theta} \text{ is an estimator of } \theta \Rightarrow \mathsf{E}\left(\hat{\Theta}\right) - \theta \text{ is the bias between } \hat{\Theta}, \theta$$

$$\begin{cases} \hat{\Theta} \text{ is an estimator of } \theta \\ \mathsf{E}\left(\hat{\Theta}\right) = \theta \end{cases} \Leftrightarrow \hat{\Theta} \text{ is an unbiased estimator of } \theta$$

$$\begin{cases} \hat{\Theta} \text{ is an estimator of } \theta \\ \mathsf{E}\left(\hat{\Theta}\right) \neq \theta \end{cases} \Leftrightarrow \hat{\Theta} \text{ is a biased estimator of } \theta$$

7.1.3 不偏最小變異性

Definition 7.1.3. $\begin{cases} \hat{\Theta}_1, \hat{\Theta}_2 \text{ are unbiased estimators of } \theta \\ \mathsf{V}\left(\hat{\Theta}_1\right) < \mathsf{V}\left(\hat{\Theta}_2\right) \end{cases} \Leftrightarrow \hat{\Theta}_1 \text{ is a better unbiased estimators of } \theta \text{ than } \hat{\Theta}_2$

Definition 7.1.4. unbiased minimum variance estimator = minimum variance unbiased estimator

$$\exists\hat{\Theta}^{*}\text{ is a biased estimator of }\theta\left(\forall\hat{\Theta}\text{ is a biased estimator of }\theta\left(\mathsf{V}\left(\hat{\Theta}\right)\geq\mathsf{V}\left(\hat{\Theta}^{*}\right)\right)\right)$$
$$\Downarrow$$
$$\hat{\Theta}^{*}\text{ is an unbiased minimum variance estimator of }\theta$$

Theorem 7.1.1. $\hat{\Theta}_1^{*},\hat{\Theta}_2^{*}$ are unbiased minimum variance estimators of $\theta\Rightarrow\hat{\Theta}_1^{*}=\hat{\Theta}_2^{*}$

Proof. □

7.1.4 Rao-Cramér不等式

Definition 7.1.5. regular condition = regularity condition

$$\begin{cases}f_X\left(x;\theta\right):X\left(\Omega\right)\to\left(\mathbb{R}^{+}\cup\{0\}\right)\\ \theta\notin X\left(\Omega\right)\\ \theta\in\left(r_l,r_u\right)\subseteq\mathbb{R}\\ \begin{cases}\frac{\partial}{\partial\theta}\sum_{x\in X(\Omega)}f_X\left(x;\theta\right)=\sum_{x\in X(\Omega)}\frac{\partial}{\partial\theta}f_X\left(x;\theta\right)=0 & X\text{ is discrete}\\ \frac{\partial}{\partial\theta}\int_{x\in X(\Omega)}f_X\left(x;\theta\right)\,\mathsf{d}x=\int_{x\in X(\Omega)}\frac{\partial}{\partial\theta}f_X\left(x;\theta\right)\,\mathsf{d}x=0 & X\text{ is continuous}\end{cases}\end{cases}$$
$$\Leftrightarrow\quad f_X\left(x;\theta\right)\text{ is regular}$$

Theorem 7.1.2. Rao-Cramér-Fréchet inequality

$$\begin{cases}\mathop{;}\limits_{i=1}^{n}X_i\text{ is a random sample}\sim\mathsf{d}\left(\theta\right) & [1]\\ f_{X_i}\left(x_i;\theta\right)\text{ is regular} & [2]\\ \hat{\Theta}\text{ is an unbiased estimator of }\theta & [3]\\ L=L\left(\mathop{;}\limits_{i=1}^{n}x_i;\theta\right)=\prod_{i=1}^{n}f_{X_i}\left(x_i;\theta\right) & [4]\\ \boldsymbol{L}=L\left(\mathop{;}\limits_{i=1}^{n}X_i;\theta\right)=\prod_{i=1}^{n}f_{X_i}\left(X_i;\theta\right)\end{cases}$$
$$\Rightarrow\quad\mathsf{V}\left(\hat{\Theta}\right)\geq\frac{1}{}$$

Proof.

$$[4]\Rightarrow\begin{cases}L\geq0\\ \int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}L\prod_{i=1}^{n}\mathsf{d}x_i=1\end{cases}$$

$$\mathsf{E}\left(\hat{\Theta}\right)=\int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}\hat{\theta}\cdot L\prod_{i=1}^{n}\mathsf{d}x_i\overset{[3]}{=}\theta$$

$$\begin{aligned}\frac{\partial\left(\hat{\theta}\cdot L\right)}{\partial\theta}=\frac{\partial\left(\hat{\theta}L\right)}{\partial\theta} &= \frac{\partial\hat{\theta}}{\partial\theta}L+\hat{\theta}\frac{\partial L}{\partial\theta}\\ &\overset{\hat{\theta}\text{ is a statistic}\Rightarrow\frac{\partial\hat{\theta}}{\partial\theta}=0}{=} 0\cdot L+\hat{\theta}\frac{\partial L}{\partial\theta}\\ &= \hat{\theta}\frac{\partial L}{\partial\theta}\end{aligned}\tag{7.1.1}$$

$$\begin{cases}\int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}\frac{\partial L}{\partial\theta}\prod_{i=1}^{n}\mathsf{d}x_i\overset{[2]}{=}\frac{\partial}{\partial\theta}\int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}L\prod_{i=1}^{n}\mathsf{d}x_i\overset{[4]}{=}\frac{\partial}{\partial\theta}1=0 & [5]\\ \int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}\frac{\partial\left(\hat{\theta}\cdot L\right)}{\partial\theta}\prod_{i=1}^{n}\mathsf{d}x_i\overset{[2]}{=}\frac{\partial}{\partial\theta}\int\cdots\int_{\left(\mathop{;}\limits_{i=1}^{n}x_i\right)\in\mathop{\times}\limits_{i=1}^{n}X_i(\Omega)}\hat{\theta}\cdot L\prod_{i=1}^{n}\mathsf{d}x_i\overset{[3]}{=}\frac{\partial}{\partial\theta}\theta=1 & [6]\end{cases}$$

$$
\begin{aligned}
\frac{\partial \ln L}{\partial \theta} &= \frac{\partial \ln L}{\partial L}\frac{\partial L}{\partial \theta} = \frac{1}{L}\frac{\partial L}{\partial \theta} \\
\Leftrightarrow \quad \frac{\partial L}{\partial \theta} &= L\frac{\partial \ln L}{\partial \theta} = \left(\frac{\partial \ln L}{\partial \theta}\right) L = \frac{\partial \ln L}{\partial \theta} \cdot L
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{E}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right) &= \underset{\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \in \mathop{\times}\limits_{i=1}^{n} X_i(\Omega)}{\int \cdots \int} \frac{\partial \ln L}{\partial \theta} \cdot L \prod_{i=1}^{n} \mathrm{d}x_i \\
&= \underset{\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \in \mathop{\times}\limits_{i=1}^{n} X_i(\Omega)}{\int \cdots \int} \frac{\partial L}{\partial \theta} \prod_{i=1}^{n} \mathrm{d}x_i \overset{[5]}{=} 0
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{E}\left(\hat{\Theta}\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right) &= \underset{\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \in \mathop{\times}\limits_{i=1}^{n} X_i(\Omega)}{\int \cdots \int} \hat{\theta}\frac{\partial \ln L}{\partial \theta} \cdot L \prod_{i=1}^{n} \mathrm{d}x_i \\
&= \underset{\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \in \mathop{\times}\limits_{i=1}^{n} X_i(\Omega)}{\int \cdots \int} \hat{\theta}\frac{\partial L}{\partial \theta} \prod_{i=1}^{n} \mathrm{d}x_i \\
&\overset{7.1.1}{=} \underset{\left(\mathop{,}\limits_{i=1}^{n} x_i\right) \in \mathop{\times}\limits_{i=1}^{n} X_i(\Omega)}{\int \cdots \int} \frac{\partial \left(\hat{\theta} L\right)}{\partial \theta} \prod_{i=1}^{n} \mathrm{d}x_i \overset{[6]}{=} 1
\end{aligned}
$$

$$
\begin{aligned}
1 \overset{4.5.1}{\geq} \mathsf{R}\left(\hat{\Theta}, \frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right) &= \frac{\mathsf{V}\left(\hat{\Theta}, \frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}{\sqrt{\mathsf{V}\left(\hat{\Theta}\right)}\sqrt{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}} \\
&= \frac{\mathsf{E}\left(\hat{\Theta}\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right) - \mathsf{E}\left(\hat{\Theta}\right)\mathsf{E}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}{\sqrt{\mathsf{V}\left(\hat{\Theta}\right)}\sqrt{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}} \\
&= \frac{1 - \theta \cdot 0}{\sqrt{\mathsf{V}\left(\hat{\Theta}\right)}\sqrt{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}} = \frac{1}{\sqrt{\mathsf{V}\left(\hat{\Theta}\right)}\sqrt{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}} \\
\sqrt{\mathsf{V}\left(\hat{\Theta}\right)} &\geq \frac{1}{\sqrt{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)}} \\
\mathsf{V}\left(\hat{\Theta}\right) &\geq \frac{1}{\mathsf{V}\left(\frac{\partial \ln \boldsymbol{L}}{\partial \theta}\right)} = \frac{1}{\mathsf{V}\left(\frac{\partial}{\partial \theta} \ln \prod_{i=1}^{n} f_{X_i}\left(X_i; \theta\right)\right)} \\
&= \frac{1}{\mathsf{V}\left(\frac{\partial}{\partial \theta} \sum_{i=1}^{n} \ln f_{X_i}\left(X_i; \theta\right)\right)} = \frac{1}{\mathsf{V}\left(\sum_{i=1}^{n} \frac{\partial}{\partial \theta} \ln f_{X_i}\left(X_i; \theta\right)\right)}
\end{aligned}
$$

$$
\begin{aligned}
& \mathsf{V}\left(\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\right) \\
\overset{4.5.4}{=} \; & \sum_{i=1}^{n}\mathsf{V}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\right) \\
& +2\sum_{1\le i<j\le n}\mathsf{V}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right),\frac{\partial}{\partial\theta}\ln f_{X_j}\left(X_j;\theta\right)\right) \\
= \; & \sum_{i=1}^{n}\left(\mathsf{E}\left(\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\right)^2\right)-\left(\mathsf{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\right)\right)^2\right) \\
& +2\sum_{1\le i<j\le n}\left(\mathsf{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\frac{\partial}{\partial\theta}\ln f_{X_j}\left(X_j;\theta\right)\right)\right. \\
& \left.-\mathsf{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}\left(X_i;\theta\right)\right)\mathsf{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_j}\left(X_j;\theta\right)\right)\right) \\
\overset{[1]}{=} \; &
\end{aligned}
$$

□

Theorem 7.1.3.

Proof.

□

Theorem 7.1.4.

Proof.

□

Definition 7.1.6. estimator efficiency

$$\mathsf{e}\left(\hat{\Theta}\right)=\frac{}{\mathsf{V}\left(\hat{}\right)}$$

7.1.5 MSE與相對有效性

Definition 7.1.7. mean-square error = MSE

$$\mathsf{MSE}\left(\hat{\Theta}\right)=\mathsf{E}\left(\left(\hat{\Theta}-\theta\right)^2\right)$$

- $\begin{cases}\hat{\Theta}_1,\hat{\Theta}_2 \text{ are estimators of } \theta \\ \mathsf{MSE}\left(\hat{\Theta}_1\right)<\mathsf{MSE}\left(\hat{\Theta}_2\right)\end{cases} \Leftrightarrow \hat{\Theta}_1$ is a better estimators of θ than $\hat{\Theta}_2$
- relative efficiency = RE

$$\mathsf{RE}=\mathsf{RE}_{12}=\frac{\mathsf{MSE}\left(\hat{\Theta}_1\right)}{\mathsf{MSE}\left(\hat{\Theta}_2\right)}$$

Theorem 7.1.5. $\mathsf{MSE}\left(\hat{\Theta}\right)=\mathsf{V}\left(\hat{\Theta}\right)+\mathsf{bias}^2$ $\hat{\Theta}$ is an unbiased estimator of θ $\underset{=}{}$ $\mathsf{V}\left(\hat{\Theta}\right)$

Proof.

$$
\begin{aligned}
\mathsf{MSE}\left(\hat{\Theta}\right)=\mathsf{E}\left(\left(\hat{\Theta}-\theta\right)^2\right) &= \mathsf{E}\left(\left(\left(\hat{\Theta}-\mathsf{E}\left(\hat{\Theta}\right)\right)+\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)\right)^2\right) \\
&= \mathsf{E}\left(\left(\hat{\Theta}-\mathsf{E}\left(\hat{\Theta}\right)\right)^2+2\left(\hat{\Theta}-\mathsf{E}\left(\hat{\Theta}\right)\right)\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)+\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)^2\right) \\
&= \mathsf{E}\left(\left(\hat{\Theta}-\mathsf{E}\left(\hat{\Theta}\right)\right)^2\right)+2\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)\mathsf{E}\left(\hat{\Theta}-\mathsf{E}\left(\hat{\Theta}\right)\right)+\mathsf{E}\left(\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)^2\right) \\
&= \mathsf{V}\left(\hat{\Theta}\right)+2\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)\cdot 0+\left(\mathsf{E}\left(\hat{\Theta}\right)-\theta\right)^2 \\
&= \mathsf{V}\left(\hat{\Theta}\right)+(\mathsf{bias})^2=\mathsf{V}\left(\hat{\Theta}\right)+\mathsf{bias}^2
\end{aligned}
$$

□

Corollary 7.1.1. $\mathsf{MSE}\left(\hat{\Theta}\right) \overset{\hat{\Theta}\text{ is an unbiased estimator of }\theta}{=} \mathsf{V}\left(\hat{\Theta}\right)$

Proof.

$$\hat{\Theta}\text{ is an unbiased estimator of }\theta \Rightarrow \mathsf{bias}=0$$

□

7.1.6 習題6-1

7.2 一致性

Fact 7.2.1. continuity of probability function and expectation

- probability function

$$
\begin{aligned}
\lim_{n\to\infty}\mathsf{P}\left(\hat{\Theta}=\hat{\theta}\right) &= \lim_{n\to\infty}\mathsf{P}\left(\hat{\theta}\left(\mathop{;}\limits_{i=1}^{n}X_i\right)=\hat{\theta}\right) \\
&= \mathsf{P}\left(\lim_{n\to\infty}\left(\hat{\theta}\left(\mathop{;}\limits_{i=1}^{n}X_i\right)=\hat{\theta}\right)\right) \\
&= \mathsf{P}\left(\lim_{n\to\infty}\hat{\theta}\left(\mathop{;}\limits_{i=1}^{n}X_i\right)=\hat{\theta}\right) \\
&= \mathsf{P}\left(\hat{\theta}\left(\lim_{n\to\infty}\mathop{;}\limits_{i=1}^{n}X_i\right)=\hat{\theta}\right) \\
&= \mathsf{P}\left(\hat{\theta}\left(\mathop{;}\limits_{i=1}^{\infty}X_i\right)=\hat{\theta}\right)
\end{aligned}
$$

- expectation

$$
\begin{aligned}
\lim_{n\to\infty}\mathsf{E}\left(\hat{\Theta}\right) &= \lim_{n\to\infty}\mathsf{E}\left(\hat{\theta}\left(\mathop{;}\limits_{i=1}^{n}X_i\right)\right) \\
&= \mathsf{E}\left(\lim_{n\to\infty}\hat{\theta}\left(\mathop{;}\limits_{i=1}^{n}X_i\right)\right) \\
&= \mathsf{E}\left(\hat{\theta}\left(\lim_{n\to\infty}\mathop{;}\limits_{i=1}^{n}X_i\right)\right) \\
&= \mathsf{E}\left(\hat{\theta}\left(\mathop{;}\limits_{i=1}^{\infty}X_i\right)\right)
\end{aligned}
$$

Definition 7.2.1. consistent estimator = CE

- simple consistent estimator = SCE

$$\begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathrm{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\overset{n}{\underset{i=1}{;}} X_i\right) \text{ is to estimate } \theta \\ \forall \epsilon > 0 \left(\lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\Theta}-\theta\right| > \epsilon\right) = 0\right) \Leftrightarrow \forall \epsilon > 0 \left(\lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\Theta}-\theta\right| \le \epsilon\right) = 1\right) \end{cases}$$

$$\Leftrightarrow \begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathrm{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\overset{n}{\underset{i=1}{;}} X_i\right) \text{ is to estimate } \theta \\ \lim_{n\to\infty} \mathrm{P}\left(\hat{\Theta} \ne \theta\right) = 0 \Leftrightarrow \lim_{n\to\infty} \mathrm{P}\left(\hat{\Theta} = \theta\right) = 1 \end{cases}$$

$$\overset{7.2.1}{\Leftrightarrow} \begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathrm{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\overset{n}{\underset{i=1}{;}} X_i\right) \text{ is to estimate } \theta \\ \mathrm{P}\left(\lim_{n\to\infty} \hat{\Theta} \ne \theta\right) = 0 \Leftrightarrow \mathrm{P}\left(\lim_{n\to\infty} \hat{\Theta} = \theta\right) = 1 \end{cases}$$

$$\Leftrightarrow \quad \hat{\Theta} \text{ is a SCE of } \theta$$

$$\Leftrightarrow \quad \hat{\Theta} \xrightarrow[n\to\infty]{\mathrm{P}} \theta$$

- squared-error consistent estimator = SECE

$$\begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathrm{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\overset{n}{\underset{i=1}{;}} X_i\right) \text{ is to estimate } \theta \\ \lim_{n\to\infty} \mathrm{E}\left(\left(\hat{\Theta}-\theta\right)^2\right) = 0 \end{cases}$$

$$\overset{7.2.1}{\Leftrightarrow} \begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathrm{d}(\theta) \\ \hat{\Theta} = \hat{\theta}\left(\overset{n}{\underset{i=1}{;}} X_i\right) \text{ is to estimate } \theta \\ 0 = \mathrm{E}\left(\lim_{n\to\infty}\left(\hat{\Theta}-\theta\right)^2\right) = \mathrm{E}\left(\left(\lim_{n\to\infty}\hat{\Theta}-\theta\right)^2\right) \end{cases}$$

$$\Leftrightarrow \quad \hat{\Theta} \text{ is a SECE of } \theta$$

Theorem 7.2.1.

$$\begin{cases} \hat{\Theta} \text{ is an unbiased estimator of } \theta & [1] \\ \lim_{n\to\infty} \mathrm{V}\left(\hat{\Theta}\right) = 0 & [2] \end{cases} \Rightarrow \hat{\Theta} \text{ is a SCE of } \theta$$

Proof.

$$\begin{aligned} \mathrm{P}\left(\left|\hat{\Theta}-\theta\right| > \epsilon\right) &\overset{[1]}{=} \mathrm{P}\left(\left|\hat{\Theta}-\mathrm{E}\left(\hat{\Theta}\right)\right| > \epsilon\right) \\ &= \mathrm{P}\left(\left|\hat{\Theta}-\mathrm{E}\left(\hat{\Theta}\right)\right| > \frac{\epsilon}{\sqrt{\mathrm{V}\left(\hat{\Theta}\right)}}\sqrt{\mathrm{V}\left(\hat{\Theta}\right)}\right) \\ &\overset{\begin{cases} k = \frac{\epsilon}{\sqrt{\mathrm{V}(\hat{\Theta})}} > 0 \\ 3.4.10 \end{cases}}{\le} \frac{1}{k^2} = \frac{\mathrm{V}\left(\hat{\Theta}\right)}{\epsilon^2} \\ \lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\Theta}-\theta\right| > \epsilon\right) &\le \lim_{n\to\infty} \frac{\mathrm{V}\left(\hat{\Theta}\right)}{\epsilon^2} = \frac{\lim_{n\to\infty}\mathrm{V}\left(\hat{\Theta}\right)}{\epsilon^2} \overset{[2]}{=} \frac{0}{\epsilon^2} = 0 \\ &\Downarrow \\ \hat{\Theta} &\text{ is } \text{a SCE of } \theta \end{aligned}$$

□

Example 7.2.1. 例1

Example 7.2.2. 例2

$$\begin{cases} \overset{n}{\underset{i=1}{;}} X_i \text{ is a random sample} \sim \mathsf{n}(\mu, \sigma) & [1] \\ \bar{X} = \dfrac{\sum_{i=1}^{n} X_i}{n} \\ S^2 = \dfrac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{n-1} \end{cases} \Rightarrow S^2 \text{ is a SCE of } \sigma^2$$

Proof.

$$\mathsf{E}\left(S^2\right) \overset{6.1.2}{=} \sigma^2 \Rightarrow S^2 \text{ is an unbiased estimator of } \sigma^2$$

$$\begin{aligned} \lim_{n\to\infty} \mathsf{V}\left(S^2\right) &= \lim_{n\to\infty} \mathsf{V}\left(\frac{\sigma^2}{n-1} \frac{(n-1)S^2}{\sigma^2}\right) \\ &= \lim_{n\to\infty} \left(\frac{\sigma^2}{n-1}\right)^2 \mathsf{V}\left(\frac{(n-1)S^2}{\sigma^2}\right) \\ \begin{cases} [1] \Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \\ 6.4.1 \end{cases} \\ &= \lim_{n\to\infty} \left(\frac{\sigma^2}{n-1}\right)^2 \cdot 2(n-1) \\ &= \lim_{n\to\infty} \frac{2\sigma^4}{n-1} = 0 \end{aligned}$$

$$\begin{cases} S^2 \text{ is an unbiased estimator of } \sigma^2 \\ \lim_{n\to\infty} \mathsf{V}\left(S^2\right) = 0 \end{cases} \overset{7.2.1}{\Rightarrow} S^2 \text{ is a SCE of } \sigma^2$$

□

Theorem 7.2.2.

$$\begin{cases} \lim_{n\to\infty} \mathsf{E}\left(\hat{\Theta}\right) = \theta & [1] \\ \lim_{n\to\infty} \mathsf{V}\left(\hat{\Theta}\right) = 0 & [2] \end{cases} \Rightarrow \hat{\Theta} \text{ is a SCE of } \theta$$

Proof. □

- 書中證明難以解釋，再想想
-

$$\begin{aligned} \lim_{n\to\infty} \mathsf{P}\left(\left|\hat{\Theta} - \theta\right| > \epsilon\right) &= \mathsf{P}\left(\lim_{n\to\infty} \left|\hat{\Theta} - \theta\right| > \epsilon\right) \\ &= \mathsf{P}\left(\lim_{n\to\infty} \left|\hat{\Theta} - \lim_{n\to\infty} \mathsf{E}\left(\hat{\Theta}\right)\right| > \epsilon\right) \\ &= \mathsf{P}\left(\lim_{n\to\infty} \left|\hat{\Theta} - \mathsf{E}\left(\hat{\Theta}\right)\right| > \epsilon\right) \\ &= \lim_{n\to\infty} \mathsf{P}\left(\left|\hat{\Theta} - \mathsf{E}\left(\hat{\Theta}\right)\right| > \epsilon\right) \\ &= \lim_{n\to\infty} \mathsf{P}\left(\left|\hat{\Theta} - \mathsf{E}\left(\hat{\Theta}\right)\right| > \frac{\epsilon}{\sqrt{\mathsf{V}\left(\hat{\Theta}\right)}} \sqrt{\mathsf{V}\left(\hat{\Theta}\right)}\right) \\ \begin{cases} k = \frac{\epsilon}{\sqrt{\mathsf{V}(\hat{\Theta})}} > 0 \\ 3.4.10 \end{cases} \\ &\le \frac{1}{k^2} = \frac{\mathsf{V}\left(\hat{\Theta}\right)}{\epsilon^2} \\ \lim_{n\to\infty} \mathsf{P}\left(\left|\hat{\Theta} - \theta\right| > \epsilon\right) &\le \lim_{n\to\infty} \frac{\mathsf{V}\left(\hat{\Theta}\right)}{\epsilon^2} = \frac{\lim_{n\to\infty} \mathsf{V}\left(\hat{\Theta}\right)}{\epsilon^2} \overset{[2]}{=} \frac{0}{\epsilon^2} = 0 \\ &\Downarrow \\ \hat{\Theta} \quad &\text{is} \quad \text{a SCE of } \theta \end{aligned}$$

7.2.1 機率收斂

Definition 7.2.2. convergence in probability

$$
\begin{aligned}
&\begin{cases}\{X_n\}=\{X_n\}_{n=1}^{\infty} \text{ is a sequence of RV}\\ \forall\epsilon>0\left(\lim_{n\to\infty}\mathsf{P}\left(|X_n-X|>\epsilon\right)=0\right)\Leftrightarrow\forall\epsilon>0\left(\lim_{n\to\infty}\mathsf{P}\left(|X_n-X|\le\epsilon\right)=1\right)\end{cases}\\
\Leftrightarrow\quad&\{X_n\} \text{ converges to } X \text{ in probability}\\
\Leftrightarrow\quad&X_n\xrightarrow[n\to\infty]{\mathsf{P}}X\Leftrightarrow X_n\xrightarrow{\mathsf{P}}X
\end{aligned}
$$

Theorem 7.2.3.

- $\begin{cases}X_n\xrightarrow[n\to\infty]{\mathsf{P}}X\\ g:\mathbb{R}\to\mathbb{R}\\ g\text{ is continuous}\end{cases}\Rightarrow g\left(X_n\right)\xrightarrow[n\to\infty]{\mathsf{P}}g\left(X\right)$
- $\begin{cases}\forall i\in I\left(X_{in}\xrightarrow[n\to\infty]{\mathsf{P}}X_i\right)\\ g:\mathbb{R}^I\to\mathbb{R}\\ g\text{ is continuous}\end{cases}\Rightarrow g\left(\underset{i\in I}{,}X_{in}\right)\xrightarrow[n\to\infty]{\mathsf{P}}g\left(\underset{i\in I}{,}X_i\right)$

Corollary 7.2.1. $\forall i\in I\left(X_{in}\xrightarrow[n\to\infty]{\mathsf{P}}X_i\right)\Rightarrow\begin{cases}\sum_{i\in I}c_iX_{in}\xrightarrow[n\to\infty]{\mathsf{P}}\sum_{i\in I}c_iX_i\\ \prod_{i\in I}c_iX_{in}\xrightarrow[n\to\infty]{\mathsf{P}}\prod_{i\in I}c_iX_i\\ \frac{X_{1,n}}{X_{2,n}}\xrightarrow[n\to\infty]{\mathsf{P}}\frac{X_1}{X_2} & \begin{cases}I\subseteq\mathbb{N}\\ X_{2,n}X_2\neq0\end{cases}\end{cases}$

- method of searching for estimator
  - method of moment = principal of substitution
  - maximum likelihood method
  - Bayesian method
  - least square method

7.3 動差推定量

7.4 最概法

7.5 貝氏推定量與大中取小推定量

Chapter 8

估計理論之進一步

Chapter 9

統計假設檢定

Chapter 10

線性模式導論

10.1 最小平方法簡單線性迴歸模式

Definition 10.1.1. simple linear regression equation

$$\hat{Y} = a + bX$$

- independent variable = control variable = regressor

$$X$$

- dependent variable = responsor

$$\hat{Y}$$

$$\hat{Y} = a + \frac{b}{X} \overset{X' = \frac{1}{X}}{=} a + bX'$$

$$\hat{Y} = ab^X \overset{\begin{cases} \hat{Y}' = \ln \hat{Y} \\ a' = \ln a \\ b' = \ln b \end{cases}}{\Rightarrow} \hat{Y}' = a' + b'X$$

- regression
  - univariate regression
  - multivariate regression

10.1.1 簡單線性迴歸

10.1.2 最小平方法

10.1.3 母數之推定

10.1.4 簡單迴歸模式

10.1.5 預測區間之估計

10.2 Gauss-Markov定理

10.3 隨機矩陣之性質

10.4 二次形式與Cochran定理

10.5 一般線性模式