หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

การนำเสนอข้อมูล

จากการเก็บรวบรวมข้อมูลที่ได้สามารถแสดงผลตามลำดับดังนี้คือ

3.1 แสดงวิธีการเรียกใช้แต่ละโปรแกรม

ในการแสดงวิธีการเรียกใช้แต่ละโปรแกรมจะเสนอในรูปแบบเดียวกันทั้ง 3 โปรแกรม คือ

- แนะนำโปรแกรมอย่างย่อ ๆ
- กล่าวถึงความสามารถของโปรแกรม
- การเตรียมชุดบัตรคำสั่งของโปรแกรม
- แสดงตัวอย่าง ของการจัดชุดบัตรคำสั่งตามข้อกำหนดของศูนย์คอมพิวเตอร์ต่าง ๆ

004134

โดยจะเริ่มจากโปรแกรม BMDP, SAS และ SPSS ตามลำดับ

การแสดงวิธีเรียกใช้แต่ละโปรแกรมมีคำและข้อความที่ใช้สำหรับโปรแกรมทั้ง 3 ในความหมายเดียวกันดังนี้

การดัดแปลงข้อมูล (Data transformation)

หมายถึง การเปลี่ยนรูปข้อมูล, การสร้างตัวแปรใหม่, การจัดกลุ่มข้อมูลใหม่, การเลือกตัวแปร หรือ การกำหนดรหัสใหม่

คำสังเกต (observation)

หมายถึง ตัวอย่างหรือประชากรที่นำมาใช้ในการวิเคราะห์หรือทำงานทางสถิติ โดยคำสังเกตหนึ่ง ๆ จะประกอบไปด้วยค่าของตัวแปร (variables) ที่ผู้วิจัยสนใจจะศึกษา

ค่าไม่สมบูรณ์ (missing value)

หมายถึง ค่าของตัวแปรใด ๆ ในคำสังเกตหนึ่งที่ผู้วิจัยกำหนดรหัสขึ้นมาเอง เพื่อใช้แทนความหมายบางอย่าง เช่น แทนความหมาย ไม่ทราบ, ไม่ตอบ, ไม่เข้าย่ายของแบบสอบถาม หรือเป็นค่าของตัวแปรที่ผู้วิจัยกำหนดเป็นอย่างอื่น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ค่าสมบูรณ์

หมายถึง ค่าของตัวแปรที่ได้จากการสังเกต, การวัดหรือการสอบถามโดยไม่เป็น ค่าไม่สมบูรณ์

ค่าที่ยอมรับได้

หมายถึง ค่าของตัวแปรที่เป็นค่าสมบูรณ์และอยู่ในช่วงที่ผู้วิจัยกำหนดไว้ ถ้าไม่ได้กำหนดไว้ ค่าที่ยอมรับได้คือ ค่าสมบูรณ์นั่นเอง

ค่าที่ไม่ยอมรับ

หมายถึง ค่าของตัวแปรที่เป็นค่าไม่สมบูรณ์ หรืออยู่นอกช่วงที่กำหนด

ค่าสังเกตที่สมบูรณ์ (Complete observation)

หมายถึง ค่าสังเกตใด ๆ ที่ตัวแปรทุกตัวในค่าสังเกตนั้นเป็นค่าที่ยอมรับได้

นอกจากนี้ยังมีคำหรือข้อความใด ๆ ที่ปรากฏภายในรูปแบบ [ ] ถ้าไม่มีการกำหนดเป็นอย่างอื่นจะมีความหมายต่อไปนี้คือ

1. ถ้าเป็นอักษรภาษาอังกฤษตัวพิมพ์ใหญ่ ถือว่าเป็นคำเฉพาะ (KEY WORD) ซึ่งผู้ใช้ต้องใช้ตามแบบที่เขียนไว้เท่านั้น

2. ถ้าเป็นอักษรภาษาอังกฤษตัวพิมพ์เล็กเป็นการบอกให้ผู้ใช้กำหนดเอาเองหรือตามข้อกำหนดของแต่ละโปรแกรม

3. วงเล็บ [ ] เป็นการบอกให้ทราบว่า ข้อความในเครื่องหมายวงเล็บนี้อาจใช้หรือไม่ใช้ก็ได้

4. วงเล็บ ( ) เป็นการบอกให้ทราบว่า ต้องใช้ข้อความตามที่กำหนดพร้อมกับเครื่องหมายวงเล็บนี้ด้วย

5. วงเล็บ { } เป็นการบอกให้ทราบว่า สามารถเลือกใช้คำหรือข้อความที่ปรากฏในวงเล็บนี้ โดยเลือกใช้เพียงอย่างใดอย่างหนึ่ง

คำหรือข้อความใด ๆ นอกเหนือจากนี้ขึ้นอยู่กับข้อกำหนดของแต่ละโปรแกรม

โปรแกรม BMDP

•BIOMEDICAL COMPUTER PROGRAMS•

โปรแกรม BMDP หรือโปรแกรม สำหรับการวิจัยทางการแพทย์ เป็นโปรแกรมสำเร็จรูปขนาดใหญ่ สำหรับทำ Data transformations และวิเคราะห์สถิติทุกระดับ โปรแกรม BMDP เป็นโปรแกรมที่ยังไม่ค่อยแพร่หลายในประเทศไทย มีใช้กันเพียงแห่งเดียว คือ ศูนย์คอมพิวเตอร์แห่งภูมิภาค สถาบันเทคโนโลยีแห่งเอเชีย ซึ่งได้เปลี่ยนแปลงเรื่อยมา ปัจจุบันที่มีอยู่เป็นโปรแกรม BMDP ปี ค.ศ. 1979
โปรแกรมนี้มีจุดประสงค์ เพื่อการวิเคราะห์สถิติสำหรับการวิจัยในทางการแพทย์

ความสามารถทั่ว ๆไปของโปรแกรม BMDP

1. การทำ Data transformations

2. สามารถเลือกเฉพาะบางส่วนของค่าสังเกต มาทำการวิเคราะห์และกำจัดข้อมูลหรือค่าสังเกตที่ไม่ต้องการได้

3. ไม่จำเป็นต้องระบุจำนวนค่าสังเกตเมื่อบันทึกข้อมูลลงบัตร

4. ผลลัพธ์จากการใช้โปรแกรม BMDP หนึ่ง สามารถนำไปเป็นข้อมูล (input) ของอีกโปรแกรมหนึ่งได้

5. การทำ Multiple transformation โดยการใช้ข้อมูลชุดเดียวทำงานต่อเนื่องกันไป เช่น หาค่าสถิติเบื้องต้น คือค่าเฉลี่ย, ส่วนเบี่ยงเบนมาตรฐานในขั้นแรก แล้วนำไปคำนวณหาคะแนนมาตรฐาน ค่าเฉลี่ยเร ยาคณิต ฯลฯ ในขั้นต่อไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ความสามารถของโปรแกรม BMDP ในการทำงานทางสถิติ

1. การคำนวณค่าสถิติเบื้องต้น (Descriptive statistics)

ซึ่งประกอบไปด้วย ค่าเฉลี่ย (mean), มัธยฐาน (median), ฐานนิยม (mode), ส่วนเบี่ยงเบนมาตรฐาน (standard deviation), ความแปรปรวน (variance), ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (standard error of mean), สัมประสิทธิ์ของความแปรผัน (coefficient of variation), ค่าสูงสุด ต่ำสุด, พิสัย (range), ความเบ้ (skewness), ความสูง (kurtosis), ส่วนเบี่ยงเบนควอไทล์ (Quartile deviation), กราฟแสดงตำแหน่งของค่าประมาณต่าง ๆ, ค่าเฉลี่ยเรขาคณิต (geometric mean), ค่าเฉลี่ยฮาร์โมนิค (harmonic mean), ค่าสถิติ . student's t, ค่าสถิติ F .

2. การแจกแจงความถี่ และการวัดความสัมพันธ์ระหว่างตัวแปร

ประกอบไปด้วย ตารางแจกแจงความถี่, ฮิสโตแกรมของความถี่, ความถี่สะสม การวัด และทดสอบความสัมพันธ์ (measure of associate and test statistic) โดยค่าสถิติต่อไปนี้ chi-square (X^2), likelihood ratio (G^2), contingency coefficient, Cramer's V, (Ø), Fisher's exact, Yates' corrected X^2, Goodman, Kruskal Kendall's τ_b, Stuart's T_c Pearson correlation (r), Spearman rank correlation r_s, Yule's Q and Y, cross-product ratio tetrachoric correlation r_t, Somers' D, Optimal prediction X, X^*, uncertainty coefficient, Mcnemar's test

3. การเขียนกราฟ และแผนภูมิ

ประกอบไปด้วย กราฟ normal , scatter (Bivariate), ฮิสโตแกรม (histogram), กราฟประกอบ การวิเคราะห์ทางสถิติต่าง ๆ

4. <u>การวิเคราะห์ความแปรปรวน และความแปรปรวนร่วม</u> (Analysis of variance and covariance)

สามารถทำการวิเคราะห์ความแปรปรวนและความแปรปรวนร่วม สำหรับข้อมูลที่มาจากการแจกแจงแบบทางเดียว (One-way ANOVA) แบบสองทาง (Two-ways ANOVA) และแบบหลายทาง (Multi-ways ANOVA) พร้อมทั้งการวิเคราะห์ที่มีการวัดซ้ำ ๆกัน สามารถวิเคราะห์ข้อมูลที่ได้จากแผนการทดลองต่าง ๆ เช่น Latin square, Incomplete block Fractional factorial, Split-plot, nested โดยจะใช้เฉพาะรูปแบบที่ได้จากผลกระทบชนิดคงที่ (Fixed effects model) และรูปแบบผสมทั่ว ๆไป (General Mixed model)

5. <u>การวิเคราะห์การถดถอย</u> Regression Analysis

สามารถทำการวิเคราะห์สมการถดถอยในรูปแบบต่าง ๆ ดังนี้

Simple Reg' , Multiple Reg' , Stepwise Reg' , All possible Reg' Regression on Principle component, Polynomial Reg' , Non-linear Reg' and maximum likelihood estimate, Stepwise Logistic Reg' , Mutivariate Reg' correlation and partial correlation

การคัดเลือกตัวแปรใน Stepwise Reg' มี 2 วิธี คือ Forward selection และ Back ward elimination.

6. <u>การวิเคราะห์ทางด้านตัวแปรพหุคูณ</u> (Multivaruate analysis) ซึ่งประกอบไปด้วยการวิเคราะห์ต่อไปนี้

Canonical correlation,Factoranalysis, Cluster analysis, Multivariate Reg; Stepwise Discriminant analysis, partial correlation

7. การวิเคราะห์ทางด้านสถิติที่ไม่ใช้พารามิเตอร์ (Nonparametric statistics)

ซึ่งประกอบด้วยค่าทดสอบดังนี้ Sign test, Wilcoxon signed-rank test, Kendall rank correlation coefficient, Spearman correlation coefficient, Friedman two-way ANOVA, Mann-Whitney U rank sum test, Kruskall-Wallis one-way ANOVA.

8. การวิเคราะห์ตารางชีพ (life table) ซึ่งใช้ในวงการแพทย์, การสำรวจสุขและการประกันชีวิต . ซึ่งประกอบไปด้วย

Life tables and Survival Functions estimate survival (time-to-accurrence)

9. การประมาณค่าเฉลี่ยโดยวิธีโรบัส (Robust estimation)

ซึ่งประกอบไปด้วยวิธีของ TRIMMED, HAMPEL, BIWEIGHT.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การเตรียมชุดบัตรคำสั่ง (Control Cards) ของโปรแกรม BMDP แบ่งเป็น 2 ประเภท คือ

1. System cards (Job Control cards)

เป็นชุดของบัตรที่กำหนดโดยศูนย์คอมพิวเตอร์แต่ละแห่ง บัตรประเภทนี้ ส่วนใหญ่จะมีเครื่องหมาย / (slash) อยู่ในคอลัมน์ที่ 1 และ 2 บัตรชุดนี้จะกำหนดรหัสของผู้มีสิทธิ์ใช้บริการของศูนย์คอมพิวเตอร์, บอกชื่อผู้ใช้, บอกให้เครื่องคอมพิวเตอร์ทราบว่า จะเรียกใช้โปรแกรม BMDP พร้อมทั้งระบุ ชุด (series) ซึ่งเรียกว่า P-series สำหรับวิเคราะห์หรือทำงานสถิติที่ต้องการ และบอกการจบของงาน บัตรชุดนี้ต้องใช้ปะหน้าและหลังของโปรแกรมทุกครั้ง

2. BMDP Control cards

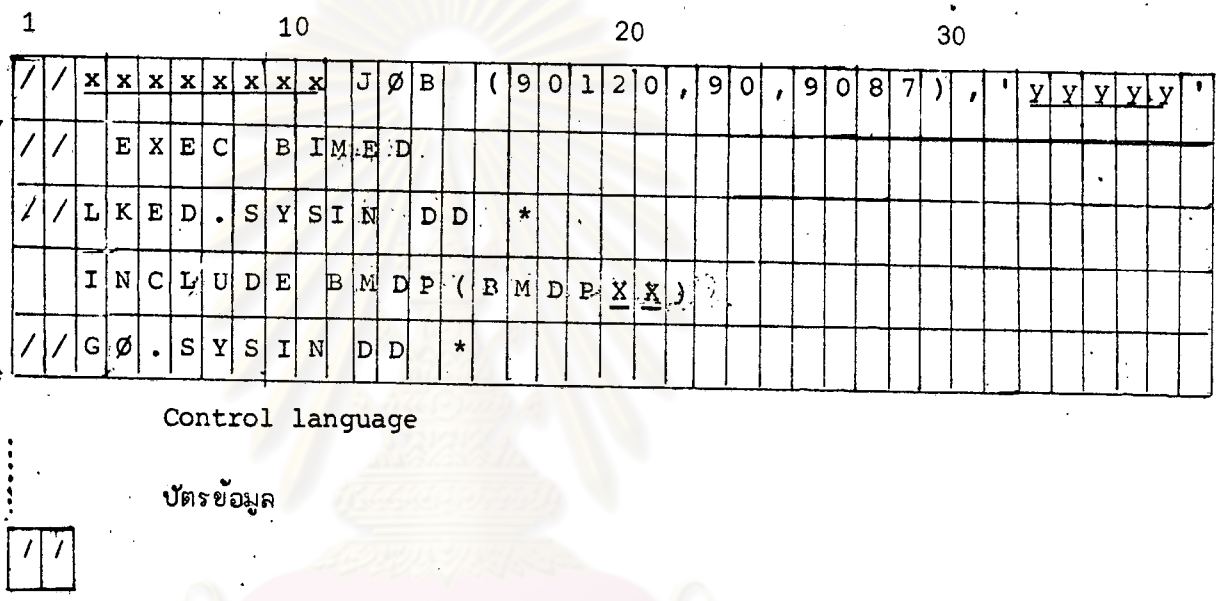
เป็นชุดของบัตรที่กำหนดโดยผู้สร้างโปรแกรม BMDP โดยเฉพาะ ซึ่งต้องเขียนตามกฎเกณฑ์ที่วางไว้ บัตรประเภทนี้ จะมีเครื่องหมาย (/) อยู่บนคอลัมน์ที่ 1 และตามด้วยคำเฉพาะของโปรแกรม BMDP (เป็น BMDP control cards ของโปรแกรมปี ค.ศ. 1979 ถ้าเป็นโปรแกรม BMDP ที่มีก่อนปี ค.ศ. 1977 จะใช้เครื่องหมาย / ไว้หลังสุดเมื่อจบคำสั่ง)[1] คำสั่งที่ใช้ในบัตรชุดนี้ใช้ภาษาอังกฤษเป็นพื้นฐาน เรียกว่า Control language บัตรชุดนี้จะบอกรายละเอียดของข้อมูล, ตำแหน่งข้อมูล, ลักษณะข้อมูล, การเปลี่ยนแปลงข้อมูล, ขยายคำสั่งเพิ่มเติมจากที่ระบุไว้ในชุดโปรแกรม (P-series) และบอกการจบของชุดคำสั่ง

---

1 W.J. Dixon, BIOMEDICAL COMPUTER PROGRAMS (California : Health Sciences Computing Facility Department of Biomathematic, School of Madicine, 1975) P.69

ตัวอย่างชุด Control cards สำหรับโปรแกรม BMDP

ของ ศูนย์คอมพิวเตอร์แห่งภูมิภาคสถาบันเทคโนโลยี แห่งเอเชีย (เอไอที)



ความหมายและสัญญลักษณ์ที่ใช้

Jobcard เป็นปัตรที่ศูนย์คอมพิวเตอร์แต่ละแห่งกำหนด

xxxxxxxx หมายถึง ให้ระบุชื่องานที่ผู้ใช้กำหนดขึ้นไม่เกิน 8 ตัวอักษร

YYYYY หมายถึง ให้ระบุชื่อ ผู้ใช้โปรแกรมนี้ไม่เกิน 5 ตัวอักษร

BMDPXX หมายถึงให้ระบุชื่อโปรแกรม BMDP และชุดที่ต้องการเรียกใช้ เช่น

BMDP1D, BMDP3S

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ความหมายของชุดโปรแกรม P-series ใน System cards

โปรแกรม BMDP แบ่งโปรแกรมเป็นชุด (series) โดยจะใช้ตัวอักษรและตัวเลขรวม 3 ตัว ในการเรียกใช้ โดยกำหนดดังนี้

1. ตัวแรกใช้ตัวอักษร P ซึ่งมาจากโปรแกรม BMDP

2. ตัวที่ 2 ใช้ตัวเลขระหว่าง 1-9 หรือ ตัวอักษร ระบุ โปรแกรมย่อย ในแต่ละชุดตามที่กำหนดไว้

3. ตัวที่ 3 ใช้ตัวอักษรซึ่งเป็นชุดของโปรแกรมที่จัดไว้มีดังนี้

ชุด D : Data description

ชุด F : Frequency tables

ชุด R : Regression Analysis

ชุด V : Analysis of Variance

ชุด M : Multivariate Analysis

ชุด L : life tables, Survival Analysis

ชุด S : Special (miscel laneous)

เวลาจะเรียกใช้เราก็ระบุชุดที่ต้องการใน System cards บางครั้งเราอาจเรียกใช้ถึง 2 ชุด เช่น เราต้องการวิเคราะห์ Multivariate Regression เราอาจต้องใช้ ชุด M (Multivariate Analysis) และชุด R (Regression Analysis)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บัตรคำสั่ง (Control cards) ประเภท BMDP control card

เป็นชุดของบัตรที่จะบอกรายละเอียดของข้อมูลที่จะอ่านเข้าไป, กำหนดชื่อตัวแปรและตำแหน่งที่บันทึกค่าของตัวแปรในบัตรข้อมูล บอกรหัสที่ใช้สำหรับค่าที่ไม่สมบูรณ์ (missing value) ขอบเขตของตัวแปร, บอกรายละเอียดที่จะเพิ่มผลลัพธ์จากการวิเคราะห์ข้อมูล ฯลฯ

บัตร Control cards ประเภทนี้ใช้ได้ตั้งแต่คอลัมน์ที่ 1 ถึง 80 โดยทั่ว ๆ ไป บัตรคำสั่งประเภท BMDP control cards จะมีโครงสร้างดังนี้

/ paragraph names commands or assignment statement

paragraph names เป็นชื่อหรือคำเฉพาะที่ใช้สำหรับโปรแกรม BMDP และเป็นส่วนประกอบของประโยค (sentence)ที่จะสั่งหรือกำหนดค่า

{ Commands or Assignment } Statement เป็นคำสั่งหรือการกำหนดค่าบางอย่าง เพื่อให้คอมพิวเตอร์ทราบ Statement ใช้ประกอบกับ Paragraph ซึ่งขึ้นอยู่กับว่า paragraph หนึ่ง ๆ จะให้ Statement นั้นใช้ได้หรือไม่ Statement อาจเป็นประโยค (sentence) ซึ่งอ่านเข้าใจได้ง่ายโดยมีพื้นฐานจากภาษาอังกฤษ Statement เดียวกัน อาจใช้กับ Paragraph ต่างกันได้ แต่ความหมายอาจจะต่างกันด้วย

ข้อกำหนดและความหมายของภาษาที่ใช้ในบัตร BMDP Control Cards.

1. paragraph name หนึ่ง ๆ ต้องนำด้วยเครื่องหมาย / บนคอลัมน์ที่ 1

2. paragraph name หนึ่ง ๆ อาจประกอบด้วยหลาย ๆ Statement ได้ โดย Statement ใหม่นั้น ต้องใช้บัตรอีก 1 ใบ ซึ่งอาจเริ่มที่คอลัมน์ใดก็ได้ (ควรจะเริ่มที่คอลัมน์เดียวกับ Statement แรก)

3. เมื่อจบ Statements หนึ่ง ๆ ต้องใส่เครื่องหมาย (.) ทุกครั้ง เพื่อแสดงการจบประโยค

4. ภายใน Statements หนึ่ง ๆ ควรเว้นวรรคหรือใช้เครื่องหมายให้เหมาะสม เพื่อแสดงการแยกคำ

5. ถ้าใน Statements หนึ่ง ๆ ต้องใช้บัตรมากกว่า 1 ใบ ก็สามารถต่อบัตรได้โดย จะเริ่มที่คอลัมน์ใดก็ได้

6. รูปแบบโดยทั่ว ๆไปของ Statement เป็นดังนี้ คือ

item name { =, is, ARE } Value(s) assigned to the Special for item

item name เป็นคำหรือข้อความเฉพาะซึ่งจะบ่งไว้ในส่วนแรกของ Statement

{ =, is, ARE } ให้เลือกใช้ได้ตามหลัง item name ตามแบบของภาษาอังกฤษ

Value (s) assigned to the Special for item เป็นค่าหรือ ชื่อตัวแปรที่กำหนดแต่ละแบบของ item หนึ่ง

7. ภาษาอังกฤษตัวพิมพ์ใหญ่และตามตัวพิมพ์เล็กที่ใช้ใน paragraph name หรือ item name หมายความว่า ต้องใช้ตามแบบนั้น โดยอาจจะใช้ชื่อเต็มที่แสดงไว้ (ทั้งตัวพิมพ์ใหญ่, พิมพ์เล็ก) หรือ อาจจะใช้แบบย่อ เฉพาะตัวพิมพ์ใหญ่ก็ได้โดยตัดตัวพิมพ์เล็กออกไป

8. ' c ' หมายถึง รูปแบบหรือข้อความที่ต้องเขียนในเครื่องหมาย

c หมายถึง ชื่อที่ผู้ใช้ตั้งขึ้นเองโดยใช้ได้ไม่เกิน 8 ตัว

หมายถึง ให้ใช้ตัวเลขบอกจำนวน

V หมายถึง ชื่อตัวแปรหรือ ชื่อตัวแปรที่มีตัวกำกับ (Sub script)

9. ชุดคำสั่งหนึ่งอาจประกอบด้วยหลาย paragraph และในการทำงานครั้งหนึ่งต้องใช้ END paragraph ทุกครั้ง มีรูปแบบ ดังนี้

```
/ END
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Paragraph name บางชนิดที่จำเป็นต้องใช้โดยทั่ว ๆ ไป เรียงตามลำดับบัตร ดังนี้

1. PROBLEM Paragraph

เป็น paragraph ที่จำเป็นต้องมีทุกครั้ง ใช้บอกชื่อการวิเคราะห์นั้น ๆ มีรูปแบบดังนี้ คือ

```
/ PROBlem        Statement
```

Statement ที่ใช้ประกอบ PROBlem paragraph มีอยู่แบบเดียว คือ

```
/ PROBlem        TITLE  = ' c '.
```

' c ' หมายถึง ให้บอกชื่อหรือข้อความที่จะบอก การวิเคราะห์หรือการทำงานทางสถิตินั้น ๆ โดยเขียนในเครื่องหมาย ' ' ซึ่งตัวอักษรและสัญญลักษณ์ในเครื่องหมาย ' ' ใช้ได้ไม่เกิน 160 ตัว จะพิมพ์ชื่อออกมาทางผลลัพธ์ อาจไม่จำเป็นต้องใช้ TITLE Statement ก็ได้ PROBlem paragraph ไม่จำเป็นต้องมี Statement ใด ๆ ตามมาก็ได้ ตัวอย่างการใช้ PROBlem paragraph

```
/ PROBlem        TITLE  IS  COMPARE  BMDP  SAS  SPSS
                      STATISTICAL  PACKAGE  PROGRAMS.
```

2. INPUT paragraph

เป็น paragraph ที่ต้องมีทุกครั้งสำหรับอธิบายรายละเอียดของข้อมูล เช่น จำนวนค่าสังเกต จำนวนตัวแปร รูปแบบของข้อมูล และบอกสื่อที่ใช้บันทึกข้อมูล

```
/ INPUT          Statement
```

Statements ที่ใช้กับ INPut paragraph มีหลาย Statement แต่จะกล่าวถึง Statement ที่จำเป็นดังนี้

2.1 | VARiable = #. |

ใช้บอกจำนวน (#) ตัวแปรที่ใช้ในการวิเคราะห์ Statement นี้ต้องใช้ทุกครั้ง

2.2 | FØRMat is '(c)'. |

ใช้อธิบาย ชนิดของข้อมูลที่จะอ่านเข้าไป พร้อมกับบอกตำแหน่งที่อยู่ของค่าตัวแปร

<u>การบอกชนิดและขอบเขตของตัวแปรที่จะอ่านเข้าไป</u>

ก. <u>แบบ F-Type</u> รูปแบบ คือ nFw,d

ใช้เมื่อค่าตัวแปรหรือข้อมูลมีค่าเป็นตัวเลขล้วน ๆ และเป็นตัวเลขที่มีค่าเป็นทศนิยม แม้ว่าสิ่งที่ใช้บันทึกจะบันทึกค่าแบบมีหรือไม่มีทศนิยมก็ตาม ซึ่งทุกโปรแกรมของ BMDP ส่วนใหญ่จะใช้แบบ F-Type ยกเว้นชุด P4D

n เป็นตัวเลขแสดงจำนวนตัวแปรที่มีรูปแบบเช่นนี้ n ตัว

w เป็นตัวเลขแสดงความกว้างของจำนวนคอลัมน์ที่ใช้

d เป็นตัวเลขแสดงจำนวนตำแหน่งหลังจุดทศนิยม

ข. <u>แบบ A-Type</u> รูปแบบคือ nAw (n,w มีความหมายเหมือน F-Type)

ใช้เมื่อค่าตัวแปรมีค่าเป็นตัวอักษร หรือ ตัวอักษรปนตัวเลขในโปรแกรม BMDP จะใช้ในชุด P4D และ จะไม่ใช้ในการคำนวณ

<u>การบอกตำแหน่งของตัวแปร</u>

ก. ใช้แบบ nX เมื่อต้องการบอกจำนวน n คอลัมน์ที่ต้องการเว้น

ข. ใช้แบบ / ซึ่งหมายถึงการขึ้นบัตรใหม่หรือบอกจำนวนบัตรที่ต้องการเว้น

2.3

```
CASE = #.
```

ใช้บอกจำนวน (#) ของค่าสังเกต (observation) ที่ต้องการจะใช้ในการทำงานทางสถิติ ไม่ต้องใช้ Statement นี้เมื่อข้อมูลมาจากแฟ้มข้อมูลของ BMDP

2.4

```
UNIT = #.
```

ใช้เมื่อข้อมูลบันทึกบนเทปแม่เหล็ก (tape) หรือจานแม่เหล็ก (disk) เพื่อบอกว่าข้อมูลที่จะใช้นั้นเก็บไว้ที่ตำแหน่งหน่วยใด โดยบ่งตัวเลข (#) ซึ่งตัวเลขเหล่านี้ต้องเป็นตัวเดียวกับที่บ่งไว้ในส่วนของ System card.

- ห้ามใช้เลข 1, 2, 6
- ถ้าใช้เลข 5 แสดงว่า ข้อมูลเก็บไว้ที่บัตร ซึ่งถ้าข้อมูลเก็บไว้ที่บัตร ไม่จำเป็นต้องใช้ UNIT Statement.

2.5

```
{ REWIND.    }
{ NO REWIND. }
```

ใช้เมื่อบันทึกข้อมูลบนเทปหรือ disk โดยทั่ว ๆไปโปรแกรม BMDP จะถือว่าเริ่มใช้ข้อมูลจากค่าสังเกตแรกที่บันทึกบนเทปหรือ disk แต่เราอาจจะไม่ใช้ทุกค่าสังเกตที่มีอยู่ เมื่อใช้ค่าสังเกตเท่าที่ต้องการทำงานอย่างหนึ่งแล้ว และเราต้องการทำงานที่สองโดยใช้ค่าสังเกตชุดเดิม หรือ เริ่มจากค่าสังเกตแรก ก็ต้องใช้ REWIND Statement เมื่อต้องการทำงานที่สอง โดยใช้ค่าสังเกตต่อมาจากค่าสังเกตแรกก็ใช้ NO REWIND Statement.

ตัวอย่างการใช้ Statement ประกอบ INPUT paragraph (ข้อมูลบันทึกบน cards)

```
/ INPUT     VARIABLES  ARE  9.
            FORMAT  IS  ' ( A4,  5F4.0,  3F4.1)'.
            CASES  ARE  188.
```

3. VARIABLE paragraph

เป็น paragraph ที่ต้องใช้ทุกครั้ง เมื่อข้อมูลไม่ได้อยู่บนแฟ้มข้อมูลของ BMDP ใช้บอกชื่อตัวแปร, ค่าสูงสุดและต่ำสุดของแต่ละตัวแปร, บอกรหัส หรือค่าที่ใช้สำหรับค่าที่ไม่สมบูรณ์ (Missing Value), เลือกตัวแปรที่จะใช้ในการวิเคราะห์ และใช้พิมพ์ชื่อหรือสัญญลักษณ์ตัวแปรชนิด A-Type มีรูปแบบดังนี้

```
/ VARiable        Statement
```

Statements ที่อาจใช้ได้กับ VARiable paragraph มีดังนี้

3.1

$$\text{NAME} = \left\{ \begin{array}{l} c_1, c_2, \ldots \\ x(1), x(2) \ldots \end{array} \right\}$$

ใช้บอกชื่อตัวแปร (c_1, c_2 ...) หรือบอกตามลำดับ Subscript x(1), x(2) โดยจะพิมพ์ชื่อเหล่านี้ทางผลลัพธ์ กำหนดชื่อตัวแปรได้ไม่เกินชื่อละ 8 คอลัมน์ และต้องขึ้นต้นชื่อด้วยตัวอักษร ถ้าไม่ได้ขึ้นต้นด้วยตัวอักษร หรือภายในชื่อนั้นมีสัญญลักษณ์อื่นปรากฏ เช่น (blank, comma, numeric, etc) ต้องเขียนภายในเครื่องหมาย Quotation-marks (' ')

3.2

```
MAXimum   =  #1, #2, ...
MINimum   =  #1, #2, ...
```

ใช้บอกช่วงของค่าตัวแปรที่จะใช้ในการทำงาน โดยบ่งค่าสูงสุด และต่ำสุดของตัวแปรที่ต้องการใช้ในรูปแบบข้างบน #1,#2,... จะเป็นค่าสูงสุดหรือต่ำสุดของตัวแปรที่ 1, 2 .. ตามลำดับ ค่าที่อยู่นอกช่วงที่กำหนด (มากกว่าค่าสูงสุด, น้อยกว่าค่าต่ำสุด) จะไม่นำมาใช้ในการทำงาน

ถ้าไม่ใช้ State นี้จะนำทุกค่าที่มีอยู่ในข้อมูลมาใช้ อาจกำหนดค่าสูงสุดและต่ำสุดของตัวแปรบางตัวได้ โดยใช้รูปแบบดังนี้

```
MAXimum  IS (n)  #1.
MAXimum  IS (n)  #2.
```

ใช้ n บ่งลำดับที่ของตัวแปรที่ต้องการกำหนด ตามลำดับที่ปรากฏใน NAME Statement

3.3

```
BLANK = { ZERO.
        { MISSING.
```

ใช้บอกรหัสที่ใช้ช่องว่าง หรือ Blank บนสื่อที่ใช้บันทึกข้อมูล ว่าเป็นรหัสที่แทนค่าศูนย์ (ZERO) หรือ ค่าไม่สมบูรณ์ (Missing) เพราะเครื่องคอมพิวเตอร์บางเครื่องไม่สามารถบอกความแตกต่างระหว่าง Blank และศูนย์ (ZERO) ได้ ถ้ารหัสที่เป็นช่องว่าง (BLANK) ใช้ในความหมายเหมือนกันตลอดโปรแกรม อาจใช้

```
BLANK  ARE  MISSING.
หรือ BLANK  ARE  ZERO.
```

3.4

```
MISSING = #1, #2, . . .
```

ใช้กำหนดรหัสหรือตัวเลข ใช้แทนค่าไม่สมบูรณ์ (missing value) ของตัวแปรแต่ละตัว โดยบ่งตามลำดับของตัวแปร เช่น

1 เป็นตัวเลขที่แทนค่าไม่สมบูรณ์ (missing value) ของตัวแปรที่ 1

2 เป็นตัวเลขที่แทนค่าไม่สมบูรณ์ (missing value) ของตัวแปรที่ 2 ตามลำดับ

3.5

$$\text{USE} = \begin{Bmatrix} v_1, v_2 \cdots \\ n_1 \text{ TO } n_2 \end{Bmatrix}$$

ใช้เลือกตัวแปรที่ต้องการใช้ในการวิเคราะห์ โดยบ่งชื่อตัวแปร หรือตัวแปรมีตัวกำกับ (Subscript) และ อาจใช้ตัวเลข (n_1, n_2) แสดงลำดับที่ของตัวแปร ประกอบกับคำเฉพาะ TO ก็ได้

3.6

$$\text{LABEL} = v_1, v_2.$$

ใช้เมื่อต้องการพิมพ์ชื่อหรือสัญญลักษณ์ของค่าสังเกตไว้ทางผลลัพธ์ โดยระบุชื่อตัวแปรที่ต้องอยู่ในแบบ A-type ได้ 2 ตัวแปร และตัวแปรแต่ละตัวต้องมีความยาวไม่เกิน 4 คอลัมน์ เช่น ตัวแปร ID ใช้เป็นชื่อและหมายเลขของพนักงาน (01SY, 02BA,...) ซึ่งเป็นแบบ A-type เมื่อต้องการให้พิมพ์ชื่อและหมายเลขทางผลลัพธ์ใช้ LABEL Statement

ตัวอย่างการใช้ Statement ประกอบ VARiable paragraph

```
/ VARiable    NAMES  ARE  ID,AGE,HEIGHT,WEIGHT,
                          BRTHPILL, CHOLSTRL, ALBUMIN,
                          CALCIUM, URICACID.
              MAXIMUM  IS  (6)  400.
              MINIMUM  IS  (6)  150.
              BLANK    ARE  MISSING.
              LABEL    IS  ID.
```

4. GROUP or CATEGORY paragraph

เป็น paragraph ที่จำเป็นต้องใช้สำหรับการทำงานบางอย่าง (โปรแกรมบางชุด) เท่านั้น ซึ่งจะใช้บ่งรหัสที่แทนค่าของตัวแปรชนิดไม่ต่อเนื่อง (discrete variable) ซึ่งอาจจะแบ่งเป็นกลุ่ม ๆ, กำหนดช่วงสำหรับตัวแปรชนิดต่อเนื่อง (Continuous variable), ให้ชื่อกลุ่มที่ใช้, รวมรหัสหรือช่วงไว้เป็นกลุ่ม

เราจะใช้ paragraph นี้ เมื่อต้องการแบ่งกลุ่ม แต่ตัวแปรที่ใช้ควรจะมีมากกว่า 10 ค่าขึ้นไปที่จะรวมเป็นกลุ่ม ควรจะใช้ paragraph นี้ เมื่อแบ่งกลุ่มตัวแปรชนิดต่อเนื่อง เช่น รายได้, ส่วนสูง, อายุ ฯลฯ รูปแบบ ของ GROUP or CATEGORY paragraph เป็นดังนี้

```
{ / GROUP      }
{   or         }   Statements
{ / CATEGORY   }
```

Statement บางชนิดที่ใช้ได้กับ GROUP paragraph มีดังนี้

4.1
```
CODE (#)  = #1, #2,...
```

#1, #2,... เป็นค่าหรือรหัสของตัวแปรลำดับที่ (#) เช่นตัวแปรตัวที่ 4 (ตามลำดับที่บันทึกไว้บนบัตรข้อมูล) ชื่อ SEX ใช้รหัส 1 แทน FEMALE รหัส 2 แทน MALE เขียน

```
/ GROUP        CODE (4)  =  1, 2
```

ใช้จัดกลุ่มค่าสังเกตตามค่าหรือรหัสของตัวแปรชนิดไม่ต่อเนื่องตามที่ระบุไว้ (#1, #2,...) ตัวแปรที่ 4 ของค่าสังเกตใด ถ้ามีค่าเป็น 1 จัดค่าสังเกตนั้นไว้ในกลุ่มที่ 1

ถ้ามีค่าเป็น 2 จัดค่าสังเกตกลุ่มที่ 2 ตามลำดับ

ถ้าตัวแปรนั้นมีค่าไม่เท่ากับที่ระบุไว้ (#1, #2,...) ก็จะไม่จัดเข้ากลุ่มใด และแต่ละค่าของตัวแปรที่จะจัดเป็นกลุ่มนั้นต้องมีค่าสังเกตจำนวน 10 ค่าสังเกตขึ้นไป

4.2

```
CUTPoint (#)  = #1, #2,...
```

ใช้จัดกลุ่มค่าสังเกตตามค่าของตัวแปรชนิดต่อเนื่อง โดยบ่งค่าสูงสุด (#1, #2,...) ของแต่ละช่วงสำหรับตัวแปรลำดับที่ (#) พิจารณาจากตัวอย่างต่อไปนี้

```
/ GROUP                CUTPoints (7)     ARE    20, 30, 40.
```

หมายถึงจะจัดกลุ่มค่าสังเกตของตัวแปรชนิดต่อเนื่องตัวที่ 7 , 4 กลุ่ม ดังนี้

กลุ่ม 1 ตัวแปรลำดับที่ 7 มีค่าน้อยกว่าหรือเท่ากับ 20 ลงมา

กลุ่ม 2 ตัวแปรลำดับที่ 7 ค่ามากกว่า 20 แต่น้อยกว่าหรือเท่ากับ 30

กลุ่ม 3 ตัวแปรลำดับที่ 7 ค่ามากกว่า 30 แต่น้อยกว่าหรือเท่ากับ 40

กลุ่ม 4 ตัวแปรลำดับที่ 7 ค่ามากกว่า 40 ขึ้นไป

จำนวนช่วงจะมีเท่ากับจำนวนค่าสูงสุดที่ระบุ(#1, #2, ...) บวกกับ 1 ตัวแปรหนึ่ง ๆ จะเลือกใช้ CODE หรือ CUTPoint Statement เพียงชนิดเดียวเท่านั้น

4.3

```
NAME (#)  = C1, C2, ...
```

ใช้บอกชื่อของกลุ่มที่แบ่งโดย CODE หรือ CUTPoint Statementโดยต้องใช้ควบคู่ไปกับ Statement ใด Statement หนึ่ง

(#) ใช้บอกลำดับที่ตัวแปร โดยใช้ตัวเลขเดียวกันกับที่ใช้ใน (#) ของ CODE or CUTPoint Statement และระบุชื่อของกลุ่ม (C_1, C_2,...) ตามกลุ่มที่แบ่งไว้ใน Statement ทั้ง 2 (C_1, C_2,...) ใช้อักษรได้ไม่เกิน 8 ตัว เช่น

```
/ GROUP                CODES(4)      ARE    1,  2.
                       NAMES(4)      ARE FEMALE, MALE.
                       CUTPOINTS(7)ARE   20, 30, 40.
                       NAMES(7)      ARE '20 ØRLESS', '21 TO 30',
                                         '31 TO 40', 'OVER 40'.
```

ใน Group paragraph อาจใช้ได้หลาย ๆ CODE หรือ CUTPOINTS Statement

4.4

RESET.

โดยปกติกลุ่มที่แบ่งหรือจัดเป็นพวกไว้โดย GROUP paragraph จะถูกเก็บไว้ในเครื่องคอมพิวเตอร์จนกว่าจะมี การจัดกลุ่มใหม่ แต่เราอาจจะใช้ลบกลุ่มที่จัดไว้ได้โดยใช้ RESET Statement.

การดัดแปลงข้อมูล (data transformation)

ในโปรแกรม BMDP เราสามารถเปลี่ยนแปลงตัวแปรดังนี้

- รวมตัวแปร 2 ตัว หรือมากกว่า 2 ตัว เป็นตัวแปรเดียว
- เปลี่ยนแปลงข้อมูล เพื่อสะดวกในการทำงาน เช่น บันทึกไว้ในรูปตัวแปร ชื่อ Time แต่การทำงานอาจใช้ในรูป 1/Time
- เปลี่ยนหรือแก้ไขค่าของตัวแปร

บางกรณีอาจจะเลือกใช้บางค่าสังเกต การเปลี่ยนแปลงตัวแปรหรือเลือกค่าสังเกต มี 2 วิธี คือ

(1) การใช้ Control language transformation

(2) การใช้ FØRTRAN transformation

(1) การแปลงข้อมูลโดยใช้ Control language transformation

เป็นการดัดแปลงข้อมูลโดยใช้ภาษาตามกฎเกณฑ์ของโปรแกรม BMDP มี paragraph ที่จำเป็นดังนี้

1. TRANSFORM paragraph

ใช้ในการสร้างตัวแปรใหม่ ๆ เปลี่ยนแปลงตัวแปร, และใช้ในการเลือกค่าสังเกตมีรูปแบบดังนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

I 16257315

```
/TRANsform                simple expressions.
```

simple expressions เป็น Statement ที่ใช้ในการกำหนดค่า, การคำนวณทางคณิตศาสตร์ การเปรียบเทียบค่าทางตรรกศาสตร์ โดยแบ่งเป็น 5 ชนิด

- assignment ใช้กำหนดค่าให้แก่ตัวแปร
- arithmetic ใช้ในการคำนวณโดยมีเครื่องหมาย (+,-,*,**,/,MOD)
- function evaluation ใช้คำนวณค่าฟังก์ชั่นต่าง ๆ เช่น (LØG, SQRT, EXP, ABS, SIN, COS, ATAN, INT, SIGN, NONB)
- logical operation (LE, LT, GE, GT, NE, EQ, AND, OR, IF)
- random no. generation ใช้กำหนดเลขสุ่มให้แก่ตัวแปร โดยการแจกแจงแบบยูนิฟอร์มและแบบปกติ

นอกจากนี้ยังมี Statement ที่ใช้กับ TRANsform paragraph โดยมีรูปแบบดังนี้

```
/ TRANsform          statements
```

Statements ที่จำเป็นต้องใช้มีดังต่อไปนี้

1.1 USE ใช้เลือกค่าสังเกตที่จะนำมาใช้งาน ซึ่งจะมีความหมายต่างจาก USE statement ใน VARIABLE paragraph(ใน VARIABLE paragraph ใช้เลือกตัวแปรที่จะนำมาวิเคราะห์) มีรูปแบบ คือ

USE = { #. / Simple expression }

ใช้ # โดยบ่งตัวเลข 3 ตัว คือ 1, 0, -1

ใช้ Simple expression ซึ่งจะให้ผล 3 อย่างคือ เป็นบวกหรือเป็นจริง
เป็นศูนย์หรือเป็นเท็จ
เป็นลบ

ซึ่งมีความหมายดังนี้

ถ้าใช้ USE = 1(ผลเป็นบวกหรือเป็นจริง) หมายถึง ใช้ค่าสังเกตนั้นในการทำงาน และลอกไว้ในแฟ้มข้อมูล BMDP

ถ้าใช้ USE = 0(ผลเป็นศูนย์หรือเป็นเท็จ) หมายถึง ไม่ใช้ค่าสังเกตในการทำงาน ลอกไว้ในแฟ้มข้อมูล BMDP

ถ้าใช้ USE =-1(ผลเป็นลบ) หมายถึง ไม่ใช้ค่าสังเกตนั้นในการวิเคราะห์ และไม่ลอกไว้ในแฟ้มข้อมูล BMDP

เช่น

```
/ TRANsform        EDUCA = STUDY GT 4.
                   EDUCB = STUDY LT 7.
                   USE   = EDUCA AND EDUCB.
```

ตัวแปรชื่อ EDUCA เป็นจริงเมื่อมีค่ามากกว่าหรือเท่ากับ 4

ตัวแปรชื่อ EDUCB เป็นจริงเมื่อมีค่าน้อยกว่าหรือเท่ากับ 7

ค่าสังเกตที่จะใช้ในการทำงาน ก็ต่อเมื่อค่าของตัวแปรใน STUDY มีค่ามากกว่า 4 และน้อยกว่า 7 หรือ ค่าสังเกตใดที่ตัวแปร STUDY มีค่าต่างจากนี้จะตัดทิ้ง

```
/ TRANsform        USE = x(1) LE x(2).
```

ค่าสังเกตที่จะใช้เมื่อตัวแปรที่ 1 มีค่าน้อยกว่าหรือเท่ากับตัวแปรที่ 2

1.2 OMIT or DELETE ใช้ยี้ค่าสังเกตที่ไม่ต้องการรวมในการทำงาน มีรูปแบบ คือ

```
{OMIT  }
{ or   }  = #1, #2,...
{DELETE}
```

บ่งลำดับที่ของค่าสังเกต (#1, #2,...) ที่ไม่ต้องการใช้ ถ้าบ่งเพียงตัวเดียว ต้องตามด้วย เครื่องหมาย (.) 2 ครั้ง เช่น OMIT = 35 ..

การใช้ OMIT มีความหมายเหมือนกับใช้ USE =0 ของแต่ละค่าสังเกต

การใช้ DELETE มีความหมายเหมือนกับใช้ USE = -1ของแต่ละค่าสังเกต

ถ้าต้องการตัดค่าสังเกตที่ไม่ต้องการออกหลายค่าควรใช้ USE Statement จะดีกว่า

2. VARIABLE paragraph

เมื่อมีการเปลี่ยนแปลงข้อมูลโดยการทำ data transformation อาจจำเป็นที่จะต้องใช้ Statement ต่อไปนี้ ใน VARiable paragraph

2.1 ADD ใช้เมื่อมีการสร้างตัวแปรใหม่จากตัวแปรเดิมที่มีอยู่แล้ว โดยต้องเพิ่มชื่อตัวแปรไว้ใน NAME Statement ของ VARiable paragraph และใช้ ADD Statement ตามหลัง NAME โดยมีรูปแบบดังนี้

```
ADD = #.
```

บ่งจำนวน (#) ตัวแปรที่เพิ่มจากเดิม (อาจสร้างจาก Control language)หรือโดย FØRTRAN transformation

จำนวนตัวแปรที่จะใช้ในการทำงานทั้งหมด เท่ากับจำนวนตัวแปรใน VARIABLE Statement ของ INPUT paragraph รวมกับจำนวนตัวแปรที่เพิ่มโดย ADD statement ของ VARiable paragraph จำนวนตัวแปรที่บ่ง (#) อาจจะใช้เครื่องหมาย (+,-) ได้ โดยมีความหมายดังนี้

ถ้า ADD เป็น (+) หรือไม่ใส่เครื่องหมาย

ตัวแปรใหม่ที่เพิ่มไปจะใช้ในการทำงานด้วย

ถ้า ADD เป็น (-) ตัวแปรที่เพิ่มไปทีหลังจะถูกตัดทิ้งไม่ใช้ในการคำนวณ

2.2 BEFORE or AFTER

เมื่อมีการเปลี่ยนแปลงข้อมูล และมีการเปลี่ยนรหัส (recode) ของค่าที่ไม่สมบูรณ์ (missing value) หรือ ค่าที่อยู่นอกช่องที่กำหนดโดย (MAX, MIN Statement) คือ XMIS ใน TRANsform paragraph ผู้ใช้อาจเลือกให้ข้อมูลที่อ่านเข้าไปใช้รหัสใหม่ดังกล่าวก่อนหรือหลังจากการแปลงข้อมูล (ทั้งจาก Control language หรือจาก FORTRAN transform) ได้โดยใช้รูปแบบดังนี้

```
BEFore.
  or
AFTer.
```

ใช้ BEFORE. เมื่อต้องการให้ข้อมูลที่อ่านเข้าไปใช้รหัสใหม่ก่อนมีการดัดแปลงข้อมูล

ใช้ AFTER. เมื่อต้องการใช้รหัสกับข้อมูลที่ดัดแปลงแล้ว

(2) การแปลงข้อมูลโดยใช้ FØRTRAN transformation

เป็นการแปลงข้อมูลโดยใช้พื้นฐานและกฎเกณฑ์ ตามภาษาฟอร์แทรน ควรจะใช้เมื่อมีข้อมูลเป็นจำนวนมากหรือหลาย ๆตัวแปร เมื่อใช้ภาษาฟอร์แทรนในการแปลงข้อมูล เราจะเรียกว่า FØRTRAN Statement ซึ่งสามารถจะใช้คำต่อไปนี้ เช่น READ, WRITE, DATA, DIMENSION, FUNCTION, SUBROUTINE. ซึ่งจะไม่กล่าวถึงในที่นี้

สรุปบัตรคำสั่งประเภท BMDP control cards โดยทั่ว ๆไป

paragraph name { commands or assign } Statement	จุดประสงค์หรือความหมายที่ใช้
/ PROBlem	ต้องใช้ทุกครั้ง
TITLE = 'c'.	ไม่จำเป็นต้องใช้; ใช้เพื่อบอกชื่องานที่จะทำไม่เกิน 160 ตัวอักษร
/ INPut	ต้องใช้ทุกครั้ง
VARiable = #.	ต้องใช้ทุกครั้ง; ใช้บอกจำนวนตัวแปรที่จะอ่านเข้าไปเป็น input.
FØRMat = 'c'.	ส่วนใหญ่ต้องใช้; ใช้บอกรูปแบบของข้อมูลที่จะอ่านไม่เกิน 800 ตัวอักษร
CASE = #.	ใช้บางครั้ง ; ใช้บอกจำนวน case
UNIT = #.	ใช้เฉพาะข้อมูลไม่ได้บันทึกบนบัตร ห้ามใช้เลข 1, 2, 6 บอกหน่วยที่เก็บแฟ้มข้อมูล
REWIND.	ใช้เฉพาะข้อมูลไม่ได้บันทึกบนบัตร
CODE = C.	ใช้ระบุชื่อรหัสของแฟ้มข้อมูลที่ใส่เข้าไปเมื่อข้อมูลอยู่ในแฟ้มข้อมูล
CONTENT = C.	ใช้ระบุข้อมูล, เมทริกซ์ ที่เก็บไว้ในแฟ้มมาใช้
LABEL = 'C'.	ต้องใช้บอกรายละเอียดของแฟ้มข้อมูลย่อย ถ้ามีหลายแฟ้มข้อมูล

/ VARiable	ถ้าข้อมูลมาจากแฟ้มข้อมูล BMDP ก็ไม่ต้องใช้
NAME = c_1, c_2,...	ใช้เมื่อต้องการบอกชื่อตัวแปรหรือใช้ตามตัวแปรลำดับ
MAXimum = #1, #2,...	ใช้กำหนดค่าสูงสุดที่ต้องการใช้ทำงานตามลำดับตัวแปร
MINimum = #1, #2,...	ใช้กำหนดค่าต่ำสุดที่ต้องการใช้ทำงานตามลำดับตัวแปร
BLANK = ZERO	ใช้บอกความหมายของช่องว่าง (BLANK) ว่าใช้แทน
MISSING.	ความหมายของศูนย์ (ZERO) หรือค่าไม่สมบูรณ์ (missing)
MISSing = #1, #2,...	ใช้กำหนดรหัสหรือค่าที่เป็นตัวเลขแทนค่าไม่สมบูรณ์
USE = V1, V2,...	ใช้กำหนดตัวแปรที่จะเลือกใช้ในการทำงาน
LABEL = V1, V2.	ใช้พิมพ์ชื่อตัวแปรที่อยู่ในรูป A-type ได้ 2 ตัวแปร
ADD = #. {0}	ใช้บอกจำนวนตัวแปรใหม่ที่เพิ่มขึ้นโดยการดัดแปลงข้อมูล
BEFore.	ใช้ตรวจสอบข้อมูลว่าเป็นค่าสมบูรณ์หรือไม่ก่อนและ
AFTer.	หลังการแปลงข้อมูล
/ GROUP or CATEGary	ใช้เมื่อต้องการจัดกลุ่มค่าลังเกต
CODE (#) = #1, #2, ...	ใช้กำหนดค่าหรือรหัสของตัวแปรแบบไม่ต่อเนื่อง เพื่อจะใช้ค่าหรือรหัสนั้นเป็นตัวจัดกลุ่ม
CUTPoint (#) = #1, #2,...	ใช้กำหนดค่าสูงสุดของแต่ละกลุ่มสำหรับค่าของตัวแปรชนิดต่อเนื่อง
NAME (#) = C1, C2,...	ใช้กำหนดชื่อของกลุ่มที่จัดไว้ตาม CODE หรือ CUTPoint
RESET. \	ใช้รวมค่าลังเกตที่จัดไว้เป็นกลุ่มตาม CODE หรือ CUTPoint

/ TRANsform	ใช้เมื่อต้องการดัดแปลงข้อมูล
simple expression	ใช้กำหนดค่า, คำนวณ, เปรียบเทียบค่า
USE = #.	ใช้เลือกค่าสังเกตที่จะใช้ในการทำงาน
simple expression	
OMIT = #1, #2, ...	ใช้ระบุลำดับที่ค่าสังเกตที่ไม่ต้องการใช้ (เหมือนใช้ USE = 0)
DELETE = #1, #2,...	ใช้ระบุลำดับที่ค่าสังเกตที่จะตัดทิ้ง (เหมือนใช้ USE = -1)
/ SAVE	ใช้เมื่อต้องการจะเก็บข้อมูลหรือผลที่ได้ไว้ในแฟ้มข้อมูล
CODE = C.	ใช้ระบุรหัสหรือชื่อเพื่อเป็นตัวชี้แฟ้มข้อมูลที่จะเก็บไว้
LABEL = 'C'.	ใช้บอกรายละเอียดของแฟ้มข้อมูลใช้ได้ไม่เกิน 40 ตัวอักษร
UNIT = #	ใช้บอกตำแหน่งหรือหน่วยที่เก็บข้อมูลบนเทปหรือจานแม่เหล็ก
NEW.	
/ END	

โปรแกรมแต่ละชุดใน BMDP อาจจะมี paragraph name เพิ่มเติมได้ตามกฎเกณฑ์ของโปรแกรมชุดนั้น ๆ

สัญญลักษณ์ต่าง ๆที่ใช้มีความหมายต่าง ๆ ดังนี้

ใช้ตัวเลขแสดงจำนวน

V ชื่อตัวแปรหรือชื่อตัวแปรที่มีตัวกำกับ

C ชื่อตัวแปรที่ใช้ได้ไม่เกิน 8 ตัวอักษร

'C' ชื่อที่ใช้บ่งในเครื่องหมาย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรม SAS

•STATISTICAL ANALYSIS SYSTEMS•

โปรแกรม SAS หรือโปรแกรมสำเร็จรูปทางสถิติวิเคราะห์ เป็นโปรแกรมสำเร็จรูปขนาดใหญ่ สำหรับทำ Data transformations และวิเคราะห์สถิติทุกระดับ เช่นเดียวกับโปรแกรม BMDP โปรแกรม SAS มีอยู่ที่ศูนย์ประมวลผลด้วยเครื่องจักร สำนักงานสถิติแห่งชาติ และศูนย์คอมพิวเตอร์แห่งภูมิภาคสถาบันเทคโนโลยีแห่งเอเชีย (เอไอที) โปรแกรมนี้มีจุดประสงค์ในการวิเคราะห์ทางสถิติโดยตรง

ความสามารถทั่ว ๆไปของโปรแกรม SAS

1. การทำ Data transformations
2. สามารถเลือกเฉพาะบางส่วนของค่าสังเกตมาทำการวิเคราะห์และสามารถกำจัดข้อมูล หรือค่าสังเกตที่ไม่ต้องการได้
3. สามารถนับข้อมูลโดยเราไม่จำเป็นต้องทราบจำนวนค่าสังเกต (observation)
4. สามารถตรวจสอบข้อมูล และให้พิมพ์เฉพาะข้อมูลที่ผิดหรือให้พิมพ์ ค่าสังเกตเฉพาะที่มีข้อมูลผิดพลาดออกมาได้ ซึ่งจะสะดวกในการแก้ไข
5. สามารถให้ทำงานซ้ำ ๆกัน กับข้อมูลหลายชุดได้ โดยไม่ต้องเขียนคำสั่งที่ให้ทำงานนั้นใหม่

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ความสามารถของโปรแกรม SAS ในการทำงานทางสถิติ

1. การคำนวณหาค่าสถิติเบื้องต้น (Descriptive statistics) ประกอบด้วย ค่าเฉลี่ย (mean), ส่วนเบี่ยงเบนมาตรฐาน (standard deviation), ความแปรปรวน(variance) ค่าสูงสุด-ต่ำสุด, สัมประสิทธิความแปรผัน (coefficient of variation), ผลรวม (SUM) ผลบวกกำลังสองของส่วนเบี่ยงเบนจากค่าเฉลี่ย (corrected sum of squares)

2. การแจกแจงความถี่และการวัดความสัมพันธ์ระหว่างตัวแปร ประกอบด้วย การแจกแจงความถี่, ตารางแจกแจงความถี่ และค่าสถิติที่ใช้วัดความสัมพันธ์ดังนี้ chi-square X^2, Kendall Tau-b τ_b, Spearman rank correlation r_s, Pearson correlation (r)

3. การเขียนกราฟและแผนภูมิ ประกอบด้วย scatter plot (Bivariate) Contour plot, กราฟประกอบการวิเคราะห์สถิติต่าง ๆ

4. การวิเคราะห์ความแปรปรวนและความแปรปรวนร่วม (Analysis of variance and covariance) สามารถทำการวิเคราะห์ข้อมูลที่มาจากการแจกแจงแบบทางเดียว (One-way ANOVA), แบบสองทาง (TWO-ways ANOVA) และแบบหลายทาง (Multi-ways ANOVA) และสามารถทำการวิเคราะห์ข้อมูลที่ได้มาจากแผนการทดลองต่าง ๆ เช่น Ramdomized Complete block, Latin Square, Factorial, Lattice.

5. การวิเคราะห์การถดถอย (Regression analysis) สามารถทำการวิเคราะห์สมการการถดถอยในรูปแบบต่าง ๆ ดังนี้
Simple Reg', Multiple Reg', Stepwise Reg' Stepwise Reg' มีวิธีการคัดเลือกตัวแปรถึง 5 วิธี คือ Forward selection, Backward elemination, Stepwise R^2 regression Maximum R^2 และ Minimum R^2 improvement การคำนวณค่าสถิติ นอกจากนี้ยังสามารถดัดแปลงข้อมูลหรือตัวแปรเพื่อใช้วิเคราะห์สมการถดถอยในรูปแบบอื่น ๆ ได้ เช่น Polynomial Reg', Non-linear Reg'.

6. การวิเคราะห์ทางด้านตัวแปรพหุคูณ (Multivariate analysis) ประกอบไปด้วยการวิเคราะห์ต่อไปนี้

Canonical correlation, Cluster analysis, Discriminat analysis, Factor analysis Multivariate ANOVA, Guttman Scaling.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การเตรียมชุดบัตรคำสั่ง (Control cards) ของโปรแกรม SAS. แบ่งเป็น 2 ประเภทใหญ่ ๆ คือ

1. Job Control Cards

ได้แก่ บัตรที่มีเครื่องหมาย / (slash) อยู่ในคอลัมน์ที่ 1 เป็นชุดของบัตรที่กำหนดโดยศูนย์คอมพิวเตอร์แต่ละแห่ง ซึ่งบัตรชุดนี้จะต้องปะหน้าและปะหลังของโปรแกรมทุกครั้ง บัตรชุดนี้จะกำหนดรหัสของผู้มีสิทธิใช้บริการของศูนย์คอมพิวเตอร์นั้น ๆ, บอกสื่อที่ข้อมูลว่าอยู่ในบัตร, เทปหรืออื่น ๆ และบอกให้เครื่องคอมพิวเตอร์ทราบว่าจะเรียกใช้โปรแกรม SAS Job Control card โดยปกติจะใช้เหมือนกันทุกครั้ง

2. SAS Control Cards

เป็นชุดของบัตรกำหนดโดยผู้สร้างโปรแกรม SAS โดยเฉพาะซึ่งบัตรชุดนี้จะแบ่งเป็น 2 ประเภทใหญ่ๆ ตามขั้นตอนการทำงาน คือ

2.1 Data Step ขั้นตอนนี้จะต้องใช้บัตรคำสั่งที่จะอธิบายรายละเอียดของข้อมูล เช่น ชื่อ ข้อมูล, สื่อที่ใช้บันทึกข้อมูล, ตำแหน่งที่อยู่ของข้อมูล, ลักษณะข้อมูลที่บันทึก ฯลฯ ซึ่งในขั้นตอนนี้อาจต้องใช้บัตรคำสั่ง เพื่อใช้ในการเปลี่ยนแปลงข้อมูล (Data transformation) เช่น จัดกลุ่มตัวแปรใหม่ กำหนดค่าตัวแปรใหม่ ฯลฯ

2.2 Proc Step ขั้นตอนนี้จะต้องใช้บัตรคำสั่ง ที่จะอธิบายรายละเอียดให้โปรแกรม SAS ทราบว่า เราต้องการอะไรเป็นการสั่งหรือเรียกโปรแกรมย่อย ที่จำเป็นต้องใช้ให้การทำงานตามที่ระบุไว้ในบัตร ซึ่งบัตรคำสั่งแต่ละคำสั่งอาจจะมีบัตรมาประกอบช่วยขยายคำสั่งที่ไปทำงาน (Procedure Information Statement) เพื่อให้ได้ผลตามที่ผู้วิจัยต้องการ

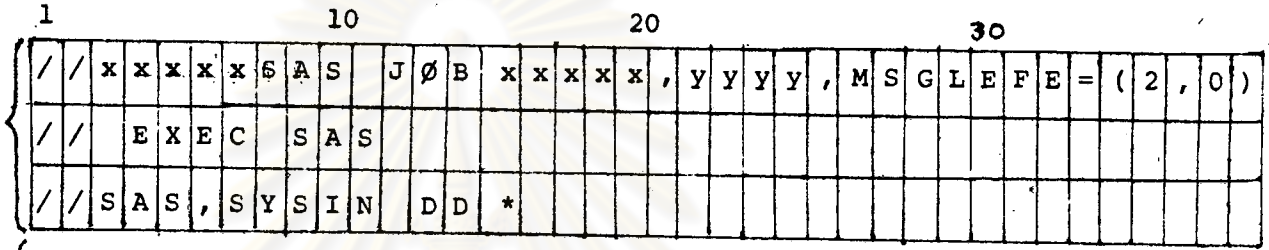
นอกจากนี้อาจมีบัตรคำสั่งสำหรับทำ data transformation หรือต้องการผลลัพธ์เพิ่มเติม เช่น พิมพ์หัวกระดาษ, พิมพ์ความหมายย่องรหัสที่ใช้ ซึ่งบัตรประเภทนี้ อาจจะแทรกอยู่ส่วนใดของโปรแกรมก็ได้

ตัวอย่างชุด Control cards สำหรับโปรแกรม SAS

ของ ศูนย์ประมวลผลด้วยเครื่องจักร สำนักงานสถิติแห่งชาติ

กรณีข้อมูลอยู่ในบัตร

1. Job Control Cards



2.1 SAS control cards — Data Step

ปัตรข้อมูล

2.2 SAS control cards — Proc Step

1. Job control cards

ความหมายของคำและสัญญลักษณ์ที่ใช้บางชนิด

xxxxx หมายถึง ให้ระบุรหัสของผู้มีสิทธิ์ใช้บริการที่กำหนดโดยศูนย์ฯ

yyyyy หมายถึง ให้ระบุผู้ใช้บริการของศูนย์ (อาจไม่ระบุก็ได้)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ข้อกำหนดเกี่ยวกับการเขียน Control Cards ของโปรแกรม SAS

การเขียนคำสั่งในบัตรชุด SAS Control Cards มีข้อกำหนดโดยทั่ว ๆ ไปดังนี้

1. เขียนคำสั่งได้ตั้งแต่ คอลัมน์ที่ 1 ถึง คอลัมน์ที่ 72
2. ถ้าคำสั่งใด ๆ ต้องใช้บัตรมากกว่า 1 ใบ ให้ต่อบัตรได้โดยจะเริ่มที่คอลัมน์ใดก็ได้ ซึ่งเราอาจเรียกสั้น ๆว่า " Free form"
3. เมื่อจบคำหนึ่ง ๆ ควรเว้นวรรค คือเว้นไว้อย่างน้อย 1 คอลัมน์ เพื่อแสดงการแยกแต่ละคำ และคำหนึ่งใช้อักษรหรือสัญญลักษณ์ไม่เกิน 8 ตัว
4. เมื่อจบคำสั่งหนึ่ง ๆ ต้องใส่เครื่องหมาย Semicolon (;) ทุกครั้ง เพื่อแสดงว่าจบคำสั่ง สำหรับรายละเอียดอื่น ๆนั้น แล้วแต่โครงสร้างของแต่ละคำสั่ง

บัตรคำสั่งประเภท SAS Control Cards

1. คำสั่งของโปรแกรม SAS ในขั้น Data Step

เป็นคำสั่งที่จะอธิบายรายละเอียดของข้อมูลมีหลายประเภทในที่นี้จะกล่าวเฉพาะที่จำเป็นต้องใช้ เรียงตามลำดับดังนี้

(1) DATA Statement

เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้ง ซึ่งจะบอกถึงการเริ่มต้นของขั้น Data Step มีรูปแบบดังนี้

```
DATA [data set name] ;
```

data set name เป็นชื่อของชุดข้อมูลที่ผู้ใช้กำหนดขึ้นเอง โดยใช้บอกชื่อชุดของข้อมูลที่จะให้ Conputerทำงานซึ่งอาจจะไม่จำเป็นต้องตั้งชื่อชุดข้อมูลนี้ก็ได้ มักตั้งชื่อเมื่อข้อมูลนั้นถูกลงสื่ออย่างอื่นที่ไม่ใช่บัตร เช่นเทปแม่เหล็ก เพื่อประโยชน์ที่จะเรียกข้อมูลชุดนี้ภายหลัง ตัวอย่าง เช่น ข้อมูลที่จะวิเคราะห์ ชื่อ EXAMPLE1 เราเขียน

```
---------------
DATA  EXAMPLE1  ;
---------------
```

ซึ่งอาจจะบอกรายละเอียดเพิ่มขึ้นได้อีกโดยเพิ่ม Option (บ่งไว้ในคู่มือการใช้ SAS) ไว้ในวงเล็บเล็ก "( )" ต่อท้ายชื่อชุดข้อมูลและจบด้วยเครื่องหมาย " ; "

(2) INPUT Statement

เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้งมีรูปแบบดังนี้ คือ

```
INPUT [DDNAME = .....]  [NREC = ......]  discriptions of data elements;
```

ใช้บอกรายชื่อตัวแปร (Variable) โดยเรียงตามลำดับของรายชื่อตัวแปรที่มีอยู่ในข้อมูล ชื่อตัวแปรนั้น ผู้ใช้กำหนดขึ้นมาเองโดยใช้เป็นตัวพิมพ์ใหญ่ไม่เกิน 8 ตัวอักษร นอกจากนั้น บัตร INPUT ยังสามารถบอกลักษณะของตัวแปรได้ด้วย

DDNAME ใช้เมื่อข้อมูลอยู่ในเทป, disk, drum เท่านั้น ซึ่ง DDNAME ต้องเหมือน DDNAME ใน JCL Statement (จะไม่กล่าวถึงในที่นี้)

NREC ใช้บอกจำนวน record หรือจำนวนบัตรที่ใช้ต่อ 1 ค่าสังเกต (observation) จะใช้ Option นี้ก็ต่อเมื่อ ตัวแปรตัวสุดท้ายที่บ่งใน INPUT Statement ไม่ได้อยู่ในบัตรใบสุดท้าย (record สุดท้าย) ของค่าสังเกต

ความหมายของ discriptions of a data elements

จะบอกรายละเอียดเกี่ยวกับตัวแปร เช่น ชื่อ, ชนิดของตัวแปร, ตำแหน่งของตัวแปร มีรูปแบบ ดังนี้

```
Variable name  [type]  [record k]  [position m-position n]
               [decimal  point]  ;
```

variable name	เป็นชื่อของตัวแปรที่เราตั้งขึ้น
type	ใช้เมื่อตัวแปรที่มีใน Variable name มีค่าเป็นอย่างอื่นที่ไม่ใช่ตัวเลข (numeric) เช่น เป็นตัวอักษร (alphabetic) หรือเป็นตัวอักษรปนตัวเลข (alpha numeric) ต้องใช้ type โดยบ่ง $ (dollar Sign) ตามหลังชื่อตัวแปรนั้น
record k	ใช้เมื่อมีหลายบัตร หรือ หลาย record ต่อ 1 ค่าสังเกต โดยบ่งตัวเลข(k) ที่แสดงว่าตัวแปรนั้นอยู่บนบัตรในที่ k
position m-position n	ใช้บ่งว่าตัวแปรนั้น อยู่ระหว่างคอลัมน์ที่เท่าไร (m) ถึงคอลัมน์เท่าไร (n) ของบัตรใช้ตามหลังชื่อตัวแปร
decimal-point	ใช้บ่งว่าตัวแปรนั้น ๆ มีทศนิยมกี่ตำแหน่ง

```
INPUT  NAME $ 1-5 A 7-10 2 B 12-15 ;
```

เป็นการบอกว่าข้อมูลชุดนี้มีตัวแปร 3 ตัว คือ

ตัวแปร ชื่อ NAME เป็นตัวแปรประเภทที่ไม่ใช่ตัวเลขล้วน อยู่ในคอลัมน์ 1-5

ตัวแปร ชื่อ A เป็นตัวแปรที่มีค่าเป็นตัวเลข อยู่ที่คอลัมน์ 7-10 เป็นทศนิยม 2 ตำแหน่ง

ตัวแปร ชื่อ B เป็นตัวแปรที่มีค่าเป็นตัวเลขอยู่ที่คอลัมน์ 12-15

(3) CARDS Statement

ใช้บัตรนี้ทุกครั้ง เมื่อข้อมูลบันทึกบนบัตร ถ้าบันทึกบนสื่ออื่น ๆ ไม่ต้องใช้บัตรนี้ บัตรนี้อยู่ในขั้นสุดท้ายของ Data Step ถัดจากบัตรนี้ไปจะเป็นบัตรข้อมูล บัตรคำสั่งนี้มี รูปแบบ คือ

```
CARDS;
```

บัตรคำสั่งนี้จะบอกว่าสื่อที่บันทึกข้อมูล คือ บัตร

บัตร Control cards ในขั้น Data Step ที่สำคัญและจำเป็นต้องใช้ทุกครั้งมีอยู่ 3 ชนิด ที่กล่าวมาแล้ว แต่ถ้ามีการดัดแปลงข้อมูลที่อ่านเข้ามาใน INPUT อาจจะเพิ่มบัตรคำสั่งที่เรียกว่า "Program Statement" ซึ่งมักจะแทรกบัตรประเภทนี้ไว้หลัง INPUT Statement และตามด้วย CARDS Statement.

(4) PROGRAM Statement

อาจแบ่งเป็นประเภทต่าง ๆ ดังนี้ ในที่นี้จะกล่าวเฉพาะที่สำคัญเท่านั้น

4.1 Assigment Statement

เป็นบัตรคำสั่งที่ใช้สร้างตัวแปรใหม่ และเปลี่ยนแปลงบางส่วนของตัวแปรที่มีอยู่แล้ว ในบัตร INPUT มีแบบที่ใช้ คือ

```
variable name = expression;
```

ความหมายของแบบที่ใช้ คือ

สร้างตัวแปรใหม่ จากผลที่ได้จากนิพจน์ทางคณิตศาสตร์ (expression) ชื่อ ตัวแปร (variable name) นั้น อาจเป็นชื่อที่ตั้งขึ้นใหม่จากรายชื่อใน INPUT หรือเป็นชื่อที่ตั้งขึ้นมาแล้วใน INPUT ถ้าเป็นชื่อตัวแปรที่ตั้งขึ้นมาใหม่ ก็จะมีค่าเท่ากับการทำงานทางฟิสิคณิต ถ้าเป็นชื่อตัวแปรที่ซ้ำกับชื่อใน INPUT ค่าเก่าของตัวแปรนี้จะถูกแทนที่ด้วยค่าใหม่ที่คำนวณได้ ภายใน expression อาจประกอบไปด้วย

<u>ชื่อตัวแปร</u> (Variable name) ซึ่งอาจใช้ชื่อใหม่หรือชื่อที่เคยใช้มาแล้วใน INPUT ก็ได้ ถ้าเป็นชื่อใหม่ จะถือว่าตัวตัวแปรนั้นมีค่าเป็นศูนย์ <u>expression operator</u> ที่ใช้ประกอบไปด้วย

ก. <u>arithmetic Operator</u> ได้แก่ +, -, *, /, **

ข. <u>comparison Operator</u> ได้แก่ =, > , <, ≠ ,

ค. <u>logical Operator</u> ได้แก่ AND, OR, NOT,

SAS function เป็นคำเฉพาะที่มีความหมายในตัวเองว่าจะสั่งให้ทำอะไร ซึ่งประกอบไปด้วย ABS, ARCOS, ARSIN, ARTAN, CELL, COS, RRP, EXP, FLØØR, GAMMA, INT, LØG, LØG10, MAX, MIN, MØD, SIGN, SIN, SQRT, TANH

4.2 IF Statement

เป็นบัตรคำสั่งที่ใช้เลือกตัวแปรที่มีค่าตามต้องการ เพื่อมาทำงานตามคำสั่งต่อมา มักจะใช้ประกอบกับ THEN Statement มีแบบที่ใช้ คือ

```
IF expression THEN Statement;
```

ความหมาย

เมื่อมีค่าเป็นจริงตาม expression แล้วจะไปทำงานตามคำสั่งใน THEN Statement ถ้าไม่เป็นจริง จะไปทำตามคำสั่งที่ตามมาต่อจาก IF Statement (ถ้ามี)

4.3 DELETE Statement

เป็นบัตรคำสั่งที่ใช้ เพื่อตัดค่าสังเกตออกไป ซึ่งมักจะใช้ประกอบกับ IF Statement แบบที่ใช้ คือ `DELETE;`

เช่น

```
IF GRADE < 3 THEN DELETE;
```

ความหมาย คือ ถ้าค่าสังเกต มีค่าตัวแปรที่ชื่อ GRADE น้อยกว่า 3 แล้ว จะตัดค่าสังเกตนั้นทิ้งไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.4 OUTPUT Statement

เป็นบัตรคำสั่งที่ใช้เก็บค่าขณะนั้นของตัวแปรทุกตัวที่มีมาก่อน OUTPUT Statement มีแบบที่ใช้ดังนี้ คือ

```
OUTPUT data set name   ;
```

data set name จะบอกชื่อตัวแปรที่จะเก็บไว้ถ้าไม่ใช้จะถือว่าเก็บไว้หมดทุกตัวแปร ซึ่งจะบรรจุในค่าสังเกต

OUTPUT Statement อาจใช้ประกอบ IF Statement ได้

4.5 LIST Statement

ใช้เมื่อต้องการพิมพ์ค่าสังเกตที่อ่านเข้าไปตาม INPUT Statement พร้อมด้วยบ่งว่าเป็นค่าสังเกตที่เท่าไร ออกมาทางผลลัพธ์ ซึ่งเรามักใช้ประกอบกับ IF-THEN Statement แบบที่ใช้ คือ `LIST ;`

ซึ่งสามารถใช้ตรวจสอบค่าสังเกตได้ เช่น เรากำหนดให้ค่าที่เรียกว่า "missing-value" เป็นเครื่องหมาย (.) ของตัวแปร ชื่อ AGE เราสามารถให้พิมพ์ค่าที่เป็น missing value ได้และบอกด้วยว่าเป็นค่าสังเกตที่เท่าไร โดยเขียนคำสั่ง ดังนี้

```
DATA;
INPUT     AB 1-9  # 5 AGE 4-6;
IF  AGE  =  .  THEN  LIST;
CARDS;
```

4.6 PUT Statement

ใช้เมื่อต้องการพิมพ์ค่าตัวแปร และเครื่องหมายสัญญลักษณ์ที่คงที่ (charactor-constant) ออกมาทางผลลัพธ์ แบบที่ใช้ คือ

```
PUT  variable name and 'charactor constant';
```

Put Statement สามารถใช้ได้แบบเดียวกันกับ INPUT Statement เช่น

```
PUT  'VALUE  FOR'  YEAR 'IS'  AMT;
```

ความหมาย ในผลลัพธ์จะพิมพ์ Charactor constant คือ 'VALUE FOR' และ 'IS' พร้อมกับค่าตัวแปร YEAR และ AMT; ในแต่ละบรรทัดของกระดาษต่อเนื่องทุกบรรทัด เช่น

```
VALUE  FOR  1973  IS  50
VALUE  FOR  1974  IS  40
```

ถ้าใช้

```
PUT  YEAR  AMT;
```

จะพิมพ์ค่าของ ตัวแปร YEAR, AMT ให้ดูทุกบรรทัด

มักใช้ PUT : Statement กับ IF-THEN Statement

4.7 ERROR Statement

ให้ผลเช่นเดียวกับ PUT Statement แต่พิมพ์ลำดับของค่าสังเกต และค่าของตัวแปรทั้งหมดของค่าสังเกตออกมาด้วย มักใช้ร่วมกับ IF-THEN Statement แบบที่ใช้ คือ

```
ERROR  variable  name  and  'charactor constant';
```

4.8 GO TO Statement

เมื่อต้องการจะเปลี่ยนขั้นตอนการทำงาน สามารถใช้คำสั่ง GO TO ตามแบบดังนี้

```
GO TO label;
```

เป็นการให้ไปทำงานที่คำสั่งซึ่งบ่งไว้ใน label ไม่ว่าจะมีขั้นตอนเดิมอย่างไร GO TO Statement อาจใช้ประกอบกับ IF-THEN Statement label เป็นชื่อเพื่อบอกตำแหน่งของคำสั่งที่ถูกอ้างถึงใน GO TO Statement โดยชื่อใน label Statement ต้องตามด้วยเครื่องหมาย (:) เสมอ

```
DATA ;
 INPUT          X Y Z ;
 IF     x = .   THEN  GO TO MISS;
 PUT    x =  ;
 MISS  :  PUT   'x Missing FØR'   -N - = ;
 CARDS ;
 ...
```

4.9 DRØP Statement

ใช้เมื่อต้องการตัดตัวแปรทิ้งไปจากทุก ๆ ค่าสังเกต โดยระบุชื่อตัวแปรนั้น ตามรูปแบบดังนี้

```
DROP variable name ;
```

ชื่อตัวแปรที่ใช้ใน variable name ต้องเป็นชื่อที่กำหนดมาแล้วใน INPUT Statement ซึ่งอาจใช้หลายตัวแปรก็ได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

นอกจาก Statement ที่กล่าวมาแล้ว หลัง THEN ใน IF Statement อาจใช้ Statement ต่อไปนี้ได้

INPUT, INFILE, SUM, RETURN, LINK, FILE, LOSTCARD, SET, MERGE, UPDATE DO ซึ่งจะไม่กล่าวในที่นี้

2. คำสั่งของโปรแกรม SAS ในขั้น PROC Step แบ่งออกเป็น 2 ส่วน ดังนี้

(1) PROCEDURE Statement

เป็นบัตรคำสั่งที่ใช้สั่งให้โปรแกรม SAS ไปทำงานตามต้องการมีรูปแบบ ดังนี้

PROC procedure name [options] [parameter] [DATA= data Set name] ;

procedure name — เป็นชื่อเฉพาะเพื่อสั่งให้โปรแกรม SAS ไปทำงานตามชื่อที่ระบุไว้ เช่น ANOVA, REGR ชื่อนี้ต้องใช้ตามที่ระบุไว้ในคู่มือ

Options — เป็นสิ่งที่เราจะใช้หรือไม่ก็ได้ เป็นข้อความหรือคำที่จะไปขยายคำสั่งใน procedure name เพื่อให้ได้รายละเอียดเพิ่มเติมทางผลลัพธ์ (Output) ซึ่งจะมีให้เลือกใช้ตาม procedure นั้น ๆ อาจจะใช้หมดทุก Options หรือ เลือกใช้บางตัวก็ได้ แต่ละ procedure จะมี Options ไม่เหมือนกัน

parameter — มีความหมายคล้าย ๆ Options.

DATA = data Set name
- เป็นการบอกว่าให้ procedure นี้ไปทำงานกับชุดข้อมูลที่ตั้งชื่อไว้ใน DATA Statement
- ถ้าให้ไปทำงานกับชุดข้อมูลที่เพิ่งอ่านเข้าไปก็ไม่จำเป็นต้องใช้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(2) Procedure Information Statement

เป็นประโยคหรือข้อความที่ใช้ขยายคำสั่งใน PROCEDURE Statement และต้องใช้ควบอยู่ไปกับ PROCEDURE Statement บัตร PROCEDURE INFORMATION Statement นั้น ต้องใช้ตามที่ PROCEDURE Statement กำหนดให้ใช้ร่วมกันได้ ซึ่งบัตร PROCEDURE INFORMATION Statement หนึ่ง ๆ อาจใช้ได้กับหลาย PROCEDURE Statementแต่ความหมายอาจแตกต่างกันได้ เช่น CLASS Statement, VARIABLES Statement

เมื่อโปรแกรม SAS ทำงานขั้น PROC Step แล้ว อาจทำการดัดแปลงข้อมูลโดยใช้ "Program Statement " ได้ นอกจากคำสั่งที่แยกไว้เป็นประเภทใหญ่ ๆ 2 ประเภท คือ

คำสั่งในขั้น Data Step และ PROC Step แล้ว อาจมีคำสั่งต่อไปนี้ ซึ่งสามารถแทรกอยู่ในส่วน Data Step หรือ PROC Step ก็ได้

(3) TITLE Statement

ใช้เมื่อต้องการให้พิมพ์ข้อความหรือชื่อการทำงานทางสถิตินั้น ๆ ไว้ ส่วนบนสุดของทุก ๆ หน้า มีรูปแบบ คือ

```
TITLE [n]  title name ;
```

เราใช้ n บอกบรรทัดที่ต้องการให้พิมพ์โดย n ใช้ได้ตั้ง 1-10 ถ้าไม่ใช้ n จะพิมพ์ข้อความไว้บรรทัดแรกของทุก ๆ หน้า title name ใช้ตัวอักษรได้ไม่เกิน 132 ตัว อาจกำหนดข้อความที่ต้องการพิมพ์ไว้ในเครื่องหมาย quotation marks (' ') ได้ และสามารถใช้ TITLE Statement ได้หลายครั้ง คือ ถ้าไม่ใช้ n เครื่องจะพิมพ์ข้อความใน TITLE Statementไว้ทุกหน้าจนมี TITLE Statement อื่นอีกจะลบล้างข้อความใน TITLE Statement แรกออกไป

(4) MACRO

ในการทำงานซ้ำ ๆ กัน หรือข้อมูลหลายชุด เราสามารถใช้ MACRO เพื่อช่วยลดการเขียนคำสั่งซ้ำ ๆ กันได้ ซึ่งมีแบบที่ใช้ คือ

```
MACRO  macroname
       test
       :
       %
```

macroname ; เป็นเรื่องของ macro ที่ผู้ใช้ตั้งขึ้นเองเพื่อเรียกใช้ภายหลัง

text ; เป็นคำสั่งหรือกลุ่มคำสั่ง ซึ่งอาจเป็นคำสั่งในขั้น Data Step หรือ PROC Step ก็ได้

% ; เป็นตัวบอกว่า จบ macro นี้

ในการใช้ MACRO นี้ คำสั่งหรือกลุ่มคำสั่งใด เราต้องกำหนดชื่อ macroname นั้นก่อน โดยใช้บัตร 1 ใบ เขียนตามแบบที่กำหนด โดยบัตรใบนี้ เมื่อจบข้อความแล้วไม่ต้องใช้ (;) แล้วตามด้วยบัตรคำสั่งหรือกลุ่มคำสั่งที่ต้องการให้ทำงานซ้ำๆ เมื่อจบกลุ่มคำสั่งนี้แล้วใช้บัตรอีก 1 ใบ บันทึกเครื่องหมาย % เพื่อบอกว่าต้องการใช้แค่นี้ (แต่ถ้ากลุ่มคำสั่งที่ต้องการทำงานซ้ำ มีบัตรข้อมูลรวมอยู่ด้วย บัตรข้อมูลใบสุดท้ายจะเป็นตัวบอกว่าจบกลุ่มคำสั่งนั้น)

เมื่อจะเรียกใช้กลุ่มคำสั่งดังกล่าว เราก็เพียงแต่ใช้บัตร 1 ใบ บันทึกชื่อที่ตั้งไว้ใน macroname แล้วใช้บัตรนี้แทรกไว้หน้าส่วนที่ต้องการให้ทำงานซ้ำ

การใช้ MACRO เราสามารถเพิ่มเติมคำสั่งใด ๆ อีกได้ งานที่ทำซ้ำอาจจะเป็นการอ่านข้อมูลในขั้น (Data Step) หรือให้ทำงานจริง ๆในขั้น PROC ก็ได้

สรุปบัตรคำสั่งประเภท SAS Control cards เท่าที่จำเป็น

	ประเภทของบัตร	จุดประสงค์หรือความหมายที่ใช้
	DATA data set name ;	ต้องใช้ทุกครั้ง ซึ่งอาจระบุชื่อของข้อมูลที่ใช้
	INPUT discription of element ;	ต้องใช้ทุกครั้ง ใช้กำหนดชื่อ, ตำแหน่ง และชนิดของตัวแปร
	program statement ;	ใช้เมื่อต้องการดัดแปลงข้อมูล(transformation)
	CARDS ;	ต้องใช้เมื่อสื่อบันทึกข้อมูล คือ บัตรเจาะรู
1	PROC procedure name options ;	ใช้ระบุงานทางสถิติที่ต้องการ
	procedure information statement	ใช้เมื่อต้องการขยายคำสั่งใน PROC
2		ระบุงานที่ 2 ที่ต้องการ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โปรแกรม SPSS

•STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES•

โปรแกรม SPSS หรือโปรแกรมสำเร็จสำหรับการวิจัยทางสังคมศาสตร์ เป็นโปรแกรมสำเร็จรูปขนาดใหญ่ สำหรับทำ Data transformations และการวิเคราะห์สถิติทุกระดับเช่นเดียวกับโปรแกรม BMDP และ SAS โปรแกรม SPSS นั้น เป็นโปรแกรมที่นิยมกันแพร่หลาย และมีที่ศูนย์คอมพิวเตอร์ทั้ง 3 แห่ง คือ ที่

สถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

ศูนย์ประมวลผลด้วยเครื่องจักร สำนักงานสถิติแห่งชาติ

สถาบันเทคโนโลยีแห่งเอเชีย

โปรแกรมนี้มีจุดประสงค์ในการวิเคราะห์สถิติ สำหรับการวิจัยทางสังคมศาสตร์โดยตรง

ความสามารถทั่ว ๆไปของโปรแกรม SPSS

1. การทำ Data Transformations
2. สามารถเลือกเฉพาะบางส่วนของค่าสังเกตมาทำการวิเคราะห์ และสามารถกำจัดข้อมูล หรือตัวแปร และค่าสังเกตที่ไม่ต้องการได้
3. กรณีข้อมูลอยู่ในสื่ออย่างอื่นที่ไม่ใช่บัตร เช่น เทป (TAPE),DISK ฯลฯ ถ้าไม่ทราบจำนวนข้อสังเกต (observation) หรือจำนวน case โปรแกรมSPSS สามารถนับได้
4. สามารถตรวจสอบความถูกต้องของบัตร control card ว่าถูกต้องหรือไม่ ถ้าผิดก็จะบอกได้ว่าผิดอย่างไร

ความสามารถของโปรแกรม SPSS ในการทำงานทางสถิติ

1. การคำนวณหาค่าสถิติเบื้องต้น (Descriptive Statistics) ซึ่งจะประกอบไปด้วยค่าเฉลี่ย (mean), ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation), ความแปรปรวน (variance) ความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (Standard error of mean), ค่าสูงสุด-ต่ำสุด, พิสัย (range) ความเบ้ (skewness), ความสูง (kurtosis), ผลรวม (sum), ค่าสถิติ student's t ค่าสถิติ F

2. การแจกแจงความถี่และการวัดความสัมพันธ์ระหว่างตัวแปร ซึ่งประกอบไปด้วยตารางแจกแจงความถี่, ฮิสโตแกรมของความถี่, ความถี่สะสม การวัดและทดสอบความสัมพันธ์โดยค่าสถิติต่อไปนี้ Chi-square $(X)^2$, Phi , Cramer's V , contingency coefficient, Fisher's test Yates' corrected X^2, Kendall's tau-b Cb, Kendall's tau-c Ce, Gamma partial and zero, Gamma for three-n-variable tables, Lambda symmetric and asymmetric, uncertainty coefficient, Somer's D , Eta

3. การเขียนกราฟและแผนภูมิ ซึ่งประกอบไปด้วย scatter plot (Bivariate), ฮิสโตแกรม (histogram) กราฟประกอบการวิเคราะห์สถิติต่าง ๆ

4. การวิเคราะห์ความแปรปรวนและความแปรปรวนร่วม (Analysis of variance and covariance) สามารถทำการวิเคราะห์ข้อมูลที่มาจากการแจกแจงแบบทางเดียว (One-way Anova) แบบสองทาง (Two-ways Anova) และแบบหลายทาง (Multi-ways Anova) และทำการวิเคราะห์การจำแนกพหู (Multiple classification Analysis)

5. การวิเคราะห์การถดถอย (Regression Analysis) สามารถทำการวิเคราะห์สมการถดถอยในรูปแบบต่าง ๆ ดังนี้

Simple Reg' , Multiple Reg' , Stepwise Reg' Stepwise Reg' นั้นใช้วิธีการคัดเลือกตัวแปรโดยวิธี Forward selection นอกจากนี้สามารถดัดแปลงข้อมูลหรือตัวแปรเพื่อใช้วิเคราะห์สมการถดถอยในรูปแบบอื่น ๆ ได้ เช่น Polynomial Reg' , Non-linear Reg'

6. <u>การวิเคราะห์ทางด้านตัวแปรพหุคูณ</u> (Multivariate Analysis) ซึ่งประกอบไปด้วย การวิเคราะห์ต่อไปนี้คือ

Canonical correlation, Discriminant Analysis, Factor analysis.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การเตรียมชุดบัตรคำสั่ง (Control cards) ของโปรแกรม SPSS แบ่งเป็น 2 ประเภทใหญ่ ๆ คือ

1. Job Control Cards

ได้แก่ บัตรที่มีเครื่องหมายดอกจันทน์ * (asterisk) หรือเครื่องหมาย / (slash) อยู่ในคอลัมน์ที่ 1 ของบัตร เป็นชุดของบัตรที่กำหนดโดยศูนย์คอมพิวเตอร์แต่ละแห่ง ซึ่งบัตรชุดนี้จะต้องใช้ปะหน้า และปะหลังของโปรแกรมทุกครั้ง บัตรชุดนี้จะกำหนดระหัสของผู้มีสิทธิใช้บริการของศูนย์คอมพิวเตอร์นั้น, บอกสื่อที่ใช้บันทึกข้อมูล และบอกให้เครื่องคอมพิวเตอร์ทราบว่าจะเรียกใช้โปรแกรมสำเร็จรูป SPSS โดยปกติ Job Control Cards จะใช้เหมือนกันทุกครั้ง

2. SPSS Control Cards

เป็นชุดของบัตรกำหนดโดยผู้สร้าง SPSS โดยเฉพาะ ซึ่งแบ่งเป็น 2 ประเภทใหญ่ ๆ คือ

2.1 Data Definition Cards

เป็นชุดของบัตรที่จะอธิบายรายละเอียดของข้อมูล เช่น ข้อมูล, ตำแหน่งที่อยู่ของข้อมูล, ลักษณะข้อมูล ฯลฯ หรือ ให้เพิ่มเติมผลลัพธ์ที่จะได้ออกมา เช่น พิมพ์หัวกระดาษ, พิมพ์ความหมายของรหัสที่ใช้ ฯลฯ

2.2 Task Definition Cards

เป็นชุดของบัตรที่สั่งให้โปรแกรม SPSS ทำงานทางสถิติที่ต้องการ ซึ่งบัตรคำสั่งให้ทำงานทางสถิตินั้น (Procedure Cards) แต่ละบัตรอาจมีบัตรมาประกอบช่วยขยายคำสั่งหรือเพิ่มผลลัพธ์ เช่น OPTIONS และ STATISTICS Card

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตัวอย่างชุด Control Cards สำหรับโปรแกรม SPSS

ของ สถาบันบริการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

1. Job Control Cards
3 บัตร

```
1                    10                   20                   30
*  $$  JØB  JNM=xxxxxxxx,CLASS=N
//  JØB  xxxx                yyyyyy
//  EXEC  PRØC=$$SPSS79
```

2.1 SPSS control cards — Data Definition cards

2.2 SPSS control cards — Task Definition cards

```
READ INPUT DATA
```

บัตรข้อมูล

1. Job Control cards
3 บัตร

```
FINISH
/*
/&
*  $$  EØJ
```

บัตร Job Control 3 ใบแรกและ 3 ใบหลัง กำหนดโดยศูนย์คอมพิวเตอร์ แต่ละแห่งต้องใช้ปะหน้าและปะหลังของโปรแกรมทุกครั้ง

ความหมายของคำและสัญญลักษณ์ที่ใช้บางชนิด

xxxx xxxx 4 หลักแรกกำหนดโดยศูนย์บริการคอมพิวเตอร์เป็นรหัสแทนผู้มีสิทธิ์ใช้บริการ

4 หลักหลัง เป็นตัวเลขที่ผู้ใช้กำหนดขึ้นเองเพื่อระบุลำดับของงาน (JØB)

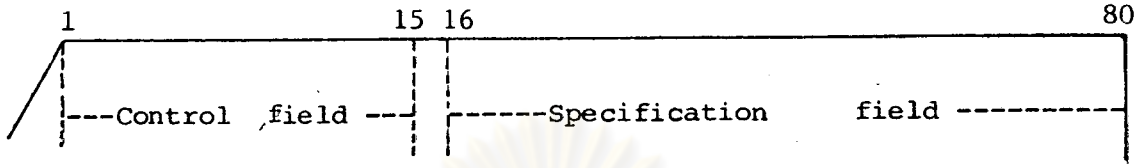
CLASS = N บอกให้คอมพิวเตอร์ทราบว่าข้อมูลที่จะอ่านบันทึกลงบัตรเจาะรู

T บอกให้คอมพิวเตอร์ทราบว่าข้อมูลที่จะอ่านบันทึกลงเทป

yyyyy เป็นชื่อของผู้ใช้บริการ

ข้อกำหนดเกี่ยวกับการเขียน Control Cards ของโปรแกรม SPSS

การเขียน Control Cards แต่ละบัตรของ SPSS (ในส่วนของ SPSS Control Cards แบ่งเป็น 2 ส่วน คือ



1. Control field คือ ส่วนที่มีขอบเขตตั้งแต่คอลัมน์ที่ 1 - คอลัมน์ที่ 15 ข้อความที่เขียนในส่วนนี้เรียกว่า Control Word ซึ่งเป็นคำเฉพาะที่ผู้ใช้ต้องเขียนตามข้อกำหนดของคู่มือ SPSS

2. Specification field คือ ส่วนที่มีขอบเขตตั้งแต่คอลัมน์ที่ 16 - คอลัมน์ที่ 80 ข้อความที่เขียนในส่วนนี้ บรรทัดแรกจะต่อเนื่องหรือสอดคล้องกับ Control Word ส่วนบรรทัดต่อไปถ้าจำเป็นต้องมีข้อความที่ต่อเนื่องกับบรรทัดแรก ส่วนที่เป็น Control field ต้องว่างไว้ คือ ข้อความในบรรทัดที่สองต้องไปขึ้นที่คอลัมน์ 16 เลย (ในกรณีใช้บัตร 1 ใบไม่พอก็ต่อบัตรใหม่ โดยเริ่มที่คอลัมน์ที่ 16 ทุกครั้ง)

การจัดทำข้อความใน Specification field มีหลักเกณฑ์ดังนี้

1. Names ชื่อตัวแปรที่ใช้อาจเป็นตัวอักษรล้วน หรือตัวอักษรปนตัวเลข กำหนดโดยผู้ใช้เอง ไม่เกิน 8 ตัว (8 คอลัมน์) แต่ในตัวแรกของชื่อที่กำหนดต้องเป็นตัวอักษรเสมอ (A - Z)

2. Values เมื่อเขียนข้อความใน Specification field ที่มีค่า ซึ่งอาจเป็นค่าที่ตัวเลขล้วน (numeric) , เป็นตัวอักษรปนตัวเลข หรือเป็นตัวอักษรล้วน (Alphabatic) มักจะใช้เป็นชื่อตัวแปร

ค่าที่กล่าวมาแล้ว ถ้าเป็นตัวเลขล้วน อาจเป็นจำนวนที่มีจุดทศนิยมต้องไม่เกิน 8 หลัก (ไม่นับทศนิยม ส่วนตัวแปรที่มีค่าเป็นตัวอักษร หรือตัวเลขปนตัวอักษร ชื่อตัวแปรที่เขียนในส่วน Specification field **ต้องเขียนภายในเครื่องหมาย** ' ' ซึ่งค่านี้จะไม่นำไปคำนวณ

3. Keyword เป็นคำเฉพาะหรือมีความหมายพิเศษ โดยเฉพาะจะมีความหมายตามที่ปรากฏใน Control Word แต่ละชนิด เช่น THRU ใช้กับค่าที่เป็นตัวเลข เช่น 0 THRU 10 TO ใช้เกี่ยวกับตัวแปร เช่น AGE TO SCØRE ซึ่งหมายถึงรวมตัวแปร AGE, SCØRE และตัวแปรทุกตัวที่อยู่ระหว่างตัวแปร AGE และ SCØRE ที่ระบุตาม VARIABLE LIST

4. Delimeters ตัวแสดงวรรคตอนใช้กรณีแยกคำ, ข้อความ ฯลฯ ที่ปรากฏในส่วน Specification field ใช้ 2 ประเภท คือ

4.1 ใช้กรณีทั่ว ๆ ไป ใช้เครื่องหมาย , หรือเว้นไว้ 1 คอลัมน์

4.2 ใช้กรณีพิเศษ คือ เครื่องหมาย ( ) , / , =

( ) ใช้เป็นกลุ่มของค่าใด ๆ เช่น V5 (0, 8, 9)

/ ใช้คั่นกลุ่มตัวแปร หรือบอกว่า ค่าของตัวแปรที่ตามหลังเครื่องหมาย / จะเริ่มบนบัตรใบถัดไป

= ใช้ในการคำนวณเป็นส่วนใหญ่

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บัตรคำสั่งประเภท SPSS Control Cards

1. Data Definition Cards

เป็นกลุ่มของบัตรที่จะอธิบายรายละเอียดเกี่ยวกับข้อมูล จะกล่าวถึงบัตรที่จำเป็นต้องใช้ในการเรียกใช้โปรแกรม SPSS เรียงตามลำดับดังนี้

(1) VARIABLE LIST Card

เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้ง มีรูปแบบดังนี้

```
1                          16
VARIABLE   LIST            list of Variable name
```

ใช้ระบุชื่อตัวแปรที่ต้องการวิเคราะห์ ชื่อตัวแปรกำหนดตามกฎของการตั้งชื่อในโปรแกรม SPSS คือ ชื่อตัวแปรมีขอบเขตไม่เกิน 8 คอลัมน์ เป็นได้ทั้งตัวอักษรล้วน, ตัวอักษรปนตัวเลข แต่ตัวแรกต้องเป็นตัวอักษรเสมอ มีได้มากที่สุด 500 ตัวแปร

ตัวแปรที่เขียนใน list of Variable name สามารถเขียนได้หลายแบบดังนี้

ก. ระบุชื่อตัวแปร ตามที่บันทึกในบัตรข้อมูล เช่น ตัวแปรมี 3 ตัว ชื่อ

```
AGE,  SEX,  EDUC
1                  16
VARIABLE  LIST     AGE,  SEX,  EDUC
```

ข. ใช้ Keyword TO เมื่อมีตัวแปรเป็นจำนวนมาก เช่น มี 200 ตัวแปร

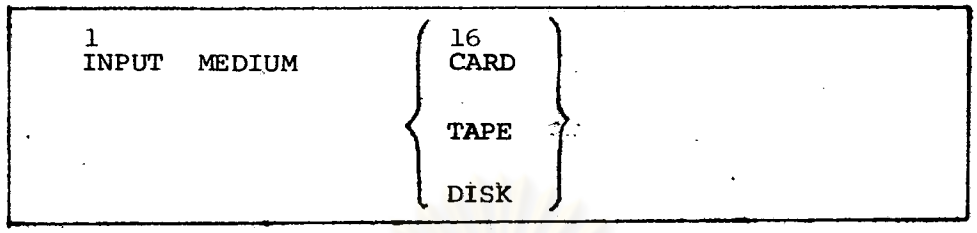
```
1                  16
VARIABLE  LIST     V1   TØ   V 200
```

ค. แบบผสมทั้งแบบ ก และ ข

```
1                  16
VARIABLE  LIST     AGE,  SEX,  V1  TØ  V 30
```

(2) INPUT MEDIUM Card

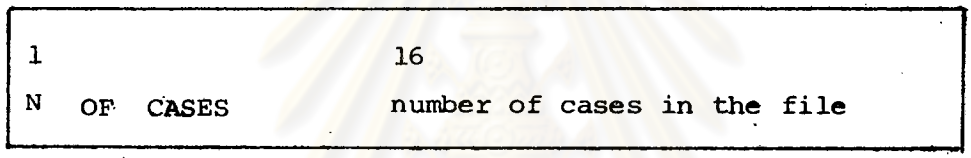
เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้ง มีรูปแบบดังนี้



ใช้ระบุสื่อที่ใช้บันทึกข้อมูลว่าเป็นบัตร (CARD), เทป (TAPE) หรือจานแม่เหล็ก (DISK)

(3) N OF CASES Card

เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้ง เมื่อไม่มีบัตร Subfile List

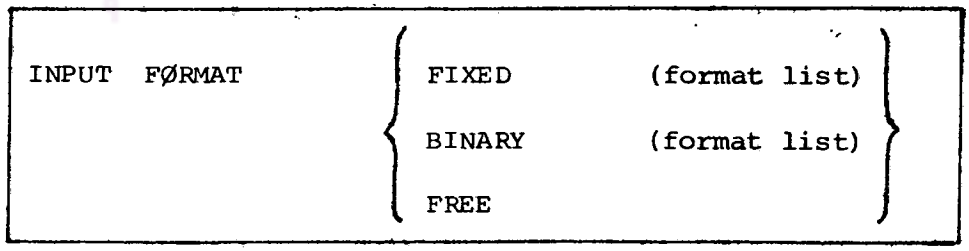


ใช้ระบุจำนวนค่าสังเกต (Observation) ที่ใช้ในการวิเคราะห์ กรณีข้อมูลบันทึกในเทป หรือจานแม่เหล็ก (disk) ถ้าเราไม่ทราบจำนวนของค่า สังเกตอาจใช้ Keyword UNKNOWN ได้

และถ้าไม่ต้องการใช้ทุกค่าสังเกตที่บันทึกในเทป หรือจานแม่เหล็ก ก็อาจระบุตัวเลขเท่าจำนวนที่ต้องการใช้

(4) INPUT FORMAT Card

เป็นบัตรคำสั่งที่ต้องใช้ทุกครั้ง มีแบบที่ใช้ดังนี้



ใช้ระบุชนิดและตำแหน่งตัวแปรแต่ละตัว เรียงกันตามที่กำหนดไว้ใน VARIABLE LIST Card ซึ่งมีเลือกใช้ 3 แบบ คือ

4.1 แบบ FIXED

หมายความว่า ค่าของตัวแปรหนึ่ง ๆ ถูกบันทึกบนตำแหน่งเดียวกัน เหมือนกันหมดทุก ๆ ค่าสังเกตคือ ค่าของตัวแปรอยู่ในตำแหน่งที่แน่นอน

4.2 แบบ BINARY

หมายความว่า ค่าของตัวแปรที่บันทึกเป็นแบบ Computer Words คือ เลขระบบฐาน 2

4.3 แบบ FREE

หมายความว่า ค่าของตัวแปรที่บันทึกบนตำแหน่งเดียวกัน อาจจะเป็นค่าของตัวแปรต่างกันก็ได้ โดยการบันทึกค่าตัวแปรแต่ละค่าบนบัตรข้อมูลนั้น อาจจะแยกกันโดยเว้นช่องว่างไว้ (Blank)

แบบ FIXED, BINARY ต้องมี format list ตามหลัง

แต่แบบ FREE ไม่ต้องมี format list ตามหลัง

แบบที่ใช้ง่าย คือ แบบ FIXED

การเขียน format list สำหรับแบบ Fixed

- การบอกชนิดและขอบเขตของตัวแปร

ก. แบบ F - type รูปแบบ คือ | n Fw.d |

ใช้เมื่อค่าตัวแปรมีค่าเป็นทศนิยม แม้สื่อที่จะใช้บันทึกจะมีหรือไม่มีทศนิยมก็ตาม

n - เป็นตัวเลขที่บอกจำนวนตัวแปรที่มีรูปแบบเช่นนี้

w - เป็นตัวเลขที่บอกความกว้างของจำนวนคอลัมน์ที่ใช้ (field, Size)

d - เป็นตัวเลขที่บอกจำนวนตำแหน่งหลังจุดทศนิยม

ข. แบบ A - type รูปแบบ คือ | n A w |

ใช้เมื่อค่าตัวแปรมีค่าเป็นตัวอักษรปนตัวเลข

n, w ความหมายเหมือนแบบ F - type

- การบอกตำแหน่งของตัวแปร โดยใช้

ก. แบบ | nX | n เป็นจำนวนที่จะบอกจำนวนคอลัมน์ที่ต้องการเว้น เช่น ต้องการเว้น 6 คอลัมน์ ก็ใช้ 6X

ข. แบบ | Tn | เป็นการบอกให้ไปอ่านคอลัมน์ที่ n นอกจากนี้ ยังมีเครื่อง / ที่ใช้ประกอบใน format list เครื่องหมาย / เป็นการบอกว่า ค่าตัวแปรที่ตามหลังเครื่องหมายนี้อยู่บนบัตรใบถัดไป เช่น

```
                1                   16
INPUT   FORMAT      FIXED (T4, 2F4.2 /////// 7X, F5.2/ 1X, A5 ////)
```

เป็นการเลือกตัวแปรมาทำการวิเคราะห์ 4 ตัวแปร เป็นแบบที่อยู่ในตำแหน่งที่แน่นอน ค่าของตัวแปรที่ 1 อยู่บนบัตรใบที่ 1 เริ่มที่คอลัมน์ที่ 4 ถึงคอลัมน์ที่ 7 เป็นตัวเลขทศนิยม 2 ตำแหน่ง

"	2	"	1	"	8	"	11	"	2 "
"	3	"	8	"	8	"	12	"	2 "
"	4	"	9	"	2	"	6	เป็นประเภทอักษรปนตัวเลข	

จำนวนเครื่องหมาย / ข้างหลังจะบอกจำนวนบัตรที่จะเว้น

จากตัวอย่างเมื่ออ่านถึงตัวแปรตัวสุดท้ายแล้วให้เว้นไป 4 บัตร

(5) **MISSING VALUE** Card

เป็นบัตรคำสั่งที่จำเป็นต้องใช้ เมื่อจะมีการแยกระหว่างค่าสมบูรณ์ และค่าที่ไม่สมบูรณ์ (missing value) เพื่อจะนำไปใช้ภายหลัง มีรูปแบบคือ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
MISSING VALUES { Variable name / or / Variable list } (missing value)/ { , } ( )
```

{Variable name / or / Variable list} หมายถึง ชื่อของตัวแปรที่มีค่าไม่สมบูรณ์

missing value เป็นตัวเลขที่เรากำหนดไว้เป็นรหัสแทนค่าที่ไม่สมบูรณ์ (missing value) ซึ่งอาจจะเป็นตัวเลขได้ไม่เกิน 3 ค่า ถ้าเป็นค่าไม่สมบูรณ์ ประเภทตัวอักษรปนตัวเลขจะเขียนในเครื่องหมาย ' ' ถ้าเป็นตัวเลขก็จะระบุได้เลย เช่น

```
1                16
MISSING VALUES  SEX  (8,9) /V1 To V4 (8,99,999)/ ID ('-','x','y')
```

ค่าที่ไม่สมบูรณ์ของตัวแปรชื่อ SEX คือ ค่าที่ใช้รหัสแทนด้วยเลข 8, 9

" V1, V2, V3, V4 ใช้รหัสเหมือนกันคือ 8, 99, 999

" ID คือ ค่าที่ใช้รหัส -, x, y ซึ่งเป็นตัวแปรที่มีค่าเป็นตัวอักษรปนตัวเลข

(6) VALUE LABELS Card

บัตรคำสั่งนี้ไม่จำเป็นต้องใช้ แต่เราจะใช้เมื่อต้องการให้พิมพ์ความหมายของรหัสที่ใช้มักใช้ในตัวแปรที่มีค่าเป็นมาตรานามบัญญัติ (Nominal Scale) มาตราเรียงลำดับ (Ordinal Scale) รูปแบบที่ใช้ คือ

```
1                 16
VALUE LABELS { Variable name / or / Variable list } (value 1) label 1 (value 2) label.../
```

{Variable name / or / Variable list} หมายถึง ชื่อตัวแปรที่ต้องการให้พิมพ์ความหมายของรหัส ชื่อตัวแปรนี้ต้องเหมือนกับที่ใช้ใน VARIABLE LIST

(value 1) หมายถึง ค่าแรกหรือรหัสแรกของตัวแปร

label 1 หมายถึง ความหมายของรหัสแรกที่ใช้

เช่น ตัวแปรชื่อ SEX มี 2 รหัส คือ เลข 1 หมายถึง MALE, เลข 2 หมายถึง FEMALE

" RACE มี 3 รหัส คือ เลข 1 หมายถึง WHITE

เลข 2 " BLACK

เลข 3 " OTHER

เราสามารถเขียนได้ดังนี้

```
1              16
VALUE LABELS SEX (1) MALE (2) FEMALE/RACE (1) WHITE (2) BLACK (3) OTHER
```

(7) VAR LABELS Card

บัตรคำสั่งนี้คล้ายกับบัตรคำสั่งที่ (6) ซึ่งจะใช้เมื่อต้องการบอกให้พิมพ์ความหมาย หรือข้อความขยายตัวแปรที่ใช้ใน VARIABLE LIST Card เพราะบางครั้งชื่อตัวแปรที่เราตั้งขึ้นใน VARIABLE LIST Card เรามักจะตั้งชื่อสั้น ๆ และไม่ให้เกิน 8 คอลัมน์ ซึ่งผู้อื่นอาจไม่รู้ความหมายบัตรนี้ จะช่วยขยายความหมายของตัวแปรที่ใช้ มีรูปแบบคือ

```
1              16
VAR LABELS     Varname 1  var label/varname 2 var label/, ...
```

Var name 1 ชื่อตัวแปรที่ 1 ที่ใช้ใน Variable List

Var label ความหมายหรือชื่อเต็มของตัวแปรที่ 1

เช่น ตัวแปรที่เราใช้ V1 ซึ่งความหมายคือ SEX

AGE " AGE OF THE RESPONDENT

EDUC " EDUCATION OF HEAD

เราสามารถเขียนได้ดังนี้

1	16
VAR LABELS	V1, SEX/AGE, AGE OF THE REPONDENT/
	EDUC, EDUCATION OF HEAD

2. Task Definition Cards

เป็นกลุ่มของบัตรที่ใช้เรียกโปรแกรมย่อยในโปรแกรม SPSS ไปทำงานทางสถิติที่ต้องการ และยังประกอบด้วยบัตรที่จะควบคุมการทำงานต่าง ๆ มีที่จำเป็นต้องใช้ดังนี้ คือ

(1) Procedure Card

เป็นบัตรที่ใช้สำหรับให้โปรแกรม SPSS ทำงานทางสถิติที่ต้องการ โดย Procedure Card แต่ละใบอาจประกอบไปด้วย Control Word เพียง 1 Control Word หรือหลาย ๆ Control Word ก็ได้ และต้องใช้ตามกฎเกณฑ์ที่ระบุไว้ ซึ่งเป็นคำที่เข้าใจง่าย สำหรับผู้ที่มีพื้นฐานทางสถิติอยู่แล้ว เช่น ANOVA , REGRESSION ฯลฯ ซึ่งโปรแกรม SPSS จะแบ่งเป็นโปรแกรมย่อย ๆ ตามชนิดของงานเรียกว่า SUB PRØGRAM

(2) OPTIONS Card

เป็นบัตรที่ใช้คู่กับ Procedure Card เพื่อขยายคำสั่งใน Procedure Card หรือให้ได้ผลลัพธ์เพิ่มเติม เวลาใช้ต้องเรียงบัตรถัดจาก Procedure Card มีรูปแบบดังนี้ คือ

1	16
OPTIONS	Integer number

Integer number เป็นเลขจำนวนเต็มเริ่มตั้งแต่เลข 1, 2 ... ซึ่งจะมีตัวเลขให้เลือกเป็นจำนวนไม่เท่ากัน แล้วแต่ Procedure Card นั้น ๆ และตัวเลขตัวเดียวกันที่แสดงใน Procedure Card หนึ่ง ก็อาจจะให้ผลลัพธ์ต่างจากอีก Procedure Card หนึ่ง เวลาต้องการได้ผลลัพธ์อย่างไร ก็เลือกตัวเลขแล้วระบุลงไป

(3) STATISTIC card

เป็นบัตรที่ใช้คู่กับ Procedure card เหมือน OPTIONS card ทำหน้าที่ให้เครื่องเลือกแสดงค่าสถิติที่ต้องการควบคู่ไปกับการคำนวณนั้น ๆ แบบที่ใช้ คือ

```
1
STATISTICS        { 16
                    inte ger number
                          or
                    ALL                }
```

inte ger number ความหมายคล้ายกับ inte ger number ใน OPTIONS cards

All ใช้เมื่อต้องการค่าสถิติทุกค่าที่สามารถทำได้ในการคำนวณนั้น ๆ ถ้าต้องการบางค่า ก็ระบุตัวเลขใน inte ger number

(4) READ INPUT DATA

เป็นบัตรคำสั่งที่จะบอกให้เครื่องคอมพิวเตอร์เริ่มอ่านข้อมูล ข้อมูลจากสื่อใด ๆ ที่ใช้บันทึกตามที่บ่งไว้ใน INPUT MEDIUM card บัตรนี้จะต้องใช้ทุกครั้ง ไม่ว่าข้อมูลอยู่ในสื่อใด ใช้ตามหลัง ØPTIONS หรือ STATISTICS card ของกลุ่ม Task Definitish cards กลุ่มแรกทันที แบบที่ใช้ คือ

```
1
READ  INPUT  DATA
```

นอกจากนี้ มีกลุ่มบัตรอีกกลุ่มหนึ่งที่จะช่วยในการทำงาน และบอกการเริ่มและการจบของงาน บัตรกลุ่มนี้เรียกว่า "RUN cards" มีที่จำเป็นดังนี้

(1) RUN NAME card

เป็นบัตรที่จะใช้เมื่อต้องการบอกชื่อโครงการที่จะใช้คอมพิวเตอร์ทำงาน มีแบบที่ใช้ คือ

```
1                              16
RUN  NAME                      label
```

label เป็นชื่อที่กำหนดขึ้น ใช้อักษรและสัญลักษณ์ต่าง ๆ ได้ไม่เกิน 64 ตัว เมื่อใช้ RUN NAME CARD จะพิมพ์ชื่อ label ไว้บนทุกหน้าของผลลัพธ์

(2) TASK NAME card

ใช้เมื่อต้องการบอกถึงงานย่อยที่จะทำตาม Procedure card โดยจะพิมพ์ชื่อที่ปรากฏบน TASK NAME บรรทัดถัดจาก RUN NAME ทุกหน้า จนกว่าจะมี TASK NAME ใหม่ ปกติเราจะใช้ TASK NAME card ไว้หน้า Procedure card แบบที่ใช้ คือ

```
1                          16
TASK  NAME                 label
```

label ใช้ระบุชื่องานย่อยที่คำนวณอยู่มีขอบเขต 64 ตัว

(3) FINISH card

บัตรนี้ต้องมีทุกครั้ง จะบอกให้ทราบถึงการจบของงานที่ทำ เป็นบัตรใบสุดท้ายก่อนชุด Job Control card ชุดหลัง มีแบบที่ใช้ คือ

```
1
FINISH
```

(4) PAGESIZE card

บัตรนี้อาจใช้หรือไม่ก็ได้ การใช้บัตรนี้จะช่วยประหยัดเวลาและกระดาษที่ใช้ เพราะโดยปกติโปรแกรม SPSS จะพิมพ์ 55 บรรทัดต่อ 1 หน้า เมื่อได้ผลลัพธ์แล้ว แม้จะไม่ครบ 55 บรรทัด ก็จะเหลือที่ว่างไว้ทำให้เปลืองกระดาษ เราสามารถจะกำหนดให้พิมพ์กี่บรรทัดใน 1 หน้า โดยกำหนดเอาเอง มีแบบที่ใช้ คือ

```
1                          16
PAGESIZE                   { n       }
                           { or      }
                           { NOEJECT }
```

n = จำนวนบรรทัดที่ต้องการให้พิมพ์ใน 1 หน้า

NOEJECT = บอกให้พิมพ์ต่อไปเลย เมื่อได้ผลลัพธ์หนึ่ง ๆ แล้ว

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

การดัดแปลงข้อมูล Data transformation

กลุ่มบัตรที่จะช่วยในการทำ Date transformation เรียกว่า "Data Modification cards" มีที่จำเป็นดังนี้

(1) RECODE card

ใช้บัตรนี้เพื่อจัดกลุ่มตัวแปรใหม่ หรือเปลี่ยนค่าใหม่ มีแบบที่ใช้ คือ

```
1                16
RECODE    { Var name }  (value list = new list).../{ } ( ) /
          {    or    }
          { Var list }
```

{Variable name / or / Variable list} = ชื่อของตัวแปรที่ปรากฏใน VARIABLE LIST ที่ต้องการทำ RECODE

Value list = ค่าของตัวแปร ซึ่งอาจจะมีหลายค่าต้องการจัดเป็นกลุ่ม

new list = ค่าตัวแปรที่กำหนดมาใหม่

การเขียนในส่วน (value list = new list) อาจเขียนได้หลายแบบ และมีคำเฉพาะประกอบ

1. เปลี่ยนค่าของตัวแปรเป็นค่าใหม่ เช่น (6 = 1) หมายความว่า ถ้าตัวแปรมีค่าเป็น 6 ให้เปลี่ยนใหม่เป็น 1
2. จัดกลุ่มใหม่ของตัวแปร เช่น (4, 6, 8 = 1) หมายความว่า ถ้าตัวแปรมีค่าเป็น 4,6,8 ให้เปลี่ยนใหม่เป็น 1
3. ใช้คำเฉพาะ "THRU" เป็นค่าตัวแปรจากค่าที่ต่อเนื่องเป็นค่าไม่ต่อเนื่อง 0 THRU 500 = 1 หมายความว่า ถ้าตัวแปรมีค่าตั้งแต่ 0-500 ให้มีค่าเป็น 1

ข้อควรระวัง คือ ถ้าตัวแปรที่มีค่าต่อเนื่อง ควรจะเขียน RECODE ให้คลุมทุก ๆ ค่า เช่น (0 THRU 500 = 1) (500 THRU 1000 = 2) หมายความว่า ถ้าตัวแปรมีค่าตั้งแต่ 0-500 ให้มีค่าเป็น 1 และค่าที่เกิน 500 - 1,000 จะมีค่าเป็น 2

4. ใช้คำเฉพาะ "LOWEST, HIGHEST, ELSE" ใช้ประกอบคำเฉพาะ "THRU" ใช้เมื่อไม่ทราบค่าสูงสุด และต่ำสุดของตัวแปร ซึ่งคำเฉพาะที่ใช้อาจใช้คำย่อก็ได้ เช่น

```
1                16
RECODE           AGE (LO THRU 15 = 1) (16 THRU 29 = 3)
                 (30 THRU 35 = 4) (36 THRU HI = 5) (ELSE = 9)
```

คำเฉพาะ ELSE (ELSE = value) ใช้กำหนดค่าของตัวแปรที่ไม่ได้กำหนดรหัสใด ๆ ไว้ให้มีค่าเท่ากับค่าที่กำหนดใน ELSE จากตัวอย่างถือว่าค่าอื่น ๆ นอกเหนือจากที่กำหนดไว้จะมีค่าเป็น 9

5. โดยปกติเครื่องคอมพิวเตอร์บางเครื่อง เช่น IBM 360/370 ไม่สามารถแยก BLANK และศูนย์ โดยถือว่า BLANK หรือช่องว่างที่ปรากฏบนบัตรข้อมูลนั้นเราใช้ในความหมายอะไร เช่น เราใช้ BLANK หมายถึง ค่าไม่สมบูรณ์ โดยค่าไม่สมบูรณ์ เราใช้รหัส 9 สามารถเขียนได้ดังนี้

```
1                16
RECODE           INCOME (BLANK = 9) (0 THRU 5000 = 1)
                 (5000 THRU 10000 = 2) (10000 THRU HI = 3)

VALUE LABELS     INCOME (1) LOW (2) MEDIUM (3) HIGH (4)
                 NOT ASCERTAINED
```

ข้อสังเกต จะต้องกำหนด BLANK ก่อนที่จะกำหนด 0 เป็นค่าอื่น ๆ เพราะถ้ากำหนดที่หลัง 0 จะถือว่า BLANK มีค่าตามศูนย์ (0)

คำเฉพาะ THRU, LOWEST, HIGHEST และ BLANK ใช้ได้กับตัวแปรที่เป็นตัวเลขเท่านั้น คำเฉพาะ ELSE ใช้ได้ทั้งตัวแปรที่เป็นตัวเลขและตัวอักษร

การเปลี่ยนข้อมูลประเภทที่รับเข้ามาในรูป A - type. (อยู่ภายในเครื่องหมาย ' ')

ก. เปลี่ยนเป็นค่าใหม่ใช้รูปแบบ 'value list' ='new value' เช่น

```
1.              16
RECODE          NAME    ('A', 'B', 'C', '3', '2', = 'N')
```

ข. เปลี่ยนให้เป็นตัวเลขในรูปแบบ 'value list' = new value เช่น

```
1               16
RECODE          ID      ('X', 'Y', '1' = 1) ('A', 'B', '2', '3' = 2)
                        ('XA' = 3) ('BZ' = 4)
```

ค. เปลี่ยนข้อมูลประเภทตัวเลข, ตัวอักษร สัญลักษณ์พิเศษอื่น ๆ alphanumeric ซึ่งมีความยาว 1 ตัวอักษรให้เป็นตัวเลข โดยใช้คำเฉพาะ CONVERT ต่อท้ายชื่อตัวแปร (variable name or variable list) โดยจะเปลี่ยนดังนี้

รูปแบบข้อมูลที่บันทึก (alphanumeric)	ผลที่ได้จากการใช้ (CONVERT NUMERIC)
'0' ถึง '9'	เปลี่ยนเป็น 0 ถึง 9 ตามลำดับที่สมนัยกัน
'8' และ '+'	" 12
'_'	" 11
ช่องว่าง (Blank) และสัญลักษณ์อื่นตลอดจนอักษร 'A' ถึง 'Z'	" 13

ข้อควรระวัง ใช้เฉพาะตัวอักษรหรือตัวเลข และสัญลักษณ์อื่น ๆ ที่มีความยาวเพียง 1 คอลัมน์เท่านั้น เมื่อข้อมูลที่บันทึกไว้มีทั้งตัวอักษร ตัวเลขและสัญลักษณ์อื่น ๆ ปนกัน ซึ่งความยาวมากกว่า 1 คอลัมน์ สามารถเปลี่ยนให้เป็นตัวเลขที่ต้องการได้ โดยใช้แบบข้อ ข ร่วมกับคำเฉพาะ CONVERT เช่น

```
1                       16
RECODE                  V2 ('A21'= 1) ('B21'= 2) ('_8_ _' =3) ('6_ _ _'= 9)
```

เมื่อใช้ RECODE card จะมีผลตลอดทั้งโปรแกรมทุกงานที่มีชื่อตัวแปรระบุไว้ใน RECODE card ถ้าต้องการใช้เพียงครั้งเดียว หรือให้ทำงานชนิดเดียว สามารถใช้รูปแบบดังนี้

1	16
*RECODE	รูปแบบและการใช้เหมือนกับ RECODE CARD

ซึ่งจะมีผลเฉพาะงานที่สั่งให้ทำโดย Procedure card ที่ตามหลัง *RECODE นี้เท่านั้น

(2) COMPUTE card

ใช้บัตรนี้เพื่อสร้างตัวแปรใหม่จากตัวแปรที่มีอยู่แล้ว หรือดัดแปลงตัวแปรที่มีอยู่แล้วในบัตร VARIABLE LIST card

1	16
COMPUTE	computed variable = arith metic expression

computed variable หมายถึง ชื่อของตัวแปร ซึ่งอาจจะเป็นตัวแปรที่สร้างขึ้นมาใหม่ หรือชื่อตัวแปรเดิมที่มีอยู่แล้ว และต้องการเปลี่ยนแปลงค่า

arithmetic expression เป็นการใช้ตัวแปรกับตัวเลข, เครื่องหมายในการคำนวณ หรือฟังก์ชั่นทางคณิตศาสตร์ ดังมีรายละเอียดดังนี้

1. arithmetic Operator คือ เครื่องหมายที่ใช้ในการคำนวณ ซึ่งประกอบไปด้วย (+, -, *, /, **)

2. mathematical function คือ ฟังก์ชั่นที่ใช้หาค่าทางคณิตศาสตร์ ซึ่งประกอบไปด้วย SQRT, LN, LG10, EXP, SIN, CØS, ATAN

3. คำเฉพาะบางคำ ซึ่งประกอบไปด้วย RND, ABS, TRUNC, MOD10

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ซึ่งลำดับการทำงานของ arithmetic expression เป็นดังนี้

1. ฟังก์ชั่นพิเศษใน mathematical function
2. ยกกำลัง
3. คูณ, หาร
4. บวก, ลบ

ถ้ามีหลายฟังก์ชั่น หรือเครื่องหมายที่ใช้ในการคำนวณมีลำดับที่เท่ากันรวมอยู่ใน arithmetic expression จะทำจากซ้ายไปขวา และถ้ามีเครื่องหมายวงเล็บ จะทำในวงเล็บก่อน

การใช้ COMPUTE card จะมีผลตลอดทั้งโปรแกรม

ถ้าต้องการใช้เพียงงานเดียวสามารถใช้ *COMPUTE เช่นเดียวกับ *RECODE ซึ่งจะมีผลเฉพาะงานที่สั่งให้ทำโดย Procedure card ที่ตามหลัง *COMPUTE เท่านั้น

ข้อควรระวัง

บัตร COMPUTE card จะต่างจากบัตรอื่น ๆ เพราะจะเขียนข้อความที่ใช้ดัดแปลงข้อมูลได้เพียงข้อความเดียว คือ กำหนดตัวแปรจาก Compute variable ได้เพียงตัวเดียว ถ้าต้องการตัวแปรใหม่ ต้องเขียนบัตร COMPUTE card ใหม่ เช่น

```
1
COMPUTE         STATE   =  TRUNC     (PLACE/100)
COMPUTE         COUNTY  =  TRUNC     (PLACE/10 - STATE * 10)
COMPUTE         TOWN    =  MOD 10    (PLACE)
```

(3) IF card

ใช้บัตรนี้เพื่อสำหรับเปลี่ยนแปลงค่าของตัวแปร หรือสร้างตัวแปรใหม่

1	16
IF	(logical expression) computed variable = arithmetic expression

computed variable และ arithmetic expression ใช้เหมือนกันกับบัตร COMPUTE card

<u>logical expression</u> เป็นการใช้ตัวแปรกับคำเฉพาะและตัวเลข เพื่อใช้เปรียบเทียบค่าทางตรรกศาสตร์ระหว่างตัวแปร หรือจำนวน 2 จำนวน ตัวแปรใหม่หรือตัวแปรเดิมใน computed variable จะถูกคำนวณ โดย arithmetic expression ที่กำหนด เมื่อ logical expression เป็นจริง ถ้าไม่เป็นจริงจะทำคำสั่งต่อไป

คำเฉพาะ (Keyword) ที่ใช้เปรียบเทียบใน logical expression มีดังนี้

GE, LE, GT, LT, EQ, NE

รูปแบบของ logical expression เป็นดังนี้

16
(arithmetic expression 1) keyword
(arithmetic expression 2)

ตัวอย่างการใช้ logical expression

<u>ตัวอย่างที่ 1</u>
```
    16
VAR 1   EQ  100
```

<u>ตัวอย่างที่ 2</u>
```
    16
VAR 1    * VAR 2  GT 4 * (VAR 3 - 7)
```

เปรียบเทียบผลคูณของ VAR 1, VAR 2 กับผลคูณของ 4 กับ (VAR 3 - 7)

ตัวอย่างการใช้ IF card

```
1           16
IF          (VAR1 * VAR2  GT 4 * (VAR3-7) NEWVAR  = VAR 1+4

IF          (VAR1 * VAR2  EG 4 * (VAR3-7) NEWVAR  = VAR 2+4

IF          (VAR1 * VAR2  LE 4 * (VAR3-7) NEWVAR  = VAR 3+4
```

จาก IF card ใบแรกหมายความว่า เมื่อ logical expression เป็นจริง compute variable คือ NEWVAR จะทำการคำนวณ arithmetic expression คือ VAR 1+4 ถ้า logical expression ไม่เป็นจริง จะไปทำตามคำสั่งในบัตรต่อมา

(การใช้บัตร IF card จะคล้ายกับ COMPUTE card คือ ถ้าต้องการเพิ่มข้อความใหม่ต้องใช้บัตรอีก 1 ใบ)

การใช้คำเฉพาะ OR, AND, NOT ใน IF card เมื่อมีหลาย ๆ logical expression ซึ่งมีการใช้คำเฉพาะ (keyword) ที่แสดงการเปรียบเทียบมากกว่า 1 คำ เราสามารถใช้ OR, AND เชื่อม logical expression ได้ โดยใช้รูปแบบดังนี้

```
                16
(logical expression 1) { OR  } (logical expression 2)
                       { AND }
```

เมื่อใช้ OR ผลของทั้งข้อความเป็นจริง เมื่อค่า logical expression 1 หรือ 2 เป็นจริง หรือเป็นจริงทั้ง 2 ค่า

เมื่อใช้ AND ผลของทั้งข้อความเป็นจริง เมื่อค่า logical expression ทั้ง 2 เป็นจริง จะใช้ NOT นำหน้า logical expression จะทำให้ผลที่ได้จาก logical นั้น ๆ กลับเป็นผลที่ตรงกันข้าม

ตัวอย่างการใช้ OR, AND, NOT IF CARD

```
1          16
IF         (VAR 1 GT VAR 2  AND  VAR 3 GT VAR 4)   VAR 5 = 1

IF         (VAR 1 GT VAR 2  OR   VAR 3 GT VAR 4)   VAR 5 = 2

IF         (NOT (VAR L GT   AND  VAR 2 LT VAR 3)   VAR 5 = 3
```

และถ้าตัวแปรที่ใช้เปรียบเทียบ (เฉพาะตัวแปรที่อยู่ซ้ายมือของ Keyword) เป็นตัวเดียวกัน อาจเขียนแบบย่อได้ เช่น

<u>ตัวอย่าง</u> IF (VAR 1 EQ 7 ØR VAR 1 GT 8) POK = 9

อาจเขียนได้เป็น

```
IF   (VAR 1  EQ 7  ØR  GT 8)      POK = 9

IF   (VAR 1  EQ 7  ØR  (VAR 1 LE 12 AND VAR 1 GE 9) )
        OMH = 99
```

เขียนย่อได้เป็น

```
IF   (VAR 1  EQ 7  OR  (LE 12  AND  GE 9)     OMH = 9
```

สามารถใช้ IF CARD ในรูปแบบ *IF ได้ในความหมายเช่นเดียวกันกับ RECØDE และ CØMPUTE

(4) SELECT IF Card

ใช้บัตรนี้เพื่อเลือกค่าสังเกตที่ต้องการมาทำงานมีรูปแบบดังนี้ คือ

```
1                16
SELECT  IF       (logical expression)
```

เมื่อใช้บัตร SELECT IF จะได้ค่าสังเกตที่เป็นจริง หรือสอดคล้องกับเงื่อนไขใน logical expression

ข้อกำหนดในการใช้ SELECT IF

1. ต้องใช้นำหน้า Task - definition card ใบแรก
2. ถ้าใช้ SELECT IF มากกว่า 1 บัตร จะได้ค่าสังเกตที่เป็นจริง หรือสอดคล้องกับเงื่อนไขทุกเงื่อนไขที่ระบุในแต่ละบัตร SELECT IF (เหมือนกับว่าเชื่อมเงื่อนไขของแต่ละบัตร SELECT IF ด้วย
3. เงื่อนไขที่ระบุใน SELECT IF นั้น จะใช้ตลอดโปรแกรมเมื่อต้องการใช้เพียงชั่วคราว ให้มีผลเฉพาะบัตร Procedure ที่ตามหลังมาเท่านั้น สามารถใช้ *SELECT IF ได้ ซึ่งจะใช้นำหน้า task - definition (Procedure card) ใดก็ได้ เมื่อใช้ SELECT IF และ *SELECT IF จะได้ค่าสังเกตที่เป็นจริงตามเงื่อนไขของ SELECT IF และ *SELECT IF เฉพาะคำสั่งตาม Procedure ที่อยู่หลัง *SELECT IF

ตัวอย่างการใช้ SELECT IF

```
1                    16
SELECT  IF           (AGE  GT 20 AND  AGE LE 50)
```

ผลที่ได้จะเลือกเฉพาะค่าสังเกต ซึ่งตัวแปร AGE มีค่ามากกว่า 20 และต้องมีค่าน้อยกว่าหรือเท่ากับ 50 คือ อยู่ระหว่าง 20 ถึง 50

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สรุปบัตรคำสั่งประเภท SPSS control cards เรียงตามลำดับเท่าที่จำเป็น

ประเภทของบัตร ชื่อบัตรที่ใช้ (Contral words)	จุดประสงค์หรือความหมายที่ใช้
RUN card	จะใช้หรือไม่ใช้ก็ได้ นอกจาก FINISH ต้องใช้ทุกครั้ง
1. PAGESIZE	ใช้เมื่อต้องการกำหนดการพิมพ์ผลลัพธ์ออกมาทางกระดาษ
2. RUN NAME	ใช้เมื่อต้องการบอกชื่อโครงงานที่จะทำ
Data Definition cards	บัตรที่ต้องใช้ทุกครั้ง VARIABLE LIST และ INPUT MEDIUM
3. FILE NAME	ใช้เมื่อต้องการบอกชื่อของแฟ้มข้อมูลที่ใช้ทำงาน
4. VARIABLE LIST	ต้องใช้ทุกครั้ง เพื่อระบุชื่อตัวแปรที่ต้องการใช้ทำงาน
5. INPUT MEDIUM	ต้องใช้ทุกครั้ง เพื่อระบุสื่อที่ใช้บันทึกข้อมูล
6. SUBFILE LIST	ใช้เมื่อต้องการบอกโครงสร้างของแฟ้มข้อมูลย่อย
6. N OF CASES	ต้องใช้ทุกครั้ง ถ้าไม่มี SUBFILE LIST เพื่อบอกจำนวนค่าสังเกตที่ใช้
7. INPUT FØRMAT	ต้องใช้ทุกครั้ง ถ้าใช้ VARIABLE LIST
7. MISSING VALUES	ใช้เมื่อต้องการกำหนดค่าไม่สมบูรณ์ของตัวแปร
7. VAR LABELS	ใช้เมื่อต้องการให้พิมพ์ข้อความขยายชื่อตัวแปร
7. VALUE LABELS	ใช้เมื่อต้องการให้พิมพ์ความหมายของรหัสที่ใช้แสดงค่าตัวแปร
7. PRINT FØRMAT	ใช้เมื่อต้องการให้พิมพ์รูปแบบของตัวแปรชนิด 'A - type'

Task Definition cards 1	ใช้สั่งให้โปรแกรม SPSS ไปทำงานที่ 1
8. TASK NAME	ใช้เมื่อต้องการบอกชื่องานย่อยทางสถิติที่จะทำ
9. procedure card	ใช้ระบุชื่อคำเฉพาะของงานทางสถิติตามข้อกำหนดของโปรแกรมใน Control field
10. OPTIONS	ใช้เมื่อต้องการขยายคำสั่งใน procedure cards
11. STATISTICS	ใช้เมื่อต้องการค่าสถิติตามที่ procedure cards ทำได้
12. READ INPUT DATA	ต้องใช้ทุกครั้ง เพื่อให้เครื่องรับข้อมูล
บัตรบันทึกข้อมูล	ถ้าใช้สื่ออย่างอื่นก็ไม่ใส่ส่วนนี้
Task Definition cards 2	ใช้สั่งให้โปรแกรม SPSS ไปทำงานที่ 2 ระบุโดย Procedure card ซึ่งจะใช้บัตรชนิดต่าง ๆ เหมือนงานที่ 1
	เมื่อต้องการทำงานที่ 3, 4...ระบุใน Procedure card
FINISH	ต้องมีทุกครั้ง เป็นบัตรสุดท้ายบอกการจบของทุกงาน

ตัวเลขข้างหน้าชื่อบัตร แสดงลำดับที่จัดไว้ในโปรแกรม

สำหรับบัตรประเภท Data Motification cards ซึ่งใช้ในการดัดแปลงข้อมูล เช่น RECODE, COMPUTE, IF, และ SELECT IF แทรกไว้ระหว่าง Data Detinition cards และ Task - Definition card คือ ไว้หน้า TASK NAME แรก

แต่บัตร *RECODE, *COMPUTE, *IF, *SELECT IF อาจไว้หน้า TASK NAME ใดก็ได้

ตัวอย่างการใช้ฟังก์ชั่นทางคณิตศาสตร์บางชนิดของโปรแกรม SPSS

Mnemonic	Meaning	Example
SQRT	Square Root	VARX = SQRT (VARA)
LN	Natural or Naperian logarithm	VARX = LN (VARA)
LG10	Base 10 logarithm	VARX = LG10 (VARA)
EXP	Exponential (e^{arg})	VARX = EXP (VARA)+ VARC)
SIN	Sine	VARX = SIN (VARA + VARC)
GOS	Cosine	VARX = COS (VARA)
ATAN	Anctangent	VARX = ATAN (VARA)
RND	Round result to (whole number	VARX = RND (VARA + VARC/6)
ABS	Absolute Value	VARX = ABS (VARA)
TRUNC	Truncate Value (whole number without rounding)	VARX = TRUNC (VARA)
MOD10	Result is remainder of division by 10	VARX = MOD10 (VARA)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.2 แสดงความสามารถในการทำงานทางสถิติแต่ละชนิดของทุกโปรแกรม

จะแสดงความสามารถของโปรแกรมทั้งสามในการทำงานทางสถิติแต่ละชนิดตามลำดับดังนี้คือ

(1) แสดงการเปรียบเทียบความสามารถของโปรแกรมทั้งสาม

โดยจะนำเสนอในรูปตาราง แบ่งเป็น 3 ช่องมีข้อความที่ปรากฏบนหัวตารางจากซ้ายไปขวาดังนี้คือ.

ความสามารถของโปรแกรม

จะระบุคำสถิติ, การทดสอบ หรือการทำงานทางสถิติที่จะบอกถึงความสามารถของโปรแกรม โดยระบุไว้เป็นข้อๆ ตามลำดับ เป็นตัวเลขในเครื่องหมายวงกลมเช่น ① , ②

โปรแกรมสำเร็จรูป

ระบุชื่อโปรแกรมสำเร็จรูป 3 ชนิด คือ BMDP, SAS และ SPSS และใช้เครื่องหมาย (√) แสดงไว้ภายใต้ชื่อโปรแกรมตรงกับข้อที่แสดงว่าโปรแกรมนั้นๆ สามารถทำได้ ถ้าไม่มีเครื่องหมาย (√) ปรากฏภายใต้โปรแกรมใดๆ แสดงว่าโปรแกรมนั้นไม่สามารถให้ผลตรงกับข้อที่ระบุได้

ความหมาย

ช่องนี้จะระบุคำศัพท์, สัญลักษณ์, สูตรหรือข้อความที่ขยายความในช่องแรกของตาราง

(2) แสดงตัวอย่างการเขียนและเตรียมชุดคำสั่งของแต่ละโปรแกรม

หลังจากแสดงการเปรียบเทียบความสามารถของโปรแกรมทั้ง 3 แล้วส่วนที่สองนี้จะแสดงตัวอย่างการเขียนคำสั่งและการเรียงบัตรคำสั่งของแต่ละโปรแกรมสำหรับงานทางสถิติหนึ่งๆ ที่ผู้วิจัยได้ทดลองใช้มาแล้วโดยใช้ข้อมูลจากคู่มือการใช้โปรแกรมสำเร็จรูป BMDP ข้อมูลที่นำมาใช้ในการวิจัยนี้มีอยู่ 5 ชุด ซึ่งผู้วิจัยได้ตั้งชื่อไว้ตามลำดับที่ใช้ เช่น " DATA SAMPLE 1"

ข้อมูลชุดหนึ่งๆ อาจจะใช้กับงานทางสถิติหลายชนิดในการใช้ข้อมูลชุดใด นั้น

จะระบุไว้ในการเขียนคำสั่งซึ่งชุดคำสั่งเหล่านี้จะเขียนไว้ในแบบฟอร์มการเขียน (Coding form) แล้วจึงนำไปเจาะบนบัตรข้อมูล

(3) <u>แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องในการใช้ชุดคำสั่งของแต่ละโปรแกรม</u> จากการทดลองใช้ชุดคำสั่งที่ได้เขียนไว้ในส่วนที่สองจะได้ผลลัพธ์ออกมาทางกระดาษต่อเนื่อง ซึ่งจะแสดงไว้ในส่วนที่สามนี้แต่ไม่สามารถแสดงผลทั้งหมดที่ได้ ดังนั้นจึงนำเสนอเพียงบางส่วนเท่านั้น นอกจากนี้ยังได้แสดงตัวเลขในเครื่องหมายวงกลมซึ่งจะมีความหมายตรงกับความสามารถหรือค่าสถิติต่างๆ ที่ระบุไว้ในตารางการเปรียบเทียบความสามารถข้อ (1) และตัวอักษรภาษาไทยในเครื่องหมายวงกลมทางด้านขวามือ หมายถึงผลลัพธ์ที่ได้จากการใช้บัตรคำสั่งที่แสดงด้วยตัวอักษรเช่นเดียวกัน ทางด้านซ้ายมือของตารางการเขียนคำสั่งบนแบบฟอร์มการเขียนในส่วนที่สอง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT 1 การคำนวณหาค่าสถิติเบื้องต้น

(DESCRIPTIVE STATISTICS)

ตารางที่ 1 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 1

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(1) ค่าเฉลี่ย	✓	✓	✓	
(2) ค่าส่วนเบี่ยงเบนมาตรฐาน	✓	✓	✓	
(3) ค่าความคลาดเคลื่อนมาตรฐาน - ของค่าเฉลี่ย	✓	✓	✓	
(4) ค่าสัมประสิทธิ์ความแปรผัน	✓	✓		
(5) ค่าต่ำสุด, ค่าสูงสุด	✓	✓	✓	
(6) พิสัย	✓		✓	
(7) ความแปรปรวน	✓	✓	✓	
(8) ผลรวม		✓	✓	
(9) ผลบวกกำลังสองของส่วนเบี่ยงเบน - จากค่าเฉลี่ย		✓		
(10) ค่ามัธยฐาน	✓		✓	
(11) ค่าฐานนิยม	✓		✓	
(12) ค่าความเบ้	✓		✓	
(13) ค่าความสูง	✓			
(14) ค่าส่วนเบี่ยงเบนควอไทล์ ควอไทล์ที่ 1 (Q_1) ควอไทล์ที่ 2 (Q_2)	✓			
(15) ช่วงความเชื่อมั่นของการประมาณ - ค่าเฉลี่ยซึ่งประกอบไปด้วย ขอบเขตทางขวาและซ้ายของช่วง	✓			

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(16) กราฟแสดงตำแหน่งของ - -ค่าประมาณต่างๆ	✓			
(17) ค่าประมาณโดยวิธีโรบัส	✓			(17) ค่าประมาณโดยวิธีโรบัส มีอยู่ 3 ชนิดคือ - trimmed mean - Hampel - biweight
(18) การจัดกลุ่มค่าสังเกตโดย -ค่าของตัวแปรแล้วหาค่าสถิติ- -เบื้องต้นภายในแต่ละกลุ่ม	✓	✓	✓	
(19) การทดสอบความมีนัยสำคัญของ -ความแปรปรวน - -2 กลุ่มโดยค่าสถิติ F	✓		✓	
(20) การทดสอบความมีนัยสำคัญของ -ค่าเฉลี่ย 2 กลุ่ม- -โดยค่าสถิติ T	✓		✓	
(21) การทดสอบความมีนัยสำคัญของ -ค่าเฉลี่ย 2 กลุ่มหลายๆ ตัวแปร -พร้อมกันโดยค่าสถิติ Hotelling's T^2 Mahalanobis D^2 และค่าสถิติ F	✓			(21) $D^2 = (\underset{\sim}{\bar{X}}_1 - \underset{\sim}{\bar{X}}_2)' \text{Spool}^{-1} (\underset{\sim}{\bar{X}}_1 - \underset{\sim}{\bar{X}}_2)$ $T^2 = D^2 / (\frac{1}{N_1} + \frac{1}{N_2})$ $F = \frac{T^2(N_1+N_2-P-1)}{P(N_1+N_2-2)}$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

DATA SAMPLE1

ID Number	Age	Height	Weight	Birthpill[2]	Cholesterol	Albumin[3]	Calcium[3]	Uric Acid[3]
2381	22	67	144	1	200	43	98	54
1946	22	64	160	2	600	35		72
1610	25	62	128	1	243	41	104	33
1797	25	68	150	2	50	38	96	30
561	19	64	125	1	158	41	99	47
2519	19	67	130	2	255	45	105	83
225	20	64	118	1	210	39	95	40
2420	20	65	119	2	192	38	93	50
1649	21	60	107	1	246	42	101	52
3108	21	65	135	2	245	34	106	48
1375	21	63	100	1	208	38	98	54
2936	21	64	120	2	260	47	106	38
998	21	67	134	1	204	40	108	34
408	21	67	145	2	192	39	95	49
913	21	63	138	1	280	41	102	41
2373	21	64	113	2	230	39	99	38
736	21	63	160	1	215	39	96	39
2334	21	64	115	2	225	44	105	44
3035	21	68	125	1	165	48	105	28
1883	21	62	106	2	200	38	95	40
2729	21	68	150	1	220	47	102	75
1848	21	64	130	2	255	34	102	40
1890	22	62	135	1	263	43	98	47
266	22	62	110	2	173	42	97	37
2125	22	57	105	1	170	46	98	45
2092	22	64	120	2	290	37	98	59
1291	22	64	115	1	263	42	102	47
1922	22	59	94	2	220	47	105	46
307	22	67	125	1	200	43	100	44
1790	22	62	97	2	192	38	95	43
239	22	58	100	1	247	42	104	52
3096	22	66	130	2	175	44	106	58
1068	22	60	100	1	155	41	96	45
2730	22	65	135	2	215	40	93	43
51	22	60	95	1	200	47	99	34
2648	22	67	124	2	247	44	102	45
906	23	63	125	1	220	32	92	42
97	23	64	105	2	207	42	100	40
649	23	63	125	1	266	42	103	47
3014	23	63	120	2	240	43	101	39
962	24	68	125	1	195	49	106	52
1912	24	64	130	2	250	39	103	46
957	24	65	130	1	225	50	108	39
338	24	65	148	2	200	37	104	49
1819	24	64	135	1	180	37	96	49
420	24	71	156	2	240	42	102	51
494	25	62	107	1	330	48	101	53
2779	25	67	175	2	175	39	93	51
477	25	66	112	1	205	46	101	33
2870	25	63	120	2	235	44	103	40
3	54	62	120	1	227		86	25
2077	54	67	127	2	260	44	106	57
3095	25	67	135	1	295	46	106	47
2363	25	67	141	2	230	38	101	52
2698	26	66	135	1	240	48	103	51
3006	26	64	118	2	238	40	99	46
796	26	65	125	1	198	44	96	43
2351	26	65	120	2	196	38	95	43
1627	27	64	120	1	172	43	98	60
152	27	64	180	2	317	37	98	84
2305	27	69	137	1	195	46	101	42
2121	27	64	125	2	185	36	94	54
883	27	63	125	1	168	42	97	41
1882	27	64	124	2	200	40	96	52
416	27	60	140	1	250	36	98	68
16	27	65	155	2	280	42	103	52
1637	28	65	108	1	260	48	106	51
3138	28	62	110	2	250	44	105	38
457	28	65	120	1	175	48	100	47
2830	28	66	113	2	305	41	93	24
1083	28	62	135	1	200	43	97	37
3359	28	65	160	2	235	42	101	41
1255	29	61	142	1	177	39	99	46
3256	29	61	115	2	235	45	98	47
1207	29	68	155	1	226	38	94	43
3210	29	65	118	2	230	44	99	44
2871	30	66	143	1	198	45	107	65
1775	30	63	110	2	295	45	98	46
1274	30	61	99	1	230	43	99	39
2365	30	63	132	2	200	37	96	34
949	30	62	125	1	230	46	104	48
3519	30	63	110	2	262	33	99	41
609	30	64	135	1	174	40	95	35
3021	30	66	112	2	250	44	100	35
1668	30	64	160	1	217	35	95	31
2877	30	68		2	212	38	100	66
3349	31	65	125	1	250	43	98	39
2454	31	66	120	2	237	34	91	49
2098	31	65	115	1	270	41	111	64
1916	31	63	110	2	280	44	99	49
782	31	66	123	1	238	37	96	33
2947	31	67	136	2	218	38	95	42
2996	32	67	132	1	185	39	103	37
441	32	68	203	2	235	38	99	37
3506	32	62	155	1	262	37	99	43
2617	32	65	126	2	160	41	97	40
2522	32	63	125	1	189	40	94	40
2400	32	71	170	2	205	37	90	60
1249	32	62	120	1	260	43	107	38
2367	32	62	145	2	240	45	108	42
57	32	66	140	1	197	44	106	58
2267	32	68	133	2	180	32	95	40
63	54	67	145	1	320	39		
234	54	67	140	2	245	39	104	56
3334	33	64	115	1	205	47	100	54
480	33	60	118	2	260	38	99	38
828	33	67	137	1	243	41	106	55
3019	33	68	130	2	195	40	95	58
61	33	65	130	1	203	44	101	48
2062	33	69	138	2	222	40	104	42
1770	34	62	112	1	197	37	93	44
3208	34	63	125	2	245	38	95	41
1140	35	62	115	1	180	40	91	59
3194	35	67	125	2	223	40	100	37
215	35	66	138	1	254	39	107	41
2501	35	66	140	2	245	39	105	56
1309	36	62	135	1	247	34	90	44
2942	36	67	120	2	175	46	103	39
819	36	66	112	1	215	43	104	42
3277	36	65	121	2	270	43	98	35
2456	37		141	1	237		105	45
2254	37	67	125	2	200	45	99	66
46	37	65	116	1	270	42	100	48
1803	37	63	129	2	230	36	91	22
2811	38	64	165	1	255	44	102	62
2341	38	65	151	2	275	38	94	46
2349	39	64	135	1	210	40	95	46
523	39	64	108	2	198	44	90	38
2675	39	63	195	1	260	40	108	42
2153	39	69	132	2	180	39	94	30
1111	39	62	100	1	210	45	91	27
1959	39	62	110	2	235	41	99	35
1701	40	63	110	1	196	39	97	42
1910	40	64	151	2	305	39	99	48
1571	40	65	145	1	170	45	100	43
2818	40	66	140	2	276	46	100	55
544	40	65	140	1	272	41	91	44
3312	40	65	137	2	315	37	96	99
383	40	67	130	1	300	40	106	52
48	40	62	117	2	290	42	99	42
263	41	62	116	1	320	44	111	61
77	41	68	215	2	255	43	105	45
564	41	64	125	1	306	45	98	62
2709	41	69	170	2	324	40	99	55
1715	42	60	105	1	240	41	101	51
1896	42	63	129	2	210	40	100	46
1326	43	66	167	1	210	40	100	52
1965	43	68	145	2	250	36	98	42
1317	43	66	138	1	335	44	105	58
1837	43	66	132	2	230	42	98	48
2220	43	64	125	1	285	45	105	50
2437	43	62	113	2	200	40	93	36
221	43	64	126	1	280	45	106	38
3286	43	65	148	2	276	41	105	50
520	55	64	124	1	275	40	98	53
22	55	64	165	2	298	36	100	63
59	44	62	118	1	253	43	94	44
2497	44	63	133	2	242	47	104	49
56	45	67	180	1	160	38	97	59
15	45	65	140	2	263	45	107	52
5	46	66		1	250	41	101	73
850	46	67	145	2	320	40	101	37
1218	46	63	138	1	257	40	90	61
132	46	62	118	2	190	38	95	43
695	46	62	103	1	230	43	102	33
541	46	65	190	2	265	41	108	85
473	47	67	135	1	297	42	100	45
2657	47	67	143	2	255	41	100	40
164	47	61	132	1	257	39	96	38
2865	47	59	94	2	257	41	103	53
554	48	62	120	1	300	39	94	51
644	48	66	143	2	225	40	100	62
952	48	67	143	1	216	40	96	47
629	48	65	134	2	248	42	102	42
1241	48	65	164	1	306	44	100	78
156	48	66	120	2	235	36	97	35
1251	48	60	125	1	195	41	95	53
64	48	64	138	2	338	37	100	58
2102	49	64	126	1	255	41	102	48
1932	49	69	158	2	217	36	106	65
44	50	69	135	1	295	43	105	63
3134	50	66	140	2	390	46	97	55
1250	52	68	150	1	250	42	97	59
1789	52	62	107	2	265	46	104	64
1149	53		178	1	227	39		50
575	53	65	140	2	220	40	107	46
2271	54	66	158	1	305	42	103	48
39	54	60	170	2	220	35	88	63

continued

continued

[1] Two values of cholesterol (cases 2 and 4) have been deliberately modified to represent extreme values to be excluded from the analysis. Eight other values were left blank to represent missing data. Each case is recorded in columns 1 to 36 of a card; each variable uses 4 columns. The appropriate format to read the data is (A4,5F4.0,3F4.1).

[2] Birth control pill user: 1 no pill; 2 pill

[3] Albumin, calcium and uric acid are recorded to the nearest 1/10 of a unit. Other measurements are to the nearest unit.

DATA SAMPLE2

AGE	FUNCTION	ACTIVE	INFARCT	ANGINA	HIGHBP
42	2	2	1	1	0
66	2	2	1	1	0
56	2	0	1	1	0
55	2	2	1	1	0
41	2	2	1	1	1
62	0	0	1	0	1
46	2	2	1	1	1
44	2	1	0	1	1
50	1	2	0	1	1
73	3	3	0	1	0
48	2	2	1	1	0
53	2	2	1	1	0
51	3	1	1	1	1
59	0	0	0	1	1
54	3	3	1	1	1
41	2	2	1	1	1
56	2	2	1	0	1
38	0	2	0	1	1
40	3	3	1	1	0
42	1	2	1	1	0
51	1	2	0	1	0
52	1	0	1	1	0
37	0	1	0	1	0
48	1	2	1	0	0
35	0	0	1	1	0
35	1	1	1	0	0
48	3	3	0	1	1
52	2	2	0	1	1
46	2	3	0	1	1
51	3	0	0	1	0
50	0	0	0	1	0
72	3	3	0	1	1
56	3	3	1	1	0
56	3	3	1	1	0
63	2	2	1	1	0
53	1	2	1	1	0
53	0	1	1	0	0
57	3	2	0	1	0
57	1	2	0	1	1
62	2	2	1	1	0
73	2	2	0	1	0
44	2	0	0	1	0
63	3	2	1	1	0
59	1	1	0	1	0
51	1	0	1	0	0
52	3	0	1	1	0
64	0	0	1	0	0
53	2	2	1	0	0
58	1	2	0	1	0
53	0	2	1	1	1
58	2	2	0	1	0
45	1	2	1	1	1
42	3	2	0	1	0
60	2	2	0	1	1
34	1	0	1	0	1
64	2	2	1	1	0
35	1	2	1	0	1
42	2	2	0	1	0
53	2	2	1	0	0
58	1	2	1	0	1
38	1	2	0	1	0
35	2	2	1	1	0
34	2	3	1	1	0
68	3	2	1	1	0
49	3	2	0	1	0
55	2	2	0	1	1
58	0	2	1	1	0
43	2	2	1	1	0
39	2	1	0	1	0
66	3	3	0	1	1
50	2	2	1	1	0
45	3	3	0	1	0
53	0	0	0	1	0
56	3	3	1	1	1
49	2	2	1	1	1
49	0	0	1	0	0
56	2	2	1	1	0
38	0	0	1	1	0
39	0	0	1	1	0
62	2	2	1	1	0
70	3	2	1	1	1
53	2	0	1	1	1
68	2	2	1	1	0
50	2	0	0	0	1
46	2	2	0	1	0
58	3	2	1	1	0
57	2	2	0	0	0
55	3	3	1	0	0
52	0	0	1	1	0
61	2	2	0	0	0
45	2	2	0	1	0
51	2	0	1	1	0
55	3	3	1	1	1
51	1	0	0	1	0
46	1	0	1	1	0
69	1	2	0	1	0
51	3	3	1	1	1
49	1	1	0	1	1
58	3	3	1	1	0
38	3	3	1	1	0
50	1	1	1	0	0
38	1	3	1	1	0
58	1	0	1	0	0
69	0	0	0	1	0
66	0	0	0	1	0
49	2	2	0	1	0
62	0	0	1	1	0
44	0	0	0	1	1
58	3	0	1	1	0
45	2	2	1	1	0
58	3	3	1	1	1
54	2	2	1	1	0
55	2	2	1	1	1
68	2	2	1	1	0
68	2	2	1	1	0
47	1	2	1	1	1
55	0	0	0	1	0

Key: AGE (in years)

FUNCTION -- functional class: 0 none, 1 minimal, 2 moderate, 3 more than moderate (3 is a combination of two groups)

ACTIVE: 0 unknown, 1 very, 2 normal, 3 limited

INFARCT -- history of past myocardial infarctions: 0 none, 1 present

ANGINA -- history of angina pectoris: 0 none, 1 present

HIGHBP -- history of high blood pressure: 0 none, 1 present

AGE is recorded in columns 3 and 4; the other variables are recorded in columns 8, 12, 16, 20 and 24.

DATA SAMPLE3

TREATMENT	DISEASE	SYSINCR
1	1	42
1	1	44
1	1	36
1	1	13
1	1	19
1	1	22
1	2	33
1	2	26
1	2	33
1	2	21
1	3	31
1	3	-3
1	3	25
1	3	25
(cont'd)		

1	3	24
2	1	28
2	1	23
2	1	34
2	1	42
2	1	13
2	2	34
2	2	33
2	2	31
2	2	36
2	3	3
2	3	26
2	3	29
2	3	32
(cont'd)		

2	3	4
2	3	16
3	1	1
3	1	29
3	1	19
3	2	11
3	2	9
3	2	7
3	2	1
3	2	-6
3	3	21
3	3	1
3	3	9
3	3	3
(cont'd)		

4	1	24
4	1	9
4	1	22
4	1	-2
4	1	15
4	2	27
4	2	12
4	2	12
4	2	-5
4	2	16
4	2	15
4	3	22
4	3	7
4	3	25
4	3	5
4	3	12

Note: TREATMNT corresponds to drugs and is coded 1 to 4. DISEASE is coded 1 to 3. SYSINCR is the increase in systolic pressure (mmHg) due to treatment. TREATMNT is recorded in column 3, DISEASE in column 6 and SYSINCR in columns 8-9.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

DATA SAMPLE4 Data from Jarvik's smoking questionnaire, administered to 110 subjects (Dr. M.E. Jarvik, unpublished)

```
3 2 1 3 2 2 1 3 2 2 3 2
4 2 5 3 5 4 5 4 3 4 4 5
5 3 4 4 5 5 4 5 3 4 3 4
4 2 4 3 5 4 4 4 3 4 3 5
4 2 4 3 4 2 4 4 2 4 3 4
3 2 3 1 3 2 3 3 2 3 3 4
4 2 4 2 4 3 3 3 2 3 3 4
3 2 3 2 4 2 3 3 2 3 3 4
3 2 3 2 4 4 3 3 2 3 3 4
3 2 4 1 4 3 4 3 2 3 3 4
5 5 1 4 1 4 1 5 4 5 5 1
1 1 2 3 3 3 3 2 1 3 1 2
2 1 3 4 4 3 4 2 1 4 1 4
2 1 3 4 4 2 3 3 1 4 1 4
2 1 4 3 4 2 4 2 1 3 1 4
3 1 3 4 4 2 4 3 1 4 1 4
3 2 4 4 4 2 4 3 3 4 2 4
2 2 4 3 4 2 4 3 3 4 2 5
3 2 4 4 5 2 5 3 3 4 2 5
3 2 4 4 5 3 5 3 3 3 2 5
3 3 5 4 5 2 4 3 2 4 2 5
3 1 3 3 4 2 2 3 2 4 2 4
3 1 3 1 3 2 3 2 2 1 1 3
3 2 4 2 4 2 4 3 2 2 3 4
3 2 5 2 4 2 5 2 2 3 2 5
3 2 5 2 5 3 5 2 2 3 3 5
4 3 5 3 5 4 5 4 3 3 3 5
3 2 5 2 5 2 5 3 2 3 3 5
4 3 5 3 5 3 5 4 3 3 3 5
3 2 5 1 5 3 5 3 2 3 3 5
3 2 5 2 5 3 5 3 2 3 3 5
4 2 5 3 5 3 5 4 2 3 3 5
2 1 5 2 5 1 5 3 1 3 2 5
1 2 2 2 2 2 1 3 2 3 3 2
3 2 4 3 4 2 4 2 3 3 2 4
2 2 3 2 4 2 4 2 2 3 2 4
3 2 2 3 4 2 4 3 2 4 2 4
```

cont'd

```
4 4 5 4 5 4 5 4 4 4 4 5
2 2 3 2 2 2 3 2 2 2 2 3
2 2 3 2 3 2 4 2 2 3 2 3
2 2 3 2 4 3 3 3 2 3 3 3
4 4 4 3 4 4 4 4 3 4 4 4
2 2 3 2 3 2 3 2 2 2 2 3
4 2 1 2 2 3 1 4 2 3 2 1
2 1 2 3 2 2 1 3 2 3 3 1
4 3 1 3 1 3 1 3 3 3 3 1
2 2 4 3 4 3 4 3 3 3 3 4
4 2 4 3 4 3 4 3 2 3 3 4
2 3 4 2 4 3 4 2 3 3 3 4
4 4 3 3 3 3 3 4 3 3 3 3
4 4 4 2 3 4 2 3 3 3 3 3
2 4 3 3 2 4 3 4 3 3 4 3
3 4 2 3 2 3 2 3 3 3 3 2
3 3 3 2 2 4 3 2 3 3 3 2
3 4 2 3 3 4 2 3 3 3 3 2
2 1 2 2 3 2 2 3 2 2 2 2
3 2 2 3 3 3 2 3 3 4 2 2
3 2 4 2 5 3 5 3 3 3 2 5
3 2 5 3 5 3 5 3 3 4 2 5
3 3 5 5 5 3 5 4 3 5 3 5
2 2 2 1 3 2 3 2 2 1 2 1
3 2 4 2 4 4 3 3 3 2 3 4
3 3 4 3 4 4 4 3 3 3 3 4
3 3 3 5 3 3 3 4 3 5 3 3
3 2 3 3 2 3 2 3 2 4 3 3
3 2 2 2 2 3 2 3 2 3 2 3
2 2 3 3 2 2 2 2 2 2 2 3
2 4 5 4 4 2 5 3 2 4 3 5
2 1 3 1 3 2 4 2 1 1 2 3
1 2 5 2 4 2 5 2 3 3 3 3
2 2 5 4 4 2 4 4 2 5 3 3
3 2 4 4 4 1 4 3 2 4 2 5
1 1 3 1 3 2 3 1 2 2 2 3
2 2 3 2 4 2 3 1 3 2 3 2
```

cont'd

```
1 1 2 3 2 1 2 3 1 4 2 1
2 1 3 3 4 2 3 3 1 3 2 4
2 2 2 2 2 2 1 2 2 2 2 1
1 1 3 1 3 1 3 1 1 2 2 3
2 1 3 2 3 1 3 1 1 2 1 3
1 1 3 1 3 1 3 1 1 1 1 3
1 1 4 1 3 1 4 1 1 1 1 4
1 1 4 3 4 1 4 1 1 3 2 4
1 1 3 1 3 1 3 1 1 1 1 3
1 1 4 1 4 1 5 1 1 1 2 4
1 1 1 2 4 1 5 1 1 2 2 5
1 1 4 3 3 1 4 2 1 4 2 4
1 1 3 3 3 1 3 1 1 3 2 2
1 1 3 2 3 1 2 1 1 2 2 1
2 2 2 2 3 2 3 2 2 2 2 2
2 2 4 3 4 3 4 3 3 4 3 4
4 4 4 4 5 4 4 4 3 4 4 5
4 5 4 3 5 4 4 4 4 4 4 4
4 4 4 4 5 4 4 4 3 4 4 5
2 3 3 2 4 3 3 3 3 3 3 3
4 3 3 4 4 3 3 4 3 4 3 4
4 4 4 3 5 4 5 5 4 4 4 5
4 3 4 4 5 3 4 5 3 5 3 4
4 3 4 3 4 4 4 4 3 4 4 5
5 1 2 4 3 1 2 5 2 4 2 3
2 2 1 1 2 2 1 2 2 3 2 1
5 2 1 3 1 1 2 4 2 4 1 1
2 2 2 1 2 1 3 2 2 2 1 2
2 2 4 3 3 1 4 2 2 3 1 5
3 2 5 4 4 2 4 2 2 4 2 5
1 1 4 1 3 2 3 2 2 2 2 2
2 1 4 1 4 2 4 2 2 2 2 4
2 2 3 2 3 2 4 2 2 2 2 4
4 1 1 4 2 1 2 4 2 4 2 2
4 2 3 3 3 2 3 3 2 3 2 4
2 2 3 1 4 2 2 2 2 3 2 1
```

cont'd

The data consist of the answers to 12 questions - each coded 1 to 5 such that a high score represents a desire to smoke. The 12 questions (and their NAMEs) are concentration (CONCENTR), annoyance (ANNOY), desire to smoke - first wording (SMOKING1), sleepiness (SLEEPY), desire to smoke - second wording (SMOKING2), tenseness (TENSE), desire to smoke - third wording (SMOKING3), alertness (ALERT), irritability (IRRITABL), tiredness (TIRED), contentedness (CONTENT), desire to smoke - fourth wording (SMOKING4). The data are recorded in columns 2, 4, 6, etc. and are read using the format (12F2.0).

DATA SAMPLE5 Fisher iris data (Fisher, 1936). Length and width of sepals (SEPALLEN and SEPALWID) and petals (PETALLEN **and** PETALWID) on 50 flowers from each of three types of iris (IRISTYPE).

1	2	3	4	5
50	33	14	02	1
64	28	56	22	3
65	28	46	15	2
67	31	56	24	3
63	28	51	15	3
46	34	14	03	1
69	31	51	23	3
62	22	45	15	2
59	32	48	18	2
46	36	10	02	1
61	30	46	14	2
60	27	51	16	2
65	30	52	20	3
56	25	39	11	2
65	30	55	18	3
58	27	51	19	3
68	32	59	23	3
51	33	17	05	1
57	28	45	13	2
62	34	54	23	3
77	38	67	22	3
63	33	47	16	2
67	33	57	25	3
76	30	66	21	3
49	25	45	17	3
55	35	13	02	1
67	30	52	23	3
70	32	47	14	2
64	32	45	15	2
61	28	40	13	2
48	31	16	02	1
59	30	51	18	3
55	24	38	11	2
63	25	50	19	3
64	32	53	23	3
52	34	14	02	1
49	36	14	01	1
54	30	45	15	2
79	38	64	20	3
44	32	13	02	1
67	33	57	21	3
50	35	16	06	1
58	26	40	12	2
44	30	13	02	1
77	28	67	20	3
63	27	49	18	3
47	32	16	02	1
55	26	44	12	2
50	23	33	10	2
72	32	60	18	3

cont'd

48	30	14	03	1
51	38	16	02	1
61	30	49	18	3
48	34	19	02	1
50	30	16	02	1
50	32	12	02	1
61	26	56	14	3
64	28	56	21	3
43	30	11	01	1
58	40	12	02	1
51	38	19	04	1
67	31	44	14	2
62	28	48	18	3
49	30	14	02	1
51	35	14	02	1
56	30	45	15	2
58	27	41	10	2
50	34	16	04	1
46	32	14	02	1
60	29	45	15	2
57	26	35	10	2
57	44	15	04	1
50	36	14	02	1
77	30	61	23	3
63	34	56	24	3
58	27	51	19	3
57	29	42	13	2
72	30	58	16	3
54	34	15	04	1
52	41	15	01	1
71	30	59	21	3
64	31	55	18	3
60	30	48	18	3
63	29	56	18	3
49	24	33	10	2
56	27	42	13	2
57	30	42	12	2
55	42	14	02	1
49	31	15	02	1
77	26	69	23	3
60	22	50	15	3
54	39	17	04	1
66	29	46	13	2
52	27	39	14	2
60	34	45	16	2
50	34	15	02	1
44	29	14	02	1
50	20	35	10	2
55	24	37	10	2
58	27	39	12	2

cont'd

47	32	13	02	1
46	31	15	02	1
69	32	57	23	3
62	29	43	13	2
74	28	61	19	3
59	30	42	15	2
51	34	15	02	1
50	35	13	03	1
56	28	49	20	3
60	22	40	10	2
73	29	63	18	3
67	25	58	18	3
49	31	15	01	1
67	31	47	15	2
63	23	44	13	2
54	37	15	02	1
56	30	41	13	2
63	25	49	15	2
61	28	47	12	2
64	29	43	13	2
51	25	30	11	2
57	28	41	13	2
65	30	58	22	3
69	31	54	21	3
54	39	13	04	1
51	35	14	03	1
72	36	61	25	3
65	32	51	20	3
61	29	47	14	2
56	29	36	13	2
69	31	49	15	2
64	27	53	19	3
68	30	55	21	3
55	25	40	13	2
48	34	16	02	1
48	30	14	01	1
45	23	13	03	1
57	25	50	20	3
57	38	17	03	1
51	38	15	03	1
55	23	40	13	2
66	30	44	14	2
68	28	48	14	2
54	34	17	02	1
51	37	15	04	1
52	35	15	02	1
58	28	51	24	3
67	30	50	17	2
63	33	60	25	3
53	37	15	02	1

Key: 1 - SEPALLEN, 2 - SEPALWID, 3 - PETALLEN, 4 - PETALWID, 5 - IRISTYPE

The three groups of iris (IRISTYPE) are Setosa, Versicolor and Virginica. All variables but IRISTYPE are recorded to 1/10 of a centimeter. The data are recorded in columns 2-3, 5-6, 8-9, 11-12 and 15.

ตารางที่ 1.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT1 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//TESTBMDP  JOB (9012Ø,9Ø,9Ø87),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
(ก) INCLUDE BMDP(BMDP1D)
    ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM    TITLE IS 'DESCRIPTIVE STATISTICS OF BMDP BY SIRICHAI'.
/INPUT      VARIABLES ARE 9.
            FORMAT IS '(A4,5F4.0,3F4.1)'.
/VARIABLES  NAMES ARE ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHOLSTRL,
                      ALBUMIN,CALCIUM,URICACID.
            MAXIMUM IS (6) 400.
            MINIMUM IS (6) 150.
            BLANKS ARE MISSING.
            LABEL IS ID.
/END
```

บัตรบันทึกข้อมูล

```
//
```

(ข) สำหรับบัตรคำสั่งในโปรแกรมชุด P2D จะใช้เหมือนกับโปรแกรมชุด P1D เพียงแต่ระบุชุดโปรแกรม P2D เท่านั้นที่ต่างกัน

ตารางที่ 1.1 (ต่อ)

```
//TESTBMDP JOB (90120,90,9087),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
 INCLUDE BMDP(BMDP3D)
  ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM   TITLE IS 'DESCRIPTIVE STATISTICS OF BMDP BY SIRICHAI'.
/INPUT     VARIABLES ARE 9.
           FORMAT IS '(A4,5F4.0,3F4.1)'.
/VARIABLE  NAMES ARE ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHOLSTRL,
                     ALBUMIN,CALCIUM,URICACID.
           MAXIMUM IS (6) 400.
           MINIMUM IS (6) 150.
           BLANKS ARE MISSING.
           LABEL IS ID.
           GROUPING IS BRTHPILL.
/GROUP     CODES(5) ARE 1, 2.
           NAMES(5) ARE NOPILL,PILL.
/END
```

บัตรบันทึกข้อมูล

```
//
```

รูปที่ 1.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT1

--- Control Language read by P1D is printed and interpreted ---

(ก)

```
NUMBER OF CASES READ. . . . . .  188

VARIABLE                STANDARD    ST.ERR.   COEFF. OF   S M A L L E S T     L A R G E S T                TOTAL
NO.  NAME      MEAN     DEVIATION   OF MEAN   VARIATION   VALUE   Z-SCORE     VALUE   Z-SCORE     RANGE        FREQUENCY

  2  AGE       33.819     10.113     0.7376     0.29904    19.000   -1.46528   55.000    2.09449    36.000      188
  3  HEIGHT    64.510      2.485     0.1822     0.03852    57.000   -3.02207   71.000    2.61155    14.000      186
  4  WEIGHT   131.671     20.661     1.5149     0.15691    94.000   -1.82335  215.000    4.03323   121.000      186
  5  BRTHPILL   1.500      0.501    (0.0366     0.33423     1.000   -0.99726    2.000    0.99743     1.000      188
  6  CHOLSTRL 236.150     42.555     3.1203     0.18020   155.000   -1.90693  390.000    3.61529   235.000      186
  7  ALBUMIN    4.111      0.358     0.0262     0.08707     3.200   -2.54565    5.000    2.48281     1.800      186
  8  CALCIUM    9.962      0.480     0.0353     0.04814     8.600   -2.84041   11.100    2.37280     2.500      185
  9  URICACID   4.771      1.157     0.0846     0.24258     2.200   -2.22129    9.900    4.43252     7.700      187
```

(1) (2) (3) (4) (5) (6)

--- Control Language read by P2D is printed and interpreted ---

(ข)

```
THREE LOCATION ESTIMATES ARE PRINTED & PLOTTED FOR COMPARISON.
WHEN ANY ESTIMATES ARE TOO CLOSE TOGETHER FOR BOTH TO APPEAR ON THE
PLOT, THE PRIORITY OF PLOTTING IS THE ORDER IN WHICH THE ESTIMATES
ARE LISTED.

Q1 AND Q3 ARE THE 1ST & 3RD QUARTILES.  S- AND S+ REPRESENT THE
QUANTITIES (MEAN-ST.DEV.) AND (MEAN+ST.DEV.).

  NUMBER OF CASES READ. . . . . .  188
```

--- analysis of variables 2 to 5 ---

(5) (6) (7) (1) (2) (14)

```
VARIABLE NUMBER . . . . . 6                 MAXIMUM      390.0000000        H   H
      NAME . . . . . CHOLSTRL               MINIMUM      155.0000000        H   H
NUMBER OF DISTINCT VALUES .  90             RANGE        235.0000000        H   H
NUMBER OF VALUES COUNTED. .  186            VARIANCE    1810.9609375        H HHH
NUMBER OF VALUES NOT COUNTED  2             ST.DEV.       42.5553894      H HHHHHHH  H
                                            (Q3-Q1)/2     30.0000000      H HHHHHHHH H
LOCATION ESTIMATES                                  ST.ERROR              HHHHHHHHHHH H
                                                                          HHHHHHHHHHHH H H
                                                                          HHHHHHHHHHHHHHHHH
                                                                      MIN--------------------MAX
```

(11) (10) (12)

```
                                                                                    DIV. BY S.E.
MEAN      236.1505280     3.1203117
MEDIAN    235.0000000     3.4641027
MODE      200.0000000
SKEWNESS    0.4703844     2.6189890
KURTOSIS    0.1034966     0.2881222
```

(14)

```
EACH "H" REPRESENTS 2.60 COUNTS

Q1= 200.0000000
Q3= 260.0000000
S-= 193.5951385
S+= 278.7058105
EACH . = 1.8217049
```

(15)

(16)

```
M                S  O                 HH        Q        S                                                M
I.....................O.....................EE.....3.....+.....................................................A
N                   D                  OA                                                               X
                    E                  IN
```

(13)

แสดงค่าประมาณโดยวิธีโรบัสเพิ่มเติมจากการใช้โปรแกรม BMDP ชุด P2D โดยเพิ่มบัตร COUNT paragraph ก่อนบัตร END paragraph ดังนี้

```
/ COUNT  ESTIMate.
```

(ข)

```
VARIABLE NUMBER . . . . . 6          MAXIMUM    390.0000000          H   H                    EACH 'H' '
        NAME . . . . . CHOLSTRL      MINIMUM    155.0000000          H   H                    REPRESENTS
NUMBER OF DISTINCT VALUES .  30      RANGE      235.0000000          H   H                      2.60
NUMBER OF VALUES COUNTED. .  186     VARIANCE  1810.9609375          H  HHH                    COUNTS
NUMBER OF VALUES NOT COUNTED  2      ST.DEV.     42.5553894          H HHHHH
                                     IQ3-Q1)/2   30.0000000        H HHHHHHH  H
LOCATION ESTIMATES                        ST.ERROR                 H HHHHHHH  H
                MEAN          236.1505280    3.1203117           HHHHHHHHHH H
                MEDIAN        235.0000000    3.4641027          HHHHHHHHHHHH H H
                MODE          200.0000000                       HHHHHHHHHHHHHHHHH
                                                          MIN----------------MAX

(17) SOME NEW LOCATION ESTIMATES                                                        Q1=  200.0000000
                HAMPEL        233.6557312                                 DIV. BY S.E.  Q3=  260.0000000
                TRIMMED       233.5522156   SKEWNESS      0.4703844       2.6189890     S-=  193.5951385
                BIWEIGHT      233.7855072   KURTOSIS      0.1034966       0.2881222     S+=  278.7058105
                                                                                  EACH . =    1.8217049

                  S  Q                        Q       S
N                 -  M                  HHH   3       +                                                  F
I.....................D.....................AEE.....................................................A
N                    D                    MDA                                                        X
                     E                    PIN
```

VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM
155.	1	0.5	0.5	200.	10	5.4	25.8	237.	2	1.1	52.2	272.	1	0.5	82.3
158.	1	0.5	1.1	203.	1	0.5	26.3	238.	2	1.1	53.2	275.	2	1.1	83.3
160.	2	1.1	2.2	204.	1	0.5	26.9	240.	5	2.7	55.9	276.	2	1.1	84.4
165.	1	0.5	2.7	205.	3	1.6	28.5	242.	1	0.5	56.5	280.	4	2.2	86.6
168.	1	0.5	3.2	207.	1	0.5	29.0	243.	2	1.1	57.5	285.	1	0.5	87.1
170.	2	1.1	4.3	208.	1	0.5	29.6	245.	4	2.2	59.7	290.	2	1.1	88.2
172.	1	0.5	4.8	210.	5	2.7	32.3	246.	1	0.5	60.2	295.	3	1.6	89.8
173.	1	0.5	5.4	212.	1	0.5	32.8	247.	3	1.6	61.8	297.	1	0.5	90.3
174.	1	0.5	5.9	215.	3	1.6	34.4	248.	1	0.5	62.4	298.	1	0.5	90.9
175.	4	2.2	8.1	216.	1	0.5	34.9	250.	8	4.3	66.7	300.	2	1.1	91.9
177.	1	0.5	8.6	217.	2	1.1	36.0	253.	1	0.5	67.2	305.	3	1.6	93.5
180.	4	2.2	10.8	218.	1	0.5	36.6	254.	1	0.5	67.7	306.	2	1.1	94.6
185.	2	1.1	11.8	220.	5	2.7	39.2	255.	6	3.2	71.0	315.	1	0.5	95.2
189.	1	0.5	12.4	222.	1	0.5	39.8	257.	3	1.6	72.6	317.	1	0.5	95.7
190.	1	0.5	12.9	223.	1	0.5	40.3	260.	6	3.2	75.8	320.	3	1.6	97.3
192.	3	1.6	14.5	225.	3	1.6	41.9	262.	2	1.1	76.9	324.	1	0.5	97.8
195.	4	2.2	16.7	226.	1	0.5	42.5	263.	3	1.6	78.5	330.	1	0.5	98.4
196.	2	1.1	17.7	227.	2	1.1	43.5	265.	2	1.1	79.6	335.	1	0.5	98.9
197.	2	1.1	18.8	230.	8	4.3	47.8	266.	1	0.5	80.1	338.	1	0.5	99.5
198.	3	1.6	20.4	235.	6	3.2	51.1	270.	3	1.6	81.7	390.	1	0.5	100.0

--- Control Language read by P3D is printed and interpreted --- (18)

```
   TEST TITLE. . . . . .WERNER BLOOD CHEMISTRY DATA
   INDEXES OF VARIABLES TO BE ANALYZED . . . . . 2  3  4  5  7  8  9
   USE COMPLETE CASES ONLY . . . . . NO
   PRINT GROUP CORRELATION MATRICES. . . . . NO
   COMPUTE HOTELLINGS T SQUARE . . . . . NO
   INDEX OF GROUPING VARIABLE. . . . . 5
   GROUPS USED IN COMPUTATIONS . . . . . 1  2

PROGRAM NOTES

1. THE T(POOLED) STATISTIC TESTS THE HYPOTHESIS THAT THE
MEANS FOR THE TWO GROUPS ARE EQUAL ASSUMING THAT THE
VARIANCES FOR THE TWO GROUPS ARE EQUAL.

2. THE T(SEPARATE) STATISTIC TESTS THE HYPOTHESIS THAT THE
MEANS FOR THE TWO GROUPS ARE EQUAL WITHOUT THE ASSUMPTION
OF EQUAL VARIANCES.  SINCE THERE IS LITTLE LOSS IN USING
THE SEPARATE VARIANCE T EVEN WHEN THE POOLED T MAY BE
APPROPRIATE AND SINCE THE ERROR IN USING THE POOLED T WHEN
VARIANCES ARE NOT EQUAL CAN BE VERY SERIOUS, THE SEPARATE
VARIANCE T SHOULD BE USED IN MOST APPLICATIONS.

3. THE F(FOR VARIANCES) STATISTIC TESTS THE HYPOTHESIS THAT
THE VARIANCES FOR THE TWO GROUPS ARE EQUAL.

4. THE P VALUES ARE THE TWO-TAIL SIGNIFICANCE LEVELS.

5. FOR ONE SAMPLE TESTS, THE F(FOR VARIANCES) SHOULD BE
IGNORED AND THE T(SEPARATE) AND T(POOLED) ARE THE SAME.

   NUMBER OF CASES READ. . . . . 188

                    BEFORE TRANSFORMATION                                    INTERVAL RANGE
VARIABLE            MINIMUM     MAXIMUM     MISSING     CATEGORY  CATEGORY   GREATER     LESS THAN
NO.  NAME           LIMIT       LIMIT       CODE        CODE      NAME       THAN        OR EQUAL TO

 5   BRTHPILL                                           1.00000   NOPILL
                                                        2.00000   PILL

FOR DIFFERENCES ON SINGLE VARIABLES

**********
*AGE     *  VARIABLE NUMBER   2          GROUP      1 NOPILL     2 PILL          1 NOPILL  (N=  94)          2 PILL    (N=  94)
**********                                MEAN         33.8188      33.8188
          STATISTICS     P VALUE   D. F.  STD DEV      10.1400      10.1400        H                          X
                                          S.E.M.        1.0459       1.0459        H                          X
(20) T (SEPARATE)      0.0     1.000  186.0  SAM SIZ      94           94           H    HH   H                X    XX   X
     T (POOLED)        0.0     1.000  186    MAXIMUM      55.0000      55.0000      HHHHHHH   H H HH  H       XXXXXXX   X X XX  X
                                             MINIMUM      19.0000      19.0000      HHHHHHHHHHHHHHHHHHH       XXXXXXXXXXXXXXXXXXX
(19) F(FOR VARIANCES)  1.00    1.000   93.   93                                  MIN--------------------MAX MIN--------------------MAX
                                                                                  AN H =    3.0 CASES       AN X =    3.0 CASES

**********
*HEIGHT  *  VARIABLE NUMBER   3          GROUP      1 NOPILL     2 PILL          1 NOPILL  (N=  92)          2 PILL    (N=  94)
**********                                MEAN         64.0866      64.3252             H                          XX
          STATISTICS     P VALUE   D. F.  STD DEV      2.4924       2.4196        H   H                          XX  X
                                          S.E.M.        0.2599       0.2496             HH HHH H                XX XXX X
T (SEPARATE)     -2.328   0.021  183.5    SAM SIZ      92           94                 HH HHH H                XX XXX XX
T (POOLED)       -2.33    0.021  184      MAXIMUM      69.0000      71.0000          H  HH HHH HH                XX XXX XX X
                                          MINIMUM      57.0000      59.0000      HH  HH HH HHH HH H             X XX XX XXX XX X X
F(FOR VARIANCES)  1.06    0.776   91.   93                                  MIN--------------------MAX MIN--------------------MAX
                                                                                  AN H =    3.0 CASES       AN X =    3.0 CASES
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 1.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT1 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//SØ3ØPØK JØB CLASS=C,MSGLEVEL=(2,0)
//  EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE1;
INPUT ID $ 1-4 AGE 5-8 HEIGHT 9-12 WEIGHT 13-16 BRTHPILL 17-20 CHØL
  CHØLSTRL 21-24 ALBUMIN 25-28 1 CALCIUM 29-31 1 URICACID 33-36 1;
CARDS;
```

บัตรบันทึกข้อมูล

```
PRØC PRINT;
TITLE 'BEFØRE SØRT';
(ก) PRØC MEANS;
TITLE 'CØMPARE BMDP SAS SPSS STATISTICAL PACKAGE PRØGRAMS';
PRØC SØRT;
 BY BRTHPILL;
PRØC PRINT;
TITLE 'AFTER SØRT';
(ข) PRØC MEANS;
 BY BRTHPILL;
TITLE 'DESCRIPTIVE STATISTICS BY BRTHPILL FØR GRØUPING';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 1.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT1

COMPARE EMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS (ก)

VARIABLE	N	MEAN	STANDARD DEV	VARIANCE	SUM	CORRECTED SS	LOW	HIGH	C.V. %
AGE	188	33.819149	10.113694	102.286563	6358.000000	19123.151064	19.000000	55.000000	29.902
HEIGHT	186	64.510753	2.49[illegible]	[illegible].175899	11999.000000	1142.478495	57.000000	71.000000	3.852
WEIGHT	186	131.672043	20.[illegible]0577	42[illegible].859450	24491.000000	7696[illegible].946624	84.000000	215.000000	15.691
BRTHPILL	188	1.500000	0.501331	0.251337	282.000000	47.000000	1.000000	2.000000	33.422
CHOLSTRL	188	237.095745	51.[illegible]6[illegible]7	268[illegible].[illegible]698	44574.000000	5019[illegible].76596	90.000000	000.000000	21.851
ALBUMIN	186	4.111290	0.3[illegible]7[illegible]	0.126142	764.700000	23.706290	3.200000	5.000000	8.707
CALCIUM	185	9.962162	0.479[illegible]	0.229974	1843.000000	42.[illegible]1351	8.600000	11.100000	4.814
URICACID	187	4.773[illegible]8[illegible]	1.1[illegible]7[illegible]31	1.[illegible]39184	892.100000	249.086235	2.200000	9.900000	24.256

(1) (2) (7) (8) (9) (5) (4)

(ข)

DESCRIPTIVE STATISTICS BY BRTHPILL FOR GROUPING
BRTHPILL=1 (18)

VARIABLE	N	MEAN	STANDARD DEV	VARIANCE	SUM	CORRECTED SS	LOW	HIGH	C.V. %
AGE	94	33.[illegible]19149	10.139642	102.816404	3179.000000	9661.925532	19.000000	55.000000	29.983
HEIGHT	92	64.0[illegible]7	2.492416	6.212136	5896.000000	565.304348	57.000000	69.000000	3.889
WEIGHT	93	130.1[illegible]	18.[illegible]9[illegible]	356.716223	12104.000000	32817.892473	95.000000	195.000000	14.512
CHOLSTRL	94	232.96[illegible]	43.4[illegible]2	1891.515100	21899.000000	175910.904255	155.000000	355.000000	18.668
ALBUMIN	92	4.1[illegible]26	0.345132	0.119116	386.200000	10.639565	3.200000	5.000000	8.222
CALCIUM	92	9.98[illegible]13	0.503491	0.253507	919.000000	23.069130	8.600000	11.100000	5.040
URICACID	93	4.741935	1.035177	1.071[illegible]92	441.000000	98.[illegible]6452	2.500000	7.600000	21.830

BRTHPILL=2

AGE	94	33.[illegible]149	10.[illegible]862	102.816404	3179.000000	9661.925532	19.000000	55.000000	29.963
HEIGHT	94	64.[illegible]532	2.[illegible]6	[illegible].804610	6103.000000	544.478723	59.000000	71.000000	3.727
WEIGHT	93	133.1[illegible]	22.[illegible]647	496.[illegible]02132	12387.000000	45720.[illegible]16129	94.000000	215.000000	16.737
CHOLSTRL	94	241.[illegible]2340	58.[illegible]3[illegible]76	3[illegible]76.[illegible]20522	22675.000000	322786.308511	90.000000	600.000000	24.423
ALBUMIN	94	4.0[illegible]4	0.[illegible]51701	[illegible].123694	378.500000	11.[illegible]3511	3.200000	4.700000	8.734
CALCIUM	93	9.93[illegible]4	0.45[illegible]79[illegible]	[illegible].207749	924.000000	19.112903	8.800000	10.800000	4.568
URICACID	94	4.75[illegible]93[illegible]	1.[illegible]71[illegible]4[illegible]	1.[illegible]12[illegible]66	451.100000	150.[illegible]45894	2.200000	9.900000	26.495

ตารางที่ 1.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT1 โดยใช้ข้อมูลจาก DATA SAMPLE1

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK          MR.SIRICHAI PONGWICHAI B2214758
// EXEC PROC=$$SPSS70
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME         'DATA SAMPLE1'
VARIABLE LIST     ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHOLSTRL,ALBUMIN,
                  CALCIUM,URICACID
INPUT MEDIUM      CARD
N OF CASES        188
INPUT FORMAT      FIXED(A4,5F4.0,3F4.1)
MISSING VALUES    AGE TO URICACID(0)
VALUE LABELS      BRTHPILL(1)NOPILL(2)PILL
RECODE            CHOLSTRL(LO THRU 140=0)(401 THRU HI=0)
(ก) TASK NAME     'COMPARE STAT1(CONDESCRIPTIVE STATISTICS)' ก
CONDESCRIPTIVE    ALL
STATISTICS        ALL
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
(ข) TASK NAME     'COMPARE DESCRIPTIVE STAT BY BRTHPILL FOR GROUPING'
BREAKDOWN         TABLES = AGE HEIGHT WEIGHT CHOLSTRL TO URICACID BY
                  BRTHPILL
OPTIONS           2,4
```

ตารางที่ 1.3 (ต่อ)

```
(ก) TASK NAME          COMPARE T-TEST
    T-TEST             GROUPS = BRTHPILL(1, 2)/VARIABLES = AGE HEIGHT WEIGHT
                       CHOLSTRL TO URICACID
    FINISH
    /*
    /&
    * $$ EOJ
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 1.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT1

VARIABLE AGE

MEAN	33.019 (1)	STD ERROR	0.738 (3)	STD DEV	10.113 (2)
VARIANCE	102.267 (7)	(13) KURTOSIS	-0.981	(12) SKEWNESS	0.411
RANGE	36.000 (6)	MINIMUM	19.000	MAXIMUM	55.000
SUM	6357.996		(5)		(5)

VALID OBSERVATIONS - (8) 198 MISSING OBSERVATIONS - 0 (n)

VARIABLE HEIGHT

MEAN	64.911	STD ERROR	0.182	STD DEV	2.485
VARIANCE	6.176	KURTOSIS	-0.023	SKEWNESS	-0.137
RANGE	14.000	MINIMUM	57.000	MAXIMUM	71.000
SUM	11998.996				

VALID OBSERVATIONS - 186 MISSING OBSERVATIONS - 2

---------- ตัวแปร WEIGHT, BRTHPILL, CHOLSTRL, ALBUMIN จะให้ผลเช่นเดียวกัน ----------

---

VARIABLE CALCIUM

MEAN	9.962	STD ERROR	0.035	STD DEV	0.480
VARIANCE	0.230	KURTOSIS	-0.363	SKEWNESS	-0.114
RANGE	2.500	MINIMUM	8.600	MAXIMUM	11.100
SUM	1843.000				

VALID OBSERVATIONS - 185 MISSING OBSERVATIONS - 3

VARIABLE URICACID

MEAN	4.771	STD ERROR	0.085	STD DEV	1.157
VARIANCE	1.339	KURTOSIS	2.499	SKEWNESS	1.076
RANGE	7.700	MINIMUM	2.200	MAXIMUM	9.900
SUM	892.100				

VALID OBSERVATIONS - 187 MISSING OBSERVATIONS - 1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(ข)

DESCRIPTION OF SUBPOPULATIONS

CRITERION VARIABLE AGE
BROKEN DOWN BY BRTHPILL (18)

FOR ENTIRE POPULATION
SUM 6358.000
MEAN 33.819
STD DEV 10.113
VARIANCE 102.267
N (188)

VARIABLE BRTHPILL

CODE 1.
NOPILL
SUM 3179.000
MEAN 33.819
STD DEV 10.140
VARIANCE 102.816
N (94)

CODE 2.
PILL
SUM 3179.000
MEAN 33.819
STD DEV 10.140
VARIANCE 102.816
N (94)

TOTAL CASES = 188

CRITERION VARIABLE URICACID
BROKEN DOWN BY BRTHPILL

FOR ENTIRE POPULATION
SUM 892.100
MEAN 4.771
STD DEV 1.157
VARIANCE 1.339
N (187)

VARIABLE BRTHPILL

CODE 1.
NOPILL
SUM 441.000
MEAN 4.742
STD DEV 1.035
VARIANCE 1.072
N (93)

CODE 2.
PILL
SUM 451.100
MEAN 4.799
STD DEV 1.271
VARIANCE 1.617
N (94)

TOTAL CASES = 188
MISSING CASES = 1 OR 0.5 PCT

ตัวแปร WEIGHT ถึงตัวแปร CALCIUM ให้ผลเช่นเดียวกัน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(A)

T - T E S T

GROUP 1 - BRTHPILL EQ 1.
GROUP 2 - BRTHPILL EQ 2.

(19) (20)

VARIABLE	NUMBER OF CASES	MEAN	STANDARD DEVIATION	STANDARD ERROR	F VALUE	2-TAIL PROB.	POOLED VARIANCE ESTIMATE T VALUE	DEGREES OF FREEDOM	2-TAIL PROB.	SEPARATE VARIANCE ESTIMATE T VALUE	DEGREES OF FREEDOM	2-TAIL PROB.
AGE												
GROUP 1	94	33.8191	10.140	1.046	1.00	1.000	0.0	186	1.000	0.0	186.00	1.000
GROUP 2	94	33.8191	10.140	1.046								
HEIGHT												
GROUP 1	92	64.0869	2.492	0.260	1.06	0.776	-2.33	184	0.021	-2.33	183.52	0.021
GROUP 2	94	64.9255	2.420	0.250								
WEIGHT												
GROUP 1	93	130.1505	18.887	1.958	1.39	0.114	-1.00	184	0.317	-1.00	179.16	0.317
GROUP 2	93	133.1935	22.293	2.312								
CHOLSTRL												
GROUP 1	94	232.9681	43.492	4.486	1.10	0.665	-1.03	184	0.304	-1.03	183.90	0.304
GROUP 2	92	239.4022	41.562	4.333								
ALBUMIN												
GROUP 1	92	4.1978	0.346	0.036	1.04	0.859	3.35	184	0.001	3.35	184.00	0.001
GROUP 2	94	4.0266	0.352	0.036								
CALCIUM												
GROUP 1	92	9.9891	0.504	0.053	1.22	0.343	0.76	183	0.449	0.76	180.81	0.449
GROUP 2	93	9.9354	0.456	0.047								
URICACID												
GROUP 1	93	4.7419	1.035	0.107	1.51	0.050	-0.34	185	0.737	-0.34	178.41	0.737
GROUP 2	94	4.7989	1.272	0.131								

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

STAT 2 การแจกแจงความถี่และการวัดความสัมพันธ์ระหว่างตัวแปร (FREQUENCY DISTRIBUTIØN AND RELATED MEASURES ØF ASSØCIATIØN)

ตารางที่ 2 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT2

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓		✓	
② การแจกแจงความถี่แบบทางเดียว	✓	✓	✓	
③ ฮิสโตแกรมของความถี่	✓		✓	
④ ตารางแจกแจงความถี่แบบสองทาง	✓	✓	✓	
⑤ ค่าสถิติสำหรับการทดสอบนัยสำคัญต่างๆดังนี้				
. Chi-square (X^2)	✓	✓	✓	
. Likelihood ratio chi-square (G^2)	✓			
. Phi, Gramer'sV (Ø)	✓		✓	
. Contingency coefficient(C)	✓		✓	
. Fisher's test	✓		✓	
. Yate's corrected X^2	✓		✓	
. Yule's Q, Y	✓			
. Cross-product ratio	✓			
. Tetrachoric correlation ($\mathcal{R}$ t)	✓			
. Gamma Γ	✓		✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
. Somer's D	✓		✓	
. Pearson correlation (ρ)	✓	✓	✓	
. Spearman rank correlation (ρ_s)	✓	✓		
. Kendall's τ b	✓	✓	✓	
. Stuart's τ c	✓		✓	
. Optimal prediction λ	✓		✓	
. Optimal prediction λ^*	✓		✓	
. Uncertainty Coefficient	✓		✓	
. Eta			✓	
(6) การเพิ่มค่าจริงให้กับค่าสังเกต- -เมื่อค่าที่คาดหวังมีค่าต่ำเกินไป	✓			
(7) เมื่อไม่ต้องการเพิ่มค่าให้ความถี่- -อาจรวมกลุ่มของตัวแปรที่มีความถี่- -น้อยเป็นกลุ่มเดียวกัน	✓			
(8) ตารางแจกแจงความถี่มากกว่า- -5 ทาง	✓		✓	
(9) ตารางแจกแจงความถี่มากกว่า- -7 ทาง			✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT2 โดยใช้ข้อมูลจาก DATA SAMPLE 2

```
//TESTBMDP  JOB (90120,90,9087),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
   INCLUDE BMDP(BMDP1F)
   ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM    TITLE IS 'FREQ. OF BMDP BY SIRICHAI'.
/INPUT      VARIABLE ARE 6.
            FORMAT IS '(6F4.0)'.
/VARIABLE   NAMES ARE FUNCTION,ACTIVE,INFARCT,ANGINA,HIGHBP.
/TABLE      COLUMN IS FUNCTION,INFARCT.
            ROW    IS ACTIVE  ,ANGINA.
            (n) MINIMUM IS 1.
/CATEGORY   CODES(2) ARE 0,  1,  2,  3.
            NAMES(2) ARE NONE,MINIMAL,MODERATE,SEVERE.
            CODES(3) ARE 1,  2,  3.
            NAMES(3) ARE VERY,NORMAL,LIMITED.
            CODES(4) ARE 0,  1.
            NAMES(4) ARE NONE,PRESENT.
            CODES(5) ARE 0,  1.
            NAMES(5) ARE NONE,PRESENT.
/PRINT      EXPECTED.
            ROWPERCENT.
```

ตารางที่ 2.1 (ต่อ)

```
/STATISTICS   COLPERCENT.
              TOTPERCENT.
              LRCHI.
              FISHER.
              CONTINGENCY.
              TETRACHORIC.
              GAMMA.
              CORRELATION.
              SPEARMAN.
              TAUS.
              LAMBDA.
              LSTAR.
              UNCERTAINTY.
/END
```

บัตรบันทึกข้อมูล

```
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 2.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT2

--- Control Language read by P1F is printed and interpreted ---

NUMBER OF CASES READ. 117

(1)

VARIABLE NO.	NAME	MEAN	FREQ	S.D.	SMALLEST	LARGEST
1	AGE	52.32	117	9.47	36.00	73.00
2	FUNCTION	1.68	117	1.01	0.0	3.00
3	ACTIVE	1.59	117	1.04	0.0	3.00
4	INFARCT	0.63	117	0.48	0.0	1.00
5	ANGINA	0.84	117	0.37	0.0	1.00
6	HIGHBP	0.31	117	0.46	0.0	1.00

TABLE NO. 1 ACTIVE VARI 3) VS FUNCTIONVARI 2)

(4) CELL FREQUENCY COUNTS

			FUNCTIONIVAR 2)				
		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00	TOTAL
ACTIVE (VAR 3)	VERY	1.00	2	4	2	1	9
	NORMAL	2.00	3	13	37	7	60
	LIMITED	3.00	0	1	2	16	19
		TOTAL	5	18	41	24	88
	NOT	COUNTED	15	6	5	3	

(5) MINIMUM NONZERO EXPECTED VALUE IS 0.511
—— ALL STATISTICS ARE COMPUTED USING THE ORIGINAL CELL FREQUENCIES
EXCLUDING ROWS AND/OR COLUMNS WHICH ARE ZERO.

STATISTIC	VALUE	D.F.	PROB.	STATISTIC	VALUE	D.F.	PROB.
PEARSON CHISQUARE	48.365	6	0.0000	LIKELIHOOD RATIO CHISQ.	43.906	6	0.0000
PHI	0.741						
CONTINGENCY COEF. C	0.596			CRAMER'S V	0.524		

STATISTIC	VALUE	ASE1	T-VALUE	DEP.	STATISTIC	VALUE	ASE1	T-VALUE	DEP.
GAMMA	0.785	0.100	5.368		STUART'S TAU-C	0.453	0.084	5.368	
					KENDALL'S TAU-B	0.536	0.086	5.368	
SOMER'S D	0.632	0.097	5.368	2	SOMER'S D	0.455	0.082	5.368	3
PRODUCT MOMENT CORRELATION	0.551	0.087	4.538		SPEARMAN RANK CORRELATION	0.574	0.089	5.537	
					LAMBDA-SYMMETRIC	0.333	0.096	2.948	
LAMBDA-ASYMMETRIC	0.340	0.085	3.484	2	LAMBDA-ASYMMETRIC	0.321	0.141	1.915	3
LAMBDA-STAR-ASYMMETRIC	0.432	0.109	3.477	2	LAMBDA-STAR-ASYMMETRIC	0.253	0.117	2.032	3
TAU-ASYMMETRIC	0.231	0.060	3.936	2	TAU-ASYMMETRIC	0.317	0.089	3.273	3
					UNCERTAINTY-NORMED	0.247	0.066	3.599	
UNCERTAINTY-ASYM.-NORMED	0.208	0.058	3.599	2	UNCERTAINTY-ASYM.-NORMED	0.302	0.079	3.599	3

แสดงค่าสถิติในการวัดความสัมพันธ์สำหรับตาราง 2X2 ของโปรแกรม BMDP ชุด P1F

TABLE NO. 2 ANGINA VAR(5) VS INFARCT VAR(4)

CELL FREQUENCY COUNTS

		INFARCT (VAR . 4)		
	EQ./EQ.	NONE 0.00	PRESENT 1.00	TOTAL
ANGINA (VAR 5) NONE	0.00	3	16	19
PRESENT	1.00	40	58	98
TOTAL		43	74	117

MINIMUM NONZERO EXPECTED VALUE IS 6.983
----- ALL STATISTICS ARE COMPUTED USING THE ORIGINAL CELL FREQUENCIES EXCLUDING ROWS AND/OR COLUMNS WHICH ARE ZERO.

STATISTIC	VALUE	D.F.	PROB.
FISHER EXACT TEST(1-TAIL)			0.0309
PEARSON CHISQUARE	4.288	1	0.0384
LIKELIHOOD RATIO CHISQ.	4.778	1	0.0288
PHI=CRAMER'S V	-0.191		
CONTINGENCY COEF, C	0.188		

STATISTIC	VALUE	D.F.	PROB.
FISHER EXACT TEST(2-TAIL)			0.0413
YATES' CORRECTED CHISQUARE	3.279	1	0.0702
MAXIMUM VALUE FOR PHI	-0.336		
MAXIMUM VALUE FOR C	0.318		

STATISTIC	VALUE	ASE1	T-VALUE	DEP.
YULE'S Q	-0.572	0.222	-2.202	
CROSS-PRODUCT RATIO	0.272			
TETRACHORIC CORRELATION	-0.396	0.166	-2.231	
GAMMA	-0.572	0.222	-2.359	
SOMER'S D	-0.250	0.097	-2.359	4
PRODUCT MOMENT CORRELATION	-0.191	0.075	-2.359	
LAMBDA-ASYMMETRIC	0.0	0.0	0.0	4
LAMBDA-STAR-ASYMMETRIC	0.0	0.0	0.0	4
TAU-ASYMMETRIC	0.037	0.028	1.286	4
UNCERTAINTY-ASYM.-NORMED	0.031	0.026	1.180	4

STATISTIC	VALUE	ASE1	T-VALUE	DEP.
YULE'S Y	-0.315	0.149	-2.420	
LOG(CROSS-PRODUCT RATIO)	-1.302	0.662	-2.505	
STUART'S TAU-C	-0.136	0.058	-2.359	
KENDALL'S TAU-B	-0.191	0.075	-2.359	
SOMER'S D	-0.146	0.062	-2.359	5
SPEARMAN RANK CORRELATION	-0.191	0.075	-2.359	
LAMBDA-SYMMETRIC	0.0	0.0	0.0	
LAMBDA-ASYMMETRIC	0.0	0.0	0.0	5
LAMBDA-STAR-ASYMMETRIC	0.0	0.0	0.0	5
TAU-ASYMMETRIC	0.037	0.029	1.188	5
UNCERTAINTY-NORMED	0.037	0.031	1.180	
UNCERTAINTY-ASYM.-NORMED	0.046	0.038	1.180	5

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

PERCENTAGES OF THE TOTAL FREQUENCY

FUNCTION(VAR 21)

		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00	TOTAL
ACTIVE (VAR 3)	VERY	1.00	2.27	4.55	2.27	1.14	10.23
	NORMAL	2.00	3.41	14.77	42.05	7.95	68.18
	LIMITED	3.00	0.0	1.14	2.27	18.18	21.59
		TOTAL	5.68	20.45	46.59	27.27	100.00

PERCENTAGES OF THE ROW TOTALS

FUNCTION(VAR 21)

		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00	TOTAL
ACTIVE (VAR 3)	VERY	1.00	22.22	44.44	22.22	11.11	100.00
	NORMAL	2.00	5.00	21.67	61.67	11.67	100.00
	LIMITED	3.00	0.0	5.26	10.53	84.21	100.00
		TOTAL	5.68	20.45	46.59	27.27	100.00

PERCENTAGES OF THE COLUMN TOTALS

FUNCTION(VAR 21)

		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00	TOTAL
ACTIVE (VAR 3)	VERY	1.00	40.00	22.22	4.88	4.17	10.23
	NORMAL	2.00	60.00	72.22	90.24	29.17	68.18
	LIMITED	3.00	0.0	5.56	4.88	66.67	21.59
		TOTAL	100.00	100.00	100.00	100.00	100.00

EXPECTED CELL VALUES

FUNCTION(VAR 21)

		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00
ACTIVE (VAR 3)	VERY	1.00	0.51	1.84	4.19	2.45
	NORMAL	2.00	3.41	12.27	27.95	16.36
	LIMITED	3.00	1.08	3.89	8.85	5.18

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

แสดงผลที่ได้เมื่อต้องการรวมกลุ่มของค่าสังเกตที่มีความถี่น้อยกว่า 1 โดยการเพิ่มบัตร MINIMUM ใน TABLE paragraph ส่วน (ก) ของโปรแกรม BMDP ชุด P1F

TABLE NO. 1 ACTIVE VAR(3) VS FUNCTIONVAR(2)

CELL FREQUENCY COUNTS

			FUNCTION(VAR 2)					
		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00	TOTAL	COLLAP. GROUP
ACTIVE (VAR 3)	VERY	1.00	2	4	2	1	9	1
	NORMAL	2.00	3	13	37	7	60	2
	LIMITED	3.00	0	1	2	16	19	3
	TOTAL		5	18	41	24	88	
	NOT COLLAP.	COUNTED GROUP	15 1	6 1	5 2	3 3		0

MINIMUM NONZERO EXPECTED VALUE IS 0.511

STATISTIC	VALUE	D.F.	PROB.
PEARSON CHISQUARE	47.014	4	0.0000

EXPECTED CELL VALUES

			FUNCTION(VAR 2)			
		EQ./EQ.	NONE 0.00	MINIMAL 1.00	MODERATE 2.00	SEVERE 3.00
ACTIVE (VAR 3)	VERY	1.00	0.51	1.84	4.19	2.45
	NORMAL	2.00	3.41	12.27	27.95	16.36
	LIMITED	3.00	1.08	3.89	9.85	5.18

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT2 โดยใช้ข้อมูลจาก DATA SAMPLE 2

```
//SØ3BOPØK JØB CLASS=C,MSGLEVEL=(2,0)
//  EXEC SAS
//SAS SYSIN DD *
DATA SAMPLE2;
INPUT AGE 3-4 FUNCTIØN 8 ACTIVE 12 INFARCT 16 ANGINA 20 HIGHBP 24;
 IF AGE < 51 THEN AGE = 1;
 IF AGE > 50 AND AGE < 61 THEN AGE = 2;
 IF AGE > 60 THEN AGE = 3;
CARDS;
```

บัตรบันทึกข้อมูล

```
(ก) PROC FREQ;
 TITLE 'FREQUENCIES 1-WAY ØF SAS BY SIRICHAI';
(ข) PROC FREQ CHISQ;
 TABLES FUNCTIØN*ACTIVE INFARCT*ANGINA AGE*FUNCTIØN*ACTIVE;
 TITLE FREQUENCIES 2,3-WAY ØF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 2.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT2

FREQUENCIES 1-WAY OF SAS BY SIRICHAI

(2)

FUNCTION	FREQUENCY	PERCENT
0	20	17.094
1	24	20.513
2	46	39.316
3	27	23.077
TOTALS	117	100.000

(ก)

---------- ตัวแปรอื่น ๆ จะให้ผลเช่นเดียวกัน ----------

FREQUENCIES 2,3-WAY OF SAS BY SIRICHAI

CHI-SQUARE TABLE FOR ANGINA VERSUS INFARCT

(ข) OBSERVED / EXPECTED / CELL CHI-SQUARE

(4)

ANGINA	INFARCT 0	INFARCT 1	TOTALS
0	9.00 6.98 3.27	10.00 12.02 1.90	19.00 19.00 5.17
1	44.00 50.02 0.44	54.00 47.98 0.26	98.00 98.00 0.70
TOTALS	53.00 53.00 3.71	64.00 64.00 2.16	117.00 117.00 5.87

(5)

TOTAL CHI-SQUARE = 4.28823 WITH 1 D.F. PROB > CHISQ =0.0362

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 2.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 2 โดยใช้ข้อมูลจาก DATA SAMPLE 2

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK                    MR.SIRICHAI PONGWICHAI B2214756
// EXEC PROC=$$SPSS79
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
VARIABLE LIST     AGE FUNCTION ACTIVE INFARCT ANGINA HIGHBP
INPUT MEDIUM      CARD
N OF CASES        117
INPUT FORMAT      FIXED(6FA.0)
MISSING VALUES    ACTIVE(0)
VALUE LABELS      FUNCTION(0)NONE(1)MINIMAL(2)MODERATE(3)SEVERE/
                  ACTIVE(1)VERY(2)NORMAL(3)LIMITED/
                  INFARCT TO HIGHBP(0)NONE(1)PRESENT/
                  AGE(1)UNDER51(2)51-60(3)OVER60
RECODE            AGE(LO THRU 50=1)(51 THRU 60=2)(61 THRU HI=3)
(n) TASK NAME     'COMPARE FREQUENCIES'
FREQUENCIES       GENERAL=ALL
OPTIONS           1,6,8,9
STATISTICS        ALL
READ INPUT DATA
```

บัตรบันทึกข้อมูล

ตารางที่ 2.3 (ต่อ)

```
TASK NAME          COMPARE TWO-WAY FREQUENCIES TABLES
                   TABLES = ACTIVE BY FUNCTION/ANGINA BY INFARCT
                          /ACTIVE BY FUNCTION BY AGE
OPTIONS            6,9;
STATISTICS         ALL
FINISH
/*
/&
* $$ EOJ
```

รูปที่ 2.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT2

```
AGE

                                                                                                    (n)
                                    (2)
CATEGORY LABEL          CODE    ABSOLUTE    RELATIVE    ADJUSTED    CUM
                                  FREQ      FREQ        FREQ        FREQ
                                            (PCT)       (PCT)       (PCT)
UNDER 51                 1.       47        40.2        40.2        40.2
51-60                    2.       48        41.0        41.0        81.2
OVER60                   3.       22        18.8        18.8       100.0
                                  -----      -----       -----
                        TOTAL     117       100.0       100.0

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'                02/05/81        PAGE    3
'COMPARE FREQUENCIES'
FILE    NONAME    (CREATION DATE = 02/05/81)

AGE

      CODE    (3)
              I
        1.    ************************************************** (    47)
              I  UNDER 51
              I
              I
        2.    ************************************************** (    48)
              I  51-60
              I
              I
        3.    ***********************                            (    22)
              I  OVER60
              I
              I.........I.........I.........I.........I.........I
              0         10        20        30        40        50
              FREQUENCY

(1)
  MEAN          1.786     STD ERR       0.068     MEDIAN        1.740
  MODE          2.000     STD DEV       0.741     VARIANCE      0.549
  KURTOSIS     -1.096     SKEWNESS      0.366     RANGE         2.000
  MINIMUM       1.000     MAXIMUM       3.000

  VALID CASES     117     MISSING CASES     0     ---------- ตัวแปรอื่น ๆ จะให้ผลเช่นเดียวกัน ----------
```

(ข)

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
'COMPARE TWO-WAY FREQUENCIES TABLES'
FILE NONAME (CREATION DATE = 29/04/81)

* C R O S S T A B U L A T I O N O F *
ACTIVE BY FUNCTION
* PAGE 1 OF 1

(4)

ACTIVE	FUNCTION: COUNT / ROW PCT / COL PCT / TOT PCT	NONE 0.	MINIMAL 1.	MODERTE 2.	SEVERE 3.	ROW TOTAL
VERY	1.	2	4	2	1	9
		22.2	44.4	22.2	11.1	10.2
		40.0	22.2	4.9	4.2	
		2.3	4.5	2.3	1.1	
NORMAL	2.	3	13	37	7	60
		5.0	21.7	61.7	11.7	68.2
		60.0	72.2	90.2	29.2	
		3.4	14.8	42.0	8.0	
LIMITED	3.	0	1	2	16	19
		0.0	5.3	10.5	84.2	21.6
		0.0	5.6	4.9	66.7	
		0.0	1.1	2.3	18.2	
COLUMN TOTAL		5	18	41	24	88
		5.7	20.5	46.6	27.3	100.0

CHI SQUARE = 48.36472 WITH 6 DEGREES OF FREEDOM SIGNIFICANCE = 0.0000
CRAMER'S V = 0.52421
CONTINGENCY COEFFICIENT = 0.59554
LAMBDA (ASYMMETRIC) = 0.32143 WITH ACTIVE DEPENDENT. = 0.34043 WITH FUNCTION DEPENDENT.
LAMBDA (SYMMETRIC) = 0.33333
UNCERTAINTY COEFFICIENT (ASYMMETRIC) = 0.30228 WITH ACTIVE DEPENDENT. = 0.20828 WITH FUNCTION DEPENDENT.
UNCERTAINTY COEFFICIENT (SYMMETRIC) = 0.24662
KENDALL'S TAU B = 0.53608 SIGNIFICANCE = 0.0000
KENDALL'S TAU C = 0.45287 SIGNIFICANCE = 0.0000
GAMMA = 0.78509
SOMERS'S D (ASYMMETRIC) = 0.45504 WITH ACTIVE DEPENDENT. = 0.63155 WITH FUNCTION DEPENDENT.
SOMERS'S D (SYMMETRIC) = 0.52896
ETA = 0.59589 WITH ACTIVE DEPENDENT. = 0.56171 WITH FUNCTION DEPENDENT.
PEARSON'S R = 0.55101 SIGNIFICANCE = 0.0000

NUMBER OF MISSING OBSERVATIONS = 29

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(ข)

CROSSTABULATION OF ANGINA BY INFARCT PAGE 1 OF 1

ANGINA	INFARCT: NONE (0.)	INFARCT: PRESENT (1.)	ROW TOTAL
NONE (0.)	3	16	19
	15.8	84.2	16.2
	7.0	21.6	
	2.6	13.7	
PRESENT (1.)	40	58	98
	40.8	59.2	83.8
	93.0	78.4	
	34.2	49.6	
COLUMN TOTAL	43	74	117
	36.8	63.2	100.0

CORRECTED CHI SQUARE = 3.27915 WITH 1 DEGREE OF FREEDOM SIGNIFICANCE = 0.0702
PHI = 0.19145
CONTINGENCY COEFFICIENT = 0.18803
LAMBDA (ASYMMETRIC) = 0.0 WITH ANGINA DEPENDENT. = 0.0 WITH INFARCT DEPENDENT.
LAMBDA (SYMMETRIC) = 0.0
UNCERTAINTY COEFFICIENT (ASYMMETRIC) = 0.04603 WITH ANGINA DEPENDENT. = 0.03105 WITH INFARCT DEPENDENT.
UNCERTAINTY COEFFICIENT (SYMMETRIC) = 0.03708
KENDALL'S TAU B = -0.19145 SIGNIFICANCE = 0.0196
KENDALL'S TAU C = -0.13617 SIGNIFICANCE = 0.0196
GAMMA = -0.57248
SOMERS'S D (ASYMMETRIC) = -0.14645 WITH ANGINA DEPENDENT. = -0.25027 WITH INFARCT DEPENDENT.
SOMERS'S D (SYMMETRIC) = -0.18477
ETA = 0.19144 WITH ANGINA DEPENDENT. = 0.19145 WITH INFARCT DEPENDENT.
PEARSON'S R = -0.19145 SIGNIFICANCE = 0.0193

--- แสดงค่าสถิติสำหรับตาราง 2X2 ของโปรแกรม SPSS -----

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT 3 การหาความสัมพันธ์ระหว่างตัวแปร

(CØRRELATIØN CØEFFICIENT)

ตารางที่ 3 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 3

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓	✓	✓	
② ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน -ในรูปเมทริกซ์	✓	✓	✓	
③ ค่าระดับนัยสำคัญที่ใช้ทดสอบ -ความสัมพันธ์ Ho: $\rho = 0$		✓	✓	
④ ค่าความแปรปรวนร่วม	✓		✓	
⑤ ค่าสัมประสิทธิ์สหสัมพันธ์ของ- Kendallในรูปเมทริกซ์ พร้อมกับ- ระดับนัยสำคัญ	✓	✓		
⑥ ค่าสัมประสิทธิ์สหสัมพันธ์ของ- -Spearman ในรูปเมทริกซ์- -พร้อมกับระดับนัยสำคัญ	✓	✓		
⑦ การประมาณค่าตัวแปรที่ไม่สมบูรณ์- -ก่อนนำมาหาความสัมพันธ์	✓			
⑧ การเลือกค่าสังเกตที่จะใช้ในการ- -คำนวณ	✓		✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 3 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//TESTBMDP JOB (9012Ø,9Ø,9Ø87),'CHULA'
// EXEC BIMED
//LKED.SYSIN DD *
  INCLUDE BMDP(BMDP8D)
  ENTRY MAIN
//GØ.SYSIN DD *
/PRØBLEM   TITLE IS 'CØRRELATIØN ØF BMDP BY SIRICHAI'.
/INPUT     VARIABLES ARE 9.
           FØRMAT IS '(A4,5F4.Ø,3F4.1)'.
/VARIABLE  NAMES ARE ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHØLSTRL,
           ALBUMIN,CALCIUM,URICACID.
           MAXIMUM IS (6) 4ØØ.
           MINIMUM IS (6) 15Ø.
           BLANKS ARE MISSING.
           LABEL IS ID.
/PRINT     MEAN.
           VARIANCE.
           CØVARIANCE.
/END
```

บัตรบันทึกข้อมูล

```
//
```

รูปที่ 3.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT3

```
WEIGHT VARIABLE . . . . . .  0
CORRELATION TYPE(S)
   USE ALL EXISTING VALUES. . . . . .  (8) NO
   USE ALL EXISTING PAIRS FOR CROSS PRODUCTS. .  NO
   USE ALL EXISTING PAIRS FOR CORRELATIONS. . .  YES
   USE ONLY COMPLETE CASES. . . . . .  NO
PRINT MATRIX OF FREQUENCIES . . . . . .  YES
PRINT MATRIX OF SUM OF WEIGHTS. . . . . .  NO
PRINT MATRIX OF MEANS . . . . . .  NO
PRINT MATRIX OF VARIANCES . . . . . .  NO
PRINT COVARIANCE MATRIX . . . . . .  NO
PRINT CORRELATION MATRIX. . . . . .  YES

NUMBER OF CASES READ. . . . . .  188
```

(1)

VARIABLE NO.	LABEL	MEAN	VARIANCE	STANDARD DEVIATION	COEFF. OF VARIATION	FREQUENCY	SUM OF WEIGHTS
2	AGE	33.81914	102.26572	10.11265	0.29902	188.00000	188.00000
3	HEIGHT	64.51074	6.17545	2.48505	0.03852	186.00000	186.00000
4	WEIGHT	131.67204	426.85327	20.66043	0.15691	186.00000	186.00000
5	BRTHPILL	1.50000	0.25133	0.50133	0.33422	188.00000	188.00000
6	CHOLSTRL	236.15053	1810.94702	42.55522	0.18020	186.00000	186.00000
7	ALBUMIN	4.11123	0.12814	0.35796	0.08707	186.00000	186.00000
8	CALCIUM	9.96208	0.22997	0.47955	0.04814	185.00000	185.00000
9	URICACID	4.77051	1.33919	1.15723	0.24258	187.00000	187.00000

FREQUENCY TABLE

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	188							
HEIGHT	3	186	186						
WEIGHT	4	186	184	186					
BRTHPILL	5	188	186	186	188				
CHOLSTRL	6	186	184	184	186	186			
ALBUMIN	7	186	185	184	186	184	186		
CALCIUM	8	185	184	183	185	184	183	185	
URICACID	9	187	185	185	187	185	185	185	187

ESTIMATES OF CORRELATIONS - CORPAIR

(2)

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	1.0000							
HEIGHT	3	0.0807	1.0000						
WEIGHT	4	0.2521	0.4759	1.0000					
BRTHPILL	5	0.0	0.1692	0.0738	1.0000				
CHOLSTRL	6	0.3678	0.0199	0.1474	0.0758	1.0000			
ALBUMIN	7	-0.0719	-0.0208	-0.2533	-0.2398	0.0534	1.0000		
CALCIUM	8	-0.0325	0.1485	0.0793	-0.0561	0.2515	0.4534	1.0000	
URICACID	9	0.1776	0.1335	0.3043	0.0247	0.2660	0.0072	0.1949	1.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ESTIMATE OF MEAN OF COLUMN VARIABLE FOR USE WITH ROW VARIABLE

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	33.819	64.511	131.672	1.500	236.151	4.111	9.962	4.771
HEIGHT	3	33.699	64.511	131.370	1.505	236.196	4.112	9.959	4.771
WEIGHT	4	33.774	64.484	131.672	1.500	236.207	4.113	9.961	4.747
BRTHPILL	5	33.819	64.511	131.672	1.500	236.151	4.111	9.962	4.771
CHOLSTRL	6	33.930	64.495	131.418	1.495	236.151	4.116	9.964	4.767
ALBUMIN	7	33.694	64.524	131.685	1.505	236.196	4.111	9.967	4.784
CALCIUM	8	33.670	64.500	131.191	1.503	235.745	4.117	9.962	4.756
URICACID	9	33.711	64.497	131.600	1.503	235.697	4.112	9.962	4.771

--- frequency table of pairs of variables ---

ESTIMATE OF VARIANCE OF COLUMN VARIABLE FOR USE WITH ROW VARIABLE

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	102.266	6.175	426.853	0.251	1810.947	0.128	0.230	1.339
HEIGHT	3	101.313	5.175	419.223	0.251	1830.273	0.129	0.230	1.353
WEIGHT	4	102.488	6.164	426.853	0.251	1826.499	0.129	0.232	1.300
BRTHPILL	5	102.256	6.175	426.853	0.251	1810.947	0.128	0.230	1.339
CHOLSTRL	6	102.183	6.175	425.233	0.251	1810.947	0.127	0.231	1.305
ALBUMIN	7	101.099	5.175	430.298	0.251	1830.276	0.128	0.221	1.325
CALCIUM	8	98.939	6.208	416.480	0.251	1791.697	0.128	0.230	1.321
URICACID	9	100.614	6.175	428.204	0.251	1782.373	0.129	0.230	1.339

--- correlation matrix ---

ESTIMATES OF COVARIANCES - COVPAIR

(4)

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	102.266							
HEIGHT	3	2.019	6.175						
WEIGHT	4	52.730	24.192	426.853					
BRTHPILL	5	0.0	0.211	0.765	0.251				
CHOLSTRL	6	158.215	2.116	129.923	1.617	1810.947			
ALBUMIN	7	-0.259	-0.019	-1.887	-0.043	0.814	0.128		
CALCIUM	8	-0.155	0.177	0.692	-0.013	5.112	0.076	0.230	
URICACID	9	2.062	0.386	7.181	0.014	12.826	0.003	0.107	1.339

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 3 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//S03800PGK JOB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE1;
INPUT ID $ 1-4 AGE 5-8 HEIGHT 9-12 WEIGHT 13-16 BRTHPILL 17-20 CHOL
 CHOLSTRL 21-24 ALBUMIN 25-28 1 CALCIUM 29-31 1 URICACID 33-36 1;
CARDS;
```

บัตรบันทึกข้อมูล

```
PROC CORR;
 TITLE 'CORRELATION OF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 3.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT3

CORRELATION OF SAS BY SIRICHAI

	N	SUM (1)	MEAN	MIN VALUE	MAX VALUE	CORRECTED SS	STANDARD DEV
AGE	188	6356.00000000	33.81914894	19.00000000	55.00000000	19123.85106383	10.11269416
HEIGHT	186	11999.00000000	64.51075269	57.00000000	71.00000000	1142.47849462	2.48506729
WEIGHT	186	24491.00000000	131.67204301	94.00000000	215.00000000	78968.99462366	20.66057672
ERTHPILL	186	252.00000000	1.35483871	1.00000000	2.00000000	47.00000000	0.50133512
CHOLSTRL	188	44574.00000000	237.09574468	80.00000000	600.00000000	501900.27659574	51.80693678
ALBUMIN	186	764.70000000	4.11129032	3.20000000	5.00000000	23.70629032	0.35796943
CALCIUM	185	1843.00000000	9.96216216	8.60000000	11.10000000	42.31513514	0.47955559
URICACID	187	892.10000000	4.77005882	2.20000000	9.90000000	249.08823529	1.15723120

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO: RHO=0 / NUMBER OF OBSERVATIONS

	AGE	HEIGHT	WEIGHT	ERTHPILL	CHOLSTRL	ALBUMIN	CALCIUM	URICACID (2)
AGE	1.000000 0.0000 188	0.080735 0.2727 186	0.252106 0.0008 186	0.000000 1.0000 186	0.271625 0.0004 188	-0.072020 0.6702 186	-0.032595 0.6646 185	0.177583 (3) 0.0143 187
HEIGHT	0.080735 0.2727 186	1.000000 0.0000 186	0.475909 0.0001 184	0.169168 0.0198 186	-0.018723 0.7955 186	-0.020773 0.7758 185	0.148517 0.0416 184	0.133464 0.0686 185
WEIGHT	0.252106 0.0008 186	0.475909 0.0001 184	1.000000 0.0000 186	0.073842 0.3176 186	0.154172 0.0334 186	-0.253354 0.0008 184	0.076292 0.6536 183	0.304327 0.0001 185
ERTHPILL	0.000000 1.0000 186	-0.169168 0.0198 186	0.073842 0.3176 186	1.000000 0.0000 186	0.079887 0.2752 188	-0.239801 0.0014 186	-0.056085 0.5452 185	0.024694 0.7369 187
CHOLSTRL	0.271625 0.0004 188	-0.018723 0.7955 186	0.154172 0.0334 186	0.079887 0.2752 188	1.000000 0.0000 188	-0.004528 0.9497 186	0.256094 0.0007 185	0.321850 0.0001 187
ALBUMIN	-0.072020 0.6702 186	-0.020773 0.7758 185	-0.253354 0.0008 184	-0.239801 0.0014 186	-0.004528 0.9497 186	1.000000 0.0000 186	0.453447 0.0001 183	0.007189 0.9196 185
CALCIUM	-0.032595 0.6646 185	0.148517 0.0416 184	0.076292 0.6536 183	-0.056085 0.5452 185	0.256094 0.0007 185	0.453447 0.0001 183	1.000000 0.0000 185	0.194862 0.0078 185
URICACID	0.177583 0.0143 187	0.133464 0.0686 185	0.304327 0.0001 185	0.024694 0.7369 187	0.321850 0.0001 187	0.007189 0.9196 185	0.194862 0.0078 185	1.000000 0.0000 187

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 3 โดยใช้ข้อมูลจาก SAMPLE 1

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK          MR.SIRICHAI PONGWICHAI 0221475
// EXEC PROC=$$SPSS79
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME         'DATA SAMPLE1'
VARIABLE LIST     ID AGE HEIGHT WEIGHT BRTHPILL CHOLSTRL ALBUMIN
                  CALCIUM URICACID
INPUT MEDIUM      CARD
N OF CASES        188
INPUT FORMAT      FIXED(A4,5F4.0,3F4.1)
MISSING VALUES    AGE TO URICACID(0)
VALUE LABELS      BRTHPILL(1)NOPILL(2)PILL
RECODE            CHOLSTRL(LO THRU 199=0)(401 THRU HI=0)
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
TASK NAME         'CORRELATION OF SPSS BY SIRICHAI'
PEARSON CORR      AGE TO URICACID
OPTIONS           3
STATISTICS        ALL
FINISH
/*
/&
* $$ EOJ
```

รูปที่ 3.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT3

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS' 29/04/81 PAGE 5
'CORRELATION OF SPSS BY SIRICHAI '
FILE 'DATA (CREATION DATE = 29/04/81) SAMPLE1'

VARIABLE	CASES	MEAN	STD DEV (1)
AGE	188	33.8191	10.1127
HEIGHT	186	64.5108	2.4851
WEIGHT	186	131.6720	20.6606
BRTHPILL	188	1.5000	0.5013
CHOLSTRL	186	236.1505	42.5555
ALBUMIN	186	4.1113	0.3580
CALCIUM	185	9.9622	0.4796
URICACID	187	4.7706	1.1572

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS' 29/04/81 PAGE 6
'CORRELATION OF SPSS BY SIRICHAI '
FILE 'DATA (CREATION DATE = 29/04/81) SAMPLE1' (4)

VARIABLES		CASES	CROSS-PROD DEV	VARIANCE-COVAR	VARIABLES		CASES	CROSS-PROD DEV	VARIANCE-COVAR (4)
AGE	HEIGHT	186	373.6022	2.0195	AGE	WEIGHT	186	9755.2258	52.7310
AGE	BRTHPILL	188	0.0	0.0	AGE	CHOLSTRL	186	29269.9570	158.2160
AGE	ALBUMIN	186	-47.9564	-0.2592	AGE	CALCIUM	185	-28.6080	-0.1555
AGE	URICACID	187	383.4117	2.0614	HEIGHT	WEIGHT	184	4427.1087	24.1919
HEIGHT	BRTHPILL	186	38.9892	0.2108	HEIGHT	CHOLSTRL	184	387.1957	2.1158
HEIGHT	ALBUMIN	185	-3.4059	-0.0185	HEIGHT	CALCIUM	184	32.4500	0.1773
HEIGHT	URICACID	185	70.9854	0.3858	WEIGHT	BRTHPILL	186	141.5000	0.7649
WEIGHT	CHOLSTRL	184	23776.0978	129.9240	WEIGHT	ALBUMIN	184	-345.4434	-1.8877
WEIGHT	CALCIUM	183	125.8580	0.6915	WEIGHT	URICACID	185	1321.2801	7.1809
BRTHPILL	CHOLSTRL	186	299.1505	1.6170	BRTHPILL	ALBUMIN	186	-7.9613	-0.0430
BRTHPILL	CALCIUM	185	-2.4811	-0.0135	BRTHPILL	URICACID	187	2.6647	0.0143
CHOLSTRL	ALBUMIN	184	148.8131	0.8132	CHOLSTRL	CALCIUM	184	935.3140	5.1110
CHOLSTRL	URICACID	185	2359.9533	12.8258	ALBUMIN	CALCIUM	183	13.8533	0.0761
ALBUMIN	URICACID	185	0.5461	0.0030	CALCIUM	URICACID	185	19.7635	0.1074

- - - - - - - - - - - - - - PEARSON CORRELATION COEFFICIENTS' - - - - - - - - - - - - -

	AGE	HEIGHT	WEIGHT	BRTHPILL	CHOLSTRL	ALBUMIN	CALCIUM	URICACID
AGE	1.0000	0.3807	0.2521	0.0	0.3678	-0.0720	-0.0326	0.1776 (2)
	(0)	(186)	(186)	(186)	(186)	(186)	(185)	(187)
	S=0.000	S=0.273	S=0.001	S=1.000	S=0.001	S=0.329	S=0.660	S=0.015
HEIGHT	0.3807	1.0000	0.4759	0.1692	0.0199	-0.0208	0.1485	0.1335 (3)
	(186)	(0)	(184)	(186)	(184)	(185)	(184)	(185)
	S=0.273	S=0.000	S=0.001	S=0.021	S=0.789	S=0.779	S=0.044	S=0.070
WEIGHT	0.2521	0.4759	1.0000	0.0738	0.1474	-0.2534	0.0703	0.3043
	(186)	(184)	(0)	(186)	(184)	(184)	(183)	(185)
	S=0.001	S=0.001	S=0.000	S=0.317	S=0.046	S=0.001	S=0.344	S=0.001
BRTHPILL	0.0	0.1692	0.0738	1.0000	0.0758	-0.2398	-0.0561	0.0247
	(186)	(186)	(186)	(0)	(186)	(186)	(185)	(187)
	S=1.000	S=0.021	S=0.317	S=0.000	S=0.304	S=0.001	S=0.448	S=0.737
CHOLSTRL	0.3678	0.0199	0.1474	0.0758	1.0000	0.0533	0.2515	0.2660
	(186)	(184)	(184)	(186)	(0)	(184)	(184)	(185)
	S=0.001	S=0.789	S=0.046	S=0.304	S=0.000	S=0.472	S=0.001	S=0.001
ALBUMIN	-0.0720	-0.0208	-0.2534	-0.2398	0.0533	1.0000	0.4534	0.0072
	(186)	(185)	(184)	(186)	(184)	(0)	(183)	(185)
	S=0.329	S=0.779	S=0.001	S=0.001	S=0.472	S=0.000	S=0.001	S=0.923
CALCIUM	-0.0326	0.1485	0.0703	-0.0561	0.2515	0.4534	1.0000	0.1949
	(185)	(184)	(183)	(185)	(184)	(183)	(0)	(185)
	S=0.660	S=0.044	S=0.344	S=0.448	S=0.001	S=0.001	S=0.000	S=0.008
URICACID	0.1776	0.1335	0.3043	0.0247	0.2660	0.0072	0.1949	1.0000
	(187)	(185)	(185)	(187)	(185)	(185)	(185)	(0)
	S=0.015	S=0.070	S=0.001	S=0.737	S=0.001	S=0.923	S=0.008	S=0.000

(COEFFICIENT / (CASES) / SIGNIFICANCE) (A VALUE OF 99.0000 IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED)

จุฬาลงกรณ์มหาวิทยาลัย

STAT 4 การเขียนกราฟเพื่อดูรูปแบบของความสัมพันธ์ระหว่างตัวแปร (BIVARIATE OR SCATTER PLØT)

ตารางที่ 4 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT4

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① กราฟแสดงความสัมพันธ์ระหว่าง 2 ตัวแปร	✓	✓	✓	
② เส้นตรงที่ได้จากสมการถดถอย -เมื่อใช้ตัวแปรในแกน X หรือ Y ทำนายตัวแปรในแกน Y หรือ X -ตามลำดับ	✓			
③ ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (r)	✓		✓	
④ ค่าสัมประสิทธิ์การตัดสินใจ (R^2)			✓	
⑤ ค่าระดับนัยสำคัญสำหรับทดสอบ -ความสัมพันธ์ของ 2 ตัวแปร [Ho: ρ=0]			✓	
⑥ ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓			
⑦ สมการถดถอยเมื่อใช้ตัวแปรใน -แกน X ทำนายตัวแปรในแกน Y -โดยแสดงค่าจุดตัดบนแกน Y (b) -และค่าสัมประสิทธิ์การถดถอย (a)	✓		✓	⑦ สมการถดถอยตัวแปร X ทำนายตัวแปร Y คือ y = ax+b
⑧ ค่าความคลาดเคลื่อนมาตรฐาน -ของการประมาณ (Sy.x)	✓		✓	
⑨ สมการถดถอยในรูปแบบเช่นเดียว -กับข้อ ⑦ - ⑧ แต่ใช้ตัวแปร Y ทำนายตัวแปร X	✓			⑨ สมการถดถอยตัวแปร Y ทำนายตัวแปร X คือ X = ay+b

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(10) กราฟแสดงความสัมพันธ์ระหว่าง 2 ตัวแปรซึ่งถูกแบ่งกลุ่มโดยตัวแปรที่ 3 ในกรอบเดียวกัน	✓			
(11) กราฟแสดงความสัมพันธ์ระหว่าง 3 ตัวแปร (Contour Plot)		✓		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDPสำหรับ STAT 4 โดยใช้ข้อมูลจาก DATA SAMPLE 1

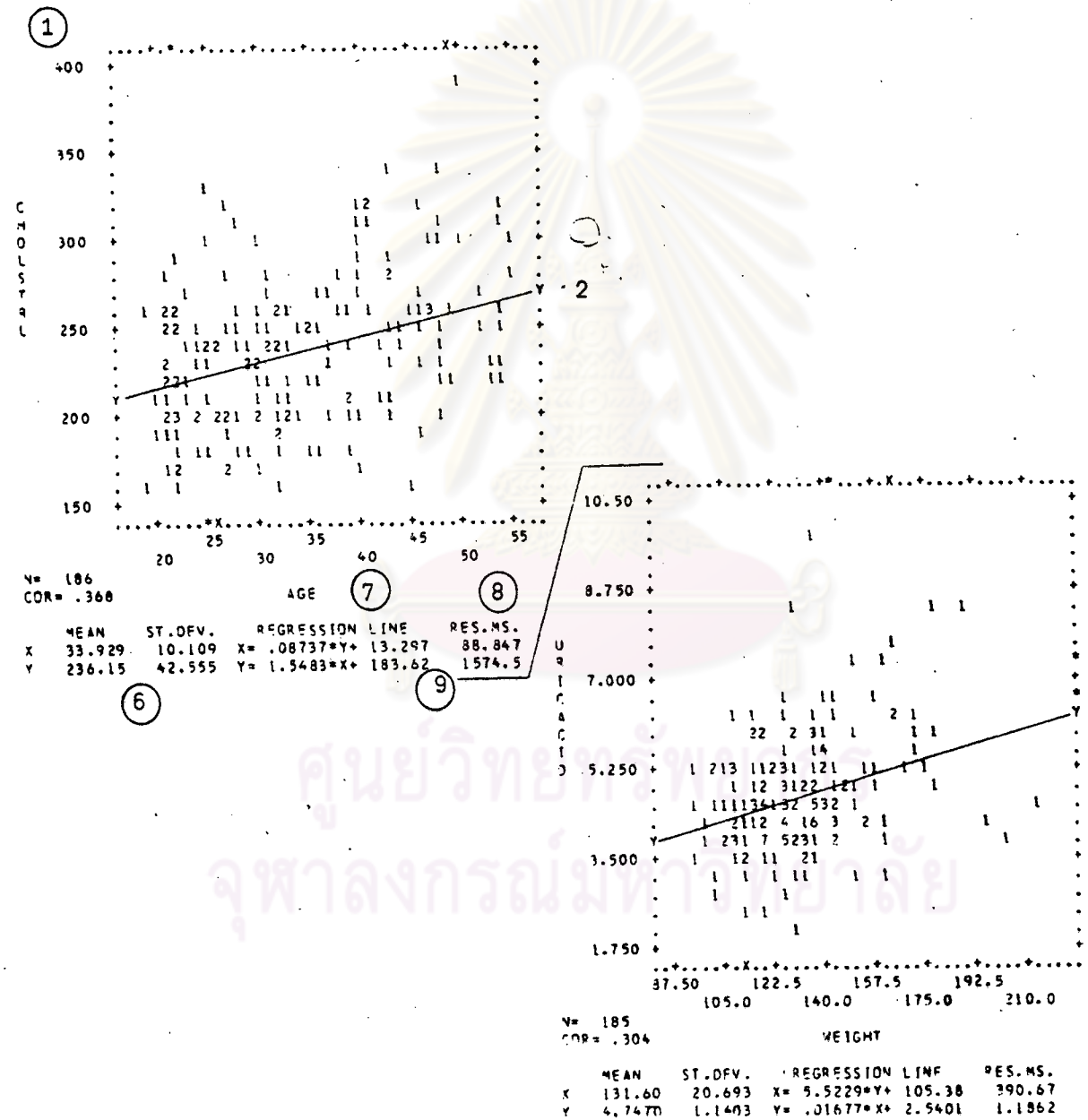
```
//TESTBMDP  JOB (9012,90,90SV),'CHULA'
//  EXEC BMED
//LKED.SYSIN DD *
   INCLUDE BMDP(BMDP6D)
   ENTRY MAIN
//GO.SYSIN  DD *
/PROBLEM    TITLE IS 'SCATTER OF BMDP BY SIRICHAI'.
/INPUT      VARIABLES ARE 9.
            FORMAT IS '(A4,5F4.0,3F4.1)'.
/VARIABLE   NAMES ARE ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHOLSTRL,
                      ALBUMIN,CALCIUM,URICACID.
            MAXIMUM IS (6) 400.
            MINIMUM IS (6) 150.
            BLANKS ARE MISSING.
            LABEL IS ID.
/PLOT       YVAR ARE CHOLSTRL,URICACID.
            XVAR ARE      AGE,WEIGHT.
            STATISTICS.
/END
```

---------- บัตรบันทึกข้อมูล ----------

```
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 4.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT4



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 4.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 4 โดยใช้ข้อมูลจาก DATA SAMPLE 1

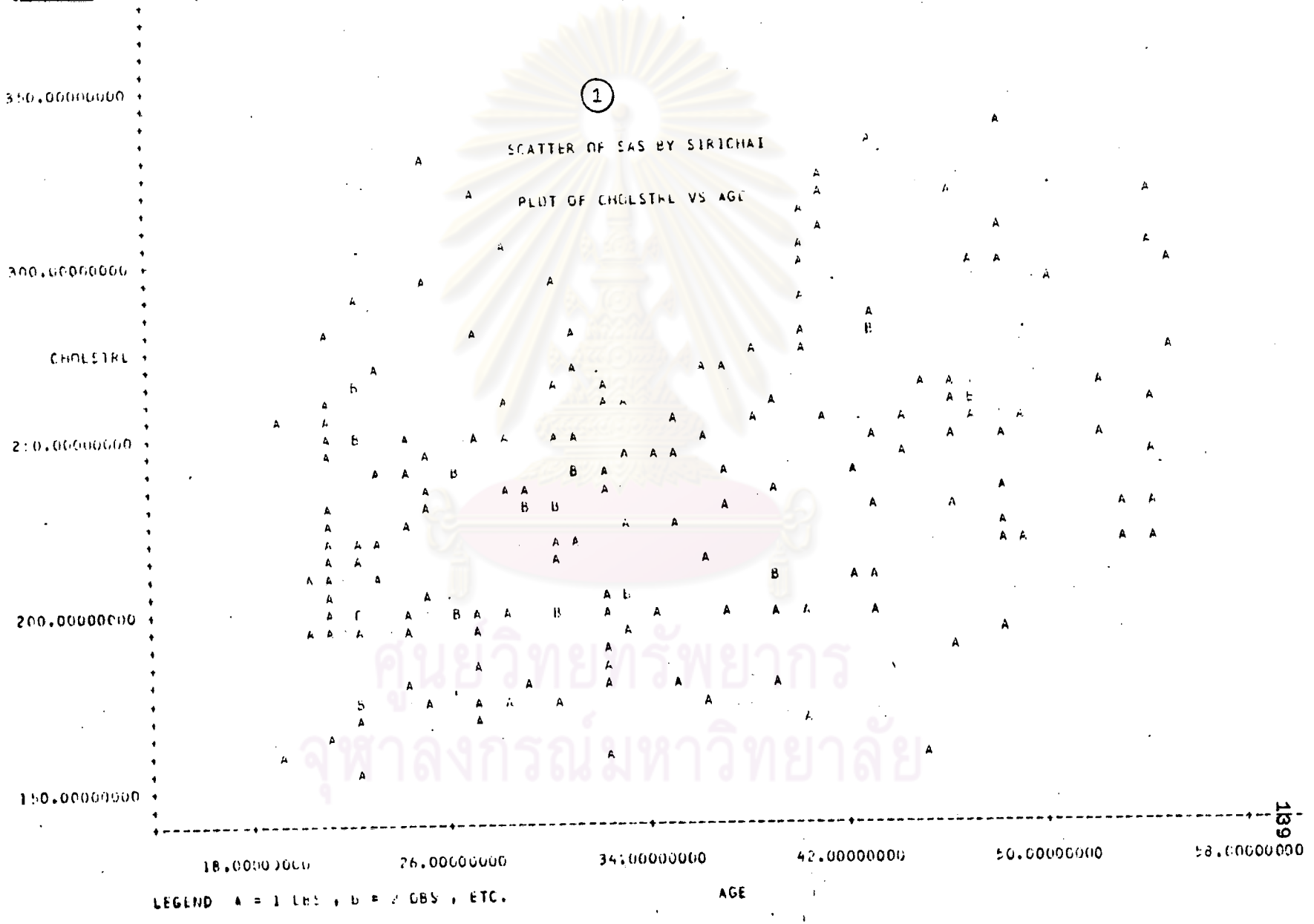
```
//S036OPOK JOB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE1;
INPUT ID $ 1-4 AGE 5-8 HEIGHT 9-12 WEIGHT 13-16 BRTHPILL 17-20
 CHOLSTRL 21-24 ALBUMIN 25-28 1 CALCIUM 29-31 1 URICACID 33-36 1;
CARDS;
```

บัตรบันทึกข้อมูล

```
PROC PLOT;
 VAR CHOLSTRL AGE URICACID WEIGHT;
 TITLE 'SCATTER OF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 4.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT4



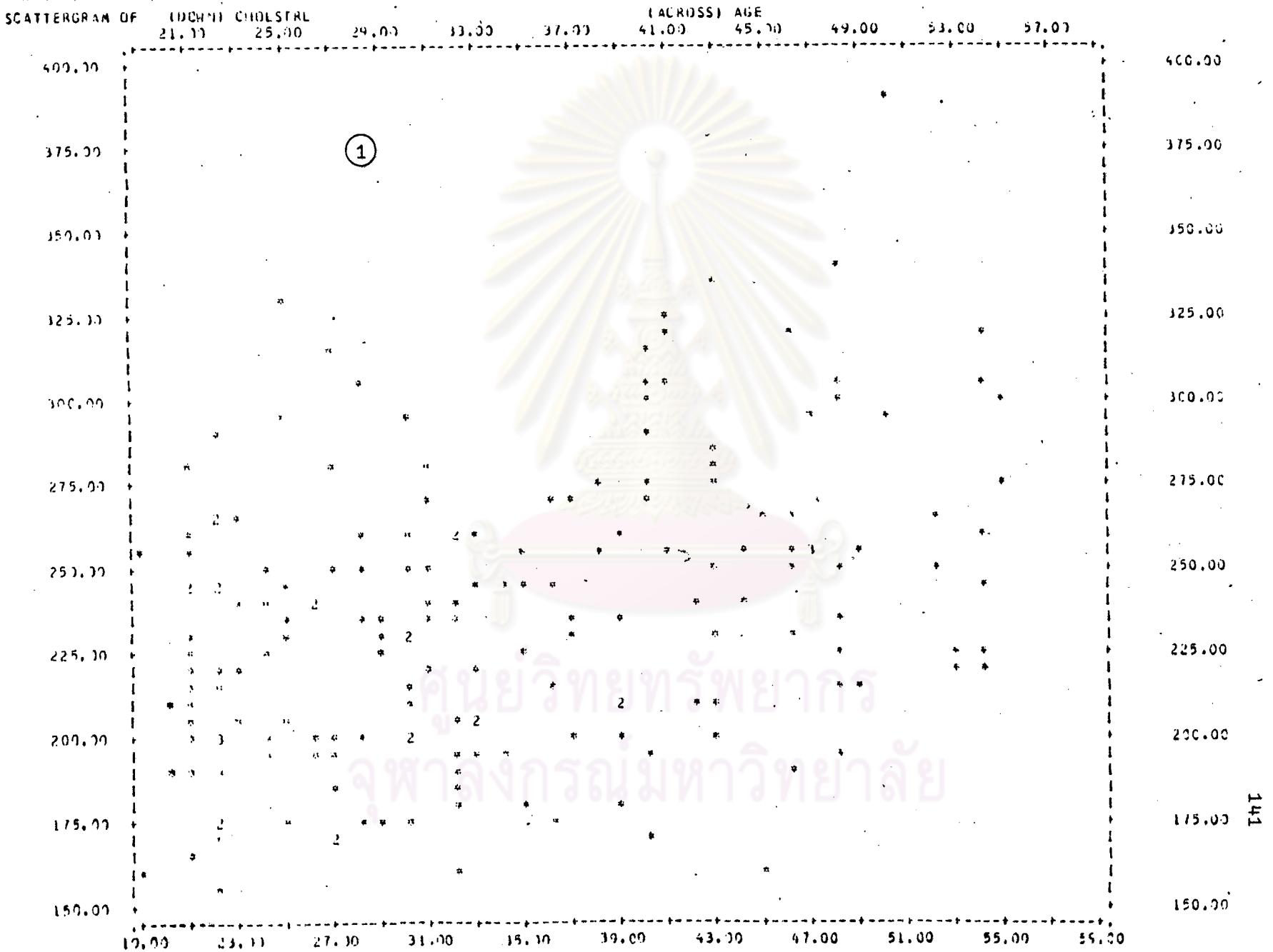
ตารางที่ 4.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 4 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK          MR. SIRICHAI PONGWICHAI B22175
// EXEC PROC=$$SPSS79
RUN NAME                'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME               'DATA SAMPLE1'
VARIABLE LIST           ID AGE HEIGHT WEIGHT BRTHPILL CHOLSTRL ALBUMIN
                        CALCIUM URICACID
INPUT MEDIUM            CARD
N OF CASES              188
INPUT FORMAT            FIXED(A4,5F4.0,3F4.1)
MISSING VALUES          AGE TO URICACID(0)
VALUE LABELS            BRTHPILL(1)NOPILL(2)PILL
RECODE                  CHOLSTRL(LO THRU 149=0)(401 THRU HI=0)
READ INPUT DATA
```

---------- บัตรบันทึกข้อมูล ----------

```
TASK NAME               'SCATTERGRAM OF SPSS BY SIRICHAI'
SCATTERGRAM             CHOLSTRL(150,400),AGE/URICACID,WEIGHT
OPTIONS                 2,4,6,7
STATISTICS              ALL
FINISH
/*
/&
* $$ EOJ
```

รูปที่ 4.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT4



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STATISTICS..

(3) CORRELATION (R) -	0.36779	(4) R SQUARED -	0.13527	(5) SIGNIFICANCE -	0.00001
(8) STD ERR OF EST -	39.68010	(7) INTERCEPT (A) -	103.61520	(7) SLOPE (B) -	1.54334
PLOTTED VALUES -	186	EXCLUDED VALUES-	0	MISSING VALUES -	2

'**********' IS PRINTED IF A COEFFICIENT CANNOT BE COMPUTED.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT5 การวิเคราะห์การถดถอย (REGRESSION ANALYSIS)

ตารางที่ 5 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 5

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① การวิเคราะห์การถดถอยเชิงเส้นสำหรับตัวแปรอิสระ 1 ตัว (Simple linear reg)	✓	✓	✓	
② การวิเคราะห์การถดถอยเชิงเส้นสำหรับตัวแปรอิสระมากกว่า 1 ตัว (Multiple linear reg)	✓	✓	✓	
③ การวิเคราะห์การถดถอยในรูปแบบโพลีโนเมียล (Polynomial reg)	✓			โปรแกรม SAS, SPSS สามารถวิเคราะห์ได้โดยดัดแปลงข้อมูลก่อนแต่โปรแกรม BMDP วิเคราะห์ได้สำหรับสมการรูปแบบดังกล่าวโดยตรง
④ การวิเคราะห์การถดถอยในรูปแบบอื่นๆ (Non-linear reg)	✓			
การคัดเลือกตัวแปรอิสระเพื่อสร้างสมการถดถอยที่ดีที่สุดโดยวิธีต่อไปนี้				② สมการถดถอยเชิงเส้นสำหรับตัวแปรอิสระ (x) มากกว่า 1 ตัว
⑤ Forward Selection	✓	✓		รูปประชากร
⑥ Backward elemination	✓	✓		$y = B_o+B_1X_1+...+B_pXp+\epsilon$
⑦ Stepwise regression	✓	✓	✓	รูปตัวอย่าง
⑧ Maximum R^2 improvement		✓		$y = b_o+b_1X_1+...b_pX_p+\epsilon$
⑨ Minimum R^2 improvement		✓		$\hat{y} = b_o+b_1X_1+..+b_pX_p$
⑩ ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓	✓	✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(11) เมทริกซ์สหสัมพันธ์หรือเมทริกซ์ความแปรปรวนร่วม	✓	✓	✓	x เป็นตัวแปรอิสระจากประชากร
(12) • ค่าสัมประสิทธิ์สหสัมพันธ์ร่วม (R)	✓	✓	✓	y เป็นตัวแปรตามจากประชากร
• ค่าสัมประสิทธิ์การตัดสินใจ (R^2)	✓	✓	✓	(12) R = Multiple--Correlation
• ค่าสัมประสิทธิ์การตัดสินใจที่ปรับแก้แล้ว (AdjustR^2)	✓		✓	R^2= Coefficient of--determination
• ค่าสัมประสิทธิ์การถดถอย (bi)	✓	✓	✓	AdjustR^2= Adjust--Coefficient of--determination
• ค่าสัมประสิทธิ์การถดถอยในรูปมาตรฐาน (biSx/Sy)	✓	✓	✓	bi = Coefficient of regression
• ค่าความคลาดเคลื่อนมาตรฐานของการประมาณค่า y(Sy.x)	✓	✓	✓	biSx/Sy = Standardized reg'Coefficient
• ค่าความคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์การถดถอย(Sbi)	✓	✓	✓	Sx = Standard deviation of x
• ค่าสถิติ T หรือ F ทดสอบความมีนัยสำคัญของ bi [Ho: βi=o]	✓	✓	✓	Sy = Standard deviation of y
• ค่าระดับนัยสำคัญของ T หรือ F	✓	✓		Sy.x = Standard error of estimate
				S(bi) = Standard error of coefficient
				T = bi/S(bi) สำหรับโปรแกรมBMDP, SAS
				F = $[bi/S(bi)]^2$ สำหรับโปรแกรม SPSS

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(13) ตารางวิเคราะห์ความแปรปรวนพร้อมกับ				$F = \frac{\text{Mean Square of reg'}}{\left(\begin{matrix}\text{Mean Square of}\\ \text{residual}\end{matrix}\right)}$
. ค่าสถิติ F ใช้ทดสอบความมีนัยสำคัญของสมการถดถอย $[Ho ; \beta_1=\beta_2=\ldots\beta_p = o]$	✓	✓	✓	$\frac{\sum(\hat{y}_i-\bar{y})^2/p}{\sum(y_i-\hat{y}_i)^2/(N-P-1)}$
. ค่าระดับนัยสำคัญของ F	✓	✓		
(14) ตารางวิเคราะห์ความแปรปรวนพร้อมกับค่าสถิติ F สำหรับทดสอบดังนี้				Sequential F-test ใช้ทดสอบสมมติฐาน Ho ; $\beta_1/\beta_1,\beta_2,\ldots$
. Sequential F-test (F-to-enter)	✓	✓		
. Partial F-test (F-to-remove)	✓	✓	✓	
(15) การวิเคราะห์ตัวแปรที่ยังไม่ได้เข้าไปในสมการถดถอย	✓		✓	
(16) ขั้นตอนที่ตัวแปรอิสระเข้าไปในสมการโดยการวิเคราะห์ทีละขั้นตอนตามวิธีของ Stepwise regression	✓		✓	
(17) ตารางสรุปการคัดเลือกตัวแปรซึ่งประกอบไปด้วยค่าสหสัมพันธ์, (r), r^2 การเปลี่ยนแปลงของค่าสัมประสิทธิ์การตัดสินใจ	✓		✓	(17) r = Simple correlation.
(18) ค่าความคลาดเคลื่อนที่เกิดจากผลต่างของค่าประมาณกับค่าสังเกต (ℓ_i)	✓	✓	✓	(18) $\ell_i = y_i - \hat{y}_i$
(19) ค่าสถิติ Durbin-Watson		✓	✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(20) ช่วงเชื่อมั่นของตัวแปรตาม (y) จากการประมาณ		✓		(20) ช่วงเชื่อมั่นของ y โปรแกรม SAS แสดงผลได้ 2 รูปแบบคือ $\mu_{y.x} = \hat{y} \pm t_{1-\alpha/2} S_{\hat{y}}$ และ $Y = \hat{y} \pm t_{1-\alpha/2} S_{y-\hat{y}}$
(21) กราฟของความสัมพันธ์ในรูปแบบต่าง ๆ	✓		✓	
(22) การเลือกสมการถดถอยโดยพิจารณาจากค่า R^2	✓	✓		
(23) การบังคับให้ตัวแปรอิสระบางตัวเข้าไปในสมการที่ต้องการวิเคราะห์	✓		✓	โดย $S_{\hat{y}} = S_{y.x} \left[\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - x)^2} \right]$
(24) การกำหนดให้ตัวแปรอิสระเข้าไปในสมการตามลำดับที่ต้องการ	✓		✓	
(25) การเลือกเกณฑ์คัดเลือกตัวแปรอิสระ	✓	✓	✓	$S_{y-\hat{y}} = S_{y.x} \left[1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right]$
(26) ผลของเมทริกซ์ที่ใช้ในการวิเคราะห์การถดถอย		✓		(26) เช่น $(x'x)$, $(x'x)^{-1}$

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 5 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//TESTBMDP JOB (9012A0,90,9087),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
    INCLUDE BMDP(BMDP1R)
    ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM    TITLE IS 'REGRESSION OF BMDP BY SIRICHAI'.
/INPUT      VARIABLES ARE 9.
            FORMAT IS '(A4,5FA.0,3FA.1)'.
/VARIABLE   NAMES ARE ID,AGE,HEIGHT,WEIGHT,BRTHPILL,CHOLSTRL,
                      ALBUMIN,CALCIUM,URICACID.
            MAXIMUM IS (6) 400.
            MINIMUM IS (6) 150.
            BLANKS ARE MISSING.
            LABEL IS ID.
(n)---------------------------------------
/REGRESS    DEPENDENT IS CHOLSTRL.
            INDEPENDENT ARE AGE,WEIGHT,URICACID.

/PRINT      DATA.
            CORRELATION.
            COVARIANCE.
            RREG.
/PLOT       RESIDUALS.
            VARIABLE IS URICACID.
```

ตารางที่ 5.1 (ต่อ)

```
/END
```

บัตรบันทึกข้อมูล

```
//
```

(ข) ถ้าต้องการทำการวิเคราะห์โดยวิธี Stepwise Regression ใช้โปรแกรมชุด P2R ระบุในบัตรใบที่ 4 แทน P1R ส่วนบัตรใบอื่นๆคงเดิม

```
/REGRESS  DEPENDENT IS CHOLSTRL.
```

(ค) ถ้าต้องการทำการวิเคราะห์โดยวิธี Forward และ Backward
ใช้โปรแกรมชุด P2R ระบุในบัตรใบที่ 4 แทน P1R และใช้บัตรต่อไปนี้แทนในส่วน ก. ส่วนบัตรใบอื่นๆ คงเดิม

```
/REGRESS  DEPENDENT IS CHOLSTRL.
          TITLE IS 'FORWARD STEPPING FOLLOWED BY BACKWARD'.
          ENTER = 1,50.
          REMOVE = 0,49.
```

รูปที่ 5.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT5

(2)

```
REGRESSION INTERCEPT. . . . . .NON-ZERO
GROUPING VARIABLE . . . . . .
WEIGHT VARIABLE . . . . . .
PRINT COVARIANCE MATRIX . . . . . YES
PRINT CORRELATION MATRIX. . . . . YES
PRINT RESIDUALS . . . . . YES
PROBIT PLOT . . . . . NO

NUMBER OF CASES READ. . . . . 188
   CASES WITH DATA MISSING OR BEYOND LIMITS . . 8
      REMAINING NUMBER OF CASES . . . . . 180
```

(10)

VARIABLE	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION	MINIMUM	MAXIMUM
2 AGE	33.53819	9.89836	0.2951[illegible]	19.00000	55.00000
3 HEIGHT	64.46597	2.48213	0.03850	57.00000	71.00000
4 WEIGHT	131.09384	20.49977	0.15637	94.00000	215.00000
5 BRTHPILL	1.50551	0.50136	0.33302	1.00000	2.00000
6 CHOLSTRL	235.83821	42.74364	0.18124	155.00000	990.00000
7 ALBUMIN	4.12052	0.35871	0.08706	3.20000	5.00000
8 CALCIUM	9.96773	0.47279	0.04743	8.80000	11.10000
9 URICACID	4.75551	1.12111	0.23575	2.20000	9.90000

(11) COVARIANCE MATRIX

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	97.9776							
HEIGHT	3	2.1922	6.1610						
WEIGHT	4	51.7972	24.0917	420.2407					
BRTHPILL	5	0.2793	0.2041	0.8236	0.2514				
CHOLSTRL	6	154.5264	1.2157	128.1221	1.9649	1827.0183			
ALBUMIN	7	-0.2802	-0.0052	-1.7253	-0.0423	0.8819	0.1287		
CALCIUM	8	-0.0404	0.1676	0.6272	-0.0149	5.1487	0.0768	0.2235	
URICACID	9	2.3143	0.3489	6.9781	0.0086	13.1294	0.0120	0.0881	1.2569

(11) CORRELATION MATRIX

		AGE 2	HEIGHT 3	WEIGHT 4	BRTHPILL 5	CHOLSTRL 6	ALBUMIN 7	CALCIUM 8	URICACID 9
AGE	2	1.0000							
HEIGHT	3	0.0892	1.0000						
WEIGHT	4	0.2553	0.4735	1.0000					
BRTHPILL	5	0.0563	0.1640	0.0801	1.0000				
CHOLSTRL	6	0.3652	0.0115	0.1462	0.0917	1.0000			
ALBUMIN	7	-0.0789	-0.0058	-0.2346	-0.2352	0.0575	1.0000		
CALCIUM	8	-0.0086	0.1428	0.0647	-0.0629	0.2548	0.6529	1.0000	
URICACID	9	0.2085	0.1254	0.3036	0.0154	0.2740	0.0299	0.1662	1.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REGRESSION TITLE. . . . . .WERNER BLOOD CHEMISTRY DATA
DEPENDENT VARIABLE. . . . . 6 CHOLSTRL
TOLERANCE 0.0100
ALL DATA CONSIDERED AS A SINGLE GROUP

(12) MULTIPLE R 0.4175 STD. ERROR OF EST. 39.1698
MULTIPLE R-SQUARE 0.1743

ANALYSIS OF VARIANCE

(13)

	SUM OF SQUARES	DF	MEAN SQUARE	F RATIO	P(TAIL)
REGRESSION	57004.242	3	19001.414	12.385	0.00000
RESIDUAL	270032.000	176	1534.273		

(12)

VARIABLE		COEFFICIENT	STD. ERROR	STD. REG COEFF	T	P(2 TAIL)
INTERCEPT		151.420				
AGE	2	1.390	0.309	0.322	4.497	0.000
WEIGHT	4	0.003	0.153	0.001	0.019	0.985
URICACID	9	7.871	2.769	0.206	2.843	0.005

CORRELATION MATRIX OF REGRESSION COEFFICIENTS

		AGE 2	WEIGHT 4	URICACID 9
AGE	2	1.0000		
WEIGHT	4	-0.2060	1.0000	
URICACID	9	-0.1422	-0.2648	1.0000

(18)

CASE LABEL	NO.	RESIDUAL	PREDICTED VALUE	VARIABLES 2 AGE 8 CALCIUM	3 HEIGHT 9 URICACID	4 WEIGHT	5 BRTHPILL	6 CHOLSTRL	7 ALBUMIN
2381	1	-24.9135	224.9135	22.0000*	67.0000*	144.0000	1.0000*	200.0000	4.3000
				9.3000	5.4000				
1946	-2	********** ***	239.1275	22.0000*	64.0000	160.0000*	2.0000	**********	3.5000*
				**********	7.2000**				
1610	3	30.4927	212.5073	25.0000	62.0000	128.0000	1.0000*	243.0000	4.1000
				10.4000	3.3000*				
1797	-4	********** ***	210.2096	25.0000	68.0000*	150.0000	2.0000	**********	3.9000
				9.6000	3.0000*				
561	5	-57.1798 *	215.1798	19.0000*	64.0000	125.0000	1.0000*	158.0000*	4.1000
				9.9000	4.7000				

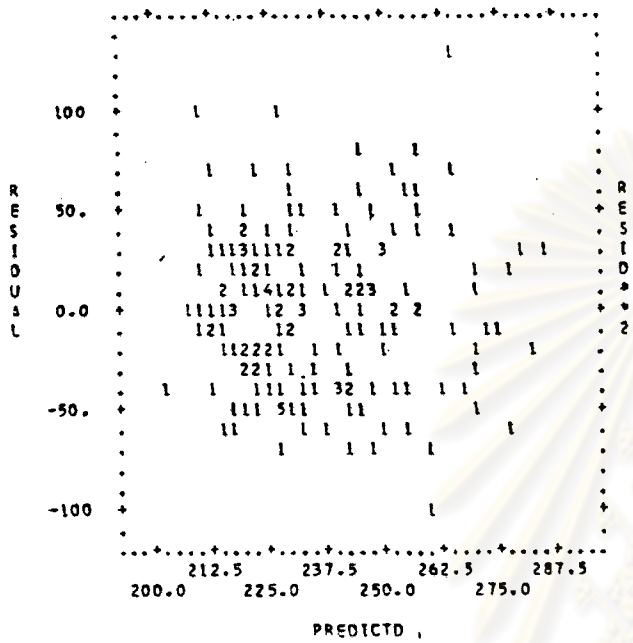
--- similar statistics for cases 6 to 185 ---

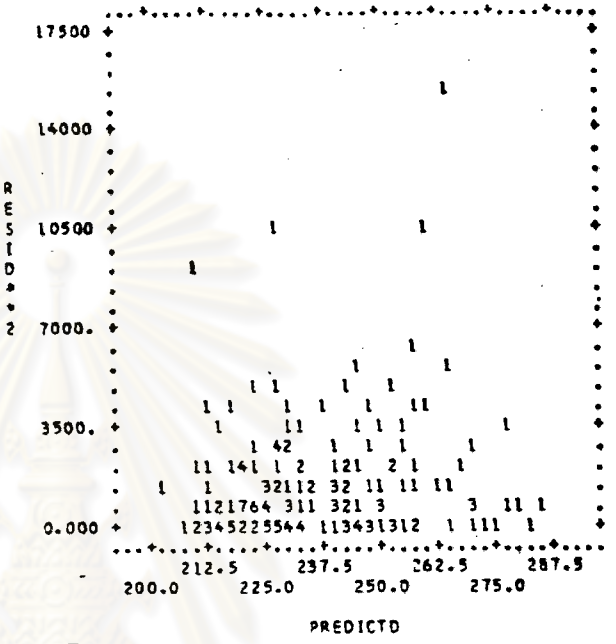
575	186	-41.6860 *	261.6860	53.0000*	55.0000	140.0000	2.0000	220.0000	4.0000
				10.7000*	4.6000				
2271	187	40.2979 *	264.7021	54.0000**	66.0000	158.0000*	1.0000*	305.0000*	4.2000
				10.3000	4.9000				
39	188	-56.5432 *	276.5432	54.0000**	60.0000*	170.0000*	2.0000	220.0000	3.5000*
				8.8000**	6.3000*				

NOTE - NEGATIVE CASE NUMBER DENOTES A CASE WITH MISSING VALUES.
THE NUMBER OF STANDARD DEVIATIONS FROM THE MEAN IS DENOTED BY UP TO 3 ASTERISKS TO THE RIGHT OF EACH RESIDUAL OR VARIABLE.
MISSING VALUES ARE DENOTED BY MORE THAN THREE ASTERISKS.

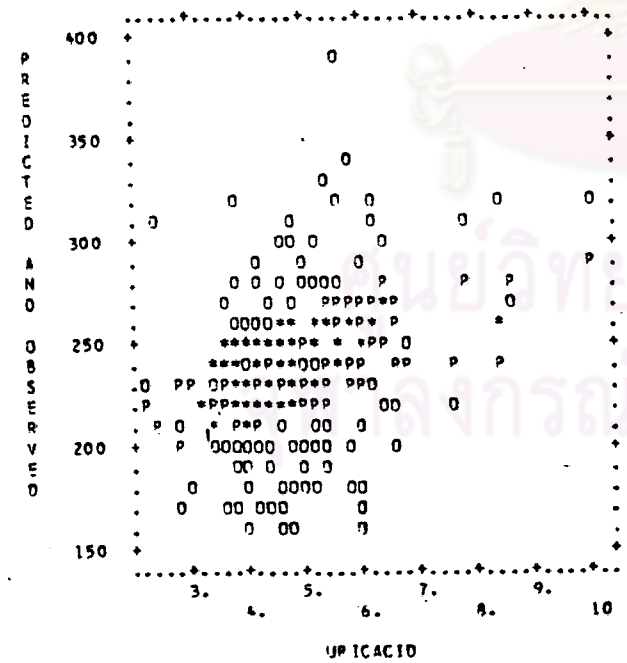
SERIAL CORRELATION OF RESIDUALS = -0.0781

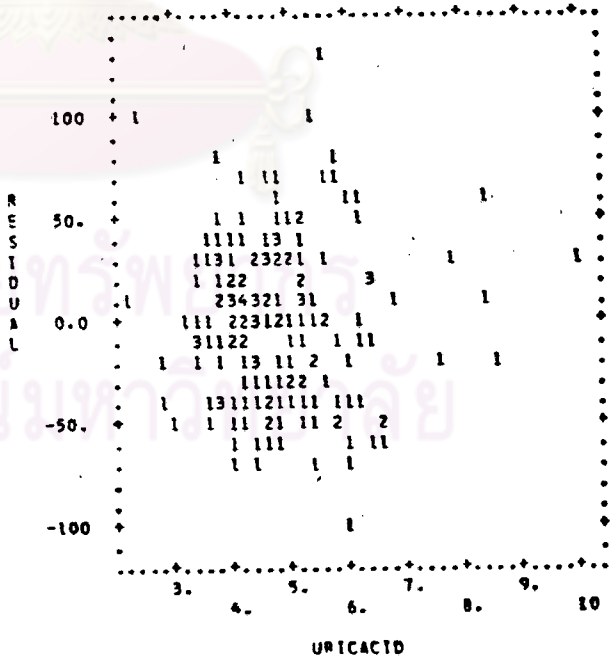
(R)


RESIDUAL
100
50.
0.0
-50.
-100
212.5
237.5
262.5
287.5
200.0
225.0
250.0
275.0
PREDICTD


RESID*2
17500
14000
10500
7000.
3500.
0.000
212.5
237.5
262.5
287.5
200.0
225.0
250.0
275.0
PREDICTD

(21)


PREDICTED AND OBSERVED
400
350
300
250
200
150
3.
5.
7.
9.
4.
6.
8.
10
URICACID


RESIDUAL
100
50.
0.0
-50.
-100
3.
5.
7.
9.
4.
6.
8.
10
URICACID

(7) BMDP2R - STEPWISE REGRESSION
HEALTH SCIENCES COMPUTING FACILITY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

PROGRAM REVISED SEPTEMBER 1977
MANUAL DATE — 1977

(ข)

--- Control Language read by P2R is printed and interpreted ---

```
REGRESSION INTERCEPT. . . . . .NON ZERO
WEIGHT VARIABLE . . . . . .
PRINT COVARIANCE MATRIX . . . . . NO
PRINT CORRELATION MATRIX. . . . . NO
PRINT ANOVA AT EACH STEP. . . . . YES
PRINT STEP OUTPUT . . . . . YES
PRINT REGRESSION COEFFICIENT SUMMARY TABLE. . . YES
PRINT PARTIAL CORRELATION SUMMARY TABLE . . . . NO
PRINT F-RATIO SUMMARY TABLE . . . . . NO
PRINT SUMMARY TABLE . . . . . YES
PRINT RESIDUALS AND DATA. . . . . NO

NUMBER OF CASES READ. . . . . 188
   CASES WITH DATA MISSING OR BEYOND LIMITS . . 9
      REMAINING NUMBER OF CASES . . . . . 180
```

(10)

VARIABLE NO.	NAME	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION	SKEWNESS	KURTOSIS	SMALLEST VALUE	LARGEST VALUE	SMALLEST STD SCORE	LARGEST STD SCORE
2	AGE	33.5382	9.8984	0.2951	0.4039	-0.9908	19.0000	55.0000	-1.4687	2.1682
3	HEIGHT	64.4660	2.4821	0.0385	-0.1185	-0.0402	57.0000	71.0000	-3.0079	2.6324
4	WEIGHT	131.0938	20.4998	0.1564	1.0572	1.9389	94.0000	215.0000	-1.8095	4.0930
5	BRTHMILL	1.5055	0.5014	0.3330	-0.0220	-2.0106	1.0000	2.0000	-1.0083	0.9863
6	CHOLSTRL	235.8382	42.7436	0.1812	0.4578	0.0640	155.0000	390.0000	-1.8912	3.6067
7	ALBUMIN	4.1205	0.3587	0.0871	-0.0678	-0.3606	3.2000	5.0000	-2.5662	2.4519
8	CALCIUM	9.9677	0.4728	0.0474	-0.0287	-0.5804	8.8000	11.1000	-2.4699	2.3949
9	URICACID	4.7555	1.1211	0.2358	1.1759	2.9185	2.2000	9.9000	-2.2794	4.5887

NOTE - KURTOSIS VALUES GREATER THAN ZERO INDICATE A DISTRIBUTION WITH HEAVIER TAILS THAN NORMAL DISTRIBUTION

```
REGRESSION TITLE. . . . . .WERNER BLOOD CHEMISTRY DATA
```

(25)

```
STEPPING ALGORITHM. . . . . .F
MAXIMUM NUMBER OF STEPS . . . . . 18
DEPENDENT VARIABLE. . . . . 6 CHOLSTRL
MINIMUM ACCEPTABLE F TO ENTER . . . . . 4.000.   4.000
MAXIMUM ACCEPTABLE F TO REMOVE. . . . . 3.900.   3.900
MINIMUM ACCEPTABLE TOLERANCE. . . . . 0.01000
```

(16)

```
STEP NO.   0

MULTIPLE R              0.0
MULTIPLE R-SQUARE       0.0
ADJUSTED R-SQUARE       0.0
STD. ERROR OF EST.     42.7436

ANALYSIS OF VARIANCE
                SUM OF SQUARES     DF    MEAN SQUARE     F RATIO
   REGRESSION  0.0                  0  0.0                  0.0
   RESIDUAL     327036.19         179  1827.018
```

VARIABLES IN EQUATION

VARIABLE		COEFFICIENT	STD. ERROR OF COEFF	STD REG COEFF	TOLERANCE	F TO REMOVE	LEVEL
(Y-INTERCEPT		235.838)					

VARIABLES NOT IN EQUATION

VARIABLE		PARTIAL CORR.	TOLERANCE	F TO ENTER	LEVEL
AGE	2	0.365	1.00000	27.399	1
HEIGHT	3	0.011	1.00000	0.023	1
WEIGHT	4	0.146	1.00000	3.889	1
BRTHMILL	5	0.092	1.00000	1.509	1
ALBUMIN	7	0.058	1.00000	0.591	1
CALCIUM	8	0.255	1.00000	12.356	1
URICACID	9	0.274	1.00000	14.[illegible]	1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STEP NO. 1
VARIABLE ENTERED 2 AGE

(12)
MULTIPLE R 0.3652
MULTIPLE R-SQUARE 0.1334
ADJUSTED R-SQUARE 0.1285
STD. ERROR OF EST. 39.9024

ANALYSIS OF VARIANCE

(13)

	SUM OF SQUARES	DF	MEAN SQUARE	F RATIO
REGRESSION	63624.603	1	63624.61	27.399
RESIDUAL	283411.63	178	1592.200	

VARIABLES IN EQUATION (14)

VARIABLE		COEFFICIENT	STD. ERROR OF COEFF	STD REG COEFF	TOLERANCE	F TO REMOVE	LEVEL
(Y-INTERCEPT		182.943)					
AGE	2	1.5772	0.3013	0.365	1.000	27.40	1

VARIABLES NOT IN EQUATION (15)

		PARTIAL CORR.	TOLERANCE	F TO ENTER	LEVEL
. HEIGHT	3	-0.023	0.99204	0.092	1
. WEIGHT	4	0.059	0.93484	0.616	1
. BRTHPILL	5	0.077	0.99683	1.043	1
. ALBUMIN	7	0.093	0.99377	1.545	1
. CALCIUM	8	0.277	0.99893	14.719	1
. URICACID	9	0.217	0.95651	8.770	1

(16)
STEP NO. 3
VARIABLE ENTERED 9 URICACID

MULTIPLE R 0.4750
MULTIPLE R-SQUARE 0.2256
ADJUSTED R-SQUARE 0.2124
STD. ERROR OF EST. 37.9329

ANALYSIS OF VARIANCE

	SUM OF SQUARES	DF	MEAN SQUARE	F RATIO
REGRESSION	73789.250	3	24596.41	17.094
RESIDUAL	253246.94	176	1438.903	

VARIABLES IN EQUATION

VARIABLE		COEFFICIENT	STD. ERROR OF COEFF	STD REG COEFF	TOLERANCE	F TO REMOVE	LEVEL
(Y-INTERCEPT		-49.738)					
AGE	2	1.4359	0.2932	0.333	0.955	23.99	1
CALCIUM	8	20.7920	6.0876	0.230	0.970	11.67	1
URICACID	9	6.3444	2.6248	0.166	0.928	5.84	1

VARIABLES NOT IN EQUATION

		PARTIAL CORR.	TOLERANCE	F TO ENTER	LEVEL
. HEIGHT	3	-0.083	0.96406	1.221	1
. WEIGHT	4	-0.005	0.86875	0.004	1
. BRTHPILL	5	0.097	0.99272	1.655	1
. ALBUMIN	7	-0.032	0.78831	0.185	1

F-LEVELS(4.000, 3.900) OR TOLERANCE INSUFFICIENT FOR FURTHER STEPPING

STEPWISE REGRESSION COEFFICIENTS

VARIABLES STEP	0 Y-INTCPT	2 AGE	3 HEIGHT	4 WEIGHT	5 BRTHPILL	7 ALBUMIN	8 CALCIUM	9 URICACID
0	235.8382*	1.5772	0.1973	0.3049	7.8169	6.8533	23.0333	10.4459
1	182.9431*	1.5772*	-0.3668	0.1182	6.0835	10.3523	23.3203	7.8848
2	-49.8298*	1.5868*	-1.0313	0.0801	7.4891	-4.5760	23.3203*	6.3444
3	-49.7384*	1.4359*	-1.2847	-0.0098	7.2891	-3.8369	20.7920*	6.3444*

NOTE-
1) REGRESSION COEFFICIENTS FOR VARIABLES IN THE EQUATION ARE INDICATED BY AN ASTERISK
2) THE REMAINING COEFFICIENTS ARE THOSE WHICH WOULD BE OBTAINED IF THAT VARIABLE WERE TO ENTER IN THE NEXT STEP

(17) SUMMARY TABLE

STEP NO.	VARIABLE ENTERED	REMOVED	MULTIPLE R	RSQ	INCREASE IN RSQ	F-TO-ENTER	F-TO-REMOVE	NUMBER OF INDEPENDENT VARIABLES INCLUDED
1	2 AGE		0.3652	0.1334	0.1334	27.3989		1
2	8 CALCIUM		0.4471	0.1999	0.0665	14.7188		2
3	9 URICACID		0.4750	0.2256	0.0257	5.3421		3

(A)

```
REGRESSION TITLE. . . . . .FORWARD STEPPING FOLLOWED BY BACKWARD STEPPING
STEPPING ALGORITHM. . . . . .F
MAXIMUM NUMBER OF STEPS . . . . . 18
DEPENDENT VARIABLE. . . . . 6 CHOLSTRL
MINIMUM ACCEPTABLE F TO ENTER . . . . . 1.000,  50.000
MAXIMUM ACCEPTABLE F TO REMOVE. . . . . 0.0  .  49.000
MINIMUM ACCEPTABLE TOLERANCE. . . . . 0.01000
```

--- results for steps 0 to 4 ---

```
STEP NO.    3
VARIABLE ENTERED    3 HEIGHT

MULTIPLE R              0.4907
MULTIPLE R-SQUARE       0.2408
ADJUSTED R-SQUARE       0.2190
STD. ERROR OF EST.     37.7742

ANALYSIS OF VARIANCE
                SUM OF SQUARES      DF     MEAN SQUARE     F RATIO
REGRESSION          78756.938        5       15751.39       11.039
RESIDUAL           248279.19       174       1426.892
```

VARIABLES IN EQUATION								VARIABLES NOT IN EQUATION				
VARIABLE		COEFFICIENT	STD. ERROR OF COEFF	STD REG COEFF	TOLERANCE	F TO REMOVE	LEVEL	VARIABLE		PARTIAL CORR.	TOLERANCE	F TO ENTER	LEVEL
(Y-INTERCEPT		21.526)											
AGE	2	1.4415	0.2929	0.334	0.948	24.22	1	WEIGHT	4	0.039	0.68591	0.259	1
HEIGHT	3	-1.5854	1.1757	-0.092	0.936	1.82	1	ALBUMIN	7	-0.014	0.74640	0.034	1
BRTHPILL	5	8.6071	5.7359	0.101	0.964	2.25	1						
CALCIUM	8	22.4556	6.1354	0.248	0.947	13.40	1						
URICACID	9	6.5983	2.6238	0.173	0.921	6.32	1						

F-LEVELS(1.000, 0.0) OR TOLERANCE INSUFFICIENT FOR FURTHER STEPPING

--- results for steps 6 to 9 ---

SUMMARY TABLE

STEP NO.	VARIABLE ENTERED	VARIABLE REMOVED	MULTIPLE R	RSQ	INCREASE IN RSQ	F-TO-ENTER	F-TO-REMOVE	NUMBER OF INDEPENDENT VARIABLES INCLUDED
1	2 AGE		0.3652	0.1334	0.1334	27.3989		1
2	8 CALCIUM	(5)	0.4471	0.1999	0.0665	14.7188		2
3	9 URICACID		0.4750	0.2256	0.0257	5.8421		3
4	5 BRTHPILL		0.4826	0.2329	0.0073	1.6554		4
5	3 HEIGHT		0.4907	0.2408	0.0079	1.8183		5
6		3 HEIGHT	0.4826	0.2329	-0.0079		1.8183	4
7		5 BRTHPILL	0.4750	0.2256	-0.0073		1.6554	3
8		9 URICACID	0.4471	0.1999	-0.0257		5.8421	2
9	(6)	8 CALCIUM	0.3652	0.1334	-0.0665		14.7187	1

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 5 โดยใช้ข้อมูลจาก DATA SAMPLE 2

```
//S03SOPOK JOB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE1;
INPUT ID $ 1-4 AGE 5-8 HEIGHT 9-12 WEIGHT 13-16 BRTHPILL 17-20
    CHOLSTRL 21-24 ALBUMIN 25-28 1 CALCIUM 29-31 1 URICACID 33-36 1;
CARDS;
```

บัตรบันทึกข้อมูล

(ก)
```
PROC REGR S USSCP SSCP COV COR;
 MODEL CHOLSTRL=AGE WEIGHT URICACID/X S I P CLM ALPHA=.01 CONF YPY;
 ID AGE WEIGHT;
 TITLE 'REGRESSION OF SAS BY SIRICHAI';
```

(ข)
```
PROC STEPWISE FORWARD BACKWARD MAXR MINR;
 VARIABLES AGE HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
           CHOLSTRL;
```

(ค)
```
PROC RSQUARE;
 VARIABLES AGE HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
           CHOLSTRL;
 TITLE 'BEST REGRESSION OF SAS BY SIRICHAI';
/*
//
```

รูปที่ 5.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT 5

```
**********************************************************************************
*                                                                                *
*  PROC REGR    :  REGRESSION OF SAS BY SIRICHAI                                 *
*                                                                                *
*  DATA SET     :  DATA0001      NUMBER OF VARIABLES =  4      NUMBER OF CLASSES =  0
*                                                                                *
(2)*  VARIABLES   :  AGE WEIGHT CHOLSTRL URICACID                                *
*                                                                                *
**********************************************************************************
```

(n)

WARNING: 5 OUT OF THE 188 RECORDS ON THE DATA SET HAVE MISSING VALUES FOR ONE OR MORE VARIABLES NEEDED BY THE PROCEDURE AND HENCE WILL NOT BE INCLUDED (EXCEPT FOR PREDICTED VALUES) IN THE ANALYSIS TO FOLLOW.

REGRESSION OF SAS BY SIRICHAI

N = 183

(10) SIMPLE STATISTICS

VARIABLE	SUM	MEAN	UNCORRECTED SS	CORRECTED SS	VARIANCE	STANDARD DEV
AGE	6181.00000000	33.77595628	227105.00000000	18335.81420765	100.74623191	10.03724225
WEIGHT	24036.00000000	131.34426230	3234624.00000000	77633.31147541	426.55606646	20.65324800
CHOLSTRL	43142.00000000	235.74863388	10497862.00000000	327194.43715847	1797.77163274	42.40013718
URICACID	861.00000000	4.74316940	4347.24000000	230.16896175	1.26466462	1.12457309

REGRESSION OF SAS BY SIRICHAI

N = 183

UNCORRECTED SUMS OF SQUARES & CROSS PRODUCTS

	AGE	WEIGHT	CHOLSTRL	URICACID
AGE	227105.00000000	821819.00000000	1484476.00000000	29690.00000000
WEIGHT	821819.00000000	3234624.00000000	5689086.00000000	115290.40000000
CHOLSTRL	1484476.00000000	5689086.00000000	10497862.00000000	206997.20000000
URICACID	29690.00000000	115290.40000000	206997.20000000	4347.24000000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REGRESSION OF SAS BY SIRICHAI

(26) THE X'X MATRIX

	INTERCPT	AGE	WEIGHT	URICACID
INTERCPT	183.00000000	6181.00000000	24036.00000000	868.00000000
AGE	6181.00000000	227105.00000000	821819.00000000	29690.00000000
WEIGHT	24036.00000000	821819.00000000	3234624.00000000	115290.40000000
URICACID	868.00000000	29690.00000000	115290.40000000	4347.24000000

REGRESSION OF SAS BY SIRICHAI

(26) THE SWEPT-OUT X'X MATRIX

	INTERCPT	AGE	WEIGHT	URICACID
INTERCPT	183.00000000	6181.00000000	24036.00000000	868.00000000
AGE	0.0	18345.81420765	9980.11475410	372.46994536
WEIGHT	0.0	0.0	72201.17252718	1080.84633662
URICACID	0.0	0.0	0.0	206.42249258

REGRESSION OF SAS BY SIRICHAI

(26) THE X'X INVERSE MATRIX , RANK = 4

	INTERCPT	AGE	WEIGHT	URICACID
INTERCPT	0.27153066	-0.00085107	-0.00139293	-0.01146215
AGE	-0.00085107	0.00005936	-0.00000666	-0.00005894
WEIGHT	-0.00139293	-0.00000666	0.00001494	-0.00007252
URICACID	-0.01146215	-0.00005894	-0.00007252	0.00484443

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

N = 183

CORRECTED SUMS OF SQUARES & CROSS PRODUCTS

	AGE	WEIGHT	CHOLSTRL	URICACID
AGE	18335.81420765	9980.11475410	27313.69398907	372.46994536
WEIGHT	9980.11475410	77633.31147541	22631.83606557	1283.58032787
CHOLSTRL	27313.69398907	22631.83606557	327194.43715847	2367.38579235
URICACID	372.46994536	1283.58032787	2367.38579235	230.14896175

REGRESSION OF SAS BY SIRICHAI

N = 183

(11) COVARIANCE MATRIX

	AGE	WEIGHT	CHOLSTRL	URICACID
AGE	100.74623191	54.83579535	150.07524170	2.04653816
WEIGHT	54.83579535	426.55665646	124.35074761	7.05263916
CHOLSTRL	150.07524170	124.35074761	1797.77163274	13.00761424
URICACID	2.04653816	7.05263916	13.00761424	1.26466462

REGRESSION OF SAS BY SIRICHAI

N = 183

(11) CORRELATION COEFFICIENTS / PROB > |R| UNDER HO: RHO=0

	AGE	WEIGHT	CHOLSTRL	URICACID
AGE	1.000000 0.0000	0.264522 0.0006	0.352637 0.0001	0.181308 0.0134
WEIGHT	0.264522 0.0006	1.000000 0.0000	0.142001 0.0521	0.303652 0.0001
CHOLSTRL	0.352637 0.0001	0.142001 0.0521	1.000000 0.0000	0.272799 0.0004
URICACID	0.181308 0.0134	0.303652 0.0001	0.272799 0.0004	1.000000 0.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REGRESSION OF SAS BY SIRICHAI

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR DEPENDENT VARIABLE CHOLSTRL

(2) (13) (12) (n)

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB > F	R-SQUARE	C.V.
REGRESSION	3	55462.09522710	18487.36507403	12.17830	0.0001	0.16950606	16.52704 %
ERROR	179	271732.34193637	1518.05777618				
CORRECTED TOTAL	182	327194.43715847				STD DEV	CHOLSTRL MEAN
						38.96226092	235.74863

(14) (14)

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB > F	PARTIAL SS	F VALUE	PROB > F
AGE	1	40687.46939077	26.80232	0.0001	29850.82338163	19.66383	0.0001
WEIGHT	1	835.12081139	0.55012	0.5340	16.02096166	0.01055	0.9149
URICACID	1	13939.50501995	9.18246	0.0032	13939.50501995	9.18246	0.0032

SOURCE	B VALUES	T FOR HO:B=0	PROB > |T|	STD ERR B	STD B VALUES
INTERCEPT	153.84287068	7.57746	0.0001	20.30269035	0.0
AGE	1.33112523	4.43439	0.0001	0.30018218	0.31511281
WEIGHT	-0.01546888	-0.10273	0.9149	0.15057703	-0.00753494
URICACID	8.21760321	3.03026	0.0032	2.71184986	0.21799437

(18)

OBS NUMBER	AGE	WEIGHT	OBSERVED VALUE	PREDICTED VALUE	RESIDUAL	LOWER 99% CL FOR MEAN	UPPER 99% CL FOR MEAN
1	22	144	200.00000000	225.27516450	-25.27516450	210.97667627	239.57365273
---------- ค่าสังเกตที่ 2 ถึงค่าสังเกตที่ 187 จะให้ผลเช่นเดียวกัน ----------							
188	54	170	220.00000000	274.86482398	-54.86482398	254.33600358	295.39364439

SUM OF RESIDUALS	=	-0.00000000
SUM OF SQUARED RESIDUALS	=	271732.34193636
SUM OF SQUARED RESIDUALS - ERROR SS	=	-0.00000001
FIRST ORDER AUTOCORRELATION OF RESIDUALS	=	-0.11550013
(19) DURBIN-WATSON D	=	2.21601845

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REGRESSION OF SAS BY SIRICHAI

(5) FORWARD SELECTION PROCEDURE FOR DEPENDENT VARIABLE CHOLSTRL (๒)

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.13338717	AGE
2	0.19991924	AGE CALCIUM
3	0.22562377	AGE CALCIUM URICACID
4	0.23289065	AGE BRTHPILL CALCIUM URICACID
5	0.24081475	AGE HEIGHT BRTHPILL CALCIUM URICACID

NO OTHER VARIABLES MET THE REQUIRED 0.5000 SIGNIFICANCE LEVEL FOR ENTRY INTO THE MODEL

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR THE ABOVE MODEL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB GT F	R-SQUARE	C.V.
REGRESSION	5	78755.65273878	15751.13054776	11.03841	0.0001	0.24081475	16.01707
ERROR	174	248282.67503900	1426.91192551				
CORRECTED TOTAL	179	327038.32777778					

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB GT F	PARTIAL SS	F VALUE	PROB GT F
AGE	1	43622.71572232	30.57141	0.0001	34556.22619976	24.21749	0.0001
CALCIUM	1	21750.53916017	15.24869	0.0003	19114.65325985	13.39502	0.0006
URICACID	1	8406.36664378	5.89130	0.0154	9023.96616434	6.32417	0.0123
BRTHPILL	1	2373.27694636	1.66323	0.1959	3213.04451282	2.25175	0.1313
HEIGHT	1	2594.75426614	1.81844	0.1758	2594.75426614	1.81844	0.1758

SOURCE	B VALUES	T FOR HO B=0	PROB GT ABS(T)	STD ERR B	STD B VALUES
MEAN	21.52790042				
AGE	1.44154062	4.92113	0.0001	0.29292196	0.33381192
CALCIUM	22.45554890	3.66003	0.0006	6.13534582	0.24638108
URICACID	6.59834086	2.51478	0.0123	2.62382340	0.17306196
BRTHPILL	8.60719675	1.50058	0.1313	5.73590491	0.10095826
HEIGHT	-1.58540837	-1.34850	0.1758	1.17568672	-0.09206408

REGRESSION OF SAS BY SIRICHAI

(6) BACKWARD ELIMINATION PROCEDURE FOR DEPENDENT VARIABLE CHOLSTRL (ข)

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
7	0.24195431	AGE HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
6	0.24194844	AGE HEIGHT WEIGHT BRTHPILL CALCIUM URICACID
5	0.24081475	AGE HEIGHT BRTHPILL CALCIUM URICACID
4	0.23288065	AGE BRTHPILL CALCIUM URICACID
3	0.22562377	AGE CALCIUM URICACID

THE VARIABLES IN THE ABOVE MODEL HAVE ALL BEEN DEEMED SIGNIFICANT AT THE 0.1000 SIGNIFICANCE LEVEL

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR THE ABOVE MODEL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB GT F	R-SQUARE	C.V.
REGRESSION	3	73787.62152627	24595.87384209	17.09323	0.0001	0.22562377	16.08425
ERROR	176	253250.70625151	1438.92446734				
CORRECTED TOTAL	179	327038.32777778					

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB GT F	PARTIAL SS	F VALUE	PROB GT F
AGE	1	43622.71572232	30.31620	0.0001	34517.05811103	23.98481	0.0001
CALCIUM	1	21758.53916017	15.12139	0.0003	16785.72055130	11.66546	0.0012
URICACID	1	8406.36664378	5.84212	0.0158	8406.36664378	5.84212	0.0158

SOURCE	B VALUES	T FOR HO B=0	PROB GT ABS(T)	STD ERR B	STD B VALUES
MEAN	-49.74100180				
AGE	1.43595548	4.89782	0.0001	0.29318251	0.33251859
CALCIUM	20.77184036	3.41547	0.0012	6.08754152	0.27958748
URICACID	6.34438915	2.41705	0.0158	2.62485096	0.16640514

REGRESSION OF SAS BY SIRICHAI

(7) STEPWISE REGRESSION PROCEDURE FOR DEPENDENT VARIABLE CHOLSTRL

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.13338717	AGE
2	0.19991924	AGE CALCIUM
3	0.27562377	AGE CALCIUM URICACID
4	0.23288065	AGE BRTHFILL CALCIUM URICACID
3	0.22562377	AGE CALCIUM URICACID

THE VARIABLES IN THE ABOVE MODEL HAVE ALL BEEN DEEMED SIGNIFICANT AT THE 0.1000 SIGNIFICANCE LEVEL

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR THE ABOVE MODEL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB GT F	R-SQUARE	C.V.
REGRESSION	3	73787.62152627	24595.87384209	17.09323	0.0001	0.22562377	16.08435
ERROR	176	253250.70625151	1438.92446734				
CORRECTED TOTAL	179	327038.32777778					

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB GT F	PARTIAL SS	F VALUE	PROB GT F
AGE	1	43622.71572232	30.31620	0.0001	34517.85811583	23.98865	0.0001
CALCIUM	1	21758.53916017	15.12139	0.0003	16785.72059130	11.66546	0.0017
URICACID	1	8406.36664378	5.84212	0.0158	8406.36664378	5.84212	0.0158

SOURCE	B VALUES	T FOR HO B=0	PROB GT ABS(T)	STD ERR B	STD B VALUES
MEAN	-48.74100180				
AGE	1.43595548	4.89782	0.0001	0.29318651	0.33251859
CALCIUM	20.79186036	3.41547	0.0012	6.08754152	0.22998249
URICACID	6.34438915	2.41705	0.0158	2.62485096	0.16640514

REGRESSION OF SAS BY SIRICHAI

(ข)

(8) MAXIMUM R-SQUARE IMPROVEMENT FOR DEPENDENT VARIABLE CHOLSTRL

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.13338717	AGE

THE ABOVE MODEL IS THE 'BEST' 1 VARIABLE MODEL FOUND BY THE MAXIMUM R-SQUARE IMPROVEMENT PROCEDURE

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR THE ABOVE MODEL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB GT F	R-SQUARE	C.V.
REGRESSION	1	43622.71572232	43622.71572232	27.39737	0.0001	0.13338717	16.91946
ERROR	178	283415.61205546	1592.22253964				
CORRECTED TOTAL	179	327038.32777778					

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB GT F	PARTIAL SS	F VALUE	PROB GT F
AGE	1	43622.71572232	27.39737	0.0001	43622.71572232	27.39737	0.0001

SOURCE	B VALUES	T FOR HO B=0	PROB GT ABS(T)	STD ERR B	STD B VALUES
MEAN	182.94192569				
AGE	1.57718293	5.23425	0.0001	0.30131979	0.36522201

---------- จะทำการวิเคราะห์สมการถดถอยที่มีตัวแปรอิสระเพิ่มขึ้นอีกครั้งละ 1 ตัวแปรจนกระทั่งไม่มี ----------

ตัวแปรอิสระใดถูกคัดเลือกเข้าไปในสมการได้อีก

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REGRESSION OF SAS BY SIRICHAI

(9) MINIMUM R-SQUARE IMPROVEMENT FOR DEPENDENT VARIABLE CHOLSTRL

(ข)

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.00013000	HEIGHT
1	0.00330541	ALBUMIN
1	0.00840410	BETHILL
1	0.02137767	WEIGHT
1	0.06490102	CALCIUM
1	0.07506295	URICACID
1	0.13338717	AGE

THE ABOVE MODEL IS THE 'BEST' 1 VARIABLE MODEL FOUND BY THE MINIMUM R-SQUARE IMPROVEMENT PROCEDURE

ANALYSIS OF VARIANCE TABLE , REGRESSION COEFFICIENTS , AND STATISTICS OF FIT FOR THE ABOVE MODEL

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB GT F	R-SQUARE	C.V.
REGRESSION	1	43622.71572232	43622.71572232	27.39737	0.0001	0.1333P717	16.91946
ERROR	178	283415.61205546	1592.22253964				
CORRECTED TOTAL	179	327038.32777778					

SOURCE	DF	SEQUENTIAL SS	F VALUE	PROB GT F	PARTIAL SS	F VALUE	PROB GT F
AGE	1	43622.71572232	27.39737	0.0001	43622.71572232	27.39737	0.0001

SOURCE	B VALUES	T FOR HO B=0	PROB GT ABS(T)	STD ERR B	STD B VALUES
MEAN	182.94192569				
AGE	1.57718293	5.23425	0.0001	0.30131975	0.36522[illegible]

---------- จะทำการวิเคราะห์สมการถดถอยที่มีตัวแปรอิสระเพิ่มขึ้นอีกครั้งละ 1 ตัวเช่นเดียวกับเทคนิค MAXIMUM ----------

(22) ALL POSSIBLE REGRESSION MODELS FOR DEPENDENT VARIABLE CHLSTRL

NUMBER IN MODEL	R-SQUARE	VARIABLES IN MODEL
1	0.00013000	HEIGHT
1	0.00330541	ALBUMIN
1	0.00840410	BRTHPILL
1	0.02137767	WEIGHT
1	0.04490102	CALCIUM
1	0.07506295	URICACID
1	0.13338717	AGE
2	0.00344215	HEIGHT ALBUMIN
2	0.00841766	HEIGHT BRTHPILL
2	0.01501893	BRTHPILL ALBUMIN
2	0.02568657	HEIGHT WEIGHT

-------- ละให้ค่า R-SQUARE สำหรับตัวแปรอิสระ 3,4,5 ในรูปแบบเช่นเดียวกัน ---------

5	0.24081475	AGE HEIGHT BRTHPILL CALCIUM URICACID
6	0.14457493	HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
6	0.19490462	AGE HEIGHT WEIGHT BRTHPILL ALBUMIN URICACID
6	0.21864245	AGE HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM
6	0.23271958	AGE HEIGHT WEIGHT ALBUMIN CALCIUM URICACID
6	0.23315419	AGE WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
6	0.24096554	AGE HEIGHT BRTHPILL ALBUMIN CALCIUM URICACID
6	0.24194844	AGE HEIGHT WEIGHT BRTHPILL CALCIUM URICACID
7	0.24195431	AGE HEIGHT WEIGHT BRTHPILL ALBUMIN CALCIUM URICACID

ตารางที่ 5.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 5 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK        MR.SIRICHAI PONGWICHAI B2810178
// EXEC PROC=$$SPSS79
RUN NAME              'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME             'DATA SAMPLE1'
VARIABLE LIST         ID AGE HEIGHT WEIGHT BRTHPILL CHOLSTRL ALBUMIN
                      CALCIUM URICACID
INPUT MEDIUM          CARD
N OF CAGES            188
INPUT FORMAT          FIXED(A1,5F1.0,3F1.1)
MISSING VALUES        AGE TO URICACID(0)
VALUE LABELS          BRTHPILL(1)NOPILL(2)PILL
RECODE                CHOLSTRL(LO THRU 149=0)(201 THRU HI=0)
READ INPUT DATA
```

(ก)

```
TASK NAME             'MULTIPLE REGRESSION OF SPSS BY SIRICHAI'
REGRESSION            VARIABLES = AGE WEIGHT CHOLSTRL URICACID/
                      REGRESSION = CHOLSTRL WITH AGE WEIGHT URICACID(1)
                                   RESID = .10
OPTIONS               2
```

(ข)

```
STATISTICS            1,5,6,7
```

ตารางที่ 5.3 (ต่อ)

```
(ย) TASK NAME      'BEST REGRESSIØN ØF SPSS BY SIRICHAI'
REGRESSIØN         VARIABLES = AGE TØ URICACID/
                   REGRESSIØN = CHØLSTRL (3,0.001,0.01) WITH AGE TØ
                   BRTHDILL. ALBUMIN TØ URICACID(5) RESID = 0.10/
ØPTIØNS            2
STATISTICS         1,5,6,7
FINISH
/*
/&
* $$ EØJ
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 5.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT 5 (n)

CORRELATION COEFFICIENTS

A VALUE OF 99.00000 IS PRINTED
IF A COEFFICIENT CANNOT BE COMPUTED.

LOWER TRIANGLE: CORRELATION COEFFICIENTS
UPPER TRIANGLE: N OF CASES FOR CORRELATION

(11)

	AGE	WEIGHT	CHOLSTRL	URICACID
AGE	188.	186.	186.	187.
WEIGHT	0.25211	186.	184.	185.
CHOLSTRL	0.36779	0.14742	186.	185.
URICACID	0.17758	0.30433	0.26598	187.

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS' 04/05/81 PAGE 23
'MULTIPLE REGRESSION OF SPSS BY SIRICHAI'
FILE 'DATA (CREATION DATE = 04/05/81) SAMPLE1'

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * VARIABLE LIST 1
REGRESSION LIST 1

(2) DEPENDENT VARIABLE.. CHOLSTRL

VARIABLE(S) ENTERED ON STEP NUMBER 1.. AGE
WEIGHT
URICACID

(12)

MULTIPLE R	0.42053
R SQUARE	0.17685
ADJUSTED R SQUARE	0.16313
STANDARD ERROR	38.93305

(13)

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	3.	58609.01818	19536.33939	12.89060
RESIDUAL	180.	272798.82315	1515.54902	

---------- VARIABLES IN THE EQUATION ----------

VARIABLE	B	BETA	STD ERROR B	F
AGE	1.391968	0.33078	0.29595	22.137
WEIGHT	0.21837350-02	0.00106	0.14960	0.000
URICACID	7.608906	0.20691	2.62639	8.393
(CONSTANT)	152.4833			

(12)

---------- VARIABLES NOT IN THE EQUATION ----------

VARIABLE	BETA IN	PARTIAL	TOLERANCE	F

ALL VARIABLES ARE IN THE EQUATION

STATISTICS WHICH CANNOT BE COMPUTED ARE PRINTED AS ALL NINES.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(n)

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * VARIABLE LIST 1
REGRESSION LIST 1

DEPENDENT VARIABLE.. CHOLSTRL

(17) SUMMARY TABLE

VARIABLE	MULTIPLE R	R SQUARE	RSQ CHANGE	SIMPLE R	B	BETA
AGE	0.36779	0.13527	0.13527	0.36779	1.391968	0.33078
WEIGHT	0.37211	0.13847	0.00319	0.14742	0.21837350-02	0.00106
URICACID	0.42053	0.17685	0.03838	0.26598	7.608906	0.20691
(CONSTANT)					152.4889	

* M U L T I P L E R E G R E S S I O N *

DEPENDENT VARIABLE: CHOLSTRL FROM VARIABLE LIST 1
REGRESSION LIST 1

(18) (21)

SEQNUM	OBSERVED CHOLSTRL	PREDICTED CHOLSTRL	RESIDUAL	PLOT OF STANDARDIZED RESIDUAL -2.0 -1.0 0.0 1.0 2.0
1	200.0000	224.5147	-24.51471	* |
2	MISSING**	230.2457	MISSING**	|
3	243.0000	212.6770	30.32301	| *

---------- ค่าสังเกตที่ 4-180 จะให้ผลเช่นเดียวกัน ----------

181	215.0000	234.8017	-19.80170	* |
182	270.0000	229.4951	40.50487	| *
183	237.0000	238.5397	-1.539673	* |
184	200.0000	254.4834	-54.48343	|
185	270.0000	240.7677	29.23225	| *
186	230.0000	221.0130	8.987017	* |
187	255.0000	252.9192	2.080810	* |
188	275.0000	240.7144	34.28563	|

DURBIN-WATSON TEST OF RESIDUAL DIFFERENCES COMPARED BY CASE ORDER (SEQNUM).

VARIABLE LIST 1, REGRESSION LIST 1. DURBIN-WATSON TEST 2.10332

(19)

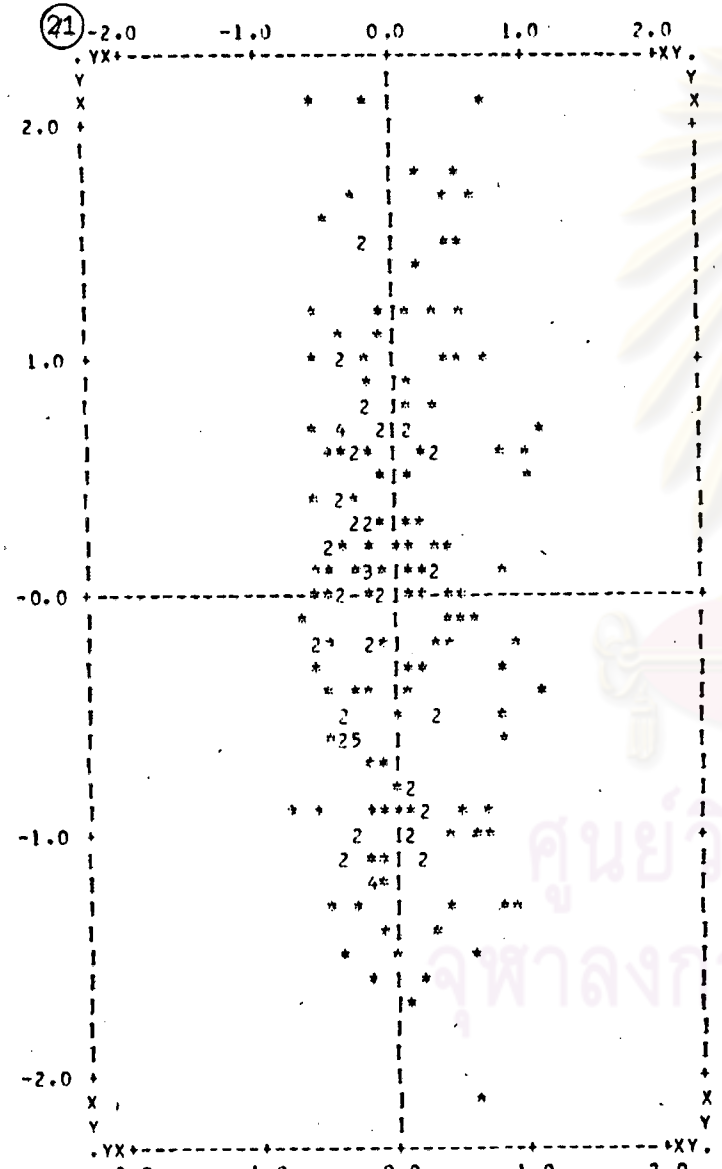
ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

* * * * * * * PLOT: STANDARDIZED RESIDUAL (DOWN) -- PREDICTED STANDARDIZED DEPENDENT VARIABLE (ACROSS) * * * * * * *

(n)

DEPENDENT VARIABLE: CHOLSTRL VARIABLE LIST 1
REGRESSION LIST 1

(21)



ROWS,COLUMNS Y: VALUES OUTSIDE (-3.0,3.0)

ROWS,COLUMNS X: VALUES IN (-3.0,-2.05) OR (2.05,3.0)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

'BEST REGRESSION OF SPSS BY SIRICHAI'

FILE 'DATA (CREATION DATE = 04/05/81) SAMPLE1'

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * * * * * VARIABLE LIST 1
REGRESSION LIST 1

DEPENDENT VARIABLE.. CHOLSTRL (7)

VARIABLE(S) ENTERED ON STEP NUMBER 1.. AGE

(12)

MULTIPLE R	0.36779
R SQUARE	0.13527
ADJUSTED R SQUARE	0.13049
STANDARD ERROR	39.68188

(13)

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	1.	44584.92288	44584.92288	28.31415
RESIDUAL	181.	285011.94664	1574.65164	

------------------ VARIABLES IN THE EQUATION ------------------

VARIABLE	B	BETA	STD ERROR B	F
AGE	1.547717	0.36779	0.29086	28.314
(CONSTANT)	183.8081			

(15)

------------- VARIABLES NOT IN THE EQUATION -------------

VARIABLE	BETA IN	PARTIAL	TOLERANCE	F
HEIGHT	-0.00986	-0.01056	0.99348	0.020
WEIGHT	0.05841	0.06078	0.93644	0.668
ORTHPILL	0.07580	0.08151	1.00000	1.204
ALBUMIN	0.08025	0.08608	0.99481	1.344
CALCIUM	0.26376	0.28349	0.99894	15.730
URICACID	0.20720	0.21927	0.96846	9.092

* *

(16)

VARIABLE(S) ENTERED ON STEP NUMBER 2.. CALCIUM

MULTIPLE R	0.45251
R SQUARE	0.20477
ADJUSTED R SQUARE	0.19593
STANDARD ERROR	38.15947

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	2.	67490.68862	33745.34431	23.17443
RESIDUAL	180.	262106.18090	1456.14545	

------------------ VARIABLES IN THE EQUATION ------------------

VARIABLE	B	BETA	STD ERROR B	F
AGE	1.583895	0.27639	0.27985	32.033
CALCIUM	23.40607	0.26376	5.90145	15.730
(CONSTANT)	-50.59054			

------------- VARIABLES NOT IN THE EQUATION -------------

VARIABLE	BETA IN	PARTIAL	TOLERANCE	F
HEIGHT	-0.05116	-0.05652	0.97061	0.574
WEIGHT	0.03654	0.03952	0.93027	0.280
ORTHPILL	0.09088	0.10175	0.99685	1.872
ALBUMIN	-0.04948	-0.04935	0.79111	0.437
URICACID	0.15917	0.17196	0.92816	5.455

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(9)

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * VARIABLE LIST 1
REGRESSION LIST 1

DEPENDENT VARIABLE.. CHOLSTRL

VARIABLE(S) ENTERED ON STEP NUMBER 3.. URICACID

MULTIPLE R	0.47779
R SQUARE	0.22828
ADJUSTED R SQUARE	0.21535
STANDARD ERROR	37.69588

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F
REGRESSION	3.	75241.50274	25080.50091	17.65015
RESIDUAL	179.	254355.36678	1420.97970	

---------- VARIABLES IN THE EQUATION ----------

VARIABLE	B	BETA	STD ERROR B	F
AGE	1.460561	0.34708	0.28145	26.930
CALCIUM	20.56888	0.23179	5.95498	11.931
URICACID	5.853373	0.15917	2.50626	5.455
(CONSTANT)	-46.07890			

---------- VARIABLES NOT IN THE EQUATION ----------

VARIABLE	BETA IN	PARTIAL	TOLERANCE	F
WEIGHT	-0.06630	-0.07403	0.96212	0.981
HEIGHT	-0.00556	-0.00589	0.86624	0.006
BRTHPILL	0.08525	0.09683	0.99546	1.685
ALBUMIN	-0.03551	-0.03583	0.78573	0.229

MAXIMUM STEP REACHED

STATISTICS WHICH CANNOT BE COMPUTED ARE PRINTED AS ALL NINES.

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * VARIABLE LIST 1
REGRESSION LIST 1

DEPENDENT VARIABLE.. CHOLSTRL

(17) SUMMARY TABLE

VARIABLE	MULTIPLE R	R SQUARE	RSQ CHANGE	SIMPLE R	B	BETA
AGE	0.36779	0.13527	0.13527	0.36779	1.460561	0.34708
CALCIUM	0.45251	0.20477	0.06950	0.25149	20.56888	0.23179
URICACID	0.47779	0.22828	0.02352	0.26598	5.853373	0.15917
(CONSTANT)					-46.07890	

STAT 6 การวิเคราะห์ความแปรปรวนและความแปรปรวนร่วม

(ANALYSIS ØF VARIANCE AND COVARIANCE)

ตารางที่ 6 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 6

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① การวิเคราะห์ความแปรปรวนแบบทางเดียว	✓	✓	✓	
② การวิเคราะห์ความแปรปรวนแบบสองทาง	✓	✓	✓	
③ การวิเคราะห์ความแปรปรวนแบบหลายทาง	✓	✓	✓	
④ การวิเคราะห์ความแปรปรวนแบบที่ได้จากแผนการทดลองต่างดังนี้				
. Latin Square	✓	✓		
. Nested design	✓	✓		
. Factorial design	✓	✓	✓	
. Splet plot design	✓	✓		
⑤ การวิเคราะห์ความแปรปรวนร่วม	✓	✓	✓	
⑥ การวิเคราะห์การจำแนกพหุ (Multiple Classification Analysis ; MCA)			✓	
⑦ ค่าสถิติเบื้องต้นของตัวแปร	✓	✓	✓	
⑧ เมทริกซ์สหสัมพันธ์ เมทริกซ์ความแปรปรวนความแปรปรวนร่วม เมทริกซ์ความสัมพันธ์อื่นๆ	✓			

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(9) ตารางวิเคราะห์ความแปรปรวนซึ่งประกอบไปด้วยค่าสถิติ F พร้อมกับระดับนัยสำคัญสำหรับการทดสอบความเท่ากันของค่าเฉลี่ย	✓	✓	✓	(9) สำหรับ k ประชากร Ho: $\mu_1 = \mu_2 = \ldots = \mu_k$
(10) การกำหนดค่าสัมประสิทธของค่าเฉลี่ยในการทดสอบความมีนัยสำคัญ	✓		✓	(10) Ho: $C_1\mu_1 + C_2\mu_2 + \ldots C_k\mu_k = 0$
(11) ค่าสถิติ T สำหรับทดสอบความเท่ากันของค่าเฉลี่ยทุก 2 กลุ่มที่เป็นไปได้	✓		✓	(11) Ho: $\mu_1 = \mu_2 ; \mu_2 = \mu_3 ; \ldots \mu_{k-1} = \mu_k$
(12) การเปรียบเทียบความแตกต่างของค่าเฉลี่ยระหว่าง 2 กลุ่มโดยค่าสถิติต่อไปนี้				
. least significant-difference (LSD)		✓	✓	
. Duncan's multiple range test			✓	
. Student-Newman-Kenls				
. Tukey				
. Scheffc				
(13) แสดงค่า 95 % ของช่วงการประมาณค่าเฉลี่ยดังนี้				
• ช่วงการประมาณค่าเฉลี่ยแต่ละกลุ่ม				
• ช่วงการประมาณค่าเฉลี่ยรวม			✓	
• ช่วงการประมาณค่าภายใต้รูปแบบผลกระทบชนิดกลุ่ม และผลกระทบชนิดคงที่				<u>รูปแบบผลกระทบชนิดคงที่</u> (Fixed effects model or Model I) i=1,2...k
(14) ค่าสถิติซึ่งใช้ทดสอบความเท่ากันของความแปรปรวนระหว่างกลุ่ม			✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(15) การทดสอบแนวโน้มของรูปแบบสมการว่าจะเป็นสมการโพลีโนเมียล กำลังที่เท่าใดโดยค่าสถิติ F	✓		✓	$Yij = \mu+\alpha i+Eij \quad j=1,2 \ldots Ni$ $Ho: \mu_1=\mu_2=\ldots=\mu_k$ $Eij \sim N(0, \sigma^2)$
(16) การทดสอบความมีนัยสำคัญของค่าเฉลี่ยรวม [Ho; $\mu = 0$]	✓			<u>รูปแบบผลกระทบชนิดสุ่ม</u> (Random effects model or Model II)
(17) การวิเคราะห์ความแปรปรวนร่วมโดยแสดง . ค่าสัมประสิทธิการถดถอยของตัวแปรร่วม ($\hat{\beta}m$) พร้อมกับค่าความคลาดเคลื่อนมาตรฐาน ($S_{\hat{\beta}m}$)	✓		✓	$Yij = \mu+ Ai+ Eij$ $Ai \sim N(0, \sigma^2 A)$ $Ho: \sigma^2 A = 0$
. ค่าสถิติ T หรือ F ทดสอบ Ho ; $\beta m = 0$	✓		✓	$Eij \sim N(0, \sigma^2)$ <u>รูปแบบผลกระทบชนิดผสม</u> (Mixed effects model or Model III)
(18) ค่าเฉลี่ยแต่ละกลุ่มปรับแก้แล้ว ($\hat{\mu}i$) พร้อมค่าสถิติ T และระดับนัยสำคัญทดสอบความเท่ากัน	✓			$Xijk = \mu+\alpha i+Bj+Cij+Eijk$ หรือ $Yijk = \mu+ Ai+Bj+Cij+Eijk$
(19) การวิเคราะห์ความแปรปรวนร่วมซึ่งมีรายละเอียดดังนี้				(17) <u>Model</u>
. ค่าสถิติ F ทดสอบความเท่ากันของค่าเฉลี่ย	✓	✓	✓	$Yij = \mu i+\beta_1(X_{1ij}-\bar{X}_1)+\beta_2(X_{2ij}-\bar{X}_2)+\ldots$
. ค่าสถิติ F ทดสอบค่าสัมประสิทธิ $\hat{\beta}m$	✓		✓	$\hat{\mu}i = \bar{y}i+\sum m\hat{\beta}m(\bar{X}_{mn}-\bar{X}_{mi})$
. ค่าสถิติ F ทดสอบความเท่ากันของค่าสัมประสิทธิการถดถอยในตัวแปรร่วมประหว่างกลุ่ม	✓			i = แสดงจำนวนกลุ่ม m = แสดงจำนวนตัวแปรร่วม j = แสดงจำนวนตัวอย่าง $T = \hat{\beta}m/S_{\hat{\beta}m}$

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(20) ค่าสัมประสิทธิ์การถดถอยของตัวแปรร่วมในแต่ละกลุ่ม ($\hat{\beta}mi$)	✓			(19) Ho ; $\beta 1 = \beta 2 \ldots \beta m = 0$ Ho ; $\beta 11 = \beta 12 = \ldots \beta 1i$ $\beta 21 = \beta 22 = \ldots \beta 2i$ $\beta m1 = \beta m2 = \ldots \beta mi$
(21) กราฟของความสัมพันธ์ระหว่างตัวแปรต่างๆ กับความคลาดเคลื่อน	✓			
(22) การวิเคราะห์ความแปรปรวนภายใต้รูปแบบผสม (Model III)	✓			$Yij = \mu i + \beta_{1i}(X_{1ij} - \bar{X}_1) + \beta_{2i}(X_2 - \bar{X}_2) + \ldots$
(23) การเลือกนำผลร่วม (interactims) มาวิเคราะห์หรือตัดทิ้ง	✓	✓	✓	
(24) การเลือกลำดับการวิเคราะห์ตามแหล่งความผันแปร			✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 6 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//TESTBMDP  JOB (9012,90,9097),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
 INCLUDE BMDP(BMDP1V)
 ENTRY MAIN
//GO.SYSIN  DD *
/PROBLEM    TITLE IS 'ANOVA AND ANOCOVA OF BMDP BY GIRICHAI'.
            VARIABLES ARE 9.
            FORMAT IS '(A4,5F4.0,3F4.1)'.
/VARIABLE   NAME ARE ID, AGE, HEIGHT, WEIGHT, BRTHPILL, CHOLSTRL,
                     ALBUMIN, CALCIUM, URICACID.
            MAXIMUM IS (6) 400.
            MINIMUM IS (6) 150.
            BLANKS ARE MISSING.
            LABEL IS ID.
            ADD IS 1.
            GROUPING IS BRTHPILL.
/GROUP      CODES(5) ARE 1, 2.
            NAMES(5) ARE NOPILL, PILL.
/TRANSFORM  LOGCHOL = LOG (CHOLSTRL).
/DESIGN     DEPENDENT ARE LOGCHOL CALCIUM.
            INDEPENDENT ARE AGE, HEIGHT, WEIGHT.
/END
.....บัตรบันทึกข้อมูล----------------------------------------
//
```

รูปที่ 6.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT 6

(ก)

CONTROL LANGUAGE TRANSFORMATIONS ARE

(1) LOGCHOL = LOG (CHOLSTRL) .

GROUPING VARIABLE IS. . . . . BRTHPILL

NUMBER OF CASES READ. 198
CASES WITH DATA MISSING OR BEYOND LIMITS . . 8
REMAINING NUMBER OF CASES 180
NUMBER OF GROUPS FOUND. 2

VARIABLE NO.	NAME	BEFORE TRANSFORMATION MINIMUM LIMIT	MAXIMUM LIMIT	MISSING CODE	CATEGORY CODE	CATEGORY NAME	INTERVAL RANGE GREATER THAN	LESS THAN OR EQUAL TO
5	BRTHPILL				1.00000	NOPILL		
					2.00000	PILL		

NUMBER OF CASES PER GROUP

NOPILL	89.
PILL	91.
TOTAL	180.

(7) ESTIMATES OF MEANS

		NOPILL 1	PILL 2	TOTAL 3
CALCIUM	9	9.9977	9.9384	9.9678
LOGCHOL	10	2.3576	2.3734	2.3656

ONE WAY ANALYSIS OF VARIANCE FOR VARIABLE LOGCHOL

(9) ANALYSIS OF VARIANCE

SOURCE OF VARIANCE	D.F.	SUM OF SQ.	MEAN SQ.	F-VALUE	TAIL AREA PROBABILITY
EQUALITY OF CELL MEANS	1	0.0111	0.0111	1.8265	0.1783
ERROR	178	1.0857	0.0061		

(11) T-TEST MATRIX FOR GROUP MEANS ON 178 DEGREES OF FREEDOM

		NOPILL 1	PILL 2
NOPILL	1	0.0	
PILL	2	1.3515	0.0

(11) PROBABILITIES FOR THE T-VALUES ABOVE

		NOPILL 1	PILL 2
NOPILL	1	1.0000	
PILL	2	0.1783	1.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ESTIMATES OF MEANS

		NOPILL 1	PILL 2	TOTAL 3
AGE	2	32.9772	34.0876	33.5385
HEIGHT	3	64.0559	64.8678	64.4663
WEIGHT	4	129.4379	132.7140	131.0941
CALCIUM	8	9.9977	9.9384	9.9678
LOGCHOL	10	2.3576	2.3736	2.3656

(5) DEPENDENT VARIABLE IS LOGCHOL

(17)

COVARIATE	REG.COEFF.	STD.ERR.	T-VALUE
AGE	0.00269	0.00057	4.68296
HEIGHT	-0.00251	0.00253	-0.98909
WEIGHT	0.00033	0.00031	1.04947

(18)

GROUP	N	GRP.MEAN	ADJ.GRP.MEAN	STD.ERR.
NOPILL	89.	2.35763	2.35865	0.00783
PILL	91.	2.37336	2.37236	0.00774

(19) ANALYSIS OF VARIANCE

SOURCE OF VARIANCE	D.F.	SUM OF SQ.	MEAN SQ.	F-VALUE	TAIL AREA PROBABILITY
EQUALITY OF ADJ. CELL MEANS	1	0.0082	0.0082	1.5294	0.2179
ZERO SLOPE	3	0.1453	0.0484	9.0126	0.0000
ERROR	175	0.9404	0.0054		
EQUALITY OF SLOPES	3	0.0163	0.0054	1.0113	0.3891
ERROR	172	0.9241	0.0054		

SLOPE WITHIN EACH GROUP

(20)

		NOPILL 1	PILL 2
AGE	2	0.0035	0.0019
HEIGHT	3	-0.0003	-0.0053
WEIGHT	4	-0.0001	0.0007

T-TEST MATRIX FOR ADJUSTED GROUP MEANS ON 175 DEGREES OF FREEDOM

(18)

		NOPILL 1	PILL 2
NOPILL	1	0.0	
PILL	2	1.2366	0.0

PROBABILITIES FOR THE T-VALUES ABOVE

(18)

		NOPILL 1	PILL 2
NOPILL	1	1.0000	
PILL	2	0.2179	1.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.1 (ต่อ) ใช้ข้อมูลจาก DATA SAMPLE 3

```
//TESTBMDP  JOB (9012,90,90,M),'CHULA'
//   EXEC BIMED
//LKED.SYSIN DD *
(ข) INCLUDE BMDP(BMDP2V)
    ENTRY MAIN
//GO.SYSIN  DD *
/PROBLEM    TITLE IS 'TWO-WAY ANOVA WITH INTERACTION'.
/INPUT      VARIABLES ARE 3.
            FORMAT IS '(3F3.0)'.
/VARIABLE   NAMES ARE TREATMNT,DISEASE,SYSINCR.
/DESIGN     DEPENDENT IS SYSINCR.
            GROUPING ARE TREATMNT,DISEASE.
/GROUP      CODES(1) ARE 1, 2, 3, 4.
            NAMES(1) ARE DRUG1,DRUG2,DRUG3,DRUG4.
            CODES(2) ARE 1, 2, 3.
            NAMES(2) ARE DISEASE1,DISEASE2,DISEASE3.
/END
```

บัตรบันทึกข้อมูล

```
//
```

DESIGN SPECIFICATIONS

(2)

GROUP = 1 2
DEPEND = 3

VARIABLE NO.	NAME	BEFORE TRANSFORMATION MINIMUM LIMIT	MAXIMUM LIMIT	MISSING CODE	CATEGORY CODE	CATEGORY NAME	INTERVAL RANGE GREATER THAN	LESS THAN OR EQUAL TO
1	TREATMNT				1.00000	DRUG1		
					2.00000	DRUG2		
					3.00000	DRUG3		
					4.00000	DRUG4		
2	DISEASE				1.00000	DISEASE1		
					2.00000	DISEASE2		
					3.00000	DISEASE3		

NUMBER OF CASES READ. 58

GROUP STRUCTURE

TREATMNT	DISEASE	COUNT
DRUG1	DISEASE1	6.
DRUG1	DISEASE2	4.
DRUG1	DISEASE3	5.
DRUG2	DISEASE1	5.
DRUG2	DISEASE2	4.
DRUG2	DISEASE3	6.
DRUG3	DISEASE1	3.
DRUG3	DISEASE2	5.
DRUG3	DISEASE3	4.
DRUG4	DISEASE1	5.
DRUG4	DISEASE2	6.
DRUG4	DISEASE3	5.

CELL MEANS FOR 1-ST DEPENDENT VARIABLE

(7)

TREATMNT= / DISEASE =	DRUG1 DISEASE1	DRUG1 DISEASE2	DRUG1 DISEASE3	DRUG2 DISEASE1	DRUG2 DISEASE2	DRUG2 DISEASE3	DRUG3 DISEASE1	DRUG3 DISEASE2	DRUG3 DISEASE3	DRUG4 DISEASE1
SYSINCR	29.33333	28.25000	20.39999	28.00000	33.50000	18.16666	16.33333	4.40000	8.50000	13.60000
COUNT	6	4	5	5	4	6	3	5	4	5

TREATMNT= / DISEASE =	DRUG4 DISEASE2	DRUG4 DISEASE3	MARGINAL
SYSINCR	12.83333	14.20000	18.87930
COUNT	6	5	58

STANDARD DEVIATIONS FOR 1-ST DEPENDENT VARIABLE

TREATMNT= / DISEASE =	DRUG1 DISEASE1	DRUG1 DISEASE2	DRUG1 DISEASE3	DRUG2 DISEASE1	DRUG2 DISEASE2	DRUG2 DISEASE3	DRUG3 DISEASE1	DRUG3 DISEASE2	DRUG3 DISEASE3	DRUG4 DISEASE1
SYSINCR	13.01793	9.85235	13.37161	10.97725	2.08167	12.52863	14.18919	6.91375	9.00000	10.54988

TREATMNT= / DISEASE =	DRUG4 DISEASE2	DRUG4 DISEASE3
SYSINCR	10.34247	8.92748

ANALYSIS OF VARIANCE FOR 1-ST DEPENDENT VARIABLE - SYSINCR

(9)

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROBABILITY
MEAN	20037.49219	1	20037.49219	181.41	0.000
T	2997.42969	3	999.14307	9.05	0.000
D	415.86719	2	207.93359	1.88	0.164
TD	707.24609	6	117.87434	1.07	0.396
(23) ERROR	5080.77734	46	110.45168		

หอสมุดกลาง จุฬาลงกรณ์มหาวิทยาลัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 6 โดยใช้ข้อมูลจาก DATA SAMPLE 1

```
//SØ38OPØK JØB CLASS=C,MSGLEVEL=(2,0)
//  EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE1;
INPUT ID $ 1-4 AGE 5-8 HEIGHT 9-12 WEIGHT 13-16 BRTHPILL 17-20
  CHØLSTRL 21-24 ALBUMIN 25-28 1 CALCIUM 29-31 1 URICACID 33-36 1;
  LØGCHØL = LØG10(CHØLSTRL);
CARDS;
```

บัตรบันทึกข้อมูล

```
PRØC ANØVA;
 CLASS BRTHPILL;
 MEAN  BRTHPILL;
 MØDEL LØGCHØL CALCIUM LØGCHØL*AGE LØGCHØL*HEIGHT LØGCHØL*WEIGHT
       =BRTHPILL
 PØØL 'EXP-ERRØR' RESIDUAL/BRTHPILL;
 TEST BRTHPILL BY 'EXP-ERRØR';
 TITLE 'ØNE-WAY ANØVA AND ANØCØVA ØF SAS BY SIRICHAI';
/*
//
```

รูปที่ 6.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT 6

(1) ONE-WAY ANOVA AND ANOVA OF SAS BY SIRICHAI

DATA SET DATA0001

CLASSES	VALUES
BRTHPILL	1 2

ONE-WAY ANOVA AND ANOVA OF SAS BY SIRICHAI

(7) MEANS

BRTHPILL	LOGCHOL	CALCIUM	AGE	HEIGHT	WEIGHT
1	2.35979304 94	9.9[illegible]13043 92	33.[illegible]9 94	64.[illegible] 92	130.[illegible]8 93
2	2.37287155 92	9.935483[illegible]7 93	33.[illegible] 94	64.[illegible] 94	133.[illegible] 93
OVERALL MEANS	2.36[illegible]19[illegible] 186	9.962[illegible] 185	33.[illegible] 188	64.[illegible]27 186	131.[illegible] 186

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(n)

ONE-WAY ANOVA AND ANOVA OF SAS BY SIRICHAI

ANALYSIS OF VARIANCE FOR VARIABLE LOGOHOL MEAN 2.3c62619c C.V. 3.2c394093 %

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	LSD .01	LSD .05	DIVISOR
BRTHPILL	1	0.0079527c	0.0079527[illegible]	(12)		
EXP-ERROR	184	1.11105139	[illegible]	[illegible]	0.0294021565	75
RESIDUAL	184	1.11105139	[illegible]			
CORRECTED TOTAL	185	1.11900417	[illegible]			

TESTS (9) SOURCE

NUM RATOR: BRTHPILL

DENOMINATOR: EXP-ERROR

DF	SUM OF SQUARES	MEAN SQUARE	F VALUE	PROB > F
1	0.0079[illegible]	[illegible]	1.[illegible]	0.[illegible]
184	1.11105139	0.[illegible]		

(19) ONE-WAY ANOVA AND ANOVA OF SAS BY SIRICHAI

(5)

CROSS PRODUCTS ANALYSIS FOR VARIABLE LOGOHOL BY AGE

SOURCE	DF	SUM OF PRODUCTS	MEAN PRODUCT	CORRELATION
BRTHPILL	1	0.13e4102[illegible]	0.13e4102[illegible]	1.0000[illegible]
EXP-ERROR	184	52.[illegible]5022	0.2863[illegible]	0.[illegible]
RESIDUAL	184	52.[illegible]5022	0.2863[illegible]	0.[illegible]
CORRECTED TOTAL	185	52.[illegible]7124	0.2[illegible]3040	0.[illegible]

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.2 (ต่อ) ใช้ข้อมูลจาก DATA SAMPLE 3

```
//SØ3ØPØK JØB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE3;
INPUT TREATMNT 3 DISEASE 6 SYSINCR 8-9;
CARDS;
```

บัตรบันทึกข้อมูล

(ข)

```
PRØC ANØVA;
 CLASS TREATMNT DISEASE;
 MEANS TREATMNT DISEASE;
 MØDEL SYSINCR = TREATMNT DISEASE TREATMNT*DISEASE;
 PØØL ERRØR-T RESIDUAL/TREATMNT;
 PØØL ERRØR-D RESIDUAL/DISEASE;
 TEST TREATMNT DISEASE TREATMNT*DISEASE BY RESIDUAL;
 TITLE 'TWØ-WAY ANØVA ØF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(2) TWO-WAY ANOVA OF SAS BY SYSTOLIC

DATA SET SAMPLE3

CLASSES	VALUES
TREATMNT	1 2 3 4
DISEASE	1 2 3

TWO-WAY ANOVA OF SAS BY SYSTOLIC

(7) MEANS

DISEASE	N	SYSINCR
1	19	[illegible]
2	19	[illegible]
3	20	[illegible]
TREATMNT		
1	15	[illegible]
2	15	[illegible]
3	12	[illegible]
4	10	[illegible]
OVERALL MEANS	56	[illegible]

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

TWO-WAY ANOVA OF SAS BY SYSINCR (9)

ANALYSIS OF VARIANCE FOR VARIABLE SYSINCR MEAN 10.8795103 C.V. 55.0075004 %

SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	LSD .01 (12)	LSD .05	DIVISOR
TREATMNT	3	3133.23051	1044.41284			
DISEASE	2	488.03930	244.31969			
TREATMNT*DISEASE (23)	6	637.40082	106.2344			
ERROR T	46	5080.81667	110.45254	16.136379	7.7068700	15
ERROR D	46	5080.81667	110.45254	9.1012940	6.0051015	19
RESIDUAL	46	5080.81667	110.45254			
CORRECTED TOTAL	57	9340.15517	163.86237			

WARNING: THERE ARE UNEQUAL CELL FREQUENCIES FOR SOME CELLS.
A WEIGHTED ANALYSIS HAS BEEN COMPUTED.

CHECK YOUR DEGREES OF FREEDOM AND CELL FREQUENCIES TO VERIFY THE CORRECTNESS OF THE ANALYSIS.

TEST: SOURCE	DF	SUM OF SQUARES	MEAN SQUARE	F VALUE (9)	PROB > F
NUMERATOR: TREATMNT	3	3133.23051	1044.41284	9.4517	0.0002
DENOMINATOR: RESIDUAL	46	5080.81667	110.45254		
NUMERATOR: DISEASE	2	488.03930	244.31969	2.2119	0.1191
DENOMINATOR: RESIDUAL	46	5080.81667	110.45254		
NUMERATOR: TREATMNT*DISEASE	6	637.40082	106.2344	0.96184	0.5376
DENOMINATOR: RESIDUAL	46	5080.81667	110.45254		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 6 โดยใช้ข้อมูลจาก DATA SAMPLE 6

```
*$$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK        MR.SIRICHAI PONGWICHAI B2214756
// EXEC PROC=$$SPSS79
RUN NAME        'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME       'DATA SAMPLE1'
VARIABLE LIST   ID AGE HEIGHT WEIGHT BRTHPILL CHOLSTRL ALBUMIN
                CALCIUM URICACID
INPUT MEDIUM    CARD
N OF CASES      198
INPUT FORMAT    FIXED(A4,5F4.0,3F4.1)
MISSING VALUES  AGE TO URICACID(0)
VALUE LABELS    BRTHPILL(1)NOPILL(2)PILL
RECODE          CHOLSTRL(LO THRU 149=0)(401 THRU HI=0)
COMPUTE         LOGCHOL = LG10(CHOLSTRL)
READ INPUT DATA
```

บัตรบันทึกข้อมูล

(ก)

```
TASK NAME       'ONE-WAY ANOVA OF SPSS BY SIRICHAI'
ONEWAY            LOGCHOL CALCIUM BY BRTHPILL(1,2)/
                  RANGES = LSD(0.01)/
OPTIONS         2,2
STATISTICS      ALL
```

ตารางที่ 6.3 (ต่อ)

(ข)

```
TASK NAME                'ANOVA OF SPSS BY SIRICHAI'
ANOVA                    LOGCHOL CALCIUM BY BRTHPILL(1,2) WITH
                         AGE HEIGHT WEIGHT/
OPTIONS                  9
STATISTICS               ALL
FINISH
/*
/&
* $$ EOJ
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 6.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT 6

(ก)

'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS' 24/06/81 PAGE 33
'ONE-WAY ANOVA OF SPSS BY STEICHAI'
FILE 'DATA (CREATION DATE = 24/06/81) SAMPLE1'

- O N E W A Y -

(1) VARIABLE LOGCHOL

(9) ANALYSIS OF VARIANCE

SOURCE	D.F.	SUM OF SQUARES	MEAN SQUARES	F RATIO	F PROB.
BETWEEN GROUPS	1	0.0656	0.0656	1.005	0.3174
WITHIN GROUPS	186	12.1326	0.0652		
TOTAL	187	12.1982			

(7) (13)

GROUP	COUNT	MEAN	STANDARD DEVIATION	STANDARD ERROR	MINIMUM	MAXIMUM	95 PCT CONF INT FOR MEAN
GRP01	94	2.3598	0.2814	0.0384	2.1903	2.8250	2.3441 TO 2.3765
GRP02	94	2.3224	0.3619	0.0363	0.0	2.5911	2.2503 TO 2.3945
TOTAL	188	2.3411	0.2554	0.0186	0.0	2.5911	2.3043 TO 2.3773
	FIXED EFFECTS MODEL		0.2554	0.0186			2.3043 TO 2.3778
	RANDOM EFFECTS MODEL		0.0264	0.0187			2.1028 TO 2.5784

RANDOM EFFECTS MODEL - ESTIMATE OF BETWEEN COMPONENT VARIANCE 0.0330

TESTS FOR HOMOGENEITY OF VARIANCES

(14) COCHRANS C = MAX. VARIANCE/SUM(VARIANCES) = 0.9492, P = 0.000 (APPROX.)
BARTLETT-BOX F = 152.508, P = 0.000
MAXIMUM VARIANCE / MINIMUM VARIANCE = 18.697

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(ย)

* * * * * * * * * * A N A L Y S I S O F V A R I A N C E * * * * * * * * * *

(5) LOGCHOL
BY BRTHPLL
WITH AGE
HEIGHT
WEIGHT

* *

(17) (19)

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES	0.353	3	0.117	3.269	0.023
AGE	0.277	1	0.277	7.779	0.006
HEIGHT	0.058	1	0.058	1.615	0.205
WEIGHT	0.000	1	0.000	0.004	0.952
MAIN EFFECTS	0.002	1	0.002	0.046	0.830
BRTHPLL	0.002	1	0.002	0.046	0.830
EXPLAINED	0.351	4	0.088	2.464	0.047
RESIDUAL	6.310	177	0.036		
TOTAL	6.662	181	0.037		

(19)

COVARIATE	RAW REGRESSION COEFFICIENT
AGE	0.004
HEIGHT	-0.008
WEIGHT	-0.000

188 CASES WERE PROCESSED.
6 CASES (3.2 PCT) WERE MISSING.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 6.3 (ต่อ) ใช้ข้อมูลจาก DATA SAMPLE 3

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK        MR.SIRICHAI PONGWICHAI B2211475
// EXEC PROC=$$SPSS79
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME         'DATA SAMPLE3'
VARIABLE LIST     TREATMNT DISEASE SYSINCR
INPUT MEDIUM      CARD
N OF CASES        58
INPUT FORMAT      FIXED(3F3.0)
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
TASK NAME         'TWO-WAY OF SPSS BY SIRICHAI'
ANOVA             SYSINCR BY TREATMNT(1,4) DISEASE(1,3)
OPTIONS           9
STATISTICS        ALL
FINISH
/*
/&
* $$ EOJ
```

* * * * * * * * * * A N A L Y S I S   O F   V A R I A N C E * * * * * * * * *

(2) SYSINCE
BY TREATMNT
DISEASE

* *

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	(9) F	SIGNIF OF F
MAIN EFFECTS	3490.953	5	699.991	6.326	0.000
TREATMNT	2997.473	3	999.157	9.046	0.000
DISEASE	415.873	2	207.937	1.883	0.164
2-WAY INTERACTIONS	707.266	6	117.878	1.067	0.396
TREATMNT DISEASE	707.266	6	117.878	1.067	0.394
EXPLAINED	4259.336	11	387.212	3.505	0.001
RESIDUAL	5081.776	46	110.451		
TOTAL	934[illegible]	57	163.861		

58 CASES WERE PROCESSED.
0 CASES (0.0 PCT) WERE MISSING.

* * * M U L T I P L E   C L A S S I F I C A T I O N   A N A L Y S I S * * *

(6) SYSINCE
BY TREATMNT
DISEASE

* *

GRAND MEAN = 33.18

VARIABLE + CATEGORY	N	UNADJUSTED DEV'N	UNADJUSTED ETA	ADJUSTED FOR INDEPENDENTS DEV'N	ADJUSTED FOR INDEPENDENTS BETA	ADJUSTED FOR INDEPENDENTS + COVARIATES DEV'N	ADJUSTED FOR INDEPENDENTS + COVARIATES BETA
TREATMNT							
1	15	7.19		6.99			
2	15	6.65		6.30			
3	12	-12.13		-11.11			
4	16	-5.33		-5.49			
			0.59		0.58		
DISEASE							
1	19	[illegible]		2.93			
2	19	[illegible]		0.70			
3	20	-2.93		-3.51			
			0.23		0.21		

MULTIPLE R SQUARED	0.382
MULTIPLE R	0.617

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT 7 ... การวิเคราะห์ความสัมพันธ์ระหว่างสองกลุ่มตัวแปร

(CANØNICAL CØRRELATIØN ANALYSIS)

ตารางที่ 7 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 7

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓		✓	
② เมทริกซ์สหสัมพันธ์ของแต่ละตัวแปร $[\hat{R}(x)]$ หรือ เมทริกซ์ความแปรปรวนร่วม $[S]$ หรือ เมทริกซ์ของ $[SSCP]$	✓	✓	✓	② $[\hat{R}(x)]$ = Sample correlation matrix $[Dia(s)]^{-1/2} S [Dia(S)]^{-1/2}$ $[S]$= Sample variance--covariance matrix $[SSCP]$ / (n-1) $[SSCP]$ = Corrected sum of square and cross products matrix
③ ค่าสหสัมพันธ์กำลังสอง (R^2) ของแต่ละตัวแปรกับตัวแปรอื่นๆในกลุ่มเดียวกันแยกแต่ละกลุ่ม	✓			
④ ค่าไอเกน (Ci,λi)	✓		✓	④ ci = eigen value or characteristic root
⑤ ค่าเฉลี่ยของตัวแปรคาโนนิคอลในแต่ละกลุ่ม		✓		ui,vi = ตัวแปรคาโนนิคอลที่ i ของกลุ่มที่ 1 และ 2 ที่สัมพันธ์กับ ci
⑥ ค่าสหสัมพันธ์คาโนนิคอล ($\sqrt{Ci}$)	✓	✓	✓	
⑦ ค่าสถิติ Wilk's lambda ใช้ทดสอบความสัมพันธ์เชิงเส้นของความคลาดเคลื่อนระหว่างตัวแปรคาโนนิคอล			✓	$\underset{\sim}{a}i, \underset{\sim}{b}i$ = เวคเตอร์สัมประสิทธิ์ของตัวแปรในกลุ่มที่ 1 และ 2 ที่สัมพันธ์กับ Ci

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
⑧ ค่าสถิติ Chi-square พร้อมกับ--ระดับนัยสำคัญ เพื่อใช้ทดสอบ--ความสัมพันธ์ของตัวแปรคาโนนิคอล	✓	✓	✓	Uij, Vij = ค่าตัวแปรคาโนนิคอลที่ i ของค่าสังเกตที่ j สำหรับตัวแปรกลุ่มที่ 1 และ 2
⑨ ค่า Normalized ของเวคเตอร์--ไอเกนที่ได้จาก SSCP ของแต่ละ--กลุ่ม (λ [SSCP] λ = 1)		✓		$\underset{\sim}{a}'i = [ai1\ ai2..aik..aip]$
⑩ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรคาโนนิคอล (Ui,Vi) กับตัวแปรเดิม (Xjk, Xjm)แยกแต่ละ--กลุ่ม	✓	✓		$\underset{\sim}{b}'i = [bi1\ bi2...bim..biq]$
⑪ ค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอลของทุกตัวแปรในแต่ละกลุ่ม($\underset{\sim}{a}i$,$\underset{\sim}{b}i$)	✓		✓	p,q = จำนวนตัวแปรกลุ่มที่ 1 และ 2: $p \leq q$
⑫ ค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอลเมื่อตัวแปรเดิมอยู่ในรูปคะแนน--มาตรฐานปกติ	✓			i = 1,2,.... p
				j = 1,2,.... n
				$Uij = \sum_{k=1,2,..p} aikXjk$
⑬ ค่าตัวแปรคาโนนิคอลแต่ละตัวของ--ทั้งสองกลุ่มในทุกๆ ค่าสังเกต	✓	✓	✓	$Vij = \sum_{m=1,2,..q} bimXjm$
⑭ ค่าความแปรปรวนของแต่ละตัวแปรคาโนนิคอล	✓			Xjk = ตัวแปรที่ k ในกลุ่มที่ 1 จากค่าสังเกตที่ j
⑮ กราฟของความสัมพันธ์ระหว่างตัว--แปรคาโนนิคอลทั้ง 2 กลุ่ม	✓	✓		Xjm = ตัวแปรที่ m ในกลุ่มที่ 2 จากค่าสังเกตที่ j

ตารางที่ 7.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 7 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
//TESTBMDP  JOB (9012O,90,9089),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
  INCLUDE BMDP(BMDP6M)
  ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM   TITLE IS 'CANONICLA OF BMDP BY SIRICHAI';
/INPUT     VARIABLES ARE 12.
           FORMAT IS '(12F2.0)'.
/VARIABLE  NAMES ARE CONCENTR,ANNOY,SMOKING1,SLEEPY,SMOKING2,TENSE,
                     SMOKING3,ALERT,IRRITABL,TIRED,CONTENT,SMOKING4.
/CANONICAL FIRST ARE SMOKING1,SMOKING2,SMOKING3,SMOKING4.
           SECOND ARE CONCENTR,ANNOY,SLEEPY,TENSE,ALERT,IRRITABL,
                      TIRED,CONTENT.
/PRINT     MATRICES ARE CORR,CANV,COEF,LOAD.
/PLOT      XVARS    ARE CNVRS1,CONCENTR.
           YVARS    ARE CNVRF1,CNVRS1.
/END
```

บัตรบันทึกข้อมูล

```
//
```

รูปที่ 7.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT 7

```
FIRST SET OF VARIABLES
        3  SMOKING1      5  SMOKING2      7  SMOKING3     12  SMOKING4

SECOND SET OF VARIABLES
        1  CONCENTR      2  ANNOY         4  SLEEPY        6  TENSE        8  ALERT
        9  IRRITABL     10  TIRED        11  CONTENT

NUMBER OF VARIABLES IN FIRST SET. . . . . .      4
NUMBER OF VARIABLES IN SECOND SET . . . . .      8
TOTAL NUMBER OF VARIABLES USED. . . . . .       12
MAXIMUM NUMBER OF CANONICAL VARIABLES . . . .    4
MINIMUM CANONICAL CORRELATION TO BE USED. . .  0.0
WEIGHT VARIABLE . . . . . .
PRECISION . . . . . . DOUBLE
TOLERANCE FOR MATRIX INVERSION. . . . . 0.0100000
EIGENVALUE LIMIT. . . . . .  0.0
APPROX. NUMBER OF VARIABLES WHICH CAN BE ANALYZED  120


DATA AFTER TRANSFORMATIONS FOR FIRST    5 CASES
CASES WITH ZERO WEIGHTS AND MISSING DATA NOT INCLUDED.

     CASE
LABEL    NUMBER      WEIGHT    3 SMOKING1    5 SMOKING2    7 SMOKING3   12 SMOKING4   1 CONCENTR    2 ANNOY       4 SLEEPY
                  6 TENSE     8 ALERT       9 IRRITABL    10 TIRED     11 CONTENT
           1       1.00000      1.00000       2.00000       1.00000      2.00000       3.00000       2.00000       3.00000
                   2.00000      3.00000       2.00000       2.00000      3.00000
           2       1.00000      5.00000       5.00000       5.00000      5.00000       4.00000       2.00000       3.00000
                   4.00000      4.00000       3.00000       4.00000      4.00000
           3       1.00000      4.00000       5.00000       4.00000      4.00000       5.00000       3.00000       4.00000
                   5.00000      5.00000       3.00000       4.00000      3.00000
           4       1.00000      4.00000       5.00000       4.00000      5.00000       4.00000       2.00000       3.00000
                   4.00000      4.00000       3.00000       4.00000      3.00000
           5       1.00000      4.00000       4.00000       4.00000      4.00000       4.00000       2.00000       3.00000
                   2.00000      4.00000       2.00000       4.00000      3.00000

 NUMBER OF CASES READ. . . . . .  110
```

(1) UNIVARIATE SUMMARY STATISTICS

VARIABLE	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION	SMALLEST VALUE	LARGEST VALUE	SMALLEST STANDARD SCORE	LARGEST STANDARD SCORE	SKEWNESS	KURTOSIS
3 SMOKING1	3.36364	1.13111	0.336275	1.00000	5.00000	-2.09	1.45	-0.36	-0.57
5 SMOKING2	3.58182	1.06126	0.296291	1.00000	5.00000	-2.43	1.34	-0.40	-0.63
7 SMOKING3	3.42727	1.16098	0.338747	1.00000	5.00000	-2.09	1.35	-0.42	-0.63
12 SMOKING4	3.50000	1.27610	0.364601	1.00000	5.00000	-1.96	1.18	-0.54	-0.79
1 CONCENTR	2.69091	1.07298	0.398744	1.00000	5.00000	-1.58	2.15	0.10	-0.81
2 ANNOY	2.11818	0.97427	0.459957	1.00000	5.00000	-1.15	2.96	0.89	0.33
4 SLEEPY	2.60909	1.02353	0.392295	1.00000	5.00000	-1.57	2.34	0.01	-0.80
6 TENSE	2.44545	0.99158	0.405480	1.00000	5.00000	-1.46	2.58	0.26	-0.79
8 ALERT	2.80909	1.01814	0.362445	1.00000	5.00000	-1.78	2.15	0.02	-0.48
9 IRRITABL	2.21818	0.78263	0.352825	1.00000	4.00000	-1.56	2.28	0.06	-0.63
10 TIRED	3.09091	0.95346	0.308473	1.00000	5.00000	-2.19	2.00	-0.31	-0.25
11 CONTENT	2.45455	0.84198	0.343027	1.00000	5.00000	-1.73	3.02	0.19	-0.16

VALUES FOR KURTOSIS GREATER THAN ZERO INDICATE DISTRIBUTIONS
WITH HEAVIER TAILS THAN THE NORMAL DISTRIBUTION.

หอสมุดกลาง จุฬาลงกรณ์มหาวิทยาลัย

CORRELATIONS

(2)

		SMOKING1 3	SMOKING2 5	SMOKING3 7	SMOKING4 12	CONCENTR 1	ANNOY 2	SLEEPY 4	TENSE 6	ALERT 8	IRRITABL 9	TIRED 10	CONTENT 11
SMOKING1	3	1.000											
SMOKING2	5	0.785	1.000										
SMOKING3	7	0.810	0.816	1.000									
SMOKING4	12	0.775	0.813	0.845	1.000								
CONCENTR	1	0.086	0.200	0.041	0.228	1.000							
ANNOY	2	0.144	0.119	0.060	0.122	0.562	1.000						
SLEEPY	4	0.140	0.211	0.126	0.277	0.457	0.360	1.000					
TENSE	6	0.222	0.301	0.120	0.214	0.579	0.705	0.273	1.000				
ALERT	8	0.101	0.223	0.039	0.201	0.802	0.578	0.606	0.594	1.000			
IRRITABL	9	0.189	0.221	0.108	0.156	0.595	0.796	0.337	0.725	0.605	1.000		
TIRED	10	0.199	0.276	0.139	0.271	0.512	0.413	0.798	0.364	0.698	0.429	1.000	
CONTENT	11	0.239	0.235	0.100	0.171	0.492	0.739	0.240	0.711	0.605	0.897	0.394	1.000

SQUARED MULTIPLE CORRELATIONS OF EACH VARIABLE IN SECOND SET WITH ALL OTHER VARIABLES IN SECOND SET

(3)

VARIABLE NUMBER	NAME	R-SQUARED
1	CONCENTR	0.68425
2	ANNOY	0.72147
4	SLEEPY	0.67121
6	TENSE	0.63869
8	ALERT	0.79529
9	IRRITABL	0.70704
10	TIRED	0.71877
11	CONTENT	0.67799

SQUARED MULTIPLE CORRELATIONS OF EACH VARIABLE IN FIRST SET WITH ALL OTHER VARIABLES IN FIRST SET

VARIABLE NUMBER	NAME	R-SQUARED
3	SMOKING1	0.71036
5	SMOKING2	0.74227
7	SMOKING3	0.79150
12	SMOKING4	0.76658

(4) EIGENVALUE | CANONICAL CORRELATION (5)

EIGENVALUE	CANONICAL CORRELATION
0.27278	0.52229
0.14128	0.37588
0.05779	0.24040
0.01882	0.13719

(8) NUMBER OF EIGENVALUES | BARTLETT'S TEST FOR REMAINING EIGENVALUES

NUMBER OF EIGENVALUES	CHI-SQUARE	D.F.	SIGNIFICANCE
0	56.31	32	0.00502
1	23.66	21	0.30975
2	8.05	12	0.78129
3	1.95	5	0.85637

BARTLETT'S TEST ABOVE INDICATES THE NUMBER OF CANONICAL VARIABLES NECESSARY TO EXPRESS THE DEPENDENCY BETWEEN THE TWO SETS OF VARIABLES. THE NECESSARY NUMBER OF CANONICAL VARIABLES IS THE SMALLEST NUMBER OF EIGENVALUES SUCH THAT THE TEST OF THE REMAINING EIGENVALUES IS NON-SIGNIFICANT. FOR EXAMPLE, IF A TEST AT THE .01 LEVEL WERE DESIRED, THEN 1 VARIABLES WOULD BE CONSIDERED NECESSARY. HOWEVER, THE NUMBER OF CANONICAL VARIABLES OF PRACTICAL VALUE IS LIKELY TO BE SMALLER.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CANONICAL VARIABLE LOADINGS (CORRELATIONS OF CANONICAL VARIABLES WITH ORIGINAL VARIABLES)

(10)

		CNVRF1 1	CNVRF2 2	CNVRF3 3	CNVRF4 4
SMOKING1	3	-0.445	-0.534	-0.659	0.296
SMOKING2	5	-0.728	-0.394	-0.151	0.548
SMOKING3	7	-0.289	-0.270	-0.673	0.787
SMOKING4	12	-0.639	0.056	-0.565	0.518

		CNVRS1 1	CNVRS2 2	CNVRS3 3	CNVRS4 4
CONCENTR	1	-0.721	0.355	-0.013	-0.311
ANNOY	2	-0.303	-0.141	-0.384	-0.401
SLEEPY	4	-0.600	0.347	-0.377	0.260
TENSE	6	-0.700	-0.333	0.001	-0.178
ALERT	8	-0.730	0.156	0.149	-0.761
IRRITABL	9	-0.457	-0.337	-0.115	-0.072
TIRED	10	-0.691	0.027	-0.256	0.072
CONTENT	11	-0.532	-0.443	-0.308	-0.598

COEFFICIENTS FOR CANONICAL VARIABLES FOR FIRST SET OF VARIABLES

(11)

		CNVRF1 1	CNVRF2 2	CNVRF3 3	CNVRF4 4
SMOKING1	3	0.378543E-01	-0.976451E 00	-0.965493E 00	-0.900841E 00
SMOKING2	5	-0.109321E 01	-0.646535E 00	0.134105E 01	0.132999E 00
SMOKING3	7	0.119115E 01	-0.173899E 00	-0.333693E-01	0.143194E 01
SMOKING4	12	-0.704060E 00	0.128569E 01	-0.660196E 00	-0.214955E 00

(12)

STANDARDIZED COEFFICIENTS FOR CANONICAL VARIABLES FOR FIRST SET OF VARIABLES
(THESE ARE THE COEFFICIENTS FOR THE STANDARDIZED VARIABLES - MEAN ZERO, STANDARD DEVIATION ONE.)

		CNVRF1 1	CNVRF2 2	CNVRF3 3	CNVRF4 4
SMOKING1	3	0.043	-1.104	-1.092	-1.019
SMOKING2	5	-1.160	-0.686	1.423	0.194
SMOKING3	7	1.383	-0.202	-0.039	1.686
SMOKING4	12	-0.898	1.641	-0.842	-0.274

COEFFICIENTS FOR CANONICAL VARIABLES FOR SECOND SET OF VARIABLES

		CNVRS1 1	CNVRS2 2	CNVRS3 3	CNVRS4 4
CONCENTR	1	-0.441692E 00	0.749510E 00	-0.470381E 00	-0.163811E 00
ANNOY	2	0.801410E 00	0.461495E 00	-0.603503E 00	-0.739549E 00
SLEEPY	4	-0.250790E 00	0.581216E 00	-0.685988E 00	0.615867E 00
TENSE	6	-0.692552E 00	-0.389734E 00	0.421877E 00	0.448775E 00
ALERT	8	0.140028E 00	0.204741E 00	0.150139E 01	-0.685341E 00
IRRITABL	9	0.900002E-01	-0.795294E 00	0.425982E 00	0.113746E 01
TIRED	10	-0.327905E 00	-0.616256E 00	-0.246355E 00	0.172116E 00
CONTENT	11	-0.402041E 00	-0.599032E 00	-0.971468E 00	-0.795208E 00

STANDARDIZED COEFFICIENTS FOR CANONICAL VARIABLES FOR SECOND SET OF VARIABLES
(THESE ARE THE COEFFICIENTS FOR THE STANDARDIZED VARIABLES - MEAN ZERO, STANDARD DEVIATION ONE.)

		CNVRS1 1	CNVRS2 2	CNVRS3 3	CNVRS4 4
CONCENTR	1	-0.474	0.800	-0.505	-0.176
ANNOY	2	0.781	0.450	-0.590	-0.721
SLEEPY	4	-0.257	0.595	-0.702	0.630
TENSE	6	-0.687	-0.378	0.418	0.445
ALERT	8	0.143	0.208	1.529	-0.698
IRRITABL	9	0.070	-0.622	0.333	0.890
TIRED	10	-0.313	-0.589	-0.235	0.164
CONTENT	11	-0.339	-0.501	-0.819	-0.670

CANONICAL VARIABLES (CASE NUMBERS REFER TO DATA BEFORE DELETION OF CASES)

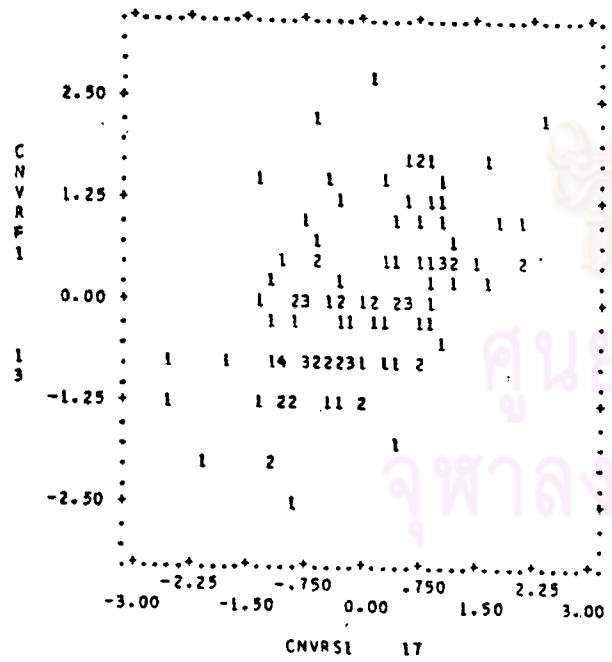
(13)

CASE LABEL	NO.	WEIGHT	CNVRF1	CNVRF2	CNVRF3	CNVRF4	CNVRS1	CNVRS2	CNVRS3	CNVRS4
	1	1.0000	-0.1954	1.8242	1.2321	-1.3620	0.1247	1.1330	-0.5973	-0.9229
	2	1.0000	-0.6712	-0.8597	-0.7208	0.7465	-2.5299	-1.3010	0.2394	-0.1330
	3	1.0000	-1.1961	-0.9950	0.9382	0.4104	-2.5714	0.9062	1.3725	0.3831
	4	1.0000	-1.9002	0.2907	0.2780	0.1954	-2.1278	-0.7060	1.2109	0.6072
	5	1.0000	-0.1029	-0.3485	-0.4028	0.2274	-0.8327	0.8508	-0.0598	-1.4278
				--- canonical variables for cases 6 to 105 ---						
	106	1.0000	-0.1029	-0.3485	-0.4028	0.2274	0.5286	-0.8461	1.3204	0.2293
	107	1.0000	0.9525	1.2745	-0.7784	0.9452	1.0792	0.1966	0.0299	0.1056
	108	1.0000	0.9958	1.6503	1.1987	0.0899	-0.7903	1.9462	0.4103	0.2740
	109	1.0000	-0.2387	1.4484	-0.7450	-0.5067	-0.2428	1.8573	-0.3426	-0.1194
	110	1.0000	-0.4109	-2.8813	2.6100	-1.1308	1.0021	-1.0009	0.4685	-0.3382

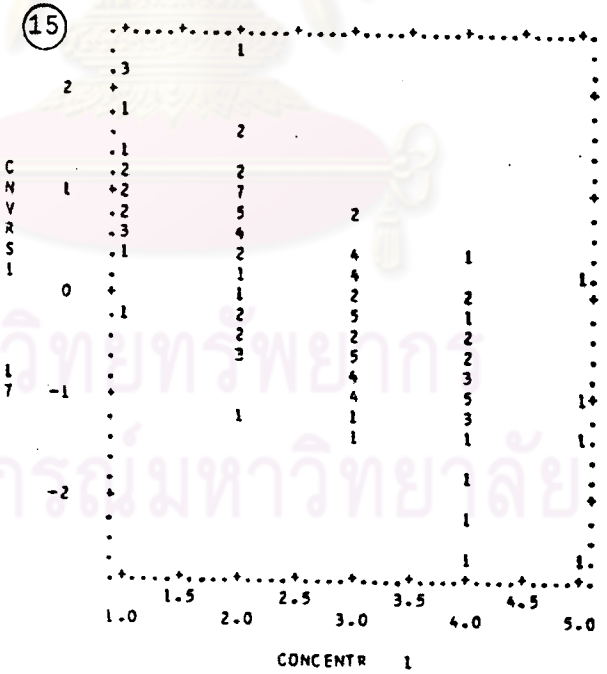
NUMERICAL CONSISTENCY CHECK

(14) THE FOLLOWING VARIANCES OF CANONICAL VARIABLES SHOULD ALL BE EQUAL TO ONE

CANONICAL VARIABLE	VARIANCE	RELATIVE ERROR
CNVRF1	0.1000000 01	0.1151860-14
CNVRF2	0.1000000 01	0.1873500-14
CNVRF3	0.1000000 01	0.1484920-14
CNVRF4	0.1000000 01	0.151268D-14
CNVRS1	0.1000000 01	-0.1776360-14
CNVRS2	0.1000000 01	0.1193490-14
CNVRS3	0.1000000 01	0.2650660-14
CNVRS4	0.1000000 01	-0.1332270-14



(15)



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 7.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 7 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
//SØ3ØPØK JØB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE4;
INPUT CØNCENTR 2 ANNØY 4 SMØKING1 6 SLEEPY 8 SMØKING2 10 TENSE 12
  SMØKING3 14 ALERT 16 IRRITABL 18 TIRED 20 CØNTENT 22 SMØKING4 24;
CARDS;
```

บัตรบันทึกข้อมูล

```
PRØC CANCØRR NG1=4 LIST=4 PLØT=1;
  VAR SMØKING1 SMØKING2 SMØKING3 SMØKING4
    CØNCENTR ANNØY SLEEPY TENSE ALERT IRRITABL TIRED CØNTENT;
TITLE 'CANØNICAL CØRRELATIØN ØF SAS BY SIRICHAI';
/*
//
```

รูปที่ 7.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT 7

CANONICAL CORRECTION OF SAS BY SIRICHAI

CANONICAL CORRELATION ANALYSIS

(2) GROUP 1 BY GROUP 2 CROSS PRODUCT MATRIX

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
SMOKING1	11.36363636	17.27272727	17.63636364	27.18181818	12.63636364	18.27272727
	TIRED	CONTENT				
	23.36363636	24.81818182				

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
SMOKING2	24.78181818	13.43636364	25.01818182	34.49090909	26.21818182	20.03636364
	TIRED	CONTENT				
	30.18181818	22.90909091				

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
SMOKING3	5.52727273	7.44545455	16.37272727	15.06363636	4.97272727	10.74545455
	TIRED	CONTENT				
	16.72727273	10.63636364				

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
SMOKING4	34.00000000	16.50000000	39.50000000	29.50000000	28.50000000	17.00000000
	TIRED	CONTENT				
	36.00000000	20.00000000				

(2)

GROUP 1 SS & CP MATRIX

	SMOKING1	SMOKING2	SMOKING3	SMOKING4
SMOKING1	139.45454545	102.72727273	115.90909091	122.00000000
SMOKING2	102.72727273	122.76363636	109.65454545	120.00000000
SMOKING3	115.90909091	109.65454545	146.91818182	136.50000000
SMOKING4	122.00000000	120.00000000	136.50000000	177.50000000

CANONICAL CORRELTION OF SAS BY SIRICHAI

CANONICAL CORRELATION ANALYSIS

GROUP 2 SS & CP MATRIX

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
CONCENTR	125.49090909	64.01818182	54.70909091	67.14545455	95.50909091	54.41818182
	TIRED	CONTENT				
	97.09090909	46.45454545				

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
ANNOY	64.01818182	103.46363636	39.08181818	74.20909091	62.48181818	66.16363636
	TIRED	CONTENT				
	41.81818182	66.09090909				

---------- ตัวแปร SLEEPY ถึงตัวแปร TIRED จะให้ผลในรูปแบบเช่นเดียวกัน ----------

	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
CONTENT	48.45454545	66.09090909	22.54545455	64.72727273	56.54545455	50.09090909
	TIRED	CONTENT				
	34.45454545	77.27272727				

CANONICAL CORRELTION OF SAS BY SIRICHAI

CANONICAL CORRELATION ANALYSIS

CANONICAL VARIABLE	(4) MEAN OF GROUP 1 CANONICAL VARIABLE	MEAN OF GROUP 2 CANONICAL VARIABLE	(6) CANONICAL CORRELATION	(8) CHI-SQUARE	DF	PROB > CHI-SQ
1	0.2[illegible]	0.31094139	0.58220803	[illegible]	5	0.0044
2	-0.1[illegible]	-0.09437324	0.37587589	[illegible]	4	0.2900
3	0.0[illegible]	0.12350256	0.24039742	0.1756	3	0.7761
4	0.1[illegible]	-0.01960819	0.13710581	1.9[illegible]	2	0.6541

(9) NORMALIZED CHARACTERISTIC VECTOR ASSOCIATED WITH GROUP 1 (VECTOR' * GROUP 1 SSCP MATRIX * VECTOR = I)

CANONICAL VARIABLE	SMOKING1	SMOKING2	SMOKING3	SMOKING4
1	-0.[illegible]	0.10473106	-0.11409166	0.0[illegible]
2	-0.[illegible]	-0.0[illegible]	-0.01665647	0.[illegible]
3	0.09.[illegible]	-0.[illegible]	0.00319620	0.[illegible]
4	-0.[illegible]	0.[illegible]	0.13967106	-0.[illegible]

NORMALIZED CHARACTERISTIC VECTOR ASSOCIATED WITH GROUP 2 (VECTOR' * GROUP 2 SSCP MATRIX * VECTOR = I)

CANONICAL VARIABLE	NERVOUS	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
1	[illegible]	-0.[illegible]	0.02402139	0.[illegible]	-0.[illegible]	-0.00662046
	TIRED	CONTENT				
	[illegible]	0.0561[illegible]				

CANONICAL VARIABLE ที่ 2,3 จะให้ผลในรูปแบบเดียวกัน

	NERVOUS	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL
4	-[illegible]	-0.07063594	0.05898932	0.04296[illegible]	-0.[illegible]	0.10[illegible]
	TIRED	CONTENT				
	[illegible]	-0.[illegible]				

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CANONICAL CORRELTION OF SAS BY SIRICHAI

CANONICAL CORRELATION ANALYSIS

(10)

CORRELATION COEFFICIENTS BETWEEN EACH CANONICAL VARIABLE OF GROUP 1 AND THE VARIABLES OF GROUP 1

CANONICAL	SMOKING1	SMOKING2	SMOKING3	SMOKING4
VAR # 1	0.444916	0.727603	0.289146	0.639460
VAR # 2	-0.534421	-0.384385	-0.269677	0.055806
VAR # 3	0.659347	0.150696	0.473153	0.565104
VAR # 4	0.285642	0.547578	0.787271	0.518300

CORRELATION COEFFICIENTS BETWEEN EACH CANONICAL VARIABLE OF GROUP 2 AND THE VARIABLES OF GROUP 2

CANONICAL	CONCENTR	ANNOY	SLEEPY	TENSE	ALERT	IRRITABL	TIRED	CONTENT
VAR # 1	0.720577	0.302594	0.600499	0.699798	0.730238	0.457458	0.691455	0.532340
VAR # 2	0.355638	-0.141141	0.346766	-0.333112	0.155530	-0.337083	0.026856	-0.440079
VAR # 3	0.013119	0.383854	0.376627	-0.001228	-0.149114	0.114798	0.256453	0.306136
VAR # 4	-0.310695	-0.400694	0.259539	-0.178341	-0.360705	-0.072344	0.071892	-0.557621

CANONICAL CORRELTION OF SAS BY SIRICHAI

CANONICAL CORRELATION ANALYSIS

(13) CANONICAL VARIABLE LISTING

OBS NUMBER	GROUP 1 CAN VAR # 1	GROUP 2 CAN VAR # 1	GROUP 1 CAN VAR # 2	GROUP 2 CAN VAR # 2	GROUP 1 CAN VAR # 3	GROUP 2 CAN VAR # 3
1	0.22657799	0.29699297	0.01225706	0.01414992	-0.03475523	0.18071658
2	0.27211089	0.55325933	-0.24481603	-0.21899018	0.15229605	0.10056872

---------- ค่าสังเกตที่ 3 ถึงค่าสังเกตที่ 108 จะให้ผลเช่นเดียวกัน ----------

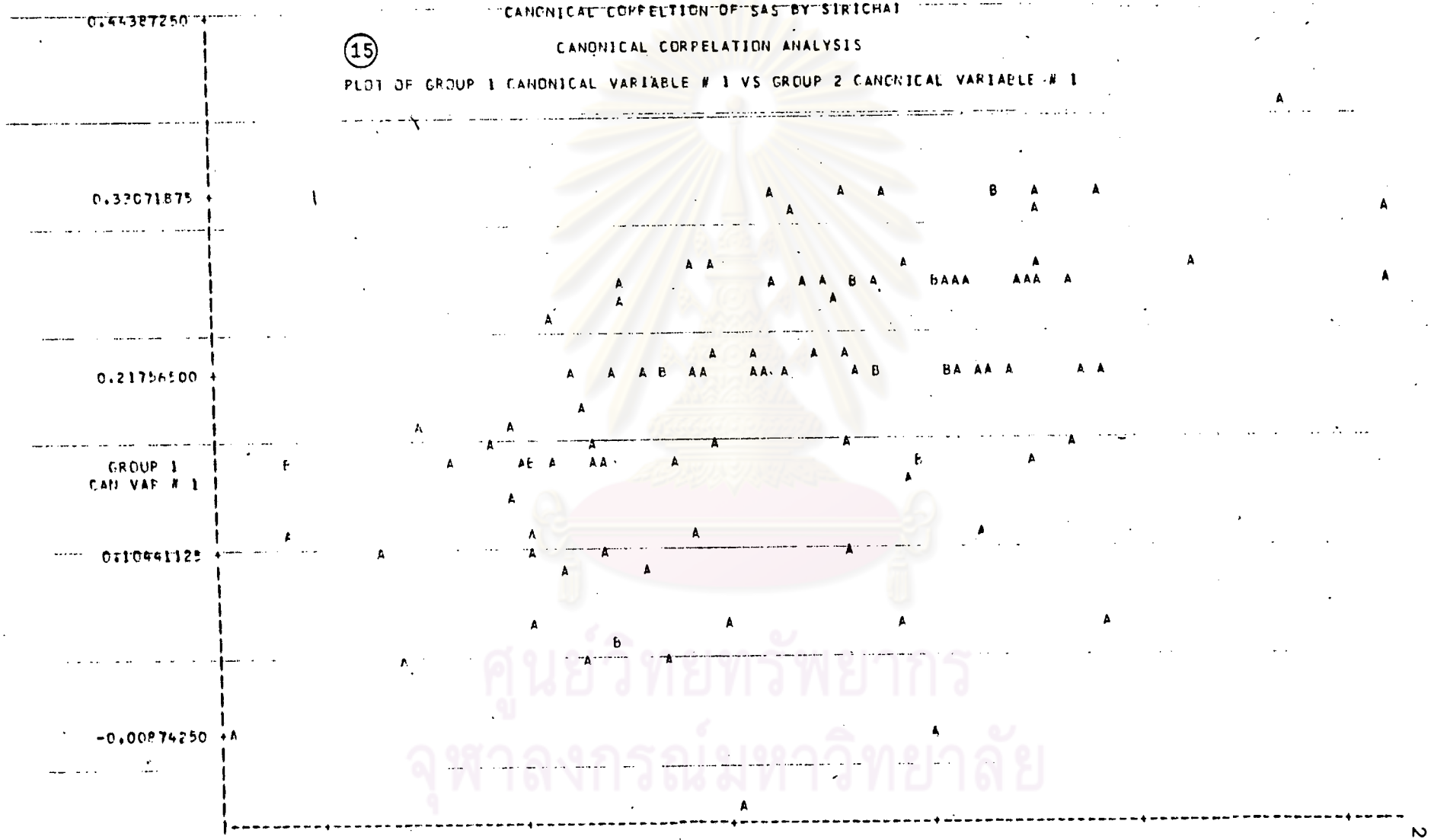
109	0.23077726	0.33419841	-0.02374240	0.08352368	0.15461296	0.15631772
110	0.24721992	0.21495524	-0.43645454	-0.19024143	-0.16673900	0.07862384

CANONICAL CORRELTION OF SAS BY SIRICHAI

(15)

CANONICAL CORRELATION ANALYSIS

PLOT OF GROUP 1 CANONICAL VARIABLE # 1 VS GROUP 2 CANONICAL VARIABLE # 1



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 7.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 7 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK        MR.SIRICHAI PONGWICHAI B221475
// EXEC PROC=$$SPSS79
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME         'DATA SAMPLE4'
VARIABLE LIST     CONCENTR ANNOY SMOKING1 SLEEPY SMOKING2 TENSE
                  SMOKING3 ALERT IRRTABL TIRED CONTENT SMOKING4;
INPUT MEDIUM      CARD
N OF CASES        110
INPUT FORMAT      FIXED(1&F2.0)
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
TASK NAME         'CANONICAL CORRELATION OF SPSS BY SIRICHAI'
CANCORR           VARIABLES = CONCENTR TO SMOKING4/
                  RELATE = SMOKING1 SMOKING2 SMOKING3 SMOKING4
                  WITH CONCENTR ANNOY SLEEPY TENSE ALERT IRRTABL
                  TIRED CONTENT/
OPTIONS           2,6,7
STATISTICS        ALL
FINISH
/*
/&
* $$ EOJ
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 7.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT 7

VARIABLE	(1) MEAN	STANDARD DEV	CASES
CONCENTR	2.6909	1.0730	110
ANNOY	2.1182	0.9743	110
SMOKING1	3.3636	1.4311	110
SLEEPY	2.6091	1.0215	110
SMOKING2	3.5318	1.3013	110
TENSE	2.4455	0.9916	110
SMOKING3	3.4273	1.1610	110
ALERT	2.8091	1.1181	110
IRRITABL	2.2182	0.7826	110
TIRED	3.0909	0.9535	110
CONTENT	2.4545	0.8429	110
SMOKING4	3.5909	1.2761	110

CORRELATION COEFFICIENTS (2)

A VALUE OF 99.00000 IS PRINTED
IF A COEFFICIENT CANNOT BE COMPUTED.

	CONCENTR	ANNOY	SMOKING1	SLEEPY	SMOKING2	TENSE	SMOKING3	ALERT	IRRITABL	TIRED
CONCENTR	1.00001	0.56183	0.08590	0.45702	0.10966	0.57999	0.04071	0.80238	0.59452	0.31197
ANNOY	0.56183	1.00000	0.14380	0.35955	0.11922	0.70473	0.06039	0.57788	0.79608	0.41300
SMOKING1	0.08590	0.14381	1.00000	0.13976	0.78512	0.22234	0.80977	0.10067	0.18537	0.19875
SLEEPY	0.45702	0.35955	0.13976	1.00000	0.21130	0.27258	0.12641	0.60561	0.33651	0.79022
SMOKING2	0.10966	0.11922	0.78512	0.21130	1.00000	0.30079	0.81650	0.22261	0.22132	0.27365
TENSE	0.57999	0.70473	0.22234	0.27258	0.30079	1.00000	0.12005	0.59390	0.72479	0.36433
SMOKING3	0.04071	0.06039	0.80977	0.12641	0.81650	0.12005	1.00000	0.03860	0.10850	0.13863
ALERT	0.80238	0.57788	0.10067	0.60561	0.22261	0.59390	0.03860	1.00000	0.60540	0.69849
IRRITABL	0.59452	0.79608	0.18537	0.33651	0.22132	0.72479	0.10850	0.60540	1.00000	0.42807
TIRED	0.31197	0.41300	0.19875	0.79022	0.27365	0.36433	0.13863	0.69849	0.42807	1.00000
CONTENT	0.49206	0.73915	0.23908	0.24001	0.23521	0.71127	0.09983	0.60515	0.69739	0.39375
SMOKING4	0.22781	0.12176	0.77543	0.27745	0.81292	0.21389	0.84527	0.20124	0.15616	0.27145

	CONTENT	SMOKING4
CONCENTR	0.49206	0.22781
ANNOY	0.73915	0.12176
SMOKING1	0.23908	0.77543
SLEEPY	0.24001	0.27745
SMOKING2	0.23521	0.81292
TENSE	0.71127	0.21389
SMOKING3	0.09983	0.84527
ALERT	0.60515	0.20124
IRRITABL	0.69739	0.15616
TIRED	0.39375	0.27145
CONTENT	1.00000	0.17277
SMOKING4	0.17277	1.00000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

- - - - - - - - - - - - - - - - - C A N O N I C A L C O R R E L A T I O N - - - - - - - - - RELATE LIST 1

NUMBER	(4) EIGENVALUE	(6) CANONICAL CORRELATION	(7) WILK S LAMBDA	(8) CHI-SQUARE	D.F.	SIGNIFICANCE
1	0.27278	0.52229	0.57731	56.31004	32	0.005
2	0.14128	0.37588	0.79336	23.66133	21	0.310
3	0.05779	0.24040	0.92448	8.34903	12	0.781
4	0.01882	0.13719	0.98118	1.04742	5	0.856

(11)

COEFFICIENTS FOR CANONICAL VARIABLES OF THE SECOND SET

	CANVAR 1
CONCENTR	0.47393
ANNOY	-1.78080
SLEEPY	0.25669
TENSE	0.60673
ALERT	-0.14257
IRRITABL	-1.07344
TIRED	0.31269
CONTENT	0.33851

COEFFICIENTS FOR CANONICAL VARIABLES OF THE FIRST SET

	CANVAR 1
SMOKING1	-0.04282
SMOKING2	1.16119
SMOKING3	-1.39291
SMOKING4	0.89846

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT 8 การวิเคราะห์เพื่อพิจารณาความแตกต่างระหว่างกลุ่มตัวแปร (DISCRIMINANT ANALYSIS)

ตารางที่ 8 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 8

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① การเลือกความน่าจะเป็นในการจัดกลุ่ม				②$[Si]$ / = เมทริกซ์ความแปรปรวนจากกลุ่ม i
. ความน่าจะเป็นที่เป็นสัดส่วนกับจำนวนค่าสังเกตในกลุ่ม (pi)	✓	✓	✓	$[Sp]$ / = เมทริกซ์ของความแปรปรวนร่วมจากทุกกลุ่ม (pool-covarinance-matrix)
. ความน่าจะเป็นที่เท่ากันทุกกลุ่ม	✓	✓	✓①	
. ความน่าจะเป็นที่กำหนดอย่างอื่น	✓		✓②	
② ค่าสถิติเบื้องต้นของแต่ละตัวแปรแยกตามกลุ่ม	✓	✓	✓	$D_i^2(x)$ = generalized-square distance-from X to-mean of group i
③ เมทริกซ์สหสัมพันธ์ของแต่ละตัวแปร $[\hat{R}(\underset{\sim}{X})]$				$D_i^2(x) = g_1(x,i)+g_2(x,i)$
เมทริกซ์ของความแปรปรวนร่วม $[Si]$	✓	✓	✓	$g_1(X,i) = (\underset{\sim}{X}-\underset{\sim}{\bar{X}}_i)'Si^{-1}$
เมทริกซ์รวมของสหสัมพันธ์ $[\hat{R}(\underset{\sim}{X})]$ p				$(\underset{\sim}{X}-\underset{\sim}{\bar{X}}_i)+\ln\|Si\|$
เมทริกซ์รวมของความแปรปรวนร่วม $[Sp]$				หรือ $= (\underset{\sim}{X}-\underset{\sim}{\bar{X}}i)\ Sp^{-1}(\underset{\sim}{X}-\underset{\sim}{\bar{X}}i)$
④ ค่า Rank ของ $[Si]$, $[Sp]$ ค่า ln $\|Si\|$ และ ln $\|Sp\|$		✓		$g_2(X,i) = -2\ln pi$; $Pi = \frac{ni}{\sum_{i=1}^{k} Ni}$
⑤ การทดสอบความเท่ากันของ เมทริกซ์ความแปรปรวนระหว่างกลุ่มโดยค่าสถิติ V.RHO หรือ $MC \overset{-1}{\sim} X^2$				หรือ $= 0$
Ho ; $\Sigma_1 = \Sigma_2 = \ldots = \Sigma_k$		✓	✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
⑥ ค่าสถิติ Wilk's lambda -(U-Statistic) และค่าสถิติ F -ใช้ทดสอบความเท่ากันระหว่างค่า -เฉลี่ยของตัวแปรใดๆ ในแต่ละกลุ่ม	✓		✓	$Ci = Ci1V_1+Ci2V_2+\ldots CijVj\ldots+CipV_p+CiO$ Ci = ค่าของฟังก์ชั่นการจัดกลุ่มสำหรับกลุ่มที่ i (classification function)
⑦ การคัดเลือกตัวแปรที่เหมาะสมเพื่อ -สร้างฟังก์ชั่นของความแตกต่างโดย -วิธี Stepwise	✓		✓	Cij= ส.ป.ส. การจัดกลุ่ม-สำหรับกลุ่มที่ i ของตัว-แปรที่ j
⑧ ค่าสัมประสิทธิ์การจัดกลุ่มสำหรับแต่ -ละตัวแปรในแต่ละกลุ่ม (Cij) -พร้อมกับค่าคงที่ (Cio)	✓		✓	Vj = ค่าของตัวแปรที่ j Cio= ค่าคงที่ของฟังก์ชั่นการ-จัดกลุ่มสำหรับกลุ่มที่ i
⑨ ตารางแสดงจำนวนค่าสังเกตที่ถูก -จัดเข้าไปในแต่ละกลุ่มโดยพิจารณา -จากฟังก์ชั่นการจัดกลุ่ม	✓	✓	✓	$\tilde{Ci} = \bar{\tilde{Xj}}\ Sp^{-1} = [Ci1\ Ci2\ \ldots\ Cip]$ $Cio = -0.5\ \bar{\tilde{Xi}}'\ Sp^{-1}\bar{\tilde{Xi}} + lnpi$
⑩ ค่าระยะห่างระหว่างเวคเตอร์ของ -ค่าเฉลี่ยแต่ละกลุ่มทุกคู่ที่เป็นไปได้ -$[D^2(i/j)]$		✓		$\tilde{X} = [V_1\ V_2 \ldots V_p]$ $Ci = \tilde{Ci}\ \tilde{X} - 0.5\ \tilde{Ci}\tilde{Xi} + lnpi$
⑪ ค่าสัมประสิทธิ์สำหรับฟังก์ชั่นความ -แตกต่างของตัวแปรเดิม (d_{ij}) -และตัวแปรในรูปคะแนนมาตรฐาน -ปกติ (aij)	✓		✓	⑩ $D^2(i/j) = (Xi-Xj)\ Sp^{-1}(Xi-Xj)$ หรือ $= (\bar{X}i-\bar{X}j)'\ \bar{S}^1i(\bar{X}i-\bar{X}j') - 2\ lnpi$
⑫ การจัดแต่ละค่าสังเกตเข้าไปในแต่ -ละกลุ่มโดยฟังก์ชั่นความแตกต่าง -(Di) ดังนี้				
. ค่า $D^2i(x)$ ของแต่ละค่าสังเกต	✓	✓		
. ค่าความน่าจะเป็น $P(\underset{\sim}{x}j/i)$			✓	$P(i/X_j) = \dfrac{\exp\left[-.5D^2i(\underset{\sim}{x})\right]}{\sum_{j=1}^{k}\exp\left[-.5D^2_j(x)\right]}$
. ค่าฟังก์ชั่นความแตกต่างของแต่ละค่า -สังเกต	✓		✓	
. ค่าความน่าจะเป็น P (i/xj)	✓	✓	✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(13) ตารางแสดงรายละเอียดเกี่ยวกับฟังก์ชั่นความแตกต่างดังนี้				P(i/xj) = ค่าความน่าจะเป็นที่ค่าสังเกตที่ j จะมาจากกลุ่ม i เมื่อทราบเวคเตอร์ Xj ของตัวแปร
ค่าไอเกน, ค่าสหสัมพันธ์-คาโนนิคอล, ค่าสถิติ	✓		✓	
Wilk's X, Chi-square ใช้-พิจารณาฟังก์ชั่นความแตกต่าง			✓	P(xj/i) = ค่าความน่าจะเป็นที่ค่าสังเกตที่ j จะอยู่ในกลุ่มที่ i เมื่อทราบกลุ่มที่ i
(14) ค่าเฉลี่ยของ Di ในแต่ละกลุ่ม	✓		✓	
(15) กราฟความสัมพันธ์ของฟังก์ชั่น-ความแตกต่าง (Di) แต่ละกลุ่ม-ในรูปเดียวกัน	✓		✓	Di = Discriminant function
(16) แสดง TERRITORIAL MAP ของ-ค่าฟังก์ชั่นความแตกต่าง			✓	Di = $di1V_1 + di2V_2 + \ldots dipV_p + dio$
(17) ค่าร้อยละของความถูกต้องในการ-จัดกลุ่ม	✓		✓	หรือ = $ai1Z_1 + ai2Z_2 + \ldots aijZj + \ldots aipZp$
				Di = ค่าฟังก์ชั่นความแตกต่างที่ i
				dij, aij = ส.ป.ส. ของตัวแปรที่ j (Vj) และของตัวแปรมาตรฐานปกติที่ j (Zj) ของฟังก์ชั่นความแตกต่างที่ i

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 8 โดยใช้ข้อมูลจาก DATA SAMPLE 5

```
//TESTBMDP  JOB (9012A0,90,9087),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
   INCLUDE BMDP(BMDP7M)
   ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM   TITLE IS 'DISCRIMINANT OF BMDP BY SIRICHAI'.
/INPUT     VARIABLES ARE 5.
           FORMAT IS '(4F3.1,F3.0)'.
/VARIABLE  NAMES ARE SEPALLEN,SEPALWID,PETALLEN,IRISTYPE.
           GROUPING IS IRISTYPE.
/GROUP     CODES(5) ARE      1,      2,      3.
           NAMES(5) ARE SETOSA,VERSICOL,VIRGINIC.
/END
```

บัตรบันทึกข้อมูล

```
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 8.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDPสำหรับ STAT 8

```
TOLERANCE. . . . . .    0.010
F-TO-ENTER . . . . . .  4.000    4.000
F-TO-REMOVE. . . . . .  3.996    3.996
METHOD . . . . . .         1
MAXIMUM FORCED LEVEL . . .  0
MAXIMUM NUMBER OF STEPS. .  10
GROUPING VARIABLE. . . . .  5
NUMBER OF GROUPS . . . . .  3
PRIOR PROBABILITIES. . . .  0.33333  0.33333  0.33333
```

(1)

VARIABLE NO.	NAME	BEFORE TRANSFORMATION MINIMUM LIMIT	MAXIMUM LIMIT	MISSING CODE	CATEGORY CODE	CATEGORY NAME	INTERVAL RANGE GREATER THAN	LESS THAN OR EQUAL TO
5	IRISTYPE				1.00000	SETOSA		
					2.00000	VERSICOL		
					3.00000	VIRGINIC		

NUMBER OF CASES READ. 150

(2)

MEANS

GROUP = VARIABLE	SETOSA	VERSICOL	VIRGINIC	ALL GPS.
1 SEPALLEN	5.00600	5.93600	6.58800	5.84333
2 SEPALWID	3.42800	2.77000	2.97400	3.05733
3 PETALLEN	1.46200	4.26000	5.55200	3.75800
4 PETALWID	0.24600	1.32600	2.02600	1.19933
5 IRISTYPE	1.00000	2.00000	3.00000	2.00000
COUNTS	50.	50.	50.	150.

--- standard deviations and coefficients of variation are printed in separate panels ---

(7)

****

STEP NUMBER 0

VARIABLE	F TO REMOVE	FORCE LEVEL	TOLERANCE	*	VARIABLE	F TO ENTER	FORCE LEVEL	TOLERANCE
	DF= 2 148			*		DF= 2 147		
				*	1 SEPALLEN	119.264	1	1.000000
				*	2 SEPALWID	49.160	1	1.000000
				*	3 PETALLEN	1180.160	1	1.000000
				*	4 PETALWID	960.007	1	1.000000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

------- STEP NUMBER 1-3 จะให้ผลในรูปแบบเช่นเดียวกันกับ STEP NUMBER 0,4 ----------

```
STEP NUMBER   4
VARIABLE ENTERED    1 SEPALLEN

  VARIABLE          F TO  FORCE  TOLERANCE  *      VARIABLE          F TO  FORCE   TOLERANCE
                    REMOVE LEVEL             *                        ENTER  LEVEL
                DF=   2  144                 *                    DF=   2  143
    1 SEPALLEN       4.721   1    0.347993   *
    2 SEPALWID      21.936   1    0.608860   *
    3 PETALLEN      35.590   1    0.365126   *
    4 PETALWID      24.904   1    0.649314   *

U-STATISTIC OR WILKS' LAMBDA    0.0234386     DEGREES OF FREEDOM   4    2     147
APPROXIMATE F-STATISTIC           199.145     DEGREES OF FREEDOM      8.00   288.00

 F - MATRIX         DEGREES OF FREEDOM =     4  144

             SETOSA    VERSICOL
VERSICOL    550.19
VIRGINIC   1098.27     105.31

CLASSIFICATION FUNCTIONS

          GROUP =    SETOSA        VERSICOL      VIRGINIC
VARIABLE
  1 SEPALLEN        23.54416       15.69820      12.44584
  2 SEPALWID        23.58786        7.07252       3.68529
  3 PETALLEN       -16.43063        5.21145      12.76655
  4 PETALWID       -17.39839        6.43422      21.07909

CONSTANT           -86.30843      -72.85257    -104.36826

CLASSIFICATION MATRIX

GROUP       PERCENT    NUMBER OF CASES CLASSIFIED INTO GROUP -
            CORRECT
                         SETOSA    VERSICOL VIRGINIC
 SETOSA      100.0         50          0        0
 VERSICOL     96.0          0         48        2
 VIRGINIC     98.0          0          1       49

TOTAL         98.0         50         49       51

JACKKNIFED CLASSIFICATION

GROUP       PERCENT    NUMBER OF CASES CLASSIFIED INTO GROUP -
            CORRECT
                         SETOSA    VERSICOL VIRGINIC
 SETOSA      100.0         50          0        0
 VERSICOL     96.0          0         48        2
 VIRGINIC     98.0          0          1       49

TOTAL         98.0         50         49       51
```

(6) (8) (9) (17)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	F VALUE TO ENTER OR REMOVE	NUMBER OF VARIABLES INCLUDED	U-STATISTIC	APPROXIMATE F-STSTISTIC	DEGREES OF FREEDOM	
1	3 PETALLEN		1180.1597	1	0.0586	1180.161	2.00	147.00
2	2 SEPALWID		43.0353	2	0.0369	307.104	4.00	292.00
3	4 PETALWID		34.5686	3	0.0250	257.303	6.00	290.00
4	1 SEPALLEN		4.7211	4	0.0234	199.145	8.00	288.00

(12) INCORRECT CLASSIFICATIONS

MAHALANOBIS D-SQUARE AND POSTERIOR PROBABILITY FOR GROUP -

GROUP SETOSA	SETOSA	VERSICOL	VIRGINIC
CASE			
1	0.2 1.000	90.7 0.000	181.6 0.000
6	1.3 1.000	84.0 0.000	170.1 0.000
10	2.3 1.000	113.7 0.000	210.0 0.0
18	2.8 1.000	87.5 0.000	165.7 0.000
26	4.0 1.000	113.2 0.000	210.2 0.0

--- similar statistics for remaining cases in SETOSA ---

GROUP VERSICOL		SETOSA	VERSICOL	VIRGINIC
CASE				
3		105.3 0.000	2.2 0.996	13.1 0.004
8		131.7 0.000	8.4 0.960	14.8 0.040
9	VIRGINIC	130.9 0.000	6.7 0.253	6.5 0.747
11		99.2 0.000	1.3 0.998	13.8 0.002
12	VIRGINIC	149.0 0.000	8.4 0.143	4.9 0.857

--- similar statistics for remaining cases in VERSICOL ---

GROUP VIRGINIC		SETOSA	VERSICOL	VIRGINIC
CASE				
2		208.6 0.0	27.3 0.000	1.9 1.000
4		207.9 0.0	31.7 0.000	4.5 1.000
5	VERSICOL	133.1 0.000	5.3 0.729	7.2 0.271
7		173.2 0.000	26.6 0.000	11.0 1.000
13		159.0 0.000	12.8 0.003	1.2 0.997

--- similar statistics for remaining cases in VIRGINIC ---

EIGENVALUES

(13) 32.19192 0.28539

CANONICAL CORRELATIONS

0.98482 0.47120

VARIABLE (11) COEFFICIENTS FOR CANONICAL VARIABLES

1 SEPALLEN	0.82938	0.02410
2 SEPALWID	1.53447	2.16452
3 PETALLEN	-2.20121	-0.93192
4 PETALWID	-2.81046	2.83919
CONSTANT	2.10510	-6.66147

GROUP	CANONICAL VARIABLES EVALUATED AT GROUP MEANS	
SETOSA	7.60760	0.21514
VERSICOL	-1.82505	-0.72790
VIRGINIC	-5.78255	0.51277

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

POINTS TO BE PLOTTED

GROUP (14)	MEAN COORDINATES		SYMBOL FOR CASES	SYMBOL FOR MEAN
SETOSA	7.61	0.22	A	1
VERSICOL	-1.83	-0.73	B	2
VIRGINIC	-5.78	0.51	C	3

GROUP SETOSA

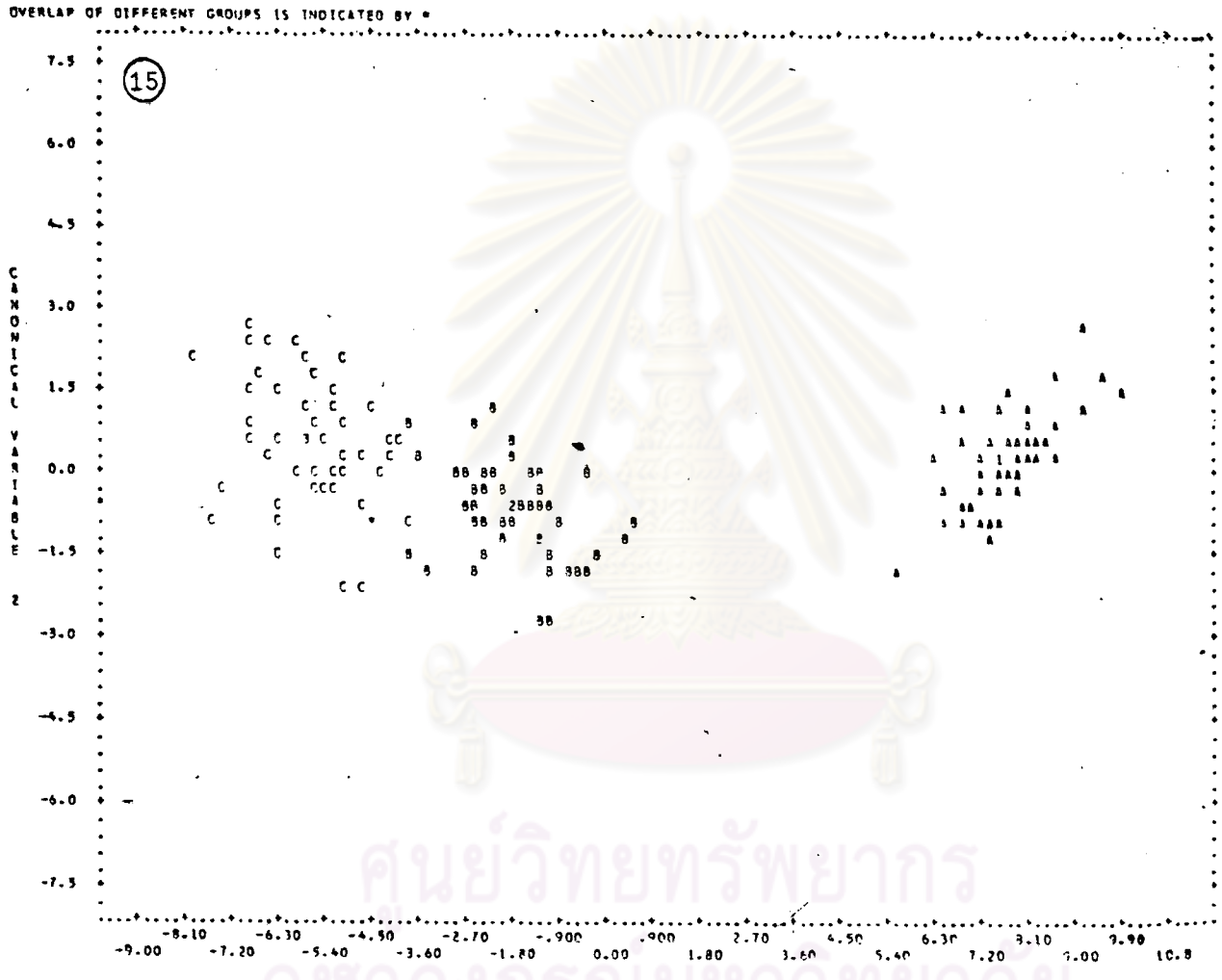
CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y
1	7.67	-0.13	44	6.93	-0.71	64	7.13	-0.79	92	7.70	1.46	126	7.78	0.58
6	7.21	0.36	47	5.83	-0.54	65	8.06	0.30	96	7.61	-0.01	135	7.22	-0.11
10	8.68	0.88	51	6.76	-0.51	68	6.82	0.46	97	6.56	-1.32	136	7.33	-1.07
18	6.25	0.44	52	8.08	0.76	69	7.19	-0.36	101	7.49	-0.27	137	5.66	-1.93
26	8.61	0.40	54	6.56	-0.39	72	8.16	2.74	102	5.81	-0.67	139	8.08	0.97
31	6.76	-0.76	55	6.77	-0.97	73	8.13	0.51	107	7.69	-0.01	140	8.02	1.14
36	7.99	0.09	56	7.96	-0.18	79	7.37	0.57	108	7.92	0.68	144	7.50	-0.19
37	8.33	0.23	59	7.57	-0.81	80	8.13	1.22	113	7.34	-0.95	145	7.59	1.21
40	7.24	-0.27	60	8.85	1.50	88	8.47	1.83	116	8.40	0.65	146	7.92	0.21
42	6.41	1.25	61	6.86	1.05	89	7.06	-0.66	125	8.58	1.43	150	8.31	0.64

GROUP VERSICOL

CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y
3	-2.55	-0.47	30	-1.00	-0.49	71	0.31	-1.32	100	-0.90	-0.90	121	0.48	-0.80
8	-3.50	-1.68	33	-1.11	-1.75	77	-1.62	-0.47	104	-1.42	-0.55	122	-1.55	-0.59
9	-3.72	1.04	38	-2.93	0.03	85	-0.22	-1.58	106	-1.86	0.32	129	-2.67	-0.64
11	-2.29	-0.33	43	-1.27	-1.21	86	-2.01	-0.91	110	-1.16	-2.64	130	-0.38	0.09
12	-4.50	-0.88	48	-2.40	-1.59	87	-1.18	-0.54	114	-2.14	0.09	131	-2.42	-0.09
14	-1.09	-1.63	49	-0.29	-1.80	93	-1.75	-0.82	115	-2.48	-1.94	134	-1.96	-1.15
19	-2.43	-0.97	62	-1.20	0.08	94	-1.96	-0.35	117	-1.33	-0.16	141	-2.26	-1.59
22	-2.45	0.80	66	-2.77	0.03	95	-2.10	1.19	118	-3.34	-1.41	142	-1.44	-0.13
28	-1.46	0.03	67	-0.78	-1.66	98	-1.19	-2.63	119	-2.26	-1.43	143	-2.46	-0.94
29	-1.80	0.48	70	-2.59	-0.17	99	-0.61	-1.94	120	-1.26	-0.55	148	-3.52	0.16

GROUP VIRGINIC

CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y	CASE	X	Y
2	-6.80	0.58	23	-6.85	2.43	46	-4.37	-0.12	81	-6.29	0.47	112	-5.33	-1.38
4	-6.65	1.81	24	-7.42	-0.17	50	-5.29	-0.04	82	-5.00	0.19	123	-6.85	0.83
5	-3.82	-0.94	25	-4.68	-0.50	53	-4.08	0.52	83	-3.94	0.61	124	-5.20	1.14
7	-5.11	1.99	27	-5.65	1.68	57	-5.11	-2.13	84	-5.61	-0.34	127	-6.95	2.72
13	-4.97	0.82	32	-4.68	0.33	58	-6.52	0.30	90	-9.17	-0.75	128	-4.44	1.35
15	-5.07	-0.03	34	-5.18	-0.36	63	-4.08	0.19	91	-4.76	-2.16	132	-5.45	-0.21
16	-5.51	-0.04	35	-5.81	2.01	74	-6.80	0.86	103	-6.27	1.65	133	-5.66	0.83
17	-6.80	1.46	39	-5.22	1.47	75	-6.52	2.45	105	-6.23	-0.71	138	-5.96	-0.09
20	-5.89	2.35	41	-5.72	1.29	76	-5.51	-0.04	109	-5.36	0.65	147	-6.76	1.60
21	-6.61	1.75	45	-7.58	-0.98	78	-4.58	-0.86	111	-6.32	-0.97	149	-7.84	2.14



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 8 โดยใช้ข้อมูลจาก DATA SAMPLE 5

```
//SØ38ØPØK JØB CLASS=C,MSGLEVEL=(2,0)
//  EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE5;
INPUT SEPALLEN 2-3 SEPALWID 5-6 PETALLEN 8-9 PETALWID 11-12
      IRISTYPE 15;
CARDS;
```

บัตรบันทึกข้อมูล

```
PRØC DISCRIM S PØØL = TEST LIST;
 CLASS IRISTYPE;
 VAR SEPALLEN SEPALWID PETALLEN PETALWID IRISTYPE;
 TITLE 'DISCRIMINANT ANALYSIS ØF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 8/.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STST 8

IRISTYPE	FREQUENCY	PRIOR PROBABILITY
1	50	0.33333333
2	50	0.33333333
3	50	0.33333333
TOTAL	150	1.00000000

DISCRIMINANT ANALYSIS SIMPLE STATISTICS

IRISTYPE = 1

VARIABLE	N	SUM	MEAN	VARIANCE	STANDARD DEVIATION
SEPALLEN	50	2503.00000000	50.06000000	12.42489796	[illegible]
SEPALWID	50	1714.00000000	34.28000000	14.36897959	[illegible]
PETALLEN	50	731.00000000	14.62000000	3.01591837	[illegible]
PETALWID	50	123.00000000	2.46000000	1.11061224	[illegible]
IRISTYPE	50	50.00000000	1.00000000	0.00000000	0.00000000

IRISTYPE = 2

VARIABLE	N	SUM	MEAN	VARIANCE	STANDARD DEVIATION
SEPALLEN	50	2968.00000000	59.36000000	26.64326531	[illegible]
SEPALWID	50	1385.00000000	27.70000000	9.84693878	[illegible]
PETALLEN	50	2130.00000000	42.60000000	22.06163265	[illegible]
PETALWID	50	663.00000000	13.26000000	3.91061224	[illegible]
IRISTYPE	50	100.00000000	2.00000000	0.00000000	0.00000000

IRISTYPE = 3

VARIABLE	N	SUM	MEAN	VARIANCE	STANDARD DEVIATION
SEPALLEN	50	3294.00000000	65.88000000	40.43428571	[illegible]
SEPALWID	50	1490.00000000	29.80000000	10.00000000	[illegible]
PETALLEN	50	2776.00000000	55.52000000	30.45877551	[illegible]
PETALWID	50	1013.00000000	20.26000000	7.54326531	[illegible]
IRISTYPE	50	150.00000000	3.00000000	0.00000000	0.00000000

DISCRIMINANT ANALYSIS OF SAS BY SIRICHAI

(5)

DISCRIMINANT ANALYSIS TEST OF HOMOGENIETY OF WITHIN COVARIANCE MATRICES

NOTATION:

K = NUMBER OF GROUPS

P = NUMBER OF VARIABLES

N = TOTAL NUMBER OF OBSERVATIONS

N(I) = NUMBER OF OBSERVATIONS IN THE I'TH GROUP

$$V = \frac{\prod |\text{WITHIN SS MATRIX(I)}|^{N(I)/2}}{|\text{POOLED SS MATRIX}|^{N/2}}$$

$$RHO = 1.0 - \left(\text{SUM} \frac{1}{N(I)-1} - \frac{1}{N-K} \right) \frac{2P^2 + 3P - 1}{6(P+1)(K-1)}$$

$$DF = .5(K-1)P(P+1)$$

UNDER NULL HYPOTHESIS: $-2 \text{ RHO LN} \left| \frac{N^{PN/2} \; V}{\prod N(I)^{PN(I)/2}} \right|$ IS DISTRIBUTED APPROXIMATELY AS CHI-SQUARE(DF)

TEST CHI-SQUARE VALUE = -14.17477239 WITH 30 DF PROB > CHI-SQ = 1.0000

SINCE THE CHI-SQUARE VALUE IS NOT SIGNIFICANT AT THE 0.1000 LEVEL, THE POOLED COVARIANCE MATRIX WILL BE USED IN THE DISCRIMINANT FUNCTION.

REFERENCE: KENDALL,M.G. AND STUART,A. THE ADVANCED THEORY OF STATISTICS VOL.3, P266 & 282.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

DISCRIMINANT ANALYSIS OF SAS BY SIRICHAI

DISCRIMINANT ANALYSIS PAIRWISE SQUARED GENERALIZED DISTANCES BETWEEN GROUPS

$$\text{WHERE: } D^2(I|J) = (\bar{X}_I - \bar{X}_J)' \, COV^{-1} \, (\bar{X}_I - \bar{X}_J)$$

(10) GENERALIZED SQUARED DISTANCE TO IRISTYPE

FROM IRISTYPE	1	2	3
1	0.00000000	89.37575966	178.23967660
2	89.37575966	0.00000000	17.10783041
3	178.23967660	17.10783041	0.00000000

DISCRIMINANT ANALYSIS CLASSIFICATION RESULTS FOR EACH OBSERVATION
GIVING GENERALIZED SQUARED DISTANCE TO EACH GROUP / POSTERIOR PROBABILITY OF MEMBERSHIP IN EACH GROU

$$\text{WHERE: } D^2_J(X) = (X - \bar{X}_J)' \, COV^{-1} \, (X - \bar{X}_J) \quad \text{AND} \quad PR(J|X) = EXP(-.5\, D^2_J(X)) \,/\, SUM\, EXP(-.5\, D^2_K(X))$$

(12) GENERALIZED SQUARED DISTANCE TO IRISTYPE

IRISTYPE	FROM IRISTYPE	CLASSIFIED INTOIRISTYPE	1	2	3
1	1	1	0.246085 1.0000	90.177982 0.0000	180.451152 0.0000
3	3	3	207.542653 0.0000	27.221247 0.0000	1.931802 1.0000
2	2	2	104.702703 0.0000	2.236486 0.9956	13.070972 0.0044
3	3	3	206.960447 0.0000	31.673055 0.0000	4.479743 1.0000
1	1	1	1.249066 1.0000	104.708544 0.0000	198.025350 0.0000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

DISCRIMINANT ANALYSIS OF SAS BY SIRICHAI

DISCRIMINANT ANALYSIS SUMMARY OF CLASSIFICATION PERFORMANCE USING GENERALIZED SQUARED DISTANCE

WHERE: $D^2_J(X) = (X - \bar{X}_J)' COV^{-1} (X - \bar{X}_J)$

(9) NUMBER OF OBSERVATIONS CLASSIFIED INTO IRISTYPE

FROM IRISTYPE	1	2	3
1	50	0	0
2	0	48	2
3	0	1	49

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 8.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 8 โดยใช้ข้อมูลจาก DATA SAMPLE 5

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK        MR. SIRICHAI PONGWICHAI B221475
// EXEC PROC=$$SPSS79
RUN NAME          'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME         'DATA SAMPLE5'
VARIABLE LIST     SEPALLEN SEPALWID PETALLEN PETALWID IRISTYPE
INPUT MEDIUM      CARD
N OF CASES        150
INPUT FORMAT      FIXED(4F3.1,F3.0)
VALUE LABELS      IRISTYPE(1)SETOSA(2)VERSICOL(3)VIRGINIC
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
TASK NAME         'DISCRIMINANT ANALYSIS OF SPSS BY SIRICHAI'
DISCRIMINANT      GROUPS=IRISTYPE(1,3)/VARIABLES=SEPALLEN TO PETALWID/
                  ANALYSIS=SEPALLEN TO PETALWID(3)/METHOD=WILKS/
                  TOLERANCE=0.01/FIN=1.0/FOUT=3.9/
OPTIONS           5,6,7,10,11,12,13,15
STATISTICS        ALL
FINISH
/*
/&
* $$ EOJ
```

รูปที่ 8.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT 8

GROUP COUNTS

	GROUP 1 SETOSA	GROUP 2 VERSICOL	GROUP 3 VIRGINIC	TOTAL
COUNT	50.0000	50.0000	50.0000	150.0000

MEANS (2)

	GROUP 1 SETOSA	GROUP 2 VERSICOL	GROUP 3 VIRGINIC	TOTAL
SEPALLEN	5.0060	5.9360	6.5880	5.8433
SEPALWID	3.4280	2.7700	2.9800	3.0593
PETALLEN	1.4620	4.2600	5.5520	3.7580
PETALWID	0.2460	1.3260	2.0260	1.1993
IRISTYPE	1.0000	2.0000	3.0000	2.0000

STANDARD DEVIATIONS

	GROUP 1 SETOSA	GROUP 2 VERSICOL	GROUP 3 VIRGINIC	TOTAL
SEPALLEN	0.3524	0.5164	0.6362	0.8281
SEPALWID	0.3791	0.3138	0.3162	0.4342
PETALLEN	0.1737	0.4699	0.5520	1.7654
PETALWID	0.1054	0.1978	0.2747	0.7622
IRISTYPE	0.0	0.0	0.0	0.8192

WILKS' LAMBDA (U-STATISTIC) AND UNIVARIATE F-RATIO WITH 2 AND 147 DEGREES OF FREEDOM

VARIABLE	WILKS' LAMBDA (6)	F
SEPALLEN	0.3033	119.5214
SEPALWID	0.5965	49.6035
PETALLEN	0.0586	1180.3325
PETALWID	0.0711	961.2454
IRISTYPE	0.0333**********	**********

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

D I S C R I M I N A N T   A N A L Y S I S

ANALYSIS NUMBER 1

F FOR INCLUSION 4.00000

F FOR DELETION 1.90000

TOLERANCE LEVEL 0.01000

MAXIMUM STEPS 10

SOLUTION METHOD - STEPWISE. SELECT VARIABLE WHICH WILL MINIMIZE WILKS' LAMBDA.

NOTE: THE ENTRY CRITERIA CURRENTLY BEING PRINTED ARE (USABLE) INTERMEDIATE RESULTS.

PRIOR PROBABILITIES -

(1)

GROUP 1 SETOSA	GROUP 2 VERSICUL	GROUP 3 VIRGINIC
0.33333	0.33333	0.33333

---

(7)

VARIABLE ENTERED ON STEP NUMBER 1.. PETALLEN

				DEGREES OF FREEDOM		SIGNIFICANCE
WILKS' LAMBDA	0.05862	APPROXIMATE F	1180.33130	2	147.00	0.0
RAO'S V	2360.41113	CHANGE IN V	2360.41113	2		0.000

F MATRIX - DEGREES OF FREEDOM: 1, 147

	GROUP 1	GROUP 2
GROUP 2	1056.91602	
GROUP 3	2259.34346	225.35432

-------- VARIABLES IN THE ANALYSIS ---------

VARIABLE	ENTRY CRITERION	F TO REMOVE
PETALLEN	1180.33229	1180.33228

----------- VARIABLES NOT IN THE ANALYSIS ------------

VARIABLE	TOLERANCE	F TO ENTER	ENTRY CRITERION
SEPALLEN	0.42758	34.61652	119.52135
SEPALWID	0.86160	43.17400	49.69350
PETALWID	0.76606	24.83459	950.24536

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

- D I S C R I M I N A N T A N A L Y S I S -

SUMMARY TABLE

STEP NUMBER	VARIABLE ENTERED	VARIABLE REMOVED	F TO ENTER OR REMOVE	NUMBER INCLUDED	WILKS' LAMBDA	SIG.	RAO'S V	CHANGE IN RAO'S V	SIG. OF CHANGE
1	PETALLEN		1180.33228	1	0.05062	0.0	2360.41113	2360.41113	0.000
2	SEPALWID		43.17400	2	0.03684	0.0	3191.56958	831.15845	0.000
3	PETALWID		34.55560	3	0.02495	0.0	4446.14063	1254.57104	0.000
4	SEPALLEN		4.75984	4	0.02340	0.0	4746.96875	300.82813	0.000

(8)

CLASSIFICATION FUNCTION COEFFICIENTS

	GROUP 1 SETOSA	GROUP 2 VERSICOL	GROUP 3 VIRGINIC
SEPALLEN	23.41666	15.56596	12.24103
SEPALWID	23.01419	7.42238	4.22912
PETALLEN	-16.21759	5.31473	12.08926
PETALWID	-17.50044	6.26483	20.79236
CONSTANT	-85.42160	-71.05094	-103.46883

(13)

DISCRIMINANT FUNCTION	EIGENVALUE	RELATIVE PERCENTAGE	CANONICAL CORRELATION	:	FUNCTIONS DERIVED	WILKS' LAMBDA	CHI-SQUARE	DF	SIGNIFICANCE
				:					
				:	0	0.0234	546.057	8	0.000
1	32.11060	99.09	0.985	:	1	0.7738	37.312	3	0.000
2	0.29233	0.91	0.476	:					

REMAINING COMPUTATIONS WILL BE BASED ON 2 DISCRIMINANT FUNCTION(S)

(11)

STANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	FUNC 1	FUNC 2
SEPALLEN	0.12145	0.11209
SEPALWID	0.11565	-0.85727
PETALLEN	-0.68071	1.36032
PETALWID	-0.37406	-1.05211

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

UNSTANDARDIZED DISCRIMINANT FUNCTION COEFFICIENTS

	FUNC 1	FUNC 2
SEPALLEN	0.14663	0.01459
SEPALWID	0.26593	-1.97440
PETALLEN	-0.38559	0.77038
PETALWID	-0.49179	-2.42985
CONSTANT	0.36859	5.07412

CENTROIDS OF GROUPS IN REDUCED SPACE

(14)

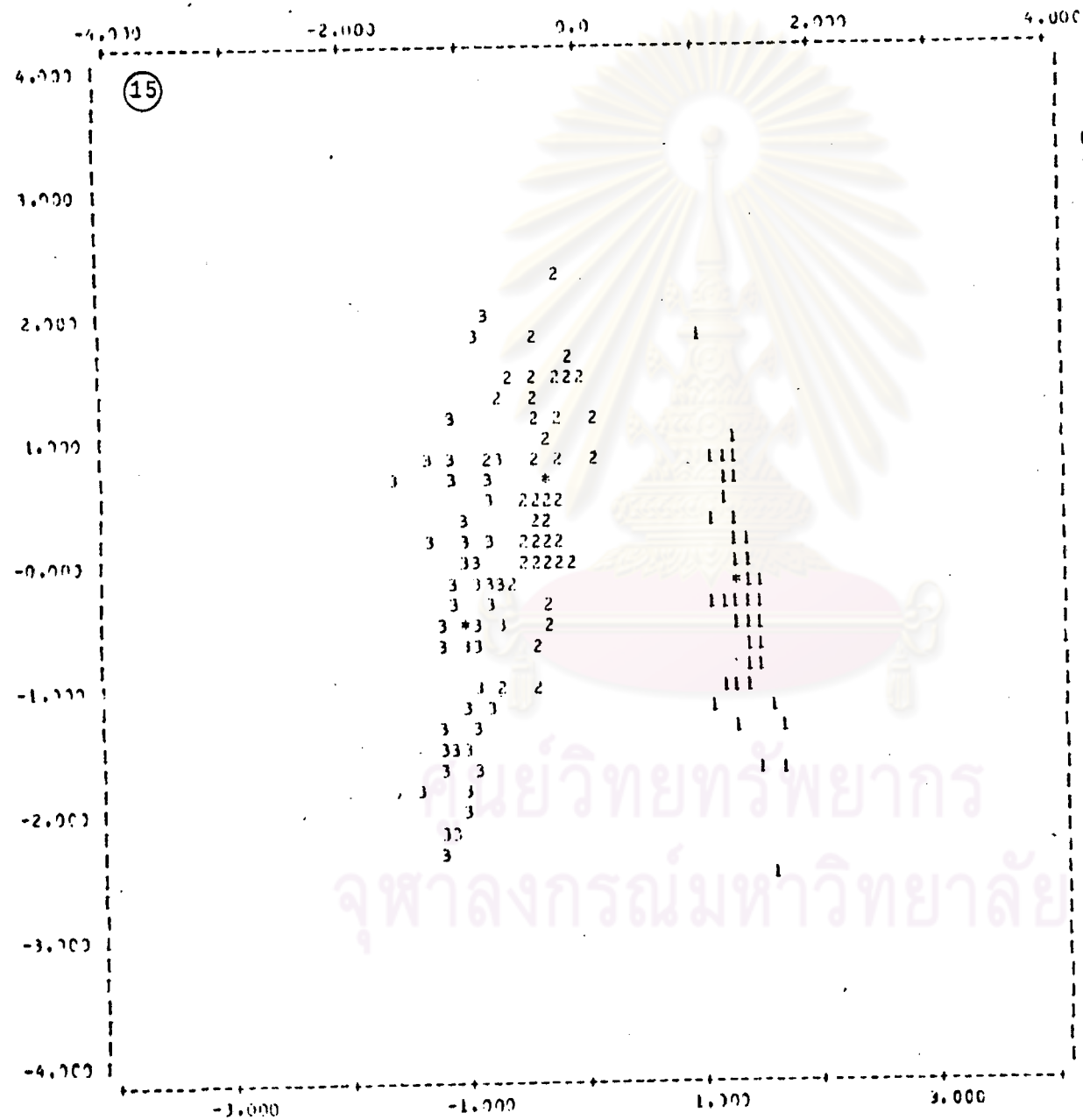
	FUNC 1	FUNC 2
GROUP 1 SETOSA	1.32949	-0.19251
GROUP 2 VERSICOL	-0.31921	0.65152
GROUP 3 VIRGINIC	-1.01029	-0.45915

(12)

CASE SUBFIL	SEQNUM	MISSING VALUES	ACTUAL GROUP		HIGHEST PROBABILITY GROUP	P(X/G)	P(G/X)	2ND HIGHEST GROUP	P(G/X)	DISCRIMINANT SCORES FUNC 1	FUNC 2
'DAT	1.	0	1		1	0.598	1.000			1.341	0.124
'DAT	2.	0	1		1	0.902	1.000			-1.100	-0.492
'DAT	3.	0	2		2	0.966	0.996	3	0.004	-0.445	0.440
'DAT	4.	0	3		3	0.674	1.000			-1.164	-1.566
'DAT	34	0	3	****	2	0.408	0.725	3	0.275	-0.667	0.822
'DAT	147.	0	3		3	0.729	1.000			-1.183	-1.372
'DAT	148.	0	2		2	0.458	0.688	3	0.312	-0.615	-0.130
'DAT	149.	0	3		3	0.139	1.000			-1.373	-1.902
'DAT	150.	0	1		1	0.952	1.000			1.453	-0.564

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

PLOT OF DISCRIMINANT SCORE 1 (HORIZONTAL) VS. DISCRIMINANT SCORE 2 (VERTICAL). * INDICATES A GROUP CENTROID.

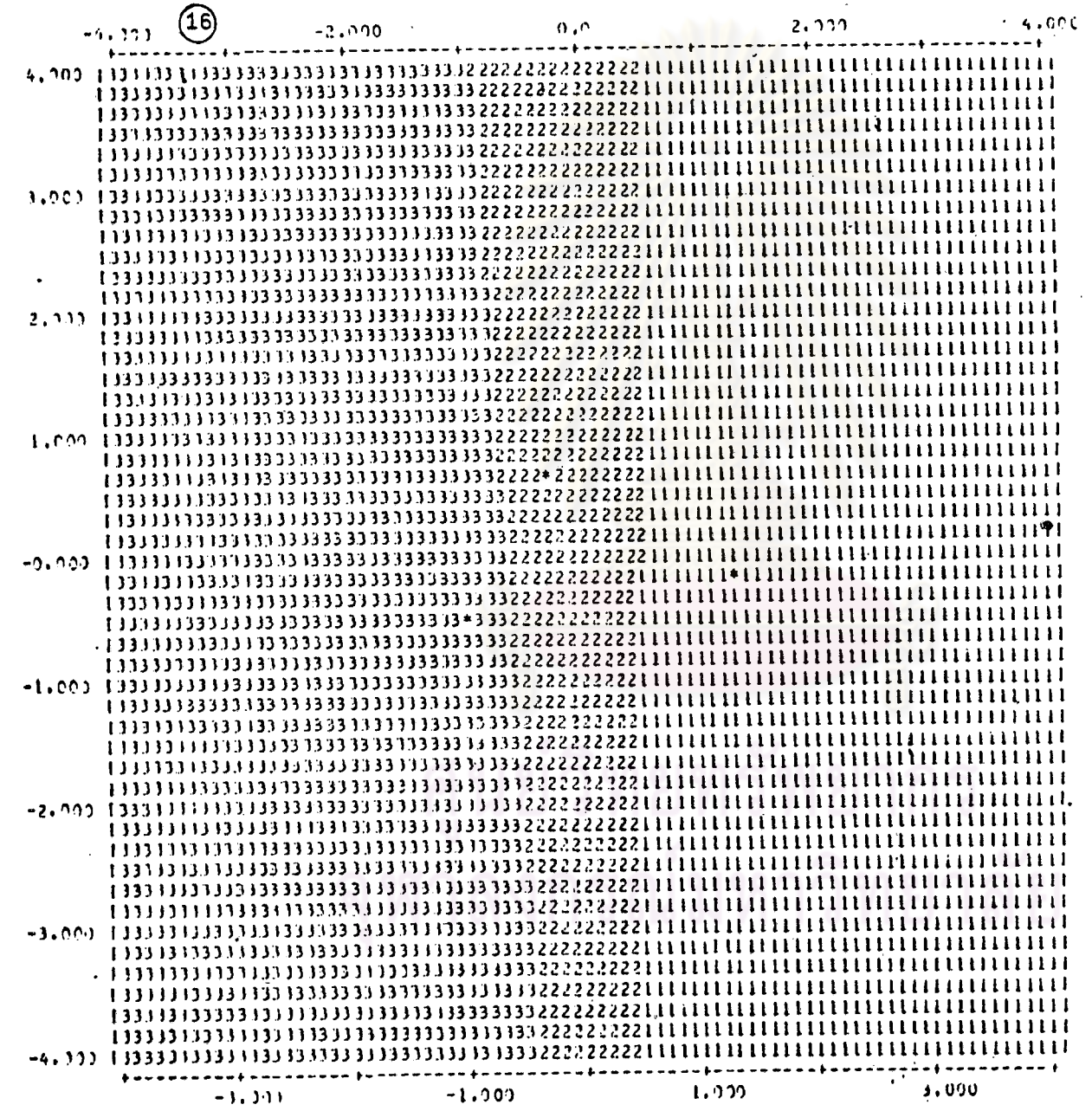


DIRECTORY OF PLOT SYMBOLS -

SYMBOL	GROUP	LABEL
1	1	SETOSA
2	2	VERSICOL
3	3	VIRGINIC
0		ALL UNGROUPED CASES
*		CENTROIDS

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

TERRITORIAL MAP OF DISCRIMINANT SCORE 1 (HORIZONTAL) VS. DISCRIMINANT SCORE 2 (VERTICAL). * INDICATES A GROUP CENTROID.



PREDICTION RESULTS -

ACTUAL GROUP	NO. OF CASES	PREDICTED GROUP MEMBERSHIP GP. 1	GP. 2	GP. 3
GROUP 1 SETOSA	50.	50. 100.0%	0. 0.0%	0. 0.0%
GROUP 2 VERSICOL	50.	0. 0.0%	48. 96.0%	2. 4.0%
GROUP 3 VIRGINIC	50.	0. 0.0%	1. 2.0%	49. 98.0%

PERCENT OF "GROUPED" CASES CORRECTLY CLASSIFIED: 98.00%

(9) (17)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

STAT 9 การวิเคราะห์ตัวประกอบ

(FACTØR ANALYSIS)

ตารางที่ 9 แสดงการเปรียบเทียบความสามารถของแต่ละโปรแกรมสำหรับ STAT 9

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
① ค่าสถิติเบื้องต้นของแต่ละตัวแปร	✓	✓	✓	SMC = Square multiple correlation R^2
② เมทริกซ์สหสัมพันธ์	✓	✓	✓	h^2 = Communulity
เมทริกซ์ความแปรปรวนร่วม	✓			h^2 = SMC ของแต่ละตัวแปรกับตัวประกอบ
③ ค่าสหสัมพันธ์พหุคูณกำลังสอง (SMC)--ของแต่ละตัวแปรกับตัวแปรอื่นๆที่เหลือ	✓			Xi = ค่าไอเกนมีจำนวนเท่ากับจำนวนตัวประกอบ(i=1,2,3...m)
④ ค่าความร่วมกัน (h^2) ระหว่าง-ตัวแปรกับตัวประกอบ	✓	✓	✓	= ความแปรปรวนของแต่ละตัวประกอบ
⑤ การเลือกประมาณค่าความร่วมกัน-ก่อนการวิเคราะห์ตัวประกอบ-โดยต่าต่างๆ	✓		✓	$= \sum_{j=1}^{n}(aij)^2$
⑥ .ค่าไอเกนของตัวประกอบ (X_i)	✓	✓	✓	aij = น้ำหนักของตัวประกอบที่ i สำหรับตัวแปรที่ j
.ค่าไอเกนสะสม ($Xi/\sum_{i=1}^{m} Xi$)	✓	✓	✓	n = จำนวนตัวแปร
.เวคเตอร์ของค่าไอเกน		✓		m = จำนวนตัวประกอบ $m \leq n$
⑦ เมทริกตัวประกอบ (amj) ที่ได้-จากการสกัดตัวประกอบโดยวิธีต่อไปนี้				$Zj = a_{1j}F_1 + a_{2j}F_2 + \ldots + a_{mj}F_m$ หรือ
.Principal Component analysis	✓	✓	✓	$Zj = a_{1j}F_1 + a_{1j}F_2 + \ldots + a_{mj}F_m + d_jU_j$
.Principal factoring analysis	✓		✓	amj = น้ำหนักของตัวประกอบที่ m สำหรับตัวแปรที่ j
.Rao's canonical factoring			✓	
.Alpha factoring			✓	
.Image factoring			✓	

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
.Maximum likelihood factoring	✓			หรือ = น้ำหนักของตัวประกอบร่วมที่ m สำหรับตัวแปรที่ n
.Kaiser's Secmd generation little jiffy	✓			Fm = ตัวประกอบที่ m
(8) เมทริกซ์ตัวประกอบที่ได้จากการหมุนแกนตัวประกอบโดยวิธีต่อไปนี้				หรือ = ตัวประกอบร่วมที่ m
				Uj = ตัวประกอบเฉพาะสำหรับตัวแปรที่ j
8.1 หมุนแกนแบบ ORTOGONAL โดยวิธี . QUARTIMAX	✓	✓	✓ ✓	dj = สัมประสิทธิ์การถดถอยของตัวแปรที่ jบนตัวประกอบเฉพาะ
. VARIMAX	✓	✓	✓	
. EQUIMAX	✓		✓	
8.2 หมุนแบบ OBLIQUE โดยวิธี	✓	✓	✓	fik = Ci₁Z₁k + Ci₂Z₂k +...+ CinZnk
. DIRECT OBLIMIN	✓			
. DIRECT QUARTIMIN	✓		✓	fik = คะแนนตัวประกอบที่ i สำหรับค่าสังเกตที่ k (i = 1,2,..m) (k = 1,2,..n)
(9) กราฟของความสัมพันธ์ระหว่างน้ำหนักของแต่ละตัวประกอบที่ได้จากเมทริกซ์ตัวประกอบ	✓			
(10) เมทริกซ์คะแนนตัวประกอบซึ่งเป็นสัมประสิทธิ์ของแต่ละตัวแปรสำหรับคะแนนตัวประกอบใดๆ (Cin)	✓		✓	Cin = สัมประสิทธิ์คะแนนตัวประกอบที่ i สำหรับตัวแปรที่ n
(11) เมทริกซ์สหสัมพันธ์ของตัวประกอบ (Rf) หรือ เมทริกซ์ของความแปรปรวนร่วม (Sf)	✓		✓	Znk = คะแนนมาตรฐานของตัวแปรที่ n สำหรับค่าสังเกตที่ k
(12) ค่าระยะห่าง Mahalanobis ของค่าต่างๆดังนี้ .ของแต่ละค่าสังเกต ($R^{-1}Z_k/n$) .ของแต่ละคะแนนตัวประกอบ ($R_f^{-1}fk/m$) .ของผลต่างระหว่างทั้งสอง $(R^{-1}Zk - R^{-1}ffk)/(n-m)$	✓			Ci = ci1 ci2 ci3...cin R = เมทริกซ์สหสัมพันธ์ของตัวแปร Ri คอลัมน์ที่ i ของเมทริกซ์สหสัมพันธ์ระหว่างตัวแปรกับตัวประกอบ $Ci = R^{-1} Ri$ nx1 nxn nx1

ความสามารถของโปรแกรม	โปรแกรมสำเร็จรูป			ความหมาย
	BMDP	SAS	SPSS	
(13) ค่าคะแนนตัวประกอบแต่ละตัว-ประกอบทุกๆ ค่าสังเกตุ (fik)	✓		✓	
(14) กราฟของความสัมพันธ์ระหว่าง-คะแนนตัวประกอบแต่ละตัว	✓		✓	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 9.1 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม BMDP สำหรับ STAT 9 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
//TESTBMDP  JOB (9012,90,9087),'CHULA'
//  EXEC BIMED
//LKED.SYSIN DD *
    INCLUDE BMDP(BMDP4M)
    ENTRY MAIN
//GO.SYSIN DD *
/PROBLEM    TITLE IS 'FACTOR OF BMDP BY SIRICHAI'.
/INPUT      VARIABLES ARE 12.
            FORMAT IS '(12F2.0)'.
/VARIABLE   NAMES ARE CONCENTR,ANNOY,SMOKING1,SLEEPY,SMOKING2,TENSE,
            SMOKING3,ALERT,IRRITABL,TIRED,CONTENT,SMOKING4.
/END
```

บัตรบันทึกข้อมูล

```
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 9.1 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม BMDP สำหรับ STAT 9

```
WEIGHT VARIABLE . . . . . .
UNROTATED FACTORS ARE PRINCIPAL COMPONENTS.
NUMBER OF FACTORS IS LIMITED TO THE NUMBER OF EIGENVALUES GREATER THAN     1.000
TOLERANCE LIMIT FOR MATRIX INVERSION. . . . . .   0.00010
VARIMAX ROTATION IS PERFORMED.
GAMMA . . . . . .   1.0000
MAXIMUM NUMBER OF ITERATIONS FOR ROTATION . . .   50
CONVERGENCE CRITERION FOR ROTATION. . . . . .  0.0000100
KAISER'S NORMALIZATION. . . . . .   YES

DATA AFTER TRANSFORMATIONS FOR FIRST     5 CASES
CASES WITH ZERO WEIGHTS AND MISSING DATA NOT INCLUDED.

     CASE
LABEL    NUMBER          WEIGHT      1 CONCENTR     2 ANNOY       3 SMOKING1     4 SLEEPY      5 SMOKING2     6 TENSE       7 SMOKING3
                     8 ALERT        9 IRRITABL    10 TIRED       11 CONTENT     12 SMOKING4
                1     1.00000        3.00000        2.00000        1.00000        3.00000        2.00000        2.00000        1.00000
                      3.00000        2.00000        2.00000        3.00000        2.00000
                2     1.00000        4.00000        2.00000        5.00000        3.00000        5.00000        4.00000        5.00000
                      4.00000        3.00000        4.00000        4.00000        5.00000
                3     1.00000        5.00000        3.00000        4.00000        4.00000        5.00000        5.00000        4.00000
                      5.00000        3.00000        4.00000        3.00000        4.00000
                4     1.00000        4.00000        2.00000        4.00000        3.00000        5.00000        4.00000        4.00000
                      4.00000        3.00000        4.00000        3.00000        5.00000
                5     1.00000        4.00000        2.00000        4.00000        3.00000        4.00000        2.00000        4.00000
                      4.00000        2.00000        4.00000        3.00000        4.00000

 NUMBER OF CASES READ. . . . . .   110
```

(1) UNIVARIATE SUMMARY STATISTICS

VARIABLE	MEAN	STANDARD DEVIATION	COEFFICIENT OF VARIATION	SMALLEST VALUE	SMALLEST STANDARD SCORE	FIRST CASE FOR SMALLEST	LARGEST VALUE	LARGEST STANDARD SCORE	FIRST CASE FOR LARGEST
1 CONCENTR	2.69091	1.07298	0.398744	1.0000	-1.59	12	5.0000	2.15	3
2 ANNOY	2.11818	0.97427	0.459957	1.0000	-1.15	12	5.0000	2.96	11
3 SMOKING1	3.36364	1.13111	0.336275	1.0000	-2.09	1	5.0000	1.45	2
4 SLEEPY	2.60909	1.02353	0.392295	1.0000	-1.57	6	5.0000	2.34	60
5 SMOKING2	3.58182	1.06126	0.296291	1.0000	-2.43	11	5.0000	1.34	2
6 TENSE	2.44545	0.99158	0.405480	1.0000	-1.46	33	5.0000	2.58	3
7 SMOKING3	3.42727	1.16098	0.338747	1.0000	-2.09	1	5.0000	1.35	2
8 ALERT	2.80909	1.01814	0.362445	1.0000	-1.78	73	5.0000	2.15	3
9 IRRITABL	2.21818	0.78263	0.352825	1.0000	-1.56	12	4.0000	2.28	11
10 TIRED	3.09091	0.95346	0.308473	1.0000	-2.19	23	5.0000	2.00	11
11 CONTENT	2.45455	0.84198	0.343027	1.0000	-1.73	12	5.0000	3.02	11
12 SMOKING4	3.50000	1.27610	0.364601	1.0000	-1.96	11	5.0000	1.18	2

CASE NUMBERS ABOVE REFER TO DATA MATRIX BEFORE ANY CASES HAVE BEEN DELETED DUE TO MISSING DATA.
CASES WITH ZERO WEIGHTS ARE NOT INCLUDED.

(2) CORRELATION MATRIX

		CONCENTR 1	ANNOY 2	SMOKING1 3	SLEEPY 4	SMOKING2 5	TENSE 6	SMOKING3 7	ALERT 8	IRRITABL 9	TIRED 10	CONTENT 11	SMOKING4 12
CONCENTR	1	1.000											
ANNOY	2	0.562	1.000										
SMOKING1	3	0.086	0.144	1.000									
SLEEPY	4	0.457	0.360	0.140	1.000								
SMOKING2	5	0.200	0.119	0.785	0.211	1.000							
TENSE	6	0.579	0.705	0.222	0.273	0.301	1.000						
SMOKING3	7	0.041	0.060	0.810	0.126	0.816	0.120	1.000					
ALERT	8	0.802	0.578	0.101	0.606	0.223	0.594	0.039	1.000				
IRRITABL	9	0.595	0.796	0.189	0.337	0.221	0.725	0.108	0.605	1.000			
TIRED	10	0.512	0.413	0.199	0.798	0.274	0.364	0.139	0.698	0.428	1.000		
CONTENT	11	0.492	0.739	0.239	0.298	0.240	0.235	0.711	0.100	0.605	0.397	0.394	1.000
SMOKING4	12	0.228	0.122	0.775	0.277	0.813	0.214	0.845	0.201	0.156	0.271	0.171	1.000

(3) SQUARED MULTIPLE CORRELATIONS (SMC) OF EACH VARIABLE, WITH ALL OTHER VARIABLES

		SMC
1	CONCENTR	0.70351
2	ANNOY	0.74250
3	SMOKING1	0.73312
4	SLEEPY	0.68377
5	SMOKING2	0.78201
6	TENSE	0.66472
7	SMOKING3	0.82062
8	ALERT	0.80208
9	IRRITABL	0.71437
10	TIRED	0.72627
11	CONTENT	0.69130
12	SMOKING4	0.80294

COMMUNALITIES OBTAINED FROM 3 FACTORS AFTER 1 ITERATIONS.
THE COMMUNALITY OF A VARIABLE IS ITS SQUARED MULTIPLE CORRELATION (COVARIANCE) WITH THE FACTORS.

(4)

1	CONCENTR	0.6601
2	ANNOY	0.7956
3	SMOKING1	0.8391
4	SLEEPY	0.8474
5	SMOKING2	0.8561
6	TENSE	0.7804
7	SMOKING3	0.8941
8	ALERT	0.3258
9	IRRITABL	0.7978
10	TIRED	0.8453
11	CONTENT	0.7715
12	SMOKING4	0.8698

(6)

FACTOR	VARIANCE EXPLAINED	CUMULATIVE PROPORTION OF TOTAL VARIANCE
1	5.426	0.452
2	2.997	0.702
3	1.361	0.815
4	0.560	0.862
5	0.363	0.892
6	0.302	0.917
7	0.241	0.937
8	0.200	0.954
9	0.158	0.967
10	0.146	0.979
11	0.137	0.991
12	0.110	1.000

THE VARIANCE EXPLAINED BY EACH FACTOR IS THE EIGENVALUE FOR THAT FACTOR.

TOTAL VARIANCE IS DEFINED AS THE SUM OF THE DIAGONAL ELEMENTS OF THE CORRELATION (COVARIANCE) MATRIX.

(7) UNROTATED FACTOR LOADINGS (PATTERN)
FOR PRINCIPAL COMPONENTS

		FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	1	0.742	-0.309	0.167
ANNOY	2	0.755	-0.361	-0.309
SMOKING1	3	0.491	0.763	-0.124
SLEEPY	4	0.611	-0.117	0.679
SMOKING2	5	0.561	0.735	-0.030
TENSE	6	0.770	-0.232	-0.366
SMOKING3	7	0.417	0.847	-0.055
ALERT	8	0.808	-0.337	0.244
IRRITABL	9	0.783	-0.302	-0.306
TIRED	10	0.702	-0.139	0.577
CONTENT	11	0.748	-0.254	-0.382
SMOKING4	12	0.540	0.757	0.070
	VP	5.426	2.997	1.361

(8) THE VP FOR EACH FACTOR IS THE SUM OF THE SQUARES OF THE ELEMENTS OF THE COLUMN OF THE FACTOR LOADING MATRIX CORRESPONDING TO THAT FACTOR. THE VP IS THE VARIANCE EXPLAINED BY THE FACTOR.

ORTHOGONAL ROTATION. GAMMA = 1.0000

ITERATION	SIMPLICITY CRITERION
0	-1.900373
1	-6.017688
2	-6.019559
3	-6.019557

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(8) ROTATED FACTOR LOADINGS (PATTERN)

		FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	1	0.601	0.034	0.546
ANNOY	2	0.867	0.021	0.209
SMOKING1	3	0.131	0.907	0.007
SLEEPY	4	0.117	0.116	0.906
SMOKING2	5	0.141	0.905	0.128
TENSE	6	0.859	0.147	0.144
SMOKING3	7	0.005	0.945	0.010
ALERT	8	0.590	0.030	0.698
IRRITABL	9	0.863	0.085	0.214
TIRED	10	0.249	0.143	0.873
CONTENT	11	0.862	0.117	0.125
SMOKING4	12	0.061	0.910	0.195
	VP	3.802	3.442	2.538

THE VP FOR EACH FACTOR IS THE SUM OF THE SQUARES OF THE ELEMENTS OF THE COLUMN OF THE FACTOR PATTERN MATRIX CORRESPONDING TO THAT FACTOR. WHEN THE ROTATION IS ORTHOGONAL, THE VP IS THE VARIANCE EXPLAINED BY THE FACTOR.

(9)

ROTATED FACTOR LOADINGS

OVERLAP IS INDICATED BY A DOLLAR SIGN. SCALE IS FROM -1 TO +1.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

SORTED ROTATED FACTOR LOADINGS (PATTERN)

		FACTOR 1	FACTOR 2	FACTOR 3
ANNOY	2	0.867	0.0	0.0
IRRITABL	9	0.863	0.0	0.0
CONTENT	11	0.862	0.0	0.0
TENSE	6	0.859	0.0	0.0
CONCENTR	1	0.601	0.0	0.546
SMOKING3	7	0.0	0.945	0.0
SMOKING4	12	0.0	0.910	0.0
SMOKING1	3	0.0	0.907	0.0
SMOKING2	5	0.0	0.905	0.0
SLEEPY	4	0.0	0.0	0.906
TIRED	10	0.0	0.0	0.873
ALERT	8	0.590	0.0	0.691
	VP	3.802	3.443	2.538

THE ABOVE FACTOR LOADING MATRIX HAS BEEN REARRANGED SO THAT THE COLUMNS APPEAR IN DECREASING ORDER OF VARIANCE EXPLAINED BY FACTORS. THE ROWS HAVE BEEN REARRANGED SO THAT FOR EACH SUCCESSIVE FACTOR, LOADINGS GREATER THAN 0.5000 APPEAR FIRST. LOADINGS LESS THAN 0.2500 HAVE BEEN REPLACED BY ZERO.

(10) FACTOR SCORE COEFFICIENTS
THESE COEFFICIENTS ARE FOR THE STANDARDIZED VARIABLES, MEAN ZERO AND STANDARD DEVIATION ONE.

		FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	1	0.09301	-0.03935	0.16307
ANNOY	2	0.27541	-0.03385	-0.09242
SMOKING1	3	0.01728	0.27403	-0.07719
SLEEPY	4	-0.17902	-0.01809	0.48035
SMOKING2	5	-0.00784	0.25672	-0.01128
TENSE	6	0.28382	0.00873	-0.13400
SMOKING3	7	-0.03190	0.29001	-0.04754
ALERT	8	0.05389	-0.04844	0.24844
IRRITABL	9	0.27036	-0.01415	-0.09178
TIRED	10	-0.12227	-0.01241	0.42845
CONTENT	11	0.29053	0.00039	-0.14409
SMOKING4	12	-0.05440	0.26711	0.04592

(11) FACTOR SCORE COVARIANCE (COMPUTED FROM FACTOR STRUCTURE AND FACTOR SCORE COEFFICIENTS)

	FACTOR 1	FACTOR 2	FACTOR 3
FACTOR 1	1.000		
FACTOR 2	-0.000	1.000	
FACTOR 3	0.000	0.000	1.000

THE DIAGONAL OF THE ABOVE MATRIX CONTAINS THE SQUARED MULTIPLE CORRELATIONS OF EACH FACTOR WITH THE VARIABLES.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ESTIMATED FACTOR SCORES AND MAHALANOBIS DISTANCES (CHI-SQUARES) FROM EACH CASE TO THE CENTROID OF ALL CASES
FOR ORIGINAL DATA (12 D.F.) FACTOR SCORES (3 D.F.) AND THEIR DIFFERENCE (9 D.F.).
EACH CHI-SQUARE HAS BEEN DIVIDED BY ITS DEGREES OF FREEDOM.
CASE NUMBERS BELOW REFER TO DATA MATRIX BEFORE DELETION OF MISSING DATA.

(12) (13)

CASE LABEL	CASE NO.	CHISQ/DF 12	CHISQ/DF 3	CHISQ/DF 9	FACTOR 1	FACTOR 2	FACTOR 3
	1	1.432	1.212	1.505	0.167	-1.899	0.014
	2	1.113	1.063	1.130	1.114	1.361	0.389
	3	1.333	1.238	1.365	1.357	0.511	1.269
	4	0.715	0.593	0.755	0.781	0.848	0.670

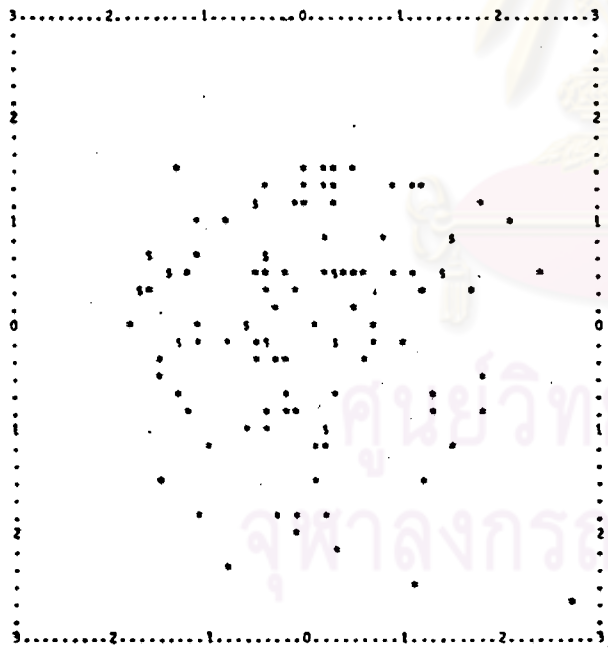
--- similar statistics for cases 5 to 105 ---

106	0.423	0.783	0.303	-0.388	0.692	-1.331
107	0.451	0.298	0.505	-0.288	0.106	-0.878
108	1.459	2.926	0.970	-1.067	-1.752	2.138
109	0.530	0.199	0.640	-0.337	-0.295	0.629
110	2.035	0.488	2.550	-0.066	-0.765	-0.935

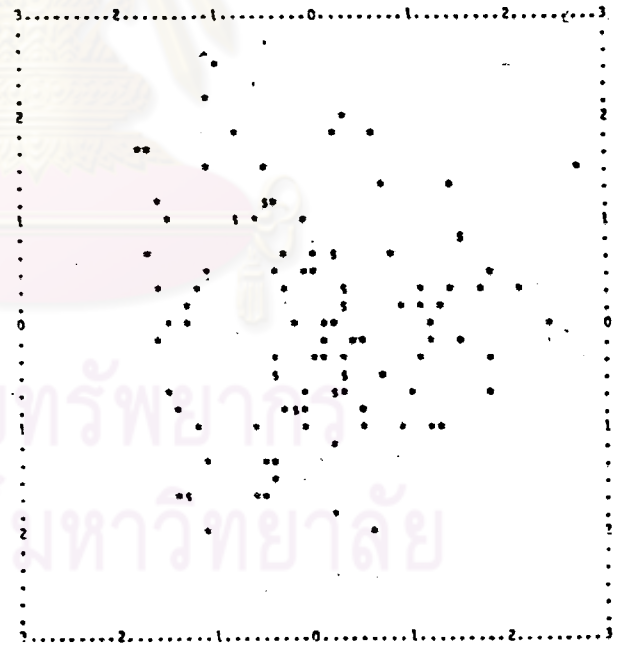
(14)

FACTOR SCORES 21

X-AXIS IS FACTOR 1, Y-AXIS IS FACTOR 2



X-AXIS IS FACTOR 1, Y-AXIS IS FACTOR 3



OVERLAP IS INDICATED BY A DOLLAR SIGN. SCALE IS FROM -3 TO +3.
FACTOR SCORES GREATER THAN 3 ARE PLOTTED AS 3. FACTOR SCORES LESS THAN -3 ARE PLOTTED AS -3.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 9.2 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SAS สำหรับ STAT 9 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
//S0396OP0K JOB CLASS=C,MSGLEVEL=(2,0)
// EXEC SAS
//SAS.SYSIN DD *
DATA SAMPLE4;
INPUT CONCENTR 2 ANNOY 4 SMOKING1 6 SLEEPY 8 SMOKING2 10 TENSE 12
 SMOKING3 14 ALERT 16 IRRITABL 18 TIRED 20 CONTENT 22 SMOKING4 24;
CARDS;
```

บัตรบันทึกข้อมูล

```
PROC FACTOR;
 TITLE 'FACTOR ANALYSIS OF SAS BY SIRICHAI';
/*
//
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 9.2 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SAS สำหรับ STAT 9

FACTOR ANALYSIS OF SAS BY SIRICHAI

VARIABLE	(1) MEAN	STANDARD DEVIATION
CONCENTR	2.09091	1.07290
ANNOY	2.11111	0.97427
SMOKING1	3.36364	1.13111
SLEEPY	2.00909	1.02355
SMOKING2	3.58182	1.06120
TENSE	2.44545	0.99158
SMOKING3	3.42727	1.16091
ALERT	2.09091	1.03814
IRRITABL	2.21818	0.78163
TIRED	3.09091	0.95346
CONTENT	2.45455	0.84198
SMOKING4	3.50000	1.27010

(2) CORRELATION MATRIX

	CONCENTR	ANNOY	SMOKING1	SLEEPY	SMOKING2	TENSE	SMOKING3	ALERT	IRRITABL	TIRED
CONCENTR	1.00000	0.56181	0.08590	0.45702	0.19966	0.57099	0.04071	0.62608	0.59452	0.51197
ANNOY	0.56181	1.00000	0.14360	0.35955	0.11422	0.70473	0.06039	0.57768	0.79608	0.41300
SMOKING1	0.08590	0.14360	1.00000	0.13970	0.76517	0.22234	0.80977	0.10067	0.10937	0.19675
SLEEPY	0.45702	0.35955	0.13970	1.00000	0.21130	0.27256	0.12041	0.60561	0.33651	0.79622
SMOKING2	0.19966	0.11422	0.76517	0.21130	1.00000	0.30070	0.81650	0.22261	0.22132	0.27365
TENSE	0.57099	0.70473	0.22234	0.27256	0.30070	1.00000	0.12005	0.59390	0.72479	0.36433
SMOKING3	0.04071	0.06039	0.80977	0.12041	0.81650	0.12005	1.00000	0.03660	0.10850	0.13663
ALERT	0.62608	0.57768	0.10067	0.60561	0.22261	0.59390	0.03660	1.00000	0.60540	0.69649
IRRITABL	0.59452	0.79608	0.10937	0.33651	0.22132	0.72479	0.10850	0.60540	1.00000	0.42867
TIRED	0.51197	0.41300	0.19675	0.79622	0.27365	0.36433	0.13663	0.69649	0.42867	1.00000
CONTENT	0.49306	0.73915	0.23968	0.24001	0.23121	0.71127	0.09963	0.60515	0.69729	0.39475
SMOKING4	0.22781	0.12170	0.77143	0.27745	0.81292	0.21369	0.84527	0.20124	0.12016	0.27145

	CONTENT	SMOKING4
CONCENTR	0.49306	0.22781
ANNOY	0.73915	0.12170
SMOKING1	0.23968	0.77143
SLEEPY	0.24001	0.27745
SMOKING2	0.23121	0.81292
TENSE	0.71127	0.21369
SMOKING3	0.09963	0.84527
ALERT	0.60515	0.20124
IRRITABL	0.69729	0.12016
TIRED	0.39475	0.27145
CONTENT	1.00000	0.17077
SMOKING4	0.17077	1.00000

CUMULATIVE PERCENTAGE OF EIGENVALUES (6)

0.45214 0.70184 0.81544

EIGEN VALUES

5.42568 3.45004 1.16052 0.96030 0.36326 0.30225 0.24080 0.19971 0.19618 0.14561
0.13074 0.11605

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(6) EIGEN VECTORS

	1	2	3
CONCENTR	0.31157	-0.17878	0.10648
ANNOY	0.32404	-0.26843	-0.26518
SMOKING1	0.21082	0.44690	-0.16016
SLEEPY	0.26419	-0.06730	0.58196
SMOKING2	0.24103	0.47451	-0.06258
TENSE	0.33040	-0.13397	-0.31401
SMOKING3	0.17905	0.48920	-0.04082
ALERT	0.34077	-0.19491	0.20906
IRRITABL	0.33600	-0.17453	-0.26675
TIRED	0.30156	-0.07902	0.49459
CONTENT	0.35124	-0.14607	-0.35749
SMOKING4	0.23166	0.43737	0.05981

FACTOR ANALYSIS OF SAS BY SPECIAL — DATA SAMPLE4 — 11:36 WEDNESDAY, JULY 29, 1981

(7) FACTOR MATRIX

	1	2	3
CONCENTR	0.74104	-0.30945	-0.11720
ANNOY	0.75478	-0.36001	-0.30931
SMOKING1	0.49110	0.76323	-0.12383
SLEEPY	0.61073	-0.11061	0.67886
SMOKING2	0.56143	0.73486	-0.02984
TENSE	0.76974	-0.23165	-0.36626
SMOKING3	0.41706	0.84485	-0.05401
ALERT	0.80773	-0.33741	0.24385
IRRITABL	0.78265	-0.30213	-0.30647
TIRED	0.70241	-0.13767	0.57669
CONTENT	0.83127	-0.25036	-0.38199
SMOKING4	0.53959	0.75747	0.06987

ITERATION CYCLE	VARIANCE
0	0.150364
1	0.501475
2	0.501629
3	0.501629
4	0.501629
5	0.501629
6	0.501629

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

(8) ROTATED FACTOR MATRIX

	1	2	3
CONCENTR	0.00013	0.03447	0.14196
ANNOY	0.86490	0.02088	0.20681
SMOKING1	0.13052	0.90664	0.00747
SLEEPY	0.11714	0.11560	0.90566
SMOKING2	0.14102	0.90549	0.12774
TENSE	0.65901	0.14701	0.14452
SMOKING3	0.00468	0.94549	0.00590
ALERT	0.56960	0.02966	0.69015
IRRITABL	0.86281	0.08155	0.11444
TIRED	0.24936	0.14264	0.87336
CONTENT	0.66157	0.11667	0.12454
SMOKING4	0.06684	0.91000	0.19459

FACTOR	VARIANCE	PERCENT
1	3.60122	30.0[illegible]
2	3.44283	28.7[illegible]
3	2.53078	21.1%

FACTOR ANALYSIS OF SAS BY SIRICHAI DATA SAMPLE4 11:56 WEDNESDAY, JULY 29, 1981

(4)

CHECK ON COMMUNALITIES

VARIABLE	ORIGINAL	FINAL	DIFFERENCE
CONCENTR	0.00014	0.00014	0.00000
ANNOY	0.79555	0.79555	0.00000
SMOKING1	0.83909	0.83909	0.00000
SLEEPY	0.84736	0.84736	0.00000
SMOKING2	0.85011	0.85011	0.00000
TENSE	0.76039	0.76039	0.00000
SMOKING3	0.89408	0.89408	0.00000
ALERT	0.82576	0.82576	0.00000
IRRITABL	0.79779	0.79779	0.00000
TIRED	0.84528	0.84528	0.00000
CONTENT	0.77155	0.77155	0.00000
SMOKING4	0.86977	0.86977	0.00000

ตารางที่ 9.3 แสดงตัวอย่างการเขียนคำสั่งของโปรแกรม SPSS สำหรับ STAT 9 โดยใช้ข้อมูลจาก DATA SAMPLE 4

```
* $$ JOB JNM=ACDBSPOK,CLASS=N
// JOB ACDBSPOK MR.SIRICHAI DONGWICHAI B221475
// EXEC PROC=$$SPSS79
RUN NAME 'COMPARE BMDP SAS SPSS STATISTICAL PACKAGE PROGRAMS'
FILE NAME 'DATA SAMPLE4'
VARIABLE LIST CONCENTR ANNOY SMOKING1 SLEEPY SMOKING2 TENSE
SMOKING3 ALERT IRRTABL TIRED CONTENT SMOKING4
INPUT MEDIUM CARD
N OF CASES 110
INPUT FORMAT FIXED(12F2.0)
READ INPUT DATA
```

บัตรบันทึกข้อมูล

```
TASK NAME 'FACTOR ANALYSIS OF SPSS BY SIRICHAI'
FACTOR VARIABLES = CONCENTR TO SMOKING4/
TYPE = PA1/
STATISTICS ALL
FINISH
/*
/&
* $$ EOJ
```

รูปที่ 9.3 แสดงผลที่ได้บางส่วนจากกระดาษต่อเนื่องของโปรแกรม SPSS สำหรับ STAT 9

VARIABLE (1)	MEAN	STANDARD DEV	CASES
CONCENTR	2.6939	1.0737	113
ANNOY	2.1152	0.9743	113
SMOKING1	2.2566	1.1351	113
SLEEPY	2.6991	1.0235	113
SMOKING2	3.6814	1.0613	113
TENSE	2.4425	0.9116	113
SMOKING3	3.4779	1.1610	113
ALERT	2.1991	1.0101	113
IRRITABL	2.1681	0.7836	113
TIRED	2.3009	0.9535	113
CONTENT	2.1549	0.8430	113
SMOKING4	3.5133	1.2761	113

CORRELATION COEFFICIENTS..

(2)

	CONCENTR	ANNOY	SMOKING1	SLEEPY	SMOKING2	TENSE	SMOKING3	ALERT	IRRITABL	TIRED
CONCENTR	1.00000	0.56103	0.30599	0.45722	0.19966	0.57399	0.04071	0.61268	0.59652	0.51157
ANNOY	0.56103	1.00000	0.14381	0.35956	0.11922	0.70473	0.06039	0.57768	0.74001	0.41131
SMOKING1	0.30599	0.14390	1.00000	0.13976	0.78512	0.22134	0.89977	0.19687	0.19875	0.23931
SLEEPY	0.45722	0.35956	0.13976	1.00000	0.21130	0.27258	0.12641	0.39361	0.33651	0.24401
SMOKING2	0.19966	0.11922	0.78512	0.21130	1.00000	0.30370	0.81650	0.24283	0.22155	0.22365
TENSE	0.57399	0.70473	0.22134	0.27258	0.30370	1.00000	0.12005	0.59359	0.72479	0.38633
SMOKING3	0.04071	0.06039	0.89977	0.12641	0.81650	0.12005	1.00000	0.03860	0.10851	0.09846
ALERT	0.61268	0.57768	0.19687	0.39361	0.24283	0.59359	0.03860	1.00000	0.61549	0.45806
IRRITABL	0.59652	0.74001	0.19875	0.33651	0.22155	0.72479	0.10850	0.61549	1.00000	0.43806
TIRED	0.51157	0.41131	0.23931	0.24401	0.22365	0.38633	0.09846	0.45806	0.43806	1.00000
CONTENT	0.49236	0.73915	0.23908	0.24001	0.22525	0.71126	0.09983	0.61515	0.69739	0.39375
SMOKING4	0.22781	0.12176	0.77543	0.27745	0.81292	0.21389	0.84527	0.20124	0.15616	0.27145

	CONTENT	SMOKING4
CONCENTR	0.49236	0.22781
ANNOY	0.73915	0.12176
SMOKING1	0.23931	0.77543
SLEEPY	0.24001	0.27745
SMOKING2	0.22525	0.81292
TENSE	0.71126	0.21389
SMOKING3	0.09983	0.84527
ALERT	0.61515	0.20124
IRRITABL	0.69739	0.15616
TIRED	0.39375	0.27145
CONTENT	1.00000	0.17077
SMOKING4	0.17077	1.00000

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

VARIABLE	EST COMMUNALITY	FACTOR	EIGENVALUE	PCT OF VAR	CUM PCT
CONCENTR	1.00000	1	5.42566	45.2	45.2
ANNOY	1.00000	2	2.95665	24.6	70.2
SMOKING1	1.00000	3	1.36051	11.3	81.5
SLEEPY	1.00000	4	0.56030	4.7	86.2
SMOKING2	1.00000	5	0.36226	3.0	89.2
TENSE	1.00000	6	0.30225	2.5	91.7
SMOKING3	1.00000	7	0.24080	2.0	93.7
ALERT	1.00000	8	0.19975	1.7	95.4
IRRITABL	1.00000	9	0.15816	1.3	96.7
TIRED	1.00000	10	0.14565	1.2	97.9
CONTENT	1.00000	11	0.13674	1.1	99.1
SMOKING4	1.00000	12	0.11023	0.9	100.0

FACTOR MATRIX USING PRINCIPAL FACTOR, NO ITERATIONS

	FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	0.74234	-0.30949	0.11729
ANNOY	0.75173	-0.36382	-0.30931
SMOKING1	0.49117	0.76333	-0.12383
SLEEPY	0.61173	-0.33693	0.67389
SMOKING2	0.56143	0.73437	-0.12984
TENSE	0.76974	-0.23183	-0.36636
SMOKING3	0.41727	0.84686	-0.15461
ALERT	0.80774	-0.32748	0.24385
IRRITABL	0.78265	-0.33014	-0.33647
TIRED	0.72248	-0.13733	0.57609
CONTENT	0.74326	-0.25630	-0.38199
SMOKING4	0.53959	0.75747	0.16783

VARIABLE	COMMUNALITY
CONCENTR	0.66015
ANNOY	0.77556
SMOKING1	0.93923
SLEEPY	0.34736
SMOKING2	0.85613
TENSE	0.73143
SMOKING3	0.89461
ALERT	0.83579
IRRITABL	0.77776
TIRED	0.84527
CONTENT	0.77153
SMOKING4	0.86071

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

VARIMAX ROTATED FACTOR MATRIX

(8)

	FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	0.61080	0.32444	0.54589
ANNOY	0.80612	0.02083	0.20870
SMOKING1	0.13057	0.90664	0.03746
SLEEPY	0.11726	0.11565	0.93567
SMOKING2	0.14139	0.90549	0.13772
TENSE	0.85934	0.14696	0.14442
SMOKING3	0.00504	0.94550	0.03989
ALERT	0.59969	0.32957	0.69178
IRRITABL	0.86235	0.20548	0.21434
TIRED	0.24967	0.16263	0.87338
CONTENT	0.86169	0.31662	0.13484
SMOKING4	0.16191	0.91937	0.19458

(11)

TRANSFORMATION MATRIX

	FACTOR 1	FACTOR 2	FACTOR 3
FACTOR 1	0.73403	0.63175	0.52417
FACTOR 2	-0.39136	0.09977	-0.19297
FACTOR 3	-0.55494	-0.76348	0.82946

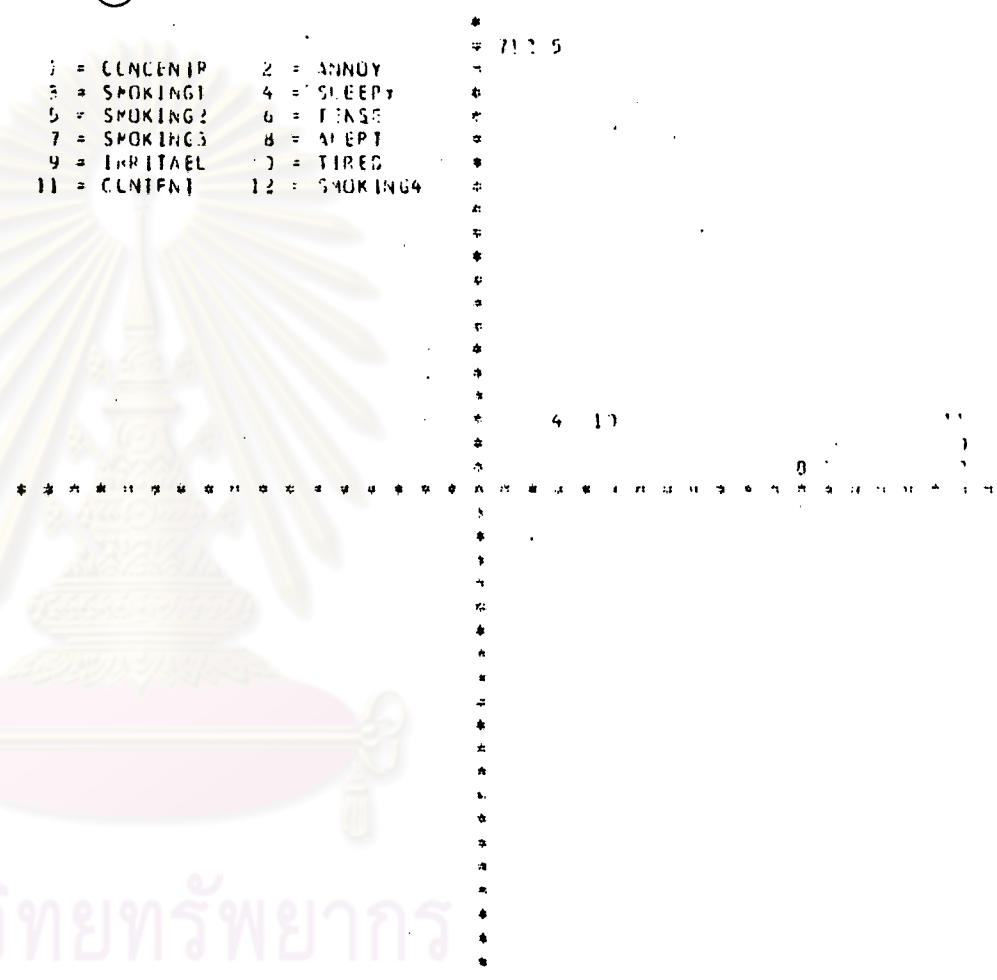
FACTOR SCORE COEFFICIENTS

(10)

	FACTOR 1	FACTOR 2	FACTOR 3
CONCENTR	0.09361	-0.03935	0.16307
ANNOY	0.27347	-0.13215	-0.09247
SMOKING1	0.01729	0.27432	-0.07723
SLEEPY	-0.17332	-0.17409	0.43036
SMOKING2	-0.03714	0.26671	-0.01128
TENSE	0.28381	0.10473	-0.13400
SMOKING3	-0.13103	0.29091	-0.04754
ALERT	0.05339	-0.04344	0.24844
IRRITABL	0.27136	-0.11414	-0.09178
TIRED	-0.13227	-0.01241	0.42945
CONTENT	0.29053	0.10039	-0.14409
SMOKING4	-0.03443	0.26711	-0.04592

(14) HORIZONTAL FACTOR 1 VERTICAL FACTOR 2

1 = CONCENTR 2 = ANNOY
3 = SMOKING1 4 = SLEEPY
5 = SMOKING2 6 = TENSE
7 = SMOKING3 8 = ALERT
9 = IRRITABL 10 = TIRED
11 = CONTENT 12 = SMOKING4



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย