



Sample size determination using G*Power

Associate Professor Pongdech Sarakarn, PhD

Department of Epidemiology and Biostatistics, Faculty of Public Health, Khon Kaen University





Topics of presentation



ประเด็นในการนำเสนอ

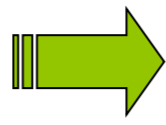
- ①. แนวคิดและหลักการในการประมาณค่า
ขนาดตัวอย่างในงานวิจัย
- ②. แนะนำโปรแกรม G*Power 3.1
- ③. กรณีศึกษา
 - >> การเปรียบเทียบค่าเฉลี่ยและค่าสัดส่วน
 - >> การวิเคราะห์ความแปรปรวน (ANOVA)
 - >> สมการถดถอยพหุคูณ (multiple linear regression)
 - >> สมการถดถอยลอจิสติก (logistic regression)



Concept of sample size determination



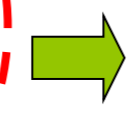
1. แนวคิดและหลักการในการประมาณค่าขนาดตัวอย่าง ในงานวิจัย



- ◎ คนให้ทั่วหม้อ
- ◎ ตักชิมบางส่วน



“หากต้องการเจาะเลือดเพื่อตรวจหาเชื้อ HIV”



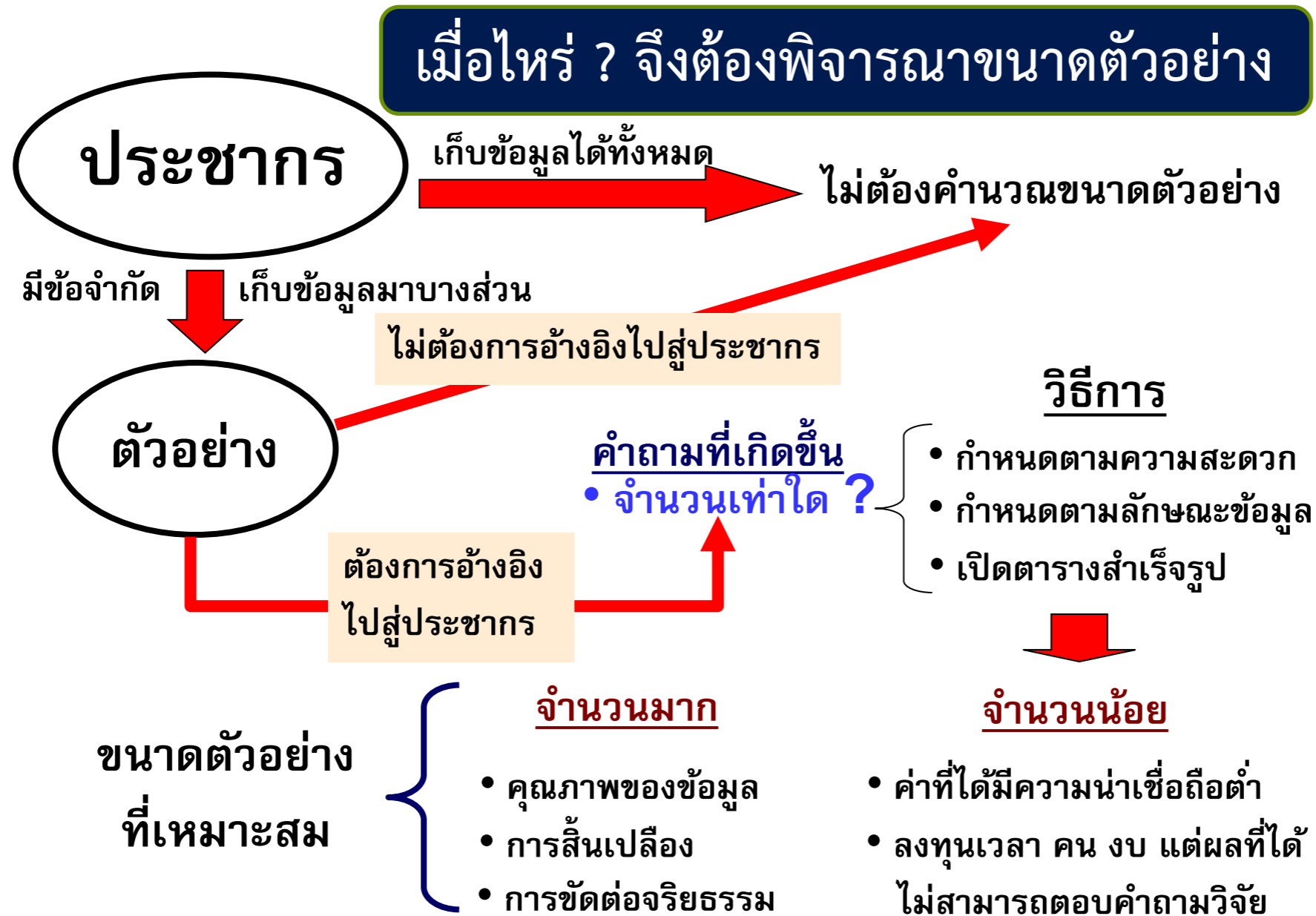
มีเหตุผลยืนยันว่า ระบบเลือดมีการไหลเวียน
ไปทั่วร่างกาย



การเจาะเลือดเอามาตรวจเพียงส่วนหนึ่ง ก็น่าจะบอกได้ว่า มีเชื้อ HIV หรือไม่



Concept of sample size determination





Concept of sample size determination



อำนาจการทดสอบคืออะไร ? ทำไม ? จึงมีความสำคัญ

การทดสอบสมมติฐานทางสถิติ

Null hypothesis (H_0)

ไม่แตกต่าง

Alternative hypothesis (H_A)

แตกต่าง

ไม่ปฏิเสธ H_0

“แตกต่าง อย่างไม่มีนัยสำคัญทางสถิติ (not statistical significance)”

ความเป็นจริงที่เป็นไปได้ในทางปฏิบัติ

ไม่แตกต่างจริง

แตกต่าง แต่ขนาดตัวอย่างไม่เพียงพอในการบ่งชี้ความแตกต่าง



Concept of sample size determination



อำนาจการทดสอบคืออะไร ? ทำไม ? จึงมีความสำคัญ

อำนาจการทดสอบ (Power of the test)

ความสามารถของสถิติที่นำมาใช้ทดสอบ ในการบ่งชี้ความสัมพันธ์ หรือ ความแตกต่าง



ความน่าจะเป็นที่จะปฏิเสธ H_0 เมื่อ H_0 ไม่จริง $[1-\beta]$

ในทางปฏิบัติจะยอมรับที่ 80% หรือ เมื่อ $\beta=0.2$



Concept of sample size determination



ขนาดผลกระทบ (Effect size)

ขนาดผลกระทบ (effect size)



ตัวบ่งชี้ขนาดของความสัมพันธ์ หรือ ความแตกต่าง

เช่น ถ้าหาความสัมพันธ์ระหว่างตัวแปรต่อเนื่องสองตัว



r

ถ้าเป็น multiple linear regression



b_0, b_1, \dots

ถ้าเป็น mean difference of t-test



Cohen's D

ถ้าเป็น ANOVA



Eta

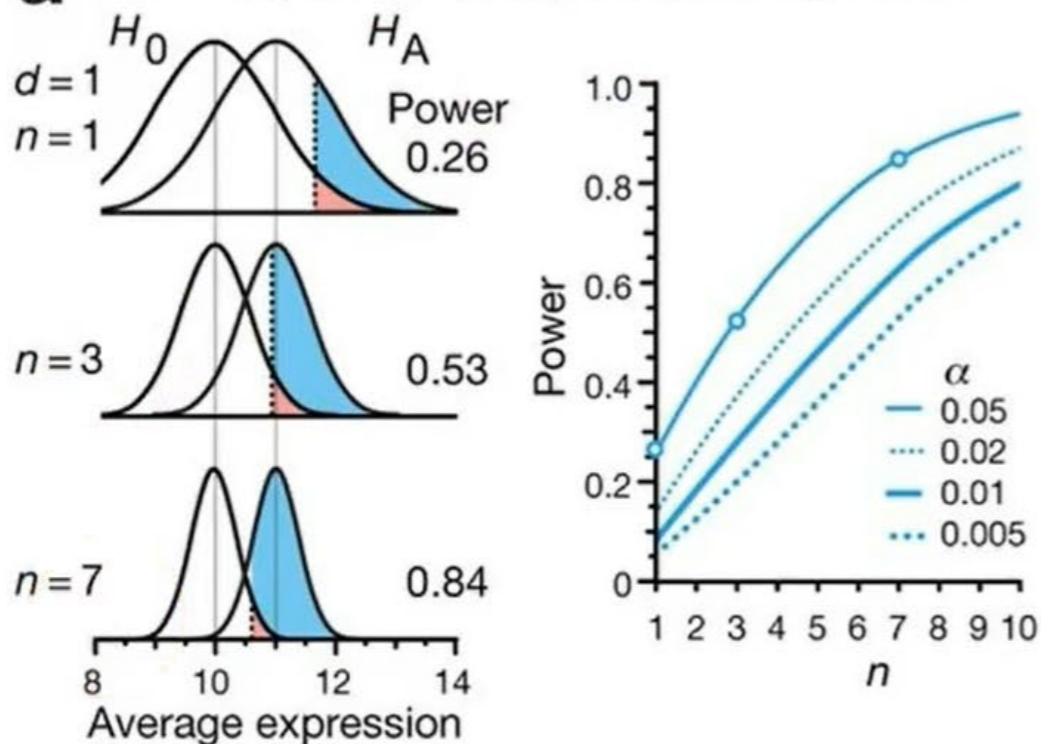


Concept of sample size determination

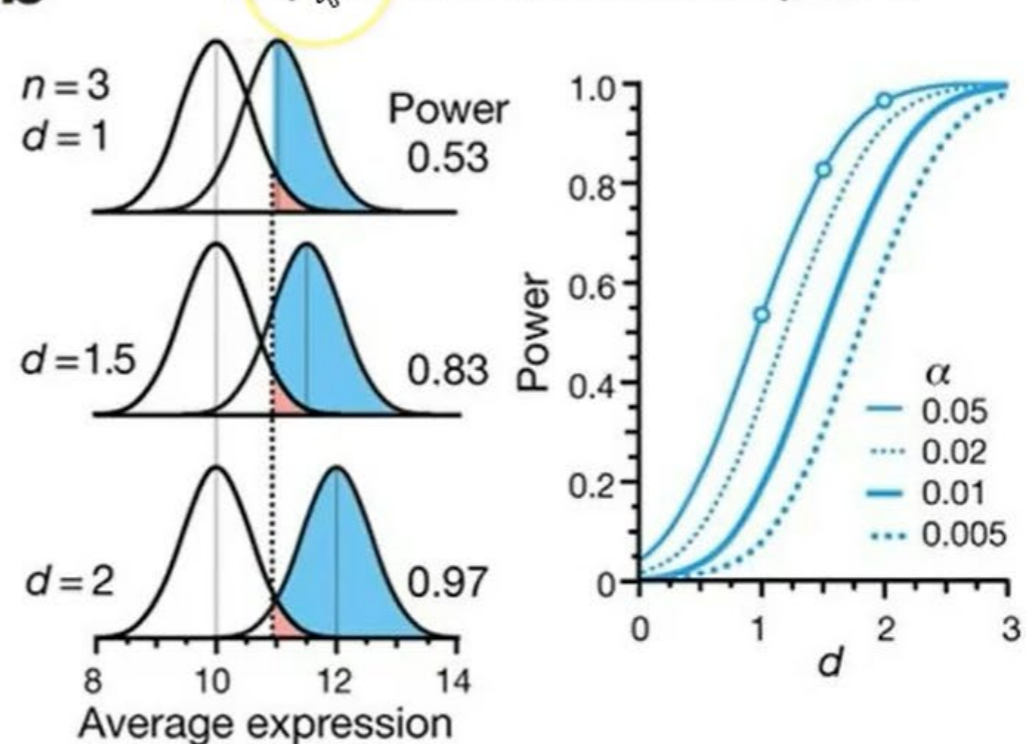


ความสัมพันธ์ระหว่างขนาดตัวอย่าง, ขนาดผลกระทบ
ที่มีผลต่ออำนาจการทดสอบ

a Impact of sample size on power



b Impact of effect size on power

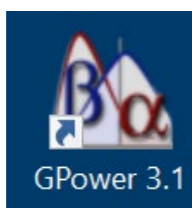




Introduction of G*Power (free software)



๒. แนะนำโปรแกรม G*Power



สามารถ download ได้ฟรี :

<https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>



Allgemeine Psychologie und Arbeitspsychologie ▾

G*Power

Statistical Power Analyses for Mac and Windows

G*Power is a tool to compute statistical power analyses for many different t tests, F tests, χ^2 tests, z tests and some exact tests. G*Power can also be used to compute effect sizes and to display graphically the results of power analyses.



Screenshots (click to enlarge)



Download

By downloading G*Power you agree to these terms of use:

1. G*Power is free for everyone. Commercial distribution is strictly prohibited.
2. G*Power is distributed from this website. If you wish to distribute G*Power in some other way, then you need to seek permission from the authors. Please send us an e-mail in which you specify how and for what purpose you intend to distribute G*Power.
3. You may use screenshots of G*Power without asking for permission.
4. Considerable effort has been put into program development and evaluation, but there is no warranty whatsoever.

เลื่อนลงมา



Link for download



- 🔗 [Download G*Power 3.1.9.7 for Windows XP, Vista, 7, 8, and 10 \(about 20 MB\)](#). Please make sure to choose "unpack with folders" in your unzip tool.
- 🔗 [Download G*Power 3.1.9.6 for Mac OS X 10.7 to 11 \(about 2 MB\)](#).



Introduction of G*Power (free software)



กลุ่มสถิติทดสอบ

- Exact
- F tests
- t tests
- χ^2 tests
- z tests

วิธีการทางสถิติในการทดสอบ

- Correlation: Bivariate normal model
- Linear multiple regression: Random model
- Proportion: Difference from constant (binomial test, one sample case)
- Proportions: Inequality, two dependent groups (McNemar)
- Proportions: Inequality, two independent groups (Fisher's exact test)
- Proportions: Inequality, two independent groups (unconditional)
- Proportions: Inequality (offset), two independent groups (unconditional)
- Proportion: Sign test (binomial test)
- Generic binomial test

ส่วนที่ใช้ในการป้อนข้อมูลพารามิเตอร์ที่เกี่ยวข้อง

ส่วนที่แสดงผลลัพธ์ของพารามิเตอร์ที่เกี่ยวข้อง

ชนิดของการวิเคราะห์อำนาจการทดสอบ

- A priori: Compute required sample size - given α , power, and effect size
- Compromise: Compute implied α & power - given β/α ratio, sample size, and effect size
- Criterion: Compute required α - given power, effect size, and sample size
- Post hoc: Compute achieved power - given α , sample size, and effect size
- Sensitivity: Compute required effect size - given α , power, and sample size

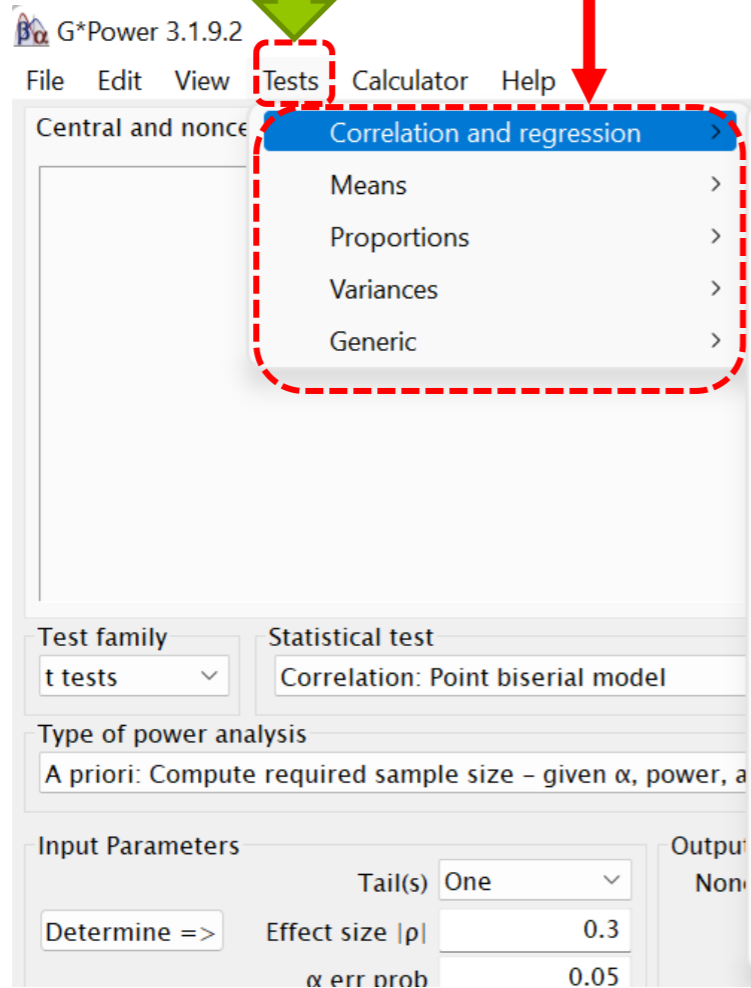


Introduction of G*Power (free software)



การเลือกวิธีตามแนวทางการวิเคราะห์ในงานวิจัย

กรณี การวิเคราะห์สหสัมพันธ์และสมการถดถอย



วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง

- Correlation: Bivariate normal model
- Correlation: Point biserial model
- Correlation: Tetrachoric model
- Correlations: Two dependent Pearson r's (common index)
- Correlations: Two dependent Pearson r's (no common index)
- Correlations: Two independent Pearson r's
- Linear bivariate regression: One group, size of slope
- Linear bivariate regression: Two groups, difference between intercepts
- Linear bivariate regression: Two groups, difference between slopes
- Linear multiple regression: Fixed model, R^2 deviation from zero
- Linear multiple regression: Fixed model, R^2 increase
- Linear multiple regression: Fixed model, single regression coefficient
- Linear multiple regression: Random model
- Logistic regression
- Poisson regression



Introduction of G*Power (free software)



การเลือกวิธีตามแนวทางการวิเคราะห์ในงานวิจัย

กรณี การวิเคราะห์เกี่ยวกับค่าเฉลี่ย

วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง

The screenshot shows the G*Power 3.1.9.2 interface. The 'Tests' menu is open, showing options: Central and noncentral, Correlation and regression, Means, Proportions, Variances, and Generic. The 'Means' option is selected. Below the menu, the 'Statistical test' is set to 'Correlation: Point biserial model'. The 'Type of power analysis' is 'A priori: Compute required sample size - given alpha, power, and effect size'. The 'Input Parameters' section shows 'Effect size |rho|' as 0.3, 'alpha err prob' as 0.05, and 'Power (1-beta err prob)' as 0.95. The 'Output' section is partially visible, showing 'Non...'. On the right, a list of statistical tests is displayed, including: One group: Difference from constant; One group: Wilcoxon (non-parametric); Two dependent groups (matched pairs); Two dependent groups (matched pairs): Wilcoxon (non-parametric); Two independent groups; Two independent groups: Wilcoxon (non-parametric); Many groups: ANCOVA: Main effects and interactions; Many groups: ANOVA: One-way (one independent variable); Many groups: ANOVA: Main effects and interactions (two or more independent variables); Repeated measures: Between factors, ANOVA-approach; Repeated measures: Between factors, MANOVA-approach; Repeated measures: Within factors, ANOVA-approach; Repeated measures: Within factors, MANOVA-approach; Repeated measures: Within-between interactions, ANOVA-approach; Repeated measures: Within-between interactions, MANOVA-approach; Multivariate: Hotelling T², one group; Multivariate: Hotelling T², two groups; Multivariate: MANOVA: Global effects; Multivariate: MANOVA: Special effects and interactions.



Introduction of G*Power (free software)



การเลือกวิธีตามแนวทางการวิเคราะห์ในงานวิจัย

กรณี การวิเคราะห์เกี่ยวกับค่าสัดส่วน

The screenshot shows the G*Power 3.1.9.2 interface. A green arrow points to the 'Tests' menu, and a red arrow points to the 'Proportions' option. A red dashed box highlights the 'Proportions' menu and the resulting list of tests. The list includes:

- One group: Difference From Constant
- One group: Sign Test
- Two dependent groups: Inequality, McNemar test
- Two independent groups: Inequality, Fisher's exact test
- Two independent groups: Inequality, unconditional exact
- Two independent groups: Inequality with offset, unconditional exact
- Two independent groups: Inequality, z-Test
- Multigroup: Goodness-of-Fit

At the bottom of the interface, the 'Test family' is set to 't tests' and the 'Statistical test' is 'Correlation: Point biserial model'.

วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง



Introduction of G*Power (free software)



การเลือกวิธีตามแนวทางการวิเคราะห์ในงานวิจัย

กรณี การวิเคราะห์เกี่ยวกับค่าความแปรปรวน

วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง

The screenshot shows the G*Power 3.1.9.2 software interface. The 'Tests' menu is open, and 'Variances' is selected. The 'One group' and 'Two groups' options are also visible. A green arrow points to the 'Tests' menu, and a red arrow points to the 'Variances' option. A blue box contains the text 'กรณี การวิเคราะห์เกี่ยวกับค่าความแปรปรวน' (Case: Analysis of variance). A red dashed box highlights the 'Tests' menu and its sub-options. Another red dashed box highlights the 'One group' and 'Two groups' options. The text 'วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง' (Statistical methods in testing related) is written next to the 'One group' and 'Two groups' options.



Introduction of G*Power (free software)



การเลือกวิธีตามแนวทางการวิเคราะห์ในงานวิจัย

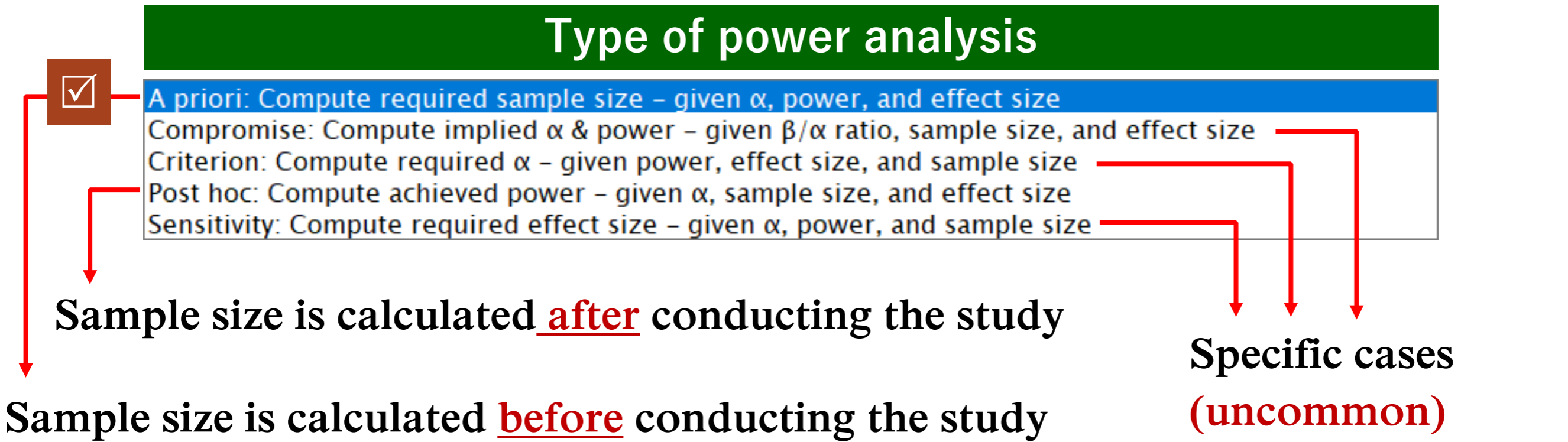
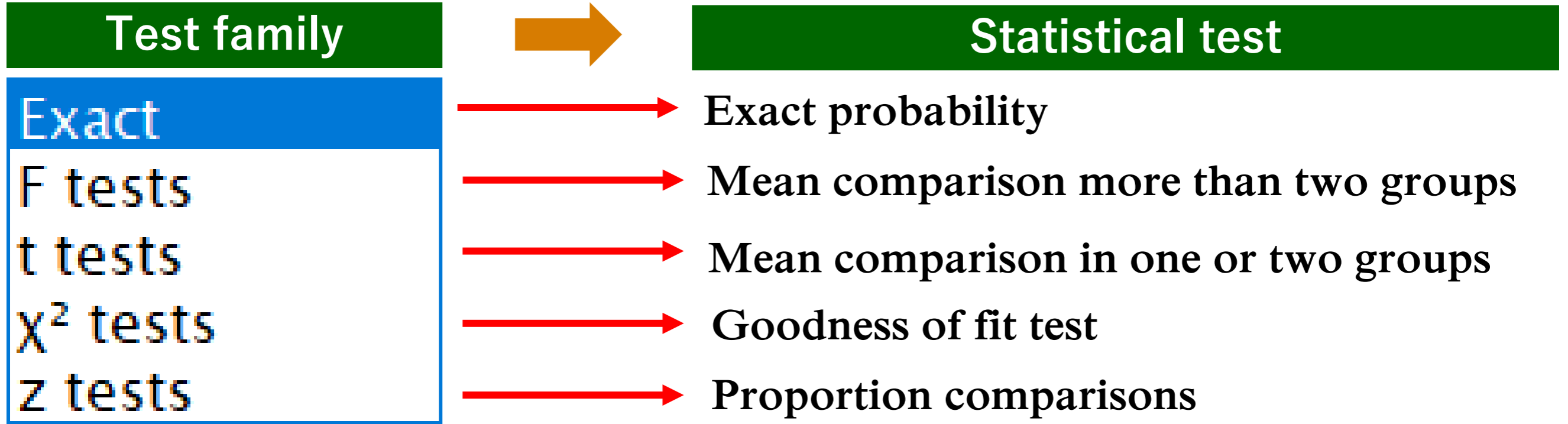
กรณี การวิเคราะห์เกี่ยวกับ Generic

วิธีการทางสถิติในการทดสอบที่เกี่ยวข้อง

- Generic binomial test
- Generic F
- Generic t
- Generic Chi²
- Generic z



Important components





Important components (*cont...*)



Input parameters

Input Parameters

	Tail(s)	One
Determine =>	Effect size d	0.5
	α err prob	0.05
	Power (1 - β err prob)	0.95
	Allocation ratio N2/N1	1

One
Two

$H_A : > \text{ or } <$
 $H_A : \neq$

Significant level
(normally, $\alpha=0.05$)

Power of the test
(normally, $\beta=0.2, 1-\beta=0.80$)

Balanced two samples
(normally, this ratio is one)

Effect size conventions
w = .10 - small
w = .30 - medium
w = .50 - large

<input type="radio"/>	n1 != n2	
Mean group 1	0	
Mean group 2	1	
SD σ within each group	0.5	
<input checked="" type="radio"/>	n1 = n2	
Mean group 1	0	
Mean group 2	1	
SD σ group 1	0.5	
SD σ group 2	0.5	
Calculate	Effect size d	?
Calculate and transfer to main window		
Close		

(recommend)

(not recommend)



Examples of calculating



- 1. For comparison between two groups**
(independent t-test two samples)
- 2. For multivariable analysis in continuous outcome**
(multiple linear regression)
- 3. For multivariable analysis in binary outcome**
(multiple logistic regression or binary logistic regression)

Example

①. For comparison between two groups

(Independent t-test two samples)



Examples of calculating



1. For comparison between two groups (independence)

Sample size for independent t-test two samples

Ex. A researcher would like to determine the sample size for his study as this research question

“Patients between group which is received drug A and group is not received drug A , after 2 month, the level of uric acid will be different or not ? ”

$$H_0 : \mu_{\text{received}} = \mu_{\text{not received}}$$

$$H_A : \mu_{\text{received}} \neq \mu_{\text{not received}}$$

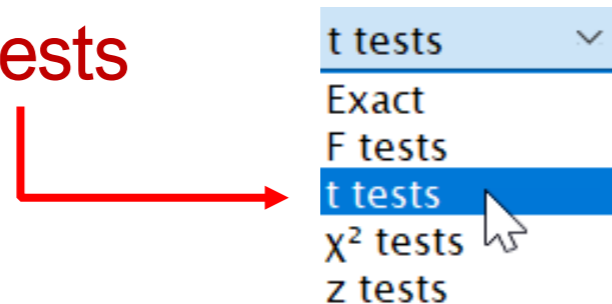


Sample size for t-test (independent 2 groups)



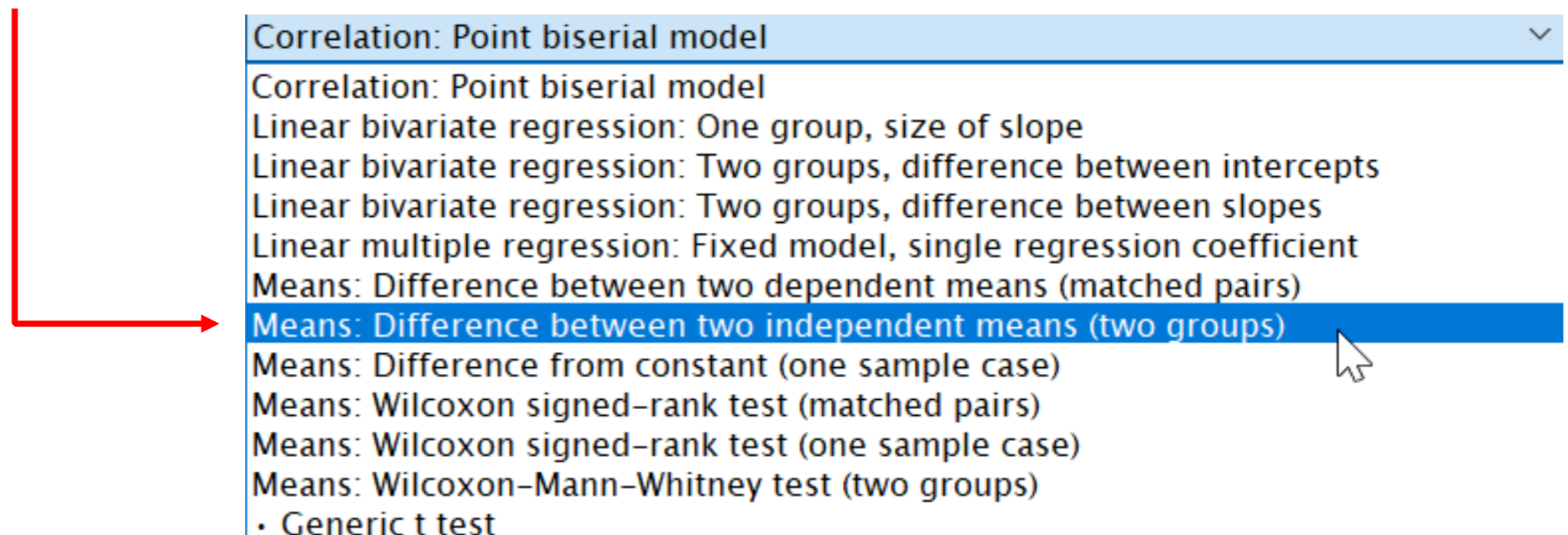
Step1 : Test family

- Selecting **t-tests**



Step2 : Statistical test

- Selecting **Means: Difference between two independent means (two groups)**



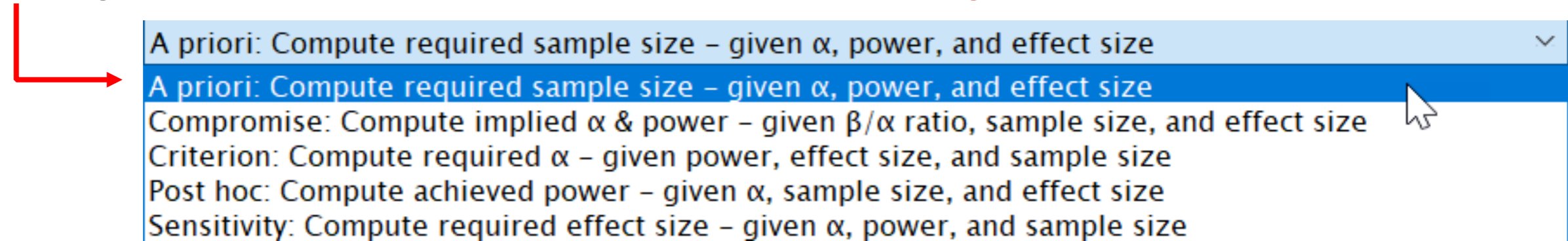


Sample size for t-test (independent 2 groups)



Step3 : Type of power analysis

- Selecting **A priori**: Compute required sample size – given α , power, and effect size



Step4 : Input parameters

- Input tail (s)'s box : **Two** [$H_A : \neq$]
- Input α err prob's box : **0.05**
- Input power (1 - β err prob)'s box : **0.80**
- Input allocation ratio N2/N1's box : **1**

Input Parameters

Tail(s)	Two
Effect size d	
α err prob	0.05
Power (1- β err prob)	0.80
Allocation ratio N2/N1	1

Determine =>

Click and input some parameters to calculate the effect size



Sample size for t-test (independent 2 groups)



Determine =>



To search the previous study or pilot study which related with his study as below;

Patient (received A)

Patient (not received A)

n	Mean	SD
12	5.34	2.92
12	8.91	4.77

n1 != n2

Mean group 1: 0

Mean group 2: 1

SD σ within each group: 0.5

n1 = n2

Mean group 1: 5.34

Mean group 2: 8.91

SD σ group 1: 2.92

SD σ group 2: 4.77

Calculate → Effect size d: 0.9027234

Calculate and transfer to main window

Close



Sample size for t-test (independent 2 groups)



Step 5 :

Calculation : click on

Calculate

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical t = 2.02108

Test family: t tests

Statistical test: Means: Difference between two independent means (two groups)

Type of power analysis: A priori: Compute required sample size - given alpha, power, and effect size

Input Parameters

Tail(s): Two

Effect size d: 0.9027234

alpha err prob: 0.05

Power (1-beta err prob): 0.80

Allocation ratio N2/N1: 1

Output Parameters

Noncentrality parameter delta: 2.9251581

Critical t: 2.0210754

Df: 40

Sample size group 1: 21

Sample size group 2: 21

Total sample size: 42

Actual power: 0.8144223

Determine =>

X-Y plot for a range of values

Calculate

GPower - Plot

File Edit View

Graph Table

t tests - Means: Difference between two independent means (two groups)
Tail(s) = Two, Allocation ratio N2/N1 = 1, alpha err prob = 0.05, Effect size d = 0.902723

Total sample size

Power (1-beta err prob)

Plot Parameters

Plot (on y axis): Total sample size

as a function of: Power (1-beta err prob)

Plot: 1 graph(s) interpolating points

with: Effect size d at 0.9027234

and: alpha err prob at 0.05

Draw plot



Sample size for t-test (independent 2 groups)



Step6 : Output : click on tab Protocol of power analyses

Central and noncentral distributions **Protocol of power analyses**

Analysis:	A priori: Compute required sample size	
Input:	Tail(s)	= Two
	Effect size d	= 0.9027234
	α err prob	= 0.05
	Power (1 - β err prob)	= 0.80
	Allocation ratio N2/N1	= 1
Output:	Noncentrality parameter δ	= 2.9251581
	Critical t	= 2.0210754
	Df	= 40
	Sample size group 1	= 21
	Sample size group 2	= 21
	Total sample size	= 42
	Actual power	= 0.8144223



The results of sample size determination for reporting in the proposal

Example

- ②. **For multivariable analysis in continuous outcome**
(Multiple linear regression)



Examples of calculating



2. For multivariable analysis in continuous outcome

Sample size for multiple linear regression

Ex. A researcher would like to determine the sample size for his study as this research question

“Factor A, B, C and D can affect (or predict) the level of uric acid or not ? ”

Factor A, B, C and D as **independent variables**

Level of uric acid (Y) as **outcome or dependent variable**

$$H_0 : \rho^2_{Y.A.B.C.D} = 0$$

$$H_A : \rho^2_{Y.A.B.C.D} > 0$$

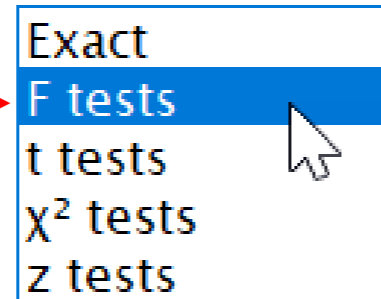


Sample size for multiple linear regression



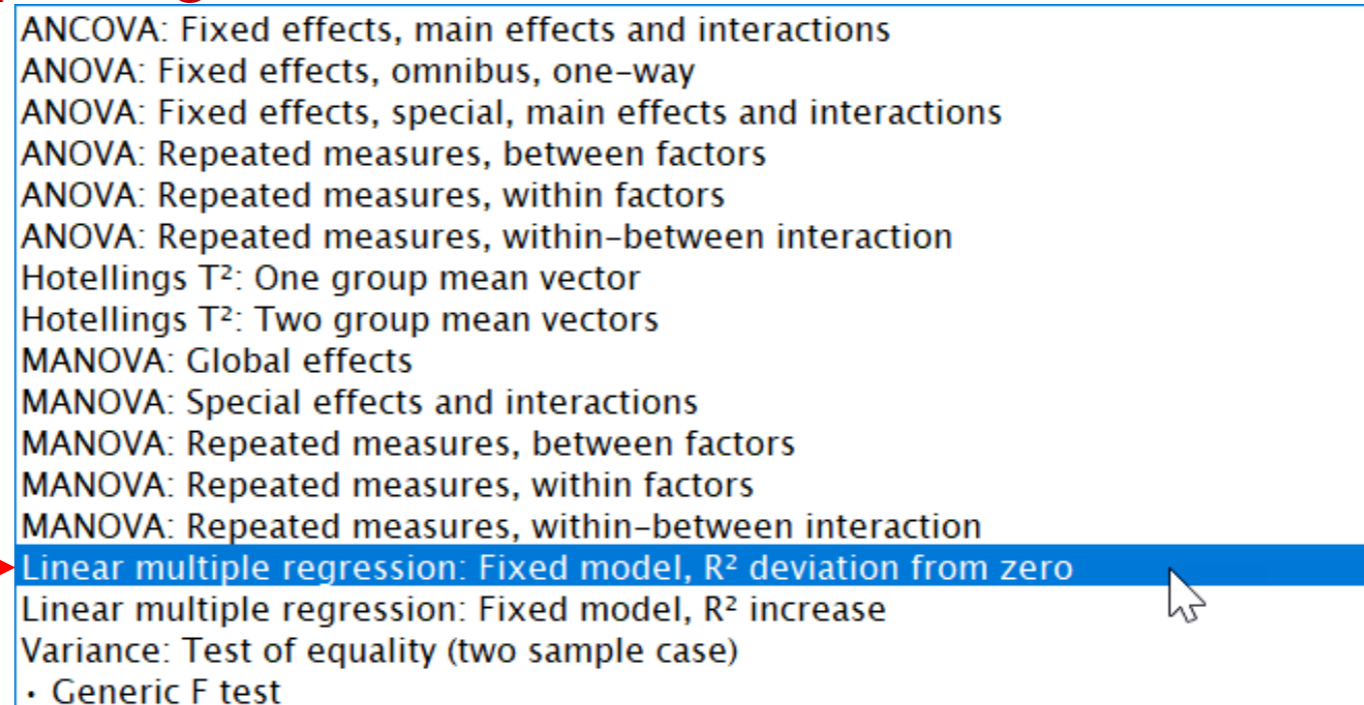
Step1 : Test family

- Selecting **F-tests**



Step2 : Statistical test

- Selecting **Linear multiple regression: Fixed model, R^2 deviation from zero**





Sample size for multiple linear regression

Step3 : Type of power analysis

- Selecting a priori: Compute required sample size – given α , power, and effect size

A priori: Compute required sample size – given α , power, and effect size

A priori: Compute required sample size – given α , power, and effect size

Compromise: Compute implied α & power – given β/α ratio, sample size, and effect size

Criterion: Compute required α – given power, effect size, and sample size

Post hoc: Compute achieved power – given α , sample size, and effect size

Sensitivity: Compute required effect size – given α , power, and sample size

Step4 : Input parameters

- Input α err prob's box : **0.05**
- Input power (1 - β err prob)'s box : **0.80**
- Input number of predictors 's box : **4**

Input Parameters

Determine =>	Effect size f^2	
	α err prob	0.05
	Power (1- β err prob)	0.80
	Number of predictors	4

Click and input some parameters to calculate the effect size



Sample size for multiple linear regression



Determine =>

1. Using R – square (R^2)

$$f^2 = \frac{R^2_{Y.A.B.C.D}}{1 - R^2_{Y.A.B.C.D}}$$

Effect size

We would like to know R^2 from previous study which related with our study

In this example, we suppose that $R^2 = 0.12$

From correlation coefficient
 Squared multiple correlation ρ^2

From predictor correlations
 Number of predictors
 Squared multiple correlation ρ^2

Specify matrices



Sample size for multiple linear regression



Step 5 :

Calculation : click on

Calculate

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical F = 2.47528

Test family: F tests

Statistical test: Linear multiple regression: Fixed model, R² deviation from zero

Type of power analysis: A priori: Compute required sample size - given alpha, power, and effect size

Input Parameters		Output Parameters	
Determine =>	Effect size f ²	0.1363636	Noncentrality parameter lambda
	alpha err prob	0.05	Critical F
	Power (1-beta err prob)	0.80	Numerator df
	Number of predictors	4	Denominator df
			Total sample size
			Actual power

X-Y plot for a range of values

Calculate

GPower - Plot

File Edit View

Graph Table

F tests - Linear multiple regression: Fixed model, R² deviation from zero
Number of predictors = 4, alpha err prob = 0.05, Effect size f² = 0.136364

Power (1-beta err prob)	Total sample size
0.60	63
0.61	64
0.62	65
0.63	67
0.64	68
0.65	69
0.66	70
0.67	72
0.68	73
0.69	74
0.70	76
0.71	77
0.72	79
0.73	80
0.74	82
0.75	84
0.76	85
0.77	87
0.78	89
0.79	90
0.80	92
0.81	94
0.82	96
0.83	99
0.84	101
0.85	103
0.86	106
0.87	109
0.88	111
0.89	114
0.90	118
0.91	121
0.92	125
0.93	130
0.94	135
0.95	141

Plot Parameters

Plot (on y axis): Total sample size

as a function of: Power (1-beta err prob)

Plot 1 graph(s) interpolating points

with Effect size f² at 0.1363636

and alpha err prob at 0.05

Draw plot



Sample size for multiple linear regression



Step6 : Output : click on tab Protocol of power analyses

Central and noncentral distributions Protocol of power analyses

F tests – Linear multiple regression: Fixed model, R^2 deviation from zero

Analysis: A priori: Compute required sample size

Input: Effect size f^2 = 0.1363636

α err prob = 0.05

Power ($1 - \beta$ err prob) = 0.80

Number of predictors = 4

Output: Noncentrality parameter λ = 12.6818148

Critical F = 2.4752774

Numerator df = 4

Denominator df = 88

Total sample size = 93

Actual power = 0.8030926



The results of sample size determination for reporting in the proposal



Sample size for multiple linear regression



Determine =>

2. Using predictors correlation matrix

We would like to know predictors correlation matrix from previous study which related with our study

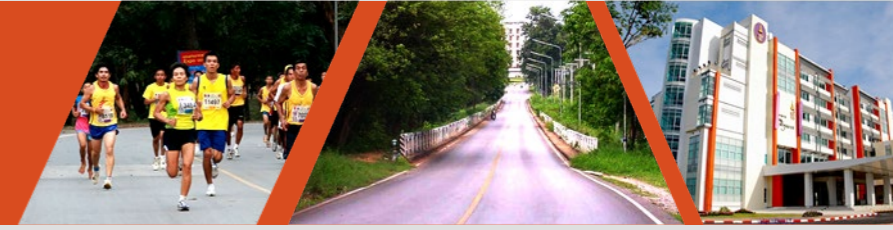
In this example, we suppose that

	Y	A	B	C	D
Y	1.0000				
A	-0.0244	1.0000			
B	0.4057	-0.3775	1.0000		
C	-0.1762	0.1706	0.0449	1.0000	
D	-0.0125	0.3788	0.0402	0.2666	1.0000

predictor	P 1	P 2	P 3	P 4
corr with outcome Y	-0.0244	0.4057	-0.1762	-0.0125



Sample size for multiple linear regression



Input predictor correlations

Corr between predictors and outcome | Corr between predictors

Number of predictors: 4

predictor	P 1	P 2	P 3	P 4
corr with outcome Y	-0.0244	0.4057	-0.1762	-0.0125

Calc ρ^2 Coefficient ρ^2 0.196391 Accept values Cancel

From correlation coefficient

Squared multiple correlation ρ^2 0.5

From predictor correlations

Number of predictors 4

Squared multiple correlation ρ^2 0.1963905

Specify matrices

Calculate → Effect size f^2 0.2443856

Calculate and transfer to main window

Close



Sample size for multiple linear regression



Step 5 :

Calculation : click on

Calculate

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical F = 2.56112

Test family: F tests

Statistical test: Linear multiple regression: Fixed model, R² deviation from zero

Type of power analysis: A priori: Compute required sample size - given alpha, power, and effect size

Input Parameters		Output Parameters	
Determine =>	Effect size f ²	0.2443856	Noncentrality parameter lambda
	alpha err prob	0.05	Critical F
	Power (1-beta err prob)	0.80	Numerator df
	Number of predictors	4	Denominator df
			Total sample size
			Actual power

X-Y plot for a range of values

Calculate

GPower - Plot

File Edit View

Graph Table

F tests - Linear multiple regression: Fixed model, R² deviation from zero
Number of predictors = 4, alpha err prob = 0.05, Effect size f² = 0.244386

Power (1-beta err prob)	Total sample size
0.60	37
0.61	38
0.62	39
0.63	39
0.64	40
0.65	41
0.66	42
0.67	42
0.68	43
0.69	44
0.70	45
0.71	45
0.72	46
0.73	47
0.74	48
0.75	49
0.76	50
0.77	51
0.78	52
0.79	53
0.80	54
0.81	55
0.82	56
0.83	57
0.84	59
0.85	60
0.86	61
0.87	63
0.88	64
0.89	66
0.90	68
0.91	70
0.92	72
0.93	75
0.94	78
0.95	81

Plot Parameters

Plot (on y axis): Total sample size

as a function of: Power (1-beta err prob)

Plot 1 graph(s): interpolating points

with: Effect size f² at 0.2443856

and: alpha err prob at 0.05

Draw plot



Sample size for multiple linear regression



Step6 : Output : click on tab Protocol of power analyses

Central and noncentral distributions Protocol of power analyses

F tests – Linear multiple regression: Fixed model, R^2 deviation from zero

Analysis: A priori: Compute required sample size

Input:

Effect size f^2	=	0.2443856
α err prob	=	0.05
Power ($1 - \beta$ err prob)	=	0.80
Number of predictors	=	4

Output:

Noncentrality parameter λ	=	13.1968224
Critical F	=	2.5611240
Numerator df	=	4
Denominator df	=	49
Total sample size	=	54
Actual power	=	0.8018363



The results of sample size determination for reporting in the proposal

Example

- ③. **For multivariable analysis in binary outcome**
(Binary logistic regression)



Examples of calculating



3. For multivariable analysis in binary outcome

Sample size for binary logistic regression

Ex. A researcher would like to determine the sample size for his study as this research question

“Sex can be associated with Coronary Heart Disease (CHD) or not ? when we use age, smoking and ht to control confounders ”

sex as **main risk factor**

age, smoking, ht as **covariates to control confounders**

CHD as **outcome or dependent variable**

$$p_{\text{male}} = \frac{75}{239} = 0.31$$

$$p_{\text{female}} = \frac{105}{761} = 0.14$$

$$\text{Odds}_{\text{male}} = \frac{75}{164} = 0.46$$

$$\text{Odds}_{\text{female}} = \frac{105}{656} = 0.16$$



	Alive	Death	Total
Male	164	75	239
Female	656	105	761
Total	820	180	1000

$$\text{Odds ratio}_{\text{male}} = \frac{0.46}{0.16} = 2.86$$



Sample size for binary logistic regression



Step1 : Test family

- Selecting **z tests**

Exact
F tests
t tests
 χ^2 tests
z tests

Step2 : Statistical test

- Selecting **Logistic regression**

Correlation: Tetrachoric model
Correlations: Two dependent Pearson r's (common index)
Correlations: Two dependent Pearson r's (no common index)
Correlations: Two independent Pearson r's
Logistic regression
Poisson regression
Proportions: Difference between two independent proportions
• Generic z test



Sample size for binary logistic regression

Step3 : Type of power analysis

- Selecting a priori: Compute required sample size – given α , power, and effect size

A priori: Compute required sample size – given α , power, and effect size

A priori: Compute required sample size – given α , power, and effect size

Compromise: Compute implied α & power – given β/α ratio, sample size, and effect size

Criterion: Compute required α – given power, effect size, and sample size

Post hoc: Compute achieved power – given α , sample size, and effect size

Sensitivity: Compute required effect size – given α , power, and sample size

Step4 : Input parameters

Input Parameters	
Tail(s)	Two
Determine =>	
Odds ratio	2.86
Pr(Y=1 X=1) H0	0.14
α err prob	0.05
Power (1- β err prob)	0.80
R ² other X	0.20
X distribution	Binomial
X parm π	0.5

$$\text{Odds ratio}_{\text{male}} = \frac{0.46}{0.16} = 2.86$$

$$p_{\text{female}} = \frac{105}{761} = 0.14$$

Correlation of “sex” with other covariates

Proportion of sex



Sample size for binary logistic regression



Step 5 :

Calculation : click on

Calculate

G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical z = 1.95996

Test family: z tests

Statistical test: Logistic regression

Type of power analysis: A priori: Compute required sample size - given alpha, power, and effect size

Input Parameters:

- Tail(s): Two
- Determine =>
- Odds ratio: 2.86
- Pr(Y=1|X=1) H0: 0.14
- alpha err prob: 0.05
- Power (1-beta err prob): 0.80
- R² other X: 0.20
- X distribution: Binomial
- X parm pi: 0.5

Output Parameters:

- Critical z: 1.9599640
- Total sample size: 223
- Actual power: 0.8013923

Options: X-Y plot for a range of values

Calculate

GPower - Plot

File Edit View

Graph Table

z tests - Logistic regression

Tail(s) = Two, Pr(Y=1|X=1) H0 = 0.14, R² other X = 0.20, X distribution = Binomial, X parm pi = 0.5, alpha err prob = 0.05, Odds ratio = 2.86

Power (1-β err prob)	Total sample size
0.6	142
0.61	145
0.62	148
0.63	151
0.64	155
0.65	158
0.66	162
0.67	165
0.68	169
0.69	173
0.70	176
0.71	180
0.72	185
0.73	189
0.74	193
0.75	197
0.76	202
0.77	207
0.78	212
0.79	217
0.80	222
0.81	228
0.82	234
0.83	240
0.84	246
0.85	253
0.86	260
0.87	268
0.88	276
0.89	285
0.90	294
0.91	305
0.92	317
0.93	330
0.94	344
0.95	362

Plot Parameters:

- Plot (on y axis): Total sample size
- with markers
- and displaying the values in the plot
- Show 0 digits
- as a function of: Power (1-β err prob)
- from 0.6 in steps of 0.01 through to 0.95
- Plot 1 graph(s) interpolating points
- with Odds ratio at 2.86
- and alpha err prob at 0.05

Draw plot



Sample size for binary logistic regression



Step 6 : Output : click on tab Protocol of power analyses

Central and noncentral distributions **Protocol of power analyses**

Options:	Large sample z-Test, Demidenko (2007) with var corr	
Analysis:	A priori: Compute required sample size	
Input:	Tail(s)	= Two
	Odds ratio	= 2.86
	Pr(Y=1 X=1) H0	= 0.14
	α err prob	= 0.05
	Power (1- β err prob)	= 0.80
	R ² other X	= 0.20
	X distribution	= Binomial
	X parm π	= 0.5
Output:	Critical z	= 1.9599640
	Total sample size	= 223
	Actual power	= 0.8013923



The results of sample size determination for reporting in the proposal



Sample size for binary logistic regression



In practice, we have two choices to input effect size for binary logistic regression

1. Using two probabilities to calculate odds ratio

Input effect size as ...

Odds ratio

Two probabilities



Pr(Y=1|X=1) H1

Pr(Y=1|X=1) H0

Calculate Odds ratio ?

Calculate and transfer to main window

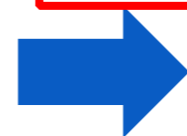
Close

2. Using odds ratio and $\Pr[Y=1|X=0]$ to calculate $\Pr[Y=1|X=1]$

Input effect size as ...

Odds ratio

Two probabilities



The same

Odds ratio

Pr(Y=1|X=1) H0

Calculate Pr(Y=1|X=1) H1

Calculate and transfer to main window

Close

Example

④. For ANOVA analysis

(One way ANOVA)

(One-way repeated measures ANOVA)



Examples of calculating



ตัวอย่างที่ 1

การเปรียบเทียบคะแนนระหว่างกลุ่ม A, B, C

นักวิจัยต้องการศึกษาคะแนนการเรียน

ในรายวิชาระเบียบวิธีวิจัยด้วยวิธี A (on-site), B (online) และ C (ทั้งแบบ on-site และ online)

คนที่	A	B	C
1	13	10	16
2	12	11	15
3	13	9	17
4	14	12	17

ถูกนำมาใช้เพื่อคำนวณหาค่า effect size



	ค่าเฉลี่ย	จำนวนตัวอย่าง
A	13	4
B	10.5	4
C	16.25	4

ข้อมูลนำร่องที่ผ่านมา

SD ของค่าในทุกกลุ่ม = 2.632



Examples of calculating



G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical F = 3.68232

One-way ANOVA

Test family: F tests

Statistical test: ANOVA: Fixed effects, omnibus, one-way

Type of power analysis: A priori: Compute required sample size - given α , power, and effect size

Input Parameters		Output Parameters		
Determine =>	Effect size f	0.8944052	Noncentrality parameter λ	14.3992919
	α err prob	0.05	Critical F	3.6823203
	Power (1- β err prob)	0.80	Numerator df	2
	Number of groups	3	Denominator df	15
			Total sample size	18
			Actual power	0.8743573

X-Y plot for a range of values Calculate

Record

Select procedure: Effect size from means

Number of groups: 3

$\Delta \sigma$ within each group: 2.632

Group	Mean	Size
1	13	4
2	10.5	4
3	16.25	4

Equal n: 4

Total sample size: 12

Calculate Effect size f: 0.8944052

Calculate and transfer to main window

Close

การใช้ข้อมูลที่ผ่านมาในการคำนวณ effect size



Examples of calculating



คลิกบน Tab นี้ เพื่อแสดงผลลัพธ์ที่ได้ ไปเขียนรายงานผลการคำนวณขนาดตัวอย่าง

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

F tests – ANOVA: Fixed effects, omnibus, one-way

Analysis: A priori: Compute required sample size

Input: Effect size f = 0.8944052

α err prob = 0.05

Power (1- β err prob) = 0.80

Number of groups = 3

Output: Noncentrality parameter λ = 14.3992919

Critical F = 3.6823203

Numerator df = 2

Denominator df = 15

Total sample size = 18

Actual power = 0.8743573



Examples of calculating



ตัวอย่างที่ 2

นักวิจัยต้องการศึกษาประสิทธิผลของโปรแกรมลดความเครียดในกลุ่มผู้ป่วยที่มีภาวะซึมเศร้าแบบหนึ่งกลุ่ม มีการวัดคะแนนระดับความเครียด 3 ครั้ง ได้แก่ ก่อนได้รับโปรแกรม, หลังได้รับโปรแกรม 1 อาทิตย์และหลังได้รับโปรแกรม 1 เดือน อยากทราบว่า ต้องใช้ขนาดตัวอย่างเท่าใด ?



One-way repeated measures ANOVA

ข้อมูลสำคัญที่จำเป็นต้องใช้

ค่า Partial eta square (Partial η^2)

ค่า Epsilon correction (Greenhouse-Geisser)



นำไปใช้คำนวณ effect size



นำไปใช้ระบุประกอบการคำนวณ
ขนาดตัวอย่างในช่อง :

Nonsphericity correction ϵ



หาได้จากงานวิจัยที่เกี่ยวข้องที่ผ่านมา



Examples of calculating



G*Power 3.1.9.2

File Edit View Tests Calculator Help

Central and noncentral distributions Protocol of power analyses

critical F = 3.75319

One-way repeated measures ANOVA

Test family: F tests

Statistical test: ANOVA: Repeated measures, within factors

Type of power analysis: A priori: Compute required sample size - given α , power, and effect size

Input Parameters		Output Parameters	
Determine =>	Effect size f	0.3333333	Noncentrality parameter λ
	α err prob	0.05	Critical F
	Power (1- β err prob)	0.80	Numerator df
	Number of groups	1	Denominator df
	Number of measurements	3	Total sample size
	Corr among rep measures	0.5	Actual power
	Nonsphericity correction ϵ	0.7	

Epsilon correction (Greenhouse-Geisser)

Options X-Y plot for a range of values Calculate

Record

เป็นค่าความสัมพันธ์ระหว่างการวัดซ้ำ ซึ่งจริงๆ ก็คือ ค่า assumption ของ compound symmetry ซึ่งเมื่อผ่านข้อตกลงนี้ ก็มักแสดงเป็น 0.5

การใช้ข้อมูลที่ผ่านมาในการคำนวณ effect size

From variances

Variance explained by effect	1.0
Variance within group	2.0

Direct Partial η^2 0.1

Calculate Effect size f 0.3333333

Calculate and transfer to main window

Close



Examples of calculating



ค่าต่างๆ ที่สามารถนำไปเขียนประกอบกรรายงานผลการคำนวณขนาดตัวอย่าง

F tests – ANOVA: Repeated measures, within factors

Analysis: A priori: Compute required sample size

Input:	Effect size f	= 0.3333333
	α err prob	= 0.05
	Power (1- β err prob)	= 0.80
	Number of groups	= 1
	Number of measurements	= 3
	Corr among rep measures	= 0.5
	Nonsphericity correction ϵ	= 0.7
Output:	Noncentrality parameter λ	= 9.7999980
	Critical F	= 3.7531914
	Numerator df	= 1.4000000
	Denominator df	= 28.0000000
	Total sample size	= 21
	Actual power	= 0.8142931



Thank you!

Contact Address:
Email: spong@kku.ac.th
www.kku.ac.th

