ตำรา อนุมานสถิติ

(STATISTICAL INFERENCE)

โดย

อาจารย์ อรุณ อินทรปาลิต

ภาควิชาวิศวกรรมชลประทาน

คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเกษตรศาสตร์

ตำราอนุมานสถิติเล่มนี้ผู้สอนได้แปลและเรียบเรียงจากตำราภาษาอังกฤษชื่อ INTRODUCTION TO STATISTICAL INFERENCE ซึ่ง Dr. Jerome C.R.Li เป็นผู้เขียนไว้

Dr. Jerome C.R.Li เคยเป็น Chaiman, Department of Statistics, Oregon State University หลายปี และเป็นอาจารย์สอนวิชา Statistical Inference และ General Regression Analysis ของผู้สอนในขณะที่กำลังศึกษาอยู่ที่ Oregon State University, U.S.A. พ.ศ. 2503 – 2505

Dr. Jerome C.R.Li เป็นผู้เชี่ยวชาญวิชา statistics อย่างยอดเยี่ยม เป็นผู้วางรากฐาน statistics และก่อตั้ง Department of Statistics ขึ้นใน Oregon State University จนเป็นที่ยอมรับนับถือจากบุคคลทั่วไปว่าท่านคือบิดาแห่ง statistics ของ Oregon State University ตำราของท่านเล่มนี้เป็นตำราดีมีประโยชน์มากเล่มหนึ่งซึ่งใช้เป็นตำราเรียนของมหาวิทยาลัยหลายแห่งในสหรัฐอเมริกา

ตำราของ Dr. Jerome C.R.Li อ่านเข้าใจง่ายและไม่ใช้คณิตศาสตร์ชั้นสูงแต่อย่างใด ท่านเขียนตำราเล่มนี้ขึ้นเพื่อใช้สอนนักเรียนชั้น senior และปริญญาโทซึ่งมาจาก department ต่าง ๆ ของมหาวิทยาลัย นักเรียนเหล่านี้ย่อมมีพื้นฐานความรู้ทางคณิตศาสตร์ไม่เท่ากัน จึงจำเป็นต้องทำให้เป็นตำราที่ทุกคนพอจะศึกษาได้ อนึ่ง Dr. Jerome C.R.Li ได้เขียนตำราเล่มนี้จาก sampling experiments ชุดหนึ่งซึ่งใช้พิสูจน์ความเป็นจริงของ theorems ต่าง ๆ

ความจริงตำราอนุมานสถิติเล่มนี้มี 24 chapters และแบ่งสอนเป็น 3 courses ต่อเนื่องกันคือ St 421, 422 และ 423, Statistical Inference (G) แต่ผู้สอนได้แปลและเรียบเรียงไว้เพียง 13 chapters และใช้เป็นตำราวิชาสถิติ 421 อนุมานสถิติ I สำหรับสอนนิสิตคณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

ในการแปลและเรียบเรียงตำราอนุมานสถิติเล่มนี้ผู้สอนพยายามรักษาเนื้อความเดิมในภาษาอังกฤษเอาไว้แต่พยายามใช้ถ้อยคำสำนวนเป็นภาษาไทยให้มากที่สุด จึงไม่ใช่เป็นการแปลตามตัวภาษาอังกฤษ ทั้งนี้เพื่อให้ผู้อ่านเข้าใจง่ายขึ้น และเนื่องจากตำราเล่มนี้ได้จัดทำขึ้นเป็นครั้งที่ 2 (ครั้งที่ 1 จัดทำเมื่อเดือนสิงหาคม 2508 และมีเพียง 9 chapters) จึงมีการแก้ไขข้อความหลายแห่งและได้เพิ่ม chapter 10 ถึง 13 ขึ้นอีก 4 chapters ด้วย

ผู้สอนได้พบกับ Dr. Jerome C.R.Li ครั้งสุดท้ายที่เมือง Tai Chung, Taiwan เมื่อวันที่ 22 กุมภาพันธ์ 2506 ขณะที่ผู้สอนเดินทางกลับจากสหรัฐอเมริกาและแวะศึกษาและดูงานชลประทานในประเทศนั้น ท่านบอกผู้สอนว่ากำลังปรับปรุงตำราของท่านให้ดีขึ้นและจะเขียนตำรา Statistical Inference II ต่อไปเพื่อให้ตำราชุดนี้จบลงอย่างสมบูรณ์ (ตำรา Statistical Inference II ของ Dr. Jerome C.R.Li ได้พิมพ์ออกจำหน่ายแล้วใน ค.ศ. 1964) ท่านได้แสดงความยินดีเมื่อทราบว่าผู้สอนจะนำความรู้ที่ได้รับจากท่านมาสอนเพื่อประโยชน์สำหรับการศึกษาและการค้นคว้าทดลองในมหาวิทยาลัยเกษตรศาสตร์

ผู้สอนได้ทราบข่าวด้วยความสลดใจจากท่านผู้หนึ่งซึ่งเชื่อถือได้ว่า Dr. Jerome C.R.Li ได้ล่วงลับไปแล้วเมื่อไม่กี่ปีมานี้เอง ถ้าหากตำราอนุมานสถิติเล่มนี้จะมีความดีและเป็นประโยชน์บ้างแล้ว ขอให้ความดีนั้นจงเป็นของ Dr. Jerome C.R.Li ผู้ล่วงลับไปแล้วทั้งสิ้น ผู้สอนขอรับผิดในข้อบกพร่องต่าง ๆ ถ้าหากจะมีขึ้นในการแปลและเรียบเรียงแต่ผู้เดียว และขอได้โปรดเข้าใจว่าผู้สอนจัดทำตำราเล่มนี้ขึ้นเพื่อความสะดวกในการสอนและการเรียนของนิสิตและใช้เฉพาะภายในมหาวิทยาลัยเกษตรศาสตร์เท่านั้น ผู้สอนไม่มีเจตนาจะจัดทำขึ้นเพื่อประโยชน์อื่นใดทั้งสิ้น

19 ม.ค. 19

คณะวิศวกรรมศาสตร์

มหาวิทยาลัยเกษตรศาสตร์

สารบัญ

Statistical Inference

Chapter 1

Introduction

วิชาอนุมานสถิติ (statistical inference) เป็นวิชาหนึ่งซึ่งน่าสนใจศึกษามาก เพราะในปัจจุบันวิชาอนุมานสถิติมีบทบาทสำคัญในงานทั่วไป โดยเฉพาะอย่างยิ่งในการวิจัยและการค้นคว้าทดลองทางวิทยาศาสตร์ทุกสาขา เราทราบดีแล้วว่าการวิจัยและการค้นคว้าทดลองสำหรับการเกษตรนั้นเป็นสิ่งจำเป็นต้องทำอยู่ตลอดเวลาเพื่อพัฒนาวิชาการด้านนี้ให้เจริญก้าวหน้าไปอย่างไม่หยุดยั้ง และใช้ความรู้ที่ได้รับจากการวิจัยเข้าช่วยเหลือการผลิตทางเกษตรให้ได้ผลดียิ่งขึ้นและแน่นอนตลอดไปด้วย ในวงงานวิศวกรรมต่าง ๆ ก็จำเป็นต้องใช้วิชาอนุมานสถิติเช่นเดียวกัน เช่นในงานวิศวกรรมชลประทานก็มีการวิจัยค้นคว้าอยู่เสมอในเรื่องความสัมพันธ์ระหว่าง น้ำ ดินและพืช การทดลองการใช้ปุ๋ยกับพืชเพื่อตรวจสอบผลผลิต ในงานวิศวกรรมโยธาก็มีการปฏิบัติการทดสอบวัสดุและดินทางวิศวกรรม ในงานวิศวกรรมอุตสาหการและโรงงานอุตสาหกรรมทั่วไปก็ต้องใช้วิชาอนุมานสถิติเข้าควบคุมการผลิตของโรงงาน ฯลฯ

เพราะความสำคัญของวิชาอนุมานสถิติดังได้กล่าวมานี้ ภาควิชาวิศวกรรมชลประทานจึงได้นำวิชาอนุมานสถิติมาสอนนิสิต โดยมีความมุ่งหมายจะให้นิสิตได้ใช้วิชานี้ให้เกิดประโยชน์ในโอกาสที่จะต้องออกไปปฏิบัติงานหรือการศึกษาเพิ่มเติม

ถ้าจะพูดอย่างกว้าง ๆ คำว่าสถิติ (statistics) หมายถึงการรวบรวมและการจัดทำตารางข้อมูลต่าง ๆ ตลอดจนการหา conclusions จากข้อมูลเหล่านั้น การรวบรวมและการจัดทำตารางข้อมูลตลอดจนการคำนวณต่าง ๆ เช่นการหาค่าเฉลี่ย (average) และจำนวนเปอร์เซ็นต์นี้ เรียกว่า descriptive statistics แต่ระบบวิธีการหา conclusion จากข้อมูลต่าง ๆ เรียกว่า statistical inference ซึ่งนิสิตจะได้เรียนในวิชานี้

ก่อนอื่นนิสิตควรจะทราบว่าวิชา statistical inference มีถ้อยคำและสัญลักษณ์ที่ใช้เฉพาะอยู่มากซึ่งผู้สอนไม่ประสงค์จะแปลถ้อยคำเหล่านี้เป็นภาษาไทย จึงขอทิ้งไว้เป็นภาษาอังกฤษตามเดิม

ความจริงคำต่าง ๆ ในวิชาสถิติก็เป็นคำธรรมดานี่เอง แต่อาจจะมีความหมายกว้างขวางขึ้นหรือมีความหมายพิเศษขึ้นบ้างเท่านั้น เช่น คำว่า observations ในทางสถิติคือการจดบันทึกเรื่องราว เช่น ความสูงของคน น้ำหนักของสัตว์ อุณหภูมิของอากาศ ฯลฯ เหล่านี้เป็นต้น ดังนั้น observations จึงมีค่าเป็นตัวเลขเสมอ และคำว่า population ก็คือชุดของ observations นั่นเอง เช่น ความสูงของนิสิต–

วิศวกรรมศาสตร์ทุกคนประกอบกันเป็นหนึ่ง population เรียกว่า population ของความสูงของนิสิตวิศวกรรมศาสตร์ ให้สังเกตคำว่า population ในวิชาสถิติ เช่น population ของความสูงของนิสิตวิศวกรรมศาสตร์ดังกล่าวนี้ หมายถึงการเอาความสูงของนิสิตแต่ละคนมาบันทึกรวบรวมไว้แต่ไม่ได้หมายถึงตัวบุคคลคือนิสิตเลย คำต่อไปที่จะต้องพบเสมอคือ sample ซึ่งหมายถึงกลุ่มของ observations ซึ่งจะมีก observations กี่ได้ที่แยกออกมาจาก population สมมติว่ามีนิสิตคณะวิศวกรรมศาสตร์ 500 คน population ของความสูงของนิสิตวิศวกรรมศาสตร์จึงประกอบด้วย observations ต่าง ๆ ซึ่งเป็นความสูงของนิสิตแต่ละคน 500 observations ถ้าเราแยก 10 observations ออกมาจาก 500 observations กลุ่มของ 10 observations นี้เรียกว่า sample ในเรื่อง sample นี้ยังมีอีกคำหนึ่งที่จำเป็นต้องทราบคือ random sample ซึ่งหมายถึง sample ที่ได้มาโดยวิธีซึ่ง observations ทุกตัวของ population นั้นมีโอกาสที่จะได้รับเลือกเท่าเทียมกัน วิธีการของ statistical inference เกิดขึ้นเมื่อเราจะหา conclusion ที่เกี่ยวกับ population จาก sample หนึ่งที่เราทราบจาก population นั้น

ถ้าเราทราบ population โดยสมบูรณ์ เราจะให้ conclusion ที่เกี่ยวกับ population ซึ่งเราต้องการทราบได้โดยไม่จำเป็นต้องหา sample ออกมาเลย เช่น ถ้าเราทราบ population ของความสูงของนิสิตวิศวกรรมศาสตร์อย่างแน่นอน คือทราบความสูงของนิสิตแต่ละคนทั้ง 500 คนแล้ว เราจะให้ conclusion เช่น บอกค่าเฉลี่ยของความสูงของนิสิตวิศวกรรมศาสตร์ได้โดยเอาความสูงของนิสิตทั้งหมดมารวมกันแล้วหารด้วย 500 แต่ถ้าเราไม่ทราบหรือไม่มีทางจะทราบ population โดยสมบูรณ์ เราจะให้ conclusion ที่เกี่ยวกับ population ได้เหมือนกันแต่จะต้องใช้วิธีการอย่างหนึ่งซึ่งจำเป็นต้องอาศัย sample อันเป็นส่วนหนึ่งของ population นั้น นักวิทยาศาสตร์ทั่วไปเรียกวิธีการดังกล่าวนี้ว่า "induction" แต่นักสถิติเรียกว่า "statistical inference"

Chapter 2

Descriptive Statistics

2.1 Mathematical Notations

โดยปกติวิชาสถิติมักจะใช้อักษรย่อหรืออักษรกรีกตัวใหญ่แทนความหมายต่าง ๆ เช่น

Σ (อ่านว่า sigma) หมายถึง "ผลบวกของ"

y's ใช้แทนความหมายของ observations ต่าง ๆ ดังนั้น:

y_1 คือ observation ที่ 1

y_2 คือ observation ที่ 2

และต่อ ๆ ไปตามลำดับ

Σy คือผลบวกของ observations (the sum of observations) ถ้า 5 observations มีค่า 3, 5, 7, 3, 2 ตามลำดับ ดังนั้น

$$y_1 = 3$$
$$y_2 = 5$$
$$y_3 = 7$$
$$y_4 = 3$$
$$y_5 = 2$$

เพราะฉะนั้น
$$\begin{aligned}\Sigma y &= y_1 + y_2 + y_3 + y_4 + y_5 \\ &= 3 + 5 + 7 + 3 + 2 \\ &= 20\end{aligned}$$

Σy^2 คือผลบวกของ observations ที่ยกกำลังสองแล้ว (the sum of the squared observations) ดังนั้น จากตัวอย่างข้างบน

$$\begin{aligned}\Sigma y^2 &= y_1^2 + y_2^2 + y_3^2 + y_4^2 + y_5^2 \\ &= (3)^2 + (5)^2 + (7)^2 + (3)^2 + (2)^2 \\ &= 9 + 25 + 49 + 9 + 4 \\ &= 96\end{aligned}$$

$(\Sigma y)^2$ คือ ค่ายกกำลังสองของผลบวกของ observations (the square of the sum of observations) ดังนั้น จากตัวอย่างที่แล้ว

$$(\Sigma y)^2 = (y_1 + y_2 + y_3 + y_4 + y_5)^2$$
$$= (3 + 5 + 7 + 3 + 2)^2$$
$$= (20)^2$$
$$= 400$$

2.2 Mean

mean ของ population คือค่าเฉลี่ยของ observations ทั้งหมดของ population นั้น ถ้าอายุของนิสิต 5 คน เป็น 22, 22, 21, 24 และ 21 ปี ตามลำดับ ดังนั้น mean หรืออายุเฉลี่ยของนิสิต 5 คนนี้คือ $\frac{22+22+21+24+21}{5} = \frac{110}{5} = 22$ ปี กล่าวได้ว่าโดยทั่วไป mean ก็คือผลบวกของ observations หารด้วยจำนวน observations เราจะใช้อักษรกรีก μ (อ่านว่า mu) แทน mean ดังนั้น

$$\mu = \frac{y_1 + y_2 + y_3 + y_4 + \cdots\cdots + y_N}{N}$$

$$= \frac{\Sigma y}{N}$$

ในเมื่อ N = จำนวน observations

2.3 Variance and Standard Deviation

variance ของ population เป็นมาตรวัดการแปรผันของ observations ต่าง ๆ ภายใน population นั้น

ถ้าเรามี 4 populations และแต่ละ population มี 4 observations ดังแสดงไว้ข้างล่างนี้

population ที่ 1 ประกอบด้วย 5, 5, 5, 5

population ที่ 2 ประกอบด้วย 6, 6, 4, 4

population ที่ 3 ประกอบด้วย 8, 8, 2, 2

population ที่ 4 ประกอบด้วย 10, 10, 0, 0

จาก populations ทั้งสี่นี้ถึงแม้ว่าแต่ละ population จะมี mean คือ 5 เท่ากันทุก population ก็ตาม แต่จำนวนการแปรผันของ observations ต่าง ๆ ภายในแต่ละ population จะไม่เท่ากัน ดังจะเห็นได้ว่า population ที่ 1 ไม่มีการแปรผันในระหว่าง observations เลย population ที่ 2 มีการแปรผันในระหว่าง observations population ที่ 3 มีการแปรผันในระหว่าง observations มากขึ้น และ population ที่ 4 มีการแปรผันในระหว่าง observations มากที่สุด

การวัดการแปรผันของ observations ทำได้หลายวิธี แต่วิธีที่ดีที่สุดคือการวัดด้วย variance ในที่นี้เราจะใช้อักษรกรีก σ^2 (sigma squared) แทน variance และเราจะหาค่าของ variance ได้โดยสมการต่อไปนี้

$$\sigma^2 = \frac{(y_1 - \mu)^2 + (y_2 - \mu)^2 + (y_3 - \mu)^2 + \ldots\ldots + (y_N - \mu)^2}{N}$$

$$= \frac{\Sigma(y - \mu)^2}{N} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

ในเมื่อ

σ^2 = variance

μ = mean ของ observations

N = จำนวน observations เหล่านั้น

จากสมการ (1) เราจะคำนวณค่าของ variance ของ populations ทั้งสี่ในตัวอย่างที่แล้วได้ดังนี้

population ที่ 1 : $\sigma^2 = \frac{(5-5)^2 + (5-5)^2 + (5-5)^2 + (5-5)^2}{4}$

$= 0$

population ที่ 2 : $\sigma^2 = \frac{(6-5)^2 + (6-5)^2 + (4-5)^2 + (4-5)^2}{4}$

$= 1$

population ที่ 3 : $\sigma^2 = \frac{(8-5)^2 + (8-5)^2 + (2-5)^2 + (2-5)^2}{4}$

$= 9$

population ที่ 4 : $\sigma^2 = \frac{(10-5)^2 + (10-5)^2 + (0-5)^2 + (0-5)^2}{4}$

$= 25$

ค่าของ square root ของ variance เรียกว่า standard deviation ฉะนั้นเราจึงใช้ σ (sigma) แทน standard deviation และ standard deviation นี้ จะใช้วัดการแปรผันของ observations ด้วยเหมือนกัน ให้สังเกตไว้อย่างหนึ่งว่า ถ้า variance มีค่าเท่ากับศูนย์แล้ว standard deviation ก็จะเท่ากับศูนย์ด้วย

2.4 Effect of Change in Observations on Mean and Variance

เนื่องจากเราหาค่าของ mean และ variance ได้จาก observations ดังนั้นถ้าเราเปลี่ยนค่าของ observations เหล่านี้ ค่าของ mean และ variance ก็จะเปลี่ยนตามไปด้วย

สมมติว่าเรามี 3 observations คือ 10, 12, 14

$$\text{mean} = \frac{10 + 12 + 14}{3} = 12$$

ถ้าเอา 10 บวกเข้ากับ observations เดิมทุกตัว observations ใหม่จะกลายเป็น 20, 22 และ 24 ตามลำดับ

ดังนั้น:

$$\text{mean ใหม่} = \frac{20 + 22 + 24}{3} = 22$$

ซึ่งมากกว่า mean เดิมเท่ากับ 22 — 12 = 10

คราวนี้ถ้าเอา 5 ลบออกจาก observations เดิมทุกตัว observations ใหม่จะกลายเป็น 5, 7 และ 9

$$\text{mean ใหม่} = \frac{5 + 7 + 9}{3} = 7$$

ซึ่งน้อยกว่า mean เดิมเท่ากับ 12 — 7 = 5

แต่ว่าการเปลี่ยนค่าของ observations โดยวิธีนี้จะไม่กระทบกระเทือนถึงค่าของ variance เลย และยังผลให้ค่าของ standard deviation ไม่เปลี่ยนไปด้วยเหมือนกัน ความจริงข้อนี้อาจพิสูจน์ให้เห็นได้ดังต่อไปนี้

จาก observations 10, 12, 14 ซึ่งมี mean = 12 นั้น ค่าของ variance คือ

$$\sigma^2 = \frac{(10-12)^2 + (12-12)^2 + (14-12)^2}{3}$$

$$= \frac{8}{3}$$

เมื่อเอา 10 บวกเข้ากับ observations เดิมทุกตัวก็จะได้ observations ใหม่คือ 20, 22, 24 ซึ่งมี mean ใหม่ = 22 ดังนั้นค่าของ variance ใหม่คือ

$$\sigma^2 \text{ ใหม่} = \frac{(20-22)^2+(22-22)^2+(24-22)^2}{3}$$

$$= \frac{8}{3}$$

ซึ่งเท่ากับ variance เดิม

ถ้าเอา 5 ลบออกจาก observations เดิมทุกตัว observations ใหม่จะกลายเป็น 5, 7 9 ซึ่งมี mean ใหม่ = 7 ดังนั้นค่าของ variance ใหม่คือ

$$\sigma^2 \text{ ใหม่} = \frac{(5-7)^2+(7-7)^2+(9-7)^2}{3}$$

$$= \frac{8}{3}$$

ซึ่งเท่ากับ variance เดิม

จากความจริงดังกล่าวนี้ เราอาจนำผลมาสรุปเป็น theorem ได้ดังต่อไปนี้

Theorem 2.4 a: If a fixed amount is added to or subtracted from each of the observations, the mean will be increased or decreased by that amount, but the variance and standard deviation are not affected.

เราได้ทราบผลที่เกิดขึ้นกับ mean และ variance เมื่อเพิ่มหรือลดค่า observations ทุกตัวด้วยจำนวนคงที่แล้ว ต่อไปนี้เราจะพิจารณาถึงผลที่เกิดขึ้นกับ mean และ variance เมื่อคูณหรือหาร observations ทุกตัวด้วยจำนวนคงที่บ้าง

จากตัวอย่างที่แล้วเรามี 3 observations คือ 10, 12, 14 ซึ่งมี mean = 12, variance $= \frac{8}{3}$ และ standard deviation $= \sqrt{\frac{8}{3}}$ ถ้าเอา 10 คูณ observations เดิมทุกตัวจะได้ observations ใหม่คือ 100, 120, 140 ดังนั้น

$$\text{mean ใหม่} = \frac{100+120+140}{3}$$

$$= 120$$

$$= 10(12)$$

$$= 10(\text{mean เดิม})$$

$$\text{variance ใหม่} = \frac{(100-120)^2+(120-120)^2+(140-120)^2}{3}$$

$$= \frac{800}{3}$$

$$= 100(\frac{8}{3})$$

$$= 10^2(\frac{8}{3})$$

$$= 10^2(\text{variance เดิม})$$

$$\text{standard deviation ใหม่} = \sqrt{100(\frac{8}{3})}$$

$$= 10\sqrt{\frac{8}{3}}$$

$$= 10\ (\text{standard deviation เดิม})$$

สำหรับการเอาจำนวนคงที่ไปหาร observations ทุกตัว ผลที่เกิดขึ้นกับ mean, variance และ standard deviation จะเป็นไปตามหลักนี้ เพราะการหาร observations ทุกตัวด้วย 10 ก็เท่ากับคูณ observations ทุกตัวด้วย $\frac{1}{10}$ นั่นเอง

จากความจริงดังกล่าวนี้ เราอาจนำผลมาสรุปเป็น theorem ได้ดังต่อไปนี้

Theorem 2.4 b: If each of the observations is multiplied by a fixed quantity m, the new mean is m times the old mean, the new variance is m^2 times the old variance, and the new standard deviation is m times the old standard deviation.

2.5 Frequency Table

ในข้อนี้จะกล่าวถึงวิธีจัด observations ต่าง ๆ ลงในตาราง สมมติว่านิสิตคณะวิศวกรรมศาสตร์ 10 คน มีอายุ 21, 20, 21, 22, 20, 22, 21, 20, 23 และ 21 ปี จะเห็นว่า

นิสิตอายุ 20 ปี มี 3 คน

นิสิตอายุ 21 ปี มี 4 คน

นิสิตอายุ 22 ปี มี 2 คน

นิสิตอายุ 23 ปี มี 1 คน

เราอาจแบ่งนิสิต 10 คนนี้ออกได้เป็น 4 กลุ่มตามอายุ กลุ่มหนึ่ง ๆ เรียกว่า "class" หนึ่ง กล่าวโดยทั่วไป frequency จะหมายถึงจำนวนครั้งซึ่ง observations ที่มีค่าเดียวกันเกิดขึ้นใน class ตารางแสดง class frequencies นี้เรียกว่า "frequency table" และจำนวน observations ทั้งหมดซึ่งเป็นผลบวกของ class frequency ทุก class เรียกว่า "total frequency" ตัวอย่าง frequency table ของอายุนิสิตข้างบนนี้ได้แสดงไว้ใน Table 2.5 a

Table 2.5 a

Age (y)	frequency (f)	relative frequency (r.f.)	relative cumulative frequency (r.c.f.)
20	3	30 %	30 %
21	4	40 %	70 %
22	2	20 %	90 %
23	1	10 %	100 %
Total	10	100 %	

ใน Table 2.5 a จะเห็นว่ามี frequencies อีก 2 อย่างคือ relative frequency และ relative cumulative frequency

relative frequency (r.f.) ของ class ใด คือ frequency ของ class นั้นที่แสดงเป็นเปอร์เซนต์ของ total frequency เช่นใน Table 2.5 a frequencyของ class อายุ 21 ปีคือ 4 ฉะนั้น relative frequency ของ class นี้คือ 40 % (4 ส่วนจาก 10 ส่วน)

relative cumulative frequency (r.c.f.) ของจุดใดจุดหนึ่งคือผลบวกของ relative frequency ต่าง ๆ ตั้งแต่แรกมาจนถึงจุดนั้น

เราอาจหา mean จาก frequency table ได้ด้วย เพราะว่าตามหลัก mean คือผลบวกของ observations หารด้วยจำนวน observations ดังนั้นจากตัวอย่าง อายุนิสิต 10 คนนี้ ค่าเฉลี่ย (mean) ของอายุนิสิตจะหาได้ดังนี้

$$\mu = \frac{\Sigma y}{N}$$

$$= \frac{21+20+21+22+20+22+21+20+23+21}{10}$$

$$= \frac{211}{10}$$

$$= 21.1 \text{ ปี}$$

และจาก Table 2.5 a จะเห็นว่า observations ที่มีค่า 20 เกิดขึ้น 3 ครั้งจึงต้องนับ 3 ครั้ง observations ที่มีค่า 21 เกิดขึ้น 4 ครั้งก็ต้องนับ 4 ครั้ง และ observations ที่มีค่า 22 และ 23 ก็ต้องนับ 2 ครั้ง และ 1 ครั้งตามลำดับเช่นเดียวกัน เพราะฉะนั้นผลบวกของ observations ทั้งหมดจึงเป็น $(20 \times 3) + (21 \times 4) + (22 \times 2) + (23 \times 1) = 211$ ปี ส่วนจำนวน observations ทั้งหมดหรือ total frequency นั้นเท่ากับ 10 ดังนั้น mean $= \frac{211}{10} = 21.1$ ปี เช่นเดียวกัน ถ้าจะเขียนเป็นสมการสำหรับใช้ทั่วไปในการหา mean จาก frequency table ก็จะได้ดังนี้

$$\mu = \frac{\Sigma yf}{N} \qquad \text{.........................} \quad (1)$$

โดยหลักเดียวกันเราอาจหา variance จาก frequency table ได้ด้วย จากหลักนี้เราจะต้องคูณ $(y-\mu)^2$ ทุกจำนวนด้วย frequency ของมัน ดังนั้นถ้าจะเขียนเป็นสมการสำหรับใช้ทั่วไปในการหา variance จาก frequency table ก็จะได้ดังนี้

$$\sigma^2 = \frac{\Sigma (y-\mu)^2 f}{N} \qquad \text{.................} \quad (2)$$

รายละเอียดการคำนวณ mean และ variance ของอายุนิสิต 10 คนได้แสดงไว้ใน Table 2.5 b ต่อไปนี้

Table 2.5 b

y	f	yf	$(y-\mu)$	$(y-\mu)^2$	$(y-\mu)^2 f$
20	3	60	- 1.1	1.21	3.63
21	4	84	- 0.1	0.01	0.04
22	2	44	0.9	0.81	1.62
23	1	23	1.9	3.61	3.61
sum	$10 = N$	$211 = \sum yf$			$8.90 = \sum (y-\mu)^2 f$

$$\mu = \frac{\sum yf}{N} = \frac{211}{10} = 21.1$$

$$\sigma^2 = \frac{\sum (y-\mu)^2 f}{N} = \frac{8.90}{10} = 0.89$$

$$\sigma = \sqrt{0.89} = 0.943$$

2.6 Histogram and Frequency Curve

histogram คือกราฟของ frequency table ดังนั้นจาก frequency table 2.5 a เราจะเขียนเป็น histogram ได้ดังรูปข้างล่างนี้

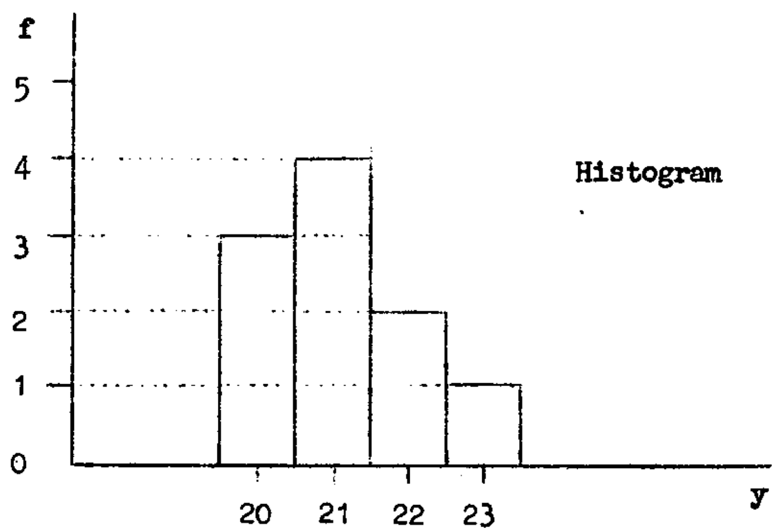


Fig. 2.6 a

จาก Fig. 2.6 a จะสังเกตเห็นว่ารูปสี่เหลี่ยมมุมฉากทุกรูปตั้งอยู่บนฐานซึ่งแทนช่วงกว้างของ observations สะเกลด้านตั้งซึ่งแทน frequency จะเริ่มด้วยศูนย์เสมอ

ความสูงของรูปสี่เหลี่ยมมุมฉากจะเท่ากับ frequency ของ class ของมัน และเนื่องจากฐานของรูปสี่เหลี่ยมมุมฉากเหล่านี้ยาวเท่ากัน ฉะนั้นเนื้อที่ของรูปสี่เหลี่ยมมุมฉากจึงเป็นปฏิภาคกับ frequencies ด้วย ตัวอย่างเช่น frequency ของ class อายุ 21 ปี เป็น 2 เท่าของ class อายุ 22 ปี เนื้อที่ของรูปสี่เหลี่ยมมุมฉากของ class อายุ 21 ปี ก็เป็น 2 เท่าของ class อายุ 22 ปีเช่นเดียวกัน เรโชของเนื้อที่ของรูปสี่เหลี่ยมมุมฉากของ class หนึ่งต่อเนื้อที่ทั้งหมดของรูปสี่เหลี่ยมมุมฉากทุกรูปของทุก class ก็คือ relative frequency ของ class นั้น ตัวอย่างเช่นเนื้อที่ของรูปสี่เหลี่ยมมุมฉากของ classอายุ 23 ปี เท่ากับ 10 % ของเนื้อที่ทั้งหมด

เราจะสมมติต่อไปว่าถ้ามีนิสิตอายุระหว่าง 20 ปี ถึง 23 ปี อยู่เป็นจำนวนมาก และถ้าเราพิจารณาช่วงอายุของนิสิตห่างกันช่วงละ 1 ปีแล้ว เราจะได้ histogram ซึ่งประกอบด้วยรูปสี่เหลี่ยมมุมฉาก 4 รูป สมมติว่าเป็นรูปดังแสดงไว้ข้างล่างนี้

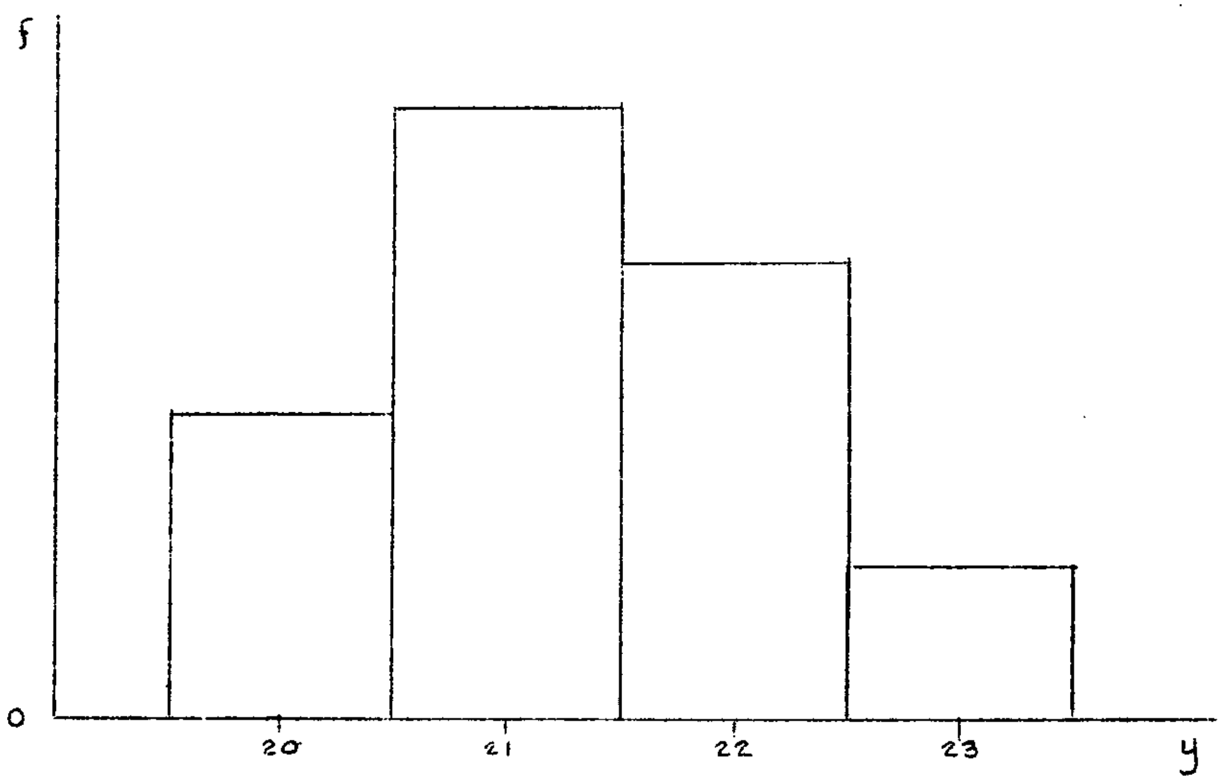


Fig. 2.6 b

ถ้าเราคิดช่วงอายุเป็นเดือน ทุก ๆ **year-class** จะมี 12 **month-classes** ดังนั้น **histo-gram** จะประกอบด้วยรูปสี่เหลี่ยมมุมฉาก 12 x 4 หรือ 48 รูป ดัง **Fig. 2.6 c**

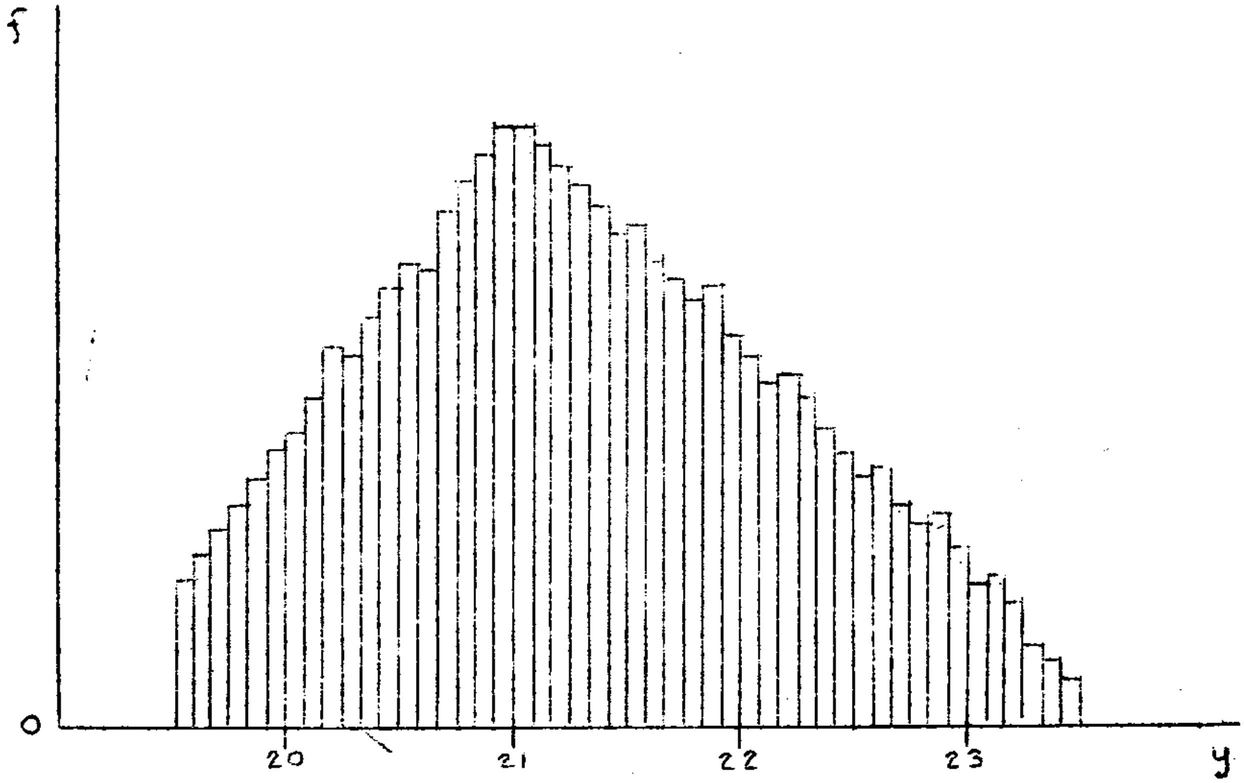


Fig. 2.6 c

ถ้าเราคิดช่วงอายุเป็นวันแล้ว ในระยะ 4 ปีนี้จะมี (4 x 365) + 1 = 1461 วัน และ **histo-gram** ก็จะมีรูปสี่เหลี่ยมมุมฉากทั้งหมดถึง 1461 รูป

จึงเห็นได้ว่า เมื่อหน่วยการวัดเล็กลงทุกที ความกว้างของรูปสี่เหลี่ยมมุมฉากก็จะยิ่งน้อยลง ส่วนจำนวนรูปสี่เหลี่ยมมุมฉากจะกลับมีมากขึ้นเรื่อย ๆ แต่เนื้อที่ทั้งหมดของรูปสี่เหลี่ยมมุมฉากจะยังคงเท่าเดิมอยู่ (ดู **Fig. 2.6 b** และ **2.6 c**)

ด้วยวิธีการดังกล่าวนี้ ยอดของรูปสี่เหลี่ยมมุมฉากต่าง ๆ จะกลายเป็นเส้นโค้งเข้าไปทุกที เส้นโค้งที่ได้นี้เรียกว่า **frequency curve** และให้สังเกตว่าไม่มีส่วนใดของ **frequency curve** อยู่ต่ำกว่าแกนราบ เพราะว่า **frequency** เป็นจำนวน **observations** จึงมีค่าลบไม่ได้

เราใช้เนื้อที่ซึ่งอยู่ระหว่างเส้นโค้งกับแกนราบแทน **relative frequency (r.f.)** และเนื้อที่ทั้งหมดใต้เส้นโค้งก็คือ **relative frequency** ทั้งหมดซึ่งเท่ากับ 100 %

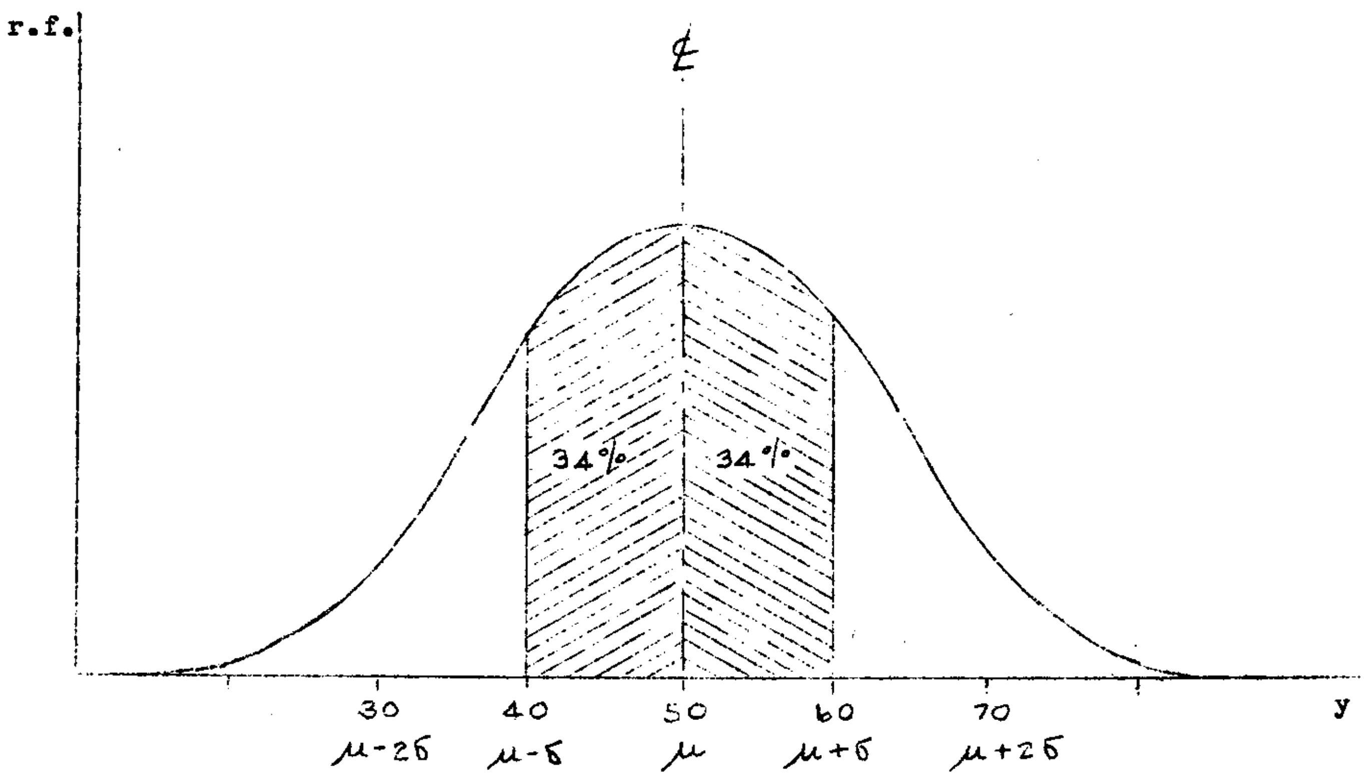
Chapter 3

The Normal Distribution

frequency curves ในทางสถิตินั้นมีมากมายนับไม่ถ้วน แต่ที่สำคัญที่สุดคือ noral curve คำว่า "normal" ในที่นี้เป็นการระบุลักษณะโดยเฉพาะอย่างหนึ่งของ curve และที่เราเรียก frequency curve ชนิดหนึ่งเป็น normal curve นั้น ก็ไม่ได้หมายความว่า frequency curves ชนิดอื่นจะเป็น abnormal curves เพราะฉะนั้น normal curve จึงเป็น frequency curve ชนิดหนึ่งซึ่งมีความแตกต่างอย่างเด่นชัดจาก frequency curves ชนิดอื่น

3.1 Some Properties of The Normal Curve

Fig. 3.1 a ข้างล่างนี้แสดงให้เห็นกราฟของ normal curve เส้นหนึ่ง



$\mu = 50$

$\sigma = 10$

Fig. 3.1 a

จาก Fig. 3.1 a จะเห็นว่า normal curve มีรูปร่างเหมือนกับระฆัง (bell-shaped) วางคว่ำ อย่างไรก็ดี curves ชนิดอื่นก็อาจมีรูปร่างเหมือนกับระฆังคว่ำได้เหมือนกัน รูปร่างของ normal curve นี้ ถ้าเราพิจารณาจากเส้นตั้งฉากกับแกนราบตรงจุด mean (μ) ออกไปทั้งสองข้างแล้วจะเห็นว่ามันเหมือนกันทุกประการ ดังนั้น ถ้าเราพับ normal curve ใน Fig. 3.1 a ตามแนวแกนกลาง (℄) แล้ว เส้นโค้งทางด้านขวามือจะทับเส้นโค้งทางด้านซ้ายมือสนิท ความจริง normal curve ไม่ใช่เส้นโค้งเส้นเดียว แต่มีได้มากมายหลายเส้น เรียกว่าเป็น family of curves Fig. 3.1 a แสดงให้เห็น normal curve เส้นหนึ่งเท่านั้น และเป็น normal curve ที่มี mean = 50 และ standard deviation = 10

ก่อนที่จะกล่าวถึง normal curve โดยจำเพาะเส้นหนึ่งนั้นเราจะต้องระบุ mean (μ) และ standard deviation (σ) ลงไปให้แน่ชัด ถ้าพิจารณา normal curves 3 เส้นใน Fig. 3.1b

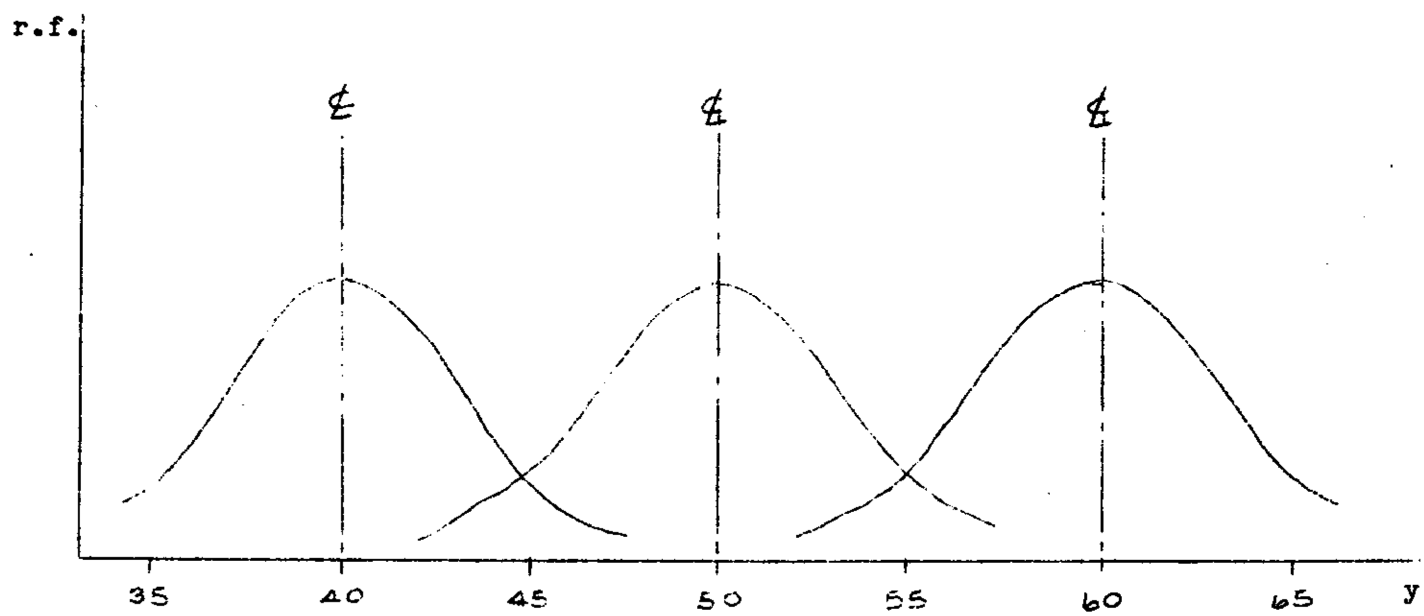


Fig. 3.1 b

จะเห็นว่า normal curves ทั้ง 3 เส้นมี means เท่ากับ 40, 50 และ 60 ตามลำดับ รูปร่างของ normal curves ทั้ง 3 เส้นเหมือนกันทุกประการ แต่ความแตกต่างของมันคือที่ตั้งบนแกนราบ normal curve เส้นที่มี mean = 50 ตั้งอยู่ทางขวามือของเส้นที่มี mean = 40 และเส้นที่มี mean = 60 ตั้งอยู่ทางขวามือของเส้นที่มี mean = 50 อีกทีหนึ่ง เพราะฉะนั้นเมื่อเรากำหนด mean ของ curve ใดลงไปแล้วตำแหน่งที่ตั้งของ curve นั้นก็จะอยู่ตายตัว

ส่วน standard deviation จะเป็นสิ่งแสดงถึงการแผ่ขยายของ curve ถ้าพิจารณา normal curves 3 เส้นใน Fig. 3.1 c

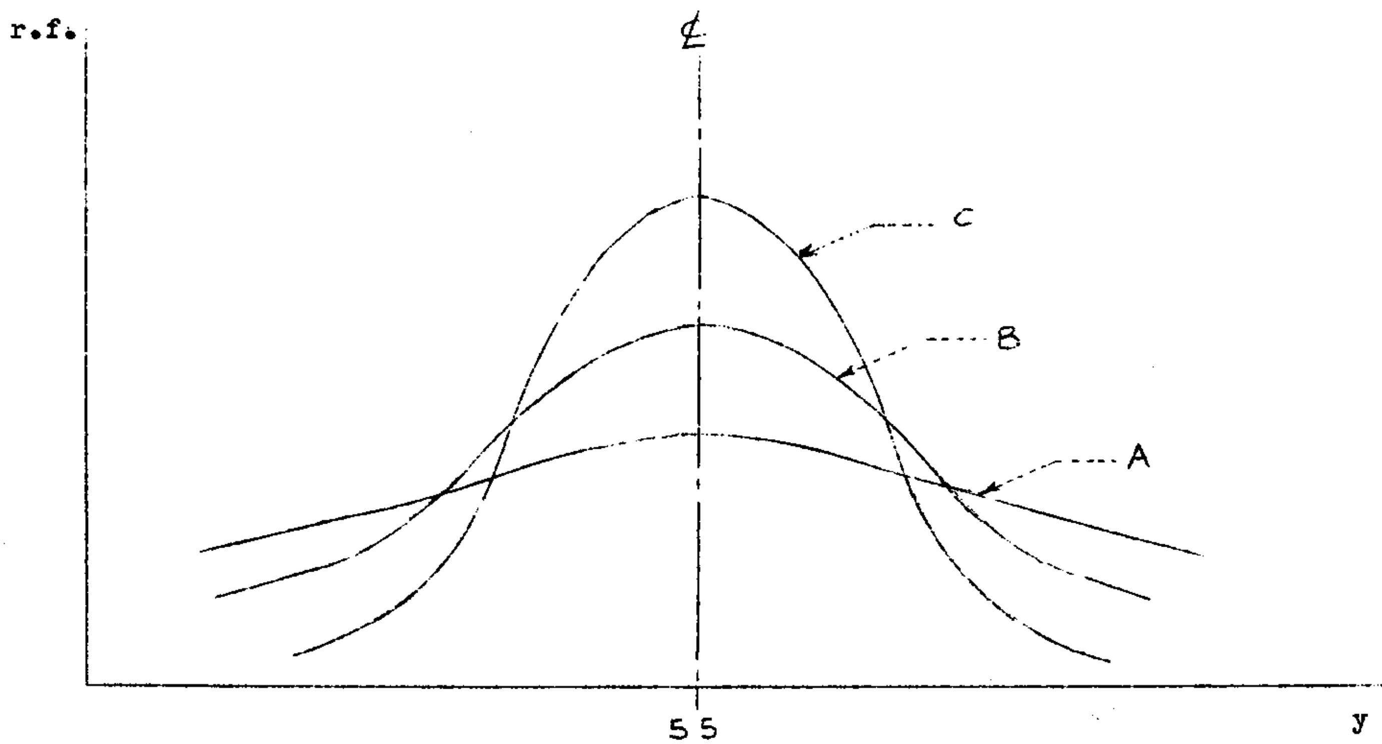


Fig. 3.1 c

จะเห็นว่า normal curves ทั้ง 3 เส้นมี mean เท่ากันคือ 55 แต่มี standard deviation ไม่เท่ากัน curve A ซึ่งเป็น curve ที่แผ่กว้างมากที่สุดนั้นมี standard deviation มากที่สุด และ curve c มี standard deviation น้อยที่สุด แต่ normal curves ทั้ง 3 เส้นใน Fig. 3.1 b มี standard deviation เท่ากัน curves จึงแผ่กว้างเท่า ๆ กัน

ได้กล่าวมาแล้วในข้อ 2.6 ว่า เนื้อที่ใต้ normal curve ทั้งหมดเท่ากับ 100 % symmetric property ของ normal curve (ดู Fig. 3.1 a) จะแสดงให้เห็นว่า 50 % ของจำนวน observations ทั้งหมดจะตกอยู่ต่ำกว่า mean และอีก 50 % จะตกอยู่เหนือ mean ช่วงระหว่าง μ ถึง $\mu + \sigma$ จะมีจำนวน observations ประมาณ 34 % ตัวอย่างเช่นใน Fig. 3.1 a ซึ่งมี mean (μ) = 50 และ standard deviation (σ) = 10 นั้น 34 % ของจำนวน observations จะมีค่ามากกว่า 50 แต่น้อยกว่า 60 โดยทำนองเดียวกันในช่วงระหว่าง μ ถึง $\mu - \sigma$ ก็จะมีจำนวน observations ที่มีค่ามากกว่า 40 แต่น้อย

กว่า 50 อยู่ 34 % ดังนั้นในช่วงระหว่าง $\mu - \sigma$ ถึง $\mu + \sigma$ หรือจาก observations ที่มีค่า 40 ถึง 60 จะมีจำนวน observations อยู่ 34 % + 34 % = 68 % และเนื่องจาก 50 % ของจำนวน observations ตกอยู่ต่ำกว่า mean (μ) และช่วงระหว่าง μ ถึง $\mu + \sigma$ ก็มีจำนวน observations อยู่ 34 % เพราะฉะนั้น 84 % ของจำนวน observations จะตกอยู่ต่ำกว่า $\mu + \sigma$ จากตัวอย่างนี้จะทราบได้ว่า 84 % ของจำนวน observations ทั้งหมดมีค่าน้อยกว่า 60 และอีก 16 % ของจำนวน observations ทั้งหมดมีค่ามากกว่า 60

3.2 Table of The Normal Curve

จากตัวอย่าง normal curve ซึ่งมี mean = 50 และ standard deviation = 10 ในข้อที่แล้ว relative frequency ของ observations ที่มีค่าระหว่าง 50 ถึง 60 จะเป็น 34 % จำนวนเปอร์เซ็นต์ดังกล่าวนี้ไม่ได้หามาจากกราฟ แต่หาได้จากการคำนวณซึ่งค่อนข้างจะยุ่งยาก

จากผลที่ได้กล่าวมาแล้วทำให้สามารถสร้างตารางขึ้นใช้กันทั่วไปได้อย่างสะดวก ดังนั้นเมื่อต้องการทราบ relative frequency ของ observations ใดว่าจะเป็นกี่เปอร์เซ็นต์ก็จะหาได้จากตาราง (ดู Table 3, Appendix) แต่ Table 3, Appendix นี้ทำไว้สำหรับ normal curve ซึ่งมี mean = 0 และ standard deviation = 1 โดยเฉพาะ อย่างไรก็ดีเราอาจใช้ Table 3, Appendix นี้กับ normal distributions หรือ normal curves ซึ่งมี mean และ standard deviation มีค่าเท่าใดก็ได้ แต่ก่อนใช้จะต้องเปลี่ยนค่าของ observations ทั้งหมดเสียก่อน (ดู Theorem 2.4 a และ 2.4 b) ตัวอย่างเช่น distribution ของ observations ของ population หนึ่งเป็น normal curve ซึ่งมี mean = 50 และ standard deviation = 10 ถ้าเอา 50 ไปลบ observations ทุกตัวแล้ว observation แต่ละตัวจะกลายเป็น (y - 50) ดังนั้น mean ของ observations (y - 50) ทั้งหมดจะมีค่าเท่ากับศูนย์ แต่ standard deviation ของ observations (y - 50) ทั้งหมดจะมีค่าเท่ากับ 10 ตามเดิม (ตาม Theorem 2.4 a) ต่อไปถ้าเอา 10ไปหาร observations (y - 50) ทุกตัวอีกทีหนึ่ง observation แต่ละตัวจะกลายเป็น $(\frac{y-50}{10})$ และ mean ของ observations $(\frac{y-50}{10})$ ทั้งหมดก็จะกลายเป็น $\frac{0}{10} = 0$ และ standard deviation ของ observations $(\frac{y-50}{10})$ ทั้งหมดจะเท่ากับ $\frac{10}{10} = 1$ ดังนั้น distribution ของ observations $(\frac{y-50}{10})$ ทั้งหมดจึงเป็น normal curve ซึ่งมี mean = 0 และ standard deviation = 1

observations $(\frac{y-50}{10})$ ซึ่งเปลี่ยนรูปมาจาก observations (y) เดิมตามตัวอย่างนี้ เรียกว่า transformed observations และใช้แทนด้วยอักษร u

ดังนั้น $u = (\frac{y - 50}{10})$

ถ้าจะเขียนเป็นสมการทั่วไปให้ใช้ได้กับทุก normal distribution แล้ว จะเป็นดังนี้

$$u = \frac{y - \mu}{\sigma} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

ซึ่ง distribution ของ u จะเป็น normal curve ที่มี mean = 0 และ standard deviation = 1 เสมอไป

จากตัวอย่างที่แล้ว ค่าของ u ของ observation ที่มีค่า 60 คือ

$$u = \frac{60 - 50}{10} = 1$$

และค่าของ u ของ observation ที่มีค่า 45 คือ

$$u = \frac{45 - 50}{10} = -0.5$$

เราจึงพูดได้อีกอย่างหนึ่งว่า 60 เป็นหนึ่ง standard deviation เหนือ mean 50 และ 45 เป็นครึ่ง standard deviation ต่ำกว่า mean 50

Table 3, Appendix ให้เนื้อที่ใต้ curve จาก $-\infty$ ถึงค่าต่าง ๆ ของ u เนื้อที่นี้แทน relative cumulative frequency (r.c.f.) เพราะเริ่มต้นมาแต่แรกคือ $-\infty$ จนถึงจุดต่าง ๆ ที่กำหนด relative cumulative frequencies ของ u = 1 และ u = - 0.5 นั้น ได้กำหนดไว้เท่ากับ 84.13% และ 30.85 % ตามลำดับ ซึ่งหมายความว่า relative cumulative frequency จาก $-\infty$ ถึง u = 1 หรือหนึ่ง standard deviation เหนือ mean คือ 84.13 % และจาก $-\infty$ ถึง u = - 0.5 หรือครึ่ง standard deviation ต่ำกว่า mean คือ 30.85 % เพราะฉะนั้น relative frequency (r.f.) ของ observations ซึ่งตกในช่วง u = 1 และ u = - 0.5 คือ 84.13 − 30.85 = 53.28 % relative frequencies (r.f.) เหล่านี้จะบอกให้ทราบถึง distribution ของ observations ที่เป็น normal distribution หรือ normal curve เพราะว่า frequency curves อื่น ๆ อีกมากอาจมีรูปเป็น bell-shaped curves เหมือนกัน ดังนั้นการพิจารณาดูแต่กราฟของ frequency curve จึงไม่เพียงพอที่จะบอกได้ว่า frequency curve นั้นเป็น normal curve หรือไม่

3.3 Normal Probability Graph Paper

การพิจารณาว่า distribution ของ observations ชุดใดเป็น normal distribution หรือไม่นั้นจะทำได้โดยการพล็อตค่าของ observations กับ relative cumulative frequencies (r.c.f.) ลงในกระดาษกราฟพิเศษชนิดหนึ่ง

ถ้าเราพล็อตค่าต่าง ๆ ของ u และ area (r.c.f.) ใน Table 3, Appendix ลงในกระดาษกราฟธรรมดาแล้วจะมีรูปดังนี้

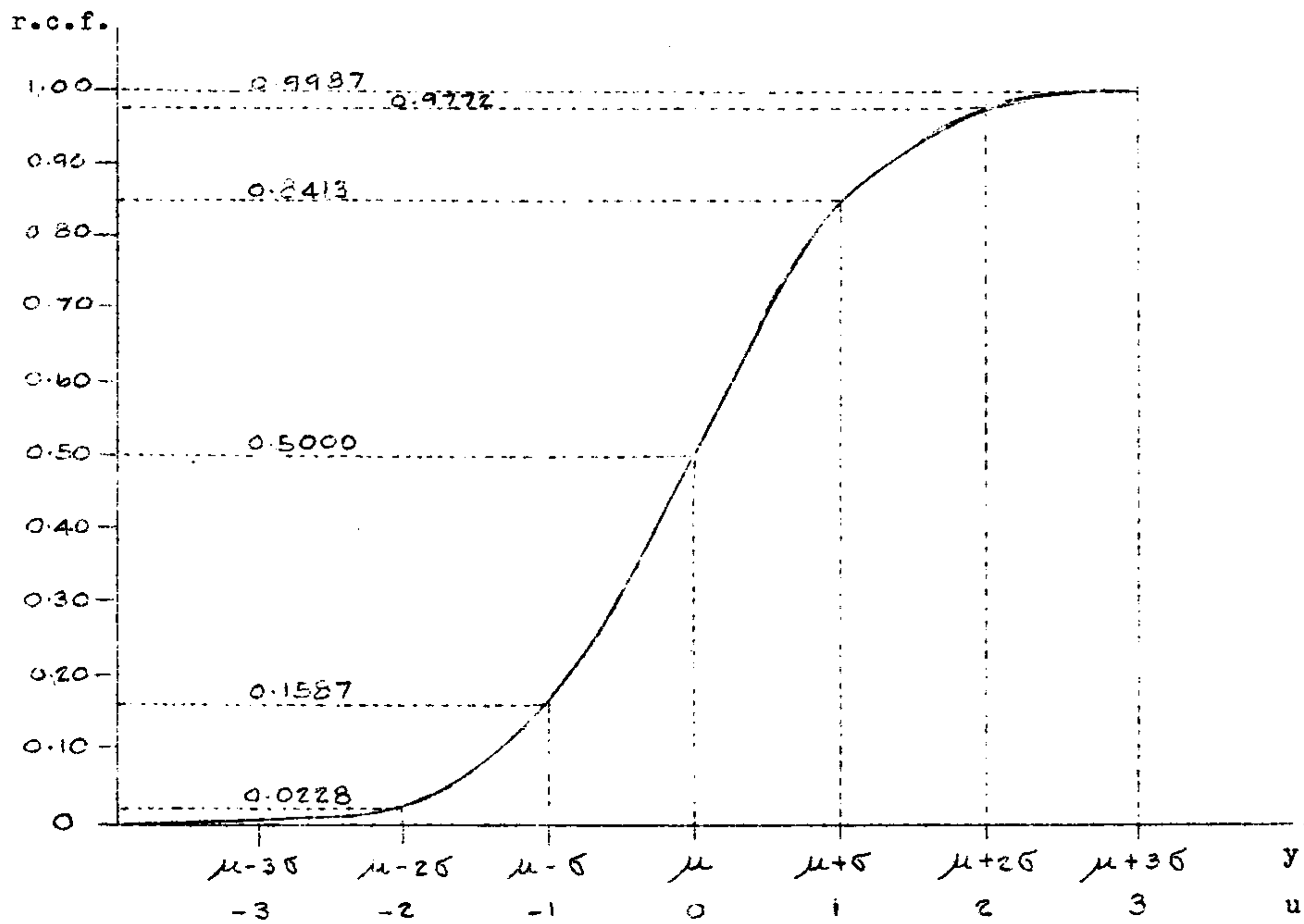


Fig. 3.3 a

ใน Fig. 3.3 a สะเกลด้านราบแสดงค่าต่าง ๆ ของ y หรือ u สะเกลด้านตั้งแสดงค่าของ relative cumulative frequencies ตามค่าของ y หรือ u เส้นโค้งที่พล็อตได้จะมีรูปเป็นตัว S (S-shaped curve)

ถ้าพล็อตเส้นโค้งนี้ลงบนแผ่นยางเราอาจยืดเส้นโค้งนี้ออกเป็นเส้นตรงได้ ดัง Fig. 3.3 b

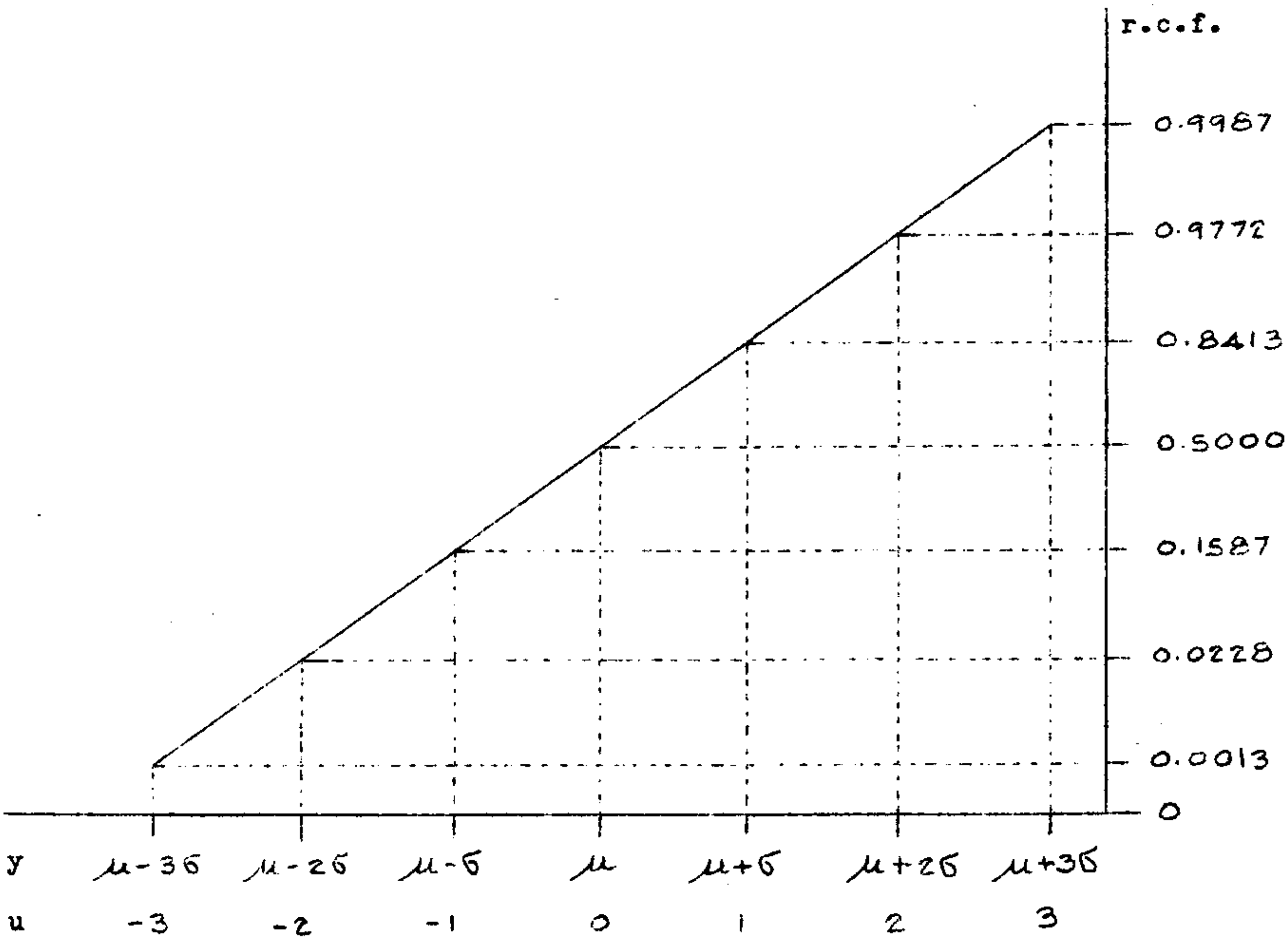


Fig. 3.3 b

แต่เมื่อยืดเส้นโค้งออกเป็นเส้นตรงแล้ว สะเกลด้านตั้งจะผิดความจริงมาก เพราะฉะนั้นเขาจึงสร้างกระดาษกราฟพิเศษขึ้นชนิดหนึ่งซึ่งมีสะเกลด้านตั้งผิดธรรมดา เรียกว่า normal probability graph paper

เมื่อเราพล็อตจุดต่าง ๆ ที่ค่าของ y หรือ u ตัดกับ relative cumulative frequencies ของมันลงบน normal probability graph paper แล้วลองเอาไม้บรรทัดทาบดู ถ้าจุดต่าง ๆ ที่พล็อตอยู่บนเส้นตรงเดียวกันแล้ว distribution ของชุด observations นั้นจะเป็น normal distribution และเราจะอ่านค่าของ mean และ standard deviation ได้จากกราฟนั้นเลยทีเดียว

ใน Fig. 3.3 b จุด 50 % บนสะเกลด้านตั้ง (r.c.f.) เป็นไปตาม mean (μ) บนสะเกลด้านราบ ดังนั้นถ้าเราลากเส้นราบจากจุด 50 % บนสะเกลด้านตั้งมาตัดเส้นกราฟแล้วลากเส้นดิ่งลงไปตัดสะเกลด้านราบ ก็จะทราบค่า mean (μ) ได้ทันที โดยทำนองเดียวกันจุด 84.13 % บนสะเกลด้านตั้งเป็นไปตาม

$\mu + \sigma$ บนสะเกลด้านราบถ้าเราลากเส้นราบจากจุด 84.13 % บนสะเกลด้านตั้งมาตัดเส้นกราฟแล้วลากเส้นดิ่งลงไปตัดสะเกลด้านราบก็จะทราบค่า mean บวกหนึ่ง standard deviation ($\mu + \sigma$) ได้เช่นเดียวกัน เมื่อเอาค่า μ ไปลบค่า $\mu + \sigma$ ก็จะทราบค่าของ standard deviation (σ) ได้ทันที

Chapter 4

Sampling Experiments

ในบทต่าง ๆ ต่อจากบทที่ 4 นี้ไปนิสิตจะได้เรียนรู้ถึง theorems ต่าง ๆ และเราจะพิสูจน์ให้เห็นความเป็นจริงของ theorems เหล่านี้โดย sampling experiments แต่เพื่อไม่ให้เสียเวลาที่จะต้องกล่าวถึงรายละเอียดของ experiments ซ้ำทุกบท เราจะได้นำอุปกรณ์และวิธีการที่ใช้ในชุดของ sampling experiments ทั้งหมดมากล่าวไว้ในที่นี้ ส่วนผลที่ต้องการเพื่อพิสูจน์ให้เห็นความเป็นจริงของ theorems ต่าง ๆ จะได้กล่าวไว้ในบทต่อ ๆ ไปตามลำดับ

4.1 Description of Population

สำหรับ sampling experiments นี้ เราจะใช้แผ่นกระดาษแข็งกลมซึ่งมีโลหะหุ้มริมจำนวน 500 แผ่นใส่ไว้ในตะกร้าเป็นเครื่องมือ บนแผ่นกระดาษแข็งแต่ละแผ่นจะมีจำนวนเลขซึ่งประกอบด้วยเลขสองตัวเขียนไว้ จำนวนเลขแต่ละจำนวนคือ observation ดังนั้นจะมีจำนวนเลขทั้งหมด 500 จำนวนหรือ 500 observations ประกอบกันเป็นหนึ่ง population จำนวนเลขบนแผ่นกระดาษแข็งเหล่านี้มีค่าต่าง ๆ กันตั้งแต่ 20 ถึง 80 แต่บางค่าก็ไม่มี และบางค่าก็มีซ้ำกัน observation (y) คือจำนวนเลขที่ประกอบด้วยเลขสองตัวที่เขียนไว้บนแผ่นกระดาษแข็ง และ frequency (f) ก็คือจำนวนแผ่นกระดาษแข็งที่มีเลขจำนวนนั้น frequency table และ histogram ของมันได้แสดงไว้ใน Table 4.1 a และ Fig. 4.1 a แล้ว

Table 4.1 a

y	f	y	f	y	f	y	f
20	1	–	–	–	–	–	–
21	0	36	7	51	20	66	6
22	0	37	9	52	19	67	5
23	1	38	10	53	19	68	4
24	1	39	11	54	18	69	3
25	1	40	12	55	18	70	3
26	1	41	13	56	17	71	2
27	1	42	14	57	16	72	2
28	2	43	16	58	14	73	1
29	2	44	17	59	13	74	1
30	3	45	18	60	12	75	1
31	3	46	18	61	11	76	1
32	4	47	19	62	10	77	1
33	5	48	19	63	9	78	0
34	6	49	20	64	7	79	0
35	6	50	20	65	6	80	1

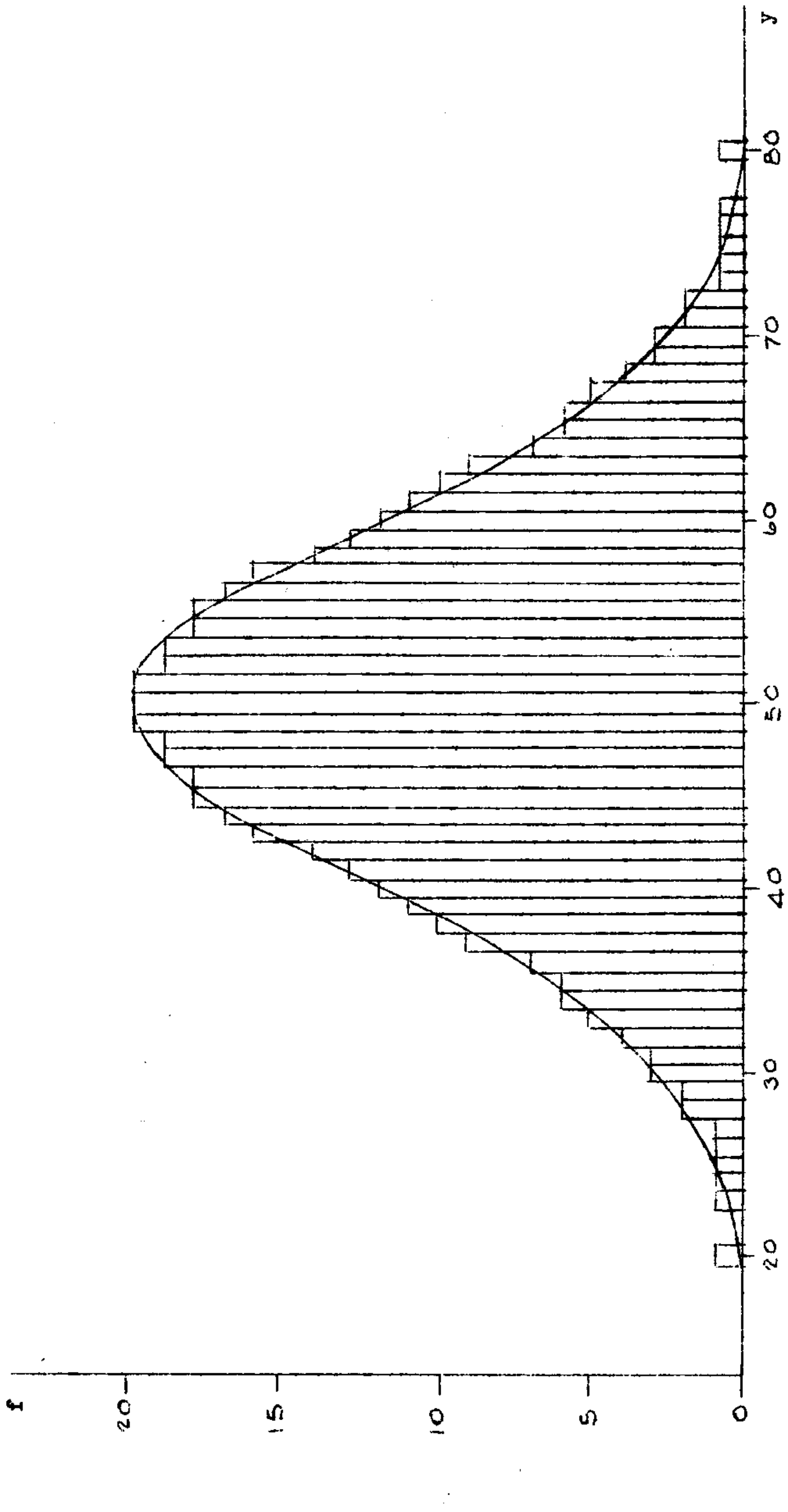


Fig. 4.1 a

observation (y) ใน Table 4.1a คือจำนวนเลขซึ่งเขียนไว้บนแผ่นกระดาษแข็ง และ frequency (f) คือจำนวนแผ่นกระดาษแข็งที่มีเลขจำนวนนั้น

population ที่ประกอบด้วย 500 observations ดังกล่าวนี้ได้ทำขึ้นให้เป็นไปตาม normal distribution โดยประมาณซึ่งมี mean = 50 และ standard deviation = 10 (ดู Fig.4.1 a) เราอาจย่อ Table 4.1a ซึ่งเป็น frequency table ลงได้อีกโดยแสดง relative cumulative frequency ไว้ด้วยดัง Table 4.1 b

Table 4.1 b

y	f	c.f.	r.c.f. (%)
below 30.5	13	13	2.6
30.5 to 35.5	24	37	7.4
35.5 to 40.5	49	86	17.2
40.5 to 45.5	78	164	32.8
45.5 to 50.5	96	260	52.0
50.5 to 55.5	94	354	70.8
55.5 to 60.5	72	426	85.2
60.5 to 65.5	43	469	93.8
65.5 to 70.5	21	490	98.0
above 70.5	10	500	100.0
	500		

เราอาจพลอต relative cumulative frequencies (r.c.f.) ตามค่าของ observations ต่าง ๆ คือ 30.5, 35.5, ..., 70.5 ลงบน normal probability graph paper ได้ดัง Fig. 4.1 b

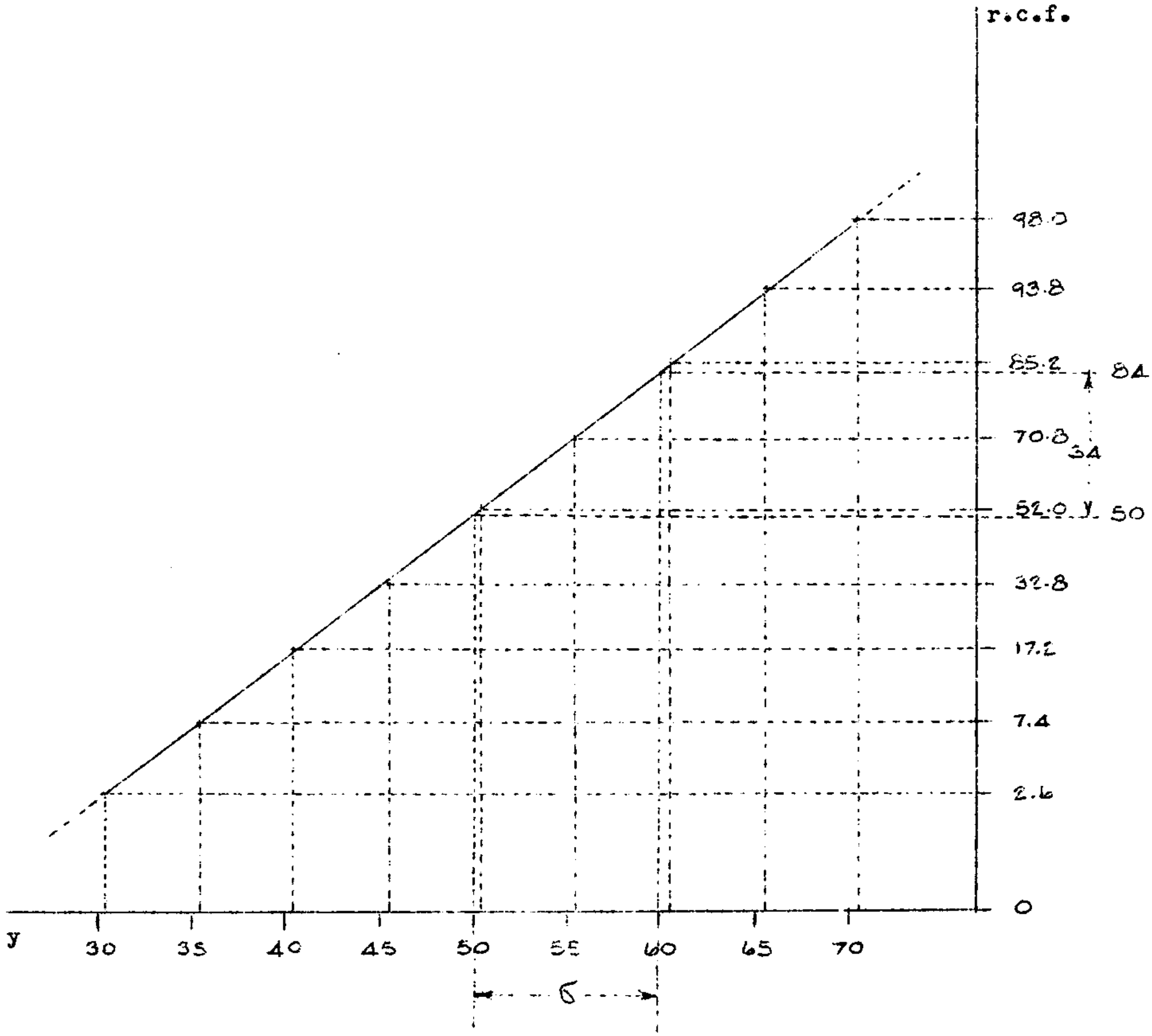


Fig. 4.1 b

การที่จุดที่พล็อตทั้ง 9 จุดอยู่บนเส้นตรงเดียวกันนั้นแสดงว่า population นี้เป็น normal จุด 50 % บนสะเกลด้านตั้งเป็นไปตาม **y** = 50 บนสะเกลด้านราบซึ่งแสดงว่า mean (μ) = 50 และจุด 84 % บนสะเกลด้านตั้งเป็นไปตาม **y** = 60 บนสะเกลด้านราบก็แสดงว่า $\mu + \sigma = 60$ เพราะฉะนั้น $\sigma = (\mu + \sigma) - (\mu) = 60 - 50 = 10$

4.2 **Drawing of Samples**

เมื่อเราเขย่าตะกร้าบรรจุแผ่นกระดาษแข็งกลมซึ่งมีจำนวนเลขเขียนติดไว้ทั้ง 500 แผ่นให้คลุกเคล้าปนกันดีแล้ว เราหยิบแผ่นกระดาษแข็ง (เรียกว่า tag) ขึ้นมาหนึ่งแผ่น แล้วจดบันทึกจำนวนเลขบน tag นั้นไว้ ใส่ tag ที่หยิบขึ้นมาคืนลงตะกร้าตามเดิมแล้วเขย่าตะกร้าให้ tags ปนกันดีแล้วจึงหยิบ tag ขึ้นมาอีกหนึ่งแผ่น จดบันทึกจำนวนเลขบน tag ใหม่นี้ไว้อีก จำนวนเลขแต่ละจำนวนที่จดบันทึกไว้คือ observation ทำโดยวิธีนี้ต่อ ๆ ไปจนได้ 5,000 observations

ในขณะที่เราหยิบ tags ขึ้นมาเราจดบันทึกจำนวนเลขบน tags เหล่านี้ไว้เป็นกลุ่ม กลุ่มละ 5 จำนวน ดังนั้นแต่ละกลุ่มจะมี 5 observations เรียกว่าหนึ่ง sample เพราะฉะนั้นจาก 5,000 observations เราจะได้ 1,000 samples ความประสงค์ที่หยิบ tag ขึ้นมาทีละแผ่นโดยเขย่าให้ปนกันดีก่อนนั้นก็เพื่อให้เชื่อได้ว่าเป็นการหยิบอย่างสุ่มจริง ๆ เพราะการเขย่าตะกร้าให้ tags ทั้งหมดปนกันดีก่อนหยิบ tag แผ่นหนึ่งขึ้นมาจะไม่มี observation ใดตกอยู่ในอิทธิพลของ observation อื่นเลย และถ้าเราไม่ใส่ tag ที่หยิบขึ้นมาคืนลงตะกร้าตามเดิมก่อนที่จะหยิบ tag แผ่นต่อไปแล้ว tag แผ่นนี้ก็จะหมดโอกาสที่จะถูกหยิบขึ้นมาอีก ดังนั้นค่าของ observation ต่อไปจะถูกจำกัด และถึงแม้ว่าเราจะใส่ tag ที่หยิบขึ้นมาคืนลงตะกร้าตามเดิมแล้วก็ตาม ถ้าไม่เขย่าตะกร้าให้ tags ทั้งหมดปนกันดีก่อนจะหยิบ tag แผ่นต่อไปแล้ว tag แผ่นที่ใส่คืนลงไปจะหล่นลงไปอยู่ข้างบน เราอาจจะหยิบมันขึ้นมาอีก ทำให้ observations ที่อยู่ใกล้เคียงกันมีค่าเหมือนกัน เพราะฉะนั้นการใส่คืน tag และการเขย่าตะกร้าให้ tags ปนกันดีก่อนหยิบ tag ทุกครั้งก็เพื่อให้เชื่อได้ว่าในการหยิบ tag ทุกครั้ง tags ของ population ทั้ง 500 แผ่นมีโอกาสเท่าเทียมกันที่จะถูกหยิบขึ้นมา observations ต่าง ๆ ที่ได้มาโดยวิธีนี้เรียกว่า "independent observations" และ sample ที่ประกอบด้วย independent observations เรียกว่า "random sample"

โดยการหยิบ tag ขึ้นมาตามวิธีนี้ samples ต่าง ๆ ซึ่งประกอบด้วย 5 observations ต่างก็จะเป็นอิสระซึ่งกันและกัน การใส่ tag คืนลงตะกร้าและการเขย่าตะกร้าให้ tags ปนกันดีจะป้องกัน observations ของ sample หนึ่งจากอิทธิพลของ observations ของ samples อื่น ๆ theorems ทั้งหมดที่ได้มาจาก sampling experiments และได้ให้ไว้ในบทต่อ ๆ ไปนี้ต้องการ samples ที่เป็น random samples และเป็นอิสระทั้งสิ้น เพราะฉะนั้นจึงเป็นสิ่งสำคัญที่จะต้องทราบ sampling scheme ไว้ในขั้นนี้ด้วย ตัวอย่างของ 4 random samples ซึ่งแต่ละ sample ประกอบด้วย 5 observations ได้แสดงไว้ใน Table 4.2

Table 4.2

	Sample No.	1	2	3	4
explanations appear in	observations (y)	50	55	67	61
		57	44	57	52
		42	37	71	68
		63	40	55	50
	calculation	32	52	46	46
chapter 5	Σy	244	228	296	277
	$\bar{y}$	48.8	45.6	59.2	55.4
chapter 7	$(\Sigma y)^2$	59,536	51,984	87,616	76,729
	$(\Sigma y)^2/n$	11,907.2	10,396.8	17,523.2	15,245.8
	Σy^2	12,506	10,634	17,920	15,665
	SS	598.8	237.2	396.8	319.2
	Σu^2	6.06	3.34	8.20	4.65
chapter 8	s^2	149.7	59.3	99.2	79.8
	s^2/n	29.94	11.86	19.84	15.96
	$\sqrt{s^2/n}$	5.472	3.444	4.454	3.995
	$\bar{y} - \mu$	-1.2	-4.4	9.2	5.4
	t	-0.219	-1.278	2.066	1.352
chapter 9	F		2.52		1.24
	$(SS_1 + SS_2)/\sigma^2$		8.360		7.160
chapter 10	$\bar{y}_1 - \bar{y}_2$		3.2		3.8
	s_p^2		104.5		89.5
	$\sqrt{2s_p^2/n}$		6.465		5.983
	t		0.495		0.635
chapter 11	$\bar{y} - 8.8$	40.0	36.8	50.4	46.6
	$\bar{y} + 8.8$	57.6	54.4	68.0	64.2
	$2.7764\sqrt{s^2/n}$	15.2	9.6	12.4	11.1
	$\bar{y}-2.7764\sqrt{s^2/n}$	33.6	36.0	46.8	44.3
	$\bar{y}+2.7764\sqrt{s^2/n}$	64.0	55.2	71.6	66.5
chapter 12	F		0.24		0.40

4.3 Computation

จำนวนเลขต่าง ๆ เช่นผลบวกของ observations ($\sum y$) ใน sample และ sample mean ($\bar{y}$) (ดู Table 4.2) นั้น ได้คำนวณไว้ทุก sample ทั้ง 1,000 samples แล้ว แต่จำนวนเลขอื่น ๆ ที่คำนวณไว้ใน Table 4.2 จะยังไม่อธิบายในบทนี้ จะเอาไว้อธิบายในบทต่อ ๆ ไปเพราะเราต้องใช้มันพิสูจน์ความเป็นจริงของ theorems ในบทต่าง ๆ ตามที่ได้แสดงไว้ในช่องที่ 1 ของ Table นี้

4.4 Parameter and Statistic

parameter หมายถึงเลขจำนวนหนึ่งซึ่งคำนวณออกมาจาก population เช่น mean, variance และ standard deviation ของ population เป็นต้น ดังนั้น mean 50 และ standard deviation 10 ของ tag population ที่ได้กล่าวมาแล้วในข้อ 4.1 ล้วนแต่เป็น parameter ทั้งนั้น

statistic หมายถึงเลขจำนวนหนึ่งซึ่งคำนวณออกมาจาก sample หนึ่ง เช่น sample mean เป็นต้น ดังนั้น means ของ 4 samples ใน Table 4.2 คือ 48.8, 45.6, 59.2 และ 55.4 ต่างก็เป็น statistic ทั้งสิ้น

เราใช้ μ แทน population mean แต่ใช้ $\bar{y}$ แทน sample mean

parameter เช่น population mean (μ) ก็ดี population variance (σ^2) ก็ดี หรือ population standard deviation (σ) ก็ดี เป็นจำนวนเลขที่มีค่าตายตัว แต่ statistic เช่น sample mean ($\bar{y}$) นั้น เป็นจำนวนเลขที่มีค่าเปลี่ยนไปได้ต่าง ๆ กันตาม samples

จาก population หนึ่งเราอาจหา samples ออกมาได้หลาย samples แต่ละ sample ก็มี mean ($\bar{y}$) ของมัน ค่าของ observations ต่าง ๆ ในแต่ละ sample ไม่เหมือนกัน เพราะฉะนั้น sample mean ($\bar{y}$) จึงมีค่าเปลี่ยนไปตาม samples การขึ้น ๆ ลง ๆ ในค่าของ statistic ตัวหนึ่ง (เช่น sample mean) จาก sample หนึ่งไปยังอีก sample หนึ่งเป็น concept สำคัญอย่างหนึ่ง ตัวอย่างของการขึ้น ๆ ลง ๆ ในค่าของ statistic ตัวหนึ่งดังกล่าวนี้จะเห็นได้จาก Table 4.2 คือ sample means ทั้ง 4 ตัวมีค่าไม่เหมือนกันเลย

4.5 Purpose of Sampling Experiments

ความประสงค์สำคัญที่สุดของสถิติ คือการหา conclusion ที่เกี่ยวกับ population จากเรื่องที่ได้รับจาก sample หนึ่งนั้น จะกระทำได้สำเร็จก็ด้วยการบรรลุถึงความรู้อันเชื่อได้ในความสัมพันธ์ระหว่าง population หนึ่งกับ samples ต่าง ๆ ของมัน ซึ่งไม่ใช่การเดา ขั้นแรกของการพัฒนาวิธีทางสถิติคือการพิจารณากำหนดว่า samples ชนิดอะไรที่เราจะได้จาก population หนึ่ง เมื่อทราบแล้วเราก็หวังว่า sample ที่กำหนดให้หนึ่ง sample นี้จะช่วยให้เราสามารถเข้าถึง conclusion ที่เกี่ยวกับ population นั้นได้

ดังนั้นความประสงค์ในเรื่อง sampling experiment ก็คือการแสดงให้เห็นความสัมพันธ์ระหว่าง population และ samples ต่าง ๆ ของมัน Table 4.2 แสดงให้เห็นว่าไม่มี sample mean ตัวใดใน 4 ตัวมีค่าเท่ากับ 50 เลย แม้ว่า population mean จะมีค่าเท่ากับ 50 ก็ตาม การขึ้น ๆ ลง ๆ ในค่าของ sample means นี้จะเป็นไปตามแบบอย่างอันหนึ่ง ดังจะแสดงไว้ในบทที่ 5

Chapter 5

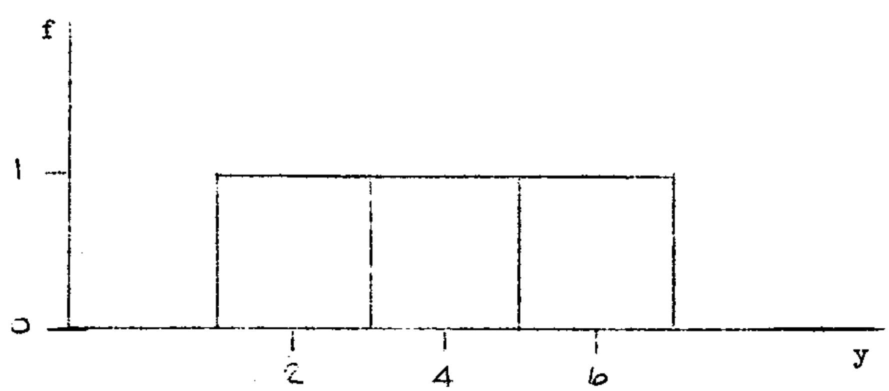
Sample Means

เรื่องสำคัญที่สุดอย่างหนึ่งซึ่งสถิติศาสตร์เข้ามาเกี่ยวข้องด้วยก็คือความสัมพันธ์ระหว่าง population หนึ่งกับ samples ต่าง ๆ ของมัน ความสัมพันธ์ดังกล่าวนี้อาจเกิดจาก population หนึ่งออกไปสู่ samples ต่าง ๆ ของมัน หรือโดยนัยกลับกันอาจเกิดจาก sample หนึ่งเข้ามาสู่ population ต้นกำเนิดของมันก็ได้ บทที่ 5 นี้ เกี่ยวกับความสัมพันธ์จาก population หนึ่งออกไปสู่ samples ต่าง ๆ และจะพิจารณาถึงลักษณะของ means ของ samples ต่าง ๆ ที่ได้มาจาก population หนึ่งที่กำหนดให้ factor สำคัญอย่างหนึ่งในการพิจารณาคือจำนวน observations ใน sample หนึ่งซึ่งเพื่อความสะดวกจะเรียกว่า the size of the sample และใช้แทนด้วยอักษร n ดังนั้นถ้า sample หนึ่งประกอบด้วย 5 observations, the size of the sample คือ 5 หรือ n = 5

5.1 Sampl~~e~~ing Scheme

จาก population หนึ่งที่กำหนดเราอาจหา samples ได้หลาย samples ความสัมพันธ์ระหว่าง population หนึ่งกับ samples ต่าง ๆ ของมันเป็นความสัมพันธ์กับ samples ทั้งหมด ไม่ใช่กับบาง samples ที่ได้จาก population นั้น ทั้งนี้จะไม่คำนึงว่าจำนวน samples ทั้งหมดจะมีมากหรือน้อยเพียงไร เราจะหา samples ทั้งหมดออกมาจาก population หนึ่งได้อย่างไรนั้นจะแสดงให้เห็นโดยตัวอย่างต่อไปนี้

ถ้าเราพิจารณาให้ 3 observations คือ 2, 4 และ 6 เป็น population หนึ่ง histogram หรือกราฟของ frequency table จะมีรูปดังนี้



ถ้า sample หนึ่งประกอบด้วยหนึ่ง observation (n = 1) เราจะได้จำนวน samples ทั้งหมดจาก population นี้เพียง 3 samples เท่านั้น samples ทั้งสามนี้คือ 2, 4 และ 6 และ means ของ 3 samples ก็คือ 2, 4 และ 6 ด้วยเช่นเดียวกัน กล่าวได้ว่าถ้าแต่ละ sample ประกอบด้วยหนึ่ง observation แล้ว mean ของ sample นั้นก็คือตัว observation นั่นเอง ถ้าแต่ละ sample ประกอบด้วย 2 observations (n = 2) เราจะได้ samples ทั้งหมด 9 samples จาก population นี้ observation ตัวที่หนึ่งของ sample หนึ่งอาจเป็น 2, 4 หรือ 6 ตัวใดตัวหนึ่ง เมื่อได้ observation ตัวที่หนึ่งแล้ว observation ตัวที่สองอาจเป็น 2, 4 หรือ 6 ตัวใดตัวหนึ่งได้อีก ดังนั้น 9 samples นี้คือ 2, 2; 2,4; 2,6; 4,2; 4,4; 4,6; 6,2; 6,4; และ 6,6 samples เหล่านี้รวมทั้งผลบวกของ observations (Σy) และ sample mean ($\bar{y}$) ได้แสดงไว้ใน Table 5.1 a แล้ว

Table 5.1 a

1 st obs.	2nd obs.	sample	Σy	$\bar{y}$
	2	2, 2	4	2
2	4	2, 4	6	3
	6	2, 6	8	4
	2	4, 2	6	3
4	4	4, 4	8	4
	6	4, 6	10	5
	2	6, 2	8	4
6	4	6, 4	10	5
	6	6, 6	12	6

จาก Table นี้จะเห็นได้ว่าแต่ละตัวของ 3 observations แรกแตกออกไปเป็น 3 สาขา จำนวนสาขาทั้งหมดจึงมี 3 x 3 หรือ 3^2 หรือ 9 สาขา sampling scheme นี้คล้ายการทอดลูกเต๋ามากกว่าการดึงไพ่ ถ้าเราเขียน observations 2, 4, 6 แยกกันลงบนไพ่ 3 ใบแล้วหา samples ต่าง ๆ ที่ประกอบด้วยไพ่ 2 ใบออกมาก็จะมีเพียง 3 possible samples เท่านั้น แทนที่จะเป็น 9 samples เหล่านี้คือ 2,4;

4,6 ; 2,6 ที่เป็นเช่นนี้ก็เพราะว่าเมื่อเราดึงไพ่ออกมา 2 ใบจะเหลือไพ่อยู่หนึ่งใบเสมอ เนื่องจากการเหลือไพ่หนึ่งใบนี้มีทางที่จะเป็นได้อยู่ 3 ทางจึงต้องมี possible samples ที่ประกอบด้วยไพ่ 2 ใบเพียง 3 samples เท่านั้น เช่นถ้าดึงไพ่ observations 2 และ 4 ออกมาก็จะเหลือไพ่ observation 6 ถ้าดึงไพ่ observations 4 และ 6 ออกมาก็จะเหลือไพ่ observation 2 และถ้าดึงไพ่ observations 2 และ 6 ออกมาก็จะเหลือไพ่ observation 4 ดังนี้

แต่ในการทอดลูกเต๋าสถานะการณ์จะแตกต่างไปจากการดึงไพ่ กล่าวคือในการทอดครั้งที่หนึ่งแต้มของลูกเต๋าที่ขึ้นหรือ observation ที่หนึ่งอาจเป็น 1, 2, 3, 4, 5 หรือ 6 แต้มใดแต้มหนึ่ง เมื่อจดบันทึก observation ที่หนึ่งแล้ว ในการทอดครั้งที่สอง observation ที่สองก็อาจเป็น 1, 2, 3, 4, 5 หรือ 6 ได้อีก ดังนั้น combinations 1,1; 2,2; ฯลฯ จึงเป็นไปได้ที่จะเกิดขึ้น ผลลัพธ์ก็คือจะได้ 6 x 6 หรือ 6^2 หรือ 36 possible samples <u>เราเรียก sampling scheme แบบทอดลูกเต๋านี้ว่า sampling with replacement</u> ซึ่งจะนำมาใช้ในวิชานี้ อนึ่ง ควรจะทราบว่า sampling scheme แบบนี้ทำให้เชื่อแน่ได้ว่ามีความอิสระในระหว่าง observations ต่าง ๆ (ดูข้อ 4.2) observation ทุกตัวของ population ยังมีโอกาสเท่าเทียมกันที่จะถูกเอาออกมาในการ draw ครั้งที่สองโดยไม่ต้องคำนึงว่า observation ที่หนึ่งของ sample นั้นจะเป็นอะไร

ถ้าเราใช้ sampling scheme แบบทอดลูกเต๋า จาก population ที่ประกอบด้วย 3 observations คือ 2, 4 และ 6 นั้น เมื่อ n = 3 เราจะได้ 3^3 หรือ 27 possible samples ดังแสดงไว้ใน Table 5.1 b

Table 5.1 b

1st. obs.	2nd. obs.	3rd. obs.	Sample	Σy	$\bar{y}$
		2	2, 2, 2	6	2.00
	2	4	2, 2, 4	8	2.67
		6	2, 2, 6	10	3.33
		2	2, 4, 2	8	2.67
2	4	4	2, 4, 4	10	3.33
		6	2, 4, 6	12	4.00
		2	2, 6, 2	10	3.33
	6	4	2, 6, 4	12	4.00
		6	2, 6, 6	14	4.67
		2	4, 2, 2	8	2.67
	2	4	4, 2, 4	10	3.33
		6	4, 2, 6	12	4.00
		2	4, 4, 2	10	3.33
4	4	4	4, 4, 4	12	4.00
		6	4, 4, 6	14	4.67
		2	4, 6, 2	12	4.00
	6	4	4, 6, 4	14	4.67
		6	4, 6, 6	16	5.33
		2	6, 2, 2	10	3.33
	2	4	6, 2, 4	12	4.00
		6	6, 2, 6	14	4.67
		2	6, 4, 2	12	4.00
6	4	4	6, 4, 4	14	4.67
		6	6, 4, 6	16	5.33
		2	6, 6, 2	14	4.67
	6	4	6, 6, 4	16	5.33
		6	6, 6, 6	18	6.00

observations 2, 4, 6 แต่ละตัวมี 3 สาขา และแต่ละสาขาก็มี 3 สาขาย่อยลงไปอีก จึงทำให้มี $3 \times 3 \times 3$ หรือ 3^3 หรือ 27 samples และถ้า $n = 4$ แล้วก็จะมี 3^4 หรือ 81 samples เราควรระลึกไว้ว่า 3 เป็นจำนวน observations ใน population และเลขกำลัง 4 เป็นจำนวน observations ใน sample หนึ่ง กล่าวโดยทั่วไป จำนวน samples ทั้งหมดคือ N^n ในเมื่อ N เป็นจำนวน observations ใน population และ n เป็นจำนวน observations ใน sample หนึ่ง หรือ the size of the sample

จำนวน N^n จะเพิ่มขึ้นตามค่าของ N หรือ n อย่างรวดเร็วมาก ใน sampling experiments ที่ได้บรรยายมาแล้วในบทที่ 4 นั้นเรามี N = 500 และ n = 5 เลขสองจำนวนนี้ไม่ใช่เลขที่มีค่ามากเลย แต่จำนวน samples ทั้งหมดจะมีถึง $(500)^5$ หรือ 31,250,000,000,000 samples

5.2 Distribution of Sample Means

ข้อที่แล้วมาแสดงว่าถ้าเราใช้ sample size เท่ากับ 2 เราจะ draw 9 possible samples ออกมาจาก population 2, 4, 6 ได้ และถ้าใช้ samples size เท่ากับ 3 ก็จะได้ 27 possible samples เราอาจคำนวณ sample mean ($\bar{y}$) สำหรับแต่ละ sample ของ samples เหล่านี้ได้ sample means สองชุดดังกล่าวนี้ได้ให้ไว้ใน Table 5.1 a และ 5.1 b ตามลำดับแล้ว สำหรับกรณี n = 4 จะมี 3^4 หรือ 81 sample means และสำหรับกรณี n = 8 จะมีถึง 3^8 หรือ 6,561 sample means (สำหรับกรณี n = 4 และ n = 8 นั้นไม่ได้แสดงให้เห็นตัว samples ไว้ในหนังสือนี้ แต่นิสิตอาจจะพิจารณาหาได้โดยทำเป็นแบบฝึกหัด) frequency tables ของ sample means ต่าง ๆ สำหรับ n = 1, 2, 4 และ 8 ได้แสดงไว้ใน Table 5.2

Table 5.2

n = 1		n = 2		n = 4		n = 8	
$\bar{y}$	f	$\bar{y}$	f	$\bar{y}$	f	$\bar{y}$	f
2	1	2	1	2.0	1	2.00	1
						2.25	8
				2.5	4	2.50	36
						2.75	112
		3	2	3.0	10	3.00	266
						3.25	504
				3.5	16	3.50	784
						3.75	1,016
4	1	4	3	4.0	19	4.00	1,107
						4.25	1,016
				4.5	16	4.50	784
						4.75	504
		5	2	5.0	10	5.00	266
						5.25	112
				5.5	4	5.50	36
						5.75	8
6	1	6	1	6.0	1	6.00	1
No. of Samples	3		9		81		6,561

และ histograms ก็ได้แสดงไว้ใน Fig. 5.2 a, b, c, d แล้ว

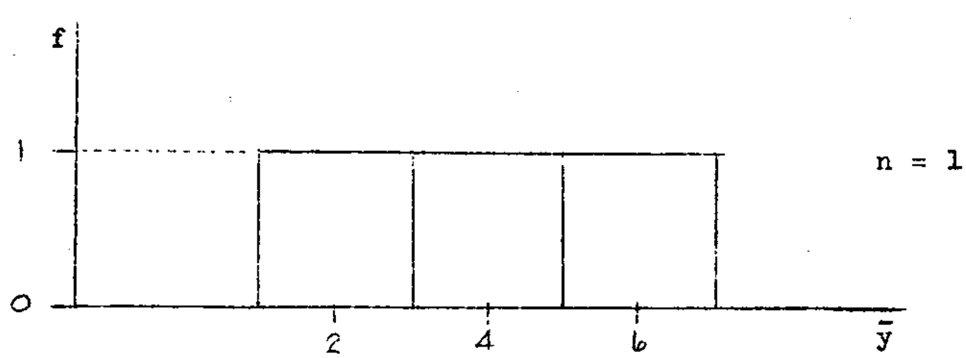


Fig. 5.2 a

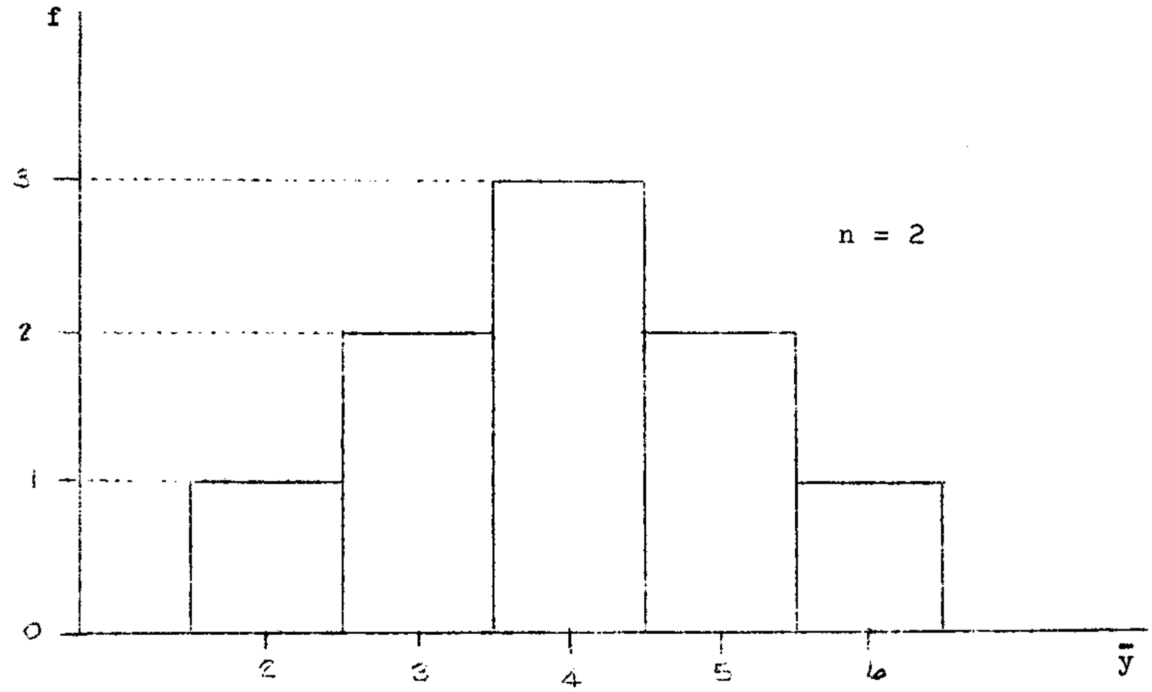


Fig. 5.2 b

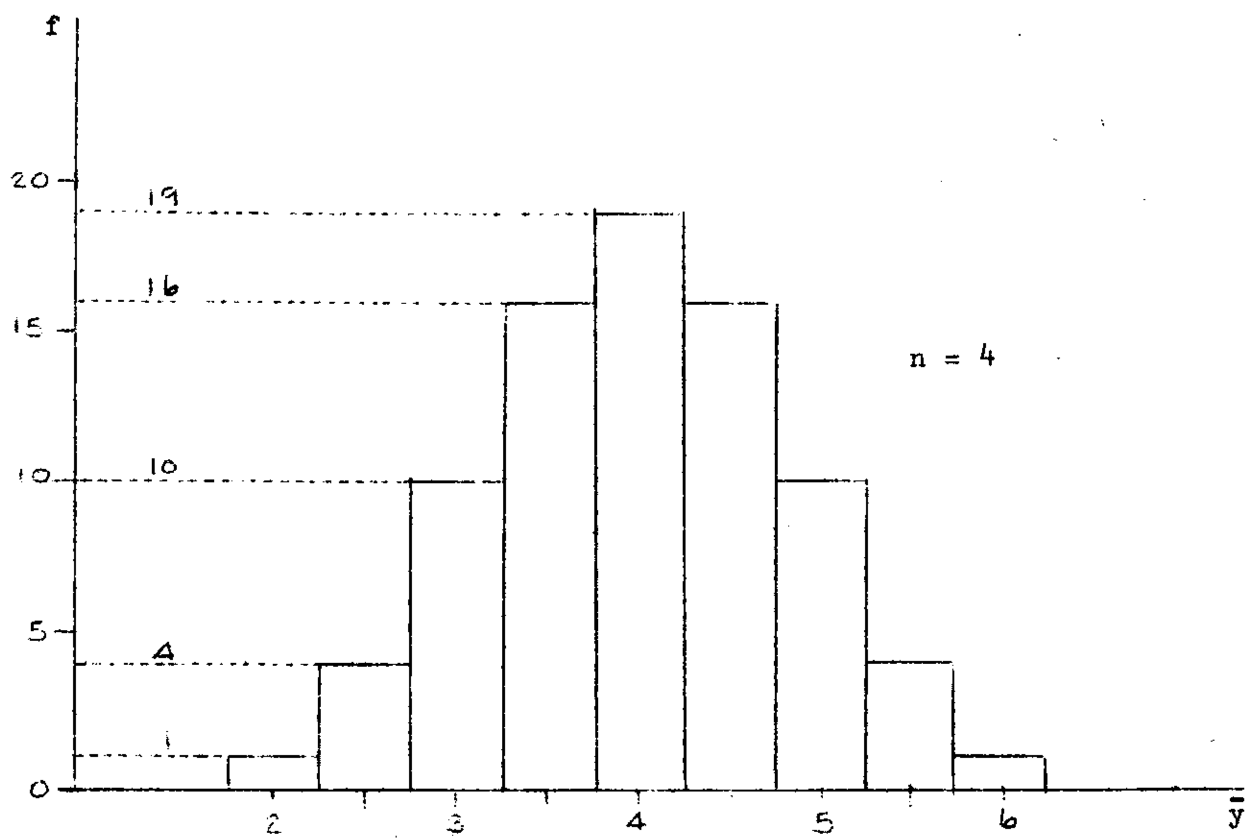


Fig. 5.2 c

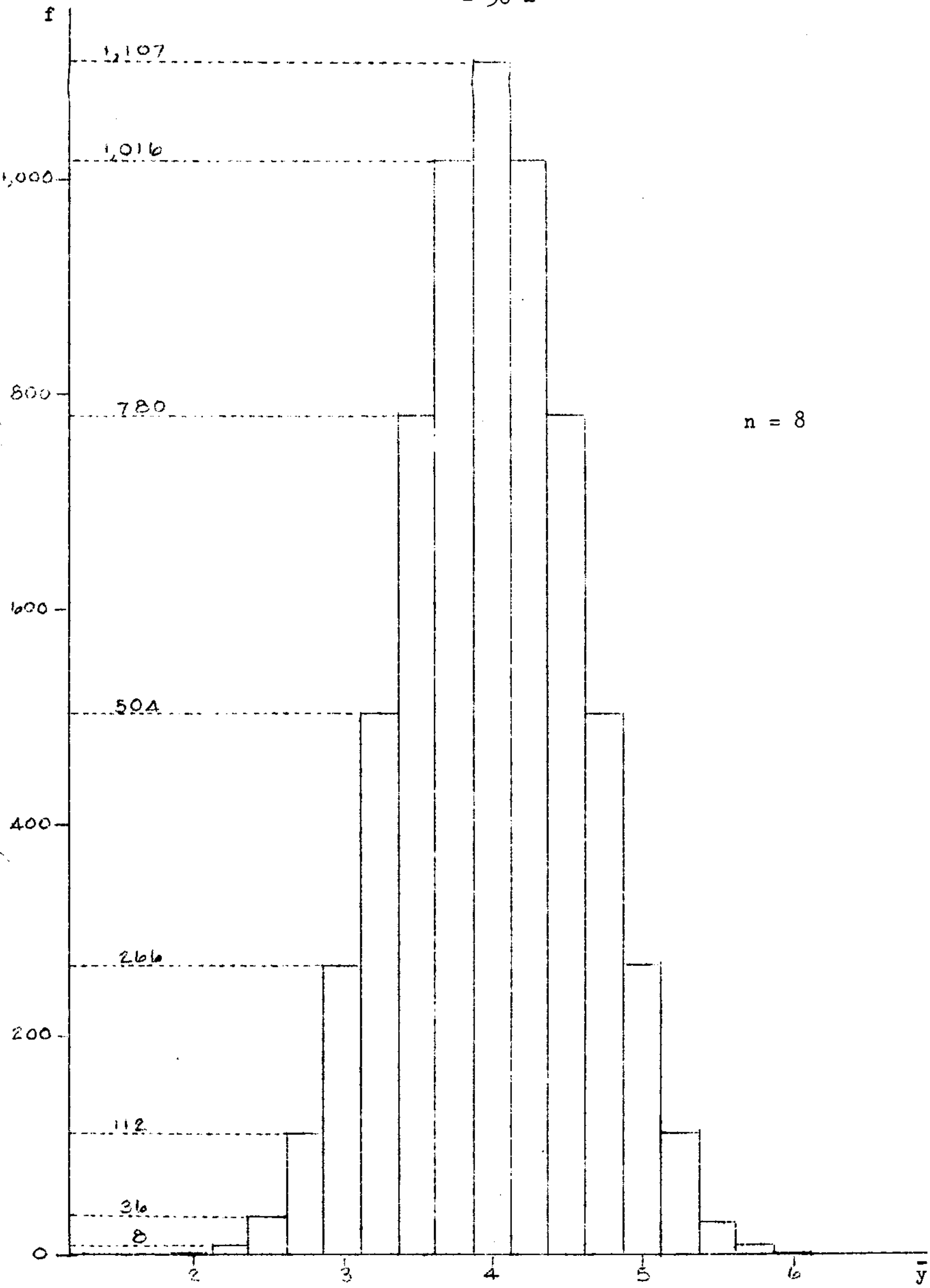

f
1,107
1,016
1,000
800
780
n = 8
600
504
400
266
200
112
36
8
0
2
3
4
5
6
$\bar{y}$

Fig. 5.2 d

จาก histograms เหล่านี้เราจะสังเกตได้ว่าเมื่อเราเพิ่มค่าของ n มากขึ้นรูปของ histogram จะคล้าย normal curve เข้าไปทุกที ปรากฏการณ์นี้จะได้นำมากล่าวไว้ใน Theorem 5.2 a

<u>Theorem 5.2 a:</u> As the size of the sample increases, the distribution of the means of all possible samples of the same size drawn from the same population becomes more and more like a normal distribution provided that the population has a finite variance.

theorem นี้เรียกว่า Central Limit Theorem ซึ่งใช้กับ population ใด ๆ ก็ได้ แต่ถ้า population เป็น normal แล้ว สถานะดังกล่าวนี้ก็จะง่ายขึ้นและผลลัพธ์จะเป็นดังที่ได้แสดงไว้ใน Theorem 5.2 b ต่อไปนี้

<u>Theorem 5.2 b:</u> If the population is normal, the distribution of sample means follows the normal distribution exactly, regardless of the size of the sample.

เราจะแสดงให้เห็นความเป็นจริงของ Theorem 5.2 b โดย sampling experiment ในข้อ 5.6

5.3 Mean and Variance of Sample Means

ในข้อที่แล้วเราได้พิจารณารูปร่างของ distribution curve ของ sample means กันไปแล้ว ในข้อนี้เราจะพิจารณาเรื่อง mean และ variance ของ sample means เหล่านี้บ้าง เราจะใช้ population ที่มี 3 observations คือ 2, 4 และ 6 เป็นตัวอย่างอีกครั้งหนึ่ง จาก population นี้ mean ของ population คือ 4 และ variance ของ population คือ

$$\sigma^2 = \frac{(2-4)^2 + (4-4)^2 + (6-4)^2}{3}$$

$$= \frac{4+0+4}{3}$$

$$= \frac{8}{3}$$

ถ้า sample size เป็น 2 จะมี 9 possible samples ดังแสดงไว้ใน Table 5.1 a frequency table ของ 9 sample means ได้ให้ไว้ใน Table 5.2 รายละเอียดของการคำนวณ mean และ

variance ของ sample means เหล่านี้ได้แสดงไว้ใน Table 5.3

Table 5.3

$\bar{y}$	f	$\bar{y}f$	$(\bar{y}-\mu)$	$(\bar{y}-\mu)^2$	$(\bar{y}-\mu)^2 f$
2	1	2	– 2	4	4
3	2	6	– 1	1	2
4	3	12	0	0	0
5	2	10	1	1	2
6	1	6	2	4	4
sum	9	36			12

$$\mu_{\bar{y}} = \frac{36}{9} = 4 = \mu \;;\quad \sigma^2_{\bar{y}} = \frac{12}{9} = \frac{4}{3} = \frac{8/3}{2} = \frac{\sigma^2}{n}$$

mean ของ sample means ทั้งหมดเท่ากับ $\frac{36}{9}$ หรือ 4 ซึ่งเท่ากับ population mean variance ของ sample means ทั้งหมดเท่ากับ $\frac{12}{9}$ หรือ $\frac{4}{3}$ ซึ่งเท่ากับ population variance $\frac{8}{3}$ หารด้วย sample size 2 mean และ variance ของ sample means ใช้แทนด้วย $\mu_{\bar{y}}$ และ $\sigma^2_{\bar{y}}$ ตามลำดับ ตัวอย่างที่ให้ไว้ในที่นี้จึงชี้ให้เห็นความจริงของ Theorem 5.3 คือ

<u>Theorem 5.3:</u> The mean of the means of all possible samples of the same size drawn from the same population is equal to the mean of that population; that is,

$$\mu_{\bar{y}} = \mu \; . \;.................. \; (1)$$

The variance of these sample means is equal to population variance divided by the size of the sample; that is,

$$\sigma^2_{\bar{y}} = \frac{\sigma^2}{n} \quad .... \; (2)$$

Equation (1) ให้ความสัมพันธ์ระหว่าง mean ของ sample means ทั้งหมดกับ mean ของ population, Equation (2) ให้ความสัมพันธ์ระหว่าง variance ของ sample means ทั้งหมดกับ variance ของ population ความสัมพันธ์ระหว่าง standard deviation ของ sample means ทั้งหมดกับ standard deviation ของ population ก็จะหาได้โดยถอด square root ทั้งสองข้างของ Equation (2) ผลลัพธ์ที่ได้คือ

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

ซึ่งพูดได้ว่า the standard deviation of the means of all possible samples of the same size drawn from the same population is equal to the population standard deviation devided by the square root of the size of the sample. standard deviation ของ means ของ all possible samples ซึ่งมี size เดียวกันและได้จาก population เดียวกันนี้ เรียกว่า <u>standard error of the mean</u>

5.4 Notations

นับแต่ต้นมาจนถึงขณะนี้เราได้พิจารณาเรื่อง means ไปแล้ว 3 ชนิดคือ

1. population mean
2. sample mean
3. mean of sample means

และ variances 2 ชนิดได้แก่

1. population variance
2. variance of sample means

ต่อไปเราจะเพิ่ม sample variance (s^2) เข้ามาอีกชนิดหนึ่ง เพื่อที่จะไม่ให้เกิดความสับสนขึ้นในภายหน้า เราจะใช้อักษรต่าง ๆ กันแทน means, variances, และ standard deviations เหล่านี้ ดังแสดงไว้ในตารางต่อไปนี้

	mean	variance	standard deviation	number of items
population	μ	σ^2	σ	N (size of population)
sample	$\bar{y}$	s^2	s	n (size of sample)
distribution of sample means	$\mu_{\bar{y}}$	$\sigma^2_{\bar{y}}$	$\sigma_{\bar{y}}$	N^n (No. of all possible samples)

5.5 Reliability of Sample Means

standard error of the mean ($\sigma_{\bar{y}}$) หรือ variance of sample means ($\sigma^2_{\bar{y}}$) ต่างก็ใช้เป็นมาตราวัด reliability ของ sample means ได้ คำ "reliability" นี้มีความหมายหลายอย่าง แต่เมื่อใช้กับ sample means แล้วจะหมายถึงความใกล้ชิดของ sample means กับ population mean เนื่องจาก mean ของ sample means ทั้งหมดเท่ากับ population mean sample means ต่าง ๆ จึงเกาะเป็นกลุ่มรอบ population mean และเนื่องจาก variance ของ sample means วัดการแปรผันในระหว่าง sample means การลด variance ของ sample means ลงจะทำให้ sample means ต่าง ๆ โอบรอบ population mean ใกล้ชิดยิ่งขึ้น ความประสงค์ที่มี sample mean ก็เพื่อจะประมาณค่า population mean ดังนั้น variance ของ sample means จึงควรจะลดลงเมื่อสามารถจะทำได้ แต่เนื่องจาก

$$\sigma^2_{\bar{y}} = \frac{\sigma^2}{n}, \quad \ldots\ldots\ldots\ldots\ldots (1)$$

การลด variance ของ sample means จึงทำได้โดยการเพิ่ม sample size (n) หรือลด population variance (σ^2) อย่างใดอย่างหนึ่ง หรือทั้งเพิ่ม sample size (n) และลด population variance (σ^2) ทั้งสองอย่างก็ได้

ควรเป็นที่สังเกตว่าข้อความ "reliability of sample means" หมายถึง sample means รวมกันทั้งหมดมากกว่าจะหมายถึง sample mean ตัวหนึ่งโดยเฉพาะ Table 5.2 แสดงให้เห็น distributions ของ sample means จาก population ซึ่งประกอบด้วย 3 observations คือ

2, 4 และ 6 ที่มี sample sizes ต่าง ๆ กัน ในทุกกรณีไม่ว่า n จะเท่ากับ 1, 2, 4 หรือ 8 จะมี sample mean ตัวหนึ่งของ sample means ทั้งหมดมีค่าเท่ากับ 2 (ดู Table 5.2) sample mean ตัวนี้ประมาณค่าของ population mean 4 ต่ำไปในทางน้อยที่สุด แต่ในขณะที่ sample size เพิ่มขึ้น relative frequency ของ sample mean 2 จะลดลงจาก $\frac{1}{3}$ สำหรับ n = 1 ไปถึง $\frac{1}{6,561}$ สำหรับ n = 8 หรืออีกนัยหนึ่ง sample means ที่ไม่พึงปรารถนาดังกล่าวนี้จะถูก draw ออกมาได้น้อยมากเมื่อ sample size เพิ่มขึ้น อย่างไรก็ดีมันอาจเป็นไปได้ว่า sample mean ตัวหนึ่งโดยเฉพาะที่ได้จากหนึ่ง observation อาจมีค่าใกล้ population mean ได้มากกว่า sample mean ตัวหนึ่งโดยเฉพาะที่ได้จาก 8 observations เราจะสังเกตได้จาก Table 5.2 ว่า sample mean 4 ในกรณี n = 1 ให้ค่าประมาณอันสมบูรณ์ของ population mean ในขณะที่ sample mean 2 ในกรณี n = 8 เป็นค่าประมาณต่ำที่สุดซึ่งเราอาจจะได้มันมาก็ได้

5.6 Experimental Verification of Theorems

ในข้อนี้จะพูดถึงการสอบโดยการทดลองให้เห็นความเป็นจริงของ theorems ต่าง ๆ ที่ได้กล่าวมาแล้ว Theorem 5.3 ชี้ให้เห็นว่า mean ของ all possible sample means ซึ่งมี size เดียวกันซึ่งได้จาก population ใด ๆ จะเท่ากับ mean ของ population นั้น และ variance ในระหว่าง sample means จะเท่ากับ population variance หารด้วย sample size, Theorems 5.2 a และ 5.2 b ยังชี้ให้เห็นชัดขึ้นอีกว่า distribution ของ sample means ที่ได้จาก population ที่ไม่เป็น normal จะเข้าไปใกล้ normal distribution ทุกทีในขณะที่ sample size เพิ่มขึ้น และ distribution ของ sample means ที่ได้จาก normal population จะเป็น normal distribution เสมอ โดยไม่คำนึงถึง sample size ว่าจะเป็นเท่าไร การบอกเป็นนัยของ theorems เหล่านี้ก็คือ relative frequency ของ sample means ภายในช่วงใดช่วงหนึ่งโดยเฉพาะจะถูกทราบก่อนที่จะ draw sample หนึ่งออกมาจาก population ที่กำหนดให้จึงทำให้ทราบ mean และ variance ของ population นั้น ตัวอย่างเช่นถ้าเรา draw all possible samples ซึ่งมี size 5 ออกมาจาก normal population ที่มี $\mu = 50$ และ $\sigma = 10$ distribution ของ sample means จะเป็น normal distribution อย่างแท้จริงซึ่งมี mean ($\mu_{\bar{y}}$) เท่ากับ 50 และ standard deviation ($\sigma_{\bar{y}}$) เท่ากับ $\frac{10}{\sqrt{5}}$ หรือ 4.47 และ 95 % ของ sample means จะตกอยู่ภายในช่วง 50 − 1.96 (4.47) ถึง 50 + 1.96 (4.47) หรือภายในช่วง 41.2 ถึง 58.8 เราจะสอบให้เห็นความเป็นจริงในเรื่องเหล่านี้ได้โดย sampling experiment ที่ได้บอกรายละเอียดไว้แล้วในบทที่ 4 กล่าวโดยย่อคือ

experiment นี้ประกอบด้วย 1,000 random samples ซึ่งแต่ละ sample มี 5 observations ได้มาจาก normal population หนึ่งที่มี mean เท่ากับ 50 และ standard deviation เท่ากับ 10 จากทุก sample เราคำนวณ sample mean ไว้ frequency distribution ของ 1,000 sample means ได้แสดงไว้ใน Table 5.6

Table 5.6

sample means	theoretical r.f. (%)	observed r.f. (%)	observed r.c.f. (%)
below 39.5	1.0	0.7	0.7
39.5 to 42.5	3.7	4.0	4.7
42.5 to 45.5	11.0	12.6	17.3
45.5 to 48.5	21.2	19.5	36.8
48.5 to 51.5	26.2	24.2	61.0
51.5 to 54.5	21.2	22.2	83.2
54.5 to 57.5	11.0	12.3	95.5
57.5 to 60.5	3.7	3.7	99.2
above 60.5	1.0	0.8	100.0
	100.0	100.0	

theoretical relative frequencies ในตารางนี้เป็น relative frequencies ถ้า draw all possible samples ซึ่งมี size 5 ออกมาจาก population และเป็น frequencies ที่ได้มาจาก table ของ normal distribution หนึ่งเช่น Table 3 in the Appendix ส่วน observed relative frequencies เป็น relative frequencies ที่ได้จากการ draw 1,000 samples เท่านั้น ตัวอย่างเช่น observed relative frequency สำหรับ class 39.5 ถึง 42.5 เป็น 4 % ซึ่งแสดงว่า 40 จาก 1,000 sample means ตกอยู่ภายในช่วง 39.5 ถึง 42.5 และใน class เดียวกันนี้ theoretical relative frequency จะเป็น 3.7 %

Table 5.6 แสดงให้เห็นว่า theoretical และ observed relative frequencies มีค่าใกล้เคียงกันมาก แต่มันก็ไม่เป็นค่าเดียวกันอย่างแท้จริง เพราะว่า theoretical relative frequency มีมูลฐานจาก samples ทั้งหมด ในขณะที่ observed relative frequency มีมูลฐานจากบาง samples เท่านั้น (ในกรณีนี้คือ 1,000 samples)

เมื่อเราพล็อต observed relative cumulative frequencies ใน Table 5.6 ลงบน normal probability graph paper จะได้ดัง Fig. 5.6

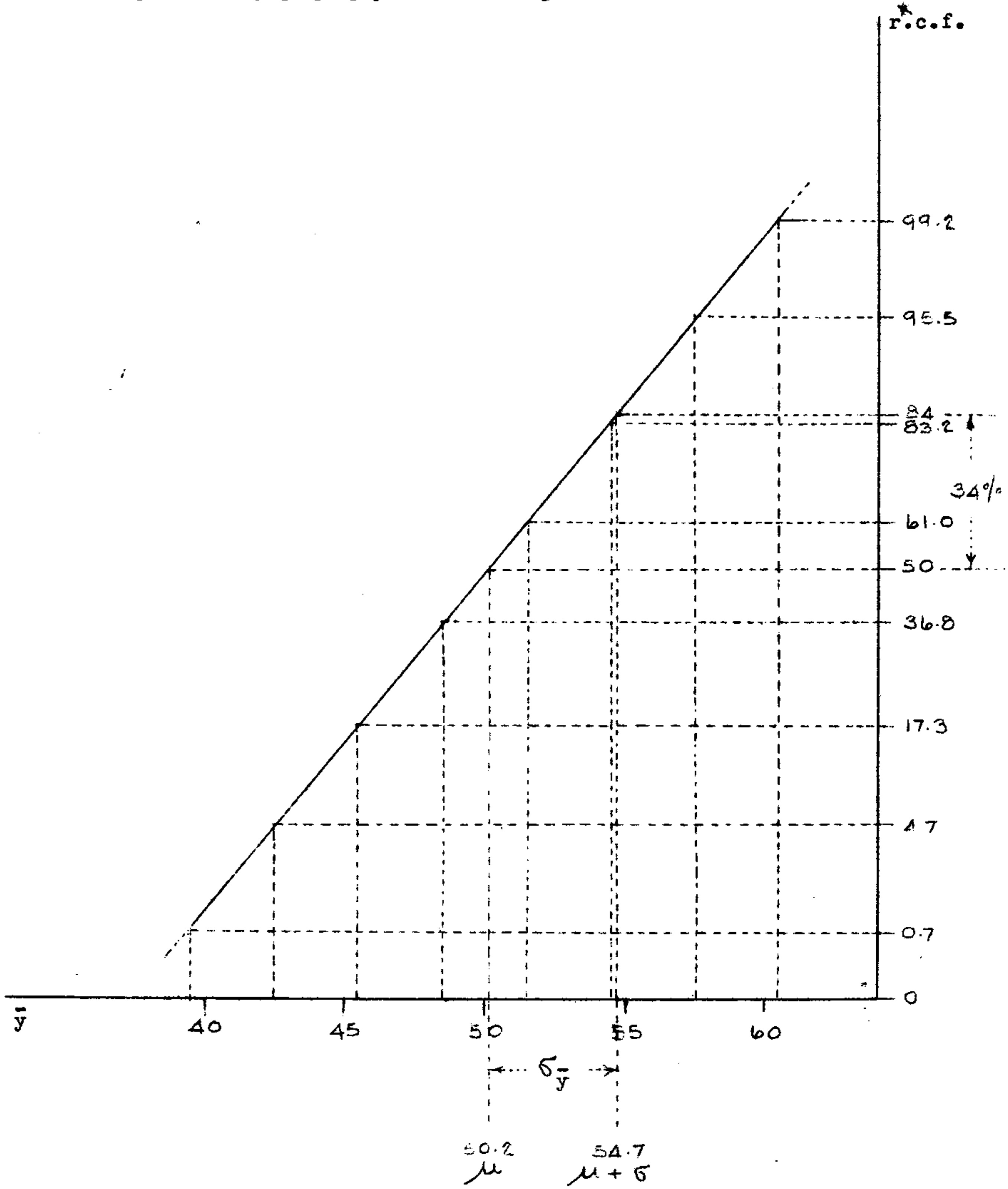


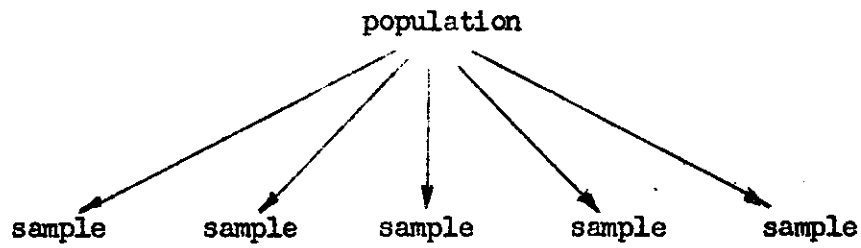
Fig. 5.6

การที่จุดต่าง ๆ ที่พล็อตเกือบจะอยู่บนเส้นตรงเดียวกันนั้นแสดงว่า **distribution** ของ 1,000 **sample means** เป็น **normal distribution** โดยประมาณ **mean** ของ **sample means** ที่อ่านได้จากกราฟคือ 50.2 เปรียบเทียบกับ **mean** ของ **sample means** ทางทฤษฎีคือ 50 และค่าของ $\bar{y}$ บนสะเกลด้านราบตามจุด 84 % ของ **r.c.f.** คือ 54.7 เพราะฉะนั้น **standard deviation** ของ **sample means** เท่ากับ 54.7 – 50.2 หรือ 4.5 เปรียบเทียบกับ **standard deviation** ของ **sample means** ทางทฤษฎีคือ $\frac{\sigma}{\sqrt{n}}$ หรือ $\frac{10}{\sqrt{5}}$ หรือ 4.47 การสอบความเป็นจริงของ **theorems** ต่าง ๆ จึงสมบูรณ์แล้วทุกประการ

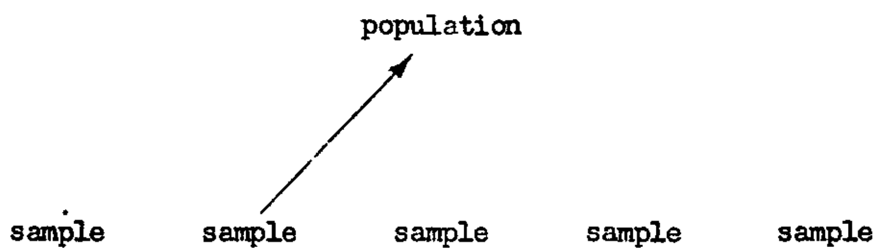
Chapter 6

Test of Hypothesis

ตามที่กล่าวมาทั้งหมดได้แสดงให้เห็นแล้วว่าสถิติศาสตร์เกี่ยวข้องกับความสัมพันธ์ระหว่าง population หนึ่งกับ samples ของมัน ความสัมพันธ์ดังกล่าวนี้อาจเป็นหนึ่งในสองชนิดตามแนวทางของความสัมพันธ์ กล่าวคือ ถ้าแนวทางของความสัมพันธ์ออกจาก population ไปสู่ sample ดังรูป



เราอาจทราบ all possible samples ซึ่งได้จาก population หนึ่งที่กำหนดให้ได้ กระบวนการดังว่านี้เรียกว่า "deduction" ซึ่งเป็นกระบวนการของการหาเหตุผลจากเรื่องทั่วไปไปสู่เรื่องเฉพาะ ในทางตรงกันข้ามถ้าแนวทางของความสัมพันธ์ออกจาก sample หนึ่งเข้ามาสู่ population ของมัน ดังรูป



เราอาจลงความเห็นหรือตัดสินในเรื่องที่เกี่ยวกับ population ได้จากมูลฐานที่ว่าเราทราบอะไรที่เกี่ยวกับ sample แล้วบ้าง กระบวนการดังว่านี้เรียกว่า "induction" ซึ่งเป็นกระบวนการของการหาเหตุผลจากเรื่องเฉพาะไปสู่เรื่องทั่วไป

กล่าวได้อีกอย่างหนึ่งว่าถ้าเราทราบ sample หนึ่งและพยายามจะบอกลักษณะของ population โดยอาศัยความรู้จาก sample เรียกว่า เรากำลังหาเหตุผลโดยกระบวนการ induction แต่ถ้าเราทราบ

population และพยายามจะบอกลักษณะของ all possible samples โดยอาศัยความรู้จาก population เรียกว่า เรากำลังหาเหตุผลโดยกระบวนการ deduction ความแตกต่างกันระหว่างกระบวนการทั้งสองนี้อาจจำได้ง่ายโดยให้นึกถึงคำ prefix "de" ซึ่งหมายความถึง "จาก" หรือ "ออกจาก" (population) และคำ prefix "in" ซึ่งหมายความถึง "ใน" หรือ "เข้ามาข้างใน" (population)

เนื้อหาของบทที่ 5 ได้เน้นหนักในเรื่อง deduction คือ การหาเหตุผลจาก population ไปสู่ all possible samples ในบทที่ 6 นี้จะเน้นหนักในเรื่อง induction คือการหาเหตุผลจาก sample หนึ่งไปสู่ population ต้นกำเนิดของมัน กล่าวให้ชัดลงไปอีกก็คือบทที่ 5 นั้นเกี่ยวกับลักษณะของ means ทั้งหลายของ all possible samples ซึ่งมี size เดียวกันและได้มาจาก population หนึ่งที่กำหนดให้ แต่บทที่ 6 นี้ เกี่ยวกับการ draw conclusion ที่เกี่ยวกับ population โดยอาศัยความรู้ที่ได้จาก sample เดียวซึ่งประกอบด้วย observations เพียงบางตัวไม่ใช่ observations ทั้งหมดของ population นั้น

6.1 Hypothesis

สมมติฐาน (hypothesis) คือ contention ซึ่งมีรากฐานบนการสังเกตเบื้องต้นของสิ่งที่ปรากฏเป็นข้อเท็จจริงซึ่งอาจจะจริงหรือไม่จริงก็ได้

การทดสอบสมมติฐาน (the test of hypothesis) ก็คือการเปรียบเทียบ contention ที่ตั้งไว้กับข้อเท็จจริงที่รวบรวมขึ้นใหม่อันตรงกับวัตถุประสงค์

ถ้าข้อเท็จจริงที่รวบรวมขึ้นใหม่ได้ถูกแสดงให้เห็นว่าสอดคล้องหรือพ้องกับ contention แล้ว contention นั้นก็ใช้ได้ คือเรายอมรับสมมติฐาน ถ้า contention กับข้อเท็จจริงไม่พ้องกันก็จะต้องตัด contention ทิ้งไป คือ เราปฏิเสธสมมติฐานนั้น

ในเรื่องสมมติฐานนี้อาจจะมีสมมติฐานเป็นทางเลือก (alternative hypotheses) ได้หนึ่งหรือสองอย่าง เช่น เราตั้งสมมติฐานว่าผู้ชายและผู้หญิงเป็นคนขับรถที่ดีเท่ากัน ในกรณีสอง alternative hypotheses, alternative hypothesis อย่างหนึ่งคือ โดยเฉลี่ยแล้วผู้ชายเป็นคนขับรถที่ดีกว่า และ alternative hypothesis อีกอย่างหนึ่งคือ ผู้หญิงเป็นคนขับรถที่ดีกว่า ภายหลังที่สมมติฐานถูกตรวจสอบกับข้อเท็จจริงแล้ว และถ้าหลักพยานมีเหตุผลดีและเพียงพอเราอาจยอมรับสมมติฐานและได้ conclusion ว่าโดยเฉลี่ยผู้ชายและผู้หญิงเป็นคนขับรถที่ดีเท่ากัน แต่ถ้าเราปฏิเสธสมมติฐาน เราอาจจะเลือกเอาหนึ่งในสองของ alternative hypotheses และได้ conclusion ว่าผู้ชายเป็นคนขับรถที่ดีกว่า หรือผู้หญิงเป็นคนขับรถที่ดีกว่า ทั้งนี้แล้วแต่หลักพยานที่ได้สังเกตไว้ อย่างไรก็ดี ถ้าเป็นที่ทราบกันก่อนแล้วว่าโดยเฉลี่ยผู้หญิง

จะเป็นคนขับรถที่ดีกว่าผู้ชายไม่ได้ สมมติฐานยังคงแสดงว่าผู้ชายและผู้หญิงเป็นคนขับรถที่ดีเท่ากัน ดังนั้น **alternative hypothesis** จึงมีเพียงอย่างเดียวคือผู้หญิงเป็นคนขับรถที่เลวกว่าผู้ชาย การตั้ง **alternative hypothesis** ว่าจะเป็นหนึ่งหรือสองอย่างจึงต้องเป็นไปตามข้อเท็จจริงที่ทราบไม่ใช่เป็นไปตามความเคียดฉันท์

6.2 Two Kinds of Errors

ถ้าเราทราบข้อเท็จจริงทั้งหมด (หมายถึงทราบ population) โดยละเอียดแล้วการตรวจสอบ **contention** กับข้อเท็จจริงจะเป็นเรื่องง่าย ๆ ธรรมดา แต่มันจะเป็นปัญหาสำคัญยิ่งถ้าเราทราบข้อเท็จจริงแต่เพียงบางส่วน (หมายถึงทราบ sample) เท่านั้น ความประสงค์ของการทดสอบสมมติฐานก็เพื่อตรวจสอบ **contention** กับข้อเท็จจริงบางส่วนนั้น เพราะฉะนั้น **conclusion** ที่ได้จึงไม่ถูกต้องเสมอไป คือ อาจจะเกิด **error** ชนิดหนึ่งในสองชนิดซึ่งเรียกว่า **Type I error** และ **Type II error** ขึ้นได้ ตัวอย่างของ **errors** 2 ชนิดอาจแสดงให้เห็นได้ดังต่อไปนี้

สมมติว่านาย ก เป็นเพื่อนนาย ข ทั้งสองคนชวนกันไปดื่มกาแฟแล้วโยนเหรียญเสี่ยงทายว่าใครจะเป็นผู้จ่ายค่ากาแฟ สมมติต่อไปว่าการโยนเหรียญวันแรกนาย ก เป็นผู้แพ้และต้องจ่ายค่ากาแฟ วันที่สองนาย ก ก็เป็นฝ่ายแพ้อีก แต่นาย ก อาจไม่สงสัยว่านาย ข เล่นโกง เพราะการที่นาย ก เล่นพนันกับนาย ข นั้น แสดงว่านาย ก เชื่อนาย ข หรือกล่าวได้อีกอย่างหนึ่งว่า นาย ก เล่นพนันบนสมมติฐานที่ว่านาย ข เป็นคนซื่อ แต่ถ้านาย ก เป็นฝ่ายแพ้พนันและต้องจ่ายค่ากาแฟถึง 10 ครั้งติดต่อกันแล้ว นาย ก อาจสงสัยว่านาย ข เล่นโกงและตัดสินใจเลิกใช้การหาเจ้ามือจ่ายค่ากาแฟโดยวิธีนี้ ตอนนี้เองที่นาย ก อาจปฏิเสธสมมติฐานที่ว่านาย ข เป็นคนซื่อเช่นแล้ว ถ้าสมมติฐานเดิมเป็นจริง มันก็อาจเป็นได้ที่นาย ก ต้องแพ้พนันถึง 10 ครั้งติดต่อกัน แต่ก็เพราะนาย ก ต้องแพ้พนันถึง 10 ครั้งติดต่อกันนี้เองที่ทำให้เขาปฏิเสธสมมติฐาน

การปฏิเสธสมมติฐานอาจมีเหตุผลสมควร นาย ข อาจเป็นคนไม่ซื่อจริง ๆ ในทางตรงกันข้ามสมมติฐานเดิมอาจถูกต้อง มันเป็นไปได้ตามทฤษฎีว่านาย ข อาจชนะพนันถึง 10 ครั้งติดต่อกันโดยไม่ได้เล่นโกงเลย เพราะฉะนั้น ถ้านาย ข เป็นคนซื่อ แต่นาย ก ตัดสินใจจากหลักพยานว่านาย ข ไม่ซื่อ นาย ก กำลังรับเอา **Type I error** คือปฏิเสธสมมติฐานที่เป็นจริง

สถานะการณ์ที่แตกต่างกับที่กล่าวมาแล้วอาจเกิดขึ้นได้ด้วย กล่าวคือในการโยนเหรียญเสี่ยงทายเพื่อจ่ายค่ากาแฟนั้น นาย ข ชนะ 7 ครั้งใน 10 ครั้ง นาย ก จะไม่สงสัยในความซื่อของนาย ข เลย หรืออีกนัยหนึ่ง นาย ก ยอมรับสมมติฐานว่า นาย ข เป็นคนซื่อ **conclusion** นี้อาจถูกต้อง แต่มันก็อาจเป็นไปได้ว่า นาย ข เป็นคนไม่ซื่อและวางแผนไว้แล้วว่าจะจงใจให้แพ้พนันบางบางครั้ง โดยจะเอาชนะพนัน

ด้วยการโกงให้มากครั้งกว่าที่แพ้ ถ้านาย ข เป็นคนไม่ซื่อ แต่นาย ก ยังเชื่อว่าซื่ออยู่เรื่อยไปแล้ว นาย ก กำลังรับเอา Type II error คือการยอมรับเอาสมมติฐานที่ผิดอย่างแท้จริงว่าเป็นสมมติฐานที่เป็นจริง

สถานะการณ์ทั้งหมดที่ได้กล่าวมาแล้วอาจนำมารวมไว้ในตารางข้างล่างนี้

	acceptance	rejection
true hypothesis	correct conclusion	Type I error
false hypothesis	Type II error	correct conclusion

จากตารางนี้ให้สังเกตว่า error ทั้งสองชนิดไม่อาจเกิดขึ้นพร้อมกันได้ ถ้าเรายอมรับสมมติฐานแล้วมี error เกิดขึ้น error นั้นต้องเป็น Type II error ถ้าเราปฏิเสธสมมติฐานแล้วมี error เกิดขึ้น error นั้นต้องเป็น Type I error

6.3 Level of Significance

ความน่าจะเป็นของการเกิดมี Type I error นี้ เรียกว่า level of significance ซึ่งจะแสดงให้เห็นได้โดยการทดสอบสมมติฐานว่า mean ของ population หนึ่งเท่ากับค่าที่กำหนดให้ สมมติฐานซึ่งอาจถูกหรือไม่ถูกนี้จะต้องถูกตรวจสอบกับข้อเท็จจริง ข้อเท็จจริงเหล่านี้คือ n observations ซึ่ง draw ออกมาจาก population นั้น ตัวอย่างเช่น สมมติฐานคือ population mean เท่ากับ 50 นั่นคือ $\mu_o = 50$ μ_o นี้ถูกใช้ในความหมายของ "hypothetical population mean" ซึ่งแตกต่างจาก population mean จริง alternative hypothesis อย่างหนึ่งคือ population mean น้อยกว่า 50 และ alternative hypothesis อีกอย่างหนึ่งคือ population mean มากกว่า 50 เพื่อทำให้ปัญหานี้ง่ายขึ้น จะสมมติว่าเราทราบ standard deviation ของ population เท่ากับ 10 เรา draw random sample ซึ่งประกอบด้วย 16 observations ออกมาจาก population และคำนวณค่า sample mean ไว้ ในขณะนี้ปัญหาที่เกิดขึ้นก็คือการตัดสินใจว่า population mean เท่ากับ 50 หรือน้อยกว่า 50 หรือมากกว่า 50 การตัดสินใจจะต้องอาศัยความรู้ที่มีคือ $\sigma = 10$, $n = 16$ และค่าของ sample mean ($\bar{y}$) อย่างไรก็ดี inductive inference เกี่ยวกับ population mean จะ draw จากหนึ่ง sample โดยปราศจากความรู้เรื่อง distribution ของ sample means ทั้งหมดไม่ได้ แต่ความรู้เรื่อง

distribution ของ sample means นี้อาจหาออกมาได้จาก theorems ต่าง ๆ ในบทที่ 5 โดยปราศจากการ draw sample หนึ่งออกมาจริง ๆ เลย เราทราบดีแล้วว่า

ก. mean ของ means ของ all possible samples ซึ่งมี size เดียวกันเท่ากับ population mean (Theorem 5.3) หรือ

$$\boxed{\mu_{\bar{y}} = \mu}$$

ข. standard deviation ของ all sample means เท่ากับ population standard deviation หารด้วย square root ของ sample size (Theorem 5.3) หรือ

$$\boxed{\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}}$$

ค. sample means follow the normal distribution (Theorem 5.2 a และ 5.2b) เพราะฉะนั้น ถ้าสมมติฐานเป็นจริง และ all possible samples ซึ่งมี size 16 ถูก draw ออกมาจาก population นี้ sample means ควรจะ follow the normal distribution ซึ่งมี mean ($\mu_{\bar{y}}$) เท่ากับ 50 และ standard deviation ($\sigma_{\bar{y}}$) เท่ากับ $\frac{\sigma}{\sqrt{n}}$ หรือ $\frac{10}{\sqrt{16}}$ หรือ 2.5 distribution ของ sample means นี้ได้แสดงให้เห็นใน Fig. 6.3 a

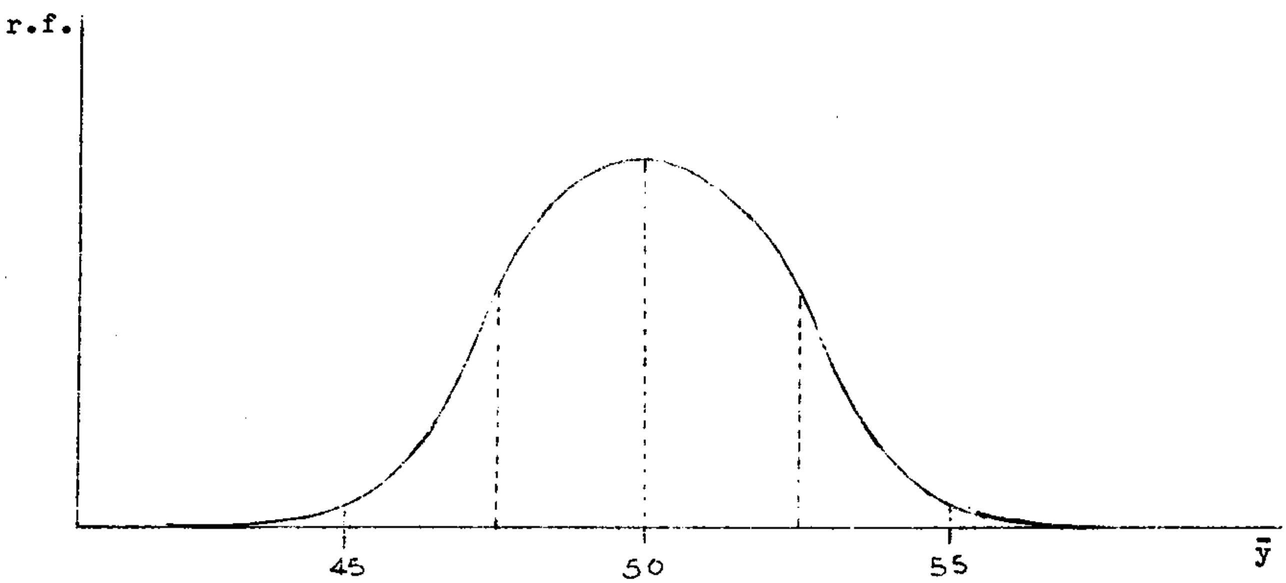


Fig. 6.3 a

ถ้าค่าของ sample mean ($\bar{y}$) ออกมาเป็น 60 การปฏิเสธสมมติฐานย่อมจะมีเหตุผลดี เพราะว่าถ้า population mean เท่ากับ 50 แล้ว มันอาจเป็นไปได้แต่ก็ไม่ค่อยจะเกิดขึ้นสำหรับ mean ของ random sample หนึ่งเท่ากับ 60 (ดู Fig. 6.3 a) การไม่พ้องกันระหว่างข้อเท็จจริงและสมมติฐานจะนำไปสู่การปฏิเสธสมมติฐานนั้น แล้ว conclusion ก็คือ population mean จริงมากกว่า 50 (มันมากกว่า 50 เท่าไรนั้นเป็นปัญหาอีกอย่างหนึ่งซึ่งจะได้พิจารณาในบทที่ 11) โดยทำนองเดียวกัน ถ้า sample mean ($\bar{y}$) ออกมาเป็น 40 conclusion ก็คือ population mean จริงน้อยกว่า 50 อย่างไรก็ดี ถ้า sample mean ($\bar{y}$) ออกมาเป็น 50.5 ซึ่งใกล้เคียงกับ 50 มาก (ดู Fig. 6.3 a) ซึ่งเราจะยอมรับสมมติฐาน และ conclusion ก็คือ population mean เท่ากับ 50 ในกรณีนี้เรายอมรับสมมติฐานโดยเหตุผลที่ว่าหลักพยานที่ได้จาก sample ไม่ได้พิสูจน์ให้เห็นว่าสมมติฐานไม่ถูกต้อง ในขณะนี้เราเห็นแล้วว่าถ้า sample mean เท่ากับ 50.5 เรายอมรับสมมติฐาน และถ้า sample mean เท่ากับ 40 หรือ 60 เราปฏิเสธสมมติฐาน การตัดสินยอมรับหรือปฏิเสธนี้ดูจะทำได้ง่ายเพราะมันเป็นค่าที่มากกว่ากันหรือน้อยกว่ากันอย่างเห็นได้ชัด แต่การตัดสินควรจะทำอย่างไร ถ้า sample mean เท่ากับ 51, 52, 53 ฯลฯ เราจะต้องขีดเส้นขึ้นเส้นหนึ่งตรงที่ใดที่หนึ่งซึ่งถ้าเลยเส้นนี้ออกไปแล้วเราจะปฏิเสธสมมติฐาน ปลายหางทั้งสองข้างของ distribution ของ sample means จึงถูกแบ่งออกเพื่อความประสงค์ดังกล่าวนี้ ส่วนที่ถูกแบ่งออกไปทั้งสองข้างเรียกว่า "critical regions" เมื่อ sample mean หนึ่งตกอยู่ภายใน critical region ข้างใดข้างหนึ่งเราปฏิเสธสมมติฐานนั้น

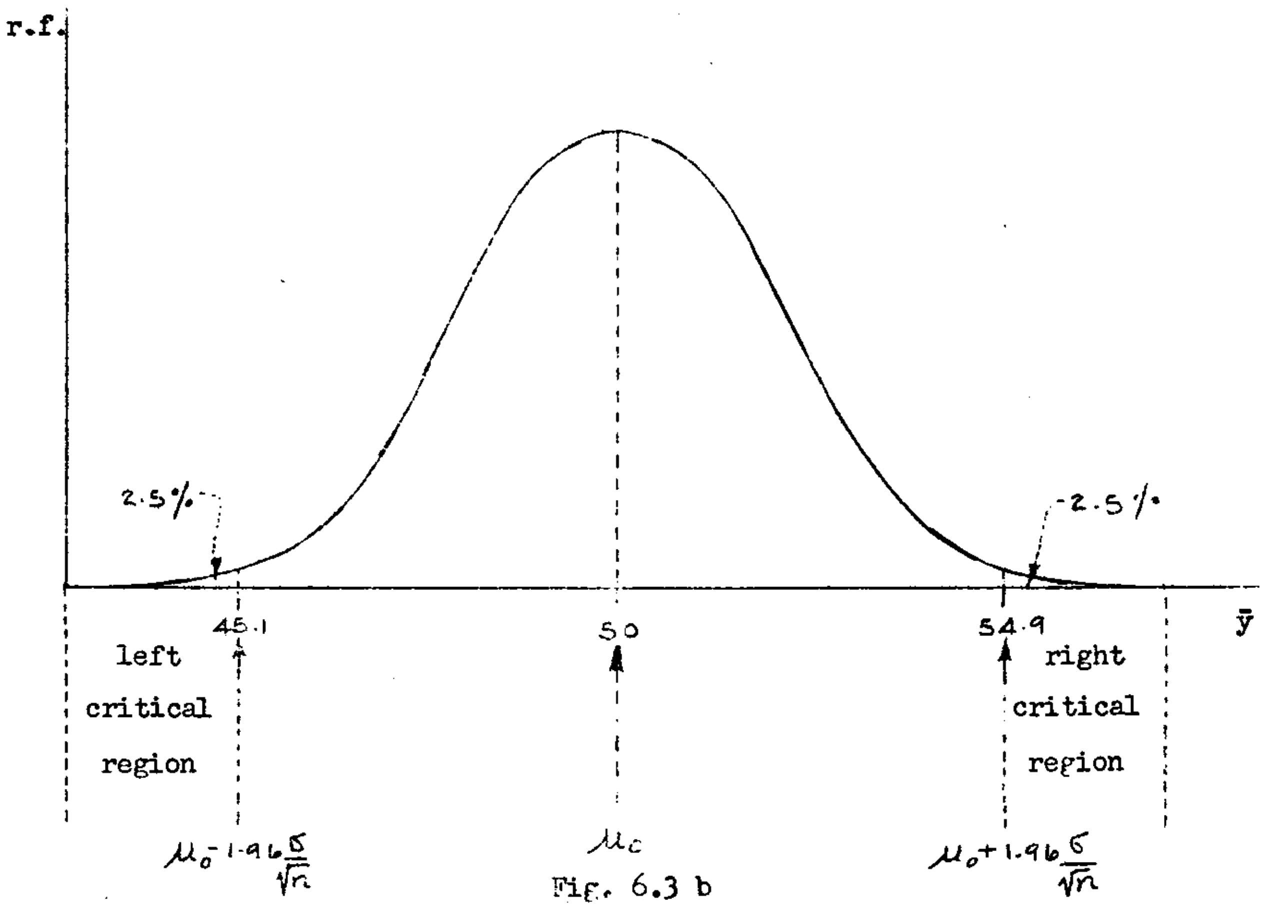


Fig. 6.3 b

คำว่า "ภายใน" และ "ภายนอก" เมื่อใช้ในการอ้างอิงเกี่ยวกับ critical regions มักจะก่อให้เกิดความสับสนในเวลาทดสอบสมมติฐานได้ ถ้า sample mean ตัวหนึ่งตกอยู่ที่ปลายหางข้างใดข้างหนึ่งของ distribution curve เราเรียกว่า sample mean ตัวนั้นอยู่ภายใน critical region แต่ถ้ามันตกอยู่ตรงส่วนกลางของ distribution curve เราเรียกว่ามันอยู่ภายนอก critical regions

ขนาดของ critical regions จะเลือกใช้ได้ตามความพอใจ critical regions อาจถูกแบ่งออกโดยให้ 2.5 % ของ sample means ตกอยู่ภายใน critical region ข้างซ้าย และอีก 2.5 % ของ sample means ตกอยู่ภายใน critical region ข้างขวา (ดู Fig. 6.3 b) sample mean ที่มีค่าน้อยกว่า $\mu_0 - 1.96\frac{\sigma}{\sqrt{n}}$ หรือ $50 - 1.96\,(\frac{10}{\sqrt{16}})$ หรือ 45.1 จะอยู่ภายใน critical region ข้างซ้าย และ sample mean ที่มีค่ามากกว่า $\mu_0 + 1.96\frac{\sigma}{\sqrt{n}}$ หรือ $50 + 1.96\,(\frac{10}{\sqrt{16}})$ หรือ 54.9 จะอยู่ภายใน critical region ข้างขวา เมื่อ sample mean ตกอยู่ภายใน critical region ข้างซ้าย conclusion คือ population mean มีค่าน้อยกว่า 50 และเมื่อมันตกอยู่ภายใน critical region ข้างขวา conclusion คือ population mean มีค่ามากกว่า 50 แต่เมื่อ sample mean ตกอยู่ภายนอก critical regions, conclusion คือ population mean มีค่าเท่ากับ 50 เราต้องทราบเสียก่อนว่า distribution ของ sample means ใน Fig. 6.3 a นั้น ได้ทำขึ้นบนสภาพซึ่งสมมติฐานถูกต้อง นั่นคือ population mean เท่ากับ 50 จริง ดังนั้น ถ้า critical regions ถูกแบ่งออกดังกล่าวข้างต้น 5 % ของ all possible samples จะตกอยู่ภายใน critical regions และจะนำไปสู่การปฏิเสธสมมติฐานที่เป็นจริงอย่างไม่ถูกต้อง (คือเกิด Type I error ขึ้น) เปอร์เซนต์ (r.f.) ของ all possible samples ซึ่งนำไปสู่การเกิด Type I error นี้เรียกว่า "significance level" เพราะฉะนั้น significance level ก็คือความน่าจะเป็น (ข้อ 3.4) ซึ่ง Type I error อาจเกิดขึ้นบนพื้นฐานของ sample เพียง sample เดียว

significance level 5 % ไม่ได้บอกเป็นนัยว่า 5 % ของ conclusions ทั้งหมดของการทดสอบสมมติฐานนั้นไม่ถูกต้อง การพิจารณาในข้อนี้ใช้กับกรณีสมมติฐานถูกต้องเท่านั้น กรณีสมมติฐานผิดจะได้พิจารณาในข้อต่อไป

ตามทฤษฎี significance level จะถูกเลือกใช้ได้ตามใจชอบ แต่โดยปกติในทางปฏิบัตินั้น เรามักจะใช้ 5 % และ 1 % significance level ในตัวอย่างที่แล้วเราอาจเลือกใช้ significance level ได้หลายอย่าง ถ้าเราใช้ค่า 2.576 แทนค่า 1.960 ในการพิจารณา critical regions, significance level ก็จะเปลี่ยนจาก 5 % เป็น 1 % ค่าต่าง ๆ ใน Table 3, Appendix ได้ให้การ

เลือกใช้ significance level ไว้หลายอย่าง แต่สำหรับปัญหาอื่นที่นำมากล่าวในบทต่อ ๆ ไปนั้น tables ต่าง ๆ จะให้การเลือกอันจำกัด ในการทดสอบสมมติฐานเป็นอันมากเราไม่สามารถใช้ 4.9 % significance level ได้นอกจากเราจะทำ table ดังกล่าวนี้ขึ้นใช้เอง ทฤษฎีของสถิติศาสตร์ให้อิสระเต็มที่ในการเลือกใช้ significance level แต่ tables ต่าง ๆ ที่มีอยู่กำหนดขีดจำกัดในการเลือกไว้ ดังจะเห็นว่าตามปกติเราใช้ 5 % และ 1 % เท่านั้น และ significance levels อื่น ๆ เช่น 3.8 %, 4.1 % ฯลฯ ถึงแม้ว่าจะเคยใช้กันแต่ก็มีน้อยมาก ในข้อต่อไปเราจะพิจารณาถึงผลที่เกิดขึ้นจากการเลือกใช้ significance level สูงและต่ำ

6.4 Type II Error

Type II error ที่ได้กล่าวมาแล้วนั้นเป็น error ที่เกิดขึ้นโดยการยอมรับสมมติฐานที่ไม่ถูกต้อง การเกิด error ชนิดนี้และความสัมพันธ์ของมันกับ level of significance จะถูกพิจารณาโดยการใช้ตัวอย่างเดียวกันกับที่ให้ไว้ในข้อที่แล้ว สมมติฐานคือ population mean เท่ากับ 50 และทราบค่าของ population standard deviation ว่าเท่ากับ 10 ถ้า mean ของ 16 observations ตกอยู่ภายนอก critical regions คือ มีค่ามากกว่า 45.1 และน้อยกว่า 54.9 (Fig. 6.4) เราก็ยอมรับสมมติฐาน Type II error จะเกิดขึ้นเฉพาะเมื่อ population mean ไม่เท่ากับ 50 แต่เท่ากับค่าอื่น เพื่อช่วยการพิจารณา เราให้ population mean จริงเท่ากับ 58 ในขณะที่ hypothetical mean ยังคงเท่ากับ 50 ในขณะนี้จะมี distribution curves ของ sample means สองเส้น เส้นหนึ่งเป็น hypothetical curve ซึ่งผิด และอีกเส้นหนึ่งเป็น curve จริง curve B ใน Fig. 6.4 เป็น distribution จริงของ sample means ซึ่งอยู่รอบ ๆ ค่า 58 ตรงกึ่งกลาง curve A เป็น hypothetical curve แต่เป็น distribution curve ที่ผิด

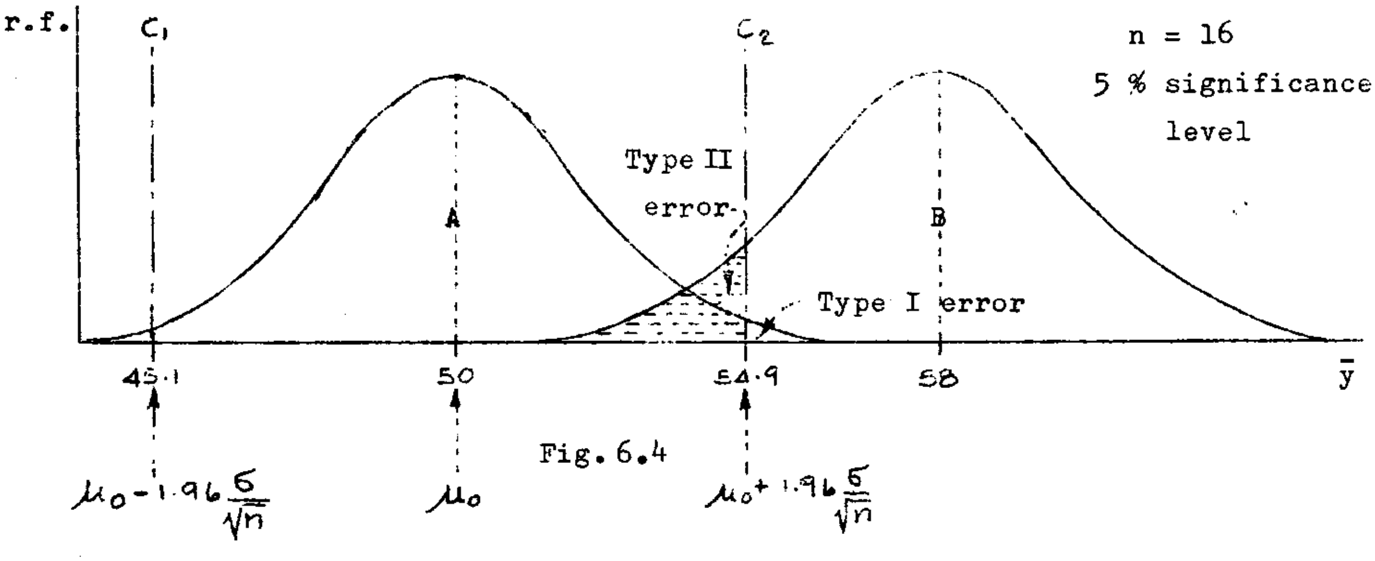


Fig. 6.4

จาก Fig. 6.4 จะสังเกตได้ว่า มีจำนวนเปอร์เซ็นต์ของ sample means ของ curve B ซึ่งมีค่าน้อยกว่า 54.9 รวมอยู่ด้วย ถ้าเราได้ sample mean ตัวหนึ่งในจำนวนนี้มา เราก็จะยอมรับเอาสมมติฐานที่ว่า population mean เท่ากับ 50 ไปอย่างผิด ๆ เพราะฉะนั้น จำนวนเปอร์เซ็นต์ (r.f.) ของ all possible samples นี้คือความน่าจะเป็น (ข้อ 3.4) ที่จะเกิด Type II error สำหรับ sample เพียง sample เดียวซึ่งแสดงไว้ด้วยส่วนของ curve B ที่ขีดเส้นประไว้ทั้งหมดทางด้านซ้ายของเส้น C_2 ใน Fig. 6.4

จาก Fig. 6.4 จะเห็นได้ว่า true mean (μ) ยิ่งห่างจาก hypothetical mean (μ_0) มากขึ้นเท่าไร ความน่าจะเป็นที่จะเกิด Type II error จะยิ่งลดน้อยลงเท่านั้น กล่าวได้อีกอย่างหนึ่งว่า ถ้า population mean จริง และ hypothetical mean อยู่ใกล้กัน เราก็มักจะยอมรับเอาสมมติฐานที่ผิดและเกิด Type II error ขึ้น ถ้ามันอยู่ห่างกันเราก็มักจะไม่ยอมรับสมมติฐานที่ผิด

เรายังเห็นอีกด้วยว่าถ้า significance level (ความน่าจะเป็นที่จะเกิด Type I error) ถูกลดลงจาก 5 % เป็น 1 % แนว C_1 จะเคลื่อนที่ไปทางซ้าย และแนว C_2 จะเคลื่อนที่ไปทางขวา เมื่อเป็นเช่นนี้ความน่าจะเป็นที่จะเกิด Type II error จะเพิ่มมากขึ้น ความพยายามใด ๆ ที่จะลด significance level ลงจะทำให้เพิ่มความน่าจะเป็นที่จะเกิด Type II error มากขึ้น จึงกล่าวได้ว่าถ้าความน่าจะเป็นที่จะเกิด error ชนิดหนึ่งถูกเพิ่มขึ้น ความน่าจะเป็นที่จะเกิด error อีกชนิดหนึ่งจะถูกลดลงและโดยกลับกัน

6.5 Sample Size

Sample size มีบทบาทสำคัญในการทดสอบสมมติฐาน เพราะว่า Type I error และ Type II error ทั้งสองชนิดนี้จะลดน้อยลงได้โดยการเพิ่ม sample size จาก Fig. 6.4 เราจะสังเกตได้ว่า ภายหลังที่เราเลือกใช้ significance level และกำหนด critical regions ลงไปแล้ว ความน่าจะเป็นของการเกิด Type II error จะขึ้นอยู่กับบริเวณที่ curve A และ curve B ล้ำกันอยู่ ถ้า sample size ถูกเพิ่มขึ้นจาก 16 เป็น 100 standard deviation ของ sample means ($\sigma_{\bar{y}}$) ก็จะถูกลดลงจาก $\frac{10}{\sqrt{16}}$ หรือ 2.5 เป็น $\frac{10}{\sqrt{100}}$ หรือ 1 แล้วการแผกระจายของแต่ละ curve ก็จะถูกลดลง (ดู Fig. 6.5)

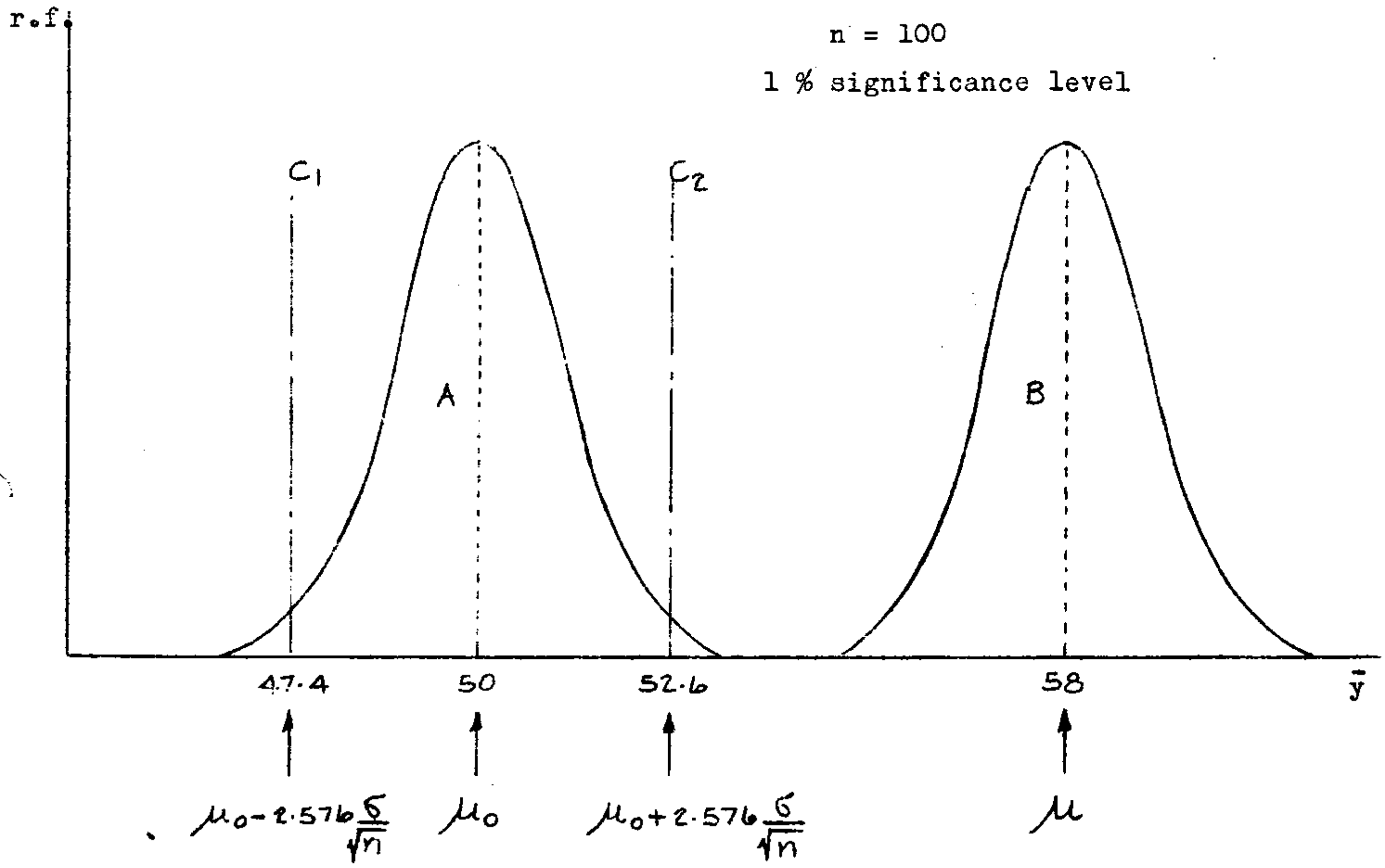


Fig. 6.5

และยังผลให้บริเวณที่ curve A และ curve B ล้ำกันอยู่ถูกลดลงด้วย ถ้าเราใช้ 5 % significance level แนว C_2 จะอยู่ที่ $50 + 1.960 \frac{\sigma}{\sqrt{n}}$ หรือ $50 + 1.960 \frac{10}{\sqrt{100}}$ หรือ 51.960 แต่ถ้าเราใช้ 1 % significance level แนว C_2 จะเลื่อนไปอยู่ที่ $50 + 2.576 \frac{\sigma}{\sqrt{n}}$ หรือ $50 + 2.576 \frac{10}{\sqrt{100}}$ หรือ 52.576 ค่า 52.576 นี้จะอยู่ต่ำกว่า true mean 58 มากกว่า 5 standard deviations ($\sigma_{\bar{y}} = 1$ สำหรับ n = 100) ดังนั้นความน่าจะเป็นของการเกิด Type II error จึงเกือบเป็นศูนย์แม้ว่าเราจะใช้ 1 % significance level ก็ตาม (Fig. 6.5) กล่าวได้อีกอย่างหนึ่งว่า sample ใหญ่ช่วยลด errors ทั้งสองชนิดลง ถ้า significance level (ความน่าจะเป็นของการเกิด Type I error) ไม่ถูกเปลี่ยนแปลง sample ใหญ่จะยังผลให้เกิดการลดลงของความน่าจะเป็นของการเกิด Type II error

การพิจารณาข้างบนนี้เป็นการพิจารณาที่เราสมมติไว้ก่อนว่า สมมติฐานนั้นผิดเพราะว่า Type II error จะเกิดขึ้นต่อเมื่อสมมติฐานผิดเท่านั้น ถ้าสมมติฐานถูกต้อง significance level ที่เลือกใช้ตาม

ขอบใจนั้นจะแสดงถึงเปอร์เซนต์อะไรของ sample means ที่นำไปสู่การปฏิเสธสมมติฐานที่ถูกต้อง หรือความน่าจะเป็นของการเกิด Type I error ถ้า significance level ไม่ถูกเปลี่ยนแปลงและสมมติฐานที่ทดสอบเป็นจริง sample ใหญ่จะไม่มีประโยชน์เหนือกว่า sample เล็กเลย เพราะว่าจะเกิด Type II error ในสมมติฐานที่เป็นจริงไม่ได้ เพราะฉะนั้นความประสงค์ของการใช้ sample ใหญ่ก็เพื่อจะลดความน่าจะเป็นของการเกิด Type II error ถ้า significance level คงที่

การพิจารณาข้างบนนี้ขึ้นอยู่กับข้อเท็จจริงที่ว่า standard deviation ของ distribution ของ sample means จะถูกลดลงถ้า sample size ถูกเพิ่มขึ้น และส่วนของ curve A และ B ที่ล้ำกันอยู่ใน Fig. 6.5 ก็จะลดลงด้วย ทั้งนี้เนื่องจาก

$$\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}} \qquad \ldots\ldots\ldots\ldots\ldots (1)$$

จะเห็นได้ว่า $\sigma_{\bar{y}}$ จะถูกลดลงได้โดยการเพิ่ม n หรือลด population standard deviation (σ) การเพิ่ม n หรือการลด σ มีผลอย่างเดียวกันต่อการทดสอบสมมติฐาน การลด σ อาจทำได้สำเร็จโดยตัวนักวิทยาศาสตร์ผู้ทำการทดลอง

ภายหลังที่กำหนด significance level แล้ว การลดความน่าจะเป็นของการเกิด Type II error จะทำได้สำเร็จโดยการเพิ่ม sample size (n) หรือลด population standard deviation (σ) ความสำคัญของการลดความน่าจะเป็นนี้จะเน้นหนักเกินไปไม่ได้เนื่องจากธรรมชาติของการทดสอบสมมติฐาน ความคิดทั้งหมดของการทดสอบสมมติฐานเป็นความพยายามที่จะแสดงหลักพยานเพื่อพิสูจน์ให้เห็นว่าสมมติฐานนั้นไม่ถูกต้อง ถ้าสมมติฐานไม่ถูกปฏิเสธมันอาจเป็นเพราะว่าหลักพยานซึ่งพิสูจน์ให้เห็นว่าสมมติฐานไม่ถูกต้องไม่ได้ถูกแสดงออกมา การขาดหลักพยานดังกล่าวนี้อาจเป็นผลจาก sample มีขนาดโตไม่พอหรือจากการทดลองที่มีข้อผิดพลาดมากก็ได้

6.6 Summary

ความสัมพันธ์อย่างแน่นแฟ้นในระหว่าง significance level (Type I error) Type II error และ sample size อาจสรุปได้ดังต่อไปนี้

1. ความแตกต่างระหว่าง true mean (μ) และ hypothetical mean (μ_0) ยิ่งมีมากเท่าใดจะยังผลให้ความน่าจะเป็นของการเกิด Type II error น้อยลงไปเท่านั้น ถ้าใช้ sample size ขนาดเดียวกัน และ significance level เดียวกัน

2. สำหรับ sample size ขนาดเดียวกัน การลด significance level (Type I error) เช่น จาก 5 % ลงเป็น 1 % จะเป็นผลในการเพิ่มความน่าจะเป็นของการเกิด Type II error

3. แนวทางที่จะลดความน่าจะเป็นของการเกิด errors ทั้งสองชนิดคือการเพิ่ม sample size หรือลด population standard deviation หรือทำทั้งสองอย่าง

4. ถ้า significance level ถูกกำหนดตายตัว การปรับปรุงเทคนิคการทดลอง และการเพิ่ม sample size จะลดความน่าจะเป็นของการเกิด Type II error ลงได้ เทคนิคของการทดลองที่ไม่ดีเมื่อรวมกับจำนวน observations ที่มีไม่พอมักจะทำให้เรายอมรับเอาสมมติฐานที่ทดสอบไม่ว่าสมมติฐานนั้นจะจริงหรือไม่จริง

6.7 The u-test

การทดสอบสมมติฐานว่า population mean เท่ากับค่าที่กำหนดให้ โดยทราบค่า population standard deviation แล้วนั้นในหนังสือเล่มนี้จะเรียกว่า "u - test" การทดสอบวิธีนี้ไม่มีชื่อที่ใช้กันทั่วไป และที่เรียกง่าย ๆ ว่า u - test ก็เพื่อความสะดวกในการอ้างอิง ในข้อที่แล้วมาตำแหน่งที่อยู่ของ sample mean ใน distribution curve ต้องถูกแสดงในเทอมของจำนวน standard deviation ที่ห่างจาก population mean จำนวน standard deviation นี้ คือ u ซึ่ง

$$u = \frac{\bar{y} - \mu_{\bar{y}}}{\sigma_{\bar{y}}}$$

$$= \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{................ (1)}$$

ตัวอย่างเช่น ให้ sample หนึ่งซึ่งมี size (n) = 25 และ mean ($\bar{y}$) = 53 ได้มาจาก population หนึ่งซึ่งมี mean (μ) = 50 และ standard deviation (σ) = 10 standard deviation ของ sample means ทั้งหมด ($\sigma_{\bar{y}}$) = $\frac{\sigma}{\sqrt{n}}$ หรือ $\frac{10}{\sqrt{25}}$ หรือ 2 ระยะห่างจาก $\bar{y}$ ถึง μ คือ 53 — 50 หรือ 3 ซึ่งเท่ากับ 1.5 standard deviation ของ sample means หรือ 1.5 $\sigma_{\bar{y}}$ และค่า 1.5 นี้ก็คือค่าของ u ดังจะเห็นได้ว่า

$$u = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{53 - 50}{\frac{10}{\sqrt{25}}}$$

$$= \frac{3}{2}$$

$$= 1.5$$

อย่างแท้จริง คงจะจำได้ว่าเราเคยพูดถึง u มาแล้วครั้งหนึ่งใน Equation (1) ของข้อ 3.2 คือ

$$u = \frac{y - \mu}{\sigma} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

ใน Equation (2) นี้ observations (y) follow the normal distribution ซึ่งมี mean เท่ากับ μ และ standard deviation เท่ากับ σ ภายหลังที่ y แต่ละตัวถูกเปลี่ยนรูปเป็น u แล้ว observations ที่เปลี่ยนรูปไปหรือ u นี้จะ follow the normal distribution ซึ่งมี mean เท่ากับ 0 และ standard deviation เท่ากับ 1 (ข้อ 3.2) ใน Equation (1) sample means ($\bar{y}$) follow the normal distribution ซึ่งมี mean เท่ากับ $\mu_{\bar{y}}$ (หรือ μ ตาม Theorem 5.3) และ standard deviation เท่ากับ $\sigma_{\bar{y}}$ (หรือ $\frac{\sigma}{\sqrt{n}}$ ตาม Theorem 5.3) ภายหลังที่มีการเปลี่ยนรูป $\bar{y}$ เป็น u คือ

$$u = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

ทำนองเดียวกันแล้ว sample means ($\bar{y}$) ที่เปลี่ยนรูปไปเป็นนี้จะ follow the normal distribution ซึ่งมี mean เท่ากับ 0 และ standard deviation เท่ากับ 1 เหมือนกัน u สองตัวใน Equations (1) และ (2) ไม่ใช่จำนวนเลขเดียวกัน แต่ที่ใช้อักษร u ด้วยกันก็เพราะว่ามันมี distribution curve อย่างเดียวกัน อย่างไรก็ดี Equation (2) ถือว่าเป็นกรณีพิเศษของ Equation (1) คือถ้า n = 1 sample mean ($\bar{y}$) ก็คือตัว observation ใน sample นั้นเอง แล้ว Equation (1) ก็จะกลายรูปเป็น Equation (2) ไปดังนี้

$$u = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{y - \mu}{\frac{\sigma}{\sqrt{1}}}$$

$$= \frac{y - \mu}{\sigma}$$

ถ้าเราใช้ 5 % significance level, critical regions คือ $\bar{y} < \mu_o - 1.96 \frac{\sigma}{\sqrt{n}}$ และ $\bar{y} > \mu_o + 1.96 \frac{\sigma}{\sqrt{n}}$ critical regions นี้อาจแสดงให้เห็นได้ง่ายอีกอย่างหนึ่งคือ $u < -1.96$ และ $u > 1.96$ inequality $u < -1.96$ บอกเป็นนัยว่า

$$\frac{\bar{y} - \mu_o}{\frac{\sigma}{\sqrt{n}}} < -1.96$$

เมื่อเอา $\frac{\sigma}{\sqrt{n}}$ คูณทั้งสองข้างของ inequality ข้างบน inequality ใหม่ที่ได้คือ

$$\bar{y} - \mu_o < -1.96 \frac{\sigma}{\sqrt{n}}$$

ถ้าเอา μ_o บวกเข้าทั้งสองข้างของ inequality ใหม่อีกครั้งหนึ่งก็จะได้

$$\bar{y} < \mu_o - 1.96 \frac{\sigma}{\sqrt{n}}$$

เพราะฉะนั้น $u < -1.96$ จึงเป็น critical region อย่างเดียวกับ $\bar{y} < \mu_o - 1.96 \frac{\sigma}{\sqrt{n}}$ โดยทำนองเดียวกัน $u > 1.96$ ก็เป็น critical region อย่างเดียวกับ $\bar{y} > \mu_o + 1.96 \frac{\sigma}{\sqrt{n}}$ ในสถานะเดียวกันต่อจากนี้ไปถ้าใช้ 5 % significance level และ 2 alternative hypotheses ค่าของ u จะถูกคำนวณจาก sample และ critical regions จะถูกแสดงในรูป $u < -1.96$ และ $u > 1.96$ alternative hypothesis ตัวหนึ่งย่อมเป็นไปตาม critical region หนึ่ง การทดสอบสมมติฐานซึ่งมี 2 alternative hypotheses และ 2 critical regions เรียกว่า <u>two-tailed test</u> แต่ถ้ามีเพียง alternative hypothesis เดียว (ข้อ 6.1) และหนึ่ง critical region เรียกว่า <u>one-tailed test</u>

6.8 Assumptions

ข้อสมมติ (assumptions) คือสภาพที่จะทำให้การทดสอบสมมติฐานสมบูรณ์ ในตัวอย่างที่ใช้อธิบายในข้อ 6.3 และ 6.4 นั้น assumption สำคัญที่สุดคือ sample เป็น random ถ้า sample ถูกเลือกออกมาอย่างจงใจทำให้ sample mean ($\bar{y}$) ใกล้หรือไกลจาก hypothetical mean (μ_o) เพื่อที่จะให้ยอมรับหรือปฏิเสธสมมติฐานแล้วภาวะแห่งวัตถุประสงค์ที่หมายและความสมบูรณ์ของการทดสอบจะถูกทำลายโดยสิ้นเชิง random sample เป็น sample ที่ draw มาจาก population โดย observation ทุกตัวใน population มีโอกาสเท่าเทียมกันที่จะได้รับเลือกออกมา

assumption อีกข้อหนึ่งก็คือ population เป็น normal และ normal population เท่านั้นที่ทำให้ distribution ของ sample means เป็น normal (Theorem 5.2 b) แต่ขนาดของ sample เท่ากับ 16 ก็ใหญ่พอที่จะทำให้เราเชื่อได้ว่า distribution ของ sample means เป็น normal โดยประมาณหรือใกล้เคียง normal มากที่สุด (Theorem 5.2 a) แม้ว่า population จะไม่เป็น normal ก็ตาม เพราะฉะนั้น assumption ข้อนี้เมื่อเปรียบเทียบกับ assumption ที่ว่า sample เป็น random แล้วจะด้อยกว่าหรือมีความสำคัญน้อยมาก

6.9 Procedures

วิธีดำเนินการ (procedures) ของการทดสอบสมมติฐานอาจแสดงให้เห็นโดยการทดลองในการทดสอบสมมติฐานว่า population mean เท่ากับ 145 ที่ 1 % significance level เราจะใช้หนึ่ง random sample ซึ่งมี 25 observations และทราบแล้วว่า population standard deviation เท่ากับ 20 วิธีดำเนินการทำได้ดังนี้

1. Hypothesis: สมมติฐานคือ population mean เท่ากับ 145 นั่นคือ $\mu_o = 145$
2. Alternative Hypotheses: alternative hypotheses คือ
   - ก. population mean น้อยกว่า 145 และ
   - ข. population mean มากกว่า 145
3. Assumptions: ข้อสมมติคือ
   - ก. sample เป็น random (สำคัญ)
   - ข. population เป็น normal (ไม่สู้สำคัญ) และ
   - ค. ทราบ population standard deviation

4. Level of Significance: significance level ที่เลือกใช้คือ 1 %

5. Critical Regions: critical regions อยู่ที่

ก. $u < -2.576$

ข. $u > 2.576$ (Table 3, Appendix)

6. Computation of Statistic

$$\mu_o = 145$$

$$n = 25 \quad \text{(กำหนดให้)}$$

$$\sum y = 3{,}471 \quad \text{(กำหนด 25 observations ให้แต่ไม่ได้นำมาแสดงไว้ในที่นี้)}$$

$$\bar{y} = \frac{3{,}471}{25} = 138.84$$

$$\bar{y} - \mu_o = 138.84 - 145 = -6.16$$

$$\sigma = 20 \quad \text{(กำหนดให้)}$$

$$\sqrt{n} = \sqrt{25} = 5$$

$$\frac{\sigma}{\sqrt{n}} = \frac{20}{5} = 4$$

$$u = \frac{\bar{y} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$$

$$= \frac{-6.16}{4}$$

$$= -1.54$$ ซึ่งอยู่ภายนอก critical regions

7. Conclusion: population mean เท่ากับ 145

6.10 Remarks

การทดสอบสมมติฐานทางสถิติใช้ได้กับข้อมูลเลขจำนวนซึ่งเป็นการวัดเท่านั้น เช่น อุณหภูมิของห้อง ความสูงของคน ฯลฯ หรือการนับ เช่น จำนวนแมลงบนใบไม้หรือจำนวนหนังสือบนหิ้ง ก่อนที่จะใช้วิธีทางสถิติได้นั้น informations ที่ได้มาจะต้องถูกแสดงเป็นเลขจำนวน ปัญหาข้อแรกของนักวิทยาศาสตร์

คือแนวทางการวางแผนอุปกรณ์ของการวัด ก่อนที่ thermometer จะถูกประดิษฐ์ขึ้นใช้นั้นอุณหภูมิจะถูกบรรยายเป็นคำพูด เช่น ร้อน อุ่น หนาว และเย็น เช่นนี้เป็นต้น สิ่งที่ไม่มีตัวตนก็ถูกทำให้มีตัวตนขึ้นมาโดย thermometer I.Q. ก็เป็นความพยายามของนักจิตวิทยาที่จะทำให้สติปัญญาของมนุษย์มีตัวตนซึ่งมองเห็นได้ ในขณะความรู้ของมนุษย์ก้าวหน้าไปคุณภาพของสิ่งต่าง ๆ ที่ไม่มีตัวตนได้ถูกทำให้เป็นคุณภาพที่มีตัวตนซึ่งมองเห็นได้ อย่างไรก็ดีการทดสอบสมมติฐานซึ่งไม่มีการแสดง information เป็นเลขจำนวนก็ใช้กันอย่างแพร่หลาย เช่น วิธีดำเนินการตามกฎหมายในสหรัฐและประเทศอื่น ๆ

Chapter 7

Sample Variance, X^2- Distribution

บทที่ 5 และบทที่ 6 ที่แล้วมาเป็นเรื่องเกี่ยวกับความสัมพันธ์ในเชิง deductive และ inductive ระหว่าง population หนึ่งกับ samples ต่าง ๆ ของมัน ความสัมพันธ์ในเชิง deductive นั้นได้กล่าวไว้ในบทที่ 5 ซึ่งบรรยายถึงลักษณะของ sample means ที่ได้มาจาก population หนึ่ง แนวทางของความสัมพันธ์จึงออกจาก population ไปสู่ samples ในทางตรงกันข้ามบทที่ 6 ซึ่งแสดงให้เห็นว่า sample เดี่ยวเพียง sample เดียวถูกใช้เพื่อทดสอบสมมติฐานที่เกี่ยวกับ population ได้อย่างไรนั้น ได้ชี้ให้เห็นความสัมพันธ์ในเชิง inductive ระหว่าง population หนึ่งกับ samples ของมัน แนวทางของความสัมพันธ์จึงออกจาก sample หนึ่งเข้ามาสู่ population

จุดมุ่งหมายของการพิจารณาในบทที่ 5 และบทที่ 6 คือ mean ในบทที่ 7 นี้เราจะกล่าวซ้ำถึงเรื่องต่าง ๆ ที่ได้กล่าวมาแล้วทั้งหมดอีกครั้งหนึ่ง แต่จุดมุ่งหมายของการพิจารณาจะเปลี่ยนจาก mean เป็น variance

7.1 Purposes of Studying Sample Variance

ความประสงค์ในการศึกษาเรื่อง sample variance มี 2 ประการคือ:

1. เพื่อให้ได้ความรู้ที่เกี่ยวกับ population variance
2. เพื่อให้ข้อจำกัดในการใช้ u-test หมดไป

u-test ซึ่งกล่าวไว้ในข้อ 6.7 นั้น ใช้สำหรับทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนด ในเมื่อ

$$u = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

ใน Equation (1) ข้างบนจะเห็นได้ว่าเราต้องทราบ population standard deviation (σ) เสียก่อนจึงจะใช้การทดสอบนี้ได้ แต่ตามปกติแล้วเราจะไม่ทราบ σ เลย เพราะฉะนั้นประโยชน์ของ u-test จึงมีจำกัด เพื่อให้ข้อจำกัดในการใช้ u-test หมดไป เราจึงต้องหาวิธีประมาณค่า σ หรือ σ^2 จาก sample

เพราะฉะนั้นไม่ว่าความสนใจของเราจะอยู่ที่ population mean หรือ population variance ก็ตาม เรื่อง sample variance เป็นสิ่งจำเป็นที่เราจะต้องทราบอย่างแท้จริง

7.2 Sample Variance

ปัญหาข้อแรกในการศึกษาเรื่อง sample variance ก็คือการพิจารณาว่า sample variance คืออะไร เพื่อจะได้ให้ค่าประมาณที่ดีของ population variance เรื่องนี้ sample mean จะบอกเป็นนัยให้เราทราบได้

เราทราบแล้วว่า

$$\text{population mean } (\mu) = \frac{\sum y}{N} = \frac{y_1 + y_2 + \ \ldots\ldots\ + y_n}{N} \quad \ldots\ldots\ (1)$$

$$\text{population variance } (\sigma^2) = \frac{\sum (y - \mu)^2}{N}$$

$$= \frac{(y_1 - \mu)^2 + (y_2 - \mu)^2 + \ \ldots\ + (y_N - \mu)^2}{N} \quad \ldots(2)$$

ในการประมาณค่า μ เรา draw sample ออกมาหนึ่ง sample ซึ่งมี n observations จาก N observations ของ population เมื่อบวก n observations เข้าด้วยกันแล้วหารด้วย n ก็จะได้ sample mean คือ

$$\bar{y} = \frac{y_1 + y_2 + \ \ldots\ldots\ + y_n}{n} \quad \ldots\ldots\ldots\ldots\ (3)$$

ตัวอย่างเช่น tag population (ข้อ 4.1) ประกอบด้วย 500 observations (N) mean (μ) คือผลบวกของ 500 observations ($\sum y$) หารด้วย 500 (N) mean ของ sample ($\bar{y}$) ที่ประกอบด้วย 10 observations (n) ก็คือ mean ของ 10 observations เท่านั้นที่เอาออกมาจาก 500 observations ในการประมาณค่า σ^2 เราจะใช้ $(y - \mu)^2$ จำนวน n เทอมรวมกันแล้วหารด้วย n กล่าวได้อีกอย่างหนึ่งว่า sample variance คือ

$$V_1 = \frac{(y_1 - \mu)^2 + (y_2 - \mu)^2 + \ \ldots\ldots\ + (y_n - \mu)^2}{n} \quad \ldots\ldots\ldots\ (4)$$

ในเมื่อ

V_1 = sample variance

y's = n observations ที่ได้มาจาก population

การประมาณค่า σ^2 ด้วย V_1 โดยวิธีนี้เป็นวิธีถูกต้องอย่างหนึ่ง และความสมบูรณ์ของมันก็จะถูกแสดงให้เห็นจริงได้โดยตัวอย่างของ population ซึ่งประกอบด้วย 3 observations คือ 2, 4 และ 6 (ข้อ 5.1) mean และ variance ของ population นี้เท่ากับ 4 และ $\frac{8}{3}$ ตามลำดับ ถ้า sample size เท่ากับ 2 หรือ n = 2 จะมี 9 possible samples ซึ่งได้จาก population นี้ samples เหล่านี้คือ 2,2; 2,4; 2,6; ....; 6,6 ทั้ง 9 samples ปรากฏอยู่ในคอลัมน์ (1) ของ Table 7.2 แล้ว variance ของ sample ที่หนึ่งคือ

$$V_1 = \frac{(y_1-\mu)^2+(y_2-\mu)^2}{n}$$

$$= \frac{(2-4)^2+(2-4)^2}{2}$$

$$= 4$$

และ variance ของ sample ที่สองคือ

$$V_1 = \frac{(2-4)^2+(4-4)^2}{2}$$

$$= 2$$

ค่าของ sample variances ทั้ง 9 ได้แสดงไว้ในคอลัมน์ (3) ของ Table 7.2 ผลบวกของ sample variances ทั้ง 9 นี้เท่ากับ 24 และค่าเฉลี่ยของมันเท่ากับ $\frac{24}{9}$ หรือ $\frac{8}{3}$ ซึ่งเป็นค่าของ population variance (σ^2) นั่นเอง ให้สังเกตคอลัมน์ (3) ของ Table 7.2 ว่าไม่มีค่าใดใน 9 ค่าของ V_1 เป็นค่าประมาณที่ถูกต้องของ σ^2 เลย นี่ก็ทำนองเดียวกันกับการใช้ $\bar{y}$ ประมาณค่า μ sample mean แต่ละตัวไม่จำเป็นต้องเท่ากับ population mean (μ) แต่ค่าเฉลี่ยของ sample means ทั้งหมดเท่ากับ μ (คอลัมน์ (2) ของ Table 7.2) ในที่นี้ V_1 แต่ละตัวก็ไม่จำเป็นต้องเท่ากับ σ^2 แต่ค่าเฉลี่ยของค่าของ V_1 ทั้งหมดเท่ากับ σ^2 เพราะฉะนั้นวิธีประมาณค่า population variance อย่างนี้จึงเป็นวิธีที่ถูกต้อง

อย่างไรก็ดีแม้ว่าวิธีประมาณค่า population variance ด้วย V_1 จะเป็นวิธีที่ถูกต้องแต่ก็ไร้ประโยชน์ Equation (4) แสดงให้เห็นว่าในการคำนวณ V_1 ตัวใดนั้นเราจะต้องทราบ population mean (μ) และในการทดสอบสมมติฐานว่า population mean เท่ากับค่าที่กำหนดให้ เราก็จะต้องทราบ population variance หรือ population standard deviation เสียก่อน วิธีที่จะแห่วกวงล้อมซึ่งมองเห็นได้ก็คือการใช้ $\bar{y}$ แทน μ ใน Equation (4) ดังนั้น sample variance ที่แก้ไขแล้วจะเป็น

$$V_2 = \frac{\sum (y - \bar{y})^2}{n}$$

$$= \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots\ldots + (y_n - \bar{y})^2}{n} \quad \ldots\ldots\ldots\ldots\ldots \quad (5)$$

variance ของ sample ที่หนึ่งซึ่งแก้ไขแล้วคือ

$$V_2 = \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2}{n}$$

$$= \frac{(2 - 2)^2 + (2 - 2)^2}{2}$$

$$= 0$$

และของ sample ที่สองคือ

$$V_2 = \frac{(2 - 3)^2 + (4 - 3)^2}{2}$$

$$= 1$$

Table 7.2

sample No.	(1) samples	(2) $\bar{y}$	(3) V_1	(4) V_2	(5) s^2	(6) s
1	2, 2	2	4	0	0	0.000
2	2, 4	3	2	1	2	1.414
3	2, 6	4	4	4	8	2.828
4	4, 2	3	2	1	2	1.414
5	4, 4	4	0	0	0	0.000
6	4, 6	5	2	1	2	1.414
7	6, 2	4	4	4	8	2.828
8	6, 4	5	2	1	2	1.414
9	6, 6	6	4	0	0	0.000
	total	36	24	12	24	11.312
	average	36/9 = 4	24/9 = 8/3	12/9 = 4/3	24/9 = 8/3	1.257
	parameter	$\mu = 4$	$\sigma^2 = 8/3$	$\sigma^2 = 4/3$	$\sigma^2 = 8/3$	$\sigma = 1.633$

ค่าต่าง ๆ ของ V_2 ของ 9 samplesได้ให้ไว้ในคอลัมน์ (4) ของ Table 7.2 เราจะเห็นได้ว่า V_2 ไม่ใช่ค่าประมาณที่ดีของ σ^2 ไม่มี V_2 ตัวใดเลยที่มีค่าเท่ากับ σ^2 และแม้แต่ค่าเฉลี่ยของมันก็ไม่เท่ากับ σ^2 ด้วย ค่าเฉลี่ยของ V_2 9 ตัวเท่ากับ $\frac{12}{9}$ หรือ $\frac{4}{3}$ ในขณะที่ $\sigma^2 = \frac{8}{3}$ หรืออีกนัยหนึ่ง V_2 ประมาณค่า σ^2 ต่ำไป เพราะฉะนั้นวิธีนี้จึงดูเหมือนจะบกพร่อง อย่างไรก็ดีเราจะได้ค่าของ V_2 ง่าย เพราะเราไม่ต้องการ population mean ในการคำนวณ ข้อเท็จจริงที่ว่ามันประมาณค่า σ^2 ต่ำไป ก็ไม่ได้ทำให้มันไร้ประโยชน์เสียทีเดียว ข้อบกพร่องในการประมาณค่าต่ำไปนี้อาจแก้ไขได้ การแก้ไขจะทำได้สำเร็จโดยการใช้ $\boxed{n - 1 \text{ แทน } n}$ เป็นตัวหารใน Equation (5) เหตุผลที่ใช้ตัวหารมีค่าน้อยลงก็เพื่อที่จะเบ่งค่า V_2 ขึ้นจนมันไม่ประมาณค่า σ^2 ต่ำไปอีก

V_2 ที่แก้ไขแล้วคือ

$$s^2 = \frac{\sum (y - \bar{y})^2}{n - 1}$$

$$= \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots\ldots + (y_n - \bar{y})^2}{n - 1} \quad \ldots\ldots\ldots\ldots\ldots (6)$$

ในเมื่อ

$y_1, y_2, \ldots\ldots\ldots, y_n$ เป็น n observationsของ sample หนึ่ง และ $\bar{y}$ เป็น sample mean ของ sample นั้น

s^2 ของ sample ที่หนึ่งของ Table 7.2 คือ

$$s^2 = \frac{(2 - 2)^2 + (2 - 2)^2}{2 - 1}$$

$$= \frac{0}{1}$$

$$= 0$$

และของ sample ที่สองคือ

$$s^2 = \frac{(2 - 3)^2 + (4 - 3)^2}{2 - 1}$$

$$= \frac{2}{1}$$

$$= 2$$

ค่าต่าง ๆ ของ s^2 ของ 9 samples ได้ให้ไว้ในคอลัมน์ (5) ของ Table 7.2 ค่าเฉลี่ยของ s^2 9 ตัวนี้เท่ากับ $\frac{24}{9} = \frac{8}{3}$ ซึ่งเป็นค่าของ population variance การใช้ n - 1 เป็นตัวหารแทน n (Equations (5) และ (6)) ได้แก้ไขการประมาณค่าต่ำไปให้ถูกต้องแล้ว ต่อแต่นี้ไปในหนังสือนี้เราจะใช้ s^2 เท่านั้นเป็น sample variance ความได้เปรียบของ s^2 คือค่าเฉลี่ยของ sample variance ทั้งหมดเท่ากับ σ^2 และเราไม่ต้องการทราบ population mean เลย theorem ต่อไปนี้จะเป็นประโยชน์สำหรับการอ้างอิงในภายหน้า

ซึ่งทำให้แต่ละ sample สามารถผลิตค่าประมาณของ population variance (σ^2) mean ของ sample variancesทั้งหมดเหล่านี้ (คอลัมน์ (5) ของ Table 7.2) เท่ากับ population variance (σ^2) เพราะฉะนั้นโดยเฉลี่ย sample variances เป็น unbiased แต่ sample variance (s^2) แต่ละตัวอาจเท่าหรือไม่เท่ากับ population variance (σ^2) ก็ได้ ไม่มี sample variance (s^2) ตัวใดที่แสดงไว้ในคอลัมน์ (5) ของ Table 7.2 เท่ากับ population variance (σ^2) เมื่อเราใช้คำ "bias" ในสถิติศาสตร์ มันจะถูกใช้ในความหมายดังกล่าวนี้ เมื่อเราเรียกค่าประมาณตัวหนึ่งว่า unbiased ก็เพราะเหตุที่ว่าค่าเฉลี่ยของค่าประมาณต่าง ๆ ที่ถูกผลิตขึ้นโดย all possible samples ซึ่งมี size อย่างหนึ่งที่กำหนดให้เท่ากับ parameter มากกว่าจะเป็นเพราะค่าประมาณที่ผลิตโดย sample หนึ่งโดยเฉพาะเท่ากับ parameter

7.4 Computing Method of Sample Variance

sample variance ได้ถูกนิยามไว้แล้วใน Equation (6) ข้อ 7.2 คือ

$$s^2 = \frac{(y_1-\bar{y})^2 + (y_2-\bar{y})^2 + \ldots\ldots + (y_n-\bar{y})^2}{n-1}$$

$$= \frac{\Sigma(y-\bar{y})^2}{n-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Equation นี้ใช้แสดงความหมายของ sample variance มากกว่าจะใช้คำนวณค่าของ s^2 ในข้อนี้จะกล่าวถึงการคำนวณค่าของ s^2 โดยวิธีลัดซึ่งใช้เครื่องคำนวณเลขได้สะดวก งานอันน่าเบื่อที่เกี่ยวกับการคำนวณค่าของ s^2 คือการหาจำนวนที่เป็นเศษหรือ $\Sigma(y-\bar{y})^2$ ใน Equation (1) นั่นเอง เพราะฉะนั้นวิธีลัดจึงเกี่ยวข้องกับจำนวนที่เป็นเศษซึ่งเป็นผลบวกของกำลังสอง (sum of squares) ของความแปรผันของ observations ต่าง ๆ จาก mean ของมัน ในหนังสือนี้เราจะใช้อักษรย่อ SS แทน sum of squares หรือจำนวนที่เป็นเศษ ดังนั้น

$$SS = \Sigma(y-\bar{y})^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

วิธีลัดในการคำนวณ SS คือ

$$\boxed{SS = \Sigma y^2 - \frac{(\Sigma y)^2}{n}} \quad \ldots\ldots\ldots (3)$$

Theorem 7.2: If all possible samples of size n are drawn from a population, the mean of all sample variances (s^2), where

$$\boxed{s^2 = \frac{\Sigma (y - \bar{y})^2}{n-1}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

is equal to the population variance (σ^2).

7.3 Unbiased Estimate

ถ้า mean ของ possible values ทั้งหมดของ statistic ตัวหนึ่งเท่ากับ parameter ตัวหนึ่ง (ข้อ 4.4) แล้ว เราเรียก statistic ตัวนั้นว่าเป็น unbiased estimate ของ parameter ตัวนั้น ตัวอย่างเช่น sample mean ($\bar{y}$) เป็น unbiased estimate ของ population mean (μ) เพราะว่า mean ของ means ของ all possible samples ซึ่งมี size ที่กำหนดให้เท่ากับ population mean (Theorem 5.3) unbiased estimate อีกตัวหนึ่งคือ s^2 เพราะว่า mean ของ variances (s^2) ของ all possible samples ซึ่งมี size ที่กำหนดให้เท่ากับ population variance (σ^2) (Theorem 7.2)

ให้สังเกตว่าวิธีคำนวณ sample mean ($\bar{y}$) คือ

$$\bar{y} = \frac{y_1 + y_2 + \ldots\ldots + y_n}{n}$$

$$= \frac{\Sigma y}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

ซึ่งทำให้แต่ละ sample สามารถผลิตค่าประมาณของ population mean (μ) mean ของ means ของ all possible samples ($\mu_{\bar{y}}$) เท่ากับ population mean (μ) เพราะฉะนั้นโดยเฉลี่ย sample means เป็น unbiased แต่ sample mean ($\bar{y}$) แต่ละตัวอาจเท่า หรือไม่เท่ากับ population mean (μ) ก็ได้ เราจะพบสถานะที่คล้ายคลึงกันนี้ใน sample variance (s^2) วิธีคำนวณ s^2 คือ

$$s^2 = \frac{(y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots\ldots + (y_n - \bar{y})^2}{n - 1}$$

$$= \frac{\Sigma (y - \bar{y})^2}{n - 1} \quad \ldots\ldots\ldots\ldots\ldots \quad (2)$$

ความจริงที่ว่าการคำนวณ SS โดยสองวิธีนี้ให้ผลลัพท์เหมือนกันจะถูกแสดงให้เห็นโดย 5 observations 3, 2, 1, 3, 1 พร้อมด้วยรายละเอียดของการคำนวณที่ได้แสดงไว้ใน Table 7.4

Table 7.4

(1)	(2)	(3)	(4)
y	$y - \bar{y}$	$(y - \bar{y})^2$	y^2
3	1	1	9
2	0	0	4
1	–1	1	1
3	1	1	9
1	–1	1	1
10	0	4	24
Σy	$\Sigma(y - \bar{y})$	$\Sigma(y - \bar{y})^2$	Σy^2

$$n = 5; \qquad \bar{y} = \frac{\Sigma y}{n} = 2$$

$$\Sigma(y - \bar{y})^2 = 4; \qquad s^2 = \frac{\Sigma(y - \bar{y})^2}{n - 1} = 1$$

$$\Sigma y^2 - \frac{(\Sigma y)^2}{n} = 24 - \frac{(10)^2}{5} = 24 - 20 = 4$$

ค่าของ SS ที่คำนวณโดย Equation (2) เท่ากับ 4 และที่คำนวณโดย Equation (3) ก็เท่ากับ 4 เหมือนกันคือ

$$\begin{aligned} SS &= \Sigma y^2 - \frac{(\Sigma y)^2}{n} \\ &= 24 - \frac{(10)^2}{5} \\ &= 24 - 20 \\ &= 4 \end{aligned}$$

ดังนั้น $$\Sigma(y - \bar{y})^2 = \Sigma y^2 - \frac{(\Sigma y)^2}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

Equation (4) นี้เป็น algebraic identity ซึ่งเป็นความจริง สำหรับชุดของจำนวนเลขใด ๆ ก็ได้ เพราะฉะนั้นในบทหลัง ๆ เราจึงใช้ SS ทั้งสองที่ได้แสดงไว้ใน Equations (2) และ (3) แทนกันอยู่เสมอ การพิสูจน์ทางพีชคณิตว่า $\Sigma(y-\bar{y})^2 = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$ จะได้ให้ไว้ในข้อ 7.8

ภายหลังที่เราคำนวณค่า SS แล้ว เราจะได้ค่า sample variance โดยเอา n - 1 ไปหาร SS สำหรับตัวอย่างนี้ $s^2 = \frac{4}{4} = 1$

เราควรสนใจสังเกตคอลัมน์ (2) ของ Table 7.4 ว่า ผลบวกของความแปรผันใน observations ต่าง ๆ จาก mean ของมันเท่ากับศูนย์ นั่นคือ

$$\Sigma(y-\bar{y}) = (y_1-\bar{y}) + (y_2-\bar{y}) + \ldots\ldots + (y_n-\bar{y}) = 0 \quad \ldots\ldots\ldots (5)$$

Equation (5) นี้ก็เป็น algebraic identity ด้วยเช่นเดียวกัน ซึ่งจะเป็นความจริงสำหรับชุดของจำนวนเลขใด ๆ ก็ได้ ในบทหลัง ๆ เราจะใช้ identity นี้บ่อย ๆ เพื่อพัฒนาวิธีคำนวณต่าง ๆ โดยทางลัด การพิสูจน์ทางพีชคณิตของมันจะได้ให้ไว้ในข้อ 7.8

7.5 X^2- Distribution

normal distribution ที่ได้บรรยายไว้ในบทที่ 3 เป็น distribution ที่สำคัญที่สุดอย่างหนึ่งในวิชาสถิติ ในข้อนี้จะได้บรรยายถึง frequency distribution อีกอย่างหนึ่งซึ่งเรียกว่า X^2-distribution, distribution นี้เกี่ยวข้องกับ normal distribution อย่างใกล้ชิด

ถ้า all possible samples ซึ่งมี size n ได้มาจาก normal population หนึ่งซึ่งมี mean เท่ากับ μ และ variance เท่ากับ σ^2 จากทุก sample จะคำนวณ sample mean ($\bar{y}$) ได้ Theorem 5.2 b กล่าวว่า distribution ของ sample means เหล่านี้เป็น normal distribution อย่างไรก็ดีไม่แต่เพียง sample mean เท่านั้นที่เราคำนวณได้จากทุก sample แต่เรายังคำนวณ statistics อื่น ๆ เช่น ผลบวกของ observations (Σy) และ sample variance (s^2) ได้อีกด้วย สำหรับแต่ละ sample ถ้าเราคำนวณ statistic

$$\Sigma u^2 = \left(\frac{y_1-\mu}{\sigma}\right)^2 + \left(\frac{y_2-\mu}{\sigma}\right)^2 + \ldots\ldots + \left(\frac{y_n-\mu}{\sigma}\right)^2$$

$$= \frac{\Sigma(y-\mu)^2}{\sigma^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

(Equation(1) ข้อ 3.2) ออกมา ค่าของ Σu^2 จะเปลี่ยนไปต่าง ๆ กันตาม sample เช่นเดียวกับค่าของ statistic อื่น ๆ การขึ้น ๆ ลง ๆ ในค่าของ Σu^2 จะเห็นได้จาก 4 random samples ที่ได้ให้ไว้ใน Table 4.2 samples เหล่านี้ได้มาจาก tag population ซึ่งเป็น normal population ที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 (ข้อ 4.1) sample ที่หนึ่งประกอบด้วย observations 50, 57, 42, 63 และ 32 และค่าของ Σu^2 ของ sample นี้คือ

$$\begin{aligned}
\Sigma u^2 &= \left(\frac{50-50}{10}\right)^2+\left(\frac{57-50}{10}\right)^2+\left(\frac{42-50}{10}\right)^2+\left(\frac{63-50}{10}\right)^2+\left(\frac{32-50}{10}\right)^2 \\
&= \frac{(50-50)^2+(57-50)^2+(42-50)^2+(63-50)^2+(32-50)^2}{100} \\
&= \frac{0+49+64+169+324}{100} \\
&= \frac{606}{100} \\
&= 6.06
\end{aligned}$$

sample ที่สองประกอบด้วย observations 55, 44, 37, 40, และ 52 และค่าของ Σu^2 คือ

$$\begin{aligned}
\Sigma u^2 &= \frac{(55-50)^2+(44-50)^2+(37-50)^2+(40-50)^2+(52-50)^2}{100} \\
&= \frac{25+36+169+100+4}{100} \\
&= \frac{334}{100} \\
&= 3.34
\end{aligned}$$

ถ้าเราหา all possible samples จาก normal population หนึ่ง แต่ละ sample ก็จะมีค่า Σu^2 ของมัน <u>distribution ของค่าของ Σu^2 เหล่านี้เรียกว่า X^2- distribution</u>

ค่าของ Σu^2 ไม่แต่เพียงจะอยู่ในอิทธิพลของความเปลี่ยนแปลงในค่าของ observations ตาม sample เท่านั้น แต่จะอยู่ในอิทธิพลของ sample size (n) ด้วย ตัวอย่างเช่น ถ้า sample ที่หนึ่งประกอบ 2 observations แรกของมันคือ 50 และ 57 แล้ว ค่าของ Σu^2 คือ

$$\Sigma u^2 = \frac{(50-50)^2+(57-50)^2}{100}$$

$$= \frac{0 + 49}{100}$$

$$= 0.49$$

แทนที่จะเป็น 6.06 (ในกรณีที่มี 5 observations) และถ้า sample ที่สองประกอบด้วย 2 observations แรกของมันคือ 55 และ 44 แล้ว ค่าของ Σu^2 คือ

$$\Sigma u^2 = \frac{(55-50)^2+ (44-50)^2}{100}$$

$$= \frac{25 + 36}{100}$$

$$= 0.61$$

แทนที่จะเป็น 3.34 (ในกรณีที่มี 5 observations) โดยเฉลี่ยแล้วค่าของ Σu^2 จะมากขึ้นถ้า sample size โตขึ้น เพราะฉะนั้นแต่ละ sample size ก็จะมี X^2 – distribution ต่างกัน ซึ่ง mean ของ distribution จะเพิ่มขึ้นตาม sample size ผลที่เกิดขึ้นคือ X^2 – distribution ไม่ใช่ frequency curve เดียวเพียง curve เดียว แต่เป็น family of curves สิ่งที่บ่งให้เห็นอย่างเด่นชัดของ curve เฉพาะเส้นหนึ่งคือ mean ของ distribution เราเรียก mean ของ distribution ที่ถูกเขียนแทนด้วย ν (nu) นี้ว่า the number of degrees of freedom (ซึ่งใช้ตัวย่อว่า d.f. หรือ D.F.) ของ X^2– distribution curves ต่าง ๆ ของ X^2 – distribution with 1, 4 และ 5 d.f. ได้แสดงไว้ใน Fig. 7.5 a

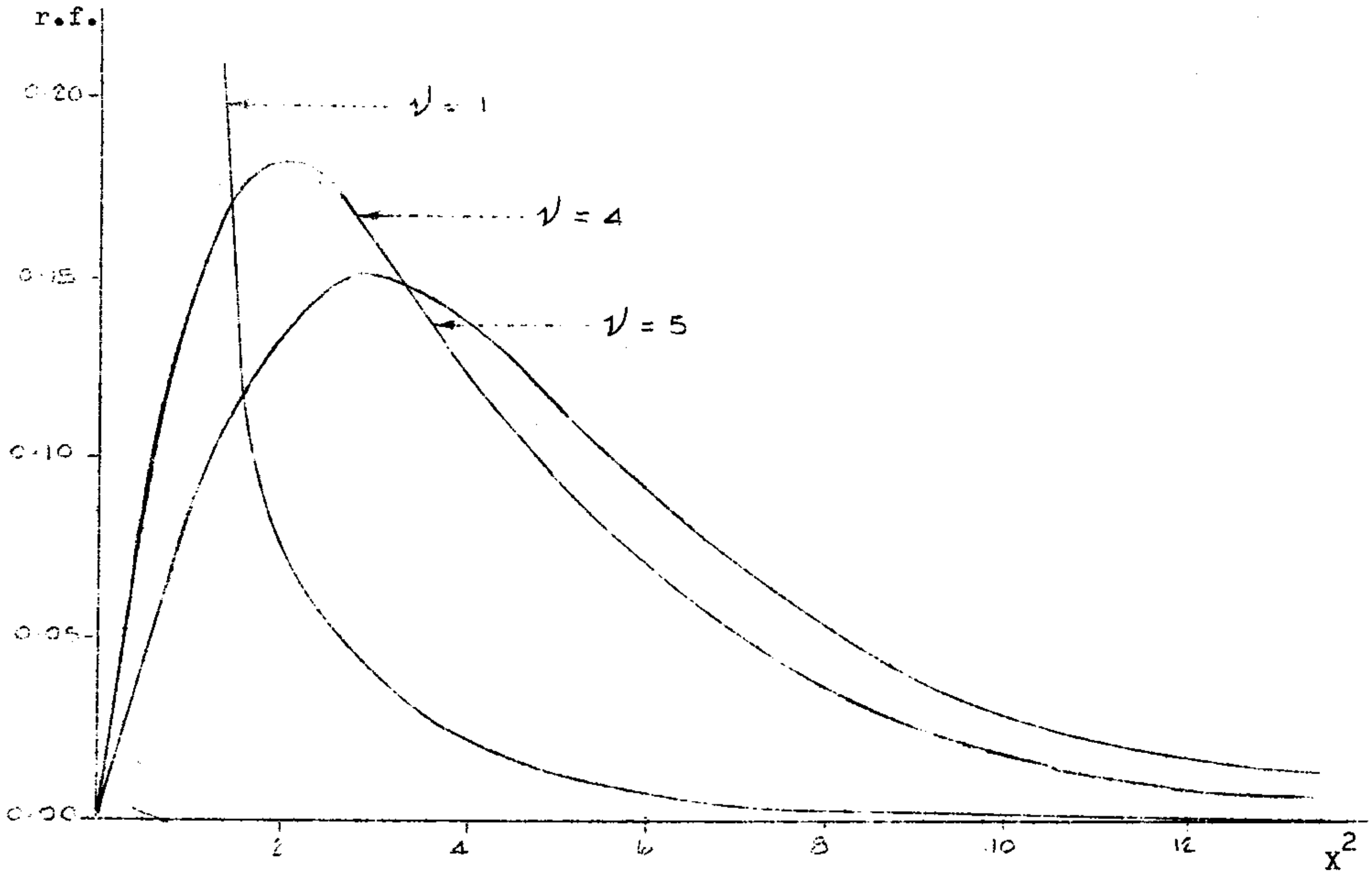


Fig. 7.5 a

การอภิปรายในข้อนี้อาจรวบรวมไว้ใน theorem ต่อไปนี้

<u>Theorem 7.5:</u> If all possible samples of size n are drawn from a normal population with mean equal to μ and variance equal to σ^2, and for each sample Σu^2 is computed, where

$$\Sigma u^2 = \Sigma \left(\frac{y-\mu}{\sigma}\right)^2 = \frac{\Sigma (y-\mu)^2}{\sigma^2}, \ldots\ldots\ldots\ldots \quad (2)$$

the frequency distribution of Σu^2 follows the X^2- distribution with n degrees of freedom (that is, ν = n).

เราจะต้องใช้คณิตศาสตร์อย่างละเอียดและยุ่งยากจึงจะได้ X^2- distributionจาก normal population อย่างไรก็ดีเราอาจแสดงความเป็นจริงของ theorem โดยการทดลองได้ตาม sampling experiment ที่ได้กล่าวไว้ในบทที่ 4 กล่าวโดยย่อเรามี 1,000 samples แต่ละ sample ประกอบด้วย 5 observations ได้มาจาก tag population ซึ่งเป็น normal population ที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 จากแต่ละ sample เราคำนวณ Σu^2เอาไว้ ค่าของ Σu^2 ของ 4 samples ดังกล่าวได้แสดงไว้ใน Table 4.2 frequency table ของค่าของ Σu^2 1,000 ค่าได้ให้ไว้ใน Table 7.5 ซึ่งได้แสดงทั้ง theoretical และ observed relative frequencies ไว้ด้วยทั้งสองอย่าง

Table 7.5

X^2 or Σu^2	observed frequency		theoretical	midpoint	mf
	f	r.f. (%)	r.f. (%)	m	
0 – 1	35	3.5	3.7	0.5	17.5
1 – 2	109	10.9	11.3	1.5	163.5
2 – 3	148	14.8	14.9	2.5	370.0
3 – 4	171	17.1	15.1	3.5	598.5
4 – 5	139	13.9	13.4	4.5	625.5
5 – 6	106	10.6	11.1	5.5	583.0
6 – 7	77	7.7	8.6	6.5	500.5
7 – 8	64	6.4	6.4	7.5	480.0
8 – 9	53	5.3	4.7	8.5	450.5
9 – 10	28	2.8	3.4	9.5	266.0
10 – 11	18	1.8	2.4	10.5	189.0
11 – 12	20	2.0	1.7	11.5	230.0
12 – 13	11	1.1	1.1	12.5	137.5
13 – 14	6	0.6	0.8	13.5	81.0
14 – 15	8	0.8	0.5	14.5	116.0
over 15	7	0.7	1.0	18.1	126.7
total	1,000	100.0	100.1		4,935.2

$$\text{mean of } X^2 = \frac{\Sigma mf}{\Sigma f} = \frac{4,935.2}{1,000} = 4.9352$$

theoretical relative frequency เป็น frequency ที่จะเกิดขึ้น ถ้าเราได้ all possible samples ซึ่งมี size 5 จาก tag population และ observed frequency เป็น frequency ที่ได้จาก 1,000 samples เท่านั้น จาก Table 7.5 เราจะเห็นได้ว่า theoretical และ observed frequencies ใกล้เคียงกัน แต่มันไม่เท่ากันอย่างแท้จริง theoretical frequency ขึ้นอยู่กับ all possible samples ซึ่งมี size 5 แต่ observed frequency ขึ้นอยู่กับ 1,000 samples เพราะฉะนั้นเราจึงไม่หวังว่า theoretical และ observed frequencies จะพ้องกันอย่างสมบูรณ์ ความพ้องกันอย่างใกล้ชิดระหว่าง theoretical และ observed frequencies ยังเห็นได้อีกใน Fig. 7.5 b ซึ่งแสดง histogram ของ distribution ของ Σu^2 1,000 ค่า และ theoretical frequency curve ของ X^2- distribution with 5 d.f. ด้วย

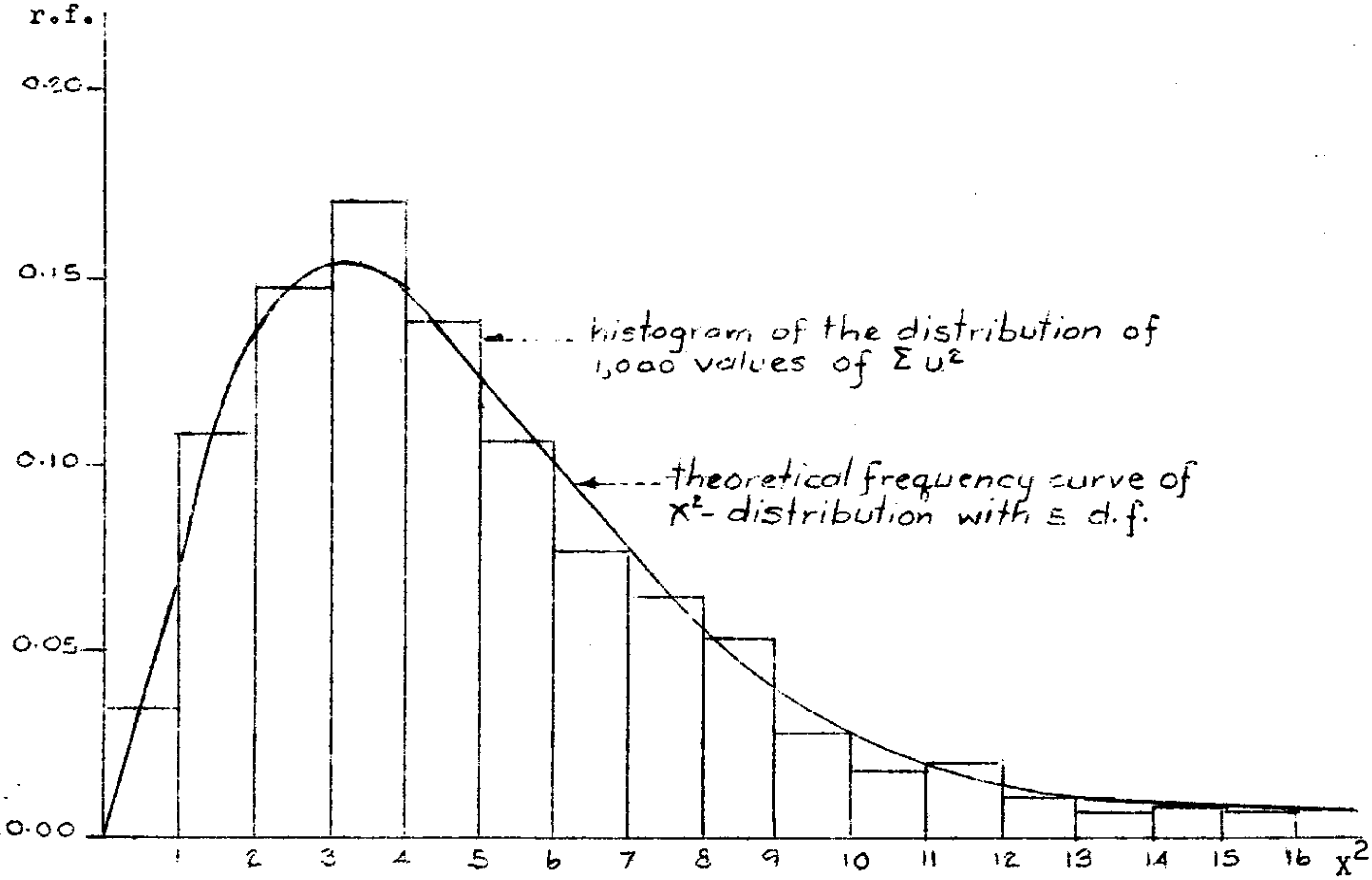


Fig. 7.5 b

ข้อเท็จจริงที่ว่า mean ของ Σu^2 เท่ากับ sample size (n) จะถูกแสดงให้เห็นได้โดย Table 7.5 mean ของ Σu^2 1,000 ค่าจะหาได้ง่าย อย่างไรก็ดีภายหลังที่ทำ Table 7.5 นี้ขึ้นจะไม่ปรากฏค่า Σu^2 แต่ละค่า ดังนั้นเราจะหา mean โดยประมาณได้จาก frequency table โดยพิจารณาว่าค่าของ Σu^2 ใด ๆ ใน class 0 – 1 เป็น 0.5 ใน class 1 – 2 เป็น 1.5 และต่อ ๆ ไป จุดกึ่งกลางของ classes ต่าง ๆ เหล่านี้ใช้แทนด้วยอักษร m ใน Table 7.5 แต่สำหรับ class "over 15" ได้ใช้ mean 18.1 ของ Σu^2 7 ค่าของ class นั้นเป็น m ดังนั้น mean โดยประมาณของ Σu^2 คือ

$$\frac{\Sigma mf}{\Sigma f} = \frac{4,935.2}{1,000} = 4.9352 \approx 5$$

ซึ่งเท่ากับ sample size 5 โดยประมาณ ถ้าเรา draw all possible samples ออกมาจาก normal population เราอาจพิสูจน์ทางคณิตศาสตร์ได้ว่า mean ของ distribution เท่ากับ 5 อย่างแท้จริง การแสดงให้เห็นจริงโดยการทดลองของ Theorem 7.5 จึงสมบูรณ์แล้ว

เช่นเดียวกับตารางแสดง relative cumulative frequency (r.c.f.)สำหรับ normal distribution (ข้อ 3.2) เราอาจทำตารางสำหรับ X^2– distribution ขึ้นใช้ได้ ตารางโดยย่อของ X^2– distribution ได้แสดงไว้ใน Table 4, Appendix แต่ละบรรทัดของตารางแทน d.f. ต่าง ๆ กัน เช่น 1, 2, ..., 100 ซึ่งแสดงไว้ในคอลัมน์ซ้ายสุด แต่ละคอลัมน์แสดงค่าเปอร์เซนต์ต่าง ๆ กัน ตัวอย่างเช่นค่าในตารางตาม 10 d.f. และ 5 % คือ 18.3070 ซึ่งหมายความว่า 5 % ของค่าของ X^2 with 10 d.f.มากกว่า 18.3070 ค่าของ X^2 ตาม 5 d.f. และ 5 % คือ 11.0705 ซึ่งหมายความว่า 5 % ของค่าของ X^2 with 5 d.f.มากกว่า 11.0705 ค่า 11.0705 นี้อาจถูกเปรียบเทียบกับค่าที่ได้รับโดย sampling experiment ได้ จาก Table 7.5 เราจะเห็นได้ว่า 5.2 % (คือ 2.0 +1.1 +0.6 +0.8+ 0.7) ของค่าของ Σu^2 1,000 ค่ามากกว่า 11, 5.2 % นี้เท่ากับ 5 % โดยประมาณตามที่คาดหมายไว้

7.6 Distribution of u^2

จาก Theorem 7.5 เรายังทราบต่อไปอีกว่า u^2 (คือ Σu^2 เมื่อ n = 1) follows the X^2– distribution with 1 d.f. distribution curves ของ u และ u^2 ได้แสดงไว้ใน Fig. 7.6 a และ 7.6 b

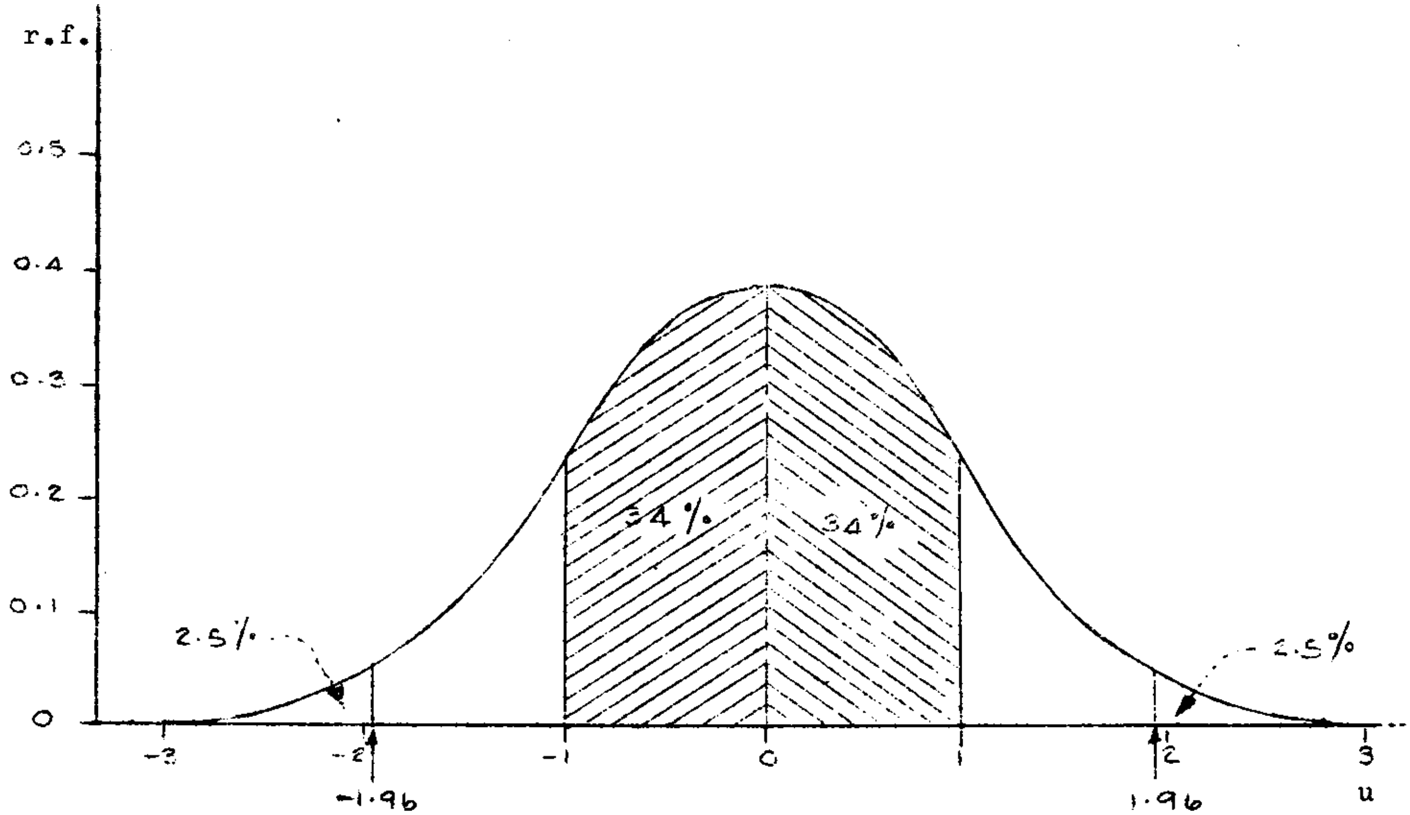


Fig. 7.6 a

distribution ของ u เป็น normal distribution ซึ่งมี mean เท่ากับศูนย์ (0) และ variance เท่ากับ 1 distribution ของ X^2 with 1 d.f. คือ u^2 เป็นการยกกำลังสองของ u (Fig. 7.6 a และ 7.6 b) เพราะว่า $(-u)^2 = u^2$

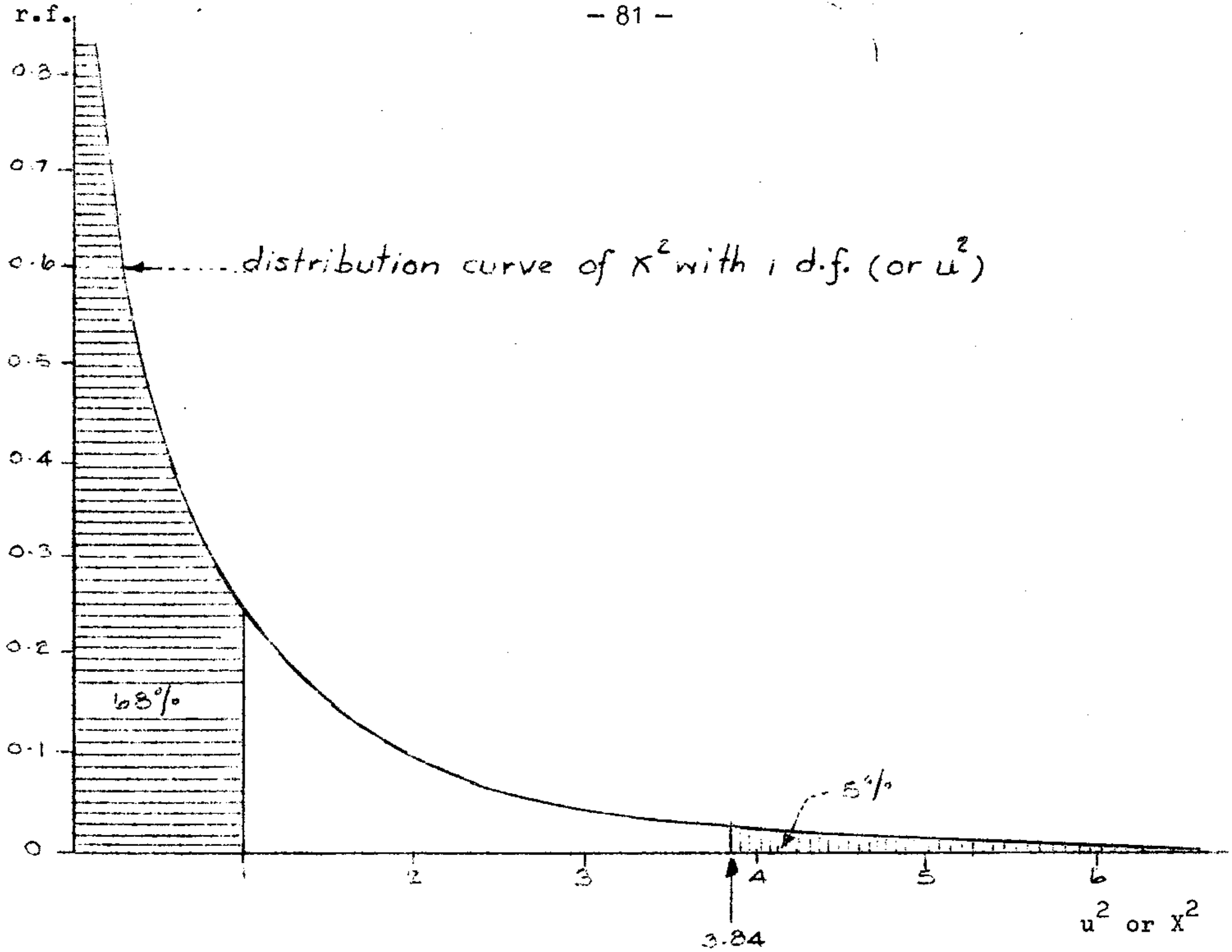


Fig. 7.6 b

กำลังสองของค่าใด ๆ ระหว่าง 0 และ 1 จะเป็นค่าระหว่าง 0 และ 1 และเช่นเดียวกันกำลังสองของค่าใด ๆ ระหว่าง 0 และ − 1 จะเป็นค่าระหว่าง 0 และ 1 ด้วย ตัวอย่างเช่น $(0.5)^2 = 0.25$, $(-0.5)^2 = 0.25$ 68 % ของค่าต่าง ๆ ของ u ตกอยู่ระหว่าง − 1 และ 1 ยังผลให้เกิด 68 % ของค่าต่าง ๆ ของ u^2 ตกอยู่ระหว่าง 0 และ 1 กำลังสองของค่าใด ๆ ที่มากกว่า 1.96 จะมากกว่า $(1.96)^2$ หรือ 3.84 และกำลังสองของค่าใด ๆ ที่น้อยกว่า − 1.96 จะมากกว่า 3.84 ด้วย ตัวอย่างเช่น 2 มากกว่า 1.96 และ − 2 น้อยกว่า − 1.96 แต่ $(2)^2 = 4$ มากกว่า $(1.96)^2$ หรือ 3.84 และ $(-2)^2 = 4$ ก็มากกว่า $(-1.96)^2$ หรือ 3.84 ด้วย เนื่องจากมี 2.5 % ของค่าของ u

มากกว่า 1.96 และ 2.5 % ของค่าของ u น้อยกว่า − 1.96 5 % ทั้งหมดของ u^2 จึงมากกว่า 3.84 เพราะฉะนั้นจะมี 68 % ของค่าของ X^2 ระหว่าง 0 และ 1 และ 5 % ของค่าของ X^2 มากกว่า 3.84

จากรูป 7.6 a และ 7.6 b ควรสังเกตว่าส่วนกลางของ curve ของ u กลายเป็นหางข้างซ้ายของ curve ของ X^2 with 1 d.f. ข้อเท็จจริงที่ว่า u^2 follows the X^2- distribution with 1 d.f. จะถูกพูดถึงบ่อย ๆ ในบทต่อ ๆ ไป การอภิปรายข้างต้นเกี่ยวกับ u และ u^2 นั้น อาจรวมรวมไว้ใน Theorem ต่อไปนี้

Theorem 7.6: If a statistic u follows the normal distribution with mean equal to zero and variance equal to 1, u^2 follows the X^2- distribution with 1 degree of freedom (that is $\nu = 1$).

7.7 Distribution of SS/σ^2

ใน 2 ข้อที่แล้วมา เราทราบแล้วว่า

$$\Sigma u^2 = \frac{\Sigma (y-\mu)^2}{\sigma^2} = \frac{(y_1-\mu)^2 + (y_2-\mu)^2 + \ldots + (y_n-\mu)^2}{\sigma^2} \quad \ldots (1)$$

follows the X^2- distribution with n degrees of freedom และ u^2 follows the X^2- distribution with 1 degree of freedom ในข้อนี้จะแสดงให้เห็นว่า SS/σ^2 follows the X^2- distribution with n - 1 degrees of freedom ในเมื่อ

$$\frac{SS}{\sigma^2} = \frac{\Sigma (y-\bar{y})^2}{\sigma^2} = \frac{(y_1-\bar{y})^2 + (y_2-\bar{y})^2 + \ldots + (y_n-\bar{y})^2}{\sigma^2} \quad \ldots\ldots\ldots (2)$$

Σu^2 และ $\frac{SS}{\sigma^2}$ สองตัวนี้ไม่ใช่ตัวเดียวกัน เราสังเกตจาก Equations (1) และ (2) ได้ว่า Σu^2 เกี่ยวข้องกับความแปรผันของ observations ของ sample หนึ่งจาก population mean แต่ $\frac{SS}{\sigma^2}$ เกี่ยวข้องกับความแปรผันของ observations ของ sample หนึ่งจาก sample mean ของมัน ดังนั้นมันจึงไม่ใช่ตัวเดียวกัน แต่มันมีความสัมพันธ์กันดังนี้

$$\boxed{\frac{\Sigma (y-\mu)^2}{\sigma^2} = \frac{\Sigma (y-\bar{y})^2}{\sigma^2} + \frac{n(\bar{y}-\mu)^2}{\sigma^2}} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

การพิสูจน์ทางพีชคณิตของ identity ข้างบนนี้จะได้ให้ไว้ในข้อ 7.8 ในข้อนี้ความสัมพันธ์ดังกล่าวจะถูกแสดงให้เห็นจริงเป็นเลขจำนวน ตัวอย่างเช่น sample หนึ่งมี 5 observations คือ 50, 57, 42, 63, 32 ได้มาจาก normal population หนึ่งที่มี mean (μ) เท่ากับ 50 และ variance (σ^2) เท่ากับ 100 แต่ละจำนวนของ 3 จำนวนใน Equation (3) เราคำนวณค่าของมันได้ดังรายละเอียดของการคำนวณที่ได้แสดงไว้ใน Table 7.7 a

Table 7.7 a

y	$y-\mu$	$(y-\mu)^2$	$y-\bar{y}$	$(y-\bar{y})^2$
50	0	0	1.2	1.44
57	7	49	8.2	67.24
42	− 8	64	− 6.8	46.24
63	13	169	14.2	201.64
32	− 18	324	− 16.8	282.24
244 Σy		606 $\Sigma(y-\mu)^2$	0 $\Sigma(y-\bar{y})$	598.80 SS
$n = 5;\ \bar{y} = 48.8;\ \mu = 50;\ (\bar{y}-\mu)^2 = 1.44$				

ผลการคำนวณ คือ

$$\frac{606}{\sigma^2} = \frac{598.80}{\sigma^2} + \frac{5(1.44)}{\sigma^2}$$

เนื่องจาก 606 = 598.80 + 7.20 ดังนั้น sample นี้ได้แสดงให้เห็นจริงใน identity ที่แสดงไว้ใน Equation (3) แล้ว

ข้างซ้ายของ Equation (3) คือ $\frac{\Sigma(y-\mu)^2}{\sigma^2}$ หรือ Σu^2 นั้นได้ถูกแสดงให้เห็นว่า follows the X^2- distribution with n degrees of freedom ในข้อ 7.5 เทอมสุดท้ายทางขวาของ Equation (3) คือ

$$\frac{n(\bar{y} - \mu)^2}{\sigma^2} = \frac{(\bar{y} - \mu)^2}{\frac{\sigma^2}{n}} = \left(\frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 = u^2 \text{ .... } (4)$$

(Equation (1) ข้อ 6.7) ก็ถูกแสดงให้เห็นว่า follows the X^2- distribution with 1 degree of freedom ในข้อ 7.6 ดังนั้นจึงมีเหตุผลที่จะคาดหมายได้ว่าเทอมกลาง

$$\frac{\Sigma(y - \bar{y})^2}{\sigma^2} = \frac{SS}{\sigma^2} \text{ .... } (5)$$

ของ Equation (3) follows the X^2- distribution with n - 1 degrees of freedom ความคาดหมายนี้จะถูกแสดงให้เห็นจริงโดย sampling experiment ที่ได้บรรยายไว้ในบทที่ 4 กล่าวโดยย่อ เรามี 1,000 samples แต่ละ sample ประกอบด้วย 5 observations ได้มาจาก normal population หนึ่งที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 สำหรับแต่ละ sample เราคำนวณค่าของ SS ไว้โดยวิธีที่ให้ไว้ในข้อ 7.4 ค่าของ SS ของ 4 samples ดังกล่าวได้ให้ไว้ใน Table 4.2 เนื่องจาก σ^2 เท่ากับ 100 จึงได้ค่า $\frac{SS}{\sigma^2}$ โดยง่าย frequency table ของ $\frac{SS}{\sigma^2}$ 1,000 ค่าได้ให้ไว้ใน Table 7.7 b

Table 7.7 b

SS/σ^2	observed frequency		theoretical	midpoint	mf
	f	r.f.(%)	r.f.(%)	m	
0 – 1	93	9.3	9.0	0.5	46.5
1 – 2	181	18.1	17.4	1.5	271.5
2 – 3	189	18.9	17.8	2.5	472.5
3 – 4	152	15.2	15.2	3.5	532.0
4 – 5	116	11.6	11.9	4.5	522.0
5 – 6	86	8.6	8.8	5.5	473.0
6 – 7	64	6.4	6.3	6.5	416.0
7 – 8	35	3.5	4.4	7.5	262.5
8 – 9	38	3.8	3.1	8.5	323.0
9 – 10	21	2.1	2.1	9.5	199.5
11 – 11	8	0.8	1.4	10.5	84.0
11 – 12	7	0.7	0.9	11.5	80.5
12 – 13	3	0.3	0.6	12.5	37.5
over 13	7	0.7	1.1	16.4	114.8
total	1,000	100.0	100.0		3,835.3

mean of $\frac{SS}{\sigma^2} = \frac{\Sigma\, mf}{\Sigma\, f} = \frac{3,835.3}{1,000} = 3.8353$ หรือประมาณ 4
ซึ่งเท่ากับ n - 1 หรือ 5 - 1

histogram ของ $\frac{SS}{\sigma^2}$ 1,000 ค่า และ X^2- curve with 4 d.f. ได้ให้ไว้ใน Fig. 7.7 ต่อไปนี้

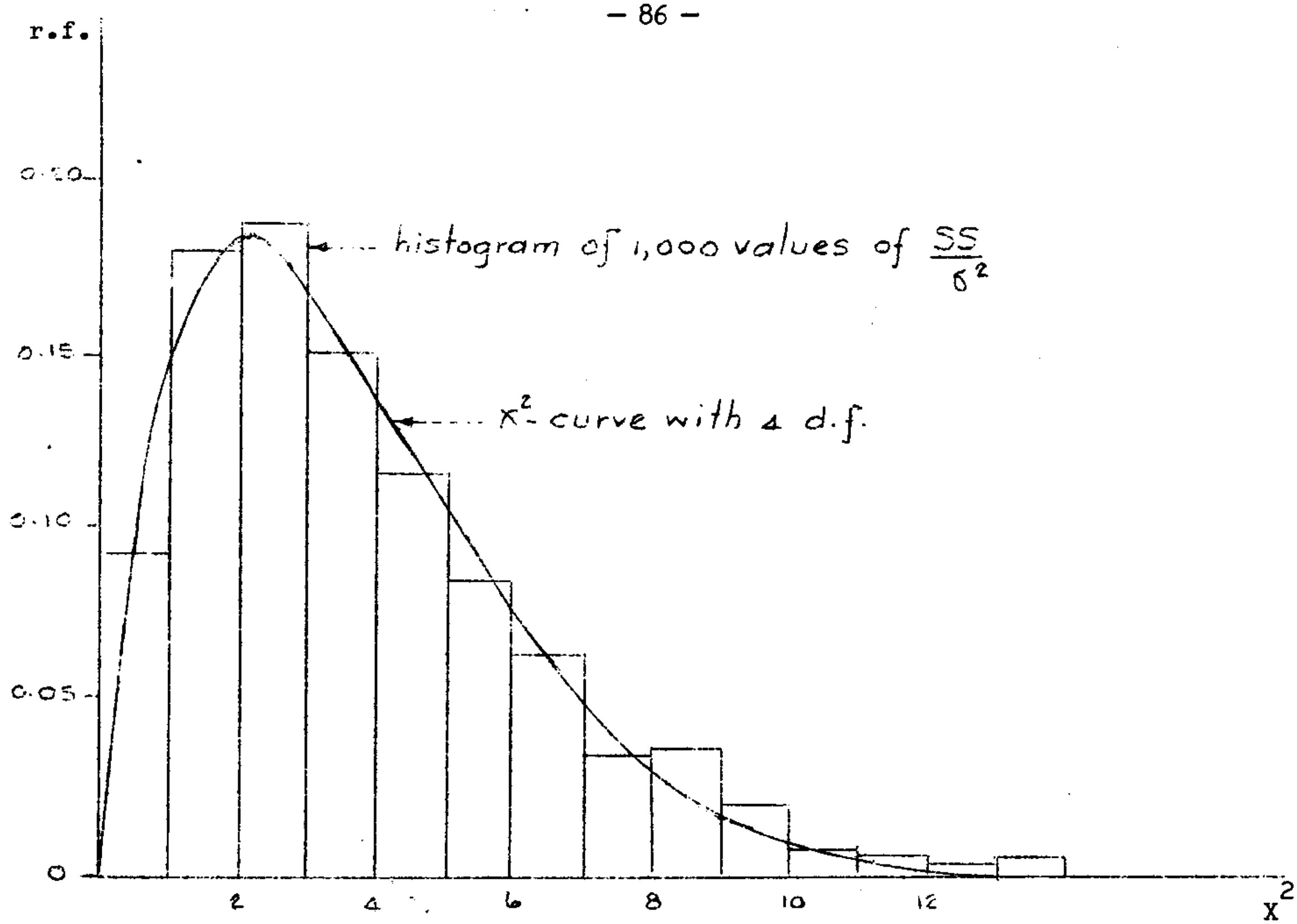


Fig. 7.7

จาก Table 7.7 b หรือ Fig. 7.7 จะเห็นได้ว่า observed และ theoretical frequencies ใกล้เคียงกันมาก แสดงว่าค่าของ $\frac{SS}{\sigma^2}$ ที่เราคำนวณได้จาก 1,000 samples ซึ่งแต่ละ sample ประกอบด้วย 5 observations นั้น follow the X^2- distribution with 4 d.f. ดังนั้นจึงแสดงถึงความเป็นจริงแล้วว่า $\frac{SS}{\sigma^2}$ follows the X^2- distribution with n - 1 d.f.

mean โดยประมาณของค่าของ $\frac{SS}{\sigma^2}$ 1,000 ค่าก็จะหาได้เท่ากับ 3.8353 (Table 7.7 b) ซึ่งใกล้เคียงกับ d.f. 4 แสดงให้เห็นความเป็นจริงยิ่งขึ้นอีกว่า $\frac{SS}{\sigma^2}$ follows the X^2- distribution with n - 1 d.f. theorem ต่อไปนี้ได้รวบรวมการอภิปรายที่แล้วมาไว้ดังนี้

<u>Theorem 7.7 a:</u> If all possible samples of size n are drawn from a normal population with variance equal to σ^2, and for each sample the value $\frac{\sum(y - \bar{y})^2}{\sigma^2}$ is computed, the values of $\frac{\sum(y - \bar{y})^2}{\sigma^2}$ follow the X^2-distribution with n - 1 degrees of freedom (that is, $\nu = n - 1$).

ในทางปฏิบัติเรามักจะพูดว่า $\sum (y - \bar{y})^2$ หรือ SS มี n - 1 d.f. เนื่องจาก d.f. n - 1 นี้ ถูกใช้เป็นตัวหารในการหา s^2 ด้วย คือ

$$\boxed{s^2 = \frac{SS}{n-1} = \frac{SS}{\nu}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

(Equation (6) ข้อ 7.2) ดังนั้น sample variance (s^2) ก็คือ SS หารด้วย d.f. ของมันนั่นเอง

เราทราบแล้วว่า SS เป็นผลบวกของกำลังสองของ $(y - \bar{y})$, n จำนวน นั่นคือ

$$SS = \sum (y - \bar{y})^2 = (y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots\ldots + (y_n - \bar{y})^2 \ldots\ldots (7)$$

แต่ผลบวกของ $(y - \bar{y})$, n จำนวนเท่ากับศูนย์ หรือ $\sum (y - \bar{y}) = 0$ (Equation (5) ข้อ 7.4) เพราะฉะนั้นเมื่อเราทราบ $(y - \bar{y})$, n - 1 จำนวน $(y - \bar{y})$ ที่เหลืออีกจำนวนหนึ่งก็จะทราบได้โดยอัตโนมัติ ตัวอย่างเช่น mean ของ 5 observations 3, 2, 1, 3, 1 คือ 2 และ $(y - \bar{y})$, 5 จำนวนคือ 1, 0, -1, 1, -1 ถ้าเราทราบ $(y - \bar{y})$, ใด ๆ รวม 4 จำนวนแล้ว $(y - \bar{y})$ ที่เหลือก็จะทราบได้โดยอัตโนมัติ เพราะว่าผลบวกของ $(y - \bar{y})$, 5 จำนวนนี้เท่ากับศูนย์

d.f. นี้จะถูกใช้เกี่ยวข้องกับทุกวิธีซึ่งจะนำมากล่าวในบทต่อ ๆ ไป เมื่อใดที่ d.f. ถูกใช้มันจะถูกใช้โดยตรงหรือโดยอ้อมซึ่งเกี่ยวกับ X^2- distribution เสมอ ถ้าเรากล่าวว่า s^2 มี ν d.f. ก็หมายความว่า statistic $\boxed{\frac{\nu s^2}{\sigma^2} = \frac{SS}{\sigma^2}}$ follows the X^2- distribution with ν d.f. ตัวอย่างเช่นถ้าเรากล่าวว่า s^2 มี 8 d.f. ก็หมายความว่า $\frac{8s^2}{\sigma^2}$ follows the X^2- distribution with 8 d.f. ในบทต่อ ๆ ไป the number of degrees of freedom จะถูกใช้บ่อย ๆ โดยไม่มีการอ้างอิงถึง X^2- distribution ซึ่งแม้ว่า X^2- distribution จะเป็นจุดต้นกำเนิดก็ตาม

จาก distribution ของ $\frac{SS}{\sigma^2}$ นี้เรายังได้ distribution ของ $\frac{s^2}{\sigma^2}$ ออกมาอีกด้วยเพราะว่า $s^2 = \frac{SS}{\nu}$ ในกรณี n = 5 หรือ $\nu = 4$, $\frac{SS}{4} = s^2$ พูดได้อีกอย่างหนึ่งว่า SS โตเป็น 4 เท่า

ของ s^2 เมื่อค่าของ SS ค่าหนึ่งตกอยู่ระหว่าง 0 และ 4 ค่าของ s^2 ซึ่งเป็นไปตาม SS นั้นจะตกอยู่ระหว่าง 0 และ 1 จาก Table 7.7 b จะเห็นได้ว่า 93 samples จาก 1,000 samples มีค่าของ $\frac{SS}{\sigma^2}$ ตกอยู่ระหว่าง 0 และ 1 และจาก 93 samples นี้เราก็จะได้ค่าของ $\frac{s^2}{\sigma^2}$ ตกอยู่ระหว่าง 0 และ 0.25 ด้วย หรืออีกนัยหนึ่ง ถ้า classes ต่าง ๆ ของ Table 7.7b ถูกเปลี่ยนเป็น 0 – 0.25, 0.25 – 0.50, และต่อ ๆ ไปแล้ว frequency table ใหม่ซึ่งประกอบด้วย classes ต่าง ๆ ที่เปลี่ยนใหม่นี้จะเป็น frequency table ของ $\frac{s^2}{\sigma^2}$ ผลก็คือ statistic $\frac{s^2}{\sigma^2}$ follows the distribution of $\frac{\chi^2}{\nu}$

Table 4, Appendix แสดงให้เห็น relative cumulative frequencies (r.c.f.) ของ χ^2- distribution ซึ่งเป็น distribution ของ $\frac{SS}{\sigma^2}$ ถ้าค่าต่าง ๆ ของทุก ๆ บรรทัดในTable นี้ถูกหารด้วย ν หรือ the number of degrees of fredom ประจำบรรทัดของมัน table ที่ได้ใหม่คือ Table 5, Appendix จะให้ relative cumulative frequencies ของ $\frac{\chi^2}{\nu}$ ซึ่งเป็น distribution ของ $\frac{s^2}{\sigma^2}$

mean ของ distribution $\frac{s^2}{\sigma^2}$ จะหาได้จาก mean ของ distribution ของ $\frac{SS}{\sigma^2}$ เพราะว่า

$$\frac{SS}{\nu} = s^2$$

$$SS = \nu s^2$$

$$\frac{SS}{\sigma^2} = \nu \frac{s^2}{\sigma^2}$$

$\frac{SS}{\sigma^2}$ จึงโตเป็น ν เท่าของ $\frac{s^2}{\sigma^2}$ ดังนั้น mean ของ $\frac{SS}{\sigma^2}$ จึงโตเป็น ν เท่าของ mean ของ $\frac{s^2}{\sigma^2}$ ด้วย แต่ mean ของ $\frac{SS}{\sigma^2} = \nu$ เพราะฉะนั้น mean ของ $\frac{s^2}{\sigma^2} = \frac{\nu}{\nu} = 1$

การพิสูจน์ว่า mean ของ $\frac{s^2}{\sigma^2} = 1$ อาจทำได้อีกวิธีหนึ่งคือ Theorem 7.2 ชี้ให้เห็นว่า mean ของ s^2 ของ all possible samples เท่ากับ σ^2 ถ้าเอา σ^2 ไปหาร s^2 ทุกตัวแล้ว mean ของ $\frac{s^2}{\sigma^2}$ จะเท่ากับ $\frac{\sigma^2}{\sigma^2} = 1$ (Theorem 2.4 b)

เพื่อสะดวกในการอ้างอิงในอนาคต เราจะได้นำผลที่ได้จากการอภิปรายข้างบนมาแสดงไว้ใน Theorem ต่อไปนี้

<u>Theorem 7.7 b:</u> If all possible samples of the same size are drawn from a given population with variance equal to σ^2, and the variance (s^2) is computed for each sample, <u>the mean of all the ratios $\frac{s^2}{\sigma^2}$ is equal to 1.</u>

ภายหลังที่เราพิจารณา mean ของ $\frac{s^2}{\sigma^2}$ แล้ว ก็ควรจะได้สนใจต่อไปว่า sample size (n) จะกระทบกระเทือนถึงความแปรผันในค่าของ $\frac{s^2}{\sigma^2}$ อย่างไรบ้าง <u>ถ้า sample size เพิ่มขึ้น sample variance (s^2) ก็จะกลายเป็นค่าประมาณที่เชื่อได้ของ population variance (σ^2) ยิ่งขึ้นทุกที</u> และค่าของ s^2 ของ all possible samples ซึ่งมี size เดียวกันก็จะอยู่เป็นกลุ่มรอบค่าของ population variance (σ^2) ใกล้ชิดยิ่งขึ้น หรือค่าของ $\frac{s^2}{\sigma^2}$ ของ all possible samples โอบรอบค่า 1 ใกล้ชิดมากขึ้น ในขณะที่ sample size (n) เข้าไปใกล้ infinity, s^2 ทุกตัวจะกลายเป็น σ^2 ดังนั้นค่าของ $\frac{s^2}{\sigma^2}$ จะกลายเป็น 1 ในทุก sample ปรากฏการณ์ดังกล่าวนี้จะสังเกตได้ใน Table 5, Appendix ตัวอย่างเช่น ทั้ง 97.5 % point และ 2.5 % point ค่าของ $\frac{\chi^2}{\nu}$ หรือ $\frac{s^2}{\sigma^2}$ จะพุ่งไปหา 1 เข้าทุกที ในเมื่อ the number of degrees of freedom เพิ่มขึ้น

7.8 Algebraic Identities

เราอาจพิสูจน์ algebraic identities ที่ให้ไว้ใน Equation (4) ข้อ 7.4 Equation (5) ข้อ 7.4 และ Equation (3) ข้อ 7.7 ได้ตามลำดับดังนี้

$$
\begin{aligned}
\Sigma (y - \bar{y})^2 &= (y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \ldots\ldots + (y_n - \bar{y})^2 \\
&= (y_1^2 - 2y_1\bar{y} + \bar{y}^2) + (y_2^2 - 2y_2\bar{y} + \bar{y}^2) + \ldots\ldots + \\
&\quad (y_n^2 - 2y_n\bar{y} + \bar{y}^2) \\
&= \Sigma y^2 - 2\bar{y}\,\Sigma y + n\bar{y}^2 \\
&= \Sigma y^2 - 2\left(\frac{\Sigma y}{n}\right)\Sigma y + n\left(\frac{\Sigma y}{n}\right)^2 \\
&= \Sigma y^2 - \frac{2(\Sigma y)^2}{n} + \frac{(\Sigma y)^2}{n} \\
&= \Sigma y^2 - \frac{(\Sigma y)^2}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)
\end{aligned}
$$

$$\Sigma\,(y - \bar{y}) = (y_1 - \bar{y}) + (y_2 - \bar{y}) + \ldots\ldots + (y_n - \bar{y})$$

$$= \Sigma\, y - n\bar{y}$$

$$= \Sigma\, y - n\frac{\Sigma\, y}{n}$$

$$= \Sigma\, y - \Sigma\, y$$

$$= 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (2)$$

$$\Sigma\,(y - \mu)^2 = \Sigma\left[y - \bar{y} + \bar{y} - \mu\right]^2$$

$$= \Sigma\left[(y - \bar{y}) + (\bar{y} - \mu)\right]^2$$

$$= \left[(y_1 - \bar{y})^2 + 2(y_1 - \bar{y})(\bar{y} - \mu) + (\bar{y} - \mu)^2\right] +$$

$$\left[(y_2 - \bar{y})^2 + 2(y_2 - \bar{y})(\bar{y} - \mu) + (\bar{y} - \mu)^2\right] + \ldots\ldots +$$

$$\left[(y_n - \bar{y})^2 + 2(y_n - \bar{y})(\bar{y} - \mu) + (\bar{y} - \mu)^2\right]$$

$$= \Sigma\,(y - \bar{y})^2 + 2(\bar{y} - \mu)\,\Sigma\,(y - \bar{y}) + n(\bar{y} - \mu)^2$$

$$= \Sigma\,(y - \bar{y})^2 + 2(\bar{y} - \mu)\ 0 + n(\bar{y} - \mu)^2$$

$$= \Sigma\,(y - \bar{y})^2 + n(\bar{y} - \mu)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (3)$$

7.9 Analysis of Variance

เราได้แสดงให้เห็นจริงโดยเลขจำนวน (ข้อ 7.7) และโดยการพิสูจน์ทางพีชคณิต (ข้อ 7.8) มาแล้วว่า

$$\Sigma\,(y - \mu)^2 = \Sigma\,(y - \bar{y})^2 + n(\bar{y} - \mu)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (1)$$

และได้ทราบแล้วว่า:

$\frac{\Sigma (y - \mu)^2}{\sigma^2}$ follows the X^2- distribution with n d.f. (Theorem 7.5)

$\frac{\Sigma (y - \bar{y})^2}{\sigma^2}$ หรือ $\frac{SS}{\sigma^2}$ follows the X^2- distribution with n - 1 d.f. (Theorem 7.7 a) และ $\frac{\Sigma(y-\bar{y})^2}{\sigma^2(n-1)}$ หรือ $\frac{\chi^2}{\sigma^2}$ — n — $\frac{\chi^2}{\nu}$ — n - 1

$\frac{n (\bar{y} - \mu)^2}{\sigma^2}$ หรือ u^2 follows the X^2- distribution with 1 d.f.(Theorem 7.6)

กล่าวได้อีกอย่างหนึ่งว่า ผลบวกของกำลังสอง (ของความแปรผันของ observations ต่าง ๆ จาก population mean) บนข้างซ้ายของ Equation (1) ถูกแบ่งออกเป็นสองส่วนประกอบบนข้างขวาของ Equation นั้น วิธีแบ่งผลบวกของกำลังสองออกเป็นส่วนประกอบต่าง ๆ อย่างนี้เรียกว่า analysis of variance การแบ่งนี้เป็นกระบวนการทางพีชคณิตบริสุทธิ์ แต่ข้อเท็จจริงที่ว่าส่วนประกอบแต่ละส่วนและผลบวกของส่วนประกอบทั้งสอง follow the X^2- distribution นั้นเป็นสิ่งสำคัญมากในวิธีการต่าง ๆ ที่ได้ให้ไว้ในบทต่อ ๆ ไป ไม่แต่เพียงผลบวกของกำลังสองจะถูกแบ่งออกเป็นส่วนประกอบต่าง ๆ เท่านั้น d.f. ของมันก็จะถูกแบ่งออกไปด้วย ดังจะเห็นได้ว่า $\Sigma (y - \mu)^2$ มี n d.f. ส่วนประกอบของมัน คือ $\Sigma (y - \bar{y})^2$ และ $n(\bar{y} - \mu)^2$ มี n - 1 และ 1 d.f. ตามลำดับ และ $n = (n - 1) + 1$

แต่ละเทอมของ 3 เทอมใน Equation (1) นั้น เมื่อถูกหารด้วย the number of degrees of freedom ของมันแล้วจะเป็น unbiased estimate (ข้อ 7.3) ของ σ^2 ทั้งสิ้น กล่าวคือ ผลบวกของกำลังสองบนข้างซ้ายของ Equation (1) เมื่อถูกหารด้วย d.f. n ก็คือ (Equation (4) ข้อ 7.2)

$$V_1 = \frac{\Sigma (y - \mu)^2}{n} \quad .... \quad (2)$$

ซึ่งได้ถูกแสดงให้เห็นมาแล้วว่าเป็น unbiased estimate ของ σ^2 (ข้อ 7.2) ส่วนประกอบ $\Sigma (y - \bar{y})^2$ หรือ SS เมื่อถูกหารด้วย d.f. ของมันคือ n - 1 ก็คือ

$$s^2 = \frac{\Sigma (y - \bar{y})^2}{n - 1} \quad .... \quad (3)$$

ซึ่งเป็น unbiased estimate ของ σ^2 (ข้อ 7.3) แต่ $n(\bar{y}-\mu)^2$ นั้นดูเหมือนกับว่าจะเป็นประเภทที่แตกต่างออกไปเพราะว่ามันเกี่ยวข้องกับ means มากกว่า observations แต่เราอาจพิจารณาให้ all possible sample means เป็นหนึ่ง population ได้เรียกว่า population ของ $\bar{y}$ แล้ว mean ของ population ของ $\bar{y}$ คือ $\mu_{\bar{y}}$ จะเท่ากับ μ และ variance คือ $\sigma^2_{\bar{y}}$ จะเท่ากับ $\frac{\sigma^2}{n}$ ถ้าเรา draw all possible samples ซึ่งมี size = 1 ออกมาจาก population ของ $\bar{y}$ แต่ละ sample เราคำนวณค่าของ V_1 ได้คือ

$$V_1 = \frac{\Sigma(\bar{y}-\mu_{\bar{y}})^2}{n}$$

$$= \frac{\Sigma(\bar{y}-\mu)^2}{1}$$

$$= \frac{(\bar{y}-\mu)^2}{1}$$

ค่าเฉลี่ยของ V_1 หรือ $\frac{(\bar{y}-\mu)^2}{1}$ ทุกตัวจะเท่ากับ $\sigma^2_{\bar{y}}$ หรือ $\frac{\sigma^2}{n}$ นั่นคือ $\frac{(\bar{y}-\mu)^2}{1}$ เป็น unbiased estimate ของ $\frac{\sigma^2}{n}$ เพราะฉะนั้น $\frac{n(\bar{y}-\mu)^2}{1}$ จะเป็น unbiased estimate ของ $n\frac{\sigma^2}{n}$ หรือ σ^2 จึงกล่าวได้ว่าแต่ละเทอมของ 3 เทอมใน Equation (1) เมื่อถูกหารด้วย d.f. ของมันแล้วจะให้ unbiased estimate ของ σ^2 เหมือนกันทั้งสิ้น อย่างไรก็ดี เฉพาะเทอมกลางเท่านั้นที่จะใช้ประโยชน์ได้เพราะไม่ได้ประกอบด้วย μ ซึ่งไม่ทราบค่า เทอม $\frac{\Sigma(y-\bar{y})^2}{n-1}$ หรือ $\frac{SS}{n-1}$ นี้ก็คือ s^2 ที่เรียนมาแล้ว

ข้อเท็จจริงที่ว่า $\Sigma(y-\mu)^2$ มี n d.f. และ $\Sigma(y-\bar{y})^2$ มี n - 1 d.f. นั้นมักจะถูกพูดเสมอว่า "ผลบวกของกำลังสองมี d.f.หายไป 1 d.f." การหายไปของ d.f. นี้เป็นเพราะเราใช้ sample mean ($\bar{y}$) แทน population mean (μ) ในผลบวกของกำลังสอง

analysis of variance เป็นวิธีที่ถูกใช้บ่อย ๆ ในบทต่อ ๆ ไป จึงได้นำมากล่าวล่วงหน้าไว้ก่อน เรื่องที่ให้ไว้ในข้อนี้เป็นกรณีธรรมดาที่สุดของ analysis of variance ประโยชน์ในการใช้ analysis of variance จะได้กล่าวไว้ในข้อ 12.6

7.10 Test of Hypothesis

เราอาจใช้ distribution ของ $\frac{SS}{\sigma^2}$ (ข้อ 7.7) ในการทดสอบสมมติฐานได้ว่า population variance เท่ากับค่าที่กำหนดให้ ตัวอย่างเช่น ค่าของ SS ของ random sample หนึ่งซึ่งมี 5

observations เป็น 480 ปัญหาก็คือการทดสอบสมมติฐานว่า population variance เท่ากับ 100 นั่นคือ σ_0^2 100 ในเมื่อ σ_0^2 เป็น hypothetical variance เราทราบจาก Theorem 7.7 a แล้วว่า $\frac{SS}{\sigma^2}$ follows the X^2- distribution with n - 1 d.f. ถ้าสมมติฐานเป็นจริง คือ $\sigma_0^2 = \sigma^2 = 100$ แล้ว จาก Table 4, Appendix จะพบว่า 2.5 % ของ X^2- values with 4 d.f. น้อยกว่า 0.484419 และ 2.5 % ของ X^2- values มากกว่า 11.1433 เพราะฉะนั้นถ้าเรา เลือกใช้ 5 % significance level, critical regions จะอยู่ที่ $X^2 < 0.484419$ และ $X^2 > 11.1433$ (Fig. 7.10 a)

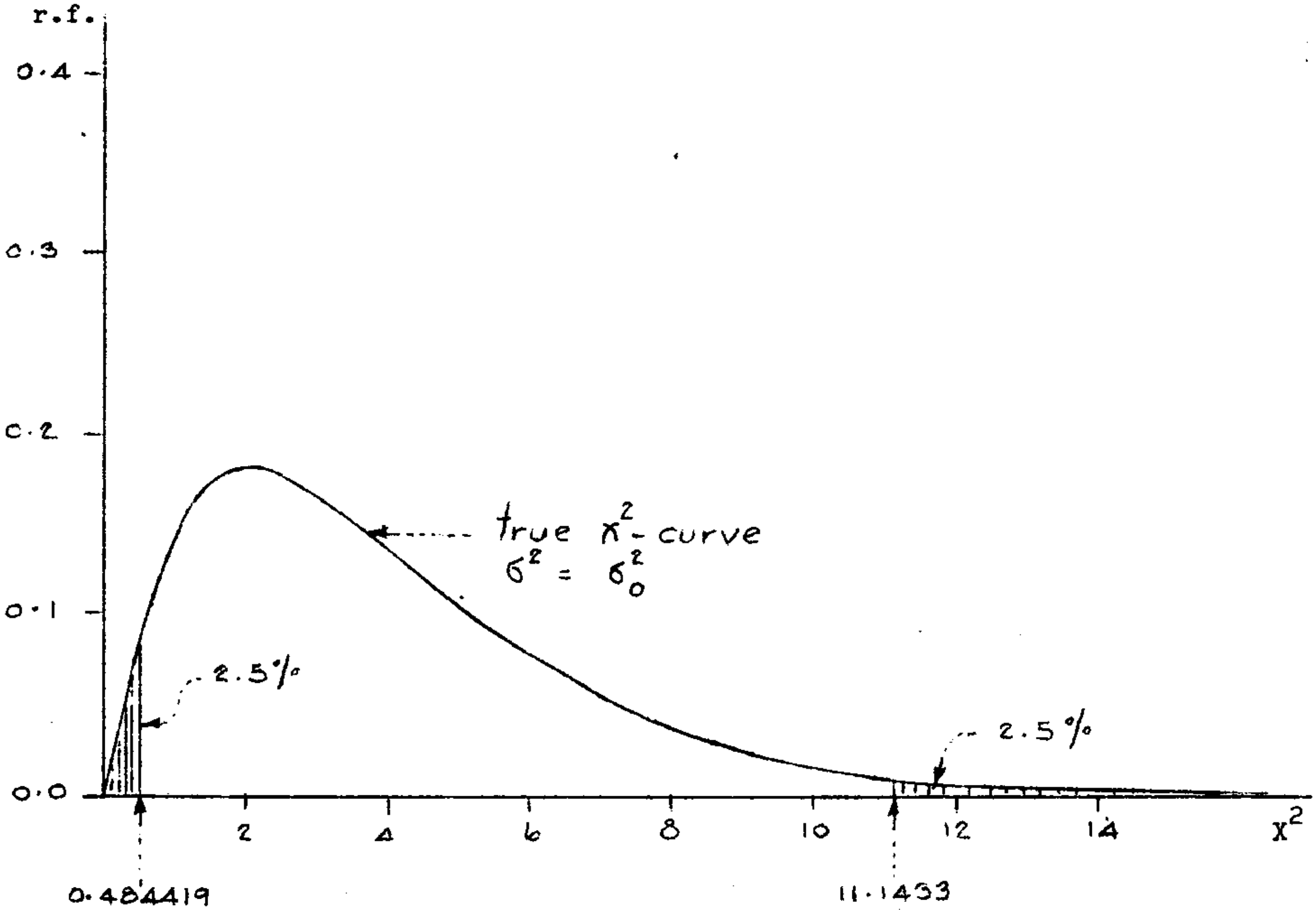


Fig. 7.10 a

statistic $\frac{SS}{\sigma_o^2}$ เท่ากับ $\frac{480}{100}$ หรือ 4.80 with 4 d.f. ค่า 4.80 นี้ตกอยู่ภายนอก critical regions เพราะฉะนั้น เราจึงยอมรับเอาสมมติฐาน และ conclusion คือ $\sigma^2 = 100$ ถ้าค่าของ $\frac{SS}{\sigma_o^2}$ ตกอยู่ภายใน critical region ข้างซ้าย conclusion จะเป็น population variance จริงน้อยกว่า 100 แต่ถ้าค่าของ $\frac{SS}{\sigma_o^2}$ ตกอยู่ภายใน critical region ข้างขวา conclusion จะเป็น population variance จริงมากกว่า 100 ถ้าสมมติฐานเป็นจริง และวิธีดำเนินการทดสอบสมมติฐานเป็นไปตามนี้ 5 % ของ all possible samples จะนำไปสู่ conclusion ผิดได้ว่า population variance ไม่เท่ากับ 100 เพราะฉะนั้น Type I error (บริเวณที่ขีดเส้น ใน Fig. 7.10a) จึงเป็น 5 %

distribution ของ $\frac{s^2}{\sigma^2}$ ใช้ในการทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนดให้ได้ด้วย เราจะใช้ sample เดียวกันที่มี SS = 480 และ n = 5 มาเป็นตัวอย่างอีกครั้ง เนื่องจาก $s^2 = \frac{480}{4} = 120$, statistic $\frac{s^2}{\sigma_o^2} = \frac{120}{100} = 1.20$ เราอาจหา critical regions จาก Table 5, Appendix ได้คือ critical regions จะอยู่ที่ $\frac{\chi^2}{\nu} < 0.121105$ และ $\frac{\chi^2}{\nu} > 2.7858$ ในเมื่อ ν เป็น the number of degrees of freedom ค่า 0.121105 และ 2.7858 สองค่านี้เป็น $\frac{1}{4}$ ของค่า 0.484419 และ 11.1433 ที่ได้จาก χ^2- table และ statistic $\frac{s^2}{\sigma_o^2}$ ก็จะเป็น $\frac{1}{4}\left(\frac{SS}{\sigma_o^2}\right)$ ด้วย เพราะฉะนั้น conclusion ที่ได้โดยวิธีการทั้งสองนี้จะเป็นอย่างเดียวกันเสมอ อนึ่งเราจะใช้ statistic $\frac{SS}{\sigma_o^2}$ หรือ $\frac{s^2}{\sigma_o^2}$ ตัวใดก็ได้ในการทดสอบสมมติฐานโดยไม่มีข้อจำกัดไว้ เหตุผลที่นำ statistic $\frac{s^2}{\sigma_o^2}$ มาศึกษาก็เพราะว่ามันให้แนวทางแห่งสามัญสำนึกในการมองเห็นการทดสอบสมมติฐานว่า population variance เท่ากับ σ_o^2 statistic s^2 เป็นค่าประมาณของ population variance จริง (σ^2) ถ้า $\frac{s^2}{\sigma_o^2}$ มากกว่า 1 มากแล้วจะชี้ให้เห็นว่า population variance จริงมากกว่า hypothetical population variance ถ้า $\frac{s^2}{\sigma_o^2}$ น้อยกว่า 1 มากก็ชี้ให้เห็นว่า population variance จริงน้อยกว่า hypothetical population variance ที่ถูกทดสอบ

7.11 Procedures of Test of Hypothesis

วิธีดำเนินการทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนดให้จะถูกแสดงโดย sample หนึ่งดังนี้ สมมติให้ sample หนึ่งมี 10 observations ซึ่งได้มาจาก population หนึ่ง observations เหล่านี้คือ

4.8, 3.2, 3.6, 4.8, 6.1, 5.6, 4.7, 5.3, 5.1, 7.6

ปัญหาก็คือการพิจารณาว่า variance ของ population นี้จะเท่ากับ 4 หรือไม่ วิธีดำเนินการทดสอบทำดังนี้

1. Hypothesis: hypothesis คือ population variance เท่ากับ 4 นั่นคือ

$\sigma^2 = \sigma_o^2 = 4$

2. Alternative hypotheses: alternative hypotheses คือ

   ก. population variance น้อยกว่า 4 นั่นคือ $\sigma^2 < 4$ และ

   ข. population variance มากกว่า 4 นั่นคือ $\sigma^2 > 4$

3. Assumptions: sample เป็น random sample ซึ่งได้มาจาก normal population

4. Level of significance: significance level ที่เลือกใช้คือ 5 %

5. Critical regions: critical regions อยู่ที่

   ก. $X^2 < 2.70039$ และ

   ข. $X^2 > 19.0228$

   (ค่าทั้งสองนี้ได้จาก Table 4, Appendix; X^2 with 9 d.f. และการทดสอบนี้เป็นแบบ two-tailed test สำหรับการทดสอบแบบ one-tailed test จะได้ให้ไว้ในข้อ 7.12)

6. Computation of statistic

$$\sigma_o^2 = 4 \text{ (กำหนดให้)}$$

$$n = 10$$

$$\Sigma y = 4.8 + 3.2 + 3.6 + 4.8 + 6.1 + 5.6 + 4.7 + 5.3 + 5.1 + 7.6$$

$$= 50.8$$

$$(\Sigma y)^2 = (50.8)^2$$

$$= 2,580.64$$

$$\frac{(\Sigma y)^2}{n} = \frac{2,580.64}{10}$$

$$= 258.064$$

$$\Sigma y^2 = (4.8)^2 + (3.2)^2 + (3.6)^2 + (4.8)^2 + (6.1)^2 + (5.6)^2 + (4.7)^2 + (5.3)^2 + (5.1)^2 + (7.6)^2$$

$$= 271.80$$

$$SS = \Sigma y^2 - \frac{(\Sigma y)^2}{n} \quad \text{.................. (ข้อ 7.4)}$$

$$= 271.80 - 258.064$$

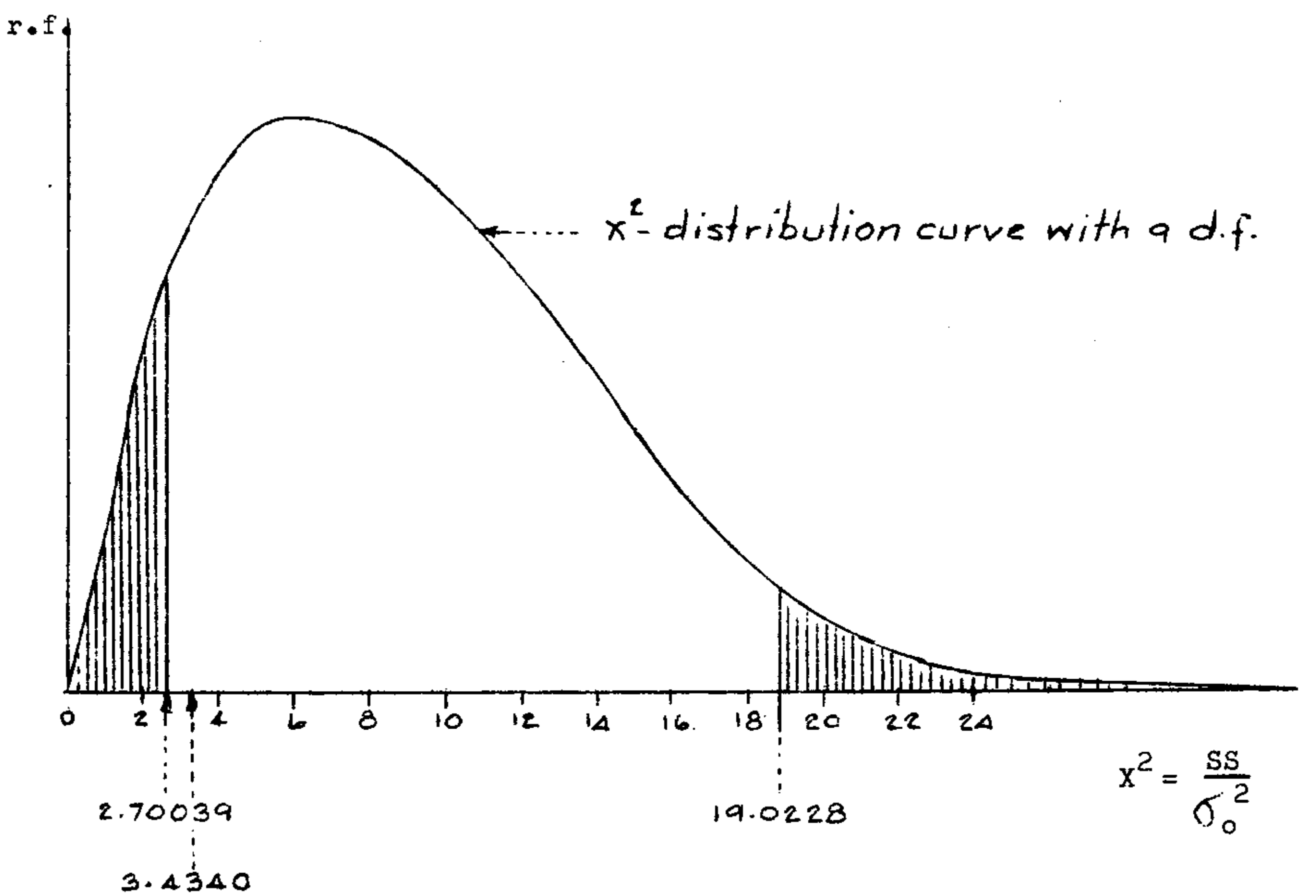
$$= 13.736$$

$$X^2 = \frac{SS}{\sigma_o^2}$$

$$= \frac{13.736}{4}$$

$$= 3.4340 \quad \text{with 9 d.f.}$$

7. Conclusion: เนื่องจาก X^2- valueที่คำนวณได้เท่ากับ 3.4340 นั้น อยู่ภายนอก critical regions



จึงยอมรับเอาสมมติฐาน conclusion คือ population variance เท่ากับ 4 (ถ้า X^2- value ที่คำนวณได้น้อยกว่า 2.70039 จะตกอยู่ใน critical region ข้างซ้าย conclusion จะเป็น population variance น้อยกว่า 4 แต่ถ้า X^2- value ที่คำนวณได้มากกว่า 19.0228 จะตกอยู่ใน critical region ข้างขวา conclusion จะเป็น population variance มากกว่า 4)

7.12 Applications

การทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนดให้นี้ได้ใช้กันมากในการอุตสาหกรรม ตามปกติโรงงานต่าง ๆ ต้องการให้สินค้าที่ผลิตมีความสม่ำเสมอในความยาว, น้ำหนัก ฯลฯ มากที่สุด ดังนั้น variance (หรือ standard deviation) ในขนาดของสินค้าที่ยอมให้มีได้จะถูกระบุไว้เป็นมาตรฐานรายการรายละเอียด (specification standard) เขาจะนำ samples ของสินค้าที่ผลิตได้ออกมาเป็นคราว ๆ และทดสอบสมมติฐานว่า มาตรฐานของการผลิตนั้นยังคงใช้ได้อยู่ ตัวอย่างเช่น น้ำหนักเฉลี่ยของผลองุ่นกระป๋องหนึ่งกระป๋องที่รินน้ำทิ้งแล้ว กำหนดให้เท่ากับ 12 ออนซ์ และกำหนดให้ standard deviation เท่ากับ $\frac{1}{4}$ ออนซ์ ตลอดเวลาที่บรรจุผลองุ่นทั้งหมดลงกระป๋องนั้นแทบจะเป็นไปไม่ได้เลยว่าจะสามารถทำให้น้ำหนักของผลองุ่นในทุกกระป๋องที่รินน้ำทิ้งแล้วเท่ากับ 12 ออนซ์ กระป๋องที่มีผลองุ่นมากก็จะหนักมากกว่า 12 ออนซ์ และกระป๋องที่มีผลองุ่นน้อยก็จะหนักน้อยกว่า 12 ออนซ์ เพราะฉะนั้นจึงหวังได้ว่าน้ำหนักของผลองุ่นกระป๋องที่รินน้ำทิ้งแล้วจะเปลี่ยนไปต่าง ๆ กัน ปัญหาที่จะต้องพิจารณาก็คือการป้องกันความแปรผันไม่ให้มีมากเกินไป random sample หนึ่งซึ่งประกอบด้วยองุ่นกระป๋อง n กระป๋อง, สมมติว่า 10 กระป๋อง, จะต้องถูกตรวจสอบเป็นคราว ๆ ไป ผลองุ่นในกระป๋องจะถูกรินน้ำทิ้งและชั่งน้ำหนักไว้ น้ำหนักของผลองุ่นในกระป๋องต่าง ๆ คือ observations (y) hypothetical variance คือ $(\frac{1}{4})^2$ หรือ $\sigma_o^2 = \frac{1}{16}$ การทดสอบสมมติฐานนี้ใช้วิธีที่ให้ไว้ในข้อ 7.11 ได้ ถ้าวัตถุประสงค์คือการป้องกันไม่ให้ population variance มีค่ามากเกินไป เราควรจะใช้การทดสอบโดยวิธี one-tailed test กล่าวได้คืออย่างหนึ่งว่าจะมีหนึ่ง critical region บนหางข้างขวาของ X^2- distribution เท่านั้น 2 critical regions ตรงที่ $X^2 < 2.70039$ และ $X^2 > 19.0228$ ตามที่ได้ให้ไว้ในข้อที่แล้วควรจะถูกแทนด้วยหนึ่ง critical region ตรงที่ $X^2 > 16.9190$ ถ้าเรายอมรับเอาสมมติฐานภายหลังการทดสอบทางสถิติก็หมายความว่ามาตรฐานการผลิตยังคงใช้ได้อยู่และไม่ต้องไปทำอะไร แต่ถ้าสมมติฐานถูกปฏิเสธ conclusion ก็คือ น้ำหนักของผลองุ่นที่รินน้ำทิ้งแล้วในกระป๋องต่าง ๆ จะแปรผันมากกว่า specified standard ที่ยอมให้มีได้ ดังนั้น จะต้องแก้ไขสถานการณ์เสียให้ถูกต้องดังเดิม ประโยชน์ของการ

ทดสอบสมมติฐานเป็นคราว ๆ ก็คือ ความผิดพลาดในกรรมวิธีการผลิต จะปรากฏให้เห็นก่อนที่จะเกิดความเสียหายอย่างใหญ่หลวงขึ้น

ตัวอย่างซึ่งเรียกว่า "quality control" ดังกล่าวมานี้เป็นการแสดงให้เห็นหลักทั่วไปของการประยุกต์วิชาสถิติในโรงงานอุตสาหกรรม พึงสังเกตว่า "quality control" ที่ใช้ในที่นี้ที่จริงนั้นเป็น "quantity control" เพราะว่าไม่มีอะไรจะต้องกระทำเกี่ยวกับคุณภาพของผลองุ่นเลย

7.13 Remarks

เช่นเดียวกับบทที่แล้ว ความประสงค์ของบทที่ 7 นี้ก็เพื่อนำเอาหลักเบื้องต้นบางอย่างของวิชาสถิติมากล่าวไว้ แม้ว่าการทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนดให้จะใช้กันมากในโรงงานอุตสาหกรรมเพื่อควบคุมคุณภาพของสินค้าก็ตาม แต่ประโยชน์ของมันก็มีอยู่อย่างจำกัด หลักต่าง ๆ ที่ให้ไว้ในบทนี้จะได้นำมาใช้ซ้ำอีกในการดัดแปลงให้เป็นวิธีที่ให้ประโยชน์ได้ดียิ่งขึ้น X^2- test นี้จะนำไปใช้ในบทที่ 21, 22 และ 24 สำหรับการทดสอบสมมติฐานหลายชนิดมากกว่าที่ได้กล่าวไว้ในบทนี้

Chapter 8

Student's t-Distribution

ในบทนี้จะได้กล่าวถึง frequency distribution สำคัญอีกอย่างหนึ่งซึ่งเรียกว่า Student's t-distribution ชื่อ Student's t-distribution นี้ได้ถูกตั้งขึ้นเพื่อเป็นเกียรติแก่ W.G. Gosset ผู้เขียนเรื่องทางสถิติคนหนึ่งซึ่งใช้นามแฝงว่า "Student" เริ่มแรกทีเดียว Gosset เป็นผู้คิด frequency นี้ขึ้นเมื่อประมาณต้นคริสต์ศตวรรษที่ 20 แล้วต่อมาภายหลัง R.A. Fisher เป็นผู้ดัดแปลงแก้ไขให้ดีขึ้น ในบทนี้จะกล่าวถึง Student's t-distribution ที่ได้ดัดแปลงแก้ไขแล้วเท่านั้น

8.1 Description of t-Distribution

u-test ที่ได้กล่าวไว้ในข้อ 6.7 ใช้สำหรับทดสอบสมมติฐานว่า population mean มีค่าเท่ากับ μ_o ที่กำหนดให้

ในเมื่อ

$$u = \frac{\bar{y} - \mu_o}{\sqrt{\frac{\sigma^2}{n}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

อย่างไรก็ดี การใช้ u-test นี้มีขอบเขตจำกัดเพราะตามปกติเราไม่ทราบค่า population variance (σ^2) ดังนั้นจึงได้พัฒนา t-distribution ขึ้นมาเพื่อขจัดความลำบากในการทดสอบสมมติฐานให้หมดไป กล่าวคือ ถ้าเราเอา sample variance (s^2) ไปใส่แทน population variance (σ^2) ใน Equation (1) ข้างบนนี้แล้ว statistic ใหม่ที่ได้จะเป็น

$$\boxed{t = \frac{\bar{y} - \mu_o}{\sqrt{\frac{s^2}{n}}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

เพราะฉะนั้นความประสงค์ที่เอา t มาใช้ก็เพื่อขจัดข้อขัดข้องในการทดสอบสมมติฐานที่ว่าเราจะต้องทราบ population variance เสียก่อนนั่นเอง sample variance (s^2) ใน Equation (2) เราคำนวณจาก sample ได้ ดังนั้นถึงแม้ว่า u จะมีประโยชน์ในทางปฏิบัติอย่างจำกัด แต่มันก็เป็นแนวทางที่ทำให้เราทราบถึง t ได้

เนื่องจาก mean ของ all possible sample means ซึ่งมี sample size อย่างเดียวกัน ($\mu_{\bar{y}}$) มีค่าเท่ากับ population mean (μ) เราจึงพอจะคาดหมายได้ว่า mean ของ t เท่ากับศูนย์ เช่นเดียวกับ mean ของ u และเรายังพอจะคาดหมายได้อีกด้วยว่า variance ของ t จะมากกว่า variance ของ u (variance ของ u = 1, ข้อ 6.7) statistic u เกิดจาก 4 elements คือ $\bar{y}$, μ, σ^2, และ n ใน 4 elements นี้ $\bar{y}$ ตัวเดียวเท่านั้นที่มีค่าเปลี่ยนไปได้ต่าง ๆ กันตาม samples แต่ statistic t นั้น ถึงแม้จะเกิดจาก 4 elements คือ $\bar{y}$, μ, s^2 และ n เช่นเดียวกันก็จริง แต่ใน 4 elements นี้มี $\bar{y}$ และ s^2 รวม 2 ตัวที่มีค่าเปลี่ยนไปได้ต่าง ๆ กันตาม samples ผลที่เกิดขึ้นก็คือ t จะมีค่าสูง ๆ ต่ำ ๆ ไปตาม samples ต่าง ๆ ได้มากกว่า u เพราะฉะนั้นเราจึงเชื่อว่า variance ของ t ต้องมากกว่า 1 ซึ่งเป็น variance ของ u

t-distribution ไม่ได้เป็น frequency curve เดียวเพียงเส้นเดียว แต่มันเป็น family of curves, d.f. ของ s^2 จะระบุถึง t-curve เส้นหนึ่งโดยเฉพาะ ถ้า d.f. เพิ่มขึ้นความแปรผันในค่าของ s^2 ระหว่าง samples ต่าง ๆ จะลดลง ผลที่เกิดขึ้นก็คือ เมื่อ d.f. ของ s^2 เพิ่มขึ้นความแปรผันในค่าของ t จะลดน้อยลง และเมื่อ d.f. เข้าไปใกล้ ∞, s^2 จะมีค่าใกล้ σ^2 เข้าไปทุกที ทำให้ค่าของ t เข้าไปใกล้ค่าของ u เข้าทุกทีด้วย ดังนั้น u จึงกลายเป็นกรณีพิเศษของ t อนึ่ง เราเรียก d.f. ของ s^2 เป็น d.f. ของ t ด้วย หรืออีกนัยหนึ่ง t-distribution อันหนึ่งโดยเฉพาะนั้นจะถูกระบุให้ทราบโดย d.f. ของ s^2 ใน Equation (2) จากกราฟของ t-distributions with 1, 4 และ ∞ d.f. ที่ให้ไว้ใน Fig. 8.1 นั้น จะเห็นได้ว่า t-curve เป็นเส้นโค้งรูประฆังวางคว่ำและดูคล้ายกับ normal curve มากทีเดียว เพราะฉะนั้นการดูจากกราฟเพียงผิวเผินจะไม่อาจบอกได้ว่า t-curve และ normal curve แตกต่างกันตรงไหน relative frequency เท่านั้นที่แสดงถึงความแตกต่างของ frequency curve อย่างหนึ่งจาก frequency curve อย่างอื่น t-distribution with ∞ d.f. ที่แสดงไว้ใน Fig. 8.1 นั้น ก็คือ u-distribution ซึ่งเป็น normal distribution ที่มี mean = 0 และ variance = 1

การพิจารณาที่แล้วมาทั้งหมดอาจสรุปเป็น theorem ได้ดังต่อไปนี้

Theorem 8.1 a: If all possible samples of size n are drawn from a normal population with mean equal to μ, and for each sample the statistic t, where

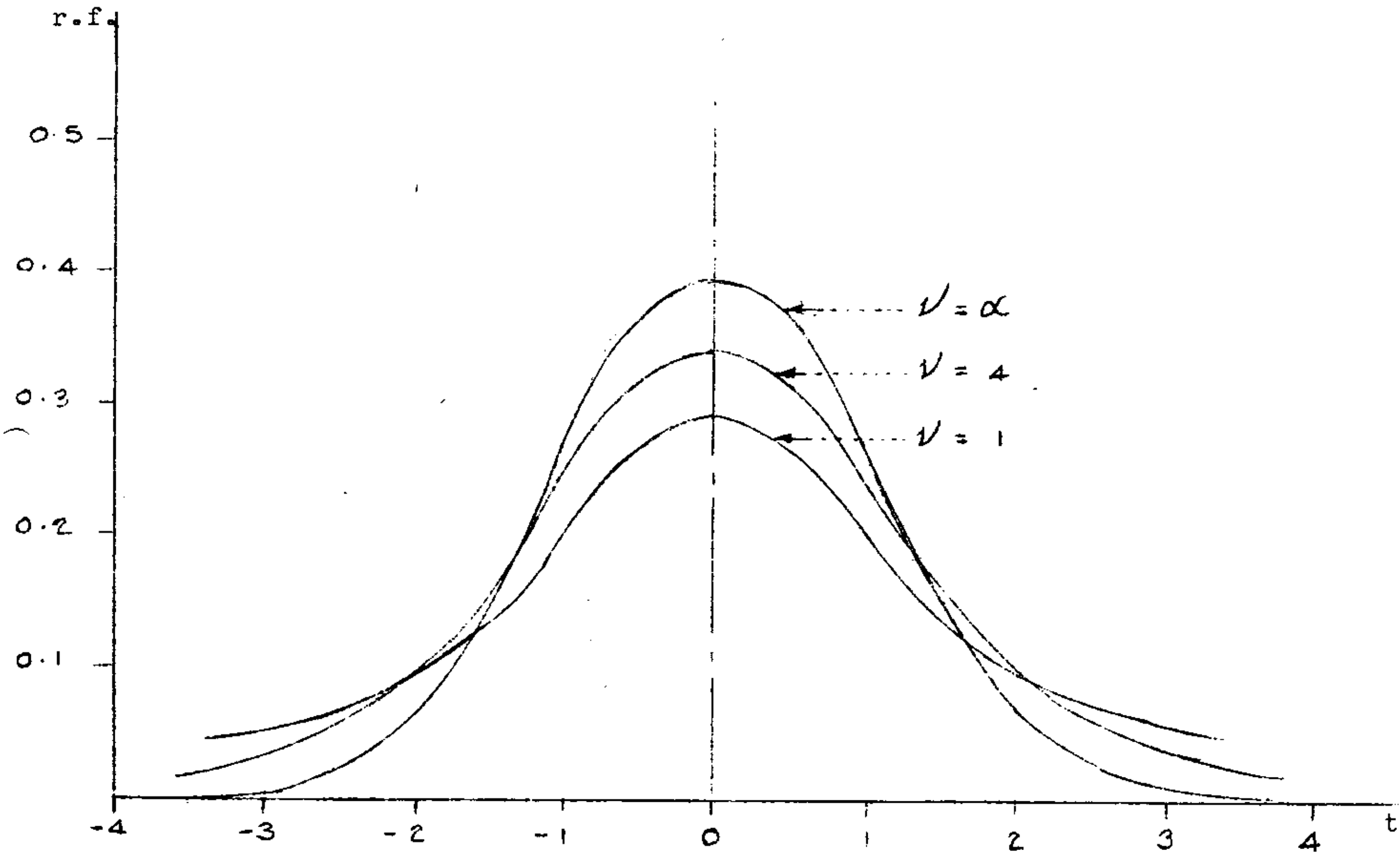


Fig. 8.1

$$t = \frac{\bar{y} - \mu}{\sqrt{\frac{s^2}{n}}} , \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

is calculated, the frequency distribution of the t-values follows the Student's t-distribution with ν d.f., where ν is the d.f. of s^2 (ν = n - 1 in this case).

<u>Theorem 8.1 b:</u> As the number of degrees of freedom of s^2 approaches infinity, the Student's t-distribution approaches the normal distribution with mean equal to zero and variance equal to 1, that is, t approaches u, as ν approaches infinity.

การแสดงให้เห็นความเป็นจริงของ Theorem 8.1 a จะได้กล่าวไว้ในข้อต่อไปนี้

8.2 Experimental Verification of t-Distribution

รายละเอียดของ sampling experiment ที่จะนำมาแสดงให้เห็นความเป็นจริงของ Theorem 8.1 a นั้นได้ให้ไว้แล้วในบทที่ 4 กล่าวโดยย่อเรามี 1,000 samples แต่ละ sample ประกอบด้วย 5 observations ซึ่งได้มาจาก tag population ซึ่งเป็น normal population ที่มี mean = 50 และ variance = 100 เราคำนวณค่าของ t ของทุก sample เอาไว้ ตัวอย่างการคำนวณค่าของ t ของ sample ที่ประกอบด้วย 5 observations คือ 50, 57, 42, 63 และ 32 นั้น จะแสดงให้เห็นดังต่อไปนี้

$$n = 5$$

$$\Sigma y = 50 + 57 + 42 + 63 + 32 = 244$$

$$\bar{y} = \frac{244}{5} = 48.8$$

$$(\Sigma y)^2 = (244)^2 = 59,536$$

$$\frac{(\Sigma y)^2}{n} = \frac{59,536}{5} = 11,907.2$$

$$\Sigma y^2 = (50)^2 + (57)^2 + (42)^2 + (63)^2 + (32)^2 = 12,506$$

$$SS = 12,506 - 11,907.2 = 598.8 \quad \text{............ (ข้อ 7.4)}$$

$$s^2 = \frac{598.8}{4} = 149.7 \quad \text{........................ (ข้อ 7.4)}$$

$$\frac{s^2}{n} = \frac{149.7}{5} = 29.94$$

$$\sqrt{\frac{s^2}{n}} = \sqrt{29.94} = 5.472$$

$$\bar{y} - \mu = 48.8 - 50 = -1.2$$

$$t = \frac{-1.2}{5.472} = -0.219$$

ทุกๆ sampleใน 1,000 samples นี้ เราคำนวณค่าของ t ไว้ตามวิธีที่แสดงข้างต้น ดังนั้นเราจะได้ค่าของ t ทั้งหมด 1,000 ค่า ค่าของ t สำหรับ 4 samples ได้ให้ไว้แล้วใน Table 4.2 เนื่องจาก s^2 ในกรณีนี้มี n-1 d.f. หรือ 4 d.f., t จึงมี 4 d.f. ด้วย frequency table ของค่าของ t 1,000 ค่าได้ให้ไว้ใน Table 8.2

Table 8.2

t	observed frequency		theoretical	midpoint	mf
	f	r.f. (%)	r.f. (%)	m.	
below − 4.5	8	.8	.5	− 5	− 40
− 4.5 to − 3.5	6	.6	.7	− 4	− 24
− 3.5 to − 2.5	23	2.3	2.1	− 3	− 69
− 2.5 to − 1.5	85	8.5	7.1	− 2	− 170
− 1.5 to − 0.5	218	21.8	21.8	− 1	− 218
− 0.5 to 0.5	325	32.5	35.6	0	0
0.5 to 1.5	219	21.9	21.8	1	219
1.5 to 2.5	80	8.0	7.1	2	160
2.5 to 3.5	25	2.5	2.1	3	75
3.5 to 4.5	4	.4	.7	4	16
above 4.5	7	.7	.5	5	35
total	1,000	100.0	100.0		− 16

$$\text{mean of } t = \frac{\Sigma mf}{\Sigma f} = \frac{-16}{1,000} = -0.016$$

theoretical frequency ที่ได้ไว้ใน Table 8.2 เป็นของ t-distribution with 4 d.f., histogram ของค่าของ t 1,000 ค่า ซึ่งเขียนทับลงบน t-curve with 4 d.f. ได้แสดงไว้ใน Fig. 8.2

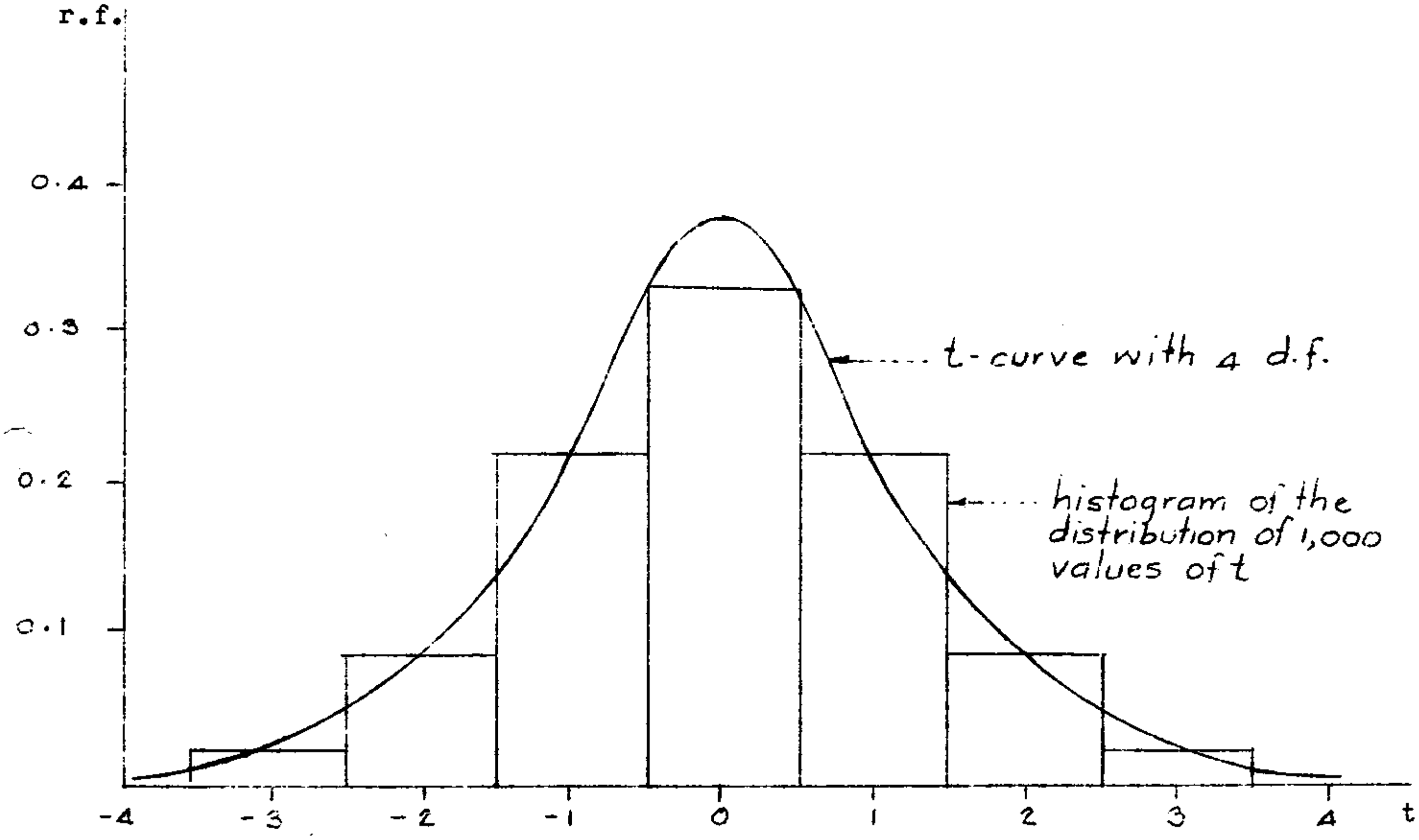


Fig. 8.2

จาก Table 8.2 หรือ Fig. 8.2 เราจะเห็นได้ว่า observed frequency และ theoretical frequency เกือบจะเท่ากัน observed frequency ของ t ขึ้นอยู่กับค่าของ t ของ 1,000 samples แต่ theoretical frequency ขึ้นอยู่กับค่าของ t ของ all possible samples ที่มี size = 5 ดังนั้น observed frequency จึงไม่เท่ากับ theoretical frequency

mean ของค่าของ t 1,000 ค่าจะหาได้ง่ายถ้าเรามีค่าของ t ครบทุกค่า แต่เมื่อเราทำ frequency table คือ Table 8.2 ขึ้นมาแล้วค่าของ t แต่ละค่าจะหายไป ค่าของ mean ของ t โดยประมาณอาจหาได้โดยใช้ midpoint (m) ของ class หนึ่งแทนค่าของ t ทั้งหมดของ class นั้น ตัวอย่างเช่น class – 0.5 ถึง 0.5 ใช้ 0 เป็นค่าของ t ทั้งหมดของ class และ class 0.5 ถึง 1.5 ใช้ 1 เป็นค่าของ t ทั้งหมดของ class class หัวและ class ท้ายไม่มีขอบเขตแน่นอนซึ่งเราจะกำหนดให้ midpoints ของ 2 classes นี้เป็น – 5 และ 5 ตามลำดับ ดังนั้น mean ของ t 1,000 ค่าโดยประมาณคือ

$$\frac{\Sigma\ mf}{\Sigma\ f} = \frac{-16}{1,000} = -0.016$$

ซึ่งมีค่าใกล้เคียงกับ 0 มากดังที่เราคาดหมายไว้

จาก Table 8.2 เรายังสังเกตได้อีกว่า variance ของ t มีค่ามากกว่า variance ของ u ตัวอย่างเช่น relative frequency ของ u ที่เลยจุด – 3.5 และ 3.5 ออกไปแล้วเกือบจะเท่ากับศูนย์ (ดู Table 3, Appendix) แต่ใน Table 8.2 นี้เราจะเห็นว่า 1.4 % (0.8 + 0.6) ของค่าของ t 1,000 ค่าน้อยกว่า – 3.5 และ 1.1 % (0.4 + 0.7) ของค่าของ t 1,000 ค่ามากกว่า 3.5 นี่แสดงให้เห็นว่า t-curve แผกว้างกว่า u-curve เพราะฉะนั้น variance ของ t ต้องมากกว่า variance ของ u หรือมากกว่า 1 อย่างแน่นอน

experiment นี้ได้แสดงให้เห็นความเป็นจริงของ t-distribution และยืนยันความคาดหมาย (ข้อ 8.1) ที่ว่า mean ของ t มีค่าเท่ากับ 0 และ <u>variance ของ t มีค่ามากกว่า 1 แล้ว</u>

8.3 t-Table

relative cumulative frequency (r.c.f.) ของ t-distribution สำหรับ d.f. ต่าง ๆ กันนั้นได้ให้ไว้แล้วใน Table 6, Appendix แต่ละบรรทัดของ t-table แทน d.f. อย่างหนึ่งโดยเฉพาะ ตัวอย่างเช่น สำหรับ 4 d.f., 2.5 % ของค่าของ t จะมากกว่า 2.776 เนื่องจาก t-curve เป็น symmetrical ค่าของ t ในตารางนี้จึงชี้ให้เห็นด้วยว่า 2.5 % ของค่าของ t นั้นน้อยกว่า – 2.776 ในขณะที่ d.f. เพิ่มขึ้นค่าต่าง ๆ ในคอลัมน์ 2.5 % ของ t-table จะลดน้อยลงทุกทีจนถึง 1.960 ซึ่งสุดเขตของมันที่มี d.f. เป็น ∝ นี่แสดงให้เห็นว่า t จะมีค่าเข้าไปใกล้ u เข้าทุกทีเมื่อ d.f. เข้าไปใกล้ infinity (เพราะว่า 2.5 % ของค่าของ u มากกว่า 1.960 และน้อยกว่า – 1.960) และยังแสดงให้เห็นอีกด้วยว่า variance ของ t จะลดน้อยลงเมื่อ d.f. เพิ่มขึ้น

8.4 Test of Hypothesis

ข้อ 8.1, 8.2 และ 8.3 ที่แล้วมาเป็นเรื่อง deductive relation ระหว่าง population หนึ่งกับ samples ต่าง ๆ ของมัน หรือพูดให้เจาะจงลงไปอีกก็คือเป็นเรื่อง distribution ของค่าของ t ของ all possible samples ซึ่งมี size เดียวกันและได้จาก normal population หนึ่งที่กำหนด

ข้อ 8.4 นี้เป็นเรื่องการ draw inductive inference ที่เกี่ยวกับ population จาก sample หนึ่งที่กำหนด หรือพูดเจาะจงลงไปให้ชัดก็คือเป็นเรื่องการทดสอบสมมติฐานว่า population mean เท่ากับค่าที่กำหนด เป็นการจำเป็นที่เราจะต้องสร้าง t-distribution ขึ้นก่อนที่การทดสอบสมมติฐานจะถูกพิจารณา เพราะว่าค่าของ t ต่าง ๆ ซึ่งเราต้องการเพื่อสร้าง critical regions นั้นย่อมได้มาจาก t-table ที่ทำขึ้นตาม distribution ของค่าของ t ต่าง ๆ ของ all possible samples ซึ่งมี size เดียวกัน และได้จาก normal population เดียวกัน

ประโยชน์ของ t นั้นคล้ายกับประโยชน์ของ u ซึ่งใช้ทดสอบสมมติฐานว่า population mean เท่ากับค่าที่กำหนด ความแตกต่างกันมีเพียงว่าใน t-test เราใช้ sample variance (s^2) แต่ใน u-test เราใช้ population variance (σ^2) เนื่องจาก distribution ของ t และของ u ไม่ใช่ distribution เดียวกัน critical regions ของมันจึงแตกต่างกัน ถ้าเราใช้ 5 % significance level ค่าทั้งสองตรงเขต critical regions สำหรับ two-tailed u-test คือ -1.960 และ 1.960 แต่สำหรับ t-test นั้นค่าทั้งสองตรงเขต critical regions จะเป็นไปตาม corresponding values ใน t-table และตาม d.f. ของมัน ตัวอย่างเช่น สำหรับ t with 4 d.f. นั้น ถ้าเราใช้ 5 % significance level ค่าทั้งสองตรงเขต critical regions คือ -2.776 และ 2.776 (ดู Table 6, Appendix)

การพิจารณาทั้งหมดในเรื่อง significance level, Type II error, และ sample size ซึ่งเกี่ยวกับ u-test ที่ได้ให้ไว้ในข้อ 6.4, 6.5, 6.6 และ 6.7 นั้นย่อมใช้ได้กับ t-test เพื่อไม่ให้ต้องกล่าวซ้ำมากนักเราจะพูดเฉพาะเรื่องตอนแรกเท่านั้น การทดสอบสมมติฐานจะถูกอธิบายได้โดย sampling experiment ของข้อ 8.2 สิ่งสำคัญที่เราต้องเข้าใจก็คือ ในการแสดงความเป็นจริงของ t-distribution with 4 d.f. นั้น true population mean 50 ถูกใช้ในการคำนวณค่าของ t 1,000 ค่า นั่นคือ

$$t = \frac{\bar{y} - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{\bar{y} - 50}{\sqrt{\frac{s^2}{5}}}$$

อย่างไรก็ดี ในการทดสอบสมมติฐานว่า population mean เท่ากับ 50 critical regions จะอยู่ที่ $t < -2.776$ และ $t > 2.776$ (ถ้าใช้ 5 % significance level) เนื่องจาก 5 % (r.f.)

ของ all possible samples ซึ่งมี size 5 ให้ค่าของ t ต่าง ๆ ตกอยู่ภายใน critical regions ทั้งสองนี้และนำไปสู่ conclusion ผิดว่า population mean ไม่เท่ากับ 50 หรือความน่าจะเป็นของหนึ่ง random sample ที่จะเกิด Type I error นั้นเป็น 0.05

8.5 Procedures

วิธีดำเนินการทดสอบสมมติฐานโดย t-test อาจแสดงให้เห็นโดยตัวอย่างซึ่งจะใช้ observations ที่มีค่าเป็นเลขหลักหน่วยเพื่อทำให้วิธีการคำนวณง่ายต่อการติดตาม observations ของ sample หนึ่งที่กำหนดคือ 5, 3, 1, 4, 2 เราจะใช้ two-tailed test และ 5 % significance level เพื่อทดสอบสมมติฐานว่า population mean เท่ากับ 5

1. Hypothesis: hypothesis คือ population mean เท่ากับ 5 นั่นคือ

$$\mu_o = 5$$

2. Alternative hypotheses: alternative hypotheses คือ

ก. population mean < 5 หรือ

ข. population mean > 5

3. Assumptions: sample ที่กำหนดให้นี้เป็น random sample ซึ่งได้มาจาก normal population

4. Level of significance: เลือกใช้ 5 % significance level

5. Critical regions: critical regions อยู่ที่

$t < -2.776$ และ

$t > 2.776$

6. Computation of t:

$$\mu_o = 5$$

$$n = 5$$

$$\sum y = 5+3+1+4+2 = 15$$

$$\bar{y} = \frac{15}{5} = 3$$

$$(\Sigma y)^2 = (15)^2 = 225$$

$$\frac{(\Sigma y)^2}{n} = \frac{225}{5} = 45$$

$$\Sigma y^2 = (5)^2 + (3)^2 + (1)^2 + (4)^2 + (2)^2 = 55$$

$$SS = 55 - 45 = 10$$

$$s^2 = \frac{10}{5 - 1} = 2.5$$

$$\frac{s^2}{n} = \frac{2.5}{5} = 0.5$$

$$\sqrt{\frac{s^2}{n}} = \sqrt{0.5} = 0.7071$$

$$\bar{y} - \mu_o = 3 - 5 = -2$$

$$t = \frac{\bar{y} - \mu_o}{\sqrt{\frac{s^2}{n}}} = \frac{-2}{0.7071} = -2.83 \text{ with 4 d.f.}$$

7. **Conclusion:** เนื่องจากค่าของ t อยู่ภายใน critical region ข้างซ้าย conclusion คือ population mean น้อยกว่า 5 (ถ้าค่าของ t อยู่ระหว่าง − 2.776 ถึง 2.776, conclusion คือ population mean เท่ากับ 5 แต่ถ้าค่าของ t มากกว่า 2.776, conclusion คือ population mean มากกว่า 5)

ให้สังเกตว่า t ไม่มีหน่วยเช่นเดียวกับ u ถ้า observations (y) มีหน่วยเป็นนิ้ว $\bar{y}$ จะมีหน่วยเป็นนิ้ว μ_o มีหน่วยเป็นนิ้ว s^2 มีหน่วยเป็น $(\text{นิ้ว})^2$ ส่วน n นั้นไม่มีหน่วย เพราะฉะนั้น หน่วยของ t คือ

$$\frac{\bar{y}\ \text{นิ้ว} - \mu_o\ \text{นิ้ว}}{\sqrt{\frac{s^2\ (\text{นิ้ว})^2}{n}}}$$

$$= \frac{\text{จำนวนนิ้ว}}{\text{จำนวนนิ้ว}}$$

= เลขโดด ๆ ไม่มีหน่วย

ผลในเรื่องนี้ก็คือหน่วยที่ใช้กับ observations ไม่มีความสัมพันธ์ในค่าของ t เลย

โดยทำนองเดียวกันถ้าเราเอาเลขจำนวนหนึ่งลบออกหรือบวกเข้ากับ observations ทั้งหมด ค่าของ t จะไม่ถูกกระทบกระเทือนอีกด้วย เช่น ถ้าเอา 32 บวกเข้ากับ observations แต่ละตัวในตัวอย่างที่แล้ว $\bar{y}$ จะมีค่าเพิ่มขึ้นอีก 32, μ_o ก็มีค่าเพิ่มขึ้นอีก 32 ด้วย ดังนั้นค่าของ $\bar{y} - \mu_o$ จะไม่เปลี่ยนไปแต่อย่างใด ค่าของ sample variance (s^2) ก็ไม่ถูกกระทบกระเทือน (ข้อ 2.4) เพราะฉะนั้นค่าของ t จึงไม่ถูกกระทบกระเทือนเลย แต่ถ้าการบวกเข้าหรือลบออกนี้ทำให้หน่วยของ observations เปลี่ยนไป ค่าของ $\bar{y} - \mu_o$, และ s^2 จะเปลี่ยนไป แต่ค่าของ t จะไม่เปลี่ยน ตัวอย่างเช่น sample ของอุณหภูมิของห้องหนึ่งเป็น 40°C, 50°C, 60°C, 70°C และ 80°C ต้องการทดสอบสมมติฐานว่าอุณหภูมิเฉลี่ยของห้องเท่ากับ 50°C ในกรณีนี้

$$\mu_o = 50$$

$$\Sigma y = 40 + 50 + 60 + 70 + 80 = 300$$

$$\bar{y} = \frac{300}{5} = 60$$

$$(\Sigma y)^2 = (300)^2 = 90{,}000$$

$$\frac{(\Sigma y)^2}{n} = \frac{90{,}000}{5} = 18{,}000$$

$$\Sigma y^2 = 1{,}600 + 2{,}500 + 3{,}600 + 4{,}900 + 6{,}400 = 19{,}000$$

$$SS = 19{,}000 - 18{,}000 = 1{,}000$$

$$s^2 = \frac{1{,}000}{4} = 250$$

$$\frac{s^2}{n} = \frac{250}{5} = 50$$

$$\sqrt{\frac{s^2}{n}} = \sqrt{50}$$

$$\bar{y} - \mu_o = 60 - 50 = 10$$

$$t = \frac{\bar{y} - \mu_o}{\sqrt{\frac{s^2}{n}}}$$

$$= \frac{10}{\sqrt{50}} = \frac{10}{7.07} = 1.41 \text{ with 4 d.f.}$$

ถ้าเปลี่ยนหน่วยของอุณหภูมิเป็น fahrenheit (F = 1.8 C + 32) ดังนั้นอุณหภูมิของห้องจะเป็น 104°F, 122°F, 140°F, 158°F และ 176°F ตามลำดับ ค่าของ t จะหาได้ดังนี้

$$\mu_o = 122$$

$$\Sigma y = 104 + 122 + 140 + 158 + 176 = 700$$

$$\bar{y} = \frac{700}{5} = 140$$

$$(\Sigma y)^2 = (700)^2 = 490,000$$

$$\frac{(\Sigma y)^2}{n} = \frac{490,000}{5} = 98,000$$

$$\Sigma y^2 = 10,816 + 14,884 + 19,600 + 24,964 + 30,976$$

$$= 101,240$$

$$SS = 101,240 - 98,000 = 3,240$$

$$s^2 = \frac{3,240}{4} = 810 \text{ .... (เปลี่ยน)}$$

$$\frac{s^2}{n} = \frac{810}{5} = 162$$

$$\sqrt{\frac{s^2}{n}} = \sqrt{162}$$

$$\bar{y} - \mu_o = 140 - 122 = 18 \text{ .... (เปลี่ยน)}$$

$$t = \frac{\bar{y} - \mu_o}{\sqrt{\frac{s^2}{n}}}$$

$$= \frac{18}{\sqrt{162}}$$

$$= \frac{18}{12.7} = 1.41 \text{ with 4 d.f.}$$

เท่ากับค่าของ t ซึ่งมีหน่วยเป็น centigrade

8.6 Applications

t-test นี้อาจใช้ในการควบคุมคุณภาพของสินค้าทางอุตสาหกรรมได้ ตัวอย่างเช่นองุ่นกระป๋อง เมื่อรินน้ำออกแล้วกำหนดให้หนักกระป๋องละ 12 ออนซ์ เป็นมาตรฐานการผลิต (ข้อ 7.12) sample ขององุ่นกระป๋องหนึ่ง sample ซึ่งประกอบด้วย n กระป๋อง (สมมติว่า 10 กระป๋อง) ที่ผลิตนั้นจะถูกนำมาตรวจเป็นคราว ๆ องุ่นในแต่ละกระป๋องจะถูกรินน้ำทิ้งและชั่งน้ำหนักแล้วบันทึกน้ำหนักไว้ โดย sample ที่ประกอบด้วย 10 observations นี้ จะทดสอบสมมติฐานได้ว่า population mean เท่ากับ 12 ออนซ์ ถ้าสมมติฐานถูกยอมรับ conclusion ก็คือมาตรฐานการผลิตยังคงเป็นไปตามที่กำหนดไว้ ถ้าค่าของ t อยู่ภายใน critical region ข้างซ้ายจะชี้ให้เห็นว่าน้ำหนักเฉลี่ยขององุ่นในกระป๋องต่ำกว่ามาตรฐานและต้องดำเนินการแก้ไข ถ้าค่าของ t อยู่ภายใน critical region ข้างขวาจะชี้ให้เห็นว่าน้ำหนักเฉลี่ยขององุ่นในกระป๋องสูงกว่ามาตรฐาน ตัวอย่างนี้แสดงให้เห็นหลักเบื้องต้นของการควบคุมค่าเฉลี่ยของสินค้าเฉพาะอย่างในโรงงานอุตสาหกรรม อย่างไรก็ดีในทางปฏิบัติมักไม่ค่อยจะได้ใช้ t-test เพื่อความประสงค์ดังกล่าวนี้ เพราะว่าการคำนวณ t นั้นเสียเวลา เขาจึงใช้วิธีอื่นซึ่งง่ายกว่า

8.7 Paired Observations

การใช้ t-test ซึ่งได้กล่าวในข้อที่แล้วไม่เกี่ยวกับการทดลองของนักวิทยาศาสตร์อย่างกว้างขวาง ประโยชน์โดยตรงสำหรับนักวิทยาศาสตร์ก็คือการใช้ t-test กับ observations ที่จับคู่กัน (paired observations) ซึ่งจะแสดงให้เห็นชัดโดยตัวอย่างต่อไปนี้

การทดลองนี้ได้ทำขึ้นในรัฐ Oregon ภาคตะวันออกใน ค.ศ. 1950 เพื่อพิจารณาถึงผลของปุ๋ยไนโตรเจนที่มีต่อผลผลิตของหัวบีทหวาน ไร่ทดลองแห่งหนึ่งถูกแบ่งออกเป็น 10 blocks ให้แต่ละ block มีเนื้อที่เท่ากัน และแต่ละ block ถูกแบ่งออกเป็น 2 plots ให้แต่ละ plot มีเนื้อที่เท่ากันอีกชั้นหนึ่ง ดังนั้นจึงมี 10 คู่ของ plots

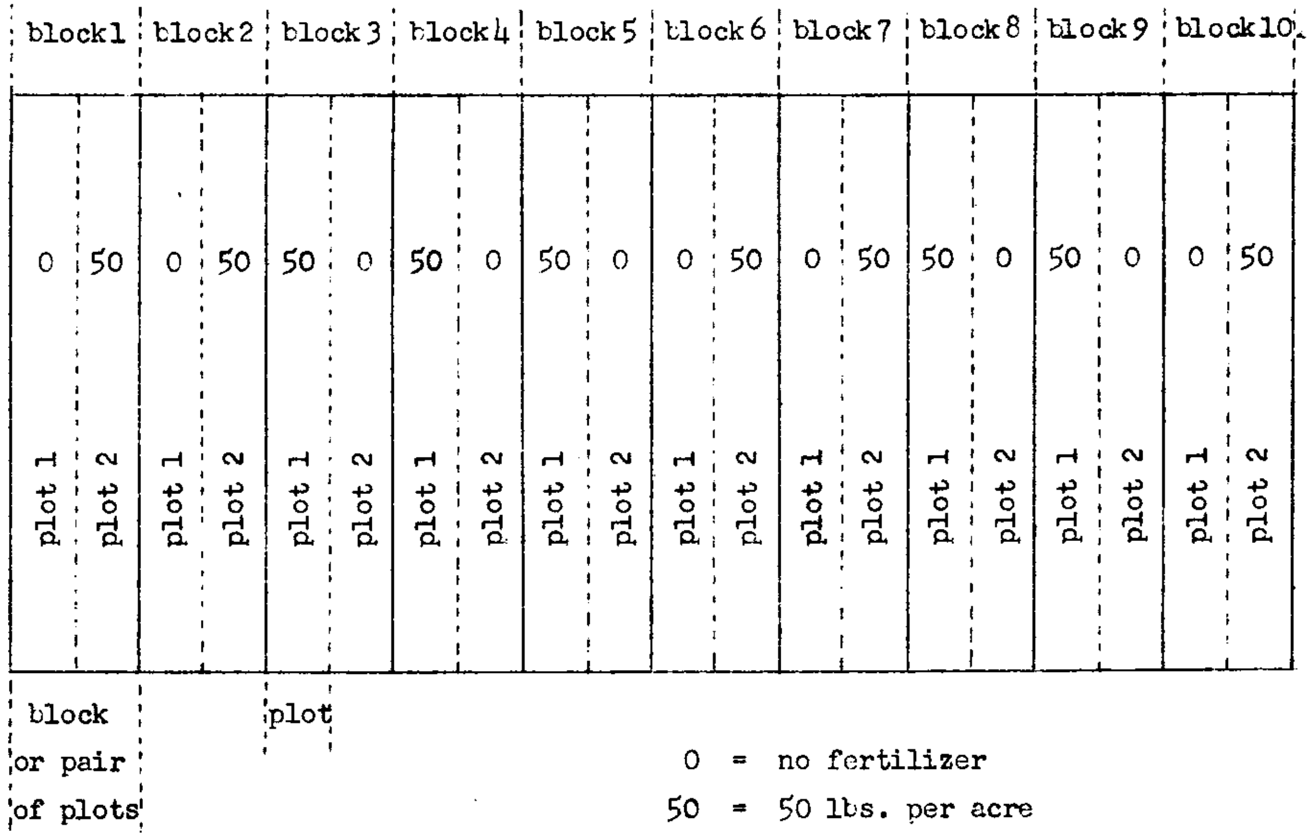


Fig. 8.7

plot หนึ่งของแต่ละคู่ (block) จะถูกเลือกโดยการเสี่ยงเช่นการโยนหัวโยนก้อยและใส่ปุ๋ยลงใน plot นั้น ในอัตรา 50 ปอนด์ของไนโตรเจนต่อเอเคอร์ แต่อีก plot หนึ่งที่เหลือของแต่ละคู่ไม่ใส่ปุ๋ยเลย แผนที่ไร่ทดลองที่มี 20 plots ได้แสดงไว้ใน Fig. 8.7 ผลผลิตของหัวบีทหวานเป็นปอนด์ที่ได้จาก 20 plots นี้ได้ให้ไว้ใน Table 8.7

Table 8.7

block No.	fertilizer (a) 0 lbs.	fertilizer (b) 50 lbs.	y difference (b) - (a)
1	140.4	170.5	30.1
2	174.7	207.4	32.7
3	170.2	215.9	45.7
4	174.6	209.0	34.4
5	154.5	171.6	17.1
6	185.0	201.2	16.2
7	118.9	209.9	91.0
8	169.8	213.3	43.5
9	174.7	184.1	9.4
10	176.7	220.4	43.7
Total	1,639.5	2,003.3	363.8

μ_o	0	$(\Sigma y)^2$	132,350.44	s^2	526.473
n	10	$\frac{(\Sigma y)^2}{n}$	13,235.04	$\frac{s^2}{n}$	52.6473
Σy	363.8	Σy^2	17,973.30	$\sqrt{\frac{s^2}{n}}$	7.2558
$\bar{y}$	36.38	SS	4,738.26		

$$t = \frac{\bar{y} - \mu_o}{\sqrt{\frac{s^2}{n}}} = \frac{36.38 - 0}{7.2558} = 5.014 \quad \text{with 9 d.f.}$$

plots ทั้งสองของ block ตั้งอยู่ชิดกันจนความแปรผันในความอุดมสมบูรณ์ของดินตามธรรมชาติจะมีผลต่อผลผลิตจาก plots ที่ใส่ปุ๋ยและไม่ใส่ปุ๋ยได้น้อยที่สุด plot ที่ใส่ปุ๋ยนั้นถูกเลือกโดยการเสี่ยงเพื่อทำให้เป็นไปตาม assumption ของ random sampling (ข้อ 8.5, item 3) เพราะว่าแต่ละ block ที่มี 2 plots เป็น experiment โดยตัวของมันเอง จึงเท่ากับว่า experiment ถูกทำถึง 10 ครั้ง ผลต่าง (y)

ระหว่างผลผลิตของ plot ที่ใส่ปุ๋ย และ plot ที่ไม่ได้ใส่ปุ๋ยของแต่ละ block ใน 10 blocks ได้ให้ไว้ใน Table 8.7 ให้สังเกตว่าผลต่างนี้จะเปลี่ยนไปตาม blocks ต่าง ๆ ทำให้เกิดเป็นชุดของ 10 ผลต่างขึ้น 10 ผลต่างนี้ประกอบกันเป็นหนึ่ง sample ซึ่งมี 10 observations ที่ได้จาก population ซึ่งประกอบด้วย observations มากมาย เพราะฉะนั้นปัญหาก็คือการ draw an inference เกี่ยวกับ population จาก sample ที่ประกอบด้วย 10 observations นั่นเอง hypothesis ที่จะถูกทดสอบคือ population mean เท่ากับ 0 หรือโดยเฉลี่ยแล้วปุ๋ยไม่ได้เพิ่มผลผลิตของหัวบีทหวาน วิธีดำเนินการทดสอบสมมติฐานก็คงเป็นเช่นเดียวกับที่ได้กล่าวไว้ในข้อที่แล้ว การทดสอบนี้ใช้แบบ two-tailed test เพราะมันเป็นไปได้ว่าปุ๋ยอาจเพิ่มหรือลดผลผลิตได้ทั้งสองอย่าง ถ้าเราใช้ 5 % significance level, critical regions จะอยู่ที่ $t < -2.262$ และ $t > 2.262$ (9 d.f.) การคำนวณค่าของ t ได้แสดงไว้ในครึ่งล่างของ Table 8.7 ค่าของ t คือ 5.014 with 9 d.f. เนื่องจากค่าของ t อยู่ภายใน critical region ข้างขวา conclusion คือการใช้ปุ๋ยไนโตรเจนในอัตรา 50 ปอนด์ต่อเอเคอร์ จะเพิ่มผลผลิตของหัวบีทหวาน

ประโยชน์ของ paired observations ไม่ได้ถูกจำกัดสำหรับการทดลองในไร่นาเท่านั้น เราอาจจะใช้กับการทดลองใด ๆ ที่ประกอบด้วย 2 treatments ได้ทั้งสิ้น เช่น อาหารสัตว์ 2 ชนิดจะถูกนำมาเลี้ยงสัตว์ 2 ฝูง เพื่อพิจารณาถึงคุณค่าของอาหารซึ่งจะแสดงให้เห็นโดยน้ำหนักของสัตว์ที่เพิ่มขึ้นภายหลังระยะเวลาเลี้ยง ก่อนที่จะเริ่มต้นการทดลองจะต้องนำสัตว์มาเข้าคู่กัน สัตว์เหล่านี้อาจเป็นลูกหมู 2 ตัวจากครอกเดียวกันหรือลูกวัว 2 ตัวที่มีน้ำหนักและอายุเท่ากัน สัตว์ตัวหนึ่งในคู่ของมันจะถูกกำหนดให้กับ treatment หนึ่งโดยวิธีเสี่ยง และสัตว์อีกตัวหนึ่งก็จะถูกกำหนดให้กับอีก treatment หนึ่ง

ตัวอย่างอื่นก็ยังมีอีกเช่นการใช้ paired observations ในการเปรียบเทียบวิธีสอน 2 วิธี เด็กกลุ่มหนึ่งเป็นต้นว่า 40 คนจะถูกแบ่งออกเป็น 2 ชั้น ๆ ละ 20 คน แต่ก่อนแบ่งเด็กจะต้องจัดเด็กให้เข้าคู่กันเป็นคู่ ๆ โดยอาศัยหลักบางอย่างเช่น I.Q. ของเด็กเป็นต้น เด็กแต่ละคนในคู่ของตนซึ่งมี I.Q. คล้ายกันจะถูกกำหนดให้อยู่ในชั้นหนึ่งโดยวิธีเสี่ยง เด็กที่เหลืออีกคนหนึ่งก็จะอยู่ในอีกชั้นหนึ่ง ทั้งสองชั้นนี้จะสอนวิชาเดียวกันโดยครูคนเดียวกันแต่ใช้วิธีสอน 2 วิธีซึ่งต่างกัน เป็นต้นว่าวิธีหนึ่งเป็นการสอนซึ่งใช้ทัศนาการเข้าช่วย (visual aids) อีกวิธีหนึ่งไม่ใช้ทัศนาการเข้าช่วย เมื่อหมดภาคเรียนจะสอบเด็กทั้ง 2 ชั้นโดยข้อสอบเดียวกัน จากคะแนนสอบของเด็ก 40 คนจะได้ 20 ผลต่างของคะแนนสอบจากเด็ก 20 คู่ เพราะฉะนั้น sample size จะเป็น 20 ผลต่างระหว่างคะแนนสอบของเด็กคู่หนึ่งอาจเป็นค่าบวกหรือค่าลบก็ได้ เครื่องหมายบวก (+) หรือลบ (−) จะทิ้งไปไม่ได้ การทดลองวิธีนี้ใช้พิจารณาว่าการสอนวิธีใดใน 2 วิธีนี้จะเหมาะแก่ครูผู้หนึ่งโดยเฉพาะกว่าอีกวิธีหนึ่ง มากกว่าจะพิจารณาถึงคุณลักษณะของวิธีสอน 2 วิธีนั้น

พึงจำไว้ว่าวิธี paired observations นี้จะใช้ได้กับ 2 treatments เท่านั้น ถ้ามีมากกว่า 2 treatments เช่น อาหารสัตว์ 4 ชนิด หรือวิธีสอน 3 วิธี เกี่ยวข้องกับการทดลองแล้วจะต้องใช้ experimental design ที่แตกต่างออกไปซึ่งเรียกว่า "randomized blocks" (บทที่ 14) randomized block design นี้จะใช้กับกี่ treatments ก็ได้ วิธี paired observations จึงเป็นกรณีพิเศษของ randomized block design (ข้อ 14.6) ดังนั้น ถ้ามี 2 treatments เราอาจใช้วิธี paired observations หรือวิธี randomized block design ก็ได้ conclusions ที่ได้จาก 2 วิธีนี้จะต้องเป็นอย่างเดียวกันเสมอ ความแตกต่างใน conclusion ต้องเกิดจากความผิดในการคำนวณอย่างแน่นอน

ข้อดีของการจับคู่เด็ก, คู่ plots, หรือคู่สัตว์ นั้น ก็เพื่อลด experimental error ลง ความแปรผันจากคู่หนึ่งไปยังคู่อื่น ๆ จะถูกขจัดให้หมดไปโดยการจับคู่กันนี้ ตัวอย่างเช่น ถ้าเอา 10 บวกเข้ากับ 2 observations ใน block 1 ของ Table 8.7 เอา 20 บวกเข้ากับ 2 observations ใน block 2, เอา 30 บวกเข้ากับ 2 observations ใน block 3, และต่อ ๆ ไปตามลำดับ ค่าของ t จะไม่ถูกกระทบกระเทือนเลยเพราะว่า 10 ผลต่างนั้นยังมีค่าคงเดิม ถ้าความแปรผันในความอุดมสมบูรณ์ของดินทำให้เกิดการเพิ่มขึ้นหรือลดลงของผลผลิตได้แล้ว ความแม่นยำของการทดลองจะไม่ถูกกระทบกระเทือนแต่อย่างไร โดยทำนองเดียวกัน I.Q. ของเด็กอาจเปลี่ยนไปต่าง ๆ กันได้มากจากเด็กคู่หนึ่งไปยังเด็กคู่อื่น ๆ แต่ความแม่นยำของการทดลองก็จะไม่ถูกกระทบกระเทือน ยิ่งกว่านั้น ถ้าใช้วิธีสอนหลายวิธีกับเด็กที่มี I.Q. ต่างกัน การทดลองย่อมมีข้อดีที่มี inductive basis กว้างขวาง conclusion ที่ได้จากเด็กที่มี I.Q. ต่าง ๆ นั้นควรจะเป็นจริงโดยทั่วไปได้มากกว่า conclusion ที่ได้จากเด็กที่มี I.Q. เฉพาะ

8.8 Remarks

วิวัฒนาการที่สำคัญอย่างหนึ่งของบทนี้ก็คือการใช้ sample mean ($\bar{y}$) และ sample variance (s^2) เป็น elements ในวิธี draw an inference เกี่ยวกับ population mean (μ) t-test เกี่ยวข้องกับ 4 elements คือ $\bar{y}$, μ_o, s^2 และ n ความช่วยเหลือร่วมกันของ elements เหล่านี้ทำให้เราสามารถเข้าถึง conclusion ได้ elements $\bar{y}$ และ s^2 ไม่ได้ถูกคำนวณขึ้นมาเพื่อประมาณค่าของ μ และ σ^2 เป็นจุดมุ่งหมายสุดท้าย แต่ $\bar{y}$ และ s^2 เป็นตัวกลางในการคำนวณค่าของ t ซึ่งทำให้เราสามารถเข้าถึง conclusion ที่เกี่ยวกับ population mean ได้

Chapter 9

Variance Ratio, F-Distribution

ในการพิจารณาลักษณะของ sample variance (s^2) ในบทที่ 7 นั้น เราได้หาทั้ง deductive และ inductive ralations ระหว่าง population variance (σ^2) และ sample variance (s^2) มาแล้ว ในบทนี้เราจะขยายเรื่องของ sample variance ออกไปอีก คือ เราจะไม่กล่าวถึง population หนึ่งกับ samples ต่าง ๆ ของมันเหมือนอย่างบทที่ 7 แต่จะกล่าวถึง populations คู่หนึ่ง และ samples ต่าง ๆ ของมันแต่ละ population ด้วย กล่าวให้เจาะจงยิ่งขึ้นก็คือในบทนี้เราจะไม่เน้นหนักในเรื่อง variance (s^2) ของหนึ่ง sample แต่จะเน้นหนักในเรื่องเรโช (s_1^2/s_2^2) ของคู่ของ sample variances, s_1^2 และ s_2^2 คู่หนึ่ง

ยิ่งกว่านั้น ในบทนี้จะได้นำ frequency distribution ใหม่อีกอย่างหนึ่งซึ่งเรียกว่า F-distribution มากล่าวด้วย R.A. Fisher เป็นผู้ค้นคิด frequency distribution ขึ้นอย่างหนึ่งเรียกว่า Z-distribution ซึ่งต่อมา Snedecore เป็นผู้ดัดแปลงแก้ไขให้ดีขึ้นและตั้งชื่อใหม่ว่า F-distribution เพื่อเป็นเกียรติแก่ Fisher แต่เราจะไม่กล่าวถึง Z-distribution ดั้งเดิมในที่นี้

9.1 Description of F-Distribution

การพิจารณาในข้อนี้เกี่ยวกับ 2 populations และเพื่อไม่ให้สับสนใน 2 populations และ samples ต่าง ๆ ของมันเราจะใช้เลข 1 และ 2 กำกับไว้กับ notations ต่าง ๆ เช่น σ^2, s^2, n, และ ν ทุกตัว เช่น σ_1^2 คือ variance ของ population ที่หนึ่ง และ σ_2^2 คือ variance ของ population ที่สอง ดังนี้เป็นต้น

เราอาจอธิบาย F-distribution ได้โดย sampling experiment จาก population หนึ่งซึ่งมี variance เท่ากับ σ_1^2 เราอาจหา all possible samples ซึ่งมี size เท่ากับ n_1 ออกมาได้ และแต่ละ sample จะคำนวณค่า variance (s_1^2) ได้ จากอีก population หนึ่งซึ่งมี variance เท่ากับ σ_2^2 เราอาจหา all possible samples ซึ่งมี size เท่ากับ n_2 ออกมาได้ และแต่ละ sample จะคำนวณค่า variance (s_2^2) ได้ ดังนั้นจะมีค่าของ s^2 2 ชุด แต่ละชุดได้มาจากแต่ละ population ของ 2 populations นั้น s_1^2 ทุกตัวของชุดที่หนึ่งจะถูกหารด้วย s_2^2 ทุกตัวของชุดที่สองเพื่อทำให้เกิดเป็น variance ratio (s_1^2/s_2^2) ถ้ามี 9 possible samples (ข้อ 5.1) ซึ่งได้

มาจาก population ที่หนึ่ง และ 25 possible samples ซึ่งได้มาจาก population ที่สองก็จะมี 9 x 25 หรือ 225 possible pairs of samples และมี 225 variance ratios เนื่องจากค่าของ s_1^2 และ s_2^2 ทั้งสองนี้เปลี่ยนไปได้ต่าง ๆ กันตาม samples ค่าของ variance ratio (s_1^2/s_2^2) จึงเปลี่ยนไปได้ต่าง ๆ กันตามคู่ของ samples ด้วย ถ้า 2 populations นั้นเป็น normal และมี variances เท่ากันแล้ว frequency distribution ของ variance ratios เหล่านี้เรียกว่า F-distribution F-distribution มี 2 d.f. คือ ν_1 และ ν_2 ซึ่งเป็น d.f. ของ s_1^2 และ s_2^2 ตามลำดับ d.f. ตัวที่หนึ่งของคู่ของ d.f. ของ F หมายถึง d.f. ของ s_1^2 ซึ่งเป็นเศษ (numerator) เสมอ และ d.f. ตัวที่สองเป็น d.f. ของ s_2^2 ซึ่งเป็นส่วน (denominator) เสมอเช่นเดียวกัน F-distribution เป็น family of frequency curves, F-distribution เฉพาะเส้นหนึ่งจะถูกระบุให้เห็นแตกต่างกับ F-distributions อื่น ๆ โดยคู่ของ d.f. กราฟของ 3 F-curves with 1 and 4, 4 and 4, 4 and 25 d.f. ได้แสดงไว้ใน Fig. 9.1

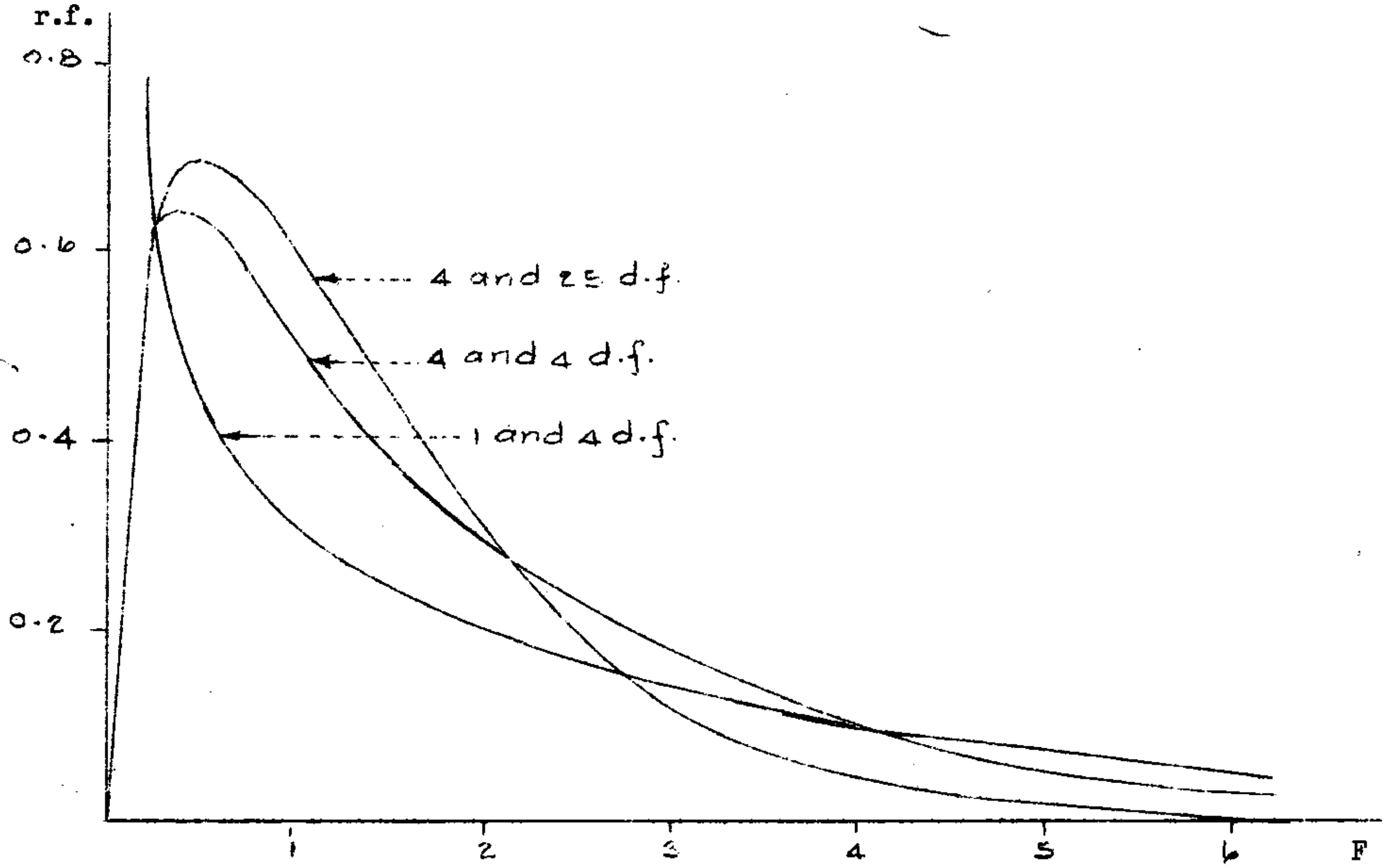


Fig. 9.1

การพิจารณาทั้งหมดอาจสรุปกล่าวไว้ใน theorem ต่อไปนี้

<u>Theorem 9.1</u> : From a normal population with variance equal to σ_1^2, all possible samples of size n_1 are drawn and, for each sample, the variance s_1^2 with $\nu_1 = n_1 - 1$ degrees of freedom is computed. From another normal population with variance equal to σ_2^2, all possible samples of size n_2 are drawn and, for each sample, the variance s_2^2 with $\nu_2 = n_2 - 1$ degrees of freedom is computed. The frequency distribution of all possible ratios

$$F = \frac{s_1^2}{s_2^2} \quad \dots\dots\dots\dots\dots\dots\dots (1)$$

follows the F-distribution with ν_1 and ν_2 degrees of freedom, <u>if $\sigma_1^2 = \sigma_2^2$</u>.

คำว่า "all possible ratios" ตามที่กล่าวไว้ใน theorem ข้างบนนี้บอกเป็นนัยว่า แต่ละ s_1^2 ของ all possible samples ที่ได้มาจาก population ที่หนึ่งมีโอกาสเท่ากันที่จะถูกหารโดย s_2^2 ทุกตัวของ all possible samples ที่ได้มาจาก population ที่สอง นี่เป็นสภาพสำคัญซึ่งจะทำให้ theorem นี้เป็นจริง เมื่อสภาพนี้ถูกต้องแล้วจะพูดได้ว่า variances, s_1^2 และ s_2^2, ถูกแจกแจงอย่างอิสระ Theorem 9.1 จะถูกแสดงให้เห็นจริงโดย sampling experiment ในข้อต่อไปนี้

9.2 Experimental Verification of F-Distribution

รายละเอียดของ experiment ได้ให้ไว้แล้วในบทที่ 4 กล่าวโดยย่อ เรามี 1,000 samples ซึ่งได้มาจาก tag population ซึ่งเป็น normal population ที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 แต่ละ sample ของ 1,000 samples ที่เราคำนวณ s^2 ไว้ ต่อไปนี้เราจะพิจารณาว่า 500 samples ของ 1,000 samples ที่มีนั้นได้มาจาก normal population หนึ่ง และ 500 samples ที่เหลือได้มาจากอีก normal population หนึ่ง เนื่องจาก 2 populations ที่คิดขึ้นนี้ความจริงเป็น tag population เดียวกัน มันต้องมี variance อย่างเดียวกัน ดังนั้นสภาพของ

Theorem 9.1 จึงถูกต้องสมบูรณ์แล้ว เราจะได้ 500 variance ratios โดยการหาร variance ของ sample ที่หนึ่งด้วย variance ของ sample ที่สอง variance ของ sample ที่สามด้วย variance ของ sample ที่สี่ และต่อ ๆ ไปตามลำดับ ตัวอย่างของค่าของ F 2 ค่าดังกล่าวนี้ได้ให้ไว้ใน Table 4.2 แล้ว frequency distribution ของ 500 variance ratios ได้ให้ไว้ใน Table 9.2

Table 9.2

F	observed frequency		theoretical
	f	r.f. (%)	r.f. (%)
0 – 1	241	48.2	50.0
1 – 2	125	25.0	24.1
2 – 3	50	10.0	10.3
3 – 4	27	5.4	5.2
4 – 5	16	3.2	3.0
5 – 6	10	2.0	1.9
6 – 7	8	1.6	1.2
7 – 8	5	1.0	.9
8 – 9	4	.8	.6
9 – 10	2	.4	.5
10 – 11	2	.4	.3
11 – 12	2	.4	.3
12 – 13	1	.2	.2
13 – 14	0	0	.2
14 – 15	0	0	.2
15 – 16	1	.2	.1
over 16	6	1.2	1.0
total	500	100.0	100.0

theoretical frequency ซึ่งได้แสดงไว้ใน Table 9.2 นี้ เป็น **frequency** ของ **F-distribution with 4 and 4 d.f.** **histogram** ของค่าของ F 500 ค่าซึ่งเขียนทับลงบน **F-distribution with 4 and 4 d.f.** ได้แสดงไว้ใน Fig. 9.2

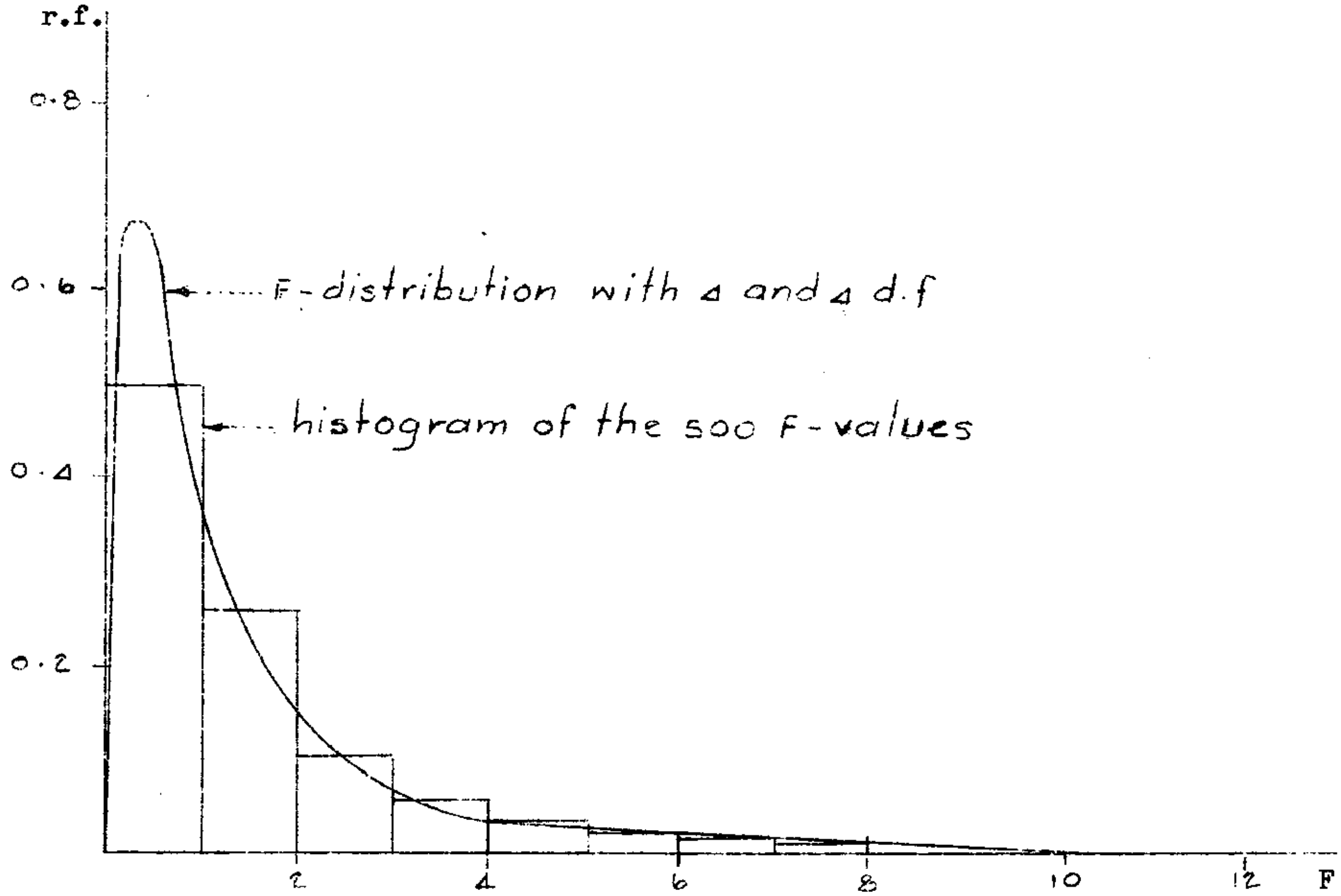


Fig. 9.2

theoretical frequency และ **observed frequency** ทั้งใน **histogram** และ **table** นั้นใกล้เคียงกันมาก นี่แสดงให้เห็นความจริงของ Theorem 9.1 แล้ว

ใน **tag population** ซึ่งมี 500 observations (ข้อ 4.1) นั้น จะมี **possible samples** ที่มี **size** 5 (ข้อ 5.1) ทั้งหมดถึง $(500)^5$ samples หนึ่ง **population** นี้ได้ให้ **sample variances** (s_1^2) เป็นจำนวนมาก **population** ที่คล้ายกันอีกหนึ่ง **population** ก็จะให้ **sample** variances (s_2^2) เป็นจำนวนเท่ากัน ถ้า s_1^2 ทุกตัวถูกหารด้วย s_2^2 ทุกตัวแล้ว จะได้ variance ratios ทั้งหมดถึง $(500)^{10}$ หรือ $9,765,625 \times 10^{20}$ ตัว variance ratios 500 ตัวที่เกี่ยวข้องกับ sampling experiment นี้เป็นแต่เพียงเศษส่วนเล็ก ๆ ของ all possible variance ratios เท่านั้น

พึงสังเกตว่าใน sampling experiment นั้น samples ต่าง ๆ เป็น independent samples (ข้อ 4.2) นี่เป็นสภาพสำคัญอย่างหนึ่งของ Theorem 9.2

9.3 F-Table

percentage points ของ F-distribution ที่มีคู่ของ d.f. ต่าง ๆ กันนั้นได้ให้ไว้ใน Table 7 ของ Appendix แล้ว d.f. ที่เขียนไว้บนหัวตารางเป็น d.f. ของ s_1^2 ของ variance ratio F และ d.f. ที่เขียนไว้ใน column ที่หนึ่งเป็น d.f. ของ s_2^2 Table 7 นี้ประกอบด้วย 4 tables แยกกัน หนึ่ง table สำหรับหนึ่ง percentage point คือ 5 %, 2.5 %, 1 % และ 0.5 % ใน 5 % F-table สำหรับ 4 and 4 d.f. ค่าในตารางคือ 6.3883 จึงกล่าวได้ว่า 5 % ของ F-values with 4 and 4 d.f. นั้นมากกว่า 6.3883 สำหรับ 1 % F-table สำหรับ 5 and 10 d.f. ค่าในตารางคือ 5.6363 ก็หมายความว่า 1 % ของ F-values with 5 and 10 d.f. นั้นมากกว่า 5.6363 และ 1 % F-table ยังแสดงให้เห็นว่า 1 % ของ F-values with 4 and 4 d.f. มากกว่า 15.977 sampling experiment (Table 9.2) ก็แสดงว่า 1.2 % ของ 500 F-values มากกว่า 16 จะเห็นได้ว่า percentages ทั้งสองคือ 1 และ 1.2 นี้ใกล้เคียงกันมากที่สุด

แต่ละ percentage point ของ 4 percentage points ของ F-table จะอยู่บนหางข้างขวาของ F-distribution ถึงแม้ว่า percentage points ต่าง ๆ บนหางข้างซ้ายของ F-distribution เช่น 99.5 % และ 97.5 % points จะไม่ได้ให้ไว้ใน table แต่ก็จะหาออกมาได้จาก table ที่มีโดยการคำนวณธรรมดา 97.5 % point ของ F with ν_1 and ν_2 d.f. จะเท่ากับส่วนกลับของ 2.5 % point ของ F with ν_2 and ν_1 d.f. ตัวอย่างเช่น 97.5 % point ของ F with 6 and 4 d.f. คือ $\frac{1}{6.2272}$ หรือ 0.16059 ในเมื่อ 6.2272 คือ 2.5 % point ของ F with 4 and 6 d.f. ความสัมพันธ์ที่คล้ายคลึงกันนี้จะเกิดขึ้นระหว่าง 99.5 % และ 0.5 % points ด้วยเช่นเดียวกัน เช่น 99.5 % point ของ F with 5 and 10 d.f. คือ $\frac{1}{13.618}$ หรือ 0.073432 ในเมื่อ 13.618 คือ 0.5 % point ของ F with 10 and 5 d.f.

9.4 Test of Hypothesis

ใน 3 ข้อที่แล้วมาเราได้พิจารณาเรื่อง deductive relation จาก 2 populations ไปสู่ samples ต่าง ๆ ของมัน ในข้อนี้เราจะพิจารณาเรื่อง inductive relation จาก 2 samples เข้ามาสู่ 2 populations ของมันบ้าง

เราอาจใช้ Theorem 9.1 ในการทดสอบสมมติฐานว่า 2 population variances, σ_1^2 และ σ_2^2 มีค่าเท่ากัน จาก 2 normal populations เรา draw random sample ออกมา population ละหนึ่ง sample แล้วคำนวณ variances, s_1^2 และ s_2^2 ไว้ และพิจารณา d.f. ของมันด้วย variance ratio F คือ s_1^2/s_2^2 with ν_1 and ν_2 d.f. ถ้า F-value ใกล้กับ 1 แสดงว่า $\sigma_1^2 = \sigma_2^2$ ถ้า F-value เล็กกว่า 1 มาก แสดงว่า σ_1^2 น้อยกว่า σ_2^2 และถ้า F-value โตกว่า 1 มากก็แสดงว่า σ_1^2 มากกว่า σ_2^2 F-value เล็กเท่าไรจึงจะเรียกว่าเล็กมาก หรือโตเท่าไรจึงจะเรียกว่าโตมากนั้นขึ้นอยู่กับ significance level ตัวอย่างเช่น sample variance (s_1^2) เท่ากับ 700 with 4 d.f. และ sample variance (s_2^2) เท่ากับ 50 with 5 d.f. ในการทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$ นั้น ค่าของ F ที่ได้คือ $s_1^2/s_2^2 = 700/50 = 14$ with 4 and 5 d.f. ถ้าเราใช้ 5 % significance level, critical region ข้างขวาจะอยู่ที่ $F > 7.3879$ (2.5 % point ของ F-distribution with 4 and 5 d.f.) และ critical region ข้างซ้ายจะอยู่ที่ $F < \frac{1}{9.3645} = 0.10679$ (9.3645 คือ 2.5 % point ของ F-distribution with 5 and 4 d.f.) เนื่องจากค่าของ F = 14 นี้อยู่ใน critical region ข้างขวา conclusion คือ σ_1^2 มากกว่า σ_2^2 แต่โดยทางทฤษฎีนั้น 2.5 % ของ F-values มากกว่า 7.3879 ถ้าสมมติฐานที่ทดสอบเป็นจริงคือ $\sigma_1^2 = \sigma_2^2$ ในเมื่อ F ตกอยู่ภายใน critical region สมมติฐานจะถูกปฏิเสธ การตัดสินใจนี้ได้ถูกกระทำเนื่องจากความอาจจะเป็นไปได้ว่าสมมติฐานนั้นผิด และ σ_1^2 มากกว่า σ_2^2

การพิจารณาทั้งหมดในเรื่องความสัมพันธ์อย่างแน่นแฟ้นระหว่าง Type I error, Type II error และ sample size ที่ได้ให้ไว้ในข้อ 6.3, 6.4, 6.5 และ 6.6 นั้น จะนำมาใช้ในการทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$

9.5 Procedures

วิธีดำเนินการทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$ อาจแสดงให้เห็นโดยตัวอย่างต่อไปนี้ observations ของ sample ที่หนึ่งคือ 2, 3, 7 และของ sample ที่สองคือ 8, 6, 5, 1 วิธีดำเนินการทดสอบทำดังนี้

1. Hypothesis : hypothesis คือ 2 population variances มีค่าเท่ากัน นั่นคือ

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

2. Alternative hypotheses : alternative hypotheses คือ

   ก. $\sigma_1^2 < \sigma_2^2$ และ

   ข. $\sigma_1^2 > \sigma_2^2$

3. Assumptions : samples ที่กำหนดให้เป็น random samples ซึ่งได้มาจาก normal populations

4. Level of significance : เลือกใช้ 5 % significance level

5. Critical regions : critical regions สำหรับ F with $\overset{3}{2}$ and $\overset{2}{3}$ d.f. อยู่ที่

   $F < 0.025533$ $\left(0.025533 = \frac{1}{39.165}\right.$ และ 39.165 คือ 2.5 % point ของ F-distribution with $\underset{2}{\cancel{3}}$ and $\underset{3}{\cancel{2}}$ d.f.)

   และ $F > 16.044$

6. Computation of F : รายละเอียดของการคำนวณค่าของ F ได้ให้ไว้ใน Table 9.5

Table 9.5

sample No.	1	2
observations	2 3 7	8 6 5 1
$\sum y$	12	20
$(\sum y)^2$	144	400
$\frac{(\sum y)^2}{n}$	48	100
$\sum y^2$	62	126
SS	14	26
s^2	7	8.667

$$SS = \sum y^2 - \frac{(\sum y)^2}{n}$$

$$s^2 = \frac{SS}{n-1}$$

ค่าของ F คือ $\frac{7}{8.667}$ 0.8077 with 2 and 3 d.f.

7. Conclusion : เนื่องจาก F อยู่ภายนอก critical regions, conclusion คือ population variances มีค่าเท่ากัน นั่นคือ $\sigma_1^2 = \sigma_2^2$
(ถ้า F น้อยกว่า 0.025533, conclusion จะเป็น $\sigma_1^2 < \sigma_2^2$
ถ้า F มากกว่า 16.044 conclusion จะเป็น $\sigma_1^2 > \sigma_2^2$)

พึงสังเกตว่า F เป็นเลขไม่มีหน่วย ถ้า observations มีหน่วยเป็นนิ้ว s^2 จะเป็นจำนวนเลขที่มีหน่วยเป็น (นิ้ว)^2 ดังนั้น s_1^2/s_2^2 ซึ่งเป็นจำนวนเลขที่มีหน่วยเป็น (นิ้ว)^2 หารด้วย จำนวนเลขที่มีหน่วยเป็น (นิ้ว)^2 เหมือนกัน ผลลัพธ์ที่ได้คือค่าของ F จะเป็นเลขไม่มีหน่วย

9.6 Weighted Mean of Sample Variances

ในตัวอย่างของข้อ 9.5 เราได้ $s_1^2 = 7$ with 2 d.f. และ $s_2^2 = 8.667$ with 3 d.f. conclusion ที่ได้จากการทดสอบสมมติฐาน คือ 2 population variances มีค่าเท่ากัน ดังนั้น s_1^2 และ s_2^2 ทั้งสองนี้จึงเป็นค่าประมาณของ variance (σ^2) ซึ่งเป็น variance รวมของ populations ทั้งสอง ปัญหาก็คือการผสม s_1^2 และ s_2^2 เข้าด้วยกันเพื่อทำให้เป็นค่าประมาณเดียวของ σ^2 ค่าเฉลี่ยของ s_1^2 และ s_2^2 จะใช้เป็นค่าประมาณของ σ^2 ดังกล่าวนี้ได้ แต่ค่าเฉลี่ยของ s_1^2 และ s_2^2 ในรูป $\frac{s_1^2 + s_1^2}{2}$ นั้นได้ทิ้งความจริงไปอย่างหนึ่ง คือ s_2^2 with 3 d.f. ยอมเป็นค่าประมาณของ σ^2 ที่แม่นยำมากกว่า s_1^2 with 2 d.f. เพราะฉะนั้น เพื่อที่จะให้ค่าประมาณแม่นยำมากขึ้นเราจะใช้ weighted average (s_p^2) เป็นค่าเฉลี่ยของ s_1^2 และ s_2^2 โดยใช้ d.f. ของ s_1^2 และ s_2^2 เป็น weights ของมัน ดังนั้น

$$\boxed{s_p^2 = \frac{\nu_1 s_1^2 + \nu_2 s_2^2}{\nu_1 + \nu_2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

ในเมื่อ s_p^2 = weighted average ของ s_1^2 และ s_2^2

ν_1 = d.f. ของ s_1^2

ν_2 = d.f. ของ s_2^2

weighted average (s_p^2) ของ s_1^2 และ s_2^2 นี้ยังมีชื่ออีกอย่างหนึ่งว่า the pooled estimate of σ^2

ตัวอย่าง : ถ้า $s_1^2 = 7$ with 2 d.f. และ $s_2^2 = 8.667$ with 3 d.f.

the pooled estimate of σ^2 คือ

$$s_p^2 = \frac{(2 \times 7) + (3 \times 8.667)}{2 + 3}$$

$$= \frac{14 + 26}{5}$$

$$= 8$$

พึงสังเกตว่า

$$s^2 = \frac{SS}{\nu} \quad \text{หรือ}$$

$$SS = \nu s^2$$

เพราะฉะนั้น the pooled estimate of σ^2 นั้น ที่แท้ก็คือ

$$\boxed{s_p^2 = \frac{SS_1 + SS_2}{\nu_1 + \nu_2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

with ($\nu_1 + \nu_2$) d.f. ในเมื่อ SS_1 เป็น SS ของ sample ที่ 1 with ν_1 d.f. และ SS_2 เป็น SS ของ sample ที่ 2 with ν_2 d.f. จาก Table 9.5 จะเห็นว่า $SS_1 = 14$ และ $SS_2 = 26$ จำนวนเลขที่เป็นเศษของ s_p^2 คือ (14 + 26) หรือ 40 และจำนวนเลขที่เป็นส่วนคือ (2 + 3) หรือ 5 ดังนั้น the pooled variance (s_p^2) จึงเท่ากับ $\frac{40}{5} = 8$ with (2 + 3) หรือ 5 d.f.

<u>Theorem 9.6:</u> If the statistic $\frac{SS_1}{\sigma^2}$ follows the X^2- distribution with ν_1 degrees of freedom and the statistic $\frac{SS_2}{\sigma^2}$ follows the X^2- distribution with ν_2 degrees of freedom and $\frac{SS_1}{\sigma^2}$ and $\frac{SS_2}{\sigma^2}$ are obtained from independent samples, the statistic

$$\frac{SS_1 + SS_2}{\sigma^2}$$

follows the X^2- distribution with ($\nu_1 + \nu_2$) degrees of freedom.

เราอาจอธิบาย theorem นี้ และแสดงให้เห็นจริงโดย sampling experiments ได้อีกด้วย ในข้อ 7.7 ได้แสดงให้เห็นแล้วว่าค่าของ $\frac{SS}{\sigma^2}$ ของ 1,000 samples follow the X^2- distribution with 4 d.f. จาก 1,000 samples นี้เราทำให้เป็น 500 คู่ของ samples ได้โดยวิธีให้ sample ที่หนึ่งกับ sample ที่สองเป็นคู่หนึ่ง sample ที่สามกับ sample ที่สี่เป็นคู่หนึ่ง และต่อ ๆ ไป ตามลำดับ ตัวอย่างค่าของ SS ของ 4 random samples ซึ่งแต่ละ sample ประกอบด้วย 5 observations ได้ให้ไว้ใน Table 4.2 แล้ว ค่าของ SS ทั้งสี่คือ 598.8, 237.2, 396.8 และ 319.2 ตาม

ลำดับ ค่าของ $\frac{(SS_1 + SS_2)}{\sigma^2}$ สำหรับคู่ที่หนึ่งของ samples คือ $\frac{(598.8 + 237.2)}{100}$ หรือ 8.360 และสำหรับคู่ที่สองของ samples คือ $\frac{(396.8 + 319.2)}{100}$ หรือ 7.160 เราจึงได้ค่าของ $\frac{(SS_1+SS_2)}{\sigma^2}$ ดังกล่าวจากแต่ละคู่ของ samples 500 คู่นี้ พึงสังเกตว่า 1,000 samples นี้เป็น independent samples (ข้อ 4.2) นี่เป็นสภาพสำคัญสำหรับความเป็นจริงของ Theorem 9.6

frequency distribution ของค่าของ $\frac{(SS_1 + SS_2)}{\sigma^2}$ 500 ค่าได้ให้ไว้ใน Table 9.6

Table 9.6

$\frac{SS_1 + SS_2}{\sigma^2}$	observed frequency		theoretical	midpoint	mf
	f	r.f. (%)	r.f. (%)	m	
0 — 1	1	0.2	0.2	.5	.5
1 — 2	6	1.2	1.7	1.5	9.0
2 — 3	29	5.8	4.7	2.5	72.5
3 — 4	32	6.4	7.7	3.5	112.0
4 — 5	54	10.8	10.0	4.5	243.0
5 — 6	54	10.8	11.0	5.5	297.0
6 — 7	56	11.2	11.1	6.5	364.0
7 — 8	58	11.6	10.3	7.5	435.0
8 — 9	52	10.4	9.1	8.5	442.0
9 — 10	30	6.0	7.7	9.5	285.0
10 — 11	29	5.8	6.3	10.5	304.5
11 — 12	31	6.2	5.1	11.5	356.5
12 — 13	19	3.8	3.9	12.5	237.5
13 — 14	12	2.4	3.0	13.5	162.0
14 — 15	13	2.6	2.3	14.5	188.5
15 — 16	5	1.0	1.7	15.5	77.5
16 — 17	7	1.4	1.2	16.5	115.5
17 — 18	3	.6	.9	17.5	52.5
18 — 19	2	.4	.6	18.5	37.0
19 — 20	3	.6	.5	19.5	58.5
over 20	4	.8	1.0	23.1	92.4
total	500	100.0	100.0		3,942.4

mean of $\frac{SS_1 + SS_2}{\sigma^2} = \frac{\Sigma mf}{\Sigma f} = \frac{3,942.4}{500} = 7.9$

theoretical frequency ที่ให้ไว้ใน table เป็น frequency ของ X^2- distribution with 8 d.f., histogram ของ 500 ค่าของ $\frac{SS_1+SS_2}{\sigma^2}$ ซึ่งเขียนทับลงบน X^2- curve with 8 d.f. ได้แสดงไว้ใน Fig. 9.6

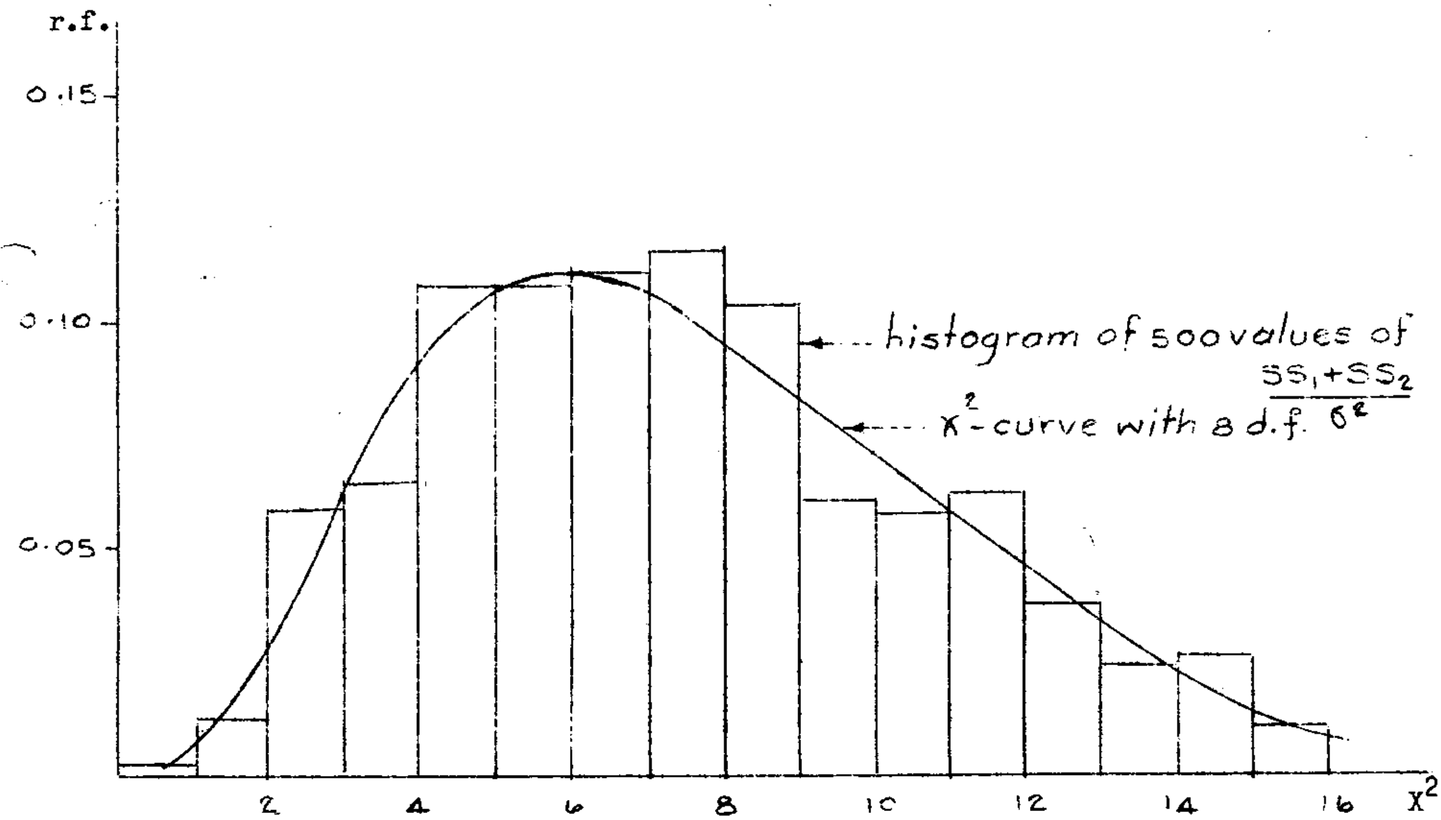


Fig. 9.6

ทั้ง table และกราฟแสดงให้เห็นว่า observed และ theoretical frequencies ใกล้เคียงกันมาก

เนื่องจาก mean ของ X^2- distribution คือ d.f. ของมัน เราจึงหวังได้ว่า mean ของ 500 ค่าของ $\frac{SS_1+SS_2}{\sigma^2}$ จะมีค่าใกล้เคียงกับ 8 mean โดยประมาณของค่า $\frac{SS_1+SS_2}{\sigma^2}$ เหล่านี้จะหาได้จาก Table 9.6 โดยใช้จุดกึ่งกลาง (m) ของ class หนึ่งแทนค่า $\frac{SS_1+SS_2}{\sigma^2}$ ทั้งหมดของ class นั้น เช่น ใช้ 0.5 แทนค่า $\frac{SS_1+SS_2}{\sigma^2}$ ทั้งหมดของ class 0 – 1 ใช้ 1.5 แทนค่า $\frac{SS_1+SS_2}{\sigma^2}$ ทั้งหมดของ class 1 – 2 และต่อ ๆ ไปตามลำดับ สำหรับ class มากกว่า 20 ไม่มีเขตจำกัด แต่เราใช้ 23.1 ซึ่งเป็นค่าเฉลี่ยของ $\frac{SS_1+SS_2}{\sigma^2}$ 4 ค่า ซึ่งตกอยู่ภายใน class นี้เป็นจุดกึ่งกลางของ class, mean โดย

ประมาณของ 500 ค่าของ $\frac{SS_1 + SS_2}{\sigma^2}$ คือ

$$\frac{\Sigma\, mf}{\Sigma\, f} = \frac{3,942.4}{500} = 7.9$$

ซึ่งใกล้เคียงกับ 8 มากดังที่หวังไว้ การแสดงให้เห็นจริงของ Theorem 9.6 จึงสมบูรณ์ และการกำหนดให้ $(\nu_1 + \nu_2)$ เป็น d.f. ของ the pooled variance (s_p^2) ในเมื่อ ν_1 และ ν_2 เป็น d.f. ของ s_1^2 และ s_2^2 ตามลำดับนั้นก็ถูกต้องแล้ว

ถ้า 2 populations มี variance (σ^2) อย่างเดียวกัน the pooled estimate of σ^2 จะหาได้จาก Equation (2) เท่านั้น แต่ถ้า 2 populations มี mean (μ) อย่างเดียวกันและมี variance (σ^2) อย่างเดียวกันด้วยแล้ว observations ของ 2 samples นั้นควรจะรวมเข้าด้วยกันเพื่อทำให้เป็น sample เดียวที่มี $n_1 + n_2$ observations และ s^2 ของ sample รวมที่มี $(n_1 + n_2 - 1)$ d.f. นี้จะเป็น pooled estimate of σ^2

the pooled estimate (s_p^2) ของ population variance (σ^2) จะหาได้จาก sample variances ที่ตัวก็ได้ การใช้วิธี weighted average ไม่จำกัดเฉพาะ 2 samples เท่านั้น กล่าวโดยทั่วไป ถ้ามี k populations ซึ่งมี variance (σ^2) อย่างเดียวกัน และเราได้ sample มาจาก populations เหล่านี้ population ละหนึ่ง sample, the pooled estimate of σ^2 ซึ่งขึ้นอยู่กับ k samples คือ

$$s_p^2 = \frac{\text{sum of k SS - values}}{\text{sum of k numbers of d.f.}} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

9.7 Relation Between F-Distribution and X^2-Distribution

ในข้อ 7.7 เราทราบแล้วว่า statistic $\frac{SS}{\sigma^2}$ follows the X^2-distribution with ν $(\nu = n - 1)$ d.f. และในข้อ 9.2 เราก็ทราบว่า statistic $\frac{s_1^2}{s_2^2}$ follows the F-distribution with ν_1 and ν_2 d.f. เนื่องจาก s^2 และ SS มีความสัมพันธ์กันคือ $s^2 = \frac{SS}{\nu}$ ดังนั้น F-distribution และ X^2-distribution จะต้องสัมพันธ์กันด้วย ความจริง 2 distributions นี้สัมพันธ์กันหลายทาง ความสัมพันธ์ที่นำมาสาธิตในข้อนี้มีความสำคัญในทางปฏิบัติมากที่สุด

$\frac{SS}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2}$

เนื่องจาก s^2 เป็นค่าประมาณของ σ^2 s^2 จะมีค่าเข้าไปใกล้ σ^2 ทุกทีในขณะที่ d.f. ของมันเข้าไปใกล้ infinity หรืออีกนัยหนึ่ง ถ้า sample หนึ่งมี observations มากมายนับไม่ถ้วน sample นั้นจะกลายเป็น population ไป และ s^2 จะกลายเป็น σ^2 เพราะฉะนั้น statistic $F = \frac{s_1^2}{s_2^2}$ จะกลายเป็น $\frac{s_1^2}{\sigma^2}$ ถ้า ν_2 เข้าไปใกล้ ∞ (infinity) แต่ statistic $\frac{s^2}{\sigma^2}$ นี้ follows the $\frac{X^2}{\nu}$- distribution ด้วย (ข้อ 7.7) ความสัมพันธ์ระหว่าง X^2 และ F จึงถูกแสดงให้เห็นได้ใน theorem ต่อไปนี้

<u>Theorem 9.7:</u> If a statistic X^2 follows the X^2- distribution with ν degrees of freedom, the statistic X^2/ν follows the F- distribution with ν and ∞ degrees of freedom.

ความสัมพันธ์ที่กล่าวใน Theorem 9.7 นี้จะสังเกตได้จาก Table 5 และ 7 ของ Appendix ตัวอย่างเช่น 5 % point ของ $\frac{X^2}{\nu}$ with 10 d.f. คือ 1.8307 ซึ่งเป็น 5 % point ของ F with 10 and ∞ d.f. ด้วย ความจริงทั้ง column ของ 5 % points ของ $\frac{X^2}{\nu}$ ที่มี d.f. ต่าง ๆ ก็คือบรรทัดล่างสุดของ 5 % F- table และ column 1 % points ของ $\frac{X^2}{\nu}$ คือบรรทัดล่างสุดของ 1 % F- table นั่นเอง ความสัมพันธ์นี้จะเป็นไปตาม percentage points ทั้งหมด และด้วยความสัมพันธ์อันนี้เอง<u>การทดสอบสมมติฐานว่า population variance เท่ากับค่าที่กำหนด (ข้อ 7.10) จึงอาจทำได้ต่าง ๆ กันถึง 3 วิธี คือ</u>

1. $X^2 = \frac{SS}{\sigma^2}$ with n - 1 d.f. หรือ
2. $\frac{X^2}{n-1} = \frac{s^2}{\sigma^2}$ with n - 1 d.f. หรือ
3. $F = \frac{s^2}{\sigma^2}$ with n - 1 and ∞ d.f.

statistic ตัวใดตัวหนึ่งใน 3 ตัวข้างบนนี้ใช้ทดสอบสมมติฐานดังกล่าวได้ทั้งสิ้น ที่จริงการทดสอบสมมติฐาน 3 วิธีซึ่งดูเหมือนจะเป็นวิธีที่แตกต่างกันนี้เป็นวิธีเดียวกันโดยแท้ และ <u>conclusions ที่ได้โดยวิธีเหล่านี้จะเป็นอย่างเดียวกันเสมอ</u>

9.8 Remarks

การทดสอบสมมติฐานว่า 2 population variances มีค่าเท่ากันอาจใช้ในการเปรียบเทียบความสม่ำเสมอของผลิตภัณฑ์อย่างหนึ่งที่ผลิตโดยกรรมวิธีสองอย่างซึ่งต่างกัน ตัวอย่างเช่น ในการผลิตอะไหล่เครื่องจักรซึ่งกำหนดความยาวไว้เท่ากับ 1.5 นิ้วนั้น ใช้กรรมวิธีในการผลิต 2 วิธี เนื่องจากจะต้องผลิตอะไหล่เครื่องจักรให้มีความยาวโดยเฉลี่ยเท่ากับ 1.5 นิ้วนี้เอง กรรมวิธีซึ่งสามารถผลิตอะไหล่ที่มีความยาวสม่ำเสมอได้มากกว่าจะเป็นกรรมวิธีที่ต้องการ เพราะถ้าอะไหล่ยาวไปหรือสั้นไปก็จะไม่พอดีกับเครื่องจักร กล่าวได้อีกอย่างหนึ่งว่า variance น้อยเป็นคุณภาพของอะไหล่เครื่องจักรที่ต้องการ การเปรียบเทียบ variances ของความยาวของอะไหล่เครื่องจักรที่ผลิตโดยกรรมวิธีทั้งสองจะต้องหา random sample ของอะไหล่ออกมากรรมวิธีละหนึ่ง sample และวัดความยาวของอะไหล่เครื่องจักรแต่ละชิ้นไว้แล้วทำตามวิธีที่ให้ไว้ในข้อ 9.5 ก็จะทดสอบสมมติฐานได้ว่า $\sigma_1^2 = \sigma_2^2$ ถ้าเรายอมรับสมมติฐาน conclusion คือ 2 กรรมวิธีสามารถผลิตอะไหล่เครื่องจักรที่มีขนาดสม่ำเสมอได้พอกัน ถ้าเราปฏิเสธสมมติฐาน conclusion คือกรรมวิธีที่ผลิตอะไหล่เครื่องจักรที่มี variance น้อยเป็นกรรมวิธีที่เราต้องการมากกว่า

ผู้ทำการทดลองค้นคว้าทางวิทยาศาสตร์อาจใช้การทดสอบที่คล้ายกันนี้ในการเปรียบเทียบความละเอียดของเทคนิคของการทดลองสองอย่าง เทคนิคซึ่งผลิต observations ที่มี variance น้อยเป็นเทคนิคที่มีความละเอียดมากกว่า

ความประสงค์สำคัญของการนำ F-distribution มาให้ทราบในขั้นนี้ก็เพื่อเตรียมไว้สำหรับเรื่องที่จะต้องศึกษาต่อไปข้างหน้า คือ analysis of variance (บทที่ 12) และไม่ใช่การทดสอบสมมติฐานว่า 2 population variances มีค่าเท่ากัน

F-table ไม่สะดวกที่จะใช้สำหรับการทดสอบแบบ two-tailed test เพราะว่า percentage points ของหางข้างซ้ายของ F-distribution ไม่ได้ทำตารางไว้ อย่างไรก็ดี F-table ไม่ได้ถูกทำขึ้นเพื่อความประสงค์นี้ โดยมากจะใช้เกี่ยวกับ analysis of variance ซึ่ง F-test ของมันเป็น one-tailed test เรื่องที่จะศึกษาต่อไปในหนังสือนี้จะไม่ใช้ two-tailed F-test เลย

Chapter 10

Difference Between Sample Means

บทที่ 5, 6 และ 8 ที่ได้กล่าวมาแล้วนั้นเป็นเรื่องเกี่ยวกับ deductive และ inductive relations ระหว่าง population หนึ่งกับ means ของ samples ต่าง ๆ ของมัน บทที่ 5 ซึ่งเกี่ยวกับ deductive relation ได้กล่าวถึงลักษณะของ sample means ที่ได้มาจาก population หนึ่ง และบทที่ 6 ซึ่งเกี่ยวกับ inductive relation ก็ได้กล่าวถึงการทดสอบสมมติฐานที่เกี่ยวกับ population โดยการใช้หนึ่ง sample และ u-test ส่วนบทที่ 8 ซึ่งเป็นบทสำคัญบทหนึ่งได้กล่าวซ้ำบทที่ 5 และ 6 อีกครั้งหนึ่ง แต่ใช้ sample variance (s^2) แทน population variance (σ^2) ใน statistic u และทำให้เราทราบเรื่อง t-test ในบทที่ 10 นี้จะกล่าวซ้ำทุกเรื่องที่ได้บรรยายมาแล้วในบทที่ 5, 6 และ 8 อีกครั้งหนึ่งโดยการพิจารณาจะไม่มุ่งไปยัง sample mean ตัวเดียว แต่จะมุ่งไปยังผลต่างระหว่าง 2 sample means

10.1 Distribution of Difference Between Sample Means

ปัญหาที่จะต้องพิจารณาในเรื่องนี้เป็นปัญหาที่เกี่ยวกับ 2 populations เพื่อไม่ให้เกิดความสับสนเราจะใช้เลข 1 และ 2 กำกับ notations ต่าง ๆ ได้แก่ μ , σ^2, $\bar{y}$, s^2, N และ n ตัวอย่างเช่น μ_1 คือ mean ของ population ที่หนึ่ง และ μ_2 คือ mean ของ population ที่สอง ปัญหาที่ต้องพิจารณาจะแสดงให้เห็นโดยตัวอย่างต่อไปนี้

ให้ population ที่หนึ่งประกอบด้วย 3 observations ($N_1 = 3$) คือ 2, 4 และ 6 ดังนั้น mean (μ_1) คือ 4 และ variance (ข้อ 2.3) คือ

$$\sigma_1^2 = \frac{(2-4)^2+(4-4)^2+(6-4)^2}{3} = \frac{8}{3} \quad \ldots\ldots\ldots\ldots (1)$$

ให้ population ที่สองประกอบด้วย 2 observations ($N_2 = 2$) คือ 3 และ 6 ดังนั้น mean (μ_2) คือ 4.5 และ variance คือ

$$\sigma_2^2 = \frac{(3-4.5)^2+(6-4.5)^2}{2} = 2.25$$

information ที่เกี่ยวกับ populations ทั้งสองอาจรวบรวมไว้ได้ดังนี้

population 1	population 2
2, 4, 6	3, 6
$N_1 = 3$	$N_2 = 2$
$\mu_1 = 4$	$\mu_2 = 4.5$
$\sigma_1^2 = \frac{8}{3}$	$\sigma_2^2 = 2.25$

จาก population ที่หนึ่งเราหา all possible samples ซึ่งมี size 2 ($n_1 = 2$) ออกมา (ข้อ 5.1) และแต่ละ sample เราคำนวณ sample mean ($\bar{y}_1$) ไว้ จาก population ที่สองเราหา all possible samples ซึ่งมี size 3 ($n_2 = 3$) ออกมาและแต่ละ sample เราคำนวณ sample mean ($\bar{y}_2$) ไว้ samples 2 ชุด และ means ของมันได้ให้ไว้ใน Table 10.1 a

Table 10.1 a

from population 1		from population 2	
samples	$\bar{y}_1$	samples	$\bar{y}_2$
2, 2	2	3, 3, 3	3
2, 4	3	3, 3, 6	4
2, 6	4	3, 6, 3	4
4, 2	3	3, 6, 6	5
4, 4	4	6, 3, 3	4
4, 6	5	6, 3, 6	5
6, 2	4	6, 6, 3	5
6, 4	5	6, 6, 6	6
6, 6	6		
sum	36	sum	36
mean	4	mean	4.5

frequency table ของ sample means 2 ชุดนี้ได้แสดงไว้ใน Table 10.1 b

Table 10.1 b

$\bar{y}_1$	f	$\bar{y}_2$	f
2	1	3	1
3	2	4	3
4	3	5	3
5	2	6	1
6	1		
total	9	total	8
$n_1 = 2$ $\mu_{\bar{y}_1} = \mu_1 = 4$ $\sigma^2_{\bar{y}_1} = \frac{\sigma^2_1}{n_1} = \frac{8/3}{2} = \frac{4}{3}$		$n_2 = 3$ $\mu_{\bar{y}_2} = \mu_2 = 4.5$ $\sigma^2_{\bar{y}_2} = \frac{\sigma^2_2}{n_2} = \frac{2.25}{3} = \frac{3}{4}$	

mean ของ 9 sample means ที่ได้มาจาก population ที่หนึ่งคือ 4 ซึ่งเท่ากับ μ_1 และ mean ของ 8 sample means ที่ได้มาจาก population ที่สองคือ 4.5 ซึ่งเท่ากับ μ_2 variance ของ sample means แต่ละชุดจะได้จาก Theorem 5.3 คือ

$$\sigma^2_{\bar{y}} = \frac{\sigma^2}{n} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

ดังนั้น variance ของ sample means ชุดที่หนึ่งคือ

$$\sigma^2_{\bar{y}_1} = \frac{\sigma^2_1}{n_1} = \frac{8/3}{2} = \frac{4}{3} \quad \ldots\ldots\ldots (3)$$

และ variance ของ sample means ชุดที่สองคือ

$$\sigma^2_{\bar{y}_2} = \frac{\sigma^2_2}{n_2} = \frac{2.25}{3} = 0.75 = \frac{3}{4} \quad \ldots\ldots\ldots (4)$$

ทั้งหมดที่กล่าวมานี้ไม่มีอะไรเป็นเรื่องใหม่เลย ทั้ง sampling scheme และลักษณะของ sample means ได้ให้ไว้แล้วในข้อ 5.1 และ 5.2 แต่ขณะนี้ได้มีปัญหาใหม่เกิดขึ้นอย่างหนึ่งคือ <u>เราจะพิจารณาลักษณะของผลต่างระหว่าง 2 sample means ($\bar{y}_1 - \bar{y}_2$) นี้ได้อย่างไร?</u> จาก population ที่หนึ่งเรามี 9 sample means และจาก population ที่สองเรามี 8 sample means ถ้า sample mean แต่ละตัวในชุดที่หนึ่งถูกเปรียบเทียบกับ sample mean ทุกตัวในชุดที่สองแล้ว จะมีผลต่างระหว่าง 2 sample means รวมทั้งหมด 9 x 8 หรือ 72 ผลต่าง ผลต่าง ($\bar{y}_1 - \bar{y}_2$), 72 จำนวนนี้ได้ให้ไว้ใน Table 10.1 c แล้ว

Table 10.1 c

$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_1 - \bar{y}_2$	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_1 - \bar{y}_2$
2	3	-1	5	3	2
2	4	-2	5	4	1
2	4	-2	5	4	1
2	5	-3	5	5	0
2	4	-2	5	4	1
2	5	-3	5	5	0
2	5	-3	5	5	0
2	6	-4	5	6	-1
3	3	0	4	3	1
3	4	-1	4	4	0
3	4	-1	4	4	0
3	5	-2	4	5	-1
3	4	-1	4	4	0
3	5	-2	4	5	-1
3	5	-2	4	5	-1
3	6	-3	4	6	-2
4	3	1	5	3	2
4	4	0	5	4	1
4	4	0	5	4	1
4	5	-1	5	5	0
4	4	0	5	4	1
4	5	-1	5	5	0
4	5	-1	5	5	0
4	6	-2	5	6	-1
3	3	0	6	3	3
3	4	-1	6	4	2
3	4	-1	6	4	2
3	5	-2	6	5	1
3	4	-1	6	4	2
3	5	-2	6	5	1
3	5	-2	6	5	1
3	6	-3	6	6	0
4	3	1			
4	4	0			
4	4	0			
4	5	-1			
4	4	0			
4	5	-1			
4	5	-1			
4	6	-2			

frequency table ของผลต่างเหล่านี้ได้ให้ไว้ใน Table 10.1 d

Table 10.1 d

$(\bar{y}_1 - \bar{y}_2)$	f	$(\bar{y}_1-\bar{y}_2)f$	$(\bar{y}_1-\bar{y}_2+0.5)$	$(\bar{y}_1-\bar{y}_2+0.5)^2$	$(\bar{y}_1-\bar{y}_2+0.5)^2f$
-4	1	-4	-3.5	12.25	12.25
-3	5	-15	-2.5	6.25	31.25
-2	12	-24	-1.5	2.25	27.00
-1	18	-18	-0.5	0.25	4.50
0	18	0	0.5	0.25	4.50
1	12	12	1.5	2.25	27.00
2	5	10	2.5	6.25	31.25
3	1	3	3.5	12.25	12.25
Total	72	-36			150.00

$$\mu_{\bar{y}_1-\bar{y}_2} = \frac{-36}{72} = -0.5 = 4 - 4.5 = \mu_1 - \mu_2$$

$$\sigma^2_{\bar{y}_1-\bar{y}_2} = \frac{150}{72} = \frac{25}{12} = \frac{4}{3} + \frac{3}{4} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

ในการอ่าน 2 tables นี้ ต้องพิจารณาผลต่าง $(\bar{y}_1 - \bar{y}_2)$ เป็นเลขหนึ่งจำนวน ดังนั้น mean และ variance ของ 72 ผลต่างจะคำนวณได้ รายละเอียดของการคำนวณได้ให้ไว้ใน Table 10.1 d ข้างบนนี้แล้ว mean ของ 72 ผลต่าง $(\bar{y}_1 - \bar{y}_2)$ คือ

$$\mu_{\bar{y}_1-\bar{y}_2} = \frac{\sum(\bar{y}_1 - \bar{y}_2)f}{72} = \frac{-36}{72} = -0.5 \quad \ldots\ldots\ldots\ (5)$$

ซึ่ง −0.5 นี้ คือผลต่างระหว่าง 2 population means $(\mu_1 - \mu_2)$ หรือ 4 − 4.5 นั่นเอง ความสัมพันธ์ระหว่าง mean ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ และ population means ทั้งสอง (μ_1 และ μ_2) จึงอาจแสดงให้เห็นได้ใน theorem ต่อไปนี้

<u>Theorem 10.1 a:</u> The mean of the differences between two sample means is equal to the difference between the means of the two populations from which the two sets of samples are drawn, that is,

$$\boxed{\mu_{\bar{y}_1-\bar{y}_2} = \mu_1 - \mu_2} \quad \ldots\ldots\ldots\ldots\ldots \ (6)$$

จากตัวอย่างที่แล้ว variance ของ 72 ผลต่างระหว่าง sample means $(\bar{y}_1 - \bar{y}_2)$ คือ

$$\sigma^2_{\bar{y}_1-\bar{y}_2} = \frac{\sum\left[(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)\right]^2 f}{72}$$

$$= \frac{\sum\left[(\bar{y}_1 - \bar{y}_2) - (-0.5)\right]^2 f}{72}$$

$$= \frac{\sum(\bar{y}_1 - \bar{y}_2 + 0.5)^2 f}{72}$$

$$= \frac{150}{72}$$

$$= \frac{25}{12}$$

แต่ผลบวกของ variances ของ sample means 2 ชุดก็เท่ากับ $\frac{25}{12}$ ด้วย คือ

$$\sigma^2_{\bar{y}_1} = \frac{\sigma^2_1}{n_1}$$

$$\sigma^2_{\bar{y}_2} = \frac{\sigma^2_2}{n_2}$$

$$\sigma^2_{\bar{y}_1} + \sigma^2_{\bar{y}_2} = \frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}$$

$$= \frac{8/3}{2} + \frac{2.25}{3}$$

$$= \frac{4}{3} + \frac{3}{4}$$

$$= \frac{25}{12}$$

ควรเน้นไว้ ณ ที่นี้ว่า ความสัมพันธ์ระหว่าง variance ของผลต่างระหว่าง sample means ($\sigma^2_{\bar{y}_1 - \bar{y}_2}$) และ variances ของ sample means ทั้งสอง ($\sigma^2_{\bar{y}_1}$ และ $\sigma^2_{\bar{y}_2}$) เป็นผลที่เกิดขึ้นโดยตรงของ sample scheme ซึ่งบ่งชัดลงไปว่า all possible samples ได้มาจากแต่ละ population ของ 2 populations และชุดของผลต่าง ($\bar{y}_1 - \bar{y}_2$) ประกอบด้วยผลต่างทั้งหมด (all possible differences) เมื่อเราใช้ sample scheme นี้ จะพูดได้ว่า sample means $\bar{y}_1$ และ $\bar{y}_2$ ถูกแจกแจงอย่างอิสระ (independently distributed) คำ "independence" นี้ยอมหมายถึงความจริงที่ว่า sample mean ทุกตัวของชุดที่หนึ่งมีโอกาสเท่ากันที่จะถูกเปรียบเทียบกับ sample mean ทุกตัวของชุดที่สอง

การพิจารณาเรื่อง variance ของผลต่างระหว่าง sample means อาจสรุปไว้ได้ใน theorem ต่อไปนี้

<u>Theorem 10.1 b:</u> The variance of the differences between two sets of independent sample means is equal to the sum of the variances of the two respective sets of sample means, that is,

$$\sigma^2_{\bar{y}_1 - \bar{y}_2} = \sigma^2_{\bar{y}_1} + \sigma^2_{\bar{y}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}, \quad \ldots\ldots\ldots\ldots (7)$$

and consequently, the standard error of the difference between two independent sample means is

$$\sigma_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

histogram ของ 2 populations และของ distribution ของผลต่างระหว่าง sample means ได้ให้ไว้ใน Fig. 10.1

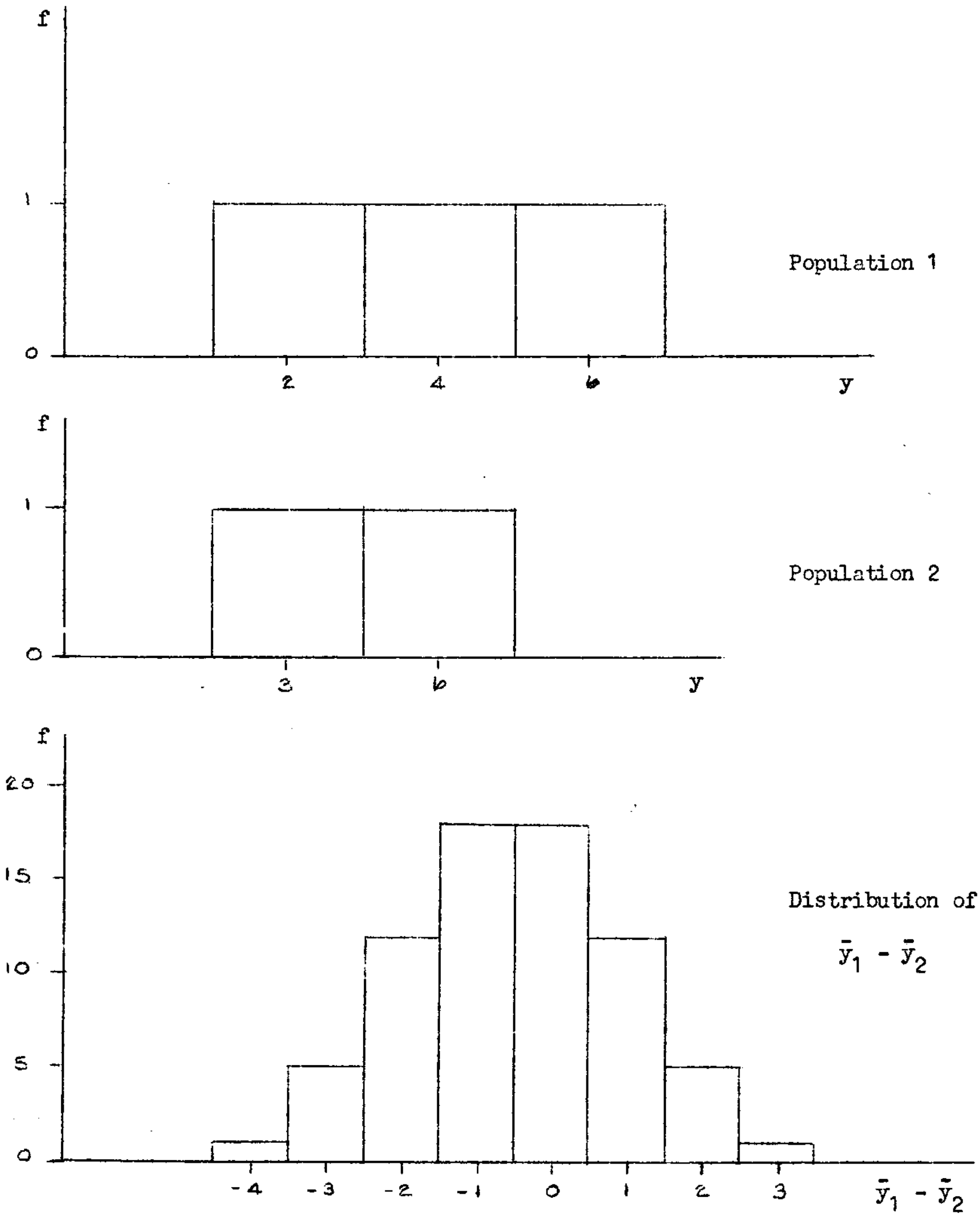

f
1
0
2
4
6
y
Population 1
f
1
0
3
6
y
Population 2
f
20
15
10
5
0
-4
-3
-2
-1
0
1
2
3
$\bar{y}_1 - \bar{y}_2$
Distribution of
$\bar{y}_1 - \bar{y}_2$

Fig. 10.1

เราจะสังเกตได้ว่า ผลต่าง ($\bar{y}_1 - \bar{y}_2$) แสดงแนวโน้มที่จะเป็นไปตาม normal curve แม้ว่า population จะไม่เป็น normal และ sample size จะไม่ใหญ่ก็ตาม ความจริงนี้จะได้แสดงไว้ใน theorem ต่อไปนี้

<u>Theorem 10.1 c:</u> The distribution of the differences $(\bar{y}_1 - \bar{y}_2)$ between sample means approaches the normal distribution as the sample sizes n_1 and n_2 increase, if the two populations from which the samples are drawn have finite variances. If the two populations are normal at the outset, the differences $(\bar{y}_1 - \bar{y}_2)$ follow the normal distribution exactly regardless of the sample sizes.

Theorems 10.1 a, b, c ได้ถูกแสดงให้เห็นจริงแล้วสำหรับ 2 populations ซึ่งประกอบด้วย observations 2, 4, 6 และ 3, 6 ตามลำดับ theorems เหล่านี้ จะถูกแสดงให้เห็นจริงอีกครั้งหนึ่งในข้อต่อไปโดย sampling experiment

10.2 Experimental Verification of Distribution of Difference Between Sample Means

Theorems 10.1 a, b และ c อาจถูกแสดงให้เห็นจริงได้โดย sampling experiment รายละเอียดของ experiment ได้บรรยายไว้แล้วในบทที่ 4 กล่าวโดยย่อเรามี 1,000 random samples แต่ละ sample ประกอบด้วย 5 observations ซึ่งได้มาจากnormal population หนึ่งที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 ในแต่ละ sample เราคำนวณ mean ($\bar{y}$) เอาไว้ ตัวอย่างของ means ของ 4 samples ดังกล่าวนี้ได้ให้ไว้ใน Table 4.2 ต่อไปเราจับคู่ sample means เข้าด้วยกันเพื่อให้ได้ผลต่างของมัน sample means ตัวที่หนึ่งและตัวที่สองจับคู่กันเป็นคู่หนึ่ง sample means ตัวที่สามและตัวที่สี่จับคู่กันเป็นคู่หนึ่ง และต่อ ๆ ไปตามลำดับ ดังนั้นจาก 1,000 sample means ที่เรามีจะได้ sample means 500 คู่ และเราคำนวณผลต่างระหว่าง sample means ($\bar{y}_1 - \bar{y}_2$) ของแต่ละคู่เอาไว้ ณ บัดนี้เรามี 500 ผลต่างระหว่าง sample means แล้ว ตัวอย่าง means ของ 4 samples ที่ให้ไว้ใน Table 4.2 คือ 48.8, 45.6, 59.2 และ 55.4 ดังนั้น 2 ผลต่างของ sample means 2 คู่นั้นคือ

$$48.8 - 45.6 = 3.2 \quad \text{และ}$$

$$59.2 - 55.4 = 3.8$$

ตามลำดับ ผลต่าง 3.2 และ 3.8 ทั้งสองนี้มีค่าบวก แต่ถ้า mean ตัวที่หนึ่งของ sample means คู่หนึ่งมีค่าน้อยกว่า mean ตัวที่สองแล้ว ผลต่างของมันจะมีค่าลบ

1,000 samples ที่เรามีนี้เป็น independent samples (ข้อ 4.2) sample ที่หนึ่งของแต่ละคู่ของ samples จะถูกมองในแง่เป็น sample ที่ได้มาจาก population ที่หนึ่ง และ sample ที่สองได้มาจาก population ที่สอง เนื่องจาก 2 populations ของ sampling experiment นี้ความจริงเป็น population เดียวกัน เพราะฉะนั้น mean ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ ควรจะเท่ากับ $\mu_1 - \mu_2$ $= 50 - 50 = 0$ (Theorem 10.1 a) และ variance ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ ก็ควรจะเท่ากับ (Theorem 10.1 b)

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{100}{5} + \frac{100}{5} = 40$$

เนื่องจาก populations เป็น normal, distribution ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ จะ follow the normal distribution (Theorem 10.1 c), frequency table ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ ทั้ง 500 ได้ให้ไว้ใน Table 10.2

Table 10.2

$\bar{y}_1 - \bar{y}_2$	f	c.f.	r.c.f. (%)
below - 10.5	24	24	4.8
- 10.5 to - 7.5	33	57	11.4
- 7.5 to - 4.5	58	115	23.0
- 4.5 to - 1.5	85	200	40.0
- 1.5 to 1.5	83	283	56.6
1.5 to 4.5	95	378	75.6
4.5 to 7.5	63	441	88.2
7.5 to 10.5	33	474	94.8
above 10.5	26	500	100.0
total	500		

เราพล็อต relative cumulative frequency (r.c.f.) ลงใน probability graph paper ได้ดัง Fig. 10.2

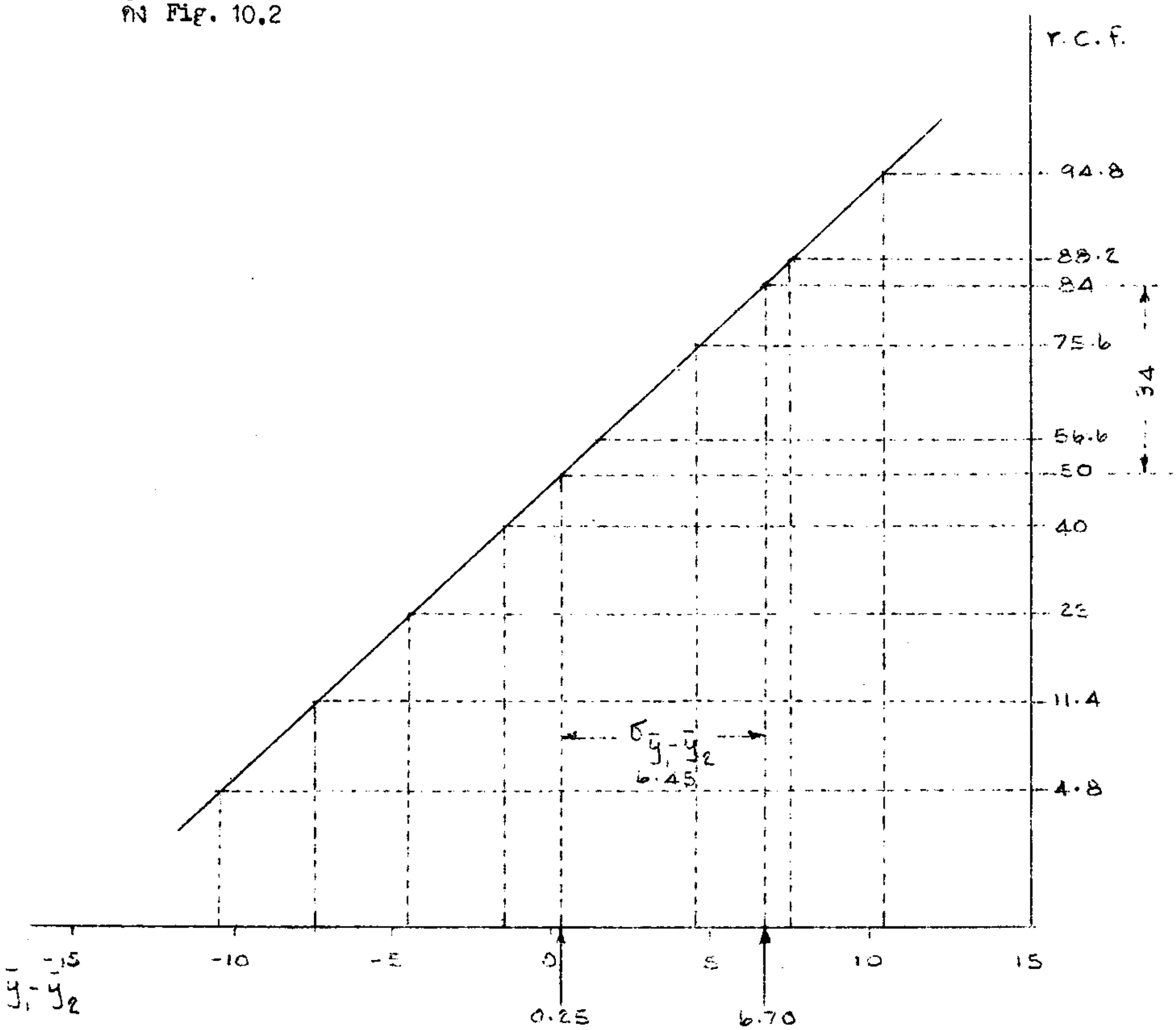


Fig. 10.2

ความเป็นจริงที่จุดต่าง ๆ เกือบอยู่บนเส้นตรงเดียวกันแสดงว่า ผลต่าง ($\bar{y}_1 - \bar{y}_2$) follow the normal distribution (ข้อ 3.3) mean ซึ่งเป็น 50 % point ของ distribution ตามที่อ่านได้จากกราฟคือ 0.25 ซึ่งเกือบเท่ากับศูนย์ดังที่หวังไว้ mean บวกหนึ่ง standard error ซึ่งเป็น 84 % point ตามที่อ่านได้จากกราฟคือ 6.70 ดังนั้น standard error ($\sigma_{\bar{y}_1 - \bar{y}_2}$) คือ 6.70 – 0.25 หรือ

6.45 ซึ่งใกล้เคียงกับ $\sqrt{40}$ หรือ 6.3245 ความเป็นจริงของ Theorems 10.1 a, b และ c ที่ให้ไว้ในข้อ 10.1 จึงสมบูรณ์แล้ว

10.3 u - Distribution

Theorems 10.1 a, b, c แสดงว่าผลต่าง $(\bar{y}_1 - \bar{y}_2)$ **follow the normal distribution** ที่มี **mean** เท่ากับ $(\mu_1 - \mu_2)$ และ **variance** เท่ากับ $(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$ ดังนั้น **statistic**

$$u = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

follows the normal distribution ที่มี **mean** เท่ากับ 0 และ **variance** เท่ากับ 1

ในข้อ 6.7 เราทราบแล้วว่า

$$u = \frac{\bar{y} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

และในข้อ 3.2 เราก็ทราบว่า

$$u = \frac{y - \mu}{\sigma} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(3)}$$

u ทั้ง 3 ตัวที่ให้ไว้ข้างบนนี้ไม่ใช่จำนวนเดียวกัน แต่มัน **follow distribution** เดียวกัน คือ **normal distribution** ที่มี **mean** เท่ากับ 0 และ **variance** เท่ากับ 1 ด้วยเหตุผลนี้เองเราจึงใช้อักษร u อย่างเดียวกัน

statistic แต่ละตัวคือ $(\bar{y}_1 - \bar{y}_2)$, $\bar{y}$ และ y ใน Equations (1), (2) และ (3) **follows the normal distribution** ในแต่ละกรณี ค่า u จะได้จากการเอา **mean** ของ **statistic** ไปลบตัว **statistic** แล้วหารผลต่างที่ได้ โดย **standard deviation** ของ **statistic** นั้น ด้วยความคล้ายคลึงกันหลายอย่างในระหว่าง u ทั้ง 3 ตัวนี้ ทำให้เราคิดไปได้ว่ามันจะต้องสัมพันธ์กันแน่

Equation (3) นั้นอาจถือเป็นกรณีพิเศษของ Equation (2) และ Equation (2) ก็เป็นกรณีพิเศษของ Equation (1) กล่าวคือ เมื่อ n_2 เข้าไปใกล้ infinity, $\bar{y}_2$ ก็จะใกล้ μ_2 และ $\frac{\sigma_2^2}{n_2}$ ก็จะใกล้ศูนย์ ดังนั้น Equation (1) จะกลายเป็น Equation (2) และเมื่อ n = 1, $\bar{y}$ จะกลายเป็น y ตัวของมันเองแล้ว Equation (2) ก็จะกลายเป็น Equation (3) จึงสรุปได้ว่า Equation (3) เป็นกรณีพิเศษของ Equation (2) ในเมื่อ n = 1 และ Equation (2) เป็นกรณีพิเศษของ Equation (1) ในเมื่อ $n_2 = \infty$

ถ้าเราทราบ 2 population variances (σ_1^2 และ σ_2^2) แล้ว เราอาจใช้ statistic u ในการทดสอบสมมติฐานว่าผลต่างระหว่าง 2 population means ($\mu_1 - \mu_2$) เท่ากับค่าที่กำหนด เนื่องจากเราไม่ค่อยจะทราบค่าของ population variances ทั้งสอง จึงต้องใช้ sample variances แทน population variances การใช้แทนกันนี้จะนำเราไปสู่ t-distribution (ข้อ 8.1)

10.4 Student's t-Distribution

statistic u ที่ให้ไว้ใน Equation (1) ข้อ 10.3 คือ

$$u = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

นั้นมีการใช้จำกัดมาก เพราะว่าโดยปกติแล้วเราไม่ทราบค่าของ 2 population variances แต่เมื่อเราใช้ sample variances แทน population variances u ก็กลายเป็น t ไป (ข้อ 8.1) t นี้จะมีได้ 2 ลักษณะขึ้นอยู่กับว่า population variances ทั้งสองจะเท่ากันหรือไม่เท่ากัน แต่ในที่นี้เราจะพิจารณาเฉพาะในกรณี population variances ทั้งสองเท่ากันเท่านั้น

ถ้า population variances ทั้งสองเท่ากัน คือ $\sigma_1^2 = \sigma_2^2$ แล้ว เราอาจใช้ σ^2 แทน variances ทั้งสองนี้ได้ จึงไม่ต้องใช้เลข 1 หรือ 2 กำกับอีกต่อไป ดังนั้น statistic u ใน Equation ข้างบนจะกลายเป็น

$$u = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2(\frac{1}{n_1} + \frac{1}{n_2})}} \quad \text{............ (1)}$$

เราอาจประมาณค่า σ^2 ได้โดย pooled variance (s_p^2) (ข้อ 9.6) ถ้า 2 samples มี n_1 และ n_2 observations ตามลำดับ s_1^2 จะมี $n_1 - 1$ d.f. และ s_2^2 จะมี $n_2 - 1$ d.f. ดังนั้น s_p^2 จะมี $(n_1 - 1) + (n_2 - 1)$ หรือ $(n_1 + n_2 - 2)$ d.f. เมื่อเราใช้ s_p^2 แทน σ^2 ใน Equation (1) statistic ที่ได้จะ follow Student's t-distribution with $n_1 + n_2 - 2$ d.f. พึงสังเกตว่า d.f. ของ t ต้องเป็น d.f. ของ s^2 ที่เข้ามาเกี่ยวข้องกับ t เสมอ (ข้อ 8.1)

ตามที่ได้พิจารณามาแล้วเราอาจสรุปเป็น theorem ได้ดังนี้.-

Theorem 10.4 a: If two given populations are normal and have the same variance, the statistic

$$t = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \quad \text{............. (2)}$$

$s_p^2 = \frac{SS_1 + SS_2}{\nu_1 + \nu_2}$

follows the Student's t-distribution with $n_1 + n_2 - 2$ degrees of freedom, where s_p^2 is the pooled estimate of the common variance of the two populations and $\bar{y}_1$ and $\bar{y}_2$ are the means of two independent samples (เราจะแสดงให้เห็นความจริงของ theorem นี้ โดย sampling experiment ในข้อ 10.5)

10.5 Experimental Verification of t-Distribution

รายละเอียดของ experiment ได้ให้ไว้แล้วในบทที่ 4 กล่าวโดยย่อเรามี 1,000 samples แต่ละ sample ประกอบด้วย 5 observations ซึ่งได้มาจาก tag population หนึ่งซึ่งเป็น normal ที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 ตัวอย่างของ samples ดังกล่าวนี้ได้ให้ไว้ใน Table 4.2 จากแต่ละ sample เราคำนวณ $\bar{y}$ และ SS ของมันไว้ ดังนั้นจาก 1,000 samples นี้ เราทำเป็น samples 500 คู่ได้โดยเอา samples ที่ 1 และที่ 2 ประกอบกันเป็นคู่หนึ่ง samples ที่ 3

และที่ 4 เป็นคู่หนึ่ง และต่อ ๆ ไปตามลำดับ pooled variance (s_p^2) ของ 2 samples ก็คือผลบวกของ 2 SS-values หารด้วย $[(n_1 - 1) + (n_2 - 1)]$ นั่นคือ

$$s_p^2 = \frac{SS_1 + SS_2}{(n_1 - 1) + (n_2 - 1)}$$

(ข้อ 9.6) ดังนั้น สำหรับ samples คู่ที่ 1 ของ Table 4.2

$$s_p^2 = \frac{598.8 + 237.2}{8}$$

$$= \frac{836.0}{8}$$

$$= 104.5$$

ค่าของ t สำหรับ samples คู่ที่ 1 นี้คือ

$$t = \frac{(\bar{y}_1 - \bar{y}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(48.8 - 45.5) - (50 - 50)}{\sqrt{104.5\left(\frac{1}{5} + \frac{1}{5}\right)}}$$

with 8 d.f. เราคำนวณค่าของ t สำหรับทุก ๆ คู่ของ sample 500 คู่ เอาไว้ frequency table ของค่าของ t 500 ค่าได้ให้ไว้ใน Table 10.5

Table 10.5

t	observed frequency		theoretical	mid-pt.	mf
	f	r.f.(%)	r.f. (%)	m	
below -4.5	0	0.0	0.1	-5	0
-4.5 to -3.5	1	0.2	0.3	-4	-4
-3.5 to -2.5	3	0.6	1.4	-3	-9
-2.5 to -1.5	37	7.4	6.8	-2	-74
-1.5 to -0.5	112	22.4	23.0	-1	-112
-0.5 to 0.5	178	35.6	36.8	0	0
0.5 to 1.5	124	24.8	23.0	1	124
1.5 to 2.5	37	7.4	6.8	2	74
2.5 to 3.5	5	1.0	1.4	3	15
3.5 to 4.5	2	0.4	0.3	4	8
over 4.5	1	0.2	0.1	5	5
total	500	100.0	100.0		27

mean of $t = \frac{\Sigma mf}{\Sigma f} = \frac{27}{500} = 0.054$

theoretical frequency ของ Student's t-distribution with 8 d.f. ได้ให้ไว้ใน Table นี้ด้วย histogram ของค่าของ t 500 ค่าซึ่งเขียนทับลงบน t-curve ได้แสดงไว้ใน Fig. 10.5

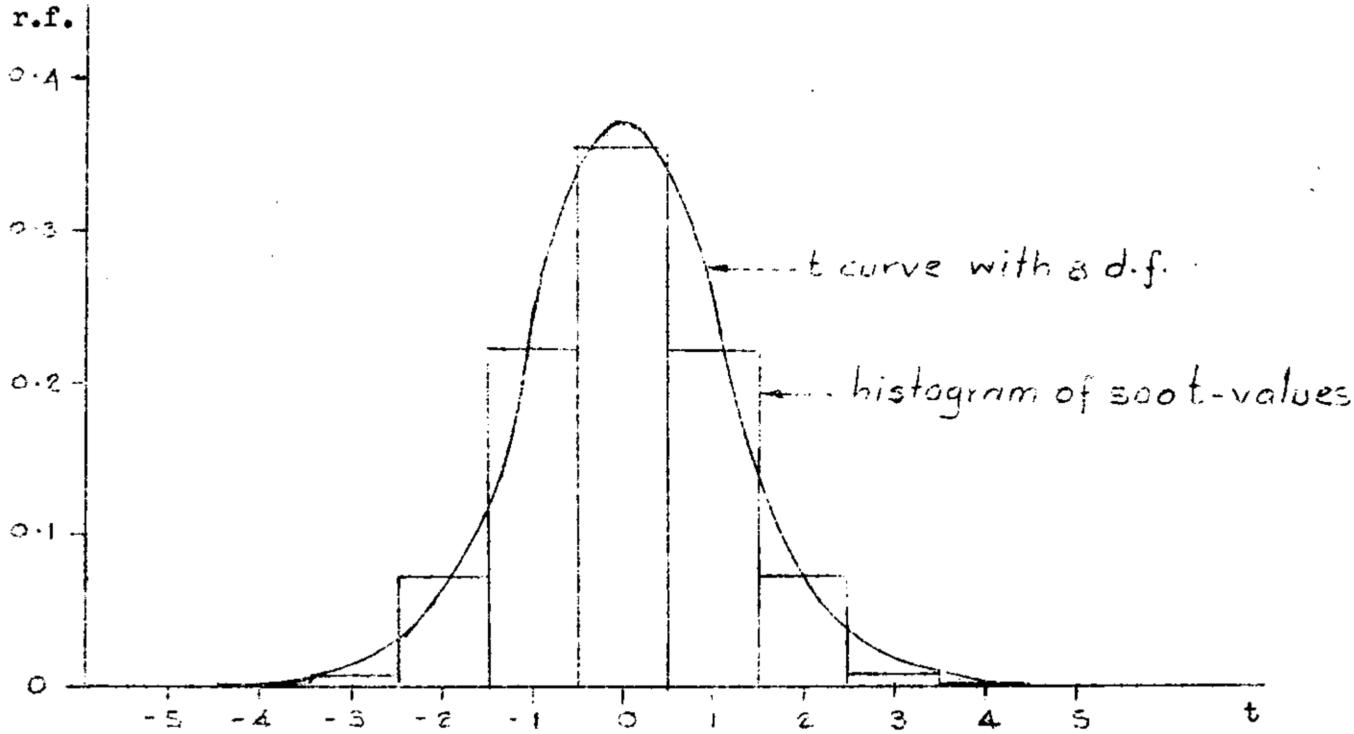


Fig. 10.5

ทั้ง frequency table และ histogram แสดงให้เห็นว่า observed และ theoretical frequencies นั้นใกล้เคียงกัน

ค่าของ mean โดยประมาณของ t 500 ค่าได้มาโดยการพิจารณาว่าค่าของ t ทั้งหมดใน class หนึ่งเท่ากับ midpoint (m) ของ class นั้น ดังนั้น mean ของ t คือ

$$\frac{\Sigma\, mf}{\Sigma\, f} = \frac{27}{500} = 0.054$$

ซึ่งใกล้เคียงกับศูนย์ดังที่หวังไว้ ความเป็นจริงของ Theorem 10.4 a จึงสมบูรณ์แล้ว

10.6 Test of Hypothesis - Procedures

5 ข้อแรกของบทนี้เกี่ยวกับ deductive ralation ระหว่าง 2 populations และ samples ต่าง ๆ ของมัน เราได้ t-distribution จาก all possible samples ซึ่งมี sizes ที่กำหนดให้ ซึ่งได้มาจาก 2 populations ณ บัดนี้ เราจะพิจารณาการ draw inductive inference ที่เกี่ยวกับ 2 population means จาก 2 random samples

Theorem 10.4 a อาจใช้ในการทดสอบสมมติฐานว่า ผลต่างระหว่าง 2 population means เท่ากับค่าที่กำหนด และแทบทั้งหมดเป็นการทดสอบสมมติฐานว่า ผลต่างระหว่าง 2 population means เท่ากับศูนย์ นั่นคือ $\mu_1 - \mu_2 = 0$ หรือ $\mu_1 = \mu_2$ การใช้ Student's t-distribution ในการทดสอบสมมติฐานได้พิจารณากันมาแล้วในรายละเอียดในบทที่ 8

ในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันนั้น statistic t (Theorem 10.4 a) จะกลายรูปเป็น

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

with $(n_1 + n_2 - 2)$ d.f. ถ้าค่าของ $t = 0$ โดยประมาณแสดงว่า $\mu_1 = \mu_2$ ถ้าค่าของ t โตกว่า 0 มาก แสดงว่า μ_1 มากกว่า μ_2 และถ้าค่าของ t เล็กกว่า 1 มาก ก็แสดงว่า μ_1 น้อยกว่า μ_2 ค่าของ t โตเท่าไรจึงจะเรียกว่าโตมากและเล็กเท่าไรจึงจะเรียกว่าเล็กมาก จะถูกพิจารณาโดย significance level

วิธีดำเนินการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันนั้นจะแสดงให้เห็นโดยตัวอย่างต่อไปนี้ ให้ sample หนึ่งประกอบด้วย observations 2, 7, 3 และอีก sample หนึ่งประกอบด้วย observations 9, 7 ปัญหาที่พิจารณาคือ 2 population means มีค่าเท่ากันหรือไม่ วิธีการทดสอบสมมติฐานทำดังนี้

1. Hypothesis : 2 population means มีค่าเท่ากัน นั่นคือ

$$\mu_1 = \mu_2$$

2. Alternative hypotheses : alternative hypotheses คือ

ก. $\mu_1 < \mu_2$ และ

ข. $\mu_1 > \mu_2$

3. Assumptions : 2 samples เป็น random samples ได้มาจาก normal populations ทั้งสองของมัน และ 2 populations นั้นมี variances อย่างเดียวกัน

4. Level of significance : เลือกใช้ 5 % significance level

5. Critical regions : critical regions สำหรับ t with 3 (คือ $n_1 + n_2 - 2$) d.f. อยู่ที่

$$t < -3.1825 \text{ และ}$$

$$t > 3.1825$$

6. Computation of t : รายละเอียดของการคำนวณค่า t ได้ให้ไว้ใน Table 10.6

Table 10.6

	sample No. 1	sample No. 2	combination	explanation
observations	2 7 3	9 7		
Σy	12	16		
n	3	2		
$\bar{y}$	4	8	-4	$\bar{y}_1 - \bar{y}_2$
$(\Sigma y)^2$	144	256		
$(\Sigma y)^2/n$	48	128		
Σy^2	62	130		
SS	14	2	16	pooled SS
d.f.	2	1	3	pooled d.f.
s^2			5.3333	$s_p^2 = \frac{16}{3}$
$\frac{1}{n}$	0.3333	0.5000	0.8333	$\frac{1}{n_1} + \frac{1}{n_2}$
			4.444	$s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$
			2.108	$\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$
			-1.90	t

ค่าของ t คือ -1.90 with 3 d.f.

7. Conclusion : เนื่องจาก t อยู่ภายนอก critical regions conclusion คือ 2 population means มีค่าเท่ากัน (ถ้า t อยู่ภายใน critical region ข้างซ้าย conclusion จะเป็น $\mu_1 < \mu_2$ ถ้า t อยู่ภายใน critical region ข้างขวา conclusion จะเป็น $\mu_1 > \mu_2$ ให้สังเกตว่า t เป็นเลขไม่มีหน่วย และจะไม่ถูกอิทธิพลโดยการเปลี่ยนหน่วยของการวัดเลย)

10.7 Advantages of Equal Sample Size

ในการกำหนดแนวการทดลองนั้นเราควรพยายามทำให้ sample size เท่ากัน ข้อดีประการหนึ่งดังที่ได้กล่าวมาแล้วในข้อ 10.4 คือ การลดลงของผลที่เกิดจากความไม่เท่ากันของ 2 population variances ใน t-test ถ้า sample sizes เท่ากันเราอาจใช้ Theorem 10.4 a ในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันแม้ว่า 2 population variances จะไม่เท่ากันก็ตาม

เหตุผลอีกประการหนึ่งของการทำให้ sample size เท่ากันมีความสำคัญมากกว่าที่ได้กล่าวมาแล้วข้างต้น variance ของผลต่างระหว่างชุดของ sample means 2 ชุด (Theorem 10.1 b) คือ

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

ถ้า $\sigma_1^2 = \sigma_2^2 = \sigma^2$, variance นี้จะกลายเป็น

$$\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

สำหรับจำนวน observationsทั้งหมดที่กำหนดใน 2 samples นั้น ค่าของ

$$\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

จะน้อยที่สุดเมื่อ $n_1 = n_2$ ตัวอย่างเช่น จำนวน observations ทั้งหมดสำหรับ 2 samples เท่ากับ 100 นั่นคือ $n_1 + n_2 = 100$ sample sizes ทั้งสองคือ n_1 และ n_2 อาจเป็น 1 และ 99, 2 และ 98, ..., 50 และ 50 สำหรับกรณี $n_1 = 1$, $n_2 = 99$, ค่าของ variance คือ

$$\sigma^2 \left(\frac{1}{1} + \frac{1}{99} \right) = 1.0101 \ \sigma^2$$

สำหรับกรณี $n_1 = n_2 = 50$, ค่าของ variance คือ

$$\sigma^2 \left(\frac{1}{50} + \frac{1}{50} \right) = 0.0400 \ \sigma^2$$

ในกรณี $n_1 = 1$ และ $n_2 = 99$ variance จะโตกว่า 25 เท่าในกรณี $n_1 = n_2 = 50$ ซึ่งแม้ว่าจำนวน observations ทั้งหมดของทั้งสองกรณีจะเท่ากับ 100 ก็ตาม variance ที่โตกว่าย่อมแสดงว่าผลต่าง $(\bar{y}_1 - \bar{y}_2)$ มีค่าสูง ๆ ต่ำ ๆ จากคู่หนึ่งของ samples ไปยังคู่อื่น ๆ ของ samples ได้มากกว่า ผลที่เกิดขึ้นก็คือในกรณี sample sizes ไม่เท่ากันนั้นความน่าจะเป็นของการเกิด Type II error จะมากกว่าในกรณี sample sizes เท่ากัน ถึงแม้ว่าจำนวน observations ทั้งหมดใน 2 samples ยังคงเดิม และ significance level ก็ยังคงเดิม เพราะฉะนั้นจึงต้องพยายามทำให้ sample sizes เท่ากัน ถ้าหากเราไม่อาจจะทำให้ sample sizes เท่ากันได้อย่างแท้จริงแล้ว ก็ควรจะทำให้มันใกล้กันมากที่สุดเท่าที่จะทำได้

ถ้าเรากำหนด n_1 ตายตัว การเพิ่มขึ้นของ n_2 จะไม่ช่วยลดความน่าจะเป็นของการเกิด Type II error ได้เลย ตัวอย่างเช่น ถ้า $n_1 = 25$, variance ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ คือ

$$\sigma^2 \left(\frac{1}{25} + \frac{1}{n_2} \right)$$

ซึ่งโตกว่า 0.04 σ^2 โดยไม่ต้องคำนึงถึง size (n_2) เลย หรืออีกนัยหนึ่ง ถ้ากำหนด $n_1 = 25$ ตายตัว variance ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ จะโตกว่าในกรณี $n_1 = n_2 = 50$ เสมอ โดยไม่ต้องคำนึงถึง size ของ sample ที่สองเลย variance ของผลต่าง $(\bar{y}_1 - \bar{y}_2)$ จึงลดลงต่อไปอีกไม่ได้ นอกจากเราจะเพิ่ม size ของ sample ที่หนึ่งขึ้น

10.8 Applications

การทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันนั้นได้ใช้กันทั่วไปในงานวิทยาศาสตร์ต่าง ๆ

ในการเปรียบเทียบคุณค่าของอาหารเลี้ยงวัว 2 ชนิด เราอาจแบ่งวัวทั้งหมดเช่น 50 ตัวโดยวิธีเสี่ยงออกเป็น 2 ฝูง ๆ ละ 25 ตัว วัวแต่ละฝูงจะถูกเลี้ยงด้วยอาหารคนละชนิด วัวแต่ละตัวถูกชั่งน้ำหนัก

2 ครั้ง คือตอนเริ่มทดลองและหมดระยะเวลาทดลองเลี้ยง น้ำหนักของวัวแต่ละตัวที่เพิ่มขึ้นระหว่างระยะเวลาทดลองเลี้ยง คือ observation ดังนั้น แต่ละฝูงจึงมี 25 observations (คือ $n_1 = n_2 = 25$) ชุดของ observations 2 ชุดนี้ต้องถือว่าเป็น 2 samples เพราะว่า observations อาจเป็นได้ต่าง ๆ กันถ้าการทดลองเลี้ยงนี้ถูกทำซ้ำกับวัวตัวอื่น ๆ ปัญหาเรื่องนี้ก็คือการพิจารณาว่าจะมีความแตกต่างอย่างแท้จริงในระหว่างความสามารถที่ทำให้วัวอ้วนของอาหาร 2 ชนิดนี้หรือไม่ กล่าวได้อีกอย่างหนึ่งว่าปัญหาก็คือการพิจารณาว่าโดย 2 samples ที่กำหนดให้นี้ 2 population means จะมีค่าเท่ากันหรือไม่ สมมติฐานที่ถูกทดสอบคือ 2 population means มีค่าเท่ากัน วิธีดำเนินการทดสอบได้ให้ไว้แล้วในข้อ 10.6 ถ้ายอมรับสมมติฐาน conclusion คืออาหารทั้ง 2 ชนิดดีพอกัน ถ้าปฏิเสธสมมติฐาน conclusion คืออาหารชนิดหนึ่งดีกว่าอีกชนิดหนึ่ง

t-test (Theorem 10.4a) ที่บรรยายในบทนี้และ t-test ของ paired observations (ข้อ 8.7) ใช้กับปัญหาแบบเดียวกัน ความแตกต่างระหว่างการทดสอบ 2 อย่างนี้อยู่ที่วิธี randomizing สิ่งที่ทดลอง ดังเช่นวัว 50 ตัวที่กล่าวมาแล้วข้างต้น ในวิธีของ paired observations เราจะต้องจัดวัว 50 ตัวนี้เข้าคู่กันเป็นคู่ ๆ โดยจงใจ ดังนั้น เราจะได้วัว 25 คู่ ซึ่งวัว 2 ตัวในคู่หนึ่งโดยเฉพาะมีความคล้ายคลึงกัน วัว 2 ตัวในคู่หนึ่งจะถูกเลี้ยงด้วยอาหาร 2 ชนิดซึ่งต่างกันโดยวิธีเสี่ยง หรืออีกนัยหนึ่งในทุกคู่วัว การกำหนดว่าวัวตัวใดจะเลี้ยงด้วยอาหารชนิดใดจะต้องทำโดยวิธีเสี่ยง ในทางตรงกันข้าม t-test ที่บรรยายในบทนี้ใช้ในกรณีซึ่งวัวจะถูกแบ่งโดยวิธี completely randomized โดยไม่มีการเลือกวัวให้เข้าคู่กัน ดังนั้นวิธีของ randomization จึงเป็นการพิจารณากำหนดว่าจะใช้ t-test แบบไหน (paired observations หรือ difference between sample means.)

10.9 Randomization

ข้อ 10.8 ได้เน้นถึงความจริงว่าวิธี randomizing สิ่งที่ทดลองทำให้เราพิจารณาได้ว่า t-test แบบไหนจึงจะเหมาะแก่การใช้ ข้อนี้จะอธิบายถึงว่าเราจะทำ randomization ได้อย่างไร เราจะใช้ตัวอย่างวัว 50 ตัวมาแสดงให้เห็นอีกครั้งหนึ่ง ในการแบ่งวัว 50 ตัวแบบ completely randomized นั้นจะทำได้โดยให้หมายเลขประจำวัวทุกตัวเสียก่อน เช่น 1, 2, 3, ..., 50 เลขเหล่านี้จะเป็นสิ่งระบุถึงวัวเหล่านั้น แล้วเราก็อ่าน random number table (Table 2, Appendix) ตามแนวดิ่ง, แนวราบ, หรือแนวทแยงก็ได้ และเลือกเลข 2 ตำแหน่ง (two-digit numbers) ใด ๆ ที่มีค่าระหว่าง 1 และ 50 เอาไว้ เลขที่เลือกไว้อาจมีค่า 03, 16, 12, 33, 18, ฯลฯ เหล่านี้เป็นต้น เลขใดที่เลือก

ไว้แล้วหรือที่มีค่าเกิน 50 จะต้องทิ้งไป การเลือกเลข 2 ตำแหน่งจะทำเรื่อยไปจนกว่าเราจะได้เลข 2 ตำแหน่งที่มีค่าต่าง ๆ กัน 25 จำนวน วัวตัวที่มีหมายเลขตรงกับเลขที่เราเลือกไว้นี้จะถูกเลี้ยงด้วยอาหารชนิดหนึ่ง ส่วนวัวที่เหลืออีก 25 ตัวจะถูกเลี้ยงด้วยอาหารอีกชนิดหนึ่ง วิธีการเลือก **random numbers** ที่ได้บรรยายข้างบนนี้ต้องการความเปลี่ยนแปลงบ้างบางครั้ง กล่าวคือจากค่า 00 ถึง 99 จะมีเลข 2 ตำแหน่งอยู่ 100 จำนวน แต่ในวิธีที่ได้บรรยายนั้น เลขใดที่มีค่ามากกว่า 50 จะต้องถูกทิ้งไป เพราะฉะนั้นค่าที่จะใช้ได้จึงมี 50 จำนวนเท่านั้น อีก 50 จำนวนจะเสียไปเปล่า ๆ อย่างไรก็ดี ค่าที่เสียไปนี้อาจใช้ประโยชน์ได้โดยเอา 50 ไปลบออกจากค่าที่เกิน 50 นั้น ๆ ตัวอย่างเช่น เราจะถือว่า 51 เป็น 1, 96 เป็น 46 และ 00 เป็น 50 โดยวิธีนี้แต่ละ **random number** จะใช้ได้ทั้งสิ้น ถ้าจำนวนวัวทั้งหมดเป็น 38 เลขที่มีค่าจาก 1 ถึง 38 จะใช้ได้ และเลขที่มีค่า 39 ถึง 76 ก็จะใช้ได้ด้วยโดยเอา 38 ไปลบออกจากแต่ละค่า ส่วนที่เหลืออีก 24 จำนวนก็จะต้องทิ้งไป อย่างไรก็ดีเพื่อให้ง่ายต่อการลบ เราจะใช้ค่า 51 ถึง 88 แทน 39 ถึง 76

Chapter 11

Confidence Interval

ในบทต่าง ๆ ที่แล้วมาเราได้พิจารณาปัญหาสถิติโดยวิธีทดสอบสมมติฐานแต่เพียงอย่างเดียวเท่านั้น แต่ในบทนี้จะกล่าวถึงวิธีพิจารณาปัญหาสถิติที่แตกต่างออกไปอีกอย่างหนึ่งคือการประมาณค่าของ parameter ตัวหนึ่งโดย interval

11.1 Inequality

inequality ซึ่งเป็นเรื่องหนึ่งที่ใช้ในคำราพีชคณิตเบื้องต้นแทบทั้งหมดนั้นจะได้นำมาใช้ในบทนี้ เราอาจสรุปหลักบางประการซึ่งเกี่ยวกับ **inequality** เพื่อสะดวกต่อการอ้างอิงได้ดังต่อไปนี้:

1. ถ้าเอาเลขจำนวนหนึ่งบวกเข้าหรือลบออกจากทั้งสองข้างของ **inequality** ทิศทางของเครื่องหมายของ **inequality** จะยังคงเป็นอย่างเดิม เช่น 2 น้อยกว่า 3 หรือ $2 < 3$ ถ้าเอา 10 บวกเข้าทั้งสองข้างของ **inequality** นี้ **inequality** ใหม่ที่ได้คือ $12 < 13$ ซึ่งยังคงเป็นจริงอยู่ตามเดิม และถ้าเอา 10 ลบออกจากทั้งสองข้างของ **inequality** $2 < 3$ **inequality** ใหม่ที่ได้คือ $-8 < -7$ ซึ่งก็ยังคงเป็นจริงอีก

2. ถ้าเอาเลขจำนวนหนึ่งซึ่งมีค่าบวก (**positive number**) คูณหรือหารทั้งสองข้างของ **inequality** ทิศทางของเครื่องหมายของ **inequality** จะยังคงเป็นอย่างเดิม เช่น เมื่อเอา 5 คูณทั้งสองข้างของ **inequality** $2 < 4$ **inequality** ใหม่ที่ได้คือ $10 < 20$ ซึ่งยังคงเป็นจริงอยู่ตามเดิม และถ้าเอา 2 หารทั้งสองข้างของ **inequality** $2 < 4$ **inequality** ใหม่ที่ได้คือ $1 < 2$ ซึ่งก็ยังคงเป็นจริงอีก

3. ถ้าเอาเลขจำนวนหนึ่งซึ่งมีค่าลบ (**negative number**) คูณหรือหารทั้งสองข้างของ **inequality** ทิศทางของเครื่องหมายของ **inequality** จะกลับทาง เช่นเมื่อเอา -1 คูณทั้งสองข้างของ **inequality** $2 < 4$ **inequality** ใหม่ที่ได้คือ $-2 > -4$ และถ้าเอา -2 หารทั้งสองข้างของ **inequality** $2 < 4$ **inequality** ใหม่ที่ได้คือ $-1 > -2$

11.2 Estimation by Interval

ปัญหาเรื่องการประมาณค่าของ parameter ตัวหนึ่งเราได้พิจารณากันมาแล้วในบทก่อน ๆ ตัวอย่างเช่นเราใช้ sample mean ($\bar{y}$) ประมาณค่าของ population mean (μ) แต่ sample mean ตัวหนึ่งไม่ค่อยจะมีค่าเท่ากับ population mean, samples ทั้งสี่ซึ่งแต่ละ sample ประกอบด้วย 5 observations ดังแสดงไว้ใน Table 4.2 นั้นเป็น random samples ซึ่งได้มาจาก tag population ที่มี mean (μ) = 50 แต่ไม่มี sample mean ($\bar{y}$) ตัวใดใน 4 sample means มีค่า = 50 เลย จึงเห็นได้ว่าโดยทั่วไปนั้น sample mean ตัวหนึ่งจะไม่เป็นค่าประมาณเพียงพอของ population mean เพราะฉะนั้นในบทนี้จะได้นำแนวทางใหม่อย่างหนึ่งมากล่าว เพื่อให้เราเข้าถึงปัญหาของการประมาณค่า นั่นคือ การใช้ interval ประมาณค่าของ population mean

ความจริงการใช้ interval ในการประมาณค่านี้เราได้ใช้กันทั่วไปในชีวิตประจำวันอยู่แล้ว เช่นการประมาณอายุคน เราอาจกล่าวว่าอายุของนาย ก. ราว 50 ปี นั่นคืออายุของนาย ก. อาจจะอยู่ตรงที่ใดที่หนึ่งระหว่าง 50 และ 60 ปี เราจะไม่ประมาณอายุของนาย ก. ว่า 52 ปี 5 เดือน กับ 14 วัน เช่นนี้เป็นแน่ ค่าประมาณโดย interval 50 ถึง 60 ปีเป็นค่าประมาณหยาบแต่ค่อนข้างจะถูกต้อง ส่วนค่าประมาณ 52 ปี 5 เดือน กับ 14 วัน เป็นค่าประมาณละเอียดแน่นอนแต่มักจะผิด

ถ้าเราใช้ interval ยิ่งยาวโอกาสที่จะประมาณค่าได้ถูกต้องก็จะยิ่งมีมาก เช่น เราอาจประมาณอายุของนาย ก. ว่าอยู่ระหว่าง 0 และ 200 ปีก็ได้ interval นี้จะครอบคลุมเอาอายุที่ถูกต้องของนาย ก. ไว้ด้วย เรามี 100 % confidence ในความถูกต้องของค่าประมาณนี้ แต่ interval ที่ยาวมากจะไร้ประโยชน์ เพราะว่าอายุของใครก็ตามจะต้องอยู่ในระหว่าง 0 และ 200 ปีทั้งสิ้น ถ้าเราทำให้ interval สั้นเช่นประมาณอายุของนาย ก. ว่าอยู่ระหว่าง 52 และ 53 ปี confidence ในความถูกต้องของค่าประมาณจะไม่มากเหมือนกับที่เราประมาณอายุของนาย ก. ว่าอยู่ระหว่าง 0 และ 200 ปี เพราะฉะนั้นความยาวของ interval และ degree of confidence ในความถูกต้องของค่าประมาณจึงเป็นสิ่งที่ควรสนใจอย่างมากในการประมาณค่าของ population mean โดย interval

11.3 Confidence Interval and Confidence Coefficient

ในข้อนี้ เราจะใช้การประมาณค่าของ population mean แสดงให้เข้าใจคำ "confidence interval" และ "confidence coefficient" เราทราบแล้วว่า statistic

$$u = \frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

follows the normal distribution ที่มี mean = 0 และ standard deviation = 1 (ข้อ 6.7). ถ้าเราหา all possible samples ซึ่งมี size n ออกมาจาก population หนึ่ง และคำนวณค่าของ u ไว้ทุก sample แล้ว 95 % ของค่าของ u เหล่านี้จะตกอยู่ในระหว่าง - 1.96 และ 1.96 นั่นคือ inequality

$$-1.96 < \frac{\bar{y} - \mu}{\sqrt{\frac{\sigma^2}{n}}} < 1.96 \text{ ....(1)}$$

ได้คลุมเอา 95 % ของ all possible samples ซึ่งมี size n ไว้ จาก Inequality (1) เราอาจหาได้ 2 intervals กล่าวคือเมื่อเอา $\sqrt{\frac{\sigma^2}{n}}$ คูณ element แต่ละตัวของ Inequality (1) แล้วเอา μ บวกเข้ากับ element แต่ละตัวอีกครั้งหนึ่ง interval ที่แสดงให้เห็นโดย inequality ใหม่ที่ได้จะเป็น

$$\mu - 1.96\sqrt{\frac{\sigma^2}{n}} < \bar{y} < \mu + 1.96\sqrt{\frac{\sigma^2}{n}} \text{ ....(2)}$$

เมื่อเอา $-\sqrt{\frac{\sigma^2}{n}}$ คูณ element แต่ละตัวของ Inequality (1) inequality ที่ได้คือ

$$1.96\sqrt{\frac{\sigma^2}{n}} > -\bar{y} + \mu > -1.96\sqrt{\frac{\sigma^2}{n}}$$

แล้วเอา $\bar{y}$ บวกเข้ากับ element แต่ละตัวของ inequality ข้างบนนี้ interval ที่แสดงให้เห็นโดย inequality ใหม่ที่ได้จะเป็น

$$\bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}} > \mu > \bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}}$$

หรือ

$$\bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}} < \mu < \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}} \quad \ldots\ldots\ldots(3)$$

intervals (2) และ (3) ทั้งสองนี้ได้มาจาก Inequality (1) ด้วยกัน เพราะฉะนั้นจึงคลุม 95 % ของ all possible samples ซึ่งมี size n เท่านั้น เนื่องจาก 2 intervals นี้ได้มาจาก inequality เดียวกันและมีรูปร่างโดยทั่วไปเป็นอย่างเดียวกันด้วย จึงทำให้ผู้เริ่มต้นศึกษาวิชาสถิติเกิดความสับสนได้บ่อย ๆ ในขั้นนี้เราต้องเข้าใจว่ามันต่างกัน ในระหว่างจำนวนต่าง ๆ ที่เกี่ยวข้องกันใน Inequalities (2) และ (3) $\bar{y}$ ตัวเดียวเท่านั้นที่มีค่าเปลี่ยนไปได้ตาม samples ใน Inequality (2) ขอบเขต 2 ข้างของ interval คือ

$$\mu - 1.96\sqrt{\frac{\sigma^2}{n}} \quad \text{และ} \quad \mu + 1.96\sqrt{\frac{\sigma^2}{n}}$$

และขอบเขตนี้จะไม่เปลี่ยนแปลงไปตาม samples ต่าง ๆ สำหรับ sampling experiment ที่ได้กล่าวไว้ในบทที่ 4 เมื่อ $\mu = 50$, $\sigma^2 = 100$ และ $n = 5$ ขอบเขต 2 ข้างของ interval คือ

$$50 - 1.96\sqrt{\frac{100}{5}} \quad \text{และ} \quad 50 + 1.96\sqrt{\frac{100}{5}}$$

หรือ $\quad 50 - 8.8 \quad$ และ $\quad 50 + 8.8$

หรือ $\quad 41.2 \quad$ และ $\quad 58.8$

จาก 4 sample meansที่ให้ไว้ใน Table 4.2 คือ 48.8, 45.6, 59.2 และ 55.4 นั้นจะเห็นว่ามี sample mean ตัวที่ 3 คือ 59.2 ตัวเดียวเท่านั้นที่ตกอยู่ข้างนอก interval 41.2 และ 58.8 นี้ ถ้าเราคำนวณ sample mean ($\bar{y}$) ของทุก sample จาก all possible samples ซึ่งมี size 5 ออกมาแล้ว 95 % ของ sample means เหล่านี้จะตกอยู่ข้างใน interval

Interval (3) ไม่ใช่ interval เดี่ยวเพียง interval เดียว แต่เป็นกลุ่มของ intervals ขอบเขต 2 ข้างของมันคือ

$$\bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}} \quad \text{และ} \quad \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}}$$

$$\text{หรือ} \quad \bar{y} - 8.8 \quad \text{และ} \quad \bar{y} + 8.8$$

นั้น จะเปลี่ยนไปตาม samples ต่าง ๆ ทั้งนี้เพราะ $\bar{y}$ มีค่าเปลี่ยนไปตาม samples ผลที่เกิดขึ้นคือแต่ละ sample จะมีหนึ่ง interval สำหรับ sample ที่หนึ่งใน Table 4.2 interval ของมันคือ 48.8 - 8.8 ถึง 48.8 + 8.8 หรือ 40.0 ถึง 57.6 intervals ต่าง ๆ ของอีก 3 samples ใน Table 4.2 คือ 36.8 ถึง 54.4, 50.4 ถึง 68.0 และ 46.6 ถึง 64.2 intervals ซึ่งเราคำนวณได้จาก samples ต่าง ๆ และใช้ประมาณค่าของ parameter ตัวหนึ่งนี้เรียกว่า confidence intervals interval 50.4 ถึง 68.0 จะจับค่าของ population mean 50 พลาดไป แต่อีก 3 intervals จะจับค่าของ population mean 50 ไว้ได้ ถ้าเราหา all possible samples ซึ่งมี size 5 จาก population และจากแต่ละ sample เราคำนวณค่า interval ของมันไว้ 95 % ของ intervals ต่าง ๆ เหล่านี้จะรวมเอาค่าของ population mean 50 เข้าไว้ด้วย จำนวนเปอร์เซ็นต์ของ all possible samples (ซึ่งมี size ที่กำหนด) ซึ่งให้ confidence intervals ที่จับค่าของ parameter ไว้ได้นี้เรียกว่า confidence coefficient ตามตัวอย่างนี้ confidence coefficient คือ 95 % ขอบเขตหรือจุดปลาย 2 ข้างของ interval เช่น

$$\bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}} \quad \text{และ} \quad \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}}$$

นั้นเรียกว่า confidence limits ความยาวของ confidence interval จึงเท่ากับผลต่างของ confidence limits นี่เอง ตามตัวอย่างนี้ ความยาวของ interval คือ

$$\bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}} - (\bar{y} - 1.96\sqrt{\frac{\sigma^2}{n}}) = \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}} - \bar{y} + 1.96\sqrt{\frac{\sigma^2}{n}}$$

$$= 2(1.96)\sqrt{\frac{\sigma^2}{n}}$$

$$= 2(1.96)\sqrt{\frac{100}{5}}$$

$$= 17.5$$

ณ บัดนี้ ควรจะเห็นได้ชัดถึงความแตกต่างระหว่าง intervals (2) และ (3) แล้ว ใน interval (2) นั้น ขอบเขตหรือจุดปลาย 2 ข้างของมันถูกจำกัดตายตัว แต่ element ตรงกลางคือ $\bar{y}$ มีค่าเปลี่ยนไปได้ตาม samples และ 95 % ของ sample means จะตกอยู่ข้างใน interval นี้ ในกรณีนี้ sample means ต่าง ๆ เปรียบเสมือนลูกบาสเกตบอลซึ่งอาจจะหล่นลงห่วงหรือหล่นนอกห่วงที่ตรึงแน่นอยู่กับที่ซึ่งคล้ายกับ interval (2) ก็ได้ ในทางตรงกันข้าม interval (3) เป็นกลุ่มของ confidence intervals ซึ่งแต่ละ interval คำนวณได้จากแต่ละ sample ดังนั้น confidence limits จึงเปลี่ยนไปได้ต่าง ๆ กันตาม samples แต่ element ตรงกลางคือ μ นั้นเป็นค่าคงที่ ดังนั้นบาง confidence intervals จะจับค่าของ population mean ไว้ได้ แต่บาง confidence intervals ก็จับค่าของ population mean พลาดไป ในความหมายนี้ confidence intervals ต่าง ๆ จึงเปรียบเสมือนเหล็กเกือกม้าที่เราโยนไปคล้องหรือพลาดหลักที่ปักแน่นอยู่กับที่ซึ่งคล้ายกับ population mean นั่นเอง confidence intervals ต่าง ๆ ที่คำนวณได้จาก sample ที่หนึ่ง, ที่สอง และที่สี่ใน Table 4.2 นั้นจับค่าของ population mean 50 ไว้ได้ แต่ confidence interval ที่คำนวณได้จาก sample ที่สามจับค่าของ population mean 50 นี้พลาดไป ดังนั้นแต่ละ confidence interval จึงอาจจับค่าของ population mean ไว้หรือพลาดไปก็ได้ จำนวนเปอร์เซ็นต์ของ all possible samples ซึ่งมี size เดียวกันและมี confidence intervals ซึ่งจับค่าของ population mean ไว้ได้นี้คือ confidence coefficient

เราเลือกใช้ confidence coefficient ได้ตามใจชอบ จึงไม่จำเป็นจะต้องเท่ากับ 95 % เสมอไป ถ้าเราใช้ confidence coefficient 99 % ก็จะต้องใช้ 2.576 แทน 1.96 ในการคำนวณ confidence limits และความยาวของ interval จะกลายเป็น

$$2(2.576)\sqrt{\frac{\sigma^2}{n}} = 2(2.576)\sqrt{\frac{100}{5}}$$
$$= 23.0$$

ณ บัดนี้เราจะเห็นได้ว่าการเพิ่ม confidence coefficient จะขยายความยาวของ confidence interval ออกไปโดยอัตโนมัติ เมื่อ confidence coefficient เพิ่มขึ้นจาก 95 % เป็น 99 % ความยาวของ confidence interval จะขยายจาก 17.5 เป็น 23.0 ทางเดียวที่จะให้ confidence coefficient สูงและมี confidence interval สั้นก็โดยการลด standard error of the mean ($\sqrt{\frac{\sigma^2}{n}}$) ลง ซึ่งจะทำได้โดยการลด population standard deviation (σ) (ข้อ 5.5) หรือเพิ่ม sample size (n) หรือทำทั้ง 2 อย่าง

พึงสังเกตว่า confidence interval อาจจับค่าของ population mean ไว้หรือพลาดไปก็ได้ การจับไว้ได้หรือพลาดไปนี้เป็นเรื่องธรรมดา จึงไม่มี Type I error และ Type II error เข้ามาเกี่ยวข้องกับ confidence interval เพราะ errors ทั้งสองนี้เกิดขึ้นในการทดสอบสมมติฐาน เมื่อ confidence interval ไม่เกี่ยวกับสมมติฐาน errors ทั้งสองนี้จึงเกิดขึ้นไม่ได้

11.4 Confidence Interval of Mean

วิธีหา confidence interval ของ population mean ในข้อ 11.3 นั้น เป็นตัวอย่างซึ่งทำให้เราทราบเรื่อง confidence interval และ confidence coefficient แต่วิธีนี้มี u-distribution เป็นพื้นฐานจึงเกือบจะไม่มีประโยชน์เลย เพราะตามปกติเราไม่ทราบค่าของ population variance (σ^2) วิธีที่ดีกว่าและสะดวกกว่าสำหรับใช้หา confidence interval ของ population mean เป็นวิธีที่ได้มาจาก t-distribution

เราทราบแล้วว่าถ้าหา all possible samples ซึ่งมี size n ออกมาจาก normal population ที่มี mean เท่ากับ μ และเราคำนวณค่าของ t ของแต่ละ sample ไว้ 95 % ของค่าของ t ทั้งหมดจะตกอยู่ในระหว่าง $-t_{0.025}$ และ $t_{0.025}$ ในเมื่อ $t_{0.025}$ เป็น 2.5 % point ของ Student's t-distribution with n - 1 d.f. (Theorem 8.1 a)

กล่าวได้อีกอย่างหนึ่งว่า inequality

$$-t_{0.025} < \frac{\bar{y} - \mu}{\sqrt{\frac{s^2}{n}}} < t_{0.025} \quad \ldots\ldots\ldots\ldots(1)$$

เป็นความจริงสำหรับ 95 % ของ all possible samples ซึ่งมี size n เมื่อเอา $-\sqrt{\frac{s^2}{n}}$ คูณแต่ละ term ใน 3 terms ของ inequality ข้างบนนี้ inequality ที่ได้คือ

$$t_{0.025}\sqrt{\frac{s^2}{n}} > -\bar{y} + \mu > -t_{0.025}\sqrt{\frac{s^2}{n}}$$

และเมื่อเอา $\bar{y}$ บวกเข้ากับ inequality ที่ได้นี้อีก inequality ใหม่จะกลายเป็น

$$\bar{y} + t_{0.025}\sqrt{\frac{s^2}{n}} > \mu > \bar{y} - t_{0.025}\sqrt{\frac{s^2}{n}}$$

หรือ

$$\bar{y} - t_{0.025}\sqrt{\frac{s^2}{n}} < \mu < \bar{y} + t_{0.025}\sqrt{\frac{s^2}{n}} \quad \ldots\ldots\ldots\ldots (2)$$

Inequality (2) บ่งถึง confidence interval ของ μ โดยใช้ confidence coefficient 95 % และควรจะเข้าใจไว้ด้วยว่า confidence interval ที่แสดงโดย Inequality (2) นั้นไม่ใช่หนึ่ง interval แต่เป็นได้หลาย intervals จากแต่ละ sample ซึ่งประกอบด้วย n observations เราคำนวณ $\bar{y}$ และ s^2 ได้ และเราได้ค่าของ $t_{0.025}$ จาก t- table ดังนั้นเราจึงคำนวณ confidence interval ของ μ ได้ ถ้าเราหา all possible samples ซึ่งมี size n ออกมาจาก normal population 95 % ของ samples ทั้งหมดนี้จะให้ค่าของ confidence intervals ซึ่งมีค่าของ population mean รวมอยู่ด้วย เราอาจใช้ sampling experiment ที่ได้บรรยายไว้ในบทที่ 4 สาธิตความหมายของ confidence interval และ confidence coefficient ของมันได้ samples ทั้งสี่ซึ่งแต่ละ sample ประกอบด้วย 5 observations ที่ได้ให้ไว้ใน Table 4.2 นั้นเป็น random samples ซึ่งได้มาจาก normal population ที่มี mean (μ) = 50 ค่าของ $\bar{y}$ และ $\sqrt{\frac{s^2}{n}}$ ของแต่ละ sample ได้คำนวณไว้แล้ว ค่าของ $t_{0.025}$ คือ 2.7764 สำหรับ sample ที่หนึ่ง confidence interval ของ population mean (เมื่อใช้ confidence coefficient 95%) คือ

$$48.8 - (2.7764 \times 5.472) < \mu < 48.8 + (2.7764 \times 5.472)$$

หรือ

$$33.6 < \mu < 64.0$$

เนื่องจากเราทราบแล้วว่า population mean (μ) = 50 sample นี้จึงให้ confidence interval ซึ่งรวมเอาค่าของ population mean เข้าไว้ด้วย confidence interval ที่คำนวณได้จากอีก 3 samples คือ 36.0 ถึง 55.2, 46.8 ถึง 71.6 และ 44.3 ถึง 66.5 ตามลำดับ แต่ละ confidence interval ใน 3 confidence intervals นี้ ได้รวมเอาค่าของ population mean 50 เข้าไว้ด้วยเหมือนกัน ถ้าเราหา all possible samples ซึ่งมี size 5 ออกมา และจากแต่ละ sample เราคำนวณ confidence interval ไว้ 95 % ของ intervals ทั้งหมดจะรวมเอาค่าของ population mean 50 ไว้ด้วย

เราอาจใช้ confidence interval ของ population mean กับปัญหาชนิดเดียวกันซึ่งได้กล่าวไว้ในข้อ 8.6 และ 8.7 ได้ ในข้อเหล่านั้นปัญหาต่าง ๆ เป็นการทดสอบสมมติฐาน แต่ในข้อนี้ไม่มีสมมติฐาน และปัญหาคือการประมาณค่าของ population mean การทดสองเรื่องบีทหวานที่ได้กล่าวไว้ในข้อ 8.7 นั้นอาจถูกใช้แสดงให้เห็นความแตกต่างในปัญหา 2 ชนิดนี้ได้ ในการทดสอบสมมติฐานว่า population mean = 0 นั้น วัตถุประสงค์ก็เพื่อพิจารณาว่าการใช้ปุ๋ยทำให้เพิ่มผลผลิตของหัวบีทหวานหรือไม่ แต่ในการหา confidence interval ของ population mean วัตถุประสงค์นั้นเพื่อพิจารณาว่าผลผลิตของหัวบีทหวานที่เพิ่มขึ้นโดยการใช้ปุ๋ยนั้นเป็นจำนวนเท่าไร

ในการปฏิบัติเราต้องการเพียงหนึ่ง sample (ซึ่งอาจจะประกอบด้วยหลาย observations) ก็เพียงพอและเราจะคำนวณเพียงหนึ่ง interval เท่านั้น เราต้องเลือกใช้ confidence coefficient ก่อนที่จะคำนวณ confidence interval ถ้าเราใช้ 95 % confidence coefficient เราก็ใช้ $t_{0.025}$ สำหรับคำนวณ interval แต่ถ้าใช้ 99 % confidence coefficient ก็จะต้องใช้ $t_{0.005}$ สำหรับคำนวณ interval confidence interval ที่คำนวณได้จาก sample หนึ่งที่กำหนดให้อาจจะรวมหรือไม่รวมเอาค่าของ population mean เข้าไว้ก็ได้ เนื่องจากเราไม่ทราบค่าของ population mean จึงไม่มีทางที่จะพิจารณาได้ว่า interval จะรวมเอาค่าของ population mean ไว้จริงหรือไม่

วิธีคำนวณ confidence limits ของ population mean นั้นง่ายมาก และวิธีคำนวณ $\bar{y}$ และ $\frac{s^2}{n}$ ก็ได้ให้ไว้แล้วในข้อ 8.5 เมื่อเราได้ค่าของ $\bar{y}$ และ $\frac{s^2}{n}$ ก็จะหา intervals ได้โดย Inequality (2) confidence intervals ต่าง ๆ ของ population mean ได้คำนวณไว้แล้วจาก 4 samples ที่ให้ไว้ใน Table 4.2 นิสิตควรจะฝึกหัดคำนวณให้คล่องโดยคำนวณ intervals เหล่านี้ซ้ำอีก

ความยาวของ confidence interval ที่ให้ไว้ใน Inequality (2) คือ

$$2t_{0.025}\sqrt{\frac{s^2}{n}}$$

นั้นจะเปลี่ยนไปได้ตาม samples ทั้งนี้เป็นเพราะค่าของ s^2 เปลี่ยนไปตาม samples แต่ความยาวเฉลี่ยของ confidence intervals ที่คำนวณได้จาก all possible samples ซึ่งมี size n จะหดสั้นลงโดยการเพิ่ม sample size

11.5 Confidence Interval of Difference Between Means

เราอาจหา confidence interval ของผลต่างระหว่าง 2 population means ได้จาก Theorem 10.4 a limits ของ 95 % confidence interval ของ $\mu_1 - \mu_2$ คือ

$$(\bar{y}_1 - \bar{y}_2) \pm t_{0.025}\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \ldots\ldots\ldots\ldots\ldots(1)$$

ถ้าต้องการใช้ confidence coefficient 99 % ก็ต้องเอา 0.5 % point ของ t- distribution with $n_1 + n_2 - 2$ d.f. ไปใช้แทน 2.5 % point วิธีคำนวณ $(\bar{y}_1 - \bar{y}_2)$ และ $\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ ได้ให้ไว้แล้วใน Table 10.6 หลังจากเราคำนวณ $(\bar{y}_1 - \bar{y}_2)$ และ $\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ ได้แล้วก็จะหา confidence limits ได้ง่ายมาก 95 % confidence interval ซึ่งพิจารณาโดยคู่ของ samples ที่ให้ไว้ใน Table 10.6 คือ

$$(\bar{y}_1 - \bar{y}_2) - t_{0.025}\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})} < \mu_1 - \mu_2 < (\bar{y}_1 - \bar{y}_2) + t_{0.025}\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$$

$$(4 - 8) - 3.1825\sqrt{\frac{16}{3}(\frac{1}{3} + \frac{1}{2})} < \mu_1 - \mu_2 < (4 - 8) + 3.1825\sqrt{\frac{16}{3}(\frac{1}{3} + \frac{1}{2})}$$

$$-4 - 3.1825\sqrt{4.444} < \mu_1 - \mu_2 < -4 + 3.1825\sqrt{4.444}$$

$$-4 - (3.1825 \times 2.108) < \mu_1 - \mu_2 < -4 + (3.1825 \times 2.108)$$

$$-4 - 6.7 < \mu_1 - \mu_2 < -4 + 6.7$$

$$-10.7 < \mu_1 - \mu_2 < 2.7$$

กล่าวได้อีกอย่างหนึ่งว่าผลต่างระหว่าง 2 population means อยู่ตรงที่หนึ่งระหว่าง - 10.7 และ 2.7

confidence interval และ การทดสอบสมมติฐานทั้ง 2 อย่างนี้ใช้กับ Theorem 10.4 a เดียวกัน แต่วัตถุประสงค์ของการใช้ต่างกัน วัตถุประสงค์ของการทดสอบสมมติฐานดังที่ได้อธิบายไว้ในข้อ 10.6 และ 10.8 นั้นคือการพิจารณาว่า 2 population means มีค่าอย่างเดียวกันหรือไม่ แต่วัตถุประสงค์ของการหา confidence interval ของ $\mu_1 - \mu_2$ นั้นคือการประมาณขนาดของผลต่างระหว่าง 2 population means

ได้แสดงไว้แล้วในข้อ 10.7 ว่า สำหรับจำนวน observations ทั้งหมดที่กำหนดให้ variance ของผลต่างระหว่าง 2 sample means

$$\sigma^2 (\frac{1}{n_1} + \frac{1}{n_2})$$

จะมีค่าน้อยที่สุดเมื่อ sample sizes ทั้งสองเท่ากัน จะเห็นได้จาก Inequality (1) ว่า ข้อได้เปรียบของการทำให้ sample sizes เท่ากันคือการทำให้ความยาวเฉลี่ยของ confidence intervals ต่าง ๆ ของผลต่างระหว่าง 2 population means สั้นลง .

Chapter 12

Analysis of Variance, One-Way Classification

ในข้อ 7.9 ได้เคยเกริ่นไว้ว่า analysis of variance เป็นวิธีแตก sum of squares ออกเป็นส่วน ๆ วัตถุประสงค์ของวิธีการนี้คือการทดสอบสมมติฐานว่า population means ต่าง ๆ มีค่าเท่ากัน วิธีทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันนั้นได้ให้ไว้แล้วในบทที่ 10 เพราะฉะนั้นเราจึงถือว่า analysis of variance เป็นการขยายเรื่อง t-test ให้กว้างขวางยิ่งขึ้น

12.1 Mechanics of Partition of Sum of Squares

ในข้อนี้จะบรรยายเฉพาะกลวิธีการแตก sum of squares เท่านั้น ส่วนการอธิบายความหมายต่าง ๆ จะกล่าวในข้อต่อไป ปัญหานี้เกี่ยวข้องกับหลาย samples หรือ k samples ซึ่งแต่ละ sample ประกอบด้วย n observations samples เหล่านี้เราจะเรียกชื่อว่า sample ที่หนึ่ง, sample ที่สอง, ..., และ sample ที่ k ตามลำดับ sample means ของ samples เหล่านั้นคือ $\bar{y}_1$, $\bar{y}_2$, ..., และ $\bar{y}_k$ กลวิธีต่าง ๆ จะได้อธิบายให้ทราบโดยตัวอย่างซึ่งแสดงไว้ใน Table 12.1 a ตัวอย่างนี้เกี่ยวข้องกับ 3 samples ($k = 3$) ซึ่งแต่ละ sample ประกอบด้วย 5 observations ($n = 5$) ผลบวกทั้งหมดของ n observations ของ sample หนึ่งแสดงด้วย T แทน Σy ที่เคยใช้ทั้งนี้เพื่อทำให้ algebraic expressions ง่ายเข้า ดังนั้น T_1 T_2, ... และ T_k จึงเป็นผลบวกของ observations ต่าง ๆ ในแต่ละ sample ของ sample ที่หนึ่ง, sample ที่สอง, ..., และ sample ที่ k ตามลำดับ ตามตัวอย่างที่แสดงไว้ใน Table 12.1 a นั้นผลบวกของ observations ต่าง ๆ ของแต่ละ sample ใน 3 samples คือ 25, 45 และ 20 <u>ผลบวกของ observations ทั้งหมดของ k samples เรียกว่า grand total ซึ่งแสดงด้วย G</u> ดังนั้น $G = \Sigma T = T_1 + T_2 + \ldots + T_k$ ตามตัวอย่างนี้ $G = 25 + 45 + 20 = 90$ ให้สังเกตว่า G เป็นผลบวกของ 15 observations ด้วย ($kn = 15$)

Table 12.1 a

sample No.	1	2	3	
observations	3 7 7 6 2	9 12 11 8 5	1 2 6 4 7	
total, T mean, $\bar{y}$	25 5	45 9	20 4	90 = G 6 = $\bar{\bar{y}}$

ณ บัดนี้จะเห็นได้ว่า ถ้า summation signs 2 ตัวมีความหมายต่างกัน เช่น $T = \Sigma y$ และ $G = \Sigma T$ ดังนี้แล้ว summation sign (Σ) ก็จะแสดงความหมายไม่พอ ในกรณี $T = \Sigma y = y_1 + y_2 + \ldots\ldots + y_n$ เครื่องหมาย Σ แสดงถึง ผลบวกของ n terms แต่ในกรณี $G = \Sigma T = T_1 + T_2 + \ldots\ldots + T_k$ เครื่องหมาย Σ แสดงถึงผลบวกของ k terms เพื่อทำให้มันต่างกันเราจะต้องเขียนเครื่องหมายกำกับไว้ที่เครื่องหมาย Σ เพื่อแสดงจำนวน terms ทั้งหมดที่จะบวกเข้าด้วยกัน ตัวอย่างเช่น

$$T = \overset{n}{\Sigma}\, y$$

และ

$$G = \overset{k}{\Sigma}\, T = \overset{kn}{\Sigma}\, y$$

เราเรียก mean ของ kn observations ว่า general mean ซึ่งแสดงด้วย $\bar{\bar{y}}$ ดังนั้น

$$\bar{\bar{y}} = \frac{G}{kn} = \frac{90}{15} = 6$$

ให้สังเกตว่า general mean ($\bar{\bar{y}}$) นี้เป็น mean ของ k sample means ด้วย นั่นคือ

$$\bar{\bar{y}} = \frac{\bar{y}_1 + \bar{y}_2 + \ldots\ldots + y_k}{k}$$

$$= \frac{5 + 9 + 4}{3}$$

$$= 6$$

observation แต่ละตัวใน 15 observations (kn = 15) อาจแตกออกได้เป็น 3 ส่วน ดังนี้

$$y = \bar{\bar{y}} + (\bar{y} - \bar{\bar{y}}) + (y - \bar{y}) \text{ .... (1)}$$

Equation (1) ข้างบนนี้เป็น algebraic identity ถ้าเราทำให้เป็นผลสำเร็จแล้ว Equation (1) ก็จะเป็น y = y ตัวอย่างเช่น observation ตัวแรกใน Table 12.1 a คือ 3, sample mean ตัวที่หนึ่งคือ 5 และ general mean คือ 6 ดังนั้นโดย Equation (1)

$$\begin{aligned} 3 &= 6 + (5 - 6) + (3 - 5) \\ &= 6 - 1 - 2 \end{aligned}$$

จำนวน (-1) ข้างบนนี้คือผลต่างของ sample mean ตัวที่หนึ่งจาก general mean และจำนวน (-2) คือผลต่างของ observation จาก sample mean ตัวที่หนึ่ง ส่วนประกอบของแต่ละ observation ของ 15 observations ได้แสดงไว้ใน Table 12.1 b

Table 12.1 b

sample No.	1	2	3
components of observations $\bar{\bar{y}} + (\bar{y} - \bar{\bar{y}}) + (y - \bar{y})$	6 - 1 - 2 6 - 1 + 2 6 - 1 + 2 6 - 1 + 1 6 - 1 - 3	6 + 3 + 0 6 + 3 + 3 6 + 3 + 2 6 + 3 - 1 6 + 3 - 4	6 - 2 - 3 6 - 2 - 2 6 - 2 + 2 6 - 2 + 0 6 - 2 + 3
sample mean, $\bar{y}$	6 - 1 + 0	6 + 3 + 0	6 - 2 + 0

วัตถุประสงค์ของการแตก observation แต่ละตัวออกเป็นส่วน ๆ ก็เพื่ออธิบาย algebraic identity ว่า

$$\sum^{kn}(y - \bar{\bar{y}})^2 = \sum^{kn}(\bar{y} - \bar{\bar{y}})^2 + \sum^{kn}(y - \bar{y})^2 \text{ .... (2)}$$

$$\sum^{kn}(y - \bar{\bar{y}})^2 = \sum^{kn}(\bar{y} - \bar{\bar{y}})^2 + \sum^{kn}(y - \bar{y})^2 + 2\sum^{kn}(\bar{y} - \bar{\bar{y}})(y - \bar{y})$$

sum of squares หรือ $\Sigma(y - \bar{\bar{y}})^2$ ใน Equation (2) นี้เรียกว่า <u>total SS</u> ซึ่งเป็น SS ของ sample รวมทั้ง kn หรือ 15 observations สำหรับตัวอย่างที่ให้ไว้ (Table 12.1 a)

$$\sum^{kn}(y - \bar{\bar{y}})^2 = (3-6)^2 + (7-6)^2 + \ldots\ldots\ldots + (7-6)^2$$

$$= 148$$

term กลางของ Equation (2) เป็น sum of squares ของ components ตัวกลางของ observations (Equation (1) และ Table 12.1 b) นั่นคือ

$$\sum^{kn}(\bar{y} - \bar{\bar{y}})^2 = 5\left[(-1)^2 + (3)^2 + (-2)^2\right]$$

$$= 70$$

หรืออีกนัยหนึ่ง

$$\sum^{kn}(\bar{y} - \bar{\bar{y}})^2 = n\sum^{k}(\bar{y} - \bar{\bar{y}})^2 \ldots\ldots\ldots \quad (3)$$

term สุดท้ายของ Equation (2) เป็น sum of squares ของ components ตัวสุดท้ายของ observations นั่นคือ

$$\sum^{kn}(y - \bar{y})^2 = (-2)^2 + 2^2 + \ldots\ldots\ldots + 3^2$$

$$= 78$$

ให้สังเกตว่า $148 = 70 + 78$ ตัวอย่างที่เป็นตัวเลขนี้ได้แสดงให้เห็นความเป็นจริงของ Equation (2) ด้วยแล้ว

การพิสูจน์ Equation (2) โดยพีชคณิตทำได้ง่ายมาก อย่างไรก็ดีนักสถิติซึ่งมีพื้นความรู้ทางคณิตศาสตร์จำกัดอาจรู้สึกว่าการใช้สัญลักษณ์อันยุ่งเหยิงในการพิสูจน์ identity นั้น ทำให้เกิดความฉงนได้มากกว่าจะเข้าใจอย่างแจ่มแจ้ง ดังนั้นการอธิบายการพิสูจน์ต่อไปนี้อาจเป็นประโยชน์มากกว่า

ค่าของ $(\bar{y} - \bar{\bar{y}})$ และ $(y - \bar{y})$ ของแต่ละ observation ของ 15 observations ($kn = 15$) ได้ให้ไว้ใน Table 12.1 b จำนวน $(\bar{y} - \bar{\bar{y}})$ นี้เป็น component ตัวกลางและ $(y - \bar{y})$ เป็น component ตัวสุดท้าย เพราะฉะนั้น Equation (2) จึงหมายถึง total SS เท่ากับ sum of squares ของ components ตัวกลางบวกด้วย sum of squares ของ components ตัวสุดท้ายของ 15 observations

เมื่อเอา $\bar{\bar{y}}$ ลบ 2 ข้างของ Equation (1) Equation ใหม่ที่ได้คือ

$$(y - \bar{\bar{y}}) = (\bar{y} - \bar{\bar{y}}) + (y - \bar{y}) \quad \ldots\ldots\ldots\ldots (4)$$

และเมื่อยกกำลังสองทั้งสองข้างของ Equation (4) จะได้

$$(y - \bar{\bar{y}})^2 = \left[(\bar{y} - \bar{\bar{y}}) + (y - \bar{y}) \right]^2$$

$$= (\bar{y} - \bar{\bar{y}})^2 + (y - \bar{y})^2 + 2(\bar{y} - \bar{\bar{y}})(y - \bar{y})$$

สำหรับ observation ตัวแรกของ sample แรก Equation นี้คือ

$$(3 - 6)^2 = (-1)^2 + (-2)^2 + 2(-1)(-2)$$

$$9 = 1 + 4 + 4$$

และสำหรับ observation ตัวที่สองของ sample แรก Equation นี้คือ

$$(7 - 6)^2 = (-1)^2 + 2^2 + 2(-1)(2)$$

$$1 = 1 + 4 - 4$$

ดังนั้นแต่ละ observation ใน 15 observations ($kn = 15$) จะมี Equation ดังกล่าว เมื่อรวม 15 equations ทั้งหมดเข้าด้วยกัน ผลลัพธ์ที่ได้คือ

$$\Sigma (y - \bar{\bar{y}})^2 = \Sigma (\bar{y} - \bar{\bar{y}})^2 + \Sigma (y - \bar{y})^2 + 2 \Sigma (\bar{y} - \bar{\bar{y}})(y - \bar{y})$$

ความแตกต่างระหว่าง Equation ข้างบนกับ Equation (2) คือ term $2\sum(\bar{y}-\bar{\bar{y}})(y-\bar{y})$ ฉะนั้น ปัญหาที่เกิดขึ้นก็คือต้องแสดงให้เห็นว่า $\sum(\bar{y}-\bar{\bar{y}})(y-\bar{y}) = 0$ เนื่องจาก $(\bar{y}-\bar{\bar{y}})$ และ $(y-\bar{y})$ เป็น components ตัวที่สองและที่สามของ observation ผลบวกของผลคูณ (sum of products) คือ $\sum(\bar{y}-\bar{\bar{y}})(y-\bar{y})$ นั้นหาได้จาก Table 12.1 b เลขจำนวนนี้เป็นผลบวกของผลคูณต่าง ๆ ของ components ตัวที่สองและที่สามของ 15 observations (kn = 15) นั่นคือ

$$\begin{aligned}\sum(\bar{y}-\bar{\bar{y}})(y-\bar{y}) &= (-1)(-2)+(-1)(+2)+\ldots\ldots+(-2)(+3)\\ &= (-1)(-2+2+2+1-3)+(3)(0+3+2-1-4)+\\ &\quad (-2)(-3-2+2+0+3)\end{aligned}$$

แต่ผลบวกของผลต่างของ observations จาก mean ของมัน = 0 (ข้อ 7.4, 7.8) เพราะฉะนั้น

$$\begin{aligned}\sum(\bar{y}-\bar{\bar{y}})(y-\bar{y}) &= (\bar{y}_1-\bar{\bar{y}})(0)+(\bar{y}_2-\bar{\bar{y}})(0)+\ldots\ldots+(\bar{y}_k-\bar{\bar{y}})(0)\\ &= 0\end{aligned}$$

จึงพิสูจน์ identity

$$\sum^{kn}(y-\bar{\bar{y}})^2 = \sum^{kn}(\bar{y}-\bar{\bar{y}})^2+\sum^{kn}(y-\bar{y})^2$$

โดยสมบูรณ์แล้ว และเราอาจเขียน Equation ข้างบนนี้ได้อีกแบบหนึ่ง (ทำนองเดียวกับ Equation(3)) คือ

$$\sum^{kn}(y-\bar{\bar{y}})^2 = n\sum^{k}(\bar{y}-\bar{\bar{y}})^2+\sum^{kn}(y-\bar{y})^2 \ \ldots\ldots\ldots\ldots(5)$$

sum of squares ทั้งสามข้างบนวัดการแปรผันของ kn observations ในทางต่างกัน total SS ($\sum^{kn}(y-\bar{\bar{y}})^2$) ซึ่งเป็น SS ของ sample รวมที่มี kn observations วัดการแปรผันทั้งหมดของ kn observations total SS จะเท่ากับศูนย์ ในเมื่อ kn observations มีค่าอย่างเดียวกัน (Table 12.1 c) sum of squares ($n\sum^{k}(\bar{y}-\bar{\bar{y}})^2$)ซึ่งเรียกว่า among-sample SS วัดการแปรผันในระหว่าง k sample means among-sample SS จะเท่ากับศูนย์ ในเมื่อ k sample means มีค่าอย่างเดียวกัน แต่ observations ต่าง ๆ ภายใน sample หนึ่งไม่จำเป็นต้องมีค่าอย่างเดียวกัน (Table 12.1 d)

sum of squares ($\sum^{kn}(y-\bar{y})^2$)ซึ่งเรียกว่า within-sample SS วัดการแปรผันของ observations ต่าง ๆ ภายใน samples within-sample SS จะเท่ากับศูนย์ ในเมื่อ observations ทุกตัวภายในแต่ละ sample ของ k samples มีค่าอย่างเดียวกัน แต่ k sample means ไม่จำเป็นต้องมีค่าอย่างเดียวกัน (Table 12.1 e) within-sample SS นี้คือผลบวกของค่าของ SS ต่าง ๆ ของ k samples โดยแท้ ดังนั้นจึงเรียกว่า pooled SS ได้อีกอย่างหนึ่งด้วย (ข้อ 9.6) นั่นคือ

$$\sum^{kn}(y-\bar{y})^2 = \sum^{n}(y-\bar{y}_1)^2+\sum^{n}(y-\bar{y}_2)^2+\ldots\ldots+\sum^{n}(y-\bar{y}_k)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots(6)$$

หรือ within-sample SS $= SS_1 + SS_2 + \ldots\ldots + SS_k$

$= \text{pooled SS} \quad \ldots\ldots\ldots\ldots\ldots\ldots(7)$

Table 12.1 c

1	2	3
6	6	6
6	6	6
6	6	6
6	6	6
6	6	6

Table 12.1 d

	1	2	3
	4	6	3
	8	9	4
	8	8	8
	7	5	6
	3	2	9
T	30	30	30
$\bar{y}$	6	6	6

Table 12.1 e

	1	2	3
	5	9	4
	5	9	4
	5	9	4
	5	9	4
	5	9	4
T	25	45	20
$\bar{y}$	5	9	4

12.2 Statistical Interpretation of Partition of Sum of Squares

ในข้อ 12.1 ได้อธิบายเรื่องกลวิธีการแตก sum of squares มาแล้ว กล่าวคือ total SS ซึ่งเป็น SS ของ sample รวมของ k samples ซึ่งแต่ละ sample ประกอบด้วย n observations นั้น แตกตัวออกได้เป็น 2 components เรียกว่า among-sample SS และ within-sample SS ในข้อนี้จะได้อธิบายทางสถิติเกี่ยวกับการแตกตัวดังกล่าวนี้

เราพิจารณา k samples ว่าเป็น random samples ซึ่งได้มาจาก normal population เดียวกัน เพราะฉะนั้นจึงมีเหตุผลที่จะรวม kn observations ทั้งหมดของ k samples เข้าด้วยกันให้เป็น sample รวม sample เดียว ถ้า all possible samples ซึ่งแต่ละ sample มี kn observations ได้มาจาก normal population หนึ่ง และจากแต่ละ sample เราคำนวณค่าของ SS ไว้ distribution ของ $\frac{SS}{\sigma^2}$ follows the χ^2- distribution with kn - 1 d.f. (Theorem 7.7 a) หรืออีกนัยหนึ่ง $\frac{\text{total SS}}{\sigma^2}$ follows the χ^2- distribution with kn - 1 d.f.

distribution ของ among-sample SS ก็อาจหาได้จาก Theorem 7.7 a distribution ของ means ($\bar{y}$) ของ all possible samples ซึ่งมี size n นั้น อาจถือเป็น population หนึ่งได้ ดังนั้น population นี้จึงมี $\bar{y}$ ต่าง ๆ เป็น observations แทนที่จะเป็น y เนื่องจาก population ต้นกำเนิดของมันเป็น normal เพราะฉะนั้น sample means ($\bar{y}$) จึง follow the normal distribution (Theorem 5.2 b) mean ($\mu_{\bar{y}}$) ของ population ของ sample means นี้เท่ากับ population mean (μ) และ variance ($\sigma^2_{\bar{y}}$)เท่ากับ $\frac{\sigma^2}{n}$ (Theorem 5.3) ดังนั้น k sample means คือ $\bar{y}_1$, $\bar{y}_2$, ..., $\bar{y}_k$ จึงกลายเป็น sample หนึ่งซึ่งมี k observations ที่ได้มาจาก population ของ sample means นี้ ถ้า all possible samples ซึ่งแต่ละ sample ประกอบด้วย k sample means ได้มาจาก population ของ sample means แล้ว เราก็อาจคำนวณค่าของ SS ของแต่ละ sample ซึ่งประกอบด้วย k sample means ได้ distribution ของ

$$\frac{\sum^{k} (\bar{y}-\bar{\bar{y}})^2}{\frac{\sigma^2}{n}} = \frac{n\sum^{k} (\bar{y}-\bar{\bar{y}})^2}{\sigma^2}$$

$$= \frac{\text{among-sample SS}}{\sigma^2}$$

ต้อง follow the χ^2- distribution with k - 1 d.f. (Theorem 7.7 a) distribution ของ within-sample SS ก็อาจจะหาได้จาก Theorems ต่าง ๆ ที่ได้กำหนดไว้ในบทต่าง ๆ ที่แล้วมา within-sample SS นี้คือ pooled SS ของ k samples นั่นเอง เราทราบแล้วว่า statistic $\frac{SS}{\sigma^2}$ follows the χ^2- distribution with n-1 d.f. (Theorem 7.7 a) ถ้า k samples เป็น independent samples แล้ว statistic

$$\frac{\text{pooled SS}}{\sigma^2} = \frac{SS_1 + SS_2 + \ldots\ldots\ldots\ldots + SS_k}{\sigma^2}$$

$$= \frac{\text{within-sample SS}}{\sigma^2}$$

follows the χ^2- distribution with $(n_1 - 1) + (n_2 - 1) + \ldots\ldots\ldots\ldots + (n_k - 1)$ d.f. (ข้อ 9.6) เนื่องจากในกรณีนี้ sample size ของทุก sample เท่ากัน ดังนั้น d.f. ของ pooled SS หรือ within-sample SS จะกลายเป็น $k(n-1)$

ณ บัดนี้เราจะเห็นได้ว่าไม่เพียงแต่ total SS จะแตกตัวออกเป็น components ต่าง ๆ เท่านั้น d.f. ของ total SS ก็แตกตัวออกไปด้วย ความสัมพันธ์ระหว่างค่าของ SS ทั้งสามและ d.f. ของมันอาจสรุปได้ดังนี้

SS : total = among-sample + within-sample

d.f. : $kn - 1 = k - 1 + k(n-1)$

เนื่องจาก sample variance (s^2) ซึ่งเท่ากับ SS หารด้วย d.f. ของมันเป็นค่าประมาณของ variance ของ population (σ^2) ซึ่งเราได้ sample นั้นมา ดังนั้น variance ของ sample ที่ประกอบด้วย k sample means คือ

$$s^2_{\bar{y}} = \frac{\sum^{k} (\bar{y} - \bar{\bar{y}})^2}{k - 1} \quad \ldots\ldots\ldots\ldots (1)$$

จึงเป็นค่าประมาณของ variance ของ population ของ sample means หรือ $s^2_{\bar{y}}$ เป็นค่าประมาณของ $\sigma^2_{\bar{y}}$ หรือ $\frac{\sigma^2}{n}$ ดังนั้น $n\,s^2_{\bar{y}}$ จะเป็นค่าประมาณของ $n\,\frac{\sigma^2}{n}$ หรือ σ^2 นั่นคือ

$$n\, s^2_{\bar{y}} = \frac{n \sum^{k} (\bar{y} - \bar{\bar{y}})^2}{k - 1}$$

$$= \frac{\text{among-sample SS}}{k - 1} \quad \ldots\ldots\ldots\ldots\ldots \quad (2)$$

เป็นค่าประมาณของ σ^2

จากข้อ 9.6 เราทราบแล้วว่า pooled SS หารด้วย pooled d.f. คือ s^2_p เป็นค่าประมาณของ population variance นั่นคือ

$$s^2_p = \frac{\text{pooled SS}}{k\,(n - 1)}$$

$$= \frac{\sum^{kn} (y - \bar{y})^2}{k\,(n - 1)}$$

$$= \frac{\text{within-sample SS}}{k\,(n - 1)} \quad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

เป็นค่าประมาณของ σ^2 เพราะฉะนั้นจึงกล่าวได้ว่าเรโช $\dfrac{n\, s^2_{\bar{y}}}{s^2_p}$ follows the F-distribution with $k-1$ and $k(n-1)$ d.f. จินตนาการเรื่องนี้ปรากฏว่าถูกต้อง ความจริงที่ว่าเรโชนี้ follows the F-distribution อาจแสดงให้เห็นได้โดย sampling experiment ซึ่งมีรายละเอียดที่ได้กล่าวไว้ในบทที่ 4

เราจัด 1,000 samples ซึ่งแต่ละ sample ประกอบด้วย 5 observations ให้เข้าคู่กันเป็น 500 คู่ โดยเอา sample ที่หนึ่งและ sample ที่สองเป็นคู่หนึ่ง sample ที่สามและ sample ที่สี่เป็นคู่หนึ่ง และต่อ ๆ ไปตามลำดับ จากแต่ละคู่ของ samples เหล่านี้เราคำนวณค่าของ among-sample SS และ within-sample SS ไว้ among-sample SS จะมี $k-1$ หรือ 1 d.f. และ within-sample SS จะมี $k\,(n-1)$ หรือ $2(5-1)$ หรือ 8 d.f. ตัวอย่างของ samples 2 คู่ได้ให้ไว้แล้วใน Table 4.2 สำหรับ sample คู่แรก means ของมันคือ 48.8 และ 45.6 ตามลำดับ และค่าของ SS คือ 598.8 และ 237.2 ตามลำดับ general mean $(\bar{\bar{y}})$ คือ $\frac{1}{2}(48.8 + 45.6) = 47.2$

ดังนั้น among-sample SS ซึ่งมี 1 d.f. คือ

$$n\sum^{k}(\bar{y}-\bar{\bar{y}})^2 = 5\left[(48.8-47.2)^2+(45.6-47.2)^2\right]$$

$$= 5\left[(1.6)^2+(-1.6)^2\right]$$

$$= 25.6$$

within-sample SS ซึ่งมี 8 d.f. เป็นผลบวกของค่าของ SS 2 ตัว คือ 598.8 + 237.2 = 836.0 ดังนั้น

$$n\,s_{\bar{y}}^2 = \frac{n\sum^{k}(\bar{y}-\bar{\bar{y}})^2}{k-1}$$

$$= \frac{25.6}{1}$$

$$= 25.6$$

และ

$$s_p^2 = \frac{\sum^{kn}(y-\bar{y})^2}{k(n-1)}$$

$$= \frac{836.0}{8}$$

$$= 104.5$$

variance ratio

$$\frac{n\,s_{\bar{y}}^2}{s_p^2} = \frac{25.6}{104.5}$$

$$= 0.245 \text{ with 1 and 8 d.f.}$$

เราคำนวณค่าของเรโชดังกล่าวนี้ไว้ทุกคู่ของ samples ทั้ง 500 คู่ frequency table ของ 500 variance ratios ได้แสดงไว้ใน Table 12.2

Table 12.2

$F = \dfrac{n\,s_{\bar{y}}^2}{s_p^2}$	observed frequency		theoretical
	f	r.f. (%)	r.f. (%)
0 - 1	332	66.4	65.1
1 - 2	72	14.4	15.4
2 - 3	37	7.4	7.4
3 - 4	16	3.2	4.1
4 - 5	11	2.2	2.5
5 - 6	15	3.0	1.6
6 - 7	5	1.0	1.1
7 - 8	2	0.4	0.7
8 - 9	4	0.8	0.5
9 - 10	1	0.2	0.4
10 - 11	0	0	0.3
11 - 12	1	0.2	0.1
over 12	4	0.8	0.8
total	500	100.0	100.0

theoretical frequency ที่ให้ไว้ใน Table เป็น frequency ของ F-distribution with 1 and 8 d.f. histogram ของ 500 variance ratios และ F-curve with 1 and 8 d.f. ซึ่งเขียนทับกันนั้นได้แสดงไว้ใน Fig. 12.2

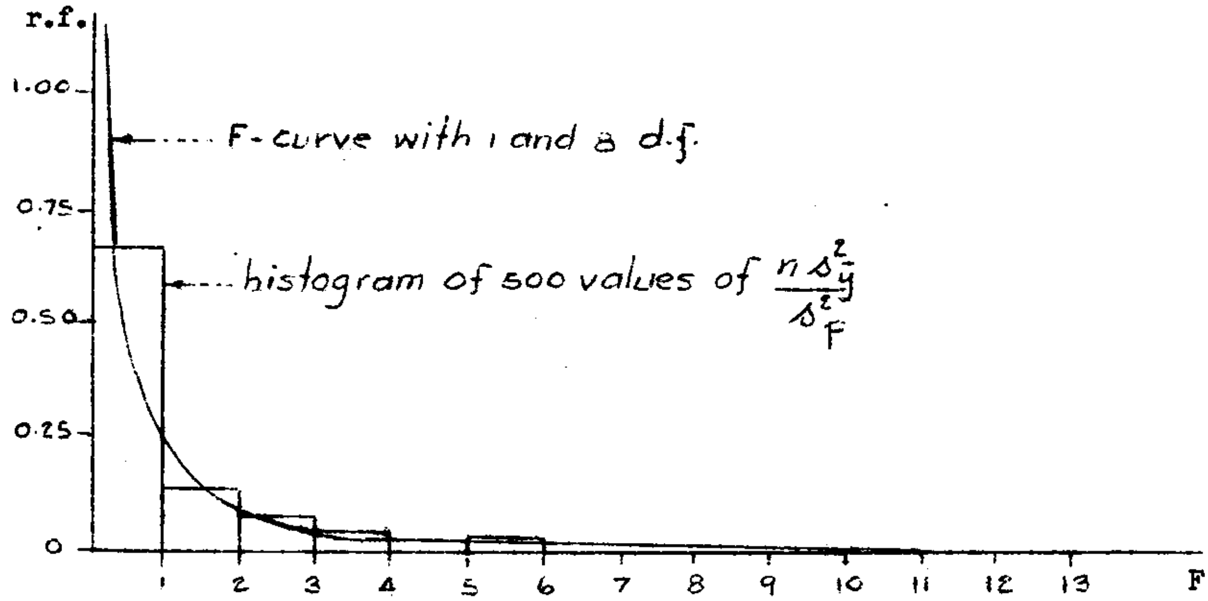


Fig. 12.2

เราจะสังเกต frequency table หรือ histogram อย่างใดอย่างหนึ่งได้ว่า observed frequency และ theoretical frequency นั้นใกล้เคียงกันมาก experiment นี้ได้แสดงให้เห็นความเป็นจริงที่ว่า statistic

$$F = \frac{\dfrac{\text{among-sample SS}}{k-1}}{\dfrac{\text{within-sample SS}}{k(n-1)}}$$

$$= \frac{\dfrac{n\sum^{k}(\bar{y}-\bar{\bar{y}})^2}{k-1}}{\dfrac{\sum^{kn}(y-\bar{y})^2}{k(n-1)}}$$

$$= \frac{n\,s^2_{\bar{y}}}{s^2_p} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

follows the F-distribution with $k-1$ and $k(n-1)$ d.f. การใช้ประโยชน์ของผลที่ได้จากการพิจารณาจะได้กล่าวไว้ในข้อหลัง ๆ

variance ($n\,s_{\bar{y}}^2$) นี้เรียกว่า among-sample mean square และ pooled variance (s_p^2) ของ k samples เรียกว่า within-sample mean square ไม่มี variance ใดของ variances ทั้งสองนี้มีความหมายใหม่เพราะชื่อของมันเลย แต่ที่เราให้ชื่อของมันใหม่ก็เพื่อสะดวกต่อการอ้างอิงและมักจะใช้ MS แทนคำว่า mean square

เรายังเรียก among-sample SS นี้ว่า treatment SS และเรียก within-sample SS ว่า error SS ได้อีกด้วย ชื่อเหล่านี้ได้มาเพราะการใช้ analysis of variance ซึ่งจะได้พิจารณากันต่อไปในข้อ 12.8

ความหมายทางฟิสิกส์ของ d.f. นั้นเราอาจสังเกตได้ใน Table 12.1 b among-sample SS มีค่าเป็น n เท่าของผลบวกของกำลังสองของ ($\bar{y}-\bar{\bar{y}}$) (Equation (3), ข้อ 12.1) ค่าของ ($\bar{y}-\bar{\bar{y}}$) เหล่านี้คือ -1, 3 และ -2 ซึ่งรวมกันแล้วเท่ากับ 0 เพราะฉะนั้นถ้าเราทราบค่าเหล่านี้เพียง 2 ค่าเราย่อมทราบค่าที่สามได้โดยอัตโนมัติ ดังนั้น d.f. ของ among-sample SS คือ 2 กล่าวโดยทั่วไปถ้าเราทราบค่าของ $k-1$ จำนวนใน k จำนวนแล้ว อีกหนึ่งจำนวนที่เหลือก็จะทราบค่าได้โดยอัตโนมัติ เพราะฉะนั้น d.f. ของมันจึงเป็น $k-1$ ส่วน within-sample SS นั้นเป็นผลบวกของกำลังสองของ ($y-\bar{y}$) ของ k samples แต่สำหรับแต่ละ sample นั้น $\sum (y-\bar{y}) = 0$ (Table 12.1b) เพราะฉะนั้นถ้าเราทราบค่าของ ($y-\bar{y}$), $n-1$ จำนวนใน n จำนวนแล้ว ค่าของ ($y-\bar{y}$) อีกหนึ่งจำนวนที่เหลือก็จะทราบได้โดยอัตโนมัติ $\sum^{n} (y-\bar{y})^2$ จึงมี $n-1$ d.f. สำหรับแต่ละ sample ดังนั้น pooled value ($\sum^{kn} (y-\bar{y})^2$) สำหรับ k samples จึงมี $k(n-1)$ d.f.

12.3 Computing Method

กลวิธีของการแตก sum of squares ได้แสดงไว้โดยละเอียดในข้อ 12.1 แล้ว ในข้อนั้นเราจงใจเลือกใช้ตัวอย่างเกี่ยวกับจำนวนเลขที่เป็นเลขหลักหน่วยแทบทั้งหมด ทั้งนี้เพื่อไม่ให้การคำนวณยุ่งยาก วัตถุประสงค์ของข้อนั้นคือการแสดงความหมายของการแตก total SS ออกเป็น among-sample SS และ within-sample SS แต่วิธีทำนั้นยุ่งยากมาก ในข้อนี้จะได้นำวิธีคำนวณค่าของ total SS, among-sample SS, within-sample SS และค่าของ F โดยทางลัดมากล่าวไว้ วิธีคำนวณทางลัดนี้ได้มาจาก identity (Equations (3) และ (4) ข้อ 7.4) คือ

$$SS = \sum^{n} (y - \bar{y})^2$$

$$= \sum y^2 - \frac{(\sum y)^2}{n} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

และวิธีคำนวณก็ทำให้เหมาะกับการใช้เครื่องคำนวณเลขเป็นส่วนใหญ่

notations ต่าง ๆ ที่ใช้ในข้อนี้คงเช่นเดียวกับที่ใช้ในข้อ 12.1 คือ

k = จำนวน samples

n = จำนวน observations ของแต่ละ sample

y = observation

T = ผลบวกของ observations ทั้งหมดของแต่ละ sample (sample total) ได้แก่ $T_1, T_2, \ldots\ldots, T_k$

$\bar{y}$ = sample mean ได้แก่ $\bar{y}_1, \bar{y}_2, \ldots\ldots, \bar{y}_k$

G = ผลบวกของ observations ทั้งหมดของ k samples (grand total)

$\bar{\bar{y}}$ = general mean

เนื่องจาก general mean ($\bar{\bar{y}}$) เป็น mean ของ kn observations ดังนั้น total SS จึงเป็น SS ของ sample รวมที่มี kn observations จาก Equation (1) จะเห็นได้ว่า

$$\text{total SS} = \sum^{kn} (y - \bar{\bar{y}})^2$$

$$= \sum^{kn} y^2 - \frac{G^2}{kn} \quad \ldots\ldots\ldots\ldots\ldots (2)$$

และเพราะว่า $\bar{\bar{y}}$ ก็เป็น mean ของ k sample means ด้วย การใช้ Equation (1) กับ sample means ทำให้เราได้

$$\text{among-sample SS} = n \sum^{k} (\bar{y} - \bar{\bar{y}})^2$$

$$= n \left[\sum^{k} \bar{y}^2 - \frac{(\sum \bar{y})^2}{k} \right]$$

แต่ $\bar{y} = \frac{T}{n}$ เพราะฉะนั้น

$$\text{among-sample SS} = n\left[\sum^{k} \left(\frac{T}{n}\right)^2 - \frac{(\sum \frac{T}{n})^2}{k}\right]$$

$$= n \sum \frac{T^2}{n^2} - n \frac{(\sum T)^2}{kn^2}$$

$$= \frac{\sum T^2}{n} - \frac{(\sum T)^2}{kn}$$

$$= \frac{\sum T^2}{n} - \frac{G^2}{kn} \quad \dots\dots\dots\dots\dots\dots (3)$$

SS ของแต่ละ sample คือ

$$\sum^{n} y^2 - \frac{(\sum y)^2}{n} = \sum^{n} y^2 - \frac{T^2}{n}$$

pooled SS ของ k samples คือผลบวกของค่าของ SS k ค่า นั่นคือ

$$\text{within-sample SS} = \sum^{kn} y^2 - \frac{\sum^{k} T^2}{n} \quad \dots\dots\dots\dots\dots\dots (4)$$

ณ บัดนี้การคำนวณ total SS, among-sample SS และ within-sample SS จะทำได้ง่ายขึ้นโดยเพียงแต่คำนวณเลข 3 จำนวนเท่านั้น คือ

I $\quad \frac{G^2}{kn}$

II $\quad \frac{\sum^{k} T^2}{n}$

III $\quad \sum^{kn} y^2$

ดังนั้นจาก Equations (2), (3) และ (4) จะเห็นได้ว่า

total SS = III - I .... (5)

among-sample SS = II - I .... (6)

within-sample SS = III - II .... (7)

และจาก equations ข้างบนนี้จะเห็นได้ชัดอีกว่า total SS เท่ากับ among-sample SS บวกด้วย within-sample SS

หลักอันเป็นพื้นฐานของวิธีคำนวณทางลัดคือการใช้ totals แทน means ในทุกขั้นตอนของการคำนวณ เราใช้ grand total (G) แทน general mean ($\bar{\bar{y}}$) และใช้ sample total (T) แทน sample mean ($\bar{y}$) ส่วน observation (y) คงไว้ตามเดิม ถ้าเรายึดหลักนี้ไว้ก็จะจำได้ง่ายว่า SS ตัวไหนประกอบด้วย combination อะไรของ 3 จำนวนคือ I, II และ III total SS ซึ่งเป็น sum of squares ของการเบี่ยงเบน ($y - \bar{\bar{y}}$) นั้นเท่ากับ III - I จำนวน III เกี่ยวข้องกับ y และจำนวน I เกี่ยวข้องกับ G ซึ่งเกี่ยวโยงไปถึง $\bar{\bar{y}}$ among-sample SS หรือ treatment SS ซึ่งเป็น sum of squares ของการเบี่ยงเบน ($\bar{y} - \bar{\bar{y}}$) นั้นเท่ากับ II - I จำนวน II เกี่ยวข้องกับ T ซึ่งเกี่ยวโยงไปถึง $\bar{y}$ และจำนวน I เกี่ยวข้องกับ G ซึ่งเกี่ยวโยงไปถึง $\bar{\bar{y}}$ within-sample SS หรือ error SS ซึ่งเป็น sum of squares ของการเบี่ยงเบน ($y - \bar{y}$) นั้นเท่ากับ III - II จำนวน III เกี่ยวข้องกับ y และจำนวน II เกี่ยวข้องกับ T ซึ่งเกี่ยวโยงไปถึง $\bar{y}$ ถ้าเราจำความสัมพันธ์เหล่านี้ได้แล้วการคำนวณทางลัดก็จะมีความหมายและง่ายต่อการจำด้วย

ใน analysis of variance เราจะต้องคำนวณค่าของ sample totals (T) และ grand total (G) เสียก่อน การคำนวณที่เหลืออาจจัดทำเป็นรูปตารางดังแสดงไว้ใน Table 12.3 a ให้สังเกตว่าในการคำนวณ total SS และ components ต่าง ๆ ของมันนั้นเราไม่ต้องการ sample means ($\bar{y}$) และ general mean ($\bar{\bar{y}}$) เลย

Table 12.3 a

preliminary calculations				
1	2	3	4	5
type of total	total of squares	No. of items squared	No. of observations per squared item	total of squares per observation $2 \div 4$
grand	G^2	1	kn	I
sample	ΣT^2	k	n	II
observation	Σy^2	kn	1	III
analysis of variance				
source of variation	sum of squares SS	degrees of freedom	mean square MS	F
among-sample	II - I	k - 1	$ns_{\bar{y}}^2$	$\frac{ns_{\bar{y}}^2}{s_p^2}$
within-sample	III - II	kn - k	s_p^2	
total	III - I	kn - 1		

ค่าของ

$$\Sigma T^2 = T_1^2 + T_2^2 + \ldots\ldots\ldots + T_k^2$$

จะหาออกมาได้โดยการใช้เครื่องคำนวณเลขครั้งเดียวติดต่อกันและ Σy^2 ก็คำนวณได้โดยวิธีเดียวกัน

ค่าของ SS ทั้งสามในครึ่งล่างของ Table 12.3 a ได้มาจากการเอา item หนึ่งไปลบ item อื่น ๆ ใน column 5 ของครึ่งบนของ table นี้ d.f. ต่าง ๆ ก็อาจจะหาได้โดยการลบกันของ items ต่าง ๆ ใน column 3 ของครึ่งบนของ table ตัวอย่างเช่น total SS จะได้จากการเอา I

ไปลบ III และ d.f. ของมันคือ kn - 1 ก็ได้จากการเอา item แรกคือ 1 ไปลบ item ที่สามคือ kn ใน column 3

total SS และ components ของมันของตัวอย่างที่ให้ไว้ใน Table 12.1 a นั้นเราได้ทราบค่าของมันแล้วในข้อ 12.1 ในขณะนี้เราจะใช้วิธีคำนวณทางลัดกับ 15 observations เดียวกันนี้ รายละเอียดของการคำนวณได้แสดงไว้ใน Table 12.3 b

Table 12.3 b

preliminary calculations				
1	2	3	4	5
type of total	total of squares	No. of items squared	No. of observations per squared item	total of squares per observation 2 ÷ 4
grand sample observation	8,100 3,050 688	1 3 15	15 5 1	540 610 688

analysis of variance				
source of variation	sum of squares SS	degrees of freedom	mean square MS	F
among-sample within-sample total	70 78 148	2 12 14	35.0 6.5	5.38

ควรสังเกตว่าค่าของ SS ทั้งสามซึ่งได้จากการคำนวณทางลัดที่ให้ไว้ใน Table 12.3 b เป็นค่าเดียวกันกับที่คำนวณได้โดยวิธีธรรมดาในข้อ 12.1 สำหรับตัวอย่างที่ได้กล่าวมานี้เรามองไม่ค่อยเห็นข้อดีของวิธีคำนวณทางลัดเพราะมีจำนวน observations ทั้งหมดน้อย และ observations แทบทั้งหมดเป็นเลขหลักหน่วย แต่เมื่อเราเปรียบเทียบ 2 วิธีนี้ดูในการปฏิบัติทั่วไปแล้วจะเห็นว่าวิธีคำนวณทางลัดนี้มีประโยชน์สมกับชื่อของมันอย่างแท้จริง

12.4 Variance Components and Models

ในข้อ 12.2 เราทราบแล้วว่า ถ้า k samples ซึ่งมี size n ได้มาจาก population เดียวกัน variance ระหว่าง k sample means คือ

$$s_{\bar{y}}^2 = \frac{\sum^{k} (\bar{y} - \bar{\bar{y}})^2}{k - 1}$$

เป็นค่าประมาณของ $\frac{\sigma^2}{n}$ ค่าเฉลี่ยของ $s_{\bar{y}}^2$ ของ all possible sets ของ k samples ซึ่งมี size n ที่ได้มาจาก population เดียวกันจะเท่ากับ $\frac{\sigma^2}{n}$ (Theorem 7.2) จึงเป็นที่น่าสนใจที่จะรู้ถึงค่าเฉลี่ยของ $s_{\bar{y}}^2$ ถ้า k samples นั้นได้มาจาก populations ต่าง ๆ ซึ่งมี means ต่างกันแต่มี variance อย่างเดียวกัน เราอาจใช้ sampling experiment ของข้อ 12.2 พิจารณาค่าเฉลี่ยนี้ได้

ใน sampling experiment ซึ่ง k = 2, n = 5 นี้ samples ทั้งหมด รวม 1,000 samples ได้มาจาก normal population เดียวกันที่มี mean เท่ากับ 50 และ variance เท่ากับ 100 อย่างไรก็ดี samples ต่าง ๆ นั้นอาจได้มาจาก populations ต่างกัน เช่น samples แรกของแต่ละคู่ของ samples 500 คู่ อาจได้มาจาก population หนึ่งซึ่งมี mean เท่ากับ 60 และ samples ที่สองอาจได้มาจาก population หนึ่งซึ่งมี mean เท่ากับ 40 ถ้าเอา 10 บวก observation ทุกตัวของ 5 observations ของ sample แรกของแต่ละคู่ของ samples และเอา 10 ลบ observation ทุกตัวของ 5 observations ของ sample ที่สองของแต่ละคู่ของ samples แล้ว samples ต่าง ๆ ซึ่งมี observations เปลี่ยนไปแล้วนี้จะกลายเป็น samples ที่ได้มาจาก 2 normal populations ซึ่งมี means เท่ากับ 60 และ 40 ตามลำดับ แต่ variances ทั้งสองยังคงเท่ากับ 100 ตามเดิม (Theorem 2.4 a) ผลที่เกิดขึ้นกับ variance ในระหว่าง sample means เนื่องจากการเปลี่ยนค่าของ observations จะเห็นได้จากตัวอย่างของ 4 samples ที่ได้ให้ไว้ใน Table 4.2 means ของ samples คู่แรกคือ 48.8

และ 45.6 ตามลำดับ จากการเปลี่ยนค่าของ observations และจาก population means ที่เปลี่ยนไปจะทำให้ sample means ทั้งสองเปลี่ยนไปเป็น 58.8 และ 35.6 ตามลำดับ แต่ general mean ($\bar{\bar{y}} = 47.2$) นั้นจะไม่เปลี่ยนแปลง เมื่อ 2 samples ได้มาจาก population เดียวกันซึ่งมี mean เท่ากับ 50 variance ของ sample means คือ

$$s^2_{\bar{y}} = \frac{(48.8 - 47.2)^2 + (45.6 - 47.2)^2}{2 - 1}$$

$$= 5.12$$

แต่ถ้า sample แรกได้มาจาก population ซึ่งมี mean เท่ากับ 60 และ sample ที่สองได้มาจาก population ซึ่งมี mean เท่ากับ 40 variance ของ sample means จะเป็น

$$s^2_{\bar{y}} = \frac{(58.8 - 47.2)^2 + (35.6 - 47.2)^2}{2 - 1}$$

$$= 269.12$$

เพราะความแปรผันใน population means นี่เองทำให้ variance ของ sample means มีค่าเพิ่มขึ้นจาก 5.12 เป็น 269.12 โดยมีส่วนเพิ่มสุทธิถึง 264 โดยเฉลี่ยแล้วส่วนที่เพิ่มขึ้นนี้เท่ากับ variance ของ 2 population means โดยทั่วไปค่าเฉลี่ยของ $s^2_{\bar{y}}$ สำหรับ all possible sets ของ k samples ที่ได้มาจาก k populations นั้นจะเท่ากับ $\frac{\sigma^2}{n} + \sigma^2_{\mu}$ ในเมื่อ σ^2_{μ} ซึ่งเป็น variance ของ k population means มีค่าดังนี้

$$\sigma^2_{\mu} = \frac{(\mu_1 - \bar{\mu})^2 + (\mu_2 - \bar{\mu})^2 + \ldots\ldots + (\mu_k - \bar{\mu})^2}{k - 1} \quad \ldots\ldots\ldots\ldots (2)$$

notation $\bar{\mu}$ ใน equation ข้างบนเป็น mean ของ k population means คือ μ_1, μ_2, ..., μ_k เราอาจใช้ sampling experiment แสดงให้เห็นความเป็นจริงนี้ได้ กล่าวคือ 1,000 samples ของ experiment ได้มาจาก population เดียวกันซึ่งมี mean เท่ากับ 50 variance ของ sample means คือ

$$s_{\bar{y}}^2 = \frac{\sum^{k} (\bar{y} - \bar{\bar{y}})^2}{k - 1}$$

$$= \frac{(\bar{y}_1 - \bar{\bar{y}})^2 + (\bar{y}_2 - \bar{\bar{y}})^2}{2 - 1}$$

$$= (\bar{y}_1 - \frac{\bar{y}_1 + \bar{y}_2}{2})^2 + (\bar{y}_2 - \frac{\bar{y}_1 + \bar{y}_2}{2})^2$$

$$= (\frac{\bar{y}_1 - \bar{y}_2}{2})^2 + (\frac{\bar{y}_2 - \bar{y}_1}{2})^2$$

$$= \frac{1}{2} (\bar{y}_1 - \bar{y}_2)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

ถ้า 2 samples ของแต่ละคู่ของ samples 500 คู่ได้มาจาก 2 populations ซึ่งมี means เท่ากับ 60 และ 40 แล้ว ผลที่เกิดขึ้นคือ $\bar{y}_1$ จะมีค่าเพิ่มขึ้น 10 และ $\bar{y}_2$ จะมีค่าลดลง 10 ทำให้ผลต่าง $(\bar{y}_1 - \bar{y}_2)$ มีค่าเพิ่มขึ้น 20 variance ของ sample means ใหม่ทั้งสองจะเป็น

$$\frac{1}{2} \left[(\bar{y}_1 - \bar{y}_2) + 20 \right]^2 = \frac{1}{2} \left[(\bar{y}_1 - \bar{y}_2)^2 + 40(\bar{y}_1 - \bar{y}_2) + 400 \right]$$

$$= \frac{(\bar{y}_1 - \bar{y}_2)^2}{2} + 20(\bar{y}_1 - \bar{y}_2) + 200$$

$$= s_{\bar{y}}^2 + 20(\bar{y}_1 - \bar{y}_2) + 200$$

ผลที่เกิดขึ้นจาก population means เป็น 60 และ 40 แทนที่จะเท่ากันก็คือ variance ของ sample means คู่หนึ่งจะมีค่าเพิ่มขึ้นอีก $20(\bar{y}_1 - \bar{y}_2) + 200$ ค่าที่เพิ่มขึ้นนี้จะเปลี่ยนไปต่าง ๆ กันตามคู่ต่าง ๆ ของ samples เนื่องจาก $\bar{y}_1$ และ $\bar{y}_2$ มีค่าเปลี่ยนไปตาม samples นั้นเอง ตาม Theorem 10.1 a mean ของผลต่างทั้งหมดคือ

$$\mu_{\bar{y}_1 - \bar{y}_2} = \mu_1 - \mu_2$$
$$= 50 - 50$$
$$= 0$$

เพราะฉะนั้นค่าที่เพิ่มขึ้นโดยเฉลี่ยของ $s^2_{\bar{y}}$ คือ 200 และ variance ของ 2 population means ก็เท่ากัน

$$\sigma^2_{\mu} = \frac{(60 - 50)^2 + (40 - 50)^2}{2 - 1}$$
$$= 200$$

ด้วย ที่กล่าวมานี้ได้แสดงให้เห็นจริงแล้วว่าค่าเฉลี่ยของ $s^2_{\bar{y}}$ ของ all possible sets ของ k samples ซึ่งแต่ละ sample ประกอบด้วย n observations ที่ได้จาก k populations ซึ่งมี means ต่าง ๆ กันนั้นจะเท่ากับ $\frac{\sigma^2}{n} + \sigma^2_{\mu}$ เนื่องจาก among-sample MS เท่ากับ $ns^2_{\bar{y}}$ (Equation (2) ข้อ 12.2) ดังนั้น among-sample MS โดยเฉลี่ยจะเท่ากับ $n(\frac{\sigma^2}{n} + \sigma^2_{\mu})$ หรือ $\sigma^2 + n\sigma^2_{\mu}$

เราทราบแล้วว่าการเอา 10 บวกเข้ากับหรือลบออกจาก observation แต่ละตัวนั้นไม่มีผลกระทบกระเทือนถึง sample variance แต่อย่างใด (Theorem 2.4 a) ดังนั้นจึงไม่กระทบกระเทือนถึง pooled variance หรือ within-sample MS

การพิจารณาที่แล้วมาอาจสรุปเป็น equations ได้ดังนี้

(ก) average of among-sample MS = $\sigma^2 + n\sigma^2_{\mu}$ ....(4)

(ข) average of within-sample MS = σ^2 ....(5)

มีการอธิบายเกี่ยวกับชุดของ k population means คือ $\mu_1, \mu_2, \ldots\ldots\ldots, \mu_k$ ได้ 2 อย่าง อย่างหนึ่งคือ population means เหล่านั้นมีค่าคงที่ อีกอย่างหนึ่งคือ population means เหล่านั้นมีค่าเปลี่ยนจาก samples ชุดหนึ่งไปยังชุดอื่น ๆ ตัวอย่างเช่นมีฝูงวัวอยู่ 4 ฝูง (k = 4) จากแต่ละฝูงเราหา sample ออกมาฝูงละหนึ่ง sample ซึ่งประกอบด้วยวัว 20 ตัว (n = 20) แล้วหาน้ำหนัก

ของวัวแต่ละตัวไว้ ถ้าความสนใจของเรามุ่งเฉพาะแต่วัว 4 ฝูงนี้เท่านั้น ฝูงวัว 4 ฝูงนี้ก็เป็น 4 populations และน้ำหนักเฉลี่ยของวัวในฝูงคือ μ_1, μ_2, μ_3 และ μ_4 ถ้าเราหา samples ซ้ำอีกก็จะต้องหาออกมาจาก 4 populations เดียวกันนี้ ในทางตรงกันข้ามเราอาจพิจารณาว่าฝูงวัว 4 ฝูงนี้เป็นหนึ่ง sample ของฝูงวัวหลายฝูง ชุดของ 4 population means คือ μ_1, μ_2, μ_3 และ μ_4 ก็จะเป็นเพียงหนึ่ง sample ของ population means จำนวนมาก ถ้าเราหา samples ซ้ำอีกก็จะต้องหาออกมาจาก populations อื่น ๆ อีก 4 populations การอธิบายตอนแรกซึ่งเกี่ยวกับ k population means เรียกว่า linear hypothesis model หรือ fixed model of analysis of variance ส่วนการอธิบายตอนหลังซึ่งเกี่ยวกับ k population means เรียกว่า component of variance model หรือ random variable model of analysis of variance ในบทที่ 12 นี้จะพิจารณาเฉพาะ linear hypothesis model เท่านั้น กล่าวคือ ชุดต่าง ๆ ของ k samples ได้มาจาก k populations เดียวกัน เราอาจแสดงให้เห็น model นี้ได้โดย sampling experiment ซึ่งมีเพียง 2 populations ที่มี means เท่ากับ 60 และ 40 sample แรกของ samples ทุกคู่ได้มาจาก population ที่มี mean เท่ากับ 60 และ sample ที่สองของ samples ทุกคู่ได้มาจาก population ที่มี mean เท่ากับ 40 กล่าวได้อีกอย่างหนึ่งว่า samples คู่ต่าง ๆ เหล่านี้ได้มาจาก populations คู่เดียวกัน แต่ละ observation อาจแสดงออกในรูปสัญลักษณ์โดย equation ต่อไปนี้

$$y = \bar{\mu} + (\mu - \bar{\mu}) + (y - \mu) \;.............(6)$$

ในเมื่อ:

y = observation

$\bar{\mu}$ = mean ของ k population means

μ = mean ของ population ที่เอา y ออกมา

จาก sampling experiment เรามี

$$\mu_1 = 60$$

$$\mu_2 = 40$$

$$\bar{\mu} = \frac{1}{2}(60 + 40) = 50$$

ดังนั้น observation ที่มีค่า 36 ซึ่งได้มาจาก population ที่สองคือ

$$36 = 50 + (40 - 50) + (36 - 40)$$
$$= 50 + (-10) + (-4)$$

ให้สังเกตว่า $(\mu - \bar{\mu})$ เท่ากับ 10 สำหรับ population แรกและเท่ากับ -10 สำหรับ population ที่สอง และจำนวนเลขทั้งสองนี้จะไม่เปลี่ยนไปอย่างใดแม้ว่าเราจะหา samples ต่าง ๆ ซ้ำจาก 2 populations นี้อีก

ค่าของ $\bar{\mu}$ ถูกประมาณโดย general mean ($\bar{\bar{y}}$), $(\mu - \bar{\mu})$ ซึ่งเรียกว่า treatment effect จะถูกประมาณโดย $(\bar{y} - \bar{\bar{y}})$ และ $(y - \mu)$ ซึ่งเรียกว่า error จะถูกประมาณโดย $(y - \bar{y})$ จากผลอันนี้เราจึงเรียก among-sample SS ซึ่งเป็น sum of squares of the estimated treatment effects ว่า treatment SS และเรียก within-sample SS ซึ่งเป็น sum of squares of the estimated errors ว่า error SS

12.5 Test of Hypothesis-Procedure

ข้อต่าง ๆ ที่แล้วมาเกี่ยวกับ deductive relation ระหว่าง k populations และ samples ต่าง ๆ ของมัน ในข้อนี้เรื่องที่จะต้องพิจารณาคือการหา inductive inferences ที่เกี่ยวกับ k population means จาก k samples

ได้แสดงไว้ในข้อ 12.2 แล้วว่าถ้า all possible sets ของ k samples ซึ่งแต่ละ sample ประกอบด้วย n observations ได้มาจาก normal population เดียวกันแล้ว statistic

$$F = \frac{\text{among-sample MS}}{\text{within-sample MS}}$$

$$= \frac{\text{treatment MS}}{\text{error MS}}$$

follows the F-distribution with $k-1$ and $k(n-1)$ d.f. และยังได้แสดงให้เห็นต่อไปอีกว่า among-sample MS เป็นค่าประมาณของ $\sigma^2 + n\sigma^2_{\mu}$ (Equation (4) ข้อ 12.4) และ within-sample MS เป็นค่าประมาณของ σ^2 ในเมื่อ σ^2 เป็น variance รวมของ k populations และ σ^2_{μ}

เป็น variance ของ k population means (Equation(5) ข้อ 12.4) ผลที่ได้เหล่านี้อาจใช้ในการทดสอบสมมติฐานว่า k population means มีค่าเท่ากันถ้า k populations นั้นเป็น normal populations และมี variance (σ^2) อย่างเดียวกัน

เมื่อ k random samples ซึ่งแต่ละ sample ประกอบด้วย n observations ได้มาจาก k populations ค่าของ F ที่ใช้ทดสอบสมมติฐานว่า k population means มีค่าเท่ากันนั้นอาจคำนวณออกมาได้ (ข้อ 12.3) ขนาดของค่าของ F ทำให้เราสามารถตัดสินใจได้ว่าจะยอมรับหรือปฏิเสธสมมติฐาน ถ้า F มีค่าใกล้เคียง 1 ย่อมชี้ให้เห็นว่า $\sigma^2 + n\sigma_\mu^2 = \sigma^2$ หรือ $\sigma_\mu^2 = 0$ หรือพูดให้ชัดยิ่งขึ้นก็คือ k population means มีค่าเท่ากันหรือ treatment effects ทั้งหมดมีค่าเท่ากับศูนย์ ถ้า F มีค่ามากเกินไปย่อมชี้ให้เห็นว่า $\sigma^2 + n\sigma_\mu^2 > \sigma^2$ หรือ $\sigma_\mu^2 > 0$ หรือพูดให้ชัดยิ่งขึ้นก็คือ k population means มีค่าไม่เท่ากันหรือ treatment effects ไม่ทั้งหมดมีค่าเท่ากับศูนย์ เนื่องจาก variance (σ_μ^2) ของ k population means มีค่าลบไม่ได้ ดังนั้น $\sigma^2 + n\sigma_\mu^2$ ก็จะน้อยกว่า σ^2 ไม่ได้ เพราะฉะนั้น <u>สมมติฐานจะถูกปฏิเสธเพราะ F มีค่ามากเกินไปเท่านั้น แต่จะไม่ถูกปฏิเสธเพราะ F มีค่าน้อยเกินไปเลย</u> กล่าวได้อีกอย่างหนึ่งว่า <u>F-test นี้เป็น one-tailed test และ critical region จะอยู่ที่หางข้างขวาของ F-distribution</u> เรื่อง F-test เป็น one-tailed test นี้ อาจทราบได้โดยสามัญสำนึก กล่าวคือ ค่าของ F น้อยที่สุดคือศูนย์ แต่ถ้า F มีค่าเท่ากับศูนย์ among-sample SS ก็จะเท่ากับศูนย์หรือ k sample means มีค่าเท่ากัน ความจริง k sample means มีค่าเท่ากันไม่ใช่มูลฐานสำหรับปฏิเสธสมมติฐานว่า k population means มีค่าเท่ากันอย่างแน่นอน

เราเคยใช้ F-test มาแล้วในข้อ 9.4 เพื่อทดสอบสมมติฐานว่า $\sigma_1^2 = \sigma_2^2$ จำนวน ($\sigma^2 + n\sigma_\mu^2$) และ σ^2 ใน F-test ที่กำลังพูดถึงนี้ก็เช่นเดียวกับ σ_1^2 และ σ_2^2 ของ F-test อันก่อนนั้นเอง ดังนั้นใน F-test นี้ สมมติฐานคือ

$$\sigma^2 + n\sigma_\mu^2 = \sigma^2$$

นั่นคือ

$$\sigma_\mu^2 = 0$$

ควรสังเกตว่าถ้าสมมติฐานเป็นจริงคือ σ_{μ}^2 เท่ากับ 0 แล้ว $\sigma^2 + n\sigma_{\mu}^2$ ต้องเท่ากับ σ^2 โดยไม่คำนึงว่า sample size (n) จะใหญ่เท่าไร ถ้าใช้ 5% significance level 5% ของ all possible sets ของ k samples จะนำไปสู่ข้อยุติผิดได้ว่า k population means มีค่าไม่เท่ากัน โดยไม่คำนึงถึง sample size แต่ถ้าสมมติฐานไม่เป็นจริง คือ $\sigma_{\mu}^2 > 0$ แล้ว $n\sigma_{\mu}^2$ อาจมีค่ามากเท่าใดก็ได้ ถ้า n มีค่ามากพอ ดังนั้นเราอาจทำให้ $n\sigma_{\mu}^2$ มีค่ามากเท่าใดก็ได้ตามใจชอบโดยการเพิ่ม sample size (n) ถ้าเราทำให้ $\sigma^2 + n\sigma_{\mu}^2$ โตกว่า σ^2 มาก ค่าของ F โดยเฉลี่ยจะโตกว่า 1 มาก และสมมติฐานมักจะถูกปฏิเสธ เพราะฉะนั้นข้อดีของ samples ใหญ่ก็คือการทำให้เกิดการปฏิเสธสมมติฐานที่ไม่เป็นจริงได้ดีขึ้น หรือลดความน่าจะเป็นของการเกิด Type II error ลงไป ตัวอย่างเช่น ถ้า population means คือ 30, 40, 50, 60 และ 70 ตามลำดับ variance ของแต่ละ population ของ 5 populations(σ^2) เท่ากับ 100 และ variance ของ 5 population means คือ

$$\sigma_{\mu}^2 = \frac{(30-50)^2+(40-50)^2+(50-50)^2+(60-50)^2+(70-50)^2}{5-1}$$

$$= 250$$

ดังนั้น $\sigma^2 + n\sigma_{\mu}^2 = 100 + n(250)$

ถ้า $n = 3$, $\sigma^2 + n\sigma_{\mu}^2$ จะเป็น 8.5 เท่าของ σ^2 และถ้า $n = 6$, $\sigma^2 + n\sigma_{\mu}^2$ จะเป็น 16 เท่าของ σ^2 n ยิ่งโตขึ้นค่าของ F โดยเฉลี่ยจะยิ่งมากขึ้น ดังนั้นสมมติฐานที่ไม่เป็นจริงก็มักจะถูกปฏิเสธได้มากขึ้น กล่าวได้อีกอย่างหนึ่งว่า ถ้า significance level คงเดิมการเพิ่มขนาดของ sample จะลดความน่าจะเป็นของการเกิด Type II error ลงไป อย่างไรก็ดีถ้า σ_{μ}^2 ยิ่งมีค่าน้อยลงก็จะยิ่งต้องการ sample size โตขึ้นเพื่อลดความน่าจะเป็นของการเกิด Type II error ตัวอย่างเช่น ถ้า population means คือ 40, 45, 50, 55 และ 60 ตามลำดับ และ variance ของ population means คือ

$$\sigma_{\mu}^2 = \frac{(40-50)^2+(45-50)^2+(50-50)^2+(55-50)^2+(60-50)^2}{5-1}$$

$$= 62.5$$

ดังนั้น เราจะต้องการ $n = 24$ ที่จะทำให้

$$\sigma^2 + n\sigma_\mu^2 = 100 + 24\,(62.5)$$
$$= 1,600$$

จำนวน $n\sigma_\mu^2$ ยังคงเป็นอย่างเดียวกันทั้ง 2 กรณีไม่ว่าในกรณี $\sigma_\mu^2 = 250$ และ $n = 6$ หรือในกรณี $\sigma_\mu^2 = 62.5$ และ $n = 24$ สิ่งซึ่งมีอิทธิพลในความน่าจะเป็นของการเกิด Type II error คือผลคูณของ n กับ σ_μ^2 ดังนั้น ถ้า k population means มีค่าต่างกันมาก samplesขนาดเล็กก็อาจทำให้เราปฏิเสธสมมติฐานที่ว่า k population means มีค่าเท่ากันได้ แต่ถ้า k population means มีค่าเกือบจะเท่ากันแล้วเราจะปฏิเสธสมมติฐานเดียวกันนี้ได้ก็โดยการใช้ samples ขนาดใหญ่เท่านั้น

ควรจำไว้ว่า statistic

$$F = \frac{\text{among-sample MS}}{\text{within-sample MS}}$$
$$= \frac{\text{treatment MS}}{\text{error MS}}$$

follows the F-distribution เฉพาะเมื่อสมมติฐานเป็นจริง คือ $\sigma_\mu^2 = 0$ หรือ $\sigma^2 + n\sigma_\mu^2 = \sigma^2$ เท่านั้น ถ้า $\sigma^2 + n\sigma_\mu^2$ โตเป็น 16 เท่าของ σ^2 เช่นนี้ ค่าของ F โดยเฉลี่ยก็จะโตเป็น 16 เท่าของค่าของ F จริงและจะไม่ follow the F-distribution จริง F-distribution ที่ผิดส่วนและ F-distribution จริง with 4 (คือ $k - 1 = 5 - 1 = 4$) และ 25 (คือ $kn - k = 30 - 5$) d.f. ได้แสดงไว้ใน Fig. 12.5

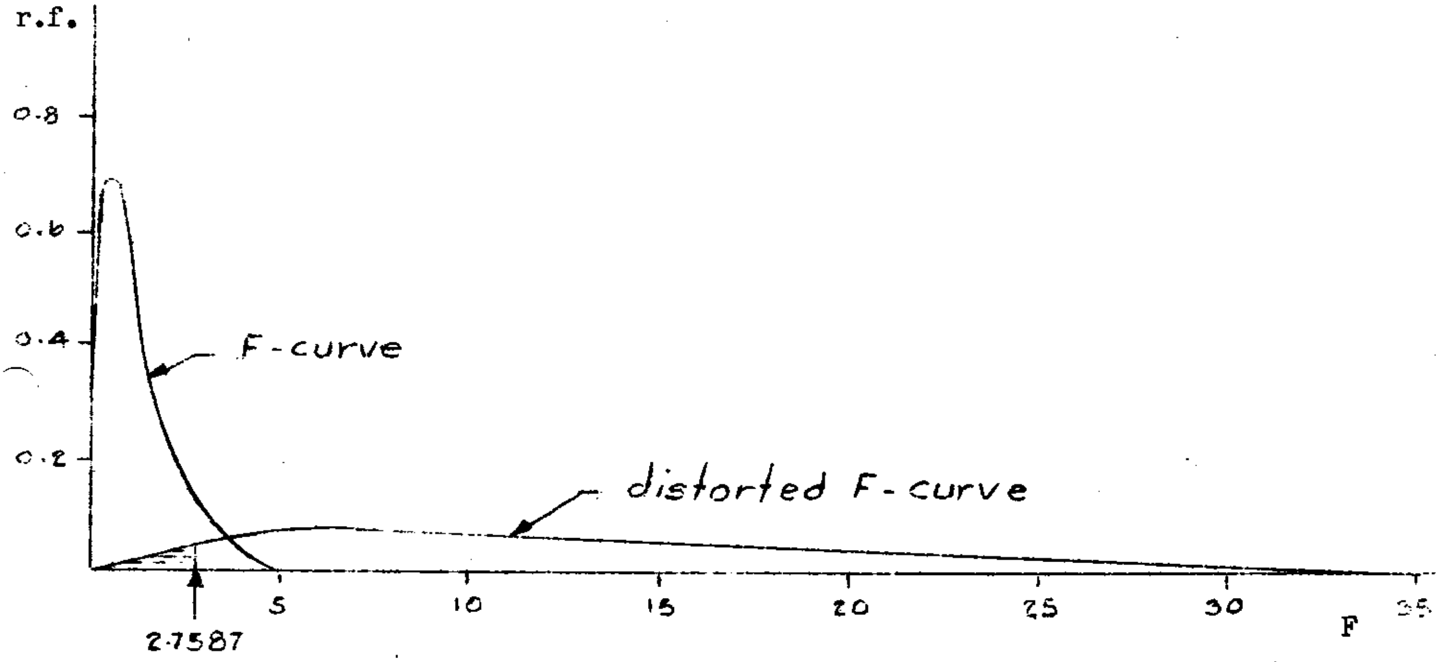


Fig. 12.5

critical region สำหรับ 5% significance level อยู่ที่ F > 2.7587 เมื่อค่าของ F < 2.7587 เรายอมรับสมมติฐานว่า k population means มีค่าเท่ากัน ดังนั้นจึงเกิด Type II error ขึ้น ใน Fig. 12.5 นั้นจะเห็นได้ว่าความน่าจะเป็นของการเกิด Type II error ซึ่งแสดงโดยเนื้อที่ซึ่งขีดเส้นต่ำของ F-distribution ที่ขีดส่วนนั้นเล็กมาก

วิธีการทดสอบสมมติฐานดังที่ได้แสดงไว้โดยตัวอย่างใน Table 12.1 a นั้น อาจสรุปได้ดังต่อไปนี้

1. Hypothesis : 3 population means มีค่าเท่ากัน นั่นคือ

$$\mu_1 = \mu_2 = \mu_3$$

สมมติฐานนี้ยังอาจแสดงได้อีกรูปหนึ่ง คือ

$$\sigma^2_{\mu} = 0$$

2. Alternative hypothesis : alternative hypothesis คือ 3 population means มีค่าไม่เป็นอย่างเดียวกันทั้งหมด นั่นคือ

$$\sigma^2_{\mu} > 0$$

3. Assumptions : 3 samples เป็น random samples ซึ่งได้มาจาก 3 normal populations ที่มี variance อย่างเดียวกัน

4. Level of significance : ใช้ 5% significance level

5. Critical region : critical region อยู่ที่

$F > 3.8853$ (เนื่องจาก $k = 3$ และ $n = 5$, d.f. ของ F จึงเป็น 2 และ 12 F-test นี้เป็น one-tailed test เพราะฉะนั้นเราจึงใช้ 5% table สำหรับ 5% significance level)

6. Computation of F : รายละเอียดของการคำนวณค่าของ F ได้แสดงไว้ใน Table 12.3 b ค่าของ F คือ 5.38 with 2 and 12 d.f.

7. Conclusion : เนื่องจากค่าของ F อยู่ภายใน critical region จึงปฏิเสธสมมติฐาน conclusion คือ 3 populations ไม่มี means อย่างเดียวกัน (ถ้าค่าของ $F < 3.8853$ conclusion จะเป็น 3 populations มี means อย่างเดียวกัน)

12.6 Relation Between t - Distribution and F - Distribution

ถ้าเรามี 2 samples เราอาจใช้ t - test ในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันได้ (ข้อ 10.6) แต่เราอาจใช้ F - test (analysis of variance with $k = 2$) เพื่อความประสงค์อย่างเดียวกันนี้ได้อีกด้วย จึงเป็นที่น่าสนใจว่าการทดสอบด้วยวิธีทั้งสองนี้จะให้ conclusion อย่างเดียวกันเสมอหรือไม่ ใน t - test นั้น ถ้า $n_1 = n_2 = n$

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{2s_p^2}{n}}} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

with $(n_1 + n_2 - 2)$ หรือ $2n - 2$ d.f. แต่ใน analysis of variance

$$F = \frac{ns_{\bar{y}}^2}{s_p^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

(Equation (4), ข้อ 12.2) with 1 ($k = 2$, $k - 1 = 1$) and $2(n - 1)$ d.f. โดยการใช้ความรู้ทางพีชคณิตเราอาจแสดงให้เห็นได้ว่า $t^2 = F$ จำนวน $s_{\bar{y}}^2$ นี้เท่ากับ $\frac{(\bar{y}_1 - \bar{y}_2)^2}{2}$ (Equation (3), ข้อ 12.4) ดังนั้น

$$F = \frac{ns_{\bar{y}}^2}{s_p^2}$$

$$= \frac{s_{\bar{y}}^2}{\frac{s_p^2}{n}}$$

$$= \frac{\frac{(\bar{y}_1 - \bar{y}_2)^2}{2}}{\frac{s_p^2}{n}}$$

$$= \frac{(\bar{y}_1 - \bar{y}_2)^2}{\frac{2s_p^2}{n}}$$

$$= \quad t^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

และ t มี 2(n-1) d.f. แต่ F จะมี 1 and 2(n-1) d.f.

t-distribution เป็น symmetrical ซึ่งกึ่งกลางของ distribution มีค่าเป็นศูนย์ (Fig. 8.1) ส่วน F-distribution เป็น asymmetrical ซึ่ง limit ทางต่ำมีค่าเป็นศูนย์ (Fig. 9.1) ค่าของ t อาจเป็นค่าลบ, ศูนย์, หรือค่าบวกอย่างใดอย่างหนึ่งก็ได้ แต่ค่าของ t^2 ต้องเป็นศูนย์หรือค่าบวกเท่านั้นเพราะกำลังสองของค่าลบจะเป็นค่าบวก สำหรับ t with 10 d.f. นั้น 2.5 % ของค่าของ t ทั้งหมดจะน้อยกว่า -2.2281 และ 2.5 % ของค่าของ t ทั้งหมดจะมากกว่า 2.2281 แต่ $(-3)^2$ และ 3^2 ต่างก็เท่ากับ 9 และมากกว่า $(2.2281)^2$ เพราะฉะนั้น 5% ของค่าของ t^2 ทั้งหมดจึงมากกว่า $(2.2281)^2$ หรือ 4.96 ค่า 4.96 นี้คือ 5% point ของ F with 1 and 10 d.f. ดังนั้น F-distribution with 1 and ν d.f. ก็คือการยกกำลังสองของ t-distribution with ν d.f. หางสองข้างของ t จะถูกพับเข้ามาเป็นหางข้างขวาของ F และส่วนกลางของ t-distribution จะกลายเป็นหางข้างซ้ายของ F-distribution กำลังสองของ 2.5% point ของ t with ν d.f. คือ 5% point ของ F with 1 and ν d.f. ความสัมพันธ์ดังกล่าวนี้เราอาจสังเกตเห็นได้จาก t-table และ F-table 2.5% points ของ t สำหรับ d.f. ต่าง ๆ คือ 12.706, 4.3027, 3.1825 ฯลฯ กำลังสองของค่าเหล่านี้คือ 161.44, 18.513, 10.128 ฯลฯ ซึ่งเป็นค่าต่าง ๆ ที่ให้ไว้ในคอลัมน์แรกของ 5% F-table เพราะฉะนั้นในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากันนั้นเราอาจใช้ two-tailed t-test หรือ one-tailed F-test (analysis of variance) อย่างใดอย่างหนึ่งได้ และการทดสอบทั้งสองนี้ย่อมให้ conclusion อย่างเดียวกันเสมอ

ความจริง $t^2 = F$ นี้อาจถูกพิสูจน์ได้อีก ในข้อ 7.9 ได้แสดงให้เห็นแล้วว่า

$$\text{SS} \; : \quad \sum (y - \mu)^2 = n(\bar{y} - \mu)^2 + \sum (y - \bar{y})^2 \quad \text{.........(4)}$$

$$\text{d.f.:} \qquad n = 1 + n - 1 \quad \text{.........(5)}$$

ดังนั้น

$$F = \frac{\frac{n(\bar{y} - \mu)^2}{1}}{\frac{\sum (y - \bar{y})^2}{n - 1}}$$

$$= \frac{n(\bar{y} - \mu)^2}{s^2}$$

$$= \frac{(\bar{y} - \mu)^2}{\frac{s^2}{n}} \quad \text{................. (6)}$$

follows the F-distribution with 1 and n-1 d.f. ใน Theorem 8.1 a ก็ได้แสดงให้เห็นแล้วว่า

$$t = \frac{\bar{y} - \mu}{\sqrt{\frac{s^2}{n}}} \quad \text{................. (7)}$$

follows the t-distribution with n-1 d.f. จึงเห็นได้จาก Equations (6) และ (7) ว่า $t^2 = F$

ความสัมพันธ์ระหว่าง t-distribution และ F-distribution อาจสรุปได้เป็น theorem ดังไปนี้:

Theorem 12.6: If a statistic t follows the Student's t-distribution with ν degrees of freedom, t^2 follows the F-distribution with 1 and ν degrees of freedom.

12.7 Assumptions

The assumptions are the conditions under which a test of hypothesis is valid. In the analysis of variance, the assumptions are the conditions under which the statistic

$$F = \frac{\text{among-sample MS}}{\text{within-sample MS}}$$

$$= \frac{\text{treatment MS}}{\text{error MS}} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

follows the F-distribution.

conditions เหล่านี้ซึ่งถูกสาธิตไว้แล้วใน sampling experiment ของข้อ 12.2 มี 3 ประการคือ :

(ก) k samples เป็น random samples ซึ่งได้มาจาก k populations

(ข) k populations เป็น normal populations

(ค) variances ของ k populations มีค่าเท่ากัน

จริงอยู่ การมีค่าเท่ากันของ k population means เป็นสภาพอันจำเป็นอีกอย่างหนึ่งซึ่งจะทำให้ statistic F follows the F-distribution แต่สภาพดังกล่าวนี้เป็นสมมติฐานที่จะถูกทดสอบ ก่อนการทดสอบจะต้องถือว่าสมมติฐานเป็นจริง แต่เมื่อทดสอบแล้วสมมติฐานนั้นอาจถูกปฏิเสธก็ได้ สภาพต่าง ๆ ทั้ง 3 ประการดังกล่าวไว้ข้างบนนี้ต้องเป็นจริง จะก่อนหลังการทดสอบก็ตาม

ถ้า assumptions ไม่สมบูรณ์ statistic F ที่ได้ให้ไว้ใน Equation (1) จะไม่ follow the F-distribution และทำให้ percentage points ต่าง ๆ ที่ได้ให้ไว้ใน F-table ไม่ถูกต้อง

ดังนั้น 5% point จะไม่ใช่ 5% point ที่แท้จริง แต่จะเป็น percentage point อันหนึ่งที่แตกต่างออกไป เพราะฉะนั้นเมื่อเราใช้ 5% point ของ F-table ในการหา critical region, significance level จริงจะไม่ใช่ 5% อย่างแท้จริง แต่จะมากกว่าหรือน้อยกว่า 5% ดังนั้นผลของการใช้ tables ที่มีอยู่โดยปราศจาก assumptions อันสมบูรณ์ซึ่งเป็นรากฐานของการทดสอบสมมติฐาน คือ significance level จะถูกกระทบกระเทือน ผลอย่างอื่นที่คิดว่าจะเกิดตามมาก็คืออาจกระทบกระเทือนถึงความน่าจะเป็นของการเกิด Type II error ด้วย

ผลที่เกิดจาก assumptions ไม่สมบูรณ์นี้มีผู้ทำการตรวจสอบหลายคน และพอจะสรุปกล่าวโดยสังเขปถึงผลของการตรวจสอบได้ดังต่อไปนี้

(ก) randomness of samples

randomness of samples อาจสัมฤทธิ์ผลได้โดย randomizing สิ่งที่ทดลอง (ข้อ 10.9) เราอาจใช้ random number table เพื่อความประสงค์นี้ได้ กล่าวได้อีกอย่างหนึ่งว่าเราขจัดผลที่เกิดจาก non-randomness โดยการทำ samples ให้เป็น random samples แทนที่จะสร้าง non-randomness ขึ้นแล้ววิตกถึงผลที่จะเกิดตามมา

(ข) normality of populations

non-normality ของ populations ไม่ทำให้เกิดความผิดพลาดอย่างร้ายแรงขึ้นใน F-test หรือ two-tailed t-test ถ้าเราใช้ F-table และ t-table ในการหา critical regions, significance level จริงจะโตกว่า significance level ที่ถูกระบุ ตัวอย่าง เช่น ถ้าเราใช้ 5% point ของ F-table ในการหา critical region, significance level จริงจะโตกว่า 5% เพราะฉะนั้นถ้าสมมติฐานเป็นจริงการปฏิเสธสมมติฐานมักจะมีได้มากกว่าที่ significance level ระบุไว้

(ค) homogeneity of variances

ถ้า variances ของ k populations ไม่ต่างกันมากเกินไป ก็จะไม่กระทบกระเทือนถึง F-test และ two-tailed t-test ร้ายแรงนัก ผลที่เกิดขึ้นจาก heterogeneity of variances จะถูกลดลงได้โดยการใช้ samples ที่มี size อย่างเดียวกัน (ข้อ 10.7)

สรุปกล่าวได้ว่าการคลาดเคลื่อนเล็กน้อยจาก assumptions จะไม่ทำให้เกิดความผิดพลาดร้ายแรงขึ้นใน F-test และ two-tailed t-test ในบทที่ 23 จะให้วิธีต่าง ๆ สำหรับแก้ไขการคลาดเคลื่อนจาก assumptions

12.8 Applications

analysis of variance เป็นเรื่องกว้างขวางมากเรื่องหนึ่งในวิชาสถิติ ที่ได้กล่าวไว้ในบทนี้เป็นเพียงกรณีสามัญที่สุดเท่านั้น แรกเริ่มทีเดียวนั้น R.A. Fisher เป็นผู้พัฒนา analysis of variance ขึ้นสำหรับให้นักเกษตรใช้วิเคราะห์ข้อมูลจากการทดลองต่าง ๆ ในสนาม ในปัจจุบันการใช้ analysis of variance ได้ขยายไปยังการทดลองทางวิทยาศาสตร์ต่าง ๆ ดังเช่นตัวอย่างที่ให้ไว้ในข้อนี้

โรงงานอุตสาหกรรมแห่งหนึ่งมีวิธีทำ fiber boards ต่าง ๆ กัน 3 วิธี (k = 3) และต้องการทราบว่ากรรมวิธีทั้งสามนี้จะผลิต fiber boards ออกมาได้แข็งแรงเท่ากันหรือไม่ จึงต้องหา random sample สมมติว่าประกอบด้วย fiber boards 20 แผ่น (n = 20) ออกมาจากกรรมวิธีละหนึ่ง sample ความแข็งแรงของ fiber board แต่ละแผ่น (observation) ของ 60 แผ่นนี้จะถูกหาไว้ เราอาจใช้ analysis of variance ในการทดสอบสมมติฐานว่าความแข็งแรงโดยเฉลี่ยของ fiber boards ที่ผลิตโดยกรรมวิธีทั้งสามนี้เป็นอย่างเดียวกัน การวิเคราะห์คงทำได้ดังต่อไปนี้

sources of variation	degrees of freedom
among processes	2 = k - 1
within processes	57 = kn - k
total	59 = kn - 1

ตัวอย่างอื่นที่จะยกมาแสดงให้เห็นอีก เช่นเราอาจเปรียบเทียบความสามารถที่ทำให้สัตว์อ้วนของอาหารสัตว์ต่าง ๆ กัน 5 อย่าง สมมติว่าเรามีสัตว์ 50 ตัว แบ่งออกเป็น 5 ฝูงโดยวิธี random ซึ่งใช้ random number table ช่วย (ข้อ 10.9) สัตว์เหล่านี้จะถูกเลี้ยงเป็นรายตัว แต่ฝูงสัตว์แต่ละฝูงซึ่งมีสัตว์ 10 ตัวจะถูกเลี้ยงด้วยอาหารต่างกัน สัตว์ทั้งหมดจะถูกชั่งน้ำหนักไว้ก่อนเริ่มการเลี้ยงและหลังการเลี้ยง น้ำหนักของสัตว์แต่ละตัวที่เพิ่มขึ้นคือ observation เราอาจใช้ analysis of variance ในการทดสอบสมมติฐานว่า สัตว์ 5 ฝูงนี้โดยเฉลี่ยจะได้น้ำหนักที่เพิ่มขึ้นเท่ากัน การวิเคราะห์คงทำได้ดังต่อไปนี้

sources of variation	degrees of freedom
among rations	4 = k - 1
within rations	45 = kn - k
total	49 = kn - 1

12.9 Specific Tests

มีการทดสอบหลายอย่างที่เกี่ยวกับ analysis of variance, F-test ที่กล่าวมาแล้วในข้อ 12.5 เป็นการทดสอบโดยทั่วไปอย่างหนึ่งซึ่งใช้แต่เพียงทดสอบสมมติฐานว่า k population means มีค่าเท่ากัน ยังมีการทดสอบอย่างอื่นที่ถูกคิดขึ้นเพื่อทดสอบสมมติฐานให้เจาะจงมากยิ่งขึ้นอันเกี่ยวกับ k population means การทดสอบเหล่านี้บางอย่างได้แสดงไว้ในข้อ 15.3, 15.4, 15.5 และ 17.8 ข้อได้เปรียบของการทดสอบเหล่านี้ซึ่งมีเหนือ F-test ทั่วไปจะถูกพิจารณาในข้อต่าง ๆ เหล่านั้น

12.10 Unequal Sample Sizes

analysis of variance ตามที่ได้พิจารณามาแล้วเป็นกรณีของ k samples ซึ่งแต่ละ sample ประกอบด้วย n observations ในข้อนี้จะพิจารณาถึงกรณี sample sizes ไม่เท่ากัน กล่าวคือ k samples ประกอบด้วย n_1, n_2, ..., n_k observations ตามลำดับ อย่างไรก็ดีสมมติฐานและ assumptions ของทั้งสองกรณีนี้คงเป็นอย่างเดียวกัน ความแตกต่างส่วนใหญ่ของมันอยู่ที่วิธีคำนวณซึ่งจะแสดงให้เห็นโดยตัวอย่างที่ให้ไว้ใน Table 12.10 a ตัวอย่างนี้มี 3 samples (k = 3) ประกอบด้วย 2, 2 และ 4 observations ตามลำดับ นั่นคือ $n_1 = 2$, $n_2 = 2$, $n_3 = 4$ และ $\Sigma n = 8$ ควรสังเกตว่าจำนวน observations ทั้งหมดในที่นี้คือ Σn แทนที่จะเป็น kn

Table 12.10 a

sample No.	1	2	3	
observations	7 9	7 5	3 4 2 3	
sample total, T sample size, n sample mean, $\bar{y}$ $\frac{T^2}{n}$	16 2 8 128	12 2 6 72	12 4 3 36	40 = grand total, G 8 = $\sum n$ 5 = general mean, $\bar{\bar{y}}$ 236 = $\sum(\frac{T^2}{n})$ *

* $\sum(\frac{T^2}{n})$ ไม่ใช่ $\frac{\sum T^2}{n}$ เพราะ sample sizes ไม่เท่ากัน

observation แต่ละตัวใน 8 observations ของ 3 samples อาจแตกออกได้เป็น 3 components คือ

$$y = \bar{\bar{y}} + (\bar{y} - \bar{\bar{y}}) + (y - \bar{y}) \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

components ของ observations ต่าง ๆ เหล่านี้ได้แสดงไว้ใน Table 12.10 b

Table 12.10 b

sample No.	1	2	3
components of observations $\bar{\bar{y}} + (\bar{y} - \bar{\bar{y}}) + (y - \bar{y})$	5 + 3 - 1 5 + 3 + 1	5 + 1 + 1 5 + 1 - 1	5 - 2 + 0 5 - 2 + 1 5 - 2 - 1 5 - 2 + 0
sample mean, $\bar{y}$	5 + 3 + 0	5 + 1 + 0	5 - 2 + 0

among-sample SS คือ

$$\Sigma\, n(\bar{y}-\bar{\bar{y}})^2 = n_1(\bar{y}_1-\bar{\bar{y}})^2 + n_2(\bar{y}_2-\bar{\bar{y}})^2 + n_3(\bar{y}_3-\bar{\bar{y}})^2 \;.......... (2)$$

$$= 2(3)^2 + 2(1)^2 + 4(-2)^2$$

$$= 36$$

with k - 1 หรือ 3 - 1 หรือ 2 d.f.

within-sample SS ซึ่งเป็น pooled SS ของ k samples คือ

$$\text{within-sample SS} = \text{pooled SS}$$

$$= \left[(-1)^2+(1)^2\right] + \left[(1)^2+(-1)^2\right] + \left[(0)^2+(1)^2+(-1)^2+(0)^2\right]$$

$$= \left[1+1\right] + \left[1+1\right] + \left[0+1+1+0\right]$$

$$= 6$$

d.f. ของ within-sample SS คือ :

$$(n_1-1) + (n_2-1) + ... + (n_k-1) = \Sigma\, n - k \;......... (3)$$

แทนที่จะเป็น kn - k หรือ k (n - 1) เพราะฉะนั้นในตัวอย่างนี้ d.f. ของ within-sample SS คือ 1 + 1 + 3 = 5

general mean ($\bar{\bar{y}}$) ในกรณีนี้จะไม่ใช่ mean ของ k sample means แต่จะเป็น weighted mean ของ k sample means ซึ่งมี sample sizes เป็น weights ของมัน นั่นคือ

$$\bar{\bar{y}} = \frac{n_1\bar{y}_1 + n_2\bar{y}_2 + ... + n_k\bar{y}_k}{n_1 + n_2 + ... + n_k}$$

$$= \frac{T_1 + T_2 + ... + T_k}{\Sigma\, n}$$

$$= \frac{G}{\Sigma\, n} \;.................... (4)$$

เราจะเห็นได้จาก Equation (4) ว่า general mean ($\bar{\bar{y}}$) คือ grand total หารด้วยจำนวน observations ทั้งหมดยังคงเป็น mean of the composite of the k samples

วิธีคำนวณค่าของ F ทางลัดนั้นเราคงคำนวณ sample totals (T) และ grand total (G) ได้เช่นเดียวกันกับกรณี sample size เท่ากัน แต่จะต้องคำนวณ $\frac{T^2}{n}$ ไว้ทุก sample (Table 12.10 a) รายละเอียดของวิธีคำนวณค่าของ F ทางลัดสำหรับตัวอย่างที่ให้ไว้นั้นได้แสดงไว้ใน Table 12.10 c และ 12.10 d แล้ว

Table 12.10 c

preliminary calculations				
1	2	3	4	5
type of total	total of squares	No. of items squared	No. of observations per squared item	total of squares per observation 2 ÷ 4
grand	G^2	1	Σn	$\frac{G^2}{\Sigma n}$ I
sample	-	-	-	$\Sigma(\frac{T^2}{n})$ II
observation	-	-	-	Σy^2 III
analysis of variance				
source of variation	sum of squares SS	degrees of freedom	mean square MS	F
among-sample	II - I	$k - 1$		
within-sample	III - II	$(\Sigma n) - k$		
total	III - I	$(\Sigma n) - 1$		

average of the within-sample MS ยังคงเป็น σ^2 ซึ่งเป็น variance รวมของ k populations (Equation (5), ข้อ 12.4) แต่ average of the among-sample MS จะไม่ใช่ $\sigma^2 + n\,\sigma^2_{\mu}$ (Equation (4), ข้อ 12.4) เพราะในกรณีนี้จะไม่มี n ในขณะนี้ sample sizes ได้กลายเป็น $n_1, n_2, \ldots\ldots, n_k$ ตามลำดับไปเสียแล้ว n ตัวนี้จึงถูกแทนด้วย n_o ในเมื่อ

$$n_o = \frac{1}{k-1}\left(\Sigma n - \frac{\Sigma n^2}{\Sigma n}\right) \quad \ldots\ldots\ldots\ldots\ldots (5)$$

Table 12.10 d $G = 40$
$\Sigma n = 8$

preliminary calculations				
1	2	3	4	5
type of total	total of squares	No. of items squared	No. of observations per squared item	total of squares per observation 2 ÷ 4
grand sample observation	1,600	1	8	200 236 242

$\Sigma y^2 = (7^2+9^2+7^2+5^2) + (3^2+4^2+7^2+3^2)$

analysis of variance

$n_1 = 3,\ n_2 = 2,\ n_3 = 4$
$k = 3$

source of variation	sum of squares SS	degrees of freedom	mean square MS	F
among-sample	36	$(3-1) = 2$ $(k-1)$	18.0	15.0
within-sample	6	$(8-3) = 5$ $(\Sigma n - k)$	1.2	
total	42	$8-1 = 7$ $(\Sigma n - 1)$		

ตามตัวอย่างที่ให้ไว้ใน Table 12.10 a นั้น $n_1 = 2$, $n_2 = 2$, $n_3 = 4$ และ $k = 3$ ดังนั้น

$$n_o = \frac{1}{3-1}\left[(2+2+4) - \frac{(2^2+2^2+4^2)}{(2+2+4)}\right]$$

$$= \frac{1}{2}\left(8 - \frac{24}{8}\right)$$

$$= 2.5 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

ค่าของ n_o เท่ากับค่าเฉลี่ยของ sample size โดยประมาณแต่น้อยกว่าเล็กน้อย ตามตัวอย่างนี้ค่าเฉลี่ยของ sample size คือ $\frac{2+2+4}{3} = 2.67$ เนื่องจากเรานำเอา Equation (5) มากล่าวโดยไม่มีการพิสูจน์ ฉะนั้น อย่างน้อยที่สุดก็ควรจะแสดงให้เห็นจริงสำหรับกรณี sample size เท่ากัน กล่าวคือถ้า sizes ของ k samples เท่ากันแล้ว n_o ก็ควรจะเท่ากับ n ดังนั้น ถ้า sample sizes เท่ากัน คือ $n_1 = n_2 = \ldots\ldots = n_k = n$ เมื่อเอา n ไปแทนค่าใน Equation (5) จะได้

$$n_o = \frac{1}{k-1}\left[kn - \frac{kn^2}{kn}\right]$$

$$= \frac{1}{k-1}\left[\frac{k^2n^2 - kn^2}{kn}\right]$$

$$= \frac{1}{k-1}\left[\frac{kn^2(k-1)}{kn}\right]$$

$$= n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

ตามที่ได้กล่าวมาแล้วทั้งหมดในข้อนี้แสดงให้เห็นว่าไม่มีความแตกต่างในหลักพื้นฐานในระหว่างกรณี sample sizes เท่ากันและไม่เท่ากันแต่อย่างใด แต่การทำให้ sample sizes เท่ากันย่อมมีข้อดีกว่า ดังจะได้พิจารณาในข้อต่อไป

12.11 Advantages of Equal Sample Size

ใน analysis of variance การใช้ sample sizes เท่ากันย่อมมีข้อดีเหนือกว่าการใช้ sample sizes ไม่เท่ากันหลายข้อ ดังจะกล่าวต่อไปนี้คือ:

1. ข้อดีที่เห็นได้ชัดที่สุดของการมี sample size เท่ากันก็คือความง่ายของการคำนวณ ถ้า sample sizes ไม่เท่ากันการคำนวณค่าของ F จะต้องใช้เวลามากขึ้นบ้างแต่ก็ไม่มากนัก การคำนวณค่าของ $\Sigma(\frac{T^2}{n})$ ดูเหมือนว่าจะต้องการเวลาบ้างแต่เมื่อได้ค่าของ T และ n แล้วก็อาจคำนวณติดต่อคราวเดียวกันได้โดยเครื่องคำนวณเลข อย่างไรก็ดีความง่ายของการคำนวณเป็นข้อดีโดยเฉพาะอย่างหนึ่งซึ่งไม่ใช่ในการคำนวณค่าของ F แต่สำหรับการทดสอบโดยเฉพาะต่าง ๆ ที่กล่าวไว้ในข้อ 12.9 ข้อดีข้อนี้จะเห็นได้ในบทหลัง ๆ ซึ่งพิจารณาถึงการทดสอบเหล่านี้ .

2. ข้อดีข้อที่ 2 คือ การเท่ากันของ sample size จะลดผลเสียที่เกิดจาก population variances ไม่เท่ากันให้น้อยลง assumption ประการหนึ่งของ analysis of variance คือ k population variances เท่ากัน (ข้อ 12.7) ถ้า sample sizes เท่ากันแล้วผลเสียซึ่งเกิดจากความไม่สมบูรณ์ของ assumption นี้จะไม่ร้ายแรงเหมือนอย่างกรณี sample sizes ไม่เท่ากัน (ข้อ 10.7)

3. ข้อดีข้อที่ 3 ของ sample size เท่ากันคือความน่าจะเป็นของการเกิด Type II error จะถูกลดลง ความน่าจะเป็นของการเกิด Type II error ใน analysis of variance นั้น อยู่ภายใต้อิทธิพลของเรโช

$$\frac{\sigma^2 + n\sigma_\mu^2}{\sigma^2} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

เรโชนี้ยิ่งมากความน่าจะเป็นของการเกิด Type II error จะยิ่งน้อยลง ผลของ n และ σ_μ^2 ใน Type II error ได้ถูกพิจารณามาแล้วในข้อ 12.5 ในขณะนี้จะได้พิจารณาถึงผลของการใช้ n_o แทน n (ข้อ 12.10) ใน analysis of variance จำนวน samples คือ k และจำนวน observations ทั้งหมดคือ Σn, distribution ของ Σn observations ในระหว่าง k samples จะกำหนดค่าของ n_o ซึ่งจะมากที่สุดเมื่อ sample sizes เท่ากัน นั่นคือ

$$n_1 = n_2 = \ldots\ldots\ldots = n_k = \frac{\Sigma n}{k}$$

ตัวอย่าง: ถ้า $k = 4, \quad \sum n = 40$

เมื่อ $n_1 = n_2 = n_3 = n_4 = 10$

$n_o = 10$

เมื่อ $n_1 = 2, \quad n_2 = 5, \quad n_3 = 15, \quad n_4 = 18$ และ $\sum n = 40$

$$n_o = \frac{1}{k-1}\left[\sum n - \frac{\sum n^2}{\sum n}\right]$$

$$= \frac{1}{4-1}\left[40 - \frac{2^2 + 5^2 + 15^2 + 18^2}{40}\right]$$

$$= 8.52 \text{ ซึ่งน้อยกว่า } 10$$

จะเห็นได้ว่าทั้ง 2 กรณีนี้ต่างก็มีจำนวน observations ทั้งหมดเป็น 40 ด้วยกัน แต่ค่าของ n_o ในกรณี sample sizes เท่ากันมากกว่าในกรณี sample sizes ไม่เท่ากัน ถ้า n_o มีค่ามากขึ้น เรโช

$$\frac{\sigma^2 + n_o \sigma^2_{\mu}}{\sigma^2}$$

ก็จะมีค่ามากขึ้น และโดยเฉลี่ยค่าของ F จะมากขึ้น ถ้า $\sigma^2_{\mu} > 0$ ดังนั้นสมมติฐานที่ว่า $\sigma^2_{\mu} = 0$ ก็มักจะถูกปฏิเสธได้มากขึ้นถ้าสมมติฐานนั้นไม่เป็นจริง เพราะฉะนั้นการทำให้ sample sizes เท่ากันจึงลดความน่าจะเป็นของการเกิด Type II error อย่างไรก็ดี ถ้าสมมติฐานเป็นจริง คือ $\sigma^2_{\mu} = 0$ แล้ว ค่าของ n_o จะไม่มีผลในเรโชซึ่งเท่ากับ 1 นี้เลย ผลที่เกิดขึ้นก็คือความน่าจะเป็นของการเกิด Type I error จะไม่ถูกกระทบกระเทือนโดยค่าของ n_o เพราะฉะนั้นการทำให้ sample sizes เท่ากันจึงมีแต่ข้อดีทั้งนั้น

เราอาจพิสูจน์โดยพีชคณิตได้ว่า

$$n_o = \frac{\sum n}{k} - \frac{s_n^2}{\sum n}$$

$$= \bar{n} - \frac{s_n^2}{\sum n} \quad \ldots\ldots\ldots\ldots\ldots \text{(2)}$$

ซึ่งแสดงว่า n_o น้อยกว่าหรือเท่ากับ sample sizes เฉลี่ย ($\bar{n}$) เสมอ ขนาดของความแตกต่างระหว่าง n_o และ $\bar{n}$ นั้นย่อมแล้วแต่ variance ของ sample sizes (s_n^2) ในเมื่อ:

$$s_n^2 = \frac{\Sigma (n - \bar{n})^2}{k - 1}$$

$$= \frac{\Sigma n^2 - \frac{(\Sigma n)^2}{k}}{k - 1} \quad \ldots\ldots\ldots\ldots\ldots (3)$$

ถ้าจำนวน observations ทั้งหมดเป็นอย่างเดียวกัน การแปรผันในระหว่าง sample sizes ยิ่งมาก ความน่าจะเป็นของการเกิด Type II error ก็จะยิ่งมาก

การพิสูจน์โดยพีชคณิตของ Equation (2) ทำได้ดังต่อไปนี้ :

$$n_o = \frac{1}{k - 1}\left[\Sigma n - \frac{\Sigma n^2}{\Sigma n}\right]$$

$$= \frac{\Sigma n}{k} - \frac{\Sigma n}{k} + \frac{1}{k - 1}\left[\frac{(\Sigma n)^2 - \Sigma n^2}{\Sigma n}\right]$$

$$= \frac{\Sigma n}{k} - \frac{(k-1)(\Sigma n)^2 - k(\Sigma n)^2 + k\Sigma n^2}{k(k-1)\Sigma n}$$

$$= \frac{\Sigma n}{k} - \frac{k\Sigma n^2 - (\Sigma n)^2}{k(k-1)\Sigma n}$$

$$= \bar{n} - \frac{s_n^2}{\Sigma n} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

เนื่องจากความเท่ากันของ sample sizes มีข้อดีหลายประการเหนือความไม่เท่ากันของ sample sizes ดังกล่าวแล้ว ในการทดลองต่าง ๆ นักวิทยาศาสตร์ผู้ทำการทดลองควรทำให้ sample sizes เท่ากันเสมอ.

Chapter 13

Review

บทนี้จะกล่าวทบทวนถึงเนื้อหาอันเป็นรากฐานและวิธีต่าง ๆ ของสถิติศาสตร์ซึ่งได้พิจารณามาแล้วในบทก่อน ๆ

บทต่าง ๆ ที่แล้วมาแสดงให้เห็นว่าสถิติศาสตร์เกี่ยวข้องกับความสัมพันธ์ระหว่าง **populations** และ **samples** ของมัน และวัตถุประสงค์สุดยอดของสถิติศาสตร์คือ การหา **inductive inference** ที่เกี่ยวกับ **population** จาก samples ที่กำหนดให้ เพราะฉะนั้นวิธีต่าง ๆ ทางสถิติ จึงอาจนับได้ว่าเป็นเครื่องมือสำหรับการตัดสินใจซึ่งตั้งอยู่บนเรื่องที่ไม่สมบูรณ์ (sample) ในการทดลองทางวิทยาศาสตร์วิธีเหล่านี้จะทำให้นักวิทยาศาสตร์เข้าถึงข้อยุติทั่วไปจากข้อมูลในการทดลอง (sample) ได้ เป็นเวลาหลายชั่วอายุคนมาแล้วที่นักวิทยาศาสตร์สาขาต่าง ๆ เป็นจำนวนมากได้ใช้คณิตศาสตร์เข้าช่วยใน **deduction** ขณะนี้เราอาจใช้สถิติศาสตร์เข้าช่วยใน induction ได้แล้ว

13.1 All Possible Samples

หลักอันเป็นรากฐานของสถิติศาสตร์อย่างหนึ่งคือ statistic ตัวหนึ่งเช่น **sample mean** ($\bar{y}$) จะมีค่าเปลี่ยนไปได้ตาม samples แต่ parameter ตัวหนึ่งเช่น population mean (μ) มีค่าตายตัว เมื่อเราหา **all possible samples** ออกมาจาก population ที่กำหนดและคำนวณค่าของ **statistic** ของแต่ละ **sample** ไว้ ก็จะได้ frequency distribution ของ **statistic** นั้น ความประสงค์ของการหา **all possible samples** ก็เพื่อสร้าง frequency **distributions** ของ **statistics** ต่าง ๆ เป็นต้นว่า **u**, X^2, **t** และ **F** จาก frequency distributions เหล่านี้เราอาจทำตารางของ **percentage points** ต่าง ๆ ได้ ดังเช่นที่ให้ไว้ใน Appendix และค่าในตารางเหล่านี้จะถูกใช้ในการคำนวณ **confidence interval** และการหา critical region ในการทดสอบสมมติฐาน **all possible samples** และ **distributions** ของ statistics ต่าง ๆ เป็นวิธีซึ่งจะนำไปสู่ผลสำเร็จ ในปัญหาต่าง ๆ ที่ถูกประยุกต์จะใช้ **percentage** points ในตารางเหล่านี้เท่านั้น ในการประมาณค่าของ **parameter** ตัวหนึ่งเราคำนวณ **confidence interval** เพียง **interval** เดียว ในการทดสอบสมมติฐานอย่างหนึ่งเราก็คำนวณค่าของ **statistic** เพียงตัวเดียว แต่ทั้ง 2 กรณีนี้ย่อมต้องการ **percentage points** ในตารางทั้งสิ้น

ความประสงค์ของการทำ sampling experiments ไม่ใช่เพื่อแสดงว่า statistics ถูกประยุกต์อย่างไร แต่เพื่อแสดงให้เห็นจริงใน distributions ของ statistics ต่าง ๆ เช่น u, X^2, t และ F และเพื่ออธิบายความหมายของ Tables ที่ให้ไว้ใน Appendix ใน sampling experiments เราหา random samples ออกมาเพียง 1,000 samples เท่านั้นแทนที่จะเป็น all possible samples เราจึงคำนวณค่าของ statistic ตัวหนึ่งเพียง 500 ค่าหรือ 1,000 ค่าเท่านั้น ค่าของ t 1,000 ค่า และค่าของ F 500 ค่า ย่อมเพียงพอที่จะแสดงให้เห็นรูปร่างทั่วไปของ distribution curves แต่จำนวนค่าเหล่านี้ไม่มากพอที่จะให้ 5 % หรือ 1 % points ของ distributions อย่างแน่นอนได้ ควรสังเกตว่า percentage points ที่ได้จาก sampling experiments ไม่เป็นอย่างเดียวกันอย่างแท้จริงกับ percentage points ที่ให้ไว้ใน tables เพราะฉะนั้น sampling experiments จึงใช้สำหรับสาธิตความหมายของ theorems มากกว่าจะใช้พิสูจน์ theorems เหล่านั้น

13.2 Relation Among Various Distributions

frequency distributions และวิธีทางสถิติต่าง ๆ ได้กล่าวไว้แล้วในบทก่อน ๆ มีบ่อยครั้งที่การทดสอบสมมติฐานอย่างเดียวกันอาจทำได้มากกว่าหนึ่งวิธี แต่ในขณะเดียวกันวิธีทดสอบวิธีหนึ่งอาจใช้ทดสอบสมมติฐานต่างกันได้ เช่นในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากัน ($\mu_1 = \mu_2$) เราอาจใช้ Student's t-test (Theorem 10.4a) หรือ F-test (analysis of variance with k = 2) วิธีใดวิธีหนึ่งก็ได้ ในทางตรงกันข้าม สมมติฐานที่ว่า 2 population variances มีค่าเท่ากัน ($\sigma_1^2 = \sigma_2^2$) และสมมติฐานที่ว่า k population means มีค่าเท่ากัน ($\mu_1 = \mu_2 = \dots\dots\dots = \mu_k$ หรือ $\sigma_{\mu}^2 = 0$) นั้น อาจถูกทดสอบด้วย F-test วิธีเดียวกันได้ distributions ต่าง ๆ ทั้งหมดซึ่งได้มาโดยการหา samples จาก normal populations ย่อมเกี่ยวข้องกันอย่างใกล้ชิด ความสัมพันธ์ระหว่าง distributions เหล่านี้มีดังต่อไปนี้:

1. ความสัมพันธ์ระหว่าง u และ X^2 ได้ให้ไว้แล้วใน Theorem 7.6 คือ "If a statistic u follows the normal distribution with mean equal to zero and variance equal to 1, u^2 follows the X^2- distribution with 1 degree of freedom".

2. ความสัมพันธ์ระหว่าง u และ t ได้ให้ไว้แล้วใน Theorem 8.1 b คือ "The t-distribution approaches the u-distribution as the number of degrees of freedom

approaches infinity".

3. ความสัมพันธ์ระหว่าง X^2 และ F ได้ให้ไว้แล้วใน Theorem 9.7 คือ "If a statistic X^2 follows the X^2- distribution with ν degrees of freedom, $\frac{X^2}{\nu}$ follows the F-distribution with ν and ∞ degrees of freedom".

4. ความสัมพันธ์ระหว่าง t และ F ได้ให้ไว้แล้วใน Theorem 12.6 คือ "If a statistic t follows the Student's t-distribution with ν degrees of freedom, t^2 follows the F-distribution with 1 and ν degrees of freedom".

5. ความสัมพันธ์ระหว่าง u และ F อาจหาได้จากความสัมพันธ์ระหว่าง u และ t และระหว่าง t และ F คือ "If u follows the t-distribution with ∞ degrees of freedom, u^2 must follow the F-distribution with 1 and ∞ degrees of freedom".

ควรสังเกตว่า distributions ของ u, t และ X^2 แต่ละ distribution เป็น special case ของ F-distribution ความสัมพันธ์ต่าง ๆ นี้ได้ถูกสรุปไว้ใน Table 13.2 ซึ่งทำขึ้นในรูปของ F-table

Table 13.2

F- table

ν_2 \ ν_1	1	2 3 ----------	∞
1 2 3 ⋮	t^2		
∞	u^2.	$\frac{X^2}{\nu_1}$	1

ถ้า Table 13.2 เป็น 5 % F-table, ตัวเลขต่าง ๆ ในคอลัมน์แรกของ table คือกำลังสองของ 2.5% points ของ t-distribution (ข้อ 12.6) ตัวเลขต่าง ๆ ในบรรทัดสุดท้ายของ table คือ 5 % points ของ $\frac{X^2}{\nu}$ (Table 5, Appendix) มุมล่างซ้ายของ table คือ u^2, 5 % point ของ F with 1 and α degrees of freedom คือ $(1.960)^2$ หรือ 3.842

4 distributions ได้ถูกนำมากล่าวตามลำดับคือ u, X^2, t และ F ตาม Table 13.2 จะเริ่มต้นจากมุมล่างซ้าย (u) แล้วขยายออกไปตามบรรทัดล่างสุดทั้งหมด จากนั้น u จะถูกขยายไปตามคอลัมน์แรกทั้งหมด และในที่สุดก็คลุม table ทั้งหมด เมื่อเปรียบเทียบกับ F-distribution แล้ว distributions ของ u, X^2 และ t ไม่สำคัญเลย แต่ก็เป็นประโยชน์ให้ทราบถึงข้อคิดทางสถิติ

เพื่อทำให้เรื่องของวิธีทางสถิติง่ายจึงต้องมีการรวบรวมวิธีต่าง ๆ เหล่านี้ วิธีหนึ่งสำหรับความประสงค์อย่างหนึ่งโดยเฉพาะ และไม่กล่าวถึงความสัมพันธ์ระหว่างวิธีต่าง ๆ ใครก็ตามที่อ่านเรื่องการค้นคว้าทดลองในสนามซึ่งใช้สถิติจะเสียเปรียบมากถ้าเขาไม่มีความรู้อยู่บ้างในเรื่องความสัมพันธ์ระหว่าง distributions ต่าง ๆ นี้ เช่น ถ้าเราทราบว่า t-test ใช้ทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากัน ($\mu_1 = \mu_2$) ได้ เราจะรู้สึกงงเมื่อเราพบว่า analysis of variance ก็ใช้เพื่อความประสงค์อย่างเดียวกันนี้ได้ด้วย

13.3 Tests of Hypotheses

ข้อ 13.2 แสดงให้เห็นแล้วว่า distributions ต่าง ๆ ได้ถูกนำมาให้ทราบอย่างมีระเบียบแบบแผน ในข้อนี้จะแสดงให้เห็นว่าวิธีทดสอบสมมติฐานต่าง ๆ ก็ถูกนำมาให้ทราบอย่างมีระเบียบแบบแผนด้วยเหมือนกัน วิธีทดสอบเหล่านี้ได้ให้ไว้ใน Table 13.3 ควรสังเกตว่าเราอาจใช้ F-test ในการทดสอบสมมติฐานทุกอย่างที่ให้ไว้ใน Table นั้น

Table 13.3

No. of parameters	hypothesis	statistic	Section No.
1	$\sigma^2 = \sigma_0^2$	χ^2	7.10 & 7.11
		$\frac{\chi^2}{\nu}$	7.10
		F	9.7
1	$\mu = \mu_0$	t	8.4 & 8.5
		F	7.9 & 12.6
2	$\sigma_1^2 = \sigma_2^2$	F	9.4 & 9.5
2	$\mu_1 = \mu_2$	t	10.6
		F	12.5
K	$\mu_1 = \mu_2 = \ldots = \mu_K$	F	12.5

statistics ทั้ง 4 ตัวคือ u, χ^2, t และ F ซึ่งกรอกไว้ใน Table 13.3 เป็นจำนวนเลขบริสุทธิ์ ค่าต่าง ๆ ของ statistics ทั้ง 4 ตัวนี้จะไม่ถูกกระทบกระเทือนเลยไม่ว่าหน่วยการวัด observations จะเป็นอะไร ความยาวอาจวัดเป็นนิ้วหรือเป็นเซ็นติเมตร ผลผลิตของพืชอาจวัดเป็นปอนด์ต่อแปลงเพาะปลูกหรือเป็น bushels ต่อเอเคอร์ และอุณหภูมิอาจวัดเป็นฟาเรนไฮท์หรือเซ็นติเกรดก็ได้ หน่วยของการวัดจะเป็นอะไรก็ตามค่าของ u, χ^2, t และ F จะไม่เปลี่ยนแปลง จากผลอันนี้ข้อยุติที่ได้โดยการทดสอบสมมติฐานจึงเป็นอย่างเดียวกันถึงแม้ว่าหน่วยของการวัดจะถูกเปลี่ยนไปต่าง ๆ กัน

สมมติฐานต่าง ๆ ที่กรอกไว้ใน Table 13.3 เป็นเรื่องง่ายมาก ถ้าเราทราบ 2 populations มันก็ไม่ยากที่จะพิจารณาว่า population ไหนมี mean โตกว่า ถ้า $\mu_1 = 49.2$ และ $\mu_2 = 49.1$ ใคร ๆ ก็เห็นได้ว่า μ_1 โตกว่า μ_2 โดยไม่เกิดมี error ชนิดใดเลย และยิ่งกว่านั้นใคร ๆ ก็อาจเห็นได้ว่า $\mu_1 - \mu_2 = 0.1$ โดยไม่ต้องใช้ confidence interval ช่วย คณิตศาสตร์ธรรมดาดังกล่าวนี้จะกลายเป็นวิธีทางสถิติอันยุ่งยากเมื่อเราไม่ทราบ population และต้องขึ้นอยู่กับ samples ถ้า $\bar{y}_1 = 49.2$ และ $\bar{y}_2 = 49.1$ จะยุติไม่ได้ว่า $\mu_1 - \mu_2 = 0.1$ เพราะว่า sample mean มีค่าเปลี่ยนไปตาม sample ข้อยุติที่เกี่ยวกับ population means ต้องได้มาโดยการใช้สถิติศาสตร์ มี 2 สภาพเท่านั้นซึ่งไม่ต้องการสถิติศาสตร์ คือ sample sizes โตมากจน samples กลายเป็นตัว populations ของมันไป หรือ population variances มีค่าเท่ากับศูนย์ เมื่อ observations ทั้งหมดของ population หนึ่งมีค่าอย่างเดียวกัน และในเวลาเดียวกัน observations ทั้งหมดของอีก population หนึ่งก็มีค่าอย่างเดียวกัน (นั่นคือ population variances ทั้งสองมีค่าเท่ากับศูนย์) ผลต่างระหว่าง μ_1 และ μ_2 จะถูกหาได้อย่างแน่นอนโดยการหา observation ออกมาจาก population ละหนึ่ง observation ตัวอย่างเช่น observations ทั้งหมดใน population หนึ่งมีค่าเท่ากับ 49.2 และ observations ทั้งหมดในอีก population หนึ่งมีค่าเท่ากับ 49.1 ดังนั้น 2 population means จะเท่ากับ 49.2 และ 49.1 ตามลำดับ จากผลอันนี้การหาผลต่างระหว่าง 2 population means จึงทำได้โดยการใช้คณิตศาสตร์ง่าย ๆ เท่านั้น และไม่จำเป็นต้องใช้วิธีทางสถิติเลย

เมื่อใดก็ตามที่ข้อยุติที่เกี่ยวกับ populations ได้มาโดยไม่ได้ใช้สถิติ สภาพอย่างหนึ่งหรือทั้งสองอย่างดังกล่าวข้างต้นต้องมีสมบูรณ์ อย่างไรก็ดีเรื่อง variances เท่ากับศูนย์นี้ไม่อยู่ในความสนใจของนักวิทยาศาสตร์ ตัวอย่างเช่น sociologists อาจสนใจที่จะทราบ income distribution ของประเทศ แต่ถ้ารายได้ของทุกคนเท่ากันแล้ว ปัญหาเรื่องนี้ก็จะอันตรธานไปหมดทันที

13.4 Significance

ในวิชาสถิติ คำ "significance" มีความหมายในทางวิชาการ โดยทั่วไปเราใช้คำนี้เกี่ยวกับการปฏิเสธสมมติฐาน ความหมายของคำ "significance" ขึ้นอยู่กับสมมติฐานที่ถูกทดสอบ นี่เป็นเหตุผลที่ว่าทำไมเราจึงกล่าวถึงการทดสอบสมมติฐานต่าง ๆ ก่อนพูดถึงคำ "significance"

ในการทดสอบสมมติฐานว่า population mean เท่ากับค่าที่กำหนด เช่น เท่ากับ 60 ถ้าสมมติฐานถูกปฏิเสธ sample mean ถูกเรียกว่าแตกต่างจาก 60 significantly แต่ถ้าสมมติฐานถูกยอมรับ sample mean ถูกเรียกว่าไม่แตกต่างจาก 60 significantly

ในการทดสอบสมมติฐานว่า 2 population means มีค่าเท่ากัน ถ้าสมมติฐานถูกปฏิเสธ 2 sample means นั้นถูกเรียกว่าแตกต่างกัน significantly นั่นคือ ถ้าข้อยุติแสดงว่า 2 population means มีค่าต่างกัน ข้อยุติที่ว่า 2 population means มีค่าต่างกันนั้นไม่ได้บอกให้ทราบว่าแตกต่างกันเท่าไร ขนาดของความแตกต่างต้องประมาณโดย confidence interval

ถ้าข้อยุติแสดงว่า k population means มีค่าไม่เท่ากัน เราเรียกผลของ analysis of variance ว่า significant

คำ "significance" นี้ใช้เฉพาะกับความเกี่ยวข้องกับ statistics เท่านั้น แต่ไม่ใช่กับ parameters เลย 2 sample means อาจแตกต่างกัน significantly หรือไม่แตกต่างกัน significantly นั้นย่อมแล้วแต่การปฏิเสธหรือการยอมรับสมมติฐานว่า 2 population means มีค่าเท่ากัน แต่คำ "significantly" ไม่ได้ใช้เพื่อขยายความแตกต่างในระหว่าง 2 population means

13.5 Sample Size

sample size มีบทบาทสำคัญในวิธีทางสถิติ ถ้ามองเพียงผิวเผินแล้วจะเห็นว่า sample size ยิ่งใหญ่ผลลัพธ์ยิ่งถูกต้องแน่นอน อย่างไรก็ดีประโยชน์โดยเฉพาะที่ได้จาก sample ใหญ่ยังมองไม่เห็นชัด

ในการทดสอบสมมติฐานนั้นย่อมมี errors 2 ชนิดเข้ามาเกี่ยวข้องด้วย การเพิ่มขึ้นในความถูกต้องแน่นอนของการทดสอบย่อมหมายถึงการลดลงของความน่าจะเป็นของการเกิด error ชนิดหนึ่งหรือทั้งสองชนิดเสมอ ความน่าจะเป็นของการเกิด Type I error เรียกว่า significance level ซึ่งอาจจะกำหนดให้ใหญ่หรือเล็กได้ตามความปรารถนาโดยไม่คำนึงถึง sample size ข้อได้เปรียบของการมี sample size ใหญ่คือการลดลงของความน่าจะเป็นของการเกิด Type II error ภายหลังที่ได้กำหนด significance level ลงไปแล้ว

ถ้าสมมติฐานที่ถูกทดสอบเป็นจริงและเกิดมี error ขึ้นก็จะเป็น Type I error เท่านั้น ตราบเท่าที่ significance level ยังคงเดิมและสมมติฐานที่ถูกทดสอบก็เป็นจริงด้วยแล้ว sample ขนาดใหญ่จะไม่ได้เปรียบ sample ขนาดเล็กเลย

ถ้าสมมติฐานที่ถูกทดสอบนั้นผิดและเกิดมี error ขึ้น ก็จะเป็น Type II error เท่านั้น ถ้า significance level ยังคงเดิมความน่าจะเป็นของการเกิด Type II error จะลดลงโดยการเพิ่ม sample size ขึ้น หรืออีกนัยหนึ่ง sample ขนาดใหญ่ค่อนข้างจะทำให้เกิดการปฏิเสธสมมติฐานที่ผิดได้ดีกว่า sample ขนาดเล็ก นี่คือข้อได้เปรียบของการใช้ sample ขนาดใหญ่ในการทดสอบสมมติฐาน

ในการประมาณค่าของ parameter ตัวหนึ่งโดย confidence interval นั้น confidence coefficient และความยาวของ interval เป็นสิ่งสำคัญมาก coefficient ยิ่งสูง interval ก็จะยิ่งจับค่าของ parameter ไว้ได้ แต่การเลือกใช้ confidence coefficient ย่อมเป็นไปตามใจชอบโดยไม่มีกฎเกณฑ์และไม่เกี่ยวกับ sample size แต่ประการใด ข้อได้เปรียบของการใช้ sample ขนาดใหญ่คือการลดความยาวของ interval ลงหลังจากที่เลือก confidence coefficient แล้ว ตัวอย่างเช่น 95 % confidence interval ของ μ คือ:

$$\bar{y} - t_{0.025}\sqrt{\frac{s^2}{n}} \qquad \text{ถึง} \qquad \bar{y} + t_{0.025}\sqrt{\frac{s^2}{n}}$$

ความยาวของ interval นี้คือ $2\,t_{0.025}\sqrt{\frac{s^2}{n}}$ ในขณะที่ $\bar{y}$ มีค่าเปลี่ยนไปตาม sample นั้น ความยาวของ interval จะเปลี่ยนไปด้วยถึงแม้ว่า sample size จะยังคงเดิมก็ตาม การเพิ่มขึ้นของ sample size จะลดความยาวโดยเฉลี่ยของ intervals ต่าง ๆ ดังกล่าวนี้

13.6 Simplified Statistical Methods

สมมติฐานและวิธีทดสอบสมมติฐานต่าง ๆ ดังที่ได้แสดงไว้ใน Table 13.3 นั้นได้กล่าวไว้แล้วในบทก่อน ๆ แต่วิธีเหล่านี้ก็ไม่ใช่วิธีที่ใช้ทดสอบสมมติฐานเหล่านั้นโดยเฉพาะเท่านั้น ในรอบ 10 ปีที่แล้วมานี้ได้มีผู้คิดวิธีง่าย ๆ ขึ้นหลายวิธีสำหรับทดสอบสมมติฐานอย่างเดียวกันเหล่านี้ บางวิธีซึ่งเรียกว่า inefficient statistics นั้น มักจะพิมพ์ไว้ในหนังสือเกี่ยวกับ industrial quality control แต่ไม่

ได้ให้ไว้ในตำราเล่มนี้ ข้อได้เปรียบของวิธีเหล่านี้คือคำนวณได้ง่าย ตัวอย่างเช่นค่าของ SS ซึ่งเราต้องการสำหรับ statistic ทุกตัวใน Table 13.3 นั้นวิธีง่าย ๆ เหล่านี้จะไม่ต้องการเลย

อย่างไรก็ดี สำหรับ observations ที่กำหนดให้อย่างเดียวกันซึ่งได้มาจาก normal population หนึ่งนั้นความน่าจะเป็นของการเกิด Type II error จะมีมากขึ้นถ้าเราใช้วิธีง่าย ๆ เหล่านี้แทนวิธีที่ให้ไว้ใน Table 13.3 นี่เป็นเหตุผลที่ว่าทำไมเราจึงเรียกวิธีง่าย ๆ เหล่านี้ว่า "inefficient statistics" เมื่อเปรียบเทียบดูแล้วเราอาจเรียกวิธีต่าง ๆ ใน Table 13.3 ได้ว่า "efficient statistics" แต่ควรเข้าใจว่าการเพิ่ม sample size ขึ้นจะลดความน่าจะเป็นของการเกิด Type II error ลง เพราะฉะนั้นความน่าจะเป็นนี้เราอาจทำให้เท่ากันได้โดยการใช้ sample sizes ของ "inefficient statistics" ให้ใหญ่กว่า sample sizes ของ "efficient statistics"

การเลือกวิธีของการทดสอบสมมติฐานนั้นขึ้นอยู่กับค่าใช้จ่ายของการคำนวณและค่าใช้จ่ายของการรวบรวม observations ถ้าค่าใช้จ่ายของการรวบรวม observations ต่ำก็อาจได้ samples ขนาดใหญ่กว่าและอาจใช้ "inefficient statistics" ได้เพื่อลดความยุ่งยากในการคำนวณลง แต่ถ้าการรวบรวม observations ต้องสิ้นเปลืองค่าใช้จ่ายมากแล้วก็จะต้องใช้อย่างรอบคอบกับ "efficient statistics" เรื่องนี้คล้ายกับเครื่องคั้นน้ำผลส้ม ถ้าส้มมีราคาโหลละ 1 เซ็นต์ก็ไม่มีความจำเป็นต้องไปซื้อเครื่องคั้นที่มีราคาแพงและมีประสิทธิภาพสูงมาคั้นส้มจนแห้ง แต่ถ้าส้มมีราคาผลละ 1 ดอลลาร์ก็จำเป็นต้องใช้เครื่องคั้นที่ดีเพื่อคั้นน้ำผลส้มทุกหยดจากส้มนั้น

นอกจากค่าใช้จ่ายแล้ว ความเฉื่อยชาของมนุษย์ยังมีบทบาทสำคัญในการเลือกวิธีทดสอบสมมติฐานอีกด้วย ตามธรรมชาตินั้นคนเรามักชอบใช้วิธีที่คุ้นเคยมากกว่าจะใช้วิธีแปลก ๆ ในปัจจุบันมักจะใช้ "inefficient statistics" กันในโรงงานอุตสาหกรรมซึ่งสามารถรวบรวม observations ได้ง่ายและเสียค่าใช้จ่ายน้อย ส่วน "efficient statistics" นั้น นักวิทยาศาสตร์จะใช้วิเคราะห์ข้อมูลที่ทดลองซึ่งตามปกติจะต้องมีการรวบรวมข้อมูลมากและเสียค่าใช้จ่ายมาก

13.7 Error

คำ "error" นี้ถูกใช้ในความหมายต่าง ๆ กันหลายอย่างในวิชาสถิติ ดังเช่น computing error, standard error, Type I error และ Type II error เหล่านี้ล้วนแต่เป็น error คนละชนิด ชนิดต่าง ๆ ของ errors จึงมีดังต่อไปนี้

1. ตามปกติเราเรียกความผิดที่เกิดขึ้นในการคำนวณว่า error ความผิดนี้อาจเกิดจากความบกพร่องของเครื่องคำนวณเลขหรือตัวผู้คำนวณก็ได้ ผู้เริ่มทำงานสถิติมักไม่ตระหนักถึงความเสียหายอย่างร้ายแรงของความผิดชนิดนี้ ความผิดในการคำนวณหากหาไม่พบแล้วอาจนำไปสู่ข้อยุติที่ผิดได้ไม่ว่าจะใช้วิธีทางสถิติกับข้อมูลที่ทดลองอย่างระมัดระวังสักเพียงใดก็ตาม

2. คำ "standard error" ไม่ได้บ่งให้ทราบว่าเกิดความผิดอย่างใดขึ้น ตามปกติ sample mean ย่อมเปลี่ยนไปตาม sample, standard error of the mean (ข้อ 5.3) วัดการเบี่ยงเบนของ sample means จาก population mean ความจริง sample means เบี่ยงเบนจาก population mean นั้นเป็นปรากฏการณ์ธรรมชาติอย่างหนึ่ง

3. คำ error ใน "error SS" ของ analysis of variance ก็ไม่ได้บ่งให้ทราบว่าเกิดความผิดอย่างใดขึ้น error ในข้อนี้คือการเบี่ยงเบนของ observation จาก sample mean

4. Type I และ Type II errors นับว่าเป็นความผิดที่แท้จริง แต่มันก็ไม่ใช่ความผิดที่เกิดจากมนุษย์หรือเครื่องคำนวณเลขแต่อย่างใด มันเป็นผลที่เกิดขึ้นเองตามธรรมชาติของการหาข้อยุติเกี่ยวกับ population ในเมื่อมีแต่เพียง samples เท่านั้น

TABLE 3

Area Under the Normal Curve

$-2 \to +\infty$

u	Area	u	Area	u	Area
-3.0	.0013	-1.0	.1587	1.0	.8413
-2.9	.0019	- .9	.1841	1.1	.8643
-2.8	.0026	- .8	.2119	1.2	.8849
-2.7	.0035	- .7	.2420	1.3	.9032
-2.6	.0047	- .6	.2743	1.4	.9192
-2.5	.0062	- .5	.3085	1.5	.9332
-2.4	.0082	- .4	.3446	1.6	.9452
-2.3	.0107	- .3	.3821	1.7	.9554
-2.2	.0139	- .2	.4207	1.8	.9641
-2.1	.0179	- .1	.4602	1.9	.9713
-2.0	.0228	0	.5000	2.0	.9772
-1.9	.0287	.1	.5398	2.1	.9821
-1.8	.0359	.2	.5793	2.2	.9861
-1.7	.0446	.3	.6179	2.3	.9893
-1.6	.0548	.4	.6554	2.4	.9918
-1.5	.0668	.5	.6915	2.5	.9938
-1.4	.0808	6	.7257	2.6	.9953
-1.3	.0968	.7	.7580	2.7	.9965
-1.2	.1151	.8	.7881	2.0	.9974
-1.1	.1357	.9	.8159	2.9	.9981
				3.0	.9987

u	Area	u	Area
-2.5758	.005	2.5758	.995
-1.9600	.025	1.9600	.975
-1.6449	.050	1.6449	.950
-0.6745	.250	0.6745	.750

TABLE 4

Percentage Points of the X^2- Distribution

ν d.f.	99.5%	97.5%	5%	2.5%	1%	.5%
1	392704×10^{-10}	982069×10^{-9}	3.84146	5.02389	6.63490	7.87944
2	0.0100251	0.0506356	5.99147	7.37776	9.21034	10.5966
3	0.0717212	0.215795	7.81473	9.34840	11.3449	12.8381
4	0.206990	0.484419	9.48773	11.1433	13.2767	14.8602
5	0.411740	0.831211	11.0705	12.8325	15.0863	16.7496
6	0.675727	1.237347	12.5916	14.4494	16.8119	18.5476
7	0.989265	1.68987	14.0671	16.0128	18.4753	20.2777
8	1.344419	2.17973	15.5073	17.5346	20.0902	21.9550
9	1.734926	2.70039	16.9190	19.0228	21.6660	23.5893
10	2.15585	3.24697	18.3070	20.4831	23.2093	25.1882
11	2.60321	3.81575	19.6751	21.9200	24.7250	26.7569
12	3.07382	4.40379	21.0261	23.3367	26.2170	28.2995
13	3.56503	5.00874	22.3621	24.7356	27.6883	29.8194
14	4.07468	5.62872	23.6848	26.1190	29.1413	31.3193
15	4.60094	6.26214	24.9958	27.4884	30.5779	32.8013
16	5.14224	6.90766	26.2962	28.8454	31.9999	34.2672
17	5.69724	7.56418	27.5871	30.1910	33.4087	35.7185
18	6.26481	8.23075	28.8693	31.5264	34.8053	37.1564
19	6.84398	8.90655	30.1435	32.8523	36.1908	38.5822
20	7.43386	9.59083	31.4104	34.1696	37.5662	39.9968
21	8.03366	10.28293	32.6705	35.4789	38.9321	41.4010
22	8.64272	10.9823	33.9244	36.7807	40.2894	42.7956
23	9.26042	11.6885	35.1725	38.0757	41.6384	44.1813
24	9.88623	12.4011	36.4151	39.3641	42.9798	45.5585
25	10.5197	13.1197	37.6525	40.6465	44.3141	46.9278
26	11.1603	13.8439	38.8852	41.9232	45.6417	48.2899
27	11.8076	14.5733	40.1133	43.1944	46.9630	49.6449
28	12.4613	15.3079	41.3372	44.4607	48.2782	50.9933
29	13.1211	16.0471	42.5569	45.7222	49.5879	52.3356
30	13.7867	16.7908	43.7729	46.9792	50.8922	53.6720
40	20.7065	24.4331	55.7585	59.3417	63.6907	66.7659
50	27.9907	32.3574	67.5048	71.4202	76.1539	79.4900
60	35.5346	40.4817	79.0819	83.2976	88.3794	91.9517
70	43.2752	48.7576	90.5312	95.0231	100.425	104.215
80	51.1720	57.1532	101.879	106.629	112.329	116.321
90	59.1963	65.6466	113.145	118.136	124.116	128.299
100	67.3276	74.2219	124.342	129.561	135.807	140.169

TABLE 5*

Percentage Points of χ^2/ν

ν d.f.	99.5%	97.5%	5%	2.5%	1%	.5%
1	392704×10^{-10}	982069×10^{-9}	3.84146	5.02389	6.63490	7.87944
2	0.0050126	0.0253178	2.99574	3.68888	4.60517	5.29830
3	0.0239071	0.0719317	2.60491	3.11613	3.78163	4.27937
4	0.0517475	0.1211048	2.37193	2.78583	3.31918	3.71505
5	0.0823480	0.1662422	2.21410	2.56650	3.01726	3.34992
6	0.1126212	0.2062245	2.09860	2.40823	2.80198	3.09127
7	0.1413236	0.2414100	2.00959	2.28754	2.63933	2.89681
8	0.1680524	0.2724663	1.93841	2.19183	2.51128	2.74438
9	0.1927696	0.3000433	1.87989	2.11364	2.40733	2.62103
10	0.2155850	0.3246970	1.83070	2.04831	2.32093	2.51882
11	0.2366555	0.3468864	1.78865	1.99273	2.24773	2.43245
12	0.2561517	0.3669825	1.75218	1.94473	2.18475	2.35829
13	0.2742331	0.3852877	1.72016	1.90274	2.12987	2.29380
14	0.2910486	0.4020514	1.69177	1.86564	2.08152	2.23709
15	0.3067293	0.4174760	1.66639	1.83256	2.03853	2.18675
16	0.3213900	0.4317288	1.64351	1.80284	1.99999	2.14170
17	0.3351318	0.4449518	1.62277	1.77594	1.96522	2.10109
18	0.3480450	0.4572639	1.60385	1.75147	1.93363	2.06424
19	0.3602095	0.4687658	1.58650	1.72907	1.90478	2.03064
20	0.3716930	0.4795415	1.57052	1.70848	1.87831	1.99984
21	0.3825552	0.4896633	1.55574	1.68947	1.85391	1.97148
22	0.3928509	0.4991955	1.54202	1.67185	1.83134	1.94525
23	0.4026270	0.5081957	1.52924	1.65547	1.81037	1.92093
24	0.4119263	0.5167125	1.51730	1.64017	1.79083	1.89827
25	0.4207880	0.5247880	1.50610	1.62586	1.77256	1.87711
26	0.4292423	0.5324577	1.49558	1.61243	1.75545	1.85730
27	0.4373185	0.5397519	1.48568	1.59979	1.73937	1.83870
28	0.4450464	0.5467107	1.47633	1.58788	1.72422	1.82119
29	0.4524517	0.5533483	1.46748	1.57663	1.70993	1.80468
30	0.4595567	0.5596933	1.45910	1.56597	1.69641	1.78907
40	0.5176625	0.6108275	1.39396	1.48354	1.59227	1.66915
50	0.5598140	0.6471480	1.35010	1.42840	1.52308	1.58980
60	0.5922433	0.6747950	1.31803	1.38829	1.47299	1.53253
70	0.6182171	0.6965371	1.29330	1.35747	1.43464	1.48879
80	0.6396500	0.7144150	1.27349	1.33276	1.40411	1.45401
90	0.6577367	0.7294067	1.25717	1.31262	1.37907	1.42554
100	0.6732760	0.7422190	1.24342	1.29561	1.35807	1.40169
∞	1.0000000	1.0000000	1.00000	1.00000	1.00000	1.00000

*This table is obtained from Table 4.

TABLE 6

Percentage Points of the t-Distribution

ν d.f.	5 %	2.5 %	1 %	.5 %
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779
27	1.703	2.052	2.473	2.771
28	1.701	2.048	2.467	2.763
29	1.699	2.045	2.462	2.756
30	1.697	2.042	2.457	2.750
40	1.684	2.021	2.423	2.704
60	1.671	2.000	2.390	2.660
120	1.658	1.980	2.358	2.617
∞	1.645	1.960	2.326	2.576

TABLE 7a *

5% Points of the F-Distribution

ν_2 \ ν_1	Degrees of Freedom for Numerator: 1	2	3	4	5	6	7	8	9
Degrees of Freedom for Denominator: 1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8868	8.8452	8.8123
4	7.7086	6.9443	6.5914	6.3883	6.2560	6.1631	6.0942	6.0410	5.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2066	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8378	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8626	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3808	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3513	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3661
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9751	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2782	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0848	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2540	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2900	2.1750	2.0867	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

TABLE 7 a (continued)

5% Points of the F-Distribution

ν_2 \ ν_1 (Degrees of Freedom for Denominator \ Degrees of Freedom for Numerator)	10	12	15	20	24	30	40	60	120	∞
1	241.88	243.91	245.95	248.01	249.05	250.09	251.14	252.20	253.25	254.32
2	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487	19.496
3	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494	8.5265
4	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6878	5.6581	5.6281
5	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3984	4.3650
6	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6688
7	3.6365	3.5747	3.5108	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	3.3472	3.2840	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	2.8536	2.7876	2.7186	2.6464	2.6090	2.5705	2.5309	2.4901	2.4480	2.4045
12	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410	2.2962
13	2.6710	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	2.6021	2.5342	2.4630	2.3879	2.3487	2.3082	2.2664	2.2230	2.1778	2.1307
15	2.5437	2.4753	2.4035	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.1040	2.0584	2.0107	1.9604
18	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	2.3779	2.3080	2.2341	2.1555	2.1141	2.0712	2.0264	1.9796	1.9302	1.8780
20	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	2.3210	2.2504	2.1757	2.0960	2.0540	2.0102	1.9645	1.9165	1.8657	1.8117
22	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.9380	1.8895	1.8380	1.7831
23	2.2747	2.2036	2.1282	2.0476	2.0050	1.9605	1.9139	1.8649	1.8128	1.7570
24	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7897	1.7331
25	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.7110
26	2.2197	2.1479	2.0716	1.9898	1.9464	1.9010	1.8533	1.8027	1.7488	1.6906
27	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7307	1.6717
28	2.1900	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6377
30	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519	1.2539
∞	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.3940	1.3180	1.2214	1.0000

TABLE 7b *

2.5% Points of the F-Distribution

ν_2 (Degrees of Freedom for Denominator) \ ν_1 (Degrees of Freedom for Numerator)	1	2	3	4	5	6	7	8	9
1	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28
2	38.506	39.000	39.165	39.248	39.298	39.331	39.355	39.373	39.387
3	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473
4	12.218	10.649	9.9792	9.6045	9.3645	9.1973	9.0741	8.9796	8.9047
5	10.007	8.4336	7.7636	7.3879	7.1464	6.9777	6.8531	6.7572	6.6810
6	8.8131	7.2598	6.5988	6.2272	5.9876	5.8197	5.6955	5.5996	5.5234
7	8.0727	6.5415	5.8898	5.5226	5.2852	5.1186	4.9949	4.8994	4.8232
8	7.5709	6.0595	5.4160	5.0526	4.8173	4.6517	4.5286	4.4332	4.3572
9	7.2093	5.7147	5.0781	4.7181	4.4844	4.3197	4.1971	4.1020	4.0260
10	6.9367	5.4564	4.8256	4.4683	4.2361	4.0721	3.9498	3.8549	3.7790
11	6.7241	5.2559	4.6300	4.2751	4.0440	3.8807	3.7586	3.6638	3.5879
12	6.5538	5.0959	4.4742	4.1212	3.8911	3.7283	3.6065	3.5118	3.4358
13	6.4143	4.9653	4.3472	3.9959	3.7667	3.6043	3.4827	3.3880	3.3120
14	6.2979	4.8567	4.2417	3.8919	3.6634	3.5014	3.3799	3.2853	3.2093
15	6.1995	4.7650	4.1528	3.8043	3.5764	3.4147	3.2934	3.1987	3.1227
16	6.1151	4.6867	4.0768	3.7294	3.5021	3.3406	3.2194	3.1248	3.0488
17	6.0420	4.6189	4.0112	3.6648	3.4379	3.2767	3.1556	3.0610	2.9849
18	5.9781	4.5597	3.9539	3.6083	3.3820	3.2209	3.0999	3.0053	2.9291
19	5.9216	4.5075	3.9034	3.5587	3.3327	3.1718	3.0509	2.9563	2.8800
20	5.8715	4.4613	3.8587	3.5147	3.2891	3.1283	3.0074	2.9128	2.8365
21	5.8266	4.4199	3.8188	3.4754	3.2501	3.0895	2.9686	2.8740	2.7977
22	5.7863	4.3828	3.7829	3.4401	3.2151	3.0546	2.9338	2.8392	2.7628
23	5.7498	4.3492	3.7505	3.4083	3.1835	3.0232	2.9024	2.8077	2.7313
24	5.7167	4.3187	3.7211	3.3794	3.1548	2.9946	2.8738	2.7791	2.7027
25	5.6864	4.2909	3.6943	3.3530	3.1287	2.9685	2.8478	2.75[illegible]	2.6766
26	5.6586	4.2655	3.6697	3.3289	3.1048	2.9447	2.8240	2.7293	2.6528
27	5.6331	4.2421	3.6472	3.3067	3.0828	2.9228	2.8021	2.7074	2.6309
28	5.6096	4.2205	3.6264	3.2863	3.0625	2.9027	2.7820	2.6872	2.6106
29	5.5878	4.2006	3.6072	3.2674	3.0438	2.8840	2.7633	2.6686	2.5919
30	5.5675	4.1821	3.5894	3.2499	3.0265	2.8667	2.7460	2.6513	2.5746
40	5.4239	4.0510	3.4633	3.1261	2.9037	2.7444	2.6238	2.5289	2.4519
60	5.2857	3.9253	3.3425	3.0077	2.7863	2.6274	2.5068	2.4117	2.3344
120	5.1524	3.8046	3.2270	2.8943	2.6740	2.5154	2.3948	2.2994	2.2217
∞	5.0239	3.6889	3.1161	2.7858	2.5665	2.4082	2.2875	2.1918	2.1136

* This table is reproduced with the permission of Professor E.S. Pearson from Biometrika, vol. 33, pp. 82-83.

TABLE 7 b (continued)

2.5% Points of the F-Distribution

ν_2 \ ν_1	Degrees of Freedom for Numerator									
Degrees of Freedom for Denominator	10	12	15	20	24	30	40	60	120	∞
1	968.63	976.71	984.87	993.10	997.25	1001.4	1005.6	1009.8	1014.0	1018.3
2	39.398	39.415	39.431	39.448	39.456	39.465	39.473	39.481	39.490	39.498
3	14.419	14.337	14.253	14.167	14.124	14.081	14.037	13.992	13.947	13.902
4	8.8439	8.7512	8.6565	8.5599	8.5109	8.4613	8.4111	8.3604	8.3092	8.2573
5	6.6192	6.5246	6.4277	6.3285	6.2780	6.2269	6.1751	6.1225	6.0693	6.0153
6	5.4613	5.3662	5.2687	5.1684	5.1172	5.0652	5.0125	4.9589	4.9045	4.8491
7	4.7611	4.6658	4.5678	4.4667	4.4150	4.3624	4.3089	4.2544	4.1989	4.1423
8	4.2951	4.1997	4.1012	3.9995	3.9472	3.8940	3.8398	3.7844	3.7279	3.6702
9	3.9639	3.8682	3.7694	3.6669	3.6142	3.5604	3.5055	3.4493	3.3918	3.3329
10	3.7168	3.6209	3.5217	3.4186	3.3654	3.3110	3.2554	3.1984	3.1399	3.0798
11	3.5257	3.4296	3.3299	3.2261	3.1725	3.1176	3.0613	3.0035	2.9441	2.8828
12	3.3736	3.2773	3.1772	3.0728	3.0187	2.9633	2.9063	2.8478	2.7874	2.7249
13	3.2497	3.1532	3.0527	2.9477	2.8932	2.8373	2.7797	2.7204	2.6590	2.5955
14	3.1469	3.0501	2.9493	2.8437	2.7888	2.7324	2.6742	2.6142	2.5519	2.4872
15	3.0602	2.9633	2.8621	2.7559	2.7006	2.6437	2.5850	2.5242	2.4611	2.3953
16	2.9862	2.8890	2.7875	2.6808	2.6252	2.5678	2.5085	2.4471	2.3831	2.3163
17	2.9222	2.8249	2.7230	2.6158	2.5598	2.5021	2.4422	2.3801	2.3153	2.2474
18	2.8664	2.7689	2.6667	2.5590	2.5027	2.4445	2.3842	2.3214	2.2558	2.1869
19	2.8173	2.7196	2.6171	2.5089	2.4523	2.3937	2.3329	2.2695	2.2032	2.1333
20	2.7737	2.6758	2.5731	2.4645	2.4076	2.3486	2.2873	2.2234	2.1562	2.0853
21	2.7348	2.6368	2.5338	2.4247	2.3675	2.3082	2.2465	2.1819	2.1141	2.0422
22	2.6998	2.6017	2.4984	2.3890	2.3315	2.2718	2.2097	2.1446	2.0760	2.0032
23	2.6682	2.5699	2.4665	2.3567	2.2989	2.2389	2.1763	2.1107	2.0415	1.9677
24	2.6396	2.5412	2.4374	2.3273	2.2693	2.2090	2.1460	2.0799	2.0099	1.9353
25	2.6135	2.5149	2.4110	2.3005	2.2422	2.1816	2.1183	2.0517	1.9811	1.9055
26	2.5895	2.4909	2.3867	2.2759	2.2174	2.1565	2.0928	2.0257	1.9545	1.8781
27	2.5676	2.4688	2.3644	2.2533	2.1946	2.1334	2.0693	2.0018	1.9299	1.8527
28	2.5473	2.4484	2.3438	2.2324	2.1735	2.1121	2.0477	1.9796	1.9072	1.8291
29	2.5286	2.4295	2.3248	2.2131	2.1540	2.0923	2.0276	1.9591	1.8861	1.8072
30	2.5112	2.4120	2.3072	2.1952	2.1359	2.0739	2.0089	1.9400	1.8664	1.7867
40	2.3882	2.2882	2.1819	2.0677	2.0069	1.9429	1.8752	1.8028	1.7242	1.6371
60	2.2702	2.1692	2.0613	1.9445	1.8817	1.8152	1.7440	1.6668	1.5810	1.4822
120	2.1570	2.0548	1.9450	1.8249	1.7597	1.6899	1.6141	1.5299	1.4327	1.3104
∞	2.0483	1.9447	1.8326	1.7085	1.6402	1.5660	1.4835	1.3883	1.2684	1.0000

TABLE 7c*

1% Points of the F-Distribution

ν_2 \ ν_1 (Degrees of Freedom for Denominator)	Degrees of Freedom for Numerator 1	2	3	4	5	6	7	8	9
1	4052.2	4999.5	5403.3	5624.6	5763.7	5859.0	5928.3	5981.6	6022.5
2	98.503	99.000	99.166	99.249	99.299	99.332	99.356	99.374	99.388
3	34.116	30.817	29.457	28.710	28.237	27.911	27.672	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1016	7.9761
7	12.246	9.5466	8.4513	7.8467	7.4604	7.1914	6.9928	6.8401	6.7188
8	11.259	8.6491	7.5910	7.0060	6.6318	6.3707	6.1776	6.0289	5.9106
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9424
11	9.6460	7.2057	6.2167	5.6683	5.3160	5.0692	4.8861	4.7445	4.6315
12	9.3302	6.9266	5.9526	5.4119	5.0643	4.8206	4.6395	4.4994	4.3875
13	9.0738	6.7010	5.7394	5.2053	4.8616	4.6204	4.4410	4.3021	4.1911
14	8.8616	6.5149	5.5639	5.0354	4.6950	4.4558	4.2779	4.1399	4.0297
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948
16	8.5310	6.2262	5.2922	4.7726	4.4374	4.2016	4.0259	3.8896	3.7804
17	8.3997	6.1121	5.1850	4.6690	4.3359	4.1015	3.9267	3.7910	3.6822
18	8.2854	6.0129	5.0919	4.5790	4.2479	4.0146	3.8406	3.7054	3.5971
19	8.1850	5.9259	5.0103	4.5003	4.1708	3.9386	3.7653	3.6305	3.5225
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567
21	8.0166	5.7804	4.8740	4.3688	4.0421	3.8117	3.6396	3.5056	3.3981
22	7.9454	5.7190	4.8166	4.3134	3.9880	3.7583	3.5867	3.4530	3.3458
23	7.8811	5.6637	4.7649	4.2635	3.9392	3.7102	3.5390	3.4057	3.2986
24	7.8229	5.6136	4.7181	4.2184	3.8951	3.6667	3.4959	3.3629	3.2560
25	7.7698	5.5680	4.6755	4.1774	3.8550	3.6272	3.4568	3.3239	3.2172
26	7.7213	5.5263	4.6366	4.1400	3.8183	3.5911	3.4210	3.2884	3.1818
27	7.6767	5.4881	4.6009	4.1056	3.7848	3.5580	3.3882	3.2558	3.1494
28	7.6356	5.4529	4.5681	4.0740	3.7539	3.5276	3.3581	3.2259	3.1195
29	7.5976	5.4205	4.5378	4.0449	3.7254	3.4995	3.3302	3.1982	3.0920
30	7.5625	5.3904	4.5097	4.0179	3.6990	3.4735	3.3045	3.1726	3.0665
40	7.3141	5.1785	4.3126	3.8283	3.5138	3.2910	3.1238	2.9930	2.8876
60	7.0771	4.9774	4.1259	3.6491	3.3389	3.1187	2.9530	2.8233	2.7185
120	6.8510	4.7865	3.9493	3.4796	3.1735	2.9559	2.7918	2.6629	2.5586
∞	6.6349	4.6052	3.7816	3.3192	3.0173	2.6020	2.6393	2.5113	2.4073

* This table is reproduced with the permission of Professor E.S. Pearson from Biometrika, vol. 33, pp. 84-85.

TABLE 7c (continued)

1% Points of the F-Distribution

ν_2 \ ν_1	Degrees of Freedom for Numerator									
	10	12	15	20	24	30	40	60	120	∞
1	6055.8	6106.3	6157.3	6208.7	6234.6	6260.7	6286.8	6313.0	6339.4	6366.0
2	99.399	99.416	99.432	99.449	99.458	99.466	99.474	99.483	99.491	99.501
3	27.229	27.052	26.872	26.690	26.598	26.505	26.411	26.316	26.221	26.125
4	14.546	14.374	14.198	14.020	13.929	13.838	13.745	13.652	13.558	13.463
5	10.051	9.8883	9.7222	9.5527	9.4665	9.3793	9.2912	9.2020	9.1118	9.0204
6	7.8741	7.7183	7.5590	7.3958	7.3127	7.2285	7.1432	7.0568	6.9690	6.8801
7	6.6201	6.4691	6.3143	6.1554	6.0743	5.9921	5.9084	5.8236	5.7372	5.6495
8	5.8143	5.6668	5.5151	5.3591	5.2793	5.1981	5.1156	5.0316	4.9460	4.8588
9	5.2565	5.1114	4.9621	4.8080	4.7290	4.6486	4.5667	4.4831	4.3978	4.3105
10	4.8492	4.7059	4.5582	4.4054	4.3269	4.2469	4.1653	4.0819	3.9965	3.9090
11	4.5393	4.3974	4.2509	4.0990	4.0209	3.9411	3.8596	3.7761	3.6904	3.6025
12	4.2961	4.1553	4.0096	3.8584	3.7805	3.7008	3.6192	3.5355	3.4494	3.3608
13	4.1003	3.9603	3.8154	3.6646	3.5868	3.5070	3.4253	3.3413	3.2548	3.1654
14	3.9394	3.8001	3.6557	3.5052	3.4274	3.3476	3.2656	3.1813	3.0942	3.0040
15	3.8049	3.6662	3.5222	3.3719	3.2940	3.2141	3.1319	3.0471	2.9595	2.8684
16	3.6909	3.5527	3.4089	3.2588	3.1808	3.1007	3.0182	2.9330	2.8447	2.7528
17	3.5931	3.4552	3.3117	3.1615	3.0835	3.0032	2.9205	2.8348	2.7459	2.6530
18	3.5082	3.3706	3.2273	3.0771	2.9990	2.9185	2.8354	2.7493	2.6597	2.5660
19	3.4338	3.2965	3.1533	3.0031	2.9249	2.8442	2.7608	2.6742	2.5839	2.4893
20	3.3682	3.2311	3.0880	2.9377	2.8594	2.7785	2.6947	2.6077	2.5168	2.4212
21	3.3098	3.1729	3.0299	2.8796	2.8011	2.7200	2.6359	2.5484	2.4568	2.3603
22	3.2576	3.1209	2.9780	2.8274	2.7488	2.6675	2.5831	2.4951	2.4029	2.3055
23	3.2106	3.0740	2.9311	2.7805	2.7017	2.6202	2.5355	2.4471	2.3542	2.2559
24	3.1681	3.0316	2.8887	2.7380	2.6591	2.5773	2.4923	2.4035	2.3099	2.2107
25	3.1294	2.9931	2.8502	2.6993	2.6203	2.5383	2.4530	2.3637	2.2695	2.1694
26	3.0941	2.9579	2.8150	2.6640	2.5848	2.5026	2.4170	2.3273	2.2325	2.1315
27	3.0618	2.9256	2.7827	2.6316	2.5522	2.4699	2.3840	2.2938	2.1984	2.0965
28	3.0320	2.8959	2.7530	2.6017	2.5223	2.4397	2.3535	2.2629	2.1670	2.0642
29	3.0045	2.8685	2.7256	2.5742	2.4946	2.4118	2.3253	2.2344	2.1378	2.0342
30	2.9791	2.8431	2.7002	2.5487	2.4689	2.3860	2.2992	2.2079	2.1107	2.0062
40	2.8005	2.6648	2.5216	2.3689	2.2880	2.2034	2.1142	2.0194	1.9172	1.8047
60	2.6318	2.4961	2.3523	2.1978	2.1154	2.0285	1.9360	1.8363	1.7263	1.6006
120	2.4721	2.3363	2.1915	2.0346	1.9500	1.8600	1.7628	1.6557	1.5330	1.3805
∞	2.3209	2.1848	2.0385	1.8783	1.7908	1.6964	1.5923	1.4730	1.3246	1.0000

Degrees of Freedom for Denominator

TABLE 7d*

0.5% Points of the F-Distribution

ν_2 \ ν_1	Degrees of Freedom for Numerator								
Degrees of Freedom for Denominator	1	2	3	4	5	6	7	8	9
1	16211	20000	21615	22500	23056	23437	23715	23925	24091
2	198.50	199.00	199.17	199.25	199.30	199.33	199.36	199.37	199.39
3	55.552	49.799	47.467	46.195	45.392	44.838	44.434	44.126	43.882
4	31.333	26.284	24.259	23.155	22.456	21.975	21.622	21.352	21.139
5	22.785	18.314	16.530	15.556	14.940	14.513	14.200	13.961	13.772
6	18.635	14.544	12.917	12.028	11.464	11.073	10.786	10.566	10.391
7	16.236	12.404	10.882	10.050	9.5221	9.1554	8.8854	8.6781	8.5138
8	14.688	11.042	9.5965	8.8051	8.3018	7.9520	7.6942	7.4960	7.3386
9	13.614	10.107	8.7171	7.9559	7.4711	7.1338	6.8849	6.6933	6.5411
10	12.826	9.4270	8.0807	7.3428	6.8723	6.5446	6.3025	6.1159	5.9676
11	12.226	8.9122	7.6004	6.8809	6.4217	6.1015	5.8648	5.6821	5.5368
12	11.754	8.5096	7.2258	6.5211	6.0711	5.7570	5.5245	5.3451	5.2021
13	11.374	8.1865	6.9257	6.2335	5.7910	5.4819	5.2529	5.0761	4.9351
14	11.060	7.9217	6.6803	5.9984	5.5623	5.2574	5.0313	4.8566	4.7173
15	10.798	7.7008	6.4760	5.8029	5.3721	5.0708	4.8473	4.6743	4.5364
16	10.575	7.5138	6.3034	5.6378	5.2117	4.9134	4.6920	4.5207	4.3838
17	10.384	7.3536	6.1556	5.4967	5.0746	4.7789	4.5594	4.3893	4.2535
18	10.218	7.2148	6.0277	5.3746	4.9560	4.6627	4.4448	4.2759	4.1410
19	10.073	7.0935	5.9161	5.2681	4.8526	4.5614	4.3448	4.1770	4.0428
20	9.9439	6.9865	5.8177	5.1743	4.7616	4.4721	4.2569	4.0900	3.9564
21	9.8295	6.8914	5.7304	5.0911	4.6808	4.3931	4.1789	4.0128	3.8799
22	9.7271	6.8064	5.6524	5.0168	4.6088	4.3225	4.1094	3.9440	3.8116
23	9.6348	6.7300	5.5823	4.9500	4.5441	4.2591	4.0469	3.8822	3.7502
24	9.5513	6.6610	5.5190	4.8898	4.4857	4.2019	3.9905	3.8264	3.6949
25	9.4753	6.5982	5.4615	4.8315	4.4327	4.1500	3.9394	3.7758	3.6447
26	9.4059	6.5409	5.4091	4.7852	4.3844	4.1027	3.8928	3.7297	3.5989
27	9.3423	6.4885	5.3611	4.7396	4.3402	4.0594	3.8501	3.6875	3.5571
28	9.2838	6.4403	5.3170	4.6977	4.2996	4.0197	3.8110	3.6487	3.5186
29	9.2297	6.3958	5.2764	4.6591	4.2622	3.9830	3.7749	3.6130	3.4832
30	9.1797	6.3547	5.2388	4.6233	4.2276	3.9492	3.7416	3.5801	3.4505
40	8.8278	6.0664	4.9759	4.3738	3.9860	3.7129	3.5088	3.3498	3.2220
60	8.4946	5.7950	4.7290	4.1399	3.7600	3.4918	3.2911	3.1344	3.0083
120	8.1790	5.5393	4.4973	3.9207	3.5482	3.2849	3.0874	2.9330	2.8083
∞	7.8794	5.2983	4.2794	3.7151	3.3499	3.0913	2.8968	2.7444	2.6210

* This table is reproduced with the permission of Professor E.S. Pearson from Biometrika, vol. 33, pp. 86-87.

TABLE 7d (continued)

0.5% Points of the F-Distribution

ν_2 \ ν_1	Degrees of Freedom for Numerator									
Degrees of Freedom for Denominator	10	12	15	20	24	30	40	60	120	∞
1	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465
2	199.40	199.42	199.43	199.45	199.46	199.47	199.47	199.48	199.49	199.51
3	43.686	43.387	43.085	42.778	42.622	42.466	42.308	42.149	41.989	41.829
4	20.967	20.705	20.438	20.167	20.030	19.892	19.752	19.611	19.468	19.325
5	13.618	13.384	13.146	12.903	12.780	12.656	12.530	12.402	12.274	12.144
6	10.250	10.034	9.8140	9.5888	9.4741	9.3583	9.2408	9.1219	9.0015	8.8793
7	8.3803	8.1764	7.9678	7.7540	7.6450	7.5345	7.4225	7.3088	7.1933	7.0760
8	7.2107	7.0149	6.8143	6.6082	6.5029	6.3961	6.2875	6.1772	6.0649	5.9505
9	6.4171	6.2274	6.0325	5.8318	5.7292	5.6248	5.5186	5.4104	5.3001	5.1875
10	5.8467	5.6613	5.4707	5.2740	5.1732	5.0705	4.9659	4.8592	4.7501	4.6385
11	5.4182	5.2363	5.0489	4.8552	4.7557	4.6543	4.5508	4.4450	4.3367	4.2256
12	5.0855	4.9063	4.7214	4.5299	4.4315	4.3309	4.2282	4.1229	4.0149	3.9039
13	4.8199	4.6429	4.4600	4.2703	4.1726	4.0727	3.9704	3.8655	3.7577	3.6465
14	4.6034	4.4281	4.2468	4.0585	3.9614	3.8619	3.7600	3.6553	3.5473	3.4359
15	4.4236	4.2498	4.0698	3.8826	3.7859	3.6867	3.5850	3.4803	3.3722	3.2602
16	4.2719	4.0994	3.9205	3.7342	3.6378	3.5388	3.4372	3.3324	3.2240	3.1115
17	4.1423	3.9709	3.7929	3.6073	3.5112	3.4124	3.3107	3.2058	3.0971	2.9839
18	4.0305	3.8599	3.6827	3.4977	3.4017	3.3030	3.2014	3.0962	2.9871	2.8732
19	3.9329	3.7631	3.5866	3.4020	3.3062	3.2075	3.1058	3.0004	2.8908	2.7762
20	3.8470	3.6779	3.5020	3.3178	3.2220	3.1234	3.0215	2.9159	2.8058	2.6904
21	3.7709	3.6024	3.4270	3.2431	3.1474	3.0488	2.9467	2.8408	2.7302	2.6140
22	3.7030	3.5350	3.3600	3.1764	3.0807	2.9821	2.8799	2.7736	2.6625	2.5455
23	3.6420	3.4745	3.2999	3.1165	3.0208	2.9221	2.8198	2.7132	2.6015	2.4837
24	3.5870	3.4199	3.2456	3.0624	2.9667	2.8679	2.7654	2.6585	2.5463	2.4276
25	3.5370	3.3704	3.1963	3.0133	2.9176	2.8187	2.7160	2.6088	2.4960	2.3765
26	3.4916	3.3252	3.1515	2.9685	2.8728	2.7738	2.6709	2.5633	2.4501	2.3297
27	3.4499	3.2839	3.1104	2.9275	2.8318	2.7327	2.6296	2.5217	2.4078	2.2867
28	3.4117	3.2460	3.0727	2.8899	2.7941	2.6949	2.5916	2.4834	2.3689	2.2469
29	3.3765	3.2111	3.0379	2.8551	2.7594	2.6601	2.5565	2.4479	2.3330	2.2102
30	3.3440	3.1787	3.0057	2.8230	2.7272	2.6278	2.5241	2.4151	2.2997	2.1760
40	3.1167	2.9531	2.7811	2.5984	2.5020	2.4015	2.2958	2.1838	2.0635	1.9318
60	2.9042	2.7419	2.5705	2.3872	2.2898	2.1874	2.0789	1.9622	1.8341	1.6885
120	2.7052	2.5439	2.3727	2.1881	2.0890	1.9839	1.8709	1.7469	1.6055	1.4311
∞	2.5188	2.3583	2.1868	1.9998	1.8983	1.7891	1.6691	1.5325	1.3637	1.0000