

Probability & Statistics (1)

# Introduction to Probability and Statistics

**Asst. Prof. Chan, Chun-Hsiang**

*Master program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan*

*Undergraduate program in Intelligent Computing and Big Data, Chung Yuan Christian University, Taoyuan, Taiwan*

*Undergraduate program in Applied Artificial Intelligence, , Chung Yuan Christian University, Taoyuan, Taiwan*

# ASBMR 2022



SEPTEMBER 9-12, 2022  
AUSTIN, TX, UNITED STATES

**+ONLINE EXPERIENCE**

# About Dr. Chan

## 目前在職

- 專任助理教授 | 智慧運算與大數據學士學位學程
- 專任助理教授 | 智慧運算與大數據碩士學位學程
- 專任助理教授 | 人工智慧應用學士學位學程
- 人工智慧分析顧問 | 台灣資安鑄造股份有限公司
- 兼任資料科學家 | 中央研究院 社會學研究所

## 主要學歷

- 博士 | 國立臺灣大學 地理環境資源學系
- 碩士 | 國立臺灣大學 地理環境資源學系
- 碩士 | 實踐大學 食品營養與保健生技學系
- 學士 | 國立臺北教育大學 社會與區域發展學系

## 主要經歷

- 兼任助理教授 | 淡江大學 人工智慧學系
- 博士後研究員 | 臺北醫學大學 醫學系 放射線學科
- 博士後研究員 | 台北市立萬芳醫院 影像醫學部
- 資料分析師 | 財團法人資訊工業策進會 資安科技研究所
- 實習生 | 行政法人國家災害防救科技中心 坡地組
- 兼任資料科學家 | 香港中文大學 新聞與傳播學院
- 研究助理 | 臺大地理系 地理計算科學研究室
- 研究助理 | 臺大地理系 遙測及空間知識實驗室
- 研究助理 | 國北社發系 土石流防災實驗室



Website

# About Dr. Chan



Website

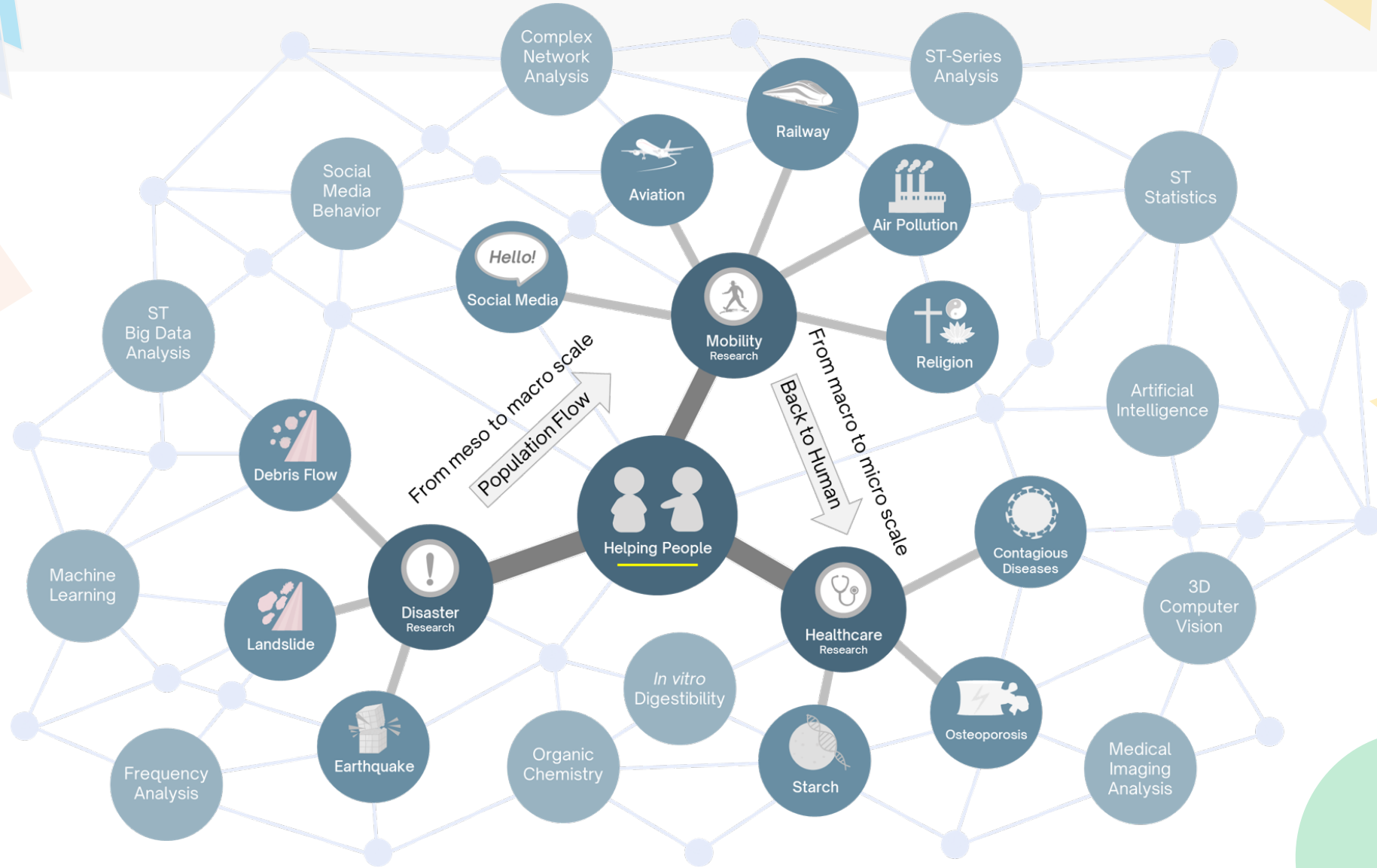
## Technical Skills

- **Computer Science:** Python, Matlab, R, C#, JavaScript, jQuery, jQueryUI, Android Developer, MySQL, Nodejs, AngularJS, MongoDB, Elasticsearch, Spark, Facebook APIs and Twitter APIs
- **Geography:** GIS (ArcGIS, QGIS, Super GIS), Spatial Statistics, Spatial Database, Complex Network Analysis, Gephi
- **Physics:** Signal Processing (in time sequence and frequency) and Electromagnetic Analysis
- **Food Chemistry:** Starch Science, Resistant Starch, Slowly Digestible Starch, *in vitro* Digestibility, SEM, XRD and HPSEC
- **Chemistry:** Organometallic synthesis, NMR, IR, HPLC, ESI-MASS and pH meter
- **Design:** Illustrator, Photoshop, Dreamwaver and Google SketchUp
- **Marketing:** Google Analysis, Facebook Marketing and Google Trend

## Interests

Emergency Medicine, Chinese Medicine, Volleyball, Sport Science, Photography, Tourism, Web and Graphic Design

# About Dr. Chan



Website

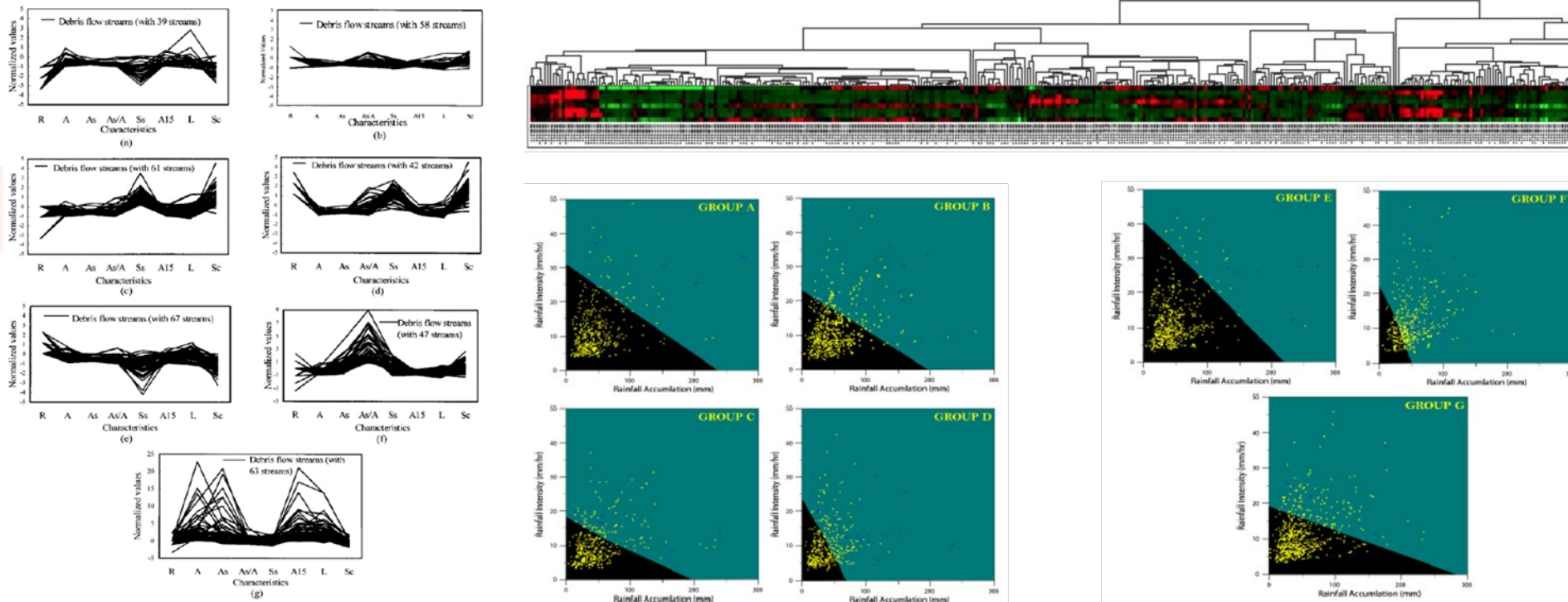
# Debris Flow Disaster | Part 1 |

Bachelor Thesis | Department of Social and Regional Development | National Taipei University of Education



Website

**Gap/ Objectives:** The Chi-Chi huge earthquake occurred in Taiwan in 1999 and it changed the critical rainfall line of debris flow streams; therefore, how to establish a new critical rainfall line for each debris flow stream with a limited data becomes a vital issue. This study conducted family competition genetic algorithm and support vector machine to establish the critical rainfall line.





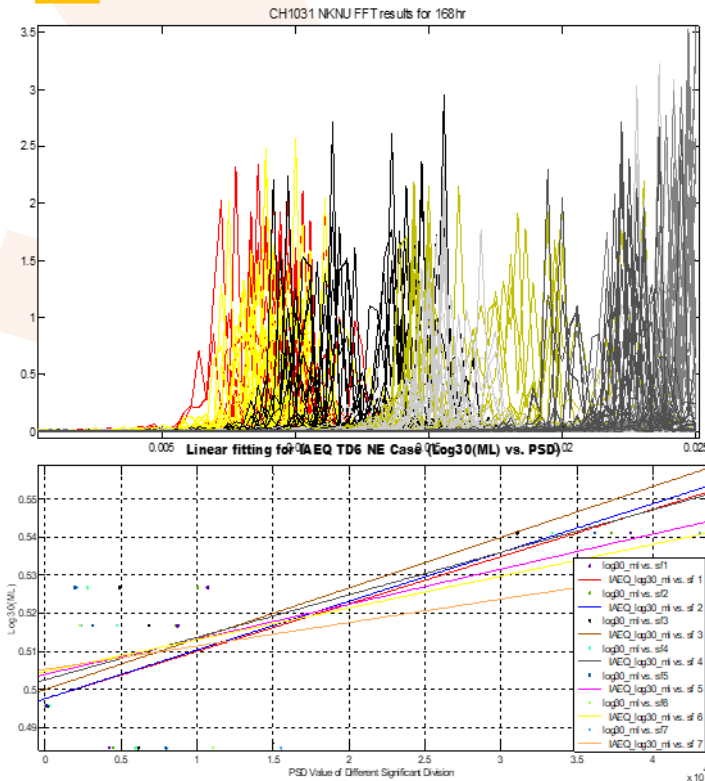
# Seismic Precursor | Part 2 |

Master Thesis | Department of Geography | National Taiwan University

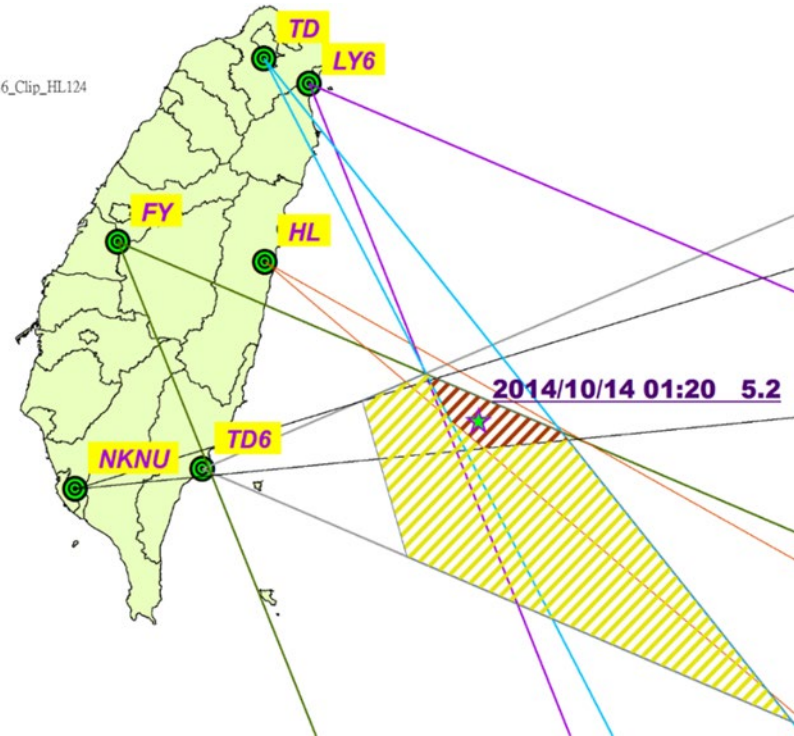


Website

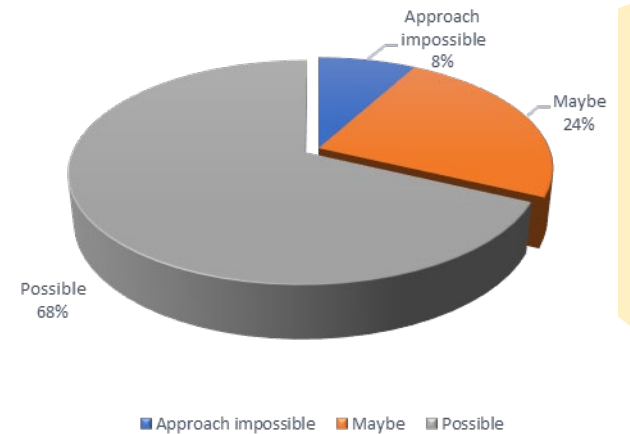
**Gap/ Objectives :** Taiwan is suffered from earthquake disasters; however, state-of-art only performed case study and empirical equation to depict seismic precursor with a low spatial and temporal accuracy. This study deployed signal processing to capture the feature frequency band, identified the epicenter, estimated break time and magnitude of an earthquake events.



- Legend
- ★ EQ005101452
  - ▨ NKNU79FYSE\_LY6SETD146\_Clip\_HL124
  - ▨ EQ005101452\_Clip
  - TD\_146
  - NKNU\_79
  - HL\_124
  - FY\_SE
  - LY6\_SE
  - TD6\_E
  - TWN\_6station\_97
  - TWN\_COUNTY



Break Time Estimation Possibility



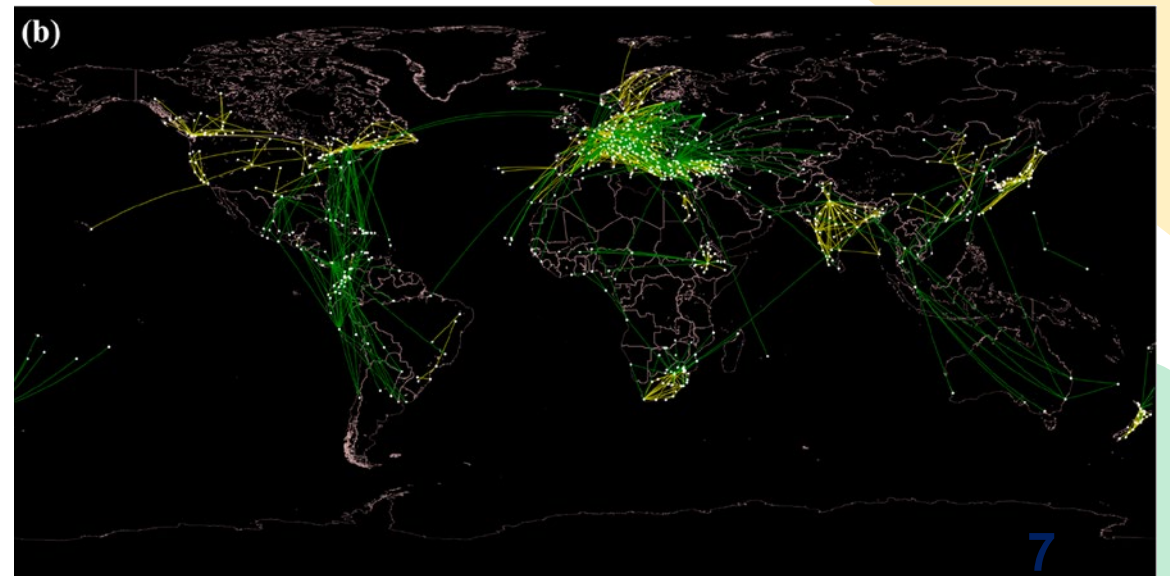
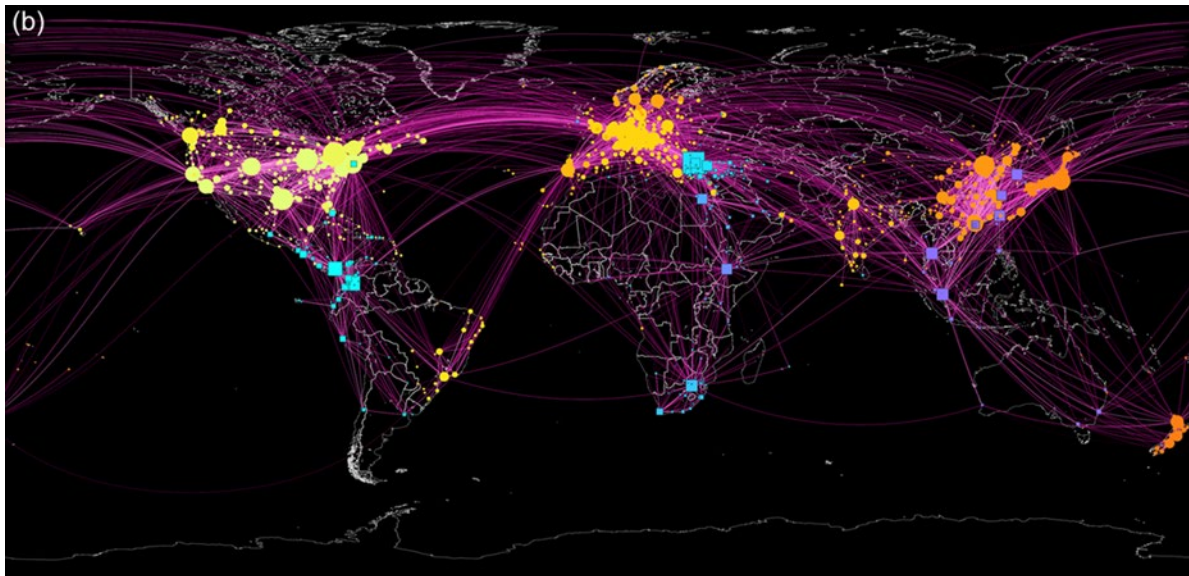
# Aviation Market | Part 3 |

Ph.D. Thesis | Department of Geography | National Taiwan University



Website

**Gap/ Objectives :** Aviation market plays an important role in the accessibility between countries, consisting of airports and routes. However, market-related studies did not consider the connectivity and airport network-related studies neglected the market characteristics. Thus, this study is to utilize the concept of airport community to understand the spatial patterns of the market region in an airline alliance and characterize the differences between airline, including regions of collaboration, competition, and dominance.





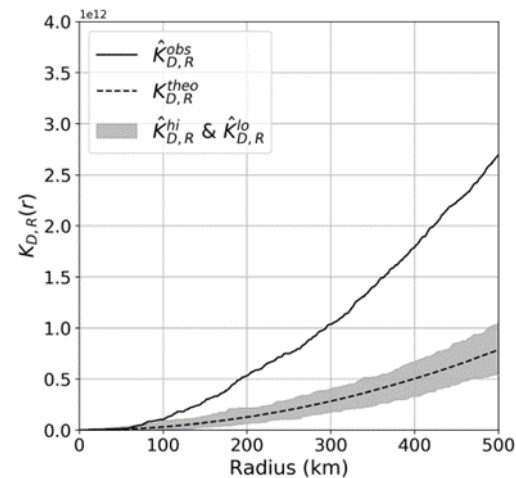
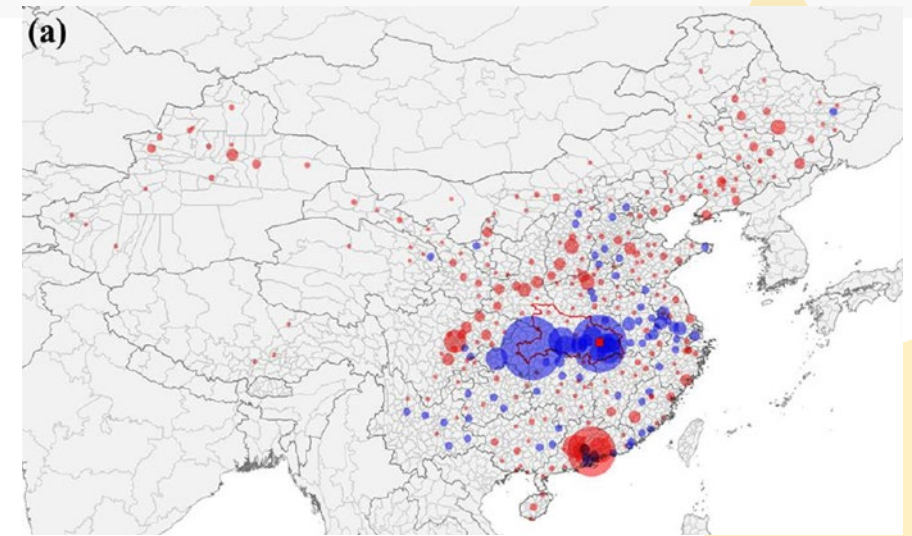
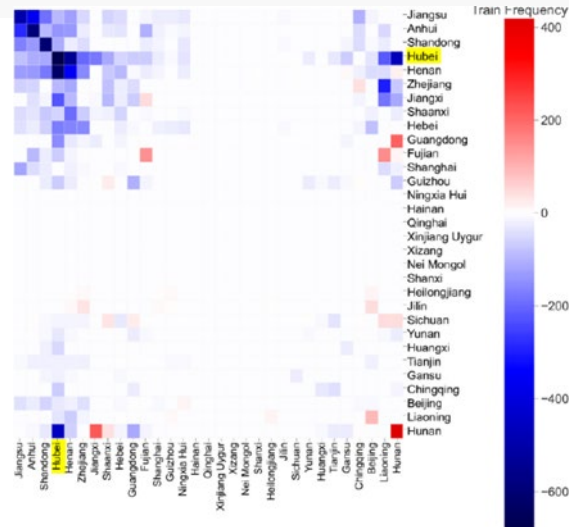
# Epidemic Transmission | Part 4 |

MOST Project | Department of Geography | National Taiwan University



Website

**Gap/ Objectives :** The possible reasons of a large number of confirmed cases concentrated in the neighboring provinces of Hubei have not been fully discussed after the Wuhan city lockdown. Therefore, this study aims to assess the changes in railway passenger transport on the early spatial transmission of COVID-19 in mainland China.



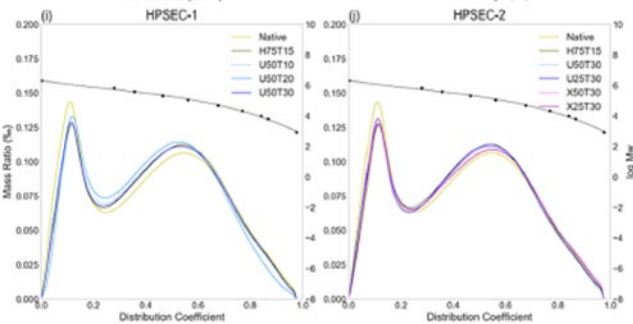
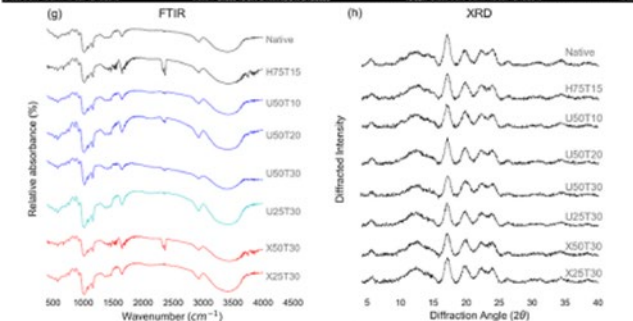
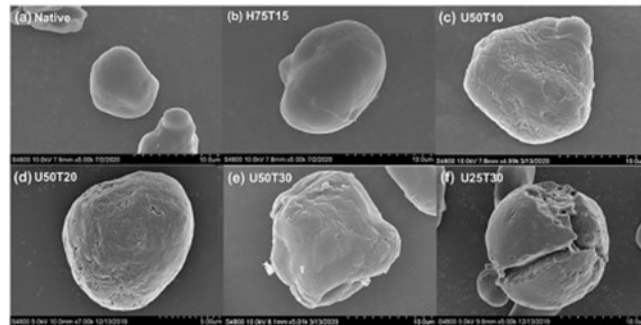
# Starch - *in vitro* digestibility

| Part 5 |

Master Thesis | Department of Food Science, Nutrition, and Nutraceutical Biotechnology | Shih Chien University



Website



**Table 3**  
The proportion of variance explained and variable loadings of each principal component.

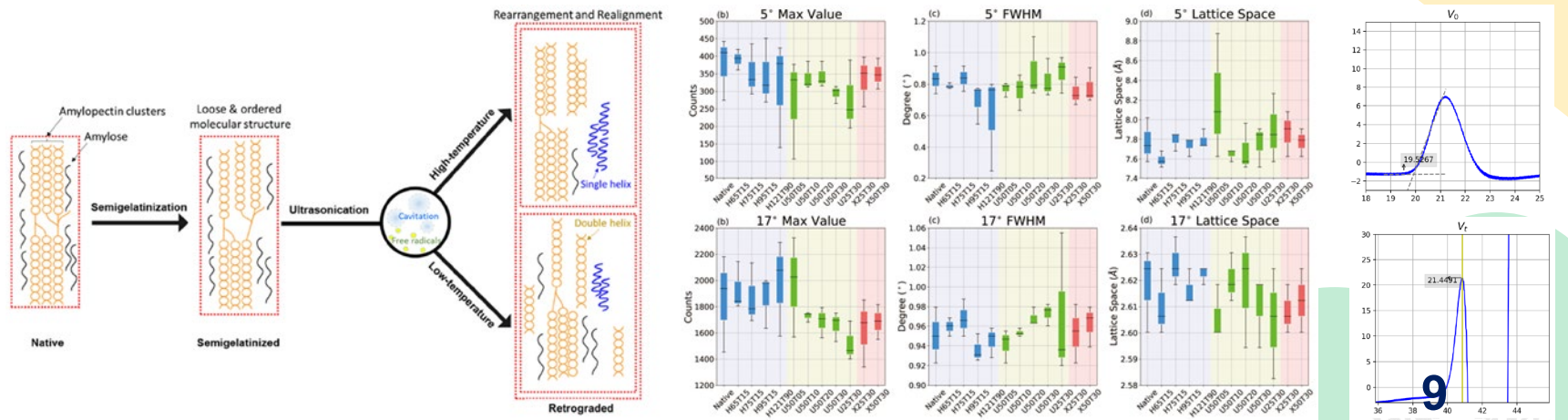
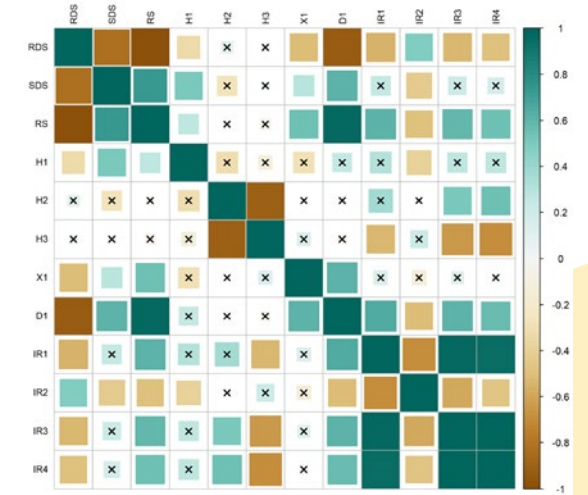
PC	PVE	Variable loadings									
		FTIR				XRD	DSC	HPSEC			AAC
		DO <sub>1047/1022</sub>	DH <sub>996/1022</sub>	AMAP <sub>927/1000</sub>	AMAP <sub>1800/1000</sub>			RC <sub>XRD</sub>	ΔH	F1	
PC1	0.49	-0.44	-0.13	-0.44	-0.43	0.13	-0.12	-0.22	-0.33	0.39	-0.27
PC2	0.25	0.00	-0.27	-0.06	-0.10	0.20	0.47	0.46	-0.37	0.26	0.46
PC3	0.14	0.06	-0.12	-0.02	-0.04	-0.75	-0.48	0.25	-0.25	0.19	0.16

PVE is the proportion of variance explained.

**Table 4**  
The regression coefficients of each principal component for SDS and RS in multivariate linear regression.

	SDS (R <sup>2</sup> = 0.77; adjusted R <sup>2</sup> = 0.71)				RS (R <sup>2</sup> = 0.93; adjusted R <sup>2</sup> = 0.91)			
	Beta	Std. Error	t value	Pr (> t )	Beta	Std. Error	t value	Pr (> t )
Intercept	14.32	0.45	31.51	<0.001***	42.13	0.82	51.22	<0.001***
PC1	-0.28	0.21	-1.32	0.211	-1.65	0.38	-4.30	0.001**
PC2	1.34	0.29	4.56	<0.001***	4.67	0.53	8.78	<0.001***
PC3	-1.62	0.40	-4.07	0.002**	-5.91	0.72	-8.20	<0.001***

Beta is the average coefficient of each independent variable.  
Std. Error is the standard error of each coefficient.  
\*\*\* indicates  $p < 0.001$ ; \*\* indicates  $p < 0.01$ ; \* indicates  $p < 0.05$ .



# Outlines

- Course Introduction
- Grading Policy
- What is Probability?
- What is Statistics?
- Why you need to take this course?
- What you will learn from this course?
- [#1] Assignment
- Question Time

# Course Introduction

In the first semester, we will cover seven parts of probability as follows. Due to the time limitation, the statistic part will be introduced in the next probability.

- 1) Combinatorial Analysis
- 2) Axioms of Probability
- 3) Conditional Probability and Independence
- 4) Random Variables
- 5) Continuous Random Variables
- 6) Joint Distributed Random Variables
- 7) Properties of Expectation



# Course Introduction

Week	Date	Content
1	Sep. 12	Introduction to Probability and Statistics (Online)
2	Sep. 19	Combinatorial Analysis (I)
3	Sep. 26	Combinatorial Analysis (II)
4	Oct. 3	Axioms of Probability (I)
5	Oct. 10	(Taiwan National Day (Holiday))
6	Oct. 17	Axioms of Probability (II)
7	Oct. 24	Conditional Probability and Independence
8	Oct. 31	Mid-term Exam
9	Nov. 7	- Mid-term Exam Week -
10	Nov. 14	Random Variables (I)
11	Nov. 21	Random Variables (II)

Week	Date	Content
12	Nov. 28	Continuous Random Variables (I)
13	Dec. 5	Continuous Random Variables (II)
14	Dec. 12	Joint Distributed Random Variables (I)
15	Dec. 19	Joint Distributed Random Variables (II)
16	Dec. 26	Final Exam
17	Jan. 2	((Holiday))
18	Jan. 9	(Final Exams)

# Before, during, after class

- Before the class, ...
  - Read the materials
  - Search online information
- During the class, ...
  - Lecture
  - Discussion
  - Lab practice
- After the class, ...
  - Assignment

# Grading Policy

- All you have to do is study hard and feel free to ask question when you do not understand.
- I believe that if you fulfill all required items, and then you will pass this course.
- Do not worry about the grade! The most important thing is what you learn from this course.

Assignments 30 %

Mid-Exam 30 %

Others 10 %

Final-Exam 30 %

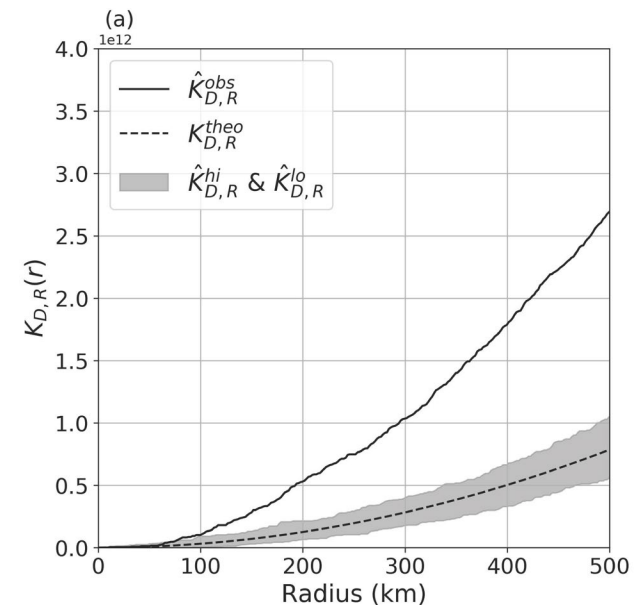
# What's the Probability?

- How does *Probability* affect your life?
  - Dice problem
  - Infectious disease → insurance
  - Disaster → insurance
  - Investment
  - Sales/ Weather/ Transportation forecasting
- In this semester, our course will concentrate on the probability part from fundamental axioms to random variables.



# What's the statistics?

- I believe you all have learned (descriptive) statistics, but do you think that could fulfill all problems? Of course not...
- So, we will learn a series of statistics from descriptive to inferential statistics in the next semester.
  - Student T test
  - ANOVA
  - Chi square
  - F statistics
  - Regression



# Why you need to take this course?

- When I was a student (just like you), I also had the same question “why I need to take this course?” Because we had learned “a lot of” knowledge of probability in the high school; however, what you had learn is the edge of iceberg. The mathematical background and physical meanings behind these terms were not covered.
- Different common courses, real life events will be involved within this course in order to expend your imagination of probability and statistic view.

# What you will learn from this course?

- After I learned this course, you may acquire various abilities of ...
  - 1) You are able to know the fundamental knowledge of probability
  - 2) You are able to design experiments for probability theories via programming
  - 3) You are able to leverage the learned knowledge into real-world problems

# [#1] Assignment

- What will you expect to learn from this course?  
Please describe your expectations with few sentences.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur at dolor maximus, varius massa eu, porttitor sapien. Vivamus facilisis malesuada faucibus. Pellentesque tristique vitae metus sed sollicitudin. Aenean massa nisl, sodales sed dolor in, placerat maximus elit. Praesent varius tortor vitae tincidunt porttitor. Duis et dui eu purus imperdiet varius. Nam posuere euismod erat at pharetra. Vestibulum in nunc ante.





# Question Time

If you have any questions, please do not hesitate to ask me.



# The End

*Thank you for your attention ))*