บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาวิธีการจัดการข้อมูลสูญหายแบบอีพีเอสเอสอี ตรวจสอบความแม่นยำและอำนาจการทดสอบที่ได้จากวิธีการจัดการข้อมูลสูญหายแบบอีพีเอสเอสอี แบบอีเอ็ม และแบบลิสท์ไวส์ และวิเคราะห์ปฏิสัมพันธ์ระหว่างวิธีการสุ่มตัวอย่าง วิธีการจัดการข้อมูลสูญหาย จำนวนข้อมูลสูญหาย และความสัมพันธ์ระหว่างตัวแปร โดยแบ่งขั้นตอนการทำวิจัยออกเป็น 3 ขั้นตอน มีรายละเอียดการดำเนินการวิจัยดังนี้

ตอนที่ 1 การพัฒนาวิธีการจัดการข้อมูลสูญหายแบบอีพีเอสเอสอี

การพัฒนาวิธีการจัดการข้อมูลสูญหายแบบอีพีเอสเอสอี มีขั้นตอนดังนี้

1. ศึกษาเอกสาร แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับวิธีการจัดการข้อมูลสูญหาย

2. วิเคราะห์แนวคิดของวิธีการจัดการข้อมูลสูญหายแบบต่าง ๆ โดยพิจารณาจากจุดเด่น จุดด้อยของวิธีการเหล่านั้นเพื่อหาแนวคิดที่จะนำเสนอวิธีการจัดการข้อมูลสูญหายแบบใหม่

3. เสนอวิธีการจัดการข้อมูลสูญหายแบบใหม่

ตอนที่ 2 การตรวจสอบคุณภาพของวิธีการจัดการข้อมูลสูญหายในด้านความแม่นยำและอำนาจการทดสอบ

การตรวจสอบคุณภาพของวิธีการจัดการข้อมูลสูญหายที่พัฒนาขึ้น คือ วิธีการจัดการข้อมูลสูญหายแบบอีพีเอสเอสอี กับวิธีการจัดการข้อมูลสูญหายที่นิยมใช้กันมาก คือ วิธีการจัดการข้อมูลสูญหายแบบอีเอ็ม และวิธีการจัดการข้อมูลสูญหายที่ใช้เป็นฐานในการเปรียบเทียบ คือ การตัดข้อมูลสูญหายออกแบบลิสท์ไวส์ มีรายละเอียดดังต่อไปนี้

1. ลักษณะข้อมูล

ข้อมูลที่ใช้ในการวิจัยครั้งนี้ได้จากการจำลองสถานการณ์โดยใช้เทคนิคมอนติคาร์โล ซิมูเลชั่น (Monte carlo Simulation Technique) เป็นข้อมูลผลสัมฤทธิ์ทางการเรียนของนักเรียนประกอบด้วยเกรดเฉลี่ยเป็นตัวแปรเกณฑ์ (Y) และตัวแปรทำนาย (X) คือ คะแนนสอบของนักเรียน การกำหนดค่าเฉลี่ย ความแปรปรวน และจำนวนประชากรของตัวแปรดังกล่าวได้จากการศึกษางานวิจัยที่ผ่านมา ส่วนสัมประสิทธิ์สหสัมพันธ์จะกำหนดให้มีค่าเท่ากับ .30 .50 และ .70 ซึ่งเป็นลักษณะของความสัมพันธ์ต่ำ ปานกลาง และความสัมพันธ์สูง

จากการศึกษางานวิจัยเกี่ยวกับเกรดเฉลี่ยของนักเรียน เช่น งานวิจัยของจันทร์เพ็ญ วรสุวัส (2545) ได้ศึกษาปัจจัยที่ส่งผลต่อความต้องการในการฝึกอบรมอาชีพระยะสั้น (9+1) ของนักเรียนมัธยมศึกษาตอนต้น ในโรงเรียนขยายโอกาสทางการศึกษา สังกัดสำนักงานการประถมศึกษาจังหวัดเลย ตัวแปรที่ศึกษา คือ ผลสัมฤทธิ์ทางการเรียนของนักเรียนจากนักเรียนจำนวน 349 คน มีค่าเฉลี่ย 2.3733 ความแปรปรวน 0.4339 และอมรรัตน์ ทิพยจันทร์ (2534) ได้ศึกษาตัวแปรที่มีผลต่อค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ในเขตการศึกษา 6 โดยใช้วิธีการพหุระดับ จำนวนกลุ่มตัวอย่าง 918 คน เกรดเฉลี่ยของนักเรียนในชั้นมัธยมศึกษาปีที่ 2 มีค่าเท่ากับ 2.47 และความแปรปรวนเท่ากับ .4096

ดังนั้นในการวิจัยครั้งนี้จึงกำหนดค่าเฉลี่ยของเกรดเฉลี่ยเท่ากับ 2.3733 และความแปรปรวนเท่ากับ .4339 เพราะเกรดเฉลี่ยของนักเรียนน่าจะอยู่ในระดับปานกลางและการกระจายค่อนข้างสูง แต่เมื่อนำลักษณะข้อมูลที่ได้จากงานวิจัยที่ผ่านมาไปแทนค่าในสมการจำลองข้อมูลได้ค่าพิสัยของเกรดเฉลี่ยมากกว่า 4 ซึ่งเป็นไปไม่ได้ในทางปฏิบัติ ผู้วิจัยจึงได้กำหนดค่าเฉลี่ยของเกรดเฉลี่ยเท่ากับ 2 และความแปรปรวนเท่ากับ .15 เพื่อให้ข้อมูลเกรดเฉลี่ยมีพิสัยอยู่ระหว่าง 0-4 ซึ่งเป็นลักษณะของเกรดเฉลี่ยโดยทั่วไป

ข้อมูลคะแนนสอบของนักเรียนจากการศึกษางานวิจัยที่ผ่านมา มีค่าเฉลี่ยอยู่ระหว่าง 46.8150 – 66.6177 ความแปรปรวนมีค่าอยู่ระหว่าง 166.3620 – 266.7056 เช่น งานวิจัยของ สมชัย วงษ์นายะ (2533. หน้า 37) ได้ศึกษาค่าประมาณพารามิเตอร์จากแบบแผนการสุ่มตัวอย่างต่างแบบ ข้อมูลที่ใช้เป็นคะแนนผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ ระดับชั้นประถมศึกษาปีที่ 6 จำนวนประชากร 7,335 คน มีค่าเฉลี่ย 46.8150 ความแปรปรวน 166.3620 ดวงใจ ปวีณอภิชาต (2535. หน้า 63) ได้ศึกษาเปรียบเทียบค่าประมาณพารามิเตอร์ของแบบแผนการสุ่มตัวอย่างแบบแบ่งชั้นที่มีตัวแปรจำแนกชั้นภูมิและการกำหนดขนาดของกลุ่มตัวอย่างย่อยที่แตกต่างกัน : กรณีศึกษาผลสัมฤทธิ์ทางการเรียน ข้อมูลที่ใช้เป็นคะแนนผลสัมฤทธิ์ทาง

การเรียนคณิตศาสตร์ จากจำนวนประชากร 9,227 คน มีค่าเฉลี่ย 58.2257 ความแปรปรวนเท่ากับ 194.8977 และงานวิจัยของ สุกัญญรัตน์ คงงาม (2539. หน้า 41-42) ได้ศึกษาคุณสมบัติของตัวประมาณค่าพารามิเตอร์ที่ได้จากการสุ่มตัวอย่างแบบหลายขั้นตอน ระหว่างวิธีสุ่มแบบง่ายกับมีระบบ ใช้ข้อมูลผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ ภาคการศึกษาปลาย ของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2536 จำนวนประชากร 7,298 คน ค่าเฉลี่ย 68.6177 ความแปรปรวนเท่ากับ 266.7056

ดังนั้นในการวิจัยครั้งนี้จึงกำหนดค่าเฉลี่ยคะแนนสอบเท่ากับ 58.2257 ความแปรปรวน เท่ากับ 194.8977 ซึ่งเป็นคะแนนสอบที่มีค่าเฉลี่ยอยู่ในระดับปานกลาง สอดคล้องกับธรรมชาติคะแนนสอบของนักเรียนโดยทั่วไปที่ค่าเฉลี่ยไม่ค่อยสูงมากนัก แต่เมื่อนำลักษณะข้อมูลที่ได้จากงานวิจัยที่ผ่านมาไปแทนค่าในสมการจำลองข้อมูลพิสัยของคะแนนมีค่ามากกว่า 100 คะแนน ซึ่งเป็นไปไม่ได้ในทางปฏิบัติ ผู้วิจัยจึงได้กำหนดค่าเฉลี่ยของคะแนนสอบเท่ากับ 50 และความแปรปรวน เท่ากับ 100 เพื่อให้ข้อมูลคะแนนสอบของนักเรียนมีพิสัยอยู่ระหว่างคะแนน 0 - 100

2. ขนาดของกลุ่มตัวอย่าง

จากการศึกษางานวิจัยทั้งในประเทศและต่างประเทศเกี่ยวกับข้อมูลสูญหายมีการศึกษากับกลุ่มตัวอย่างขนาดเล็กและกลุ่มตัวอย่างขนาดใหญ่ ในกลุ่มตัวอย่างขนาดเล็กใช้กลุ่มตัวอย่างจำนวน 20 คน ส่วนกลุ่มตัวอย่างขนาดใหญ่ใช้กลุ่มตัวอย่างจำนวน 500 คน เป็นขนาดกลุ่มตัวอย่างจำนวนมากที่สุด (พรศิริ หมื่นไชยศรี, 2529 ; สุนันทา วีระกุลเทวัญ, 2544 ; ถวัลย์ จันทร์เพ็ง, 2531 ; วารุณี ตรีบำรุงศักดิ์, 2538 ; Kromrey & Hines, 1991 ; Gleason & Staelin, 1975 ; Enders, 2001) ในจำนวนงานวิจัยทั้งหมดที่ศึกษามีงานวิจัยเชิงสำรวจอยู่ 3 เรื่อง ใช้กลุ่มตัวอย่าง 357 , 331 และ 500 คน (Gross, 1990 ; Fred & Lii, 1998 ; Zhou, 2002) ส่วนใหญ่สอดคล้องกับจำนวนกลุ่มตัวอย่างที่นักวิจัยศึกษาเชิงสำรวจ ซึ่งส่วนมากผู้วิจัยจะใช้ตารางกำหนดขนาดกลุ่มตัวอย่าง เช่น ตารางของยามาเน่ (อุทุมพร จามรมาน, 2530. หน้า 30 อ้างอิงจาก Yamane, 1973) และตารางของ เครจซี่และมอร์แกน (บุญชม ศรีสะอาด, 2532. หน้า 39-40) ประชากรขนาด 10,000 คน จะได้กลุ่มตัวอย่างเท่ากับ 370 คน ส่วนการกำหนดขนาดกลุ่มตัวอย่างโดยใช้ตารางของ ศิริชัย กาญจนวาสี (ศิริชัย กาญจนวาสี และคณะ, 2540. หน้า 124-125) ณ ระดับความเชื่อมั่น 95% ยอมให้เกิดความคลาดเคลื่อนได้ 5% ของส่วนเบี่ยงเบนมาตรฐาน ซึ่งถ้าจะเปรียบเทียบกับความคลาดเคลื่อนทั้งหมดจะไม่เกิน 1% จะได้กลุ่มตัวอย่างเท่ากับ 1,379 คน แต่ถ้าพิจารณาโดยใช้สูตร (สมหวัง พิธิยานุวัฒน์, 2525. หน้า 60 ;

อุทุมพร ทองอุไทย, 2532. หน้า 32-33) โดยกำหนดความคลาดเคลื่อนประมาณ 5% จะได้กลุ่มตัวอย่างจำนวน 667 คน ทั้งสองสูตรได้จำนวนกลุ่มตัวอย่างใกล้เคียงกัน

การกำหนดขนาดกลุ่มตัวอย่างในการวิจัยเชิงสำรวจจะต้องให้มีความเหมาะสมที่ผู้วิจัยสามารถดำเนินการวิจัยได้ซึ่งไม่ควรมากหรือน้อยเกินไป ถ้าน้อยเกินไปการสรุปอ้างอิงไปยังประชากรก็ไม่ถูกต้อง แต่ถ้ามากเกินไปผู้วิจัยก็ไม่สามารถทำได้เพราะต้องลงทุนสูง ดังนั้นเพื่อให้สามารถนำผลการวิจัยไปใช้ได้สอดคล้องกับสภาพความเป็นจริง ผู้วิจัยจึงกำหนดขนาดกลุ่มตัวอย่างเท่ากับ 350 คน ซึ่งสอดคล้องกับงานวิจัยเชิงสำรวจที่ศึกษาข้อมูลสูญหายในต่างประเทศ และสอดคล้องกับการกำหนดขนาดกลุ่มตัวอย่างโดยใช้ตารางของยามาเน่และเครจซี่และมอร์แกน

การกำหนดขนาดกลุ่มตัวอย่างย่อยในแต่ละระดับชั้นหรือในแต่ละกลุ่มใช้สูตรของนีย์แมนซึ่ง เป็นวิธีการกำหนดขนาดกลุ่มตัวอย่างที่ถือว่าค่าใช้จ่ายต่อหน่วยในการสำรวจแต่ละระดับชั้นมีค่าใกล้เคียงกันหรือไม่แตกต่างกันอย่างเห็นได้ชัดเจนหรือเท่ากัน (สุกัญญรัตน์ คงงาม, 2539. หน้า 22-23) และจากการศึกษาของ ดวงใจ ปวีณอภิชาต (2535) พบว่า การกำหนดขนาดกลุ่มตัวอย่างย่อยแบบนีย์แมน เป็นวิธีการกำหนดขนาดกลุ่มตัวอย่างที่มีประสิทธิภาพมากที่สุดในการประมาณค่าเฉลี่ยเลขคณิตและความแปรปรวนของคะแนนผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของประชากรการกำหนดขนาดตัวอย่างของแต่ละระดับชั้น มีสูตรดังนี้

$$n_h = \frac{nN_hS_h}{\sum_{h=1}^{L} N_hS_h} \quad \text{โดยที่ } h=1,2,3,\ldots,L$$

เมื่อ n_h คือ ขนาดของกลุ่มตัวอย่างในระดับชั้นที่ h

S_h คือ ส่วนเบี่ยงเบนมาตรฐานของประชากรย่อยในระดับชั้นที่ h

L คือ จำนวนระดับชั้น

n คือ ขนาดของกลุ่มตัวอย่าง

3. วิธีดำเนินการ

สร้างข้อมูลจำลองโดยใช้เทคนิคมอนติ คาร์โล ซิมูเลชั่น กำหนดให้มีค่าเฉลี่ย ความแปรปรวน และจำนวนข้อมูลจากการศึกษางานวิจัยที่ผ่านมา ส่วนสัมประสิทธิ์สหสัมพันธ์จะกำหนดให้มีค่าเท่ากับ .30 .50 และ .70 ซึ่งเป็นลักษณะความสัมพันธ์ต่ำ ปานกลาง และสูง ตามลำดับ ขั้นตอนการสร้างกลุ่มประชากรที่มีการแจกแจงแบบปกติสองตัวแปร มีดังนี้

1. สร้างตัวเลขสุ่มให้มีค่า 0 ถึง 1 โดยใช้คำสั่ง RAND() จากโปรแกรม FOXPRO

2. สร้างข้อมูลมีการแจกแจงแบบปกติ 2 ค่า คือ Y1 และ Y2 มีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 โดยการประยุกต์โปรแกรมย่อยสับรูทีน NORMAL ที่มาร์สาเกลีย และเบรย์ เป็นผู้คิดขึ้น (ถวัลย์ จันทร์เพ็ง, 2531. หน้า 26) ผู้วิจัยเขียนโปรแกรมโดยใช้คำสั่งภาษา FOXPRO ได้ดังนี้

```
*** Y=GRADE,X=SCORE OF ACHIEVEMENT
CLEAR
CLOSE ALL
SET TALK OFF
SET SAFETY OFF
SET DECIMALS TO 6
N=1
S=1
R=.7
ROUND = 10000
CREATE TABLE DATARAN1 (ID N(10),Y1 N(10,6),Y2 N(10,6),X N(10,6),;
YX1 N(10,6),YX2 N(10,6), YX3 N(10,6), STA1 N(10,6), CUSTER1 N(10,6),;
STA2 N(10,6), CUSTER2 N(10,6),STA3 N(10,6), CUSTER3 N(10,6),NID N(10))
DELETE ALL
PACK
@ 19,8 SAY 'PROGRAM "CONSTRUCTING BIVARIATE NORMAL DISTRIBUTION"
IS  WORKING'
@ 20,28 SAY '1 ST POPULATION'
@ 21,30 SAY 'PROCESS TIME'+STR(N)
DO WHILE N <= ROUND
DO WHILE  S>=1
      RN=RAND()
      V1=2*RN - 1
      RN=RAND()
      V2=2*RN - 1
      S=V1*V1 + V2*V2
ENDDO
      RNN1 = V1*SQRT ((-2*LOG(S))/S)
      RNN2 = V2*SQRT ((-2*LOG(S))/S)
```

```
Y11=RNN1
Y22=RNN2
XX=50.0000+(SQRT(100.0000)*Y11)
CRR=SQRT(.15)/SQRT(100.0000)
RR=CRR*R
YY=1.6127 + (RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
YY1=2.0000 +(RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
YY2=2.3873 +(RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
USE DATARAN1
APPEND BLANK
REPLACE ID WITH N
REPLACE Y1 WITH RNN1
REPLACE Y2 WITH RNN2
REPLACE X WITH XX
REPLACE Yx1 WITH YY1
@ 21,30 SAY 'PROCESS TIME'+STR(N)
N=N+1
S=1
ENDDO
```

3. สร้างกลุ่มประชากรมีการแจกแจงแบบปกติสองตัวแปร เกรดเฉลี่ย (Y) มีค่าเฉลี่ยเท่ากับ 2.00 ความแปรปรวนเท่ากับ .15 คะแนนผลสัมฤทธิ์ทางการเรียน (X) มีค่าเฉลี่ยเท่ากับ 50 ความแปรปรวนเท่ากับ 100 สัมประสิทธิ์สหสัมพันธ์มีค่าเท่ากับ .30 , .50 และ .70 โดยใช้สมการ ต่อไปนี้

$$X = \mu_1 + \sqrt{\sigma_1^2}\ Y1$$

$$Y = \mu_2 + (\rho\sigma_2 / \sigma_1)(X - \mu_1) + \sqrt{\sigma_2^2(1-\rho^2)}\ Y2$$

เมื่อ

μ_1, σ_1^2 เป็นค่าเฉลี่ยและความแปรปรวนของคะแนนผลสัมฤทธิ์ทางการเรียน (X)

μ_2, σ_2^2 เป็นค่าเฉลี่ยและความแปรปรวนของเกรดเฉลี่ย (Y)

ผู้วิจัยเขียนโปรแกรมได้ดังนี้

```
XX=50.0000+(SQRT(100.0000)*Y11)
```

```
CRR=SQRT(.15)/SQRT(100.0000)
RR=CRR*R
YY=1.6127 + (RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
YY1=2.0000 +(RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
YY2=2.3873 +(RR*(XX-50.0000)) + (SQRT(.15*(1-(R*R))) * Y22)
```

รายละเอียดของคำสั่งทั้งหมดอยู่ในโปรแกรมชื่อ PRANDOM1.PRG ในภาคผนวก ก

เนื่องจากในการวิจัยครั้งนี้ผู้วิจัยต้องการศึกษาว่าอำนาจการทดสอบแตกต่างกันด้วยหรือไม่ ดังนั้นจึงสร้างประชากรขึ้นมาอีก 2 กลุ่มโดยให้มีค่าเฉลี่ยของเกรดเฉลี่ยมากกว่าและน้อยกว่าค่าเฉลี่ยของเกรดเฉลี่ยจากประชากรในการจำลองครั้งแรกเท่ากับ 1σ และ -1σ

ดังนั้นกลุ่มประชากรที่จำลองขึ้นมาในแต่ละครั้งของการศึกษาจะมี 3 กลุ่ม คือ

1. กลุ่มประชากรที่มีค่าเฉลี่ยและความแปรปรวนเท่ากับค่าเฉลี่ยและความแปรปรวนที่ได้ศึกษาจากงานวิจัยที่ผ่านมา โดยผู้วิจัยกำหนดให้มีค่าเฉลี่ยของเกรดเฉลี่ยเท่ากับ 2.00 ความแปรปรวนเท่ากับ .15 และสัมประสิทธิ์สหสัมพันธ์เท่ากับ .30 .50 และ .70 กลุ่มประชากรกลุ่มนี้ใช้หาความแม่นยำและอำนาจของการทดสอบในการทดสอบความสัมพันธ์

2. กลุ่มประชากรที่สองมีความแปรปรวนเท่ากับความแปรปรวนของประชากรกลุ่มแรกและสัมประสิทธิ์สหสัมพันธ์ .30 .50 และ .70 แต่มีค่าเฉลี่ยของเกรดเฉลี่ยมากกว่ากลุ่มประชากรกลุ่มแรกเท่ากับ 1σ ใช้หาอำนาจของการทดสอบในการทดสอบที (t-test) และการทดสอบเอฟ (F-test) หรือการวิเคราะห์ความแปรปรวน

3. กลุ่มประชากรที่สามมีความแปรปรวนเท่ากับความแปรปรวนของประชากรกลุ่มแรกและสัมประสิทธิ์สหสัมพันธ์ .30 .50 และ .70 แต่มีค่าเฉลี่ยของเกรดเฉลี่ยน้อยกว่ากลุ่มประชากรกลุ่มแรกเท่ากับ -1σ ใช้หาอำนาจของการทดสอบในการทดสอบที (t-test) และการทดสอบเอฟ (F-test) หรือการวิเคราะห์ความแปรปรวน

เมื่อจำลองข้อมูลเกรดเฉลี่ยและผลสัมฤทธิ์ทางการเรียนของนักเรียนได้แล้ว กำหนดข้อมูลที่ได้จากการจำลองสถานการณ์ของแต่ละคนว่าเป็นกลุ่มใดและระดับชั้นประเภทใด มีขั้นตอนดังนี้

1. การกำหนดข้อมูลของแต่ละคนว่าเป็นกลุ่มใดทำโดยใช้หลักการว่าการแบ่งประชากรออกเป็นกลุ่มเป็นลักษณะของประชากรที่รวมกันอยู่เป็นกลุ่ม (cluster) โดยแต่ละกลุ่มประชากรประกอบด้วยสมาชิกในกลุ่มที่มีลักษณะข้อมูลหลากหลาย แต่เมื่อพิจารณาระหว่างกลุ่มประชากรจะประกอบด้วยสมาชิกโดยส่วนรวมมีลักษณะคล้ายคลึงกัน (Homogeneous) (ศิริชัย กาญจนวาสี, 2526) ในการวิจัยครั้งนี้กำหนดกลุ่มประชากร 14 กลุ่ม จากลักษณะดังกล่าวจึงดำเนินการกำหนดข้อมูลที่ได้จากการจำลองสถานการณ์เป็นกลุ่มดังต่อไปนี้

1.1 สุ่มตัวเลขขึ้นมาแบบสุ่มให้มีค่าอยู่ระหว่าง 1-14

1.2 สมาชิกในประชากรคนแรกจะถูกกำหนดให้อยู่ในกลุ่มตามหมายเลขที่สุ่มได้ ในข้อที่ 1

1.3 ทำขั้นตอนที่ 1 และ 2 ไปเรื่อย ๆ จนครบประชากรทุกคน

2. การกำหนดข้อมูลของแต่ละคนว่าเป็นระดับชั้นประเภทใดทำโดยใช้หลักการว่าการจัดข้อมูลให้มีลักษณะเป็นระดับชั้นมีลักษณะการจำแนกประชากรออกเป็นพวกหรือชั้น (Stratum) โดยให้ประชากรในชั้นเดียวกันมีลักษณะคล้ายกัน (Homogeneous) แต่ประชากรที่ต่างชั้นกันมีความแตกต่างกัน (Heterogeneous) ในการวิจัยครั้งนี้กำหนดระดับชั้น 3 ระดับชั้น จากลักษณะดังกล่าว จึงดำเนินการกำหนดข้อมูลที่ได้จากการจำลองสถานการณ์เป็นระดับชั้นดังนี้

2.1 แบ่งเกรดเฉลี่ยของนักเรียนออกเป็น 3 กลุ่ม คือ กลุ่มสูง กลุ่มปานกลาง และกลุ่มต่ำ กลุ่มสูงมีเกรดเฉลี่ยมากกว่าหรือเท่ากับ $\bar{x}+1$S.D. กลุ่มต่ำมีเกรดเฉลี่ยน้อยกว่าหรือเท่ากับ $\bar{x}-1$S.D. ส่วนที่อยู่ตรงกลางเป็นกลุ่มปานกลาง

2.2 กลุ่มสูงกำหนดให้เป็นระดับชั้นที่ 3 กลุ่มปานกลางกำหนดให้เป็นระดับชั้นที่ 2 กลุ่มต่ำกำหนดให้เป็นระดับชั้นที่ 1

ดำเนินการวิเคราะห์ข้อมูลที่ได้จากการจำลองสถานการณ์ดังต่อไปนี้

1. หาค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ โดยใช้ข้อมูลจากประชากรกลุ่มแรก

2. สุ่มตัวอย่างจากประชากรกลุ่มแรกด้วยวิธีการสุ่มแบบแบ่งชั้นจากลักษณะข้อมูลที่กำหนด ลักษณะข้อมูลในการสุ่มครั้งแรกจะมีความสัมพันธ์ต่ำ คือ มีสัมประสิทธิ์สหสัมพันธ์เท่ากับ .30 โดยใช้คอมพิวเตอร์สุ่มตามจำนวนกลุ่มตัวอย่างที่กำหนด เขียนคำสั่งคอมพิวเตอร์ลบข้อมูลเกรดเฉลี่ยเป็นการสร้างข้อมูลสูญหาย จำนวนข้อมูลสูญหายเท่ากับ 5% ลักษณะของข้อมูลสูญหายในการวิจัยครั้งนี้จะกำหนดเป็นข้อมูลสูญหายแบบสุ่ม (randomly

missing data) หมายถึง ค่าของข้อมูลสูญหายที่เกิดขึ้นไม่ได้ขึ้นอยู่กับคะแนนของข้อมูลค่าใดค่าหนึ่ง หรือไม่ได้สูญหายเฉพาะคะแนนที่สูงหรือต่ำเท่านั้น

3. จัดกระทำข้อมูลสูญหายโดยการตัดข้อมูลสูญหายออกแบบลิสท์ไวส์

4. คำนวณหาค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์

5. คำนวณความแม่นยำโดยพิจารณาจากความแตกต่างของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ ที่ได้จากกลุ่มตัวอย่างกับค่าที่คำนวณจากประชากร

6. คำนวณอำนาจการทดสอบ มีขั้นตอนดังนี้

6.1 คำนวณอำนาจการทดสอบจากการทดสอบความสัมพันธ์ใช้ข้อมูลที่ได้จากการสุ่มตัวอย่างจากประชากรกลุ่มแรกโดยหาความสัมพันธ์ระหว่างตัวแปรเกรดเฉลี่ยและคะแนนผลสัมฤทธิ์ทางการเรียน แล้วทำการทดสอบว่าตัวแปรทั้งสองตัวมีความสัมพันธ์กันหรือไม่ เนื่องจากค่าสัมประสิทธิ์สหสัมพันธ์ของประชากรมีค่าไม่เท่ากับศูนย์ ดังนั้นการปฏิเสธสมมุติฐานสูญในการทดสอบนี้ก็คืออำนาจการทดสอบ การทดสอบสมมุติฐานของการวิจัยครั้งนี้กำหนดระดับนัยสำคัญเท่ากับ .05 ($\alpha = .05$) เพราะว่า การกำหนดระดับนัยสำคัญนั้นควรพิจารณาถึงการนำผลการทดสอบไปใช้หลังการทดลองหรือผลสืบเนื่องในทางปฏิบัติที่จะตามมา วัตถุประสงค์ของการวิจัย อำนาจการทดสอบ ระดับการควบคุมตัวแปรต่าง ๆ ในแบบแผนการทดลอง ความถูกต้องของผลการทดสอบ และการทดสอบแบบ 1 หาง หรือ 2 หาง (สุชาดา บวรกิติวงศ์, 2541. หน้า 17-19) แนวทางที่ผู้วิจัยควรให้ความสำคัญมากที่สุด คือ การนำผลการทดสอบไปใช้หลังการทดลองหรือผลสืบเนื่องในทางปฏิบัติที่จะตามมา ซึ่งเกี่ยวข้องกับการควบคุมความคลาดเคลื่อนประเภทที่ 1 และประเภทที่ 2 และมีผลต่ออำนาจของการทดสอบ ในการวิจัยครั้งนี้ทำการทดสอบสมมุติฐานเกี่ยวกับค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ เช่น ถ้าทดสอบสมมุติฐานโดยใช้การทดสอบที (t-test) $H_0 : \mu_1 = \mu_2 ; H_1 : \mu_1 \# \mu_2$ ผู้วิจัยควรควบคุมความคลาดเคลื่อนประเภทที่ 1 (α) เพราะถ้าปฏิเสธสมมุติฐานสูญทั้ง ๆ ที่ความเป็นจริงค่าเฉลี่ยเท่ากันนั้นมีความเสี่ยงมากกว่าการยอมรับสมมุติฐานสูญแต่ในความเป็นจริงแล้วค่าเฉลี่ยไม่เท่ากัน แต่ถ้าค่า α มีค่าน้อย ค่า β จะมากขึ้นทำให้อำนาจของการทดสอบน้อย ดังนั้นผู้วิจัยจึงกำหนดระดับนัยสำคัญในการทดสอบสมมุติฐานให้เท่ากับ .05 เพื่อให้อำนาจการทดสอบสูงขึ้น

6.2 เปลี่ยนกลุ่มประชากรเป็นกลุ่มประชากรที่สอง ที่มีค่าเฉลี่ยของเกรดเฉลี่ยมากกว่ากลุ่มประชากรแรก 1σ

6.3 ดำเนินการตามขั้นตอนที่ 2 และ 3 คือ สุ่มตัวอย่างแบบแบ่งชั้น สร้างข้อมูลสูญหายจำนวน 5% และจัดกระทำข้อมูลสูญหายโดยการตัดข้อมูลสูญหายออกแบบลิสท์ไวส์

6.4 คำนวณอำนาจการทดสอบจากการทดสอบที (t-test) โดยใช้ข้อมูลที่ได้จากข้อ 2 , 3 และข้อ 6.3 จากประชากรแรกและประชากรที่สองมาเปรียบเทียบกัน เนื่องจากประชากรกลุ่มแรกและประชากรกลุ่มที่สองมีค่าเฉลี่ยไม่เท่ากัน ดังนั้นการปฏิเสธสมมุติฐานสูญในการทดสอบนี้ก็คืออำนาจการทดสอบ

6.5 เปลี่ยนกลุ่มประชากรเป็นกลุ่มประชากรที่สาม ที่มีค่าเฉลี่ยของเกรดเฉลี่ยน้อยกว่ากลุ่มประชากรแรก -1σ

6.6 ดำเนินการตามขั้นตอนที่ 2 และ 3 คือ สุ่มตัวอย่างแบบแบ่งชั้น สร้างข้อมูลสูญหายจำนวน 5% และจัดกระทำข้อมูลสูญหายโดยการตัดข้อมูลสูญหายออกแบบลิสท์ไวส์

6.7 คำนวณอำนาจการทดสอบจากการทดสอบที (t-test) โดยใช้ข้อมูลที่ได้จากข้อ 2,3 และข้อ 6.6 จากประชากรแรกและประชากรที่สามมาเปรียบเทียบกัน เนื่องจากประชากรแรกและประชากรกลุ่มที่สามมีค่าเฉลี่ยไม่เท่ากัน ดังนั้นการปฏิเสธสมมุติฐานสูญในการทดสอบนี้ก็คืออำนาจการทดสอบ และคำนวณอำนาจการทดสอบจากการทดสอบเอฟ (F-test) หรือการวิเคราะห์ความแปรปรวน โดยใช้ข้อมูลที่ได้จากข้อ 2,3,6.3 และ 6.6 จากประชากรแรก ประชากรที่สอง และประชากรที่สามมาเปรียบเทียบกัน เนื่องจากประชากรทั้งสามกลุ่มมีค่าเฉลี่ยไม่เท่ากัน ดังนั้นการปฏิเสธสมมุติฐานสูญในการทดสอบนี้ก็คืออำนาจการทดสอบ

7. ดำเนินการตามขั้นตอนที่ 2 ถึงขั้นตอนที่ 6 เป็นจำนวน 1,000 ครั้ง แล้วคำนวณหาค่าสถิติต่อไปนี้

7.1 ค่าเฉลี่ยกำลังสองของความแม่นยำของค่าเฉลี่ยเลขคณิต

7.2 ค่าเฉลี่ยกำลังสองของความแม่นยำของค่าความแปรปรวน

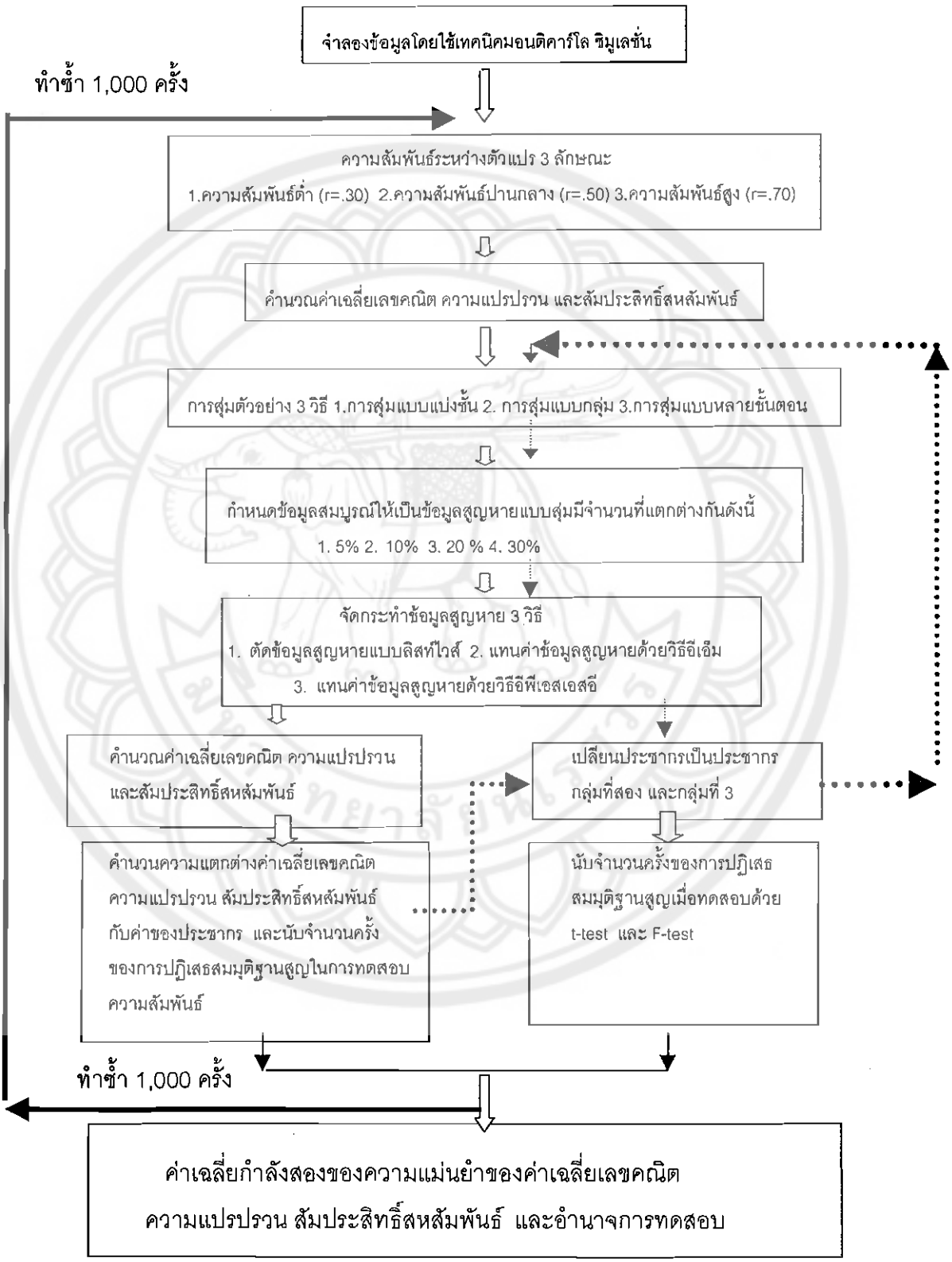
7.3 ค่าเฉลี่ยกำลังสองของความแม่นยำของค่าสัมประสิทธิ์สหสัมพันธ์

7.4 คำนวณอำนาจการทดสอบโดยพิจารณาจากจำนวนครั้งของการปฏิเสธสมมุติฐานสูญของการทดสอบที (t-test) การทดสอบเอฟ (F-test) และการทดสอบสหสัมพันธ์กับจำนวนครั้งของการทดสอบทั้งหมด สัดส่วนที่ได้ก็คือ อำนาจของการทดสอบ

ดำเนินการในลักษณะเดียวกันจากขั้นตอนที่ 2 ถึงขั้นตอนที่ 7 โดยกำหนดความสัมพันธ์ระหว่างตัวแปร วิธีการสุ่ม จำนวนข้อมูลสูญหาย และวิธีการจัดการข้อมูลสูญหาย ดังต่อไปนี้

1. ความสัมพันธ์ระหว่างตัวแปร
   1.1. ข้อมูลมีความสัมพันธ์ปานกลาง .50
   1.2 ข้อมูลมีความสัมพันธ์สูง .70
2. วิธีการสุ่ม
   2.1 การสุ่มแบบกลุ่ม
   2.2 การสุ่มแบบหลายขั้นตอน
3. จำนวนข้อมูลสูญหาย
   3.1 10% 3.2 20% 3.3 30%
4. วิธีการจัดการข้อมูลสูญหาย
   4.1 การแทนค่าข้อมูลด้วยวิธีอีเอ็ม
   4.2 การแทนค่าข้อมูลด้วยวิธีอีพีเอสเอสอี

การดำเนินการวิจัยสรุปเป็นแผนภาพได้ดังนี้



ภาพ 7 แผนภาพขั้นตอนการดำเนินการวิจัย

4. **เกณฑ์ในการเสนอผลการเปรียบเทียบ**

เกณฑ์ในการเสนอผลการเปรียบเทียบว่าวิธีการจัดการข้อมูลสูญหาย วิธีการสุ่มตัวอย่าง ความสัมพันธ์ระหว่างตัวแปร และจำนวนข้อมูลสูญหายที่แตกต่างกัน วิธีใดดีที่สุดจะใช้ค่าเฉลี่ยกำลังสองของความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้จากวิธีการจัดการข้อมูลสูญหาย วิธีการสุ่มตัวอย่าง ความสัมพันธ์ระหว่างตัวแปร และจำนวนข้อมูลสูญหายที่แตกต่างกันกับค่าพารามิเตอร์ มีค่าน้อยที่สุด ถ้าดัชนีดังกล่าวของ 2 วิธีมีค่าเท่ากัน วิธีที่ให้อำนาจการทดสอบสูงกว่าจะถือว่าเป็นวิธีที่ดีที่สุด

การศึกษาค่าประมาณพารามิเตอร์โดยใช้เทคนิคมอนติ คาร์โล ผู้วิจัยสามารถทำได้ 2 วิธี คือ 1). ศึกษาโดยใช้สถิติทดสอบว่ามีนัยสำคัญหรือไม่ อาจจะเป็นการทดสอบความแตกต่าง หรือทดสอบความสัมพันธ์ 2). ศึกษาโดยนำค่าที่วิเคราะห์ได้มาเปรียบเทียบกันโดยถือว่าค่าที่ได้แทนค่าพารามิเตอร์ตัวหนึ่ง ในทางปฏิบัติการอ้างอิงเกี่ยวกับค่าพารามิเตอร์สามารถสรุปได้จากค่าสถิติจากกลุ่มตัวอย่างโดยการทำซ้ำหลายร้อยค่าและอาจจะให้การอ้างอิงดีกว่าการตอบคำถามเชิงวิเคราะห์แบบเก่าโดยการประมาณค่าหรือการทดสอบสมมุติฐาน (Rubin, 1991. p. 245)

จากการศึกษางานวิจัยที่เกี่ยวข้องถ้าผู้วิจัยใช้สถิติทดสอบจำนวนครั้งในการทำซ้ำมีค่าตั้งแต่ 40, 100, 250 (ต่าย เชี่ยงฉี, 2534 ; สำราญ มีแจ้ง, 2534 ; อรุณี อ่อนสวัสดิ์, 2537 ; Enders, 2001) ซึ่งเป็นจำนวนครั้งค่อนข้างน้อย แต่ถ้าใช้วิธีที่ 2 คือ ศึกษาโดยการทำซ้ำแล้วนำค่าที่ ได้มาเปรียบเทียบกันวิธีนี้จำนวนครั้งในการทำซ้ำค่อนข้างสูง เช่น 100, 200,400, 1000, 4000 (Graham et al, 1996 ; วารุณี ตรีบำรุงศักดิ์, 2538 ; Gross, 1990 ; Little et al, 1998 ; Fred and Lii, 1998 ; ดวงใจ ปวีณอภิชาต, 2535 ; สุกัญญรัตน์ คงงาม, 2539 ; สมชัย วงษ์นายะ, 2533 ; Graham and Schafer, 1999 ; Brockmeier et al, 1998 ; ถวัลย์ จันทร์เพ็ง, 2531 ; Kromrey and Hines, 1994 ; Raaijmakers, 1999) ถึงแม้ว่างานวิจัยของถวัลย์ จันทร์เพ็ง ใช้การทำซ้ำถึงจำนวน 4,000 ครั้ง ก็น่าจะเป็นเพราะว่ากลุ่มตัวอย่างมีขนาดเล็กทำให้การทำซ้ำจำนวนมากครั้งเป็นไปได้สูง โดยส่วนใหญ่จะทำซ้ำจำนวน 1,000 ครั้ง ซึ่งสอดคล้องกับจำนวนการทำซ้ำ 1,000 ครั้ง เพื่อศึกษาประมาณค่าพารามิเตอร์ในการวิจัยครั้งนี้ และเพื่อให้เกิดความมั่นใจได้ว่าค่าเฉลี่ยที่นำมาเปรียบเทียบกันนั้นมีความแตกต่างกันจริง จึงได้ทำการทดสอบความแตกต่างของค่าเฉลี่ยความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์จำแนกตามวิธีการสุ่มตัวอย่าง จำนวนข้อมูลสูญหาย ความสัมพันธ์ระหว่างตัวแปรและวิธีการจัดการข้อมูลสูญหาย โดยการวิเคราะห์ความแปรปรวนแบบทางเดียว (One Way Anova) เมื่อทดสอบแล้วพบว่ามีนัยสำคัญทางสถิติได้วิเคราะห์เปรียบเทียบค่าเฉลี่ยเป็นรายคู่โดยใช้วิธีของทูกี้ (Tukey)

ตอนที่ 3 การวิเคราะห์ปฏิสัมพันธ์ระหว่างวิธีการสุ่มตัวอย่าง วิธีการจัดการข้อมูลสูญหาย จำนวนข้อมูลสูญหาย และความสัมพันธ์ระหว่างตัวแปร

การวิเคราะห์ปฏิสัมพันธ์ระหว่างวิธีการสุ่มตัวอย่าง วิธีการจัดการข้อมูสูญหาย จำนวนข้อมูลสูญหาย และความสัมพันธ์ระหว่างตัวแปร ที่มีต่อความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ เฟรดและลี (Fred & Lii, 1998) และคอร์ทินา (Fred & Lii, 1998. unpaged citing Cortina, 1993. pp. 915-922) ให้ข้อสังเกตว่าผลของการศึกษาการจัดการข้อมูลสูญหายจะปะปนกันในสภาพการที่มีปฏิสัมพันธ์กัน คำกล่าวของนักวิจัยทั้งสองท่านสอดคล้องกับงานวิจัยของ เอ็นเดอร์ (Enders, 2001. pp. 713-740) ที่ได้ศึกษาพบว่าวิธีการจัดการข้อมูลสูญหาย ระดับความสัมพันธ์ระหว่างตัวแปร และจำนวนข้อมูลสูญหายมีปฏิสัมพันธ์กัน แต่ตัวแปรในการศึกษาครั้งนี้มีตัวแปรวิธีการสุ่มตัวอย่างเพิ่มขึ้นมาอีกตัวแปรหนึ่ง ถึงแม้ว่ายังไม่มีผู้ใดศึกษาว่าวิธีการสุ่มตัวอย่างมีปฏิสัมพันธ์กับวิธีการจัดการข้อมูลสูญหาย จำนวนข้อมูลสูญหาย และระดับความสัมพันธ์ระหว่างตัวแปรหรือไม่ แต่เพื่อให้ได้ข้อค้นพบที่ชัดเจนและสามารถนำผลการวิจัยไปใช้ได้อย่างกว้างขวาง ผู้วิจัยจึงได้ศึกษาปฏิสัมพันธ์ระหว่างวิธีการสุ่มตัวอย่าง วิธีการจัดการข้อมูลสูญหาย ความสัมพันธ์ระหว่างตัวแปร และจำนวนข้อมูลสูญหายที่มีต่อความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ ประกอบกับตัวแปรตามความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ น่าจะมีความเกี่ยวข้องกันเพราะว่าการคำนวณความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์จะนำค่าเฉลี่ยเลขคณิตของตัวแปรที่ศึกษามาคำนวณด้วย ถ้าใช้เทคนิคการวิเคราะห์ความแปรปรวนทำการวิเคราะห์โดยใช้ตัวแปรตามทีละตัวลักษณะดังกล่าวทำให้การวิเคราะห์ความแปรปรวนที่ทดสอบแต่ละครั้งเป็นอิสระจากกันไม่เกี่ยวข้อง กันซึ่งผิดไปจากความเป็นจริงที่วัดตัวแปรตามเหล่านั้นจากสมาชิกกลุ่มตัวอย่างกลุ่มเดียวกันค่าที่วัดย่อมมีความสัมพันธ์กันในทางใดทางหนึ่ง ดังนั้นจึงวิเคราะห์ข้อมูลในการศึกษาครั้งนี้ด้วยการวิเคราะห์ความแปรปรวนพหุคูณ 4 ทาง (Four way multivariate analysis of variance : FOUR WAY MANOVA) การวิเคราะห์ความแปรปรวนพหุคูณจะนำเอาค่าสหสัมพันธ์ระหว่างตัวแปรตามมาพิจารณาด้วย (สำราญ มีแจ้ง, 2543. หน้า 121) ซึ่งสอดคล้องกับลักษณะของตัวแปรตามที่นำมาศึกษาในครั้งนี้ การวิเคราะห์ข้อมูล ผู้วิจัยกำหนดขนาดของกลุ่มตัวอย่างในแต่ละเซลเท่ากับ 100 ดังตาราง

ตาราง 4 ลักษณะข้อมูลในการวิเคราะห์ความแปรปรวนพหุคูณสี่ทาง

			M1	M2	M3
DA1	S1	NM1			
		NM2			
		NM3			
		NM4			
	S2	NM1			
		NM2			
		NM3			
		NM4			
	S3	NM1			
		NM2			
		NM3			
		NM4			
DA2	S1	NM1			
		NM2			
		NM3			
		NM4			
	S2	NM1			
		NM2			
		NM3			
		NM4			
	S3	NM1			
		NM2			
		NM3			
		NM4			
DA3	S1	NM1			
		NM2			
		NM3			
		NM4			
	S2	NM1			
		NM2			
		NM3			
		NM4			
	S3	NM1			
		NM2			
		NM3			
		NM4			

โดยกำหนดสัญลักษณ์ในตารางดังต่อไปนี้

วิธีการสุ่ม 3 วิธี คือ

1. การสุ่มแบบแบ่งชั้น (S1)

2. การสุ่มแบบกลุ่ม (S2)

3. การสุ่มแบบหลายขั้นตอน (S3)

วิธีการจัดการข้อมูลสูญหาย 3 วิธี คือ

1. การตัดข้อมูลออกแบบลิสท์ไวส์ (Listwise deletion) (M1)

2. การแทนค่าข้อมูลด้วยวิธีอีเอ็ม (EM algorithm or Expectation maximization) (M2)

3. การแทนค่าข้อมูลด้วยวิธีอีพีเอสเอสอี (Estimated parameter and Smallest standard error) (M3)

ความสัมพันธ์ระหว่างตัวแปร 3 ระดับ คือ

1. ข้อมูลที่มีความสัมพันธ์ต่ำ (r=.30) (DA1)

2. ข้อมูลที่มีความสัมพันธ์ปานกลาง (r=.50) (DA2)

3. ข้อมูลที่มีความสัมพันธ์สูง (r=.70) (DA3)

จำนวนข้อมูลสูญหาย

1. 5% (NM1)

2. 10% (NM2)

3. 20% (NM3)

4. 30% (NM4)

ขั้นตอนการวิเคราะห์ความแปรปรวนพหุคูณ 4 ทาง มีดังต่อไปนี้

1. หาค่าสหสัมพันธ์ภายในกลุ่มของตัวแปรตามที่กำหนดโดยหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายตามวิธีของเพียร์สัน (Pearson Product Moment Correlation) และวิเคราะห์โดยวิธีของ Bartlett 's Test of Sphericity เพื่อหาความสัมพันธ์รวมของตัวแปรตาม ความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ การกระทำดังกล่าวเพื่อตรวจสอบข้อตกลงเบื้องต้น (Assumption) ข้อแรกของการวิเคราะห์ความแปรปรวนพหุคูณ (MANOVA)

2. วิเคราะห์สถิติ F-test โดยวิธีการของ Bartlett เพื่อทดสอบความเป็นเอกพันธ์ของความแปรปรวนของข้อมูล (Homogeneity) ตามข้อตกลงข้อที่สองของการวิเคราะห์ความแปรปรวนพหุคูณ (MANOVA)

3. วิเคราะห์ความแปรปรวนพหุคูณ (MANOVA) ใช้สถิติแลมบ์ดา $(\wedge)$ ของวิลคส์ (Wilks) (Tatsuoka, 1971. p. 198 ; Lindeman & others, 1980. p.232) เป็นค่าแตกต่างระหว่างเซ็นทรอยด์ (Centroids) ของกลุ่มต่าง ๆ สถิติแลมบ์ดาเป็นสถิติที่นิยมใช้ในการทดสอบนัยสำคัญของ MANOVA ในกรณีที่การวิเคราะห์มีตัวแปรจำแนกกลุ่มจำนวน 3 กลุ่มขึ้นไป (ปัญญา ธีระวิทยเลิศ, 2534. หน้า 68) การหาค่าสถิติแลมบ์ดา มีสูตรดังนี้ (สำเริง บุญเรืองรัตน์, 2526. หน้า 141-142)

$$\wedge = \frac{|W|}{|T|}$$

เมื่อ $\underset{\sim}{T}$ คือ Total SSCP Matrix ของทุกกลุ่ม

$\underset{\sim}{B}$ คือ SSCP Matrix ของระหว่างกลุ่ม

$\underset{\sim}{W}$ คือ SSCP Matrix ของภายในกลุ่ม

ตามสมการ $\underset{\sim}{T} = \underset{\sim}{B} + \underset{\sim}{W}$ (สำเริง บุญเรืองรัตน์, 2526. หน้า 116)

ในการวิเคราะห์ความแปรปรวนพหุคูณ 2 ทาง หรือมากกว่านั้นมีตัวแปรอิสระหลายตัว ดังนั้น สูตรที่ใช้คำนวณ $\wedge$ จะเป็นดังนี้

$$\wedge_h = \frac{\left|\underset{\sim e}{S}\right|}{\left|\underset{\sim h}{S} + \underset{\sim e}{S}\right|}$$

เมื่อ $\underset{\sim e}{S}$ คือ เมตริกซ์ของ SSCP ของ error สำหรับแต่ละแบบการทดลอง

$\underset{\sim h}{S}$ คือ เมตริกซ์ของ SSCP effect ที่ต้องการทดสอบสมมุติฐาน

4. การทดสอบความมีนัยสำคัญของสถิติ $\wedge$ ใช้สถิติ F โดยการเปลี่ยนค่า $\wedge$ เป็นค่า F โดยวิธีการของ Rao ดังนี้

$$R = \frac{1 - \wedge^{1/s}}{\wedge^{1/s}} \cdot \frac{ms - pv_h / 2 + 1}{pv_h}$$

เมื่อ V_e คือ จำนวน degree of freedom ของ error

V_h คือ จำนวน degree of freedom ของ effect หรือ
จำนวน degree of freedom ของตัวแปรอิสระที่ต้องการ
ทดสอบนัยสำคัญ

P คือ จำนวนตัวแปรตาม

เมื่อผลการทดสอบมีนัยสำคัญทั้ง Main effect และ Interaction effect จะทำการวิเคราะห์ความแตกต่างระหว่างกลุ่ม (Discriminant analysis) เพื่อดูว่า Main effect และ Interaction effect นั้นแตกต่างกันที่ตัวแปรใด มีขั้นตอนการวิเคราะห์ดังนี้

1. หาค่า λ จากสมการ $\left|S_w^{-1}S_h - \lambda I\right| = 0$

2. ทดสอบความมีนัยสำคัญของ λ โดยใช้สถิติทดสอบของ Bartlett 's V statistic ดังสูตร

$$V = \left[V_h + V_e - (p + v_h + 1)/2\right]\ln(1+\lambda)$$

สถิติทดสอบของ Bartlett 's V statistic มีการแจกแจงเป็นไค-สแควร์ degree of freedom เท่ากับ $(p-1)(v_h + 1 - 2)$

3. สร้างสมการจำแนกในรูปคะแนนดิบ โดยการแทนค่า λ ในสมการ $S_w^{-1}S_h - \lambda I$ แล้วคำนวนหาค่าของเมตริกซ์นี้ หาโคแฟกเตอร์ (cofactor) ของแถวแรกหรือแถวที่สองของค่าของเมตริกซ์ $S_w^{-1}S_h - \lambda I$ แล้วทำเป็น normalized ก็จะได้ฟังก์ชันการจำแนก (discriminant function) จากค่า λ ถ้าค่า λ มีนัยสำคัญมากกว่าหนึ่งค่า จะมีฟังก์ชันการจำแนกมากกว่าหนึ่งสมการ

4. สร้างสมการจำแนกในรูปคะแนนมาตรฐานโดยการแปลงฟังก์ชันการจำแนกในรูปคะแนนดิบให้อยู่ในรูปคะแนนมาตรฐาน

การแปลความหมายฟังก์ชั่นการจำแนก (Discriminant Function)

1. จากสมการจำแนกในรูปคะแนนดิบนำค่าเฉลี่ยของตัวแปรตามความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ในกลุ่มย่อยของ Main effect หรือ Interaction effect มาแทนค่าจะได้ค่าเฉลี่ยของสมการจำแนกในกลุ่มย่อยดังกล่าว ความแตกต่างที่ปรากฏจะอธิบายปฏิสัมพันธ์ที่มีนัยสำคัญ

2. จากสมการจำแนกในรูปคะแนนมาตรฐานสามารถตัดสินได้ว่าค่าเฉลี่ยของสมการจำแนกกลุ่มย่อยที่มีค่าต่ำเกิดจากตัวแปรใด มีน้ำหนักต่ำและสูงประมาณกี่เท่า

ในการวิจัยครั้งนี้มุ่งศึกษาประเด็นหลักที่สำคัญคือ ปฏิสัมพันธ์สี่ทาง (Four-Way Interaction) เพราะว่าในการทำวิจัยแต่ละครั้งนั้นเมื่อมีข้อมูลสูญหายเกิดขึ้น มีวิธีการจัดการข้อมูลสูญหายหลายวิธี และมีตัวแปรที่สำคัญอื่น ๆ เข้ามาเกี่ยวข้องเสมอ คือ ความสัมพันธ์ระหว่างตัวแปร วิธีการสุ่มตัวอย่าง และจำนวนของข้อมูลสูญหาย พร้อม ๆ กันทั้ง 4 ตัวแปร

การวิเคราะห์ข้อมูลเพื่อให้สามารถตอบวัตถุประสงค์ข้อที่ 2 ผู้วิจัยจะนำค่าเฉลี่ยกำลังสองของความแม่นยำของค่าเฉลี่ยเลขคณิต ค่าเฉลี่ยกำลังสองของความแม่นยำของความแปรปรวน และค่าเฉลี่ยกำลังสองของความแม่นยำของสัมประสิทธิ์สหสัมพันธ์ที่ได้จากการทำซ้ำ 1,000 ครั้ง มาเปรียบเทียบกันโดยที่ไม่ทำการทดสอบความแตกต่างของค่าเฉลี่ยภายหลังเพราะการทำซ้ำจำนวนมากครั้งสามารถประมาณค่าพารามิเตอร์ได้

ในกรณีที่ข้อมูลฝ่าฝืนข้อตกลงเบื้องต้น ผู้วิจัยจะใช้การวิเคราะห์ความแปรปรวนสี่ทาง (Four way ANOVA) ของตัวแปรตามความแม่นยำของค่าเฉลี่ยเลขคณิต ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ ทีละตัวแปร