

ขนาดตัวอย่างที่เหมาะสมสำหรับการทดสอบสัมประสิทธิ์สหสัมพันธ์  
ของข้อมูลที่มีการแจกแจงแบบไขว้วาริ เอทน์อร์มอล



นางสาวกรรณิภา เลียง เจริญสิทธิ์



ศูนย์วิทยทรัพยากร  
วิทยานิพนธ์นี้ เป็นส่วนหนึ่งของการศึกษาคามหลักสูตรปริญญาสาทิติศาสตรมหาบัณฑิต  
จุฬาลงกรณ์มหาวิทยาลัย  
ภาควิชาสถิติ

บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย


พ.ศ. 2527

ISBN 974 - 563 - 486 - 7

010234

I15045699

DETERMINATION OF SAMPLE SIZE FOR TESTING  
CORRELATION COEFFICIENT OF BIVARIATE NORMAL DATA



Miss Kannika Liengcharoensit

A Thesis Submitted in Partial Fulfilment of the Requirements  
for the Degree of Master of Science

Department of Statistics

Graduate School

Chulalongkorn University


1984

หัวข้อวิทยานิพนธ์ ขนาดตัวอย่างที่เหมาะสมสำหรับการทดสอบสัมประสิทธิ์สหสัมพันธ์ของข้อมูลที่มีการแจกแจงแบบไม่วารี เอทนอนร์มอล


โดย นางสาวกรรณิกา เลียง เจริญสิทธิ์  
ภาควิชา สถิติ  
อาจารย์ที่ปรึกษา อาจารย์ ดร.อนุชิต ฉ้ายอดมรรคผล

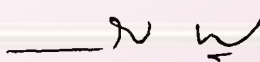



บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้บัณฑิตวิทยาลัยนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาโทบัณฑิต


  
.....คณบดีบัณฑิตวิทยาลัย  
(รองศาสตราจารย์ ดร.สุประดิษฐ์ ชุนนาค)

คณะกรรมการสอบวิทยานิพนธ์

  
.....ประธานกรรมการ  
(รองศาสตราจารย์ ส่องศรี ทิทยารัตน์)

  
.....กรรมการ  
(รองศาสตราจารย์ ดร.สรชัย พิศาลบุตร)

  
.....กรรมการ  
( อาจารย์ ดร.ธีระพร วีระถาวร )

  
.....กรรมการ  
( อาจารย์ ดร.อนุชิต ฉ้ายอดมรรคผล )

หัวข้อวิทยานิพนธ์ ขนาดตัวอย่างที่เหมาะสมสำหรับการทดสอบสมมติฐานสองทิศทางของข้อมูล  
ที่มีการแจกแจงแบบไบนารีเอทอร์มอล

ชื่อนิสิต นางสาว กรรณิกา เสียง เจริญสิทธิ์

อาจารย์ที่ปรึกษา อาจารย์ ดร. อนุชิต ล้ายอดมรรคผล

ภาควิชา สถิติ

ปีการศึกษา 2526



บทคัดย่อ

ปี ค.ศ. 1915 R.A. Fisher พบว่าลักษณะการแจกแจงของค่าสัมประสิทธิ์สหสัมพันธ์เชิงเส้นแบบเพียร์สัน ( $r$ ) เมื่อ  $\rho \neq 0$  ขึ้นอยู่กับขนาดตัวอย่าง ( $n$ ) และค่าสัมประสิทธิ์สหสัมพันธ์ของประชากร ( $\rho$ ) จึงทำการศึกษาเพิ่มและพบว่า การแปลงค่า  $r$  โดยใช้สูตร

$$Z_F = \frac{1}{2} \ln \left( \frac{1+r}{1-r} \right)$$

จะทำให้ลักษณะการแจกแจงของ  $Z_F$  เป็นแบบปกติโดยมีค่าเฉลี่ย

$$\mu_{Z_F} = \frac{1}{2} \ln \left( \frac{1+\rho}{1-\rho} \right) \quad \text{และส่วน เบี่ยงเบนมาตรฐาน}$$

$$\sigma_{Z_F} = \frac{1}{\sqrt{n-3}}$$

การแปลงค่านี้นำมาใช้อย่างกว้างขวางในการทดสอบสมมติฐาน

$$1. H_0 : \rho = \rho_0 \quad (\rho_0 \neq 0)$$

$$H_1 : \rho \neq \rho_0$$

$$2. H_0 : \rho \geq \rho_0 \quad (\rho_0 \neq 0)$$

$$H_1 : \rho < \rho_0$$

$$3. H_0 : \rho \leq \rho_0 \quad (\rho_0 \neq 0)$$

$$H_1 : \rho > \rho_0$$

เนื่องจากการสรุปของ R.A. Fisher ที่ว่าเมื่อ  $n$  มีขนาดใหญ่พอจะทำให้ผลสรุปถูกต้อง การศึกษาเรื่องนี้มีวัตถุประสงค์เพื่อศึกษาความเหมาะสมของขนาดตัวอย่างโดยอาศัยวิธีการซีมู เลชัน เพื่อหาข้อสรุปที่ชัดเจนยิ่งขึ้น ในการศึกษาจำเป็นต้องใช้วิธีการสร้างตัวแปรสุ่มแบบปกติสองตัวแปรสำหรับค่า  $\rho$  ตั้งแต่  $\rho = 0.1, 0.2, \dots, 0.9$  การศึกษานี้แสดงการทดสอบในกรณีค่า  $\rho$  เป็นค่าบวก ในกรณีที่ค่า  $\rho$  เป็นค่าลบสามารถพิสูจน์และสรุปได้ในทำนองเดียวกัน

ผลจากการศึกษาได้สร้างตารางตัวแปรสุ่มแบบปกติสองตัวแปร ณ ระดับค่า  $\rho$  ต่าง ๆ ตั้งแต่  $\rho = 0.1, 0.2, \dots, 0.9$  เพื่อนำไปใช้ประโยชน์กรณีที่ต้องการสุ่มตัวอย่างที่มีคุณสมบัติดังกล่าว

ผลการศึกษาลักษณะการแจกแจงของข้อมูลยืนยันลักษณะความเบ้ของการแจกแจงของ  $r$  กรณี  $\rho$  ไม่เท่ากับ 0 เมื่อ  $n$  มีค่าน้อยกว่า 25 แต่เมื่อ  $n$  มีค่าเท่ากับหรือมากกว่า 25 การแจกแจงของ  $r$  จะเป็นแบบปกติโดยประมาณ และยืนยันว่าเมื่อแปลงค่า  $r$  โดย Fisher's transformation แล้ว  $Z_F$  จะมีลักษณะการแจกแจงแบบปกติ โดยประมาณ

ข้อสรุปที่สำคัญที่ได้จากการศึกษาคือ ในการทดสอบสมมุติฐานกรณี  $\rho$  มีค่าอื่น ๆ ที่ไม่เท่ากับ 0 ณ ระดับ  $\alpha = 0.01$  ขนาดตัวอย่างที่เหมาะสม คือ  $n$  ควรมีค่าตั้งแต่ 9 ที่ระดับ  $\alpha = 0.05$  และที่ระดับ  $\alpha = 0.10$   $n$  ควรมีค่าตั้งแต่ 5 ขึ้นไป

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



R.A. Fisher stated that when  $n$  is large enough it will give a correct conclusion. How large is large is our problem because the sample size will be usually subjectively determined. The purpose of studying is to find the suitability of sample size by using simulation to find the exact conclusion. It is necessary to generate the bivariate normal data with parameter  $\rho = 0.1, 0.2, \dots, 0.9$ . In this thesis we only study the testing hypothesis when  $\rho$  is positive. For negative  $\rho$ , it can be proved and concluded in the same.

In studying the table of the bivariate normal data when value of  $\rho$  varies according to  $0.1, 0.2, \dots, 0.9$  is given, it will be useful for the purpose of random sampling to use a bivariate normal data by researcher who studies in a related topic.

From studying the distribution of data, we discover that when  $\rho$  is not equal to 0 and  $n$  is less than 25, the distribution of  $r$  is skewed but when  $n$  is at least 25, the distribution of  $r$  is approximately normal. By transforming  $r$  to be  $Z_F$ ,  $Z_F$  will distribute normally.

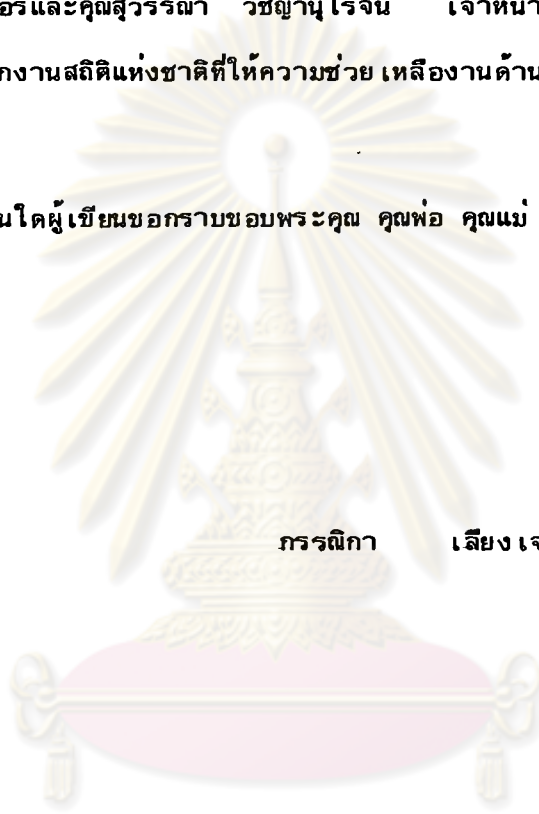
The main conclusion is that the suitability of sample size for testing hypothesis when  $\rho$  is not equal to 0 at  $\alpha = .01$ ,  $n$  will have a value of at least 9, at  $\alpha = .05$  and  $\alpha = .10$ ,  $n$  will have a value of at least 5.



กิติกรรมประกาศ

ผู้เขียนขอกราบขอบพระคุณ อาจารย์ ดร.อนุชิต ล้ายอดมรรคผล ที่กรุณาได้รับเป็นอาจารย์ที่ปรึกษาและให้คำแนะนำในการทำวิทยานิพนธ์ พร้อมทั้งตรวจและให้คำแนะนำในการแก้ไขจนวิทยานิพนธ์ฉบับนี้ เสร็จสมบูรณ์ และขอขอบคุณ คุณชนกิจ อีระกาญจน์ เจ้าหน้าที่งานพัฒนาระบบงานคอมพิวเตอร์และคุณสุวรรณา วิชญาโรจน์ เจ้าหน้าที่งานควบคุมฐานข้อมูลกองประมวลผลสถิติ สำนักงานสถิติแห่งชาติที่ให้ความช่วยเหลืองานด้านโปรแกรมคอมพิวเตอร์ ด้วยน้ำใจอันดียิ่ง

นอกเหนือสิ่งอื่นใดผู้เขียนขอกราบขอบพระคุณ คุณพ่อ คุณแม่ ที่คอย เป็นกำลังใจให้ตลอดเวลา.



กรรณิกา เลียง เจริญสิทธิ์

ศูนย์วิทยพัทยากร  
จุฬาลงกรณ์มหาวิทยาลัย





หน้า

บทคัดย่อภาษาไทย .....	ก
บทคัดย่อภาษาอังกฤษ .....	ข
กิตติกรรมประกาศ .....	ค
รายการตารางประกอบ .....	ข
รายการรูปประกอบ .....	ง
บทที่ 1 บทนำ .....	1
1.1 ความเป็นมาของปัญหา .....	1
1.2 วัตถุประสงค์ของการวิจัย .....	4
1.3 ขอบเขตของการวิจัย .....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ .....	5
1.5 คำสำคัญ .....	5
บทที่ 2 ทฤษฎีที่ใช้ในการวิจัย .....	6
2.1 การสร้างตารางข้อมูลที่มีการแจกแจงแบบปกติสองตัวแปรโดยวิธีการ - ซีมี เลชั่น .....	6
2.2 การแจกแจงแบบปกติสองตัวแปร .....	8
2.3 การแจกแจงแบบมีเงื่อนไข .....	9
2.4 การทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ ( $\rho$ ) = 0 .....	11
2.5 การทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ ( $\rho$ ) $\neq$ 0 .....	12
2.6 การทดสอบภาวะสภาวะปกติแบบ โทโม โกรอฟ .....	13

สารบัญ (ต่อ)

	หน้า
บทที่ 3 เครื่องมือที่ใช้ในการวิจัย .....	20
3.1 การผลิต เลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ .....	20
3.1.1 การผลิต เลขสุ่มโดยการโปรแกรม .....	20
3.1.2 การผลิต เลขสุ่ม โดย RND .....	21
3.2 การสุ่มตัวอย่าง .....	22
3.3 โปรแกรมที่ใช้ในงานทั้งหมด .....	24
บทที่ 4 ขั้นตอนการวิจัยและผลการวิจัย .....	25
4.1 การเตรียมข้อมูลสำหรับการวิจัย .....	25
4.2 การวิเคราะห์ลักษณะการแจกแจงของค่าสัมประสิทธิ์สหสัมพันธ์จากตัวอย่าง .....	26
4.3 การวิเคราะห์ลักษณะการแจกแจงของค่า Fisher's transformation .....	31
4.4 การหาขนาดตัวอย่างที่เหมาะสมในการทดสอบสัมประสิทธิ์สหสัมพันธ์ $(\rho) \neq 0$ .....	35
บทที่ 5 สรุปผลการวิจัย .....	48
5.1 สรุปผลการวิจัย .....	48
5.2 ข้อเสนอแนะ .....	49
บรรณานุกรม .....	51
ภาคผนวก .....	53
สัญลักษณ์ที่ใช้ .....	129
ประวัติผู้เขียน .....	133

รายการตารางประกอบ

ตารางที่	หน้า
4.2.1 แสดงผลสรุปการทดสอบการแจกแจงของค่า $x$ .....	29
4.3.1 แสดงผลสรุปการทดสอบการแจกแจงของค่า $Z_F$ .....	32
4.4.1 แสดงค่าสัดส่วนของการปฏิเสธสมมติฐาน $H_0$ ที่ระดับนัยสำคัญ 1 % .....	40
4.4.2 แสดงค่าสัดส่วนของการปฏิเสธสมมติฐาน $H_0$ ที่ระดับนัยสำคัญ 5 % .....	41
4.4.3 แสดงค่าสัดส่วนของการปฏิเสธสมมติฐาน $H_0$ ที่ระดับนัยสำคัญ 10 % .....	42
4.4.4 แสดงค่าเฉลี่ย, ส่วน เบี่ยงเบนมาตรฐานของสัดส่วนการปฏิเสธสมมติฐานและ ค่าคำนวณ $t$ ณ. ระดับนัยสำคัญ 1 % .....	44
4.4.5 แสดงค่าเฉลี่ย, ส่วน เบี่ยงเบนมาตรฐานของสัดส่วนการปฏิเสธสมมติฐานและ ค่าคำนวณ $t$ ณ. ระดับนัยสำคัญ 5 % .....	45
4.4.6 แสดงค่าเฉลี่ย, ส่วน เบี่ยงเบนมาตรฐานของสัดส่วนการปฏิเสธสมมติฐานและ ค่าคำนวณ $t$ ณ. ระดับนัยสำคัญ 10 % .....	46

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

รายการรูปประกอบ

รูปที่	หน้า
2.6.1 กราฟของ $F_0(x)$ , $S_n(x)$ และ $D_n$ .....	14
4.2.1 พังก์ชันการแจกแจงของค่า $x$	
เมื่อ $n = 10$ , $\rho = .0, .5, .8$ .....	30
4.2.2 พังก์ชันการแจกแจงของค่า $x$	
เมื่อ $n = 50$ , $\rho = .0, .5, .8$ .....	30
4.3.1 พังก์ชันการแจกแจงของค่า $Z_F$	
เมื่อ $n = 10$ , $\rho = .0, .5, .8$ .....	33
4.3.2 พังก์ชันการแจกแจงของค่า $Z_F$	
เมื่อ $n = 50$ , $\rho = .0, .5, .8$ .....	33
 แผนผังที่	
1 การสร้างตัวแปรสุ่มปกติสองตัวแปร .....	7
2 แสดงการเปรียบเทียบค่า $Z$ จากการคำนวณกับค่า $Z$ จากตารางและนับจำนวนการปฏิเสธสมมติฐาน .....	38

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย