

# สถิติในชีวิตประจำวัน

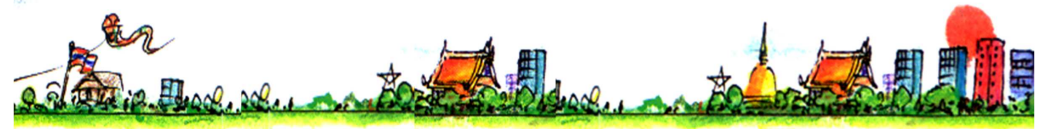
โพล และเลขดัชนี

รองศาสตราจารย์พิษณุ เจียวคุณ

ภาควิชาสถิติ คณะวิทยาศาสตร์



# ความสำคัญของสถิติในชีวิตประจำวัน



## การค้นหาคความจริง

วิธีอนุมาน (Deduction) คือการตั้งกฎเกณฑ์ไว้  
และนำความเป็นจริงที่มีอยู่มาวิเคราะห์ สังเคราะห์  
เพื่อนำไปใช้ในการอธิบายข้อเท็จจริงต่าง ๆ



## การค้นหาคความจริง

วิธีอุปมาน (Induction) คือการสังเกตข้อเท็จจริง  
ต่าง ๆ หรือทำการทดลองซ้ำ ๆ จนกระทั่งได้  
ข้อสรุปของสิ่งที่เกิดขึ้น



## ระเบียบวิธีทางวิทยาศาสตร์

1. ระบุปัญหา
2. ตั้งสมมุติฐาน
3. เก็บรวบรวมข้อมูลหรือทำการทดลอง
4. วิเคราะห์ข้อมูลเพื่อนำไปสู่การหาข้อสรุป
5. ตรวจสอบเพื่อยืนยันผลการศึกษา



## สถิติ ข้อมูลและการรวบรวมข้อมูล

- ความหมายของสถิติ (Statistics)  
มาจากคำว่า STATE ในอดีตสถิติจึงเป็นเครื่องมือที่ใช้ในการบริหารจัดการข้อมูลของรัฐ ปัจจุบันสถิติ เป็นศาสตร์ที่ว่าด้วย “ข้อมูล” ที่ประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายข้อมูล



## ข้อมูล (Data)

- คือการบันทึกข้อเท็จจริง (Fact) หรือปรากฏการณ์ต่าง ๆ ที่เกิดขึ้น ซึ่งข้อมูลเหล่านี้จะมีการเก็บรวบรวมไว้หลาย ๆ ครั้ง หรือทำการทดลองซ้ำ ๆ กัน จนกระทั่งได้ข้อเท็จจริงมาจำนวนหนึ่ง



## สารสนเทศ (Information)

- เป็นการนำข้อมูล ข้อเท็จจริงที่มีอยู่มาผ่านกระบวนการ ประมวลผลจนกระทั่งได้ข้อสรุป และสามารถนำไปใช้ในการตัดสินใจได้ และเมื่อนำใช้แล้ว ในสถานการณ์ต่างจะมีประสบการณ์เกิดขึ้น จนเป็น**ความรู้** และนำไปสู่**การจัดการความรู้ (KM)**



## ข้อมูลสถิติ และสารสนเทศทางสถิติ

การรวบรวมข้อมูลทางสถิติที่สนใจไว้ด้วยกัน และมีการนำเอาข้อมูลเหล่านั้น มาจัดการโดยจัดทำเป็นหมวดหมู่ วิเคราะห์ข้อมูลจนได้สารสนเทศทางสถิติ และสามารถนำไปใช้ในการตัดสินใจได้



## สถิติศาสตร์ (Statistics)

คือศาสตร์ที่ประกอบไปด้วย

1. การเก็บรวบรวมข้อมูล
2. การนำเสนอข้อมูล
3. การวิเคราะห์ข้อมูล และ
4. การแปลความหมายข้อมูล

ระเบียบวิธีทางสถิติ Statistical methods



## พารามิเตอร์และค่าสถิติ

พารามิเตอร์(Parameter) คือค่าที่ประมวลได้จากทุกหน่วยของประชากรที่สนใจ

ค่าสถิติ(Statistic) คือค่าที่ประมวลได้จากตัวอย่าง



## ขอบข่ายของสถิติ

- สถิติแบ่งออกเป็น 2 ลักษณะคือ
- 1. สถิติเชิงพรรณนา (Descriptive Statistics) คือ การอธิบายหรือพรรณนา ลักษณะของข้อมูล หรืออาจจัดข้อมูลให้เป็นหมวดหมู่ และอธิบายข้อมูลนั้น



## ขอบข่ายของสถิติ

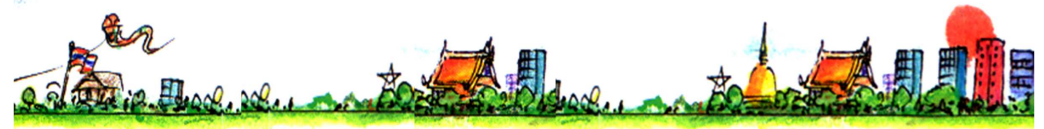
- สถิติแบ่งออกเป็น 2 ลักษณะคือ
- 2. **สถิติเชิงอนุมาน** (Inferential Statistics) คือ การหาข้อสรุปของประชากรที่สนใจศึกษา โดยอาศัยข้อมูลจากตัวอย่าง



## ธรรมชาติของข้อมูล

ข้อมูลแบ่งได้ 2 ลักษณะคือ

1. ข้อมูลเชิงคุณภาพ (Qualitative data) คือข้อมูลที่ไม่สามารถวัดค่าออกมาเป็นตัวเลขได้โดยตรง เช่น เพศ ศาสนา สีผิว หมูโลहित เป็นต้น



## ธรรมชาติของข้อมูล

ข้อมูลแบ่งได้ 2 ลักษณะคือ

1. ข้อมูลเชิงปริมาณ (Quantitative data) คือข้อมูลที่สามารถวัดค่าออกมาเป็นตัวเลขได้โดยตรง เช่น อายุ น้ำหนัก ส่วนสูง คะแนนสอบ ความเร็ว ระยะทาง เป็นต้น



## มาตรวัดของข้อมูล (Scale of measurement)

ข้อมูลแบ่งได้ 4 มาตรวัดคือ

1. มาตรฐานบัญญัติ (Nominal scale)
2. มาตรฐานเรียงลำดับ (Ordinal scale)
3. มาตรฐานंतरภาค (Interval scale)
4. มาตรฐานอัตราส่วน (Ratio scale)



## การเก็บรวบรวมข้อมูลทางสถิติ

1. การทำสำมะโน (Census) คือการเก็บข้อมูลทุกหน่วยที่ต้องการเจ้านับ

2. การสำรวจด้วยตัวอย่าง (Sample survey)

คือการรวบรวมข้อมูลเพียงบางหน่วยจากประชากร เรียกว่า หน่วยตัวอย่าง (sample unit) การได้มาของหน่วยตัวอย่าง ต้องใช้วิธีการสุ่มตัวอย่าง (sampling techniques)



## การนำเสนอข้อมูลทางสถิติ

การนำเสนอข้อมูลสามารถทำได้หลายวิธี เช่น การสร้างตารางแจกแจงความถี่ (นำเสนอข้อมูลเชิงปริมาณ) การนำเสนอข้อมูลในรูปแบบตาราง (นำเสนอข้อมูลเชิงคุณภาพ)



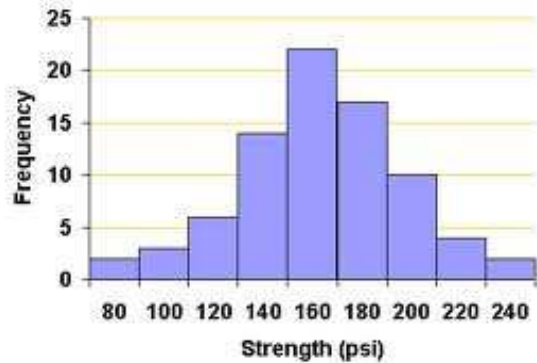
## ตารางแจกแจงความถี่

Audit Time (Days)	Frequency
10-14	4
15-19	8
20-24	5
25-29	2
30-34	1
Total	20

## การนำเสนอข้อมูลในรูปแบบตาราง

Exchange	Frequency	Percent frequency
New York Stock exchange	20	80
Nasdaq National Market	5	20
Totals	25	100

# Histogram



## การวิเคราะห์ข้อมูลเบื้องต้น

- การวัดแนวโน้มเข้าสู่ส่วนกลาง เช่น ค่าเฉลี่ย ค่าเฉลี่ยถ่วงน้ำหนัก ค่ามัธยฐาน และ ฐานนิยม
- การวัดการกระจาย เช่น ค่าเบี่ยงเบนมาตรฐาน ค่าความแปรปรวน



## ค่าเฉลี่ยเลขคณิต (Mean)

- ใช้สัญลักษณ์  $\mu$  หรือ  $\bar{X}$

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$\mu = \frac{x_1 + x_2 + \dots + x_N}{N}$$

## มัธยฐาน (Median)

- คือการหาค่าที่อยู่ตรงกลางของข้อมูลทั้งหมดที่นำมาเรียงลำดับแล้ว จากน้อยไปหามาก หรือจากมากไปหาน้อย ถ้าจำนวนข้อมูลเป็นเลขคี่ ค่ามัธยฐานจะอยู่ตรงตำแหน่งที่  $(N+1)/2$  แต่ถ้าจำนวนข้อมูลเป็นเลขคู่ มัธยฐานของข้อมูลจะอยู่ระหว่างข้อมูลตำแหน่งที่  $N/2$  กับ  $(N/2)+1$

## ฐานนิยม (Mode)

- คือข้อมูลที่เกิดขึ้นบ่อยที่สุด หรือมีความถี่สูงสุด สามารถเป็นได้ทั้งข้อมูลเชิงปริมาณและข้อมูลเชิงคุณภาพ
- ฐานนิยม อาจจะมีมากกว่า 1 ค่า หรืออาจจะไม่มีก็ได้ ถ้าข้อมูลชุดนั้นไม่มีข้อมูลซ้ำกันเลย

## ค่าเฉลี่ยถ่วงน้ำหนัก (weighted mean)

- The mean of a data set in which **not all values are equally represented**

$$\bar{x}_w = \frac{\sum x_i w_i}{\sum w_i}$$

$W_i$  = น้ำหนักถ่วงของข้อมูลแต่ละตัว ( $x_i$ )

## คุณสมบัติของค่าเฉลี่ย มัธยฐาน และฐานนิยม

### ค่าเฉลี่ยเลขคณิต

1. ค่าเฉลี่ยเลขคณิตของข้อมูลชุดหนึ่ง จะมีเพียงค่าเดียว
2. ค่าเฉลี่ยเลขคณิต เหมาะสมกับข้อมูลที่มีค่าใกล้เคียงกัน และไม่มีข้อมูลที่มีค่าผิดปกติ
3. ค่าเฉลี่ยสามารถนำไปใช้ในการคำนวณสถิติขั้นสูง

## Properties and uses of Central Tendency

### • มัธยฐาน

1. มัธยฐานของข้อมูลชุดหนึ่ง จะมีเพียงค่าเดียว
2. ถ้ามีข้อมูลที่มีค่าปกติ จะไม่มีผลกระทบต่อการหาค่ามัธยฐาน เนื่องจากมัธยฐานนำเฉพาะค่าที่อยู่ตรงกลางมาพิจารณาเท่านั้น

## Properties and uses of Central Tendency

- **ฐานนิยม**
- ฐานนิยม เหมาะสมกับข้อมูลเชิงคุณภาพ ที่มีมาตรวัดแบบนามบัญญัติ และมาตรวัดเรียงลำดับ
- ฐานนิยม ไม่เหมาะสมกับข้อมูลเชิงปริมาณ

## การวัดการกระจาย (Measures of Variation)

- The main objective of measures of variation is to measure the the spread or variability of a data set.
- วัดดูประสงค์เพื่อวัดความห่างกันของข้อมูลชุดหนึ่ง ๆ

## ความแปรปรวน (variance)

- The variance is the average of the squares of the distance each value is from the mean. The symbol for the population variance is  $\sigma^2$

$$\sigma^2 = \frac{\sum (X_i - \mu)^2}{N}$$

For sample variance

$$s^2 = \frac{\sum (X_i - \bar{x})^2}{n-1}$$

## ค่าเบี่ยงเบนมาตรฐาน (Standard deviation)

- The standard deviation is the square root of the variance

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$

For sample standard deviation

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$



## Coefficient of Variation : CV

### สัมประสิทธิ์การแปรผัน

- ใช้ในการเปรียบเทียบการกระจายของข้อมูล 2 ชุดที่  
อาจจะมีหน่วยเหมือนกันหรือต่างกัน

for populations  $CV = \frac{\sigma}{\mu}$

for samples  $CV = \frac{s}{\bar{x}}$