

ความน่าจะเป็นและสถิติสำหรับวิศวกร



Probability and Statistics
for Engineers

e-mail: reudee.ni@ssru.ac.th

Web site: http://www.elfit.ssru.ac.th/reudee_ni/

รศ.ดร.ฤดี นิยมรัตน์

คำอธิบายรายวิชา

ความน่าจะเป็น ตัวแปรสุ่ม
การทดสอบสมมติฐาน การแจกแจง
การแจกแจงแบบมีเงื่อนไข ฟังก์ชันความ
หนาแน่น ความน่าจะเป็น ค่าความคาดหวัง

Probability, random variables,
hypotheses testing, distribution,
conditional distribution, probability
density function, expectations,

variances,

statistical inferences,

regression analysis,

implementation

2 ความแปรปรวน

1 การอนุมานทางสถิติ

3 การวิเคราะห์การถดถอย

การประยุกต์ใช้งาน



กำหนดการสอน 5 ครั้ง

- การประเมินผล
1. คะแนนการร่วมกิจกรรมในชั้นเรียน
 2. คะแนนการทดสอบ รวม 30 คะแนน

หัวข้อการเรียน	การทดสอบ (คะแนน)
การอนุมานทางสถิติ	20
t-test	20
F-test (one way ANOVA)	20
การวิเคราะห์การถดถอยอย่างง่าย	20
การวิเคราะห์การถดถอยพหุคูณ	20

การทุจริต
ในการทดสอบ
มีคะแนนการ
ประเมินผล
ทั้ง 5 ครั้ง เป็น 0



สถิติเชิงอนุมาน

1

Inferential Statistics

สถิติ

สถิติเชิงพรรณนา
(Descriptive Statistics)

สถิติเชิงอนุมาน
(Inferential Statistics)

การนำข้อสรุปที่ได้จากการศึกษาข้อมูลตัวอย่าง
ไปอ้างอิงถึงลักษณะของประชากร



Sampling



Inferential Statistics

Location parameter

- Mean
- Median
- Modal value
- Sum

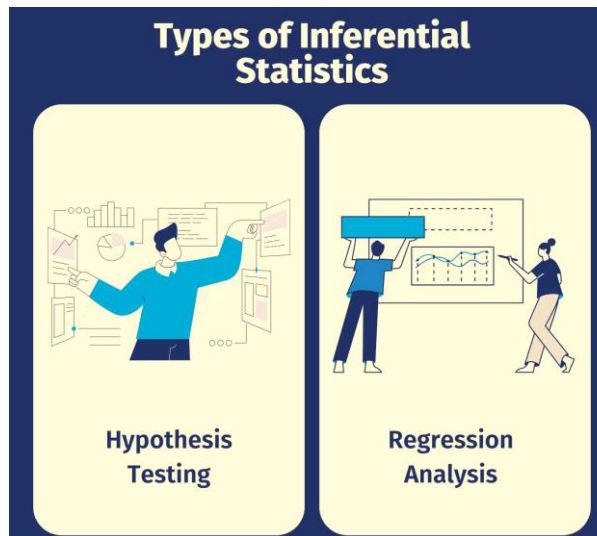
Dispersion parameter

- Standard deviation
- Variance
- Range

Descriptive statistics

2

การทดสอบสมมติฐาน
Z-test
t-test
F-test



การวิเคราะห์
การถดถอย

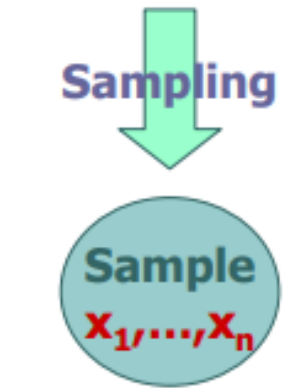
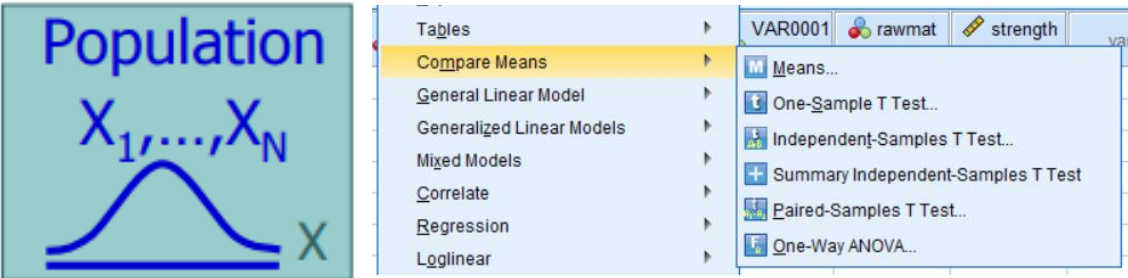
3



สถิติเชิงอนุมาน (Inferential Statistics)

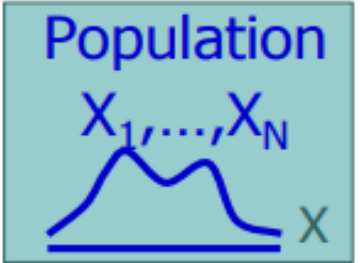
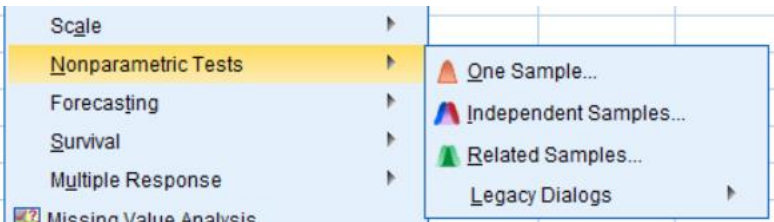
สถิติที่กำหนดรูปแบบของประชากร (Parametric Statistics)

Z-test, t-test เป็นการอนุมานค่าเฉลี่ยของประชากร
F-test เป็นการอนุมานค่าความแปรปรวนของประชากร

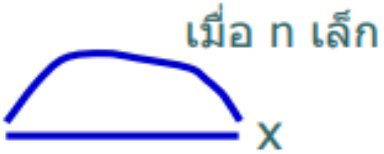


สถิติที่ไม่กำหนดรูปแบบของประชากร (Non-Parametric Statistics)

- 1-Single-Sample Case
- Binomial test
- Chi-square test
- Kolmogorov-Smirnov test
- Run Test
- 2-Two Related (DV) Sample
- 3-Two Independent Samples
- 4-k Related Samples



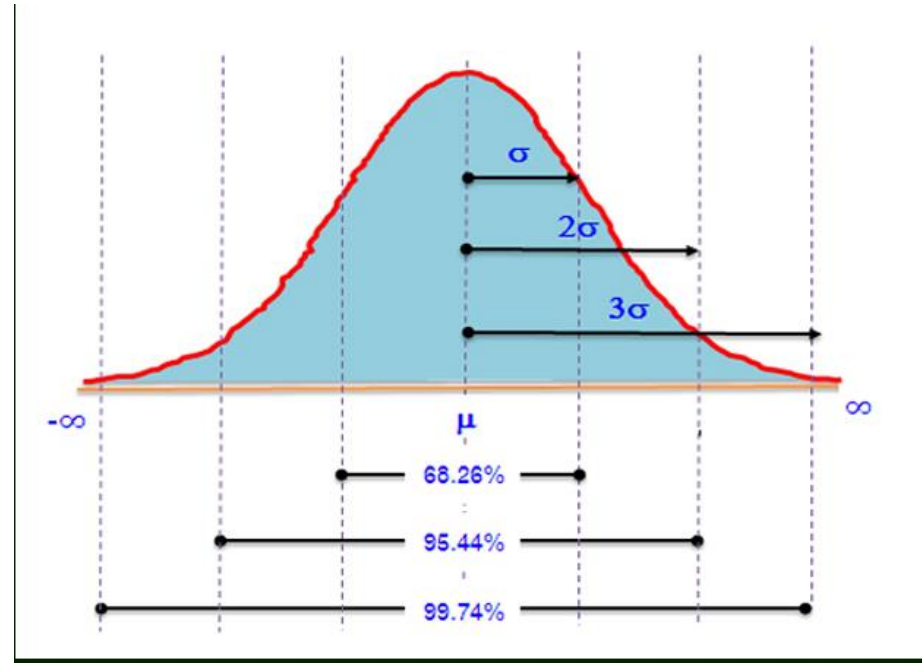
↓
Sampling



ข้อตกลงเบื้องต้นสำหรับการทดสอบ ทั้ง t-test และ F-test

1. ข้อมูลเป็นระดับ Interval Scale และ Ratio Scale

1.นามบัญญัติ (Nominal)	ชื่อ ประเภท แบ่งกลุ่ม แบ่งพวก ชนิดของสิ่งต่าง ๆ	<ul style="list-style-type: none"> -เพศ (หญิง/ ชาย) -กลุ่มเลือด (A B AB และ O) -สถานภาพสมรส (โสด คู่ หม้าย หย่า และ แยก) -บ้านเลขที่ เบอร์โทรศัพท์ หมายเลขเสื้อนักกีฬา ฯลฯ
2.เรียงอันดับ (Ordinal)	บอกลำดับข้อมูลได้ว่ามาก น้อย สูงหรือต่างกัน แต่บอกไม่ได้ว่าต่างกันเท่าใด	<ul style="list-style-type: none"> -ระดับความเจ็บปวด (มาก ปานกลาง น้อย) -ระดับความคิดเห็น (เห็นด้วย เฉย ๆ ไม่เห็นด้วย) -ระดับความพึงพอใจ (มาก น้อย ไม่พอใจ)
3.อันตรภาค (Interval)	เป็นตัวเลขต่อเนื่อง แต่เป็นศูนย์ไม่แท้ (คือที่จุด 0 ยังมีค่าอยู่)	<ul style="list-style-type: none"> -อุณหภูมิ - I.Q. -คะแนนความรู้ -คะแนนทัศนคติ -คะแนนการปฏิบัติ ฯลฯ
4.อัตราส่วน (Ratio)	เป็นตัวเลขต่อเนื่อง มีจุดศูนย์แท้ (คือค่า 0 มีความหมายว่าไม่มีค่า)	<ul style="list-style-type: none"> -น้ำหนัก -ส่วนสูง -น้ำตาลในเลือด -ความดันโลหิต -จำนวนผู้ป่วย ฯลฯ



2. ข้อมูล/ตัวแปรตามมีการแจกแจงแบบปกติเป็นโค้งปกติ (Normal Distribution Curve) >> ทดสอบ

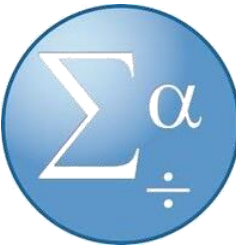
3. ประชากรแต่ละกลุ่มที่นำมาศึกษามีความแปรปรวน (Variance) เท่ากัน >> ทดสอบ



ข้อตกลงเบื้องต้นสำหรับการทดสอบ ทั้ง t-test และ F-test

1. ข้อมูลเป็นระดับ Interval Scale และ Ratio Scale
2. ข้อมูล/ตัวแปรตามมีการแจกแจงแบบปกติเป็นโค้งปกติ (Normal Distribution Curve) >> ทดสอบ
3. ประชากรแต่ละกลุ่มมีความแปรปรวน (Variance) เท่ากัน

H_0 : ข้อมูลมีการแจกแจงแบบปกติ
 H_1 : ข้อมูลไม่มีการแจกแจงแบบปกติ



SPSS Statistics Data Editor

Analyze > Nonparametric Tests > Legacy Dialogs > 1-Sample K-S...

1-Sample K-S

One-Sample Kolmogorov-Smirnov Test

ตัวแปรตาม Dependent Variable, DV: Y

Test Variable List: ผลผลิต

Test Distribution: Normal Uniform Poisson Exponential

2

IBM SPSS Statistics Viewer

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		ผลผลิต
N		21
Normal Parameters ^{a,b}	Mean	64.4762
	Std. Deviation	9.30387
Most Extreme Differences	Absolute	.185
	Positive	.185
	Negative	-.157
Test Statistic		.185
Asymp. Sig. (2-tailed)		.058 ^c

a. Test distribution is Normal.

3

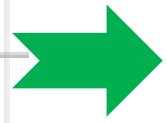
ค่า Sig. หรือค่า P-value มีค่ามากกว่าค่าอัลฟา (.05) จึงยอมรับ H_0 คือข้อมูลมีการแจกแจงแบบปกติ





การทำทศนิยม 3 ตำแหน่งใน output ของ SPSS

-1.247E-15
.333
.333
.333
-1.208E-15



*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Insert Pivot **Format** Analyze Graphs Utilities Window Help

- Cell Properties...
- Table Properties...
- TableLooks...
- Footnote Marker...
- Set Data Cell Widths...
- Autofit
- ReNUMBER Footnotes
- Rotate Inner Column Labels
- Rotate Outer Row Labels
- Breakpoints

Audio Learning Style
Kinesthetic Learning Style

a. Dependent Variable: happy



Cell Properties

Font and Background **Format Value** Alignment and Margins

Preview: 0.000

Format

Category: All, **Number**, Date, Time, Currency

Format: ##, ## ; ###E-#, #####, #####,###, ###E+##, ##.##%

Decimals: 3

OK Cancel Apply Help

.000
.333
.333
.333
.000



การติดตั้ง

Data Analysis

ของ Excel



1. Arrow pointing to the ribbon area.

2. Arrow pointing to the task pane area.

3. Arrow pointing to the 'Add-in' list.

4. Arrow pointing to 'Analysis ToolPak' in the list.

5. Arrow pointing to the 'Go' button.

6. Arrow pointing to the 'Add-in ที่มีอยู่' list.

7. Arrow pointing to the 'Check' box for 'Analysis ToolPak - VBA'.

8. Arrow pointing to the 'ข้อมูล' (Data) tab on the ribbon.

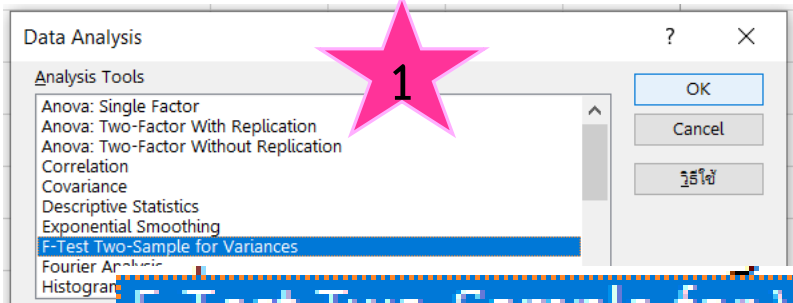
9. Arrow pointing to the 'Data Analysis' button on the ribbon.



ข้อตกลงเบื้องต้นสำหรับการทดสอบ ทั้ง t-test และ F-test

1. ข้อมูลเป็นระดับ Interval Scale และ Ratio Scale
2. ข้อมูล/ตัวแปรตามมีการแจกแจงแบบปกติเป็นโค้งปกติ (Normal Distribution Curve)
3. ประชากรแต่ละกลุ่มมีความแปรปรวน (Variance) เท่ากัน >> ทดสอบ

$H_0: \sigma_1^2 = \sigma_2^2$
 $H_1: \sigma_1^2 \neq \sigma_2^2$



F-Test Two-Sample for Variances

3

F-Test Two-Sample for Variances		
	Variable 1	Variable 2
Mean	62.7	69.62
Variance	43.505	22.672
Observations	5	5
df	4	4
F	1.918887	
P(F<=f) one-tail	0.271694	
F Critical one-tail	6.388233	

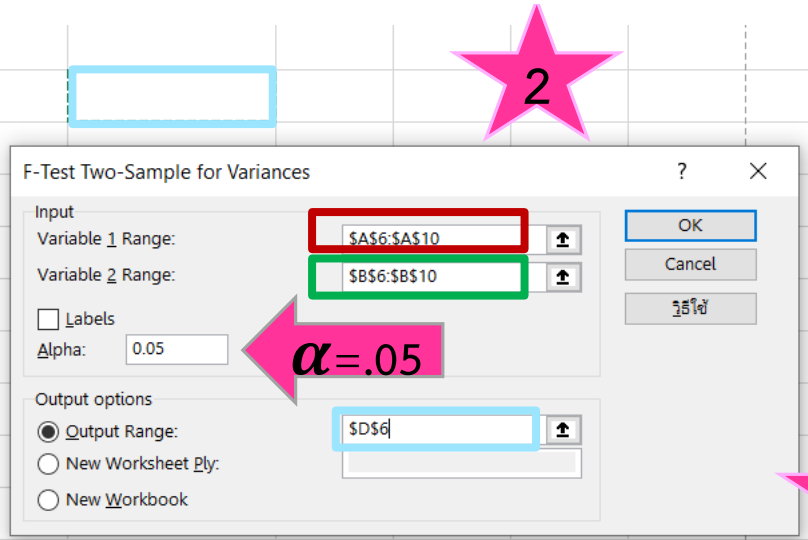
ค่า P-value \geq ค่าอัลฟา (.05)
 จึงยอมรับ H_0
 หมายถึง ความแปรปรวน
 ไม่แตกต่างกัน (เท่ากัน)

t-Test: Two-Sample Assuming Equal Variances

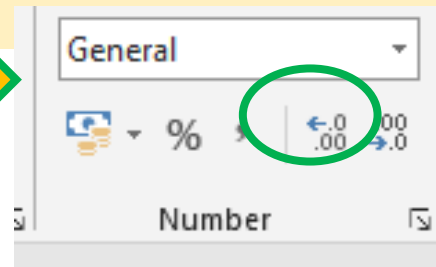
ถ้า ค่า P-value < ค่าอัลฟา (.05)
 จึงปฏิเสธ H_0 และยอมรับ H_1
 หมายถึง ความแปรปรวนแตกต่างกัน

t-Test: Two-Sample Assuming Unequal Variances

Farm 1	Farm 2
53.0	67.1
62.5	73.0
63.5	63.0
71.6	70.0
62.9	75.0



4 ทำให้เป็นทศนิยม 3 ตำแหน่ง



ความแปรปรวน



Variances

Variance² ค่าความแปรปรวน

สำหรับวัดการกระจายตัวของข้อมูล

ค่าความแปรปรวน

คือ ค่าที่ใช้วัดการกระจายตัวของข้อมูล
= ค่าเบี่ยงเบนมาตรฐานยกกำลัง 2

population

sample variance:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1} = 10$$

population

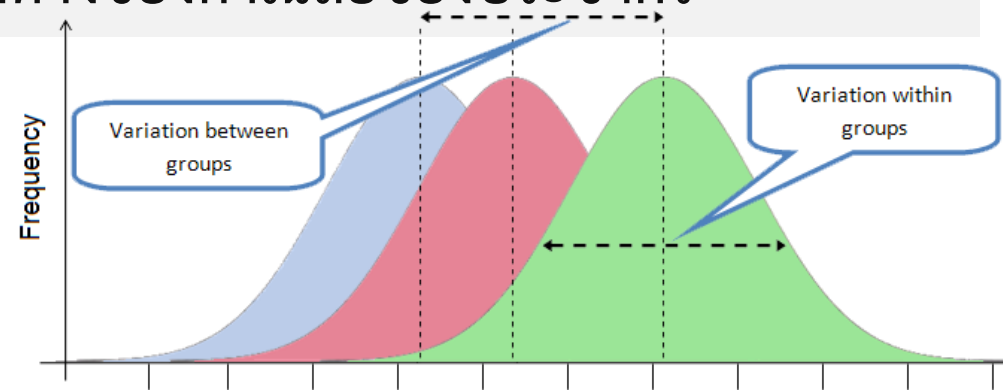
population variance:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n} = 8$$

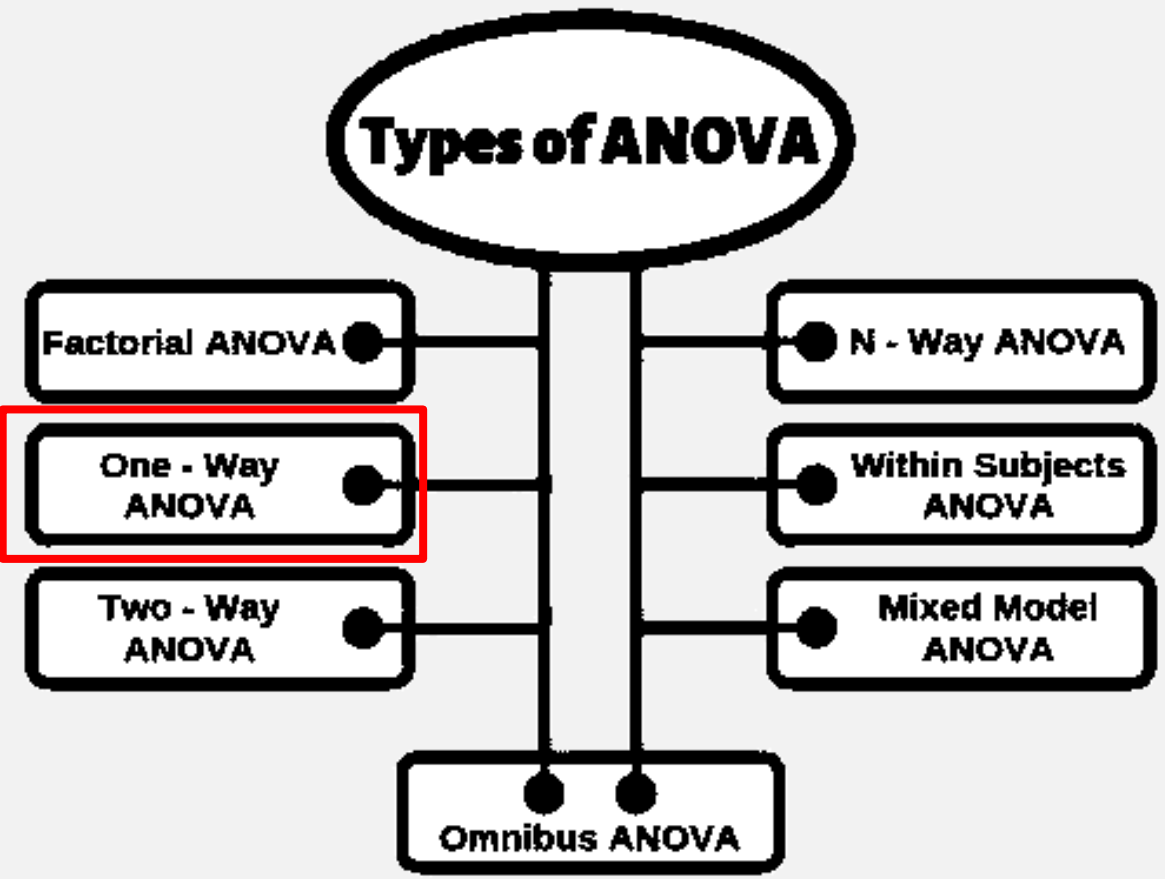
ANOVA

(ANALYSIS OF VARIANCE)

การหาความแตกต่างของค่าเฉลี่ยของประชากร



การวิเคราะห์ความแปรปรวน Analysis of Variance, ANOVA



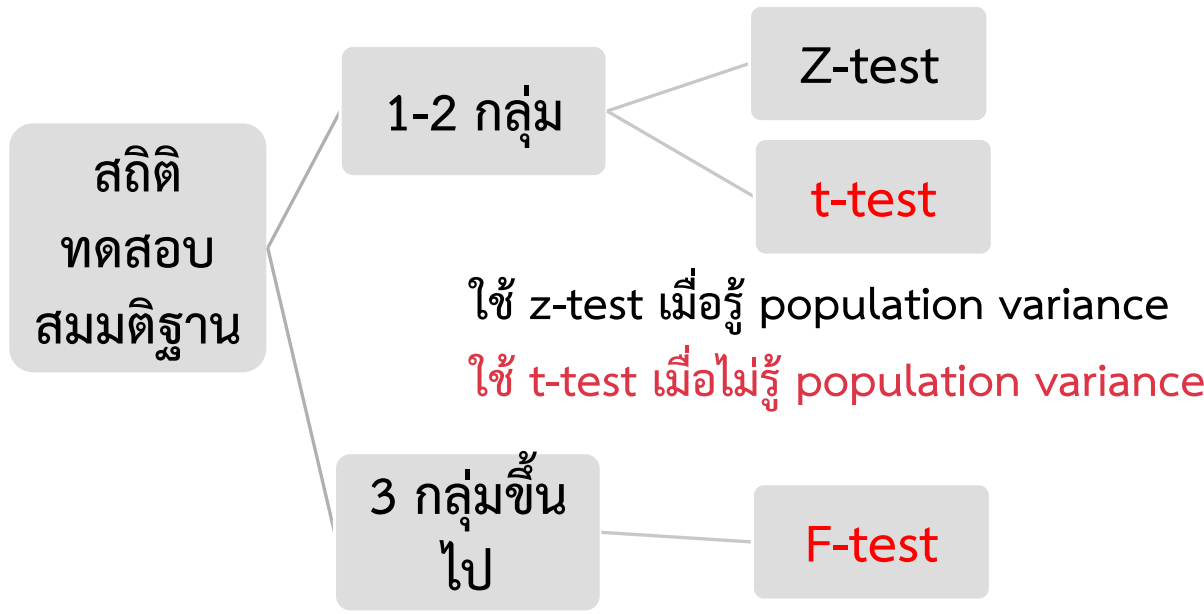
2

การทดสอบสมมติฐาน
Z-test
t-test
F-test

Types of Inferential Statistics

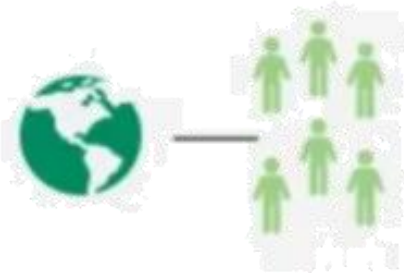
Hypothesis Testing

Regression Analysis



การทดสอบค่าที (t-test)

One sample t-test

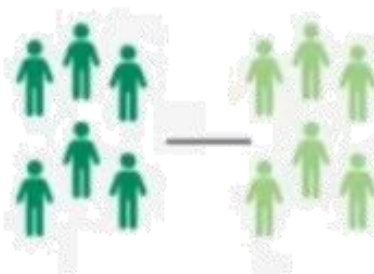


Is there a difference between a group and the population.

เปรียบเทียบข้อมูล 1 กลุ่ม
กับค่ามาตรฐานหรือค่าที่มีก่อนแล้ว

พนักงานแผนผลิต 30 คน มีผลการทดสอบ
แตกต่างกันจากค่าคะแนนสอบมาตรฐานหรือไม่

Independent samples t-test

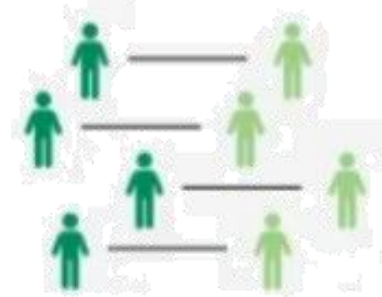


Is there a difference between two groups

เปรียบเทียบข้อมูล 2 กลุ่ม
ที่เป็นอิสระต่อกัน

พนักงานเพศชาย 100 คน และเพศหญิง 80
คน มีการเข้าใช้งาน e-learning ของบริษัท
แตกต่างกันหรือไม่

Paired samples t-test



Is there a difference in a group between two points in time

เปรียบเทียบข้อมูล 2 กลุ่ม
ที่ไม่เป็นอิสระต่อกัน

พนักงานฝ่ายผลิตมีความรู้ก่อนและหลัง
การอบรมแตกต่างกันหรือไม่



การทดสอบค่าเอฟ (F-test)

ANOVA ทดสอบความแตกต่างของค่าเฉลี่ยของประชากร ตั้งแต่ 3 กลุ่มขึ้นไป

One-Way ANOVA มีตัวแปร 1 ตัว

นักศึกษา 4 ชั้นปี
ใช้ชั่วโมงการศึกษาด้วยตนเองแตกต่างกันหรือไม่

พนักงานแผนกผลิต 3 กะ
มีผลผลิตแตกต่างกันหรือไม่

ผลของยา 4 ประเภทต่อระดับน้ำตาลในเลือด

Two-Way ANOVA มีตัวแปร 2 ตัว

นักศึกษา 4 ชั้นปี จาก 4 คณะ
ใช้ชั่วโมงการศึกษาด้วยตนเองแตกต่างกันหรือไม่

พนักงานแผนกผลิต 3 กะ จากโรงงานผลิต 3 โรงงาน
มีผลผลิตแตกต่างกันหรือไม่

ขนาดสินค้า (3 ขนาด) และสีของสินค้า (5 สี)
ที่มีผลต่อยอดขายหรือไม่



Factor ปัจจัย

Treatment ระดับของปัจจัย สิ่งหรือวิธีปฏิบัติต่อวัตถุทดลอง เพื่อใช้ในการวัดผล เปรียบเทียบผล

Experimental Unit หน่วยทดลอง วัตถุหรือสิ่งที่ใช้ในการทดลอง

Response Measurement/Observation ค่าสังเกต ค่าที่เก็บรวบรวมได้จากการทดลอง เป็นข้อมูลนำมาเปรียบเทียบ

เปรียบเทียบผลของความดัน จากวิธีการรักษา 4 วิธี โดยทดลองกับคนไข้ 40 คน ใช้ 10 คน ต่อวิธีการรักษาโรค 1 วิธี

หน่วยทดลอง (ประชากร)	คนไข้
Factor (ตัวแปรต้น)	วิธีการรักษา
Treatment	4 วิธี
Response Measurement (ตัวแปรตาม)	ผลของความดัน

ตัวแปรต้น หรือตัวแปรอิสระ (Independent Variable) คือสิ่งที่กำหนดขึ้น เพื่อศึกษาว่าส่งผลอย่างไรต่อสิ่งอื่น

ตัวแปรตาม (Dependent Variable) คือ ผลลัพธ์ หรือ ผลที่เกิดขึ้น จากการเปลี่ยนแปลงของตัวแปรต้นที่ต้องการวัดและสังเกต

	สัญลักษณ์ ค่าพารามิเตอร์	สัญลักษณ์ ค่าสถิติ
ค่าเฉลี่ย(Mean)	μ (มิว)	\bar{x}
ค่าสัดส่วน(Proportion)	π (พาย) หรือ P	p
ส่วนเบี่ยงเบนมาตรฐาน(Standard Deviation)	σ (ซิกมา)	s หรือ s.d.
ความแปรปรวน(Variance)	σ^2	s^2 หรือ s.d. ²
สัมประสิทธิ์สหสัมพันธ์(Correlation Coefficient)	ρ (โร)	r
สัมประสิทธิ์การถดถอย(Regression Coefficient)	β (เบต้า)	b



ทดสอบ 20 คะแนน

เรื่อง การเลือกใช้สถิติทดสอบสมมติฐาน



Probability and Statistics for Engineers

รศ.ดร.ฤดี นิยมรัตน์



Probability and Statistics
for Engineers