การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น ด้วยโปรแกรมสำเร็จรูป

SPSS for Windows & MATHCAD

เนื้อหาใช้ได้กับโปรแกรม

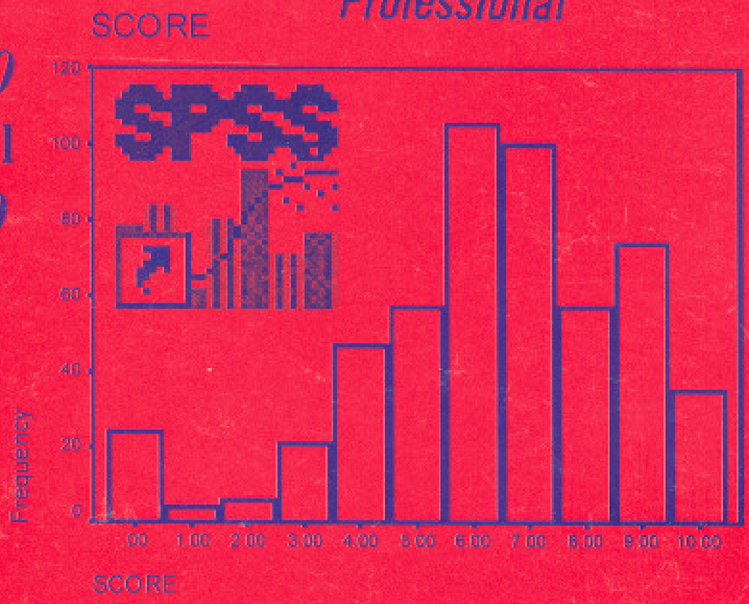
SPSS for Windows

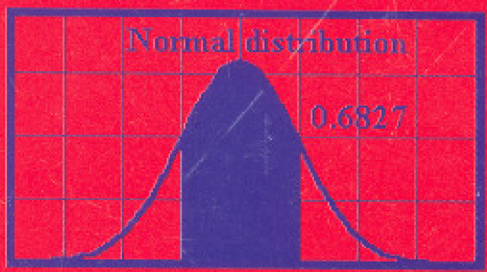
version 7.5 , 8.0 , 9.0

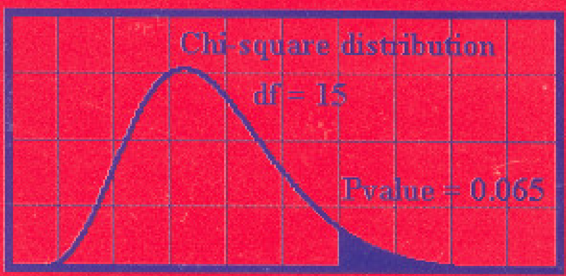
Mathcad Professional

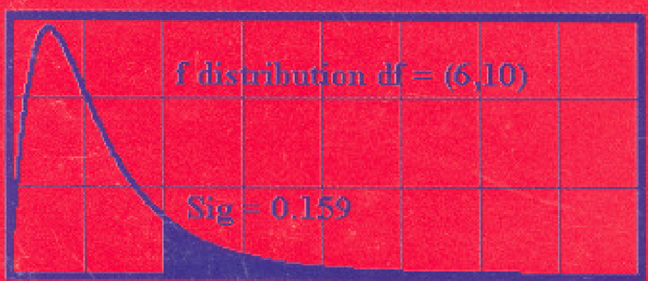
version 7 , 8 , 2000

Mathcad Professional









หนังสือเล่มนี้เหมาะสำหรับ

- ปฏิบัติการวิชาความน่าจะเป็นและสถิติ
- ปฏิบัติการ การวิเคราะห์ข้อมูลทางสถิติ
- ใช้เป็นคู่มือ SPSS for Windows 7.5
- การนำ MATHCAD ช่วยคำนวณค่าสถิติ

พิมพ์ครั้งที่ 2

ฉบับปรับปรุงเพิ่มเติม

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
จุฬาลงกรณ์มหาวิทยาลัย

ผู้เขียน **รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา**

การศึกษา วท.บ. (คณิตศาสตร์) คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

วท.ม. (คณิตศาสตร์) คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ที่ทำงาน ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ผลงานตำรา
- พีชคณิตเชิงเส้น (2542)
- พีชคณิตระดับอุดมศึกษา (2538)
- ภาษาเบสิก (2531)
- คณิตศาสตร์ขั้นสูง (2538)
- ตัวกำหนดและเมทริกซ์ :ระเบียบวิธีการคณนา (2540)
- คู่มือตัดตัวเลือกคณิตศาสตร์ ม. 4 คณิตศาสตร์ ม. ต้น
- คู่มือตัดตัวเลือกคณิตศาสตร์ GMAT และ MBA
- คู่มือตัดตัวเลือกข้อสอบคณิตศาสตร์ ม. ปลาย (ภาค 1 – ภาค 3)
- คู่มือโปรแกรมสำเร็จรูป LINDO (2538)
- คู่มือโปรแกรมสำเร็จรูป MATHCAD (2541)
- การหาผลเฉลยสมการเชิงอนุพันธ์ (2541)
- รวมปัญหาอุปนัยเชิงคณิตศาสตร์ (2541)
- สารพันปัญหาคณิตศาสตร์ (2542)
- โจทย์เสริมประสบการณ์การตัดตัวเลือกข้อสอบ Entrance ระบบใหม่

การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น ด้วยโปรแกรมสำเร็จรูป

SPSS for Windows & MATHCAD

SPSS for Windows

MATHCAD Professional

หนังสือเล่มนี้ได้รับการสนับสนุนจากโครงการส่งเสริมการผลิตตำราและหนังสือ
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2542

โปรแกรมคอมพิวเตอร์ที่สามารถใช้ประกอบกับหนังสือ

การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น

ด้วยโปรแกรมสำเร็จรูป

SPSS for Windows & MATHCAD

คือ

SPSS for Windows 7.5

SPSS for Windows 8.0

SPSS for Windows 9.0

Mathcad Professional 7

Mathcad Professional 8

Mathcad Professional 2000

การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น ด้วยโปรแกรมสำเร็จรูป

SPSS for Windows & MATHCAD

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น
ด้วยโปรแกรมสำเร็จรูป SPSS for Windows & MATHCAD

ผู้เขียน รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

พิมพ์ครั้งที่ 1 กรกฎาคม พ.ศ. 2541
พิมพ์ครั้งที่ 2 กุมภาพันธ์ พ.ศ. 2543

ISBN 974 – 639 – 311 – 1

จัดจำหน่ายโดย ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย
ถนนพญาไท กรุงเทพฯ 10330
ศาลาพระเกี้ยว โทร. 2554433 , 2187000 โทรสาร 2554441
สยามสแควร์ โทร. 2189888 , 2516141 โทรสาร 2549495
email: cubook@chula.ac.th
http://www.cubook.chula.ac.th

พิมพ์ที่โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย โทร. 2183563-4, 2153612
นายประเสริฐ ศีลพิพัฒน์ ผู้พิมพ์ผู้โฆษณา กุมภาพันธ์ 2543
4210-185/2,000(2)

คำนำ

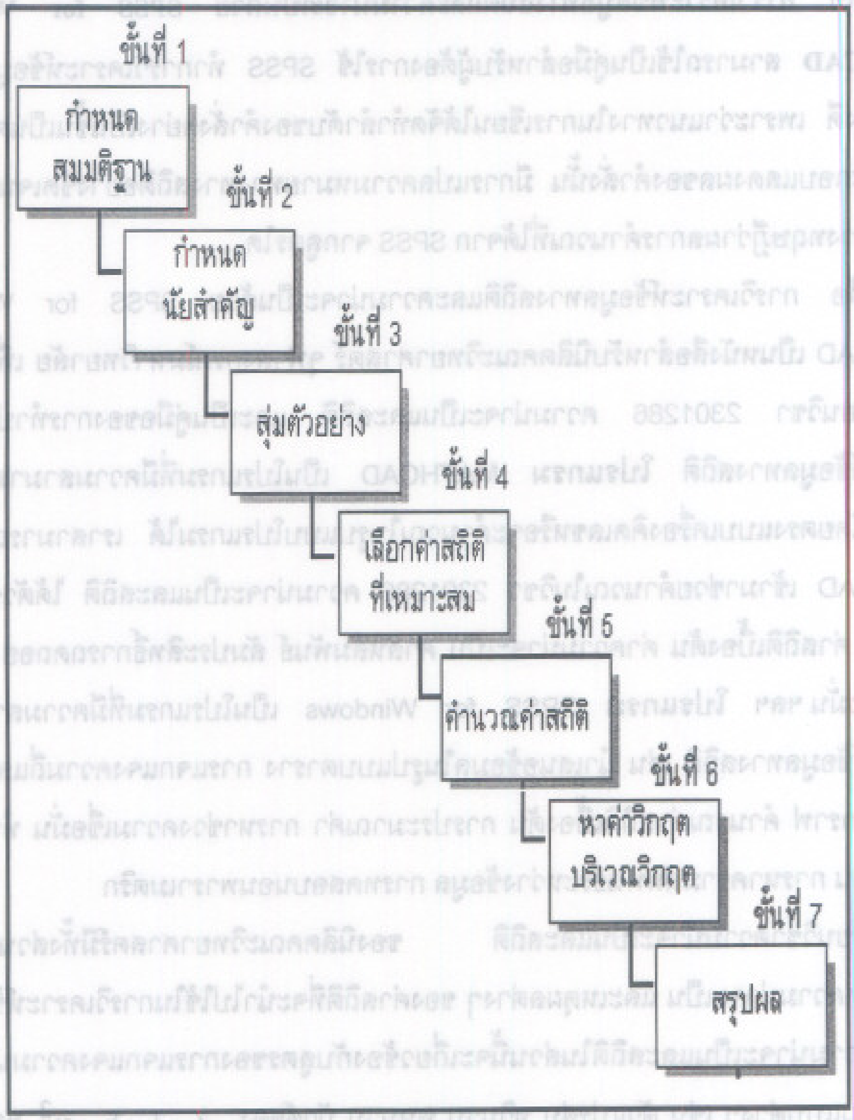
หนังสือ **การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็นด้วย SPSS for Windows & MATHCAD** สามารถใช้เป็นคู่มือสำหรับผู้ต้องการใช้ SPSS ทำการวิเคราะห์ข้อมูลทางสถิติได้เป็นอย่างดี เพราะว่าแนวทางในการเขียนได้จัดทำลำดับของคำสั่งอย่างเป็นขั้นเป็นตอน และ มีภาพประกอบแสดงผลของคำสั่งนั้น มีการแปลความหมายทางทางสถิติอย่างชัดเจน และ แสดงเหตุผลทางทฤษฎีว่าผลการคำนวณที่ได้จาก SPSS จากสูตรใด

หนังสือ การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็นด้วย SPSS for Windows & MATHCAD เป็นหนังสือสำหรับนิสิตคณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เพื่อใช้ประกอบในการเรียนวิชา 2301286 ความน่าจะเป็นและสถิติ และเป็นคู่มือของการทำปฏิบัติการการวิเคราะห์ข้อมูลทางสถิติ **โปรแกรม** MATHCAD เป็นโปรแกรมที่มีความสามารถที่จะทำการคำนวณโดยตรงแบบเครื่องคิดเลขหรือจะคำนวณในรูปแบบโปรแกรมได้ เราสามารถนำโปรแกรม MATHCAD เข้ามาช่วยคำนวณในวิชา 2301286 ความน่าจะเป็นและสถิติ ได้ตัวอย่างเช่นการคำนวณ ค่าสถิติเบื้องต้น ค่าความน่าจะเป็น ค่าสหสัมพันธ์ สัมประสิทธิ์การถดถอย คำนวณช่วงความเชื่อมั่น ฯลฯ **โปรแกรม** SPSS for Windows เป็นโปรแกรมที่มีความสามารถในการวิเคราะห์ข้อมูลทางสถิติ เช่น นำเสนอข้อมูลในรูปแบบตาราง การแจกแจงความถี่และนำเสนอข้อมูลในรูปกราฟ คำนวณค่าสถิติเบื้องต้น การประมาณค่า การหาช่วงความเชื่อมั่น ทำการทดสอบสมมติฐาน การหาความสัมพันธ์ระหว่างข้อมูล การทดสอบนอนพาราเมตริก

การเรียนวิชาความน่าจะเป็นและสถิติ ของนิสิตคณะวิทยาศาสตร์มีทั้งส่วนของหลักการทฤษฎีบทความน่าจะเป็น และเหตุผลต่างๆ ของค่าสถิติที่จะนำไปใช้ในการวิเคราะห์ข้อมูล ซึ่งเนื้อหาของความน่าจะเป็นและสถิติในส่วนนี้จะเกี่ยวข้องกับสูตรของการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่างๆ เช่น ตัวแปรสุ่ม ทวินาม พหุนาม ปัวส์ซอง , Z , T , F , χ^2 ฯลฯ เนื้อหาในส่วนนี้จะใช้โปรแกรม MATHCAD เข้ามาช่วยในการคำนวณเมื่อนิสิตมีความเข้าใจมากยิ่งขึ้น เมื่อนิสิตได้เรียนหลักการวิเคราะห์ข้อมูลทางสถิติตามเนื้อหาในวิชา ความน่าจะเป็นและสถิติ แล้วโปรแกรมสำเร็จรูป SPSS for Windows ซึ่งมีความสามารถในการวิเคราะห์ข้อมูลทางสถิติ จะทำให้นิสิตสามารถทำการวิเคราะห์ข้อมูลที่มีจำนวนมากๆ ได้สะดวกและรวดเร็วมากยิ่งขึ้น

รองศาสตราจารย์ ดำรงค์ ทิพย์โยธา

ขั้นการทดสอบสมมติฐานโดยสังเขป



หมายเหตุ ค่าสถิติที่เหมาะสมที่นิยมใช้เช่น Z T F และ χ^2

สารบัญ

หน้าที่

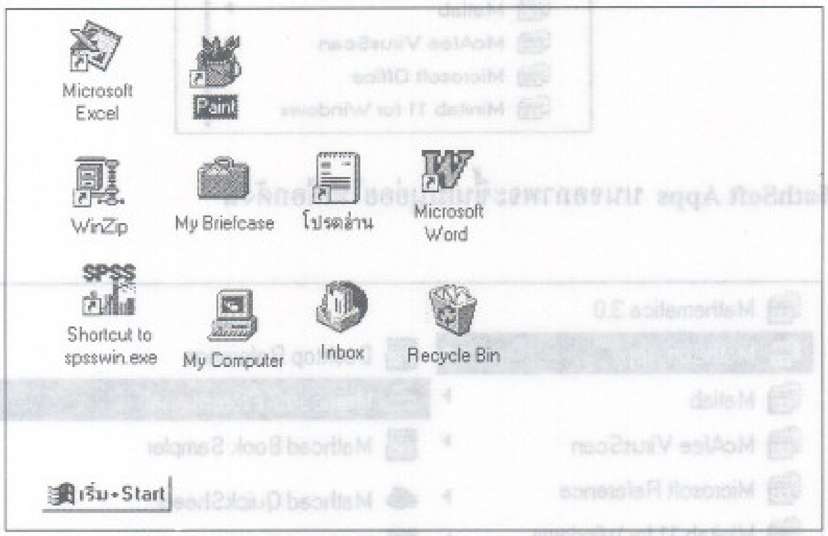
บทที่ 1

การใช้โปรแกรมสำเร็จรูป MATHCAD

บทที่ 1 นี้จะเรียนรู้เกี่ยวกับการนำโปรแกรมสำเร็จ MATHCAD เข้ามาทำงาน ทำการคำนวณค่าเบื้องต้นและประยุกต์การคำนวณเข้ากับเนื้อหาวิชาความน่าจะเป็นและสถิติ

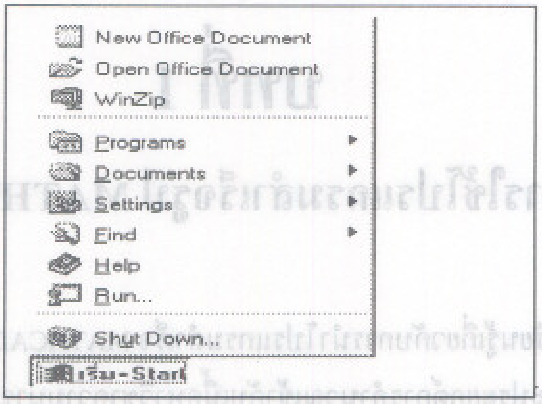
1.1 เข้าสู่การทำงานกับโปรแกรม MATHCAD

1. เปิดเครื่องคอมพิวเตอร์
2. รอจนจอภาพขึ้น

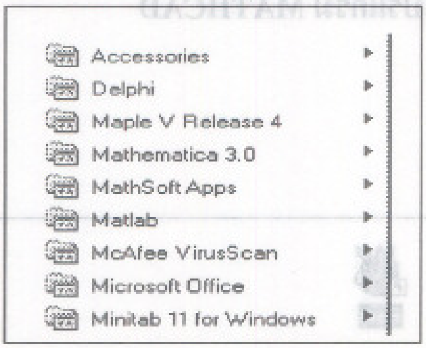


หมายเหตุ เนื้อหาที่ใช้โปรแกรม MATHCAD สามารถใช้ได้กับ Mathcad Professional varsion 7.0 , 8.0 และ Mathcad Professional 200

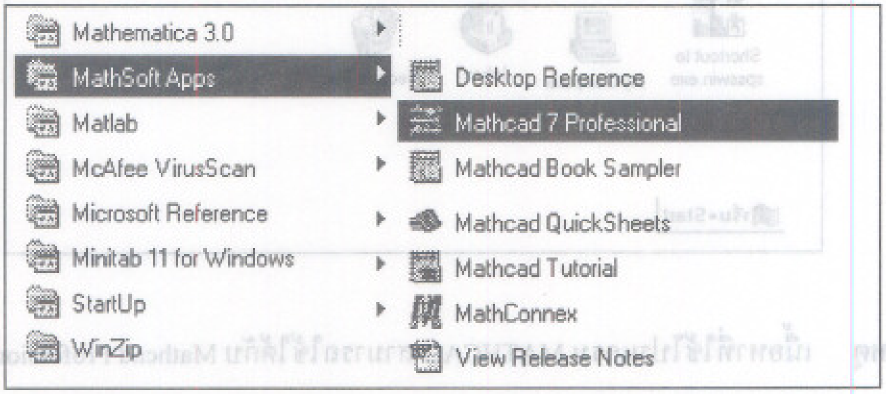
คลิกที่ปุ่ม **START** จะมีเมนูให้เลือก



คลิกที่เมนู **Program** จะมีเมนูย่อยให้เลือก



คลิกที่ **MathSoft Apps** บนจอภาพจะขึ้นเมนูย่อยให้เลือกดังนี้

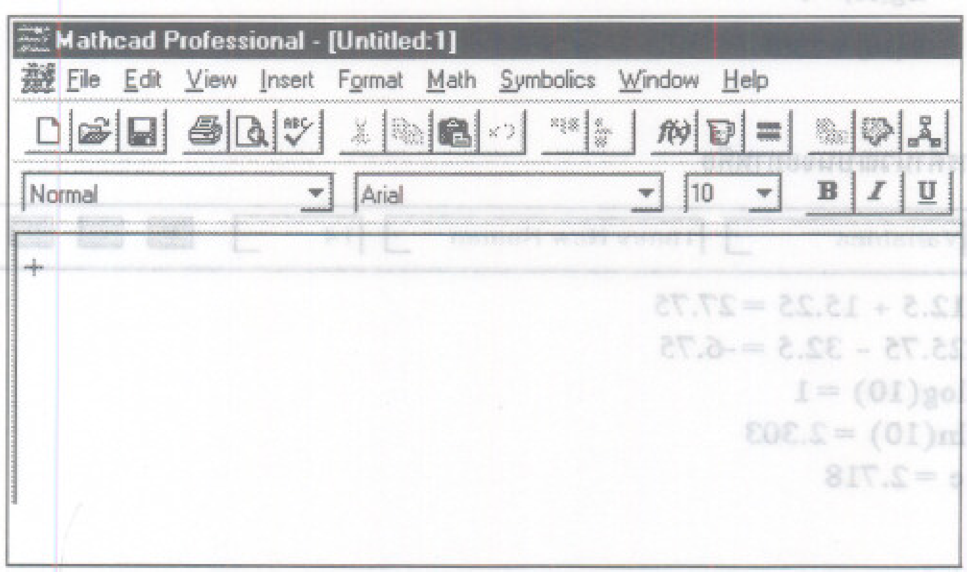


คลิกที่ **Mathcad 7 Professional** จะเข้าสู่การทำงานของโปรแกรม MATHCAD

จอภาพจะแสดง LOGO ของ **Mathcad 7 Professional**

รอจนสักครู่ คอมพิวเตอร์ก็พร้อมจะทำงานด้วยโปรแกรม **Mathcad 7 Professional**

จอภาพจะมีลักษณะดังนี้

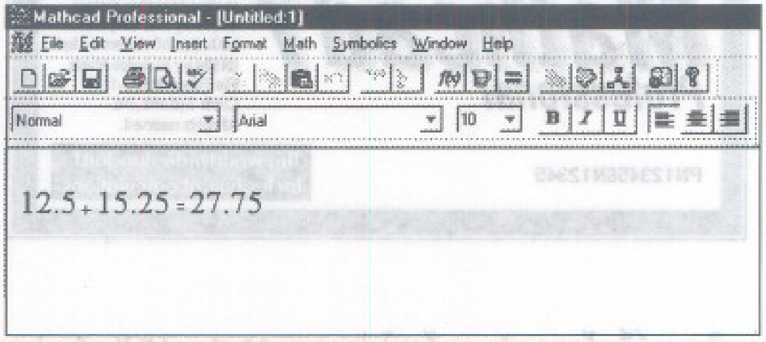


1.2 การคำนวณเบื้องต้นด้วยโปรแกรมสำเร็จรูป MATHCAD

ความสามารถที่ใช้งานง่ายที่สุดของ MATHCAD คือการคำนวณแบบเครื่องคิดเลข เช่นการหาผลบวกของ 12.5 + 15.25 ทำดังนี้

1. ให้พิมพ์ 12.5+15.25
2. แล้วพิมพ์ = จะได้ผลลัพธ์ทันที

ผลบนจอภาพ MATHCAD คือ



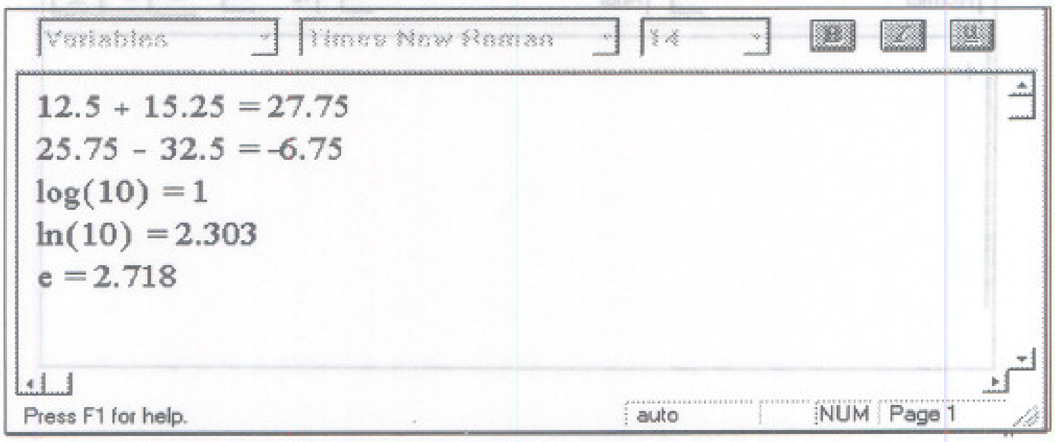
ตัวอย่างการคำนวณแบบอื่นๆ เช่น

$25.75 - 32.5 = -6.75$

$\log(10) = 1$

$\ln(10) = 2.303$

ผลการคำนวณบนจอภาพคือ



ในการคำนวณค่าต่างๆ โปรแกรม MATHCAD จะทำการจัดรูปแบบการพิมพ์บนจอภาพให้สอดคล้องกับความหมายทางคณิตศาสตร์เสมอ เช่น การหาค่า $\frac{25}{4} =$ 6.25 ขั้นตอนการพิมพ์คือ

ขั้นที่ 1 พิมพ์ 25

ขั้นที่ 2 เมื่อเรากดเครื่องหมายหาร (/) บนจอภาพจะจัดรูปแบบเป็นลักษณะของเศษส่วนทันที

$$\boxed{\frac{25}{\blacksquare}}$$

ขั้นที่ 3 ต่อไปจึงพิมพ์ 4 แล้วกด = จะได้ผลการคำนวณที่ต้องการ

$$\frac{25}{4} = 6.25$$

ดังนั้นขอให้สังเกตการจัดรูปแบบบนจอภาพของ MATHCAD ขณะคำนวณ

ตัวอย่างแบบต่างๆ ของการคำนวณด้วยโปรแกรม MATHCAD

1. การหาผลบวก 45.25 + 17.5

พิมพ์	ผลบนจอภาพ
45.25+17.5=↵	45.25 + 17.5 = 62.75

หมายเหตุ 1. การคำนวณใหม่หรือขึ้นบรรทัดใหม่ต้องกด <ENTER>

2. เครื่องหมาย = เป็นการสั่งให้ทำการคำนวณ

2. การหาผลหาร $\frac{25}{4}$

พิมพ์	ผลบนจอภาพ
25/	$\boxed{\frac{25}{\blacksquare}}$
4=↵	$\frac{25}{4} = 6.25$

3. การหาผลคูณของ 15 กับ 12

พิมพ์	ผลบนจอภาพ
15	15
*	$15\cdot\blacksquare$
12=↵	$15\cdot 12 = 180$

4. การคำนวณเลขยกกำลัง 3^4

พิมพ์	ผลบนจอภาพ
3	3
^	$3^{\blacksquare}$
4=↵	$3^4 = 81$

5. การคำนวณโดยการใช้ฟังก์ชันทางคณิตศาสตร์

พิมพ์	ผลบนจอภาพ
log(10)=↵	$\log(10) = 1$
ln(10)=↵	$\ln(10) = 2.303$
sin(	$\sin(\blacksquare$
<Ctrl>+P	$\sin(\pi$
/	$\sin\left(\frac{\pi}{\blacksquare}\right.$
6)= ↵	$\sin\left(\frac{\pi}{6}\right) = 0.5$

6. การคำนวณค่า n!

พิมพ์	ผลบนจอภาพ
1!=↵	1! = 1
2!=↵	2! = 2
n:5↵	n := 5
n!=↵	n! = 120

7. การกำหนดค่าให้กับตัวแปร , การกำหนดสูตรฟังก์ชัน

พิมพ์	ผลบนจอภาพ
x	$\boxed{x}$
:	$\boxed{x :=}$ **หมายเหตุ** การกด : จะได้เครื่องหมาย :=
4↵	x := 4
f(x)	f(x)
:	$\boxed{f(x) :=}$
x^2↵	$f(x) := x^2$
f(x)=↵	f(x) = 16
f(3)=↵	f(3) = 9

หมายเหตุ สัญลักษณ์ := หมายถึงการกำหนดค่าให้เป็น หรือ การกำหนดสูตรให้เป็น

8. การหาค่าปริพันธ์ $\int_a^b f(x)dx$ ตัวอย่างเช่น $\int_1^4 (x^2+4)dx$

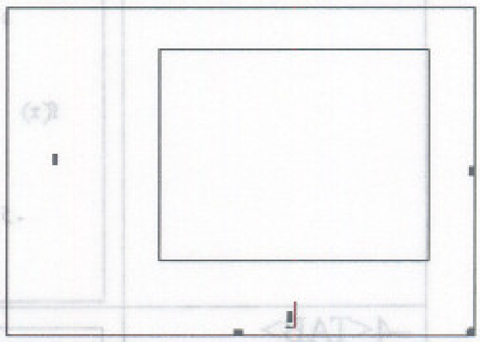
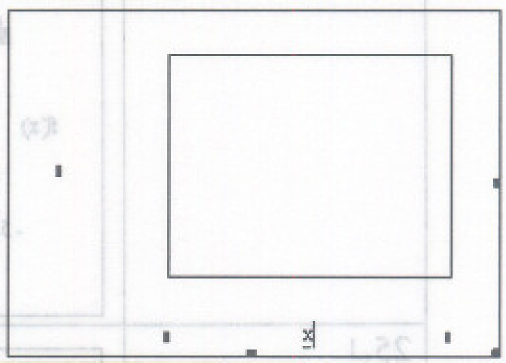
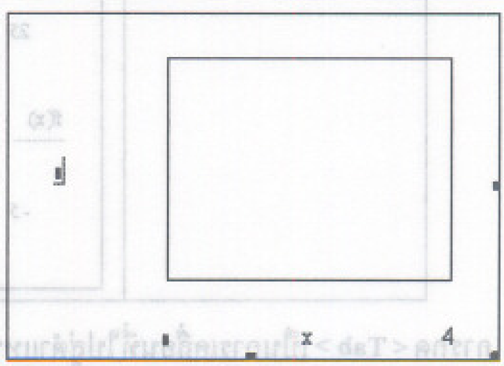
พิมพ์	ผลบนจอภาพ
f(x)	f(x)
:	$\boxed{f(x) :=}$

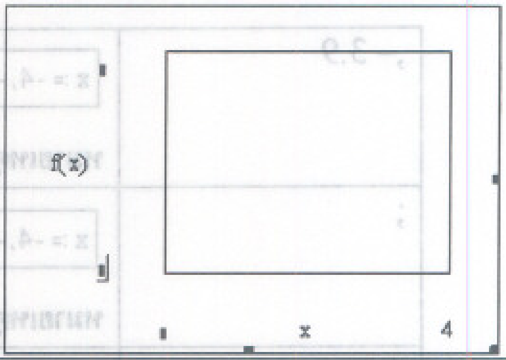
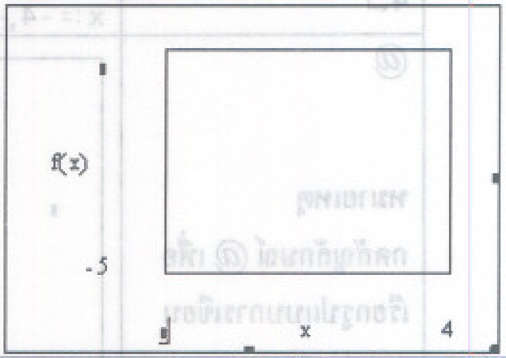
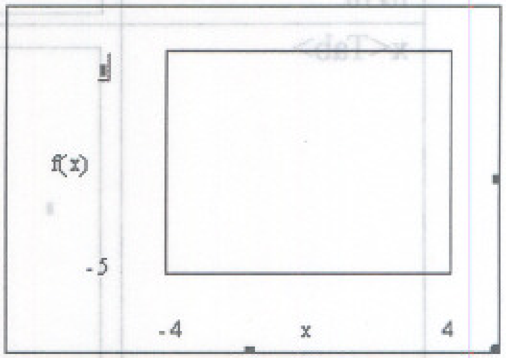
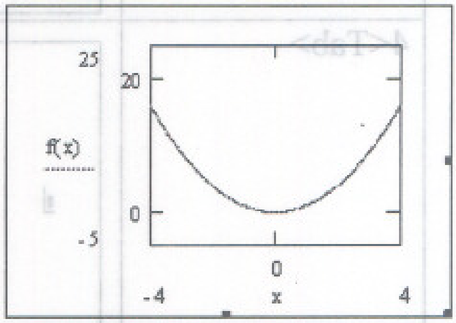
x^2	$f(x) := x^2$
<Space bar>	$f(x) := x^2$ จะเกิดกรอบตั้งฉากคลุม x^2
+4↵	$f(x) := x^2 + 4$
การกดสัญลักษณ์ปริพันธ์ &	$\int_{\blacksquare}^{\blacksquare} \blacksquare \, d\blacksquare$
f(x) <Tab>	$\int_{\blacksquare}^{\blacksquare} f(x)\, d\blacksquare$
x <Tab>	$\int_{\blacksquare}^{\blacksquare} f(x)\, dx$
1<Tab>	$\int_{1}^{\blacksquare} f(x)\, dx$
4=↵	$\int_{1}^{4} f(x)\, dx = 33$

1.3 การเขียนกราฟของฟังก์ชัน

ตัวอย่างเช่นการเขียนกราฟของ $f(x) = x^2 + 4$ บนช่วง $[-4, 4]$

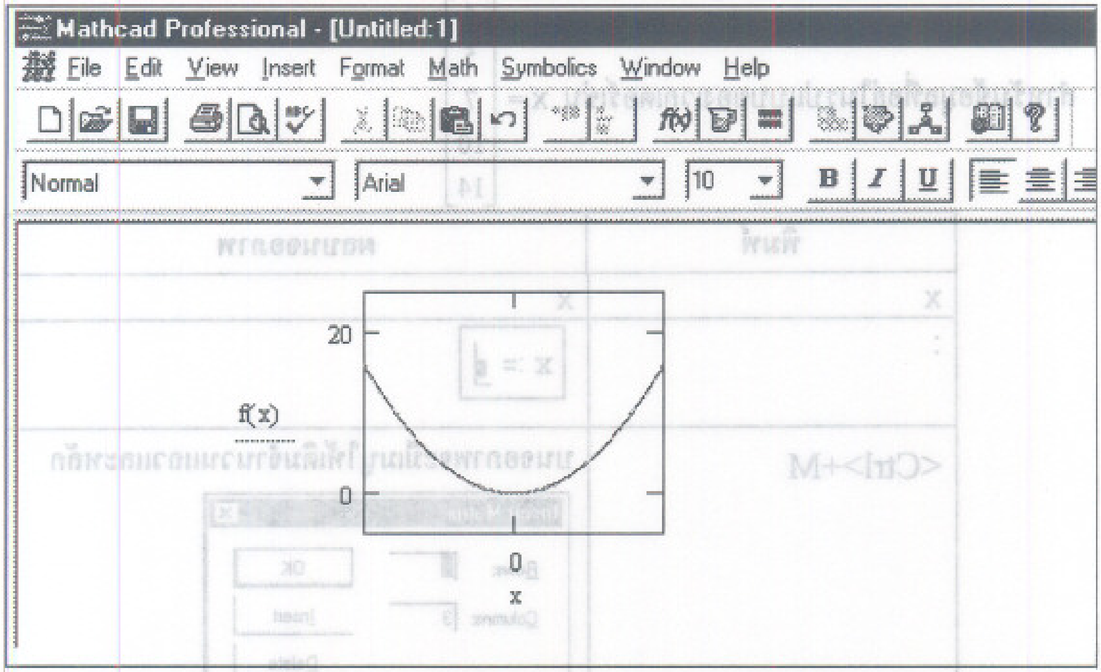
พิมพ์	ผลบนจอภาพ
f(x):x^2 <Space Bar>+4	$f(x) := x^2 + 4$
x:– 4	$x := -4$

,– 3.9	x := -4, -3.9 หมายเหตุ เป็นการกำหนดให้ค่า x เพิ่มทีละ 0.1
;	x := -4, -3.9 .. หมายเหตุ กด ; จะได้ ..
4↵	x := -4, -3.9 .. 4
@ หมายเหตุ กดสัญลักษณ์ @ เพื่อเรียกรูปแบบการเขียนกราฟ	
x<Tab>	
4<Tab>	

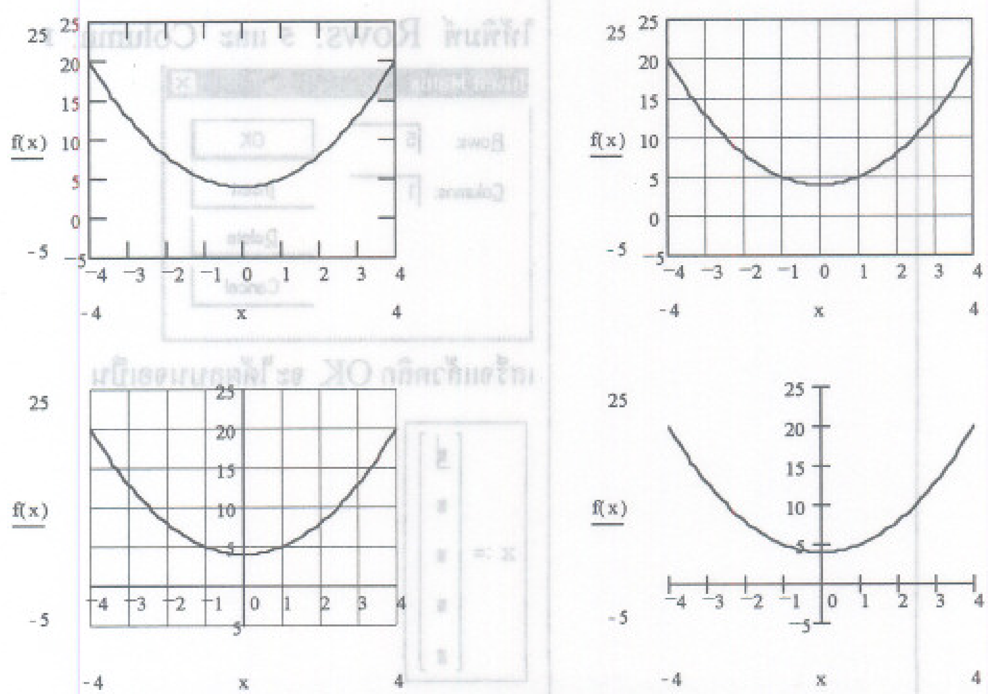
f(x)<Tab>	
–5<Tab>	
–4<TAB>	
25↵	

หมายเหตุ การกด < Tab > เป็นการเคลื่อนที่ไปสู่ตำแหน่งที่ต้องมีการกำหนดค่าที่เหมาะสมของการคำนวณ

ผลบนจอภาพทั้งหมดคือ

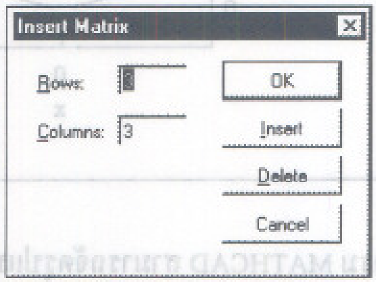
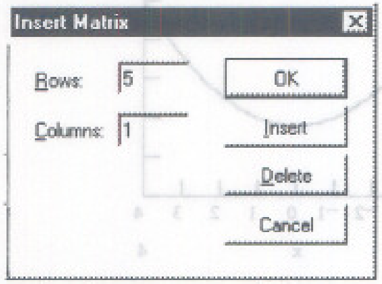


หมายเหตุ โปรแกรม MATHCAD สามารถจัดรูปแบบของกราฟได้หลายลักษณะเช่น



1.4 การกำหนดข้อมูลในรูปแบบเวกเตอร์

สำหรับข้อมูลที่อยู่ในรูปแบบของเวกเตอร์เช่น $x = \begin{bmatrix} 4 \\ 5 \\ 7 \\ 10 \\ 14 \end{bmatrix}$

พิมพ์	ผลบนจอภาพ
x	x
:	x :=
<Ctrl>+M	บนจอภาพจะมีเมนูให้เติมจำนวนแถวและหลัก  ให้พิมพ์ Rows: 5 และ Column: 1  เสร็จแล้วคลิก OK จะได้ผลบนจอเป็น 

4<Tab>	$x := \begin{bmatrix} 4 \\ \blacksquare \\ \blacksquare \\ \blacksquare \\ \blacksquare \end{bmatrix}$
5<Tab>	$x := \begin{bmatrix} 4 \\ 5 \\ \blacksquare \\ \blacksquare \\ \blacksquare \end{bmatrix}$
7<Tab>	$x := \begin{bmatrix} 4 \\ 5 \\ 7 \\ \blacksquare \\ \blacksquare \end{bmatrix}$
10<Tab>	$x := \begin{bmatrix} 4 \\ 5 \\ 7 \\ 10 \\ \blacksquare \end{bmatrix}$
14↵	$x := \begin{bmatrix} 4 \\ 5 \\ 7 \\ 10 \\ 14 \end{bmatrix}$

ใน MATHCAD มีฟังก์ชันที่ช่วยคำนวณค่าทางด้านสถิติของข้อมูลดังนี้

mean(x) = ค่าเฉลี่ยเลขคณิตของข้อมูลใน x

median(x) = ค่ามัธยฐานของข้อมูลใน x

var(x) = ค่าความแปรปรวนของข้อมูลใน x

stdev(x) = ค่าส่วนเบี่ยงเบนมาตรฐานของข้อมูลใน x

length(x) = จำนวนข้อมูลใน x

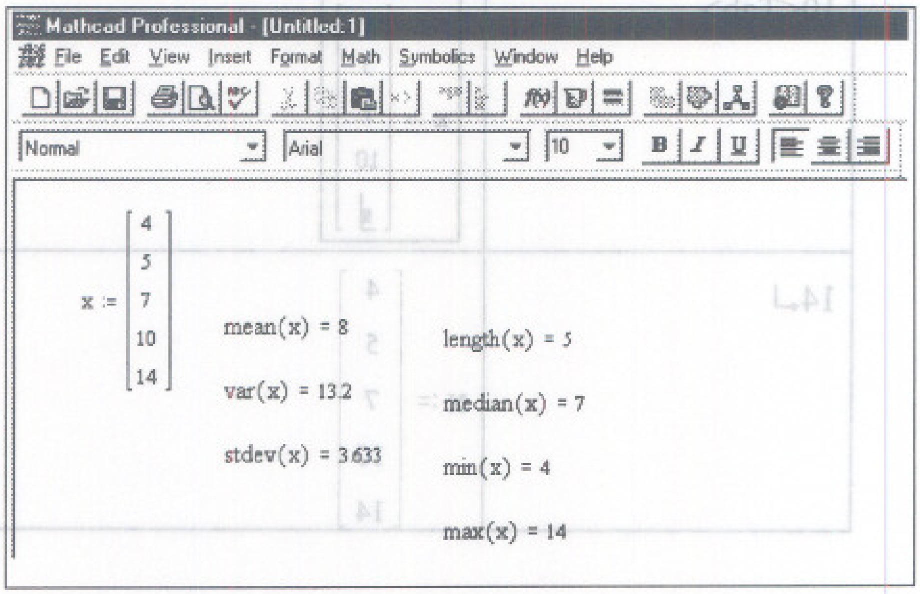
max(x) = ค่าสูงสุดของข้อมูลใน x

min(x) = ค่าต่ำสุดของข้อมูลใน x

ตัวอย่างเช่น mean(x) = 8 var(x) = 13.2 stdev(x) = 3.633

length(x) = 5 median(x) = 7 min(x) = 4 max(x) = 14

ผลการคำนวณบนจอภาพคือ



1.5 การพิมพ์สูตรที่มีรูปแบบยาก

ตัวอย่างเช่นต้องการพิมพ์สูตร

$$f(x) := \frac{x^2 + 4}{x - 3}$$

พิมพ์	ผลบนจอภาพ
f(x):	$f(x) := \blacksquare$
x^2	$f(x) := x^{2}$
<Space Bar>	$f(x) := x^{2}$
+4	$f(x) := x^2 + 4$
<Space Bar>	$f(x) := x^2 + 4$
/	$f(x) := \frac{x^2 + 4}{\blacksquare}$
x–3↵	$f(x) := \frac{x^2 + 4}{x - 3}$

ตัวอย่างการคำนวณค่าฟังก์ชัน

$f(4) = 20$ $\qquad$ $f(2) = -8$

$f(f(1)) = -1.8636$

การพิมพ์ฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม Z $\quad f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}\cdot z^{2}}$

พิมพ์	ผลบนจอภาพ
f(z):	$f(z) := \blacksquare$
1/	$f(z) := \frac{1}{\blacksquare}$
\	$f(z) := \frac{1}{\sqrt{\blacksquare}}$ หมายเหตุ การกด \ จะได้เครื่องหมาย $\sqrt{\ }$
2*	$f(z) := \frac{1}{\sqrt{2\cdot\blacksquare}}$
<Ctrl>+P	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}$
<Space Bar>	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}$
<Space Bar>	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}$
<Space Bar>	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}$

*	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot\blacksquare$
e^	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\blacksquare}$
–1/2	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}}$
<Space Bar>	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}}$
*z^2↵	$f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}\cdot z^2}$

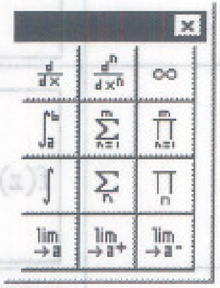
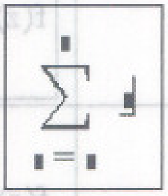
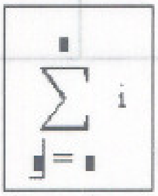
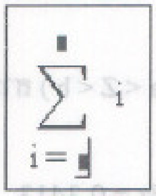
หมายเหตุ $f(z) := \frac{1}{\sqrt{2\cdot\pi}}\cdot e^{\frac{-1}{2}\cdot z^2}$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นของตัวแปรสุ่มปกติมาตรฐาน Z ความน่าจะเป็น $P(a < Z < b)$ สามารถคำนวณได้โดยง่ายดังนี้

$$P(0 < Z < 1) = \int_0^1 f(z)\,dz = 0.3413$$

$$P(0 < Z < 2) = \int_0^2 f(z)\,dz = 0.4772$$

1.6 การคำนวณค่าเกี่ยวกับผลบวกในรูปแบบ $\sum$

ตัวอย่างเช่นการหาค่าของ $\sum_{i=1}^{20} i$

พิมพ์	ผลบนจอภาพ
คลิกที่สัญลักษณ์ 	จะได้แถบเครื่องมือดังนี้ 
คลิกที่สัญลักษณ์ 	
i<Tab>	
i<Tab>	

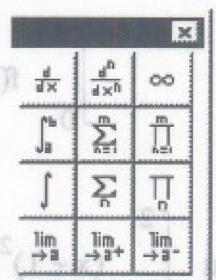
1<Tab>	$\sum_{i=1}^{\blacksquare} i$
20=↵	$\sum_{i=1}^{20} i = 210$

ตัวอย่างการคำนวณค่าแบบอื่นๆ เช่น

$$\sum_{i=1}^{5} i^3 = 225 \qquad \sum_{k=2}^{4} k^2 = 29 \qquad \left[\sum_{n=0}^{4} \left(n^2 + n + 1\right)\right] = 45$$

1.7 การคำนวณค่าปริพันธ์ $\int_a^b f(x)\,dx$

ตัวอย่างเช่น $\int_0^1 x^3\,dx$

พิมพ์	ผลบนจอภาพ
คลิกที่สัญลักษณ์ 	จะได้แถบเครื่องมือดังนี้ 

คลิกที่สัญลักษณ์ $\int_a^b$	$\int_{\blacksquare}^{\blacksquare} \blacksquare \, d\blacksquare$
x^3<Tab>	$\int_{\blacksquare}^{\blacksquare} x^3 \, d\blacksquare$
x<Tab>	$\int_{\blacksquare}^{\blacksquare} x^3 \, dx$
0<Tab>	$\int_{0}^{\blacksquare} x^3 \, dx$
1=↵	$\int_0^1 x^3 \, dx = 0.25$

หมายเหตุ ในกรณีที่เรากำหนด $f(x) = x^3$ จะทำให้การคำนวณสะดวกขึ้นดังนี้ $f(x) := x^3$

$$\int_0^1 f(x)\, dx = 0.25 \qquad \int_0^2 f(x)\, dx = 4$$

หรือจะกำหนดสูตรอื่นๆ ก็ได้เช่น $\int_0^2 (x+1)^2 \, dx = 8.667$

1.8 การหาผลบวกในรูปแบบ $\sum x_i$

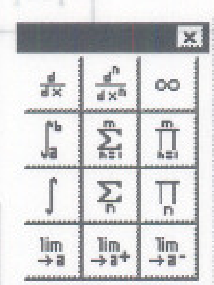
$$x := \begin{bmatrix} 2 \\ 4 \\ 5 \\ 9 \end{bmatrix} \qquad y := \begin{bmatrix} 3 \\ 5 \\ 7 \\ 15 \end{bmatrix}$$

ORIGIN := 1

$$\sum_{i=1}^{4} x_i = 20 \qquad \sum_{i=1}^{4} \left(x_i\right)^2 = 126$$

$$\sum_{i=1}^{4} y_i = 30 \qquad \sum_{i=1}^{4} x_i \cdot y_i = 196 \qquad \sum_{i=1}^{4} \left(y_i\right)^2 = 308$$

หมายเหตุ ควรจะกำหนด ORIGIN = 1

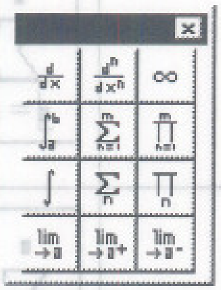
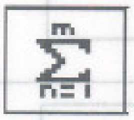
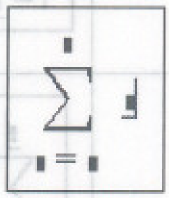
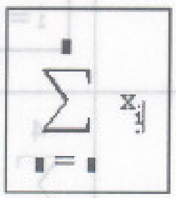
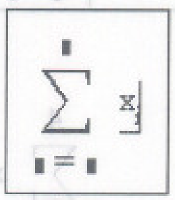
พิมพ์	ผลบนจอภาพ
กำหนดข้อมูล x และ y	$x := \begin{bmatrix} 2 \\ 4 \\ 5 \\ 9 \end{bmatrix} \qquad y := \begin{bmatrix} 3 \\ 5 \\ 7 \\ 15 \end{bmatrix}$
คลิกที่สัญลักษณ์ $\boxed{\int \frac{dy}{dx}}$	จะได้แถบเครื่องมือดังนี้ 

คลิกที่สัญลักษณ์ $\sum_{n=1}^{m}$	$\sum_{\blacksquare=\blacksquare}^{\blacksquare}\blacksquare$
x[i<Tab>	$\sum_{\blacksquare=\blacksquare}^{\blacksquare}x_i$ หมายเหตุ การกด [จะได้ เครื่องหมาย subscript
i<Tab>	$\sum_{i=\blacksquare}^{\blacksquare}x_i$
1<Tab>	$\sum_{i=1}^{\blacksquare}x_i$
4=↵	$\sum_{i=1}^{4}x_i=20$

ตัวอย่างการคำนวณอื่นๆ เช่น

$$\sum_{i=1}^{4}y_i=30 \qquad \left[\sum_{i=1}^{4}(x_i+y_i)\right]=50$$

การพิมพ์เพื่อหาผลบวก $\sum_{i=1}^{4}\left(x_i\right)^2 = 126$

พิมพ์	ผลบนจอภาพ
คลิกที่สัญลักษณ์ 	จะได้แถบเครื่องมือดังนี้ 
คลิกที่สัญลักษณ์ 	
x[i	
<Space Bar>	

^	$\sum_{\blacksquare=\blacksquare}^{\blacksquare}(x_i)^{\blacksquare}$
2<Tab>	$\sum_{\blacksquare=\blacksquare}^{\blacksquare}(x_i)^{2}$
i<Tab>	$\sum_{i=\blacksquare}^{\blacksquare}(x_i)^{2}$
1<Tab>	$\sum_{i=1}^{\blacksquare}(x_i)^{2}$
4=↵	$\sum_{i=1}^{4}(x_i)^{2} = 126$

ตัวอย่างการคำนวณอื่นๆ เช่น

$$\sum_{i=1}^{4}(y_i)^2 = 308 \qquad \sum_{i=1}^{4}x_i \cdot y_i = 196$$

1.9 การสร้างตารางฟังก์ชัน

ตัวอย่างเช่น $x := 1..4$ $f(x) := 2 \cdot x + 4$

x	f(x)
1	6
2	8
3	10
4	12

พิมพ์	ผลบนจอภาพ
x:1;4↵	$x := 1..4$
f(x):2*x+4↵	$f(x) := 2 \cdot x + 4$
x=↵	x 1 2 3 4
f(x)=	f(x) 6 8 10 12

ตัวอย่าง การสร้างตารางความน่าจะเป็นของตัวแปรสุ่มทวินาม

$n := 5$

$p := 0.2$

$$b(x,n,p) := \frac{n!}{x!\cdot(n-x)!}\cdot p^x\cdot(1-p)^{n-x}$$

$x := 0..\ 5$

x	b(x, n, p)
0	0.3277
1	0.4096
2	0.2048
3	0.0512
4	0.0064
5	0.0003

ตัวอย่าง การสร้างตารางความน่าจะเป็นของตัวแปรสุ่มปัวส์ซอง

$mue := 4$

$$p(x) := \frac{e^{-mue}\cdot mue^x}{x!}$$

$x := 0..\ 10$

x	p(x)	$\sum_{i=0}^{x} p(i)$
0	0.018316	0.018316
1	0.073263	0.091578
2	0.146525	0.238103
3	0.195367	0.43347
4	0.195367	0.628837
5	0.156293	0.78513
6	0.104196	0.889326
7	0.05954	0.948866
8	0.02977	0.978637
9	0.013231	0.991868
10	0.005292	0.99716

แบบฝึกหัด 1

จงหาค่า

1. $\frac{22}{7}$ = ... π = ...

2. $(1+\frac{1}{50})^{50}$ = ... $(1+\frac{1}{5000})^{5000}$ = ... e = ...

3. $\sin(105^\circ)$ = ... $\cos(75^\circ)$ = ... $\tan(75^\circ)$= ....

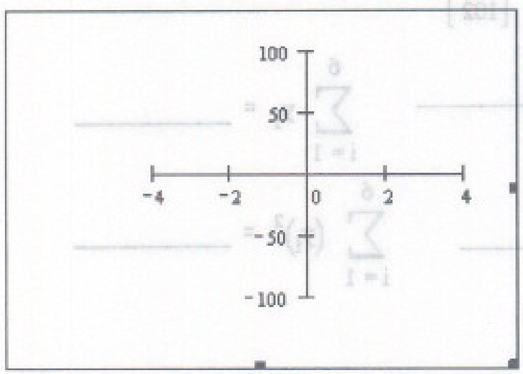
4. $\sin(\frac{4\pi}{5})$ = ... $\cos(\frac{4\pi}{15})$ = ... $\tan(\frac{7\pi}{12})$= ....

5. 9!=.......................... 12!=.......................... 15!=..........................

6. $\frac{10!}{(2!)(8!)}$ =.............. $\frac{10!}{(1!)(2!)(3!)(4!)}$ =..............

7. $\int_1^4 (x^2+x+1)dx$ = ...

8. จงเขียนกราฟของ $f(x)=x^3+3x-7$ บนช่วง $[-4, 4]$



9. $\frac{e^{-4}4^7}{7!}$ = ... $\frac{e^{-5}5^2}{2!}$ = ...

10. $\frac{10!}{(2!)(8!)}(0.5)^2(0.8)^8$ = ...

11. $\int_2^3 \frac{1}{18}(3+2x)dx$ = ...

12. $\int_{\frac{1}{4}}^{\frac{1}{2}} \int_0^1 \frac{x(1+3y^2)}{4} dxdy$ = ...

13. $\dfrac{\sum_{x=1}^{6} x}{6}$ = ...

14. $\dfrac{\sum_{x=1}^{6} (x-3.5)^2}{6}$ = = ...

15. $\binom{15}{5}(\frac{1}{4})^5(\frac{3}{4})^{10}$ = ...

16. $\sum_{x=0}^{3} \dfrac{e^{-5}5^x}{x!}$ = ...

17. จงหาผลบวกต่อไปนี้

$$x := \begin{bmatrix} 12 \\ 15 \\ 23 \\ 28 \\ 35 \\ 45 \end{bmatrix} \qquad y := \begin{bmatrix} 53 \\ 64 \\ 75 \\ 86 \\ 92 \\ 102 \end{bmatrix}$$

$\sum_{i=1}^{6} x_i$ = ____________ $\sum_{i=1}^{6} y_i$ = ____________

$\sum_{i=1}^{6} (x_i)^2$ = ______ $\sum_{i=1}^{6} (y_i)^2$ = ____________

$\sum_{i=1}^{6} x_i \cdot y_i$ = ____________

18. กำหนดข้อมูลจำนวน 15 ค่า

15 , 21 , 32 , 45 , 26 , 35 , 45 , 78 , 32 , 25 , 67 , 45 , 36 , 78 , 90

จงหาค่าของ ค่าเฉลี่ย = ... ค่าความแปรปรวน = ...

ส่วนเบี่ยงเบนมาตรฐาน = ...

ค่ามัธยฐาน = ...

บทที่ 2

เสริมการคำนวณความน่าจะเป็นและสถิติด้วย MATHCAD

บทที่ 2 นี้จะนำความสามารถของโปรแกรมสำเร็จ MATHCAD เข้ามาคำนวณค่าต่างๆ เกี่ยวกับความน่าจะเป็นและสถิติ เช่นการจัดลำดับจัดหมู่ ความน่าจะเป็นแบบทวินาม ไฮเพอร์จีออเมตริก ค่าคาดคะเน

2.1 การจัดลำดับและจัดหมู่

จากสูตร $_nP_r = \dfrac{n!}{(n-r)!}$ การคำนวณด้วย MATHCAD ทำดังนี้

$$P(n,r) := \frac{n!}{(n-r)!}$$

$n := 5$

$r := 0..n$

r	P(n,r)
0	1
1	5
2	20
3	60
4	120
5	120

จากสูตร $_nC_r = \dfrac{n!}{r!(n-r)!}$ การคำนวณด้วย MATHCAD ทำดังนี้

$$C(n,r) := \frac{n!}{r!\cdot(n-r)!}$$

$n := 5$

$r := 0..n$

r	C(n,r)
0	1
1	5
2	10
3	10
4	5
5	1

2.2 การหาค่าเฉลี่ยและความแปรปรวนของประชากร

กำหนดข้อมูลประชากร 4 , 5 , 7 , 9 , 12 , 15 , 18

$$x := \begin{bmatrix} 4 \\ 5 \\ 7 \\ 9 \\ 12 \\ 15 \\ 18 \end{bmatrix} \qquad \text{mean}(x) = 10 \qquad \text{var}(x) = 23.429 \qquad \text{stdev}(x) = 4.84$$

หมายเหตุ mean(x) ได้มาจากสูตร $\dfrac{\sum_{i=1}^{n} x_i}{n}$ var(x) ได้มาจากสูตร $\dfrac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}$

$\text{stdev}(x) = \sqrt{\text{var}(x)}$ สามารถตรวจสอบได้ดังนี้

ORIGIN := 1 n := 7

$$\frac{\sum_{i=1}^{n} x_i}{n} = 10$$

$$\frac{\sum_{i=1}^{n} (x_i - 10)^2}{n} = 23.429 \qquad \sqrt{\frac{\sum_{i=1}^{n} (x_i - 10)^2}{n}} = 4.84$$

2.3 การหาค่าคาดคะเนของตัวแปรสุ่มไม่ต่อเนื่อง

ตัวอย่าง การหาค่าคาดคะเน และ ความแปรปรวนของตัวแปรสุ่ม X

X	ความน่าจะเป็นของ X
0	$\frac{1}{35}$
1	$\frac{12}{35}$
2	$\frac{18}{35}$
3	$\frac{4}{35}$

สูตรการคำนวณค่าคาดคะเน $E(X) = \mu = \sum_{i=1}^{n} x_i f(x_i)$

สูตรการคำนวณค่าความแปรปรวน $\sigma^2 = \sum_{i=1}^{n} ((x_i - \mu)^2) f(x_i)$

การคำนวณด้วย MATHCAD

$$\text{ORIGIN} := 1 \quad n := 4 \quad x := \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} \qquad f := \begin{bmatrix} \frac{1}{35} \\ \frac{12}{35} \\ \frac{18}{35} \\ \frac{4}{35} \end{bmatrix}$$

$$\mu := \sum_{i=1}^{n} x_i \cdot f_i \qquad \mu = 1.714$$

$$\sigma := \sqrt{\sum_{i=1}^{n} \left[(x_i - \mu)^2 \cdot f_i \right]} \qquad \sigma = 0.7$$

$$\sigma^2 = 0.49$$

2.4 การหาค่าคาดคะเนของตัวแปรสุ่มต่อเนื่อง

ตัวอย่าง ตัวแปรสุ่มต่อเนื่อง x มีฟังก์ชันการแจกแจงความน่าจะเป็น

$$f(x) = \begin{cases} \dfrac{x^2}{3} & ,-1 < x < 2 \\ 0 & ,x \text{ มีค่าอื่นๆ} \end{cases}$$

สูตร $P(a < X < b) = \int_a^b f(x)dx$ ค่าคาดคะเน $E(X) = \mu = \int_{-\infty}^{\infty} xf(x)dx$

ค่าความแปรปรวน $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$

การคำนวณด้วย MATHCAD

$$f(x) := \frac{x^2}{3}$$

$$P(a,b) := \int_a^b f(x)\,dx$$

$P(0,2) = 0.889$ $P(1,2) = 0.778$

$$\mu := \int_{-1}^{2} x \cdot f(x)\,dx \qquad \mu = 1.25$$

$$\sigma := \sqrt{\int_{-1}^{2} (x-\mu)^2 \cdot f(x)\,dx} \qquad \sigma = 0.798 \qquad \sigma^2 = 0.637$$

หมายเหตุ เนื่องจาก f(x) เป็น 0 นอกช่วง (–1 , 2)

ดังนั้นการปริพันธ์เพื่อหา μ และ σ^2 คิดบนช่วง (–1 , 2)

ตัวอย่าง X และ Y เป็นตัวแปรสุ่มต่อเนื่อง มีฟังก์ชันการแจกแจงความน่าจะเป็นร่วมกันคือ

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4} & ,0 < x < 2 \text{ and } 0 < y < 1 \\ 0 & ,\text{otherwise} \end{cases}$$

การคำนวณด้วย MATHCAD

$$f(x,y) := \frac{x\cdot\left(1+3\cdot y^2\right)}{4}$$

$$P(a,b,c,d) := \int_a^b \int_c^d f(x,y)\, dy\, dx$$

$P(0,2,0,1) = 1$
$P(0,1,0,1) = 0.25$
$P(0,0.5,0,1) = 0.063$

หมายเหตุ P(a,b,c,d) = ความน่าจะเป็นที่ $a < x < b$ และ $c < y < d$

การคำนวณค่าคาดคะเน

$$Ex := \int_0^2 \int_0^1 x\cdot f(x,y)\, dy\, dx \qquad Ex = 1.333$$

$$Ey := \int_0^2 \int_0^1 y\cdot f(x,y)\, dy\, dx \qquad Ey = 0.625$$

$$Exy := \int_0^2 \int_0^1 (x\cdot y)\cdot f(x,y)\, dy\, dx \qquad Exy = 0.833$$

$$Exx := \int_0^2 \int_0^1 \left(x^2\right)\cdot f(x,y)\, dy\, dx \qquad Exx = 2$$

$$Eyy := \int_0^2 \int_0^1 \left(y^2\right)\cdot f(x,y)\, dy\, dx \qquad Eyy = 0.467$$

หมายเหตุ Ex หมายถึง E(X) Ey หมายถึง E(Y) Exy หมายถึง E(XY)

2.5 โมเมนต์อันดับที่ k รอบค่า เฉลี่ย μ ของตัวแปรสุ่มไม่ต่อเนื่อง X

ตัวแปรสุ่มไม่ต่อเนื่อง X มีค่าเป็น $X = x_1, x_2, x_3, \ldots, x_n$ และฟังก์ชันความน่าจะเป็น f(x)

โมเมนต์อันดับที่ k รอบค่า เฉลี่ย μ ของ ตัวแปร X มีค่าเท่ากับ $\sum_{i=1}^{n} (x_i - \mu)^k f(x_i)$

กำหนดตัวแปรสุ่ม X = 1, 2, 3, 4, 5, 6 และ $f(x) = \frac{1}{6}$

จงหาค่าของ μ และโมเมนต์อันดับ 0 , 1 , 2 , 3 รอบ μ

ตัวอย่างการคำนวณด้วย MATHCAD

ORIGIN := 1

$$x := \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix} \qquad f(x) := \frac{1}{6}$$

$$\mu := \sum_{i=1}^{6} x_i \cdot f(x_i) \qquad \mu = 3.5$$

$$\text{moment}(k) := \sum_{i=1}^{6} (x_i - \mu)^k \cdot f(x_i)$$

k := 0 .. 3

k	moment(k)
0	1
1	0
2	2.917
3	0

2.6 โมเมนต์อันดับที่ k รอบค่าเฉลี่ย μ ของ ตัวแปรสุ่มต่อเนื่อง X

โมเมนต์อันดับที่ k รอบจุด x = 0 ของ ตัวแปร X คือ $E(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx$

โมเมนต์อันดับที่ k รอบค่า เฉลี่ย μ ของ ตัวแปร X คือ $E((X-\mu)^k) = \int_{-\infty}^{\infty} (x-\mu)^k f(x)dx$

ค่าคาดคะเน = โมเมนต์อันดับที่ 1 รอบจุด x = 0 ของตัวแปร X

ค่าความแปรปรวน = โมเมนต์อันดับที่ 2 รอบค่า เฉลี่ย μ ของ ตัวแปร X คือ $E((X-\mu)^2)$

ตัวอย่าง กำหนด $f(x) = \begin{cases} 2x^2 + \frac{1}{3} & , 0 < x < 1 \\ 0 & , \text{otherwise} \end{cases}$

จงหาค่าของ μ และโมเมนต์อันดับ 0 , 1 , 2 , 3 รอบ μ

การคำนวณด้วย MATHCAD

$f(x) := 2 \cdot x^2 + \frac{1}{3}$

$\mu := \int_0^1 x \cdot f(x)\, dx \qquad \mu = 0.667$

$\text{moment}(k) := \int_0^1 (x - \mu)^k \cdot f(x)\, dx$

$k := 0 .. 3$

k	moment(k)
0	1
1	0
2	0.067
3	-0.013

2.7 การคำนวณค่าความน่าจะเป็นของตัวแปรสุ่มทวินาม

ลักษณะของการทดลอง ความน่าจะเป็นของผลสำเร็จแต่ละครั้ง p ทำการทดลอง n ครั้ง

X = จำนวนครั้งของผลสำเร็จ = 0, 1, 2, ... , n

ตัวแปรสุ่มทวินาม X มีฟังก์ชันการแจกแจงความน่าจะเป็น $b(x, n, p) = \binom{n}{x} p^{x}(1-p)^{n-x}$

ตัวอย่าง 2.7.1 ในการทดสอบยางรถบรรทุก พบว่า 25% ของยางรถแตกก่อนที่เสร็จสิ้นการทดสอบ จงหาความน่าจะเป็นที่จะมียางแตก 5 ถึง 10 เส้นในการทดสอบยางทั้งหมด 15 เส้น

แนวคิด จำนวนครั้งที่ทดลอง $n := 15$ ความน่าจะเป็นของผลสำเร็จแต่ละครั้ง $p := 0.25$

ความน่าจะเป็นที่จะมียางแตก 5 ถึง 10 เส้น = $\sum_{x=5}^{10} b(x,n,p) = 0.313$

การคำนวณด้วย MATHCAD

$$b(x,n,p) := \frac{n!}{x!\cdot(n-x)!}\cdot p^{x}\cdot(1-p)^{n-x}$$

$x := 0 .. 4$

x	b(x, 15, 0.25)
0	0.013
1	0.067
2	0.156
3	0.225
4	0.225

$$\sum_{x=5}^{10} b\left(x, 15, \frac{1}{4}\right) = 0.313 \qquad \sum_{x=1}^{3} b\left(x, 15, \frac{1}{4}\right) = 0.448$$

$$\sum_{x=4}^{15} b\left(x, 15, \frac{1}{4}\right) = 0.539 \qquad \sum_{x=0}^{6} b\left(x, 15, \frac{1}{4}\right) = 0.943$$

สรุป ในการทดสอบยางทั้งหมด 15 เส้น

1.1 ความน่าจะเป็นที่จะมียางแตก 1 ถึง 3 เส้น = 0.448

1.2 ความน่าจะเป็นที่จะมียางแตกอย่างน้อย 4 เส้น = 0.539

1.3 ความน่าจะเป็นที่จะมียางแตกอย่างมาก 6 เส้น = 0.943

2.8 การคำนวณค่าความน่าจะเป็นของตัวแปรสุ่มไฮเพอร์จีออเมตริก

ลักษณะทั่วไปของการทดลอง

1. ประชากรทั้งหมดมีจำนวน N
2. ประกอบด้วย ผลสำเร็จจำนวน k และ ผลไม่สำเร็จจำนวน N – k
3. สุ่มตัวอย่างขนาด n จากประชากรพร้อมกัน

X = จำนวนผลสำเร็จ = 0, 1, 2, ... , (n หรือ k)

ตัวแปรสุ่ม X เรียกว่าตัวแปรสุ่มไฮเพอร์จีออเมตริกมีฟังก์ชันการแจกแจงความน่าจะเป็น

$$h(x, N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

ตัวอย่าง 2.8.1 พฤสีกล่องหนึ่งมี 10 อันเป็นสีเหลือง 3 อันนอกนั้นเป็นสีม่วงสุ่มเลือก 4 อันแล้วยิง จงหาความน่าจะเป็นที่ทั้ง 4 อันที่เลือกมาเป็นสีม่วง

แนวคิด กำหนดสูตร $C(n,r) := \frac{n!}{r!\cdot(n-r)!}$

จำนวนของทั้งหมด N = 10 จำนวนของผลสำเร็จ k = 7 จำนวนของที่หยิบออกมา n = 4

ความน่าจะเป็นแบบไฮเพอร์จีออเมตริก $h(x,N,n,k) := \frac{C(k,x)\cdot C(N-k,n-x)}{C(N,n)}$

ความน่าจะเป็นที่ทั้ง 4 อันที่เลือกมาเป็นสีม่วง = $h(4,N,n,k) = 0.167$

การคำนวณด้วย MATHCAD

$N := 10$ $\quad k := 7$ $\quad n := 4$

$C(n,r) := \frac{n!}{r!\cdot(n-r)!}$ $\quad h(x,N,n,k) := \frac{C(k,x)\cdot C(N-k,n-x)}{C(N,n)}$

$h(4,N,n,k) = 0.167$

$x := 1..4$

x	h(x,N,n,k)
1	0.033
2	0.3
3	0.5
4	0.167

เพราะฉะนั้น ความน่าจะเป็นที่อย่างมาก 2 อัน เป็นม่วง = 0.333

ความน่าจะเป็นที่อย่างน้อย 2 อัน เป็นม่วง = 0.967

ความน่าจะเป็นที่จะไม่ได้สีม่วง = 0

2.9 การคำนวณค่าความน่าจะเป็นของตัวแปรสุ่มปัวส์ซอง

ลักษณะของตัวแปรสุ่มปัวส์ซอง

μ = จำนวนครั้งของผลสำเร็จที่เกิดในช่วงเวลา

X = จำนวนครั้งของผลสำเร็จ = 0, 1, 2, ...

X เป็นตัวแปรสุ่มปัวส์ซอง มีฟังก์ชันการแจกแจงความน่าจะเป็น $p(x,\mu) = \dfrac{e^{-\mu}\mu^{x}}{x!}$

ตัวอย่าง 2.8.1 ชายฝั่งตะวันออกของประเทศหนึ่งถูกพายุใต้ฝุ่นโดยเฉลี่ย 6 ครั้งต่อปี จงหาความน่าจะเป็นที่ในปีต่อไป ชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นน้อยกว่า 4 ครั้ง

แนวคิด ค่าเฉลี่ย $\mu = 6$ $\quad p(x,\ \mu) = \dfrac{e^{-6}6^{x}}{x!}$

การคำนวณด้วย MATHCAD

$\mu := 6 \qquad p(x,\mu) := \dfrac{e^{-\mu}\cdot\mu^{x}}{x!}$

$x := 0..4$

x	p(x,μ)
0	0.002
1	0.015
2	0.045
3	0.089
4	0.134

$\sum_{x=0}^{3} p(x,\mu) = 0.151 \qquad \sum_{x=6}^{8} p(x,\mu) = 0.402$

$1 - \sum_{x=0}^{4} p(x,\mu) = 0.715 \qquad 1 - \sum_{x=0}^{3} p(x,\mu) = 0.849$

ความน่าจะเป็นที่ในปีต่อไป ชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นน้อยกว่า 4 ครั้ง = 0.151

ความน่าจะเป็นที่ในปีต่อไปชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นระหว่าง 6 ถึง 8 ครั้ง = 0.402

ความน่าจะเป็นที่ในปีต่อไปชายฝั่งแห่งนี้จะถูกพายุอย่างน้อย 5 ครั้ง = 0.715

ความน่าจะเป็นที่ในปีต่อไปชายฝั่งแห่งนี้จะไม่ถูกพายุใต้ฝุ่นเข้าเลย = 0.002

ความน่าจะเป็นที่ในปีต่อไปชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นเข้าเกิน 3 ครั้ง = 0.849

2.10 การคำนวณค่าความน่าจะเป็นของตัวแปรสุ่มพหุนาม

การทดลอง 1 ครั้งมีผลได้ k แบบด้วยความน่าจะเป็น $p_1, p_2, p_3, \ldots, p_k$ ทำการทดลองทั้งหมด n ครั้ง จำนวนครั้งที่เกิดเหตุการณ์ E_i คือ x_i $i = 1, 2, \ldots, k$

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1!x_2!x_3!\ldots x_k!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \ldots p_k^{x_k}$$

ตัวอย่าง 2.10.1 ตามทฤษฎีของพันธุศาสตร์ ในการทดลองผสมพันธ์หนู ปรากฏว่าได้หนูเป็น สีแดง สีดำ และสีขาว ด้วยอัตราส่วน 8:4:4

จงหาความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 5 ตัว สีดำ 2 ตัว และ สีขาว 1 ตัว

แนวคิด ความน่าจะเป็นที่จะได้สีแดง $p_1 := \frac{8}{16}$

ความน่าจะเป็นที่จะได้สีดำ $p_2 := \frac{4}{16}$

ความน่าจะเป็นที่จะได้สีขาว $p_3 := \frac{4}{16}$

X_1, X_2, X_3 เป็นจำนวนสีแดง สีดำ สีขาว ตามลำดับ

การคำนวณด้วย MATHCAD

$p_1 := \frac{8}{16}$

$p_2 := \frac{4}{16}$

$p_3 := \frac{4}{16}$

$P(x1, x2, x3) := \frac{8!}{x1! \cdot x2! \cdot x3!} \cdot (p_1)^{x1} \cdot (p_2)^{x2} \cdot (p_3)^{x3}$

$P(5, 2, 1) = 0.082$

$P(6, 1, 1) = 0.055$

$P(4, 2, 2) = 0.103$

ความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 5 ตัว สีดำ 2 ตัว และ สีขาว 1 ตัว = 0.082

ความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 6 ตัว สีดำ 1 ตัว และ สีขาว 1 ตัว = 0.055

ความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 4 ตัว สีดำ 2 ตัว และ สีขาว 2 ตัว = 0.103

2.11 การหาความน่าจะเป็นของตัวแปรสุ่มปกติมาตรฐาน

ตัวแปรสุ่มปกติมาตรฐาน Z มีฟังก์ชันการแจกแจงความน่าจะเป็น $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$

$$P(a < Z < b) = \int_a^b f(z)dz$$

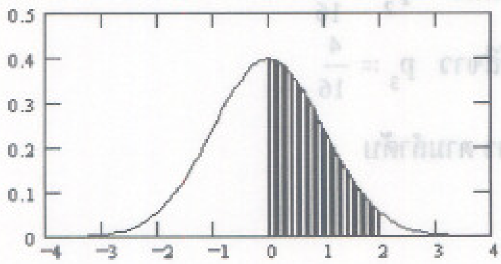
การคำนวณด้วย MATHCAD

$f(z) := \frac{1}{\sqrt{2 \cdot \pi}} \cdot e^{-\frac{1}{2} \cdot z^2}$ $\qquad P(a,b) := \int_a^b f(z)\,dz$

$P(0,1) = 0.3413$

$P(0,2) = 0.4772$

การหาพื้นที่ใต้โค้งปกติมาตรฐานที่แรเงา



พื้นที่ใต้โค้งปกติมาตรฐาน = P(0 < Z < 2) = 0.4771

หมายเหตุ cnorm(b) หมายถึง $P(-\infty < Z < b)$

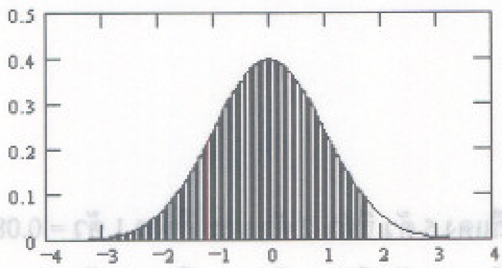
cnorm(0) = 0.5

cnorm(1) = 0.8413

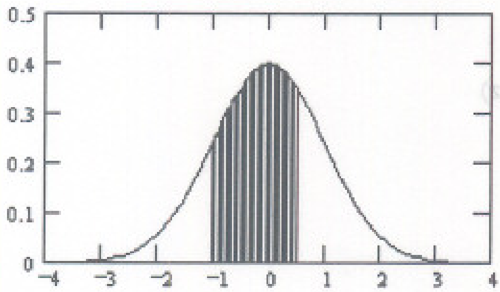
P(a, b) := cnorm(b) − cnorm(a)

P(0, 1) = 0.3413

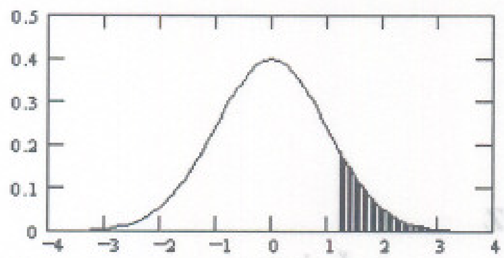
P(0, 2) = 0.4772



พื้นที่ใต้โค้งปกติมาตรฐาน = $P(-\infty < Z < 1.75)$ = cnorm(1.75) = 0.9599



พื้นที่ใต้โค้งปกติมาตรฐาน = P(-1< Z < 0.5) = cnorm(0.5) − cnorm(- 1) = 0.5328



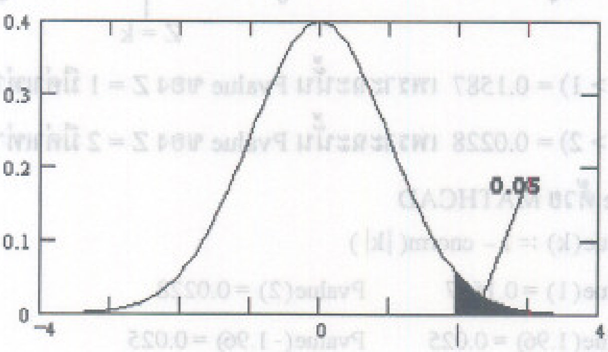
พื้นที่ใต้โค้งปกติมาตรฐาน = P(1.25 < Z < ∞) = 1 − cnorm(1.25) = 0.1056

2.12 การหาค่า z_A

ความหมายของ z_A คือค่าตัวเลขที่ทำให้ $P(Z > z_A) = A$

ตัวอย่างเช่น P(Z > 1.96) = 0.025 เพราะฉะนั้น $z_{0.025} = 1.96$

ตัวอย่าง จงหาค่าของ $z_{0.05}$



$z_{0.05}$ = ?

การคำนวณด้วย MATHCAD

$TOL := 0.000001$

$z := 1$

$Z(A) := root(A - (1 - cnorm(z)), z)$

$Z(0.1) = 1.282$

$Z(0.05) = 1.645$

$Z(0.025) = 1.96$

$Z(0.01) = 2.326$

$Z(0.005) = 2.576$

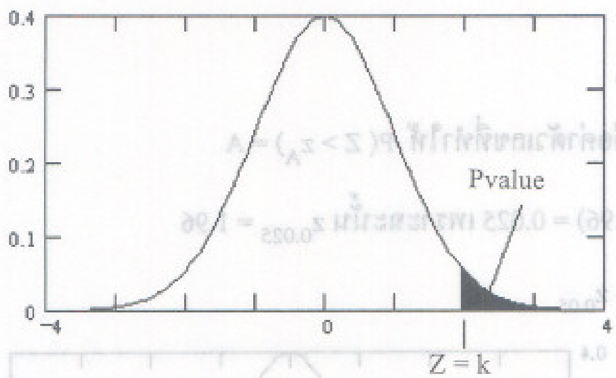
เพราะฉะนั้น ค่า $z_{0.1} = 1.282$

ค่า $z_{0.05} = 1.645$

ค่า $z_{0.025} = 1.96$

2.13 การหาค่า Pvalue ของค่าสถิติ Z

เมื่อกำหนดค่า Z = k ค่า Pvalue ของค่า Z = k หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้งปกติมาตรฐานบนช่วง $(|k|, \infty)$ นั่นคือ $Pvalue_k = P(|k| < Z < \infty)$



ตัวอย่างเช่น $P(Z > 1) = 0.1587$ เพราะฉะนั้น Pvalue ของ Z = 1 มีค่าเท่ากับ 0.1587

$P(Z > 2) = 0.0228$ เพราะฉะนั้น Pvalue ของ Z = 2 มีค่าเท่ากับ 0.0228

การคำนวณค่า Pvalue ด้วย MATHCAD

$Pvalue(k) := 1 - cnorm(|k|)$

$Pvalue(1) = 0.1587$ $Pvalue(2) = 0.0228$

$Pvalue(1.96) = 0.025$ $Pvalue(-1.96) = 0.025$

$Pvalue(1.645) = 0.05$ $Pvalue(-1.645) = 0.05$

2.14 การหาค่า One-Tailed significant และ Two-Tailed significant ของค่าสถิติ Z

กำหนดให้ Z = k

One-Tailed significant ของ Z = k คือ Pvalue(k)

Two-Tailed significant ของ Z = k คือ 2Pvalue(k)

ตัวอย่างเช่น Z = 2 มีค่า

One-Tailed significant ของ Z = 2 คือ Pvalue(2) = 0.0228

Two-Tailed significant ของ Z = 2 คือ Pvalue(2) = 0.0456

การคำนวณด้วย MATHCAD

$\mathrm{OneTailSig}(z) := 1 - \mathrm{cnorm}(|z|)$

$\mathrm{OneTailSig}(1) = 0.1587$

$\mathrm{OneTailSig}(1.96) = 0.025$

$\mathrm{TwoTailSig}(z) := 2\cdot(1 - \mathrm{cnorm}(|z|))$

$\mathrm{TwoTailSig}(1) = 0.3173$

$\mathrm{TwoTailSig}(-1.96) = 0.05$

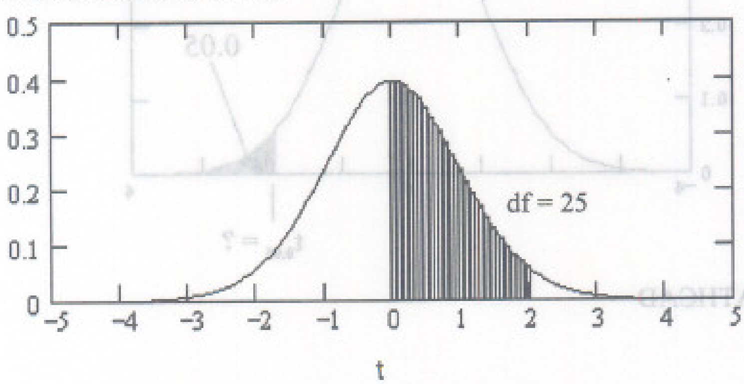
2.15 การหาความน่าจะเป็นของตัวแปรสุ่มที

ตัวแปรสุ่ม T มีองศาความเป็นอิสระ v มีฟังก์ชันการแจกแจงความน่าจะเป็น

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

ความน่าจะเป็น $P(a < T < b) = \int_a^b h(t)dt$

กำหนดองศาความเป็นอิสระ v = 25



การคำนวณด้วย MATHCAD

$v := 25$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$P(a,b) := \int_a^b h(t)\,dt$$

$P(0,2) = 0.4718$

ความน่าจะเป็นที่ $0 < T < 2$ เท่ากับ $P(0 < T < 2) = 0.4718$

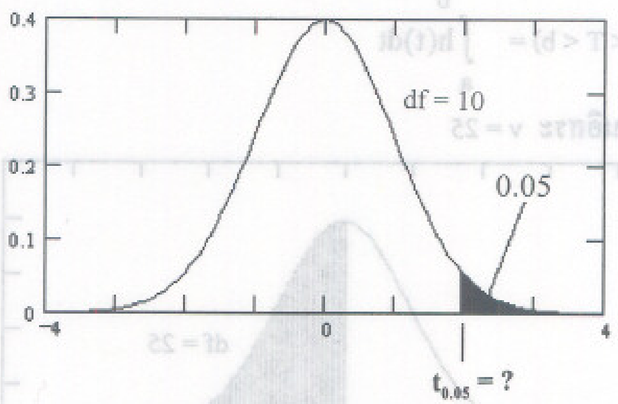
2.16 การหาค่า t_A

ความหมายของ t_A คือค่าตัวเลขที่ทำให้ $P(T > t_A) = A$

ตัวอย่างเช่น เมื่อ df = 25 $P(T > 1.316) = 0.1$ เพราะฉะนั้น $t_{0.1,\, df = 25} = 1.316$

เมื่อ df = 15 $P(T > 2.131) = 0.05$ เพราะฉะนั้น $t_{0.05,\, df = 15} = 2.131$

ตัวอย่าง จงหาค่าของ $t_{0.05,\, df = 10}$



การคำนวณด้วย MATHCAD

$TOL := 0.000001$

$v := 10$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$t := 2$

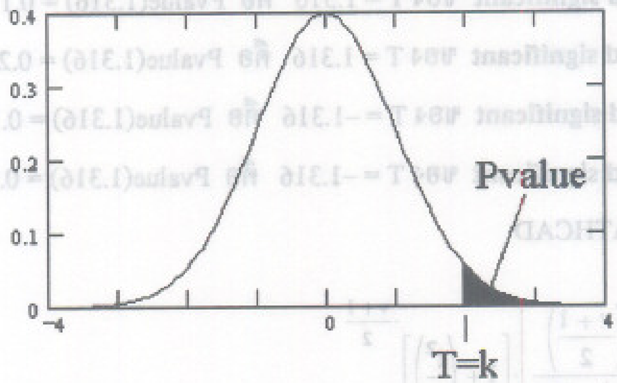
$$T(A) := \mathrm{root}\left[A - \left(0.5 - \int_0^{|t|} h(t)\,dt\right), t\right]$$

$T(0.1) = 1.372$ $T(0.05) = 1.812$ $T(0.025) = 2.228$

เพราะฉะนั้น ค่า $t_{0.05,df=10}$ เท่ากับ = 1.812

2.17 การหาค่า Pvalue ของค่าสถิติ T

เมื่อกำหนดค่า T = k ค่า Pvalue ของค่า T = k หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้ง T บนช่วง (|k| , ∞) นั่นคือ $Pvalue_k = P(T > |k|)$



ตัวอย่างเช่น กำหนด df = 25

$P(T > 1.316) = 0.01$ เพราะฉะนั้น Pvalue ของ T = 1.316 มีค่าเท่ากับ 0.1

$P(T > 2.787) = 0.005$ เพราะฉะนั้น Pvalue ของ T = 2.787 มีค่าเท่ากับ 0.005

$P(T > |-2.485|) = 0.01$ เพราะฉะนั้น Pvalue ของ T = −2.485 มีค่าเท่ากับ 0.01

การคำนวณค่า Pvalue ด้วย MATHCAD

$v := 25$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$\text{Pvalue}(k) := 0.5 - \int_0^{|k|} h(t)\,dt$$

Pvalue(1.316) = 0.1
Pvalue(2.787) = 0.005
Pvalue(- 2.485) = 0.01

2.18 การหาค่า One-Tailed significant และ Two-Tailed significant ของค่าสถิติ T

กำหนดให้ T = k

One-Tailed significant ของ T = k คือ Pvalue(k)

Two-Tailed significant ของ T = k คือ 2Pvalue(k)

ตัวอย่าง กำหนดองศาความเป็นอิสระ v = 25 ค่าสถิติ T = 1.316 มีค่า

One-Tailed significant ของ T = 1.316 คือ Pvalue(1.316) = 0.1

Two-Tailed significant ของ T = 1.316 คือ Pvalue(1.316) = 0.2

One-Tailed significant ของ T = –1.316 คือ Pvalue(1.316) = 0.1

Two-Tailed significant ของ T = –1.316 คือ Pvalue(1.316) = 0.2

การคำนวณด้วย MATHCAD

v := 25

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$\text{OneTailSig}(t) := 0.5 - \int_0^{|t|} h(t)\,dt$$

TwoTailSig(t) := 2·OneTailSig(t)
OneTailSig(1.316) = 0.1
TwoTailSig(1.316) = 0.2
OneTailSig(- 1.316) = 0.1

TwoTailSig(- 1.316) = 0.2

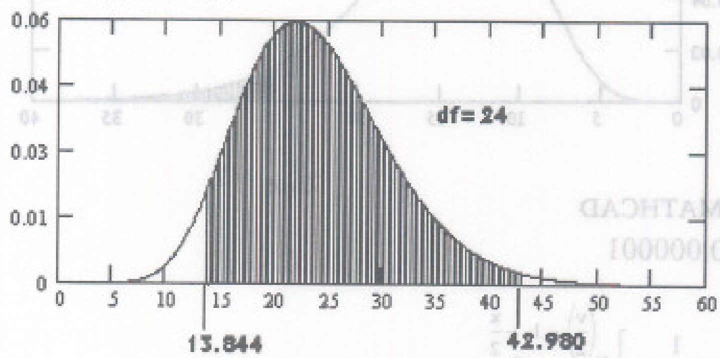
2.19 การหาความน่าจะเป็นของตัวแปรสุ่มไคสแควร์ χ^2

ตัวแปรสุ่มไคสแควร์ χ^2 ที่มีองศาความเป็นอิสระ v มีฟังก์ชันการแจกแจงความน่าจะเป็น

$$f(x) := \left[\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)}\right] \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$

ความน่าจะเป็น $P(a < \chi^2 < b) = \int_a^b f(x)dx$

กำหนดองศาความเป็นอิสระ $v = 24$



ความน่าจะเป็น $P(13.844 < \chi^2 < 42.980) = \int_{13.844}^{42.980} f(x)\, dx = 0.94$

การคำนวณด้วย MATHCAD

$v := 24$ $\qquad f(x) := \left[\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)}\right] \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$

$P(a,b) := \int_a^b f(x)\, dx$

$P(13.844, 42.980) = 0.94$

$P(9.886, 45.558) = 0.99$

$P(9.886, 10^3) = 0.995$

$P(36.415, 10000) = 0.05$

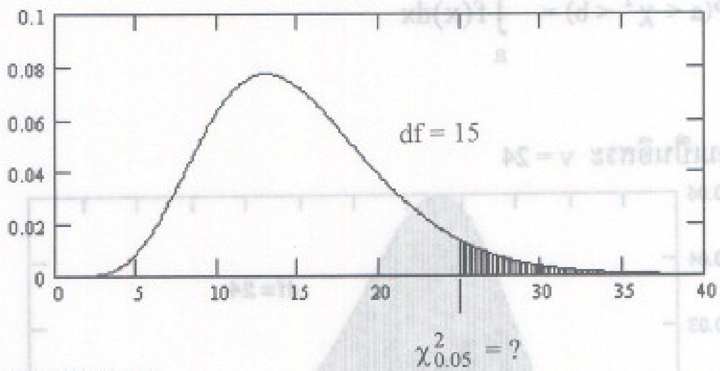
2.20 การหาค่า χ^2_A

ความหมายของ χ^2_A คือค่าตัวเลขที่ทำให้ $P(\chi^2 > \chi^2_A) = A$

ตัวอย่างเช่น เมื่อ df = 24 $P(\chi^2 > 36.415) = 0.05$ เพราะฉะนั้น $\chi^2_{0.05,df=24} = 36.415$

เมื่อ df = 24 $P(\chi^2 > 9.886) = 0.995$ เพราะฉะนั้น $\chi^2_{0.995,df=24} = 9.886$

ตัวอย่าง จงหาค่าของ $\chi^2_{0.05,df=15}$



การคำนวณด้วย MATHCAD

$TOL := 0.000001$

$v := 15$

$$f(x) := \left[\frac{1}{2^{\frac{v}{2}}\cdot\Gamma\left(\frac{v}{2}\right)}\right]\cdot x^{\left(\frac{v}{2}\right)-1}\cdot e^{-\frac{x}{2}}$$

$x := 25$

$$\text{Chisquare}(A) := \text{root}\left[\left[A - \left(1 - \int_0^x f(x)\,dx\right)\right], x\right]$$

$\text{Chisquare}(0.05) = 24.996$

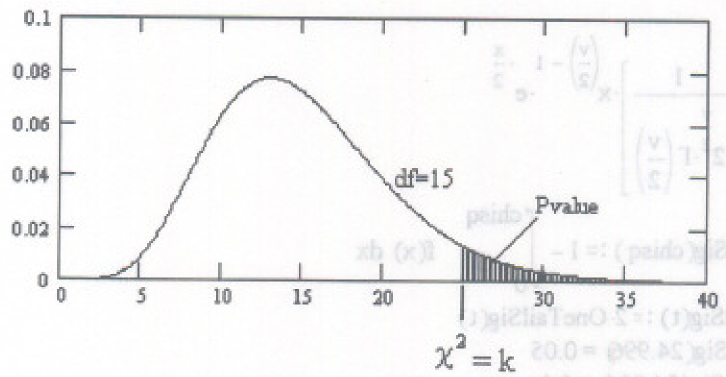
$\text{Chisquare}(0.025) = 27.488$

$\text{Chisquare}(0.005) = 32.801$

เพราะฉะนั้นค่า $\chi^2_{0.05,df=15} = 24.996$, $\chi^2_{0.025,df=15} = 27.488$, $\chi^2_{0.005,df=15} = 32.801$

2.21 การหาค่า Pvalue ของค่าสถิติ χ^2

เมื่อกำหนดค่า $\chi^2 = k$ ค่า Pvalue ของค่า $\chi^2 = k$ หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้ง χ^2 บนช่วง (k, ∞) นั่นคือ $\text{Pvalue}_k = P(\chi^2 > k)$



ตัวอย่างเช่น กำหนด df = 15

$P(\chi^2 > 24.996) = 0.05$ เพราะฉะนั้น Pvalue ของ $\chi^2 = 24.996$ มีค่าเท่ากับ 0.05

$P(\chi^2 > 27.488) = 0.025$ เพราะฉะนั้น Pvalue ของ $\chi^2 = 27.488$ มีค่าเท่ากับ 0.025

การคำนวณค่า Pvalue ด้วย MATHCAD

$v := 15$

$$f(x) := \left[\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)}\right] \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$

$$\text{Pvalue}(k) := 1 - \int_0^k f(x)\,dx$$

$\text{Pvalue}(24.996) = 0.05$
$\text{Pvalue}(27.488) = 0.025$
$\text{Pvalue}(30.578) = 0.01$

2.22 การหาค่า One-Tailed significant และ Two-Tailed significant ของค่าสถิติ χ^2

กำหนดให้ $\chi^2 = k$

One-Tailed significant ของ $\chi^2 = k$ คือ Pvalue(k)

Two-Tailed significant ของ $\chi^2 = k$ คือ 2Pvalue(k)

ตัวอย่าง กำหนดองศาความเป็นอิสระ v = 15 ค่าสถิติ $\chi^2 = 24.996$ มีค่า

One-Tailed significant ของ $\chi^2 = 24.996$ คือ Pvalue(24.996) = 0.05

Two-Tailed significant ของ $\chi^2 = 24.996$ คือ Pvalue(24.996) = 0.1

การคำนวณด้วย MATHCAD

$v := 15$

$$f(x) := \left[\frac{1}{2^{\frac{v}{2}}\cdot\Gamma\left(\frac{v}{2}\right)}\right]\cdot x^{\left(\frac{v}{2}\right)-1}\cdot e^{-\frac{x}{2}}$$

$$\text{OneTailSig(chisq)} := 1 - \int_0^{\text{chisq}} f(x)\,dx$$

$\text{TwoTailSig}(t) := 2\cdot\text{OneTailSig}(t)$

$\text{OneTailSig}(24.996) = 0.05$

$\text{TwoTailSig}(24.996) = 0.1$

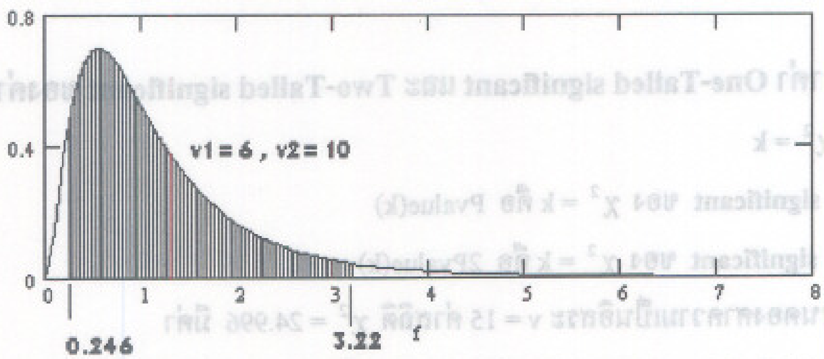
2.23 การหาความน่าจะเป็นของตัวแปรสุ่ม F

ตัวแปรสุ่ม F มีองศาความเป็นอิสระ v_1 และ v_2 มีฟังก์ชันการแจกแจงความน่าจะเป็น

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

ความน่าจะเป็น $P(a < F < b) = \int_a^b h(f)df$

กำหนดองศาความเป็นอิสระ $v_1 = 6$, $v_2 = 10$



$$P(0.246 < f < 3.22) = \int_{0.246}^{3.22} h(f)\,df = 0.9$$

การคำนวณด้วย MATHCAD

v1 := 6
v2 := 10

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

$$P(a,b) := \int_a^b h(f)\, df$$

P(0.246, 3.22) = 0.9
P(0, 3.22) = 0.95
P(3.22, 1000) = 0.05

2.24 การหาค่า f_A

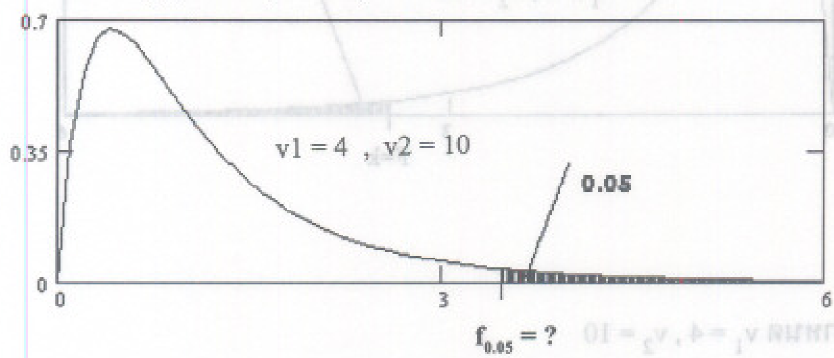
ความหมายของ f_A คือค่าตัวเลขที่ทำให้ $P(F > f_A) = A$

ตัวอย่างเช่น

เมื่อ $v_1 = 6$, $v_2 = 10$ $P(F > 3.22) = 0.05$ เพราะฉะนั้น $f_{0.05} = 3.22$

เมื่อ $v_1 = 6$, $v_2 = 10$ $P(F > 5.39) = 0.01$ เพราะฉะนั้น $f_{0.01} = 5.39$

ตัวอย่าง จงหาค่าของ $f_{0.05}$ เมื่อ $v_1 = 4$, $v_2 = 10$



การคำนวณด้วย MATHCAD

v1 := 4
v2 := 10

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

$x := 6$

$TOL := 0.00001$

$$F(A) := \mathrm{root}\left[A-\left(1-\int_0^{|x|} h(f)\,df\right), x\right]$$

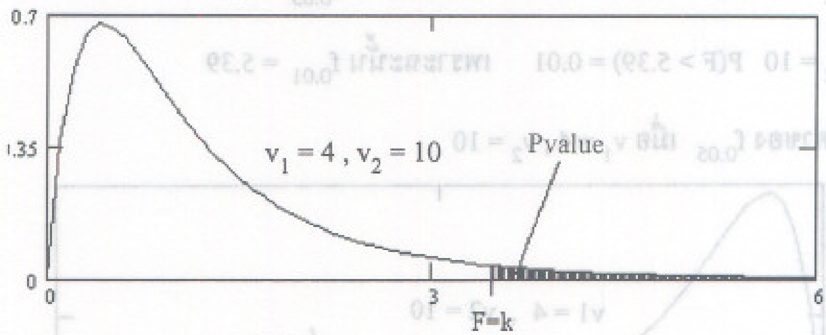
$F(0.01) = 5.994$

$F(0.05) = 3.478$

เพราะฉะนั้น เมื่อ $v_1 = 4$, $v_2 = 10$ จะได้ว่า $f_{0.05} = 3.478$ และ $f_{0.01} = 5.994$

2.25 การหาค่า Pvalue ของค่าสถิติเอฟ

เมื่อกำหนดค่า F = k ค่า Pvalue ของค่า F = k หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้ง F บนช่วง (k , ∞) นั่นคือ $Pvalue_k = P(F > k)$



ตัวอย่างเช่น กำหนด $v_1 = 4$, $v_2 = 10$

$P(F > 3.48) = 0.05$ เพราะฉะนั้น Pvalue ของ F = 3.48 มีค่าเท่ากับ 0.05

$P(F > 5.99) = 0.01$ เพราะฉะนั้น Pvalue ของ F = 5.99 มีค่าเท่ากับ 0.01

การคำนวณค่า Pvalue ด้วย MATHCAD

$v1 := 4$

$v2 := 10$

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

$$Pvalue(k) := 1 - \int_0^k h(f)\,df$$

$Pvalue(5.99) = 0.01$

$Pvalue(3.48) = 0.05$

2.26 การหาค่า One-Tailed significant และ Two-Tailed significant ของค่าสถิติ F

กำหนดให้ $F = k$

One-Tailed significant ของ $F = k$ คือ $Pvalue(k)$

Two-Tailed significant ของ $F = k$ คือ $2Pvalue(k)$

ตัวอย่าง กำหนดองศาความเป็นอิสระ $v_1 = 4$, $v_2 = 10$ ค่าสถิติ $F = 5.99$ มีค่า

One-Tailed significant ของ $F = 5.99$ คือ $Pvalue(5.99) = 0.01$

Two-Tailed significant ของ $F = 5.99$ คือ $Pvalue(5.99) = 0.02$

การคำนวณด้วย MATHCAD

$v1 := 4$ $\quad$ $v2 := 10$ $\quad$ $h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$

$OneTailSig(F) := 1 - \int_0^F h(f)\,df$ $\quad$ $TwoTailSig(F) := 2\cdot OneTailSig(F)$

$OneTailSig(5.99) = 0.01$ $\quad$ $TwoTailSig(5.99) = 0.02$

$OneTailSig(3.48) = 0.05$ $\quad$ $TwoTailSig(3.48) = 0.1$

แบบฝึกหัด 2.

จงหาค่าของ

1. $\binom{20}{5} =$ .... $\qquad \binom{25}{10} =$ ....

2. $\dfrac{e^{-4}4^7}{7!} =$ .... $\qquad \dfrac{e^{-5}5^2}{2!} =$ ....

3. $\dfrac{10!}{(2!)(8!)}(0.5)^2(0.8)^8 =$ .... $\qquad \binom{15}{5}(\frac{1}{4})^5(\frac{3}{4})^{10} =$ ....

4. $\int_2^3 \frac{1}{18}(3+2x)dx =$ .... $\qquad \int_{\frac{1}{4}}^{\frac{1}{2}} \int_0^1 \frac{x(1+3y^2)}{4} dxdy =$ ....

5. $\sum_{x=0}^{3} \frac{e^{-5}5^x}{x!} =$ ....

6. กำหนดข้อมูล $x = \begin{bmatrix} 12 \\ 25 \\ 31 \\ 42 \\ 55 \\ 61 \\ 75 \\ 46 \\ 85 \\ 32 \end{bmatrix}$ จงหาค่าของ mean(x) = ... var(x) = ... stdev(x) = ...

7. กำหนดข้อมูลประชากร 23, 35 , 41 , 36 , 52 , 16 , 74 , 55 จงหาค่าเฉลี่ย และ ความแปรปรวน

8. กำหนดข้อมูลตัวอย่าง 25, 51 , 43 , 38 , 55 , 18 , 64 , 86 จงหาค่าเฉลี่ย และ ความแปรปรวน

9. ตัวแปรสุ่มต่อเนื่อง X มีฟังก์ชันการแจกแจงความน่าจะเป็น $f(x) = \begin{cases} \dfrac{20000}{x^3} & ,x > 100 \\ 0 & ,x \le 100 \end{cases}$

จงหาค่าของ $P(100 < X < 200) =$ ....

$P(0 < X < 150) =$ ....

10. กำหนดตัวแปรสุ่ม X และฟังก์ชันการแจกแจงความน่าจะเป็น ดังนี้

X	ความน่าจะเป็นของ X
2	$\frac{2}{35}$
4	$\frac{14}{35}$
6	$\frac{16}{35}$
8	$\frac{3}{35}$

จงหาค่าคาดคะเน และ ความแปรปรวน ของตัวแปรสุ่ม X

11. กำหนด X = 1, 2, 3, 4, 5, 6 จงหา

$\dfrac{\sum_{i=1}^{n}(x_i-\mu)}{n}$ = ... $\dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$ = ...

$\dfrac{\sum_{i=1}^{n}(x_i-\mu)^3}{n}$ = ... $\dfrac{\sum_{i=1}^{n}(x_i-\mu)^4}{n}$ = ...

12. กำหนด $f(x, y) = \begin{cases} 4xy & ,0 < x < 1, 0 < y < 1 \\ 0 & , \quad \text{otherwise} \end{cases}$ จงหาค่าความน่าจะเป็น

12.1 $P(0 < X < \frac{1}{4}, \frac{1}{4} < Y < \frac{1}{2})$ 12.2 $P(0 < X < \frac{3}{4}, \frac{1}{8} < Y < \frac{1}{2})$

12.3 $P(\frac{1}{4} < Y < \frac{1}{2})$ 12.4 $P(X < Y)$

13. ในการทดสอบยางรถบรรทุก พบว่า 25% ของยางรถแตกก่อนที่เสร็จสิ้นการทดสอบ ในการทดสอบยางทั้งหมด 16 เส้น จงหา

13.1 ความน่าจะเป็นที่จะมียางแตก 1 ถึง 3 เส้น

13.2 ความน่าจะเป็นที่จะมียางแตกอย่างน้อย 4 เส้น

13.3 ความน่าจะเป็นที่จะมียางแตกอย่างมาก 6 เส้น

15. พลุสีกล่องหนึ่งมี 15 อัน เป็นสีเหลือง 6 อัน นอกนั้นเป็นสีม่วง สุ่มเลือก 4 อันแล้วยิง จงหา

15.1 ความน่าจะเป็นที่ทั้ง 4 อันที่เลือกมาเป็นสีม่วง

15.2 ความน่าจะเป็นที่อย่างมาก 2 อัน เป็นสีเหลือง

16. ชายฝั่งตะวันออกของประเทศหนึ่งถูกพายุใต้ฝุ่นโดยเฉลี่ย 4 ครั้งต่อปี จงหา

16.1 ความน่าจะเป็นที่ในปีต่อไป ชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นน้อยกว่า 4 ครั้ง

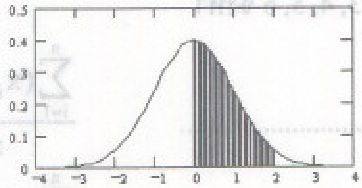
16.2 ความน่าจะเป็นที่ในปีต่อไปชายฝั่งแห่งนี้จะถูกพายุใต้ฝุ่นระหว่าง 6 ถึง 8 ครั้ง

17. ตามทฤษฎีของพันธุศาสตร์ ในการทดลองผสมพันธ์หนู ปรากฏว่าได้หนูเป็น สีแดง สีดำ และสีขาว ด้วยอัตราส่วน 8:5:3 จงหาความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 5 ตัว สีดำ 2 ตัว และ สีขาว 1 ตัว

17.1 ความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 6 ตัว สีดำ 1 ตัว และ สีขาว 1 ตัว

17.2 ความน่าจะเป็นที่ลูกหนู 8 ตัวที่ได้จะเป็น สีแดง 4 ตัว สีดำ 2 ตัว และ สีขาว 2 ตัว

18. จงหา



18.1 พื้นที่ใต้โค้งปกติมาตรฐาน = $P(0 < Z < 2)$ = ...

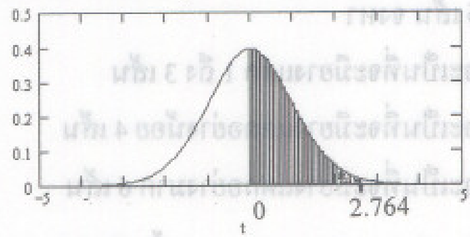
18.2 $P(1.25 < Z < 2.55)$ = ... 18.3 $P(-1.96 < Z < 1.65)$ = ...

19. จงหาค่าของ $z_{0.1}$ = ... $z_{0.05}$ = ... $z_{0.025}$ = ...

20. เมื่อกำหนดค่า k ค่า Pvalue ของค่า Z = k หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้งปกติมาตรฐานบนช่วง $(|k|, \infty)$ นั่นคือ $\text{Pvalue}_k = P(|k| < Z < \infty)$

Pvalue(Z = 1.96) = ... Pvalue(Z = 2.15) = ... Pvalue(Z = -2.15) = ...

21. จงหาค่าความน่าจะเป็นของตัวแปรสุ่ม T



กำหนด องศาความเป็นอิสระ v = 24 $P(0 < T < 2.764)$ = ...

กำหนด องศาความเป็นอิสระ v = 10 $P(0 < T < 2.764)$ = ...

กำหนด องศาความเป็นอิสระ v = 14 $P(0 < T < 2.624)$ = ...

22. กำหนด องศาความเป็นอิสระ v สัญลักษณ์ t_α หมายถึงค่าที่ทำให้ $P(t_\alpha < T < \infty) = \alpha$

กำหนด องศาความเป็นอิสระ v = 12 ค่า $t_{0.025}$ = ...

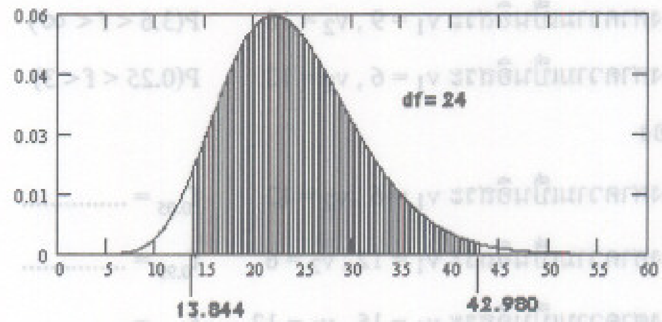
กำหนด องศาความเป็นอิสระ v = 20 ค่า $t_{0.05}$ = ...

23. กำหนดค่า k ค่า Pvalue ของค่า T = k หมายถึงพื้นที่ใต้โค้งทางด้านขวาของเส้นโค้ง T บนช่วง (|k| , ∞) นั่นคือ $\text{Pvalue}_k = P(|k| < T < \infty)$

23.1 กำหนด องศาความเป็นอิสระ v = 4 ค่า Pvalue ของ T = -1.6 มีค่า = ...

23.2 กำหนด องศาความเป็นอิสระ v = 6 ค่า Pvalue ของ T = 4.808 มีค่า = ...

24. จงหาค่าความน่าจะเป็นของตัวแปรสุ่มไคสแควร์



24.1 องศาความเป็นอิสระ v = 24 $P(13.844 < \chi^2 < 42.980)$ = ...

24.2 องศาความเป็นอิสระ v = 10 $P(2.156 < \chi^2 < 25.188)$ = ...

24.3 องศาความเป็นอิสระ v = 8 $P(2.156 < \chi^2 < 25.188)$ = ...

25. กำหนดองศาความเป็นอิสระ v สัญลักษณ์ χ^2_α หมายถึงค่าที่ทำให้ $P(\chi^2_\alpha < \chi^2 < \infty) = \alpha$

25.1 องศาความเป็นอิสระ v = 12 ค่า $\chi^2_{0.025}$ = ...

25.2 องศาความเป็นอิสระ v = 18 ค่า $\chi^2_{0.05}$ = ...

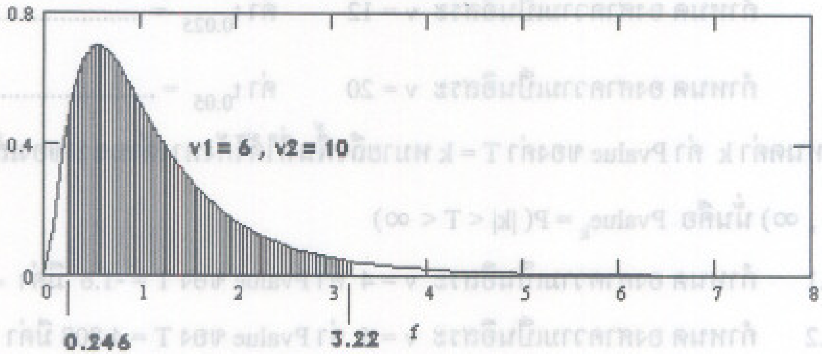
25.3 องศาความเป็นอิสระ v = 20 ค่า $\chi^2_{0.05}$ = ...

26. กำหนดค่า k ค่า Pvalue ของค่า $\chi^2 = k$ หมายถึง พื้นที่ใต้โค้งทางด้านขวาของเส้นโค้งไคสแควร์บนช่วง (k , ∞) นั่นคือ $\text{Pvalue}_k = P(k < \chi^2 < \infty)$

26.1 องศาความเป็นอิสระ v = 5 ค่า Pvalue ของ $\chi^2 = 11.070$ มีค่า = ...

26.2 องศาความเป็นอิสระ v = 2 ค่า Pvalue ของ $\chi^2 = 8.556$ มีค่า = ...

27. จงหาความน่าจะเป็นของตัวแปรสุ่ม F



27.1 องศาความเป็นอิสระ $v_1 = 6$, $v_2 = 10$ $P(0.246 < f < 3.22)$ = ...

27.2 องศาความเป็นอิสระ $v_1 = 9$, $v_2 = 18$ $P(3.6 < f < \infty)$ = ...

27.3 องศาความเป็นอิสระ $v_1 = 6$, $v_2 = 12$ $P(0.25 < f < 3)$ = ...

28. จงหาค่าของ

28.1 องศาความเป็นอิสระ $v_1 = 6$, $v_2 = 12$ $f_{0.05}$ = ...

28.2 องศาความเป็นอิสระ $v_1 = 12$, $v_2 = 6$ $f_{0.95}$ = ...

28.3 องศาความเป็นอิสระ $v_1 = 15$, $v_2 = 12$ $f_{0.01}$ = ...

28.4 องศาความเป็นอิสระ $v_1 = 12$, $v_2 = 15$ $f_{0.99}$ = ...

บทที่ 3

การวิเคราะห์ข้อมูลด้วย MATHCAD

บทที่ 3 นี้จะนำความสามารถของโปรแกรมสำเร็จ MATHCAD เข้ามาทำการวิเคราะห์ข้อมูลเช่น การหาช่วงความเชื่อมั่น หาขนาดตัวอย่าง ทดสอบสมมติฐาน หาความสัมพันธ์ระหว่างตัวแปร การสร้างเส้นโค้งปฏิบัติการ (OC-curve)

3.1 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ μ

ตัวอย่าง 3.1.1 เส้นผ่านศูนย์กลางของโลหะรูปทรงกระบอกมีการแจกแจงปกติ จากการเก็บข้อมูลมาได้ 9 ตัวคือ 1.01 0.97 1.03 1.04 0.99 0.98 0.99 1.01 1.03

จงหาช่วงความเชื่อมั่น 99% ของ μ

วิธีทำ ช่วงความเชื่อมั่น 99% ของ μ คือ $\bar{x} - t_{0.005}\dfrac{s}{\sqrt{n}} < \mu < \bar{x} + t_{0.005}\dfrac{s}{\sqrt{n}}$

การคำนวณด้วย MATHCAD

$$x := \begin{bmatrix} 1.01 \\ 0.97 \\ 1.03 \\ 1.04 \\ 0.99 \\ 0.98 \\ 0.99 \\ 1.01 \\ 1.03 \end{bmatrix}$$

$n := \text{length}(x)$ $\qquad t_{0.005} := 3.35$

$\text{xbar} := \text{mean}(x)$ $\qquad \text{xbar} = 1.006$

$s := \sqrt{\dfrac{\text{var}(x)\cdot n}{n-1}}$ $\qquad s = 0.025$

$$\text{Lowerlimit} := \text{xbar} - t_{0.005}\left(\frac{s}{\sqrt{n}}\right) \qquad \text{Lowerlimit} = 0.978$$

$$\text{Upperlimit} := \text{xbar} + t_{0.005}\left(\frac{s}{\sqrt{n}}\right) \qquad \text{Upperlimit} = 1.033$$

สรุปช่วงความเชื่อมั่น 99 % ของ μ คือ (0.978 , 1.033)

หมายเหตุ เพราะว่า var(x) = ค่าความแปรปรวนของประชากร

เพราะฉะนั้น ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง = $s = \sqrt{\frac{\text{var}(x) \cdot n}{n-1}}$

3.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ μ_D

ตัวอย่าง 3.2.1 สุ่มตัวอย่างนิสิตที่เรียนสถิติ 10 คน สอบถามคะแนนสอบย่อยครั้งที่ 1 และ 2 ได้ข้อมูลดังนี้

คนที่	1	2	3	4	5	6	7	8	9	10
ครั้งที่ 1.	76	60	85	58	91	75	82	64	79	88
ครั้งที่ 2.	81	52	87	70	86	77	90	63	85	83

จงหาช่วงความเชื่อมั่น 98 % ของผลต่างที่แท้จริงของคะแนนสอบย่อย

วิธีทำ ช่วงความเชื่อมั่น 98% ของ μ_D คือ $\bar{d} - t_{0.01}\frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{0.01}\frac{s_d}{\sqrt{n}}$

การคำนวณด้วย MATHCAD

$$x1 := \begin{bmatrix} 76 \\ 60 \\ 85 \\ 58 \\ 91 \\ 75 \\ 82 \\ 64 \\ 79 \\ 88 \end{bmatrix} \qquad x2 := \begin{bmatrix} 81 \\ 52 \\ 87 \\ 70 \\ 86 \\ 77 \\ 90 \\ 63 \\ 85 \\ 83 \end{bmatrix} \qquad d := x1 - x2 \qquad d = \begin{bmatrix} -5 \\ 8 \\ -2 \\ -12 \\ 5 \\ -2 \\ -8 \\ 1 \\ -6 \\ 5 \end{bmatrix}$$

$n := \text{length}(d)$ $\qquad t_{0.01} := 2.82$

$\text{dbar} := \text{mean}(d)$ $\qquad \text{dbar} = -1.6$

$sd := \sqrt{\frac{var(d) \cdot n}{n-1}}$ $\qquad$ $sd = 6.3805$

$Lowerlimit := dbar - t_{0.01} \cdot \left(\frac{sd}{\sqrt{n}}\right)$ $\qquad$ $Lowerlimit = -7.2919$

$Upperlimit := dbar + t_{0.01} \cdot \left(\frac{sd}{\sqrt{n}}\right)$ $\qquad$ $Upperlimit = 4.0919$

สรุปช่วงความเชื่อมั่น 98% ของ μ_D คือ (–7.2919 , 4.0919)

3.3 การสร้างโค้งปฏิบัติการแผนการสุ่มตัวอย่าง (OC–curve)

หลักการทางสถิติ $\qquad$ θ = สัดส่วนของชำรุด

และ $\qquad$ n = จำนวนครั้งที่ทำการทดลอง

จากข้อตกลงของการทดสอบถ้าพบของชำรุดไม่เกิน c ชิ้นจึงจะยอมรับสินค้าทั้งหมด

ภายใต้ข้อสมมติว่า จำนวนสินค้าที่ชำรุดมีการแจกแจงแบบทวินาม

$P(A|\theta)$ = ความน่าจะเป็นที่จะยอมรับสินค้าทั้งหมด เมื่อ สัดส่วนของชำรุด = θ

= ความน่าจะเป็นที่สุ่มตัวอย่าง n พบของชำรุดไม่เกิน c ชิ้นเมื่อ สัดส่วนของชำรุด = θ

$$= \sum_{x=0}^{c} b(x,n,\theta)$$

ตัวอย่าง 3.3.1 n = 10 , c = 2

$P(A|\theta)$ = ความน่าจะเป็นที่จะยอมรับสินค้าทั้งหมด เมื่อ สัดส่วนของชำรุด = θ

= ความน่าจะเป็นที่สุ่มตัวอย่าง n = 10 พบของชำรุดไม่เกิน 2 ชิ้นเมื่อสัดส่วนของชำรุด = θ

$$= \sum_{x=0}^{c} b(x,n,\theta)$$

การคำนวณด้วย MATHCAD

$n := 10$
$c := 2$

$$b(x,n,p) := \frac{n!}{x! \cdot (n-x)!} \cdot p^x \cdot (1-p)^{n-x}$$

$$P(\theta) := \sum_{x=0}^{c} b(x,n,\theta)$$

$P(0.1) = 0.9298$

P(0.5) = 0.0547
P(0.9) = 0

ผลการคำนวณจาก MATHCAD $P(A|\theta = 0.1) = 0.9298$

$P(A|\theta = 0.5) = 0.0547$

$P(A|\theta = 0.9) = 0$

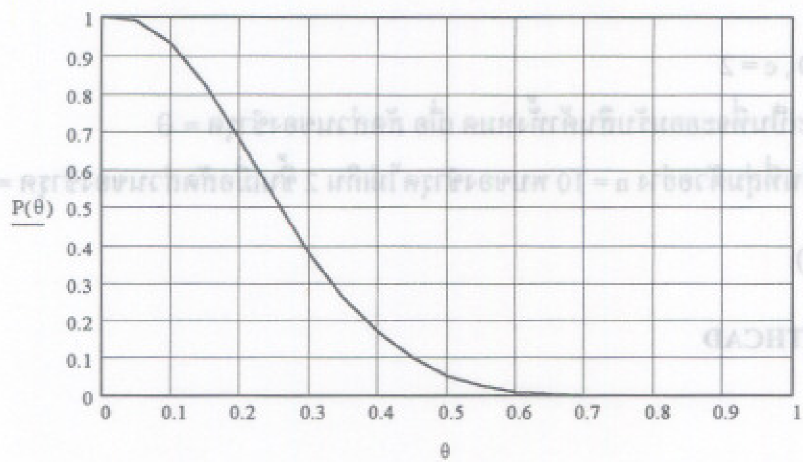
การคำนวณในรูปแบบตาราง

$\theta := 0, 0.1 .. 1$

θ	P(θ)
0	1
0.1	0.929
0.2	0.677
0.3	0.382
0.4	0.167
0.5	0.054
0.6	0.012
0.7	0.001
0.8	0.000
0.9	0
1	0

กราฟ OC-curve ที่ได้

$\theta := 0, 0.05 .. 1$



เส้นโค้ง y = P(θ) เรียกว่า โค้งปฏิบัติการแผนการสุ่มตัวอย่าง (Operating characteristics curve for sampling plan หรือ OC – curve)

ตัวอย่าง 3.3.2 ผลิตภัณฑ์ชนิดหนึ่งบรรจุกล่องละ 25 ชิ้น กำหนดแผนการสุ่มตัวอย่าง (2 , 0) โดยทำการตรวจสอบกล่องละ 2 ชิ้น ถ้าดีทั้ง 2 ชิ้นจึงจะยอมรับว่าผลิตภัณฑ์ดีทั้งกล่อง จงสร้างตารางความน่าจะเป็นของการยอมรับเมื่อสัดส่วนของชำรุด $\theta = 0, \frac{1}{25}, \frac{2}{25}, \ldots, 1$ และเขียนเส้นโค้ง OC-curve

วิธีทำ $P(A|\theta)$ = ความน่าจะเป็นที่จะยอมรับสินค้าทั้งหมด เมื่อ สัดส่วนของชำรุด $= \theta$

= ความน่าจะเป็นที่สุ่มตัวอย่าง $n = 2$ พบของชำรุด 0 ชิ้น $= \dfrac{\binom{25\theta}{0}\binom{25-25\theta}{2}}{\binom{25}{2}}$

การคำนวณด้วย MATHCAD

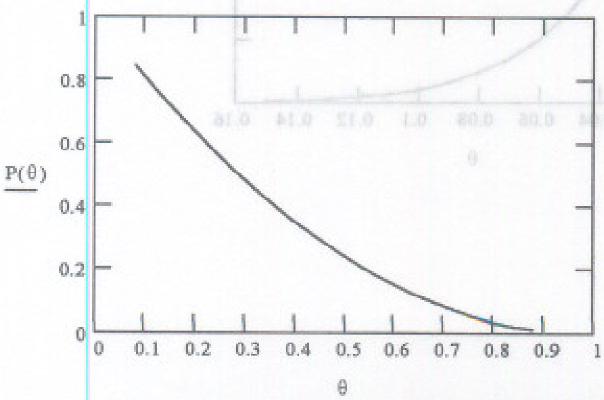
$C(n,r) := \dfrac{n!}{r!\cdot(n-r)!}$ $\qquad P(\theta) := \dfrac{C(25\theta,0)\cdot C(25-25\theta,2)}{C(25,2)}$

$P(0) = 1$ $\qquad P\left(\frac{1}{25}\right) = 0.92$ $\qquad P\left(\frac{2}{25}\right) = 0.8433$ $\qquad P\left(\frac{23}{25}\right) = 0.0033$

$\theta := \frac{3}{25}, \frac{4}{25} .. \frac{12}{25}$ $\qquad \theta := \frac{13}{25}, \frac{14}{25} .. \frac{22}{25}$

θ	P(θ)
0.12	0.77
0.16	0.7
0.2	0.633
0.24	0.57
0.28	0.51
0.32	0.453
0.36	0.4
0.4	0.35
0.44	0.303
0.48	0.26

θ	P(θ)
0.52	0.22
0.56	0.183
0.6	0.15
0.64	0.12
0.68	0.093
0.72	0.07
0.76	0.05
0.8	0.033
0.84	0.02
0.88	0.01



ตัวอย่าง 3.3.3 ในการตรวจสอบผลิตภัณฑ์จำนวนมาก

กำหนดแผนการสุ่มตัวอย่างเชิงเดี่ยว (50 , 1)

โดยทำการตรวจสอบสินค้า 50 ชิ้นจะยอมรับว่าผลิตภัณฑ์ทั้งหมดมีคุณภาพดี ถ้าพบสินค้าชำรุด ไม่มากกว่า 1 ชิ้น จงสร้างตารางความน่าจะเป็นของการยอมรับเมื่อสัดส่วนของชำรุด = θ และเขียนเส้นโค้ง OC-curve

วิธีทำ $P(A|\theta)$ = ความน่าจะเป็นที่จะยอมรับสินค้าทั้งหมด เมื่อ สัดส่วนของชำรุด = θ

= ความน่าจะเป็นที่สุ่มตัวอย่าง n = 50 พบของชำรุดไม่เกิน 1 ชิ้น

X = จำนวนสินค้าชำรุดที่พบ = 0, 1, 2, ...

เราประมาณความน่าจะเป็นด้วยความน่าจะเป็นแบบ ปัวส์ซอง $\mu = n\theta$

$$P(A|\theta) = \sum_{x=0}^{1} \frac{e^{-\mu}\mu^x}{x!}$$

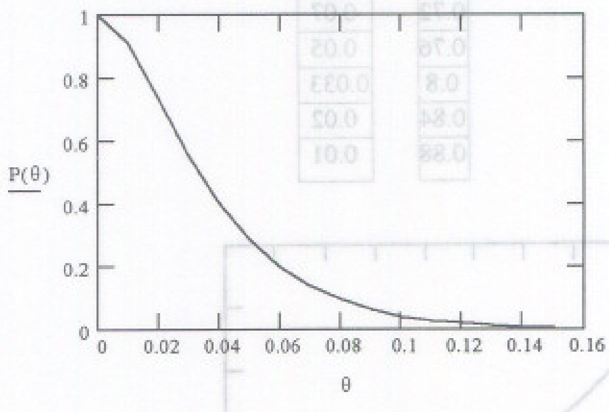
การคำนวณด้วย MATHCAD

n := 50

θ := 0, 0.01.. 0.1

$$P(\theta) := \sum_{x=0}^{1} \frac{e^{-(n\cdot\theta)}\cdot(n\cdot\theta)^x}{x!}$$

θ	P(θ)
0	1
0.01	0.909
0.02	0.735
0.03	0.557
0.04	0.406
0.05	0.287
0.06	0.199
0.07	0.135
0.08	0.091
0.09	0.061
0.1	0.040
0.11	0.026
0.12	0.017
0.13	0.011
0.14	0.007
0.15	0.004



3.4 การคำนวณเกี่ยวกับคุณภาพส่งออกโดยเฉลี่ย

จากแผนการสุ่มตัวอย่างเชิงเดี่ยว (n , c)

θ = สัดส่วนของชำรุด

$P(A|\theta)$ = ความน่าจะเป็นที่จะยอมรับสินค้าทั้งหมด เมื่อ สัดส่วนของชำรุด = θ

$AOQ(\theta) = \theta\ P(A|\theta)$ เรียกว่า **คุณภาพส่งออกโดยเฉลี่ย** (Average Outgoing Quality)

เพราะว่า $0 < \theta\ P(A|\theta) < 1$ เพราะฉะนั้น $\theta\ P(A|\theta)$ มีค่าสูงสุด บนช่วง $0 < \theta < 1$

ให้ $\theta^*\ P(A|\theta^*)$ เป็นค่าสูงสุดของคุณภาพส่งออกโดยเฉลี่ย เราเรียก $\theta^*\ P(A|\theta^*)$ ว่า **ขีดจำกัดของคุณภาพส่งออกโดยเฉลี่ย** (Average Outgoing Quality)

ตัวอย่าง 3.4.1 ในการตรวจสอบผลิตภัณฑ์จำนวนมาก

กำหนดแผนการสุ่มตัวอย่างเชิงเดี่ยว (50 , 1)

โดยทำการตรวจสอบสินค้า 50 ชิ้นจะยอมรับว่าผลิตภัณฑ์ทั้งหมดมีคุณภาพดี ถ้าพบสินค้าชำรุดไม่มากกว่า 1 ชิ้น จงเขียนเส้นโค้ง AOQ-curve และ หาค่าของ $\theta^*\ P(A|\theta^*)$

วิธีทำ $P(A|\theta) = \sum_{x=0}^{1} \frac{e^{-\mu}\mu^x}{x!}$

$AOQ(\theta) = \theta\ P(A|\theta)$

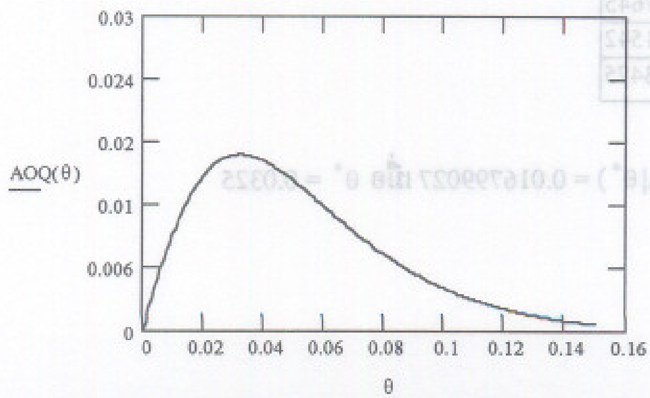
การคำนวณด้วย MATHCAD

$n := 50$

$\theta := 0, 0.001 .. 0.1$

$P(\theta) := \sum_{x=0}^{1} \frac{e^{-(n\cdot\theta)}\cdot(n\cdot\theta)^x}{x!}$

$AOQ(\theta) := \theta\cdot P(\theta)$



จากกราฟ $AOQ(\theta) = \theta\ P(A|\theta)$ มีค่ามากบนช่วง $0.02 < \theta < 0.06$

$\theta := 0.02, 0.03 .. 0.06$

θ	$AOQ(\theta)$
0.02	0.014715177
0.03	0.01673476
0.04	0.01624023
0.05	0.014364874
0.06	0.011948896

จากตาราง $AOQ(\theta) = \theta\ P(A|\theta)$ มีค่ามากบนช่วง $0.03 < \theta < 0.04$

$\theta := 0.03, 0.032 .. 0.04$

θ	$AOQ(\theta)$
0.03	0.01673476
0.032	0.016797790
0.034	0.016770347
0.036	0.016662127
0.038	0.016482461
0.04	0.01624023

จากตาราง $AOQ(\theta) = \theta\ P(A|\theta)$ มีค่ามากบนช่วง $0.032 < \theta < 0.034$

$\theta := 0.032, 0.0325 .. 0.034$

θ	$AOQ(\theta)$
0.032	0.0167977903
0.0325	0.0167990273
0.033	0.0167947645
0.0335	0.0167851542
0.034	0.0167703475

สรุป เลือก $\theta^*\ P(A|\theta^*) = 0.016799027$ เมื่อ $\theta^* = 0.0325$

3.5 การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว $\hat{y} = a + bx$

ตัวอย่าง 3.5.1 กำหนดข้อมูล X และ Y ดังนี้

x	y
1	9
3	17
5	23
9	28
12	33

จงหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว $\hat{y} = a + bx$

วิธีทำ วิธีที่ 1.

จากสูตร

$$b := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right] \cdot \left[\left[\sum_{i=1}^{n} (y_i)\right]\right]}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right]^2}$$

และ

$$a = \overline{Y} - b\overline{X}$$

การคำนวณด้วย MATHCAD

ORIGIN := 1

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix} \qquad n := 5$$

$$b := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right] \cdot \left[\left[\sum_{i=1}^{n} (y_i)\right]\right]}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right]^2} \qquad b = 2.038$$

$$a := \text{mean}(y) - b \cdot \text{mean}(x) \qquad a = 9.775$$

สรุป สมการเส้นถดถอยเชิงเส้นเชิงเดียวคือ $\hat{y} = 9.775 + 2.038x$

วิธีที่ 2. ใช้ฟังก์ชันสำเร็จรูปของ MATHCAD

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix}$$

$b := \mathrm{slope}(x, y)$ $\qquad$ $b = 2.038$

$a := \mathrm{intercept}(x, y)$ $\qquad$ $a = 9.775$

สรุป สมการเส้นถดถอยเชิงเส้นเชิงเดียวคือ $\hat{y} = 9.775 + 2.038x$

3.6 การเขียนกราฟของแผนภาพการกระจาย

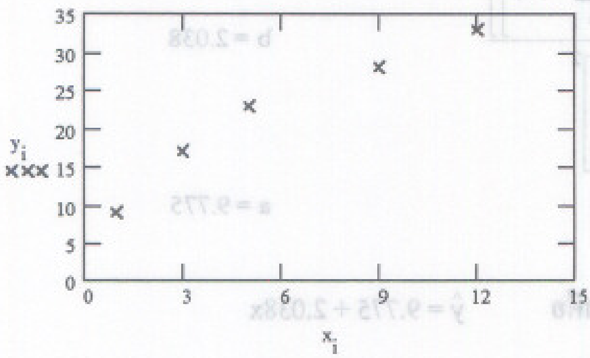
ตัวอย่าง 3.6.1 กำหนดข้อมูล X และ Y ดังนี้

x	y
1	9
3	17
5	23
9	28
12	33

จงเขียนกราฟของแผนภาพการกระจาย

วิธีทำ

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix} \qquad n := \mathrm{length}(x) \qquad i := 1..n$$



การเขียนกราฟของแผนภาพการกระจายและสมการเส้นถดถอย

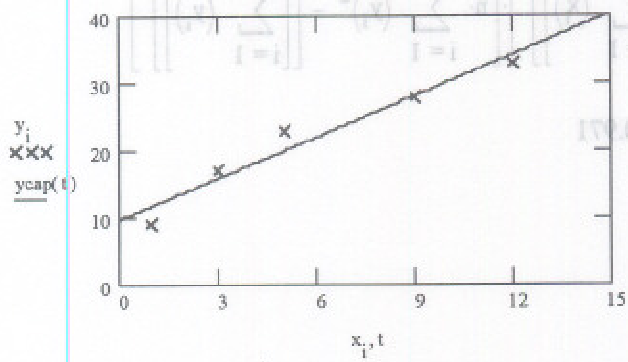
$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix}$$

$n := \text{length}(x)$ $\qquad i := 1 .. n$

$b := \text{slope}(x, y)$ $\qquad b = 2.038$

$a := \text{intercept}(x, y)$ $\qquad a = 9.775$

$\text{ycap}(t) := a + b \cdot t$ $\qquad t := 0, 0.1 .. 15$



3.7 การหาสัมประสิทธิ์สหสัมพันธ์ r

ตัวอย่าง 3.7.1 กำหนดข้อมูล X และ Y ดังนี้

x	y
1	9
3	17
5	23
9	28
12	33

จงหาสัมประสิทธิ์สหสัมพันธ์ r

วิธีทำ แบบที่ 1. ใช้สูตร $r := \dfrac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right] \cdot \left[\left[\sum_{i=1}^{n} (y_i)\right]\right]}{\sqrt{\left[n \cdot \sum_{i=1}^{n} (x_i)^2 - \left[\left[\sum_{i=1}^{n} (x_i)\right]\right]^2\right] \cdot \left[n \cdot \sum_{i=1}^{n} (y_i)^2 - \left[\left[\sum_{i=1}^{n} (y_i)\right]\right]^2\right]}}$

การคำนวณด้วย MATHCAD

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix} \qquad n := 5$$

$$r := \frac{n\cdot\left[\sum_{i=1}^{n}\left(x_i\cdot y_i\right)\right] - \left[\left[\sum_{i=1}^{n}\left(x_i\right)\right]\right]\cdot\left[\left[\sum_{i=1}^{n}\left(y_i\right)\right]\right]}{\sqrt{\left[n\cdot\sum_{i=1}^{n}\left(x_i\right)^2 - \left[\left[\sum_{i=1}^{n}\left(x_i\right)\right]\right]^2\right]\cdot\left[n\cdot\sum_{i=1}^{n}\left(y_i\right)^2 - \left[\left[\sum_{i=1}^{n}\left(y_i\right)\right]\right]^2\right]}}$$

$r = 0.971$

สรุป สัมประสิทธิ์สหสัมพันธ์ $r = 0.971$

แบบที่ 2 ใช้ฟังก์ชันสำเร็จรูป corr

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix}$$

$\mathrm{corr}(x, y) = 0.971$

สรุป สัมประสิทธิ์สหสัมพันธ์ $r = 0.971$

แบบที่ 3

$$x := \begin{bmatrix} 1 \\ 3 \\ 5 \\ 9 \\ 12 \end{bmatrix} \qquad y := \begin{bmatrix} 9 \\ 17 \\ 23 \\ 28 \\ 33 \end{bmatrix} \qquad n := 5$$

$$S_{xx} := \left[\sum_{i=1}^{n}\left(x_i\right)^2 - \frac{\left[\left[\sum_{i=1}^{n}\left(x_i\right)\right]\right]^2}{n}\right] \qquad S_{xx} = 80$$

$$S_{yy} := \left[\sum_{i=1}^{n} (y_i)^2 - \frac{\left[\left[\sum_{i=1}^{n} (y_i)\right]\right]^2}{n}\right] \qquad S_{yy} = 352$$

$$S_{xy} := \left[\left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \frac{\left[\sum_{i=1}^{n} (x_i)\right] \cdot \left[\sum_{i=1}^{n} (y_i)\right]}{n}\right] \qquad S_{xy} = 163$$

$$r := \frac{S_{xy}}{\sqrt{S_{xx} \cdot S_{yy}}} \qquad r = 0.971$$

สรุป สัมประสิทธิ์สหสัมพันธ์ $r = 0.971$

3.8 การเขียนกราฟของแผนภาพการกระจายบนสเกลแบบต่างๆ

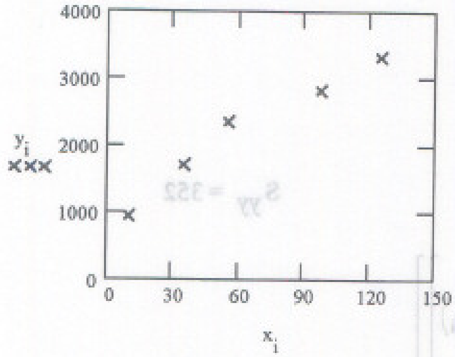
ตัวอย่าง 3.8.1 กำหนดข้อมูล X และ Y ดังนี้

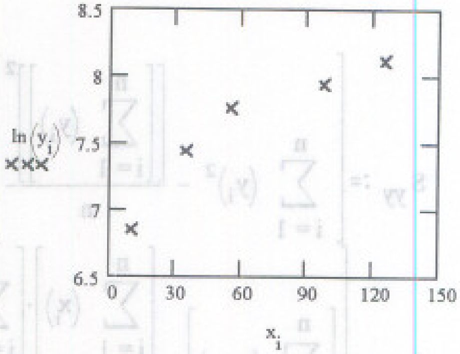
x	y
10	950
35	1700
55	2350
98	2810
125	3330

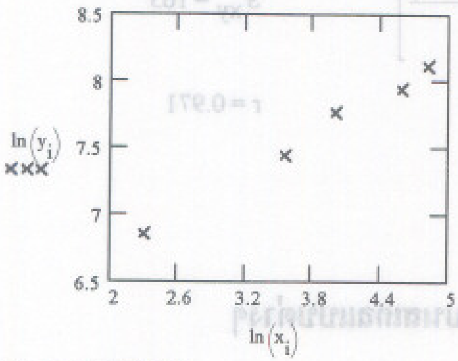
จงเขียนแผนภาพการกระจายบนสเกล (X , Y) , (X , ln(Y)) , (ln(X),Y) , (ln(X) , ln(Y))
และจงคำนวณหาค่าสหสัมพันธ์ทั้ง 4 รูปแบบ

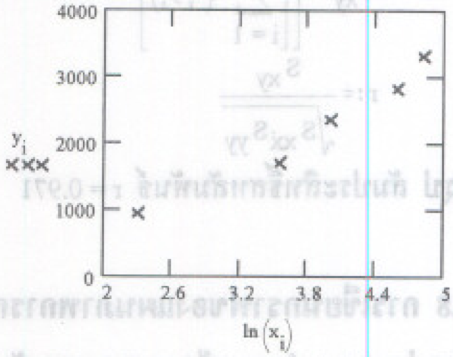
วิธีทำ

$$\text{ORIGIN} := 1 \qquad x := \begin{bmatrix} 10 \\ 35 \\ 55 \\ 98 \\ 125 \end{bmatrix} \qquad y := \begin{bmatrix} 950 \\ 1700 \\ 2350 \\ 2810 \\ 3330 \end{bmatrix}$$









$\mathrm{corr}(x, y) = 0.977615$

$\mathrm{corr}\left(x, \overrightarrow{\ln(y)}\right) = 0.931322$

$\mathrm{corr}\left(\overrightarrow{\ln(x)}, y\right) = 0.980105$

$\mathrm{corr}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right) = 0.996281$

	รูปแบบสมการ	สหสัมพันธ์
(X , Y)	y = a + b x	0.977615
(ln(X) , Y)	y = a + b ln(x)	0.931322
(X , ln(Y))	ln(y) = a + b x	0.980105
(ln(X) , ln(Y))	ln(y) = a + b ln(x)	0.996281

สรุป เลือกรูปแบบสมการ ln(y) = a + b ln(x)

$b := \mathrm{slope}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right)$ $\quad b = 0.492935$

$a := \mathrm{intercept}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right)$ $\quad a = 5.721121$

สมการเส้นถดถอยคือ ln(y) = 5.721121 + 0.492935 ln(x)

แบบฝึกหัด 3.

1. เส้นผ่านศูนย์กลางของโลหะรูปทรงกระบอกมีการแจกแจงปกติจากการเก็บข้อมูลมาได้

   1.11 0.98 1.07 1.14 0.99 0.95 0.89 1.05 1.08

   1.23 0.98 1.01

   1.1 จงหาช่วงความเชื่อมั่น 90% ของ μ

   1.2 จงหาช่วงความเชื่อมั่น 95% ของ μ

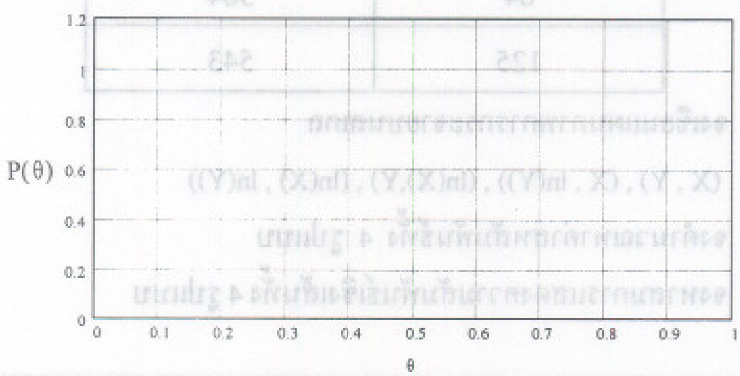
2. สุ่มตัวอย่างนิสิตที่เรียนสถิติ 8 คน สอบถามคะแนนสอบย่อยครั้งที่ 1 และ 2 ได้ข้อมูลดังนี้

คนที่	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
ครั้งที่ 1.	56	68	75	82	86	74	75	63
ครั้งที่ 2.	75	76	87	86	73	85	72	86

   2.1 จงหาช่วงความเชื่อมั่น 90 % ของผลต่างที่แท้จริงของคะแนนสอบย่อย

   2.2 จงหาช่วงความเชื่อมั่น 95 % ของผลต่างที่แท้จริงของคะแนนสอบย่อย

3. จงเขียนกราฟของ $P(\theta) := \sum_{x=0}^{1} b(x, 10, \theta)$



จงหาค่าของ P(0.1) = ... P(0.2) = ...

P(0.25) = ... P(0.5) = ...

4. กำหนดข้อมูล X และ Y ดังนี้

X	12	14	17	25	32	41	47	53	64	75
Y	23	26	35	47	64	78	92	102	115	127

จงหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว $\hat{y} = a + bx$

5. กำหนดข้อมูล X และ Y ดังนี้

X	42	56	74	86	96	105	126	225	243	268
Y	23	26	35	48	64	98	99	102	135	177

5.1 จงเขียนกราฟของแผนภาพการกระจาย

5.2 จงหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว $\hat{y} = a + bx$

5.3 จงหาสัมประสิทธิ์สหสัมพันธ์ r

6. กำหนดข้อมูล X และ Y ดังนี้

x	y
13	185
25	195
37	237
64	364
125	543

6.1 จงเขียนแผนภาพการกระจายบนสเกล

(X , Y) , (X , ln(Y)) , (ln(X),Y) , (ln(X) , ln(Y))

6.2. จงคำนวณหาค่าสหสัมพันธ์ทั้ง 4 รูปแบบ

6.3 จงหาสมการแสดงความสัมพันธ์เชิงเส้นทั้ง 4 รูปแบบ

	รูปแบบสมการ	สหสัมพันธ์
(X , Y)		
(X , ln(Y))		
(ln(X) , Y)		
(ln(X), ln(Y))		

บทที่ 4

ความรู้เบื้องต้นเกี่ยวกับ SPSS for Windows

ในการศึกษาวิชา ความน่าจะเป็นและสถิติ เราเลือกใช้โปรแกรม MATHCAD เข้ามาช่วยในการคำนวณทำให้ได้เห็นที่มาของค่าสถิติต่างๆ เช่นการคำนวณค่าเฉลี่ย ความแปรปรวน ตามเนื้อหาบทเรียนของวิชาความน่าจะเป็นและสถิติ ในการทำงานจริงจะพบว่าข้อมูลมักจะมีเป็นจำนวนมาก ดังนั้นเราจึงเลือกใช้โปรแกรม SPSS มาช่วยในการวิเคราะห์ข้อมูล โปรแกรม SPSS (Statistical Package for the Social Sciences) เป็นโปรแกรมสำเร็จรูปที่ใช้ วิเคราะห์ข้อมูลทางสถิติ ซึ่งเป็นโปรแกรม ที่มีความนิยมใช้งานกันมาก โปรแกรมที่นิยมใช้กันมากคือ SPSS for Windows version 7.5 , 8.0 และ 9.0

โปรแกรม SPSS for Windows version 7.5 สามารถนำ ข้อมูล SPSS หรือ ข้อมูลโปรแกรมเดิมที่สร้างมาจาก SPSS version 3.0 – 6.0 ทั้งในระบบ DOS และ ระบบปฏิบัติการ Windows กลับมาใช้ได้ นอกจากนั้นขั้นตอนการใช้งานต่างของ SPSS for Windows version 7.5 ยังมีคล้ายกับการใช้งาน SPSS for Windows version 8.0 จนสามารถใช้ความรู้จากการวิเคราะห์ข้อมูลจาก version 7.5 ไปใช้งานกับ version 8.0 และ 9.0 ได้ทันที

โปรแกรม SPSS for Windows version 7.5 สามารถรับข้อมูลที่สร้างจากโปรแกรมประเภทต่างๆ ได้หลายประเภท เช่น Excel MATHCAD Microsoft Word ฯลฯ นอกจากนี้โปรแกรม SPSS for Windows ยังสามารถบันทึกคำสั่งที่เกิดจากขั้นตอนการทำงานตามลำดับต่างๆ จากการใช้เมาส์เลือกเมนูของโปรแกรมที่มีอยู่ มาบันทึกเป็น ชุดคำสั่ง (Command Language) เพื่อประโยชน์ในการเรียกคำสั่งเหล่านี้มาใช้ได้อีกในครั้งต่อๆ ไปภายหลัง ผู้ที่เคยใช้โปรแกรมอื่นๆ ที่ทำงานบนวินโดวส์สามารถเรียนรู้การใช้งานโปรแกรม SPSS for Windows ได้อย่างรวดเร็ว และสามารถนำคุณสมบัติของวินโดวส์มาใช้ได้อย่างเต็มที่ เช่น copy cut paste การย้าย การคัดลอก การพิมพ์ การแลกเปลี่ยนข้อมูลระหว่างโปรแกรม ฯลฯ

4.1 คอมพิวเตอร์ที่สามารถทำงานกับโปรแกรม SPSS for Windows

ความต้องการของเครื่องคอมพิวเตอร์ฮาร์ดแวร์และซอฟท์แวร์ที่สามารถนำโปรแกรม SPSS for Windows ไปใช้ได้จะต้องมีคุณสมบัติอย่างต่ำดังต่อไปนี้

❑ เครื่องคอมพิวเตอร์ IBM PC หรือ IBM Compatible ที่ใช้ Window95 , Windows98

❑ หน่วยความจำภายใน (RAM) อย่างน้อย 16 Megabyte

❑ Hard disk มีที่ว่างอย่างน้อย 55 Mb

❑ จอภาพ (Monitor) ต้องสามารถแสดงผลทางด้านกราฟฟิกได้

❑ โปรแกรม Microsoft Windows 95 หรือ Windows 98

❑ โปรแกรม SPSS for Window version 7.5 , 8.0 ,9.0

เพื่อความสะดวกในการทำงานและการเชื่อมโยงข้อมูลน่าจะมี Excel , Microsoft Word , Mathcad

4.2 ความสามารถของโปรแกรม SPSS for Windows

4.2.1 ความสามารถในการวิเคราะห์ข้อมูล

เป็นความสามารถที่จะทำการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติดังต่อไปนี้

1. การคำนวณค่าสถิติเบื้องต้น (Descriptive Statistics) สามารถคำนวณค่าสถิติพื้นฐานทั่วๆ ไป เช่น ค่าเฉลี่ย(Mean) มัธยฐาน(Median) ฐานนิยม(Mode) พิสัย(Range) ความแปรปรวน (Variance) ส่วนเบี่ยงเบนมาตรฐาน(Standard deviation) ฯลฯ

2. การแจกแจงความถี่ (Frequency Distributions) สามารถแจกแจงค่าของตัวแปรตามจำนวนที่นับได้ทั้งแบบทางเดียวและแบบหลายทาง (Crosstabs) พร้อมทั้งแสดงค่าสถิติที่เกี่ยวข้อง เช่น ค่าเฉลี่ย(Mean) มัธยฐาน(Median) ฐานนิยม(Mode) พิสัย(Range) ความแปรปรวน (Variance) ส่วนเบี่ยงเบนมาตรฐาน(Standard deviation) เปอร์เซ็นต์ไทล์ (Percentiles) กราฟแท่งหรือค่าสถิติที่เกี่ยวข้องกับการทดสอบทางสถิติ เช่น Chi-Squares Phi

3. การเปรียบเทียบค่าเฉลี่ย (Mean Groups Comparison) สามารถเปรียบเทียบและทดสอบค่าเฉลี่ยระหว่างกลุ่ม 2 กลุ่มตัวอย่างโดยค่าสถิติ t (Student't) และสำหรับหลายกลุ่มตัวอย่างโดยค่าสถิติ F ด้วยการวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) ทั้งแบบทางเดียวและแบบหลายทาง

4. การหาความสัมพันธ์ระหว่างตัวแปร (Correlation) สามารถคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรแบบต่างๆ เช่น Pearson Kendall Spearman

5. การวิเคราะห์การถดถอย (Regression Analysis) สามารถหาความสัมพันธ์เพื่อการพยากรณ์ โดยวิธีการถดถอยเชิงเส้น (Linear Regression Analysis) ทั้งแบบ 1 ตัวแปรอิสระและแบบหลายตัวแปรอิสระ นอกจากนี้ยังสามารถดูรูปแบบความสัมพันธ์ในลักษณะอื่นๆ ที่ไม่ใช่เส้นตรง เช่น Linear Quadratic Logarithmic ฯลฯ

6. การทดสอบแบบนอนพาราเมตริก (Non – Parametric Test) สามารถวิเคราะห์ข้อมูลโดยวิธีของนอนพาราเมตริกสำหรับการทดสอบแบบต่างๆ เช่น Sign Test Wilcoxon Friedman Kolmokorov – Smirnov ฯลฯ

7. การวิเคราะห์ข้อมูลสำหรับคำตอบแบบหลายคำตอบ (Multiple Response Analysis) สามารถวิเคราะห์ข้อมูลจากแบบสอบถามที่มีตัวเลือกมาให้และผู้ตอบสามารถตอบได้มากกว่า 1 คำตอบ

4.2.2 ความสามารถในการนำเสนอข้อมูลด้วยกราฟ

โปรแกรม SPSS สำหรับวินโดวส์สามารถนำเสนอข้อมูลในรูปของกราฟ หรือชาร์ทแบบต่างๆ เช่น กราฟแท่ง (Bar, Histogram) กราฟเส้น (Line) กราฟวงกลม (Pie) และกราฟชนิดอื่นๆ (Area, High – Low)

4.2.3 ความสามารถในการทำงานด้านอื่นๆ

ในการใช้งานโปรแกรม SPSS นอกจากจะทำการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติแล้วผู้ใช้อาจจะมีการดำเนินการกับข้อมูลในลักษณะต่างๆ เช่น สร้างตัวแปรเพิ่ม เรียงลำดับข้อมูล คัดเลือกข้อมูลมาทำการวิเคราะห์ ฯลฯ ซึ่งสามารถแบ่งเป็นประเภทต่างๆ ได้ดังนี้

1. การเปลี่ยนรูปข้อมูล (Data Transformation) โดยการเปลี่ยนค่าใหม่ จัดค่าใหม่ หรือสร้างตัวแปรใหม่ด้วยฟังก์ชันพิเศษต่างๆทางคณิตศาสตร์ที่โปรแกรมให้มา

2. การจัดกลุ่มตัวแปร (Define Set of Variables) โดยการเลือกตัวแปร หรือจัดกลุ่มตัวแปรไว้เป็นชุดต่างๆ เพื่อนำมาวิเคราะห์เป็นชุดๆ ในภายหลัง

3. การเลือกข้อมูล (Select Case) โดยการเลือกข้อมูลด้วยเงื่อนไขต่างๆ ที่ต้องการ หรือการเลือกข้อมูลโดยการสุ่มตัวอย่าง

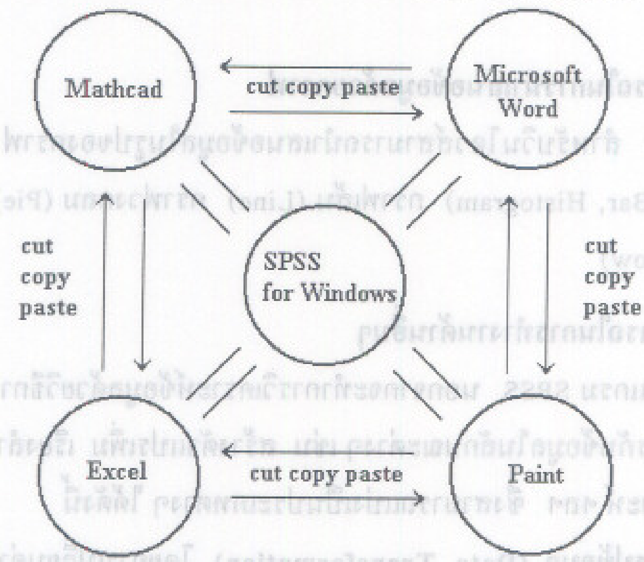
4. การสร้างข้อมูลแบบอนุกรมเวลา (Create Time Series) โดยการสร้างข้อมูลที่เกิดขึ้นตามเวลา เช่น วัน เดือน ไตรมาส ฯลฯ สำหรับการวิเคราะห์แบบอนุกรมเวลา

5. การดำเนินการกับข้อมูลในลักษณะอื่นๆ โดยการเรียงลำดับข้อมูล การให้น้ำหนักหรือความสำคัญแก่ชุดข้อมูล การสลับที่ข้อมูลระหว่างแถวและคอลัมน์

6. การจัดการกับไฟล์ข้อมูล โดยการรวมไฟล์ข้อมูลตั้งแต่ 2 ไฟล์ ด้วยวิธีการต่างๆ เช่น รวมตัวแปร รวมชุดข้อมูล ฯลฯ

4.2.4 ความสามารถในการเชื่อมโยงข้อมูลกับโปรแกรมอื่นๆ

การทำงานของโปรแกรม SPSS for Windows version 7.5 เป็นการทำงานภายใต้ระบบปฏิบัติการ Windows ดังนั้นเราสามารถใช้ความสามารถขั้นพื้นฐาน เช่น การเลือกบริเวณเพื่อ copy cut paste ฯลฯ แล้วนำข้อมูลนั้นไปใช้กับโปรแกรมอื่นๆ เช่น Excel , Micorsoft Word , Mathcad หรือนำข้อมูลจาก Excel , Micorsoft Word , Mathcad มาใช้กับ SPSS for Windows



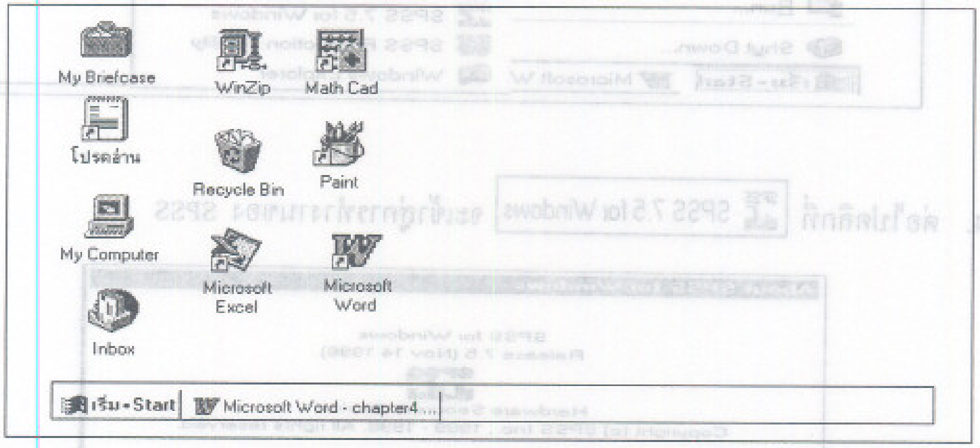
ตัวอย่างเช่น ข้อมูลในรูปแบบ column สามารถนำมาเป็นข้อมูลในรูปแบบตัวแปรของ SPSS for Windows ได้ หรือข้อมูลที่วิเคราะห์ได้จาก SPSS for Windows สามารถ copy รูปแบบตารางไปเป็นตารางของ Microsoft Word ได้ทันที

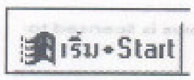
หมายเหตุ ตัวอย่างการใช้งานจะนำภาพที่เกิดจากการใช้โปรแกรม SPSS for Windows version 7.5 มาประกอบคำอธิบาย ข้อแตกต่างที่เปลี่ยนไปจาก SPSS 7.5 เป็น SPSS 8.0 และ 9.0 ได้นำไปกล่าวไว้ที่ภาคผนวกที่ 3 และ ภาคผนวกที่ 4

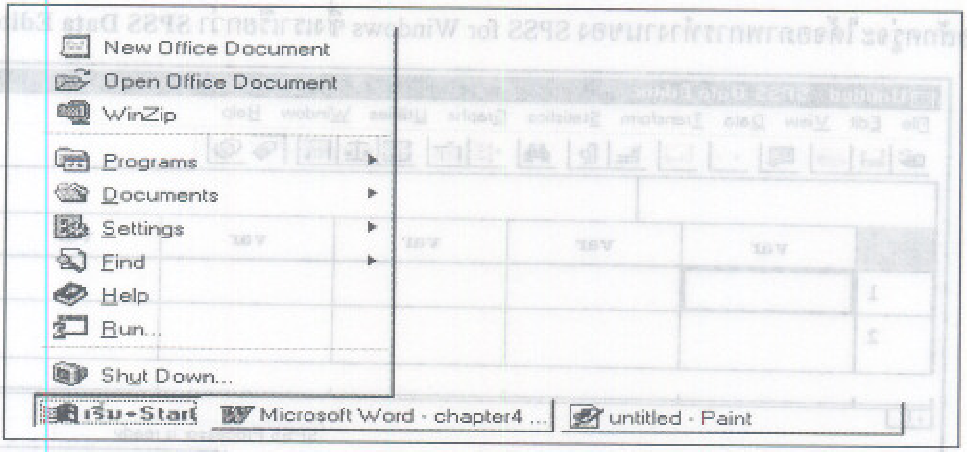
4.3 การเข้าสู่การทำงานของโปรแกรม SPSS for Windows

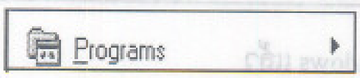
สำหรับคอมพิวเตอร์ที่ติดตั้งโปรแกรม SPSS for Windows เสร็จเรียบร้อยแล้ว การเข้าสู่การทำงานมีขั้นตอนดังนี้

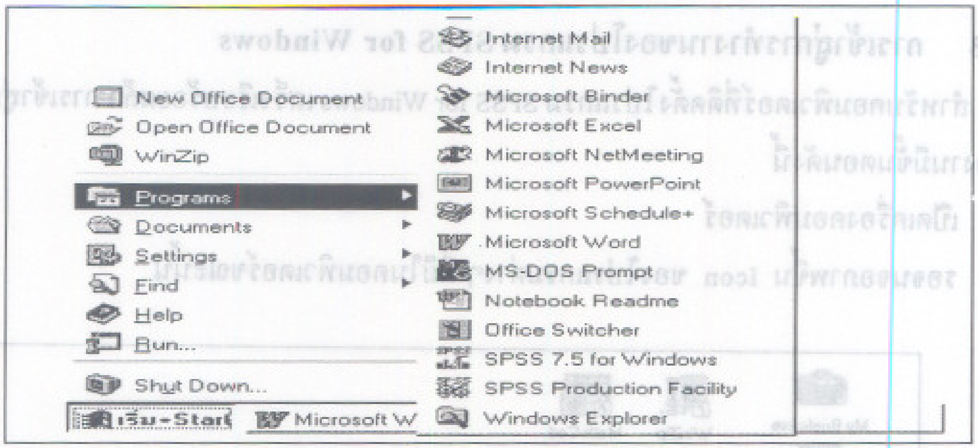
1. เปิดเครื่องคอมพิวเตอร์
2. รอจนจอภาพขึ้น Icon ของโปรแกรมต่างๆ ที่มีในคอมพิวเตอร์ขณะนั้น



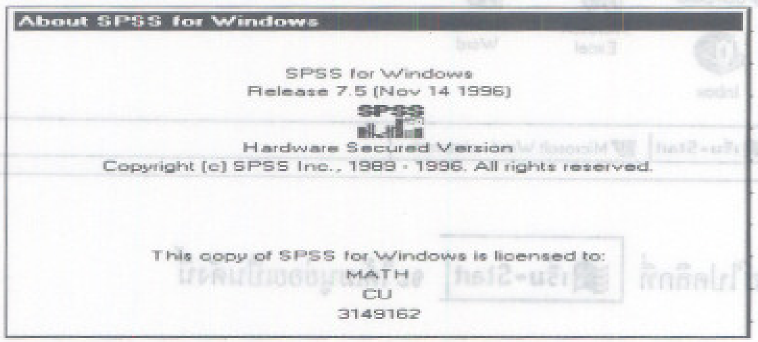
3. เลื่อนเมาส์ไปคลิกที่  จะได้เมนูย่อยเป็นดังนี้



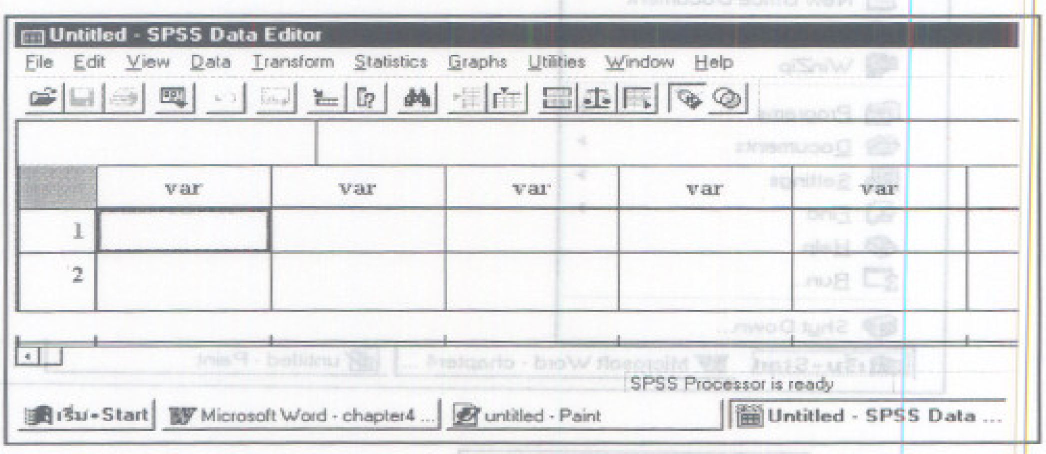
4. ต่อไปคลิกที่  จะได้เมนูย่อยเป็นดังนี้



4. ต่อไปคลิกที่ SPSS 7.5 for Windows จะเข้าสู่การทำงานของ SPSS



รอสักครู่จะได้จอภาพการทำงานของ SPSS for Windows ซึ่งเราเรียกว่า **SPSS Data Editor**



ขณะเราพร้อมที่จะทำงานกับ SPSS for Windows แล้ว

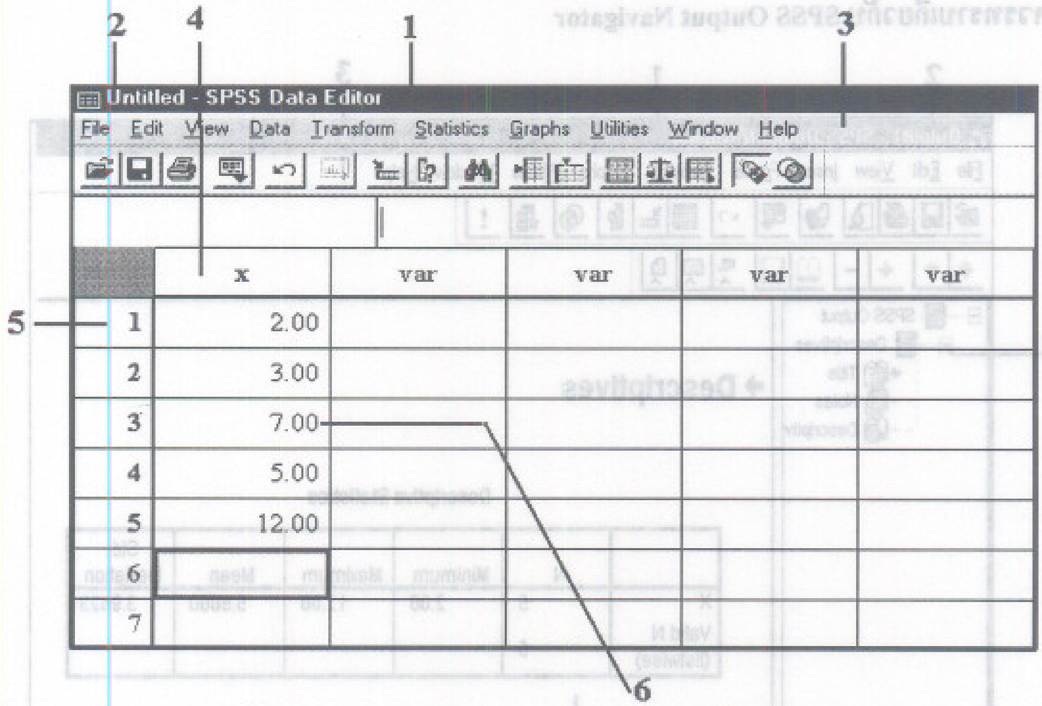
4.4 WINDOW ของการทำงานแบบต่างๆ ของ SPSS for Windows

การทำงานของโปรแกรม SPSS มีการจำแนกส่วนของ WINDOW ที่สำคัญดังนี้

1. SPSS Data Editor

เป็นวินโดวส์สำหรับเก็บไฟล์ข้อมูลที่จะนำมาวิเคราะห์ด้วยโปรแกรม SPSS ซึ่งผู้ใช้อาจจะสร้างไฟล์ข้อมูลใหม่ หรือนำข้อมูลที่สร้างจากโปรแกรมอื่นๆ เรียกเข้ามาไว้ใน Data Editor แล้วใช้งานต่อไป Data Editor จะเปิดได้ครั้งละ 1 วินโดวส์เท่านั้น

ข้อควรทราบเกี่ยวกับ SPSS Data Editor



หมายเลข 1. ชื่อบอกชนิดของ Window ใน SPSS ขณะนี้คือ **SPSS Data Editor**

หมายเลข 2. ชื่อแฟ้มข้อมูลที่กำลังใช้งาน หากยังไม่ได้ตั้งชื่อ SPSS จะใช้ชื่อว่า Untitled

หมายเลข 3. แถบเมนูของ **SPSS Data Editor**

หมายเลข 4. ชื่อตัวแปรของข้อมูล

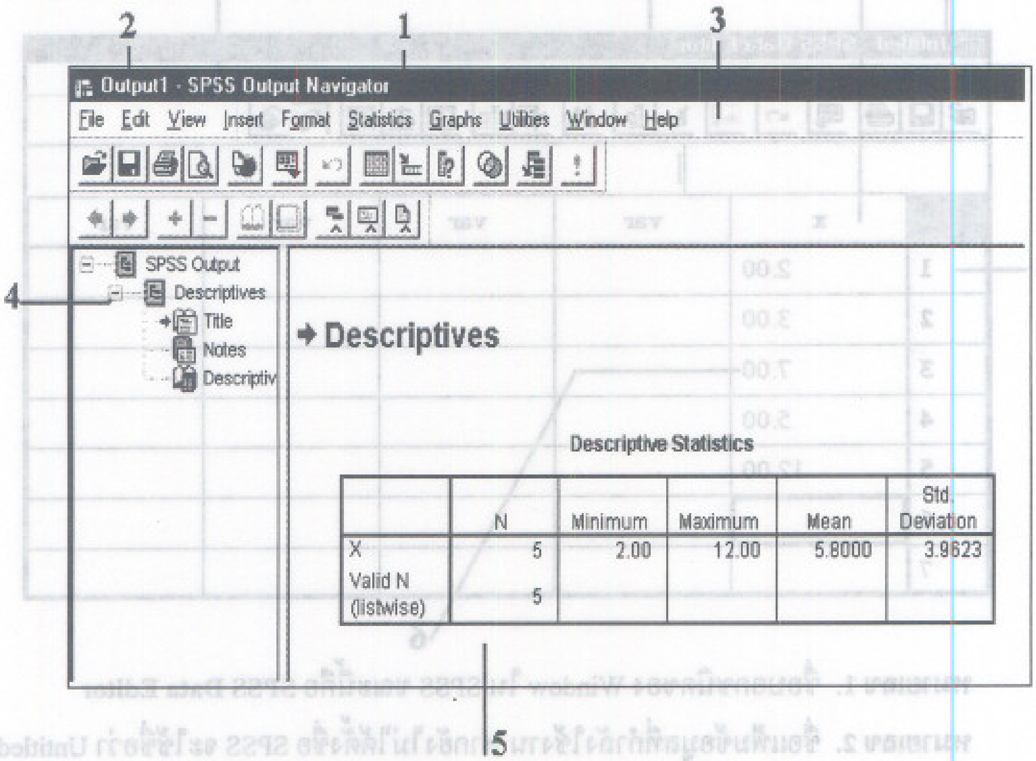
หมายเลข 5. ลำดับที่ของค่าสังเกตในแฟ้มข้อมูล

หมายเลข 6. ค่าของข้อมูล ค่าสังเกตตัวที่ 3 ของตัวแปร x

2. SPSS Output Navigator

เป็นวินโดวส์สำหรับเก็บบันทึกผลลัพธ์ของการวิเคราะห์ข้อมูล ที่เกิดขึ้นจากการใช้งานโปรแกรม SPSS โดยจะบันทึกผลลัพธ์ที่เกิดขึ้นเองทุกครั้งที่มีการใช้งานของโปรแกรม SPSS และผลลัพธ์จะถูกบันทึกอย่างต่อเนื่องจนกว่าจะมีการสั่งให้บันทึกผลลัพธ์ในวินโดวส์ Output อื่นๆ ผู้ใช้สามารถ เปิดวินโดวส์ Output ได้มากกว่า 1 วินโดวส์ Output ถ้ามีการเปิดวินโดวส์ Output มากกว่า 1 วินโดวส์ จะต้องมีการกำหนดวินโดวส์หนึ่งให้ทำหน้าที่เก็บผลลัพธ์ที่เกิดจากการประมวลผล

ข้อควรทราบเกี่ยวกับ SPSS Output Navigator



หมายเลข 1. ชื่อบอกชนิดของ Window ใน SPSS ขณะนี้คือ **SPSS Output Navigator**

หมายเลข 2. ชื่อแฟ้ม Output File ที่กำลังใช้งาน หากยังไม่ได้ตั้งชื่อจะใช้ชื่อว่า Output1

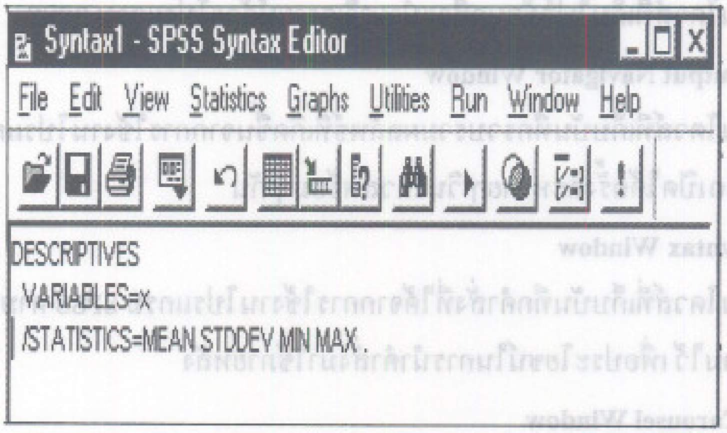
หมายเลข 3. แถบเมนูของ **SPSS Output Navigator**

หมายเลข 4. แผนภูมิต้นไม้แสดงลำดับและตำแหน่งของการแสดงผล

หมายเลข 5. ผลของการวิเคราะห์ข้อมูล

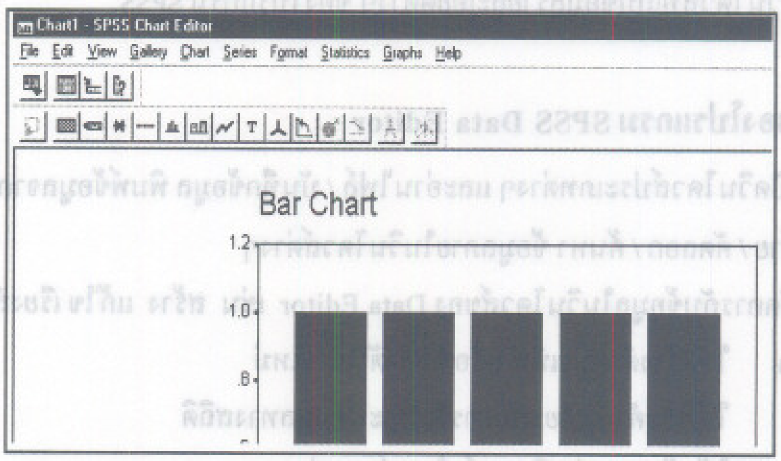
3. SPSS Syntax Editor

เป็นวินโดวส์สำหรับเก็บบันทึกคำสั่งที่ได้จากการใช้งานโปรแกรม SPSS ตามขั้นตอนต่างๆ ที่ทำของผู้ใช้ขณะนั้น ให้ผู้ใช้นำคำสั่งที่เกิดขึ้นนี้มาใช้ได้อีกโดยไม่ต้องสั่งการทำงานแบบเก่าซ้ำอีก หรือผู้ใช้สามารถเปลี่ยนแปลงแก้ไขใหม่ได้



4. SPSS Chart Editor

เป็นวินโดวส์ของกราฟหนึ่งๆ สำหรับ ให้ผู้ใช้ เปลี่ยนแปลงแก้ไขกราฟ ที่สร้างขึ้นมา เช่น เปลี่ยนรูปแบบตัวอักษร เปลี่ยนสี ฯลฯ



4.5 สรุปเนื้อหาของคำสั่งและขั้นตอนการทำงานโดยย่อของ SPSS for Windows

1. ประเภทของ Windows ในโปรแกรม SPSS For Windows V 7.5

(1) SPSS Data Editor Window

เป็นวินโดวส์ที่เก็บไฟล์ข้อมูลที่จะนำมาวิเคราะห์ด้วยโปรแกรม SPSS

(2) SPSS Output Navigator Window

เป็นวินโดวส์ที่เก็บบันทึกรวบรวมผลลัพธ์ที่เกิดขึ้นจากการใช้งานโปรแกรม SPSS สามารถเปิดได้ครั้งละหลายๆ วินโดวส์พร้อมๆ กัน

(3) SPSS Syntax Window

เป็นวินโดวส์ที่เก็บบันทึกคำสั่งที่ได้จากการใช้งานโปรแกรม SPSS ตามขั้นตอนต่างๆ มารวบรวมไว้ เพื่อประโยชน์ในการนำคำสั่งมาใช้ภายหลัง

(4) Chart Carousel Window

เป็นวินโดวส์ที่เก็บบันทึกรวบรวมกราฟ ต่างๆ ทั้งหมดที่เกิดขึ้นจากการ โปรแกรม SPSS

(5) Chart Window

เป็นวินโดวส์ของกราฟที่มีอยู่สำหรับให้ผู้ใช้เปลี่ยนแปลง แก้ไขรายละเอียดต่างๆ

(6) Help Window

เป็นวินโดวส์ที่เก็บข้อมูลรายละเอียดต่างๆ ของโปรแกรม SPSS

2. Menu ของโปรแกรม SPSS Data Editor

File ใช้เปิดวินโดวส์ประเภทต่างๆ และอ่านไฟล์ / บันทึกข้อมูล พิมพ์ข้อมูลจากวินโดวส์ต่างๆ

Edit ใช้ย้าย / คัดลอก / ค้นหา ข้อมูลภายในวินโดวส์ต่างๆ

Data ใช้จัดการกับข้อมูลในวินโดวส์ของ Data Editor เช่น สร้าง แก้ไข เรียงลำดับข้อมูล

Transform ใช้สร้างตัวแปรเพิ่ม หรือ จัดค่าตัวแปรใหม่

Statistics ใช้เรียกคำสั่งเกี่ยวกับการวิเคราะห์ข้อมูลทางสถิติ

Graphs ใช้สร้างกราฟ หรือ ชาร์ทในรูปแบบต่างๆ

Utilities ใช้เปลี่ยนแปลงลักษณะต่างๆ ของตัวอักษร กำหนดกลุ่มตัวแปรที่จะใช้งาน กำหนดรูปแบบของวินโดวส์ Data Editor

Windows ใช้จัดเรียง Windows ในรูปแบบต่างๆ การเลือกแสดงสถานะต่างๆ ของ Window การเรียกใช้ Windows ประเภทต่างๆ

Help ใช้ขอคำอธิบายการใช้โปรแกรม SPSS

3. การจัดเตรียมข้อมูลโดย SPSS Data Editor

(3.1) กำหนดชื่อตัวแปรและรายละเอียดของตัวแปร

- เลือกเซลล์หรือคอลัมน์ที่ต้องการ กำหนด หรือ เปลี่ยนชื่อตัวแปร
- **Data / Define Variable**
- กำหนดชื่อตัวแปรในบ็อกซ์ของ **Variable name**

ถ้าต้องการเปลี่ยนแปลงรายละเอียดของตัวแปรให้แตกต่างจากที่โปรแกรมกำหนดให้ สามารถเปลี่ยนแปลงได้ 4 ประเภท ดังนี้

1.1 กำหนดประเภทของตัวแปร เลือกปุ่ม **Type...**

1.2 กำหนดข้อความขยายชื่อและอธิบายค่าตัวแปร เลือกปุ่ม **Labels...**

1.3 กำหนดค่าที่ขาดหายไปหรือค่าไม่สมบูรณ์ เลือกปุ่ม **Missing Values**

1.4 กำหนดความกว้าง/จัดข้อความของคอลัมน์ เลือกปุ่ม **Column Format**

(3.2) การป้อนข้อมูล

2.1 ใช้แป้น **Enter** สำหรับการป้อนข้อมูลครั้งละ 1 ตัวแปร

2.2 ใช้แป้น ↓ → ← ↑ สำหรับการป้อนข้อมูลแล้วเลื่อนไปเซลล์ถัดไป

2.3 ใช้แป้น **Tab** สำหรับการป้อนข้อมูลครั้งละ 1 ชุด (แถว)

(3.3) การบันทึกข้อมูล

- **File / Save Data** สำหรับการบันทึกภายใต้ชื่อเดิม
- **File / Save As...** สำหรับการบันทึกภายใต้ชื่อใหม่

(3.4) การเรียกใช้ข้อมูลที่บันทึกไว้แล้ว

File / Open / Data... เลือกหรือพิมพ์ชื่อแฟ้มที่ต้องการ

(3.5) การพิมพ์ ข้อมูล คำสั่ง หรือ ผลลัพธ์ออกเครื่องพิมพ์

- เลือกวินโดวส์ที่ต้องการ (**Data Editor,Output Navigator , ...**)
- **File / Print... / OK**

4. การทำงานที่สำคัญใน SPSS Data Editor

(4.1) การค้นหาชุดข้อมูลและตัวแปร

4.1.1 การค้นหาชุดข้อมูล

- **Data / Go to Case...**
- ป้อนตำแหน่งของชุดข้อมูลที่ต้องการค้นหา

4.1.2 การค้นหาตัวแปร

- **Utilities / Variables...**
- เลือกตัวแปรที่ต้องการ

(4.2) การคัดลอก หรือ ย้ายข้อมูล

- เลือกข้อมูลที่ต้องการคัดลอก หรือ ย้าย
- **Edit / Copy หรือ Edit / Cut**
- เลือกเซลล์ซึ่งเป็นตำแหน่งที่ต้องการคัดลอกข้อมูลมาไว้
- **Edit / Paste**

(4.3) การแทรก หรือ ลบชุดข้อมูล

4.3.1 การแทรกชุดข้อมูล

- คลิก ที่หัวแถวที่ต้องการแทรกไว้ (จะแทรกไว้เหนือแถวที่เลือก)
- **Data / Insert Case**

4.3.2 การลบชุดข้อมูล

- คลิก ที่หัวแถวหรือกลุ่มของหัวแถว

(4.4) การแทรก หรือ ลบตัวแปร

4.4.1 การแทรกตัวแปร

- คลิก ที่ชื่อตัวแปรที่ต้องการแทรก (จะแทรกไว้ข้างหน้าตัวแปรที่เลือก)
- **Data / Insert Variable**

4.4.2 การลบตัวแปร

- คลิก ที่ชื่อตัวแปร หรือกลุ่มของตัวแปร
- **Edit / Clear** (หรือกดแป้น Del)

5. การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Summarize

(5.1) การแจกแจงความถี่แบบทางเดียว

- **Statistics / Summarize / Frequencies...**
- เลือกตัวแปรไว้ในบ็อกซ์ของ Variable(s)
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics , Chart หรือ Format
- เลือกปุ่ม **OK**

(5.2) การคำนวณค่าสถิติเบื้องต้น

- **Statistics / Summarize / Descriptives...**
- เลือกตัวแปรไว้ในบ็อกซ์ของ Variable(s)
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(5.3) การตรวจสอบข้อมูล

- **Statistics / Summarize / Explore...**
- เลือกตัวแปรมาไว้ในบ็อกซ์ของ Dependent List
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics , Plots หรือ Options
- เลือกปุ่ม **OK**

(5.4) การแจกแจงความถี่ตั้งแต่ 2 ทาง

- **Statistics / Summarize / Crosstabs...**
- เลือกตัวแปรอย่างน้อย 1 ตัว ที่ต้องการให้อยู่ด้านแถวไว้ในบ็อกซ์ของ Row[s]
- เลือกตัวแปรอย่างน้อย 1 ตัว ที่ต้องการให้อยู่ด้านหลักไว้ในบ็อกซ์ Column[s]
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics , Cell หรือ Format
- เลือกปุ่ม **OK**

6. การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Compare Means

(6.1) การคำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

- **Statistics / Compare Means / Means...**
- เลือกตัวแปรไว้ในบ็อกซ์ของ Dependent List และ Independent List
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(6.2) การทดสอบค่าเฉลี่ย 2 กลุ่มที่เป็นอิสระต่อกัน

- **Statistics / Compare Means / Independent-Samples T Test...**
- เลือกตัวแปรไว้ในบ็อกซ์ของ Test Variable[s] และ Grouping Variables
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(6.3) การทดสอบค่าเฉลี่ย 2 กลุ่มที่มีความสัมพันธ์กัน

- **Statistics / Compare Means / Paired-Samples T Test..**
- เลือกตัวแปรมาไว้ในบ็อกซ์ของ Paired Variables
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(6.4) การทดสอบค่าเฉลี่ยหลายกลุ่มแบบทางเดียว

- **Statistics / Compare Means / One-Way ANOVA...**
- เลือกตัวแปรอย่างน้อย 2 ตัวไว้ในบ๊อกซ์ของ Dependent List และ Factor
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Contrasts , Post Hoc , Options...
- เลือกปุ่ม **OK**

(6.5) การทดสอบค่าเฉลี่ยหลายกลุ่มแบบหลายทาง

- **Statistics / General Linear Model / Simple Factorial...**
- เลือกตัวแปรอย่างน้อย 2 ตัวไว้ในบ๊อกซ์ของ Dependent และ Factor
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่บ๊อกซ์ Covariate[S] หรือ Options
- เลือกปุ่ม **OK**

7. การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Correlate หรือ Regression

(7.1) การหาความสัมพันธ์ของข้อมูลเชิงปริมาณ

- **Statistics / Correlate /** Bivariate...
- เลือกตัวแปรไว้ในบ็อกซ์ของ Variables
- เลือกวิธีการวิเคราะห์ทางสถิติที่จะใช้ในส่วนของ Correlation Coefficients
- เลือกวิธีการทดสอบในส่วนของ Test of Significance
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(7.2) การหาความสัมพันธ์บางส่วนของข้อมูลเชิงปริมาณ

- **Statistics / Correlate / Partial...**
- เลือกตัวแปรไว้ในบ็อกซ์ของ Variables และ Controlling for
- เลือกวิธีการทดสอบในส่วนของ Test of Significance...
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(7.3) การพยากรณ์โดยวิธีวิเคราะห์การถดถอย

- **Statistics / Regression / Linear...**
- เลือกตัวแปรตามไว้ในบ็อกซ์ของ Dependent
- เลือกตัวแปรอิสระอย่างน้อย 1 ตัวไว้ในบ็อกซ์ของ Independent[s]
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม WLS, Statistics , Plot , Save , Options
- เลือกปุ่ม **OK**

(7.4) การเลือกรูปแบบของการพยากรณ์

- **Statistics / Regression / Curve Estimation...**
- เลือกตัวแปรตามไว้ในบ็อกซ์ของ Dependent
- เลือกตัวแปรอิสระไว้ในบ็อกซ์ของ Independent
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม WLS, Statistics , Plot , Save , Options
- เลือกปุ่ม **OK**

8. การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Nonparametric Tests

(8.1) การทดสอบอัตราส่วน

- **Statistics / Nonparametric Tests / Chi-Square...**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ็อกซ์ของ Test Variable List
- อาจจะกำหนดค่าความถี่ใหม่ที่ต้องการ ไว้ในส่วนของ Expected Values
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.2) การทดสอบสัดส่วน

- **Statistics / Nonparametric Tests / Binomial...**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ็อกซ์ของ Test Variables List

- อาจจะกำหนดค่าสัดส่วนใหม่ที่ต้องการไว้ในส่วนของ Test Proportion
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.3) การทดสอบความเป็นตัวอย่างสุ่ม

- **Statistics / Nonparametric Tests / Runs...**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ๊อกซ์ของ Test Variable List
- อาจจะเลือกวิธีการแบ่งกลุ่มข้อมูลเพิ่มอีกในส่วนของ Cut Point
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.4) การทดสอบรูปแบบการแจกแจงของข้อมูล

- **Statistics / Nonparametric Tests / 1-Sample K-S**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ๊อกซ์ของ Test Variable List
- อาจจะเลือกวิธีการแบ่งกลุ่มข้อมูลเพิ่มอีกในส่วนของ Cut Point
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.5) การทดสอบสำหรับข้อมูล 2 กลุ่มที่เป็นอิสระต่อกัน

- **Statistics / Nonparametric Tests / 2 Independents Samples...**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ๊อกซ์ของ Test Variable List
- เลือกตัวแปรที่ต้องการเป็นตัวแบ่งกลุ่มไว้ในบ๊อกซ์ของ Grouping Variable
- เลือกวิธีทางสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.6) การทดสอบสำหรับข้อมูล k กลุ่มที่เป็นอิสระต่อกัน

- **Statistics / Nonparametric Tests / k Independent Samples...**
- เลือกตัวแปรที่ต้องการทดสอบไว้ในบ๊อกซ์ของ Test Variables List
- เลือกตัวแปรที่ต้องการเป็นตัวแบ่งกลุ่มไว้ในบ๊อกซ์ของ Grouping Variable
- เลือกวิธีทางสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.7) การทดสอบสำหรับข้อมูล 2 กลุ่มที่มีความสัมพันธ์

- **Statistics / Nonparametric Tests / 2 Related Samples...**
- เลือกตัวแปรที่ต้องการทดสอบ 2 ตัวไว้ในบ๊อกซ์ของ Test Variable List
- เลือกวิธีทางสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Options...
- เลือกปุ่ม **OK**

(8.8) การทดสอบสำหรับข้อมูล k กลุ่มที่มีความสัมพันธ์

- **Statistics / Nonparametric Tests / k Related Samples...**
- เลือกตัวแปรอย่างน้อย 2 ตัวแปรไว้ในบ๊อกซ์ของ Test Variable List
- เลือกวิธีทางสถิติที่จะใช้ทดสอบในส่วนของ Test Type
- อาจจะกำหนดรายละเอียดเพิ่มเติมที่ปุ่ม Statistics......
- เลือกปุ่ม **OK**

9. การวิเคราะห์ข้อมูลด้วยเมนู Transform

(9.1) การเปลี่ยนค่าตัวแปรไว้ในตัวแปรเดิม

- **Transform / Recode / Recode into Same Variables**
- เลือกตัวแปรที่ต้องการเปลี่ยนค่าไว้ในบ๊อกซ์ของ Variables
- เลือกปุ่ม Old and New Values
  - ◆ กำหนดค่าที่ต้องการเปลี่ยนในบ๊อกซ์ Old Value
  - ◆ กำหนดค่าใหม่ที่จะแทนค่าเดิมในบ๊อกซ์ New Value
- เลือกปุ่ม Continue
- ถ้าต้องเปลี่ยนข้อมูลบางชุดให้เลือกที่ปุ่ม If...
- เลือกปุ่ม **OK**

(9.2) การเปลี่ยนค่าตัวแปรไว้ในตัวแปรใหม่

- **Transform / Recode / Recode into Different Variables**
- เลือกตัวแปรที่ต้องการเปลี่ยนค่าไว้ในบ๊อกซ์ของ Variables
- กำหนดชื่อตัวแปรใหม่ในบ๊อกซ์ของ Output Variable
- อาจจะกำหนดข้อความขยายชื่อตัวแปรไว้ในบ๊อกซ์ของ Label / เลือก Change
- เลือกปุ่ม Old and New Values

- ◆ กำหนดค่าที่ต้องการเปลี่ยนในบ๊อกซ์ Old Value
- ◆ กำหนดค่าใหม่ที่จะแทนค่าเดิมในบ๊อกซ์ New Value
- เลือกปุ่ม Continue
- ถ้าต้องเปลี่ยนข้อมูลบางชุดให้เลือกที่ปุ่ม If...
- เลือกปุ่ม **OK**

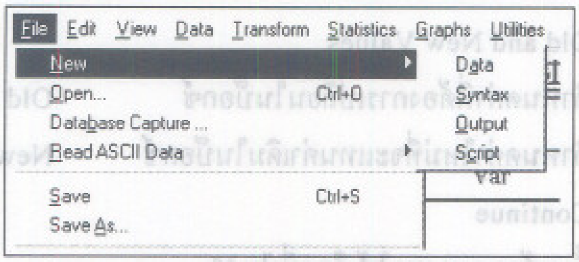
(9.3) การสร้างตัวแปรใหม่จากการคำนวณและเงื่อนไข

- **Transform / Compute**
- กำหนดชื่อตัวแปรใหม่ในบ๊อกซ์ของ Target Variable
- กำหนดนิพจน์ทางคณิตศาสตร์ไว้ในบ๊อกซ์ของ Numeric Expression
- ถ้าต้องการสร้างตัวแปรใหม่แบบมีเงื่อนไขให้เลือกที่ปุ่ม IF...
- เลือกปุ่ม **OK**

10. การเปิดวินโดวส์หลายแบบพร้อมกัน

เมื่อเริ่มใช้โปรแกรม SPSS จะปรากฏวินโดวส์ SPSS Data Editor เมื่อทำการวิเคราะห์ข้อมูลก็จะเกิด Window Output Navigator ถ้าต้องการเปิดวินโดวส์เพิ่มเติม สามารถทำได้ดังนี้

- เปิดเมนู File และเลือกรายการ New จะมีชนิดของ Window ให้เลือก 4 ชนิดคือ



⊙ **Data** สำหรับวินโดวส์ **Data Editor** ⊙ **Syntax** สำหรับวินโดวส์ **Syntax**

⊙ **Output** สำหรับวินโดวส์ **Output Navigator** ⊙ **Script** สำหรับการเขียน **script**

- คลิกที่ชนิดที่ต้องการ

11. การบันทึกข้อมูลใน Window ต่างๆ

ผู้ใช้สามารถบันทึกข้อมูลในวินโดวส์ที่ถูกเปิดขึ้นมาใช้งาน โดยบันทึกไว้ในรูปของไฟล์ ซึ่งโปรแกรม SPSS ได้จัดแบ่งประเภทของไฟล์ดังนี้

ชนิดของวินโดวส์	ส่วนขยายของไฟล์
Data Editor Window	*.sav
Output Window	*.spo
Syntax Window	*.sps
Chart Window	*.sct

การบันทึกข้อมูลที่อยู่ในวินโดวส์ใดๆ สามารถดำเนินการได้ดังนี้

● เลือกวินโดวส์ที่จะบันทึกข้อมูลโดยการใช้เมาส์คลิก บริเวณใดๆ ในวินโดวส์ที่ต้องการจะปรากฏแถบแสงที่ชื่อวินโดวส์นั้น เปิดเมนู File และเลือกรายการใดรายการหนึ่ง

⊙ Save ชื่อชนิดของวินโดวส์สำหรับบันทึกภายใต้ชื่อไฟล์เดิมที่เคยบันทึกไว้แล้ว

⊙ Save as สำหรับการบันทึกภายใต้ชื่อไฟล์ใหม่

● กำหนดชื่อ ตำแหน่งไดรฟ์ ประเภทของไฟล์

● เลือกปุ่ม **Save**

12. การเรียกไฟล์ข้อมูล

ไฟล์ข้อมูลของวินโดวส์ที่ถูกบันทึกไว้แล้วเมื่อต้องการนำมาใช้ต้องทำดังนี้

● เปิดเมนู File และเลือกรายการ Open จะปรากฏรายการให้เลือกตามชนิดของวินโดวส์ต่างๆ ในความหมายของต่อไปนี้

⊙ Data สำหรับไฟล์ข้อมูลของ SPSS Data Editor Window

⊙ Chart สำหรับไฟล์ข้อมูลของ Chart Window

⊙ SPSS Syntax สำหรับไฟล์ข้อมูลของ SPSS Syntax Windows

⊙ SPSS Output สำหรับไฟล์ข้อมูลของ SPSS Output Windows

● พิมพ์ชื่อที่ต้องการ แล้วกด **Open**

13. การบันทึกข้อมูล

เป็นการบันทึกข้อมูลในวินโดวส์

● เลือกชนิดวินโดวส์ **Data , Output , Syntax , ...** ที่ต้องการบันทึกข้อมูล

● เลือกเมนู **File**

● เลือกรายการ **Save** หรือ **Save as**

● กำหนดชื่อไฟล์ และตำแหน่งที่จะบันทึกตามความต้องการ

● เลือกปุ่ม **Save**

- **สรุปความสามารถในการวิเคราะห์ข้อมูลโดยย่อของ SPSS for Windows**

- **การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Summarize**
  - การแจกแจงความถี่แบบทางเดียว
  - การคำนวณค่าสถิติเบื้องต้น
  - การแจกแจงความถี่ตั้งแต่ 2 ทาง
- **การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Compare Means**
  - การคำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม
  - การทดสอบค่าเฉลี่ย 2 กลุ่มที่เป็นอิสระต่อกัน
  - การทดสอบค่าเฉลี่ย 2 กลุ่มที่มีความสัมพันธ์กัน
  - การทดสอบค่าเฉลี่ยหลายกลุ่มแบบทางเดียว
  - การทดสอบค่าเฉลี่ยหลายกลุ่มแบบหลายทาง
- **การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Correlate หรือ Regression**
  - การหาความสัมพันธ์ของข้อมูลเชิงปริมาณ
  - การหาความสัมพันธ์บางส่วนของข้อมูลเชิงปริมาณ
  - การพยากรณ์โดยวิธีวิเคราะห์การถดถอย
  - การเลือกรูปแบบของการพยากรณ์
- **การวิเคราะห์ข้อมูลด้วยเมนู Statistics / Nonparametric Tests**
  - การทดสอบอัตราส่วน
  - การทดสอบสัดส่วน
  - การทดสอบความเป็นตัวอย่างสุ่ม
  - การทดสอบรูปแบบการแจกแจงของข้อมูล
  - การทดสอบสำหรับข้อมูล 2 กลุ่มที่เป็นอิสระต่อกัน
  - การทดสอบสำหรับข้อมูล k กลุ่มที่เป็นอิสระต่อกัน
  - การทดสอบสำหรับข้อมูล 2 กลุ่มที่มีความสัมพันธ์
  - การทดสอบสำหรับข้อมูล k กลุ่มที่มีความสัมพันธ์

บทที่ 5

การสร้างแฟ้มข้อมูลและการคำนวณค่าสถิติเบื้องต้น

ในบทนี้จะศึกษาเกี่ยวกับการสร้างแฟ้มข้อมูล การเรียกใช้ การเปลี่ยนแปลงแก้ไข และการวิเคราะห์ค่าสถิติเบื้องต้น ดังต่อไปนี้

♠ Statistics / Summarize / Frequency.. การแจกแจงความถี่ 1 ทาง และ การวิเคราะห์ค่าสถิติ

♥ Statistics / Summarize / Descriptives.. การวิเคราะห์ค่าสถิติเบื้องต้น

♦ Statistics / Summarize / Explore.. การวิเคราะห์ค่าสถิติเบื้องต้น

♣ Statistics / Summarize / Crosstabs.. การแจกแจงความถี่ 2 ทาง และ การวิเคราะห์ค่าสถิติ

5.1 การสร้างแฟ้มข้อมูลใน SPSS Data Editor

การสร้างแฟ้มข้อมูลใน SPSS Data Editor สรุปเป็นขั้นตอน โดยย่อเป็นดังนี้

- เลือก คอลัมน์ที่ต้องการ กำหนดหรือเปลี่ยนชื่อตัวแปร
- **Data / Define Variable**
- กำหนดชื่อตัวแปรในบ็อกซ์ของ **Variable name**

ถ้าต้องการเปลี่ยนแปลงรายละเอียดของตัวแปรให้แตกต่างจากที่โปรแกรมกำหนดให้ สามารถเปลี่ยนแปลงได้ 4 ประเภท ดังนี้

1.1 กำหนดประเภทของตัวแปร เลือกปุ่ม **Type...**

1.2 กำหนดข้อความขยายชื่อและอธิบายค่าตัวแปร เลือกปุ่ม **Labels...**

1.3 กำหนดค่าที่ขาดหายไปหรือค่าไม่สมบูรณ์ เลือกปุ่ม **Missing Values**

1.4 กำหนดความกว้าง/จัดข้อความของคอลัมน์ เลือกปุ่ม **Column Format**

การป้อนข้อมูล 2.1 ใช้แป้น **Enter** สำหรับการป้อนข้อมูลครั้งละ 1 ตัวแปร

2.2 ใช้แป้น ↓ → ← ↑ สำหรับการป้อนข้อมูลแล้วเลื่อนไปเซลล์ถัดไป

2.3 ใช้แป้น **Tab** สำหรับการป้อนข้อมูลครั้งละ 1 ชุด (แถว)

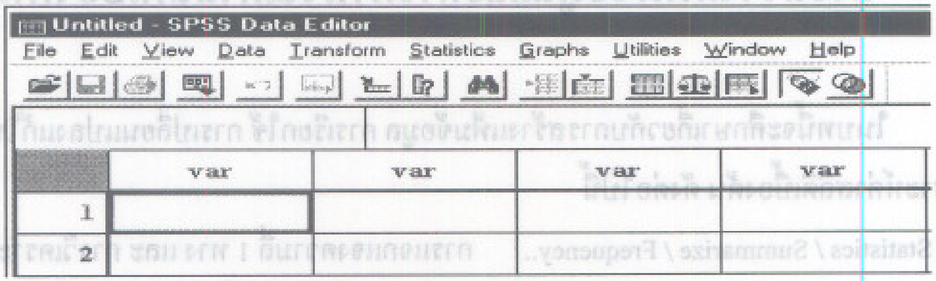
ตัวอย่าง 5.1.1 จงสร้างแฟ้มข้อมูล (กำหนดตัวแปรชื่อ score) ที่ประกอบด้วย

3, 3, 6, 4, 5, 1, 2, 3, 8, 4, 5, 6

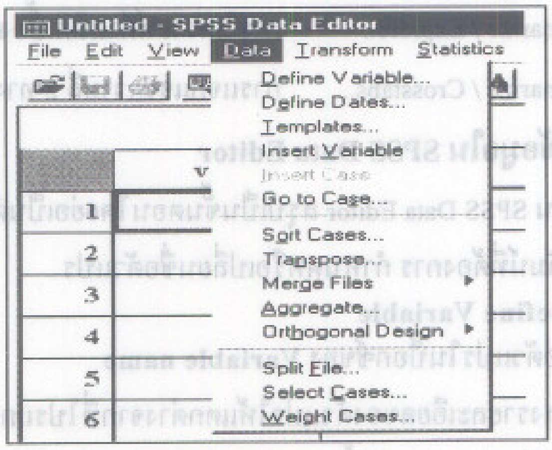
และทำการหาค่าเฉลี่ย มัธยฐาน ฐานนิยม ส่วนเบี่ยงเบนมาตรฐาน ค่าความแปรปรวน

วิธีทำ การสร้างแฟ้มข้อมูล

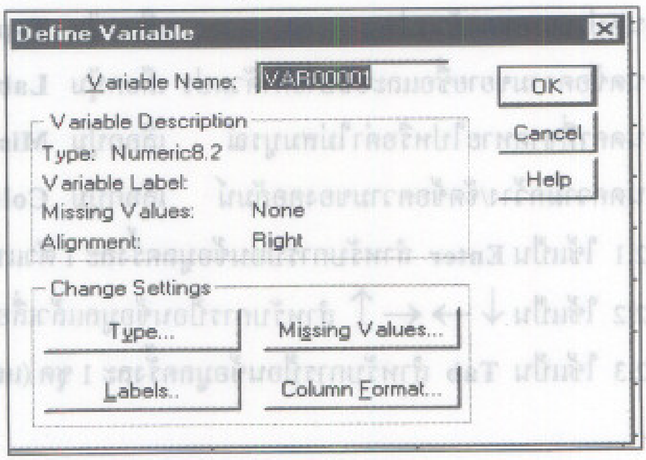
ขั้นที่ 1. ขอให้เริ่มต้นจาก SPSS Data Editor



ขั้นที่ 2. คลิกที่ Data จะได้เมนูย่อย



ขั้นที่ 3. คลิกที่ Define Variables จะได้เมนูย่อย



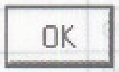
ถ้าต้องการเปลี่ยนแปลงรายละเอียดของตัวแปรให้แตกต่างจากที่โปรแกรมกำหนดให้ สามารถเปลี่ยนแปลงได้ 4 ประเภท ดังนี้

1.1 กำหนดประเภทของตัวแปร เลือกปุ่ม **Type...**

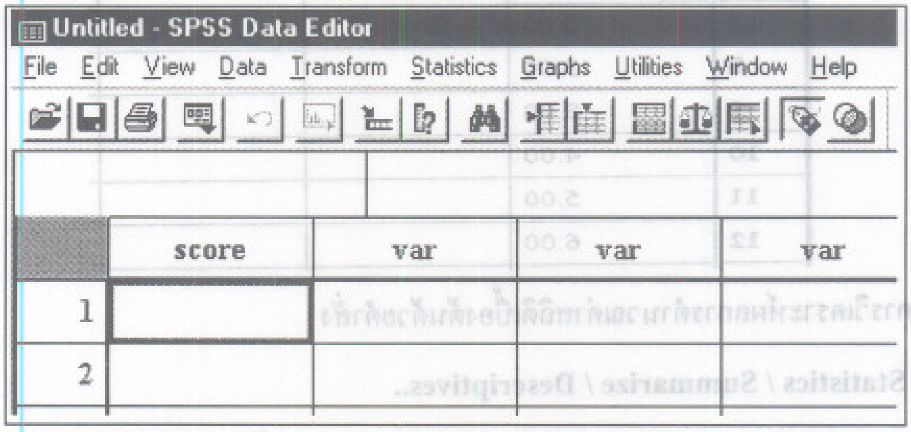
1.2 กำหนดข้อความขยายชื่อและอธิบายค่าตัวแปร เลือกปุ่ม **Labels...**

1.3 กำหนดค่าที่ขาดหายไปหรือค่าไม่สมบูรณ์ เลือกปุ่ม **Missing Values**

1.4 กำหนดความกว้าง/จัดข้อความของคอลัมน์ เลือกปุ่ม **Column Format**

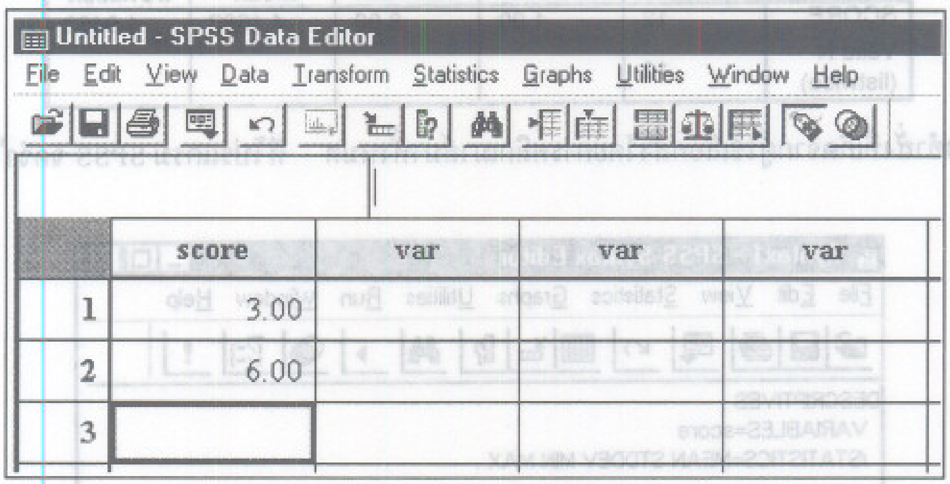
ขั้นที่ 4. พิมพ์ชื่อตัวแปร score เสร็จแล้วคลิก 

จะได้ column ที่ 1 เป็นช่องตัวแปร score



ขั้นที่ 5. พิมพ์ข้อมูลที่ละตัวดังนี้

3↓ 6↓



ใส่ข้อมูลไปจนครบทุกตัวจะได้ผลบนจอภาพดังนี้

1:score 3

	score	var	var
1	3.00		
2	3.00		
3	6.00		
4	4.00		
5	5.00		
6	1.00		
7	2.00		
8	3.00		
9	8.00		
10	4.00		
11	5.00		
12	6.00		

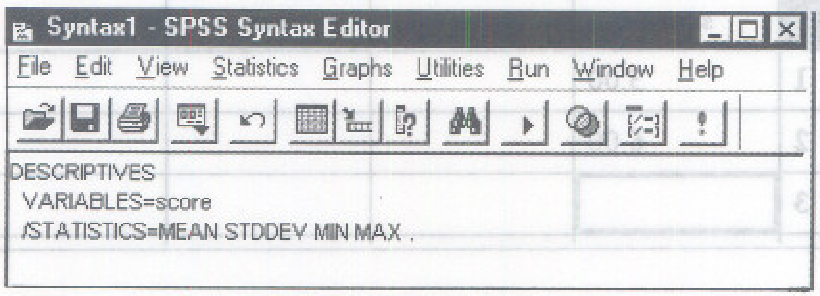
ตัวอย่างการวิเคราะห์ผลการคำนวณค่าสถิติเบื้องต้นด้วยคำสั่ง

Statistics / Summarize / Descriptives..

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
SCORE	12	1.00	8.00	4.1667	1.9462
Valid N (listwise)	12				

ชุดของคำสั่งที่เกิดจากผู้ใช้เลือกสั่งโดยการคลิกเมาส์มาทั้งหมด ที่โปรแกรม SPSS จดจำไว้ให้



ตัวอย่าง 5.1.2 การสร้างแฟ้มข้อมูล 2 ตัวแปรเช่นข้อมูลคะแนนสอบย่อย 2 ครั้งของนิสิต 10 คน

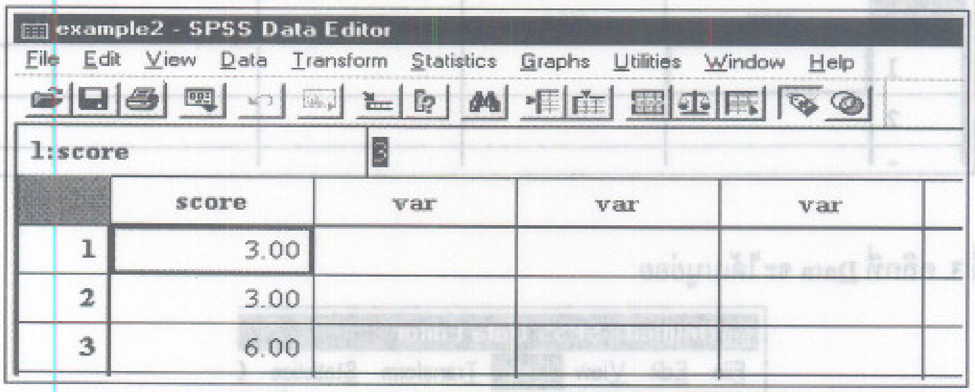
ครั้งที่ 1.	ครั้งที่ 2
76	81
60	52
85	87
58	70
91	86
75	77
82	90
64	63
79	85
88	83

และทำการบันทึกโดยใช้ชื่อแฟ้ม example2.sav

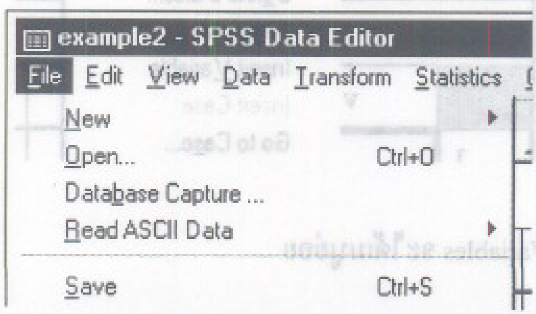
วิธีทำ การสร้างแฟ้มข้อมูล

ขั้นที่ 1. ขอให้เริ่มต้นจาก SPSS Data Editor

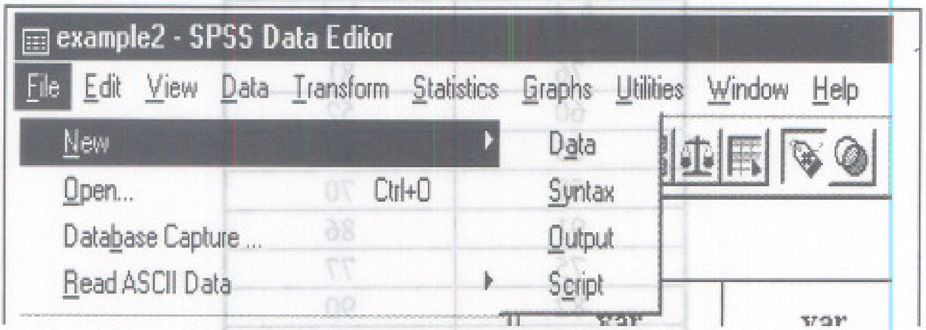
หมายเหตุ ในตัวอย่างนี้ขอสมมติว่า ใน SPSS Data Editor เดิมมีข้อมูลเก่าค้างอยู่ การทำข้อมูลใหม่ต้องใช้คำสั่ง **File / New / Data..**



ขั้นที่ 2. คลิกที่ File จะได้เมนูย่อย

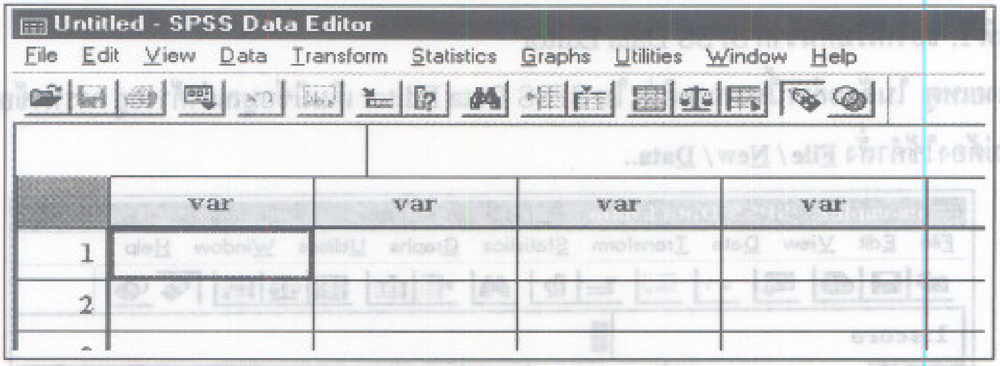


เลื่อนเมาส์มาที่ **New** จะได้เมนูย่อย

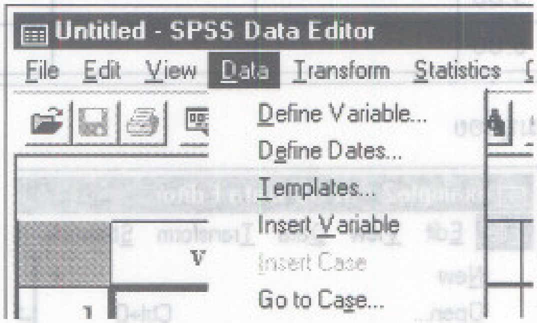


ต่อไปคลิกที่ **Data** คอมพิวเตอร์จะทำการลบข้อมูลขณะนั้นออกจากจอภาพ หรือ ถามเราว่าต้องการบันทึกข้อมูลหรือไม่

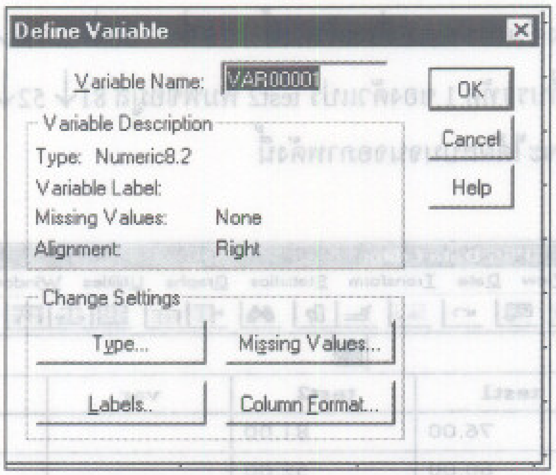
ต่อไปขอเริ่มต้นที่ SPSS Data Editor ว่างๆ ดังนี้

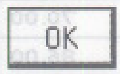


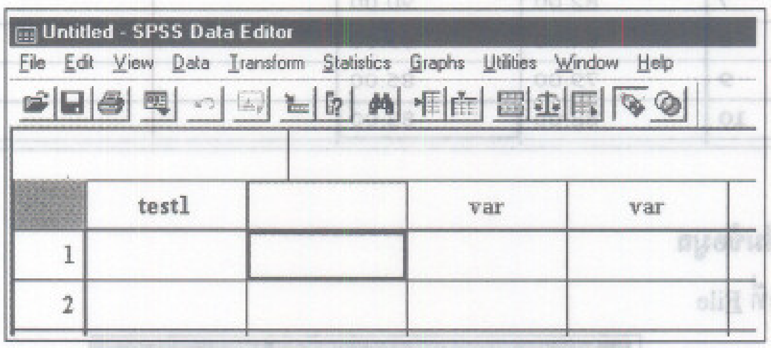
ขั้นที่ 3. คลิกที่ **Data** จะได้เมนูย่อย



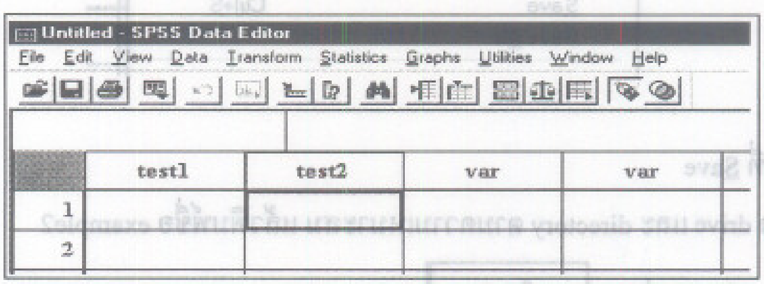
ขั้นที่ 4. คลิกที่ **Define Variables** จะได้เมนูย่อย



ขั้นที่ 5. พิมพ์ชื่อตัวแปร test1 เสร็จแล้วคลิก  จะได้ column ที่ 1 เป็นช่องตัวแปร test1

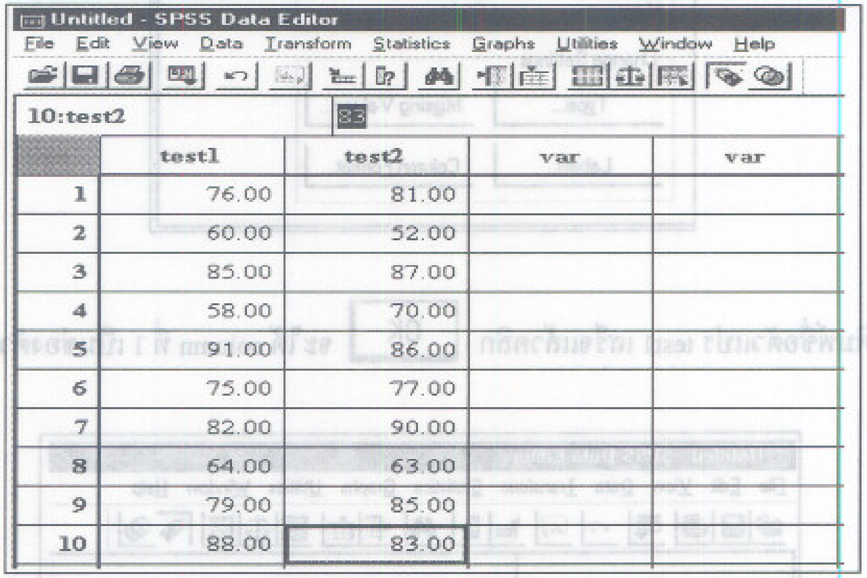


เลื่อนเมาส์ไปที่ Column 2 ทำแบบเดียวกันโดยใช้คำสั่ง Data \ Define Variables ..
และใช้ชื่อตัวแปร test2 จะได้ผลบนจอภาพเป็น



ขั้นที่ 5. พิมพ์ข้อมูลของตัวแปร test1 ที่ละตัวดังนี้ 76↓ 60↓ ... 79↓ 88↵

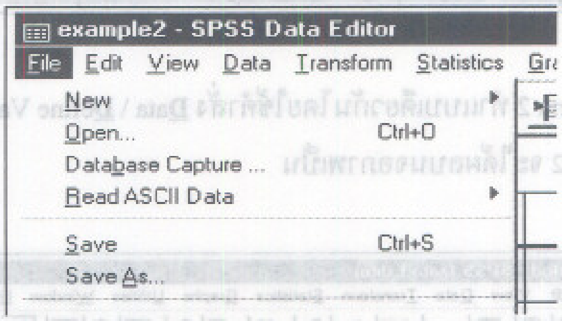
เลื่อนเมาส์มาที่บรรทัด 1 ของตัวแปร test2 พิมพ์ข้อมูล 81↓ 52↓ ... 85↓ 83↵ ใส่ข้อไปจนครบทุกตัวจะได้ผลบนจอภาพดังนี้



	test1	test2	var	var
1	76.00	81.00		
2	60.00	52.00		
3	85.00	87.00		
4	58.00	70.00		
5	91.00	86.00		
6	75.00	77.00		
7	82.00	90.00		
8	64.00	63.00		
9	79.00	85.00		
10	88.00	83.00		

การบันทึกแฟ้มข้อมูล

ขั้นที่ 1. คลิกที่ File



ขั้นที่ 2. คลิกที่ Save

ขั้นที่ 3. เลือก drive และ directory ตามความเหมาะสม แล้วพิมพ์ชื่อ example2

เสร็จแล้วคลิกที่ปุ่ม Save

5.2 การเปิดแฟ้มข้อมูลและการแสดงผลของข้อมูล

แฟ้มข้อมูลชื่อ example4.sav บันทึกไว้ในแผ่น diskette ที่ drive A: ประกอบด้วยตัวแปร 8 ตัว และมีค่าสังเกต 50 ค่า

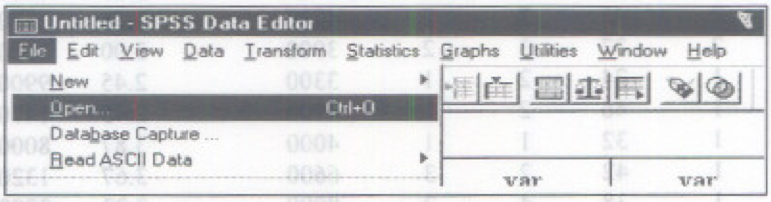
id	sex	age	educ	status	income	grade	bonus
1	1	37	2	4	5500	3.78	11000.00
2	2	29	3	1	4100	3.89	12300.00
3	2	48	1	2	5400	3.67	21600.00
4	1	99	1	2	9999	2.78	19998.00
5	2	33	2	9	9999	3.00	29997.00
6	2	45	3	4	8300	3.45	16600.00
7	2	38	1	4	7700	3.89	7700.00
8	2	23	3	1	3900	3.67	11700.00
9	1	34	2	4	4500	2.56	9000.00
10	1	50	2	2	6700	2.69	6700.00
11	2	43	2	2	4700	3.56	18800.00
12	2	37	3	2	3900	3.00	3900.00
13	1	24	2	1	3300	2.45	9900.00
14	1	46	2	2	4900	2.45	14700.00
15	1	32	1	1	4000	3.87	8000.00
16	1	42	2	3	6600	3.67	13200.00
17	1	38	4	2	8000	3.23	32000.00
18	2	41	2	3	7000	3.45	21000.00
19	2	99	1	9	2000	3.21	2000.00
20	1	54	2	2	7400	3.00	22200.00
21	2	32	3	9	6200	2.56	24800.00
22	1	43	1	2	4700	2.45	18800.00
23	2	22	1	1	3400	3.78	3400.00
24	1	40	2	2	5900	2.67	17700.00
25	1	37	4	9	7500	3.45	22500.00
26	1	28	1	1	3100	2.78	9300.00
27	1	44	3	2	6800	2.56	13600.00
28	1	56	2	2	6400	2.78	19200.00
29	1	35	3	1	5800	3.33	5800.00
30	2	42	1	2	3900	2.56	11700.00
31	1	21	2	1	4700	2.67	14100.00
32	1	39	2	2	5900	2.89	17700.00
33	1	45	1	2	4900	2.56	4900.00
34	1	31	1	2	3100	3.23	9300.00
35	1	51	2	3	5400	3.01	5400.00
36	1	23	3	1	6300	2.77	12600.00
37	1	40	3	2	7100	2.89	21300.00
38	1	47	2	3	6600	2.77	19800.00
36	1	53	2	2	7200	2.31	21600.00
40	2	27	2	1	1700	2.67	5100.00
41	1	29	4	1	5000	2.89	15000.00
42	1	40	3	2	6000	3.67	18000.00
43	2	30	1	1	3000	2.56	12000.00
44	2	53	2	2	4700	3.00	9400.00

45	1	31	1	1	2800	2.74	5600.00
46	1	45	2	2	5700	2.67	22800.00
47	1	22	2	4	4300	3.07	4300.00
48	2	34	1	1	3900	2.56	7800.00
49	2	33	3	2	6700	2.12	20100.00
50	1	54	2	2	4800	2.66	19200.00

5.2.1 ขั้นตอนการเปิดแฟ้มข้อมูล

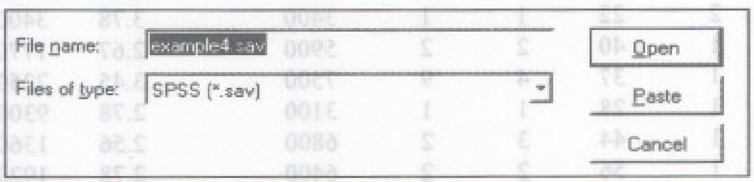
ขั้นที่ 1. นำแผ่น disk ที่มีข้อมูลใส่ drive A:

ขั้นที่ 2. เลือกคำสั่ง File

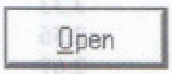


คลิกที่คำสั่ง Open

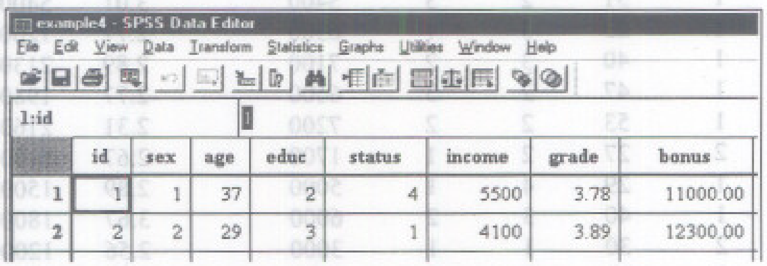
ขั้นที่ 3. พิมพ์ชื่อแฟ้มข้อมูล example4



ขั้นที่ 4. คลิกที่ปุ่ม

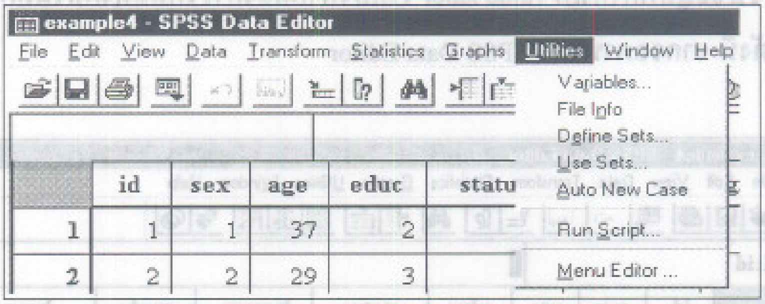


จะได้ข้อมูลบนจอภาพดังนี้

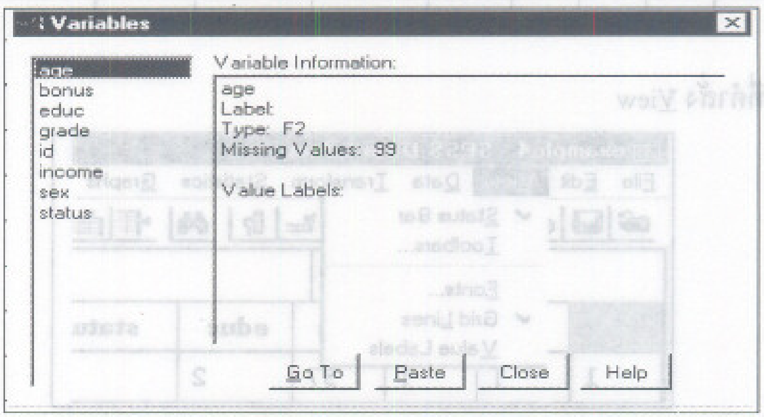


5.2.2 การดูรายละเอียดของตัวแปร

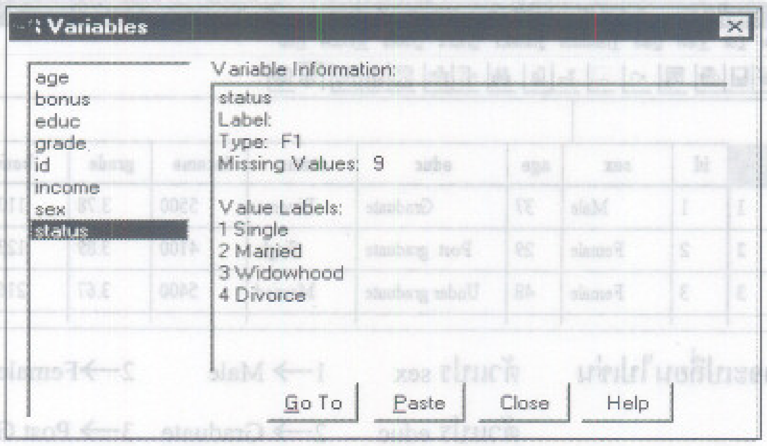
ขั้นที่ 1. คลิกที่คำสั่ง Utilities..



ขั้นที่ 2. คลิกที่คำสั่ง Variables.. จะได้เมนูย่อยดังนี้

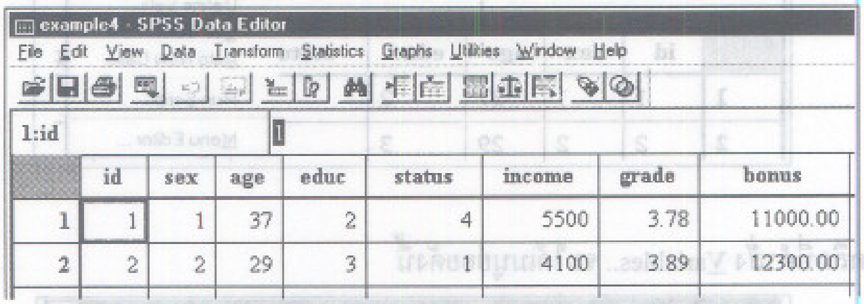


ต้องการดูรายละเอียดของตัวแปรใดให้คลิกที่ชื่อของตัวแปรที่ต้องการ เช่น



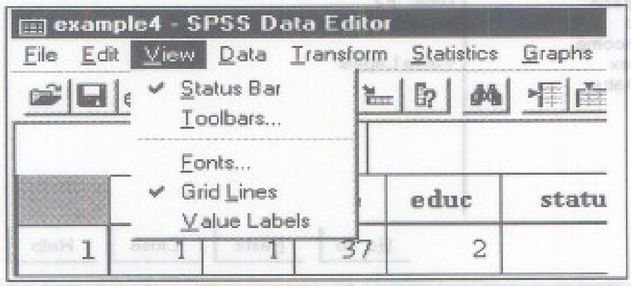
5.2.3 การสั่งให้ SPSS Data Editor แสดง value label ที่กำหนด ไว้

สำหรับข้อมูลที่กำหนด value label ไว้แล้วหากต้องการให้แสดงผลในลักษณะของ value label ต้องทำดังนี้ จากจอภาพของ SPSS Data Editor

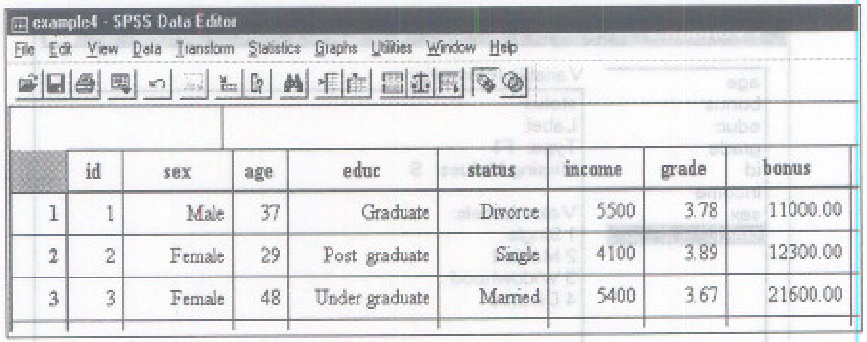


	id	sex	age	educ	status	income	grade	bonus
1	1	1	37	2	4	5500	3.78	11000.00
2	2	2	29	3	1	4100	3.89	12300.00

ขั้นที่ 1. คลิกที่คำสั่ง View



ขั้นที่ 2. คลิกที่คำสั่ง Value Labels



	id	sex	age	educ	status	income	grade	bonus
1	1	Male	37	Graduate	Divorce	5500	3.78	11000.00
2	2	Female	29	Post graduate	Single	4100	3.89	12300.00
3	3	Female	48	Under graduate	Married	5400	3.67	21600.00

การแสดงผลจะเปลี่ยนไปเช่น ตัวแปร sex 1→ Male 2→Female

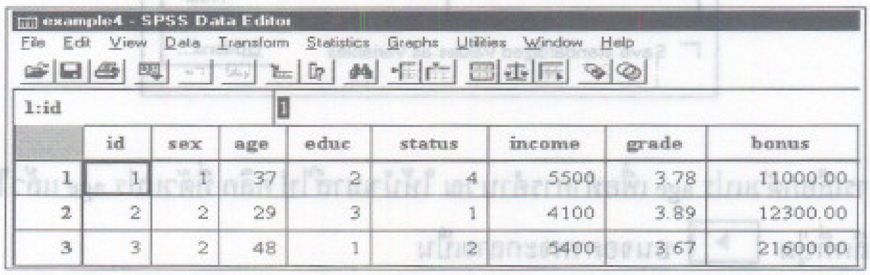
ตัวแปร educ 2→ Graduate 3→ Post Graduate

ตามที่กำหนดไว้ตอนที่สร้างแฟ้มข้อมูล

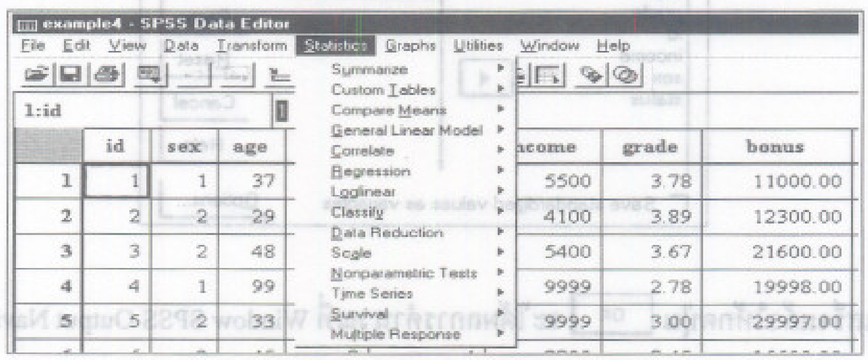
5.3 การใช้คำสั่ง Statistics / Summarize / Descriptives ..

คำสั่ง Statistics / Summarize / Descriptives .. เป็นคำสั่งใช้ในการหาค่าสถิติเบื้องต้น เช่น ค่าเฉลี่ยเลขคณิต ค่าฐานนิยม ค่ามัธยฐาน ค่าส่วนเบี่ยงเบนมาตรฐาน ค่าสูงสุด ค่าต่ำสุด การใช้คำสั่งนี้ได้ต้องมีข้อมูลใน SPSS Data Editor ก่อน

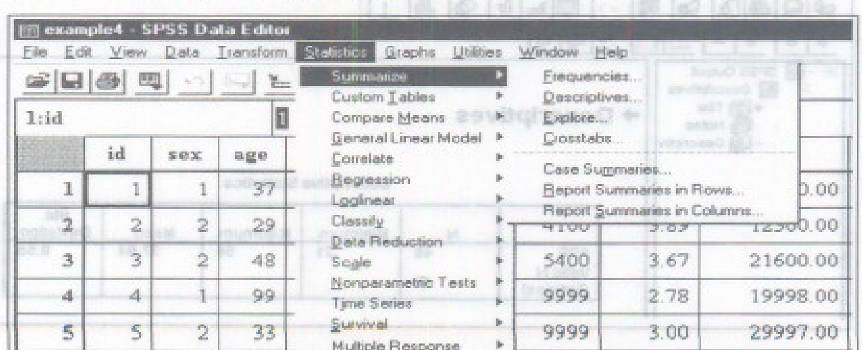
ขั้นที่ 1. เปิดแฟ้มข้อมูล example4 เข้ามาใน SPSS Data Editor



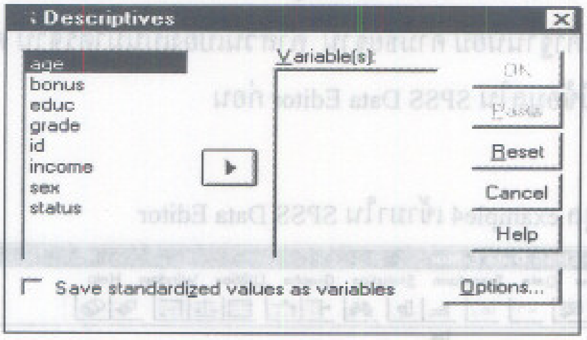
ขั้นที่ 2. เลื่อนเมาส์ไปคลิกที่ Statistics จอภาพจะเป็นดังนี้



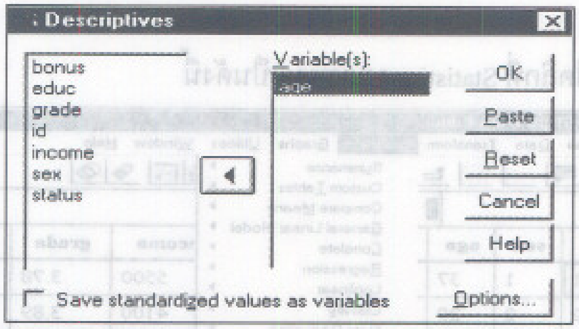
ขั้นที่ 3. เลื่อนเมาส์ไปชี้ที่ Summarize จอภาพจะเปลี่ยนแปลงเป็นดังนี้



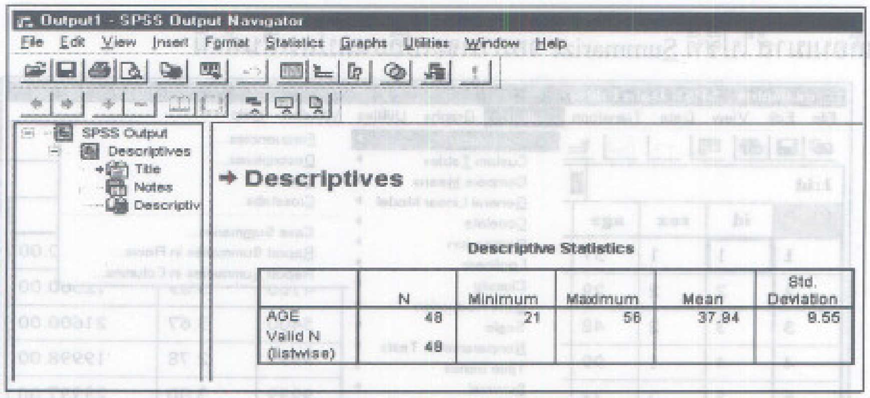
ขั้นที่ 4.. ต่อไปให้นำเมาส์ไปคลิกที่ Descriptive... จะได้เมนูย่อยเป็นดังนี้



ขั้นที่ 5..การเลือกตัวแปร age เพื่อทำการคำนวณ ให้นำเมาส์ไป คลิก ที่ตัวแปร age แล้วไปจึงนำเมาส์ไปคลิกที่ปุ่ม ▶ บนจอภาพจะกลายเป็น



ขั้นที่ 6.. เสร็จแล้วให้กดปุ่ม OK จะได้ผลการคำนวณที่ Window SPSS Output Navigator ดังนี้

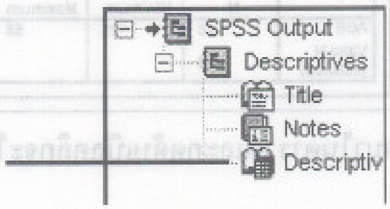


	N	Minimum	Maximum	Mean	Std. Deviation
AGE	48	21	56	37.94	9.55
Valid N (listwise)	48				

5.3.1 การเปลี่ยนรูปแบบของตารางในการแสดงผลของ Output Navigator

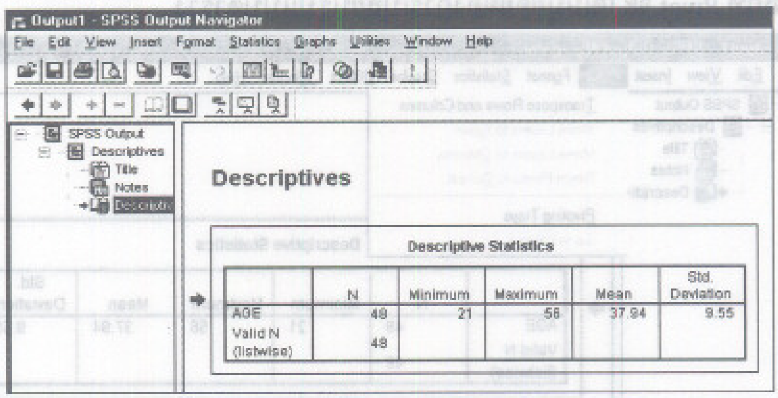
การแสดงผลของการคำนวณหาก ตารางแนวนอนมีความยาวมากจะทำให้เราไม่มีความสะดวกที่จะเห็นผลการคำนวณทั้งหมดในหน้าจอ ดังนั้นเราควรจะทำการ Transpose ให้ตารางแสดงผลในแนวตั้ง

ขั้นที่ 1. จากจอภาพใน SPSS Output Navigator ให้เลือกตารางที่ต้องการ ในที่นี้ขอให้เลือกตาราง

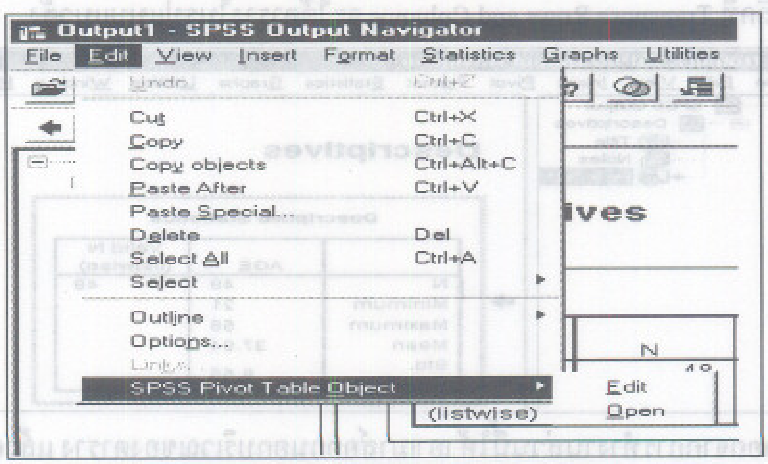


โดยการคลิกที่ชื่อของตาราง

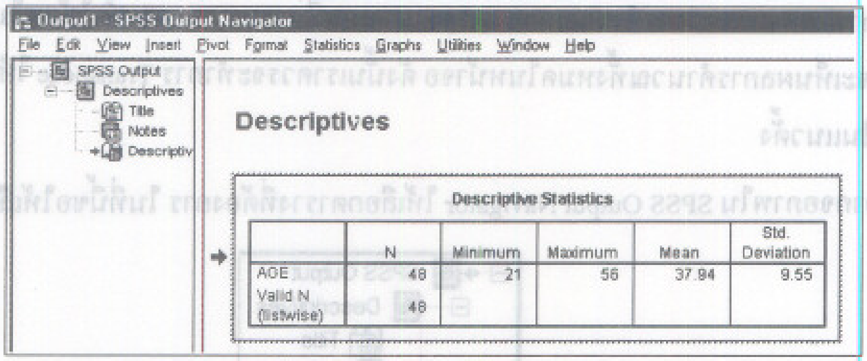
จะเห็นว่าบนจอภาพจะมีลูกศรสีแดงขึ้นที่ตาราง



ขั้นที่ 2. คลิกที่เมนู Edit และเลื่อนเมาส์ไปที่ SPSS Pivot Tables Object จะได้เมนูย่อย Edit หรือ Open

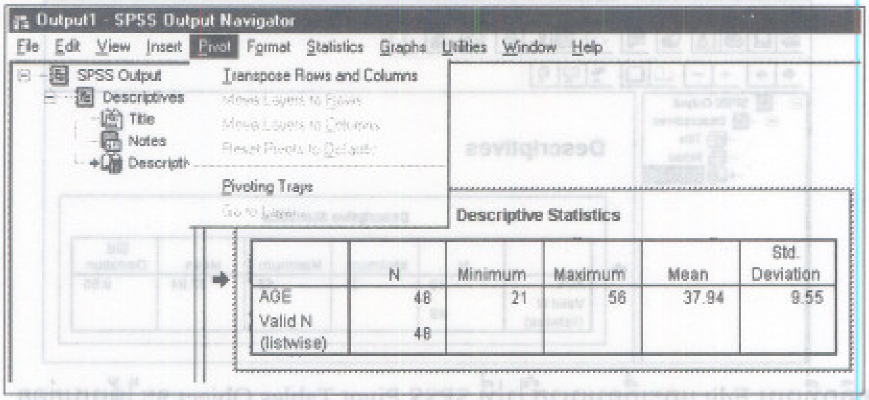


ขั้นที่ 3. ให้คลิกที่ Edit จอภาพจะเปลี่ยนแปลงโดยที่เมนูบาร์จะเป็นเมนูของการแก้ไขตาราง

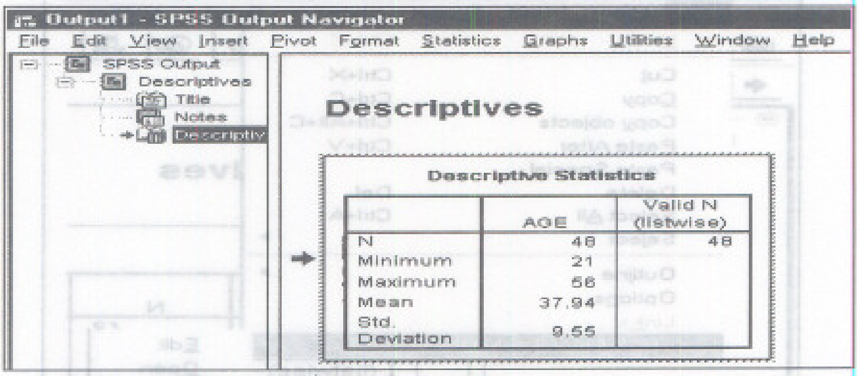


หมายเหตุ การเลื่อน pointer เข้ามาในตารางและกดดับเบิลคลิกจะให้ผลการทำงานเหมือนกับการทำตาม ขั้นที่ 1 - 3

ขั้นที่ 4. ให้คลิกที่ Pivot จะได้เมนูย่อยของการเปลี่ยนรูปแบบตาราง



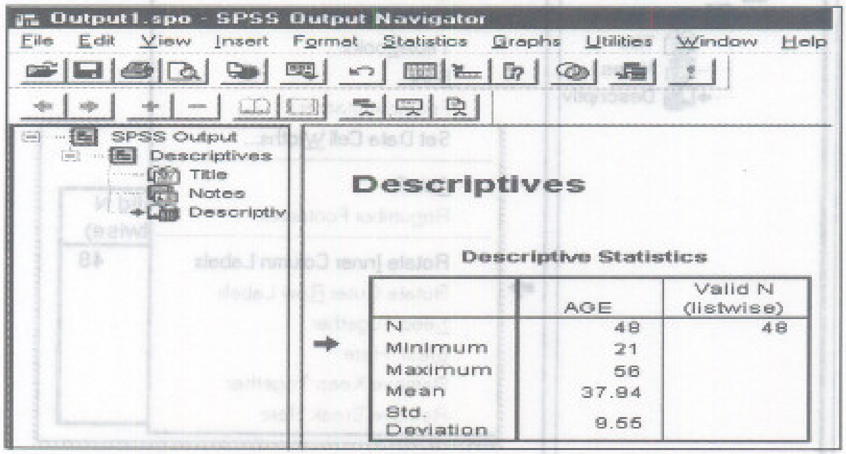
ขั้นที่ 5. ให้คลิกที่ Transpose Rows and Columns จะได้ตารางในรูปแบบแนวตั้ง



ขั้นที่ 6. การออกจากการทำงานส่วนนี้ให้ เอาเมาส์ออกนอกบริเวณของตาราง แล้วกดคลิกหนึ่งที

5.3.2 การกำหนดตำแหน่งทศนิยมของการคำนวณในตารางของ Output Navigator

ตารางผลการคำนวณที่ได้เราสามารถกำหนดการแสดงผลว่าต้องการให้แสดงผลการคำนวณเป็นทศนิยม k ตำแหน่งได้ตามความต้องการ สมมติว่าเราต้องการให้แสดงผลการคำนวณของ Mean ให้เป็นทศนิยม 6 ตำแหน่ง มีขั้นตอนการใช้คำสั่งดังนี้ จากจอภาพ



ขั้นที่ 1. จากจอภาพใน SPSS Output Navigator ให้เลือกตารางที่ต้องการเลื่อนเมาส์เข้าในบริเวณของตาราง แล้วกด ดับเบิลคลิกจะได้ผลเป็นดังนี้

Descriptive Statistics

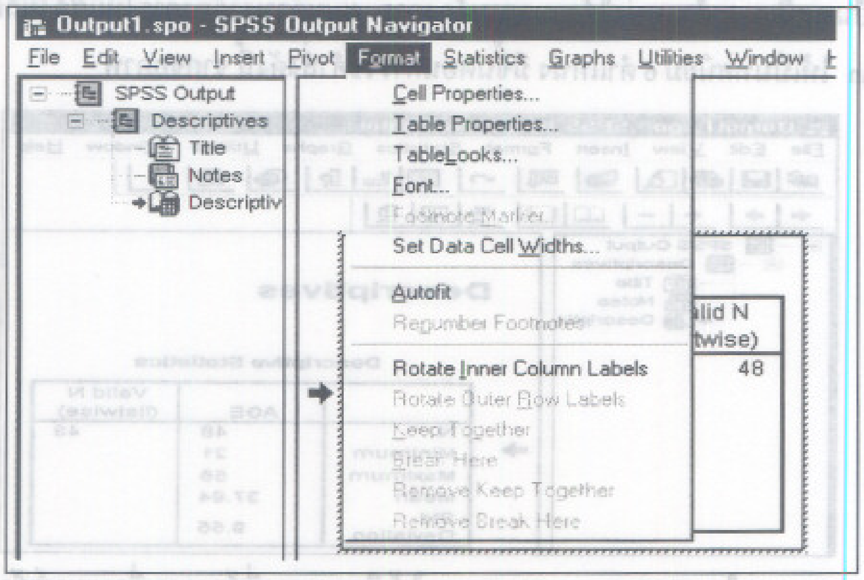
	AGE	Valid N (listwise)
N	48	48
Minimum	21	
Maximum	56	
Mean	37.94	
Std. Deviation	9.55	

ขั้นที่ 2. เอา Pointer ไปคลิกที่ 37.94 จะมีการเปลี่ยนแปลงดังนี้

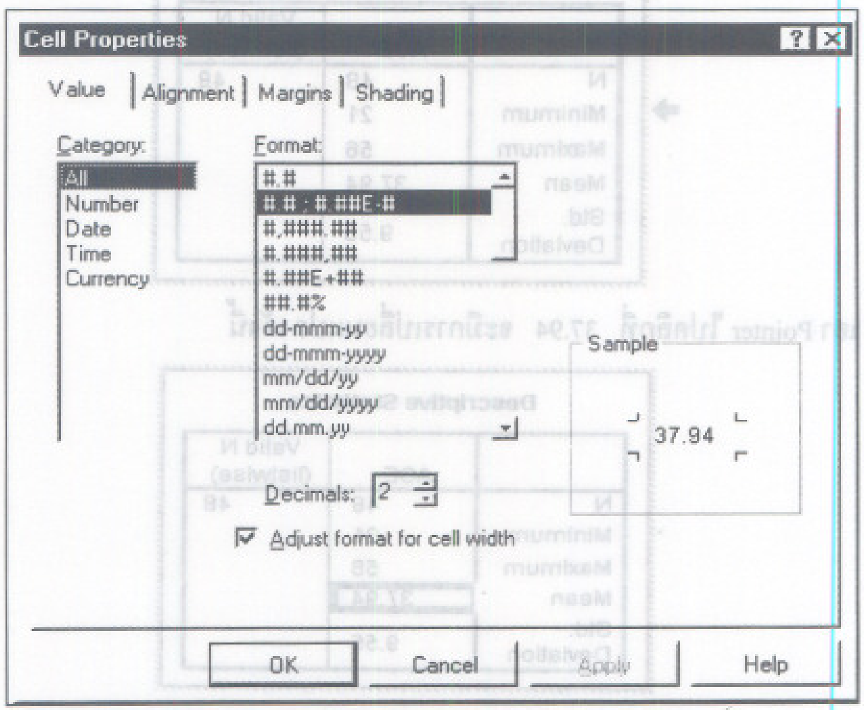
Descriptive Statistics

	AGE	Valid N (listwise)
N	48	48
Minimum	21	
Maximum	56	
Mean	37.94	
Std. Deviation	9.55	

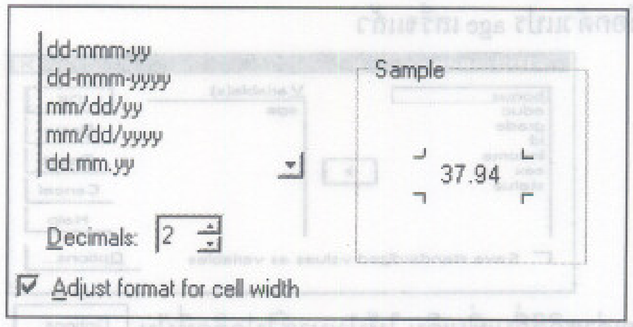
ขั้นที่ 3. คลิกที่เมนู Format จะได้เมนูย่อยเป็นดังนี้



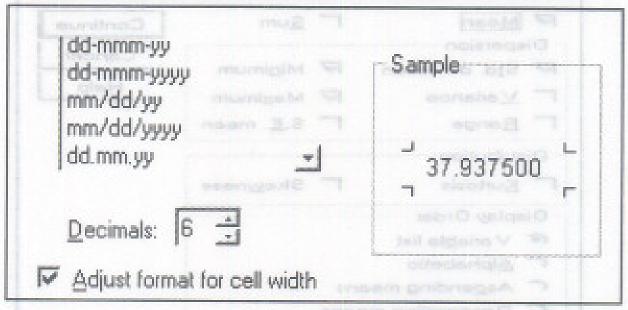
ขั้นที่ 4. เลือกคำสั่ง Cell Properties.. จะได้เมนูย่อยเป็นดังนี้



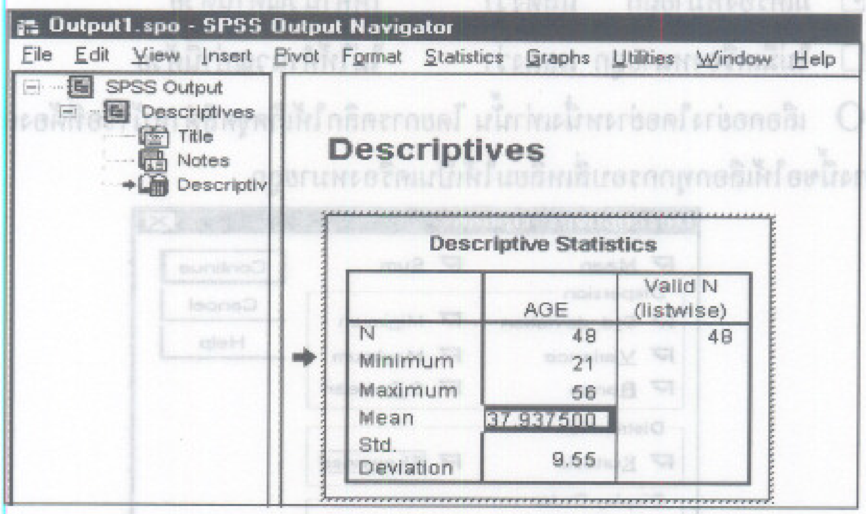
ขั้นที่ 5. ขอให้สังเกตที่



ทำการเพิ่มตำแหน่งทศนิยมโดยการลบเลข 2 ออกแล้วพิมพ์ของใหม่เป็นเลข 6



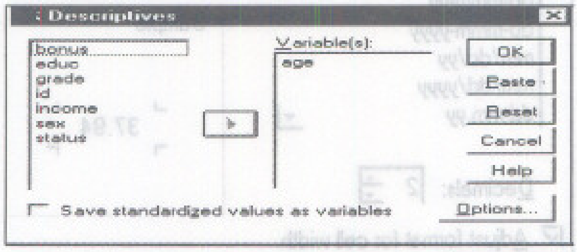
เสร็จแล้วกด OK จะได้ผลตามที่ต้องการ



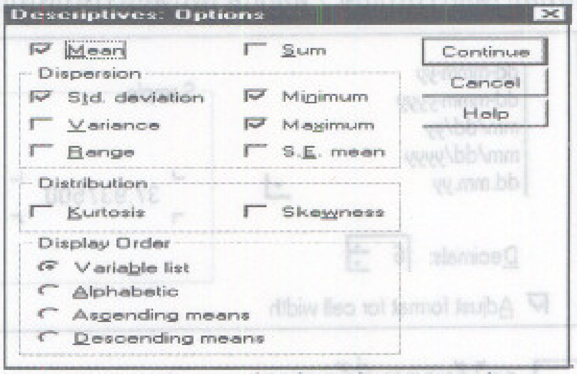
ขั้นสุดท้ายให้นำเมาส์มาคลิกนอกบริเวณตารางก็จะกลับไปที่ SPSS Output Navigator ตามปกติ

5.3.3 การคำนวณค่าสถิติเบื้องต้นอื่นๆ ด้วยคำสั่ง Descriptives...

จากขั้นตอนที่เราเลือกตัวแปร age เสร็จแล้ว



หากต้องการคำนวณค่าสถิติอื่นเพิ่มเติม ให้นำเมาส์ไปคลิกที่ปุ่ม Options... บนจอภาพจะมีเมนูย่อย ให้เลือกค่าสถิติต่างๆ เพิ่มเติม



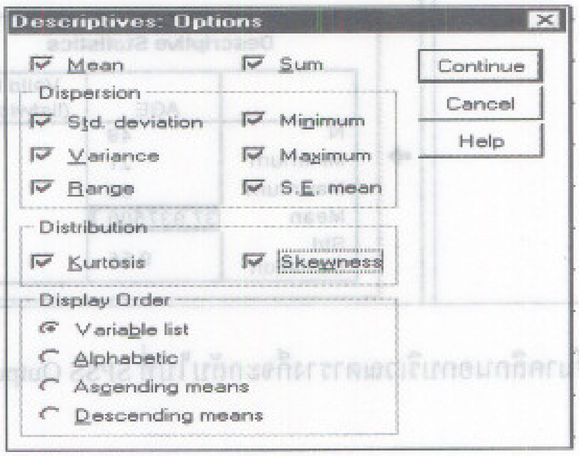
ค่าสถิติอื่นๆที่ต้องการคำนวนให้คลิกที่กรอบสี่เหลี่ยมเพื่อให้เกิดเครื่องหมายถูก

☑ มีเครื่องหมายถูก แสดงว่า ให้คำนวณค่านี้ด้วย

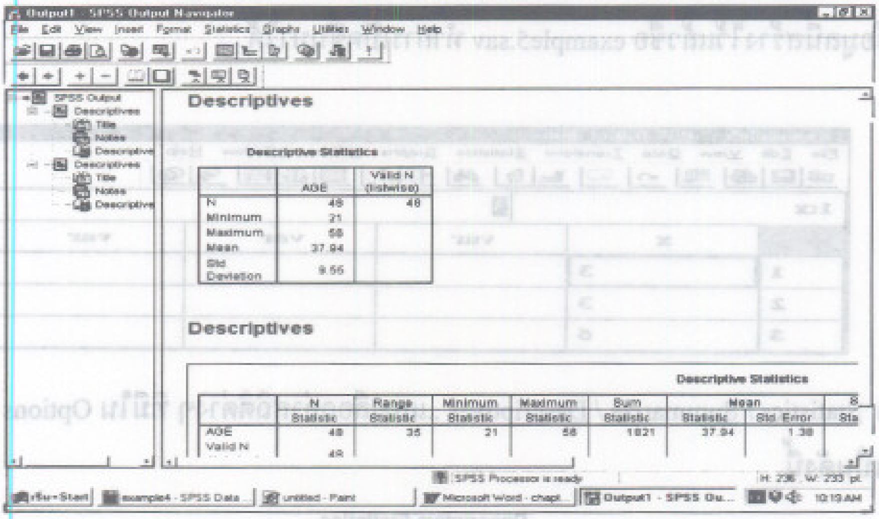
☐ ไม่มีเครื่องหมายถูก แสดงว่า ไม่ให้คำนวณค่านี้ด้วย

O เลือกอย่างใดอย่างหนึ่งเท่านั้น โดยการคลิกให้เกิดจุดสีดำหน้าข้อที่ต้องการ

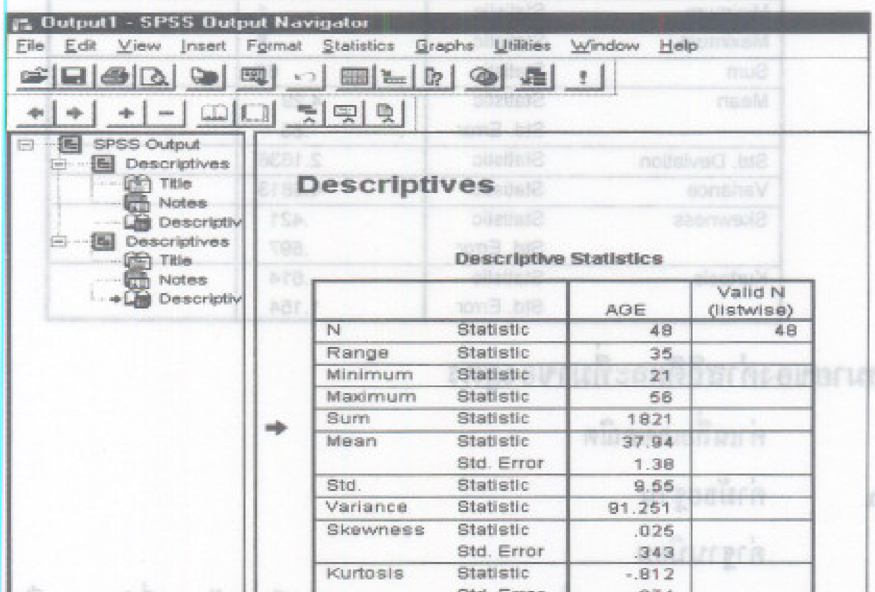
ในตัวอย่างนี้ขอให้เลือกทุกกรอบสี่เหลี่ยมให้เป็นเครื่องหมายถูก



เสร็จแล้วให้คลิก Continue และ OK ตามลำดับจะได้ผลการคำนวณดังนี้



ทำการ Transpose ตารางจะได้ผลดังนี้



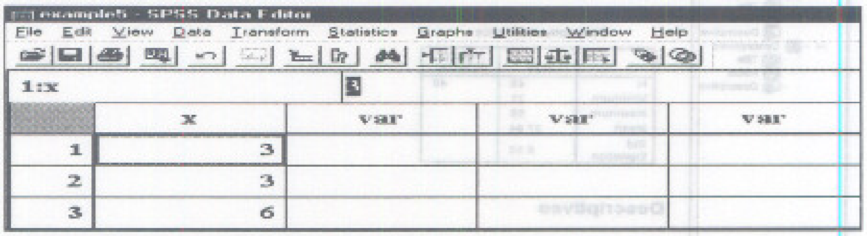
Descriptive Statistics

		AGE	Valid N (listwise)
N	Statistic	48	48
Range	Statistic	35	
Minimum	Statistic	21	
Maximum	Statistic	56	
Sum	Statistic	1821	
Mean	Statistic	37.94	
	Std. Error	1.38	
Std.	Statistic	9.55	
Variance	Statistic	91.251	
Skewness	Statistic	.025	
	Std. Error	.343	
Kurtosis	Statistic	-.812	
	Std. Error	.674	

5.4 สูตรของค่าสถิติและเปรียบเทียบการคำนวณ MATHCAD กับ SPSS

ตัวอย่าง 5.4.1 กำหนดข้อมูล 14 ตัวคือ 3 , 3 , 6 , 4 , 5 , 8 , 1 , 2 , 3 , 8 , 4 , 5 , 2 , 6 จงคำนวณค่าสิถิติเบื้องต้น ด้วยคำสั่ง Statistics / Summarize / Descriptives ..

วิธีทำ ข้อมูลนี้สร้างไว้แล้วชื่อ example5.sav ทำการเปิดขึ้นมาได้



ใช้คำสั่ง Statistics / Summarize / Descriptives ..และเลือกค่าสถิติต่างๆ ที่มีใน Options ได้ผลการคำนวณเป็นดังนี้

Descriptive Statistics

		X	Valid N (listwise)
N	Statistic	14	14
Range	Statistic	7	
Minimum	Statistic	1	
Maximum	Statistic	8	
Sum	Statistic	60	
Mean	Statistic	4.29	
	Std. Error	.58	
Std. Deviation	Statistic	2.1636	
Variance	Statistic	4.6813	
Skewness	Statistic	.421	
	Std. Error	.597	
Kurtosis	Statistic	-.614	
	Std. Error	1.154	

ความหมายของค่าสถิติและที่มาของสูตร

Mean ค่าเฉลี่ยเลขคณิต

Median ค่ามัธยฐาน

Mode ค่าฐานนิยม

Standard Deviation ค่าส่วนเบี่ยงเบนมาตรฐาน (SPSS ถือว่าข้อมูลที่คำนวณเป็น ตัวอย่าง)

Variance ค่าความแปรปรวน (ใน SPSS ถือว่าข้อมูลที่คำนวณเป็น ตัวอย่าง)

Range พิสัย = max - min

Skewness เป็นค่าที่บอกว่าข้อมูล ที่เรามีอยู่นั่นเมื่อนำไปเขียน โค้งความถี่ จะมีลักษณะความเบ้ของเส้นโค้งเป็นอย่างไร

- Skewness < 0 โค้งความถี่จะมีลักษณะ เบ้ทางด้านซ้าย หรือ เบ้ทางด้านลบ
- Skewness = 0 โค้งความถี่จะมีลักษณะสมมาตร เป็นรูประฆังคว่ำ หรือ normal curve
- Skewness > 0 โค้งความถี่จะมีลักษณะ เบ้ทางด้านขวา หรือ เบ้ทางด้านบวก

Kurtosis เป็นค่าที่บอกว่าข้อมูล ที่เรามีอยู่นั่นเมื่อนำไปเขียนโค้งความถี่ จะมีลักษณะของเส้นโค้งมีการกระจายเป็นอย่างไร

- Kurtosis < 0 ข้อมูลมีการกระจายมาก โค้งความถี่จะมีลักษณะค่อนข้างแบน
- Kurtosis = 0 ข้อมูลมีการกระจายแบบปกติ โค้งความถี่จะมีลักษณะคล้ายการแจกแจงปกติ
- Kurtosis > 0 ข้อมูลมีการกระจายน้อย โค้งความถี่จะมีลักษณะสูงโด่ง

Mean (Std. Error) เป็นค่าที่ได้มาจากสูตร $\dfrac{\text{Standard Deviation}}{\sqrt{n}}$

Minimum ค่าต่ำสุดของข้อมูล

Maximum ค่าสูงสุดของข้อมูล

Sum ผลบวกของข้อมูล

การคำนวณค่าสถิติเบื้องต้นด้วย MATHCAD

กำหนด

$$x := \begin{pmatrix} 3 \\ 3 \\ 6 \\ 4 \\ 5 \\ 8 \\ 1 \\ 2 \\ 3 \\ 8 \\ 4 \\ 5 \\ 2 \\ 6 \end{pmatrix}$$

ORIGIN := 1 $\quad$ i := 1 .. 14 $\quad \sum_{i=1}^{14} x_i = 60$

n := length(x) $\quad$ n = 14 $\quad$ min(x) = 1

max(x) = 8 $\quad$ median(x) = 4 $\quad$ mean(x) = 4.286

var(x) = 4.347 $\quad$ stdev(x) = 2.085

สูตรค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของข้อมูล

เมื่อกำหนดว่าข้อมูลนั้นคือประชากร

$$\text{variance\_population\_formula} := \frac{\sum_{i=1}^{n} (x_i - \text{mean}(x))^2}{n}$$

variance_population_formula = 4.3469

$$\text{standard\_deviation\_population\_formula} := \sqrt{\frac{\sum_{i=1}^{n} (x_i - \text{mean}(x))^2}{n}}$$

standard_deviation_population_formula = 2.0849

สูตรค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของข้อมูล

เมื่อกำหนดว่าข้อมูลนั้นคือ ตัวอย่าง

$$\text{variance\_sample\_formula} := \frac{\sum_{i=1}^{n} (x_i - \text{mean}(x))^2}{n-1}$$

variance_sample_formula = 4.6813

$$\text{standard\_deviation\_sample\_formula} := \sqrt{\frac{\sum_{i=1}^{n} (x_i - \text{mean}(x))^2}{n-1}}$$

standard_deviation_sample_formula = 2.1636

คำเตือน var(x) ของ MATHCAD เป็นค่าความแปรปรวนของประชากร

stdev(x) ของ MATHCAD เป็นค่าส่วนเบี่ยงเบนมาตรฐานของประชากร

Variance ของ SPSS เป็นค่าความแปรปรวนของตัวอย่าง

Standard Deviation ของ SPSS เป็นค่าส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง

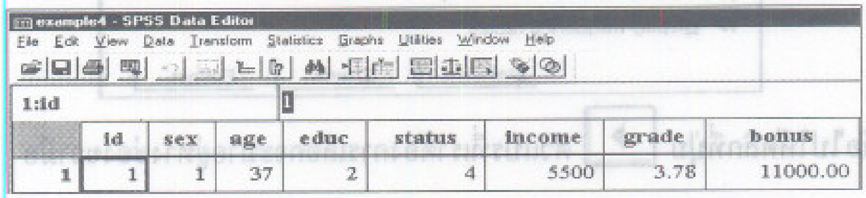
5.5 การคำนวณค่าสถิติเบื้องต้นโดยใช้คำสั่ง Statistics / Summarize / Frequencies...

คำสั่ง Statistics / Summarize / Frequencies... ใช้ในการคำนวณ

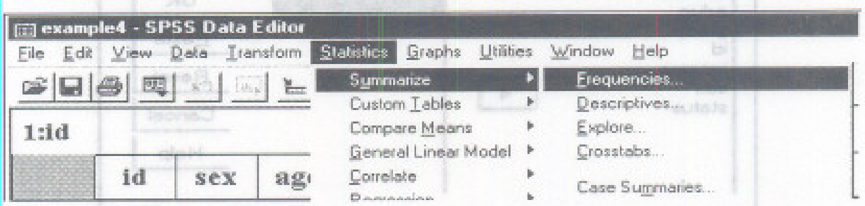
- ความถี่ข้อมูล หาจำนวนชาย หญิง หาจำนวนคนที่มี status ต่างๆ กัน
- ค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ย มัธยฐาน เปอร์เซ็นไทล์
- เขียนกราฟความถี่ของข้อมูล

ตัวอย่างการใช้คำสั่ง Statistics / Summarize / Frequencies กับตัวแปร age ในแฟ้มข้อมูล example4.sav

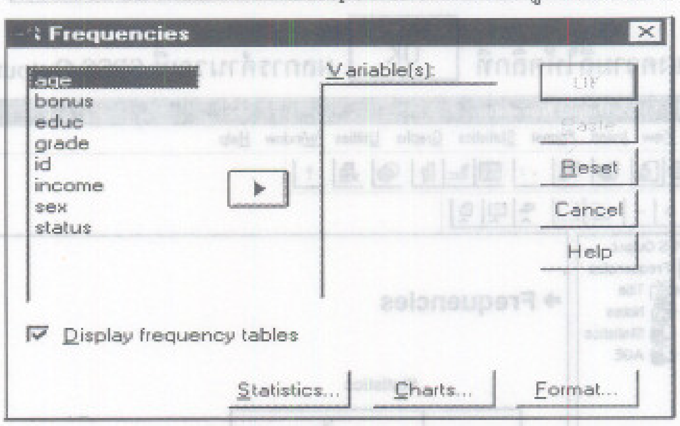
ขั้นที่ 1. นำแฟ้ม example4.sav เข้ามาใน SPSS Data Editor โดยใช้คำสั่ง File / Open ....



ขั้นที่ 2. เลือกคำสั่ง Statistics / Summarize / Frequencies บนจอภาพจะเป็นดังนี้



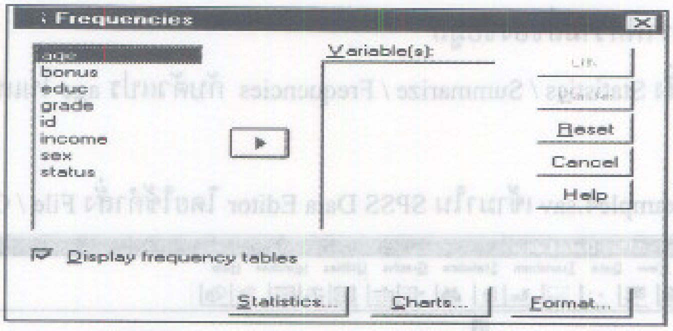
ขั้นที่ 3.. การเลือกคำสั่งให้เอาเมาส์ไปคลิกที่ Frequencies จะได้เมนูย่อยของคำสั่งดังนี้



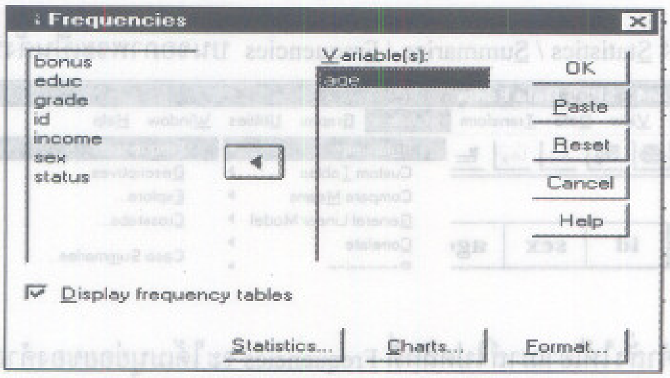
หมายเหตุ 1. ลำดับตัวแปรเรียงตามตัวอักษร

2. เมื่อเข้ามาครั้งแรกตัวแปรตัวแรกจะมีแถบสีแล้ว

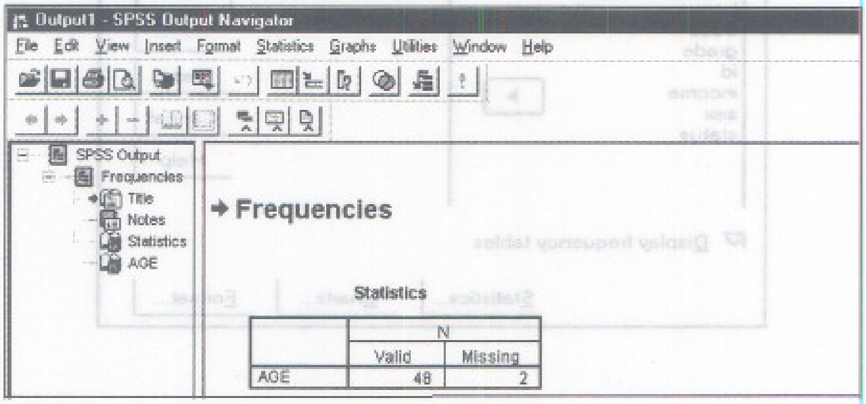
ขั้นที่ 4. เลือกตัวแปร โดยการเอาเมาส์คลิกที่ตัวแปรที่ต้องการ เช่น ตัวแปร age ขึ้นเป็นแถบสี



ขั้นที่ 5. ต่อไปให้คลิกที่ปุ่ม ▶ ตัวแปรที่เราต้องการเลือกจะมาอยู่ทางช่องขวามือ



ขั้นที่ 6. การแจกแจงความถี่ให้คลิกที่ OK ผลการคำนวณที่ SPSS Output Navigator ดังนี้



ผลการคำนวณทั้งหมดเป็นดังนี้

Frequencies

Statistics

	N	
	Valid	Missing
AGE	48	2

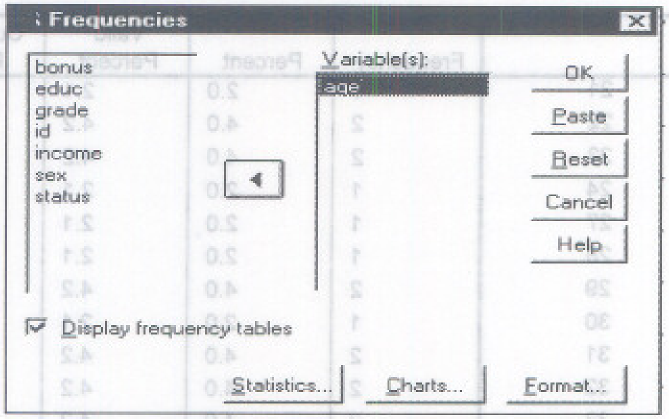
AGE

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	21	1	2.0	2.1	2.1
	22	2	4.0	4.2	6.3
	23	2	4.0	4.2	10.4
	24	1	2.0	2.1	12.5
	27	1	2.0	2.1	14.6
	28	1	2.0	2.1	16.7
	29	2	4.0	4.2	20.8
	30	1	2.0	2.1	22.9
	31	2	4.0	4.2	27.1
	32	2	4.0	4.2	31.3
	33	2	4.0	4.2	35.4
	34	2	4.0	4.2	39.6
	35	1	2.0	2.1	41.7
	37	3	6.0	6.3	47.9
	38	2	4.0	4.2	52.1
	39	1	2.0	2.1	54.2
	40	3	6.0	6.3	60.4
	41	1	2.0	2.1	62.5
	42	2	4.0	4.2	66.7
	43	2	4.0	4.2	70.8
	44	1	2.0	2.1	72.9
	45	3	6.0	6.3	79.2
	46	1	2.0	2.1	81.3
	47	1	2.0	2.1	83.3
	48	1	2.0	2.1	85.4
	50	1	2.0	2.1	87.5
	51	1	2.0	2.1	89.6
	53	2	4.0	4.2	93.8
	54	2	4.0	4.2	97.9
	56	1	2.0	2.1	100.0
	Total	48	96.0	100.0	
Missing	99	2	4.0		
	Total	2	4.0		
Total		50	100.0		

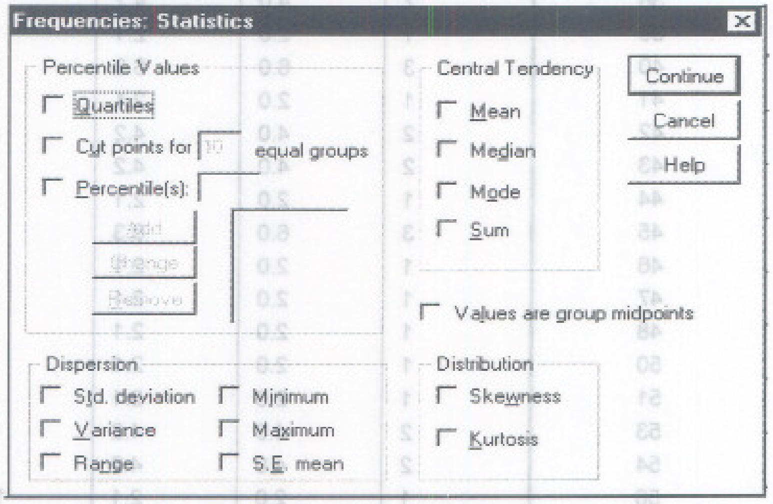
ความสามารถอื่นๆ ของคำสั่ง Frequencies.... สามารถทำได้

- หาค่าสถิติเบื้องต้น (เหมือนคำสั่ง Descriptive)
- หา เปอร์เซ็นไทล์ 1,2,3,...,99
- เขียนกราฟความถี่

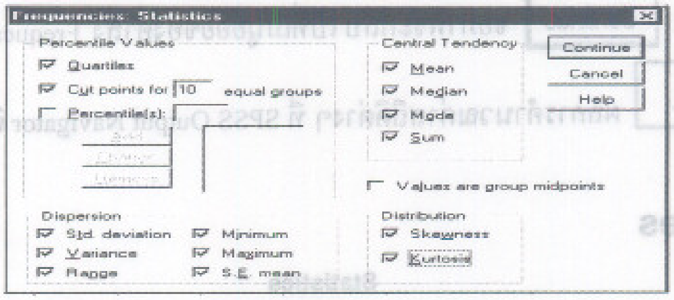
ขั้นที่ 7. จากขั้นตอนที่มีแฟ้มข้อมูลและเลือกตัวแปร age แล้ว



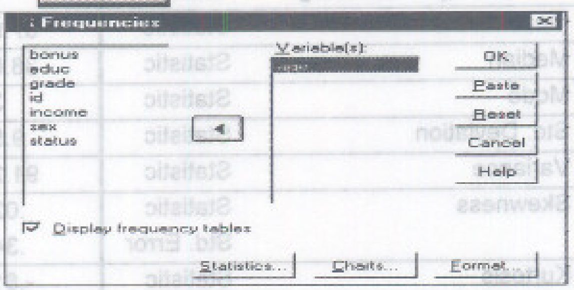
ขั้นที่ 8. ให้คลิกที่ปุ่ม Statistics... จะได้เมนูย่อยสำหรับเลือกคำนวณค่าสถิติที่ต้องการ



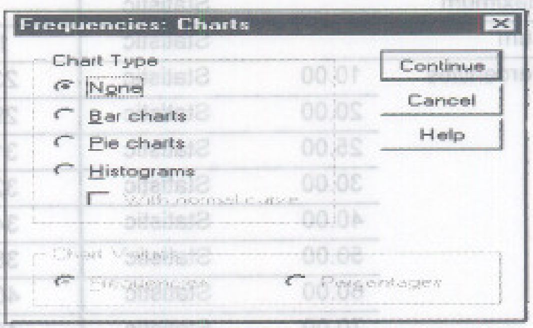
ต้องการคำนวณค่าสถิติใดให้ใส่ **เครื่องหมายถูก** ในช่องสี่เหลี่ยม (ในที่นี้ขอให้เลือกหมด ยกเว้น Percentile) ผลบนจอภาพจะเป็นดังนี้



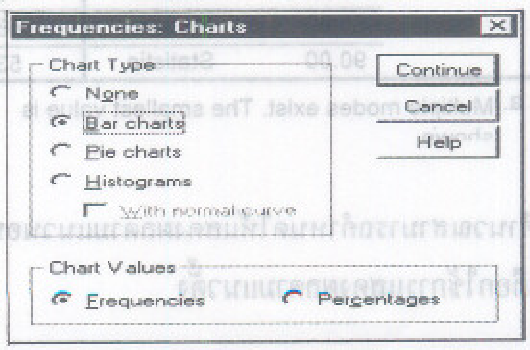
ขั้นที่ 9. เสร็จแล้วให้คลิก Continue จอภาพจะกลับไปที่เมนูย่อยของคำสั่ง Frequencies..



ขั้นที่ 10. การสั่งให้เขียนกราฟของการแจกแจงความถี่ให้คลิกที่ Charts... จะได้เมนูย่อย



ให้คลิกที่ Bar charts



เสร็จแล้วให้คลิก Continue จอภาพจะกลับไปที่เมนูย่อยของคำสั่ง Frequencies..

ให้คลิกที่ OK ผลการคำนวณค่าสถิติต่างๆ ที่ SPSS Output Navigator คือ

Frequencies

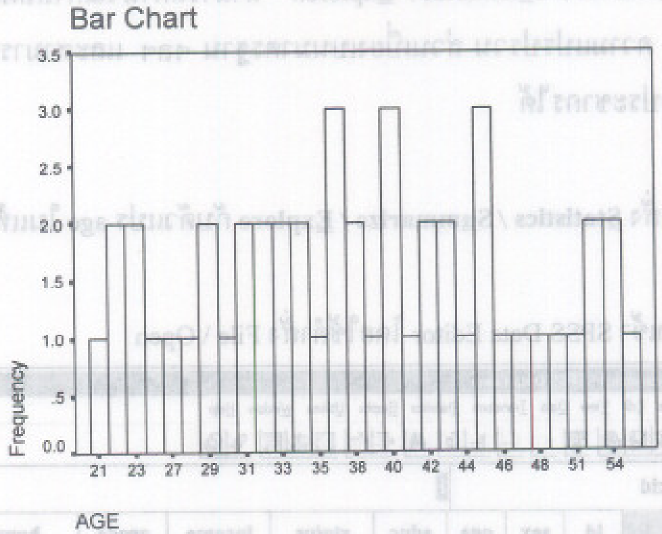
Statistics

			AGE
N	Valid	Statistic	48
	Missing	Statistic	2
Mean		Statistic	37.94
Median		Statistic	38.00
Mode		Statistic	37[a]
Std. Deviation		Statistic	9.55
Variance		Statistic	91.25
Skewness		Statistic	.025
		Std. Error	.343
Kurtosis		Statistic	-.812
		Std. Error	.674
Range		Statistic	35
Minimum		Statistic	21
Maximum		Statistic	56
Sum		Statistic	1821
Percentiles	10.00	Statistic	23.00
	20.00	Statistic	29.00
	25.00	Statistic	31.00
	30.00	Statistic	32.00
	40.00	Statistic	34.60
	50.00	Statistic	38.00
	60.00	Statistic	40.40
	70.00	Statistic	43.30
	75.00	Statistic	45.00
	80.00	Statistic	46.20
	90.00	Statistic	53.00

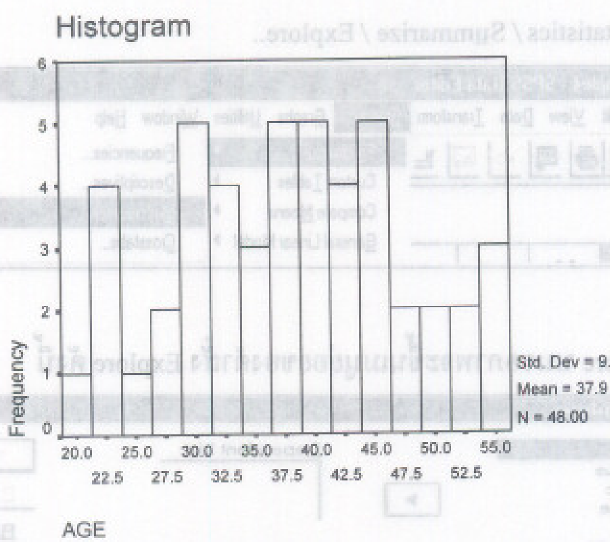
a. Multiple modes exist. The smallest value is shown

หมายเหตุ ตารางผลการคำนวณสามารถกำหนดให้แสดงผลตามแนวนอน หรือ แนวตั้งก็ได้ การแสดงควรเลือกใช้การแสดงผลตามแนวตั้ง

กราฟของความถี่แบบ Bar chart ที่ได้คือ



ในกรณีที่เลือกเป็น Histogram กราฟของความถี่แบบ Histogram ที่ได้คือ

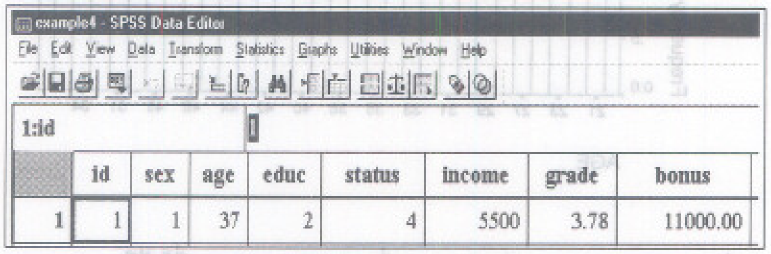


5.6 การคำนวณค่าสถิติเบื้องต้นโดยใช้คำสั่ง Statistics / Summarize / Explore

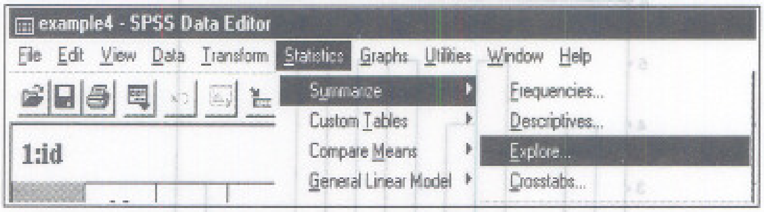
คำสั่ง Statistics / Summarize / Explore.. สามารถคำนวณค่าสถิติเบื้องต้นต่างๆ ได้เช่น ค่าเฉลี่ย มัธยฐาน ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ฯลฯ และสามารถหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยประชากรได้

ตัวอย่างการใช้คำสั่ง **Statistics / Summarize / Explore** กับตัวแปร age ในแฟ้มข้อมูล example4.sav

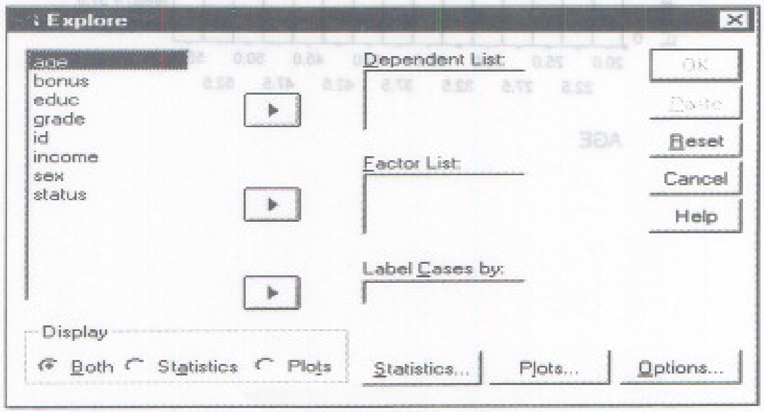
ขั้นที่ 1. นำข้อมูลเข้า SPSS Data Editor โดยใช้คำสั่ง File \ Open



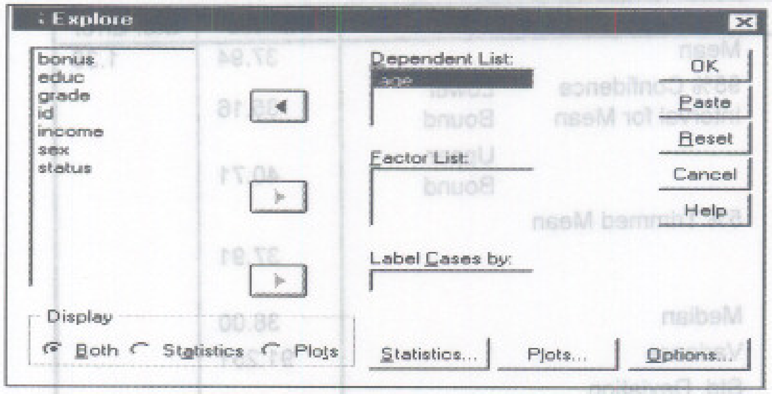
ขั้นที่ 2. เลือกคำสั่ง Statistics / Summarize / Explore..



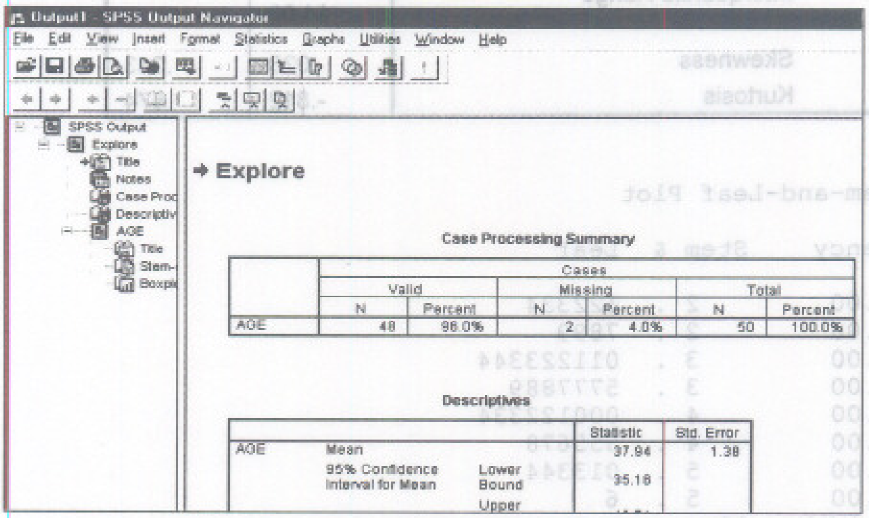
ขั้นที่ 3. คลิกที่ Explore บนจอภาพจะขึ้นเมนูย่อยของคำสั่ง Explore ดังนี้



ขั้นที่ 4. เลือกตัวแปร age มาไว้ที่ Dependent List



ขั้นที่ 5. คลิกที่ OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

Explore

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
AGE	48	96.0%	2	4.0%	50	100.0%

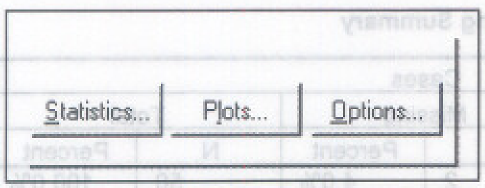
Descriptives

			Statistic	Std. Error
AGE	Mean		37.94	1.38
	95% Confidence Interval for Mean	Lower Bound	35.16	
		Upper Bound	40.71	
	5% Trimmed Mean		37.91	
	Median		38.00	
	Variance		91.251	
	Std. Deviation		9.55	
	Minimum		21	
	Maximum		56	
	Range		35	
	Interquartile Range		14.00	
	Skewness		.025	.343
	Kurtosis		-.812	.674

AGE

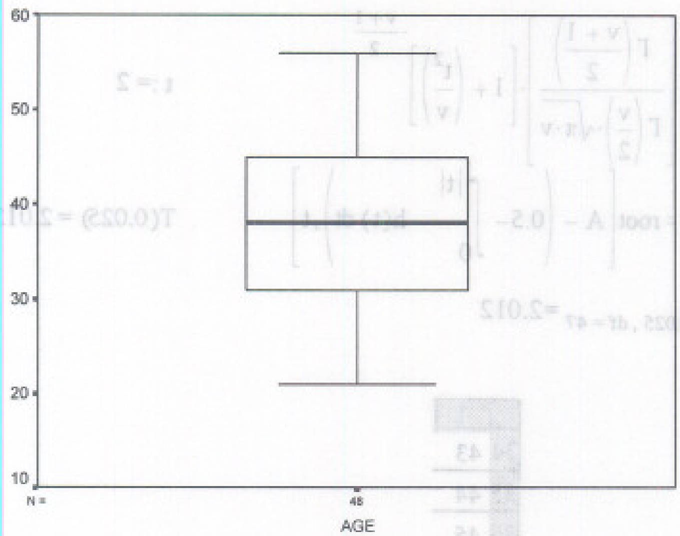
```
AGE Stem-and-Leaf Plot

 Frequency    Stem &  Leaf

     6.00        2 .  122334
     4.00        2 .  7899
     9.00        3 .  011223344
     7.00        3 .  5777889
     9.00        4 .  000122334
     6.00        4 .  555678
     6.00        5 .  013344
     1.00        5 .  6
 Stem width:    10
 Each leaf:        1 case(s)
```

หมายเหตุ ความสามารถอื่นๆ ของคำสั่ง Statistics \ Summarize \ Explore.. โดยเลือกค่าเพิ่มเติมได้

ที่ปุ่ม 

เช่นการหาเปอร์เซ็นไทล์ , เขียนกราฟของการแจกแจงความถี่

กราฟแบบ Stem-and-Leaf Plot ที่ได้



หมายเหตุ • จำนวนข้อมูลต้องคิดจาก n = 48 ซึ่งตัดค่าที่ไม่สมบูรณ์ออกไป 2 ตัว

- **Interquartile Range = 14.00**

  หมายถึงค่าได้มาจาก ควอไทล์ที่ 3 – ควอไทล์ที่ 1

- **5% Trimmed Mean = 37.91**

  หมายถึงค่าเฉลี่ยที่ได้มาจากข้อมูล 90 % ของทั้งหมดโดยการตัดค่าที่มากออกไป 5 % และตัดค่าที่น้อยออกไป 5 %

- **95% Confidence Interval for Mean Lower Bound = 35.16**

  Upper Bound = 40.71

  เป็นช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยที่แท้จริงของประชากร

  เพราะฉะนั้น ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยที่แท้จริงของประชากรคือ $35.16 < \mu < 40.71$

- สูตรช่วงความเชื่อมั่น 95% ของค่า μ คือ

$$\overline{X} - t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right) < \mu < \overline{X} + t_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right) \qquad \text{เมื่อ } df = n - 1$$

เปรียบเทียบผลการคำนวณด้วย MATHCAD

การหาค่า $t_{0.025\,,\,df=47}$ TOL := 0.000001 v := 47

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

t := 2

$$T(A) := \mathrm{root}\left[A-\left(0.5-\int_0^{|t|} h(t)\,dt\right),t\right]$$

T(0.025) = 2.012

เพราะฉะนั้น $t_{0.025\,,\,df=47} = 2.012$

x =

	1
1	21
2	22
3	22
4	23
5	23
6	24
7	27
8	28
9	29
10	29
11	30
12	31
13	31
14	32
15	32

x =

	1
34	43
35	44
36	45
37	45
38	45
39	46
40	47
41	48
42	50
43	51
44	53
45	53
46	54
47	54
48	56

n := length(x) n = 48

mean(x) = 37.938 median(x) = 38

$$\text{sample\_variance} := \frac{\mathrm{var}(x)}{n-1}\cdot n$$

sample_variance = 91.251

$$s := \sqrt{\frac{\mathrm{var}(x)}{n-1}\cdot n}$$

s = 9.553

$$\text{Lower} := \mathrm{mean}(x) - (2.012)\cdot\left(\frac{s}{\sqrt{n}}\right)$$

Lower = 35.163

$$\text{Upper} := \mathrm{mean}(x) + (2.012)\cdot\left(\frac{s}{\sqrt{n}}\right)$$

Upper = 40.712

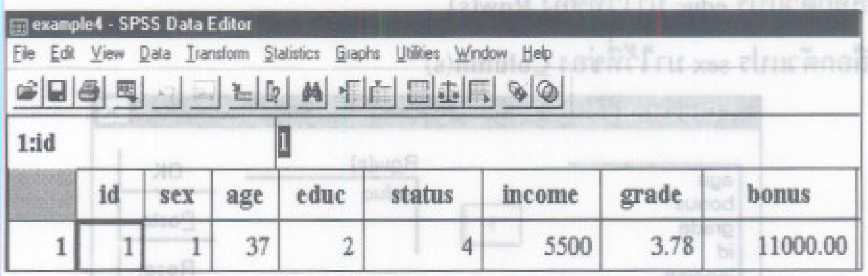
5.7 การคำนวณค่าสถิติเบื้องต้นโดยใช้คำสั่ง Statistics / Summarize / Crosstabs

คำสั่ง **Statistics / Summarize / Crosstabs** ใช้ในการคำนวณ

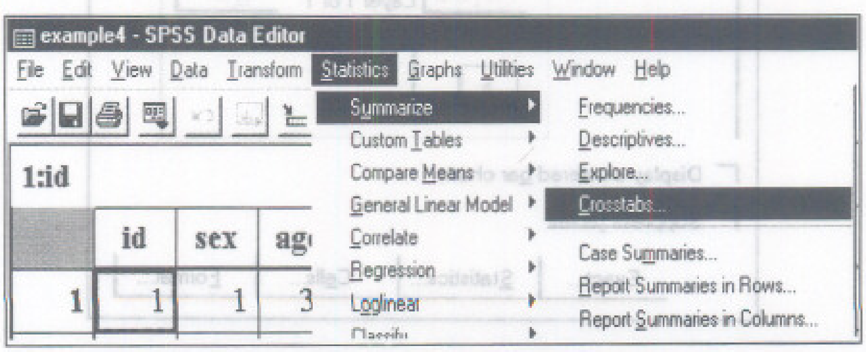
- ความถี่ข้อมูลแบบจำแนก 2 ทาง
- ค่าสถิติเบื้องต้นเช่น ค่าเฉลี่ย มัธยฐาน เปอร์เซ็นไทล์
- เขียนกราฟเปรียบเทียบความถี่ของข้อมูล
- คำนวณค่าสถิติไคสแควส์เพื่อทดสอบความเป็นอิสระของข้อมูลได้
- ฯลฯ

ตัวอย่างการใช้คำสั่ง **Statistics / Summarize / Crosstabs** กับตัวแปร educ และ sex โดยทำการแจกแจงความถี่เพื่อหาจำนวนคนจำแนกตามระดับการศึกษา และ เพศ ในแฟ้มข้อมูล example4.sav

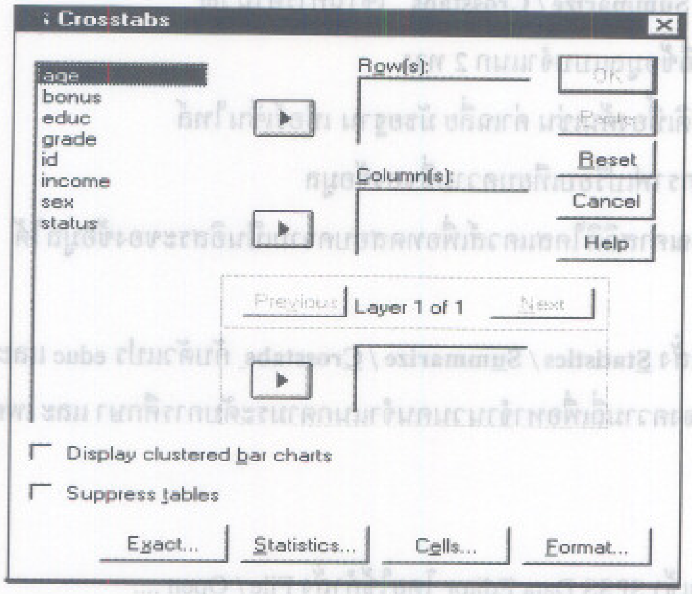
ขั้นที่ 1. นำข้อมูลเข้า SPSS Data Editor โดยใช้คำสั่ง File / Open ....



ขั้นที่ 2. เลือกคำสั่ง **Statistics / Summarize / Crosstabs..**

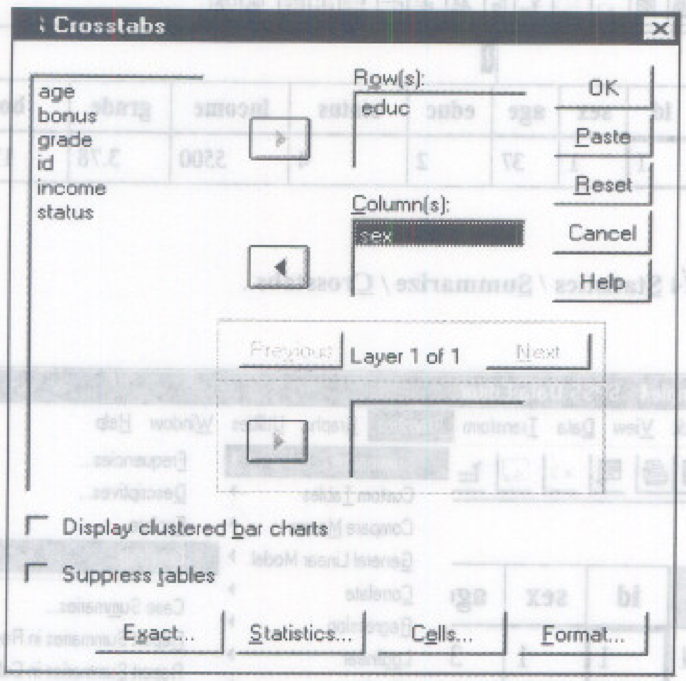


ขั้นที่ 3. ใช้เมาส์คลิกที่ **Crosstabs** บนจอภาพจะขึ้นเมนูย่อยของคำสั่ง Crosstabs

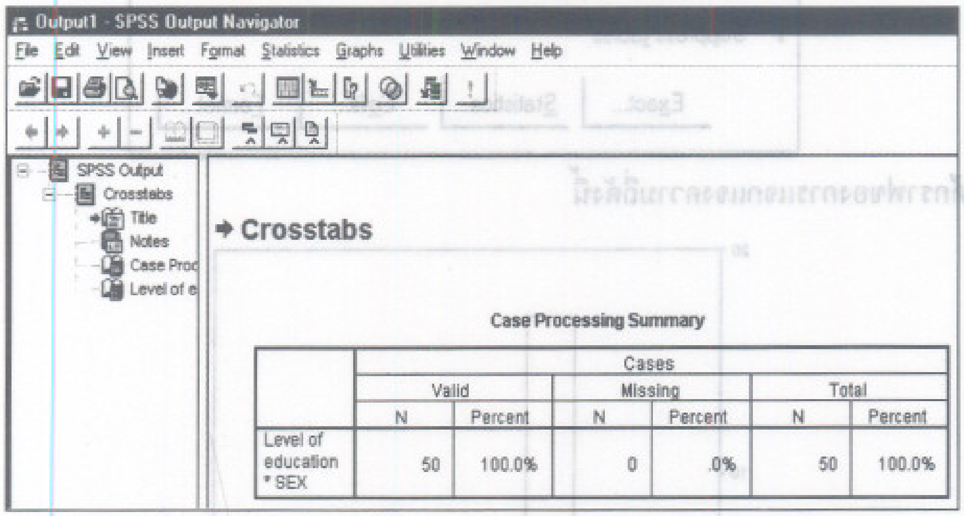


ขั้นที่ 4. เลือกตัวแปร educ มาไว้ที่ช่อง **Row(s)**

เลือกตัวแปร sex มาไว้ที่ช่อง **Column(s)**



ขั้นที่ 5. ต่อไปให้คลิกที่ OK จะได้ผลการคำนวณที่ SPSS Output Nevigator ดังนี้



ผลการคำนวณทั้งหมดคือ

Crosstabs

Case Processing Summary

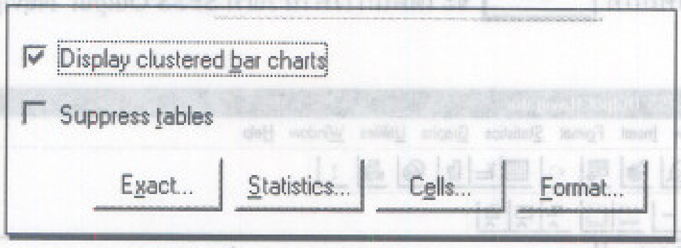
	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Level of education * SEX	50	100.0%	0	.0%	50	100.0%

Level of education * SEX Crosstabulation

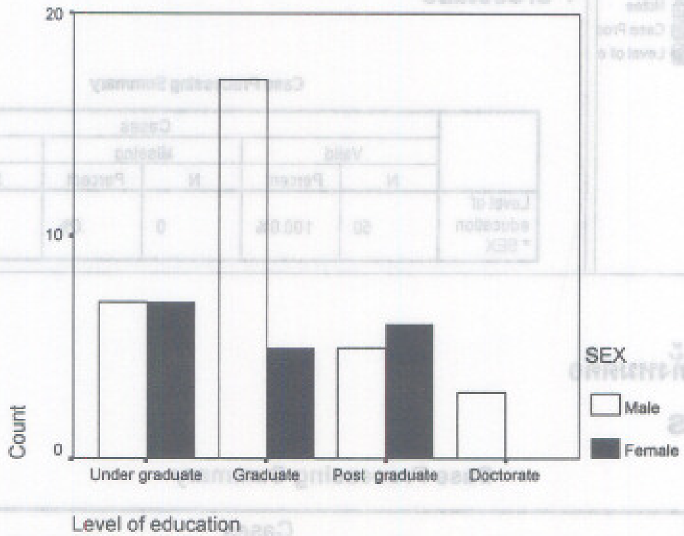
Count

		SEX		Total
		Male	Female	
Level of education	Under graduate	7	7	14
	Graduate	17	5	22
	Post graduate	5	6	11
	Doctorate	3		3
Total		32	18	50

หมายเหตุ ในกรณีที่เราเลือก Display clustered bar charts



จะได้กราฟของการแจกแจงความถี่ดังนี้



ในกรณีที่เราเลือก Statistics... และเลือกค่าสถิติ Chi-square จะได้ผลการคำนวณ

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.203[a]	3	.102
Likelihood Ratio	7.193	3	.066
Linear-by-Linear Association	.500	1	.480
N of Valid Cases	50		

a. 3 cells (37.5%) have expected count less than 5. The minimum expected count is 1.08.

ซึ่งค่าไคสแควส์ 6.203 มีประโยชน์ในการสรุปผลว่า educ กับ sex เป็นอิสระต่อกันหรือไม่

แบบฝึกหัด 5.

กำหนดข้อมูลประกอบด้วย 5 ตัวแปร และ 30 ค่าสังเกต

	sex	age	educ	status	grade
1	1	37	2	4	3.78
2	2	29	3	1	3.89
3	2	48	1	2	3.67
4	1	99	1	2	2.78
5	2	33	2	9	3.00
6	2	45	3	4	3.45
7	2	38	1	4	3.89
8	2	23	3	1	3.67
9	1	34	2	4	2.56
10	1	50	2	2	2.69
11	2	43	2	2	3.56
12	2	37	3	2	3.00
13	1	24	2	1	2.45
14	1	46	2	2	2.45
15	1	32	1	1	3.87

	sex	age	educ	status	grade
16	1	42	2	3	3.67
17	1	38	4	2	3.23
18	2	41	2	3	3.45
19	2	99	1	9	3.21
20	1	54	2	2	3.00
21	2	32	3	9	2.56
22	1	43	1	2	2.45
23	2	22	1	1	3.78
24	1	40	2	2	2.67
25	1	37	4	9	3.45
26	1	28	1	1	2.78
27	1	44	3	2	2.56
28	1	56	2	2	2.78
29	1	35	3	1	3.33
30	2	42	1	2	2.56

1. จงหาค่าสถิติเบื้องต้นของตัวแปร GRADE โดยเติมค่าที่ได้ลงในช่องว่าง

Descriptive Statistics

		GRADE	Valid N (listwise)
N	Statistic		
Range	Statistic		
Minimum	Statistic		
Maximum	Statistic		
Sum	Statistic		
Mean	Statistic		
	Std. Error		
Std.	Statistic		
Variance	Statistic		
Skewness	Statistic		
	Std. Error		
Kurtosis	Statistic		
	Std. Error		

2. จงทำการแจกแจงความถี่ของตัวแปร STATUS โดยเติมค่าที่ได้ลงในช่องว่าง

STATUS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Single				
	Married				
	Widowhood				
	Divorce				
	Total				
Missing	9				
	Total				
Total					

3. จงทำการแจกแจงความถี่ของข้อมูลที่จำแนก sex และ status โดยเติมค่าที่ได้ลงในช่องว่าง

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SEX * STATUS						

SEX * STATUS Crosstabulation

Count

		STATUS				Total
		Single	Married	Widowhood	Divorce	
SEX	Male					
	Female					
Total						

4. จงหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของ age
5. จงหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของ grade

บทที่ 6

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าพารามิเตอร์

ในบทที่ 6 จะเป็นการใช้โปรแกรม SPSS ช่วยในการหาช่วงความเชื่อมั่นของค่าพารามิเตอร์ คำสั่งสำคัญที่ใช้คือ **Statistics \ Compare Means ..**

Statistics \ Compare Means \ Means...

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

สร้างตาราง ANOVA เพื่อทดสอบค่าเฉลี่ยของประชากรหลายชุดเท่ากันหรือไม่ได้

Statistics \ Compare Means \ One-Sample T Test...

คำนวณค่าสถิติเบื้องต้น

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ

Statistics \ Compare Means \ Independent Samples T Test...

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$

Statistics \ Compare Means \ Paired-Samples T test...

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ ข้อมูลที่ไม่อิสระต่อกัน

หาค่าสหสัมพันธ์(Correlation) ได้

Statistics \ Compare Means \ One-Way ANOVA...

คำนวณค่าสถิติเบื้องต้นจำแนกตามกลุ่ม

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ จำแนกตามกลุ่มและรวมกลุ่ม

ทำตาราง ANOVA เพื่อทดสอบสมมติฐาน $H_o : \mu_1 = \mu_2 = \mu_3 = ... = \mu_k$

ทดสอบสมมติฐาน $H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = ... = \sigma_k^2$ ได้

6.1 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ

หลักการทางทฤษฎีในเนื้อหาวิชาของความน่าจะเป็นและสถิติของการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ จำแนกเป็นกรณีต่างๆ ดังนี้

1. กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง $\bar{x}$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

2. กรณีประชากรมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง $\bar{x}$ และ ค่าความแปรปรวนของตัวอย่าง s^2

2.1 $n \geq 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

2.2 $n < 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \ ; \mathrm{df} = n - 1$$

3. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n \geq 30$

คำนวณค่าเฉลี่ยของตัวอย่าง $\bar{x}$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

4. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n \geq 30$

คำนวณค่าเฉลี่ยของตัวอย่าง $\bar{x}$ และ ค่าความแปรปรวนของตัวอย่าง s^2

และประมาณค่า σ^2 ด้วย s^2 ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x} - z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ μ ด้วย SPSS For Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์จะมีการแจกแจงแบบปกติหรือไม่ก็ได้

2. สุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยของตัวอย่าง $\bar{x}$ และค่าความแปรปรวนของตัวอย่าง s^2

3. ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า μ คือ

$$\bar{x}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}<\mu<\bar{x}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \quad ; \ df = n - 1$$

หมายเหตุ โปรแกรม SPSS จะคิดว่าข้อมูลที่นำมาคำนวณเป็นข้อมูลตัวอย่างเสมอ และมีคำสั่งให้เลือกใช้หลายแบบเช่น

♥ โดยการใช้คำสั่ง Statistics / Summarize / Explore..

♠ โดยการใช้คำสั่ง Statistics / Compare Means / One-Sample T Test...

ตัวอย่าง 6.1.1 อายุหลอดไฟฟ้ามีการแจกแจงปกติ ค่าความแปรปรวนของประชากร $\sigma^2 = 1600$ สุ่มตัวอย่างหลอดไฟฟ้าจำนวน 30 หลอด หาค่าเฉลี่ยของตัวอย่างได้เท่ากับ 780 ชั่วโมง จงหาช่วงความเชื่อมั่น 95% ของค่า μ

วิธีทำ โดยใช้หลักการทางทฤษฎีของความน่าจะเป็นและสถิติ

กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

สุ่มตัวอย่างขนาด $n = 30$ ค่าเฉลี่ยของตัวอย่าง $\bar{x} = 780$ ช่วงความเชื่อมั่น 95%ของค่า μ คือ

$$\bar{x}-z_{0.025}\frac{\sigma}{\sqrt{n}}<\mu<\bar{x}+z_{0.025}\frac{\sigma}{\sqrt{n}}$$

$$780-1.96\left(\frac{40}{\sqrt{30}}\right)<\mu<780+1.96\left(\frac{40}{\sqrt{30}}\right)$$

$$765.68<\mu<794.31$$

ตัวอย่าง 6.1.2 ในการประมาณค่าเฉลี่ยของประชากรหลอดไฟฟ้า ผู้ทดลองได้ทำการสุ่มตัวอย่างหลอดไฟฟ้าจำนวน 30 หลอด ได้ข้อมูลดังนี้

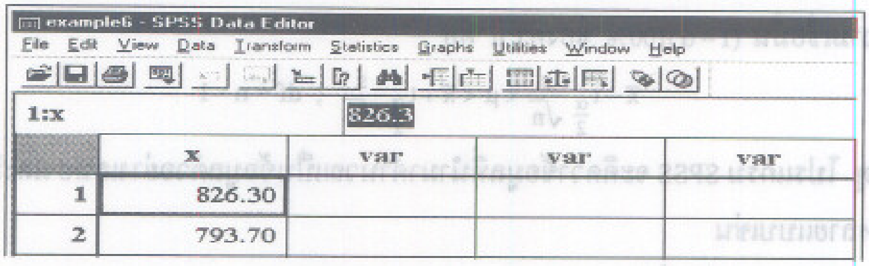
826.30	793.70	829.90	780.00	750.70	810.20	717.80	786.30	835.80	739.00
770.10	722.80	804.40	786.90	732.50	823.70	726.60	725.60	799.80	801.40
765.50	724.10	811.00	829.20	818.30	730.40	785.70	822.30	731.60	818.40

จงหาช่วงความเชื่อมั่น 95% ของค่า μ

วิธีทำ การวิเคราะห์ข้อมูลโดย SPSS

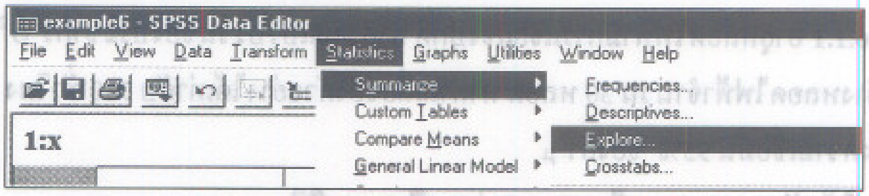
โดยการใช้คำสั่ง Statistics / Summarize / Explore..

ขั้นที่ 1. นำข้อมูลเข้าสู่โปรแกรม SPSS (ข้อมูลนี้อยู่ในแผ่น diskette ชื่อแฟ้ม example6.sav)

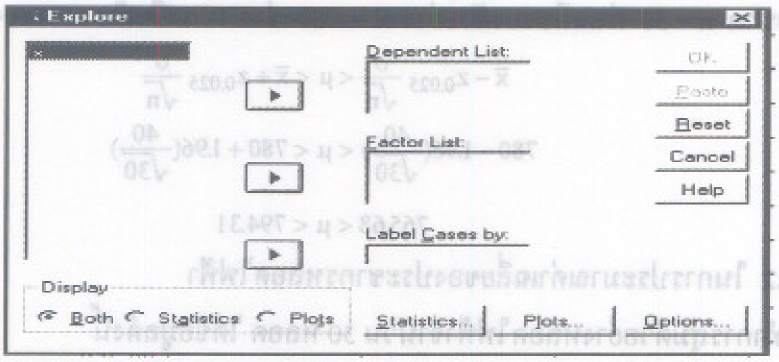


เป็นข้อมูลประกอบด้วย 1 ตัวแปรคือตัวแปร x และมี 30 ค่าสังเกต

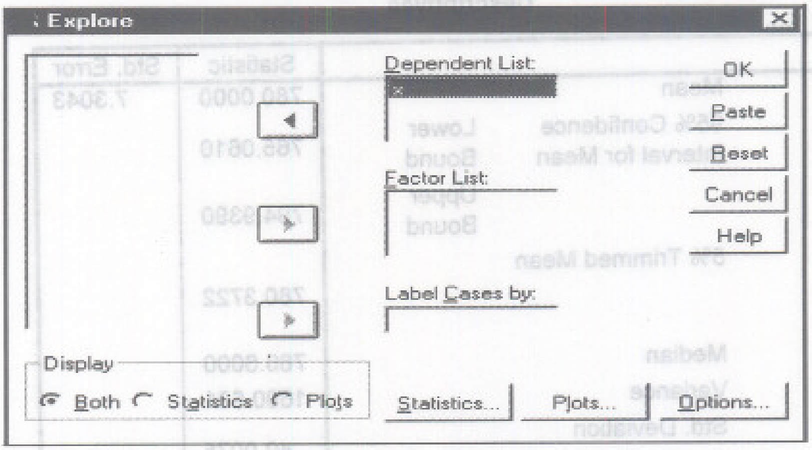
ขั้นที่ 2. ใช้คำสั่ง Statistics / Summarize / Explore..



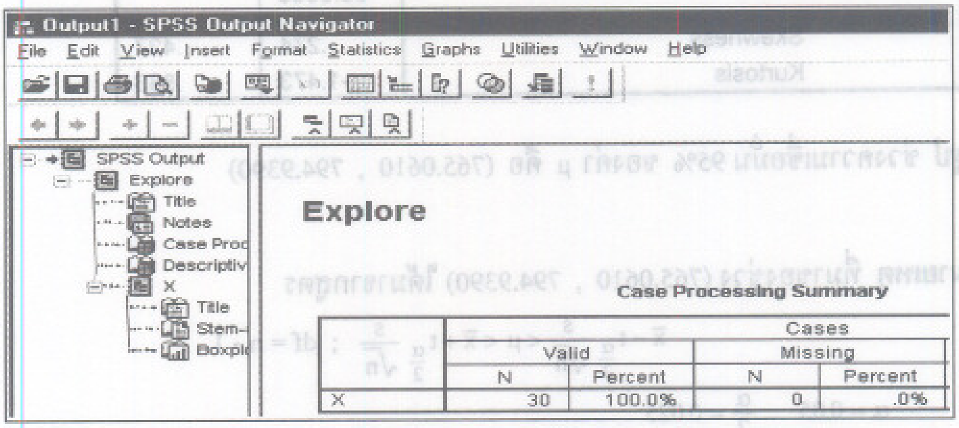
ขั้นที่ 3. คลิกที่ Explore บนจอภาพจะกลายเป็น



ขั้นที่ 4. เลือกตัวแปร x ไปทำการวิเคราะห์



ขั้นที่ 5. คลิกที่ OK จะได้ผลการคำนวณที่ SPSS Output Navigator เป็นดังนี้



ผลการคำนวณโดยละเอียดคือ

Explore

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
X	30	100.0%	0	.0%	30	100.0%

Descriptives

			Statistic	Std. Error
X	Mean		780.0000	7.3043
	95% Confidence Interval for Mean	Lower Bound	765.0610	
		Upper Bound	794.9390	
	5% Trimmed Mean		780.3722	
	Median		786.6000	
	Variance		1600.601	
	Std. Deviation		40.0075	
	Minimum		717.80	
	Maximum		835.80	
	Range		118.00	
	Interquartile Range		86.0500	
	Skewness		-.274	.427
	Kurtosis		-1.473	.833

สรุป ช่วงความเชื่อมั่น 95% ของค่า μ คือ (765.0610 , 794.9390)

หมายเหตุ ที่มาของช่วง (765.0610 , 794.9390) ได้มาจากสูตร

$$\bar{x} - t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} \quad ; \text{df} = n - 1$$

$\alpha = 0.05 \qquad \frac{\alpha}{2} = 0.025$

$df = 30 - 1 = 29$

$t_{0.025,29} = 2.045$

$\bar{x} = 780.00 \qquad s = 40.0075$

ช่วงความเชื่อมั่น 95% ของค่า μ คือ $\quad \bar{x} - t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$

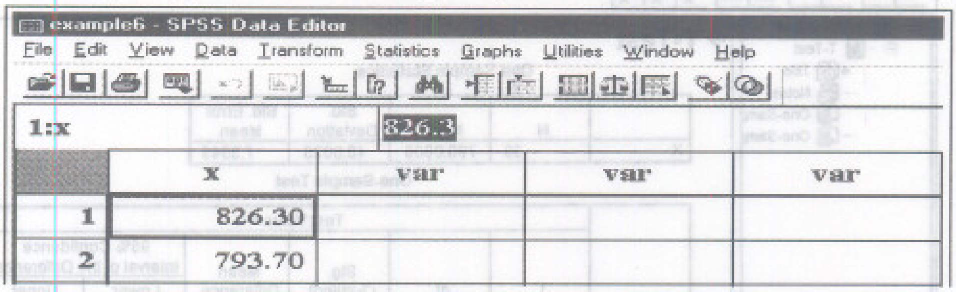
$$780.0000 - 2.045(\frac{40.0075}{\sqrt{30}}) < \mu < 780.0000 + 2.045(\frac{40.0075}{\sqrt{30}})$$

$$780.0000 - 14.937 < \mu < 780.000 + 14.937$$

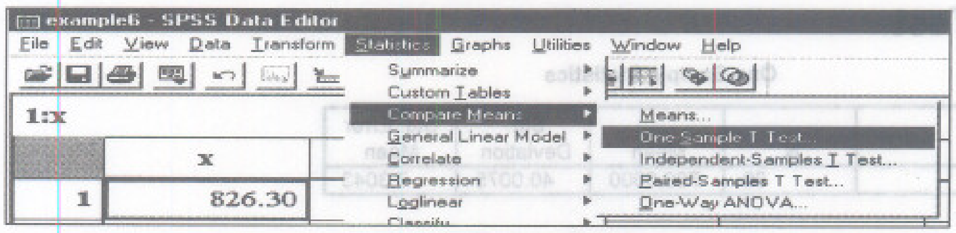
$$765.063 < \mu < 794.937$$

โดยใช้คำสั่ง Statistics / Compare Means / One-Sample T-test...

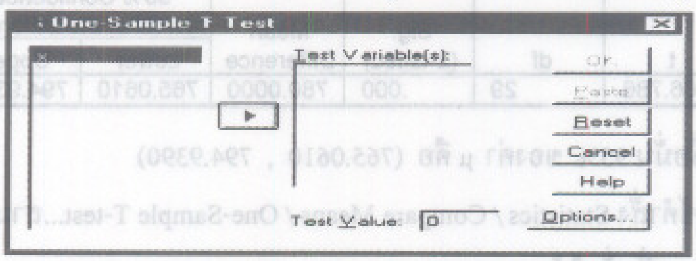
ขั้นที่ 1. นำข้อมูลเข้าสู่โปรแกรม SPSS (ข้อมูลนี้อยู่ในแผ่น diskette ชื่อแฟ้ม example6.sav)



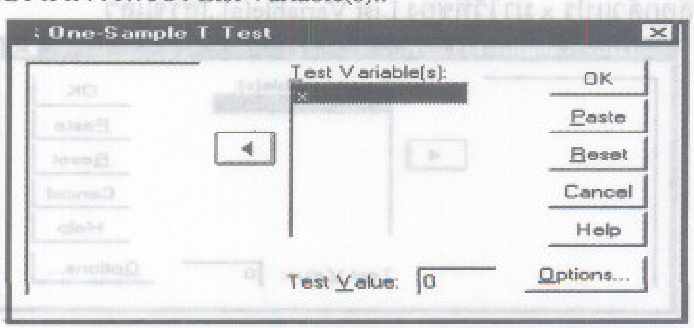
ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / One-Sample T-test...



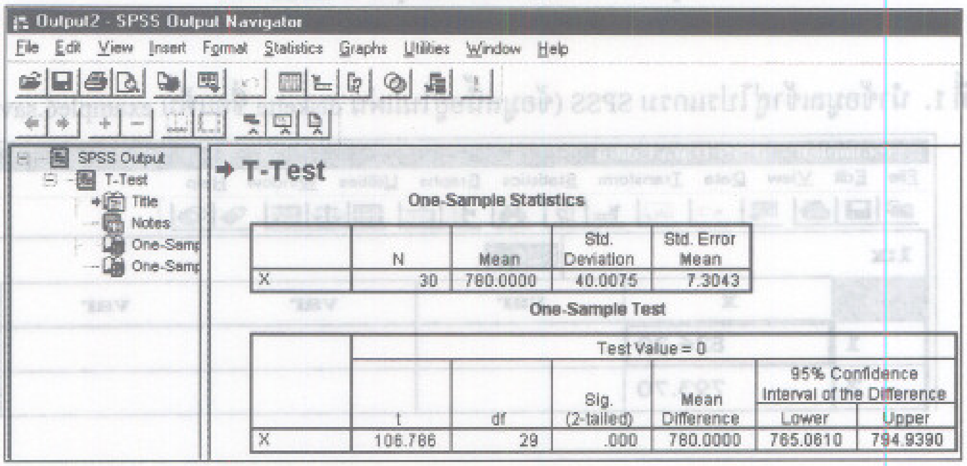
ขั้นที่ 3. คลิกที่ **One-Sample T-test** จอภาพจะขึ้นเมนูย่อยของคำสั่ง **One-Sample T-test**



ขั้นที่ 4. เลือกตัวแปร x มาไว้ที่ช่อง List Variable(s)..



ขั้นที่ 5. เสร็จแล้วคลิกที่ OK จะได้ผลที่จอ Output Navigator ดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
X	30	780.0000	40.0075	7.3043

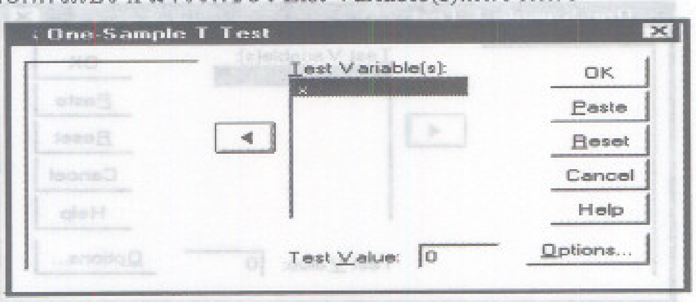
One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
X	106.786	29	.000	780.0000	765.0610	794.9390

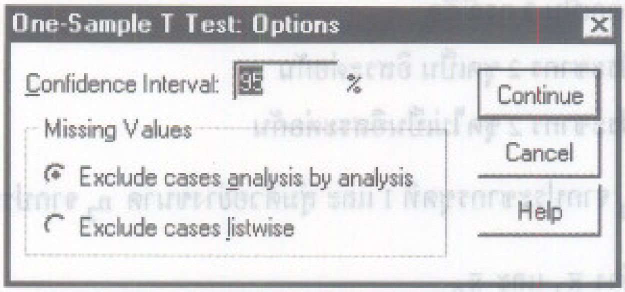
สรุป ช่วงความเชื่อมั่น 95% ของค่า μ คือ (765.0610 , 794.9390)

หมายเหตุ โดยใช้คำสั่ง Statistics / Compare Means / One-Sample T-test...สามารถเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นได้

จากขั้นที่ 4. เมื่อเลือกตัวแปร x มาไว้ที่ช่อง List Variable(s)..เสร็จแล้ว



ขั้นที่ 4.1 หากต้องการเปลี่ยนเปอร์เซ็นต์ ของช่วงความเชื่อมั่นให้คลิกที่ Options... บนจอภาพจะขึ้นเมนูย่อยเป็น



ขั้นที่ 4.2 ให้เปลี่ยนจาก 95% เป็น 99% แล้วคลิกที่ Continue และ OK ตามลำดับ จะได้ผลที่จอ SPSS Output Navigator ดังนี้

One-Sample Test

	Test Value = 0					
					99% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
X	106.786	29	.000	780.0000	759.8664	800.1336

สรุป ช่วงความเชื่อมั่น 99% ของค่า μ คือ (759.8664 , 800.1366)

หมายเหตุ ที่มาของ (759.8664 , 800.1366) ได้จากสูตร

$$\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \quad ; \ df = n-1$$

$\alpha = 0.01$ $\qquad \frac{\alpha}{2} = 0.005$

$df = 30 - 1 = 29$ $\qquad t_{0.005,29} = 2.756$

$\bar{x} = 780.00$ $\qquad s = 40.0075$

ช่วงความเชื่อมั่น 99% ของค่า μ คือ $\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$

$$780.0000 - 2.765(\frac{40.0075}{\sqrt{30}}) < \mu < 780.0000 + 2.765(\frac{40.0075}{\sqrt{30}})$$

$$780.0000 - 20.131 < \mu < 780.000 + 20.131$$

$$759.869 < \mu < 800.131$$

6.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ กรณีประชากร 2 ชุดเป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติในการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ จะจำแนกออกเป็น 2 กรณีคือ

- กรณีที่ประชากร 2 ชุดเป็น อิสระต่อกัน
- กรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน

สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2

หาค่าเฉลี่ยของตัวอย่าง $\bar{x}_1$ และ $\bar{x}_2$

1. กรณี $n_1 \geq 30$ และ $n_2 \geq 30$

1.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

1.2. กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2 และประมาณ σ_1^2 และ σ_2^2 ด้วย s_1^2 และ s_2^2

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

2. กรณี $n_1 < 30$ หรือ $n_2 < 30$

2.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

2.2 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

2.2.1 ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<(\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

เมื่อ $s_p^2 = \dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ และ $df = n_1+n_2-2$

2.2.2 ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

เมื่อ $df = \dfrac{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})^2}{(\frac{s_1^2}{n_1})^2\frac{1}{(n_1-1)}+(\frac{s_2^2}{n_2})^2\frac{1}{(n_2-1)}}$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $\mu_1-\mu_2$ ด้วย SPSS For Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์จะมีการแจกแจงแบบปกติหรือไม่ก็ได้
2. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2
3. หาค่าเฉลี่ยของตัวอย่าง $\bar{x}_1$ และ $\bar{x}_2$ หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2}}s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

เมื่อ $s_p^2 = \dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ และ $df = n_1+n_2-2$

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2}}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

เมื่อ $df = \dfrac{(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2})^2}{(\frac{s_1^2}{n_1})^2\frac{1}{(n_1-1)}+(\frac{s_2^2}{n_2})^2\frac{1}{(n_2-1)}}$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ **ของค่า** $\mu_1 - \mu_2$

โดยใช้คำสั่ง Statistics / Compare Means / Independent Samples T test..

ตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด

ตัวอย่างขนาด $n_1 = 9$ จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36	**57.76**	**71.94**	**61.77**	**58.66**	**71.61**	**71.52**	**58.67**	**62.77**

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

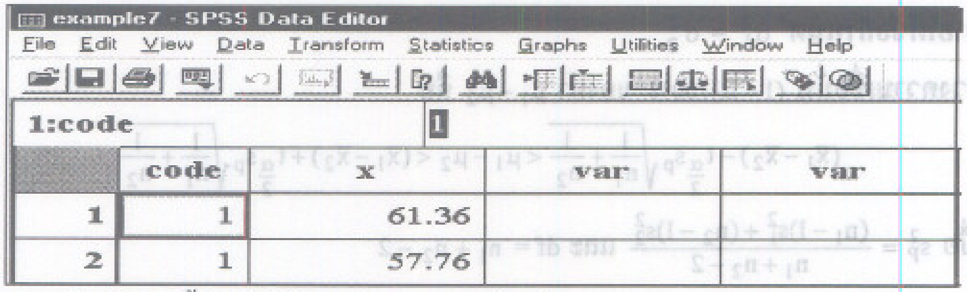
56.92	**58.30**	**67.48**	**53.96**	**62.00**	**59.61**	**52.02**	**61.60**
64.83	**58.55**	**52.53**	**64.74**	**55.51**	**66.18**	**55.51**	**54.18**

จงหาช่วงความเชื่อมั่น 95% ของค่า $\mu_1 - \mu_2$

วิธีทำ การคำนวณด้วย SPSS

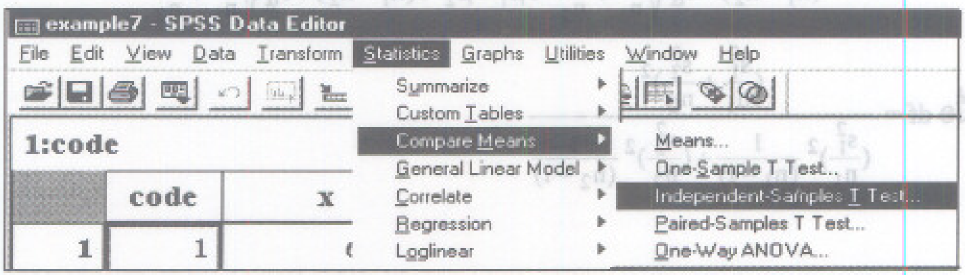
โดยใช้คำสั่ง Statistics / Compare Means / Independent Samples T test..

ขั้นที่ 1. สร้างแฟ้มข้อมูลโดยกำหนดให้มีตัวแปร 2 ตัวคือ ตัวแปรจำแนกกลุ่มตัวอย่าง (code) และ ตัวแปรข้อมูล (x)

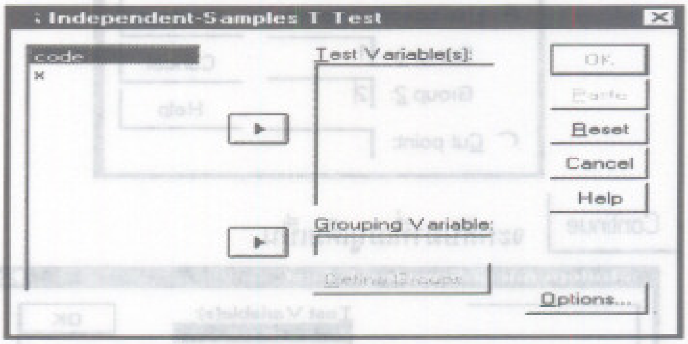


หมายเหตุ แฟ้มข้อมูลนี้บันทึกไว้ในแผ่น diskette ชื่อแฟ้ม example7.sav

ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / Independent Samples T test ..

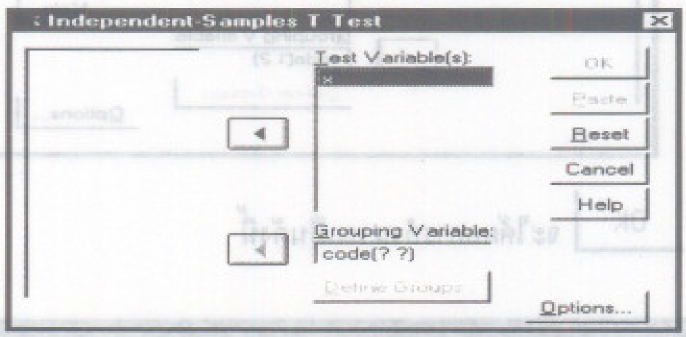


ขั้นที่ 3. คลิกที่ **Independent Sample T test** จะได้เมนูย่อยของคำสั่งดังนี้



ขั้นที่ 4. เลือกตัวแปร code มาไว้ที่ช่อง Grouping Variable

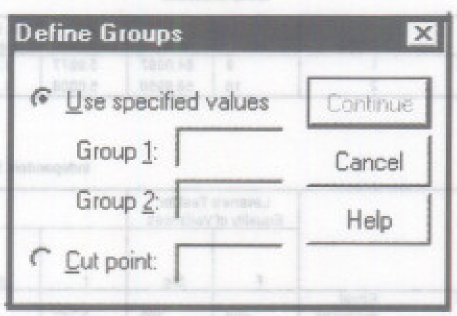
และ เลือกตัวแปร x มาไว้ที่ช่อง Test Variable(s)



ขั้นที่ 5. เลือกหมายเลขของกลุ่มในตัวแปร code ที่ต้องการวิเคราะห์ข้อมูล

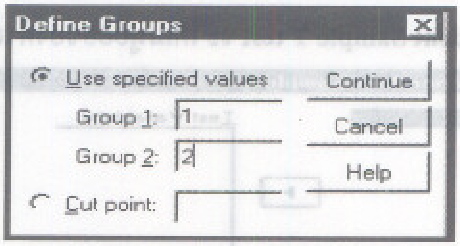
1. คลิกที่ **code(? ?)**
2. ต่อไปคลิกที่ Define Groups...

จะได้เมนูย่อยของการเลือกหมายเลขกลุ่มเป็นดังนี้

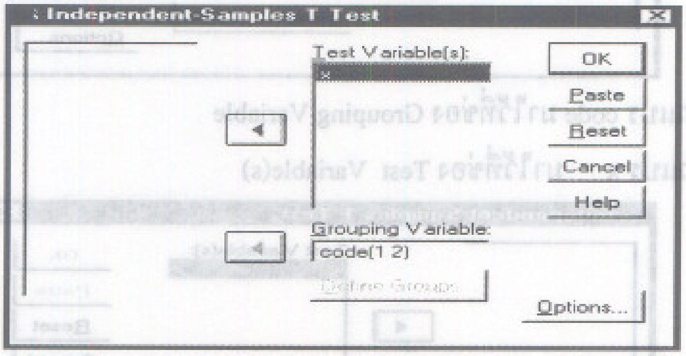


ขั้นที่ 6. ให้นำเมาส์มาคลิกที่ช่อง Group 1 และ พิมพ์หมายเลข 1 ในช่อง Group 1

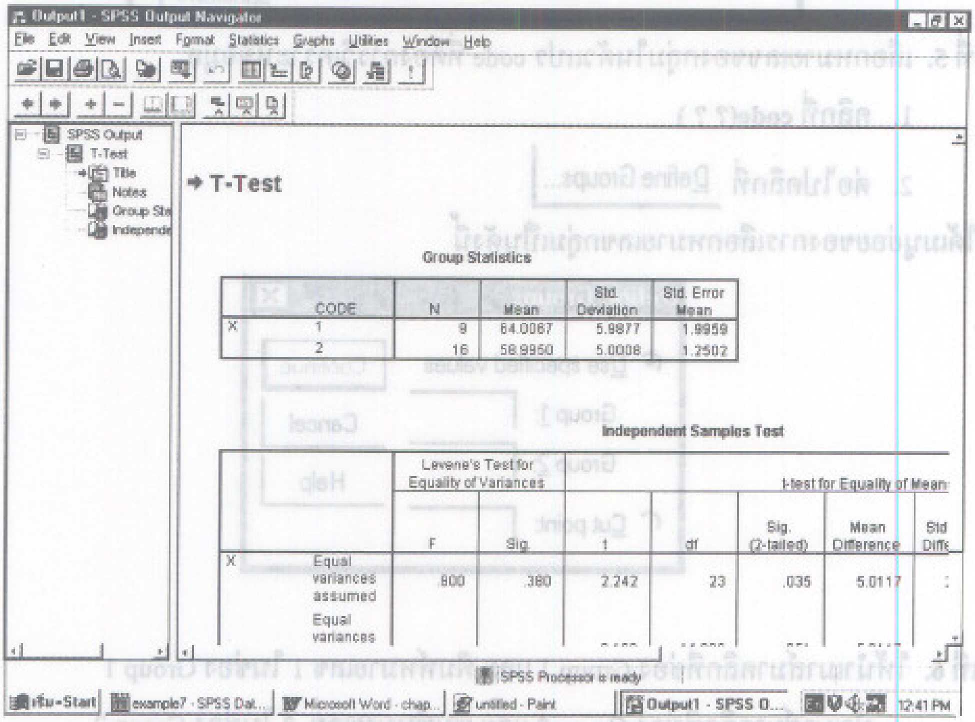
นำเมาส์มาคลิกที่ช่อง Group 2 และ พิมพ์หมายเลข 2 ในช่อง Group 2



ขั้นที่ 7. คลิกที่ Continue จะกลับมาที่เมนูเดิมเป็น



ขั้นที่ 8. คลิกที่ OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

Group Statistics

	CODE	N	Mean	Std. Deviation	Std. Error Mean
X	1	9	64.0067	5.9877	1.9959
	2	16	58.9950	5.0008	1.2502

Independent Samples Test

			X: Equal variances assumed	X: Equal variances not assumed
Levene's Test for Equality of Variances	F		.800	
	Sig.		.380	
t-test for Equality of Means	t		2.242	2.128
	df		23	14.333
	Sig. (2-tailed)		.035	.051
	Mean Difference		5.0117	5.0117
	Std. Error Difference		2.2353	2.3551
	95% Confidence Interval of the Mean	Lower	.3876	-.0286
		Upper	9.6357	10.0520

การนำผลการคำนวณของ SPSS ไปใช้งานต้องเลือกให้เหมาะสมกับข้อกำหนดของประชากร

กรณีที่ 1. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนเท่ากัน ต้องใช้ผลสรุปใน **Equal variances assumed**

ช่วงความเชื่อมั่น 95% ของค่า $\mu_1-\mu_2$ คือ $0.3876<\mu_1-\mu_2<9.6357$

กรณีที่ 2. ภายใต้เงื่อนไขว่าประชากรมีการแจกแจงปกติและมีความแปรปรวนไม่เท่ากัน ต้องใช้ผลสรุปใน **Equal variances not assumed**

ช่วงความเชื่อมั่น 95% ของค่า $\mu_1-\mu_2$ คือ $-0.0286<\mu_1-\mu_2<10.0520$

ที่มาของค่าสถิติ และสูตรการคำนวณ สามารถตรวจสอบได้ด้วย MATHCAD ดังนี้

$$\text{sample2} := \begin{bmatrix} 56.92 \\ 58.3 \\ 67.48 \\ 53.96 \\ 62 \\ 59.61 \\ 52.02 \\ 61.6 \\ 64.83 \\ 58.55 \\ 52.53 \\ 64.74 \\ 55.51 \\ 66.18 \\ 55.51 \\ 54.18 \end{bmatrix} \qquad \text{sample1} := \begin{bmatrix} 61.36 \\ 57.76 \\ 71.94 \\ 61.77 \\ 58.66 \\ 71.61 \\ 71.52 \\ 58.67 \\ 62.77 \end{bmatrix}$$

ที่มาและสูตรของค่าสถิติต่างๆ ในตาราง Group Statistics

$\text{xbar1} := \text{mean}(\text{sample1})$ $\qquad$ $\text{xbar1} = 64.0067$

$\text{xbar2} := \text{mean}(\text{sample2})$ $\qquad$ $\text{xbar2} = 58.995$

$\text{n1} := \text{length}(\text{sample1})$ $\qquad$ $\text{n1} = 9$

$\text{n2} := \text{length}(\text{sample2})$ $\qquad$ $\text{n2} = 16$

$$\text{s1} := \sqrt{\frac{\text{n1}\cdot\text{var}(\text{sample1})}{(\text{n1} - 1)}} \qquad \text{s1} = 5.9877$$

$$\text{s2} := \sqrt{\frac{\text{n2}\cdot\text{var}(\text{sample2})}{(\text{n2} - 1)}} \qquad \text{s2} = 5.0008$$

$$\text{std\_Error\_Mean1} := \frac{\text{s1}}{\sqrt{\text{n1}}} \qquad \text{std\_Error\_Mean1} = 1.9959$$

$$\text{std\_Error\_Mean2} := \frac{\text{s2}}{\sqrt{\text{n2}}} \qquad \text{std\_Error\_Mean2} = 1.2502$$

ที่มาและสูตรของค่าสถิติในตาราง Independent Samples Test

กรณีที่ความแปรปรวนประชากรทั้งสองชุดเท่ากัน

$$sp := \sqrt{\frac{(n1-1)\cdot s1^2 + (n2-1)\cdot s2^2}{n1+n2-2}} \qquad sp = 5.3647$$

$$t := \frac{xbar1 - xbar2}{sp\cdot\sqrt{\frac{1}{n1}+\frac{1}{n2}}} \qquad t = 2.242$$

$$df := n1 + n2 - 2 \qquad df = 23$$

$$v := df$$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$Sig\_2\_tailed := 2\cdot\int_t^{100} h(x)dx \qquad Sig\_2\_tailed = 0.0349$$

$$Mean\_Difference := xbar1 - xbar2 \qquad Mean\_Difference = 5.0117$$

$$Std\_Error\_Difference := sp\cdot\sqrt{\frac{1}{n1}+\frac{1}{n2}} \qquad Std\_Error\_Difference = 2.2353$$

$$TOL := 0.00001 \qquad k := 0 \quad \alpha := 0.05$$

$$t\_alpha\_divide2 := root\left(\frac{\alpha}{2} - \int_k^{100} h(x)\,dx, k\right) \qquad t\_alpha\_divide2 = 2.0686$$

$$Lower := (xbar1 - xbar2) - t\_alpha\_divide2\cdot sp\cdot\sqrt{\frac{1}{n1}+\frac{1}{n2}} \qquad Lower = 0.3876$$

$$Upper := (xbar1 - xbar2) + t\_alpha\_divide2\cdot sp\cdot\sqrt{\frac{1}{n1}+\frac{1}{n2}} \qquad Upper = 9.6357$$

กรณีความแปรปรวนประชากรทั้งสองชุดไม่เท่ากัน

$$t := \frac{xbar1 - xbar2}{\sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}}} \qquad t = 2.128$$

$$df := \frac{\left(\frac{s1^2}{n1} + \frac{s2^2}{n2}\right)^2}{\left(\frac{s1^2}{n1}\right)^2 \cdot \left(\frac{1}{n1 - 1}\right) + \left(\frac{s2^2}{n2}\right)^2 \cdot \left(\frac{1}{n2 - 1}\right)} \qquad df = 14.3325$$

$v := df$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}}\right] \cdot \left[1 + \left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$Sig\_2\_tailed := 2 \cdot \int_t^{100} h(x)dx \qquad Sig\_2\_tailed = 0.0511$$

$$Mean\_Difference := xbar1 - xbar2 \qquad Mean\_Difference = 5.0117$$

$$Std\_Error\_Difference := \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \qquad Std\_Error\_Difference = 2.3551$$

$TOL := 0.00001 \qquad k := 0$

$$t\_alpha\_divide2 := root\left(\frac{\alpha}{2} - \int_k^{100} h(x)\,dx, k\right) \qquad t\_alpha\_divide2 = 2.1401$$

$$Lower := (xbar1 - xbar2) - t\_alpha\_divide2 \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \qquad Lower = -0.0285$$

$$Upper := (xbar1 - xbar2) + t\_alpha\_divide2 \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \qquad Upper = 10.0518$$

ตัวอย่าง 6.2.2 ข้อมูลปริมาณน้ำฝน(หน่วยเป็นนิ้ว)ของตำบลที่ 1 ในช่วง 15 ปีที่ผ่านมาเป็นดังนี้

2.40 2.42 1.87 2.50 2.29 1.68 2.57 1.60 1.65 1.41

1.66 1.32 2.43 1.83 1.41

ข้อมูลปริมาณน้ำฝน (หน่วยเป็นนิ้ว) ของตำบลที่ 2 ในช่วง 10 ปีที่ผ่านมาเป็นดังนี้

0.79 1.25 0.72 0.84 1.32 1.35 1.29 0.72 0.96 1.13

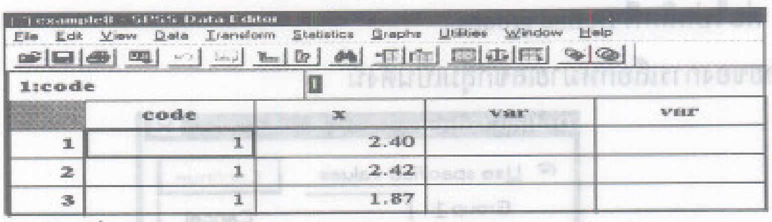
สมมติว่าข้อมูลมีการแจกแจงปกติและมีค่าความแปรปรวนประชากรแตกต่างกัน

จงหาช่วงความเชื่อมั่น 95 % ของความแตกต่างของค่าเฉลี่ยของปริมาณน้ำฝน

วิธีทำ การวิเคราะห์ข้อมูลด้วย SPSS

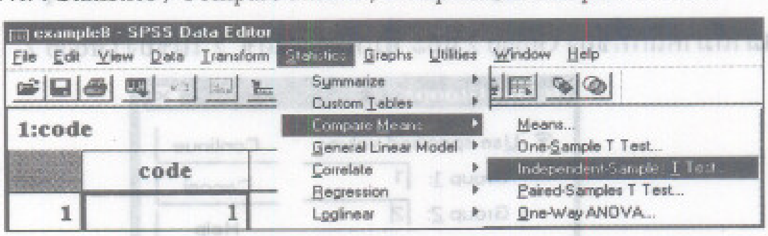
ขั้นที่ 1. สร้างแฟ้มข้อมูลที่ประกอบด้วย 2 ตัวแปรคือ

code เป็นตัวแปรจำแนกกลุ่ม x เป็นตัวแปรปริมาณน้ำฝน

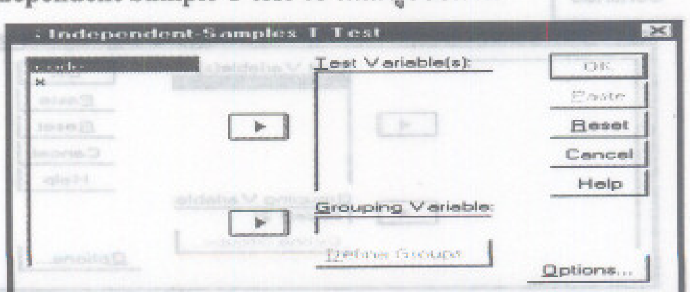


หมายเหตุ แฟ้มข้อมูลชื่อ example8.sav

ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / Independent Samples T Test..

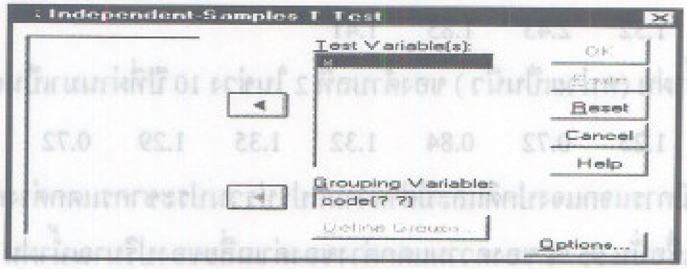


ขั้นที่ 3. คลิกที่ **Independent Sample T test** จะได้เมนูย่อยดังนี้



ขั้นที่ 4 เลือกตัวแปร code มาไว้ที่ช่อง Grouping Variable

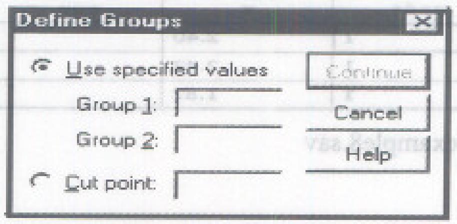
และ เลือกตัวแปร x มาไว้ที่ช่อง Test Variable(s)



ขั้นที่ 5. ต่อไปเลือกหมายเลขของกลุ่มในตัวแปร code ที่เราต้องการวิเคราะห์ข้อมูลโดยทำดังนี้

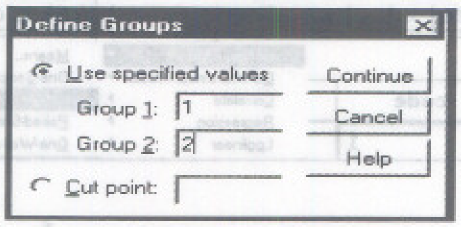
1. คลิกที่ code(? ?)
2. ต่อไปคลิกที่ Define groups..

จะได้เมนูย่อยของการเลือกหมายเลขกลุ่มเป็นดังนี้

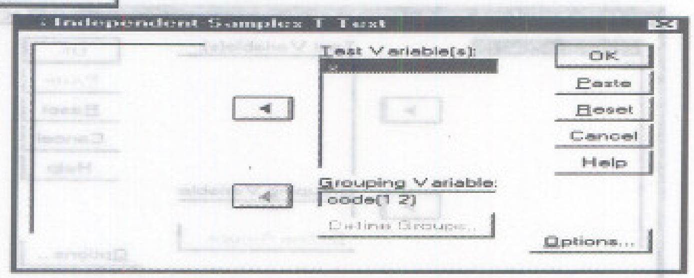


ขั้นที่ 6. นำเมาส์มาคลิกที่ช่อง Group 1 และ พิมพ์หมายเลข 1 ในช่อง Group 1

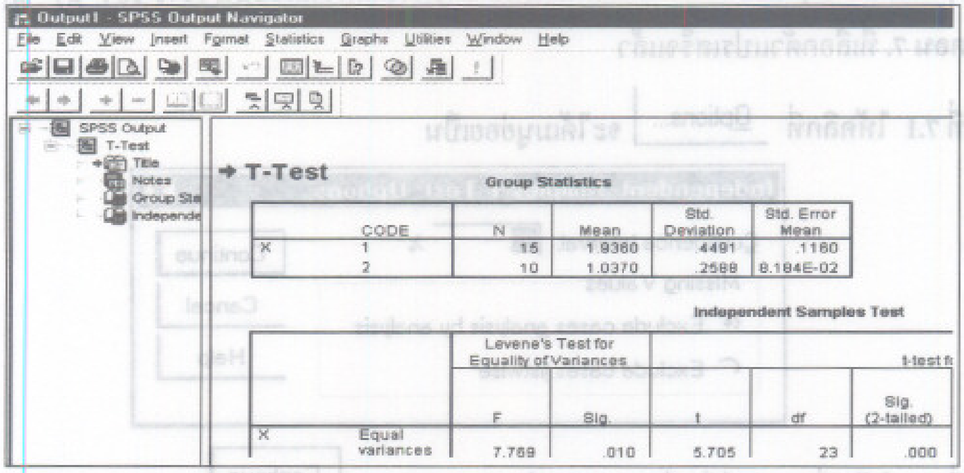
นำเมาส์มาคลิกที่ช่อง Group 2 และ พิมพ์หมายเลข 2 ในช่อง Group 2



ขั้นที่ 7. คลิกที่ 



ขั้นที่ 8. คลิกที่ OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

Group Statistics

	CODE	N	Mean	Std. Deviation	Std. Error Mean
X	1	15	1.9360	.4491	.1160
	2	10	1.0370	.2588	8.184E-02

Independent Samples Test

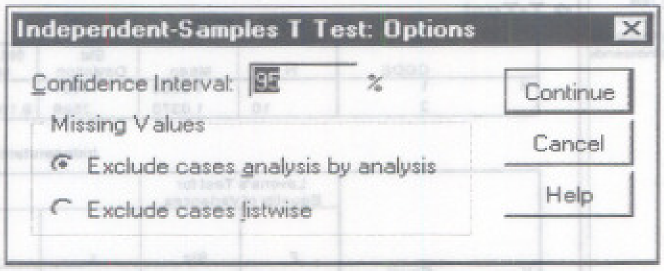
			X	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Varianc	F		7.769	
	Sig.		.010	
t-test for Equality of Means	t		5.705	6.334
	df		23	22.671
	Sig. (2-tailed)		.000	.000
	Mean Difference		.8990	.8990
	Std. Error Difference		.1576	.1419
	95% Confidence Interval of the Mean	Lower	.5730	.6052
		Upper	1.2250	1.1928

สรุป ช่วงความเชื่อมั่น 95 % ของความแตกต่างของค่าเฉลี่ยของปริมาณน้ำฝน คือ

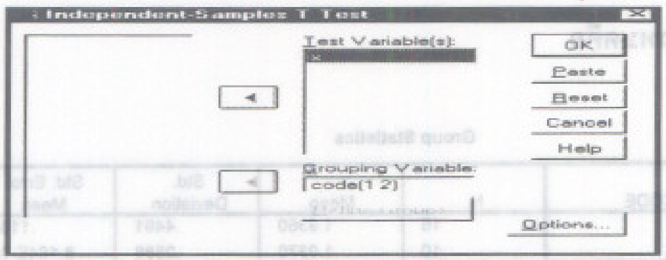
$$0.6052 < \mu_2 - \mu_1 < 1.1928$$

การเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นในการใช้งานของคำสั่ง Statistics / Compare Means / Independent Samples T Test.. ตัวอย่างเช่นต้องการหาช่วงความเชื่อมั่น 99% ของ $\mu_2 - \mu_1$ จากขั้นตอน 7. ที่เลือกตัวแปรเสร็จแล้ว

ขั้นที่ 7.1 ให้คลิกที่ Options... จะได้เมนูย่อยเป็น



ขั้นที่ 7.2 ให้เปลี่ยนเปอร์เซ็นต์ จาก 95% เป็น 99% แล้วกด Continue จอภาพจะกลับมาที่เมนู



ขั้นที่ 8. คลิกที่ OK จะได้ผลการคำนวณใหม่ในส่วนของ Independent Samples Test ดังนี้

Independent Samples Test

			x: Equal variances assumed	x: Equal variances not assumed
Levene's Test for Equality of Variances	F		7.769	
	Sig.		.010	
t-test for Equality of Means	t		5.705	6.334
	df		23	22.671
	Sig. (2-tailed)		.000	.000
	Mean Difference		.8990	.8990
	Std. Error Difference		.1576	.1419
	99% Confidence Interval of the Mean	Lower	.4566	.5000
		Upper	1.3414	1.2980

สรุป ช่วงความเชื่อมั่น 99% ของ $\mu_2 - \mu_1$ คือ (0.5000 , 1.2980)

6.3 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $\mu_1-\mu_2$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติในการหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ เมื่อประชากร 2 ชุด ไม่เป็นอิสระต่อกัน

สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2 ได้ข้อมูลเป็น

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
x_3	y_3
:	:
x_n	y_n

ขั้นตอนการคำนวณ 1. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i \quad i = 1,2,\ldots,n$

2. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง $\bar{d}$

3. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง s_d

กรณี $n \geq 30$ ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$\bar{d} - z_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}} < \mu_1-\mu_2 < \bar{d} + z_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}}$$

กรณี $n < 30$ และภายใต้ข้อสมมติว่าผลต่างของข้อมูลมีการแจกแจงปกติ

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$\bar{d} - t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}} < \mu_1-\mu_2 < \bar{d} + t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}} \quad \text{เมื่อ} \quad df = n-1$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $\mu_1-\mu_2$ ด้วย SPSS for Windows

1. ข้อมูลที่นำมาทำการวิเคราะห์มีการแจกแจงแบบปกติ
2. สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2
3. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i \quad i = 1,2,\ldots,n$
4. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง $\bar{d}$
5. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง s_d

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า $\mu_1-\mu_2$ คือ

$$\bar{d} - t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}} < \mu_1-\mu_2 < \bar{d} + t_{\frac{\alpha}{2}}\frac{s_d}{\sqrt{n}} \quad \text{เมื่อ} \quad df = n-1$$

การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของ $\mu_1 - \mu_2$ กรณีประชากร 2 ชุดไม่เป็นอิสระต่อกัน โดยใช้คำสั่ง Statistics / Compare Means / Paired – Samples T Test..

ตัวอย่าง 6.3.1 จากตัวอย่างสุ่มของนิสิตที่เรียนสถิติ 10 คน เก็บคะแนนการสอบย่อยครั้งที่ 1 และครั้งที่ 2 ของนิสิต 10 คนนั้น ปรากฏผลดังนี้

คนที่	คะแนนครั้งที่ 1.	คะแนนครั้งที่ 2.
1	76	81
2	60	52
3	85	87
4	58	70
5	91	86
6	75	77
7	82	90
8	64	63
9	79	85
10	88	83

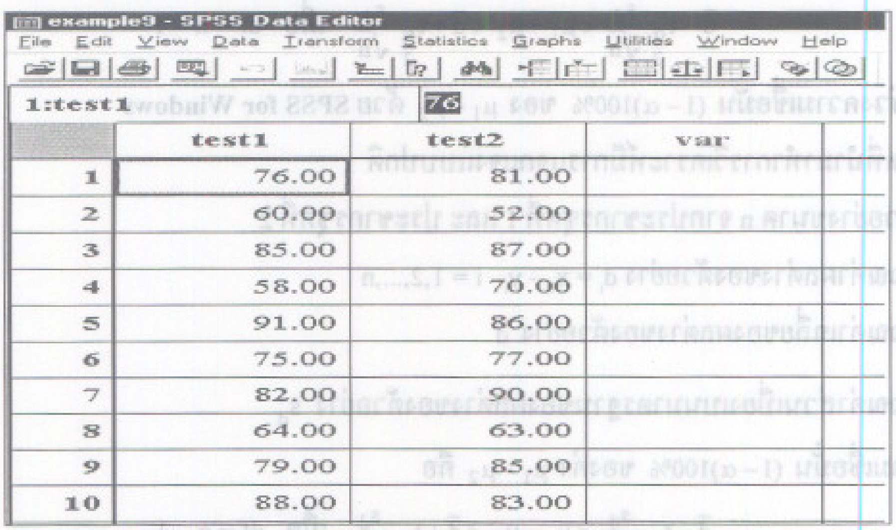
จงหาช่วงความเชื่อมั่น 98% ของค่าผลต่างที่แท้จริงในการสอบย่อย

วิธีทำ กำหนด μ_1 เป็นค่าเฉลี่ยของประชากรชุดที่ 1 และ μ_2 เป็นค่าเฉลี่ยของประชากรชุดที่ 2

การคำนวณด้วย SPSS for Windows

ขั้นที่ 1. สร้างแฟ้มข้อมูล โดยกำหนดให้มีตัวแปร 2 ตัวคือ

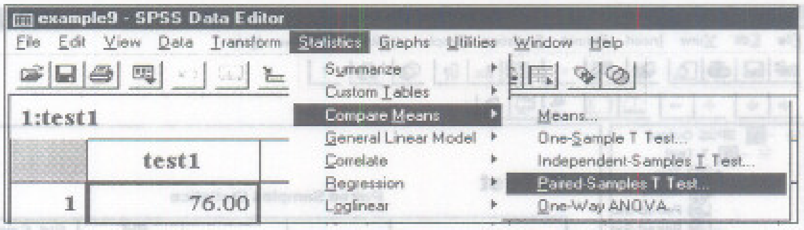
ตัวแปร test1 เป็นคะแนนสอบย่อยครั้งที่ 1 ตัวแปร test2 เป็นคะแนนสอบย่อยครั้งที่ 2



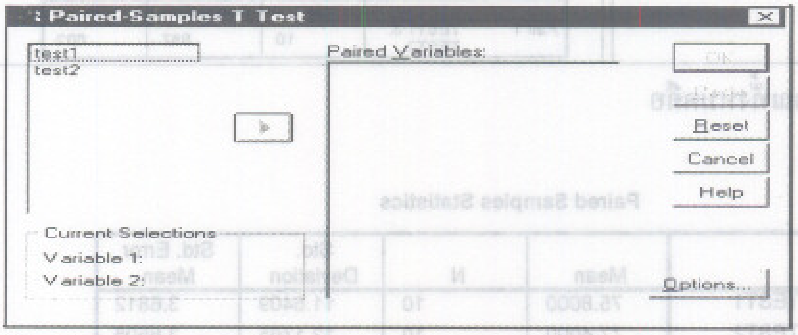
	test1	test2	var
1	76.00	81.00	
2	60.00	52.00	
3	85.00	87.00	
4	58.00	70.00	
5	91.00	86.00	
6	75.00	77.00	
7	82.00	90.00	
8	64.00	63.00	
9	79.00	85.00	
10	88.00	83.00	

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example9.sav

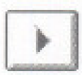
ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / Paired–Samples T Test..

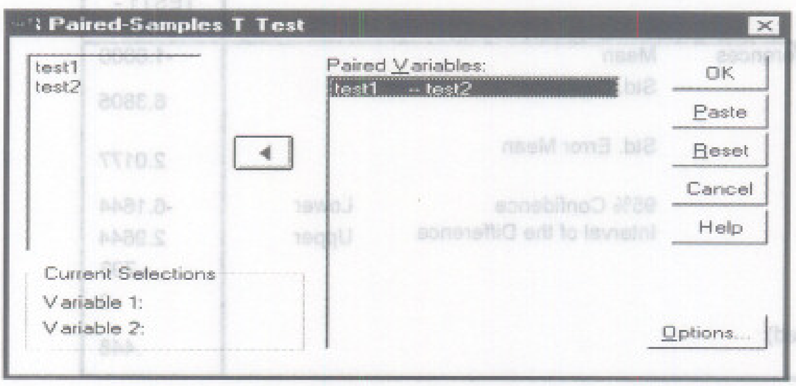


ขั้นที่ 3. คลิกที่ **Paired–Samples T Test** จะได้เมนูย่อยของคำสั่ง **Paired–Samples T Test** ดังนี้

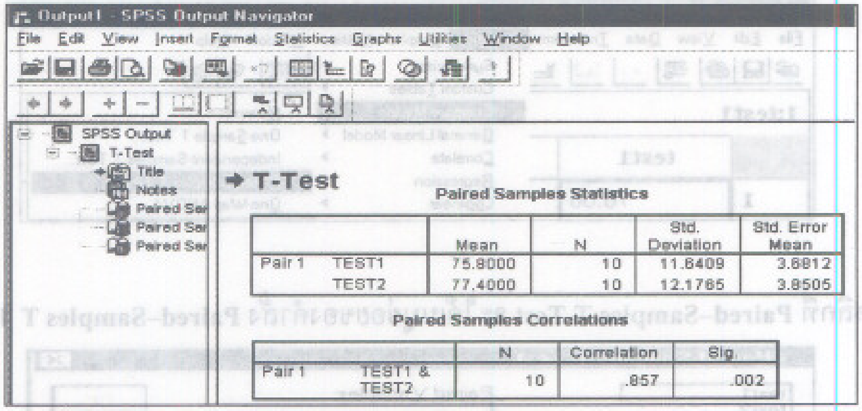


ขั้นที่ 4. การเลือกตัวแปร

1. คลิกที่ตัวแปร test1 จะได้ Variable 1: **test1**
2. คลิกที่ตัวแปร test2 จะได้ Variable 2: **test2**
3. คลิกที่  จะได้คู่ของตัวแปรที่ต้องการในช่อง Pair Variables: **test1 – test2**



ขั้นที่ 5. คลิกที่ OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	TEST1	75.8000	10	11.6409	3.6812
	TEST2	77.4000	10	12.1765	3.8505

Paired Samples Correlations

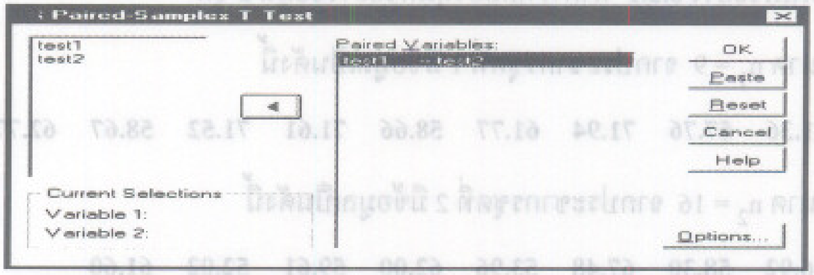
		N	Correlation	Sig.
Pair 1	TEST1 & TEST2	10	.857	.002

Paired Samples Test

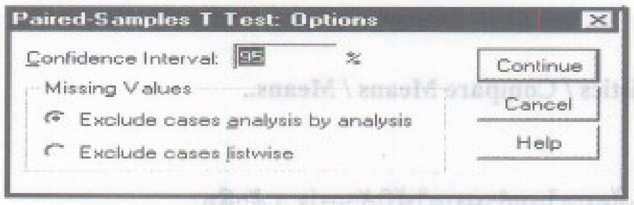
			Pair 1 TEST1 - TEST2
Paired Differences	Mean		-1.6000
	Std. Deviation		6.3805
	Std. Error Mean		2.0177
	95% Confidence Interval of the Difference	Lower	-6.1644
		Upper	2.9644
t			-.793
df			9
Sig. (2-tailed)			.448

สรุป ช่วงความเชื่อมั่น 95% ของ $\mu_1 - \mu_2$ คือ $-6.1644 < \mu_1 - \mu_2 < 2.9644$

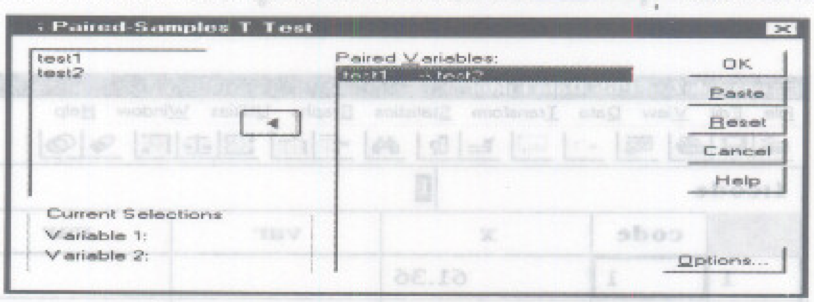
การเปลี่ยนเปอร์เซ็นต์ของช่วงความเชื่อมั่นในการใช้งานของคำสั่ง Statistics / Compare Means / Paired – Samples T Test.. จาก **ขั้นตอน 4.** ที่เลือกตัวแปรเสร็จแล้ว



ขั้นที่ 4.1 ให้คลิกที่ Options... จะได้เมนูย่อยเป็น



ขั้นที่ 4.2 ให้เปลี่ยน เปอร์เซ็นต์ จาก 95% เป็น 98% แล้วกด Continue จอภาพจะกลับมาที่เมนู



ขั้นที่ 5. คลิกที่ OK จะได้ผลการคำนวณดังนี้ในส่วนของตาราง Paired Samples Test ดังนี้

Paired Samples Test

			Pair 1 TEST1 - TEST2
Paired Differences	Mean		-1.6000
	Std. Deviation		6.3805
	Std. Error Mean		2.0177
	98% Confidence Interval of the Difference	Lower	-7.2928
		Upper	4.0928
t			-.793
df			9
Sig. (2-tailed)			.448

สรุป ช่วงความเชื่อมั่น 98% ของ $\mu_1 - \mu_2$ คือ $-7.2928 < \mu_1 - \mu_2 < 4.0928$

6.4 การวิเคราะห์ข้อมูลด้วยคำสั่ง Statistics / Compare Means / Means..

จากข้อมูลในตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด

ตัวอย่างขนาด $n_1 = 9$ จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36 57.76 71.94 61.77 58.66 71.61 71.52 58.67 62.77

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

56.92 58.30 67.48 53.96 62.00 59.61 52.02 61.60

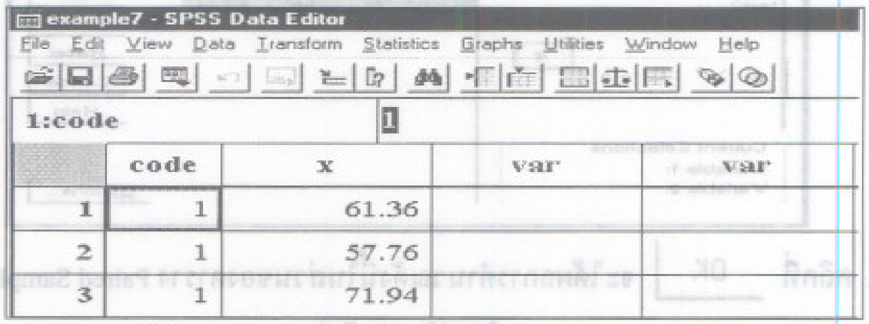
64.83 58.55 52.53 64.74 55.51 66.18 55.51 54.18

โดยใช้ คำสั่ง **Statistics / Compare Means / Means..**

วิธีทำ

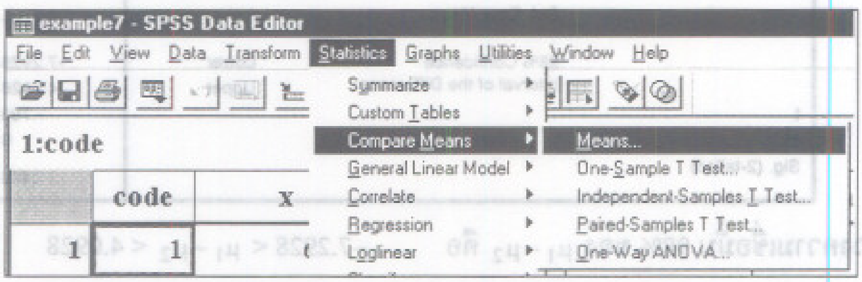
ขั้นที่ 1. สร้างแฟ้มข้อมูลโดยกำหนดให้มีตัวแปร 2 ตัวคือ

ตัวแปรจำแนกกลุ่มตัวอย่าง (code) และ ตัวแปรข้อมูล (x)

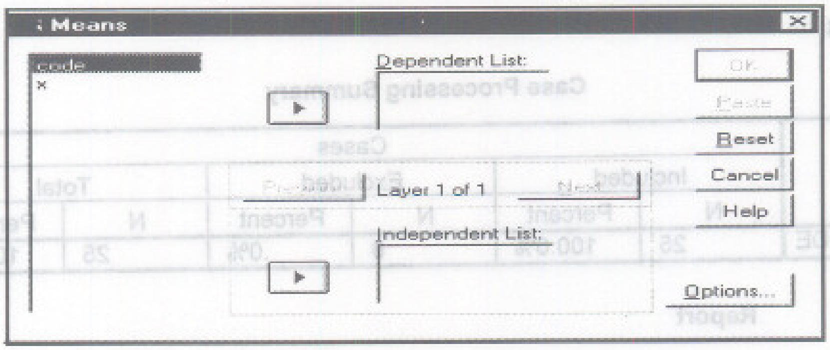


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example7.sav

ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / Means..

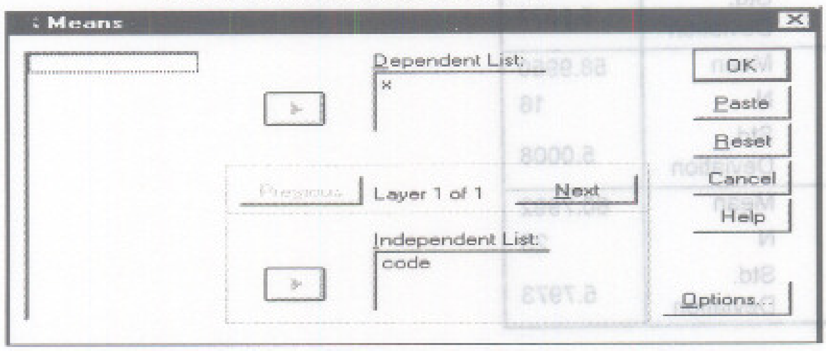


ขั้นที่ 3. คลิกที่ Mean จะได้เมนูย่อยของคำสั่ง Statistics / Compare Means / Means..

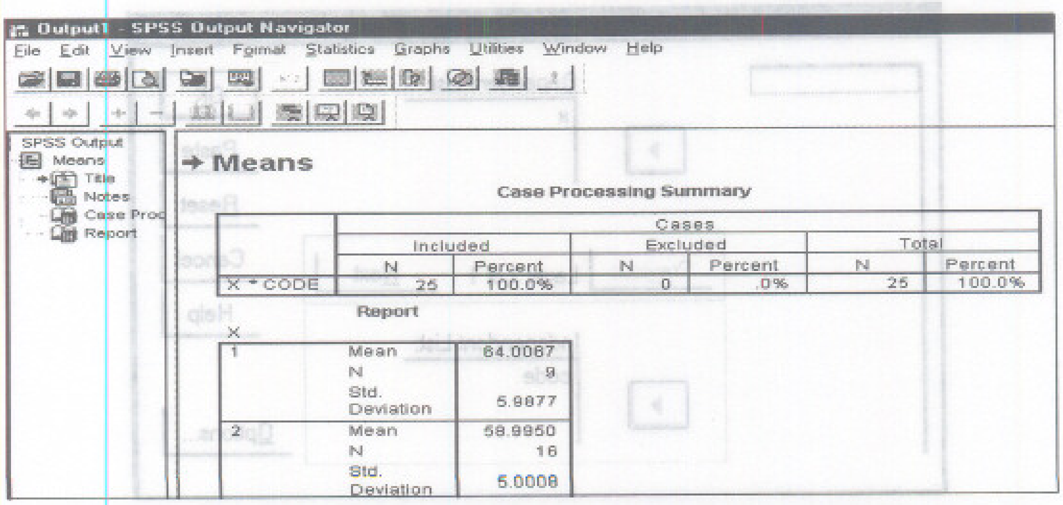


ขั้นที่ 4. เลือกตัวแปร x มาไว้ที่ช่อง Dependent List

code มาไว้ที่ช่อง Independent List



ขั้นที่ 5. เสร็จแล้วคลิกที่ OK จะได้ผลบนจอภาพเป็นดังนี้



	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
X * CODE	25	100.0%	0	.0%	25	100.0%

Report

X

CODE		
1	Mean	64.0067
	N	9
	Std. Deviation	5.9877
2	Mean	58.9950
	N	16
	Std. Deviation	5.0008

ผลการคำนวณที่ได้คือ

Means

Case Processing Summary

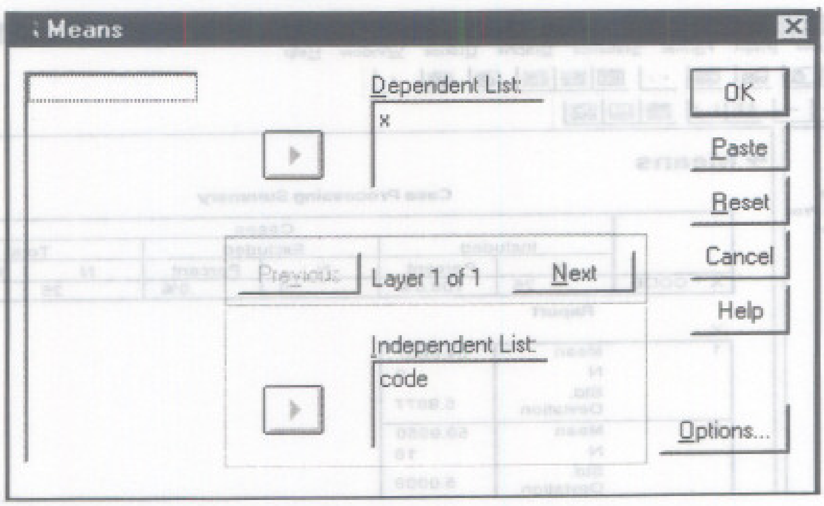
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
X * CODE	25	100.0%	0	.0%	25	100.0%

Report

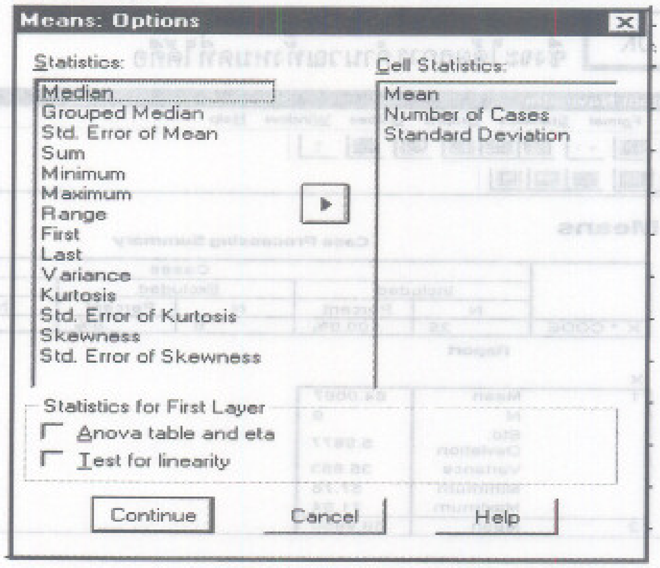
X

1	Mean	64.0067
	N	9
	Std. Deviation	5.9877
2	Mean	58.9950
	N	16
	Std. Deviation	5.0008
Total	Mean	60.7992
	N	25
	Std. Deviation	5.7973

การเพิ่มเติมความสามารถในการคำนวณค่าสถิติต่างๆ ของคำสั่ง Statistics / Compare Means / Means..จากขั้นตอนที่เลือกตัวแปรเสร็จแล้ว



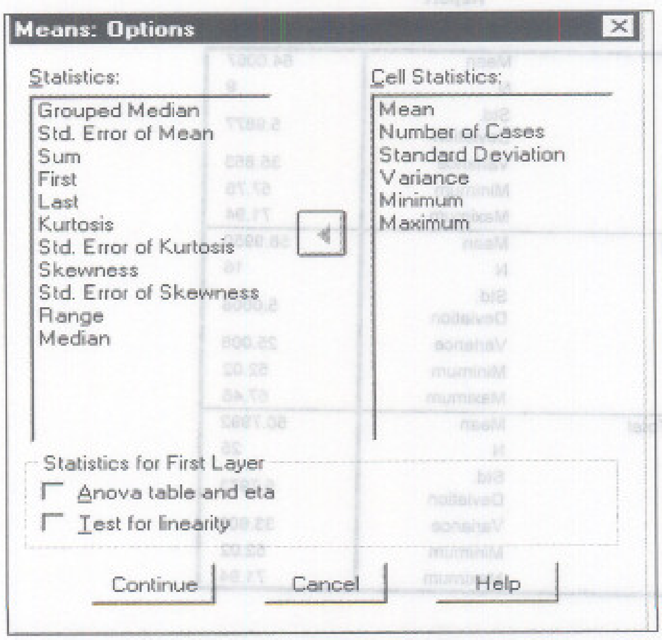
ขั้นที่ 4.1 ให้คลิกที่ Options... บนจอภาพจะขึ้นเมนูย่อยดังนี้



ขั้นที่ 4.2 ต้องการคำนวณค่าสถิติใดเพิ่มให้ทำการเลือกโดยการ

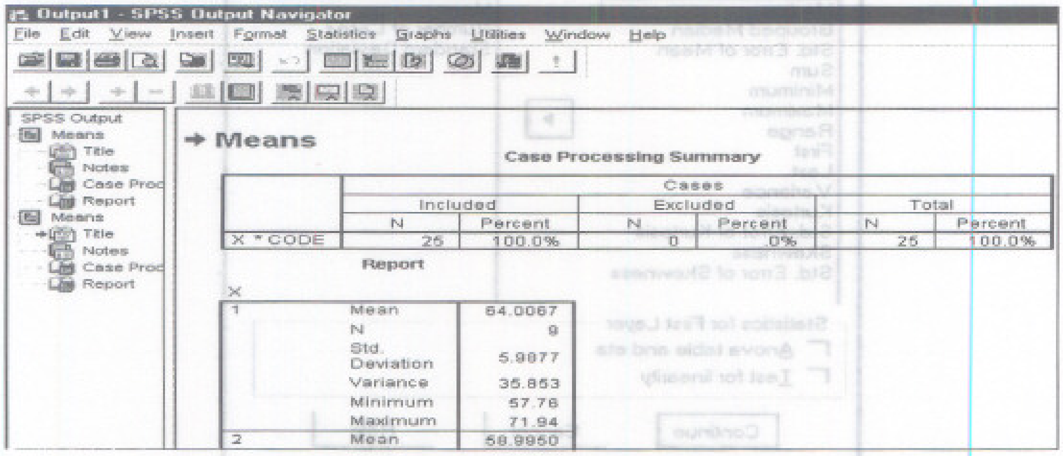
1. คลิกที่ค่าสถิติที่ช่อง Statistics เช่นค่า Variance
2. คลิกที่ ▶ จะทำให้ค่าสถิติ Variance มาอยู่ที่ช่อง Cell Statistics

ในที่นี้จะขอเลือกค่าสถิติเพิ่มจากของเดิมคือ Variance , Minimum และ Maximum



ขั้นที่ 4.3 คลิกที่ Continue เพื่อกลับไป เมนู Means

ขั้นที่ 4.4 คลิก OK ซึ่งจะได้ผลการคำนวณทั้งหมดที่ได้คือ



Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
X * CODE	25	100.0%	0	.0%	25	100.0%

Report

X

1	Mean	64.0067
	N	9
	Std. Deviation	5.9877
	Variance	35.853
	Minimum	57.76
	Maximum	71.94
2	Mean	58.9950
	N	16
	Std. Deviation	5.0008
	Variance	25.008
	Minimum	52.02
	Maximum	67.48
Total	Mean	60.7992
	N	25
	Std. Deviation	5.7973
	Variance	33.609
	Minimum	52.02
	Maximum	71.94

6.5 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่าเฉลี่ย ด้วยคำสั่ง Statistics / Compare Means / One-Way ANOVA..

จากข้อมูลในตัวอย่าง 6.2.1 ทำการทดลองสุ่มตัวอย่างข้อมูล 2 ชุด

ตัวอย่างขนาด $n_1 = 9$ จากประชากรชุดที่ 1 มีข้อมูลเป็นดังนี้

61.36 57.76 71.94 61.77 58.66 71.61 71.52 58.67 62.77

ตัวอย่างขนาด $n_2 = 16$ จากประชากรชุดที่ 2 มีข้อมูลเป็นดังนี้

56.92 58.30 67.48 53.96 62.00 59.61 52.02 61.60

64.83 58.55 52.53 64.74 55.51 66.18 55.51 54.18

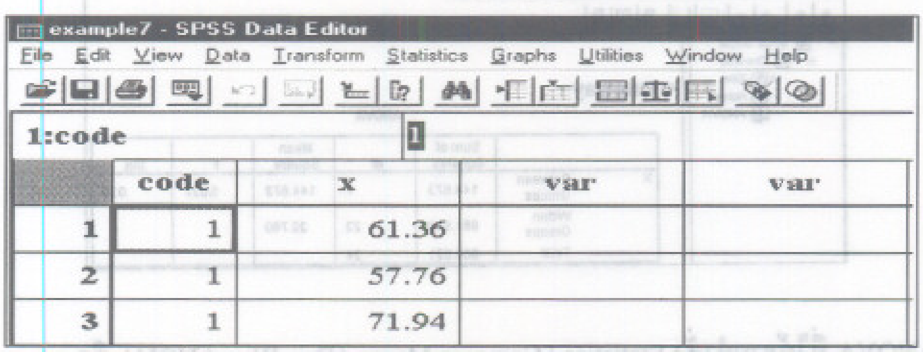
เราสามารถหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ย

โดยใช้ คำสั่ง Statistics / Compare Means / One-Way ANOVA

วิธีทำ

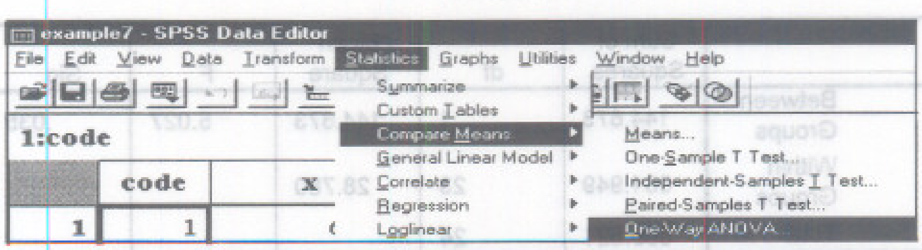
ขั้นที่ 1. สร้างแฟ้มข้อมูลโดยกำหนดให้มีตัวแปร 2 ตัวคือ

ตัวแปรจำแนกกลุ่มตัวอย่าง (code) และ ตัวแปรข้อมูล (x)

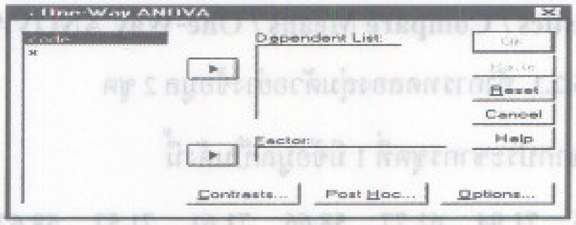


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example7.sav

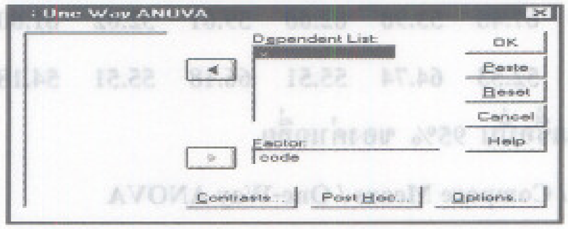
ขั้นที่ 2. เลือกคำสั่ง Statistics / Compare Means / One-Way ANOVA



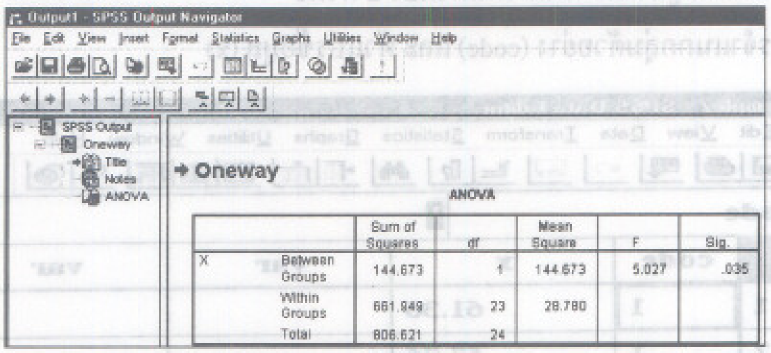
ขั้นที่ 3. คลิกที่ One-Way ANOVA จะได้เมนูย่อยของคำสั่ง One-Way ANOVA ดังนี้



ขั้นที่ 4. เลือกตัวแปร code มาไว้ที่ช่อง Factor และ เลือกตัวแปร x มาไว้ที่ช่อง Dependent List



ขั้นที่ 5. คลิก OK จะได้ผลการคำนวณเป็นดังนี้



ตาราง ANOVA ที่ได้จากคำสั่ง Statistics / Compare Means / One-Way ANOVA คือ

Oneway

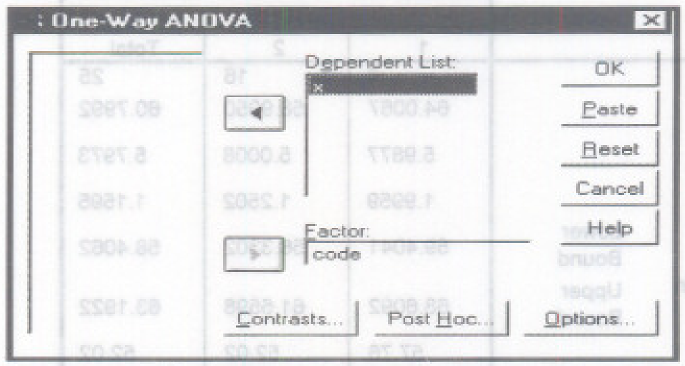
ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
X	Between Groups	144.673	1	144.673	5.027	.035
	Within Groups	661.949	23	28.780		
	Total	806.621	24			

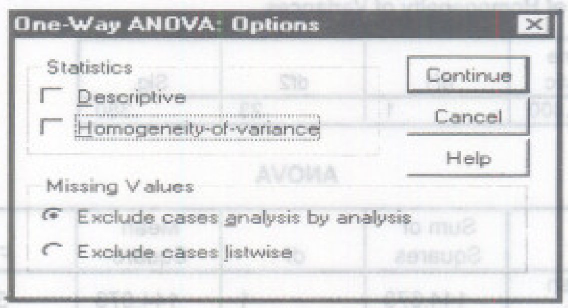
การเพิ่มเติมความสามารถในการคำนวณของคำสั่ง One-Way ANOVA..

1. การคำนวณค่าสถิติเบื้องต้น และ การหาช่วงความเชื่อมั่น
2. การทดสอบว่าค่าความแปรปรวนของประชากรเท่ากันหรือไม่

จากขั้นที่ 4. เมื่อเลือกตัวแปรเสร็จแล้ว

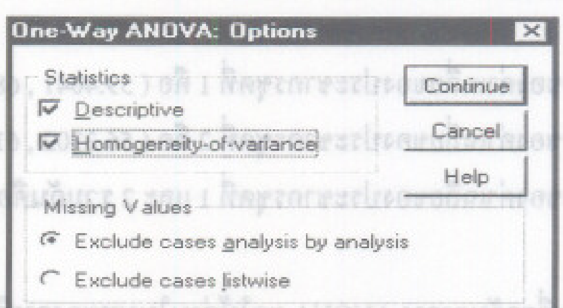


ขั้นที่ 4.1 ให้กดปุ่ม Options... จะได้เมนู ย่อยของ One-Way ANOVA : Options ดังนี้



ถ้าต้องการคำนวณค่าสถิติเบื้องต้นและการหาช่วงความเชื่อมั่นให้คลิกที่ Descriptive

ถ้าต้องการทดสอบว่าค่าความแปรปรวนเท่ากันหรือไม่ ให้คลิกที่ Homogeneity-of-variance



ขั้นที่ 4.2 เสร็จแล้วคลิก Continue

ขั้นที่ 4.3 คลิก OK ที่เมนูของ One-Way ANOVA จะได้ผลการคำนวณใหม่เป็นดังนี้

ผลการคำนวณใหม่บน SPSS Output Navigator ที่ได้คือ

Oneway

Descriptives

		X		
		CODE		
		1	2	Total
N		9	16	25
Mean		64.0067	58.9950	60.7992
Std. Deviation		5.9877	5.0008	5.7973
Std. Error		1.9959	1.2502	1.1595
95% Confidence Interval for Mean	Lower Bound	59.4041	56.3302	58.4062
	Upper Bound	68.6092	61.6598	63.1922
Minimum		57.76	52.02	52.02
Maximum		71.94	67.48	71.94

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
X	.800	1	23	.380

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
X	Between Groups	144.673	1	144.673	5.027	.035
	Within Groups	661.949	23	28.780		
	Total	806.621	24			

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 1 คือ (59.4041 , 68.6092)

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 2 คือ (56.3302 , 61.6598)

ช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยของประชากรชุดที่ 1 และ 2 รวมกันคือ (58.4062 , 63.1922)

หมายเหตุ การสรุปผลเกี่ยวกับตาราง ANOVA ขอให้อ่านในบทของ**การวิเคราะห์ความแปรปรวน**

การทดสอบสมมติฐานว่าความแปรปรวนของประชากรเท่ากันหรือไม่โดยใช้ผลการคำนวณจากตาราง Test of Homogeneity of Variance ขอให้อ่านในบทของ**การทดสอบสมมติฐาน**

แบบฝึกหัด 6.

1. ในการวัดเส้นผ่านศูนย์กลางของตัวอย่างขนาด 40 ของลูกปืนที่ผลิตได้ข้อมูลดังนี้

0.84	0.83	0.79	0.82	0.85	0.86	0.80	0.87	0.79	0.89
0.88	0.77	0.76	0.89	0.76	0.82	0.79	0.82	0.86	0.83
0.81	0.77	0.78	0.88	0.88	0.86	0.83	0.80	0.85	0.85
0.88	0.83	0.85	0.84	0.82	0.81	0.81	0.85	0.87	0.89

จงหาช่วงความเชื่อมั่น 95% ของค่าเฉลี่ยเส้นผ่านศูนย์กลางของลูกปืน

2. สุ่มเลือกตัวอย่างหลอดไฟ 25 หลอดเพื่อวัดอายุการใช้งาน(หน่วยชั่วโมง) ได้ข้อมูลดังนี้
1212 , 1275 , 1556 , 1538 , 1126 , 1747 , 1721 , 1211 , 1225 , 1368 , 1520 , 1352 , 1582 ,
1340 , 1302 , 1363 , 1652 , 1308 , 1089 , 1118 , 1777 , 1365 , 1564 , 1541 , 1402
จงหาช่วงความเชื่อมั่น 99% ของอายุเฉลี่ยของหลอดไฟ

3. ในการหาช่วงความเชื่อมั่น 95% ของปริมาณเฉลี่ยของเครื่องดื่มที่ขาย
จึงทำการสุ่มตัวอย่างของเครื่องดื่ม 36 ถ้วยได้ข้อมูลเป็นดังนี้

7.18	8.06	7.62	7.39	7.64	7.02	6.99	7.94	6.69	7.23
6.98	6.49	7.84	8.02	7.31	7.30	7.56	7.67	7.35	7.17
8.02	7.88	7.32	6.65	7.94	7.47	7.98	8.07	7.42	7.41
7.40	6.41	6.60	7.82	6.44	7.96				

จงหาช่วงความเชื่อมั่น 90% ของปริมาณเฉลี่ยของเครื่องดื่ม

4. ในการประมาณความสูงเฉลี่ยของนักศึกษา จึงทำการสุ่มเลือกมา 30 คน ได้ข้อมูลเป็นดังนี้

65.43	68.17	72.90	67.77	67.31	71.71	73.38	70.66	67.68	65.98
68.89	64.43	67.74	64.20	66.90	70.41	71.80	70.19	64.52	70.38
66.61	66.84	69.39	68.10	68.28	65.52	71.15	71.93	71.87	64.84

จงหาช่วงความเชื่อมั่น 99% ของส่วนสูงเฉลี่ยของนักศึกษา

5. สุ่มตัวอย่างจากหลอดไฟ 2 รุ่นๆ ละ 25 หลอดเพื่อตรวจสอบอายุการใช้งาน ได้ข้อมูลดังนี้
ข้อมูลรุ่น A 1884 , 1405 , 1552 , 1457 , 1611 , 1371 , 1794 , 1525 , 1294 , 1325 , 1867
1409 , 1819 , 1826 , 1719 , 1814 , 1327 , 1477 , 1448 , 1828 , 1698 , 1706 , 1572 , 1293 , 1476

ข้อมูลรุ่น B 1572 , 1347 , 1362 , 1225 , 1724 , 1301 , 1752 , 1324 , 1238 , 1142 , 1513 1261 , 1622 , 1603 , 1754 , 1666 , 1402 , 1283 , 1420 , 1474 , 1678 , 1437 , 1122 , 1190 , 1213

ภายใต้เงื่อนไขว่าความแปรปรวนของประชากรทั้ง 2 ชุด<u>เท่ากัน</u>

5.1 จงหาช่วงความเชื่อมั่น 90% ของอายุเฉลี่ยของหลอดไฟฟ้า

ภายใต้เงื่อนไขว่าความแปรปรวนของประชากรทั้ง 2 ชุด**<u>ไม่เท่ากัน</u>**

5.2 จงหาช่วงความเชื่อมั่น 90% ของอายุเฉลี่ยของหลอดไฟฟ้า

6. เพื่อจะเปรียบเทียบผลผลิตของพันธุ์ข้าวโพด 2 ชนิด ใช้ที่ดินแปลงละ 50 ไร่ 2 แปลง นำข้าวโพดทั้ง 2 ชนิดไปปลูกภายใต้สภาพการณ์เหมือนกัน ปรากฏว่า

ข้าวโพดพันธุ์ ก. จากแปลงที่ 1ให้ผลผลิต(หน่วยเป็นถัง)แต่ละไร่เป็นดังนี้

72.09 , 77.66 , 85.19 , 79.49 , 76.05 , 68.88 , 85.28 , 73.31 , 83.56 , 73.13 , 71.27 , 84.70 , 78.47

77.37 , 69.46 , 86.16 , 74.47 , 84.14 , 68.51 , 76.58 , 79.77 , 88.27 , 78.70 , 77.84 , 77.05 , 69.45

85.89 , 78.26 , 85.74 , 76.16 , 76.66 , 81.94 , 82.22 , 86.73 , 72.95 , 70.13 , 79.91 , 74.24 , 79.83

79.10 , 84.36 , 84.10 , 81.64 , 86.79 , 75.46 , 69.13 , 70.70 , 74.66 , 80.78 , 80.79

ข้าวโพดพันธุ์ ข. จากแปลงที่ 2ให้ผลผลิต(หน่วยเป็นถัง)แต่ละไร่เป็นดังนี้

80.43 , 86.15 , 92.92 , 87.88 , 76.34 , 85.06 , 97.68 , 86.22 , 90.72 , 82.11 , 75.70 , 93.03 , 94.89

83.74 , 80.63 , 91.44 , 79.93 , 94.79 , 81.68 , 86.69 , 93.82 , 93.53 , 89.38 , 86.73 , 85.81 , 76.77

96.28 , 78.68 , 92.99 , 85.11 , 78.26 , 87.15 , 94.32 , 87.78 , 75.69 , 83.57 , 98.27 , 83.35 , 91.08

89.34 , 89.59 , 97.34 , 97.34 , 87.72 , 84.77 , 77.70 , 84.31 , 86.74 , 88.12 , 90.42

ภายใต้เงื่อนไขว่าความแปรปรวนของประชากรทั้ง 2 ชุด<u>เท่ากัน</u>

6.1 จงหาช่วงความเชื่อมั่น 99% ของความแตกต่างของผลผลิตเฉลี่ย

ภายใต้เงื่อนไขว่าความแปรปรวนของประชากรทั้ง 2 ชุด**<u>ไม่เท่ากัน</u>**

6.2 จงหาช่วงความเชื่อมั่น 99% ของความแตกต่างของผลผลิตเฉลี่ย

7. มีผู้กล่าวว่า แผนควบคุมอาหารแบบใหม่สามารถลดน้ำหนักได้โดยเฉลี่ย 10 ปอนด์ ในสองสัปดาห์ จึงทำการสุ่มตัวอย่างของผู้หญิง 7 คนซึ่งปฏิบัติตามแผนควบคุมอาหาร ได้ข้อมูลดังนี้

คนที่	1	2	3	4	5	6	7
น้ำหนักก่อนคุมอาหาร	129	133	136	152	141	138	125
น้ำหนักหลังคุมอาหาร	130	121	128	137	129	132	120

จงหาช่วงความเชื่อมั่น 90% ของผลต่างของน้ำหนักก่อนและหลังคุมอาหาร

บทที่ 7

การทดสอบสมมติฐาน

การทำงานทางด้านสถิตินอกจากจะมีเรื่องการการหาช่วงความเชื่อมั่น แล้วยังมีการทดสอบสมมติฐานต่างๆ เช่น

การทดสอบสมมติฐานว่า $\mu = \mu_0$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\mu_1 = \mu_2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\mu_D = 0$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_0^2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$ จริงหรือไม่

การทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2$ จริงหรือไม่

การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันจริงหรือไม่

การทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงตามที่คาดไว้จริงหรือไม่

ในบทที่ 7 นี้เราจึงศึกษาเกี่ยวกับการทดสอบสมมติฐานแบบต่างๆ ทั้งหลักการขั้นตอนการทำงานทางทฤษฎี และ การนำ SPSS for Windows เข้ามาช่วยในการคำนวณ

7.1 การทดสอบสมมติฐานว่า $\mu = \mu_0$ จริงหรือไม่

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = \mu_0$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq \mu_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าเฉลี่ยและค่าความแปรปรวนของตัวอย่าง $\overline{x}, s^2$

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ T

ขั้นที่ 5. คำนวณค่าสถิติ $z_{\text{คำนวณ}}$ หรือ $t_{\text{คำนวณ}}$ ตามที่เลือกในขั้นตอนที่ 4 จากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$

บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า T ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ $\quad df = n - 1$

บริเวณวิกฤตคือ $T < -t_{\frac{\alpha}{2}}$ หรือ $T > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า

7.1 กรณีใช้ค่า Z ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

7.2 กรณีใช้ค่า T ปฏิเสธ H_0 ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$

เกณฑ์การเลือกค่าสถิติที่เหมาะสมของขั้นที่ 4. (ในทางทฤษฎี)

1. กรณีประชากรมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2 เลือก $Z = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{\sigma}{\sqrt{n}}\right)}$

2. กรณีประชากรมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

2.1 ตัวอย่างขนาด $n \geq 30$ แทนค่า σ ด้วย s เลือก $Z = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{s}{\sqrt{n}}\right)}$

2.2 ตัวอย่างขนาด $n < 30$ เลือก $T = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{s}{\sqrt{n}}\right)}$

3. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ^2

ตัวอย่างขนาด $n \geq 30$ เลือก $Z = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{\sigma}{\sqrt{n}}\right)}$

4. กรณีไม่ได้กำหนดว่าข้อมูลมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ^2

ต้องใช้ตัวอย่างขนาด $n \geq 30$ แทนค่า σ ด้วย s เลือก $Z = \dfrac{\bar{x} - \mu_0}{\left(\dfrac{s}{\sqrt{n}}\right)}$

หลักการและขั้นตอนของการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = \mu_0$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq \mu_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n คำนวณค่าเฉลี่ยและค่าความแปรปรวนของตัวอย่าง $\bar{x}, s^2$

ขั้นที่ 4. SPSS for Windows เลือกค่าสถิติ T เท่านั้น

ขั้นที่ 5. คำนวณค่าสถิติ $t_{คำนวณ} = \dfrac{\bar{x} - \mu_0}{(\frac{s}{\sqrt{n}})}$ และ องศาความอิสระ df = n – 1

ขั้นที่ 6. คำนวณค่า Sig (2-tailed) ของค่าสถิติ $t_{คำนวณ}$

Sig (2-tailed) = 2 เท่าของพื้นที่ใต้โค้งทางหางด้านขวาที่ระยะ $| t_{คำนวณ} |$

$= 2\,P(T > | t_{คำนวณ} |)$

ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ T จากตัวอย่าง กับ ค่าวิกฤตจากตาราง

โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{คำนวณ} < -t_{\frac{\alpha}{2}}$ หรือ $t_{คำนวณ} > t_{\frac{\alpha}{2}}$ แล้ว ปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig(2 - tailed) กับ ระดับนัยสำคัญ α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig(2 - tailed) < α แล้ว ปฏิเสธ H_0

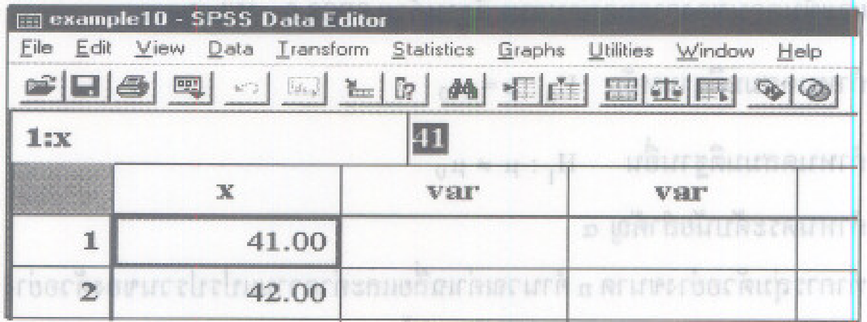
ตัวอย่าง 7.1.1. เท่าที่ผ่านมานิสิตใช้เวลาลงทะเบียนโดยเฉลี่ย 50 นาที ขณะนี้มหาวิทยาลัยกำลังทดลองให้นิสิตลงทะเบียนเรียนโดยใช้ระบบคอมพิวเตอร์ เพื่อทดสอบว่าเวลาเฉลี่ยที่ใช้ในการลงทะเบียนโดยใช้ระบบคอมพิวเตอร์มีค่าเท่ากับ 50 นาทีหรือไม่ จึงทำการสุ่มตัวอย่างการลงทะเบียนของนิสิต 12 คนได้ข้อมูลดังนี้ 41 , 42 , 47 , 41 , 54 , 26 , 26 , 65 , 34 , 49 , 29 , 50 กำหนดระดับนัยสำคัญ 0.05

วิธีทำ ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = 50$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq 50$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

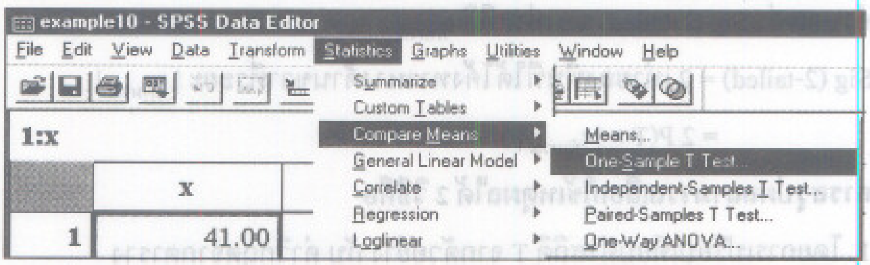
ขั้นที่ 3. นำข้อมูลเข้าสู่ SPSS



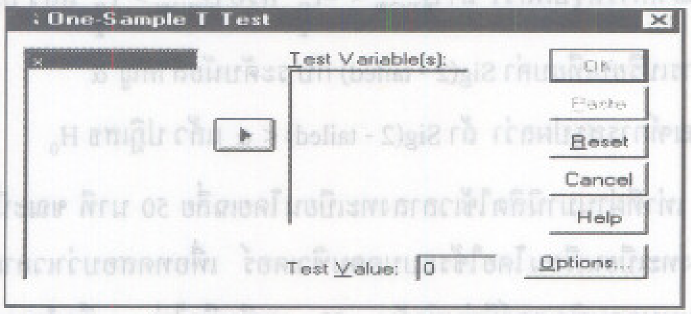
(**หมายเหตุ** แฟ้มข้อมูลนี้ชื่อ example10.sav)

ขั้นที่ 4. เลือกค่าสถิติ T

ขั้นที่ 4.1 วิเคราะห์ข้อมูลด้วยคำสั่ง Statistics / Compare Means / One Sample T Test...

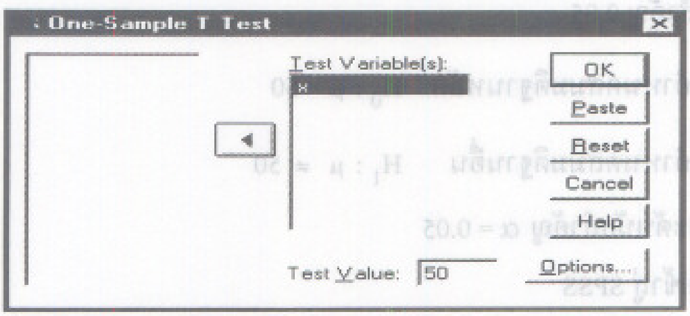


ขั้นที่ 4.2 คลิกที่ One Sample T Test..จะได้ผลบนจอภาพเป็น

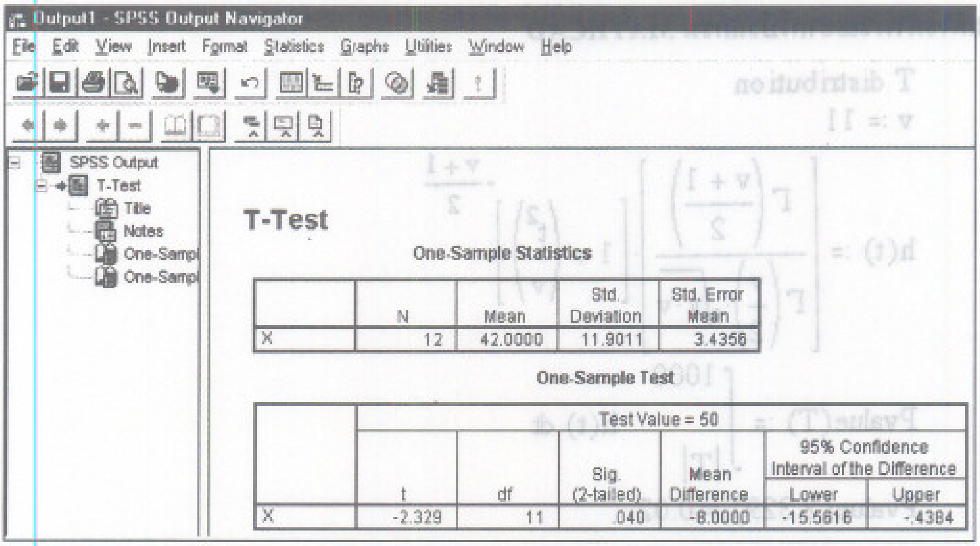


ขั้นที่ 4.3 เลือกตัวแปร x มาไว้ที่ Test Variable(s)

และ เลือกค่า Test Values เป็น 50 ตามค่าที่เราต้องการทดสอบ



ขั้นที่ 4.3 คลิก OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
X	12	42.0000	11.9011	3.4356

One-Sample Test

	Test Value = 50					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
X	-2.329	11	.040	-8.0000	-15.5616	-.4384

ขั้นที่ 5. คำนวณค่าสถิติ $T = \frac{\bar{X}-\mu_0}{(\frac{s}{\sqrt{n}})}$ และ องศาความอิสระ df = 12 − 1 = 11

ผลการคำนวณของ SPSS ได้ว่า ค่าสถิติ $t_{\text{คำนวณ}} = -2.329$ องศาความอิสระ = 11

ขั้นที่ 6. คำนวณค่า Sig (2-tailed) ของค่าสถิติ T

ผลการคำนวณของ SPSS ได้ว่า Sig (2-tailed) = 0.040

หมายเหตุ ตรวจสอบการคำนวณด้วย MATHCAD จะเห็นได้ว่าค่า

Sig (2-tailed) = 0.040 มาจาก 2 เท่าของ Pvalue(– 2.329)

ผลการคำนวณจากโปรแกรม MATHCAD

T distribution

$v := 11$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$$

$$\mathrm{Pvalue}(T) := \int_{|T|}^{1000} h(t)\,dt$$

$$\mathrm{Pvalue}(2.329) = 0.02$$

ขั้นที่ 7. สรุปผล

แบบที่ 1. โดยการเปรียบเทียบค่าสถิติ $t_{\text{คำนวณ}} = -2.329$ กับ ค่าวิกฤตจากตาราง

จากตารางสถิติจะได้ค่า $t_{0.025,df=11} = 2.201$

เพราะว่า $H_1 : \mu \neq 50$ เพราะฉะนั้นบริเวณวิกฤตคือ $T < -2.201$ หรือ $T > 2.201$

สรุป ปฏิเสธ H_0

หรือ **แบบที่ 2.** โดยการเปรียบเทียบค่า Sig (2-tailed) กับ ระดับนัยสำคัญ α

เพราะว่า Sig (2-tailed) = $0.04 < 0.05 = \alpha$

สรุปปฏิเสธ H_0

สรุปผลการทดสอบสมมติฐานจากข้อมูลที่เก็บมาได้ ต้องปฏิเสธสมมติฐานที่กล่าวว่าเวลาเฉลี่ยที่ใช้ในการลงทะเบียนโดยใช้ระบบคอมพิวเตอร์มีค่าเท่ากับ 50 นาที ที่ระดับนัยสำคัญ 0.05

7.2 การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีที่ประชากร 2 ชุดเป็นอิสระต่อกัน

หลักการและขั้นตอนการทำงานทางทฤษฎีเกี่ยวกับการทดสอบสมมติฐาน

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 หาค่าเฉลี่ยของตัวอย่าง $\bar{x}_1$

สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2 หาค่าเฉลี่ยของตัวอย่าง $\bar{x}_2$

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ T

ขั้นที่ 5. คำนวณค่าสถิติ $z_{\text{คำนวณ}}$ หรือ $t_{\text{คำนวณ}}$ ตามที่เลือกในขั้นที่ 4 จากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$

บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า T ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$

บริเวณวิกฤตคือ $T < -t_{\frac{\alpha}{2}}$ หรือ $T > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า

7.1 กรณีใช้ค่า Z ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

7.2 กรณีใช้ค่า T ปฏิเสธ H_0 ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$

เกณฑ์การเลือกค่าสถิติที่เหมาะสมของขั้นที่ 4.

1. กรณี $n_1 \geq 30$ และ $n_2 \geq 30$

1.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

เลือกใช้ค่าสถิติ $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

1.2. กรณีประชากร 2 ชุดมีการแจกแจงปกติและไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และประมาณ σ_1^2 ด้วย s_1^2

หาค่าความแปรปรวนของตัวอย่าง s_2^2 และประมาณ σ_2^2 ด้วย s_2^2

เลือกใช้ค่าสถิติ $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

2. กรณี $n_1 < 30$ หรือ $n_2 < 30$

2.1 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

เลือกใช้ค่าสถิติ $Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

2.2 กรณีประชากร 2 ชุดมีการแจกแจงปกติ และ ไม่รู้ค่าความแปรปรวน σ_1^2 และ σ_2^2

หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

2.2.1 **ภายใต้ข้อกำหนด** $\sigma_1^2 = \sigma_2^2$

เลือกใช้ค่าสถิติ $T = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

เมื่อ $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ และ $df = n_1 + n_2 - 2$

2.2.2 **ภายใต้ข้อกำหนด** $\sigma_1^2 \neq \sigma_2^2$

เลือกใช้ค่าสถิติ $T = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

เมื่อ $df = \dfrac{(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2})^2}{(\dfrac{s_1^2}{n_1})^2 \dfrac{1}{(n_1 - 1)} + (\dfrac{s_2^2}{n_2})^2 \dfrac{1}{(n_2 - 1)}}$

หลักการและขั้นตอนการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n_1 จากประชากรชุดที่ 1 และ สุ่มตัวอย่างขนาด n_2 จากประชากรชุดที่ 2 หาค่าเฉลี่ยของตัวอย่าง $\bar{x}_1$ และ $\bar{x}_2$ หาค่าความแปรปรวนของตัวอย่าง s_1^2 และ s_2^2

ขั้นที่ 4. เลือกค่าสถิติ T เท่านั้น

ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

เลือกใช้ค่าสถิติ $T = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

เมื่อ $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ และ $df = n_1 + n_2 - 2$

ภายใต้ข้อกำหนด $\sigma_1^2 \neq \sigma_2^2$

เลือกใช้ค่าสถิติ $T = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

เมื่อ $df = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\left(\dfrac{s_1^2}{n_1}\right)^2 \dfrac{1}{(n_1 - 1)} + \left(\dfrac{s_2^2}{n_2}\right)^2 \dfrac{1}{(n_2 - 1)}}$

ขั้นที่ 5. คำนวณค่าสถิติ T ซึ่งเราเรียกว่า $t_{\text{คำนวณ}}$ และ องศาความอิสระ df

ขั้นที่ 6. คำนวณค่า Sig (2–tailed) ของค่าสถิติ $t_{\text{คำนวณ}}$

Sig (2–tailed) = 2 เท่าของพื้นที่ใต้โค้ง ทางหางด้านขวาที่ระยะ $|\,t_{\text{คำนวณ}}\,|$

$= 2\,P(\,T > |\,t_{\text{คำนวณ}}\,|\,)$

ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ T จากตัวอย่าง กับ ค่าวิกฤตจากตาราง

โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้ว ปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig(2 – tailed) กับ ระดับนัยสำคัญ α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig(2 – tailed) $< \alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 7.2.1 โรงงานผลิตแป้งกระป๋องมีเครื่องจักร 2 เครื่อง

ให้ μ_1 เป็นค่าเฉลี่ยประชากรของน้ำหนักของแป้งที่บรรจุในกระป๋องที่ผลิตจากเครื่องจักรที่ 1

μ_2 เป็นค่าเฉลี่ยประชากรของน้ำหนักของแป้งที่บรรจุในกระป๋องที่ผลิตจากเครื่องจักรที่ 2

เพื่อทำการทดสอบว่าค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องเท่ากันหรือไม่ จึงทำการสุ่มตัวอย่างแป้งกระป๋องจากเครื่องจักรเครื่องที่ 1 และ 2 มาอย่างละ 100 กระป๋อง ข้อมูลเก็บไว้ที่แฟ้มข้อมูลชื่อ example11.sav ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$ จงทำการทดสอบสมมติฐานว่าค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องเท่ากันหรือไม่ ที่ระดับนัยสำคัญ 0.01

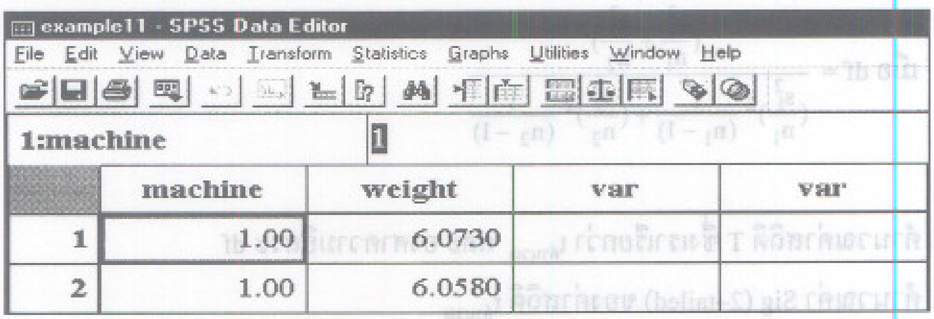
วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.01$

ขั้นที่ 3. นำข้อมูลเข้าสู่ SPSS



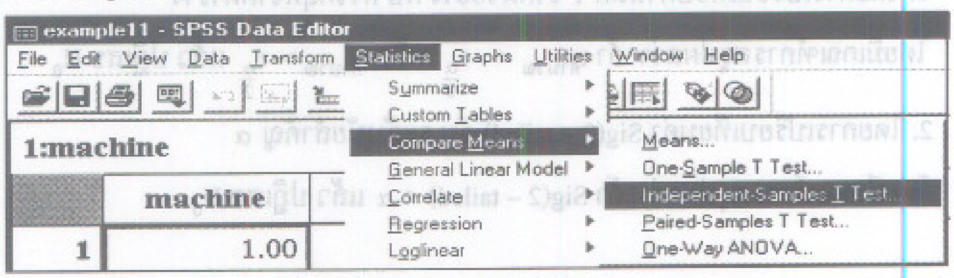
หมายเหตุ machine เป็นตัวแปรจำแนกกลุ่มของโรงงาน

weight เป็นตัวแปรเก็บค่าน้ำหนักแป้ง

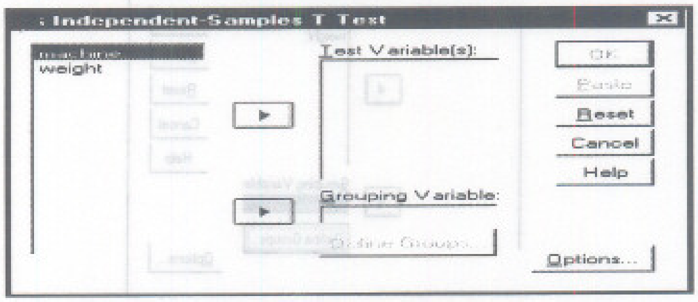
ขั้นที่ 4. เลือกค่าสถิติ T เท่านั้น

ขั้นที่ 4.1 การเข้าสู่การวิเคราะห์ข้อมูลด้วยคำสั่ง

Statistics / Compare Means / Independent Sample T Test...

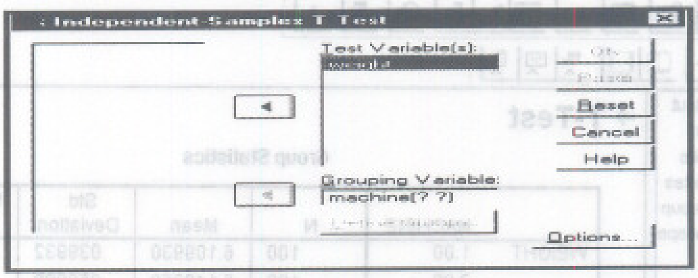


ขั้นที่ 4.2 คลิกที่ Independent Sample T Test..จะได้ผลบนจอภาพเป็น

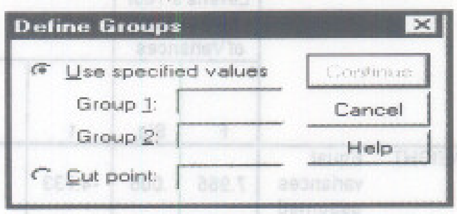


ขั้นที่ 4.3 เลือกตัวแปร weight มาไว้ที่ Test Variable(s)

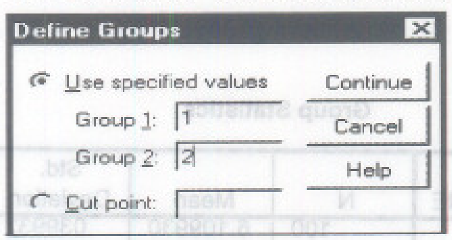
และ เลือกตัวแปร machine มาไว้ที่ Group Variable(s)



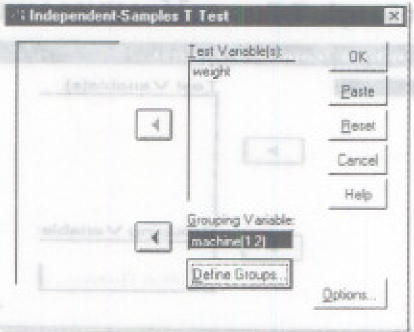
ขั้นที่ 4.4 คลิกที่ช่อง Group Variable(s) และ คลิกที่ Define Groups จอภาพจะมีเมนูย่อยเป็น



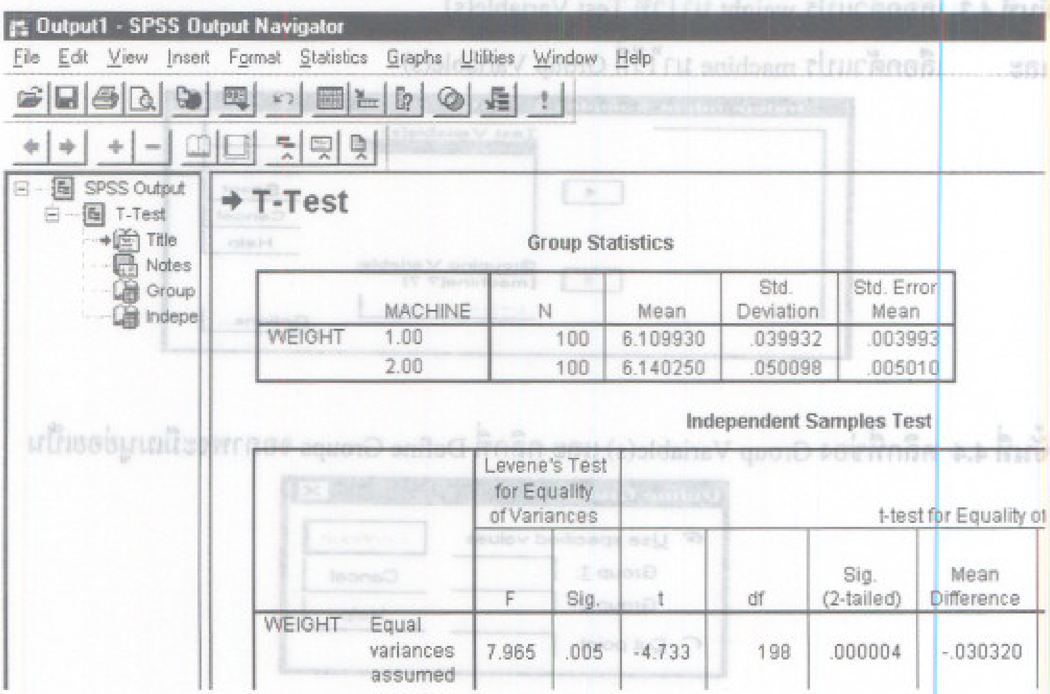
ขั้นที่ 4.5 การเลือกกลุ่มเพื่อทดสอบ ให้พิมพ์ 1 ในช่อง Group 1. และ พิมพ์ 1 ในช่อง Group 2.



ขั้นที่ 4.6 กด Continue จะกลับมาเมนูย่อย



ขั้นที่ 4.7 คลิก OK จะได้ผลการคำนวณที่ SPSS Output Navigator เป็น



การคำนวณอย่างละเอียดคือ

T-Test

Group Statistics

	MACHINE	N	Mean	Std. Deviation	Std. Error Mean
WEIGHT	1.00	100	6.109930	.039932	.003993
	2.00	100	6.140250	.050098	.005010

Independent Samples Test

			WEIGHT	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F		7.965	
	Sig.		.005	
t-test for Equality of Means	t		-4.733	-4.733
	df		198	188.620
	Sig. (2-tailed)		.0000042	.0000043
	Mean Difference		-.0303200	-.0303200
	Std. Error Difference		.0064065	.0064065
	95% Confidence Interval of the Mean	Lower	-.0429538	-.0429577
		Upper	-.0176862	-.0176823

ภายใต้ข้อกำหนด $\sigma_1^2 = \sigma_2^2$

ขั้นที่ 5. $T = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ เมื่อ $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ และ $df = n_1 + n_2 - 2$

จากการคำนวณของ SPSS ได้ว่า $t_{\text{คำนวณ}} = -4.733$, $df = 198$

ขั้นที่ 6. คำนวณค่า Sig(2–tailed) ของค่าสถิติ T

Sig (2–tailed) = 0.0000042

ขั้นที่ 7. สรุปผล

1. โดยการเปรียบเทียบค่าสถิติ $t_{\text{คำนวณ}}$ จากตัวอย่าง กับ ค่าวิกฤตจากตาราง

จากตารางสถิติจะได้ค่า $t_{0.005,df=198} = 2.6008873$

เพราะว่ากำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

เพราะฉะนั้นบริเวณวิกฤตคือ $T < -2.6008873$ หรือ $T > 26008873$ สรุปปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig (2–tailed) กับ ระดับนัยสำคัญ α

เพราะว่ากำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

เพราะฉะนั้นต้องเปรียบเทียบค่า $\alpha = 0.01$ กับค่าของ Sig (2–tailed)

เพราะว่า Sig (2–tailed) = $0.0000042 < 0.01 = \alpha$ สรุปปฏิเสธ H_0

ผลสรุปของการทดสอบคือ ค่าเฉลี่ยของน้ำหนักแป้งในกระป๋องไม่เท่ากันที่ระดับนัยสำคัญ 0.01

เสริมความรู้ของการคำนวณค่าทางสถิติด้วย MATHCAD

$$x = \begin{array}{c|c} & 1 \\ \hline 1 & 6.073 \\ 2 & 6.058 \\ 3 & 6.102 \\ 4 & 6.126 \\ 5 & 6.138 \end{array} \quad x = \begin{array}{c|c} & 1 \\ \hline 96 & 6.125 \\ 97 & 6.098 \\ 98 & 6.126 \\ 99 & 6.16 \\ 100 & 6.123 \end{array} \quad y = \begin{array}{c|c} & 1 \\ \hline 1 & 6.094 \\ 2 & 6.075 \\ 3 & 6.13 \\ 4 & 6.16 \\ 5 & 6.175 \end{array} \quad y = \begin{array}{c|c} & 1 \\ \hline 96 & 6.159 \\ 97 & 6.125 \\ 98 & 6.16 \\ 99 & 6.203 \\ 100 & 6.156 \end{array}$$

$xbar1 := \text{mean}(x)$ $\qquad xbar1 = 6.10993$

$xbar2 := \text{mean}(y)$ $\qquad xbar2 = 6.14025$

$n1 := \text{length}(x)$ $\qquad n1 = 100$

$n2 := \text{length}(y)$ $\qquad n2 = 100$

$s1 := \sqrt{\dfrac{n1 \cdot \text{var}(x)}{n1 - 1}}$ $\qquad s1 = 0.039932$

$s2 := \sqrt{\dfrac{n2 \cdot \text{var}(y)}{n2 - 1}}$ $\qquad s2 = 0.050098$

Equal variances assumed

$$sp := \sqrt{\frac{(n1 - 1)\cdot s1^2 + (n2 - 1)\cdot s2^2}{n1 + n2 - 2}}$$

$Std\_Error\_Difference := sp \cdot \sqrt{\dfrac{1}{n1} + \dfrac{1}{n2}}$ $\qquad Std\_Error\_Difference = 0.006407$

$Mean\_Difference := xbar1 - xbar$ $\qquad Mean\_Difference = -0.03032$

$t := \dfrac{(xbar1 - xbar2) - (0)}{sp \cdot \sqrt{\dfrac{1}{n1} + \dfrac{1}{n2}}}$ $\qquad t = -4.732663$

$df := n1 + n2 - 2$ $\qquad df = 198$

การหาค่า $t_{0.005, df = 198}$

$v := 198$

$TOL := 0.000001 \qquad h(t) := \left[\dfrac{\Gamma\left(\dfrac{v+1}{2}\right)}{\Gamma\left(\dfrac{v}{2}\right)\cdot\sqrt{\pi \cdot v}}\right]\cdot\left[1 + \left(\dfrac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$

$k := 2.60088 \qquad T(A) := \text{root}\left[A - \left(0.5 - \displaystyle\int_0^{k} h(t)\,dt\right), k\right] \qquad T(0.005) = 2.60088728$

เพราะฉะนั้น $t_{0.005, df = 198} = 2.60088728$

การหาค่า Sig.(2-tailed)

$$\text{Pvalue}(T) := 0.5 - \int_0^{|T|} h(t)\,dt \qquad \text{Pvalue}(-4.733) = 0.0000021$$

เพราะฉะนั้น Sig(2–tailed) = 2(0.0000021) = 0.0000042 ตามค่าในตารางจาก SPSS

Equal variances not assumed

$$\text{Std\_Error\_Difference} := \sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}} \qquad \text{Std\_Error\_Difference} = 0.006407$$

$$t := \frac{(xbar1 - xbar2) - (0)}{\sqrt{\frac{s1^2}{n1} + \frac{s2^2}{n2}}} \qquad t = -4.732663$$

$$df := \frac{\left(\frac{s1^2}{n1} + \frac{s2^2}{n2}\right)^2}{\left(\frac{s1^2}{n1}\right)^2 \cdot \left(\frac{1}{n1-1}\right) + \left(\frac{s2^2}{n2}\right)^2 \cdot \left(\frac{1}{n2-1}\right)} \qquad df = 188.620126$$

การหาค่า $t_{0.005, df = 188.62}$

$$v := 188.62 \quad TOL := 0.000001 \qquad k := 2.60088$$

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \cdot \sqrt{\pi \cdot v}}\right] \cdot \left[1 + \left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}} \qquad T(A) := \text{root}\left[A - \left(0.5 - \int_0^{|k|} h(t)\,dt\right), k\right]$$

$$T(0.005) = 2.60214332$$

เพราะฉะนั้น $t_{0.005, df = 188.620} = 2.60214332$

การหาค่า Sig.(2–tailed)

$$\text{Pvalue}(T) := 0.5 - \int_0^{|T|} h(t)\,dt \qquad \text{Pvalue}(-4.733) = 0.00000217$$

เพราะฉะนั้น Sig.(2–tailed) = 2(0.00000217) = 0.00000434 ตามค่าในตารางจาก SPSS

7.3 การทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ กรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติ กรณีที่ประชากร 2 ชุดไม่เป็นอิสระต่อกัน ในการทดสอบสมมติฐาน $H_0 : \mu_1 = \mu_2$ หรือ การทดสอบสมมติฐาน $H_0 : \mu_D = d_0$ มีขั้นตอนการทำงานดังนี้

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = d_0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq d_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n จากประชากรชุดที่ 1 และ ประชากรชุดที่ 2 ได้ข้อมูลเป็น

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
x_3	y_3
$\vdots$	$\vdots$
x_n	y_n

คำนวณ 1. คำนวณค่าผลต่างของตัวอย่าง $d_i = x_i - y_i \quad i = 1,2,\ldots,n$

2. คำนวณค่าเฉลี่ยของผลต่างของตัวอย่าง $\bar{d}$

3. คำนวณค่าส่วนเบี่ยงเบนมาตรฐานของผลต่างของตัวอย่าง s_d

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสม Z หรือ T

กรณี $n \geq 30$ เลือก $Z = \dfrac{\bar{d} - d_0}{\frac{s_d}{\sqrt{n}}}$

กรณี $n < 30$ และ ภายใต้ข้อสมมติว่า ผลต่างของข้อมูลมีการแจกแจงปกติ

เลือก $T = \dfrac{\bar{d} - d_0}{\frac{s_d}{\sqrt{n}}}$ เมื่อ $df = n - 1$

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

6.1 กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

6.2 กรณีใช้ค่า T ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $T < -t_{\frac{\alpha}{2}}$ หรือ $T > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า

7.1 กรณีใช้ค่า Z ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

7.2 กรณีใช้ค่า T ปฏิเสธ H_0 ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$

หลักการและขั้นตอนการทำงานของการวิเคราะห์ข้อมูลด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = d_0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq d_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. สุ่มตัวอย่างขนาด n คำนวณค่า $\overline{D}$ และ s_d

ขั้นที่ 4. เลือกค่าสถิติ T

ขั้นที่ 5. คำนวณค่าสถิติ T และ องศาความอิสระ df

$$T = \frac{\overline{D} - d_0}{\frac{s_d}{\sqrt{n}}} \quad \text{และ} \quad df = n - 1$$

เราเรียกค่า T ที่คำนวณได้ว่า $t_{\text{คำนวณ}}$

ขั้นที่ 6. คำนวณค่า Sig (2 – tailed) ของค่าสถิติ $t_{\text{คำนวณ}}$

Sig (2 – tailed) = 2 เท่าของพื้นที่ใต้โค้ง ทางหางด้านขวาที่ระยะ $|t_{\text{คำนวณ}}|$

$= 2P(T > |t_{\text{คำนวณ}}|)$

ขั้นที่ 7. การสรุปผลสามารถเลือกใช้เหตุผลได้ 2 วิธีคือ

1. โดยการเปรียบเทียบค่าสถิติ T จากตัวอย่าง กับ ค่าวิกฤตจากตาราง

โดยมีเกณฑ์การสรุปผลว่า ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$ แล้ว ปฏิเสธ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig(2 – tailed) กับ ระดับนัยสำคัญ α

โดยมีเกณฑ์การสรุปผลว่า ถ้า Sig(2 – tailed) < α แล้ว ปฏิเสธ H_0

ตัวอย่าง 7.3.1 จากตัวอย่างสารที่มีเหล็ก 5 ตัวอย่างนำมาวิเคราะห์หาปริมาณด้วยวิธีแบ่งออกเป็น 2 ตัวอย่างย่อยและใช้วิธีการวิเคราะห์หาปริมาณเหล็กปรากฏผลดังนี้

ตัวอย่างที่	1	2	3	4	5
วิธีวิเคราะห์ด้วยรังสีเอ็กซ์	2.0	2.0	2.3	2.1	2.4
วิธีวิเคราะห์ด้วยเคมี	2.2	1.9	2.5	2.3	2.4

ให้ μ_1 เป็นค่าเฉลี่ยประชากรของปริมาณเหล็กที่วิเคราะห์ด้วยรังสีเอ็กซ์

μ_2 เป็นค่าเฉลี่ยประชากรของปริมาณเหล็กที่วิเคราะห์ด้วยเคมี

สมมติว่าประชากรมีการแจกแจงปกติ

จงทดสอบว่าการทดสอบทั้งสองวิธีให้ผลทัดเทียมกันที่ระดับนัยสำคัญ 0.05

วิธีทำ การวิเคราะห์ด้วย SPSS

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_D = 0$

กำหนดสมมติฐานอื่น $H_1 : \mu_D \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

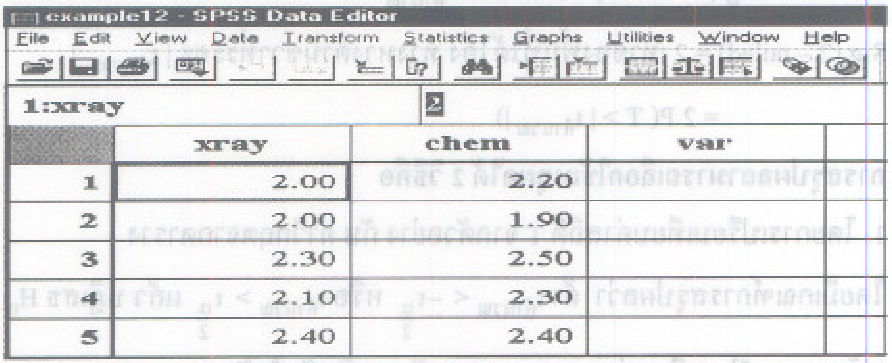
ขั้นที่ 3. สุ่มตัวอย่างขนาด n

ขั้นที่ 4. นำข้อมูลเข้ามาทำการวิเคราะห์ เพื่อหาค่าสถิติ T

ขั้นที่ 4.1 นำข้อมูลเข้าสู่ SPSS Data Editor

xray เป็นตัวแปร ปริมาณเหล็กที่วิธีวิเคราะห์ด้วยรังสีเอ็กซ์

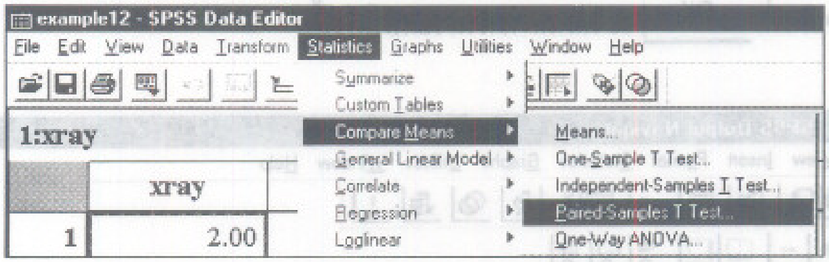
chem เป็นตัวแปร ปริมาณเหล็กที่ วิธีวิเคราะห์ด้วยเคมี



	xray	chem	var
1	2.00	2.20	
2	2.00	1.90	
3	2.30	2.50	
4	2.10	2.30	
5	2.40	2.40	

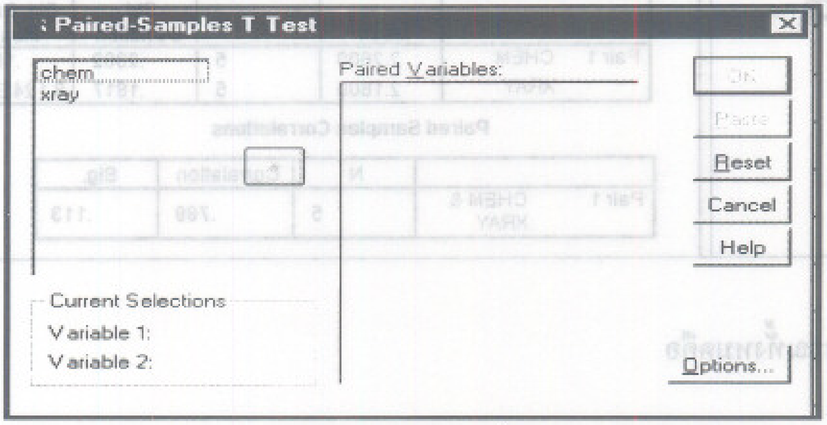
หมายเหตุ แฟ้มข้อมูลชื่อ example12.sav

ขั้นที่ 4.2 เลือกคำสั่ง Statistics / Compare Means / Paired – Samples T Test..



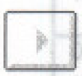
ขั้นที่ 4.3 คลิกที่ Paired – Samples T Test..

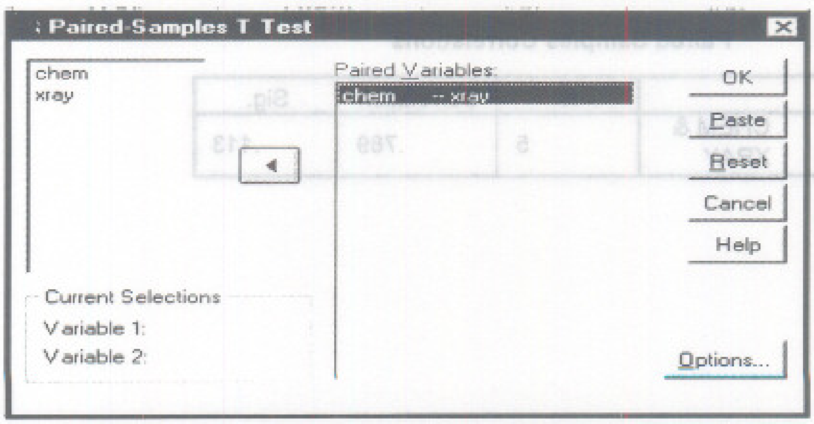
จอภาพจะขึ้นเมนูของคำสั่ง Statistics / Compare Means / Paired – Samples T Test..



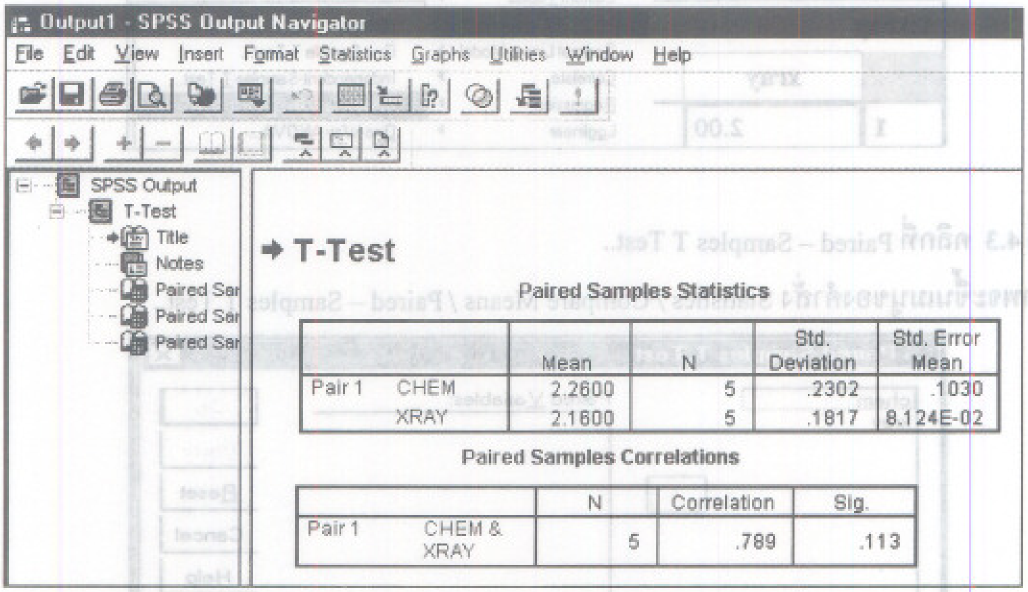
เลือกตัวแปร Variable 1 เป็น chem โดยการคลิกที่ตัวแปร chem

เลือกตัวแปร Variable 2 เป็น xray โดยการคลิกที่ตัวแปร xray

ขั้นที่ 4.4 คลิกที่  เพื่อย้ายคู่ของตัวแปร chem – xray มาไว้ที่ช่อง Paired–Variables



ขั้นที่ 4.5 คลิกที่ OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	CHEM	2.2600	5	.2302	.1030
	XRAY	2.1600	5	.1817	8.124E-02

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	CHEM & XRAY	5	.789	.113

Paired Samples Test

			Pair 1 CHEM - XRAY
Paired Differences	Mean		1.000E-01
	Std. Deviation		.1414
	Std. Error Mean		6.325E-02
	95% Confidence Interval of the Difference	Lower	-7.56E-02
		Upper	.2756
t			1.581
df			4
Sig. (2-tailed)			.189

ขั้นที่ 4. เลือกค่าสถิติ T

ขั้นที่ 5. คำนวณค่าสถิติ $t_{\text{คำนวณ}}$ และ ค่าองศาความอิสระ df

$t_{\text{คำนวณ}}$ = 1.581 และ df = 4

ขั้นที่ 6. คำนวณค่า Sig (2–tailed) ของค่าสถิติ T

Sig (2–tailed) = 0.189

ขั้นที่ 7. สรุปผล

1. จากการเปิดตาราง $t_{0.025,df=4} = 2.776$

   บริเวณวิกฤตคือ $T < -2.776$ หรือ $T > 2.776$

   เพราะว่า $t_{\text{คำนวณ}} = 1.581$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น สรุป ยอมรับ H_0

หรือ 2. โดยการเปรียบเทียบค่า Sig (2–tailed) กับ α

   เพราะว่า Sig (2–tailed) = $0.189 > 0.05 = \alpha$ เพราะฉะนั้น สรุป ยอมรับ H_0

หมายเหตุ การคำนวณด้วย MATHCAD

$$xray := \begin{bmatrix} 2.0 \\ 2.0 \\ 2.3 \\ 2.1 \\ 2.4 \end{bmatrix} \qquad chem := \begin{bmatrix} 2.2 \\ 1.9 \\ 2.5 \\ 2.3 \\ 2.4 \end{bmatrix}$$

$n := 5$

$\text{mean}(xray) = 2.16$ $\qquad$ $\text{mean}(chem) = 2.26$

$sdxray := \sqrt{\dfrac{n \cdot \text{var}(xray)}{n-1}}$ $\qquad$ $sdxray = 0.181659$

$sdchem := \sqrt{\dfrac{n \cdot \text{var}(chem)}{n-1}}$ $\qquad$ $sdchem = 0.230217$

$Correlation := \text{corr}(xray, chem)$ $\qquad$ $Correlation = 0.789076$

การคำนวณค่าผลต่าง d

$d := chem - xray$

$$d = \begin{bmatrix} 0.2 \\ -0.1 \\ 0.2 \\ 0.2 \\ 0 \end{bmatrix}$$

$dbar := \text{mean}(d)$ $\qquad$ $dbar = 0.1$

$sd := \sqrt{\dfrac{n \cdot \text{var}(d)}{n-1}}$ $\qquad$ $sd = 0.141421$

การคำนวณค่าสถิติ T

$t := \dfrac{dbar - 0}{\left(\dfrac{sd}{\sqrt{n}}\right)}$ $\qquad$ $t = 1.581139$

การคำนวณค่า Sig(2-Tailed)

$v := n - 1$ $\qquad$ $h(t) := \left[\dfrac{\Gamma\left(\dfrac{v+1}{2}\right)}{\Gamma\left(\dfrac{v}{2}\right) \cdot \sqrt{\pi \cdot v}}\right] \cdot \left[1 + \left(\dfrac{t^2}{v}\right)\right]^{-\frac{v+1}{2}}$

$Pvalue(T) := \displaystyle\int_{|T|}^{1000} h(t)\, dt$

$Pvalue(1.581) = 0.094517$

$Sig\_2tailed := 2 \cdot Pvalue(1.581)$

$Sig\_2tailed = 0.189034$

7.4 การทดสอบสมมติฐาน $H_0 : \sigma^2 = \sigma_0^2$

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma^2 = \sigma_0^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma^2 \neq \sigma_0^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n คำนวณค่าความแปรปรวนของตัวอย่าง s^2

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าสถิติไคสแควร์

$$\chi^2_{\text{คำนวณ}} = \frac{(n-1)s^2}{\sigma_0^2} \qquad df = n-1$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต

ค่าวิกฤตคือ $\chi^2_{1-\frac{\alpha}{2}}$ หรือ $\chi^2_{\frac{\alpha}{2}}$

บริเวณวิกฤตคือ $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ หรือ $\chi^2 > \chi^2_{\frac{\alpha}{2}}$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} < \chi^2_{1-\frac{\alpha}{2}}$ หรือ $\chi^2_{\text{คำนวณ}} > \chi^2_{\frac{\alpha}{2}}$

ตัวอย่าง 7.4.1 ผู้ผลิตอ้างว่าอายุการใช้งานของแบตเตอรีมีการแจกแจงปกติ และมีส่วนเบี่ยงเบนมาตรฐานเป็น 0.9 ปี เพื่อทดสอบคำกล่าวอ้างของผู้ผลิตจึงทำการสุ่มตัวอย่างแบตเตอรีออกมา 10 ลูกได้อายุใช้งานดังนี้

5.25 3.76 5.36 3.67 6.05 3.89 3.39 6.12 6.49 6.03

ท่านคิดว่า $\sigma^2 = 0.81$ ปีใช่หรือไม่ ที่ระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma^2 = 0.81$

กำหนดสมมติฐานอื่น $H_1 : \sigma^2 \neq 0.81$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด 10 คำนวณค่าความแปรปรวนของตัวอย่าง s^2

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \dfrac{(n-1)s^2}{\sigma_0^2}$

ผลการคำนวณด้วย MATHCAD

$\sigma := \sqrt{0.81}$

$$\text{data} := \begin{bmatrix} 5.25 \\ 3.76 \\ 5.36 \\ 3.67 \\ 6.05 \\ 3.89 \\ 3.39 \\ 6.12 \\ 6.49 \\ 6.03 \end{bmatrix}$$

$n := 10$ $\qquad s := \sqrt{\dfrac{n \cdot \text{var}(\text{data})}{n-1}}$

$\text{chisquare} := \dfrac{(n-1)\cdot s^2}{\sigma^2}$ $\qquad \text{chisquare} = 16.001$

เพราะฉะนั้น $\chi^2_{\text{คำนวณ}} = 16.001$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต df = 10 – 1 = 9 และ $\chi^2_{0.025} = 19.023$, $\chi^2_{0.975} = 2.7$

บริเวณวิกฤตคือ $\chi^2 < 2.7$ หรือ $\chi^2 > 19.023$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

เพราะว่า $\chi^2_{\text{คำนวณ}}$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ยอมรับ H_0

7.5 การทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2 คำนวณค่าความแปรปรวนของตัวอย่าง s_1^2, s_2^2

ขั้นที่ 4. เลือกค่าสถิติเอฟ F

ขั้นที่ 5. คำนวณค่าสถิติเอฟ $f_{คำนวณ} = \frac{s_1^2}{s_2^2}$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตเอฟ $v_1 = n_1 - 1$ และ $v_2 = n_2 - 1$

ค่าวิกฤตคือ $f_{1-\frac{\alpha}{2}}$ หรือ $f_{\frac{\alpha}{2}}$

บริเวณวิกฤตคือ $F < f_{1-\frac{\alpha}{2}}$ หรือ $F > f_{\frac{\alpha}{2}}$

ขั้นที่ 7. **สรุปผล** โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ปฏิเสธ H_0 ถ้า $f_{คำนวณ} < f_{1-\frac{\alpha}{2}}$ หรือ $f_{คำนวณ} > f_{\frac{\alpha}{2}}$

ตัวอย่าง 7.5.1 วัตถุ 5 ชิ้นได้รับการปฏิบัติแบบที่ 1 ได้ผลการทดลองเป็นดังนี้

1.024 0.972 1.004 0.986 1.015

วัตถุ 6 ชิ้นได้รับการปฏิบัติแบบที่ 2 ได้ผลการทดลองเป็นดังนี้

1.017 0.991 1.018 1.018 0.983 0.975

จงทดสอบสมมติฐานด้วยระดับนัยสำคัญ 0.1 ว่า $H_0 : \sigma_1^2 = \sigma_2^2$ แย้งกับ $H_1 : \sigma_1^2 \neq \sigma_2^2$

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.1$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2 คำนวณค่าความแปรปรวนของตัวอย่าง s_1^2, s_2^2

การคำนวณด้วย MATHCAD

$$x1 := \begin{bmatrix} 1.024 \\ 0.972 \\ 1.004 \\ 0.986 \\ 1.015 \end{bmatrix} \qquad x2 := \begin{bmatrix} 1.017 \\ 0.991 \\ 1.018 \\ 1.018 \\ 0.983 \\ 0.975 \end{bmatrix}$$

$n1 := \text{length}(x1)$ $\qquad$ $n2 := \text{length}(x2)$

$n1 = 5$ $\qquad$ $n2 = 6$

$$s1 := \sqrt{\frac{\text{var}(x1)\cdot n1}{n1-1}} \qquad s2 := \sqrt{\frac{\text{var}(x2)\cdot n2}{n2-1}}$$

ขั้นที่ 4. เลือกค่าสถิติเอฟ

ขั้นที่ 5. คำนวณค่าสถิติเอฟ $\qquad f_{\text{คำนวณ}} = \dfrac{s_1^2}{s_2^2}$

การคำนวณด้วย MATHCAD $\qquad \text{Fcompute} := \dfrac{s1^2}{s2^2}$

$\text{Fcompute} = 1.162927$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตเอฟ $v_1 = n_1 - 1 = 4$ และ $v_2 = n_2 - 1 = 5$

ค่าวิกฤตคือ $f_{0.05}(4,5) = 5.199$ และ $f_{0.95}(4,5) = \dfrac{1}{f_{0.05}(5,4)} = \dfrac{1}{6.26} = 0.1597444$

บริเวณวิกฤตคือ $F < 0.1597444$ หรือ $F > 5.199$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

เพราะว่า $0.1597444 < f_{\text{คำนวณ}} < 5.199$ เพราะฉะนั้น ยอมรับ H_0

หลักการและขั้นตอนของการวิเคราะห์ด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก $\quad H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $\quad H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2 จากประชากรแต่ละชุด

ขั้นที่ 4. เลือกค่าสถิติ Levene

ขั้นที่ 5. คำนวณค่าสถิติ Levene และค่า Sig

ขั้นที่ 6. ทำการวิเคราะห์ข้อมูล

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่า Sig กับ ค่านัยสำคัญ α

ปฏิเสธ H_0 ถ้า Sig $< \alpha$

จากตัวอย่าง 7.5.1 ต้องสร้างแฟ้มข้อมูลแบบ 2 ตัวแปรคือ ตัวแปร code จำแนกกลุ่มประชากร และตัวแปร x เก็บข้อมูลที่วัดได้จากการทดลอง

การทดสอบสมมติฐานด้วยระดับนัยสำคัญ 0.1 ว่า $H_0 : \sigma_1^2 = \sigma_2^2$ แย้งกับ $H_1 : \sigma_1^2 \neq \sigma_2^2$

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \sigma_1^2 = \sigma_2^2$

กำหนดสมมติฐานอื่น $H_1 : \sigma_1^2 \neq \sigma_2^2$

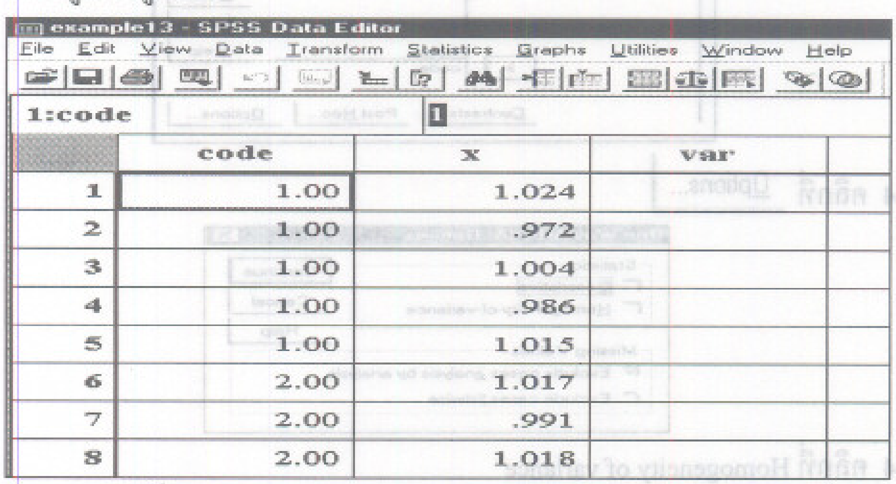
ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.1$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n_1 และ n_2

ขั้นที่ 4. เลือกค่าสถิติ Levene

ขั้นที่ 5. คำนวณค่าสถิติ Levene และค่า Sig

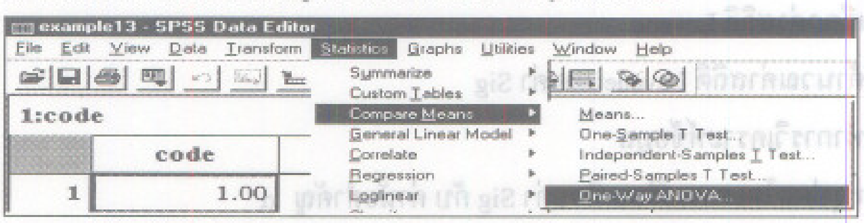
ขั้นที่ 6. นำข้อมูลเข้าสู่ SPSS



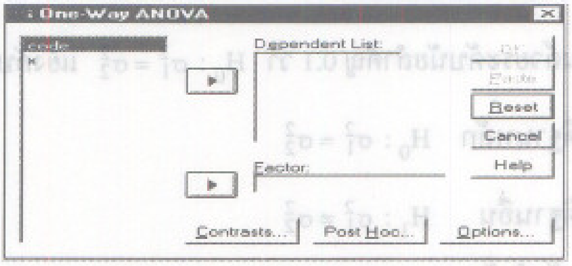
	code	x	var
1	1.00	1.024	
2	1.00	.972	
3	1.00	1.004	
4	1.00	.986	
5	1.00	1.015	
6	2.00	1.017	
7	2.00	.991	
8	2.00	1.018	

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example13.sav

ขั้นที่ 6.1 เลือกคำสั่ง Statistics / Compare Means / One-Way ANOVA

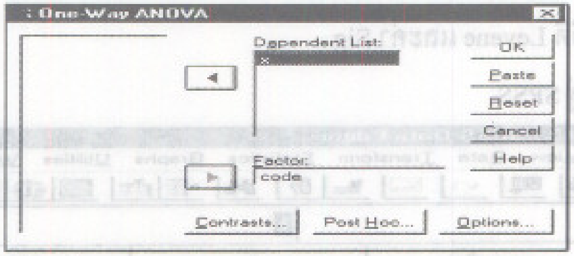


ขั้นที่ 6.2 คลิกที่คำสั่ง Statistics \ Compare Means \ One-Way ANOVA จะได้เมนูย่อยเป็น

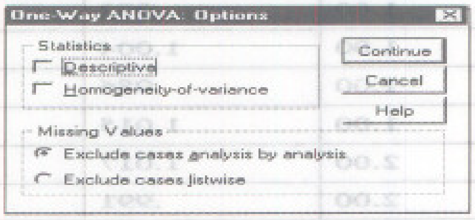


ขั้นที่ 6.3 เลือกตัวแปร x ไปไว้ที่ Dependent List

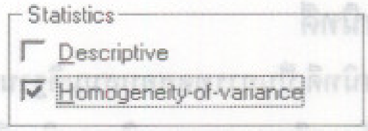
เลือกตัวแปร code ไปไว้ที่ Factor



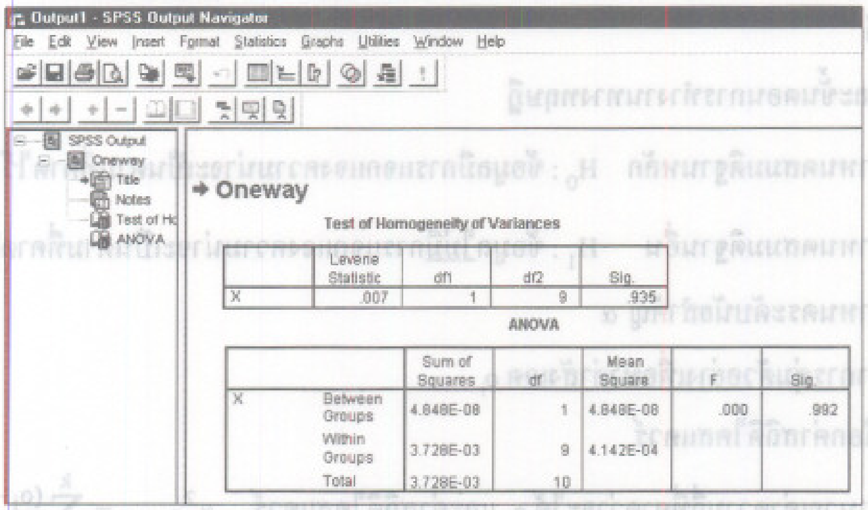
ขั้นที่ 6.4 คลิกที่ Options...



ขั้นที่ 6.4 คลิกที่ Homogeneity of variance



ขั้นที่ 6.5 คลิกที่ Continue และ OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

Oneway

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
X	.007	1	9	.935

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
X	Between Groups	4.848E-08	1	4.848E-08	.000	.992
	Within Groups	3.728E-03	9	4.142E-04		
	Total	3.728E-03	10			

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่า Sig = 0.935 กับ ค่านัยสำคัญ $\alpha = 0.1$

เพราะว่า Sig = 0.935 > $\alpha = 0.1$ เพราะฉะนั้น ยอมรับ H_0

7.6 การทดสอบภาวะสารูปสนิทดี

การทดสอบภาวะสารูปสนิทดีเป็นการทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้หรือไม่เช่น ข้อมูลมีการแจกแจงทวินามจริงหรือไม่ ข้อมูลมีการแจกแจงปกติจริงหรือไม่ ข้อมูลมีการแจกแจงปัวส์ซองจริงหรือไม่ ข้อมูลมีการแจกแจงตามอัตราส่วนที่คาดไว้จริงหรือไม่

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้

กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่มีการแจกแจงความน่าจะเป็นตามที่คาดไว้

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i และค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_α df = k – จำนวนค่าสถิติที่ใช้

บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

ตัวอย่าง 7.6.1. การทดลองโยนเหรียญ 3 อัน 240 ครั้ง ให้ x เป็นจำนวนหัวที่ได้ในการโยนเหรียญแต่ละครั้งผลการทดลองบันทึกไว้ที่แฟ้มข้อมูลชื่อ example14.sav

จงทดสอบสมมติฐานว่า เหรียญทั้งสามอันมีความเที่ยงตรง กำหนดระดับนัยสำคัญ 0.05

วิธีทำ การคำนวณโดย **MATHCAD**

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : เหรียญทั้งสามอันมีความเที่ยงตรง

กำหนดสมมติฐานอื่น H_1 : เหรียญทั้งสามอันไม่มีความเที่ยงตรง

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i

ภายใต้เงื่อนไขที่กำหนดว่าเหรียญทั้งสามอันมีความเที่ยงตรง

เพราะฉะนั้น x = 0, 1, 2, 3 มีการแจกแจงแบบทวินาม ดังนั้นค่าความถี่ที่คาดว่าจะได้คือ

X	P(X=x)	e_i
0	$\frac{1}{8}$	$\frac{1}{8}(240)= 30$
1	$\frac{3}{8}$	$\frac{3}{8}(240)=90$
2	$\frac{3}{8}$	$\frac{3}{8}(240)=90$
3	$\frac{1}{8}$	$\frac{1}{8}(240)= 30$
		240

การคำนวณโดยใช้ MATHCAD ทำได้ดังนี้

$$O := \begin{bmatrix} 24 \\ 98 \\ 95 \\ 23 \end{bmatrix} \qquad E := \begin{bmatrix} 30 \\ 90 \\ 90 \\ 30 \end{bmatrix}$$

$k := 4 \qquad i := 1..k$

$$\text{chisquare} := \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

chisquare = 3.822

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{0.05} = 7.815$ df = 3

บริเวณวิกฤตคือ $\chi^2 > 7.815$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > 7.815$

สรุปผล ยอมรับ H_0

หลักการและขั้นตอนการทำงานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลมีการแจกแจงความน่าจะเป็นตามที่คาดไว้

กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่มีการแจกแจงความน่าจะเป็นตามที่คาดไว้

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i และค่าสถิติไคสแควร์ และค่า Sig

$$\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_α $df = k -$ จำนวนค่าสถิติที่ใช้

บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

ขั้นที่ 7. สรุปผลทำได้ 2 แบบคือ

1. ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

2. ปฏิเสธ H_0 ถ้า $Sig < \alpha$

จากตัวอย่าง 7.6.1. การทดลองโยนเหรียญ 3 อัน 240 ครั้ง ให้ x เป็นจำนวนหัวที่ได้ในการโยนเหรียญแต่ละครั้งผลการทดลองบันทึกไว้ที่แฟ้มข้อมูลชื่อ example14.sav

จงทดสอบสมมติฐานว่า เหรียญทั้งสามอันมีความเที่ยงตรง กำหนดระดับนัยสำคัญ 0.05

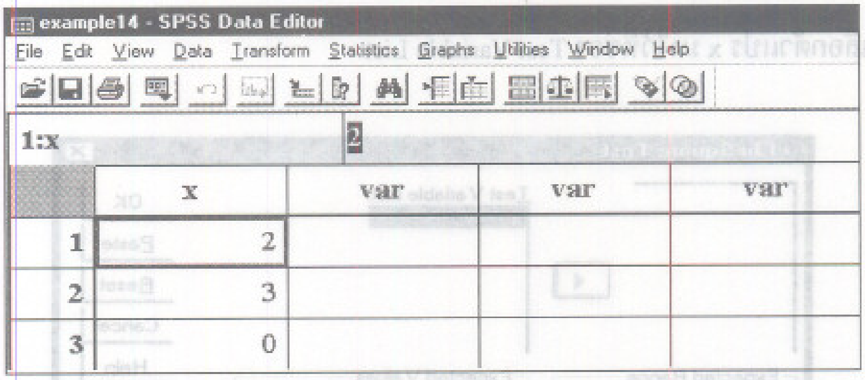
วิธีทำ การวิเคราะห์ข้อมูลด้วย SPSS

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : เหรียญทั้งสามอันมีความเที่ยงตรง

กำหนดสมมติฐานอื่น H_1 : เหรียญทั้งสามอันไม่มีความเที่ยงตรง

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

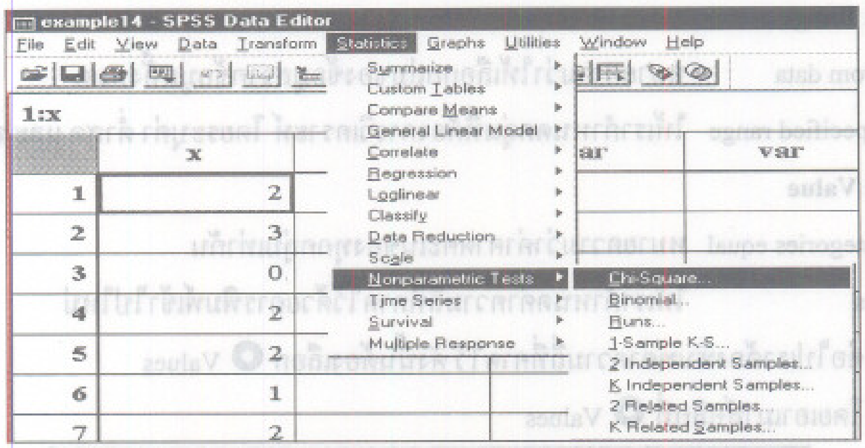
ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_i



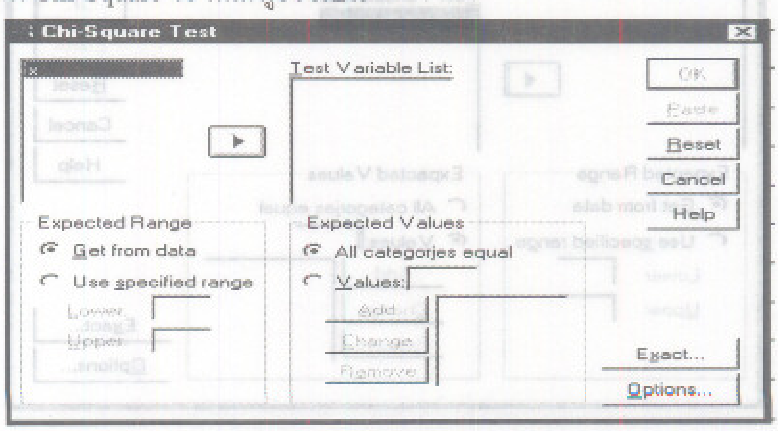
ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_i

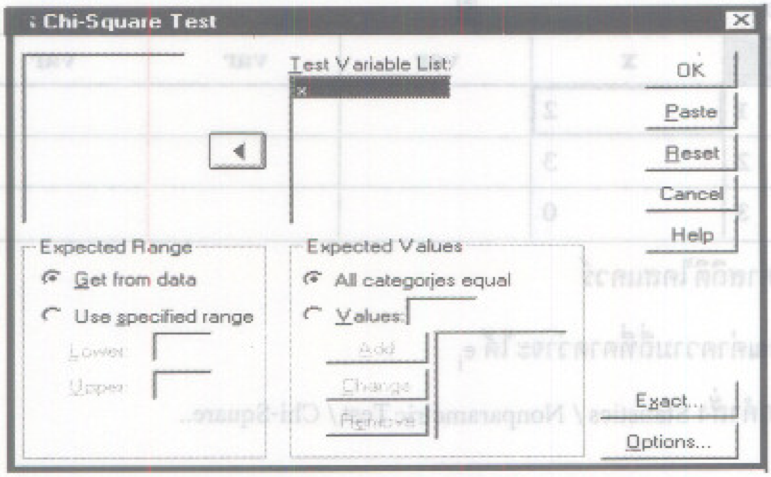
ขั้นที่ 5.1 เลือกคำสั่ง Statistics / Nonparametric Test / Chi-Square..



ขั้นที่ 5.2 คลิกที่ Chi-Square จะได้เมนูย่อยเป็น



ขั้นที่ 5.3 เลือกตัวแปร x มาไว้ที่ช่อง Test Variable List



Expected Range

○ Get from data หมายความว่าให้เลือกกลุ่มของข้อมูลจากข้อมูลทั้งหมด

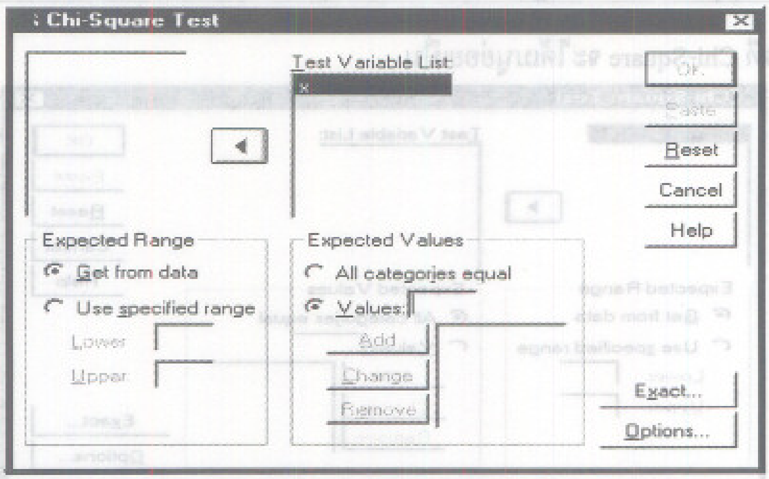
○ Use specified range ให้เรากำหนดกลุ่มที่ต้องการวิเคราะห์ โดยระบุค่า ต่ำสุด และ สูงสุด

Expected Value

○ All categories equal หมายความว่าค่าคาดคะเนของทุกกลุ่มเท่ากัน

○ Values ให้เรากำหนดค่าความถี่ที่คาดไว้ด้วยการพิมพ์เข้าไปใหม่

ขั้นที่ 5.4 ต่อไปเราต้องพิมพ์ค่าความถี่ที่คาดไว้ ดังนั้นต้องเลือก ○ Values โดยเอาเมาส์คลิกที่ ○ Values

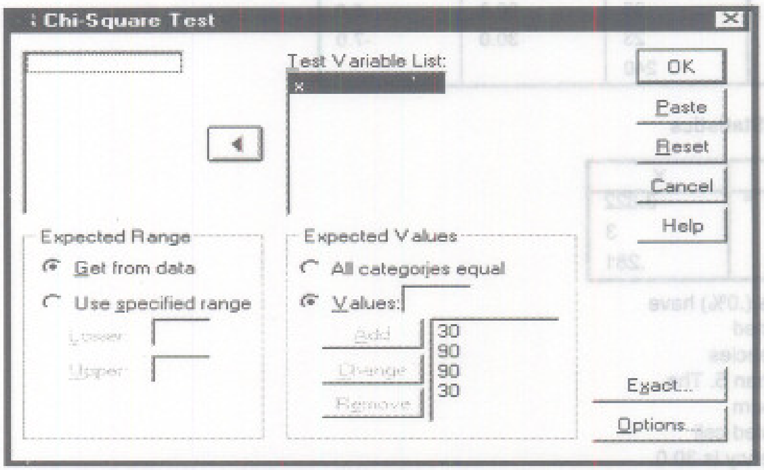


ขั้นที่ 5.5 การกำหนดค่าความถี่ที่คาดไว้ คือ 30 90 90 และ 30 ให้ทำดังนี้

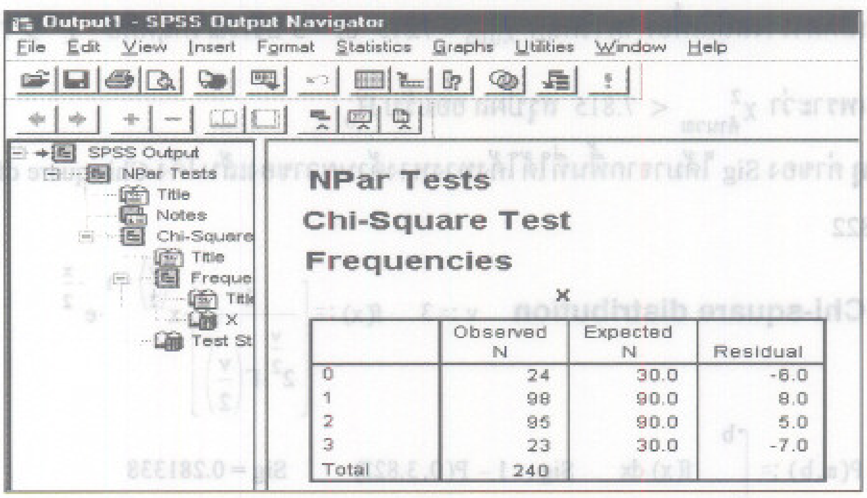
พิมพ์ 30 แล้วคลิก Add　　　พิมพ์ 90 แล้วคลิก Add

พิมพ์ 90 แล้วคลิก Add　　　พิมพ์ 30 แล้วคลิก Add

ผลบนจอภาพที่ได้คือ



เสร็จแล้วคลิก OK จะได้ผลการคำนวณที่ Output Navigator ดังนี้



	Observed N	Expected N	Residual
0	24	30.0	-6.0
1	98	90.0	8.0
2	95	90.0	5.0
3	23	30.0	-7.0
Total	240		

ผลการคำนวณทั้งหมดคือ

NPar Tests
Chi-Square Test
Frequencies

X

	Observed N	Expected N	Residual
0	24	30.0	-6.0
1	98	90.0	8.0
2	95	90.0	5.0
3	23	30.0	-7.0
Total	240		

Test Statistics

	X
Chi-Square [a]	3.822
df	3
Asymp. Sig.	.281

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 30.0.

ตัวเลขที่ได้จาก Output Navigator คือค่าสถิติไคสแควร์

$$\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = 3.822 \qquad df = 3$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{0.05} = 7.815$ df = 3 บริเวณวิกฤตคือ $\chi^2 > 7.815$

ขั้นที่ 7. เพราะว่า $\chi^2_{\text{คำนวณ}} < 7.815$ สรุปผล ยอมรับ H_0

หมายเหตุ ค่าของ Sig ได้มาจากพื้นที่ใต้โค้งทางหางด้านขวาของเส้นโค้ง Chi square df = 3 ที่ระยะ 3.822

Chi-square distribution $v := 3 \qquad f(x) := \left[\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)}\right] \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$

$P(a,b) := \int_a^b f(x)\,dx \qquad Sig := 1 - P(0, 3.822) \qquad Sig = 0.281338$

7.7 การทดสอบสมมติฐานว่าข้อมูลเป็นอิสระต่อกันหรือไม่

ในกรณีที่เราต้องการทดสอบความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวเกี่ยวข้องกันหรือไม่ ตัวอย่างเช่น

– การฉีดวัคซีนป้องกันโรคอหิวาตกโรคกับการเป็นโรคอหิวาตกโรค เกี่ยวข้องกันหรือไม่

– การนับถือศาสนา และ ถิ่นที่อยู่ เกี่ยวข้องกันหรือไม่

เราจะทำการทดสอบทางสถิติเพื่อดูว่าข้อมูลเป็นอิสระต่อกันหรือไม่

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลเป็นอิสระต่อกัน

กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่เป็นอิสระต่อกัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_{ij} และค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_α $df = (r-1)(c-1)$ บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่างกับค่าวิกฤต ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

หลักการและขั้นตอนการทดสอบสมมติฐานด้วย SPSS for Windows

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลเป็นอิสระต่อกัน

กำหนดสมมติฐานอื่น H_1 : ข้อมูลไม่เป็นอิสระต่อกัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_{ij} และค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$

และค่า Sig (ค่านัยสำคัญของค่าสถิติ $\chi^2_{\text{คำนวณ}}$)

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต χ^2_α $df = (r-1)(c-1)$ บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

ขั้นที่ 7. สรุปผลมี 2 วิธีคือ 1. ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

2. ปฏิเสธ H_0 ถ้า Sig < α

ตัวอย่าง 7.7.1. แฟ้มข้อมูล example15.sav บันทึกข้อมูลเพื่อศึกษาความสัมพันธ์ระหว่างการนับถือศาสนา และถิ่นที่อยู่

ศาสนา โปรเทศตัน คาธอลิก และ ยิว

ถิ่นที่อยู่ ฝั่งตะวันออก และ ฝั่งตะวันตก

จงทดสอบสมมติฐานว่า การนับถือศาสนา และ ถิ่นที่อยู่มีความสัมพันธ์กันหรือไม่ กำหนดนัยสำคัญ 0.05

วิธีทำ การคำนวณโดย SPSS

ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : การนับถือศาสนาและถิ่นที่อยู่อาศัย ไม่มีความสัมพันธ์กัน

กำหนดสมมติฐานอื่น H_1 : การนับถือศาสนาและถิ่นที่อยู่อาศัย มีความสัมพันธ์กัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

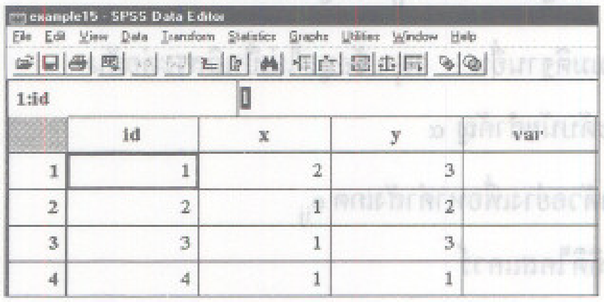
ขั้นที่ 3. ทำการสุ่มตัวอย่างเพื่อหาค่าสังเกต o_{ij}

เมื่อสุ่มตัวอย่างมาแล้วต้องสร้าง แฟ้มข้อมูลประกอบด้วย 2 ตัวแปร

x เป็นตัวแปรจำแนก ถิ่นที่อยู่ ฝั่งตะวันออก = 1 และ ฝั่งตะวันตก = 2

y เป็นตัวแปรจำแนก ศาสนา โปรเทศตัน = 1 คาธอลิก = 2 และ ยิว = 3

แฟ้มข้อมูลที่สร้างแล้วชื่อ example15.sav



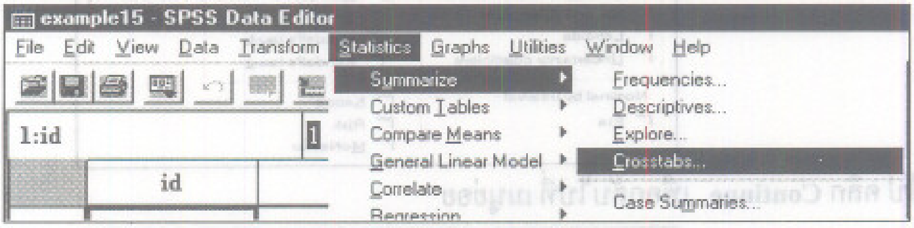
	id	x	y	var
1	1	2	3	
2	2	1	2	
3	3	1	3	
4	4	1	1	

ขั้นที่ 4. เลือกค่าสถิติไคสแควร์

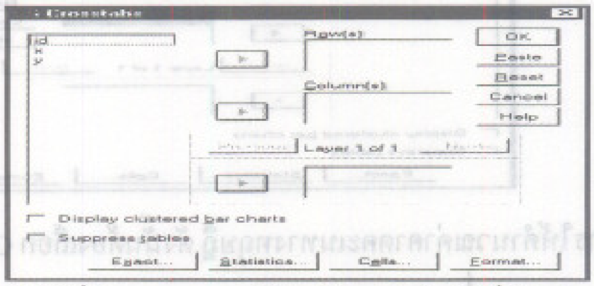
ขั้นที่ 5. คำนวณค่าความถี่ที่คาดว่าจะได้ e_{ij} และค่าสถิติไคสแควร์ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$

การคำนวณโดย SPSS

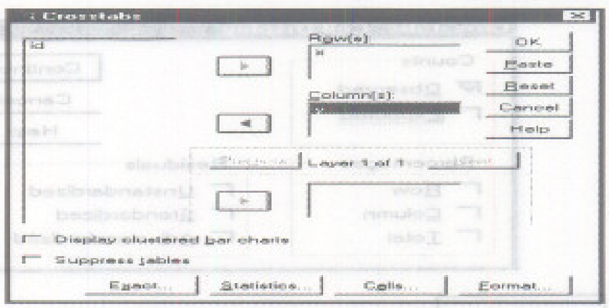
ขั้นที่ 5.1 ใช้คำสั่ง **Statistics / Summarize / Crosstabs ..**



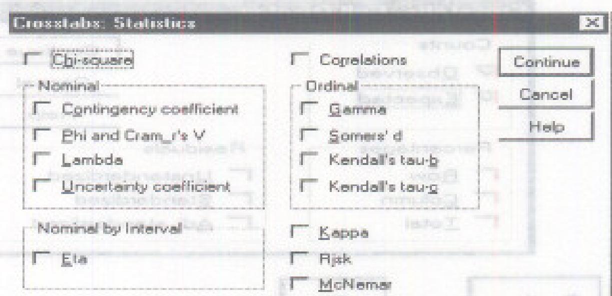
ขั้นที่ 5.2 คลิกที่ Crosstabs.. จะได้เมนูดังนี้



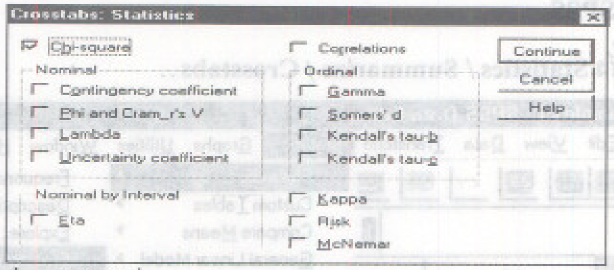
ขั้นที่ 5.3 เลือกตัวแปร x ไปไว้ที่ช่อง Row(s) เลือกตัวแปร y ไปไว้ที่ช่อง Column(s)



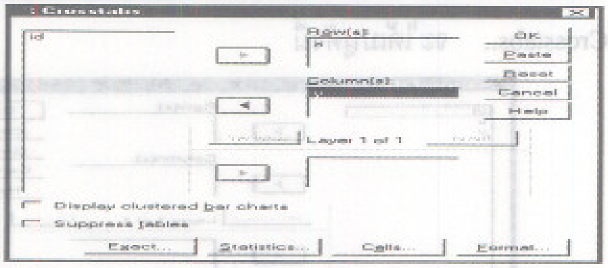
คลิกที่ **Statistics..** จะได้เมนูย่อยเป็น



คลิกที่ ☐ Chi-square ให้มีเครื่องหมายถูก ☑ Chi-square

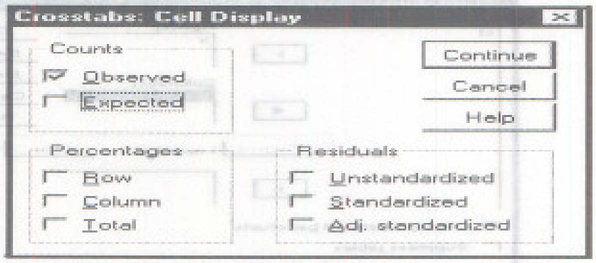


ต่อไป คลิก **Continue** เพื่อกลับไปที่ เมนูย่อย

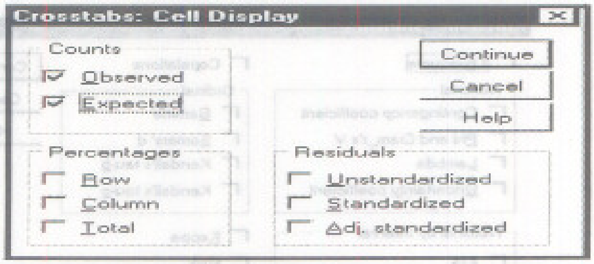


เพราะว่าเราต้องการให้คำนวณค่าคาดคะเนทางทฤษฎี ดังนั้นต้องเลือก Options **Cells**

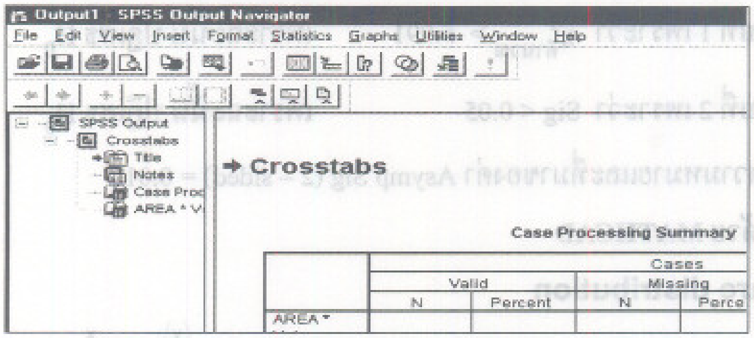
ขั้นที่ 5.4 คลิกที่ Cells... จะได้เมนูย่อยเป็น



ขั้นที่ 5.5 คลิกที่ Expected ให้เกิดเครื่องหมายถูก



ขั้นที่ 5.6 คลิกที่ Continue และ OK ตามลำดับ จะได้ผลการคำนวณที่ SPSS Output Navigator ดังนี้



ผลการคำนวณทั้งหมดคือ Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
AREA * Value Label of variable y	1000	100.0%	0	.0%	1000	100.0%

AREA * Value Label of variable y Crosstabulation

			Value Label of variable y			
			Protestant	Christ	Jew	Total
AREA	East	Count	182	215	203	600
		Expected Count	201.6	210.6	187.8	600.0
	West	Count	154	136	110	400
		Expected Count	134.4	140.4	125.2	400.0
Total		Count	336	351	313	1000
		Expected Count	336.0	351.0	313.0	1000.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.069[a]	2	.018
Likelihood Ratio	8.053	2	.018
Linear-by-Linear Association	7.774	1	.005
N of Valid Cases	1000		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 125.20.

จากผลการคำนวณจะได้ $\chi^2_{\text{คำนวณ}} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} = 8.069$ และ df = 2 และ Sig = 0.018

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต $\chi^2_{0.05} = 5.991$ df = 2 บริเวณวิกฤตคือ $\chi^2 > 5.991$

ขั้นที่ 7. โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

แบบที่ 1 เพราะว่า $\chi^2_{\text{คำนวณ}} > 5.991$ เพราะฉะนั้น ปฏิเสธ H_0

แบบที่ 2 เพราะว่า Sig < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ ความหมายและที่มาของค่า Asymp Sig (2 – sided) = 0.018

การคำนวณด้วย **MATHCAD**

Chi–square distribution

$$v := 2 \qquad f(x) := \left[\frac{1}{2^{\frac{v}{2}} \cdot \Gamma\left(\frac{v}{2}\right)}\right] \cdot x^{\left(\frac{v}{2}\right)-1} \cdot e^{-\frac{x}{2}}$$

$$P(a,b) := \int_a^b f(x)\,dx \qquad P(8.069, 1000) = 0.0176945823$$

ความหมายของ Asymp Sig (2 – sided) = 0.018 คือพื้นที่ใต้โค้งทางหางด้านขวาของเส้นโค้งไคสแควร์ที่ระยะ 8.069

การสรุปผลสามารถนำค่า Asymp Sig (2 – sided) เปรียบเทียบกับค่า α

ถ้า Asymp Sig (2–sided) < α แล้ว ปฏิเสธ H_0

หมายเหตุ ในกรณีที่ข้อมูลแจกแจงความถี่แล้ว การคำนวณโดยใช้ MATHCAD ทำได้ดังนี้

$$O := \begin{pmatrix} 182 & 215 & 203 \\ 154 & 136 & 110 \end{pmatrix} \qquad r := 2 \qquad c := 3 \qquad N := \sum_{i=1}^{r} \sum_{j=1}^{c} O_{(i,j)}$$

$$i := 1..r \qquad j := 1..c$$

$$R_i := \sum_{j=1}^{c} O_{(i,j)} \qquad C_j := \sum_{i=1}^{r} O_{(i,j)} \qquad E_{(i,j)} := \frac{R_i \cdot C_j}{N}$$

$$E = \begin{pmatrix} 201.6 & 210.6 & 187.8 \\ 134.4 & 140.4 & 125.2 \end{pmatrix}$$

$$\text{chisquare} := \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{\left[O_{(i,j)} - E_{(i,j)}\right]^2}{E_{(i,j)}}$$

$$\text{chisquare} = 8.069$$

แบบฝึกหัด 7.

1. มีการค้นพบวิธีผลิตวัสดุซีเมนต์ชนิดหนึ่งให้ทนแรงกดได้ 5000 ปอนด์ต่อตารางนิ้ว เพื่อทดสอบสมมติฐาน $H_0 : \mu = 5000$ แย้งกับสมมติฐาน $H_1 : \mu \neq 5000$ กำหนดระดับนัยสำคัญ 0.05 จึงสุ่มตัวอย่างวัสดุซีเมนต์พวกนี้มา 50 ชิ้นได้ข้อมูลเป็นดังนี้

4758 , 4839 , 5004 , 4906 , 5105 , 4831 , 5057 , 4886 , 4796 , 4820 , 5175 , 4808 , 4762 , 4982
5012 , 4828 , 4948 , 4782 , 5088 , 4977 , 5127 , 5161 , 4985 , 4953 , 5121 , 5087 , 5178 , 5016
4870 , 5112 , 4916 , 5043 , 4762 , 4874 , 5006 , 5111 , 4962 , 5071 , 4951 , 5072 , 5010 , 5068
4999 , 4822 , 4937 , 4976 , 5075 , 4829 , 4965 , 5053

2. โรงงานแห่งหนึ่งผลิตหลอดไฟฟ้าต้องการทดสอบสมมติฐาน $\mu = 800$ ชั่วโมง แย้งกับ $H_1 : \mu \neq 800$ ชั่วโมงที่ระดับนัยสำคัญ 0.01 ผลการสุ่มตัวอย่างหลอดไฟ 30 หลอดได้ข้อมูลดังนี้

782 , 827 , 726 , 767 , 814 , 851 , 798 , 837 , 794 , 837 , 815 , 836 , 811 , 748 , 789 , 803 , 838
750 , 799 , 831 , 747 , 746 , 830 , 789 , 871 , 748 , 849 , 754 , 848 , 748

3. เท่าที่ผ่านมาปรากฏว่า ส่วนสูงเฉลี่ยของนิสิตชายเป็น 68.5 นิ้ว ส่วนเบี่ยงเบนมาตรฐาน 2.7 มีเหตุผลหรือไม่ที่จะเชื่อว่าส่วนสูงเฉลี่ยเปลี่ยนไป ถ้าสุ่มตัวอย่างนิสิตใหม่ 50 คน ได้ข้อมูลดังนี้

65, 65, 70.2, 67.7, 72.8, 65.1, 71.4, 65.4, 71.3, 65.1, 70.5, 66.3, 70.1, 70.5, 64.8, 71.5, 68.5, 67.7
72.6, 68.8, 68.1, 71.6, 67.9, 72.9, 70.6, 65.5, 71.6, 68.4, 63.9, 69, 68.7 , 71.6, 69.8, 71.6, 64.7
66.6, 66.3, 65, 71.5, 69.3, 66, 63.7, 71.2, 65.6, 68.9, 64.7, 70.7, 68.8, 67.8, 70.2

จงทดสอบที่ระดับนัยสำคัญ 0.1

4. เป็นที่กล่าวอ้างว่า โดยเฉลี่ยรถแต่ละคันจะเดินทางปีละ 12,000 ไมล์ เมื่อทดสอบคำกล่าวอ้างนี้ที่ระดับนัยสำคัญ 0.01 สุ่มตัวอย่างเจ้าของรถ 60 คัน และขอให้บันทึกระยะการเดินทางตลอดปีได้ข้อมูลดังนี้

14137, 15366, 15809, 14725, 16393, 11989, 11888, 14318, 11185, 11157, 18445, 18124
12315, 13911, 15803, 14267, 15540, 17777, 15775, 15261, 11940, 15169, 12450, 15600
15425, 14640, 16784, 15736, 17339, 15351, 18271, 16670, 16149, 13018, 12998, 11281
17746, 11684, 11026, 15916, 15081, 13896, 14389, 11665, 16760, 17533, 17939, 12948
11443, 17169, 15643, 10965, 16028, 11248, 12313, 18345, 16106, 14623, 14661, 14769

ท่านจะเห็นด้วยกับคำกล่าวอ้างหรือไม่

5. สุ่มตัวอย่างบุหรี่ชนิดหนึ่ง 8 มวน ได้ข้อมูลปริมาณนิโคตินดังนี้

15.3 16.9 20.2 18.2 22.2 16.7 21.2 17.8 หน่วยเป็นมิลลิกรัม

ท่านคิดว่า $\sigma = 2$ ใช่หรือไม่ ที่ระดับนัยสำคัญ 0.05

6. สุ่มตัวอย่างขนาด $n_1 = 25$ จากประชากรชุดที่ 1 สุ่มตัวอย่างขนาด $n_2 = 36$ จากประชากรชุดที่ 2

จงทดสอบสมมติฐาน $\mu_1 = \mu_2$ แย้งกับสมมติฐาน $\mu_1 \neq \mu_2$ ที่ระดับนัยสำคัญ 0.05

ตัวอย่างจากประชากรชุดที่ 1 คือ

73, 73, 83, 78, 88, 73, 85, 74, 85, 73, 84, 76, 83, 84, 73, 86, 80, 78, 88, 80, 79, 86, 79, 88, 84

ตัวอย่างจากประชากรชุดที่ 2 คือ

72, 80, 76, 70, 77, 76, 80, 78, 80, 71, 74, 73, 72, 80, 77, 73, 70, 80, 72, 77, 71, 79, 77, 75, 78, 75, 77, 78, 76, 78, 72, 72, 76, 71, 71, 81

7. ชาวนาอ้างว่าข้าวโพดพันธุ์ ก . ให้ผลผลิตเท่ากับข้าวโพดพันธุ์ ข. เพื่อทดสอบคำกล่าวอ้างนี้ จึงทดลองปลูกข้าวโพดทั้งสองชนิดในพื้นที่ 50 ไร่ ภายใต้สภาพการอย่างเดียวกัน จงทดสอบคำกล่าวอ้างที่ระดับนัยสำคัญ 0.05

ข้อมูลข้าวโพด ก. : 93 , 93 , 94.7 , 81.5 , 80.7 , 94.3 , 93.6 , 76.7 , 79.7 , 85.3 , 94.7 , 87.3 , 86.8 , 86.9 , 94.1 , 78.9 , 81.8 , 96.3 , 90.9 , 89.4 , 91.3 , 78.5 , 83.7 , 91.5 , 76.5 , 85.3 , 80.6 , 79.1 , 86.4 , 79.7 , 80.3 , 92.1 , 87.6 , 93.8 , 87.2 , 94.1 , 80.6 , 83.6 , 76.6 , 84.2 , 82.7 , 80.1 , 78.5 , 94.7 , 80.5 , 85.6 , 84.8 , 94 , 81.5 , 90.2

ข้อมูลข้าวโพด ข. : 86.6 , 83.3 , 83.8 , 70.6 , 68.3 , 62.5 , 85.5 , 82.9 , 80.2 , 81.2 , 77.7 , 73.1 , 80.5 , 73 , 76.9 , 74.8 , 77.2 , 75.8 , 68.4 , 79.5 , 76.3 , 75.4 , 68.1 , 84.8 , 81.5 , 79.6 , 77.8 , 75.3, 86.2 , 86.7 , 79.3 , 75.8 , 77.6 , 84.9 , 83.1 , 79.4 , 73.3 , 84.9 , 73.8 , 84.8 , 71.6 , 85.2 , 68.5 , 69.1 , 62.9 , 70.2 , 77.5 , 68.9 , 74.5 , 77.9

8. สุ่มตัวอย่างคน 4 คน บันทึกน้ำหนักก่อนงดสูบบุหรี่และ หลังงดสูบบุหรี่ 5 สัปดาห์ ได้ผลดังนี้

คนที่	1	2	3	4	
น้ำหนักก่อนงดสูบบุหรี่	148	176	153	116	ปอนด์
น้ำหนักหลังงดสูบบุหรี่	154	176	151	121	ปอนด์

จงทดสอบสมมติฐานว่า น้ำหนักไม่เปลี่ยนแปลงถ้าเลิกสูบบุหรี่ ที่ระดับนัยสำคัญ 0.05

บทที่ 8

สหสัมพันธ์และการถดถอยเชิงเส้น

การทำงานทางด้านสถิติเรามักจะพบว่ามีตัวแปร 2 ตัวหรือมากกว่าอาจจะมีความสัมพันธ์กัน เช่น น้ำหนักกับอายุ รายได้กับรายจ่าย ความสัมพันธ์ของตัวแปรอาจจะเป็น

ความสัมพันธ์แบบเชิงเดียว (simple correlation) ซึ่งประกอบด้วยตัวแปรอิสระ 1 ตัว และตัวแปรตาม 1 ตัว รูปแบบของสมการความสัมพันธ์เชิงเส้นเชิงเดียวอาจมีรูปแบบเป็น

- $y = a + bx$
- $\ln y = a + b \ln x$
- $y = a + b \ln x$
- $\ln y = a + bx$

ความสัมพันธ์แบบพหุคูณ (Multiple correlation) ซึ่งประกอบด้วยตัวแปรอิสระมากกว่า 1 ตัว และตัวแปรตาม 1 ตัว

รูปแบบของสมการความสัมพันธ์พหุคูณอาจมีรูปแบบเป็น

- $y = a + b_1x_1 + b_2x_2 + b_3x_3$

8.1 การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียว (Simple Linear Regression) และสหสัมพันธ์ (Correlation)

กำหนดให้ X เป็นตัวแปรอิสระ และ Y เป็นตัวแปรตาม ความสัมพันธ์ที่แท้จริงของ X และ Y คือ

$$\mu_{Y|x} = \alpha + \beta x$$

α เรียกว่า **สัมประสิทธิ์การถดถอย** (regression coefficients)

เราต้องการประมาณความสัมพันธ์ $\mu_{Y|x} = \alpha + \beta x$ ด้วย $\hat{y} = a + bx$

ρ เป็นค่า **สหสัมพันธ์** (correlation) จากข้อมูลตัวอย่างเราจะประมาณค่า ρ ด้วย r

หมายเหตุ 1. $-1 \le \rho \le 1$

2. $|\rho|$ มีค่ามาก แสดงว่า X และ Y มีความสัมพันธ์กันมาก

3. $\rho = 0$ แสดงว่า X และ Y ไม่มีความสัมพันธ์

4. $\rho > 0$ แสดงว่า ถ้า X มีค่าเพิ่มขึ้น แล้ว Y มีค่าเพิ่มขึ้น
หรือ ถ้า X มีค่าลดลง แล้ว Y มีค่า ลดลง

5. $\rho < 0$ แสดงว่า ถ้า X มีค่าเพิ่มขึ้น แล้ว Y มีค่าลดลง
หรือ ถ้า X มีค่าลดลง แล้ว Y มีค่าเพิ่มขึ้น

6. b และ r จะมีเครื่องหมายเหมือนกัน แต่ b สามารถ บอกอัตราการเพิ่มหรือลดของตัวแปรตามเทียบกับตัวแปรอิสระ

การหาสมการเส้นถดถอยเชิงเส้นเชิงเดียวและสหสัมพันธ์

หลักการทางทฤษฎีของความน่าจะเป็นและสถิติ จากข้อมูลที่เก็บมาได้

ตัวอย่างจากประชากรชุดที่ 1.	ตัวอย่างจากประชากรชุดที่ 2.
x_1	y_1
x_2	y_2
x_3	y_3
:	:
x_n	y_n

เราต้องการหาค่า a และ b ที่ทำให้ $\hat{y} = a + bx$ และ สัมประสิทธิ์สหสัมพันธ์ r

ขั้นตอนการคำนวณ

ขั้นที่ 1. คำนวณค่า $\sum x$, $\sum y$, $\sum xy$, $\sum x^2$, $\sum y^2$

ขั้นที่ 2. คำนวณค่า

$$b := \frac{n \cdot \sum_{i=1}^{n} x_i \cdot y_i - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$a = \overline{y} - b\overline{x}$$

$$r := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{\sqrt{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n \cdot \sum_{i=1}^{n} (y_i)^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

ตัวอย่างเช่น

X	1.50	1.80	2.40	3.00	3.50	3.90	4.40	4.80	5.00
Y	4.80	5.70	7.00	8.30	10.90	12.40	13.10	13.60	15.30

เราสามารถหาสมการ $\hat{y} = ax + b$ และค่า r ตามขั้นตอนการคำนวณดังนี้

การคำนวณด้วย MATHCAD

แบบที่ 1 คำนวณค่าตามสูตร

ORIGIN := 1

$$x := \begin{bmatrix} 1.5 \\ 1.8 \\ 2.4 \\ 3.0 \\ 3.5 \\ 3.9 \\ 4.4 \\ 4.8 \\ 5.0 \end{bmatrix} \qquad y := \begin{bmatrix} 4.8 \\ 5.7 \\ 7.0 \\ 8.3 \\ 10.9 \\ 12.4 \\ 13.1 \\ 13.6 \\ 15.3 \end{bmatrix} \qquad n := \mathrm{length}(x)$$

$$b := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad b = 2.93028$$

$$a := \mathrm{mean}(y) - b \cdot \mathrm{mean}(x) \qquad a = 0.256947$$

$$r := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{\sqrt{\left[n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right] \cdot \left[n \cdot \sum_{i=1}^{n} (y_i)^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]}} \qquad r = 0.991089$$

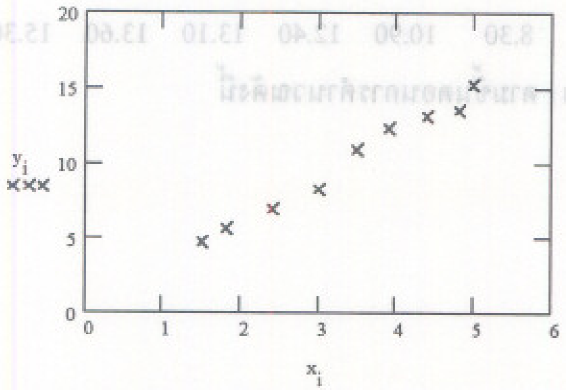
แบบที่ 2 ใช้ฟังก์ชัน slope(x,y) และ intercept(x,y) ของ MATHCAD

b := slope(x,y)   b = 2.93028

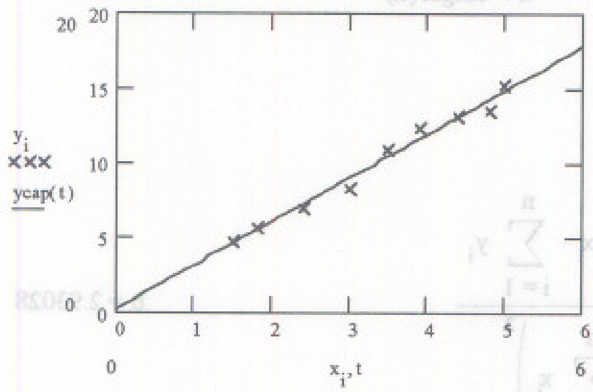
a := intercept(x,y)   a = 0.256947

r := corr(x,y)   r = 0.991089

แผนภาพการกระจายของข้อมูลเป็นดังนี้



แผนภาพการกระจายของข้อมูล และสมการเส้นถดถอย $\hat{y} = a + bx$



การคำนวณด้วย SPSS

ขั้นที่ 1 นำข้อมูลเข้าสู่ SPSS Data Editor

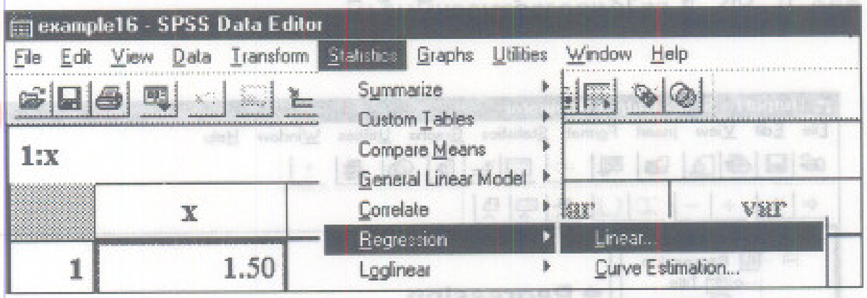
example16 - SPSS Data Editor

File Edit View Data Transform Statistics Graphs Utilities Window Help

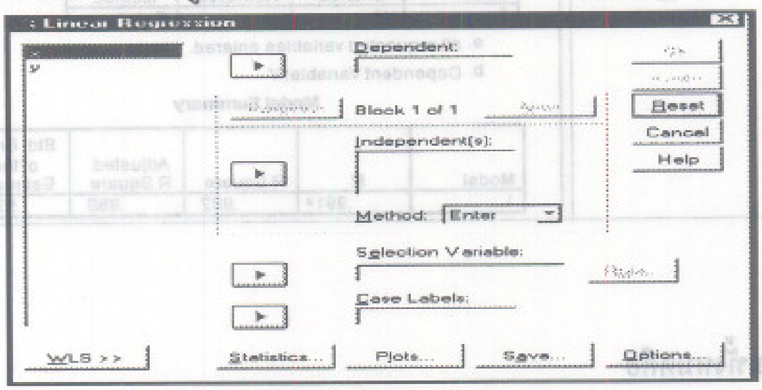
1:x 1.5

	x	y	var	var
1	1.50	4.80		
2	1.80	5.70		
3	2.40	7.00		
4	3.00	8.30		
5	3.50	10.90		
6	3.90	12.40		
7	4.40	13.10		
8	4.80	13.60		
9	5.00	15.30		

ขั้นที่ 2. ใช้คำสั่ง Statistics / Regression / Linear..

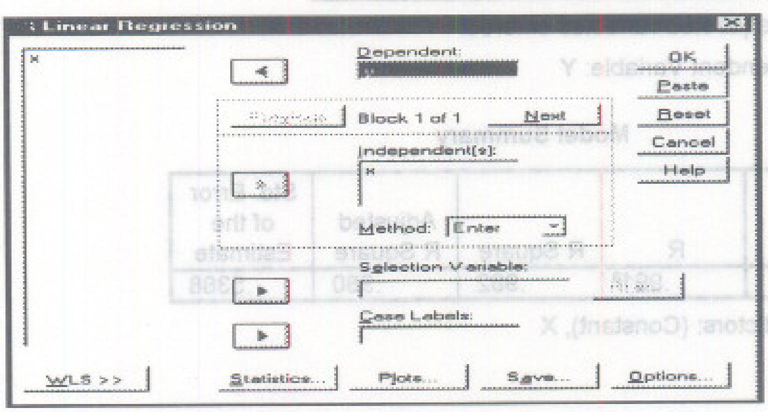


ขั้นที่ 3. คลิกที่ Linear จะได้เมนูของคำสั่งดังนี้

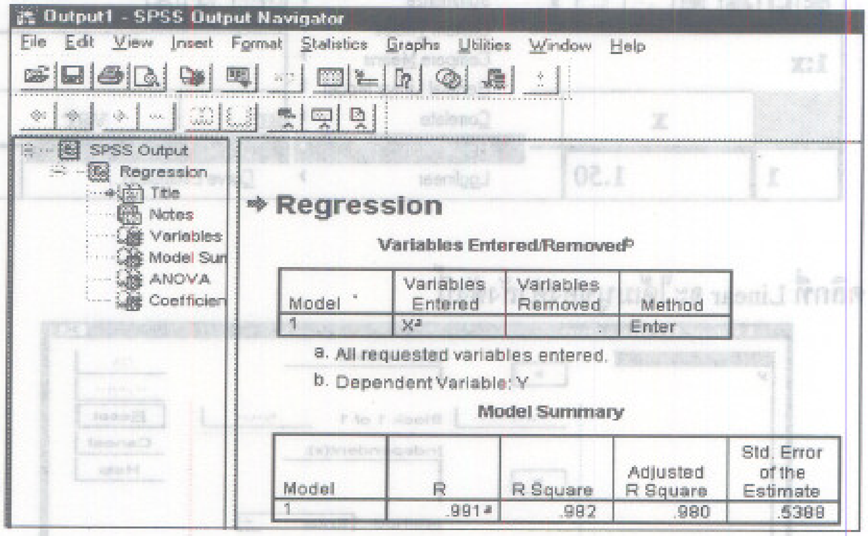


ขั้นที่ 4. เลือกตัวแปร x เป็นตัวแปรอิสระ นำไปไว้ที่ช่อง Independent(s)

เลือกตัวแปร y เป็นตัวแปรตาม นำไปไว้ที่ช่อง Dependent(s)



ขั้นที่ 5. คลิก OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	X[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.991[a]	.982	.980	.5388

a. Predictors: (Constant), X

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.484	1	112.484	387.516	.000000218[a]
	Residual	2.032	7	.290		
	Total	114.516	8			

a. Predictors: (Constant), X

b. Dependent Variable: Y

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.257	.532		.483	.644063634
	X	2.930	.149	.991	19.685	.000000218

a. Dependent Variable: Y

จากผลการคำนวณของ SPSS จะได้ $a = 0.257$, $b = 2.930$, $r = 0.991$

หมายเหตุ ที่มาของค่าสถิติในตาราง Coefficients

ค่าสถิติในช่องของ ตัวแปร X จากข้อมูล

$$x := \begin{bmatrix} 1.5 \\ 1.8 \\ 2.4 \\ 3.0 \\ 3.5 \\ 3.9 \\ 4.4 \\ 4.8 \\ 5.0 \end{bmatrix} \qquad y := \begin{bmatrix} 4.8 \\ 5.7 \\ 7.0 \\ 8.3 \\ 10.9 \\ 12.4 \\ 13.1 \\ 13.6 \\ 15.3 \end{bmatrix} \qquad n := \mathrm{length}(x)$$

Unstandardized Coefficients B คือค่าสัมประสิทธิ์ของการถดถอยเชิงเส้น b ที่คำนวณจากสูตร

$$b := \frac{n \cdot \left[\sum_{i=1}^{n} (x_i \cdot y_i)\right] - \sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n \cdot \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad b = 2.9303$$

Unstadardized Coefficients Std. Error คือส่วนเบี่ยงเบนมาตรฐานของค่าสถิติ b ที่คำนวณ

จากสูตร $\sigma_b = \dfrac{S}{\sqrt{S_{XX}}}$

โดยมีขั้นตอนการคำนวณที่สำคัญดังนี้

$$S_{xx} := \sum_{i=1}^{n} (x_i)^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} \qquad S_{xx} = 13.1$$

$$S_{xy} := \sum_{i=1}^{n} x_i \cdot y_i - \frac{\sum_{i=1}^{n} x_i \cdot \sum_{i=1}^{n} y_i}{n} \qquad S_{xy} = 38.3867$$

$$S_{yy} := \sum_{i=1}^{n} (y_i)^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \qquad S_{yy} = 114.5156$$

$$SSE := \sum_{i=1}^{n} (y_i - a - b \cdot x_i)^2 \qquad SSE = 2.0319$$

หรือ $SSE := S_{yy} - b \cdot S_{xy}$ $\qquad SSE = 2.0319$

$$\frac{SSE}{n-2} = 0.2903 \qquad S := \sqrt{\frac{SSE}{n-2}} \qquad S = 0.5388$$

$$\frac{S}{\sqrt{S_{xx}}} = 0.1489$$

สรุป $\sigma_b = 0.1489$

Standardized Coefficients Beta ในกรณีของความสัมพันธ์เชิงเส้นตรงค่าของ Standardized Coefficients Beta (X) มีค่าเท่ากับค่าสัมประสิทธิ์สหสัมพันธ์

ค่า t ได้มาจากสูตร $t = \dfrac{b}{\left(\dfrac{S}{\sqrt{S_{XX}}}\right)}$ $\qquad t = 19.68543$

ค่า Sig คือ 2 เท่าของพื้นที่ใต้โค้งทางหางด้านขวาของเส้นโค้ง t เมื่อ df = n -2 = 7

T distribution TOL := 0.0000001 v := 7

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}} \qquad Pvalue(T) := 0.5 - \int_0^{|T|} h(t)\,dt$$

Pvalue(19.68543) = 0.000000109 2·Pvalue(19.68543) = 0.0000002181

95% Confidence interval for B หมายถึงช่วงความเชื่อมั่น 95% ของค่า β มีสูตรเป็น

$$b - t_{\frac{\alpha}{2}}\frac{S}{\sqrt{S_{XX}}} < \beta < b + t_{\frac{\alpha}{2}}\frac{S}{\sqrt{S_{XX}}}$$

$$\text{Lower Bound} = b - t_{\frac{\alpha}{2}}\frac{S}{\sqrt{S_{XX}}} \qquad \text{Upper Bound} = b + t_{\frac{\alpha}{2}}\frac{S}{\sqrt{S_{XX}}}$$

talphadivide2 := 2.365

$$LowerBound := b - talphadivide2\cdot\left(\frac{S}{\sqrt{S_{xx}}}\right) \qquad LowerBound = 2.5782$$

$$UpperBound := b + talphadivide2\cdot\left(\frac{S}{\sqrt{S_{xx}}}\right) \qquad UpperBound = 3.2823$$

ค่าสถิติในช่องของ Constant

Unstandardized Coefficients B คือค่าระยะตัดแกน Y จากสมการ $\hat{y} = a + bx$

a := mean(y) − b·mean(x)

a = 0.2569

Unstandardized Coefficients Std. Error คือส่วนเบี่ยงเบนมาตรฐานของค่าสถิติ a ที่คำนวณ

จากสูตร $\sigma_a = s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}}$ $\qquad Std\_Error\_Constant := S\cdot\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\cdot S_{xx}}}$

Std_Error_Constant = 0.5324

ค่า t ได้มาจากสูตร $t = \frac{a}{s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}}}$ $\qquad$ t = 0.4827

ค่า Sig คือ 2 เท่าของพื้นที่ใต้โค้งทางหางด้านขวาของเส้นโค้ง t = 0.4827 (จากที่คำนวณได้) เมื่อ df = n − 2 = 7

T distribution TOL := 0.00000001 v := 7

$$h(t) := \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\cdot\sqrt{\pi\cdot v}}\right]\cdot\left[1+\left(\frac{t^2}{v}\right)\right]^{-\frac{v+1}{2}} \qquad \text{Pvalue}(T) := 0.5 - \int_0^{|T|} h(t)\,dt$$

Pvalue(0.4827) = 0.3220 2·Pvalue(0.4827) = 0.6440

95% Confidence interval for B หมายถึงช่วงความเชื่อมั่น 95% ของค่า α มีสูตรเป็น

$$a - t_{\frac{\alpha}{2}} s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}} < \alpha < a + t_{\frac{\alpha}{2}} s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}}$$

$$\text{Lower Bound} = a - t_{\frac{\alpha}{2}} s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}} \qquad \text{Upper Bound} = a + t_{\frac{\alpha}{2}} s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}}$$

talphadivide2 := 2.365

$$\text{LowerBound} := a - \text{talphadivide2}\cdot\left[S\cdot\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\cdot S_{XX}}}\right] \qquad \text{LowerBound} = -1.0021$$

$$\text{UpperBound} := a + \text{talphadivide2}\cdot\left[S\cdot\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n\cdot S_{XX}}}\right] \qquad \text{UpperBound} = 1.516$$

ที่มาของค่าสถิติในตาราง Model Summary

R = ค่าสัมประสิทธิ์สหสัมพันธ์ เป็นตัวเลขที่บอกระดับและทิศทางของความสัมพันธ์ระหว่างตัวแปร

คำนวณได้จากสูตร

$$r := \frac{n\cdot\left[\sum_{i=1}^{n}(x_i\cdot y_i)\right] - \sum_{i=1}^{n} x_i\cdot\sum_{i=1}^{n} y_i}{\sqrt{n\cdot\sum_{i=1}^{n}(x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\cdot\sqrt{n\cdot\sum_{i=1}^{n}(y_i)^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

หรือใช้ฟังก์ชัน corr(x,y) ของ MATHCAD R = corr(x , y) R = 0.9911

R Square เป็นค่าสัมประสิทธิ์การตัดสินใจ ได้มาจากค่า R^2 เป็นตัวเลขที่ใช้ในการอธิบายว่า สมการเส้นถดถอย $\hat{y} = a + bx$ มีความเหมาะสมที่จะนำไปใช้ในการอธิบายความสัมพันธ์ได้ดีหรือไม่ กล่าวคือ

R^2 มีค่าเข้าใกล้ 1 แสดงว่าสมการเส้นถดถอย $\hat{y} = a + bx$ มีความเหมาะสมดีมาก

R^2 มีค่าเข้าใกล้ 0 แสดงว่าสมการเส้นถดถอย $\hat{y} = a + bx$ ไม่มีความเหมาะสม

ตัวอย่างการแปรความหมาย

$R^2 = 0.1$ สมการเส้นถดถอย $\hat{y} = a + bx$ ใช้อธิบายการเปลี่ยนแปลงของค่า y ได้ 10 %

$R^2 = 0.98226$ สมการเส้นถดถอย $\hat{y} = a + bx$ ใช้อธิบายการเปลี่ยนแปลงของค่า y ได้ 98.226%

Adjusted R Squares เป็นค่าที่ใช้ในการปรับปรุงค่าของ R Squares ในกรณีที่ค่า ของ n มีน้อยๆ

สูตรของ

$$\text{Adjust R Square} = 1 - \frac{(n-1)}{(n-2)}\left[\frac{\sum_{i=1}^{n}(y_i - \hat{y}_1)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}\right]$$

ขั้นตอนการคำนวณของ MATHCAD

$$\text{ycap}(x) := a + b\cdot x \qquad \text{ycap}_i := \text{ycap}(x_i) \qquad \text{ycap} = \begin{bmatrix} 4.6524 \\ 5.5315 \\ 7.2896 \\ 9.0478 \\ 10.5129 \\ 11.685 \\ 13.1502 \\ 14.3223 \\ 14.9083 \end{bmatrix}$$

$$\text{Adjusted\_R\_Square} := 1 - \left(\frac{n-1}{n-2}\right)\cdot\frac{\sum_{i=1}^{n}(y_i - \text{ycap}_i)^2}{\sum_{i=1}^{n}(y_i - \text{mean}(y))^2}$$

$$\text{Adjusted\_R\_Square} = 0.9797$$

ที่มาของค่าสถิติในตาราง ที่มาของค่าสถิติในตาราง ANOVA
จากข้อมูล X และ Y ผลบวกต่างๆ มีสูตรเป็น

SST=Sum of Squares Total $\quad SST := \sum_{i=1}^{n} (y_i)^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} \quad SST = 114.5156$

SSR=Sum of Squares Regression $\quad SSR := b \cdot S_{xy} \quad SSR = 112.4837$

SSE=Sum of Squares Residual $\quad SSE := SST - SSR \quad SSE = 2.0319$

df ของ SSR คือ 1 df ของ SSE คือ $n - 2$ df ของ SST คือ $n - 1$

$\text{Mean\_Square\_Regression} := \frac{SSR}{1} \quad \text{Mean\_Square\_Regression} = 112.4837$

$\text{Mean\_Square\_Residual} := \frac{SSE}{n-2} \quad \text{Mean\_Square\_Residual} = 0.2903$

$F := \frac{\text{Mean\_Square\_Regression}}{\text{Mean\_Square\_Residual}} \quad F = 387.5163$

ค่า Sig เป็นค่าที่คำนวณมาจากพื้นที่ใต้โค้งทางหางด้านขวาของเส้นโค้งเอฟ $v_1 = 1$ และ $v_2 = 7$ ที่ระยะ F จากค่าในตารางที่คำนวณได้

F distribution

$v1 := 1 \quad v2 := 7$

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

$TOL := 0.000000001$

$Pvalue(F) := 1 - \int_0^F h(f)\,df \quad Pvalue(387.5163) = 0.0000002181$

ค่า F และ Sig ในตาราง ANOVA ใช้ในการทดสอบสมมติฐาน $H_O : \beta = 0$ แย้งกับ $H_1 : \beta \neq 0$ ซึ่งจะได้เรียนในหัวข้อต่อๆ ไป

8.2 การหาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า β และ α

เราสามารถประมาณค่าของ β และ α โดยใช้ช่วงความเชื่อมั่นที่มีสูตรดังนี้

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า β คือ

$$b - t_{\frac{\alpha}{2}} \frac{S}{\sqrt{S_{XX}}} < \beta < b + t_{\frac{\alpha}{2}} \frac{S}{\sqrt{S_{XX}}} \qquad (df = n-2)$$

ช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า α คือ

$$a - t_{\frac{\alpha}{2}}\, s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}} < \alpha < a + t_{\frac{\alpha}{2}}\, s\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{nS_{XX}}} \qquad (df = n-2)$$

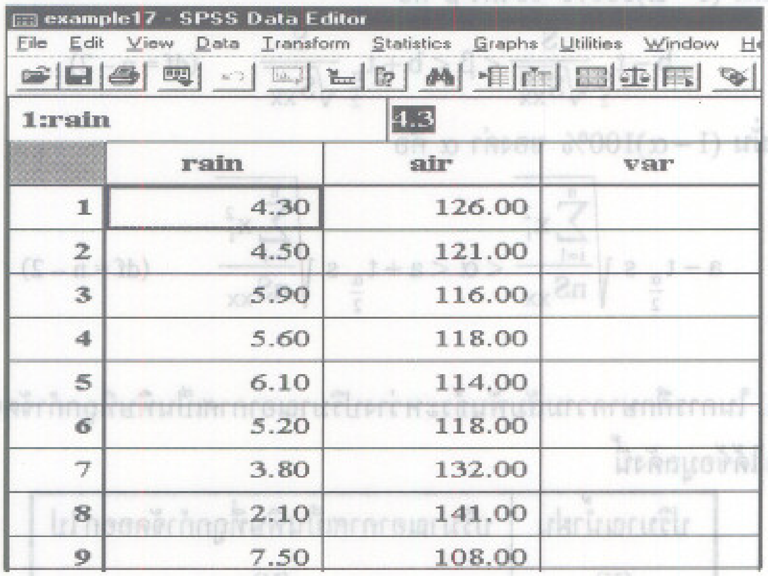
ตัวอย่าง 8.2.1. ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝนได้ข้อมูลดังนี้

ปริมาณน้ำฝน (X) (0.01 นิ้ว)	ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป (Y) (ไมโครกรัมต่อลูกบาศก์เมตร)
4.30	126.00
4.50	121.00
5.90	116.00
5.60	118.00
6.10	114.00
5.20	118.00
3.80	132.00
2.10	141.00
7.50	108.00

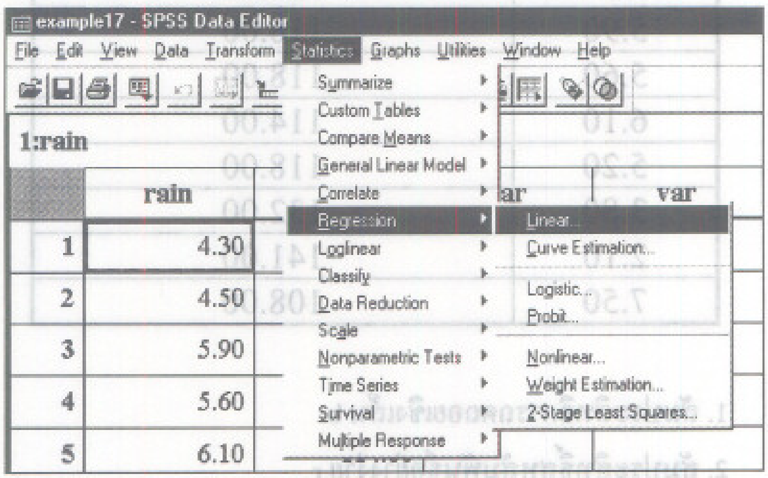
จงหาค่า
1. สัมประสิทธิ์การถดถอยเชิงเส้น b
2. สัมประสิทธิ์สหสัมพันธ์อย่างง่าย r
3. สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$
4. ช่วงความเชื่อมั่น 95% ของค่า β
5. ช่วงความเชื่อมั่น 95% ของค่า α

วิธีทำ

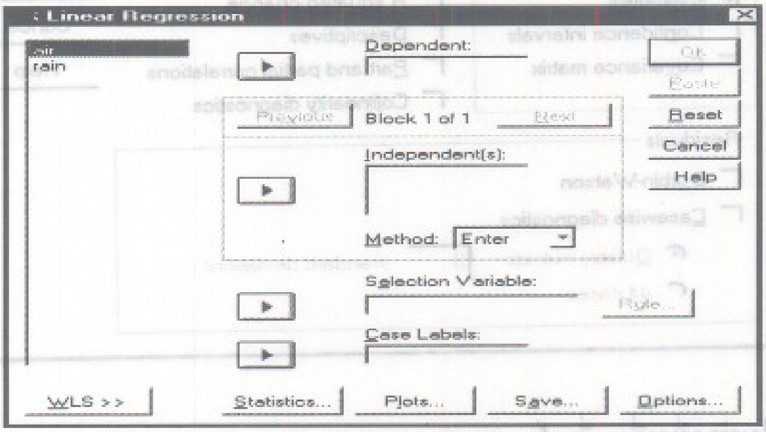
ขั้นที่ 1. นำข้อมูลเข้าสู่ SPSS Data Editor



ขั้นที่ 2 ใช้คำสั่ง Statistics / Regression / Linear...

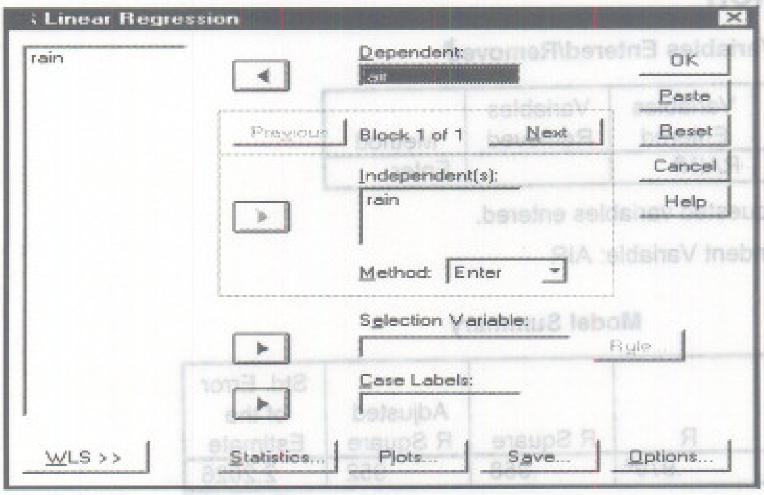


ขั้นที่ 3 คลิกที่ Linear จะได้เมนู



ขั้นที่ 4 เลือกตัวแปร air ไว้ที่ช่อง Dependent

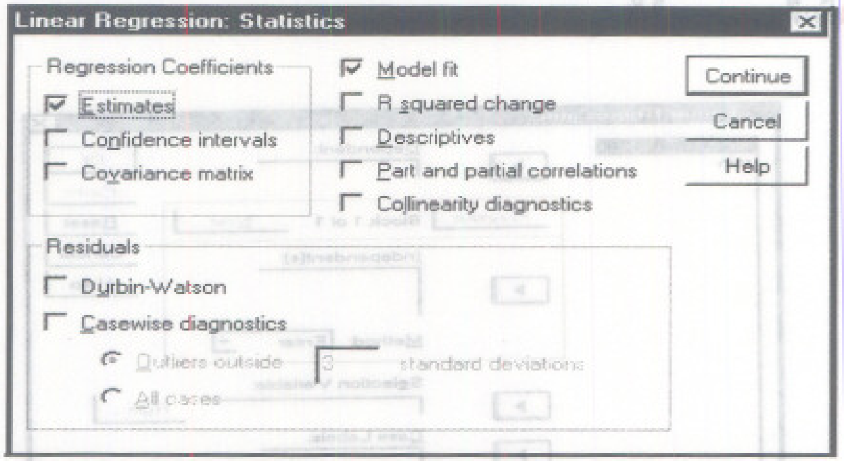
เลือกตัวแปร rain ไว้ที่ช่อง Independent



ขั้นที่ 5 เพราะว่าเราต้องการช่วงความเชื่อมั่นของ β และ α

เพราะฉะนั้นเราต้องเลือก Option **Statistics**

เมื่อคลิกที่ Statistics จะได้เมนูย่อยดังนี้



ขั้นที่ 6 คลิกในช่องสี่เหลี่ยมที่หน้า Confidence intervals

ขั้นที่ 7 คลิก Continue จะกลับไปเมนูของคำสั่ง Statistics / Regression / Linear...

ขั้นที่ 8 คลิกที่ OK จะได้ผลการคำนวณดังนี้

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	RAIN[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: AIR

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.979[a]	.958	.952	2.2026

a. Predictors: (Constant), RAIN

หมายเหตุ ค่าสัมประสิทธิ์สหสัมพันธ์ R ที่ได้จากตาราง Model Summay ต้องมีเครื่องหมายเหมือนกับค่าสัมประสิทธิ์การถดถอย b

เพราะฉะนั้น ค่าสัมประสิทธิ์สหสัมพันธ์ $r = -0.979$

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	770.262	1	770.262	158.768	.000004579292[a]
	Residual	33.961	7	4.852		
	Total	804.222	8			

a. Predictors: (Constant), RAIN

b. Dependent Variable: AIR

Coefficients[a]

		Model 1 (Constant)	Model 1 RAIN
Unstandardized Coefficients	B	153.175	-6.324
	Std. Error	2.615	.502
Standardized Coefficients	Beta		-.979
t		58.583	-12.600
Sig.		.000000000217	.000004579292
95% Confidence Interval for B	Lower Bound	146.993	-7.511
	Upper Bound	159.358	-5.137

a. Dependent Variable: AIR

สรุป
1. สัมประสิทธิ์การถดถอยเชิงเส้น $b = -6.324$
2. สัมประสิทธิ์สหสัมพันธ์อย่างง่าย $r = -0.979$
3. สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$ คือ $\hat{y} = 153.175 - 6.324x$
4. ช่วงความเชื่อมั่น 95% ของค่า β คือ $-7.511 < \beta < -5.137$
5. ช่วงความเชื่อมั่น 95% ของค่า α คือ $146.993 < \alpha < 159.358$

8.3 การทดสอบสมมติฐาน $H_0 : \rho = 0$

หลักการและขั้นตอนของการการทดสอบสมมติฐาน $H_0 : \rho = 0$

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า r

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสมคือ T

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง $t = r\sqrt{\frac{n-2}{1-r^2}}$, $df = n-2$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต

ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $T < -t_{\frac{\alpha}{2}}$ หรือ $T > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ปฏิเสธ H_0 ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$

จากตัวอย่าง 8.2.1. ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝนได้ข้อมูลดังนี้

ปริมาณน้ำฝน (X) (0.01 นิ้ว)	ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไป (Y) (ไมโครกรัมต่อลูกบาศก์เมตร)
4.30	126.00
4.50	121.00
5.90	116.00
5.60	118.00
6.10	114.00
5.20	118.00
3.80	132.00
2.10	141.00
7.50	108.00

จงทดสอบว่าตัวแปรทั้งคู่ไม่มีความสัมพันธ์กัน กำหนดระดับนัยสำคัญ 0.05

การทดสอบสมมติฐานทำดังนี้

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการคำนวณค่า $r = -0.9786$

ขั้นที่ 4. เลือกค่าสถิติ T

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง

$$t_{\text{คำนวณ}} = r\sqrt{\frac{n-2}{1-r^2}} = -0.9786\sqrt{\frac{9-2}{1-0.9786^2}} = -12.58 \quad , df = n-2 = 7$$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต

ค่าวิกฤตคือ $-t_{0.025} = -2.365$ และ $t_{0.025} = 2.365$

บริเวณวิกฤตคือ $T < -2.365$ หรือ $T > 2.365$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

เพราะว่า $t_{\text{คำนวณ}} = -12.58 < -2.365$ เพราะฉะนั้น ปฏิเสธ H_0

หลักการและขั้นตอนการทำงานเมื่อใช้ **SPSS for Windows**

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า r

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสมคือ T

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง

$$t_{\text{คำนวณ}} = r\sqrt{\frac{n-2}{1-r^2}}$$

หมายเหตุ ในกรณีที่ $\rho_0 = 0$ จะได้ว่า $t = r\sqrt{\frac{n-2}{1-r^2}}$ และ $t = \frac{b}{\left(\frac{S}{\sqrt{S_{XX}}}\right)}$ เป็นค่าเดียวกัน

ขั้นที่ 6. ใช้ค่า Sig ของค่า t ที่คำนวณได้ในการสรุปผล

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่า Sig กับ ค่า α **ถ้า Sig < α แล้วปฏิเสธ H_0**

จากตัวอย่าง 8.2.1 ในการศึกษาความสัมพันธ์ระหว่างปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝน จงทดสอบว่าตัวแปร **ปริมาณอากาศเป็นพิษที่ถูกกำจัดออกไปกับปริมาณน้ำฝน** ไม่มีความสัมพันธ์กันที่ระดับมีนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \rho = 0$

กำหนดสมมติฐานอื่น $H_1 : \rho \neq 0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและนำข้อมูลเข้าสู่การคำนวณด้วย SPSS

ขั้นที่ 4. เลือกค่าสถิติที่เหมาะสมคือ T

ขั้นที่ 5. จากผลการคำนวณของ SPSS ข้างต้น

Coefficients[a]

		Model	
		1	
		(Constant)	RAIN
Unstandardized Coefficients	B	153.175	-6.324
	Std. Error	2.615	.502
Standardized Coefficients	Beta		-.979
t		58.583	-12.600
Sig.		.000000000217	.000004579292
95% Confidence Interval for B	Lower Bound	146.993	-7.511
	Upper Bound	159.358	-5.137

a. Dependent Variable: AIR

คำนวณค่าสถิติที่ได้จากข้อมูลตัวอย่าง $t = r\sqrt{\frac{n-2}{1-r^2}} = -12.600$

โดยมีค่า Sig = 0.000004579292 df = 7

ขั้นที่ 6. Sig = 0.000004579292

ขั้นที่ 7. **สรุปผล** เพราะว่า Sig < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

8.4 การทดสอบสมมติฐาน $H_0 : \beta = \beta_0$

หลักการและขั้นตอนการทำงานทางทฤษฎี

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = \beta_0$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq \beta_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่า b, s, S_{XX}

ขั้นที่ 4. เลือกค่าสถิติ T

ขั้นที่ 5. คำนวณค่าสถิติที่เลือกจากข้อมูลตัวอย่าง $t_{\text{คำนวณ}} = \dfrac{b-\beta_0}{\left(\dfrac{s}{\sqrt{S_{XX}}}\right)}$, $df = n-2$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤต

ค่าวิกฤตคือ $-t_{\frac{\alpha}{2}}$ และ $t_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $T < -t_{\frac{\alpha}{2}}$ หรือ $T > t_{\frac{\alpha}{2}}$

ขั้นที่ 7. สรุปผลโดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

ปฏิเสธ H_0 ถ้า $t_{\text{คำนวณ}} < -t_{\frac{\alpha}{2}}$ หรือ $t_{\text{คำนวณ}} > t_{\frac{\alpha}{2}}$

หลักการและขั้นตอนการทำงานด้วย SPSS และวิธีสรุปผล

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = \beta_0$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq \beta_0$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ α

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าโดยการนำข้อมูลเข้าสู่โปรแกรม SPSS

ขั้นที่ 4. เพราะว่าผลการคำนวณของ SPSS ไม่ให้ค่า $t = \dfrac{b-\beta_0}{\left(\dfrac{s}{\sqrt{S_{XX}}}\right)}$ ออกมาโดยตรง

เพราะฉะนั้นเราจึงใช้ช่วงความเชื่อมั่นของ β ช่วยในการสรุปสมมติฐาน

ขั้นที่ 5. ให้หาช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า β

ขั้นที่ 6. ไม่มีการเปิดตารางสถิติเพื่อหาค่าวิกฤต

ขั้นที่ 7. สรุปผลโดยการดูว่า β_0 อยู่ในช่วงความเชื่อมั่น $(1-\alpha)100\%$ ของค่า β ที่หาได้หรือไม่

ถ้า β_0 อยู่ในช่วงความเชื่อมั่นที่หาได้ แล้ว ยอมรับ H_0

จากตัวอย่างข้อมูล

X	1.50	1.80	2.40	3.00	3.50	3.90	4.40	4.80	5.00
Y	4.80	5.70	7.00	8.30	10.90	12.40	13.10	13.60	15.30

มีสมการถดถอยเป็น $\hat{y} = a + bx$

จงทดสอบสมมติฐานว่า $\beta = 2.5$ แย้งกับสมมติฐาน $\beta \neq 2.5$ ที่ระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \beta = 2.5$

กำหนดสมมติฐานอื่น $H_1 : \beta \neq 2.5$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างขนาด n และ คำนวณค่าโดยการนำข้อมูลเข้าสู่โปรแกรม SPSS

ขั้นที่ 4. เพราะว่าผลการคำนวณของ SPSS ไม่ให้ค่า $t = \dfrac{b-\beta_0}{\left(\dfrac{s}{\sqrt{S_{xx}}}\right)}$ ออกมาโดยตรง

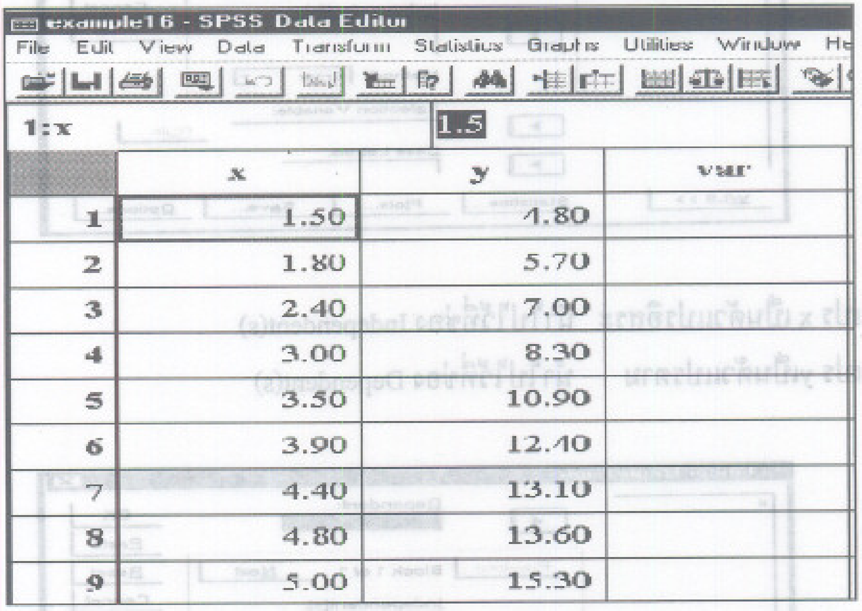
เพราะฉะนั้นเราจึงใช้ช่วงความเชื่อมั่นของ β ช่วยในการสรุปสมมติฐาน

ขั้นที่ 5. เพราะว่า $H_1 : \beta \neq \beta_0$

เพราะฉะนั้น ให้หาช่วงความเชื่อมั่น 95% ของค่า β

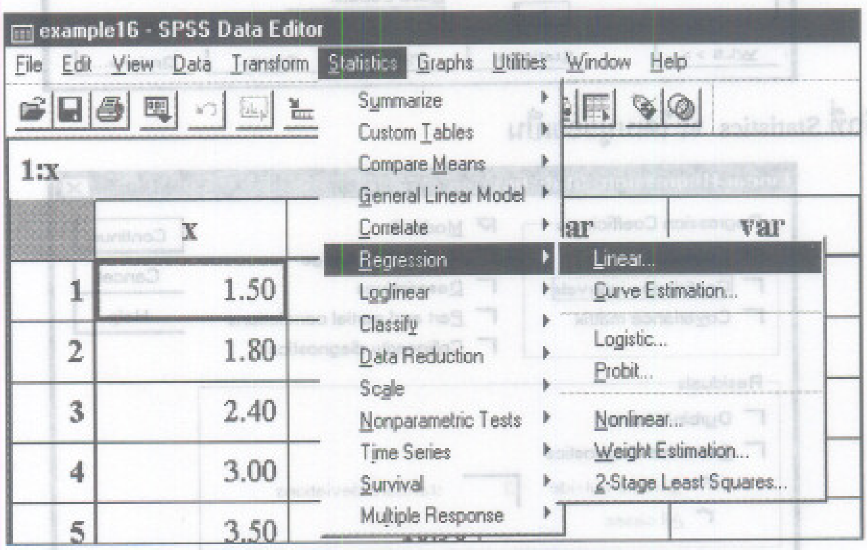
การหาช่วงความเชื่อมั่น 95% ของ β

นำข้อมูลเข้าสู่ SPSS

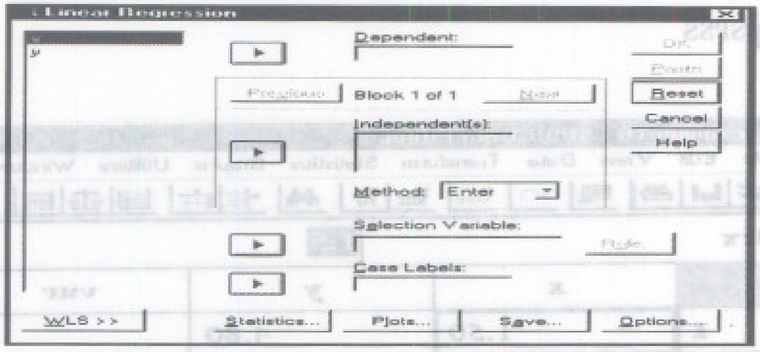


	x	y	var
1	1.50	4.80	
2	1.80	5.70	
3	2.40	7.00	
4	3.00	8.30	
5	3.50	10.90	
6	3.90	12.40	
7	4.40	13.10	
8	4.80	13.60	
9	5.00	15.30	

ใช้คำสั่ง Statistics / Regression / Linear..

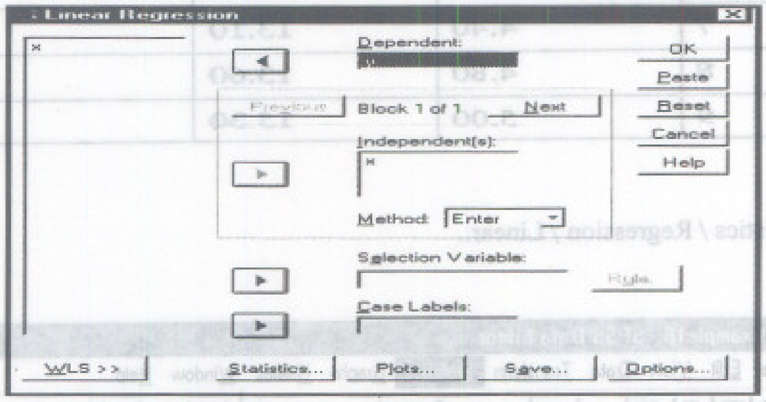


เมื่อคลิกที่ Linear จะได้เมนูของคำสั่งดังนี้

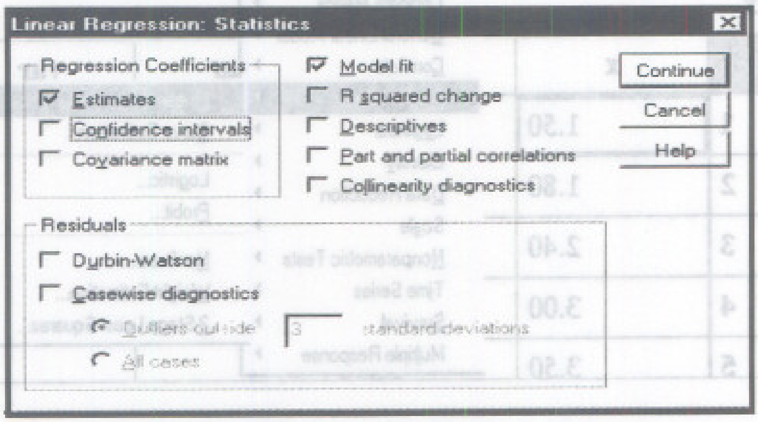


เลือกตัวแปร x เป็นตัวแปรอิสระ นำไปไว้ที่ช่อง Independent(s)

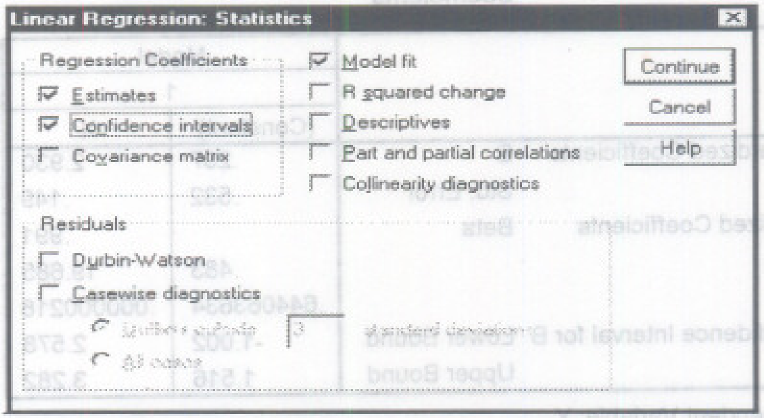
เลือกตัวแปร yเป็นตัวแปรตาม นำไปไว้ที่ช่อง Dependent(s)



ต่อไปคลิกที่ Statistics จะได้เมนูย่อยเป็น



คลิกในกรอบสี่เหลี่ยมหน้าช่อง Confidence interval



เสร็จแล้ว Continue และ OK ตามลำดับ จะได้ผลการคำนวณดังนี้

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	X[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Y

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.991[a]	.982	.980	.5388

a. Predictors: (Constant), X

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	112.484	1	112.484	387.516	.000000218[a]
	Residual	2.032	7	.290		
	Total	114.516	8			

a. Predictors: (Constant), X

b. Dependent Variable: Y

Coefficients[a]

		Model	
		1	
		(Constant)	X
Unstandardized Coefficients	B	.257	2.930
	Std. Error	.532	.149
Standardized Coefficients	Beta		.991
t		.483	19.685
Sig.		.644063634	.000000218
95% Confidence Interval for B	Lower Bound	-1.002	2.578
	Upper Bound	1.516	3.282

a. Dependent Variable: Y

ขั้นที่ 6. ไม่มีการเปิดตารางสถิติเพื่อหาค่าวิกฤต

ขั้นที่ 7. สรุปผล โดยการดูว่า $\beta_0 = 2.5$ อยู่ในช่วงความเชื่อมั่นที่หาได้หรือไม่

จากผลการคำนวณช่วงความเชื่อมั่น 95% ของ β คือ (2.578 , 3.282)

เพราะว่า $\beta_0 = 2.5$ ไม่อยู่ในช่วงความเชื่อมั่น 95% ของ β คือ (2.578 , 3.282)

เพราะฉะนั้นปฏิเสธ H_0

8.5 การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์และสมการถดถอยพหุคูณ

8.5.1 การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรหลายๆ คู่

ในกรณีที่เรามีตัวแปรหลายคู่ที่ต้องการหาค่าสัมประสิทธิสหสัมพันธ์

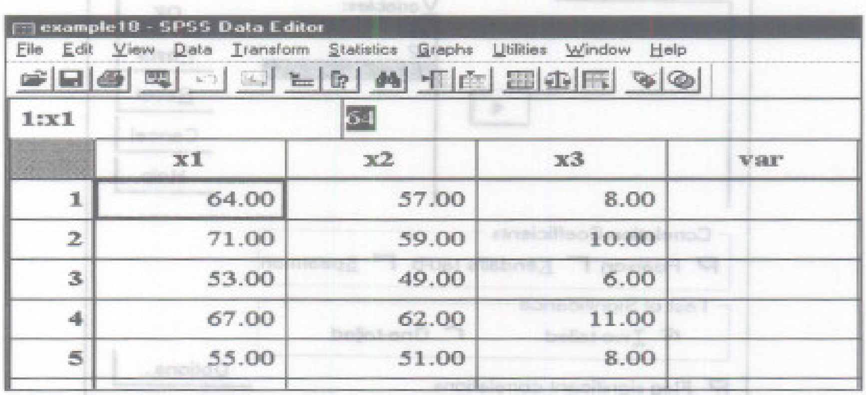
เช่นข้อมูล น้ำหนัก (X_1) , ความสูง(X_2) , อายุ(X_3)

x1	x2	x3
64.00	57.00	8.00
71.00	59.00	10.00
53.00	49.00	6.00
67.00	62.00	11.00
55.00	51.00	8.00
58.00	50.00	7.00
77.00	55.00	10.00
57.00	48.00	9.00
56.00	52.00	10.00
51.00	42.00	6.00
76.00	61.00	12.00
68.00	57.00	9.00

การหาค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรหลายคู่พร้อมกันด้วย SPSS

ขั้นที่ 1 นำข้อมูลเข้าสู่ SPSS

หมายเหตุ แฟ้มข้อมูลชื่อ example18.sav

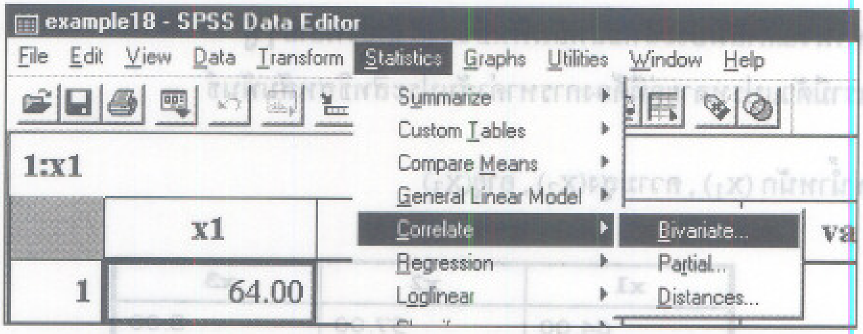

example18 - SPSS Data Editor

File Edit View Data Transform Statistics Graphs Utilities Window Help

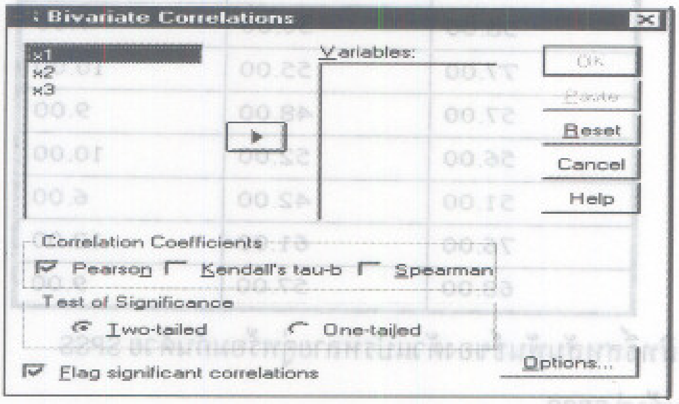
1:x1

	x1	x2	x3	var
1	64.00	57.00	8.00	
2	71.00	59.00	10.00	
3	53.00	49.00	6.00	
4	67.00	62.00	11.00	
5	55.00	51.00	8.00	

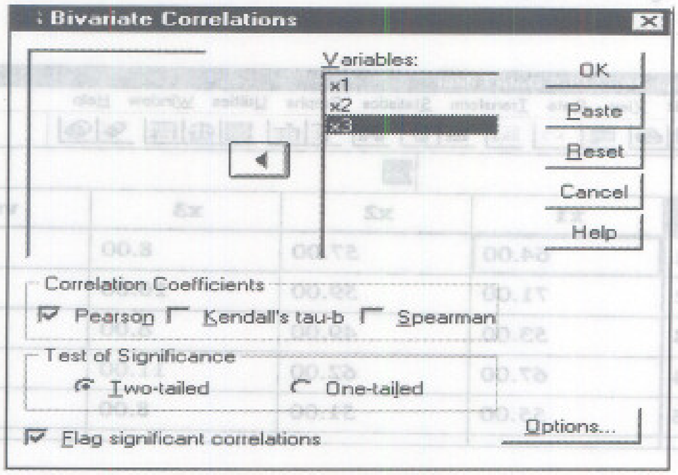
ขั้นที่ 2. เลือกคำสั่ง Statistics / Correlate / Bivariate..



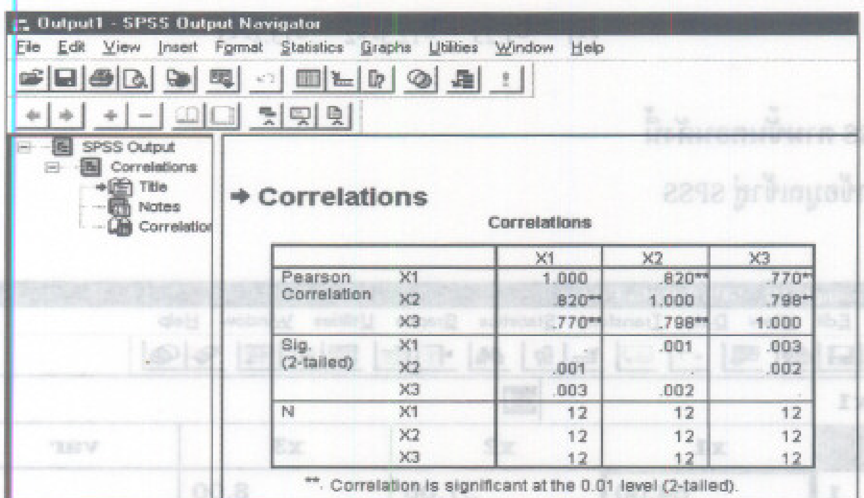
จะได้เมนูของคำสั่งดังนี้



ขั้นที่ 3. เลือกตัวแปร x_1 , x_2 และ x_3 มาไว้ที่ช่อง Variables



ขั้นที่ 4. ต่อไปคลิกที่ OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

Correlations

Correlations

		X1	X2	X3
Pearson Correlation	X1	1.000	.820**	.770**
	X2	.820**	1.000	.798**
	X3	.770**	.798**	1.000
Sig. (2-tailed)	X1	.	.001	.003
	X2	.001	.	.002
	X3	.003	.002	.
N	X1	12	12	12
	X2	12	12	12
	X3	12	12	12

**. Correlation is significant at the 0.01 level (2-tailed).

การแปลความหมาย

ค่าสหสัมพันธ์ของ น้ำหนัก (X_1) , ความสูง(X_2) เท่ากับ 0.820

ค่าสหสัมพันธ์ของ น้ำหนัก (X_1) , อายุ(X_3) เท่ากับ 0.770

ค่าสหสัมพันธ์ของ ความสูง(X_2) , อายุ(X_3) เท่ากับ 0.798

เพราะฉะนั้น น้ำหนัก (X_1) , ความสูง(X_2) มีความสัมพันธ์กันมากที่สุด

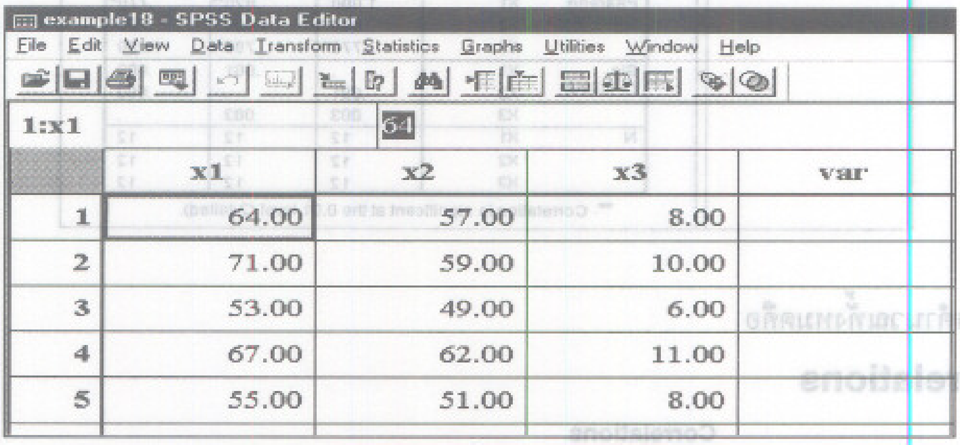
8.5.2. การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณและสมการถดถอยพหุคูณ

จากข้อมูลข้างต้น เราสามารถหาสมการ

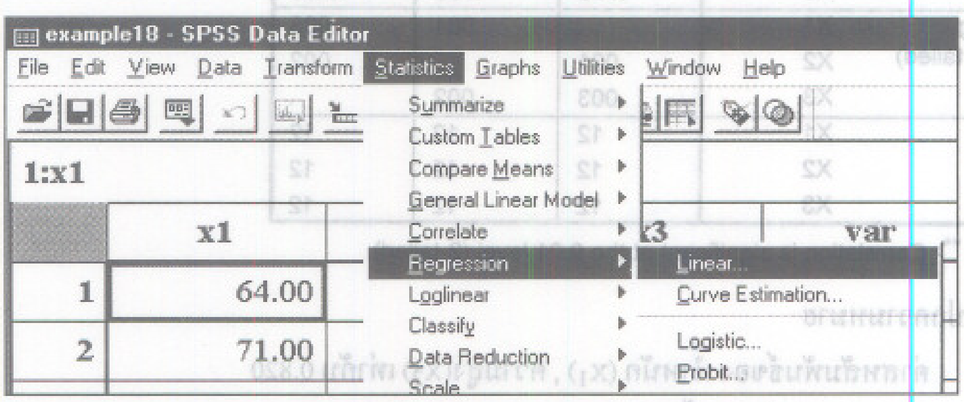
$$X_1 = b_{1.23} + b_{12.3}X_2 + b_{13.2}X_3$$

โดยใช้ SPSS ตามขั้นตอนดังนี้

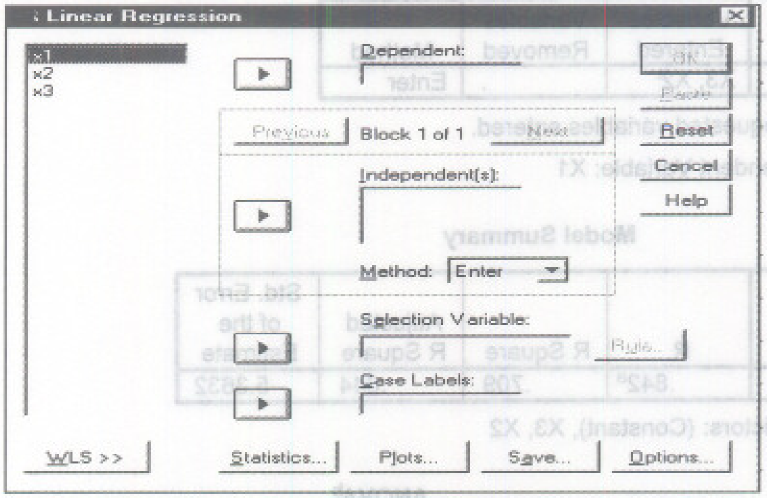
ขั้นที่ 1 นำข้อมูลเข้าสู่ SPSS



ขั้นที่ 2. เลือกคำสั่ง Statistics / Regression /Linear..

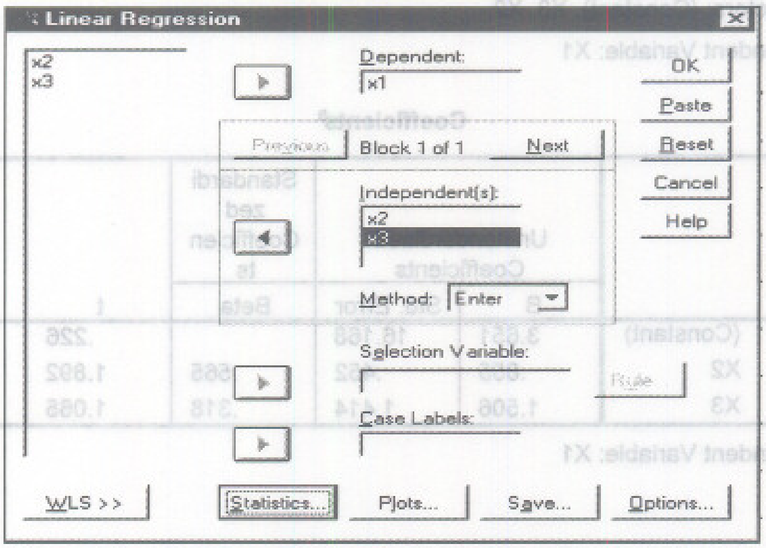


ขั้นที่ 3. คลิกที่ Liner จะได้เมนูย่อยเป็น



ขั้นที่ 4. เลือกตัวแปร x_1 ไปไว้ที่ช่อง dependent

เลือกตัวแปร x_2 ,x_3 ไปไว้ที่ช่อง Independent



ขั้นที่ 5.คลิก OK จะได้ผลการคำนวณเป็น

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	X3, X2[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: X1

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.842[a]	.709	.644	5.3632

a. Predictors: (Constant), X3, X2

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	629.373	2	314.687	10.940	.004[a]
	Residual	258.877	9	28.764		
	Total	888.250	11			

a. Predictors: (Constant), X3, X2

b. Dependent Variable: X1

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.651	16.168		.226	.826
	X2	.855	.452	.565	1.892	.091
	X3	1.506	1.414	.318	1.065	.315

a. Dependent Variable: X1

ความหมายของผลการคำนวณที่ได้คือ $X_1 = b_{1.23} + b_{12.3}X_2 + b_{13.2}X_3$

จากตาราง Coefficient จะได้ $b_{1.23} = 3.651$, $b_{12.3} = 0.855$, $b_{13.2} = 1.506$

เพราะฉะนั้นสมการถดถอยคือ $X_1 = 3.651 + 0.855\ X_2 + 1.506\ X_3$

8.6 การเลือกรูปแบบความสัมพันธ์แบบเชิงเดียวที่เหมาะสมกับข้อมูล

ความสัมพันธ์แบบเชิงเดียว ซึ่งประกอบด้วยตัวแปรอิสระ 1 ตัว และตัวแปรตาม 1 ตัว รูปแบบของสมการความสัมพันธ์เชิงเส้นเชิงเดียวอาจมีรูปแบบเป็น

- $y = a + bx$
- $\ln y = a + b \ln x$
- $y = a + b \ln x$
- $\ln y = a + bx$

เมื่อเรามีข้อมูลและต้องการรู้ว่ารูปแบบใดเหมาะสมกับข้อมูล สามารถใช้โปรแกรม MATHCAD ช่วยในการเขียนกราฟและคำนวณค่าสหสัมพันธ์ได้ดังนี้

จากตัวอย่างข้อมูล

X	Y
1.52	40.8
1.85	52.7
2.48	74.0
3.06	85.3
3.53	100.9
3.97	121.4
4.44	130.1
4.85	135.6
5.09	150.3

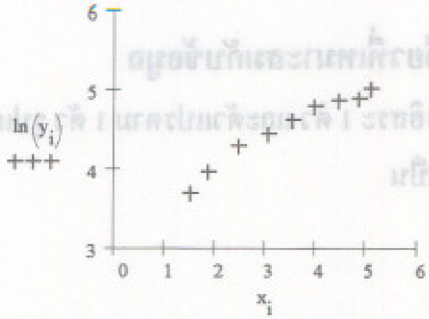
เราสามารถเขียนแผนภาพการกระจาย 4 รูปแบบ และคำนวณค่าสหสัมพันธ์ได้ ดังนี้

การคำนวณด้วย MATHCAD

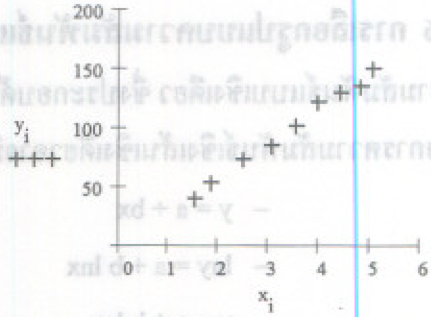
$$x := \begin{bmatrix} 1.52 \\ 1.85 \\ 2.48 \\ 3.06 \\ 3.53 \\ 3.97 \\ 4.44 \\ 4.85 \\ 5.09 \end{bmatrix} \qquad y := \begin{bmatrix} 40.8 \\ 52.7 \\ 74.0 \\ 85.3 \\ 100.9 \\ 121.4 \\ 130.1 \\ 135.6 \\ 150.3 \end{bmatrix}$$

$i := 1..9$

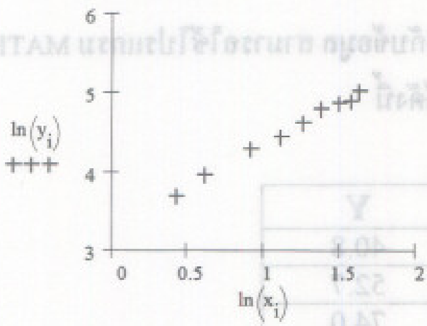
ORIGIN:= 1



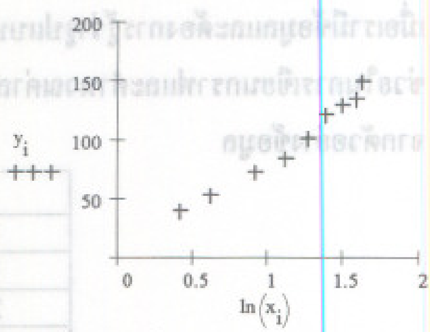
$\mathrm{corr}\left(x, \overrightarrow{\ln(y)}\right) = 0.9791$



$\mathrm{corr}(x, y) = 0.9959$



$\mathrm{corr}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right) = 0.997$



$\mathrm{corr}\left(\overrightarrow{\ln(x)}, y\right) = 0.9861$

รูปแบบความสัมพันธ์	ค่าสัมประสิทธิ์สหสัมพันธ์
y = a + bx	0.9959
lny = a + b lnx	0.9970
y = a + b lnx	0.9861
lny = a + bx	0.9791

เราควรเลือกรูปแบบที่มีค่าสัมประสิทธิ์สหสัมพันธ์มากที่สุด นั่นคือ lny = a + b lnx

$a := \mathrm{intercept}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right)$ $\qquad$ $b := \mathrm{slope}\left(\overrightarrow{\ln(x)}, \overrightarrow{\ln(y)}\right)$

$a = 3.3057$ $\qquad$ $b = 1.0461$

สรุป สมการแสดงความสัมพันธ์ที่เหมาะสมกับข้อมูลคือ

$\ln y = a + b \ln x$

$= 3.3057 + 1.0461 \ln x$

แบบฝึกหัด 8.

1. การทดลองเพื่อตรวจสอบความสัมพันธ์ระหว่างการไหล (x) กับประสิทธิภาพ (y) ของเครื่องสูบใบพัดแบบหนึ่งมีข้อมูลต่อไปนี้

x (แกลลอนต่อนาที)	3.5	7.0	10.5	14.0	17.5	21.0
y (เปอร์เซ็นต์)	1.6	3.0	4.6	5.1	6.1	7.3

จงหาค่า
1. สัมประสิทธิ์การถดถอยเชิงเส้น b
2. สัมประสิทธิ์สหสัมพันธ์อย่างง่าย r
3. สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$
4. ช่วงความเชื่อมั่น 95% ของค่า β
5. ช่วงความเชื่อมั่น 95% ของค่า α

2. บริษัทประกันชีวิตแห่งหนึ่ง ต้องการตรวจสอบการมีความสัมพันธ์เชิงเส้นเพื่อใช้พยากรณ์เงินขายประกันต่อปี (แสนบาท) ของผู้ขายประกัน 10 คน ซึ่งมีคะแนนการสอบความถนัดดังนี้

ผู้ขายประกัน	1	2	3	4	5	6	7	8	9	10
คะแนน	90	75	80	65	85	90	95	75	70	60
เงินขายได้	1.1	1.2	1.6	0.7	2.1	2.7	2.4	0.7	0.6	0.9

จงหาค่า
1. สัมประสิทธิ์การถดถอยเชิงเส้น b
2. สัมประสิทธิ์สหสัมพันธ์อย่างง่าย r
3. สมการของเส้นถดถอยเชิงเส้น $\hat{y} = a + bx$
4. ช่วงความเชื่อมั่น 95% ของค่า β
5. ช่วงความเชื่อมั่น 95% ของค่า α

3. จงคำนวณสัมประสิทธิ์สหสัมพันธ์อย่างง่ายของข้อมูล

x คะแนนคณิตศาสตร์	70	92	80	74	65	83
y คะแนนภาษาอังกฤษ	74	84	63	87	78	90

และจงทดสอบว่าคะแนนคณิตศาสตร์และคะแนนภาษาอังกฤษไม่มีความสัมพันธ์กันที่ระดับมีนัยสำคัญ 0.05

4. จงหาสัมประสิทธิ์สหสัมพันธ์เชิงเส้นของตัวแปรทุกคู่ ต่อไปนี้

X1	X2	X3
13	160	9
15	100	7
16	70	5
18	40	4
22	30	3
24	20	2

5. กำหนดข้อมูล X และ Y ดังนี้

X	Y
13	185
25	195
37	237
64	364
125	543

5.1 จงเขียนแผนภาพการกระจายบนสเกล

(X , Y) , (X , ln(Y)) , (ln(X),Y) , (ln(X) , ln(Y))

5.2. จงคำนวณหาค่าสหสัมพันธ์ทั้ง 4 รูปแบบ

5.3 จงหาสมการแสดงความสัมพันธ์เชิงเส้นทั้ง 4 รูปแบบ

	รูปแบบสมการ	สหสัมพันธ์
(X , Y)		
(ln(X) , Y)		
(X , ln(Y))		
(ln(X) , ln(Y))		

บทที่ 9

การวิเคราะห์ความแปรปรวน

การวิเคราะห์ความแปรปรวน เป็นการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากรตั้งแต่ 2 ชุดขึ้นไปเท่ากันหรือไม่

9.1 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว

ขั้นตอนการทดสอบสมมติฐาน $\mu_1 = \mu_2 = \mu_3 = \dots\dots\dots = \mu_k$

ขั้นที่ 1 กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \dots\dots\dots = \mu_k$

กำหนดสมมติฐานทางเลือก $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots\dots\dots \neq \mu_k$

ขั้นที่ 2 กำหนดนัยสำคัญ α

ขั้นที่ 3 ทำการสุ่มตัวอย่าง

ขั้นที่ 4 เลือกค่าสถิติ F

ขั้นที่ 5 คำนวณค่าสถิติ F จากตัวอย่าง (สร้างตาราง ANOVA)

ขั้นที่ 6 เปิดตารางหาค่าวิกฤต

ค่าวิกฤตคือ f_α องศาความเสรี $v_1 = k-1$, $v_2 = N-k$

N = จำนวนค่าสังเกตทั้งหมด

บริเวณวิกฤตคือ $F > f_\alpha$

ขั้นที่ 7 สรุปผล

ถ้า $f_{\text{คำนวณ}} > f_\alpha$ แล้ว ปฏิเสธ H_0

ตัวอย่าง 9.1.1 ข้อมูลการดูดความชื้นของคอนกรีต 5 ชนิดเป็นดังนี้

ชนิดที่ 1	ชนิดที่ 2	ชนิดที่ 3	ชนิดที่ 4	ชนิดที่ 5
551.00	595.00	639.00	417.00	563.00
457.00	580.00	615.00	449.00	631.00
450.00	508.00	511.00	517.00	522.00
731.00	583.00	573.00	438.00	613.00
499.00	633.00	648.00	415.00	656.00
632.00	517.00	677.00	555.00	679.00

กำหนด $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ เป็นค่าเฉลี่ยประชากรของการดูดความชื้นของคอนกรีต 5 ชนิด จงทดสอบสมมติฐานว่า $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ เท่ากัน ที่ระดับนัยสำคัญ 0.05

วิธีทำ

ขั้นที่ 1 $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5$

ขั้นที่ 2 กำหนดนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3 ทำการสุ่มตัวอย่าง

ขั้นที่ 4 เลือกค่าสถิติ F

$T_{..}$ = ผลรวมค่าสังเกตทั้งหมด

y_{ij} = ค่าสังเกตตัวที่ i หลักที่ j

$T_{i.}$ = ผลรวมค่าสังเกตใน treatment ที่ i

$n_{i.}$ = จำนวนค่าสังเกตใน treatment ที่ i

คำนวณค่า $$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{T_{..}^2}{N} = 209377$$

$$SSA = \sum_{i=1}^{k}\frac{T_{i.}^2}{n_i} - \frac{T_{..}^2}{N} = 85356$$

$$SSE = SST - SSA = 124021$$

ขั้นที่ 5 คำนวณค่าสถิติ F จากตัวอย่าง **(สร้างตาราง ANOVA)**

แหล่งการแปรผัน	**ผลบวกกำลังสอง**	**องศาเสรี**	**ค่าเฉลี่ยของผลบวกกำลังสอง**	$f_{คำนวณ}$
วิธีการปฏิบัติ (Treatment)	SSA = 85365	k – 1 = 4	21339	4.30
ความคลาดเคลื่อน (Error)	SSE = 124021	k(n – 1) = 25	4961	
ทั้งหมด (Total)	SST = 209377	nk – 1		

ขั้นที่ 6 เปิดตารางหาค่าวิกฤต

ค่าวิกฤตคือ $f_{0.05} = 2.76$ องศาความเสรี $v_1 = 4$, $v_2 = 25$

บริเวณวิกฤตคือบริเวณที่ $F > 2.76$

ขั้นที่ 7 สรุปผล

เพราะว่า $f_{คำนวณ} = 4.30 > 2.76$

เพราะฉะนั้น ปฏิเสธ H_0

การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวด้วย SPSS

ขั้นที่ 1 กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = ... = \mu_k$

กำหนดสมมติฐานทางเลือก $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq .... \neq \mu_k$

ขั้นที่ 2 กำหนดนัยสำคัญ α

ขั้นที่ 3 ทำการสุ่มตัวอย่าง

ขั้นที่ 4 เลือกค่าสถิติ F

ขั้นที่ 5 คำนวณค่าสถิติ F จากตัวอย่าง (สร้างตาราง ANOVA)

5.1 นำค่า $f_{คำนวณ}$ ไปใช้ในการสรุปผล

5.2 นำค่า Sig ไปใช้ในการสรุปผล

หมายเหตุ Sig = พื้นที่ใต้โค้งความน่าจะเป็นของตัวแปรสุ่ม F ทางหางด้านขวา ตั้งแต่ $f_{\text{คำนวณ}}$ ถึง ∞

ขั้นที่ 6 6.1 ใช้ค่าเปิดตารางหาค่าวิกฤต

ค่าวิกฤตคือ f_α องศาความเสรี $v_1 = k - 1$, $v_2 = N - k$

N = จำนวนค่าสังเกตทั้งหมด

บริเวณวิกฤตคือ $F > f_\alpha$

6.2 ใช้ค่า Sig ในการสรุปผล

ขั้นที่ 7 สรุปผล

แบบที่ 1 ถ้า $f_{\text{คำนวณ}} > f_\alpha$ แล้ว ปฏิเสธ H_0

แบบที่ 2 ถ้า Sig < α แล้ว ปฏิเสธ H_0

จากข้อมูลตัวอย่าง 9.1.1

จงทดสอบสมมติฐานว่า $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ เท่ากัน ที่ระดับนัยสำคัญ 0.05

การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวด้วย SPSS

ขั้นที่ 1 กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots\ldots = \mu_5$

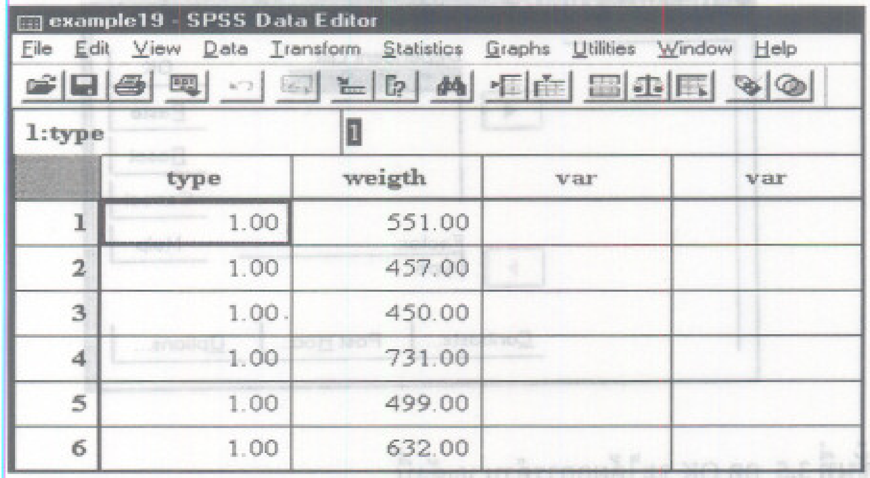
กำหนดสมมติฐานทางเลือก $H_1 : \mu_1 \neq \mu_2 \neq \mu_3 \neq \ldots\ldots \neq \mu_5$

ขั้นที่ 2 กำหนดนัยสำคัญ $\alpha = 0.05$

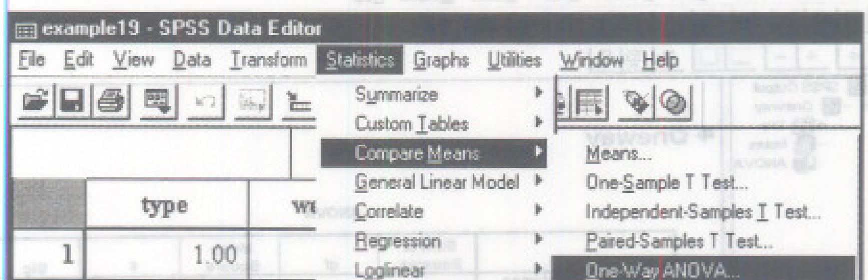
ขั้นที่ 3 ทำการสุ่มตัวอย่างและสร้างแฟ้มข้อมูล

ขั้นที่ 3.1 การทำแฟ้มข้อมูลต้องมีตัวแปร type เป็นตัวแปรจำแนกกลุ่ม และ weight เป็นตัวแปรน้ำหนักของคอนกรีต

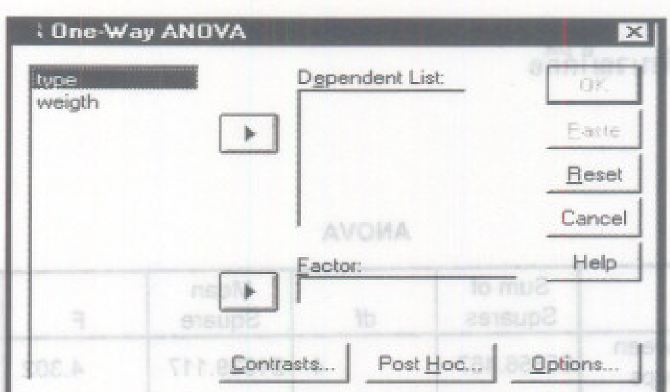
หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example19.sav



ขั้นที่ 3.2 ใช้คำสั่ง Statistics / Compare Means / One-Way ANOVA...

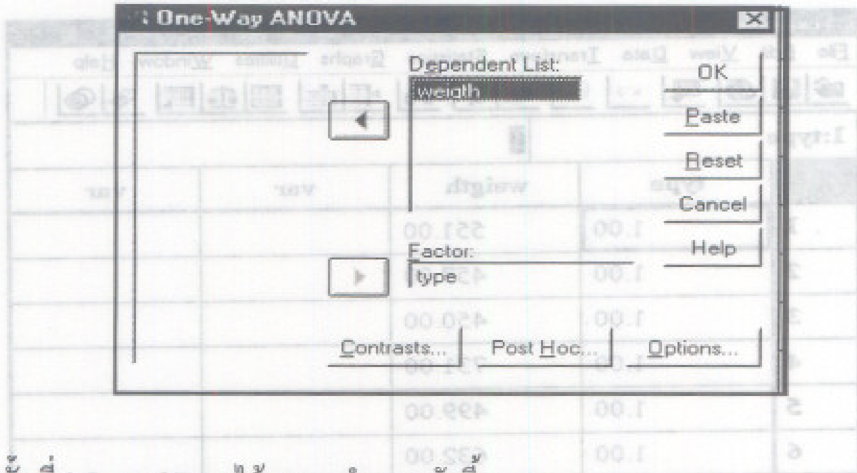


ขั้นที่ 3.3 คลิกที่ One-Way ANOVA จะได้เมนูย่อยเป็น

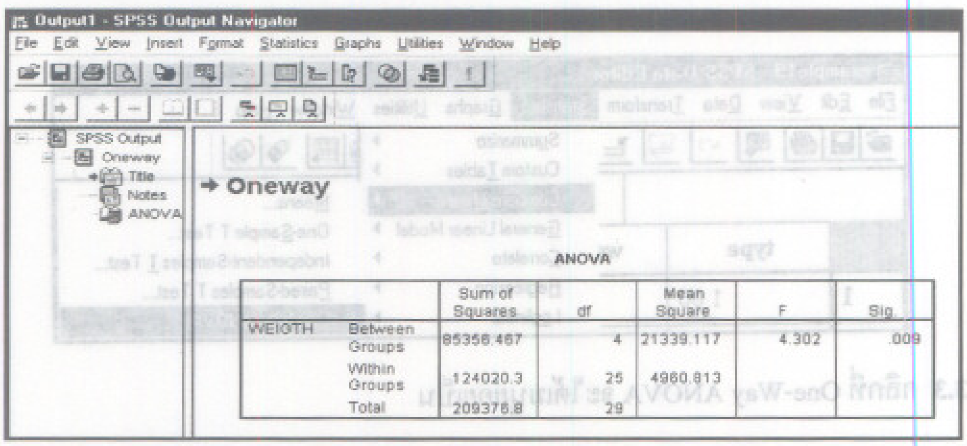


ขั้นที่ 3.4 เลือกตัวแปร type ไปไว้ที่ช่อง Factor

เลือกตัวแปร weigth ไปไว้ที่ช่อง Dependent List



ขั้นที่ 3.5 กด OK จะได้ผลการคำนวณดังนี้



ตาราง ANOVA ที่คำนวณได้คือ

Oneway

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
WEIGTH	Between Groups	85356.467	4	21339.117	4.302	.009
	Within Groups	124020.3	25	4960.813		
	Total	209376.8	29			

ขั้นที่ 4 สรุปผลโดยใช้ค่า Sig หรือ เปรียบเทียบค่า $f_{\text{คำนวณ}}$ กับค่าวิกฤต

ขั้นที่ 5 จากตาราง ANOVA $f_{\text{คำนวณ}} = 4.302$, Sig = 0.009

ขั้นที่ 6 เปิดตารางหาค่าวิกฤต

ค่าวิกฤตคือ $f_{0.05} = 2.76$ องศาความเสรี $v_1 = 4$, $v_2 = 25$

บริเวณวิกฤตคือบริเวณที่ $F > 2.76$

ขั้นที่ 7 สรุปผล

แบบที่ 1 เพราะว่า $f_{\text{คำนวณ}} = 4.302 > 2.76$ เพราะฉะนั้น ปฏิเสธ H_0

หรือ แบบที่ 2 เพราะว่า Sig = 0.009 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ 1. ในทางปฏิบัติการสรุปผลโดยดูค่า Sig มีความสะดวกกว่า

2. ที่มาของค่า Sig คือ

F distribution

$v1 := 4$

$v2 := 25$

$$h(f) := \frac{\Gamma\left(\frac{v1+v2}{2}\right)\cdot\left(\frac{v1}{v2}\right)^{\frac{v1}{2}}\cdot f^{\left(\frac{v1}{2}\right)-1}}{\Gamma\left(\frac{v1}{2}\right)\cdot\Gamma\left(\frac{v2}{2}\right)\cdot\left[1+\left(\frac{v1}{v2}\right)\cdot f\right]^{\frac{v1+v2}{2}}}$$

$$\int_{4.302}^{1000} h(f)\,df = 0.008747$$

9.2 การวิเคราะห์ความแปรปรวนแบบจำแนกสองทาง

การวิเคราะห์ความแปรปรวนเมื่อข้อมูลมีการสุ่มอย่างสมบูรณ์ในแต่ละกลุ่มลักษณะข้อมูลคือ

		วิธีการปฏิบัติ (Treatment)			
		1	**2**	**... i ...**	**k**
กลุ่ม (Block)	**1**	y_{11}	y_{21}	y_{i1}	y_{k1}
	2	y_{12}	y_{22}	y_{i2}	y_{k2}
				:	
	j	y_{1j}	y_{2j}	y_{ij}	y_{kj}
				:	
	b	y_{1b}	y_{2b}	y_{ib}	y_{kb}

จากข้อมูลที่เก็บมาได้ทำการคำนวณต่อไปดังนี้

		วิธีการปฏิบัติ (Treatment)					
		1	**2**	**.. i ...**	**k**	**รวม**	**ค่าเฉลี่ย**
กลุ่ม (Block)	**1**	y_{11}	y_{21}	y_{i1}	y_{k1}	$T_{.1}$	$\bar{y}_{.1}$
	2	y_{12}	y_{22}	y_{i2}	y_{k2}	$T_{.2}$	$\bar{y}_{.2}$
				:			
	j	y_{1j}	y_{2j}	y_{ij}	y_{kj}	$T_{.j}$	$\bar{y}_{.j}$
				:			
	b	y_{1b}	y_{2b}	y_{ib}	y_{kb}	$T_{.b}$	$\bar{y}_{.b}$
	รวม	$T_{1.}$	$T_{2.}$	$T_{i.}$	$T_{k.}$	$T_{..}$	
	ค่าเฉลี่ย	$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\bar{y}_{i.}$	$\bar{y}_{k.}$		

y_{ij} = ค่าสังเกตจาก treatment ที่ i และ Block ที่ j

$\mu_{i.}$ = ค่าเฉลี่ยประชากรของ Treatment ที่ i

$\mu_{.j}$ = ค่าเฉลี่ยประชากรของ Block ที่ j

$T_{i.}$ = ผลรวมของค่าสังเกตจาก Treatment ที่ i

$T_{.j}$ = ผลรวมของค่าสังเกตจาก Block ที่ j

$T_{..}$ = ผลรวมของค่าสังเกตทั้งหมด

ขั้นตอนของการทดสอบสมมติฐาน

ขั้นที่ 1 สมมติฐานเกี่ยวกับวิธีปฏิบัติการ (Treatment)

$H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.} = \ldots\ldots = \mu_{k.}$

$H_1 : \mu_{1.} \neq \mu_{2.} \neq \mu_{3.} \neq \ldots\ldots \neq \mu_{k.}$

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม (Block)

$H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3} = \ldots\ldots = \mu_{.b}$

$H_1 : \mu_{.1} \neq \mu_{.2} \neq \mu_{.3} \neq \ldots\ldots \neq \mu_{.b}$

ขั้นที่ 2 กำหนดนัยสำคัญ α

ขั้นที่ 3 ทำการสุ่มตัวอย่าง

ขั้นที่ 4 เลือกค่าสถิติ F

ขั้นที่ 5 คำนวณค่าสถิติ F จากตัวอย่างโดยการสร้างตาราง ANOVA

ตาราง ANOVA

แหล่งการแปรผัน	ผลบวกกำลังสอง	องศาเสรี	ค่าเฉลี่ยของผลบวกกำลังสอง	$f_{คำนวณ}$
วิธีการปฏิบัติ (Treatment)	SSA	$k-1$	$MSA = \frac{SSA}{k-1}$	$f_{treatment} = \frac{MSA}{MSE}$
กลุ่ม (Block)	SSB	$b-1$	$MSB = \frac{SSB}{b-1}$	$f_{block} = \frac{MSB}{MSE}$
ความคลาดเคลื่อน (Error)	SSE	$(k-1)(b-1)$	$MSE = \frac{SSE}{(b-1)(k-1)}$	
ทั้งหมด (Total)	SST	$kb-1$		

ขั้นที่ 6 เปิดตารางหาค่าวิกฤต

6.1 ค่าวิกฤตของการสรุปผลเกี่ยวกับ Treatment คือ f_α

โดยมีค่าองศาความเสรี $v_1 = k-1$, $v_2 = (k-1)(b-1)$

บริเวณวิกฤตคือ $F > f_\alpha$

6.2 ค่าวิกฤตของการสรุปผลเกี่ยวกับ Block คือ f_{α}

โดยมีค่าองศาความเสรี $v_1 = b - 1$, $v_2 = (k - 1)(b - 1)$

บริเวณวิกฤตคือ $F > f_{\alpha}$

ขั้นที่ 7 สรุปผล 7.1 การสรุปผลเกี่ยวกับ Treatment

7.1.1 ถ้า $f_{treatment} > f_{\alpha}$ ของ Treatment แล้ว ปฏิเสธ H_0

7.1.2 ถ้า Sig < α แล้ว ปฏิเสธ H_0

7.2 การสรุปผลเกี่ยวกับ Block

7.2.1 ถ้า $f_{block} > f_{\alpha}$ ของ Block แล้ว ปฏิเสธ H_0

7.2.2 ถ้า Sig < α แล้ว ปฏิเสธ H_0

ตัวอย่าง 9.2.1 ในการเปรียบเทียบประสิทธิภาพของเครื่องจักร 4 ชนิด และความสามารถของคนที่คุมเครื่องจักร 5 คน ข้อมูลของการทำงานหน่วยเป็น วินาที จากการสุ่มตัวอย่างคือ

	เครื่องจักร			
	1	2	3	4
เจ้าหน้าที่คนที่ 1	44	38	47	36
2	46	40	52	43
3	34	36	44	32
4	43	38	46	33
5	38	42	49	39

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ว่า เครื่องจักร 4 เครื่องมีอัตราเร็วเท่ากัน

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ว่า เจ้าหน้าที่ 5 คน ปฏิบัติการด้วยอัตราเร็วเท่ากัน

วิธีทำ

ขั้นที่ 1 สมมติฐานเกี่ยวกับวิธีปฏิบัติการ(เครื่องจักร)

$H_0 : \mu_{1.} = \mu_{2.} = \mu_{3.} = \mu_{4.}$

$H_1 : \mu_{1.} \neq \mu_{2.} \neq \mu_{3.} \neq \mu_{4.}$

สมมติฐานเกี่ยวกับความแตกต่างระหว่างกลุ่ม(เจ้าหน้าที่)

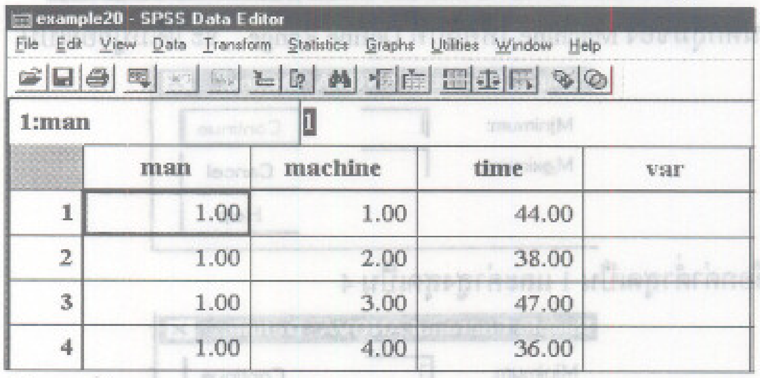
$$H_0 : \mu_{.1} = \mu_{.2} = \mu_{.3} = \mu_{.4} = \mu_{.5}$$

$$H_1 : \mu_{.1} \neq \mu_{.2} \neq \mu_{.3} \neq \mu_{.4} \neq \mu_{.5}$$

ขั้นที่ 2 กำหนดนัยสำคัญ $\alpha = 0.05$

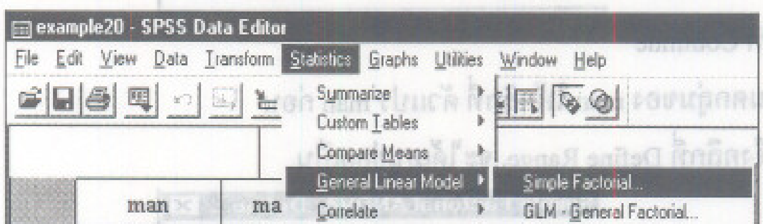
ขั้นที่ 3 นำข้อมูลเข้าสู่ SPSS

ขั้นที่ 3.1 การสร้างแฟ้มข้อมูลต้องกำหนดตัวแปร man ตัวแปรจำแนกคน machine ตัวแปรจำแนกเครื่องจักร time ตัวแปรเก็บข้อมูลที่ต้องการวิเคราะห์

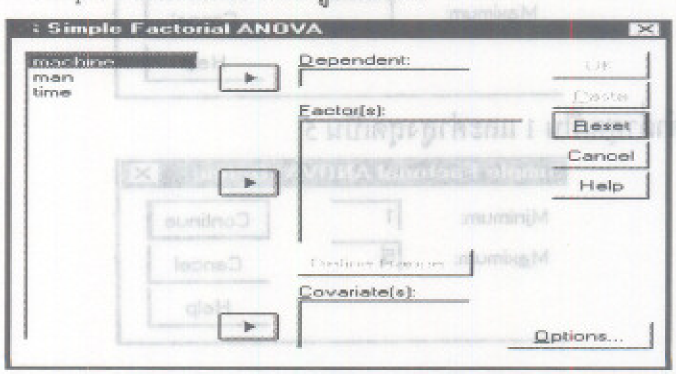


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example20.sav

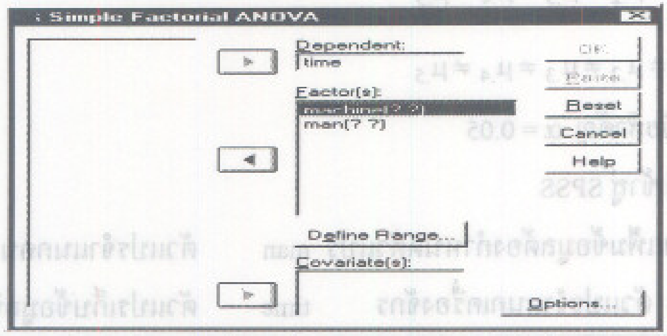
ขั้นที่ 3.2 เลือกใช้คำสั่ง Statistics \ General Linear Model \ Simple Factorial..



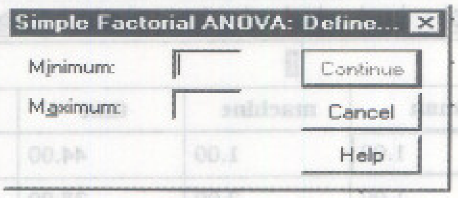
ขั้นที่ 3.3 เลือกคำสั่ง Simple Factorial จะได้เมนูย่อยดังนี้



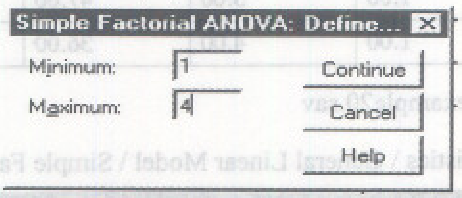
ขั้นที่ 3.4 เลือกตัวแปร time ไปไว้ที่ช่อง Dependent เลือกตัวแปร machine ไปไว้ที่ช่อง Factor(s) เลือกตัวแปร man ไปไว้ที่ช่อง Factor(s)



ขั้นที่ 3.5 กำหนดกลุ่มของ Machine ให้คลิกที่ Define Range.. จะได้เมนูย่อยเป็น

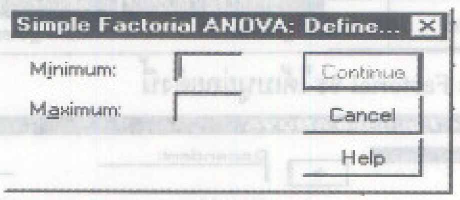


ขั้นที่ 3.6 ให้เลือกค่าต่ำสุดเป็น 1 และค่าสูงสุดเป็น 4

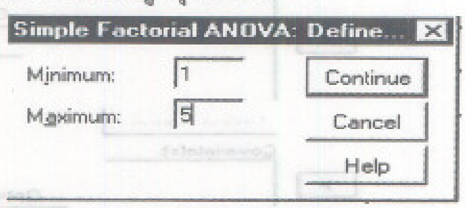


ขั้นที่ 3.7 คลิกที่ Continue

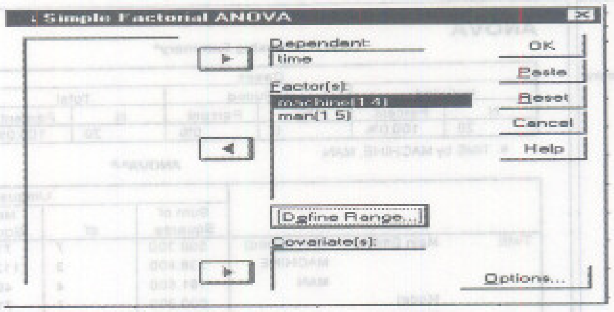
ขั้นที่ 3.8 กำหนดกลุ่มของ man ให้คลิกที่ ตัวแปร man ก่อน แล้วจึงคลิกที่ Define Range..จะได้เมนูย่อยเป็น



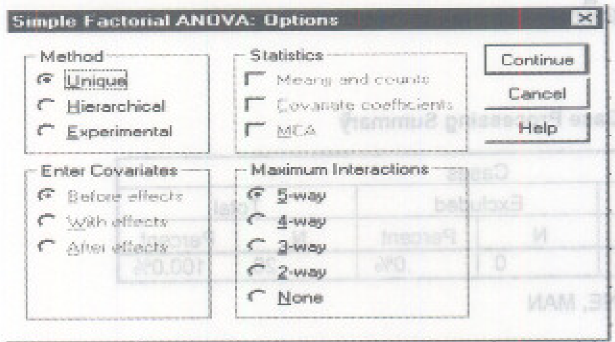
ขั้นที่ 3.9 ให้เลือกค่าต่ำสุดเป็น 1 และค่าสูงสุดเป็น 5



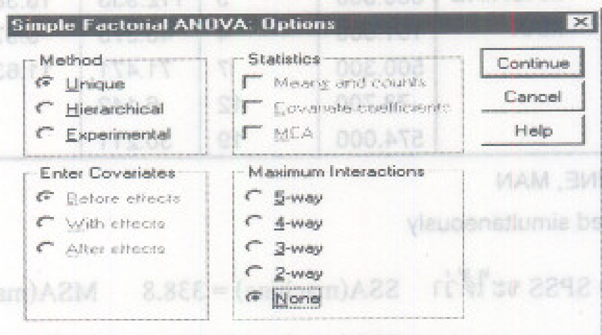
ขั้นที่ 3.10 คลิกที่ Continue จะกลับมาที่เมนูย่อย



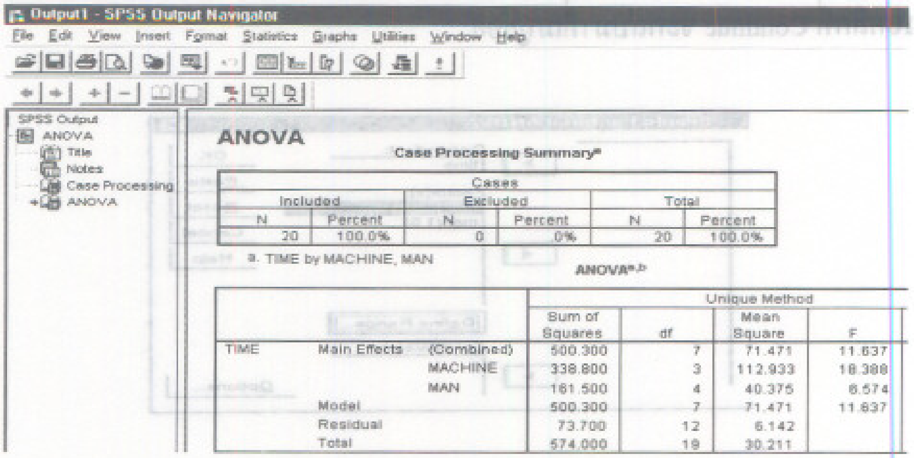
ขั้นที่ 3.11 คลิก Options..จะได้



คลิกที่ None ให้เป็น None บนจอภาพจะเป็น



ขั้นที่ 3.12 คลิก Continue และ OK ตามลำดับจะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

ANOVA

Case Processing Summary[a]

Cases					
Included		Excluded		Total	
N	Percent	N	Percent	N	Percent
20	100.0%	0	.0%	20	100.0%

a. TIME by MACHINE, MAN

ANOVA[a,b]

			Unique Method				
			Sum of Squares	df	Mean Square	F	Sig.
TIME	Main Effects	(Combined)	500.300	7	71.471	11.637	.0001669961
		MACHINE	338.800	3	112.933	18.388	.0000877779
		MAN	161.500	4	40.375	6.574	.0048466575
	Model		500.300	7	71.471	11.637	.0001669961
	Residual		73.700	12	6.142		
	Total		574.000	19	30.211		

a. TIME by MACHINE, MAN

b. All effects entered simultaneously

จากผลการคำนวณของ SPSS จะได้ว่า SSA(machine) = 338.8 MSA(machine) = 112.933

SSB(man) = 161.5 MSB(man) = 40.375

SSE = 73.7 MSE = 6.14

SST = 574

ขั้นที่ 4 เลือกค่าสถิติ F

ขั้นที่ 5 คำนวณค่าสถิติ F จากตัวอย่าง

$f_{\text{คำนวณ}}$ ของการสรุปผลเกี่ยวกับ machine $= \dfrac{MSA(machine)}{MSE} = \dfrac{112.933}{6.14} = 18.39$

$f_{\text{คำนวณ}}$ ของการสรุปผลเกี่ยวกับ man $= \dfrac{MSB(man)}{MSE} = \dfrac{40.38}{6.14} = 6.58$

ตาราง ANOVA คือ

แหล่งการแปรผัน	ผลบวกกำลังสอง	องศาเสรี	ค่าเฉลี่ยของผลบวกกำลังสอง	$f_{\text{คำนวณ}}$	Sig
วิธีการปฏิบัติ(Treatment) (machine)	338.800	3	122.933	$f_{treatment} = 18.39$	0.0000877779
กลุ่ม(Block) (man)	161.500	4	40.38	$f_{block} = 6.58$	0.004866575
ความคลาดเคลื่อน (Error)	73.700	12	6.14		
ทั้งหมด (Total)	574.000	19			

ขั้นที่ 6 เปิดตารางหาค่าวิกฤต

ค่าวิกฤตของการสรุปเกี่ยวกับ machine $f_{0.05\,(v1 = 3\,,\,v2 = 12)} = 3.49$

ค่าวิกฤตของการสรุปเกี่ยวกับ man $f_{0.05\,(v1 = 4\,,\,v2 = 12)} = 3.26$

ขั้นที่ 7 สรุปผล

การสรุปผลเกี่ยวกับ machine

1. เพราะว่า F คำนวณของ machine = 18.39 > 3.49 เพราะฉะนั้น ปฏิเสธ H_0
2. เพราะว่า Sig = 0.0000877779 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

การสรุปผลเกี่ยวกับ man

1. เพราะว่า F คำนวณของ man = 6.58 > 3.26 เพราะฉะนั้น ปฏิเสธ H_0
2. เพราะว่า Sig = 0.0048466575 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

แบบฝึกหัด 9.

1. เพื่อเปรียบเทียบกำลังต้านทานการดึงของยางที่ผลิตโดยเครื่องจักร 6 เครื่อง สุ่มเลือกตัวอย่างขนาด 4 จากแต่ละเครื่องจักร ได้กำลังต้านทานการดึงเฉลี่ยเป็นปอนด์ต่อตารางนิ้ว X 10^{-2} ดังนี้

เครื่องจักร

1	2	3	4	5	6
17.5	16.4	20.3	14.6	17.5	18.3
16.9	19.2	15.7	16.7	19.2	16.2
15.8	17.7	17.8	20.8	16.5	17.5
18.6	15.4	18.9	18.9	20.5	20.1

จงวิเคราะห์ความแปรปรวนที่ระดับนัยสำคัญ 0.05 และตรวจสอบว่าค่าเฉลี่ยของการปฏิบัติแตกต่างอย่างมีนัยสำคัญหรือไม่

2. ครูสามคนได้ทำการสอนคณิตศาสตร์เบื้องต้นแก่นักเรียน 3 กลุ่ม ผลการสอบภาคปลายการศึกษาบันทึกผลคะแนนของตัวอย่างคะแนนไว้ดังนี้

ครูคนที่หนึ่ง : 73 , 92 , 82 , 43 , 80 , 73 , 66 , 60 , 45 , 93 , 36 , 77

ครูคนที่สอง : 88 , 78 , 48 , 91 , 51 , 85 , 74 , 77 , 31 , 78 , 62 , 76 , 96 , 80 , 56

ครูคนที่สาม : 68 , 79 , 56 , 91 , 71 , 71 , 87 , 41 , 59 , 68 , 53 , 79 , 15

จงทดสอบว่าค่าเฉลี่ยของคะแนนสอบที่สอนโดยครูทั้งสามคนมีความแตกต่างกันหรือไม่ที่ระดับนัยสำคัญ 0.05

3. เพื่อตรวจสอบประสิทธิภาพในการทำงานของ นาย ก. ข. และ ค. จึงให้คนทั้งสาม ควบคุมเครื่องจักรที่ใช้ผลิตสินค้า 4 เครื่อง เครื่องละหนึ่งสัปดาห์หมุนเวียนกันไปปรากฏว่ามีสินค้าที่ไม่ได้มาตรฐานดังนี้

คนควบคุมเครื่องจักร

		นาย ก	นาย ข	นาย ค
เครื่องจักร	1.	10	14	18
	2.	13	16	22
	3.	13	19	14
	4.	16	27	18

จะสรุปได้หรือไม่ว่าประสิทธิภาพของคนทั้งสามไม่แตกต่างกัน และ ประสิทธิภาพของเครื่องจักรทั้ง 4 เครื่อง ไม่แตกต่างกัน ที่ระดับนัยสำคัญ 0.05

บทที่ 10

การทดสอบสมมติฐานแบบนอนพาราเมตริก

ในกรณีที่เราไม่ทราบการแจกแจงของประชากรและเราต้องการทดสอบสมมติฐานเกี่ยวกับลักษณะบางอย่างของประชากร เราจะทำการทดสอบสมมติฐาน**แบบนอนพาราเมตริก** (Non parametric Test) การทดสอบที่เราจะเรียนกันในบทนี้คือ

การทดสอบว่าตัวอย่างที่เราเลือกมาเป็นไปโดยสุ่มหรือไม่

การทดสอบว่าประชากรมีการแจกแจงตามที่เราคาดไว้หรือไม่

การทดสอบว่าประชากร 2 กลุ่มมีความสัมพันธ์กันหรือไม่

การทดสอบว่าค่าเฉลี่ยของประชากร k กลุ่มตัวอย่างเท่ากันหรือไม่

10.1 การทดสอบว่าตัวอย่างที่เราเลือกมาเป็นไปโดยสุ่มหรือไม่

การทดสอบว่าข้อมูลตัวอย่างที่เราเก็บรวบรวมมาได้เป็นการสุ่มจริงหรือไม่ สามารถทำการทดสอบได้โดยใช้วิธี **ทดสอบรันส์** (Runs Test)

การทดสอบสมมติฐานโดยใช้ Runs Test ของ SPSS for Windows

ตัวอย่าง 10.1.1 ข้อมูลจำนวนคนที่อยู่ในแถวเพื่อรอถอนเงินจากเครื่อง ATM ที่เก็บมาในช่วงเวลา 40 วันต่อเนื่องกันเป็นดังนี้

6	7	5	6	8	6	8	6	6	4
3	2	4	4	3	4	7	5	6	8
6	6	3	5	2	5	4	4	3	7
5	5	4	3	7	4	6	5	2	8

จงทดสอบว่าจำนวนคนที่อยู่ในแถวเป็นไปอย่างสุ่ม กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

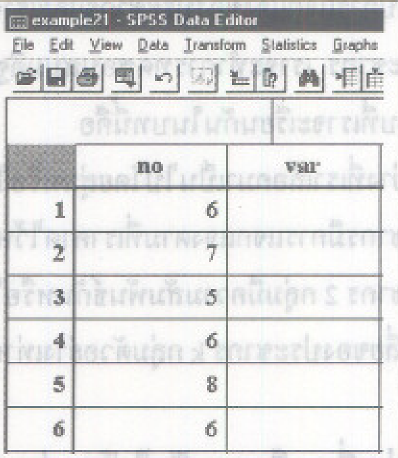
ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : จำนวนคนที่อยู่ในแถวเป็นไปอย่างสุ่ม

กำหนดสมมติฐานอื่น H_1 : จำนวนคนที่อยู่ในแถวไม่เป็นไปอย่างสุ่ม

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

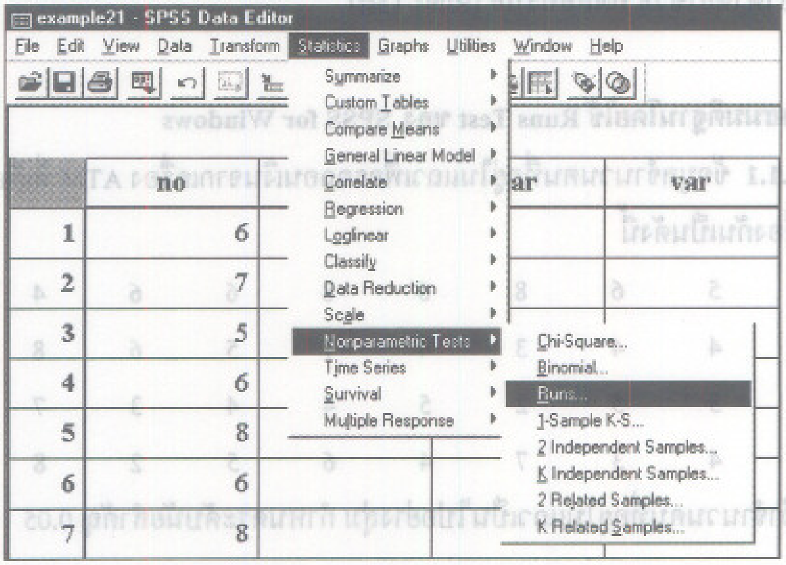
ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

ขั้นที่ 3.1 สร้างแฟ้มข้อมูล

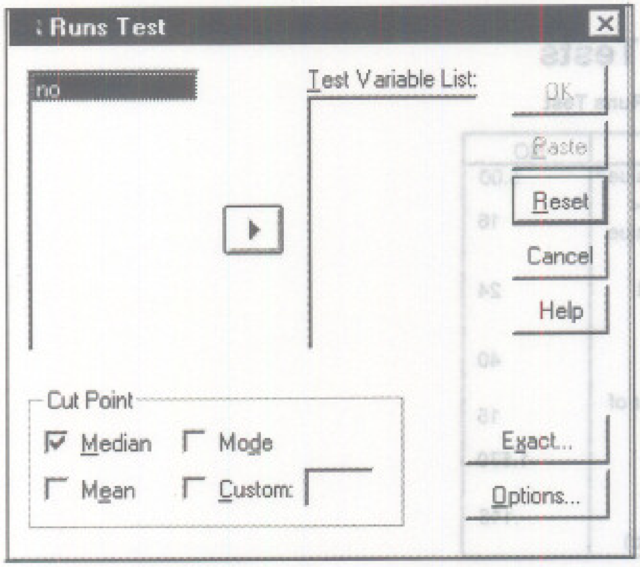


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example21.sav

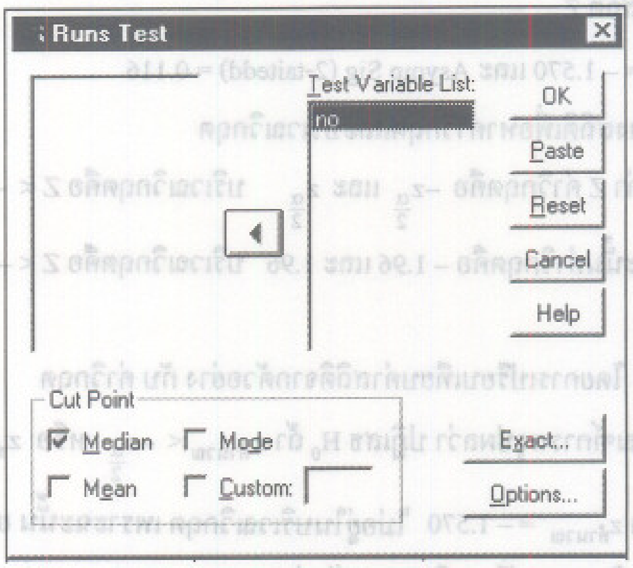
ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric Tests / Runs ..



ขั้นที่ 3.3 คลิกที่คำสั่ง Runs ..จะได้เมนูย่อย



ขั้นที่ 3.4 เลือกตัวแปร no ไปไว้ที่ช่อง Test Variable



หมายเหตุ ขณะนี้เป็นการทำ Runs Test โดยทำการเปรียบเทียบกับค่า Median เราสามารถทำการทดสอบ โดยทำการเปรียบเทียบกับค่า Mean Mode หรือค่าอื่นๆ ที่กำหนดเองได้

ขั้นที่ 3.5 กด OK จะได้ผลการคำนวณเป็น

NPar Tests

Runs Test

	NO
Test Value[a]	5.00
Cases < Test Value	16
Cases >= Test Value	24
Total Cases	40
Number of Runs	15
Z	-1.570
Asymp. Sig. (2-tailed)	.116

a. Median

ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = -1.570$ และ Asymp Sig (2-taitedd) = 0.116

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ – 1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

เพราะว่า $z_{\text{คำนวณ}} = -1.570$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ยอมรับ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig < α

เพราะว่า Sig = 0.116 > 0.05 เพราะฉะนั้น **ยอมรับ H_0**

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

การทดสอบว่าประชากรมีค่าเฉลี่ยตามที่เราคาดไว้หรือไม่

ตัวอย่าง 10.1.2 ข้อมูลของจำนวนซัลเฟอร์ออกไซด์ ที่ออกมาจากโรงงานอุตสาหกรรมในแต่ละวัน ที่เก็บมาได้ในช่วง 60 วัน เป็นดังนี้

17.00 , 15.00 , 20.00 , 29.00 , 19.00 , 18.00 , 22.00 , 25.00 , 27.00 , 9.00 , 24.00 , 20.00
17.00 , 6.00 , 24.00 , 14.00 , 15.00 , 23.00 , 24.00 , 26.00 , 19.00 , 23.00 , 28.00 , 19.00
16.00 , 22.00 , 24.00 , 17.00 , 20.00 , 13.00 , 19.00 , 10.00 , 23.00 , 18.00 , 31.00 , 13.00
20.00 , 17.00 , 24.00 , 14.00 , 28.00 , 19.00 , 16.00 , 22.00 , 24.00 , 17.00 , 20.00 , 13.00
19.00 , 10.00 , 23.00 , 18.00 , 17.00 , 15.00 , 20.00 , 29.00 , 19.00 , 18.00 , 22.00 , 25.00

จงทดสอบสมมติฐานว่า ค่าเฉลี่ยของซัลเฟอร์ออกไซด์เท่ากับ 20 ที่ระดับความมีนัยสำคัญ 0.05

วิธีทำ

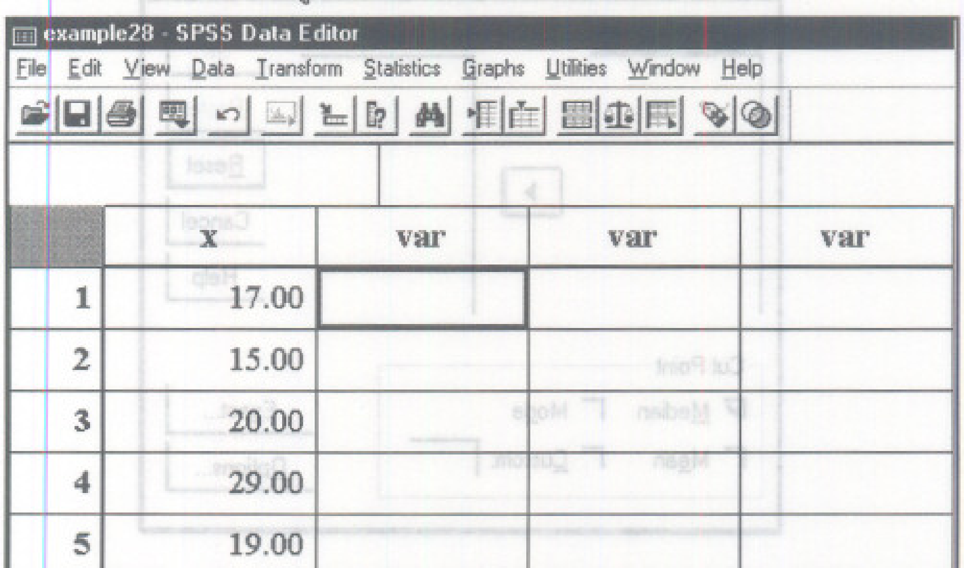
ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu = 20$

กำหนดสมมติฐานอื่น $H_1 : \mu \neq 20$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

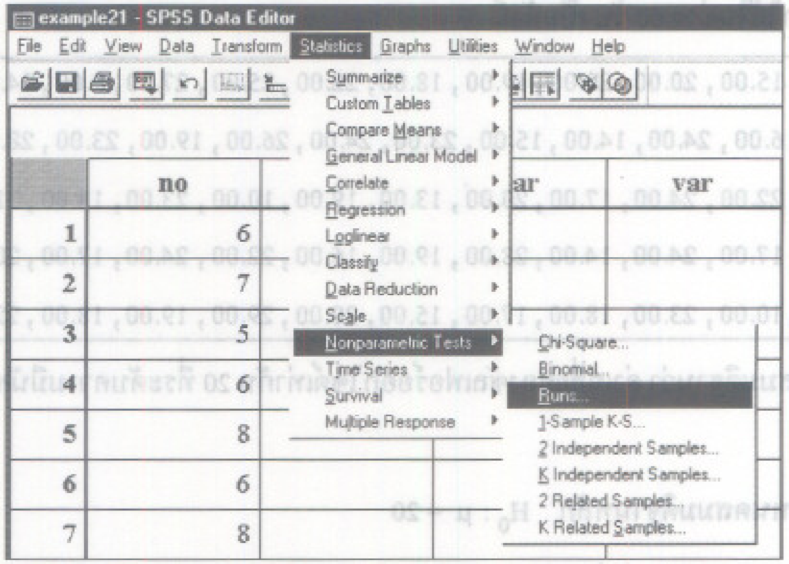
ขั้นที่ 3.1 สร้างแฟ้มข้อมูล



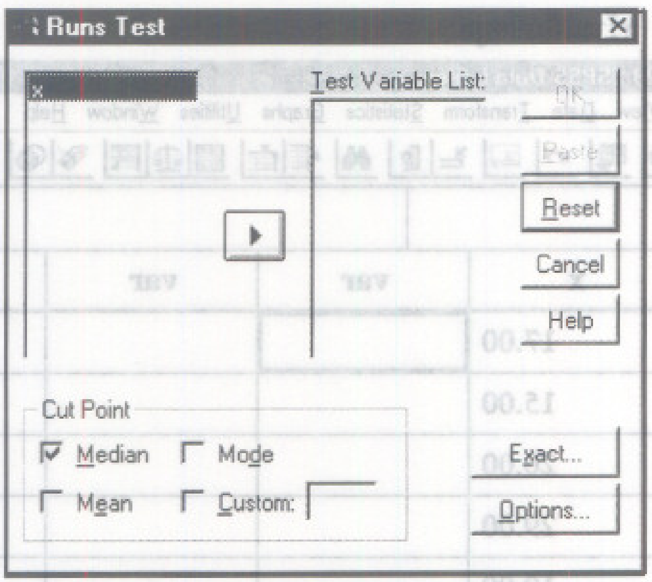
	x	var	var	var
1	17.00			
2	15.00			
3	20.00			
4	29.00			
5	19.00			

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example28.sav

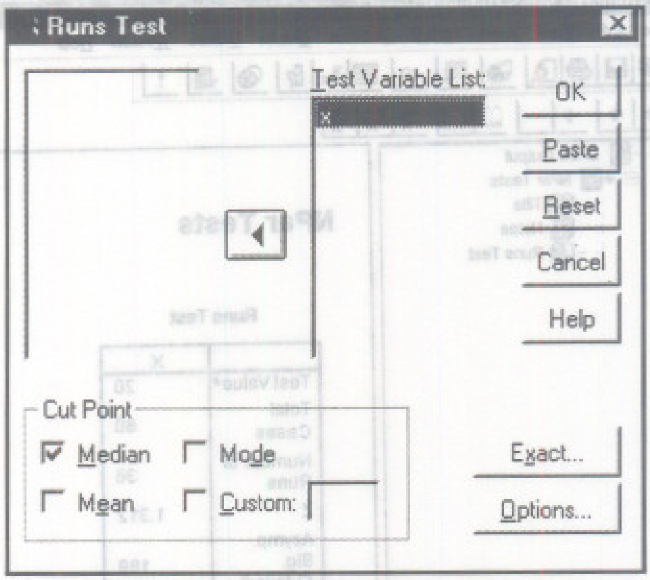
ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric Tests / Runs ..



ขั้นที่ 3.3 คลิกที่คำสั่ง Runs ..จะได้เมนูย่อย

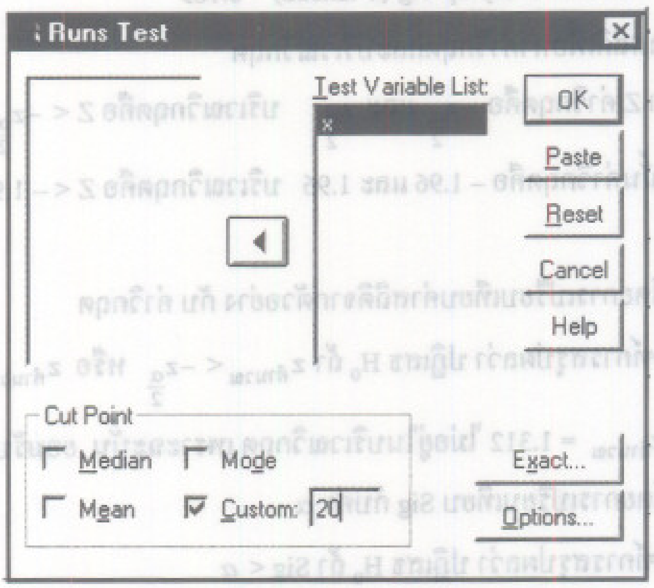


ขั้นที่ 3.4 เลือกตัวแปร x ไปไว้ที่ช่อง Test Variable

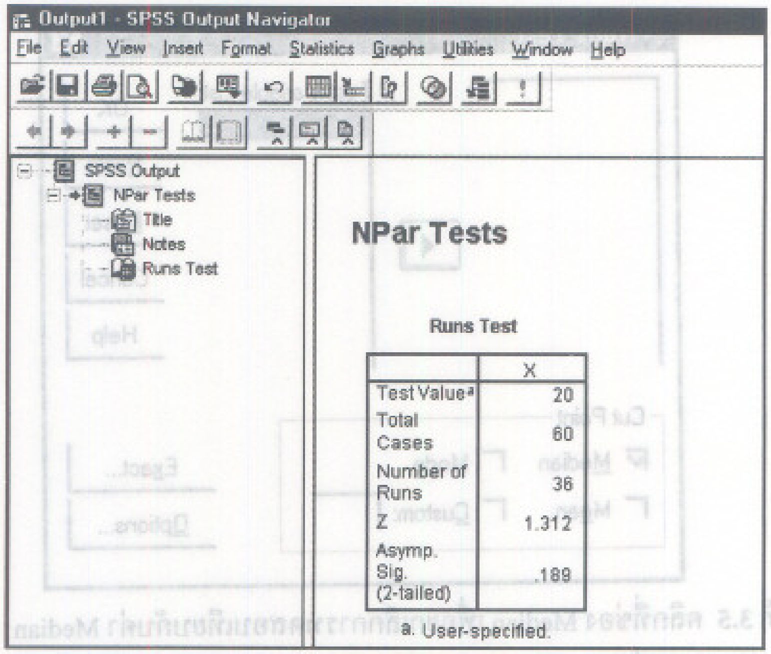


ขั้นที่ 3.5 คลิกที่ช่อง Median เพื่อยกเลิกการทดสอบเทียบกับค่า Median

คลิกที่ช่อง Custom และพิมพ์ค่า 20 ในช่อง Custom



ขั้นที่ 3.6 คลิก OK จะได้ผลการคำนวณเป็น



ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = 1.312$ และ Asymp Sig (2-taitedd) = 0.189

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ – 1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

เพราะว่า $z_{\text{คำนวณ}} = 1.312$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ยอมรับ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig < α

เพราะว่า Sig = 0.189 > 0.05 เพราะฉะนั้น ยอมรับ H_0

10.2 การทดสอบว่าประชากรมีการแจกแจงตามที่เราคาดไว้หรือไม่

การทดสอบว่าประชากรที่เราสนใจมีการแจกแจงปกติจริงหรือไม่ ประชากรที่เราสนใจมีการแจกแจง uniform จริงหรือไม่ ประชากรที่เราสนใจมีการแจกแจงปัวส์ซองจริงหรือไม่ เราสามารถทำการทดสอบแบบ Non parametric Test ได้

ตัวอย่าง 10.2.1 การทดสอบว่าน้ำหนักของนักเรียนมีการแจกแจงปกติจริงหรือไม่ จึงทำการสุ่มตัวอย่างน้ำหนักนักเรียนมา 50 คน ได้ข้อมูลดังนี้

50.00	69.00	108.00	85.00	132.00	67.00	121.00	80.00	59.00	64.00
148.00	61.00	50.00	103.00	110.00	66.00	95.00	55.00	128.00	101.00
137.00	145.00	103.00	96.00	136.00	127.00	149.00	111.00	76.00	134.00
87.00	117.00	50.00	77.00	108.00	133.00	98.00	124.00	95.00	124.00
109.00	123.00	107.00	65.00	92.00	101.00	125.00	66.00	90.00	110.00

กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

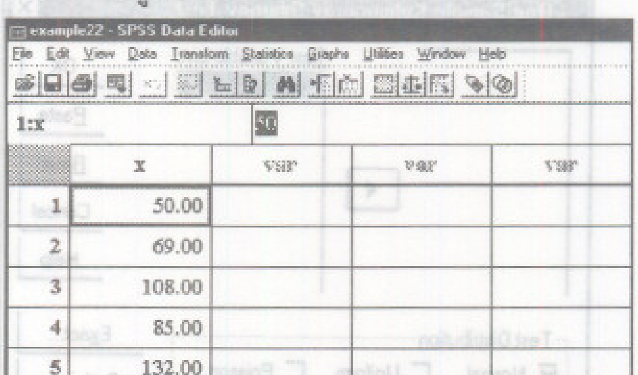
ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ข้อมูลน้ำหนักมีการแจกแจงปกติ

กำหนดสมมติฐานอื่น H_1 : ข้อมูลน้ำหนักไม่มีการแจกแจงปกติ

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

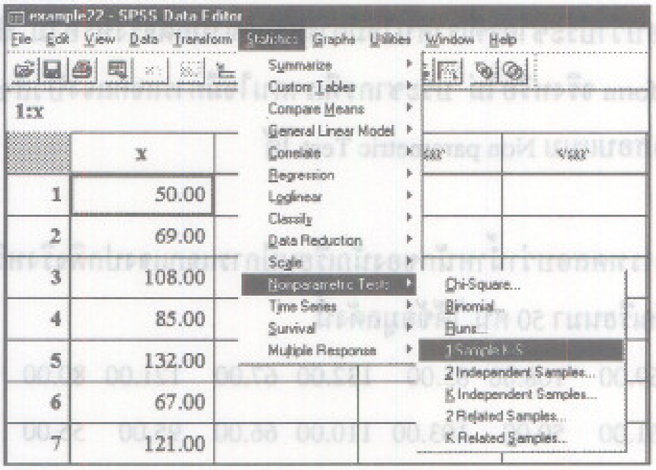
ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

ขั้นที่ 3.1 สร้างแฟ้มข้อมูล

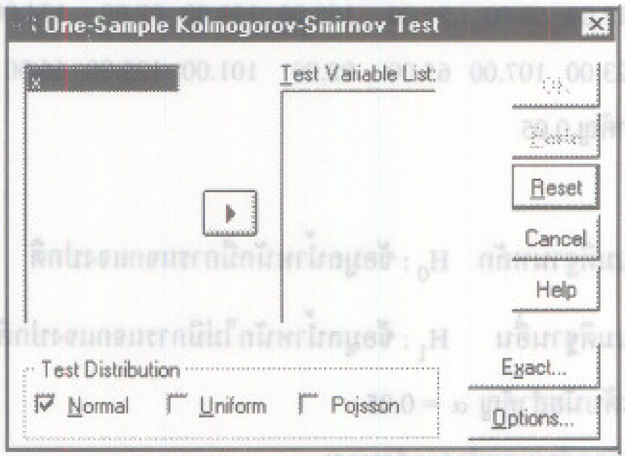


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example22.sav

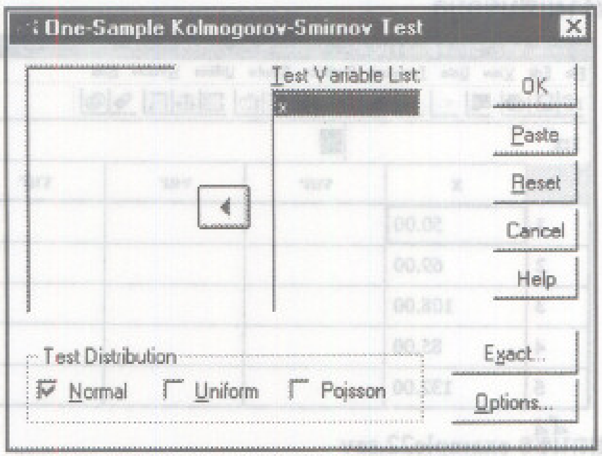
ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric Tests / Sample K-S.



ขั้นที่ 3.3 คลิกที่คำสั่ง Sample K-S .จะได้เมนูย่อย



ขั้นที่ 3.4 เลือกตัวแปร x ไปไว้ที่ช่อง Test Variable List



ขั้นที่ 3.5 กด OK จะได้ผลการคำนวณเป็น

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		X
N		50
Normal Parameters[a,b]	Mean	99.5600
	Std. Deviation	28.1259
Most Extreme Differences	Absolute	.101
	Positive	.101
	Negative	-.078
Kolmogorov-Smirnov Z		.717
Asymp. Sig. (2-tailed)		.683

a. Test distribution is Normal.

b. Calculated from data.

ขั้นที่ 4. เลือกค่าสถิติ Z (Kolmogorov-Smirnov Z)

ขั้นที่ 5. $z_{\text{คำนวณ}} = 0.717$ และ Asymp Sig (2-taitedd) = 0.683

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ – 1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

เพราะว่า $z_{\text{คำนวณ}} = 0.717$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ยอมรับ H_0

หรือ แบบที่ 2 โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig < α

เพราะว่า Sig = 0.683 > 0.005 เพราะฉะนั้น ยอมรับ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

10.3 การทดสอบว่าประชากร 2 กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่

10.3.1 ประชากร 2 ชุดไม่เป็นอิสระต่อกัน

ในกรณีที่ประชากร 2 ชุดไม่อิสระต่อกัน และ ไม่ทราบการแจกแจงของประชากร เราสามารถทำการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากร 2 ชุดเท่ากันหรือไม่ โดยใช้วิธีทดสอบ **Wilcoxon Signed Ranks Test**

ตัวอย่าง 10.3.1 การทดสอบว่าโปรแกรมการควบคุมน้ำหนักโดยใช้เวลา 30 วันจะมีผลทำให้น้ำหนักลดลง ได้ทำการเก็บข้อมูลน้ำหนักของชาย 40 คนได้ข้อมูลดังนี้

	x	y
1	147.00	137.90
2	183.50	176.20
3	232.10	219.00
4	161.60	163.80
5	197.50	193.50
6	206.30	201.40
7	177.00	180.60
8	215.40	203.20
9	147.70	149.00
10	208.10	195.40

	x	y
11	137.90	140.00
12	176.20	170.00
13	219.00	210.00
14	163.80	160.00
15	137.90	140.00
16	176.20	170.00
17	219.00	210.00
18	163.80	165.00
19	193.50	195.00
20	201.40	205.00

	x	y
21	180.60	185.00
22	203.20	195.00
23	137.90	140.00
24	176.20	170.00
25	219.00	200.00
26	163.80	155.00
27	193.50	190.00
28	201.40	200.00
29	180.60	170.00
30	137.90	140.00

	x	y
31	176.20	177.00
32	219.00	211.00
33	163.80	174.00
34	193.50	195.00
35	201.40	200.00
36	180.60	180.00
37	203.20	203.00
38	149.00	150.00
39	195.40	185.00
40	145.00	150.00

x เป็นน้ำหนักก่อนเข้าโปรแกรม y เป็นน้ำหนักหลังเข้าโปรแกรม

จงทดสอบสมมติฐานว่าโปรแกรมการควบคุมน้ำหนักไม่ทำให้น้ำหนักเปลี่ยนแปลง

กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

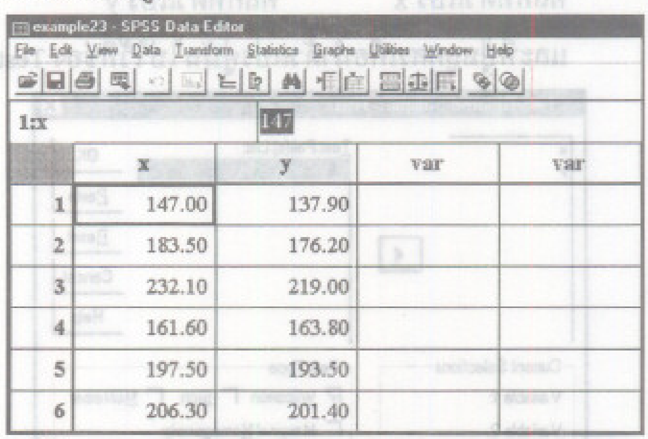
ขั้นที่ 1. กำหนดสมมติฐานหลัก H_0 : ค่าเฉลี่ยของน้ำหนักก่อนและหลังเข้าโปรแกรมเท่ากัน

กำหนดสมมติฐานอื่น H_1 : ค่าเฉลี่ยของน้ำหนักก่อนและหลังเข้าโปรแกรม**ไม่**เท่ากัน

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

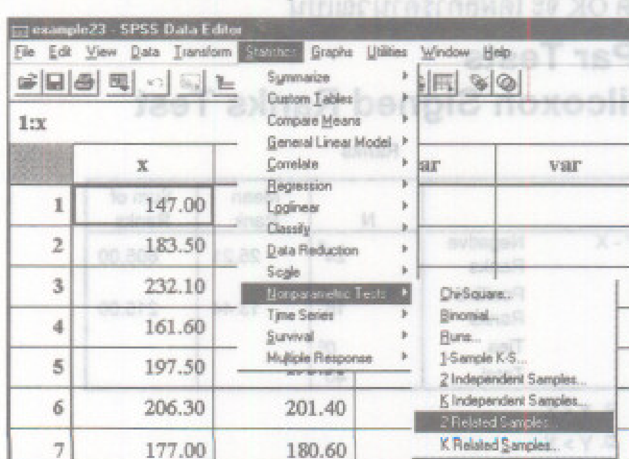
ขั้นที่ 3.1 สร้างแฟ้มข้อมูล



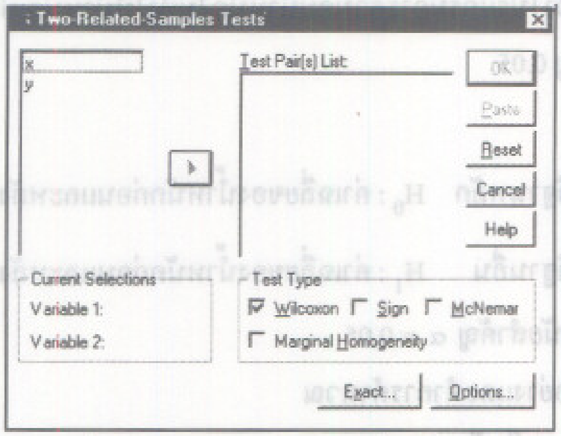
	x	y	var	var
1	147.00	137.90		
2	183.50	176.20		
3	232.10	219.00		
4	161.60	163.80		
5	197.50	193.50		
6	206.30	201.40		

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example23.sav

ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric Tests / 2 Related Samples...

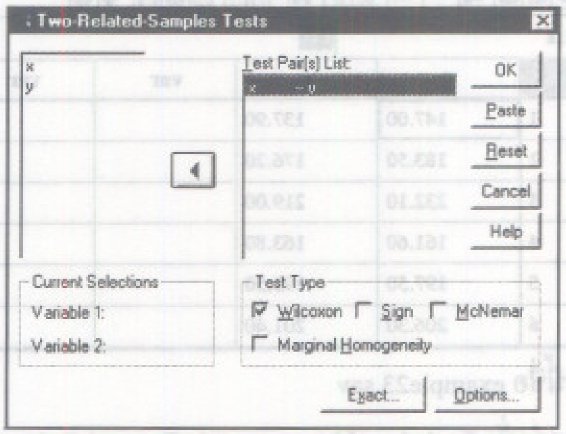


ขั้นที่ 3.3 คลิกที่คำสั่ง 2 Related Samples จะได้เมนูย่อย



ขั้นที่ 3.4 คลิกที่ตัวแปร x คลิกที่ตัวแปร y

และที่ปุ่มลูกศรเพื่อนำตัวแปรคู่นั้นไปไว้ที่ช่อง Test Pair(s) List



ขั้นที่ 3.5 กด OK จะได้ผลการคำนวณเป็น

NPar Tests
Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Y - X	Negative Ranks	24[a]	25.21	605.00
	Positive Ranks	16[b]	13.44	215.00
	Ties	0[c]		
	Total	40		

a. Y < X

b. Y > X

c. X = Y

Test Statistics[b]

	Y - X
Z	-2.622[a]
Asymp. Sig. (2-tailed)	.009

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{\text{คำนวณ}} = -2.622$ และ Asymp Sig (2-taitedd) = 0.009

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ – 1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $z_{\text{คำนวณ}} < -z_{\frac{\alpha}{2}}$ หรือ $z_{\text{คำนวณ}} > z_{\frac{\alpha}{2}}$

เพราะว่า $z_{\text{คำนวณ}} = -2.622 < -1.96$ เพราะฉะนั้น ปฏิเสธ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig $< \alpha$

เพราะว่า Sig $= 0.009 < 0.05$ เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

10.3.2 ประชากร 2 ชุดเป็นอิสระต่อกัน

ตัวอย่าง 10.3.2 ปริมาณของนิโคตินที่มีในบุหรี่ 2 ยี่ห้อคือ

ยี่ห้อ A	2.1	4.0	6.3	5.4	4.8	3.7	6.1	3.3		
ยี่ห้อ B	4.1	0.6	3.1	2.5	4.0	6.2	1.6	2.2	1.9	5.4

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าปริมาณของนิโคตินที่มีในบุหรี่ 2 ยี่ห้อเท่ากันหรือไม่

วิธีทำ

ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

ขั้นที่ 3.1 สร้างแฟ้มข้อมูล

code เป็นตัวแปรจำแนกกลุ่ม

x เป็นตัวแปรปริมาณนิโคติน

example27 - SPSS Data Editor

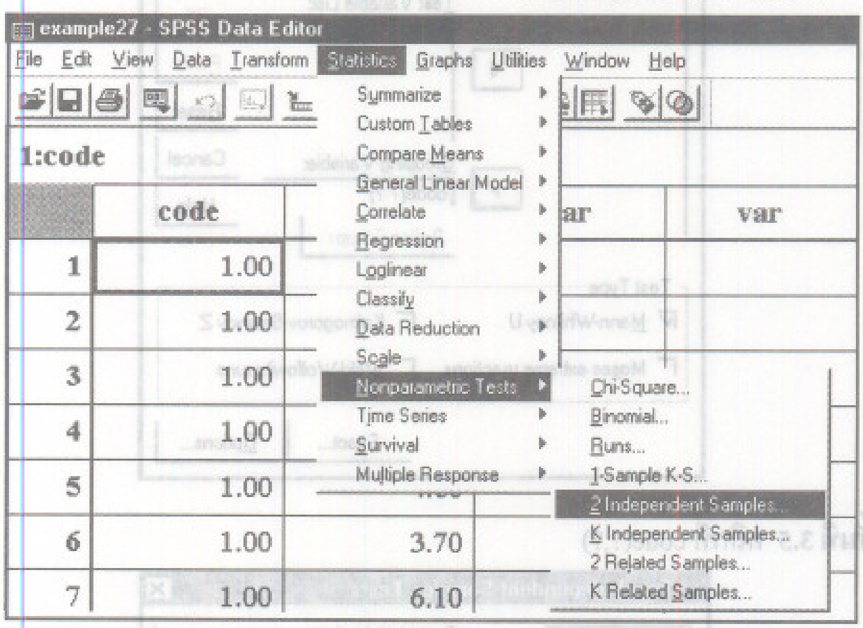
File Edit View Data Transform Statistics Graphs Utilities Window Help

1:code 1

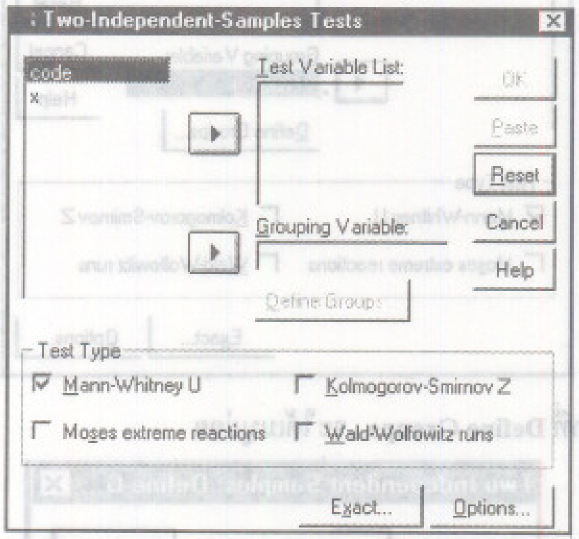
	code	x	var	var
1	1.00	2.10		
2	1.00	4.00		
3	1.00	6.30		
4	1.00	5.40		
5	1.00	4.80		
6	1.00	3.70		
7	1.00	6.10		

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example27.sav

ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric / 2 Independent Samples..

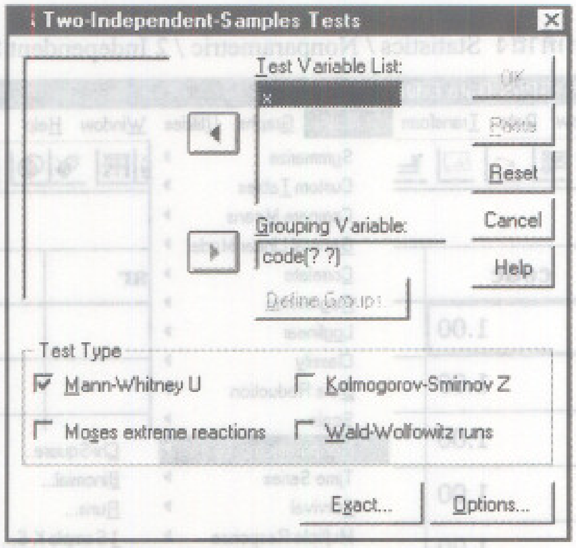


ขั้นที่ 3.3 คลิกที่ 2 Independent Samples..จะได้เมนูย่อย

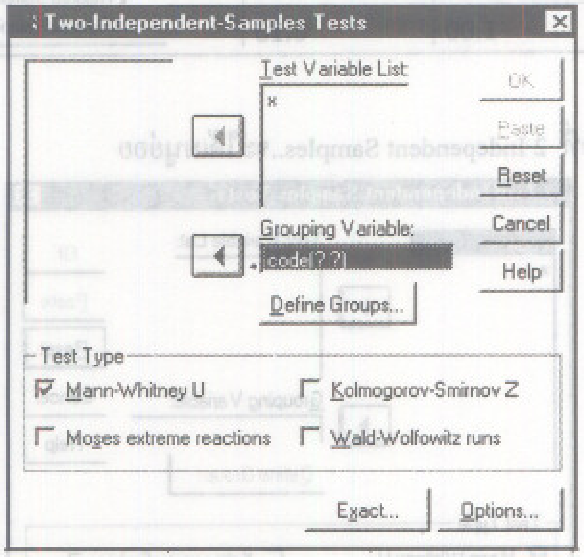


ขั้นที่ 3.4 เลือกตัวแปร x มาที่ช่อง Test Variable List

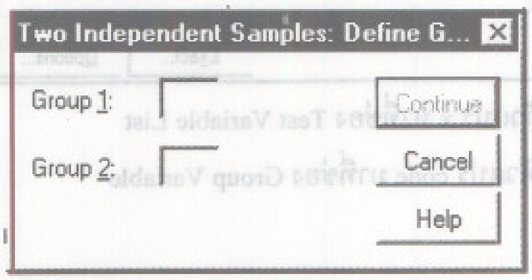
เลือกตัวแปร code มาที่ช่อง Group Variable



ขั้นที่ 3.5 คลิกที่ code(?,?)

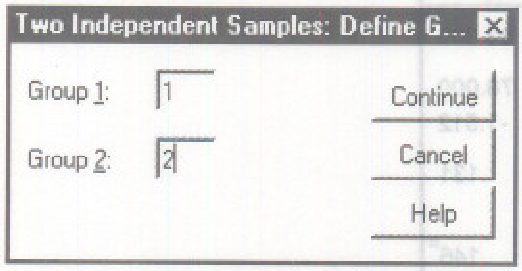


ขั้นที่ 3.6 คลิกที่ **Define Groups..** จะได้เมนูย่อย

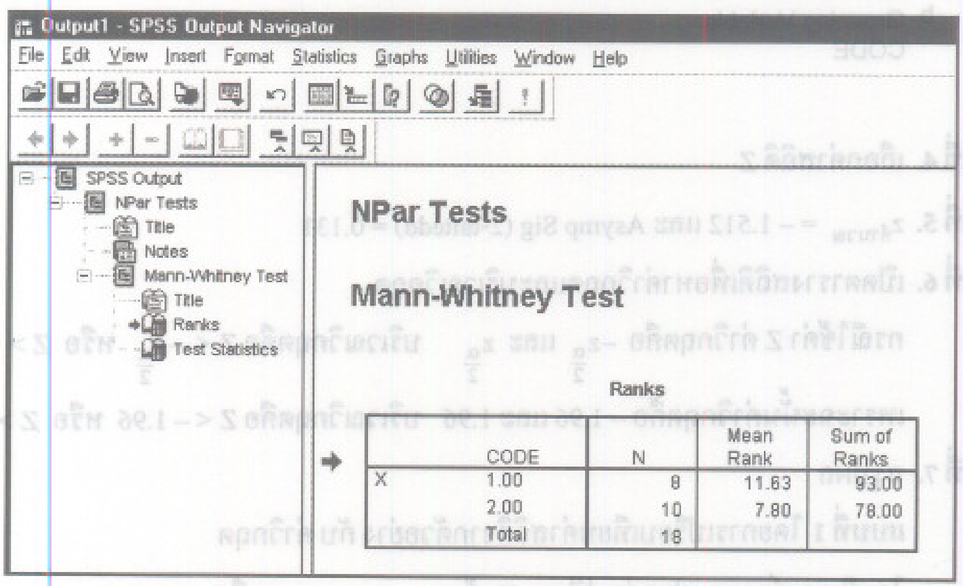


ขั้นที่ 3.7 พิมพ์ 1 ในช่อง Group 1 แล้วกด Tab

พิมพ์ 2 ในช่อง Group 2



ขั้นที่ 3.8 คลิกที่ Continue และตามด้วย OK ตามลำดับ จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

NPar Tests
Mann-Whitney Test

Ranks

	CODE	N	Mean Rank	Sum of Ranks
X	1.00	8	11.63	93.00
	2.00	10	7.80	78.00
	Total	18		

Test Statistics[b]

	X
Mann-Whitney U	23.000
Wilcoxon W	78.000
Z	-1.512
Asymp. Sig. (2-tailed)	.131
Exact Sig. [2*(1-tailed Sig.)]	.146[a]

a. Not corrected for ties.

b. Grouping Variable: CODE

ขั้นที่ 4. เลือกค่าสถิติ Z

ขั้นที่ 5. $z_{คำนวณ} = -1.512$ และ Asymp Sig (2-taitedd) = 0.131

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

กรณีใช้ค่า Z ค่าวิกฤตคือ $-z_{\frac{\alpha}{2}}$ และ $z_{\frac{\alpha}{2}}$ บริเวณวิกฤตคือ $Z < -z_{\frac{\alpha}{2}}$ หรือ $Z > z_{\frac{\alpha}{2}}$

เพราะฉะนั้นค่าวิกฤตคือ – 1.96 และ 1.96 บริเวณวิกฤตคือ $Z < -1.96$ หรือ $Z > 1.96$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $z_{คำนวณ} < -z_{\frac{\alpha}{2}}$ หรือ $z_{คำนวณ} > z_{\frac{\alpha}{2}}$

เพราะว่า $z_{คำนวณ} = -1.512$ ไม่อยู่ในบริเวณวิกฤต เพราะฉะนั้น ยอมรับ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig $< \alpha$

เพราะว่า Sig = 0.131 > 0.05 เพราะฉะนั้น ยอมรับ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

10.4 การทดสอบว่าประชากร k กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่

10.4.1 ประชากร k กลุ่มเป็นอิสระต่อกัน

ในกรณีที่ประชากร k ชุดอิสระต่อกัน และ ไม่ทราบการแจกแจงของประชากร เราสามารถทำการทดสอบสมมติฐานว่าค่าเฉลี่ยของประชากร 2 ชุดเท่ากันหรือไม่ โดยใช้วิธีทดสอบ **Kruskal - Wallis Test**

ตัวอย่าง 10.4.1 คะแนนสอบวิชาภาษาเยอรมันของนักเรียน 3 กลุ่มที่มาจากวิธีการสอนที่ต่างกัน

วิธีที่ 1	94	88	91	74	87	97	
วิธีที่ 2	85	82	79	84	63	72	80
วิธีที่ 3	89	67	72	76	69		

จงทดสอบสมมติฐานว่าวิธีการสอนทั้งสามแบบให้ผลเหมือนกัน กำหนดระดับนัยสำคัญ 0.05

วิธีทำ

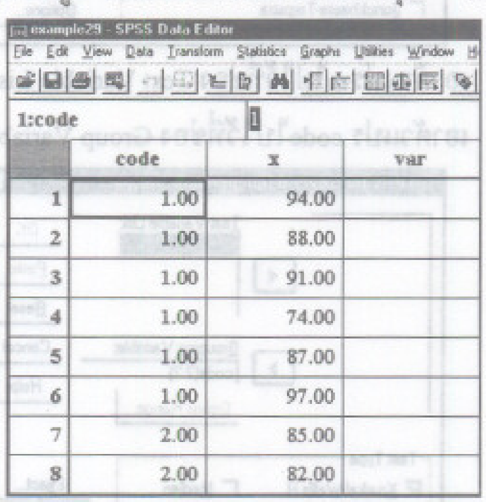
ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

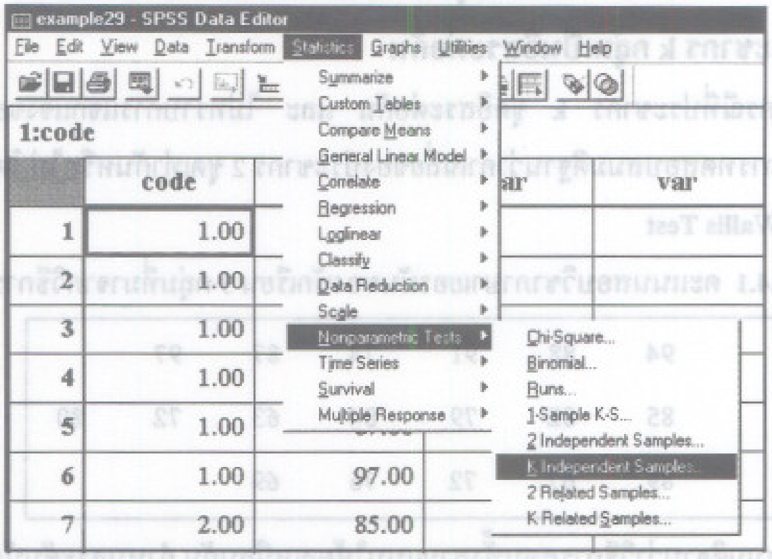
ขั้นที่ 3.1 สร้างแฟ้มข้อมูล code เป็นตัวแปรจำแนกกลุ่ม x เป็นตัวแปรเก็บคะแนน



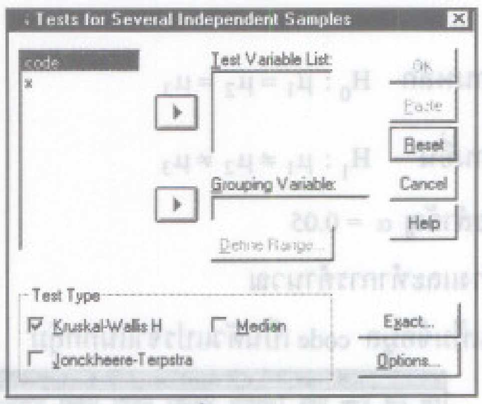
	code	x	var
1	1.00	94.00	
2	1.00	88.00	
3	1.00	91.00	
4	1.00	74.00	
5	1.00	87.00	
6	1.00	97.00	
7	2.00	85.00	
8	2.00	82.00	

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example29.sav

ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric Tests / K Independent Samples...

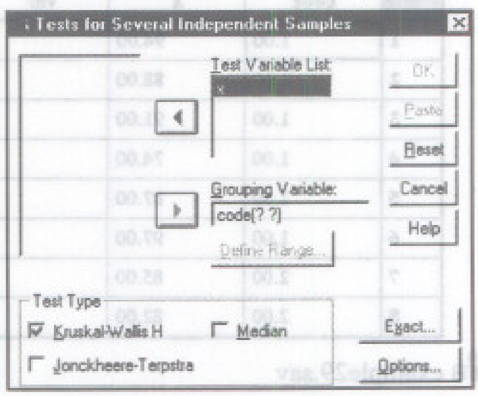


ขั้นที่ 3.3 คลิกที่คำสั่ง K Independent Samples... จะได้เมนูย่อย

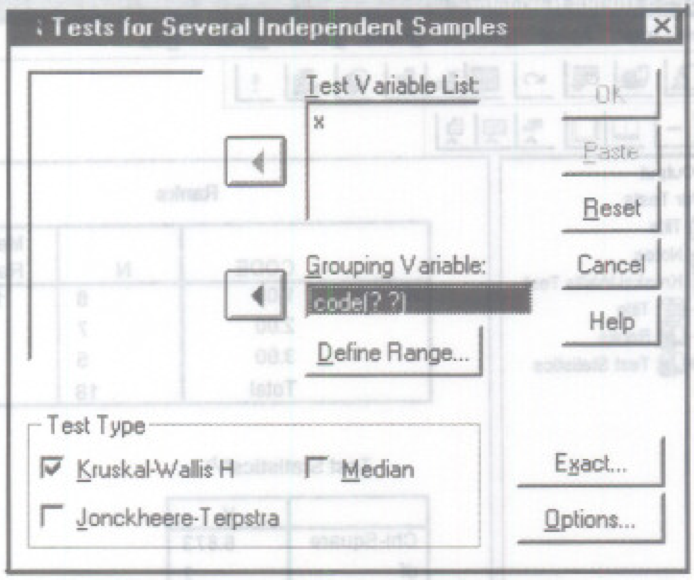


ขั้นที่ 3.4 เอาตัวแปร x ไปไว้ที่ช่อง Test Variable List

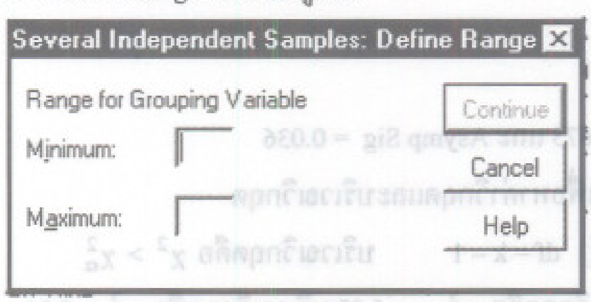
เอาตัวแปร code ไปไว้ที่ช่อง Group Variable



ขั้นที่ 3.5 คลิกที่ช่อง Group Variable

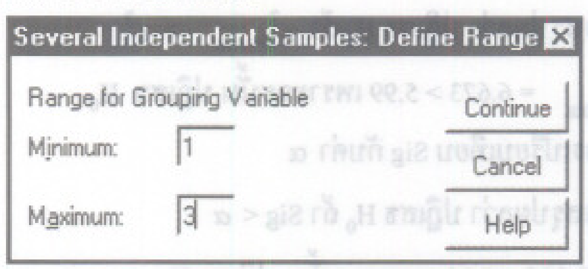


ขั้นที่ 3.6 คลิกที่ Define Range จะได้เมนูย่อย

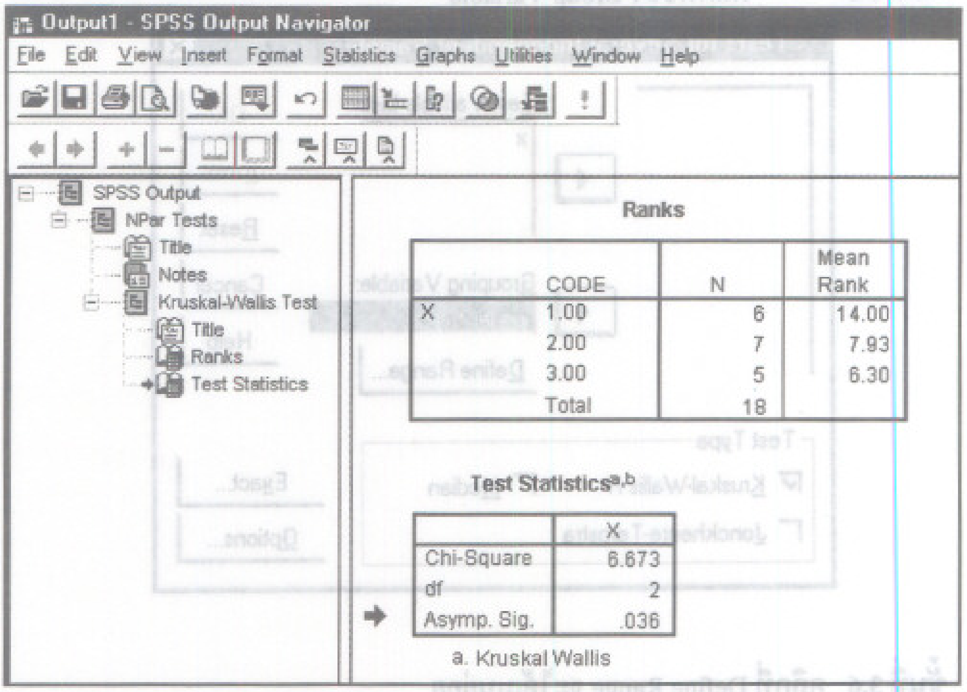


ขั้นที่ 3.7 พิมพ์ 1 ในช่อง Minimum

พิมพ์ 3 ในช่อง Maximum



ขั้นที่ 3.8 กด Continue และ OK ตามลำดับ จะได้ผลการคำนวณเป็น



Ranks

	CODE	N	Mean Rank
X	1.00	6	14.00
	2.00	7	7.93
	3.00	5	6.30
	Total	18	

Test Statistics[a,b]

	X
Chi-Square	6.673
df	2
Asymp. Sig.	.036

a. Kruskal Wallis

ขั้นที่ 4. เลือกค่าสถิติ χ^2

ขั้นที่ 5. $\chi^2_{\text{คำนวณ}} = 6.673$ และ Asymp Sig $= 0.036$

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

ค่าวิกฤตคือ χ^2_α $df = k-1$ บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

เพราะฉะนั้นค่าวิกฤตคือ $\chi^2_{0.05} = 5.99$ บริเวณวิกฤตคือ $\chi^2 > 5.99$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

เพราะว่า $\chi^2_{\text{คำนวณ}} = 6.673 > 5.99$ เพราะฉะนั้น ปฏิเสธ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig $< \alpha$

เพราะว่า Sig $= 0.036 < 0.05$ เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

10.4.2 ประชากร 2 กลุ่มไม่เป็นอิสระต่อกัน

ตัวอย่าง 10.4.2 เครื่องมือ 3 แบบ สำหรับวัดปริมาณของซัลเฟอร์มอน็อกไซด์ในบรรยากาศได้ข้อมูลเป็นดังนี้

		ปริมาณของซัลเฟอร์มอน็อกไซด์ที่วัดได้แต่ละวัน		
		เครื่องมือแบบที่ 1	**เครื่องมือแบบที่ 2**	**เครื่องมือแบบที่ 3**
วันที่	1	0.96	0.87	0.76
	2	0.82	0.74	0.85
	3	0.75	0.63	0.74
	4	0.61	0.55	0.46
	5	0.89	0.76	0.78
	6	0.64	0.70	0.81
	7	0.81	0.69	0.72
	8	0.68	0.57	0.56
	9	0.65	0.53	0.56
	10	0.84	0.88	0.74
	11	0.59	0.51	0.62
	12	0.94	0.79	0.68

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าผลการวัดของเครื่องมือทั้ง 3 แบบมีผลไม่แตกต่างกัน

วิธีทำ

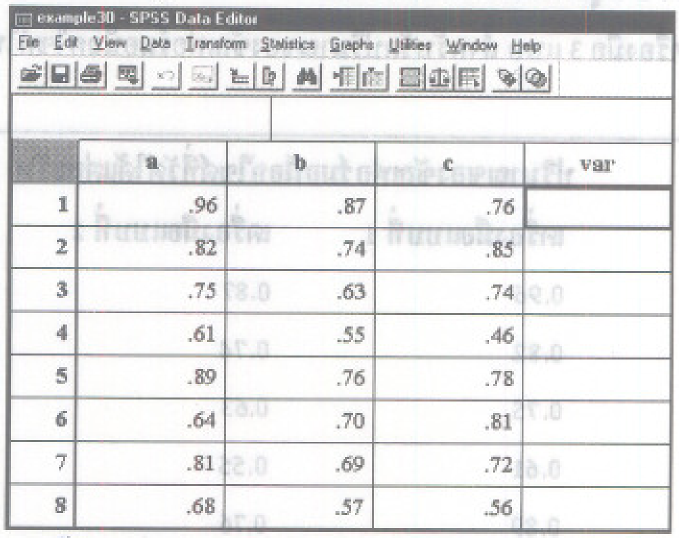
ขั้นที่ 1. กำหนดสมมติฐานหลัก $H_0 : \mu_1 = \mu_2 = \mu_3$

กำหนดสมมติฐานอื่น $H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

ขั้นที่ 2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

ขั้นที่ 3. ทำการสุ่มตัวอย่างและทำการคำนวณ

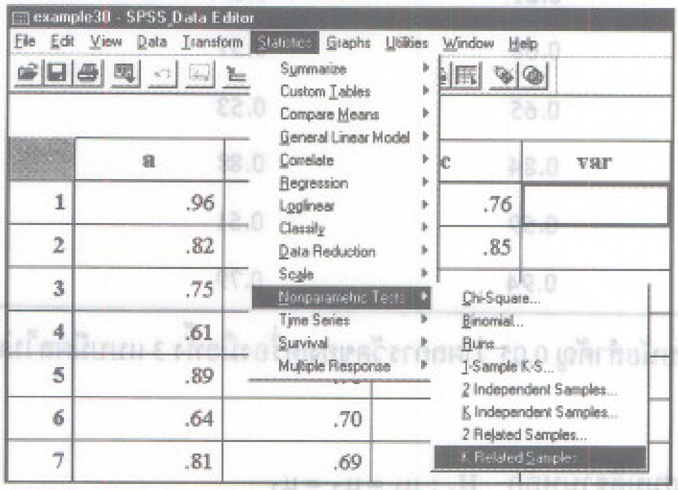
ขั้นที่ 3.1 สร้างแฟ้มข้อมูล



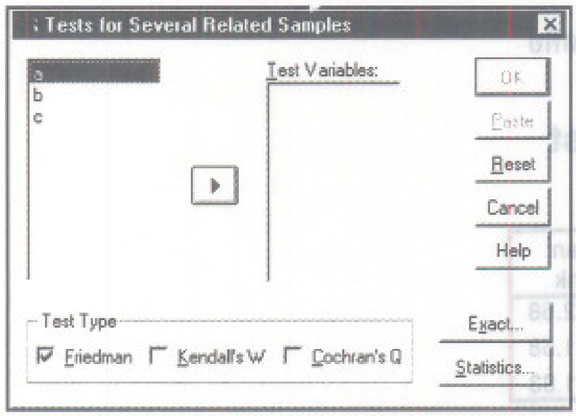
	a	b	c	var
1	.96	.87	.76	
2	.82	.74	.85	
3	.75	.63	.74	
4	.61	.55	.46	
5	.89	.76	.78	
6	.64	.70	.81	
7	.81	.69	.72	
8	.68	.57	.56	

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example30.sav

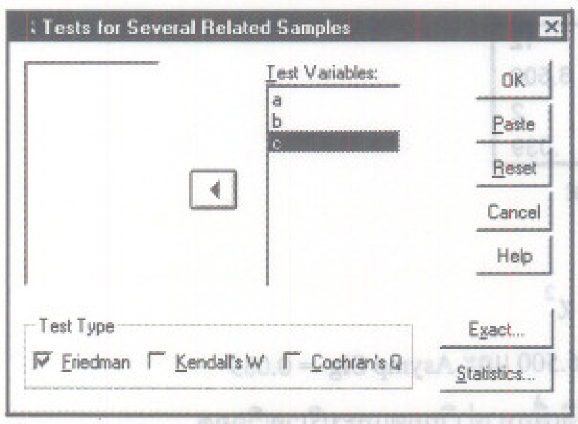
ขั้นที่ 3.2 เลือกคำสั่ง Statistics / Nonparametric / K Related Samples..



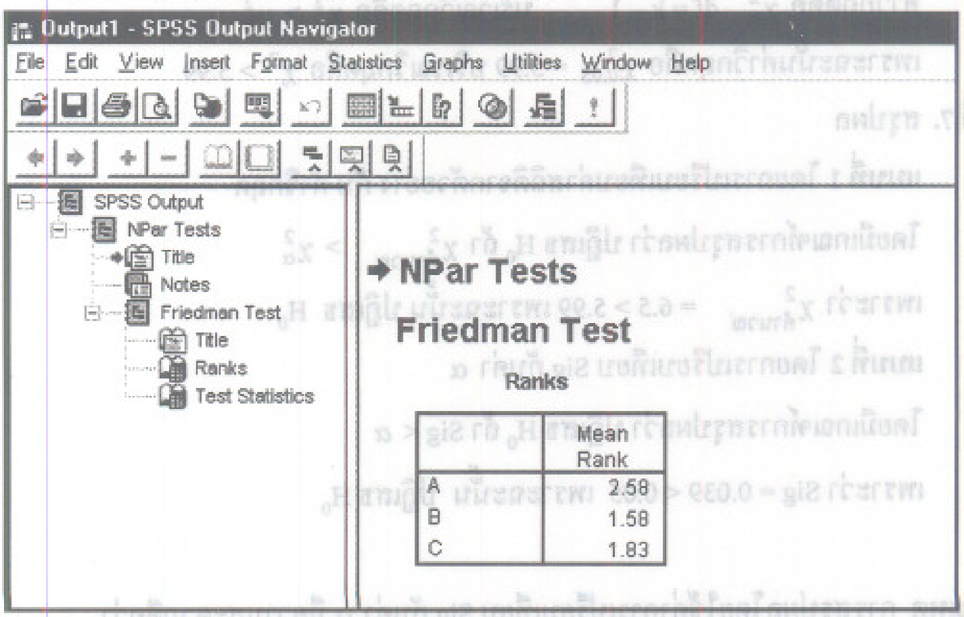
ขั้นที่ 3.3 คลิกที่ K Related Samples....จะได้เมนูย่อย



ขั้นที่ 3.4 เลือกตัวแปร a , b , c มาที่ช่อง Test Variable



ขั้นที่ 3.5 คลิกที่ OK จะได้ผลการคำนวณเป็น



ผลการคำนวณทั้งหมดคือ

NPar Tests
Friedman Test

Ranks

	Mean Rank
A	2.58
B	1.58
C	1.83

Test Statistics[a]

N	12
Chi-Square	6.500
df	2
Asymp. Sig.	.039

a. Friedman Test

ขั้นที่ 4. เลือกค่าสถิติ χ^2

ขั้นที่ 5. $\chi^2_{\text{คำนวณ}} = 6.500$ และ Asymp Sig = 0.039

ขั้นที่ 6. เปิดตารางสถิติเพื่อหาค่าวิกฤตและบริเวณวิกฤต

ค่าวิกฤตคือ χ^2_α $df = k - 1$ บริเวณวิกฤตคือ $\chi^2 > \chi^2_\alpha$

เพราะฉะนั้นค่าวิกฤตคือ $\chi^2_{0.05} = 5.99$ บริเวณวิกฤตคือ $\chi^2 > 5.99$

ขั้นที่ 7. สรุปผล

แบบที่ 1 โดยการเปรียบเทียบค่าสถิติจากตัวอย่าง กับ ค่าวิกฤต

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า $\chi^2_{\text{คำนวณ}} > \chi^2_\alpha$

เพราะว่า $\chi^2_{\text{คำนวณ}} = 6.5 > 5.99$ เพราะฉะนั้น ปฏิเสธ H_0

หรือ **แบบที่ 2** โดยการเปรียบเทียบ Sig กับค่า α

โดยมีเกณฑ์การสรุปผลว่า ปฏิเสธ H_0 ถ้า Sig < α

เพราะว่า Sig = 0.039 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

หมายเหตุ การสรุปผลโดยใช้ค่าการเปรียบเทียบ Sig กับค่า α มีความสะดวกดีกว่า

10.5 การหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ (Rank Correlation Coefficient)

การหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน (Rank Correlation Coefficient) โดยการใช้โปรแกรม SPSS

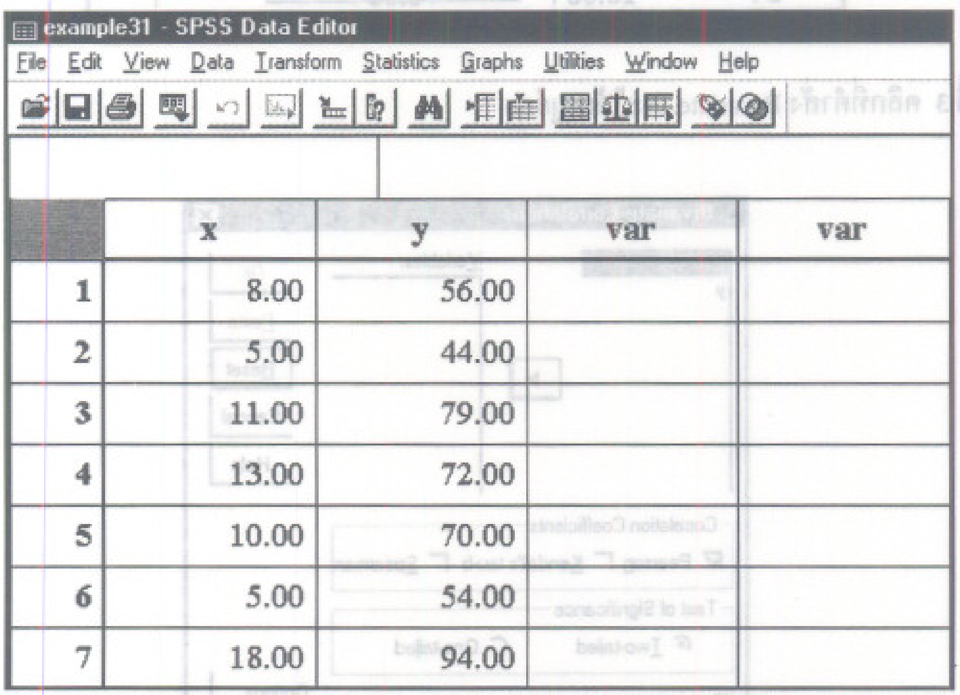
ตัวอย่าง 10.5.1 ข้อมูลของ จำนวนชั่วโมงดูหนังสือ และคะแนนสอบที่นักเรียนทำได้ เป็นดังนี้

นักเรียนคนที่	1	2	3	4	5	6	7	8	9	10
ดูหนังสือ	8	5	11	13	10	5	18	15	2	8
คะแนน	56	44	79	72	70	54	95	85	33	65

จงหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน และ

วิธีทำ

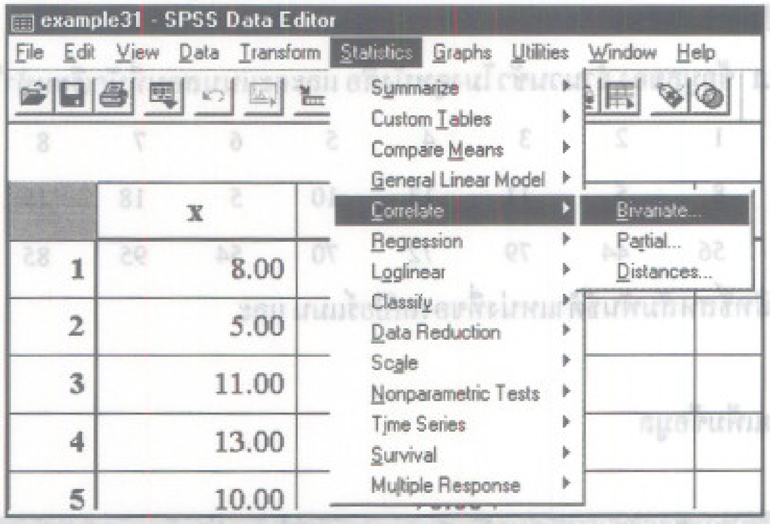
ขั้นที่ 1. สร้างแฟ้มข้อมูล



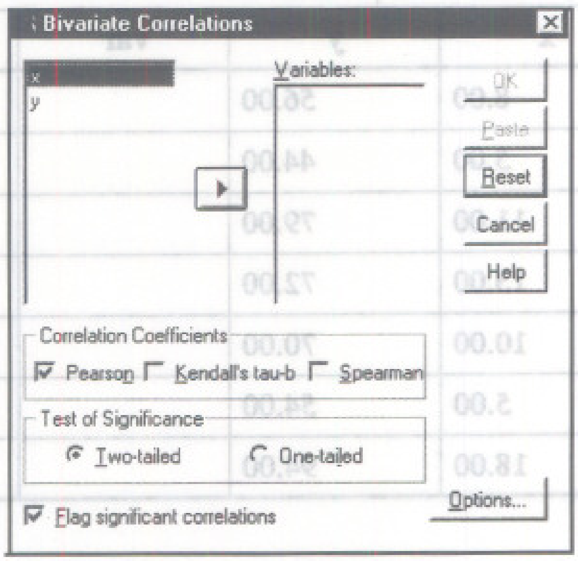
	x	y	var	var
1	8.00	56.00		
2	5.00	44.00		
3	11.00	79.00		
4	13.00	72.00		
5	10.00	70.00		
6	5.00	54.00		
7	18.00	94.00		

หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example31.sav

ขั้นที่ 2 เลือกคำสั่ง Statistics / Correlation / Bivariate..

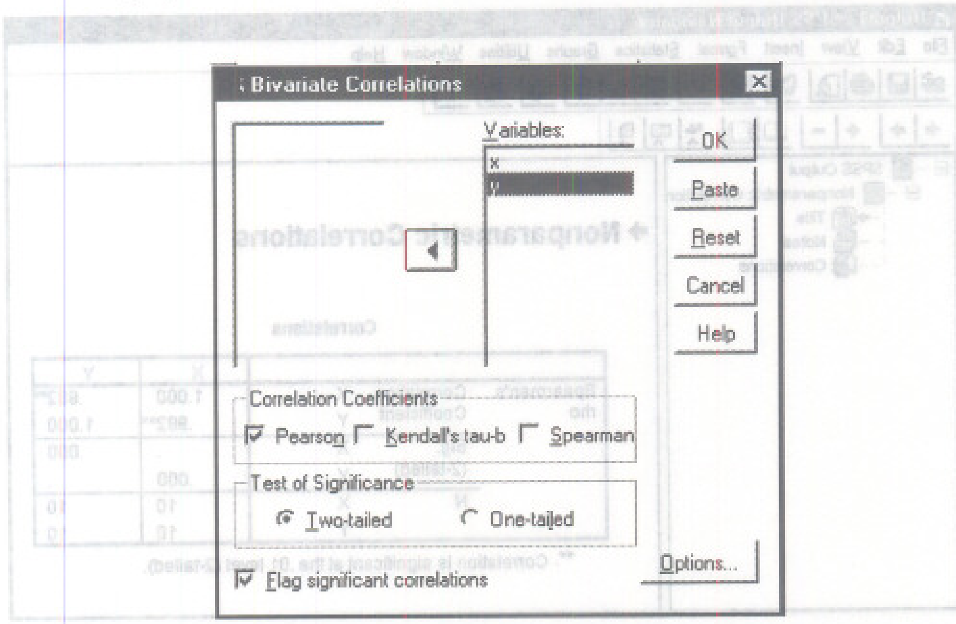


ขั้นที่ 3 คลิกที่คำสั่ง Bivariate...จะได้เมนูย่อย



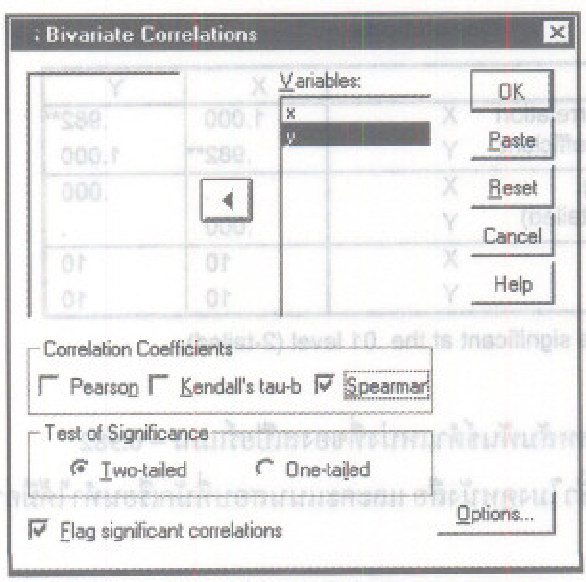
ขั้นที่ 4 คลิกที่ตัวแปร x คลิกที่ตัวแปร y

และที่ปุ่มลูกศรเพื่อนำตัวแปรคู่นั้นไปไว้ที่ช่อง Variables

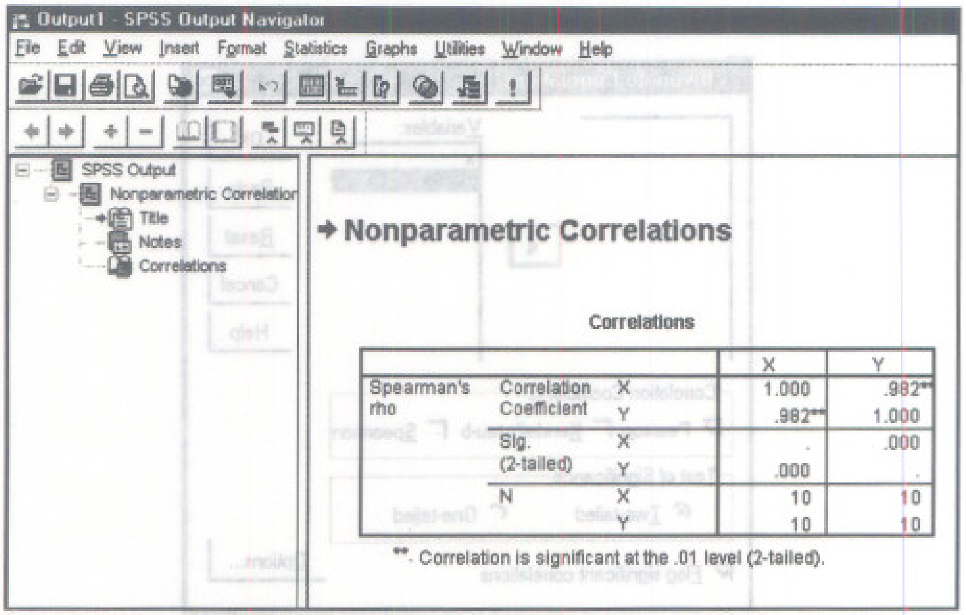


ขั้นที่ 5 คลิกที่ Pearson เพื่อยกเลิก และ

คลิกที่ Spearman เพื่อเลือกคำนวณสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน



ขั้นที่ 5 กด OK จะได้ผลการคำนวณเป็นดังนี้



ผลการคำนวณทั้งหมดคือ

Nonparametric Correlations

Correlations

			X	Y
Spearman's rho	Correlation Coefficient	X	1.000	.982**
		Y	.982**	1.000
	Sig. (2-tailed)	X	.	.000
		Y	.000	.
	N	X	10	10
		Y	10	10

**. Correlation is significant at the .01 level (2-tailed).

สรุป ค่าสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน = 0.982

เพราะฉะนั้น จำนวนชั่วโมงดูหนังสือ และคะแนนสอบที่นักเรียนทำได้มีความสัมพันธ์กันในทิศทางเดียวกัน

แบบฝึกหัด 10.

1. ข้อมูลจำนวนคนที่อยู่ในแถวเพื่อรอถอนเงินจากเครื่อง ATM ที่เก็บมาในช่วงเวลา วันที่ต่อเนื่องกันเป็นดังนี้ 6 , 7 , 5 , 6 , 8 , 6 , 8 , 6 , 6 , 4 , 3 , 2 , 4 , 4 , 3 , 4 , 7 , 5 , 6 , 8 , 6 , 6 , 3 , 5 , 2 , 5 , 4 , 4 , 3 , 7 , 5 , 5 , 4 , 3 , 7 , 4 , 6 , 5 , 2 , 8 , 5 , 6 , 8 , 7 , 9 , 4 , 5 , 1 , 1 , 3 , 2 , 9

จงทดสอบว่าจำนวนคนที่อยู่ในแถวเป็นไปอย่างสุ่ม กำหนดระดับนัยสำคัญ 0.05

2. ข้อมูลของจำนวนซัลเฟอร์ออกไซด์ ที่ออกมาจากโรงงานอุตสาหกรรมในแต่ละวัน ที่เก็บมาได้เป็นดังนี้ 18.00, 22.00, 25.00, 27.00, 9.00, 24.00, 20.00, 17.00, 6.00, 24.00, 14.00, 15.00, 23.00, 24.00, 26.00, 19.00, 23.00, 28.00, 19.00, 16.00, 22.00, 24.00, 17.00, 20.00, 13.00, 19.00, 10.00, 23.00, 18.00, 31.00, 13.00, 20.00, 17.00, 24.00, 14.00, 28.00, 19.00, 16.00, 22.00, 24.00

จงทดสอบสมมติฐานว่า ค่าเฉลี่ยของซัลเฟอร์ออกไซด์เท่ากับ 18 ที่ระดับความมีนัยสำคัญ 0.05

3. การทดสอบว่าน้ำหนักของนักเรียนมีการแจกแจงปกติจริงหรือไม่ จึงทำการสุ่มตัวอย่างน้ำหนักนักเรียนมา 80 คน ได้ข้อมูลดังนี้ กำหนดระดับนัยสำคัญ 0.05

51 , 69 , 105 , 55 , 112 , 99 , 121 , 50 , 59 , 64 , 57 , 117 , 50 , 77 , 105 , 115 , 95 , 124 , 95 , 124
119 , 121 , 107 , 65 , 92 , 101 , 129 , 66 , 90 , 110 , 117 , 115 , 101 , 96 , 126 , 127 , 149 , 111
76 , 114 , 57 , 117 , 50 , 77 , 107 , 111 , 95 , 124 , 95 , 124 , 57 , 117 , 50 , 77 , 105 , 111 , 95
124 , 95 , 124 , 109 , 121 , 107 , 65 , 95 , 101 , 125 , 66 , 90 , 110 , 117 , 145 , 101 , 96 , 116
127 , 149 , 111 , 76 , 114

4. การทดสอบว่าโปรแกรมการควบคุมน้ำหนักจะมีผลทำให้น้ำหนักลดลง ได้ทำการเก็บข้อมูลน้ำหนักของชาย 7 คน ได้ข้อมูลดังนี้

คนที่	1	2	3	4	5	6	7
น้ำหนักก่อนเข้าโปรแกรม	54	56	72	85	63	57	55
น้ำหนักหลังเข้าโปรแกรม	55	58	70	84	62	56	51

จงทดสอบสมมติฐานว่าโปรแกรมการควบคุมน้ำหนักไม่ทำให้น้ำหนักเปลี่ยนแปลง กำหนดระดับนัยสำคัญ 0.05

5. ปริมาณของนิโคตินที่มีในบุหรี่ 2 ยี่ห้อคือ

ยี่ห้อ A 2.1 , 4.0 , 6.3 , 5.4 , 4.8 , 3.7 , 6.1 , 3.3 , 2.3 , 1.3 , 2.5

ยี่ห้อ B 4.1 , 0.6 , 3.1 , 2.5 , 4.0 , 6.2 , 1.6 , 2.2 , 1.9 , 5.4 , 2.4 ,1.2 , 1.3 , 2.3

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าปริมาณของนิโคตินที่มีในบุหรี่ 2 ยี่ห้อเท่ากันหรือไม่

5. คะแนนสอบวิชาภาษาอังกฤษของนักเรียน 4 กลุ่มที่มาจากวิธีการสอนที่ต่างกัน

วิธีที่ 1 85 , 72 , 56 , 85 , 96 , 45 , 85 , 74 , 56 , 56 , 75

วิธีที่ 2 48 , 85 , 42 , 56 , 85 , 96 , 45 , 85 , 44 , 56 , 56 , 45 , 55 ,74 , 63

วิธีที่ 3 85 , 62 , 56 , 85 , 96 , 45 , 85 , 64 , 56 , 56 , 65 , 85 , 67 , 56

วิธีที่ 4 85 , 12 , 56 , 85 , 96 , 45 , 85 , 14 , 56 , 56 , 15 , 35

จงทดสอบสมมติฐานว่าวิธีการสอนทั้งสี่แบบให้ผลเหมือนกัน กำหนดระดับนัยสำคัญ 0.05

6. เครื่องมือ 4 แบบสำหรับวัดปริมาณของซัลเฟอร์มอน็อกไซด์ในบรรยากาศได้ข้อมูลเป็นดังนี้

1	2	3	4
0.74	0.76	0.76	0.82
0.46	0.85	0.85	0.74
0.78	0.74	0.74	0.63
0.81	0.46	0.46	0.55
0.72	0.78	0.78	0.76
0.56	0.81	0.81	0.70
0.56	0.72	0.72	0.74
0.69	0.57	0.56	0.46
0.57	0.53	0.56	0.78
0.84	0.88	0.74	0.81
0.59	0.51	0.62	0.72
0.94	0.79	0.68	0.57

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าผลการวัดของเครื่องมือทั้ง 4 แบบมีผลไม่แตกต่างกัน

7. ข้อมูลของ จำนวนชั่วโมงดูหนังสือ และคะแนนสอบที่นักเรียน 12 คนทำได้ เป็นดังนี้

นักเรียนคนที่	1	2	3	4	5	6	7	8	9	10	11	12
ดูหนังสือ	8	5	11	13	10	5	18	15	2	8	12	3
คะแนน	60	56	44	79	72	70	54	95	85	33	80	45

จงหาสัมประสิทธิ์สหสัมพันธ์ตำแหน่งที่ของสเปียร์แมน

ภาคผนวกที่ 1

แบบสอบถามและการสร้างแฟ้มของมูล

สิ่งที่สำคัญของผู้ที่ต้องการวิเคราะห์ข้อมูลต้องทำคือ การวางแผนเก็บข้อมูล การออกแบบรูปแบบของแบบสอบถาม การแปลความหมายของแบบสอบถามเพื่อเป็นข้อมูลของการวิเคราะห์ด้วยโปรแกรม SPSS for Windows ตัวอย่างเช่น บริษัทแห่งหนึ่งต้องการเก็บรวบรวมข้อมูลและวิเคราะห์ข้อมูลของพนักงานเกี่ยวกับ เพศ อายุ ระดับการศึกษา สถานะภาพการแต่งงาน เงินเดือน ระดับคะแนนผลการเรียน และ เงินตอบแทนประจำปี จึงทำการออกแบบของแบบสอบถามดังนี้

แบบสอบถามข้อมูลพนักงาน

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว....................................... ☐
2. เพศ ☐ ชาย ☐ หญิง ☐
3. อายุ .... ปี ☐
4. ระดับการศึกษา
   - ☐ ต่ำกว่าระดับปริญญาตรี ☐ จบระดับปริญญาตรี
   - ☐ จบระดับปริญญาโท ☐ จบระดับปริญญาเอก
5. สถานะภาพ ☐
   - ☐ โสด ☐ แต่งงานแล้ว
   - ☐ เป็นหม้าย ☐ หย่าร้าง
6. เงินเดือน .... บาท
7. ระดับคะแนน..
8. เงินตอบแทนประจำปี............................ บาท

ข้อกำหนดในการสร้างแฟ้มข้อมูล

จากแบบสอบถามที่ผู้ที่ต้องการวิเคราะห์ข้อมูล เมื่อต้องการจะทำเป็นข้อมูลสำหรับ SPSS for Windows ต้องทำการกำหนดค่าต่างๆ เช่น ชื่อแฟ้ม (file name) ชื่อตัวแปร (variable name) ชนิดของค่าตัวแปร กำหนดค่าข้อมูลที่ไม่สมบูรณ์ (missing value) คำอธิบายความหมายของชื่อตัวแปร (variable label) , คำอธิบายความหมายของค่าตัวแปร (value label)

ข้อกำหนดของแฟ้มข้อมูลที่เราต้องการเป็นดังนี้ **ชื่อแฟ้มข้อมูล Example4.sav**

1. เลขประจำตัว	กำหนดชื่อตัวแปร	id
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 3 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	ไม่มี
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี
2. เพศ	กำหนดชื่อตัวแปร	sex
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	1. Male 2. Female
3. อายุ	กำหนดชื่อตัวแปร	age
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 2 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	99
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี
4. ระดับการศึกษา	กำหนดชื่อตัวแปร	educ
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	Level of education
	คำอธิบายความหมายของค่าตัวแปร	

1. Under graduate 2. Graduate 3. Post graduate 4. Doctorate

5. **สถานะภาพ**	กำหนดชื่อตัวแปร	status
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 1 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	
	1. single 2. Married 3. Widowhood 4. Divorce	
6. **เงินเดือน**	กำหนดชื่อตัวแปร	income
	กำหนดชนิดของข้อมูล	จำนวนเต็ม 4 หลัก
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9999
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี
7. **ระดับคะแนน**	กำหนดชื่อตัวแปร	grade
	กำหนดชนิดของข้อมูล	จำนวนจริง xxx.xx
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	9.99
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี
8. **เงินตอบแทนประจำปี**	กำหนดชื่อตัวแปร	bonus
	กำหนดชนิดของข้อมูล	จำนวนจริง xxxxxx.xx
	ค่าที่กำหนดให้สำหรับข้อมูลที่ไม่สมบูรณ์	ไม่มี
	คำอธิบายความหมายของชื่อตัวแปร	ไม่มี
	คำอธิบายความหมายของค่าตัวแปร	ไม่มี

คำแนะนำสำหรับเจ้าหน้าที่พิมพ์ข้อมูล

1. เลขประจำตัว		พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม
2. เพศ	❑ ชาย	พิมพ์ข้อมูลเป็นเลข 1
	❑ หญิง	พิมพ์ข้อมูลเป็นเลข 2

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

3. อายุ พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 99

4. ระดับการศึกษา ☐ ต่ำกว่าระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 1
   ☐ จบระดับปริญญาตรี พิมพ์ข้อมูลเป็นเลข 2
   ☐ จบระดับปริญญาโท พิมพ์ข้อมูลเป็นเลข 3
   ☐ จบระดับปริญญาเอก พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

5. สถานะภาพ ☐ โสด พิมพ์ข้อมูลเป็นเลข 1
   ☐ แต่งงานแล้ว พิมพ์ข้อมูลเป็นเลข 2
   ☐ เป็นหม้าย พิมพ์ข้อมูลเป็นเลข 3
   ☐ หย่าร้าง พิมพ์ข้อมูลเป็นเลข 4

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9

6. เงินเดือน พิมพ์ข้อมูลตามค่าจริง

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9999

7. ระดับคะแนน พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

หมายเหตุ ไม่ตอบ หรือ ข้อมูลไม่สมบูรณ์ให้พิมพ์ข้อมูลเป็นเลข 9.99

8. เงินตอบแทนประจำปี พิมพ์ข้อมูลตามค่าจริงจากแบบสอบถาม

ตัวอย่างแบบสอบถามข้อมูลพนักงานที่กรอกแล้ว

สำหรับเจ้าหน้าที่กรอกข้อมูล

1. เลขประจำตัว..................................
2. เพศ ☑ ชาย ☐ หญิง [1]
3. อายุ **37** ปี
4. ระดับการศึกษา
   ☐ ต่ำกว่าระดับปริญญาตรี ☑ จบระดับปริญญาตรี [2]
   ☐ จบระดับปริญญาโท ☐ จบระดับปริญญาเอก
5. สถานะภาพ
   ☐ โสด ☐ แต่งงานแล้ว ☐ เป็นหม้าย ☑ หย่าร้าง [4]

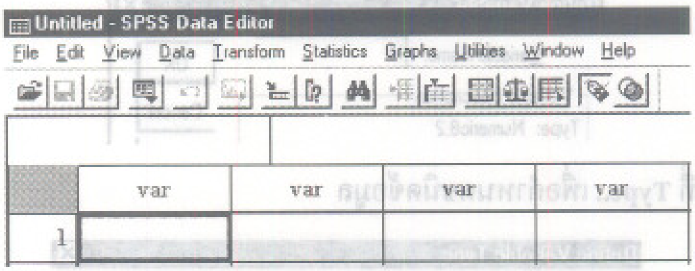
6. เงินเดือน **5500** บาท

7. ระดับคะแนน **3.78**

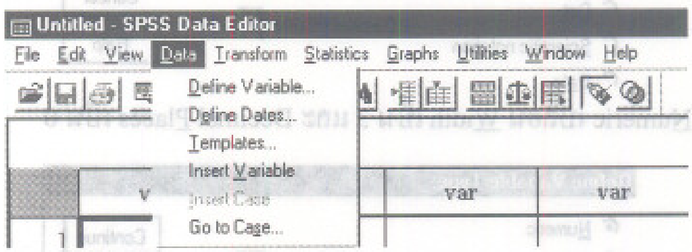
8. เงินตอบแทนประจำปี **11000.00** บาท

การสร้างแฟ้มข้อมูลใน SPSS for Windows

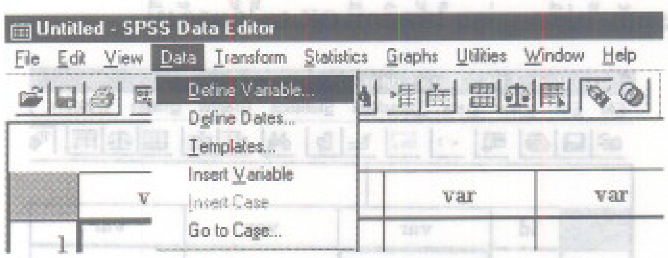
เริ่มต้นการสร้างแฟ้มข้อมูลที่ SPSS Data Editor



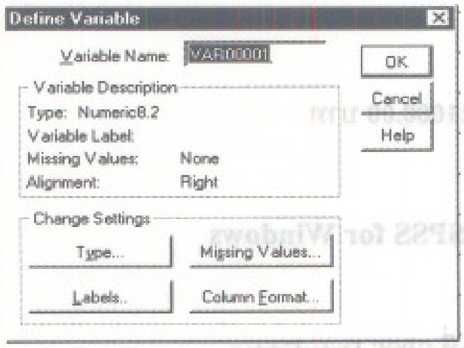
ขั้นที่ 1 คลิกที่ Data



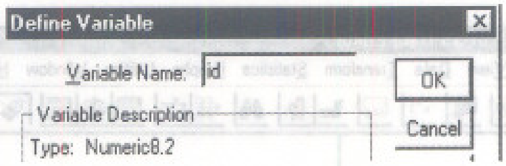
ขั้นที่ 2 คลิกที่ Define Variable..



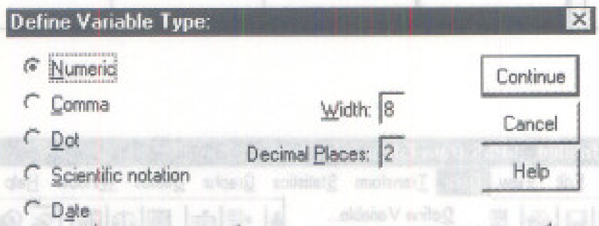
จะได้เมนูย่อยเป็น



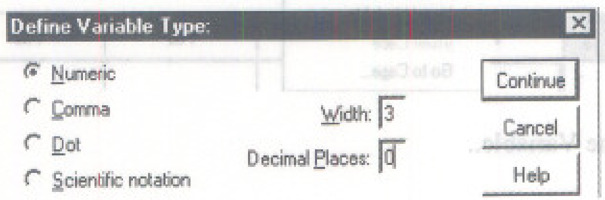
ขั้นที่ 3 พิมพ์ชื่อตัวแปร **id** ที่ช่อง **Variable Name**



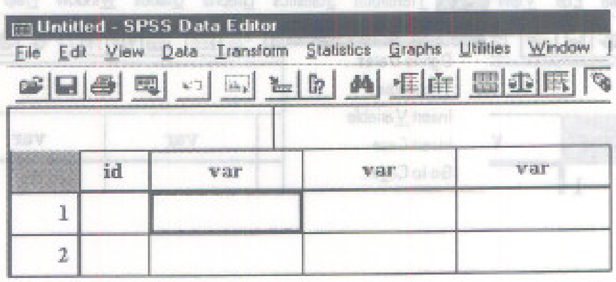
ขั้นที่ 4 คลิกที่ **Type..** เพื่อกำหนดชนิดข้อมูล



เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 3 และ Decimal Places เป็น 0

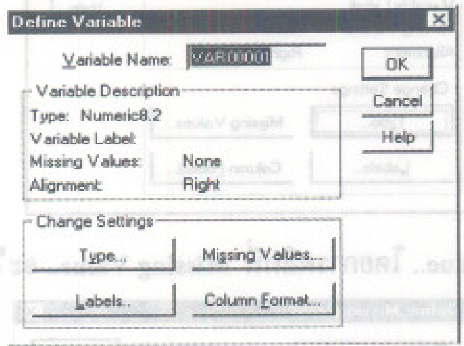


คลิก Continue จะกลับไปที่เมนูย่อย ให้คลิกที่ OK จะได้ผลดังนี้

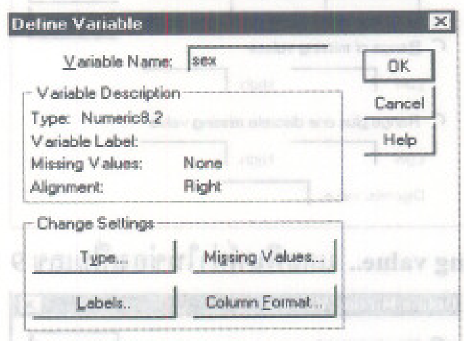


ขั้นที่ 5 การกำหนดตัวแปร sex ให้เลื่อน Pointer ไปที่ column ที่ 2

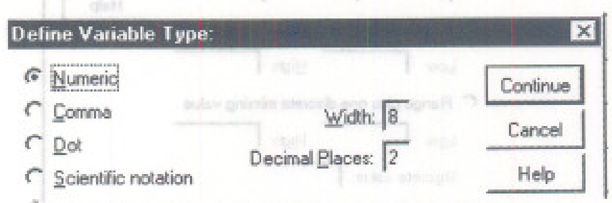
คลิกที่ Data และ คลิกที่ Define Variable.. จะได้เมนูย่อย



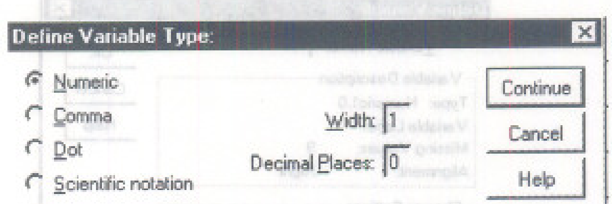
พิมพ์ชื่อตัวแปร sex



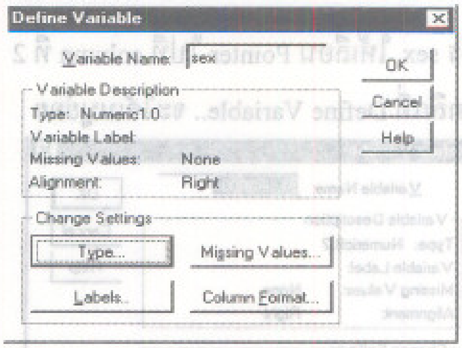
ขั้นที่ 6 กำหนดชนิดของตัวแปร sex โดยการคลิกที่ **Type..**



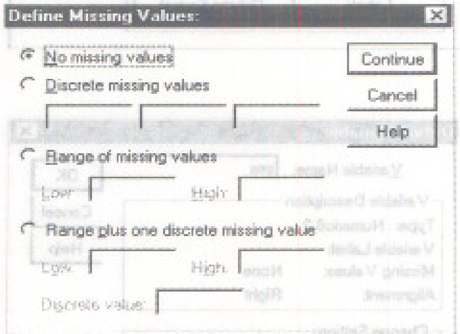
เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 1 และ Decimal Places เป็น 0



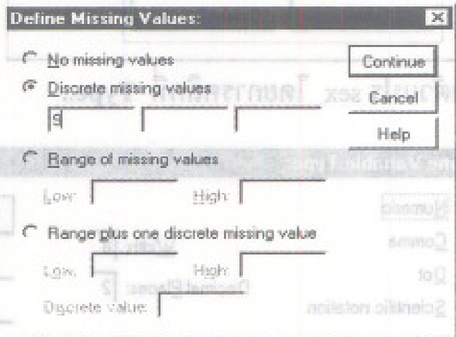
คลิก Continue จะกลับไปที่เมนูย่อย



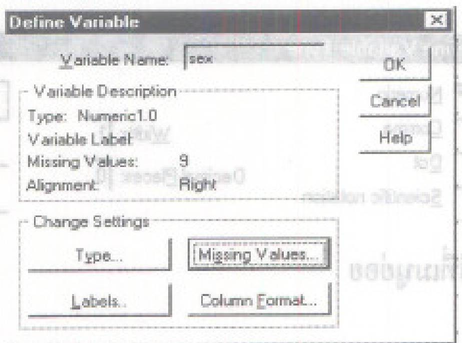
ขั้นที่ 7 กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..** จะได้เมนูย่อยเป็น



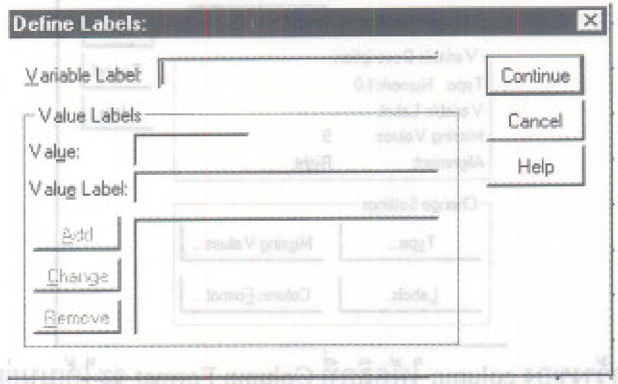
เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 9



ต่อไปคลิก Continue จะกลับไปที่เมนูย่อย

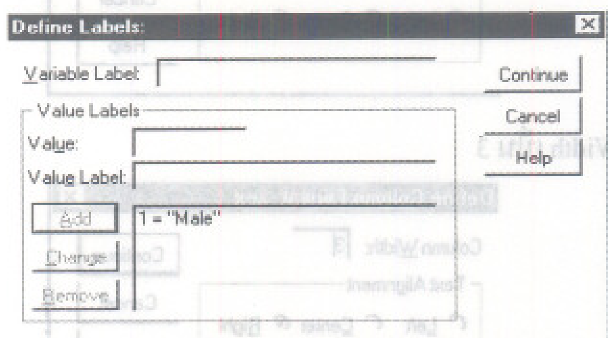


ขั้นที่ 7 การกำหนดคำอธิบายเกี่ยวกับตัวแปรให้คลิกที่ **Label..** จะได้เมนูย่อยเป็น



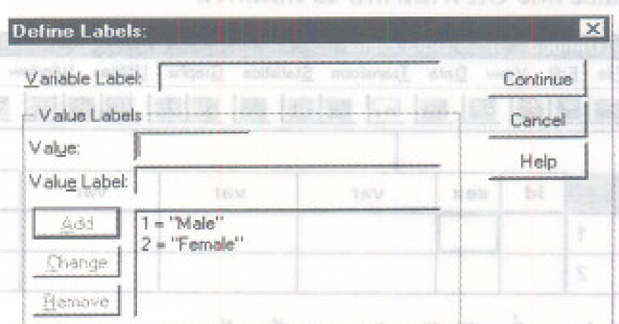
ขั้นที่ 7.1 ไปที่ช่อง Value พิมพ์ค่า 1 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label

ขั้นที่ 7.2 พิมพ์ความหมายของค่าเป็น Male เสร็จแล้วคลิกที่ Add

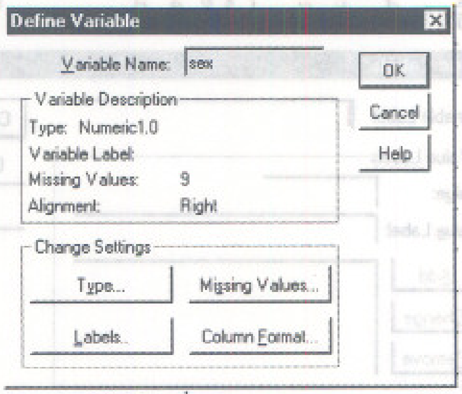


ขั้นที่ 7.3 ไปที่ช่อง Value พิมพ์ค่า 2 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label

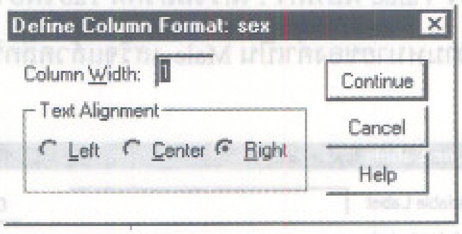
ขั้นที่ 7.4 พิมพ์ความหมายของค่าเป็น Female เสร็จแล้วคลิกที่ Add



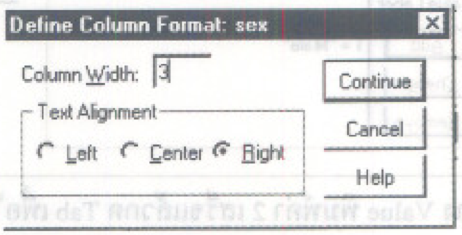
คลิก Continue จะได้เมนูย่อยเป็น



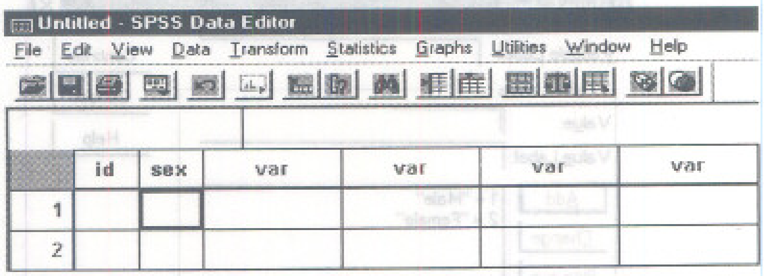
การกำหนดความกว้างของ column ให้คลิกที่ Column Format จะได้เมนูย่อยเป็น



เปลี่ยนค่า Column Width เป็น 3



เสร็จแล้วคลิก continue และ OK ตามลำดับ จะได้ผลดังนี้



ขณะนี้เรากำหนดค่าต่างๆ เกี่ยวกับตัวแปร sex เสร็จแล้ว

ในทำนองเดียวกันการกำหนดค่าเกี่ยวกับตัวแปรอื่นๆ สามารถทำได้ตามขั้นตอน โดยย่อดังนี้

การกำหนดตัวแปร อายุ age

ขั้นที่ 8 คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 8.1 พิมพ์ชื่อตัวแปร age

ขั้นที่ 8.2 กำหนดชนิดของตัวแปร age โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 2 และ Decimal Places เป็น 0

ขั้นที่ 8.3 กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..**

เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 99

ขั้นที่ 8.4 กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 3

การกำหนดตัวแปร ระดับการศึกษา educ

ขั้นที่ 9 คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 9.1 พิมพ์ชื่อตัวแปร educ

ขั้นที่ 9.2 กำหนดชนิดของตัวแปร educ โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 1 และ Decimal Places เป็น 0

ขั้นที่ 9.3 กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..**

เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 9

ขั้นที่ 9.4 การกำหนดคำอธิบายเกี่ยวกับตัวแปรให้คลิกที่ **Label..**

ขั้นที่ 9.4.1 ไปที่ช่อง Variable Label พิมพ์ Lavel of education

ขั้นที่ 9.4.2 ไปที่ช่อง Value พิมพ์ค่า 1 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Under graduate เสร็จแล้วคลิกที่ Add

ขั้นที่ 9.4.3 ไปที่ช่อง Value พิมพ์ค่า 2 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Graduate เสร็จแล้วคลิกที่ Add

ขั้นที่ 9.4.4 ไปที่ช่อง Value พิมพ์ค่า 3 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Post graduate เสร็จแล้วคลิกที่ Add

ขั้นที่ 9.4.5 ไปที่ช่อง Value พิมพ์ค่า 4 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Doctorate เสร็จแล้วคลิกที่ Add

ขั้นที่ 9.5 กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 4

การกำหนดตัวแปร สถานะภาพ status

ขั้นที่ 10. คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 10.1 พิมพ์ชื่อตัวแปร status

ขั้นที่ 10.2 กำหนดชนิดของตัวแปร status โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 1 และ Decimal Places เป็น 0

ขั้นที่ 10.3 กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..**

เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 9

ขั้นที่ 10.4 การกำหนดคำอธิบายเกี่ยวกับตัวแปรให้คลิกที่ **Label..**

ขั้นที่ 10.4.1 ไปที่ช่อง Value พิมพ์ค่า 1 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Single เสร็จแล้วคลิกที่ Add

ขั้นที่ 10.4.2 ไปที่ช่อง Value พิมพ์ค่า 2 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Married เสร็จแล้วคลิกที่ Add

ขั้นที่ 10.4.3 ไปที่ช่อง Value พิมพ์ค่า 3 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Widowhood เสร็จแล้วคลิกที่ Add

ขั้นที่ 10.4.4 ไปที่ช่อง Value พิมพ์ค่า 4 เสร็จแล้วกด Tab เพื่อไปที่ช่อง Value Label พิมพ์ความหมายของค่าเป็น Divorce เสร็จแล้วคลิกที่ Add

ขั้นที่ 10.5 กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 5

การกำหนดตัวแปร เงินเดือน income

ขั้นที่ 11 คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 11.1 พิมพ์ชื่อตัวแปร income

ขั้นที่ 11.2 กำหนดชนิดของตัวแปร income โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 4 และ Decimal Places เป็น 0

ขั้นที่ 11.3กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..**

เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 9999

ขั้นที่ 11.4กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 6

การกำหนดตัวแปร ระดับคะแนน grade

ขั้นที่ 11 คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 11.1พิมพ์ชื่อตัวแปร grade

ขั้นที่ 11.2กำหนดชนิดของตัวแปร grade โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 6 และ Decimal Places เป็น 2

ขั้นที่ 11.3กำหนด Missing Value.. โดยการคลิกที่ **Missing Value..**

เลือกชนิดเป็น **Discrete missing value..** และพิมพ์ค่าในช่องเป็นเลข 9.99

ขั้นที่ 11.4กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 6

การกำหนดตัวแปร เงินตอบแทนประจำปี bonus

ขั้นที่ 12 คลิกที่ Data และ คลิกที่ Define Variable..

ขั้นที่ 12.1พิมพ์ชื่อตัวแปร bonus

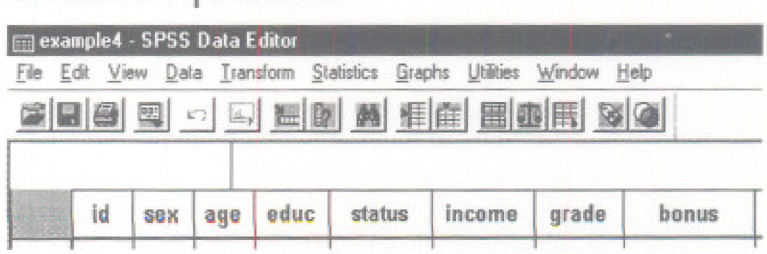
ขั้นที่ 12.2กำหนดชนิดของตัวแปร bonus โดยการคลิกที่ **Type..**

เลือกชนิดเป็น Numeric เปลี่ยน Width เป็น 9 และ Decimal Places เป็น 2

ขั้นที่ 12.3กำหนดความกว้างของ column ให้คลิกที่ Column Format

เปลี่ยนค่า Column Width เป็น 10

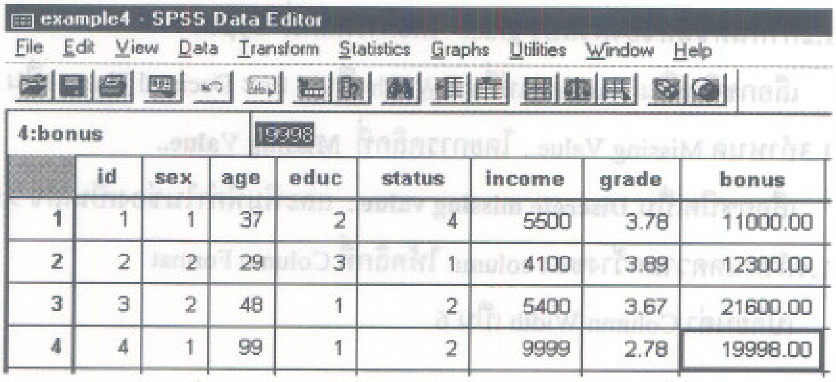
ผลบนจอภาพเมื่อกำหนดค่าต่างๆ เสร็จแล้วคือ



จากข้อมูลที่เก็บมาได้เช่น

id	sex	age	edu	status	income	grade	bonus
1	1	37	2	4	5500	3.78	11000.00
2	2	29	3	1	4100	3.89	12300.00
3	2	48	1	2	5400	3.67	21600.00
4	1	99	1	2	9999	2.78	19998.00

เมื่อพิมพ์ข้อมูลเสร็จแล้วจะได้ผลเป็น

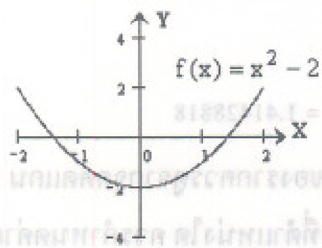


เสร็จแล้วบันทึกเป็นแฟ้มข้อมูลโดยใช้ชื่อ Example4.sav

ภาคผนวกที่ 2

การหารากสมการ f(x) = 0 และ การกำหนดข้อมูลให้กับตัวแปร

การหารากของสมการ f(x) = 0 โดยวิธีของนิวตันมีขั้นตอนดังนี้



ขั้นที่ 1 กำหนดค่าเริ่มต้น x_0 ที่เหมาะสมเพื่อเป็นประมาณของรากสมการ f(x) = 0 ค่าเริ่มต้นสามารถประมาณค่าได้จากจุดตัดแกน X ของสมการ

ขั้นที่ 2 คำนวณค่า $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$; n = 0 , 1 , 2 , 3 , ...

ขั้นที่ 3 ถ้า $\lim\limits_{n\to\infty} x_n = L$ แล้ว L เป็นรากของสมการ f(x) = 0

ตัวอย่างเช่นการหารากของสมการ $f(x) = x^2 - 2 = 0$

$f(x) := x^2 - 2$ $\quad$ $fpi(x) := 2\cdot x$ $\quad$ $x_0 := 1.5$ $\quad$ $n := 0..5$

$$x_{n+1} := x_n - \frac{f(x_n)}{fpi(x_n)}$$

n	x_n
0	1.5
1	1.41666667
2	1.41421569
3	1.41421356
4	1.41421356
5	1.41421356

ค่าประมาณของรากสมการ f(x) = 0 คือ x = 1.41421356

โปรแกรม MATHCAD มีคำสั่งที่ใช้ในการหารากของสมการ $f(x) = 0$ คือ root(f(x) ,x) โดยมีขั้นตอนการสั่งงานดังนี้

ขั้นที่ 1 กำหนดสูตร f(x)

ขั้นที่ 2 กำหนดค่าเริ่มต้น $x = x_0$

ขั้นที่ 3 ใช้คำสั่ง root(f(x) , x)

ตัวอย่างเช่น

$f(x) := x^2 - 2$

$x := 1.5$

$\mathrm{root}(f(x), x) = 1.41428818$

หมายเหตุ 1. การเลือกค่าเริ่มต้นของรากควรดูจากจุดตัดแกน X ของกราฟ

2. หากต้องการค่ารากที่ตำแหน่งใด ควรกำหนดค่าเริ่มต้น $x = x_0$ ให้ใกล้เคียงกับค่านั้น ตัวอย่างเช่น

$f(x) := x^2 - 2$

$x := -1.5$

$\mathrm{root}(f(x), x) = -1.41428818$

ความสัมพันธ์ระหว่างความถูกต้องของการคำนวณกับตัวแปร TOL

ตัวแปร TOL เป็นตัวแปรที่โปรแกรม MATHCAD กำหนดค่า default เป็น TOL = 0.001 ซึ่งเราสามารถกำหนดค่าใหม่ได้เป็น TOL = 0.1 หรือ TOL = 0.000001 ซึ่งจะมีผลต่อความถูกต้องของการคำนวณ ตัวอย่างเช่น

$f(x) := x^2 - 2$

$x := 1.5$

$TOL := 0.1$

$\mathrm{root}(f(x), x) = 1.42063492$

เมื่อเปรียบเทียบกับค่าจริง $\sqrt{2} = 1.41421356$ จะเห็นได้ว่าค่าประมาณที่ได้ถูกต้อง 1 ตำแหน่ง

$f(x) := x^2 - 2$

$x := 1.5$

$TOL := 0.001$

$\mathrm{root}(f(x), x) = 1.41428818$

เมื่อเปรียบเทียบกับค่าจริง $\sqrt{2} = 1.41421356$ จะเห็นได้ว่าค่าประมาณที่ได้ถูกต้อง 3 ตำแหน่ง

$f(x) := x^2 - 2$

$x := 1.5$

$TOL := 0.0000001$

$\mathrm{root}(f(x), x) = 1.41421356$

เมื่อเปรียบเทียบกับค่าจริง $\sqrt{2} = 1.41421356$ จะเห็นได้ว่าค่าประมาณที่ได้ถูกต้อง 8 ตำแหน่ง

การกำหนดข้อมูลให้กับตัวแปรในกรณีที่ไม่ทราบจำนวนข้อมูล

ตัวอย่างเช่น ข้อมูลคือ 12 , 15 , 23 , 17 , 28 , 23 , 26 , 32

พิมพ์	ผลบนจอภาพ
ORIGIN:1	ORIGIN := 1
i:1;20 กำหนด subscript ให้เพียงพอต่อจำนวนข้อมูลโดยต้องให้มากกว่าจำนวนข้อมูล	i := 1 .. 20
x[i:12	$x_i := 12$
,	$x_i := \begin{bmatrix} 12 \\ \blacksquare \end{bmatrix}$
15,	$x_i := \begin{bmatrix} 12 \\ 15 \\ \blacksquare \end{bmatrix}$

23,17,28,23,26,23 เมื่อพิมพ์ข้อมูลตัวสุดท้าย 32 เสร็จแล้วให้กด ENTER	$x_i :=$ $\boxed{12}$ $\boxed{15}$ $\boxed{23}$ $\boxed{17}$ $\boxed{28}$ $\boxed{23}$ $\boxed{26}$ $\boxed{32}$

ผลบนจอภาพคือ

ORIGIN := 1

i := 1 .. 20

$x_i :=$

$$\begin{array}{|c|}\hline 12\\\hline 15\\\hline 23\\\hline 17\\\hline 28\\\hline 23\\\hline 26\\\hline 32\\\hline\end{array}$$

ตัวอย่างการคำนวณเกี่ยวกับข้อมูล x เช่น

$$x = \begin{bmatrix} 12\\ 15\\ 23\\ 17\\ 28\\ 23\\ 26\\ 32 \end{bmatrix}$$

mean(x) = 22
median(x) = 23
var(x) = 41
stdev(x) = 6.40312424
length(x) = 8
min(x) = 12
max(x) = 32

หมายเหตุเราเตรียม subscript ไว้ใช้งาน 20 ตัว แต่เรา input ข้อมูลเพียง 8 ตัว เพราะฉะนั้นจำนวนข้อมูลในตัวแปร x จึงมีเพียง 8 ตัวตามที่พิมพ์เข้าไปจริง

ภาคผนวกที่ 3

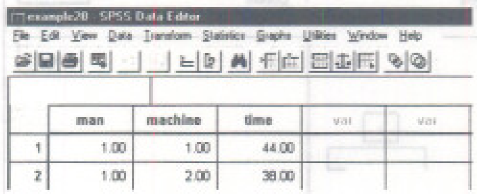
การวิเคราะห์ความแปรปรวนแบบจำแนก 2 ทางด้วย

SPSS for Windows version 8.0 และ 9.0

จากตัวอย่าง 9.2.1 หน้า 264 เมื่อวิเคราะห์ความแปรปรวนแบบจำแนก 2 ทางด้วย SPSS for Windows version 8.0 และ 9.0 ต้องทำต่อจากขั้นตอนที่ 2 หน้า 265 ดังต่อไปนี้

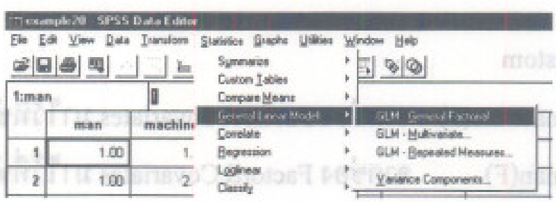
ขั้นที่ 3 นำข้อมูลเข้าสู่ SPSS

ขั้นที่ 3.1 การสร้างแฟ้มข้อมูลต้องกำหนดตัวแปร man ตัวแปรจำแนกคน machine ตัวแปรจำแนกเครื่องจักร time ตัวแปรเก็บข้อมูลที่ต้องการวิเคราะห์จาก SPSS Data Editor ของ version 8.0 หรือ 9.0

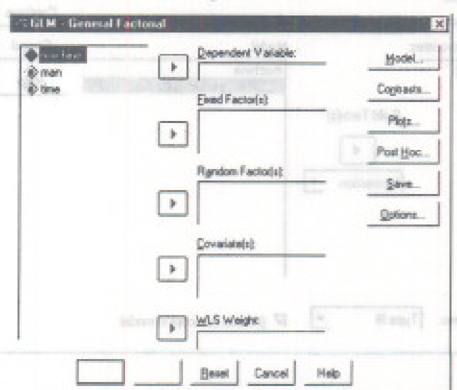


หมายเหตุ แฟ้มข้อมูลนี้ชื่อ example20.sav

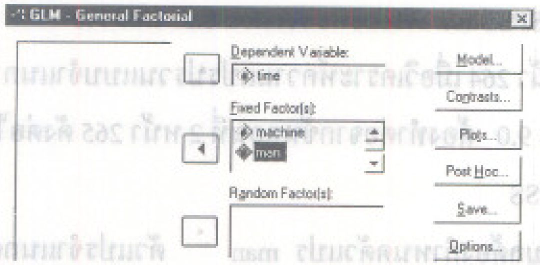
ขั้นที่ 3.2 เลือกใช้คำสั่ง Statistics / General Linear Model / GLM General Factorial..



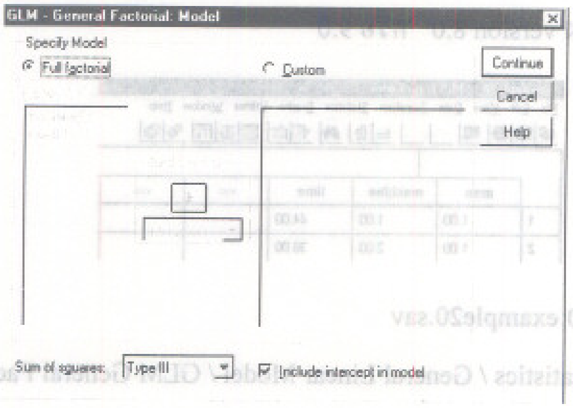
ขั้นที่ 3.3 เลือกคำสั่ง Simple Factorial จะได้เมนูย่อยดังนี้



ขั้นที่ 3.4 เลือกตัวแปร time ไปที่ช่อง Dependent Variable
เลือกตัวแปร machine ไปที่ช่อง Fixed Factor(s)
เลือกตัวแปร man ไปที่ช่อง Fixed Factor(s)

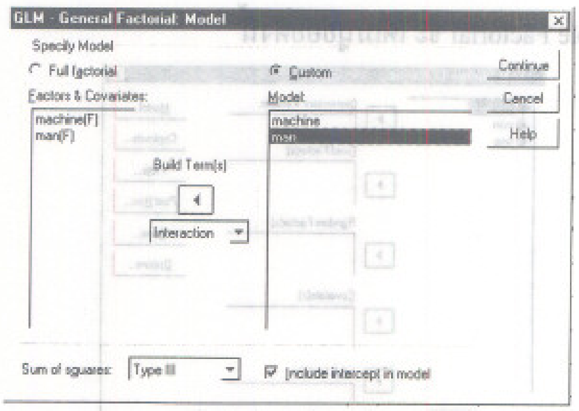


ขั้นที่ 3.5 คลิกที่ **Model** จะได้เมนูย่อย

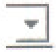


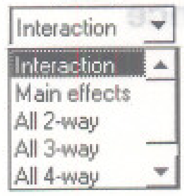
ขั้นที่ 3.6 ให้เลือก O Custom

ขั้นที่ 3.7 เลือกตัวแปร machine(F) จากช่อง Factor&Covariates มาไว้ที่ช่อง Model
เลือกตัวแปร man(F) จากช่อง Factor&Covariates มาไว้ที่ช่อง Model

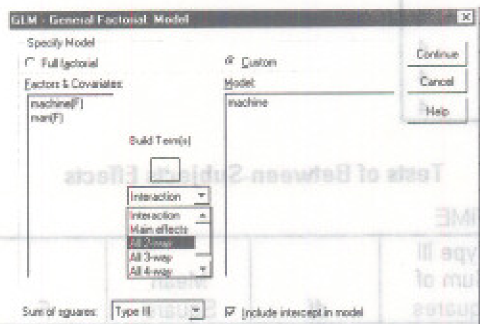


ขั้นที่ 3.8 คลิกที่ Built Term(s)

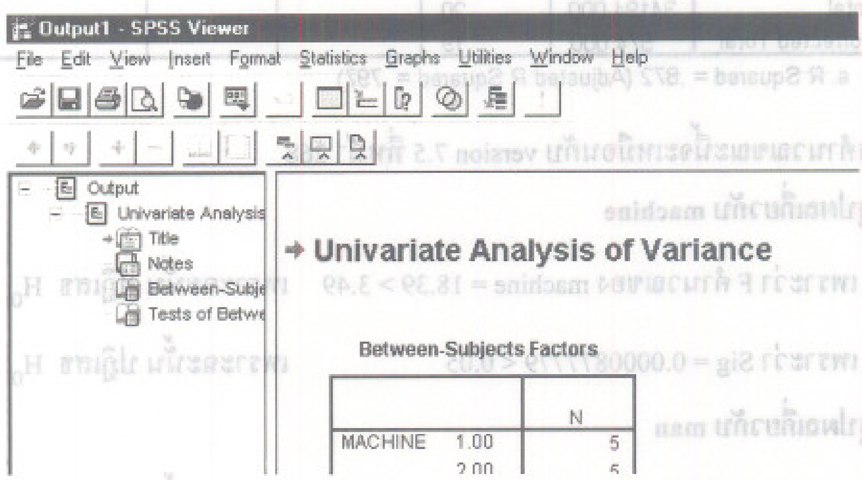
คลิกที่  จะได้



เลือก All 2–way



ขั้นที่ 3.9 คลิกที่ Continue และ OK ตามลำดับ จะได้ผลการคำนวณดังนี้



ผลการคำนวณทั้งหมดคือ

Univariate Analysis of Variance

Between-Subjects Factors

		N
MACHINE	1.00	5
	2.00	5
	3.00	5
	4.00	5
MAN	1.00	4
	2.00	4
	3.00	4
	4.00	4
	5.00	4

Tests of Between-Subjects Effects

Dependent Variable: TIME

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	500.300[a]	7	71.471	11.637	.0001669961
Intercept	33620.000	1	33620.000	5474.084	.0000000000
MACHINE	338.800	3	112.933	18.388	.0000877779
MAN	161.500	4	40.375	6.574	.0048466575
Error	73.700	12	6.142		
Total	34194.000	20			
Corrected Total	574.000	19			

a. R Squared = .872 (Adjusted R Squared = .797)

ผลการคำนวณขณะนี้จะเหมือนกับ version 7.5 ที่หน้า 268

การสรุปผลเกี่ยวกับ machine

1. เพราะว่า F คำนวณของ machine = 18.39 > 3.49 เพราะฉะนั้น ปฏิเสธ H_0

2. เพราะว่า Sig = 0.0000877779 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

การสรุปผลเกี่ยวกับ man

1. เพราะว่า F คำนวณของ man = 6.58 > 3.26 เพราะฉะนั้น ปฏิเสธ H_0

2. เพราะว่า Sig = 0.0048466575 < 0.05 เพราะฉะนั้น ปฏิเสธ H_0

ภาคผนวกที่ 4

SPSS for Windows version 7.5 , 8.0 และ 9.0

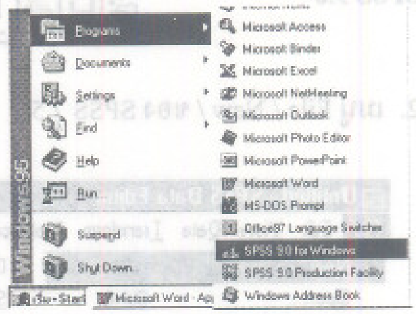
การวิเคราะห์ข้อมูลสถิติโดยใช้ SPSS for Windows version 7.5 , 8.0 และ 9.0 มีความสามารถโดยรวมใกล้เคียงกันมาก หากผู้ใช้โปรแกรม SPSS for Windows version 7.5 มาแล้ว ก็จะสามารถใช้ SPSS for Windows version 8.0 และ 9.0 ได้ทันที ในภาคผนวกที่ 4 นี้จะนำเสนอข้อแตกต่างเล็กๆ น้อยๆ ที่พบใน SPSS for Windows version 7.5 , 8.0 และ 9.0 เช่น

การเข้าสู่โปรแกรม SPSS for Windows 9.0

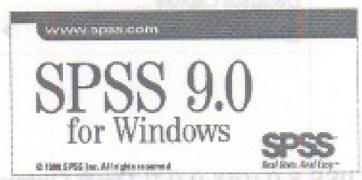
คลิก **Start**

คลิก **Programs**

คลิก **SPSS 9.0 for Windows**



จะได้ LOGO



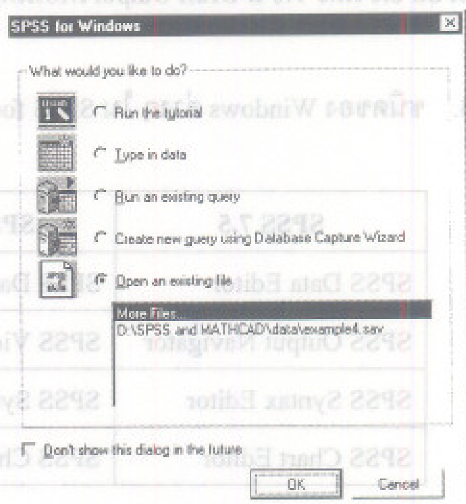
รอสักครู่จะได้

เพื่อให้เราเลือก Option ที่เหมาะสม

กับการทำงานของเราต่อไป

เสร็จแล้วคลิก OK

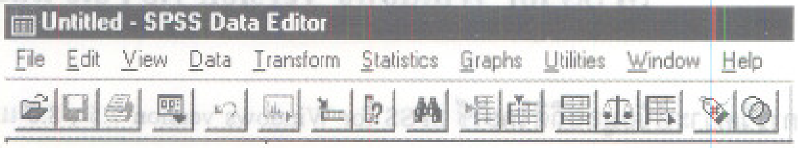
จะได้จอภาพของ **SPSS for Windows Data Editor**



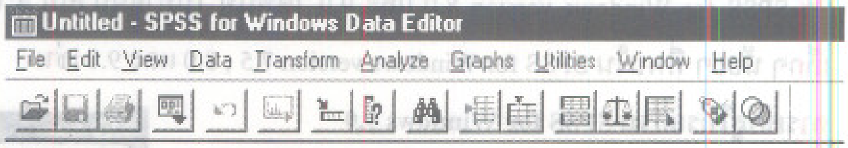
ข้อแตกต่างอื่นๆ ของ SPSS for Windows 7.5 , 8.0 และ 9.0 เช่น

1. เมนูบาร์ของ SPSS 7.5 และ 8.0 เหมือนกัน แต่ SPSS 9.0 เปลี่ยน **Statistics** เป็น **Analyze**

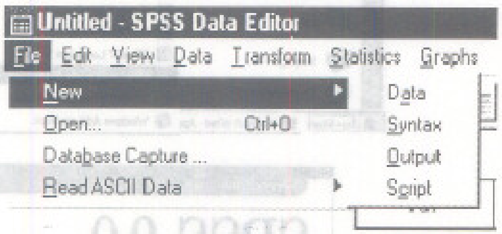
SPSS 7.5 , 8.0



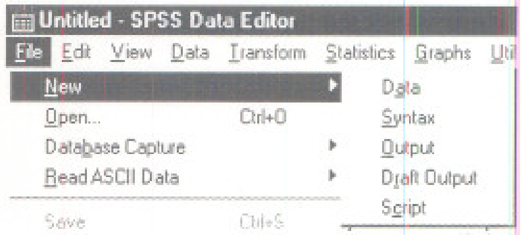
SPSS 9.0



2. เมนู File / New / ของ SPSS 7.5



เมนู File / New / ของ SPSS 8.0 และ 9.0



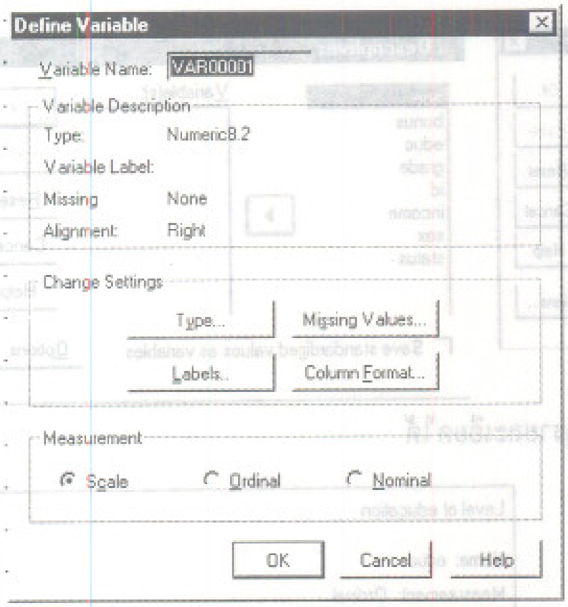
SPSS 8.0 และ 9.0 มี Draft Output เพื่อแสดง Output ในรูปแบบ Text

3. ชนิดของ Windows ต่างๆ ใน SPSS foe Windows 7.5 , 8.0 , 9.0

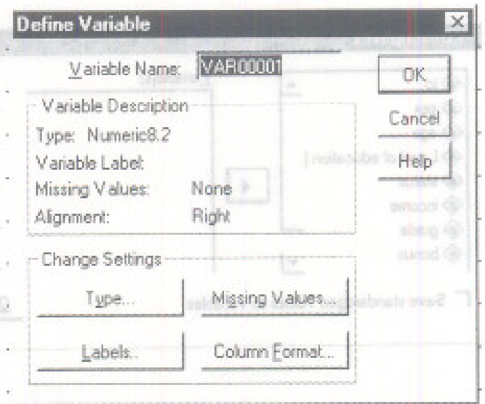
SPSS 7.5	SPSS 8.0	SPSS 9.0
SPSS Data Editor	SPSS Data Editor	SPSS for Windows Data Editor
SPSS Output Navigator	SPSS Viewer	SPSS for Windows Viewer
SPSS Syntax Editor	SPSS Syntax Editor	SPSS for Windows Syntax Editor
SPSS Chart Editor	SPSS Chart Editor	SPSS for Windows Chart Editor

4. เมนูย่อยของคำสั่ง Data / Define Variable..

SPSS 8.0 และ SPSS 9.0

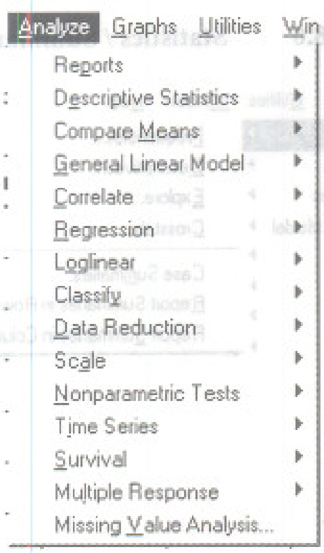


SPSS 7.5

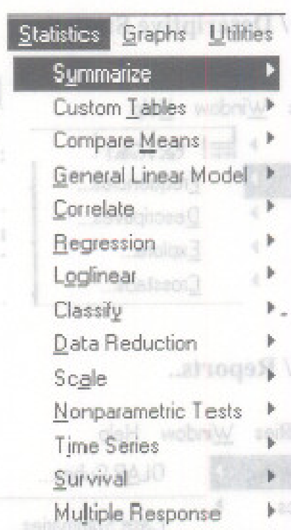


5. เมนูย่อยของ **Statistics** จาก SPSS 7.5และ 8.0 เปลี่ยนเป็น **Analyze**

SPSS 9.0

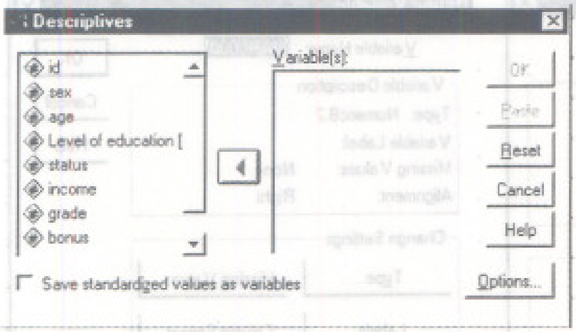


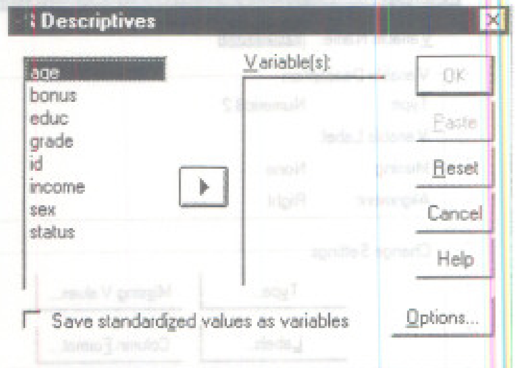
SPSS 7.5 , 8.0



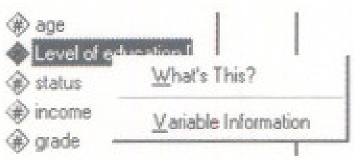
6. เมนูย่อยของการเลือกตัวแปร SPSS 8.0 และ 9.0 จะมีสัญลักษณ์บอกชนิดตัวแปร

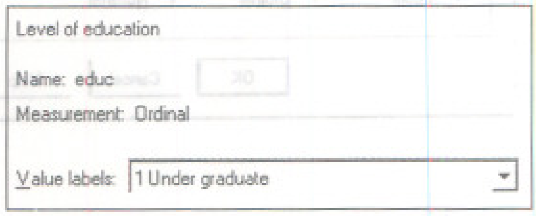
SPSS 8.0 , 9.0 **SPSS 7.5**





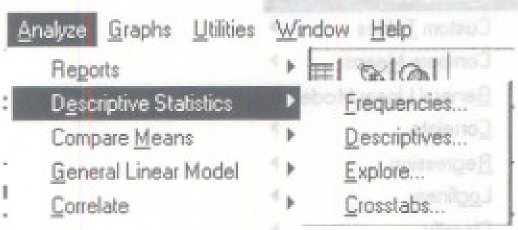
เราสามารถคลิกเมาส์ขวาที่ชื่อตัวแปรนั้นเพื่อดูรายละเอียดได้

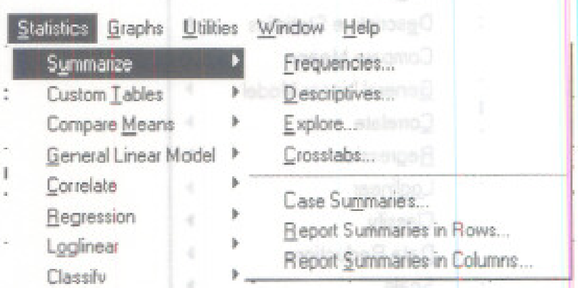




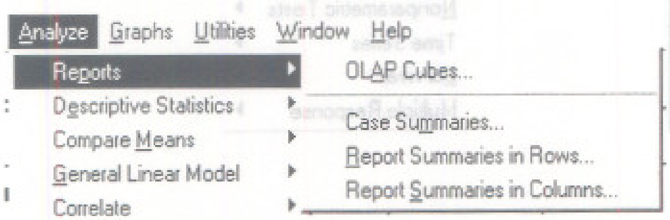
7. เมนูการวิเคราะห์ค่าสถิติเบื้องต้น

SPSS 9.0 **Analyze / Descriptive Statistics..** **SPSS 7.5 , 8.0** **Statistics / Summarize /..**





SPSS 9.0 **Analyze / Reports..**



บรรณานุกรม

Joseph G. Van Matre , Glenn H. Gilbreath , **Statistics for Business and Economics** , Third Edition ,Business Publication,Inc., Homewood, Illinois ,1987

Ronald E. Walpole , Raymond H. Myers , **Probability and Statistics for Engineers and Scientists** Third Edition , Macmillan Publishing Company , NewYork , 1985.

SPSS Base 7.5 Application Guide , SPSS Inc. USA 1997

SPSS Base 7.5 for Windows User's Guide , SPSS Inc. USA 1997

กรรณิกา ทิตาราม **สถิติเชิงคณิตศาสตร์** ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2528

คณาจารย์ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย **ความน่าจะเป็นและสถิติ** พิทักษ์การพิมพ์ กรุงเทพมหานคร 2528

ดำรงค์ ทิพย์โยธา **คู่มือ MATHCAD** โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2541

พรพรรณ แย้มกลิ่น , สุพพัดดา ปวนะฤทธิ์ **เอกสารประกอบคำบรรยาย วิชาความน่าจะเป็นและสถิติ** ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2530

ศิริชัย พงษ์วิชัย **การวิเคราะห์ข้อมูลทางสถิติด้วยคอมพิวเตอร์** พิมพ์ครั้งที่ 8 สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย กรุงเทพมหานคร 2539

ดัชนี

คำสั่งสำคัญที่ใช้คำนวณค่าสถิติของ MATHCAD

กำหนด A , B เป็นข้อมูลในรูปแบบ เวกเตอร์ หรือ เมทริกซ์

mean(A)	ค่าเฉลี่ยเลขคณิตของข้อมูลใน A
median(A)	ค่าเฉลี่ยมัธยฐานของข้อมูลใน A
length(A)	จำนวนข้อมูลในเวกเตอร์ข้อมูล A
min(A)	ค่าต่ำสุดในเวกเตอร์ข้อมูล A
max(A)	ค่าสูงสุดในเวกเตอร์ข้อมูล A
stdev(A)	ส่วนเบี่ยงเบนมาตรฐานของข้อมูลใน A (ข้อมูลใน A เป็นประชากรทั้งหมด)
Stdev(v)	ส่วนเบี่ยงเบนมาตรฐานของข้อมูลใน A (ข้อมูลใน A เป็นตัวอย่าง)
var(A)	ความแปรปรวนของข้อมูลใน A (ข้อมูลใน A เป็นประชากรทั้งหมด)
Var(A)	ความแปรปรวนของข้อมูลใน A (ข้อมูลใน A เป็นตัวอย่าง)
cvar(A,B)	ความแปรปรวนร่วมของข้อมูลใน A และ B
corr(A,B)	Pearson's correlation ของข้อมูลใน A และ B
cnorm(b)	ความน่าจะเป็น $P(-\infty < Z < b)$
slope(x,y)	สัมประสิทธิ์การถดถอยของสมการถดถอย $y = a + bx$
intercept(x,y)	ระยะตัดแกน Y ของสมการถดถอย $y = a + bx$
root(f(x),x)	หารากของสมการ $f(x) = 0$ (ต้องการกำหนดค่าเริ่มต้น $x = x_0$ ของรากด้วย)

แป้นพิมพ์และหน้าที่

F2	COPY	F3	CUT	F4	PASTE
CTRL+F9	แทรกบรรทัดว่าง	CTRL+F10	ลบบรรทัดว่าง		
ALT+O+S	แยกบริเวณสูตรที่ซ้อนกัน	CTRL+R	พิมพ์จอภาพใหม่ให้สมบูรณ์		
CTRL+P	ได้สัญลักษณ์ π	CTRL+Z	ได้สัญลักษณ์ ∞		

ตัวแปรที่กำหนดค่าไว้แล้วของ MATHCAD

TOL ช่วยกำหนดความละเอียดและถูกต้องของการคำนวณเช่น การหาราก การอินทิเกรต

ORIGIN กำหนดค่าเริ่มต้นของตัว Subscript ให้เหมาะสมกับข้อมูล

หมายเหตุ ค่าที่โปรแกรมกำหนดไว้ให้ใช้คือ TOL = 0.001 , ORIGIN = 0

Mathcad Professional

การวิเคราะห์ข้อมูลทางสถิติและความน่าจะเป็น ด้วยโปรแกรมสำเร็จรูป

SPSS for Windows & MATHCAD

เป็นหนังสือที่จะนำความสามารถของ SPSS for Windows & MATHCAD มาช่วยในการศึกษาวิชา ความน่าจะเป็นและสถิติ และ การวิเคราะห์ข้อมูลทางสถิติ ตัวอย่างของความสามารถต่าง ๆ เช่น

MATHCAD	SPSS for Windows
♠ คำนวณค่าสถิติเบื้องต้น ค่าเฉลี่ย ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน	♠ นำเสนอข้อมูลในรูปแบบตารางการแจกแจงความถี่และในรูปกราฟแบบต่าง ๆ
♥ คำนวณค่าความน่าจะเป็นของตัวแปรสุ่ม ตารางความน่าจะเป็นและเขียนโค้งความน่าจะเป็นของตัวแปรต่าง ๆ เช่น ตัวแปรสุ่ม ทวินาม ปัวส์ซอง Z T F และ ไคสแควร์	♥ คำนวณค่าสถิติเบื้องต้น ค่าเฉลี่ย มัธยฐาน ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน
♦ คำนวณค่าสหสัมพันธ์	♦ ทำการประมาณค่าพารามิเตอร์
♣ คำนวณสัมประสิทธิ์การถดถอย	♣ หาช่วงความเชื่อมั่นของค่าพารามิเตอร์
♠ คำนวณสมการถดถอยเชิงเส้น	♠ ทำการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ย ความเป็นอิสระ ภาวะสารูปสนิทดี
♥ คำนวณช่วงความเชื่อมั่น	♥ วิเคราะห์ความแปรปรวน
♦ สร้างกราฟ OC-curve	♦ หาความสัมพันธ์ระหว่างข้อมูล
♣ ฯลฯ	♣ ทำการทดสอบนอนพาราเมตริก
	♠ ฯลฯ

การวิเคราะห์ข้อมูลทางฯ
ISBN 974-639-311-1
9 789746 393119
C112
5530 250.00 บาท