การทดสอบสมมติฐานทางสถิติ

Appearances to the mind are of four kinds:
Things either are what they appear to be ,
Or they either are, nor appear to be ,
Or they are and do not appear to be ,
Or they are not. and yet appear lo be
Rightly to aim in all these cases is the wise man's task
Greek Philosopher EPICTETUS

ในวันที่ท้องฟ้าปกคลุมไปด้วยเมฆ นาย ก มองออกไปทางหน้าต่าง และพูดว่า "วันนี้ฝนจะตก" คำกล่าวที่นาย ก กล่าวนี้จะเป็นคำกล่าวที่พยากรณ์อากาศในวันนั้น และคำกล่าวนี้แหละจะเป็นสมมติฐาน (Hypothesis) เกี่ยวกับอากาศ นาย ก กำลังประสบกับปัญหาตัดสินใจในแง่ที่ว่า "เขาควรจะนำเสื้อกันฝนติดตัวไปด้วยหรือไม่" ไม่ว่านาย ก จะตัดสินใจอย่างไรก็จะมีผลอยู่ในสภาวะอย่างหนึ่งใน 4 อย่าง ดังนี้

(1) เขานำเสื้อฝนติดตัวไปด้วย และฝนตก (ตัดสินใจถูก)
(2) เขาไม่นำเสื้อกันฝนติดตัวไปด้วย และฝนก็ไม่ตก (ตัดสินใจถูกอีก)
(3) เขานำเสื้อกันฝนติดตัวไปด้วย แต่ฝนไม่ตก (ตัดสินใจผิด)
และ (4) เขาไม่นำเสื้อกันฝนติดตัวไปด้วย แต่ฝนตก (ตัดสินใจผิดอีก)

สองสภาวะสุดท้ายไม่เป็นที่ปรารถนาของเขา เพราะเป็นการตัดสินใจที่เคลื่อนคลาด โดยอุมคติแล้วนาย ก อยากจะมีเสื้อฝนติดตัวไปด้วยถ้าฝนตก และไม่นำติดตัวไปด้วยถ้าฝนไม่ตก แต่เขาไม่ทราบว่าฝนจะตกหรือไม่ เมื่อเริ่มวันใหม่ และเขาจะต้องตัดสินใจก่อนเริ่มวันใหม่ ก่อนที่เขาจะออกจากบ้าน เขาจะต้องคิดถึงองค์ประกอบในการทำนายอากาศ โดยใช้ประสบการณ์ และทำให้ผลลัพธ์ของการตัดสินใจผิดอยู่ในสมดุลย์ ถ้าเขาคิดว่าจะทำให้เกิดโรคปอดบวม เขาก็ชอบที่จะคลาดเคลื่อนในทางที่จะนำเสื้อกันฝนติดตัวไปด้วย ถึงแม้ฝนจะไม่ตกในวันนั้น ในทางตรงกันข้าม ถ้าเขามีร่างกายสมบูรณ์ แต่ชอบหลงลืมเสื้อกันฝน เขาควรจะไม่นำเสื้อกันฝนติดตัวไปด้วย

ที่กล่าวมานี้เป็นแนวความคิดในเรื่องการทดสอบสมมติฐานทางสถิติ

6.1 ความหมายของสมมติฐาน (What is a hypothesis ?)

คำว่า สมมติฐาน (Hypothesis) นั้นโดยทั่วไปจะหมายถึงทฤษฎีหรือข้อเสนอ (Proposition) ที่เสนอขึ้นมาใช้ชั่วคราวเพื่ออธิบายความจริงบางอย่าง และใช้เป็นแนวทางในการสืบสวนค้นคว้าเรื่องอื่น ๆ สมมติฐานนั้นอาจจะเป็นจริงหรือไม่ก็ได้

สำหรับ สมมติฐานทางสถิติ (Statistical Hypothesis) นั้นจะเป็นข้อเสนอหรือคำกล่าว (Statement, Assertion, or Claim) เกี่ยวกับธรรมชาติของประชากรที่ต้องการจะศึกษาและสำรวจ

โดยปกติจะเป็นคำกล่าวเกี่ยวกับ

(1) คุณลักษณะหรือพารามิเตอร์ประชากร

(2) การแจงประชากร

(3) ทั้งการแจกแจงและพารามิเตอร์

เช่นรายได้เฉลี่ยต่อครอบครัวต่อปีของชาวนาไทยประมาณ 4000 บาท (μ = 4000) ลูกเต๋าลูกนี้สมดุลย์ (π = 1/6, π เป็นโอกาสที่หน้าใดหน้าหนึ่งจะเกิดขึ้น) เปอร์เซ็นต์ของเด็กในเมืองที่ได้รับการศึกษาระดับสูง ๆ จะมากกว่าเด็กชนบท ($\pi_U > \pi_R$) อายุเฉลี่ยของคนไทยมีการแจกแจงแบบปกติ ระดับสติปัญญาของนักเรียน ป. 4 มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย 100 เป็นต้น

สมมติฐานทางสถิติที่เกี่ยวกับพารามิเตอร์ประชากรนั้น ส่วนมากเราเพียงแต่ระบุค่าของพารามิเตอร์เฉพาะตัวที่สนใจเท่านั้น อาจจะเป็นหนึ่งหรือมากกว่าก็ได้ แต่ไม่จำเป็นต้องระบุทุกตัว

วิธีการที่แน่นอนเพื่อจะค้นพบว่าสมมติฐานนั้นเป็นความจริงหรือเป็นเท็จ (Truth of Falsity) ก็อาศัยการสำรวจประชากรทั้งหมด ซึ่งวิธีการดังกล่าวนั้นก็ไม่สะดวกเสมอไป และบางครั้งการศึกษาประชากรทั้งหมดยังพบความคลาดเคลื่อนอีกด้วย ดังนั้นจึงต้องอาศัยตัวอย่างมาสรุปผล ในทางสถิติเราจึงไม่พิสูจน์สมมติฐานว่าเป็นจริงหรือเป็นเท็จแต่จะดูความเป็นไปได้ของสมมติฐานโดยอาศัยตัวอย่างมาสนับสนุน

สมมติฐานที่จะทดสอบเพื่อจุดประสงค์ของการปฏิเสธ ซึ่งเป็นถ้อยแถลงหรือข้อความที่ขึ้นอยู่กับวิธีการทดสอบ (Test Procedure) นั้นจะเรียกว่า สมมติฐานหลัก (Null Hypothesis) ซึ่งมักจะแทนด้วย Ho เช่น ถ้าสมมติฐานหลักกล่าวว่าเหรียญยุติธรรมแล้วสมมติฐานหลักจะเขียนได้เป็น

$$Ho : \pi = 1/2$$

หรือสมมติฐานหลักกล่าวว่า อายุเฉลี่ยของเครื่องจักรที่ผลิตในประเทศไทยจะมีอายุใช้งานอย่างน้อย 5 ปี แล้วจะเขียนได้เป็น

$$Ho : \mu \geq 5$$

สมมติฐานที่ขัดแย้งกับสมมติฐานหลักอันเป็นคำกล่าวอย่างอื่นที่เป็นไปได้ทั้งหมดและไม่รวมหรือคาบเกี่ยวกับสมมติฐานหลัก Ho จะเรียกว่า สมมติฐานรอง (Alternative Hypothesis) ซึ่งแทนด้วย Ha สมมติฐานรองนี้จะยอมรับเมื่อสมมติฐานหลักได้รับการปฏิเสธ จากตัวอย่างข้างบนสมมติฐานรองที่ขัดแย้งกับสมมติฐานหลักจะเป็นอย่างใดอย่างหนึ่ง ดังนี้

$$Ha : \pi \neq 1/2 \ , \quad Ha : \pi \neq 3/4$$
$$Ha : \pi > 2/3 \ , \quad Ha : \pi < 1/2 \ , \ ..$$

สำหรับปัญหาการทดสอบนั้น สมมติฐานไหนจะเป็นสมมติฐานหลัก และสมมติฐานไหนจะเป็นสมมติฐานรองก็เป็นคำถามที่สำคัญมาก และเราจะต้องทำกับหลักเหตุผลของวิธีการทดสอบทางสถิติ (Statistical test procedure) ในการสร้างแบบทดสอบสมมติฐาน แนวความคิดหลักก็คือการพิสูจน์โดยการโต้แย้ง (Proof by contradiction) ตัวอย่างเช่น ครูสงสัยว่าเด็กที่มาจากครอบครัวที่มีเศรษฐกิจดี จะเรียนหนังสือดีกว่าเด็กที่มาจากครอบครัวที่มีเศรษฐกิจด้อยกว่า นี่ก็เป็นความเชื่อของครู ซึ่งเขาสามารถสร้างสมมติฐานและทำการทดสอบได้ ภายใต้สมมติฐานหลัก Ho เราจะพูดว่า มันไม่เป็นกรณีนั้น และกล่าวว่า

Ho : เด็กที่มาจากครอบครัวที่มีเศรษฐกิจดีจะเรียนหนังสือได้ไม่ดีกว่า

Ha : เด็กที่มาจากครอบครัวที่มีเศรษฐกิจดีจะเรียนหนังสือได้ดีกว่าหรือโดยสัญญลักษณ์

$$\text{Ho}: \mu_H \leq \mu_L \quad ; \quad \text{Ha}: \mu_H > \mu_L$$

ถ้าประจักษ์พยานจากตัวอย่างนำไปสู่การปฏิเสธสมมติฐานหลัก Ho นั่นคือขัดแย้งกับสมมติฐานหลัก หรือสนับสนุนสมมติฐานรอง Ha นั่นเอง แล้วเราจะยอมรับ Ha

หลักสำคัญของการกำหนดสมมติฐานหลักและรองนั้นก็คือคำถามที่ว่า "เด็กที่มาจากครอบครัวที่มีเศรษฐกิจดีจะเรียนหนังสือดีกว่าเด็กที่มาจากครอบครัวที่มีเศรษฐกิจด้อยกว่าหรือ ?" ดังนั้นสมมติฐานหลักและสมมติฐานรองจะพิจารณาได้จากคำถามที่เป็นคำกล่าวของปัญหานั้น ๆ

สมมติฐานทางสถิติอันประกอบด้วยสมมติฐานหลัก Ho และสมมติฐานรอง Ha นั้นมีสิ่งที่น่าสนใจพิเศษ ดังนี้

(1) สมมติฐานหลัก Ho จะเป็นคำกล่าวที่มีคำว่า "เท่ากับ =" รวมไว้ด้วยเสมออันเป็นคำกล่าวที่บ่งถึงการเท่ากัน หรือการไม่แตกต่างกัน นั่นเอง

(2) สมมติฐานหลัก Ho และ Ha จะไม่รวมค่าของตัวสถิติไว้เป็นอันขาด และ

(3) สมมติฐานรอง Ha นั้น ตามปกติจะเป็นคำกล่าวที่ผู้ทำการทดสอบ หรือผู้กล่าวข้อความนั้นมีความสนใจ

ลองพิจารณาสมมติฐานหลัก Ho และรอง Ha จากคำกล่าวหรือสมมติฐานทางสถิติต่อไปนี้

ก. รายได้เฉลี่ยต่อปีของครอบครัวชาวนาไทยไม่น่าจะเท่ากับ 7,000 บาท

$$\text{Ho}: \mu = 7{,}000\ ;\ \text{Ha}: \mu \neq 7{,}000$$

ข. รายได้เฉลี่ยต่อปีของครอบครัวชาวนาไทยมากกว่า 6,000 บาท

$$\text{H0}: \mu \leq 6{,}000\ ;\ \text{Ha}: \mu > 6{,}000$$

ค. รายได้เฉลี่ยต่อปีของครอบครัวชาวนาไทยน้อยกว่า 7,000 บาท

$$Ho : \mu \geq 7{,}000 \; ; \; Ha: \mu < 7{,}000$$

สมมติฐานทางสถิตินั้น เราสามารถแบ่งออกได้เป็น 2 ประเภท ตามค่าที่ระบุไว้ของพารามิเตอร์ ว่าเป็นค่าเดี่ยว ๆ หรือเป็นช่วง ดังนี้

(1) สมมติฐานเชิงเดี่ยว (Simple Hypothesis) จะเป็นถ้อยแถลงถึงค่าเดี่ยว ๆ ของพารามิเตอร์ในประชากรที่เราสนใจ เช่นรายได้เฉลี่ยต่อเดือนของกรรมกรท่าเรือเท่ากับ 1,500 บาท นักศึกษามหาวิทยาลัยที่เป็นลูกชาวนามีเพียง 5% เชาวน์ปัญญาของนักศึกษามหาวิทยาลัยมีการแจกแจงปกติที่มีค่าเฉลี่ย 110 และส่วนเบี่ยงเบนมาตรฐาน 10 เป็นต้น

(2) สมมติฐานเชิงผสม (Composite Hypothesis) จะเป็นถ้อยแถลงถึงค่าของพารามิเตอร์ในประชากรที่สนใจเป็นแบบช่วงหรือพิสัย เช่น สัมประสิทธิ์สหสัมพันธ์ระหว่างเชาวน์ปัญญากับผลการเรียนสำหรับนักศึกษารามคำแหงน้อยกว่า 0.50 อายุเฉลี่ยของคนไทยอย่างน้อย 50 ปี ประชาชนไทยเห็นด้วยกับแนวความคิดในการวางแผนครอบครัวระหว่าง 70 ถึง 85 เปอร์เซ็นต์ รายเฉลี่ยต่อครอบครัว (ต่อปี) ของชาวนาไทยประมาณ 6,000 บาท แต่ไม่ทราบค่าส่วนเบี่ยงเบนมาตรฐาน เป็นต้น

ในบางกรณีสมมติฐานที่เราตั้งขึ้นนั้นจะไม่มีข้อสมมติ (Assumptions) เกี่ยวกับรูปแบบหรือการแจกแจงของประชากร สมมติฐานเหล่านี้เรียกว่า สมมติฐานที่เป็นอิสระจากรูปการแจกแจงหรือสมมติฐานที่ไม่เกี่ยวกับพารามิเตอร์ (Distribution-Free or Nonparametric Hypothesis) สมมติฐานแบบนี้เราจะได้ทดสอบต่อไปในบทหลัง

แบบทดสอบ (Test) ของสมมติฐานทางสถิติ นั้นจะเป็นกฎเกณฑ์ (Rule) หรือกระบวนการที่นำไปสู่การตัดสินใจว่าจะยอมรับหรือปฏิเสธสมมติฐานที่พิจารณา เมื่อได้รับค่าจากตัวอย่างที่ทดลอง (Experimental Sample Values) เกณฑ์นี้ได้กำหนดก่อนที่จะสุ่มตัวอย่าง และเรียกกันว่า "เกณฑ์ตัดสินใจ (Decision Rule)" ถ้าประจักษ์พยานที่รวบรวมจากตัวอย่างไม่สอดคล้องกับคำกล่าวภายใต้สมมติฐานหลัก Ho ก็จะปฏิเสธ Ho และสรุปว่า Ho เป็นเท็จ จงระลึกไว้เสมอว่าเมื่อปฏิเสธสมมติฐานหลัก Ho นั้น เรามิได้พิสูจน์ว่า Ho เป็นเท็จ นั่นหมายความว่าตัวอย่างที่ได้ไม่สอดคล้องกับคำกล่าวที่กำหนดในสมมติฐานหลัก Ho และประจักษ์พยานนั้นนำเราไปสู่การสรุปที่ว่าสมมติฐานเป็นเท็จ

ตัวสถิติ หรือฟังก์ชั่นของค่าสังเกตจากตัวอย่างที่ใช้เป็นเครื่องมือในการตัดสินใจเกี่ยวกับสมมติฐานหลัก Ho นั้นจะเรียกว่า ตัวสถิติทดสอบ (Test Statistic) กลุ่มหรือเซ็ทของค่าที่เป็นไปได้ทั้งหมดของตัวสถิติทดสอบ จะเรียกว่า กลุ่มผลทดลองของการทดสอบ บางค่าในกลุ่มผลทดลองนี้เกือบจะไม่มีโอกาสเกิดขึ้นเลย ถ้าสมมติฐานหลัก Ho เป็นจริง เซ็ทของค่าเหล่านี้ซึ่งเป็นเซ็ทย่อยของกลุ่มผลทดลองจะใช้เป็นค่าที่จะนำไปสู่การตัดสินใจที่จะปฏิเสธสมมติฐานหลัก และจะ

เรียกว่า เขตวิกฤต (Critical Region) หรือเขตปฏิเสธ (Rejection region) ดังนั้นสมมติฐานหลัก Ho จะได้รับการปฏิเสธถ้าค่าของตัวสถิติทดสอบอยู่ในเซ็ทย่อยที่เกือบจะไม่มีโอกาสเกิดขึ้นเลย และจะยอมรับ Ho ถ้าค่าของตัวสถิติทดสอบอยู่ในเซ็ทย่อยที่มีโอกาสเกิดขึ้นมาก หรืออยู่ในเขตยอมรับ (Acceptance Region) นั่นเอง

6.2 ความคลาดเคลื่อนในการทดสอบสมมติฐาน (Error in Hypothesis Testing)

ในการตัดสินใจที่จะยอมรับหรือปฏิเสธสมมติฐานหลัก Ho ที่ตั้งขึ้นโดยอาศัยตัวอย่างเป็นเครื่องมือนั้น เราอาจจะตัดสินใจผิดก็ได้ นั่นคือจะมีความคลาดเคลื่อน (Error) เกิดขึ้นนั่นเอง ความคลาดเคลื่อนที่เกิดขึ้นก็เนื่องจากตัวอย่างมีความคลาดเคลื่อนที่เรียกกันว่า ความคลาดเคลื่อนตัวอย่าง (Sampling error) ซึ่งเป็นคุณลักษณะประจำตัวของตัวอย่างสุ่มนั้นปะปนอยู่ด้วย บางครั้งตัวอย่างก็มีความคลาดเคลื่อนอย่างอื่นเข้ามาปะปนอยู่ด้วย ความคลาดเคลื่อนในการทดสอบสมมติฐานจึงเกิดขึ้นได้จากความเป็นไปได้ของสมมติฐานหลักที่ว่า "อาจจะเป็นจริง หรือเป็นเท็จ" ดังนี้

(1) ถ้า Ho เป็นจริง เราสามารถทำการตัดสินใจอย่างหนึ่งใน 2 อย่างคือยอมรับ Ho ว่าเป็นจริง ในกรณีนี้เราจะทำการตัดสินใจถูกต้อง หรือปฏิเสธ Ho และสรุปว่าไม่เป็นจริง ในกรณีนี้เราจะทำการตัดสินใจไม่ถูกต้อง ความคลาดเคลื่อนของการปฏิเสธสมมติฐานหลัก ทั้ง ๆ ที่เป็นจริงนี้จะเรียกว่า "ความคลาดเคลื่อนแบบ 1 (Type I Error or Alpha Error) หรือความคลาดเคลื่อนจากการปฏิเสธ (Rejection Error) " Error " ความน่าจะเป็นที่จะกระทำความคลาดเคลื่อนแบบนี้จะเรียกว่า "การเสี่ยงแบบ 1 (Alpha Risk), ระดับนัยสำคัญ (Level of Significance) หรือขนาดของการทดสอบ (Size of Test)" เรามักแทนด้วย α

$$\begin{aligned}\alpha &= P(\text{ความคลาดเคลื่อนแบบ } 1)\\ &= P(\text{ปฏิเสธ Ho ในเมื่อ Ho เป็นจริง})\end{aligned}$$

ดังนั้นเราจะได้สภาวะต่อไปนี้

Ho เป็นจริง
- ยอมรับ Ho : ตัดสินใจถูกต้อง (ความน่าจะเป็น $1 - \alpha$)
- ปฏิเสธ Ho : ตัดสินใจไม่ถูกต้อง (ความน่าจะเป็น α)

(2) ถ้า Ho เป็นเท็จ เราก็สามารถทำการตัดสินใจอย่างใดอย่างหนึ่งใน 2 อย่างเช่นกัน นั่นคือไม่ปฏิเสธ Ho ก็ยอมรับ Ho การปฏิเสธ Ho จะเป็นผลทำให้การตัดสินใจถูกต้อง แต่การ

ยอมรับ Ho จะหมายความว่า เราได้ทำความคลาดเคลื่อน (ทำการตัดสินใจไม่ถูกต้อง) ความคลาดเคลื่อนในการยอมรับ Ho เมื่อ Ho เป็นเท็จนี้เรียกว่า "ความคลาดเคลื่อนแบบ 2 (Type II Error or Beta Error) หรือ ความคลาดเคลื่อนจากการยอมรับ (Acceptance Error) ความน่าจะเป็นของความคลาดเคลื่อนแบบนี้ เรียกว่า "การเสี่ยงแบบ 2 (Beta Risk)" ซึ่งจะแทนด้วย β ดังนั้น

$$\beta = P(\text{ความคลาดเคลื่อนแบบ 2})$$
$$= P(\text{ยอมรับ Ho ในเมื่อ Ho เป็นเท็จ})$$

สำหรับสภาวะที่กล่าวมาจะเป็นดังนี้

Ho เป็นเท็จ
- ยอมรับ Ho : ตัดสินใจไม่ถูกต้อง (ความน่าจะเป็น β)
- ปฏิเสธ Ho : ตัดสินใจถูกต้อง (ความน่าจะเป็น $1-\beta$)

วิธีการทดสอบทางสถิติจะนำไปสู่ความเป็นไปได้อย่างใดอย่างหนึ่งจาก 4 อย่าง ดังกล่าวมาแล้ว โดยที่ 2 อย่าง เป็นการตัดสินใจถูกต้อง และอีก 2 อย่าง จะไม่ถูกต้องความเป็นไปได้เหล่านี้พร้อมทั้งความน่าจะเป็นสรุปได้ดังนี้

	การตัดสินใจ (Decision)	ยอมรับ Ho	ปฏิเสธ Ho (ยอมรับ Ha)
สภาวะที่แท้จริง (True Situation)	Ho เป็นจริง	No Error ($1-\alpha$) Confidence Level	Type I Error (α) Significance Level
	Ho เป็นเท็จ (Ha เป็นจริง)	Type II Error (β)	No Error ($1-\beta$) Power

ในการทดสอบสมมติฐานนั้นเรามุ่งหวังที่จะให้ความคลาดเคลื่อนทั้งสองแบบมีโอกาสเกิดขึ้นน้อย ๆ นั่นคือต้องการให้ α และ β มีค่าน้อยนั่นเอง การที่ทั้ง α และ β จะมีค่าน้อยได้ก็ต้องอาศัยตัวอย่างสุ่มขนาดใหญ่ ($n \rightarrow \infty$) และถ้าต้องการให้ทั้ง α และ β เป็นศูนย์ เราก็ต้องสำรวจทั้งประชากร แต่ในทางปฏิบัติเราไม่สามารถใช้ตัวอย่างขนาดใหญ่ได้ เราเพียงแต่กำหนดตัวอย่างขนาดหนึ่ง (n) ดังนั้นเราจึงทำให้ทั้ง α และ β มีค่าน้อยพร้อมกันไม่ได้ สำหรับ α และ β นี้ก็มีความสัมพันธ์กันในแง่ที่ว่า ถ้า α มีค่ามากแล้ว β จะมีค่าน้อย หรือถ้า α มีค่าน้อยแล้ว β จะมีค่ามาก

โดยทั่วไปในการทดสอบสมมติฐาน เรามักจะกะ α ไว้ล่วงหน้าก่อนที่จะรวบรวม ตัวอย่าง

ในทางปฏิบัติเรามักให้ α = 0.01, 0.05, หรือ 0.01 ซึ่งเป็นค่าน้อย ๆ เพราะถือว่าความคลาดเคลื่อนแบบ 1 นี้จะทำให้ผู้ทดสอบได้รับความสูญเสียมากกว่า และพยายามหาวิธีการทดสอบ (Test Procedure) หรือตัวสถิติทดสอบ (Test Statistic) ที่ดีเพื่อทำให้ β มีค่าน้อย ๆ หรือทำให้ $(1-\beta)$ มีค่ามาก ๆ นั่นเอง

สำหรับ $(1-\beta)$ นี้เรียกว่า อำนาจทดสอบ (Power of Test) ซึ่งเป็นความน่าจะเป็นที่วิธีการทดสอบจะตรวจพบสมมติฐานหลัก Ho ที่เป็นเท็จ หรือเป็นความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลักที่เป็นเท็จ นั่นคือ

อำนาจทดสอบ = $1-\beta$

= P(ปฏิเสธ Ho ในเมื่อ Ho เป็นจริง)

= P(ตัวสถิติทดสอบอยู่ในเขตวิกฤต ในเมื่อ Ha เป็นจริง)

ลองพิจารณาตัวอย่างต่อไปนี้ประกอบคำอธิบายต่าง ๆ ที่ได้กล่าวมาทั้งหมด พร้อมทั้งได้แสดงวิธีการคำนวณค่าต่างไว้ด้วย

ตัวอย่าง บริษัทผลิตสินค้าจะทำการรับมอบชิ้นส่วนสินค้าจากโรงงานผลิตแห่งหนึ่งโดยทำสัญญาส่งมอบกันคราวละ 500000 ชิ้น เจ้าหน้าที่ตรวจรับของตั้งเกณฑ์ในการรับชิ้นส่วนไว้ว่า "ถ้าค่าเฉลี่ยของเส้นรอบวงที่ผิดขนาดไปจากมาตรฐานไม่เกิน 0.03 มม. จะรับชิ้นส่วนเหล่านั้น" นั่นคือ ถ้า X เป็นความยาวของเส้นรอบวงที่ผิดขนาดไปจากมาตรฐานแล้วเจ้าหน้าที่ฝ่ายตรวจรับตั้งเกณฑ์ในการรับชิ้นส่วนไว้ดังนี้

จะยอมรับชิ้นส่วนถ้าค่าเฉลี่ย (μ) น้อยกว่า 0.03 มม.

จะไม่ยอมรับ ถ้าค่าเฉลี่ยมากกว่า 0.03 มม.

เจ้าหน้าที่ตรวจรับจะรับชิ้นส่วนหรือไม่นั้นกระทำได้โดยการวัดเส้นรอบวงของชิ้นส่วนทุกชิ้นและคำนวณหาค่าเฉลี่ย แล้วลบด้วยมาตรฐาน ถ้าผลต่างมากกว่า 0.03 มม. ก็จะไม่รับชิ้นส่วนเหล่านั้น แต่การที่จะวัดเส้นรอบวงของชิ้นส่วนทั้งหมดนั้นกระทำไม่ได้เลยในทางปฏิบัติ เพราะต้องเสียเวลาและค่าใช้จ่ายมาก

ดังนั้นวิธีหนึ่งที่กระทำได้คือใช้ตัวอย่างสุ่มหรือตัวแทนขนาดหนึ่ง (n) จากชิ้นส่วนทั้งหมดที่ส่งมา แล้ววัดเส้นรอบวงของตัวอย่างชิ้นส่วนเหล่านั้น เพื่อหาส่วนที่ผิดขนาด และคำนวณหาค่าเฉลี่ยที่ผิดขนาด ($\bar{X}$) เราหวังว่า $\bar{X}$ จะมีค่าใกล้ ๆ กับ μ ดังนั้นเราจึงสรุปได้ว่า $\mu \leq 0.03$ ก็ต่อเมื่อ $\bar{X} \leq$ ค่าคงที่ซึ่งกำหนดไว้ (ซึ่งควรจะมากกว่า 0.03) และค่าคงที่นั้นกำหนดกันอย่างไร จะพิจารณากันต่อไป

เราจะเห็นได้ว่า การตัดสินใจที่จะยอมรับหรือปฏิเสธชิ้นส่วนที่ส่งมาให้นั้น ที่แท้จริงจะขึ้นอยู่กับตัวอย่างหรือผลของการทดลอง และสังเกตค่าเฉลี่ยตัวอย่าง $\bar{X}$ ค่าเฉลี่ยตัวอย่างจะ

เป็นเกณฑ์ (Criterion) ในการตัดสินใจว่า วิธีการหรือกระบวนการ (Procedure) เช่นนั้นจะนำไปสู่การตัดสินใจที่ดี (Optimal decision) หรือไม่ ถ้าดีแล้วขนาดตัวอย่าง (n) ควรจะเป็นเท่าใด? ถ้าไม่ดีวิธีการอื่นใดจะนำไปสู่การตัดสินใจที่ดี วิธีการอื่นนั้นอาจจะเป็นค่าสังเกตที่มีค่ามากสุด (X_{max}) จากตัวอย่างสุ่มขนาด n ถ้า $X_{max} \leq$ ค่าคงที่ซึ่งกำหนดไว้ เราก็จะสรุปว่า $\mu \leq 0.03$ นั่นคือ เราจะยอมรับชิ้นส่วนสินค้า แต่ถ้าเป็นอย่างอื่นเราก็ไม่รับ สมมติว่าวิธีการเช่นนี้ดีกว่าแบบก่อน แล้วขนาดตัวอย่าง n ควรจะมีค่าเท่าใด? การเลือกวิธีการที่ดี (Optimal Procedure) นั้น เราใช้อะไรเป็นพื้นฐาน เราจะกล่าวถึงต่อไป

จากตัวอย่างนี้เราจะมีสมมติฐานทางสถิติ 2 อย่าง ดังนี้ (1) สมมติฐานแรก คือ $\mu \leq 0.03$ และ (2) สมมติฐานที่สอง คือ $\mu > 0.03$กระบวนการหรือวิธีการที่ทำให้เกิดการเลือกระหว่างสองสมมติฐานนี้เรียกว่า "การทดสอบสมมติฐานทางสถิติ"นั่นเอง

สมมติฐานที่จะทำการทดสอบ ($\mu \ngtr 0.03$) นี้เรียกว่า สมมติฐานหลัก Ho และสมมติฐานอื่น ($\mu > 0.03$) ก็เรียกว่า สมมติฐานรอง Ha นั่นเอง ดังนั้นการทดสอบสมมติฐานทางสถิติจะเกี่ยวข้องกับการปฏิเสธหรือยอมรับสมมติฐานหลัก Ho หรือกล่าวได้อีกอย่างหนึ่งว่า การทดสอบสมมติฐานทางสถิติจะเป็นการพิจารณาหรือตัดสินใจว่า สมมติฐานหลัก Ho เป็นจริง หรือเป็นเท็จ ในทางสัญญลักษณ์ เราเขียนสมมติฐานทั้งสองได้เป็น

$$\text{Ho} : \mu \leq 0.03 \; ; \; \text{Ha} : \mu > 0.03$$

การตัดสินใจที่จะยอมรับหรือปฏิเสธสมมติฐานหลัก Ho จะขึ้นอยู่กับผลของการทดลอง หรือค่าสังเกตจากตัวอย่าง สมมติว่าเลือกตัวอย่างขนาด n มาได้ และคำนวณค่าเฉลี่ย $\bar{X}$ กับตั้งเกณฑ์ตัดสินใจไว้ว่า "ยอมรับ Ho ก็ต่อเมื่อ $\bar{X} \leq 0.035$ และจะปฏิเสธ ถ้าได้ $\bar{X}$ เป็นอย่างอื่น" ตามเกณฑ์นี้เราจะปฏิเสธ Ho ก็ต่อเมื่อผลทดลองมี $\bar{X}$ มากกว่า 0.035

ตามเกณฑ์ตัดสินใจนี้เราจะปฏิเสธสมมติฐานหลัก ถ้าค่าเฉลี่ยตัวอย่าง $\bar{X}$ ตกอยู่ในเขตวิกฤตหรือเขตปฏิเสธ เขตวิกฤตนั้นกำหนดจากเซทของค่า $\bar{X}$ ที่มากกว่า 0.035 เพื่อความสะดวกในการวิเคราะห์ เราจะสมมติว่าตัวแปร $\bar{X}$ มีการแจกแจงแบบปกติที่ไม่ทราบค่าเฉลี่ย (μ) แต่ทราบส่วนเบี่ยงเบนมาตรฐาน $\sigma = 0.006$ สำหรับ X ซึ่งก็เป็นตัวแปรเชิงสุ่มโดยมีการแจกแจงแบบปกติที่มีค่าเฉลี่ย μ และส่วนเบี่ยงเบนมาตรฐาน $\sigma/\sqrt{n}$ เมื่อเราใช้ผลทดลองเป็นเครื่องมือในการตัดสินใจและใช้เกณฑ์ตัดสินใจที่กล่าวมาแล้วนั้น เราจะพบว่า

ก. ถ้าค่าเฉลี่ยของเส้นรอบวงที่ผิดพลาด (μ) นั้น จริง ๆ เท่ากับ 0.03 แล้ว

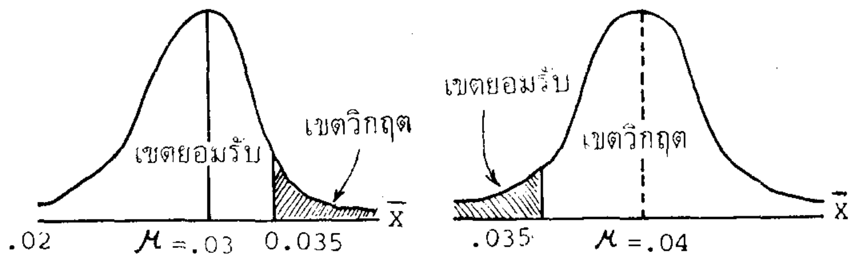
(1) สมมติฐานหลัก Ho ได้รับการยอมรับ ในเมื่อค่าของ $\bar{X}$ ไม่ได้อยู่ในเขตวิกฤต

(2) สมมติฐานหลัก Ho ได้รับการปฏิเสธ ในเมื่อค่าของ $\bar{X}$ ตกอยู่ในเขตวิกฤต

ข. ถ้าค่าเฉลี่ย (μ) นั้น จริง ๆ เป็น 0.04 แล้ว

(3) สมมติฐานหลัก Ho ได้รับการยอมรับ ในเมื่อค่าของ $\bar{X}$ ไม่อยู่ในเขตวิกฤต

(4) สมมติฐานหลัก Ho ได้รับการปฏิเสธ ในเมื่อค่าของ $\bar{X}$ อยู่ในเขตวิกฤต



ตามเกณฑ์ตัดสินใจนี้ เราจะเห็นได้ว่า กรณี (1) และ (4) นั้น เราทำการตัดสินใจถูกต้อง แต่กรณี (2) และ (4) เราจะทำการตัดสินใจผิด ผลดังกล่าวสรุปได้ดังนี้

	การตัดสินใจ	ยอมรับ Ho	ปฏิเสธ Ho
	Ho เป็นจริง	กรณี (1)	กรณี (2)
สภาวะที่แท้จริง		ตัดสินใจถูก	ตัดสินใจผิด
	Ho เป็นเท็จ	กรณี (3)	กรณี (4)
		ตัดสินใจผิด	ตัดสินใจถูก

ในกรณี (2) นี้ เราปฏิเสธชิ้นส่วนสินค้าที่ส่งมาให้ทั้ง ๆ ที่ $\mu = 0.03$ ดังนั้น จึงเกิดความคลาดเคลื่อนที่ชื่อว่า "ความคลาดเคลื่อนแบบ 1" ความน่าจะเป็นที่ความคลาดเคลื่อนแบบนี้จะเกิดขึ้นซึ่งเรียกว่า "การเสี่ยงแบบ 1, α" คำนวณได้จาก

$$\alpha = P\left\{\bar{X} > 0.035 \middle| \mu \leq 0.03\right\}$$

$$= P\left\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{0.035-0.03}{\sigma/\sqrt{n}}\right)$$

$$= P\left\{Z > \frac{0.005}{0.006/\sqrt{n}}\right\}$$

ถ้า $\bar{X}$ คำนวณจากตัวอย่างสุ่มขนาด 4 หรือ $n = 4$ แล้วการเสี่ยงแบบ 1 จะเป็น

$$\alpha = P_{I}\ z > \frac{0.005}{0.006/\sqrt{4}}\Big\} = 0.0485$$

ซึ่งหมายความว่า ถ้ากำหนดเกณฑ์ตัดสินใจดังกล่าว แล้วความคลาดเคลื่อนแบบ 1 จะมีโอกาสเกิดขึ้นสูงสุดเท่ากับ 0.0485

ส่วนกรณี (3) เรายอมรับชิ้นส่วนสินค้าที่ส่งมาให้ทั้ง ๆ ที่ μ = 0.040 ดังนั้นเราจึงทำความคลาดเคลื่อนที่ชื่อ "ความคลาดเคลื่อนแบบ 2" นั่นเอง และการเสี่ยงแบบ 2 หรือ β นี้คำนวณได้จาก

$$\beta(\mu)=P\{\bar{X}\leq 0.035|\mu > 0.03\}$$

เราจะเห็นได้ว่าการเสี่ยงแบบ 2 นี้ไม่คงที่ จะขึ้นอยู่กับค่าของ μ ถ้า μ = 0.04 เราจะได้

$$P(\mu = 0.04) = P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq \frac{0.035-0.04}{\sigma/n}\right)$$

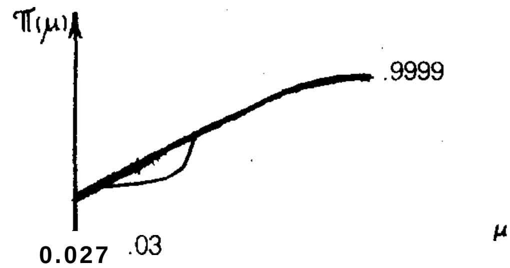
$$= P\left(Z \leq \frac{-0.005}{0.006/\sqrt{4}}\right) = 0.0485$$

สำหรับคุณลักษณะของเกณฑ์ตัดสินใจนั้นจะนิยามด้วยฟังก์ชัน π ที่เรียกว่า "อำนาจทดสอบ" อำนาจทดสอบก็คือความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลัก Ho ในเมื่อสมมติฐานหลักเป็นเท็จ นั่นคือ

$$\pi(\mu) = 1-\beta(\mu) = P\{\bar{X}>0.035|\mu > 0.03\}$$

ตารางและกราฟต่อไปนี้จะแสดงถึงค่าของฟังก์ชันอำนาจทดสอบ (Power function) สำหรับแต่ละค่าที่เป็นไปได้ของค่าเฉลี่ย μ ในเมื่อกำหนดเกณฑ์ตัดสินใจไว้ดังที่กล่าวมาแล้ว

μ	.030	033	.036	.039	.042	.045	.048
$\pi(\mu)$	.0485	.2527	.6305	.9087	.9902	.9996	.9999



จากเกณฑ์ตัดสินใจที่ว่า "จะปฏิเสธ Ho : $\mu \leq 0.03$ ถ้า $\bar{X} > 0.035$" เราจะได้

$$\alpha \leq P\left\{z > \frac{.005}{.006/\sqrt{n}}\right\} \quad \text{และ} \quad \beta(\mu=.04)=P\left\{z\leq \frac{-.005}{.006/\sqrt{n}}\right\}$$

ซึ่งทั้ง α และ β ขึ้นอยู่กับ n ถ้า n โตขึ้น แล้ว α และ β จะมีค่าลดลง นี่แสดงว่า สำหรับเกณฑ์ตัดสินใจที่กำหนดให้ ถ้าต้องการควบคุม α และ β ให้น้อย แล้วเราต้องสุ่มตัวอย่างขนาดโต

จากที่กล่าวมานี้เราเห็นได้ว่า เจ้าหน้าที่ตรวจรับของทำการทดสอบสมมติฐาน โดยอาศัยเกณฑ์ตัดสินใจว่า $\bar{X}$ จะน้อยกว่าหรือเท่ากับ 0.035 และ n = 4 วิธีการนี้บ่อยครั้งที่ไม่เหมาะสมในการทดสอบสมมติฐาน ความไม่เหมาะสมของวิธีการนี้มาจากเกณฑ์ที่ตั้งขึ้นจากเจ้าหน้าที่นั้นตายตัวมากเกินไป ถ้าเจ้าหน้าที่ตัดสินใจที่จะให้มีความน่าจะเป็นที่จะปฏิเสธชิ้นส่วนสินค้าที่ส่งมาให้ซึ่งมี $\mu \leq 0.03$ สูงสุดเพียง 0.02 แล้วเราจะหาวิธีการทดสอบที่เหมาะสมในกรณีนี้ นั่นคือเราจะทดสอบสมมติฐานต่อไปนี้

$$H_0: \mu \leq 0.03; \quad H_a: \mu > 0.03$$

โดยมี $\alpha_{max} = 0.02$ แล้วเราจะได้ว่า

$$\alpha_{max} = P\left\{\bar{x} > k / \mu = 0.03\right\}$$

ในเมื่อ k เป็นค่าคงที่ซึ่งเราจะพิจารณา และค่า k นี้ต้องสอดคล้องกับสมการต่อไปนี้

$$\begin{aligned} 0.02 &= P\left\{ \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > \frac{k-0.03}{\sigma/\sqrt{n}} \right\} \\ &= P\left\{ Z > \frac{k-0.03}{\sigma/\sqrt{n}} \right\} \\ &= 1-P\left\{ Z \leq \frac{k-0.03}{\sigma/\sqrt{n}} \right\} \end{aligned}$$

ดังนั้น $P\left\{ Z \leq \frac{k-0.03}{\sigma/\sqrt{n}} \right\} = 0.98$

จากตารางปกติมาตรฐานเราได้ $P(Z \leq 2.055) = 0.98$

เพราะฉะนั้น

$$\begin{aligned} \frac{k-0.03}{\sigma/\sqrt{n}} &= 2.055 \\ k &= 0.03+2.055\ (\sigma/\sqrt{n}) \\ &= 0.03 + 0.01233/\sqrt{n}; \quad \sigma = .006 \end{aligned}$$

นั่นคือปฏิเสธ Ho ถ้า $\bar{X} > 0.03 + 0.01233/\sqrt{n}$

สำหรับตัวอย่างขนาด n นั้นจะเป็นเท่าใด? วิธีการที่ดีในการพิจารณาขนาดตัวอย่าง (n) จะได้พูดถึงต่อไป ขอให้เรามาวิเคราะห์ผลกระทบของขนาดตัวอย่างต่อการเสี่ยงแบบ 2 หรือ β

กันก่อน

ถ้าเราให้ขนาดตัวอย่างเป็น 4, 9 หรือ 16 แล้วฟังก์ชั่นอำนาจทดสอบจะหาค่าได้ดังนี้ เมื่อ n = 4 แล้ว

$$B(\mu) = P\left\{\bar{X} \le 0.03 + \frac{0.01233}{\sqrt{4}} / \mu = 0.03\right\}$$

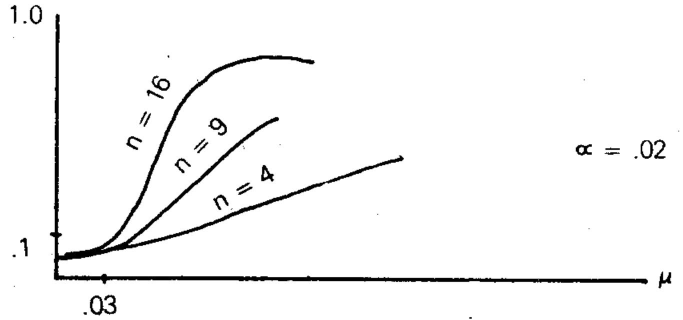
$$= P\left\{\bar{X} \le 0.03616 / \mu > 0.03\right\}$$

หรือ

$$\pi(\mu) = 1 - P\left\{\bar{X} \le 0.03616 / \mu > 0.03\right\}$$

ในทำนองเดียวกันกับ n = 4 เราสามารถคำนวณค่าของ $\pi(\mu)$ ในเมื่อ n = 9 และ n = 16 ได้ กราฟและค่าของฟังก์ชั่นอำนาจทดสอบสำหรับตัวอย่างขนาด 4, 9 และ 16 จะเป็นดังนี้

	μ	.030	.033	.036	.039	.042	.045	.048
$\pi(\mu)$	n = 4	.02	.15	.48	.82	.97	.99	1
	n = 9	.02	.29	.82	.99	1		
	n = 16	.02	.48	.97	1			



เราจะเห็นว่า การเสี่ยงแบบ 2 หรือ β จะลดลงเมื่อขนาดตัวอย่าง (n) เพิ่มขึ้น เช่นถ้าชิ้นส่วนสินค้าที่ส่งมาให้มี $\mu = 0.039$ จะมีโอกาสรับของนั้น 0.18 หรือ 18 ครั้งใน 100 ครั้ง(หรือ 01) เมื่อ n = 9 จากตารางก่อน เราพบว่าชิ้นส่วนสินค้าที่ส่งมามี $\mu = 0.039$ นั้นมีโอกาสที่จะยอมรับ 10 ครั้งใน 100 ครั้ง (β = .0913) เมื่อ n = 4 และ $\alpha_{max} = 0.0485$

นั่นคือ สำหรับ Ho : $\mu \le 0.03$; Ha : $\mu = 0.039$ และ n = 4 เมื่อ $\alpha = 0.0485$ เราได้ β = .0913 แต่เมื่อ $\alpha = 0.02$ เราได้ β = 0.18 ซึ่งหมายความว่าเมื่อ α ลดลง แล้ว β จะเพิ่มขึ้น หรือในทางกลับกัน

จากที่กล่าวมาพอจะสรุปให้เห็นความสัมพันธ์ระหว่างการเสี่ยงแบบ 1, การเสี่ยงแบบ

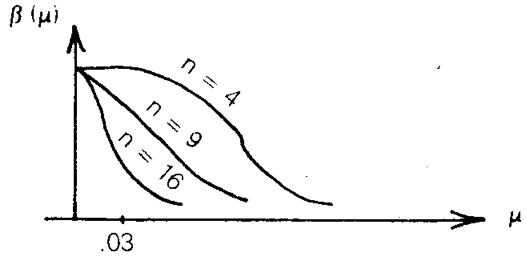
2, และขนาดตัวอย่าง (หรือ α, β, n) ได้ดังนี้

(1) ถ้าลดการเสี่ยงแบบ 1 แล้วจะมีผลทำให้การเสี่ยงแบบ 2 เพิ่มขึ้น หรือในทางกลับกัน

(2) เมื่อกำหนด α ไว้ ถ้าเพิ่มขนาดตัวอย่าง n จะทำให้ β ลดลง หรือควบคุม β ได้ เมื่อกำหนดขนาดตัวอย่างที่เหมาะสม

(3) ถ้าเพิ่มขนาดตัวอย่าง n แล้วทั้ง α และ β จะลดลง (สำหรับเกณฑ์ตัดสินใจที่กำหนดไว้)

ถ้าเราเขียนกราฟ $\beta(\mu)$ สำหรับค่าจริงของ μ ต่าง ๆ กัน โดยกำหนด $\alpha = 0.02$ และ $n = 4, 9, 16$ ให้ แล้วกราฟนั้นจะได้ชื่อว่า โค้งคุณลักษณะเชิงปฏิบัติ (Operating Characteristic Curve) หรือเรียกง่าย ๆ ว่า โค้ง OC ซึ่งแสดงได้ดังนี้



โค้ง OC สำหรับเกณฑ์ตัดสินใจที่กำหนดไว้นั้นจะพิจารณาได้จากระดับนัยสำคัญ หรือการเสี่ยงแบบ 1 และขนาดตัวอย่าง เป็นที่ประจักษ์ว่าสำหรับ α ที่กำหนด ถ้าเพิ่มตัวอย่าง n แล้วการเสี่ยงแบบ 2 หรือ β จะลดลง ในทางปฏิบัติต้องหาดุลยภาพระหว่างค่าใช้จ่าย ที่ต้องสังเกตเพิ่มเติม และประโยชน์จากการลดการเสี่ยงแบบ 2 แต่ทั่ว ๆ ไปเราไม่สามารถจะกำหนดค่าใช้จ่ายของพารามิเตอร์ที่เกี่ยวข้องกับวิธีการทดสอบอื่น ๆ และโดยที่เราไม่มีความรู้เกี่ยวกับค่าใช้จ่ายของพารามิเตอร์เหล่านี้ เกณฑ์ที่ใช้ประเมินและเปรียบเทียบวิธีการทดสอบสมมติฐาน จึงใช้โค้ง OC, $\beta(\mu)$ หรือฟังก์ชันกำนาจทดสอบ, $\pi(\mu)$

เราได้ศึกษาวิธีการตัดสินใจที่เกี่ยวข้องกับผลทดลองของตัวแปรเชิงสุ่ม $\bar{X}$, ขนาดตัวอย่าง n, และเขตยอมรับ ($\bar{X} \leq$ ค่าคงที่ซึ่งกำหนดไว้) มาแล้ว และเราจะเรียกวิธีการตัดสินใจเช่นนี้ว่า วิธีการ $\bar{X}$ เราลองพิจารณาว่า วิธีการ $\bar{X}$ นี้จะดีกว่าวิธีการ $\bar{X}_{max}$ หรือไม่ วิธีการ $\bar{X}_{max}$ ก็คือ "เลือกตัวอย่างขนาด n และสังเกตค่าของ X ที่มากสุด หรือ X_{max}" ถ้า $X_{max} \leq$ ค่าคงที่ซึ่งระบุไว้ ก็จะสรุปว่า $\mu \leq 0.03$ ถ้าเป็นอย่างอื่นก็สรุปว่า $\mu > 0.03$

ลองเปรียบเทียบสองวิธีการ $\bar{X}$ และ X_{max} นี้ โดยกำหนด α เป็น 0.02 และ $n = 4$ สำหรับทั้งสองวิธีการ และใช้โค้ง OC เป็นเกณฑ์ในการเปรียบเทียบ นั่นคือจะเปรียบเทียบทั้งสองวิธีการในเทอมของความน่าจะเป็น

เกณฑ์ตัดสินใจของวิธีการ X_{max} นี้จะยอมรับ Ho ถ้า $X_{max} \leq k$ ดังนั้น

$$P(Xmax \leq k) = 1 - \alpha = 0.98$$

โดยที่ค่าสังเกตเหล่านี้เป็นตัวแปรแบบปกติที่มีการแจกแจงเหมือนกัน และเป็นอิสระกัน เราจึงเขียนได้เป็น

$$P(X_1 \leq k)\ P(X_2 \leq k)\ P(X_k \leq k) P(X_4 \leq k) = .98$$

$$\left\{P(X \leq k)\right\}^4 = .98 \quad P(X \leq k) = (.98)^{1/4}$$

ดังนั้น $P\left\{Z \leq \frac{k - 0.03}{.006}\right\} = (0.98)^{1/4} = .995 \quad k = 0.0454$

ซึ่งหมายความว่า "ปฏิเสธ Ho เมื่อ X_{max} มากกว่า 0.0454" สำหรับการเสี่ยงแบบ 2 จะหาได้จาก

$$\beta(\mu) = P(Xmax \leq 0.0454 / \mu > 0.03)$$

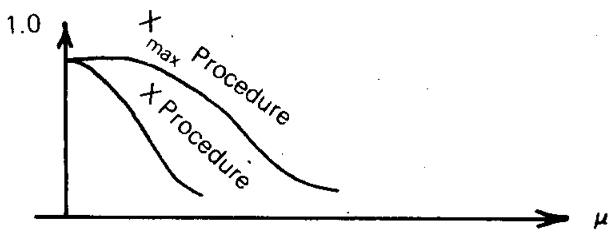
สมมติว่า μ = 0.036 แล้ว

$$\beta(0.036) = P(Xmax \leq 0.0454 / \mu = 0.036)$$
$$= \left\{P\left(Z \leq \frac{.0454 - .036}{.006}\right)\right\}^4 = 0.78$$

ในทำนองเดียวกัน เราสามารถคำนวณ β สำหรับค่าที่เป็นไปได้ของ μ ดังตารางต่อไปนี้

μ	.030	.033	.036	.039	.042	.045	.048
$\beta(\mu)$	.98	.92	.78	.53	.26	.06	.01

ดังนั้นเราจึงได้โค้ง OC สำหรับวิธีการ $\overline{X}$ และ X_{max} ดังต่อไปนี้



เราจะเห็นได้ว่า วิธีการ $\overline{X}$ ดีกว่าวิธีการ X_{max} เพราะในช่วงทั้งหมดของสมมติฐานรอง วิธีการ $\overline{X}$ ให้ β น้อยกว่า วิธีการ $\overline{X}$ นี้จะเป็นวิธีการที่ดีที่สุดไหม วิธีการที่ได้ชื่อว่า ดีที่สุด (Optimal) นั้น เขตวิกฤตจะต้องทำให้เกิดการเสี่ยงแบบ 2 หรือ β น้อยสุด สำหรับช่วงทั้งหมดของสมมติฐานรอง Ha ในเมื่อกำหนดระดับนัยสำคัญ α และขนาดตัวอย่าง n ให้ วิธีการที่ดีที่สุดนี้จะเรียกว่า "แบบ

ทดสอบที่มีอำนาจสูงสุด (Uniformly Most Powerfal Test) ในบางสถานการณ์อาจจะมีวิธีการที่ดีเกินกว่าหนึ่งวิธี สำหรับวิธีการนี้ก็เป็นวิธีการที่ดี ซึ่งสามารถพิสูจน์ได้ด้วยทฤษฎีเนย์แมน-เพียร์สัน

6.3 ขั้นตอนในการทดสอบสมมติฐานทางสถิติ

ในการทดสอบสมมติฐานทางสถิตินั้น เรามักทำตามขั้นตอนต่าง ๆ ต่อไปนี้ ซึ่งเป็นแบบหนึ่งที่ใช้กันบ่อย ๆ

ขั้นที่ 1 ตั้งสมมติฐาน (Hypothesis Formulation) สมมติฐานทดสอบที่ตั้งขึ้นนั้นจะประกอบด้วยสมมติฐานหลัก Ho และสมมติฐานรอง Ha สำหรับสมมติฐานเกี่ยวกับพารามิเตอร์ประชากรนั้นจะมีสมมติฐานหลัก Ho เป็นแบบใดแบบหนึ่งดังต่อไปนี้ ถ้า Θ_0 เป็นค่าที่ระบุไว้ของพารามิเตอร์ Θ

$$H_O: \theta = \theta_0 ; H_O: \quad \theta \leq \theta_0 \qquad H_O: \theta \geq \theta_0$$

และสมมติฐานรอง Ha จะเป็นแบบหนึ่งใน 2 แบบ ดังนี้

(ก) สมมติฐานรองทางเดียว (One-sided Alternatives) มีรูปแบบ ดังนี้

$$H_a: \theta < \theta_0 \; H_a: \theta > \theta_0$$

(ข) สมมติฐานรองสองทาง (Two-sided Alternatives) มีแบบฟอร์มเป็น

$$H_a: \theta \neq \theta_0$$

ซึ่งเป็นสมมติฐานที่กล่าวทั้ง $\Theta < \Theta_0$ และ $\Theta < \Theta_0$

ตัวอย่างของการตั้งสมมติฐาน เช่น ถ้ามีคำกล่าวว่า "ครอบครัวเกษตรกรไทยมีรายได้เฉลี่ยต่อปีเป็น 6000 บาท" สมมติฐานที่ตั้งขึ้นจะเป็น

Ho μ = 6000 บาท ; Ha : $\mu \neq$ 6000 บาท

แต่ถ้ากล่าวว่า "ครอบครัวเกษตรกรไทยมีรายได้เฉลี่ยมากกว่า 6000 บาท" สมมติฐานจะเป็น

Ho : $\mu \leq$ 6000 บาท ; Ha : $\mu >$ 6000 บาท

หรือกล่า ว่า"ครอบครัวเกษตรกรไทยมีรายได้เฉลี่ยน้อยกว่า 6000 บาท" แล้วสมมติฐานจะเป็น

Ho : $\mu \geq$ 6000 บาท ; Ha : $\mu <$ 6000 บาท

ขั้นที่ 2 กำหนดระดับนัยสำคัญ และขนาดตัวอย่าง (Specifying Significance Level and Sample Size) โดยทั่วไปเรากำหนดระดับนัยสำคัญหรือการเสี่ยงแบบ 1 เป็น α = .05 หรือ .01 ถ้าสมมติฐานหลัก Ho ได้รับการปฏิเสธ ณ ระดับนัยสำคัญ = .05 แล้วการทดสอบจะเรียกว่า

"มีนัยสำคัญ (Significant)" แต่ถ้าปฏิเสธ ณ α = .01 จะเรียกว่า "มีนัยสำคัญยิ่ง (Highly Significant)"

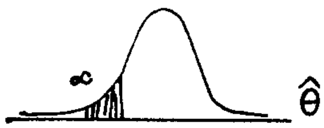
บางครั้งเรากำหนดอำนาจทดสอบ (1 - β) ไว้ด้วย ซึ่งจะทำให้สามารถพิจารณาขนาดตัวอย่างที่จะต้องเลือกสุ่มได้

ในการกำหนดขนาดตัวอย่างนั้นขึ้นอยู่กับปัจจัยต่าง ๆ นั้นจะเอื้ออำนวนให้หรือไม่ ถ้าใช้ตัวอย่างขนาดใหญ่จะลด β ลง (ในเมื่อ α ควบคุมไว้แล้ว)

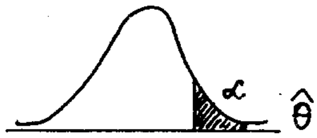
ขั้นที่ 3 เลือกตัวสถิติทดสอบ และตั้งเกณฑ์ตัดสินใจ (Selecting Test Statistic and Establishing Decision Criteria) ตัวสถิติทดสอบที่จะเลือกนั้นจะเป็นตัวสถิติที่มีการแจกแจงตัวอย่างสอดคล้องกับพารามิเตอร์ประชากรที่ระบุไว้ในคำกล่าวของสมมติฐานหลัก Ho และสมมติฐานรอง Ha ถ้าข้อกำหนดบางอย่างสอดคล้องกัน แล้วการแจกแจงตัวอย่างของตัวสถิติทดสอบจะทราบได้ ข้อกำหนดเหล่านี้มักจะเกี่ยวกับประชากรที่จะสุ่มตัวอย่าง และของตัวอย่างเอง เช่นประชากรต้องมีการแจกแจงเป็นแบบหนึ่งและทราบพารามิเตอร์อะไรบ้าง ตัวอย่างก็ต้องเป็นแบบสุ่ม เป็นต้น

จากการที่ทราบการแจกแจงตัวอย่างของตัวสถิติทดสอบ ทำให้เราสามารถกำหนดเกณฑ์ตัดสินใจซึ่งใช้เป็นแนวทางในการยอมรับหรือปฏิเสธสมมติฐานหลัก ได้ เกณฑ์ ตัดสินใจนั้นมักจะกล่าวในแบบมาตรฐานของค่าตัวสถิติทดสอบที่ใช้ ซึ่งค่าเหล่านี้หาได้จากตารางสถิติต่าง ๆ เช่น ตารางปกติมาตรฐาน Z , ตารางที่ t, ตารางไคสแควร์ χ^2, หรือ ตารางเอฟ F เป็นต้น เกณฑ์ตัดสินใจนี้จะแบ่งการแจกแจงตัวอย่างของตัวสถิติทดสอบออกเป็น 2 เขต คือ เขตปฏิเสธ และเขตยอมรับ เขตเหล่านี้จะขึ้นอยู่กับสมมติฐานรอง Ha (ซึ่งบอกทิศทางของเขตปฏิเสธ), ระดับนัยสำคัญ α (บอกขนาดของเขตปฏิเสธ), และการแจกแจงตัวอย่างของตัวสถิติทดสอบ (ใช้หาจุดวิกฤตที่จะปฏิเสธสมมติฐานหลัก Ho หรือจุดแบ่งเขตปฏิเสธกับเขตยอมรับ) เขตปฏิเสธสำหรับสมมติฐานรอง Ha แบบต่าง ๆ จะเป็นดังนี้

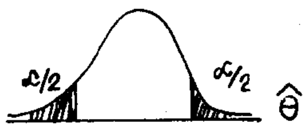
(1) $H_a: \theta < \theta_0$



(2) $H_a: \theta > \theta_0$



(3) $H_a: \theta \neq \theta_0$



ขั้นที่ 4 ทำการทดลอง และคำนวณตัวสถิติ (Performs Experiment and Doing Computations) **สามขั้นตอนที่กล่าวมาแล้วนั้นเป็นการเตรียมการ หรือเป็นขั้นวางแผนทดสอบ ขั้นที่ 4 นี้จึงเป็นงานสนาม หรืองานในห้องทดลอง เพื่อทำการทดลองสำหรับรวบรวมข้อมูลข่าวสาร จากข้อมูลข่าวสารที่ได้จึงคำนวณตัวสถิติต่าง ๆ ที่จำเป็นและคำนวณค่าของตัวสถิติทดสอบ ซึ่งกำหนดไว้ในขั้นที่** 3 ด้วย

ขั้นที่ 5 สรุปผลหรือทำการตัดสินใจ (Drawing Conclusion or Making Decisions) **ขั้นสุดท้ายนี้เป็นการสรุปผลในทางสถิติโดยใช้ค่าของตัวสถิติที่คำนวณได้ในขั้นที่ 4 มาเป็นเครื่องมือสำหรับตัดสินใจ นั่นคือถ้าค่าของตัวสถิติทดสอบตกอยู่ในเขตปฏิเสธ เราก็ปฏิเสธสมมติฐานหลักนั้น แต่ถ้าไม่อยู่ก็ยอมรับสมมติฐานหลัก** เท่านั้น

ในการสรุปผลนั้นจะต้องกล่าวให้สอดคล้องกับคำกล่าว หรือสมมติฐานนั้นด้วย มิใช่เพียงแต่บอกว่าปฏิเสธหรือยอมรับสมมติฐานหลักเท่านั้น

ตัวอย่าง นักวิจัยกล่าวว่า "บุหรี่ที่มีนิโคตินเฉลี่ยน้อยกว่า 30 มก. จะปลอดภัยจากโรคปอด" ถ้าเราต้องการจะทดสอบว่าบุหรี่กรุงทองนั้นปลอดภัยจากโรคปอดหรือไม่ เราก็ทำได้ตาม**ขั้นตอน** ดังนี้

(1) ปัญหานี้ต้องการให้เราทดสอบสมมติฐานที่ว่า

Ho : $\mu \geq 30$ มก., Ha : $\mu < 30$ มก.

ถ้าเราปฏิเสธสมมติฐานหลัก Ho เราก็แน่ใจว่าบุหรี่กรุงทองปลอดภัยจากโรคปอด

(2) ในการทดสอบสมมติฐานนี้จะใช้ระดับนัยสำคัญ α = .01 หรือ 1% และ ใช้ตัวอย่างของบุหรี่กรุงทอง 100 มวน (จะใช้น้อยหรือมากกว่า 100 มวน ก็ได้)

(3) ตัวสถิติทดสอบที่ใช้ คือ

$$z = \frac{x - \mu_0}{\sigma / \sqrt{n}}$$

ซึ่งจะมีการแจกแจงแบบปกมาตรฐาน ถ้า (ก) เลือกบุหรี่แบบสุ่ม, (ข) จำนวนนิโคตินต่อมวนมีการแจกแจงแบบปกติ, และ (ค) การแจกแจงนั้นมีค่าเฉลี่ย μ = 30 มก. และมีส่วนเบี่ยงเบนมาตรฐาน σ = 8 มก.

เราพยายามเลือกตัวอย่างเพื่อให้สอดคล้องกับ (ก) ได้โดยการเลือกบุหรี่ 1 มวน จากบุหรี่หนึ่งซอง และเลือกแต่ละซองจากบุหรี่ที่ขายอยู่ตามร้านต่าง ๆ เมื่อขนาดตัวอย่างโตมาก (ขนาด 100) เช่นนี้จะทำให้ข้อกำหนด (ข) ไม่จำเป็น (ตามทฤษฎีขีดจำกัดสู่ส่วนกลาง (Central Limit Theorem, CLT)) สำหรับข้อกำหนด (ค) ที่ว่า σ = 8 นั้นอาจจะไม่เป็นจริงก็ได้ (กรณีที่ไม่ทราบ เราใช้ตัวสถิติทดสอบอื่นได้)

เกณฑ์ตัดสินใจซึ่งกำหนดเขตวิกฤตนั้นพิจารณาได้จากตารางปกติมาตรฐาน **Z** ดังนี้ จะ

ปฏิเสธสมมติฐานหลัก Ho ถ้าค่าของตัวสถิติทดสอบ Z น้อยกว่า - 2.326 หรือค่าสัมบูรณ์ของตัวสถิติทดสอบ z มากกว่า 2.326 นั่นคือ $z < 2.326$ หรือ $|z| > 2.326$

(4) ขั้นนี้ก็ทำการทดลองโดยการสุ่มตัวอย่างของบุหรี่ 100 มวน แล้วหาจำนวนนิโคตินในบุหรี่เหล่านั้น และแล้วคำนวณค่าเฉลี่ย สมมติว่าได้เป็น $\overline{X}$ = 26 มก. เราจึงคำนวณค่าของตัวสถิติทดสอบ Z ได้เป็น

$$= \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}} = \frac{26-30}{8/\sqrt{100}} = -5.00$$

(5) เนื่องจากค่าของตัวสถิติทดสอบตกอยู่ในเขตวิกฤต จึงปฏิเสธสมมติฐานหลัก Ho ซึ่งหมายความว่าผู้สูบบุหรี่กรุงทองจะปลอดภัยจากโรคปอด

6.4 การทดสอบพารามิเตอร์ของประชากรแบบปกติ (Normal Population)

ประชากรแบบปกติชนิดตัวแปรเดียว (Univariate) มีพารามิเตอร์ค่าเฉลี่ย (μ) และส่วนเบี่ยงเบนมาตรฐาน (σ) ส่วนประชากรแบบปกติชนิดสองตัวแปร (Bivariate) นั้นนอกจากจะมีพารามิเตอร์ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานในแต่ละตัวแปร แล้วยังมีพารามิเตอร์สัมประสิทธิ์สหสัมพันธ์ (ρ) ซึ่งเป็นมาตราวัดความสัมพันธ์ระหว่างตัวแปรทั้งสองอีกด้วย

6.4.1 การทดสอบเกี่ยวกับค่าเฉลี่ย (Tests of Means)

สำหรับประชากรเดียว เราจะทดสอบค่าเฉลี่ยประชากรที่ระบุไว้ (μ_o) ส่วนสองประชากรขึ้นไปจะทดสอบเกี่ยวกับผลต่างหรือการเท่ากันของค่าเฉลี่ยต่าง ๆ

6.4.1.1 ทดสอบค่าเฉลี่ยที่ระบุไว้ ถ้าเราสนใจประชากรแบบปกติที่ไม่ทราบค่าเฉลี่ยของประชากรนั้นแล้วสมมติฐานที่จะทดสอบจะเป็นแบบใดแบบหนึ่งใน 3 แบบ ต่อไปนี้

(1) $H_O: \mu \geq \mu_o;\ H_a: \mu < \mu_o$

(2) $H_O: \mu \leq \mu_o;\ H_a: \mu > \mu_o$

(3) $H_O: \mu = \mu_o;\ H_a: \mu \neq \mu_o$

ในเมื่อ μ_o เป็นค่าที่ระบุไว้ของค่าเฉลี่ยประชากรที่สนใจ

การทดสอบสมมติฐานนี้ก็อาศัยตัวอย่างสุ่มขนาด n จากประชากรแบบปกติที่สนใจ สำหรับตัวสถิติทดสอบที่ใช้ทดสอบ สมมติฐานหลัก Ho นี้จะขึ้นอยู่กับการทราบหรือไม่ทราบ

ค่าความแปรปรวนประชากร ดังนี้

ก. ถ้าทราบค่าความแปรปรวน σ^2 ตัวสถิติทดสอบจะเป็น

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

ซึ่งมีการแจกแจงแบบปกติมาตรฐาน n (0, 1) ดังนั้นเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α จึงกำหนดได้ตามสมมติฐานรอง Ha ดังนี้

(1) สำหรับ Ha : $\mu < \mu_0$ จะปฏิเสธ Ho ถ้า $Z < -z_{\alpha}$

(2) สำหรับ Ha : $\mu > \mu_0$ จะปฏิเสธ Ho ถ้า $Z < Z_{\alpha}$

(3) สำหรับ Ha : $\mu \neq \mu_0$ จะปฏิเสธ Ho ถ้า $Z < -Z_{\alpha/2}$ หรือ $Z > Z_{\alpha/2}$

ข. ถ้าไม่ทราบความแปรปรวน σ^2 เราใช้ความแปรปรวนตัวอย่าง S^2 ไปประมาณ σ^2 ดังนั้นสถิติทดสอบจึงเป็น

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

ซึ่งมีการแจกแจงแบบที (Student t) ด้วยองศาความเป็นอิสระ n-1 สำหรับเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α จะเป็นดังนี้

(1) สำหรับ Ha : $\mu < \mu_0$ จะปฏิเสธ Ho ถ้า $T < -t_{\alpha}^{(n-1)}$

(2) สำหรับ Ha : $\mu > \mu_0$ จะปฏิเสธ Ho ถ้า $T > t_{\alpha}^{(n-1)}$

(3) สำหรับ Ha : $\mu = \mu_0$ จะปฏิเสธ Ho ถ้า $T < -t_{\alpha}^{(n-1)}$ หรือ $T > t_{\alpha/2}^{(n-1)}$

ในเมื่อ $t_{\alpha/2}^{(n-1)}$ และ $t_{\alpha}^{(n-1)}$ เป็นค่าคงที่จากตารางที ซึ่งมีองศาความเป็นอิสระ n-1 และทำให้เกิดพื้นที่ทางขวามือเท่ากับ $\alpha/2$ และ α ตามลำดับ

ตัวอย่าง เท่าที่ผ่านมาคะแนนเฉลี่ยในวิชาสถิติเป็น 50 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8 คะแนน อาจารย์ผู้หนึ่งกล่าวว่า นักศึกษาที่ผ่านคณิตศาสตร์ และสถิติเบื้องต้น มาแล้วจะมีคะแนนเฉลี่ยในวิชาสถิติมากกว่า 50 คะแนน

ในการศึกษาเพื่อทดสอบคำกล่าวนี้ ได้สุ่มนักศึกษาที่ผ่านวิชาทั้งสองมา 36 ราย ปรากฏว่าได้คะแนนเฉลี่ย 60 คะแนน และส่วนเบี่ยงเบนมาตรฐาน 10 คะแนน

ข้อมูลจากตัวอย่างนี้สนับสนุนคำกล่าวของอาจารย์หรือไม่?

$$Ho\ \mu \leq 50,\quad Ha:\ \mu > 50$$

ถ้าสมมติว่านักศึกษาที่ผ่านคณิตศาสตร์และสถิติเบื้องต้นมีส่วนเบี่ยงเบนมาตร ฐานของคะแนนวิชาสถิติเท่ากับส่วนเบี่ยงเบนมาตร ฐานของคะแนนสถิติสำหรับนักศึกษาทั้งหมด นั่นคือ $\sigma = 8$ แล้วตัวสถิติทดสอบจะเป็น

$$Z = \frac{\overline{X} - \mu_o}{\sigma/\sqrt{n}}$$

เมื่อ $\sigma = 8$, $n = 36$, $\overline{X} = 60$, และ $\mu_o = 50$ เราจึงได้ค่าของ Z เป็น

$$Z = \frac{60-50}{8/\sqrt{36}} = 7.50$$

สำหรับ $\alpha = .01$ เรามี $Z_{.01} = 2.33$ ดังนั้นจึงปฏิเสธ Ho นั่นคือคำกล่าวของอาจารย์น่าจะเป็นไปได้

แต่ถ้าสมมติว่า $\sigma \neq 8$ เราประมาณได้ด้วย $S = 10$ แล้วเราใช้ตัวสถิติ

$$T = \frac{\overline{X} - \mu_o}{S/\sqrt{n}}$$

ซึ่งจะมีค่าเป็น

$$T = \frac{60-50}{10/\sqrt{36}} = 6.00$$

สำหรับ $\alpha = .01$ เรามี $t_{.01}^{(36-1)} = 2.75$ ดังนั้นจึงปฏิเสธ Ho เช่นเดียวกัน

6.4.1.2 ทดสอบผลต่างของสองค่าเฉลี่ย ถ้าเราต้องการจะเปรียบเทียบ หรือทดสอบผลต่างค่าเฉลี่ยในสองประชากรแบบปกติว่าเป็นอย่างไร นั่นคือเราต้องการทดสอบสมมติฐานแบบหนึ่งแบบใด ดังนี้

(1) $H_o : \mu_1 - \mu_2 \geq d_o$; $H_a : \mu_1 - \mu_2 < d_o$

(2) $H_o : \mu_1 - \mu_2 \leq d_o$; $H_a : \mu_1 - \mu_2 > d_o$

(3) $H_o : \mu_1 - \mu_2 = d_o$; $H_a : \mu_1 - \mu_2 \neq d_o$

ในเมื่อ d_o เป็นค่าคงที่ ซึ่งระบุไว้ ถ้า $d_o = 0$ แล้วสมมติฐานที่ทดสอบจะเป็น

(1) $H_o : \mu_1 \geq \mu_2$; $H_a : \mu_1 < \mu_2$

(2) $H_o : \mu_1 \leq \mu_2$; $H_a : \mu_1 > \mu_2$

(3) $H_o : \mu_1 = \mu_2$; $H_a : \mu_1 \neq \mu_2$

ในการทดสอบสมมติฐานเหล่านี้จะอาศัยตัวอย่างจากประชากรที่สนใจด้วยขนาด n_1 และ n_2 สำหรับตัวอย่างนั้นอาจจะเป็นอิสระกัน หรือสัมพันธ์กันก็ได้ ดังนั้นวิธีการทดสอบผลต่างของสองค่าเฉลี่ย จึงต้องแยกกล่าวเป็น 2 กรณี ดังนี้

ก. กรณีตัวอย่างเป็นอิสระกัน เมื่อตัวอย่างที่สุ่มจากประชากรแบบปกตินั่นเป็นอิสระกัน แล้วตัวสถิติทดสอบที่ใช้ก็ยังขึ้นกับการทราบหรือไม่ทราบค่าความแปรปรวนของสองประชากรแบบปกติที่สนใจนั้นอีกด้วย ดังนี้

(1) ทราบค่าความแปรปรวนทั้งคู่ เราใช้ตัวสถิติทดสอบ

$$Z = \frac{(\bar{X}_1-\bar{X}_2)-d_0}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$$

ซึ่งมีการแจกแจง**แบบปกติมาตร**ฐาน N (0,1) ดังนั้นเกณฑ์ตัดสินใจจึงกำหนดเป็นเช่นเดียวกับการทดสอบค่าเฉลี่ยที่ระบุไว้

ตัวอย่าง ในการทดสอบสมมติฐานที่ว่า เงินเดือนเริ่มต้นของบัณฑิตทางจิตวิทยาจะมากกว่าบัณฑิตทางอื่นที่มีระยะเวลาการศึกษาเท่ากัน ได้อาศัยตัวอย่างบัณฑิตจิตวิทยา 60 ราย และบัณฑิตสาขาอื่น 100 ราย ปรากฏว่าได้ค่าสรุปเกี่ยวกับเงินเดือนเริ่มต้น ดังนี้

$$\bar{X}_1 = 2150, \quad \bar{X}_2 = 2050$$

$$S_1 = 180, \quad S_2 = 200$$

เราจะสมมติว่าความแปรปรวนของสองประชากรนั้นทราบค่า คือให้เท่ากับที่ได้จากตัวอย่างนั้น นั่นคือ $\sigma_1 = 180$ และ $\sigma_2 = 200$

$$Ho : \mu_1 \leq \mu_2 \; ; \quad Ha : \mu_1 > \mu_2$$

$$Z = \frac{(\bar{X}_1-\bar{X}_2)-do}{\sqrt{\sigma_1^2/n + \sigma_2^2/n_2}} = \frac{(2150-2050)-0}{\sqrt{180^2/60 + 200^2/100}} = 3.26$$

เมื่อ $\alpha = .05$ เราได้ $Z_{.05} = 1.645$ ซึ่งแสดงว่ายอมรับ Ho ไม่ได้ ($Z > 1.645$) นั่นคือบัณฑิตทางจิตวิทยาจะได้เงินเดือนเริ่มต้นมากกว่าบัณฑิตสาขาอื่น

(2) ไม่ทราบค่าความแปรปรวนทั้งคู่ แต่ถือว่าเท่ากัน ($\sigma_1^2 = \sigma_2^2$) ถ้าให้ S_1^2 และ S_2^2 เป็นความแปรปรวนตัวอย่างจากประชากรทั้งสอง และจะใช้เป็นตัวประมาณความแปรปรวนประชากรที่ถือว่าเท่ากันนั้น ดังนั้นตัวประมาณค่าของความแปรปรวน $\sigma_1^2 = \sigma_2^2 = \sigma^2$ จึงเป็น $S_p^2 = \hat{\sigma}^2$

$$S_p^2 = \{(n_1-)S_1^2+(n_2-1)\ S_2^2\}/(n_1+n_2-2)$$

ตัวสถิติทดสอบจึงเป็น

$$T = \{(\bar{X}_1-\bar{X}_2) \ -do\}/\sqrt{S_p^2(1/n_1+1/n_2)}$$

ซึ่งมีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ $\nu = n_1 + n_2 -2$ **แล้วเกณฑ์ตัดสินใจ ณ**

ระดับนัยสำคัญ α ตามสมมติฐานรอง Ha จะเป็นดังนี้

สำหรับ	Ha : $\mu_1-\mu_2 < d_0$	ปฏิเสธ	H_0	ถ้า	$T < -t_{\alpha}^{(\nu)}$	
สำหรับ	Ha : $\mu_1-\mu_2 > d_0$	ปฏิเสธ	H_0	ถ้า	$T > t_{\alpha}^{(\nu)}$	
สำหรับ	Ha : $\mu_1-\mu_2 \neq d_0$	ปฏิเสธ	H_0	ถ้า	$T < -t_{\alpha/2}^{(\nu)}$	หรือ $T > t_{\alpha/2}^{(\nu)}$

ในเมื่อ $\nu = n_1 + n_2 - 2$

ตัวอย่าง จากตัวอย่างเรื่องรายได้เริ่มต้นที่แล้วมา ถ้าความแปรปรวนไม่ทราบค่า แต่ถือว่าเท่ากัน แล้วเราจะประมาณความแปรปรวนได้เป็น

$$S_p^2 = \{(60-1)180^2 + (100-1)200^2\}/(60+100-2) = 37162.025$$

แล้วค่าของตัวสถิติทดสอบจะคำนวณได้เป็น

$$T = \{(2150-2050)-0\}/\sqrt{37162.025(1/60+1/100)} = 3.18$$

เมื่อ α = .05 และ ν = 60 + 100-2 = 158 เราได้ $t_{.05}^{(158)}$ = 1.645 จึงปฏิเสธ H_0 ซึ่งสรุปได้เช่นเดียวกัน

(3) ไม่ทราบความแปรปรวนทั้งคู่ และถือว่าไม่เท่ากัน ($\sigma_1^2 \neq \sigma_2^2$) ถ้าให้ S_1^2 และ S_2^2 เป็นตัวประมาณของความแปรปรวนประชากร σ_1^2 และ σ_2^2 แล้วตัวสถิติทดสอบจะเป็น

$$T = \{(\bar{X}_1-\bar{X}_2)-d_0\}/\sqrt{S_1^2/n_1 + S_2^2/n_2}$$

ซึ่งมีการแจกแจงแบบที (โดยประมาณ) สำหรับองศาความเป็นอิสระ ν จะกำหนดไว้ดังนี้

ตามวิธีของB.L.Welch

$$\nu = \left[(S_1^2/n_1+S_2^2/n_2)^2/\{(S_1^2/n_1)^2/(n_1+1)+(S_2^2/n_2)^2/(n_2+1)\}\right]-2$$

ตามวิธีของ F.E. Satterthwaite

$$\nu = (S_1^2/n_1+S_2^2/n_2)^2/\{(S_1^2/n_1)^2/(n_1-1)+(S_2^2/n_2)^2/(n_2-1)\}$$

ในการปฏิเสธ H_0 ณ ระดับนัยสำคัญ α นั้นทำได้โดยเปรียบเทียบค่าของสถิติทดสอบ T กับค่าวิกฤต $t_{\alpha}^{(\nu)}$ จากตารางที.

แต่ N.G.Cochran และ A.M. Cox ใช้เปรียบเทียบค่าของตัวสถิติทดสอบ T กับค่าเฉลี่ยของ t_1 และ t_2 หรือ

$$t = \{(S_1^2/n_1)t_1+(S_2^2/n_2)t_2\}/(S_1^2/n_1+S_2^2/n_2)$$

ในเมื่อ t_1 และ t_2 เป็นค่าจากตารางที ด้วยองศาความเป็นอิสระ n_1-1 และ n_2-1 ตามลำดับ และมีระดับนัยสำคัญ α ตามที่ระบุไว้ เราจะเห็นได้ว่าวิธีนี้ไม่ต้องคำนวณค่าขององศาความเป็นอิสระ

ตัวอย่าง นักจิตวิทยาต้องการศึกษาเวลาที่ใช้แก้ปัญหาอย่างหนึ่งของนักศึกษาชายและหญิงว่า นักศึกษาหญิงจะใช้เวลามากกว่าหรือไม่ จึงทำการทดลองโดยอาศัยนักศึกษาหญิง 5 ราย และนักศึกษาชาย 8 ราย ปรากฏว่านักศึกษาหญิงใช้เวลาเฉลี่ย 20 นาที ความแปรปรวน 8.5 ส่วนนักศึกษาชายใช้เวลาเฉลี่ย 18 นาที ความแปรปรวน 10.57

ผลของการศึกษาจะสรุปว่าอย่างไร?

$$Ho: \mu_1 \leq \mu_2 \ ; \qquad Ha: \mu_1 > \mu_2$$

ในเมื่อ μ_1 เป็นค่าเฉลี่ยจริงของเวลาที่ใช้แก้ปัญหาของนักศึกษาหญิง

เมื่อ $n_1 = 5, n_2 = 8, \bar{x}_1 = 20, \bar{x}_2\ 18, s^2_1 = 8.5$ และ $s^2_2 = 10.57$ เราจะได้ค่าของตัวสถิติทดสอบ T เป็น

$$T = \{(20-18)-0\}/\sqrt{8.5/5 + 10.57/8} = 1.1506$$

สำหรับองศาความเป็นอิสระ ν เราหาได้ดังนี้

ตามวิธี Welch

$$\nu = \left[(8.5/5+10.57/8)^2/\{(8.5/5)^2/(5+1)+(10.57/8)^2/(8+1)\}\right]-2$$
$$= 11.56 \simeq 12$$

ตามวิธี Satterthwaite

$$\nu = (8.5/5+10.57/8)^2/\{(8.5/5)^2/(5-1)+(10.57/8)^2/(8-1)\}$$
$$= 3.109 \simeq 3$$

สำหรับค่าเฉลี่ยของ t ตามวิธี Cochran - Cox เราได้

$$t = \{(8.5/5)(2.132)+(10.57/8)(1.895)\}/(8.5/5+10.57/8) = 2.028$$

ในเมื่อ $t^{(5-1)}_{.05} = 2.132 = t_1$, $t^{(8-1)}_{.05} = 1.895 = t_2$

ดังนั้นเราจะสรุปผลได้ดังนี้ ตามวิธี Welch เราปฏิเสธ H_0 ไม่ได้เพราะ $T = 1.1506 < t^{(12)}_{.05} = 1.782$ ตามวิธี Satterthewaite เราก็ปฏิเสธ H_0 ไม่ได้เพราะ $T < t^{(3)}_{.05} = 2.353$ สำหรับวิธีการของ Cochran - Cox ก็สรุปได้เช่นเดียวกันเพราะ $T < t$

เพราะฉะนั้นจึงสรุปได้ว่า เวลาที่ใช้แก้ปัญหาของนักศึกษาทั้งสองเพศจะไม่แตกต่างกัน

ข. กรณีตัวอย่างไม่เป็นอิสระกัน หรือมีข้อมูลจับคู่กัน บางครั้งค่าสังเกตในตัวอย่างทั้งสองมีความสัมพันธ์กัน เช่น สังเกตจากก่อนและหลังการทดลอง หรือค่าสังเกตได้จากการจับคู่หน่วยทดลอง ถ้าให้ $X_{11}, X_{21}, \ldots X_{n1}$, และ $X_{12}, X_{22}, \ldots X_{n2}$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มี

ความแปรปรวนเท่ากัน (σ^2) แล้วตัวสถิติทดสอบจะกำหนดไว้ดังนี้

(1) ถ้าไม่ทราบค่าความแปรปรวน σ^2 เราใช้ตัวสถิติทดสอบ

$$T = \{(\bar{X}_1 - \bar{X}_2) - do\} / \sqrt{S_1^2 + S_2^2 - 2S_{12}} \quad \text{หรือ} \quad T = (\bar{D} - do)/S_D/\sqrt{n}$$

ในเมื่อ S_{12} เป็นความแปรปรวนร่วมของค่าสังเกตทั้งสอง, $\bar{D} = \sum D_i/n$ โดยที่

$D_i = X_{i1} - X_{i2} (i=1,2,...n)$, $S_D^2 = (\sum(D-\bar{D})^2)/(n-1)$

$= (1/(n-1))(\Sigma D^2 - (\Sigma D)^2/n)$

ตัวสถิติ T นี้จะมีการแจกแจงแบบทีด้วยองศาความเป็นอิสระ n-1 ดังนั้น เกณฑ์ตัดสินใจจึงกำหนดได้เช่นเดียวกันกับตัวสถิติที่มีการแจกแจงแบบที

(2) ถ้าทราบค่าความแปรปรวน σ^2 เราใช้ตัวสถิติทดสอบ T เช่นเดียวกัน แต่การตัดสินใจที่จะปฏิเสธ H_0 จะต้องเปรียบเทียบค่า T ที่คำนวณได้กับค่า t จากตารางที่มีองศาความเป็นอิสระ ∞ หรือค่า Z จากตารางปกติ

ตัวอย่าง ครูต้องการทราบว่าข้อทดสอบคู่ขนาน 2 ฉบับ นี้มีความยากง่ายแตกต่างกันหรือไม่ จึงทำการทดลองกับนักศึกษา 10 คน ได้ผลทดลองซึ่งเป็นคะแนนดังนี้

นักศึกษา	1	2	3	4	5	6	7	8	9	10
ฉบับ ก	27	85	62	70	43	95	68	75	97	36
ฉบับ ข	35	90	60	82	47	90	72	86	94	41
ผลต่าง	-8	-5	2	-12	-4	5	-4	-11	3	-5

$$Ho: \mu_1 = \mu_2; \quad Ha: \mu_1 \neq \mu_2$$

สำหรับ $n = 10, \bar{D} = -3.9, S^2_D = 32.99$ เราจึงได้ค่าของตัวสถิติ T ดังนี้

$$T = (-3.9-0)/32.99/\sqrt{10} = -2.13$$

ในเมื่อ $\Sigma D = -8-5+2-12-4+5-4-11+3-5 = -39$

$\Sigma D^2 = (-8)^2 - (-5)^2 - \dots\dots\dots - (-5)^2 = 449$

เมื่อ $\alpha = .05$ และ $\nu = 10-1 = 9$ เราได้ $t^{(9)}_{.05/2} = 2.262$ จึงปฏิเสธ H_o ไม่ได้ แสดงว่าข้อทดสอบคู่ขนาน 2 ฉบับ นี้มีความยากง่ายไม่แตกต่างกัน

บางครั้งการจับคู่หน่วยทดลองจะยุ่งยาก จึงต้องอาศัยเกณฑ์หรือคุณลักษณะอย่างหนึ่งเป็นหลักในการกำหนดกลุ่มตัวอย่างทั้งสองโดยที่กลุ่มทั้งสองมีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานในเกณฑ์ไม่แตกต่างกัน ดังนั้นตัวสถิติทดสอบจะเป็น

$H_a: \mu_j \neq \mu; \exists_j \qquad (j = 1,2,\ldots k)$

ในการทดสอบสมมติฐานนี้ก็อาศัยตัวอย่างจากประชากรต่าง ๆ เช่นเดียวกัน และโดยที่ตัวอย่างอาจจะเป็นอิสระ หรือมีความสัมพันธ์กันก็ได้ เราจึงต้องพิจารณาเป็น 2 กรณี ดังนี้

ก. กรณีตัวอย่างเป็นอิสระกัน ตัวสถิติทดสอบสำหรับกรณีนี้จะขึ้นอยู่กับความแปรปรวนของประชากรที่สุ่มตัวอย่างมาว่าเท่ากันหมดหรือไม่

(1) ถ้าความแปรปรวนประชากรเท่ากันหมด ($\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$) ตัวสถิติที่ใช้ประมาณค่าความแปรปรวนประชากรที่ถือว่าเท่ากันหมด σ^2 นั้นจะเป็น S_a^2 และ S_w^2 ซึ่งเป็นความแปรปรวนระหว่างกลุ่ม และความแปรปรวนภายในกลุ่ม ตัวสถิติทั้งสองจะมีคุณสมบัติดังนี้

S_n^2 จะเป็นตัวประมาณค่าที่ไม่เอียงเฉของ σ^2 นั่นคือ $E(S_w^2) = \sigma^2$ แต่ S_a^2 โดยปกติจะเป็นตัวประมาณค่าที่เอียงเฉทางบวก นั่นคือ $E(S_a^2) = \sigma^2 + \Sigma n_j(\mu_j - \mu)^2/(k-1)$ และจะไม่เอียงเฉก็ต่อเมื่อสมมติฐานหลัก $H_o: \mu_j = \mu, \forall_j$ เป็นจริง

ดังนั้นตัวสถิติทดสอบจึงอาศัยตัวสถิติ S_w^2 และ S_a^2 นั้น นั่นคือจะเป็น

ของ μ

$$F = S_\alpha^2/S_W^2$$

ในเมื่อ
$$S_a^2 = \sum_{i,j}^{n_j, k} (\bar{X}_{.j} - \bar{X})^2/(k-1) = \sum n_j(\bar{X}_{.j} - \bar{X})^2/(k-1)$$

$$= (\sum X_{.j}^2/n_j - C)/(k-1) \quad ; \quad C = (\sum X_{ij})^2/n$$

$$S_W^2 = \sum (X_{ij} - \bar{X})^2/(n-k) = \sum (n_j - 1)S_j^2/(n-k)$$

$$= (\sum X_{ij}^2 - \sum X_{.j}^2/n_j))(n-k)$$

และ n_j เป็นขนาดตัวอย่างของกลุ่ม j, n เป็นขนาดตัวอย่างรวม $n = n_1 + n_2 + \ldots + n_k$, k เป็นจำนวนกลุ่มตัวอย่าง, $X_{.j}$ เป็นผลรวมของค่าสังเกตในตัวอย่าง j, $\bar{X}_{.j}$ เป็นค่าเฉลี่ยของตัวอย่าง j, S_j^2 เป็นความแปรปรวนของตัวอย่าง j

ถ้าสมมติฐานหลัก H_o เป็นจริง แล้วตัวสถิติทดสอบ F จะมีการแจกแจงแบบเอฟ (Snedecor F) ด้วยองศาความเป็นอิสระ $k-1, n-k$ เกณฑ์ตัดสินใจสำหรับระดับนัยสำคัญ α จึงกำหนดไว้ว่า "จะปฏิเสธ H_o ถ้า $F > F_\alpha^{(k-1, n-k)}$ ในเมื่อ $F_\alpha^{(k-1, n-k)}$ เป็นค่าคงที่จากตารางเอฟ"

ในการทดสอบสมมติฐานหลัก $H_o: \mu_j = \mu, \forall_j$ นี้เราใช้ตัวสถิติทดสอบอีกตัวหนึ่งซึ่งได้ชื่อว่า ตัวสถิติ Q (Studentized Range) และกำหนดไว้ดังนี้

$$T = \{(\bar{X}_1-\bar{X}_2)-do\}/\sqrt{(S_1^2/n_{.1}+S_2^2/n_2)(1-r^2_{xy})}$$

ในเมื่อ r_{xy} เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรที่สนใจ (X) กับเกณฑ์ (Y)

ตัวสถิติทดสอบ T นี้จะมีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ

$$(n_2-1)+(n_2-1)-1 = n_1+n_2-3$$

ตัวอย่าง ในการศึกษาเพื่อเปรียบเทียบความถนัดเชิงกลของนักเรียนกลุ่มวิชาการกับนักเรียนกลุ่มเทคนิค การกำหนดกลุ่มตัวอย่างทั้งสองได้ใช้สติปัญญาเป็นเกณฑ์ จากการศึกษาตัวอย่างได้ข้อมูลมาดังนี้

		กลุ่มวิชาการ	กลุ่มเทคนิค
ขนาดตัวอย่าง		125	137
คะแนนแบบทดสอบ	$\bar{Y}$	102	102.80
ระดับสติปัญญา	S	33.65	31.62
คะแนนแบบทดสอบ	$\bar{X}$	51.42	54.38
ความถนัดเชิงกล	S	6.24	7.14

สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนจากแบบทดสอบ

ระดับสติปัญญากับความถนัดเชิงกล r_{xy} = .30

กลุ่มเทคนิคจะมีความถนัดเชิงกลมากกว่ากลุ่มวิชาการหรือไม่ ?

$$Ho: \mu_1 = \mu_2 \ ; \ Ha: \ \mu_1 > \mu_2$$

ในเมื่อ μ_1 แทนคะแนนเฉลี่ยจริงของความถนัดเชิงกลในกลุ่มเทคนิค

$$T = \{(54.38-51.42)-0\}/\sqrt{(6.24^2/125+7.14^2/137)(1-30^2)} = 3.75$$

เมื่อ α = .05 เราได้ $t_{.05}^{(\nu)}$ = 1.645, ν = 125 + 137 – 3 = 259 จึงสรุปได้ว่านักเรียนกลุ่มเทคนิคมีความถนัดเชิงกลมากกว่ากลุ่มวิชาการ

บางครั้งเกณฑ์ในการกำหนดกลุ่มทั้งสองจะต้องอาศัยเกณฑ์ หลายเกณฑ์ เช่นระดับสติปัญญา, รายได้, และอื่น ๆ เราต้องใช้สัมประสิทธิ์สหสัมพันธ์เชิงซ้อน $r_{x.y_1y_2...}$ แทน

6.4.1.3 ทดสอบการเท่ากันของค่าเฉลี่ยในหลายประชากร บางครั้งเรามีประชากรที่สนใจมากกว่าสอง และต้องการเปรียบเทียบค่าเฉลี่ยประชากรทั้งหลายว่าเท่ากันหมดหรือไม่ นั่นคือจะทดสอบสมมติฐานที่ว่า

$$Ho: \mu_1 = \mu_2 = \ldots = \mu_k = \mu$$
$$Ho: \mu_j = \mu; \ \forall_j$$

$$Q = (\bar{X}_{max} - \bar{X}_{min}) / \sqrt{S_W^2 / n_o} = T_{max} - T_{min} / \sqrt{n_o S_W^2}$$

ในเมื่อ $\bar{X}_{max}$, $\bar{X}_{min}$ เป็นค่าเฉลี่ยตัวอย่างที่มีค่ามากสุด และน้อยสุด, T_{max}, T_{min} เป็นผลรวมตัวอย่างที่มีค่ามากสุดและน้อยสุด, และ n_o เป็นขนาดตัวอย่างของกลุ่มใด ๆ เมื่อขนาดตัวอย่างเท่ากัน $n_o = n_1 = n_2 \ldots = n_k$ แต่ถ้าไม่เท่ากัน $n_o = k/(1/n_1 + 1/n_2 + \ldots + 1/n_k)$

เกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญกำหนดไว้ว่า "จะปฏิเสธ H_o ถ้า $Q > Q^{k,\nu}$ ในเมื่อ $Q^{k,\nu}$ เป็นค่าจากตารางที่สอดคล้องกับ k (จำนวนตัวอย่าง) และ ν (องศาความเป็นอิสระของ S_W^2 นั่นคือ $\nu = n - k$

ตัวอย่าง ในการศึกษาเพื่อเปรียบเทียบค่าเฉลี่ยของตัวแปรทางจิตวิทยาสำหรับกลุ่มตัวอย่าง 3 กลุ่ม ซึ่งมีขนาดตัวอย่างเท่ากันหมดคือ 8 ได้ผลสรุปของข้อมูลจากการศึกษา ดังนี้

ผลรวม $X_{.1} = 576$, $X_{.2} = 578$, และ $\bar{X}_{.3} = 630$

$S_a^2 = 117.1650$, $S_w^2 = 52.9048$

$$Ho: \mu_j = \mu, \forall_j \qquad j = 1,2,3$$

$$Ha: \mu_j \neq \mu, \exists_j$$

จากผลสรุปข้อมูล เราได้ค่าของตัวสถิติทดสอบ F และ Q ดังนี้

$$F = 117.1650/52.9048 = 2.215$$

$$T = 630 - 576 / \sqrt{8(52.9048)} = 2.625$$

เนื่องจาก $F < F_{.05}^{(3-1, 24-3)} = 3.4$ จึงปฏิเสธ H_o ไม่ได้ ในทำนองเดียวกัน $Q < Q_{.05}^{(3,21)} = 3.56$ จึงสรุปได้เช่นเดียวกัน นั่นคือกลุ่มทั้งสามมีค่าเฉลี่ยของตัวแปรทางเศรษฐศาสตร์เท่ากันหมด

เมื่อสมมติฐานหลัก H_o ได้รับการปฏิเสธ หรือค่าเฉลี่ยประชากรไม่เท่ากันหมดนั้น ถ้าเราต้องการทราบว่าค่าเฉลี่ยประชากรคู่ใดบ้างที่แตกต่างกัน เราจะอาศัยช่วงเชื่อมั่น $100(1-\alpha)\%$ สำหรับผลต่างของค่าเฉลี่ยคู่นั้น ๆ ดังนี้

$$\mu_i - \mu_j = (\bar{X}_{.i} - \bar{X}_{.j}) \pm S\sqrt{S_W^2 (1/n_1 + 1/n_j)}$$

ในเมื่อ $S = \sqrt{(k-1) F_{\alpha}^{(k-1, n-k)}}$ และ $i, j = 1, 2 \ldots k; \quad (i < j)$

ในการทดสอบการเท่ากันของค่าเฉลี่ย หรือทดสอบสมมติฐานหลัก Ho : $\mu_i = \mu$ นั้น เรามีข้อตกลงว่าประชากรจะต้องมีความแปรปรวนเท่ากันหมด นั่นคือ $\sigma_1^2 = \sigma_2^2 = \ldots \sigma_k^2 = \sigma^2$ ดังนั้นตัวประมาณค่าของความแปรปรวน σ^2 จากตัวอย่างสุ่มที่เป็นอิสระกัน k ตัวอย่างที่นำมา

ทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยนี้จะเป็น

$$S_p^2 = \{(n_1-1)S_1^2 + (n_2-1)S_2^2 \ldots + (n_k-1)S_k^2\}/(n_1+n_2+\ldots n_k-k)$$
$$= \sum(n_j-1)S_j^2/(n-k) \;;\; n = n_1+n_2+\ldots n_k$$

เราจะเห็นได้ว่า S^2_p ก็คือ $S^2_{w_2}$ นันเอง ดังนั้นตัวประมาณค่า S_p^2 จึงเป็นตัวประมาณค่าที่ไม่เอียงเฉของ σ^2 นั่นคือ $E(S^2p) = \sigma^2$

จากตัวอย่างเราจะได้ค่าประมาณของ σ^2 เป็น $S_p^2 = 52.9048$

ตัวประมาณค่า S_p^2 นี้มีคุณสมบัติที่น่าสนใจอีกดังนี้

$$X^2 = (n-k)S_p^2/\sigma^2$$

จะมีการแจกแจงแบบไคสแควร์ ด้วยองศาความเป็นอิสระ n-k ทั้งนี้ต้องสอดคล้องกับข้อตกลงที่ว่า ตัวอย่างสุ่มต้องมาจากประชากรแบบปกติและเป็นอิสระกัน ดังนั้นเราสามารถประมาณช่วงเชื่อมั่น $100(1-\alpha)\%$ สำหรับ σ^2 ได้เป็น

$$(n-k)S^2_p/X^{2(n-k)}_{\alpha/2} \leq \sigma^2 \leq (n-k)S_p^2/X^{2(n-k)}_{1-\alpha/2}$$

จากตัวอย่างเราได้ช่วงเชื่อมั่น 95% สำหรับ σ^2 ดังนี้

$$(24-3)(52.9048)/35.48 \leq \sigma^2 \leq (24-3)(52.9048)/10.28$$
$$31.42 \leq \sigma^2 \leq 108.08$$

ถ้าเราสร้างช่วงเชื่อมั่นสำหรับความแปรปรวนแต่ละประชากรจากตัวอย่าง แล้วเราจะเห็นว่าช่วงเชื่อมั่นเหล่านั้นจะกว้างกว่าช่วงเชื่อมั่นข้างบนนี้ ซึ่งแสดงให้เห็นว่า S^2_p เป็นตัวประมาณค่าที่มีประสิทธิภาพของ σ^2 มากกว่าตัวประมาณค่าจากแต่ละตัวอย่าง (S^2_j) แต่อย่าลืมว่า S^2_p นั้นใช้ประมาณ σ^2 ได้ก็ต่อเมื่อประชากรมีความแปรปรวนเท่ากันหมด

(2) ถ้าความแปรปรวนไม่เท่ากันหมด $(\sigma^2_i \pm \sigma^2_j)$ เราจะใช้ตัวสถิติทดสอบชนิดถ่วงน้ำหนัก ดังนี้

$$F_o = (1/k-1)\sum \hat{W}_j^2 (\bar{X}_{.j}-\bar{X}_o)^2$$

ในเมื่อ $\hat{W}_j = n_j/S_j^2$; $\bar{X}_o = \sum \hat{W}_j\bar{X}_j / W$; $W = \sum \hat{W}_j$

ถ้าสมมติฐานหลักเป็นจริง แล้วตัวสถิติ F_o จะมีการแจกแจงแบบเอฟ (โดยประมาณ) ด้วยองศาความเป็นอิสระ $k-1, \nu_o$ ในเมื่อ $\nu_o = (k^2-1)/3A$ โดยที่ A มีค่าเป็น

$$A = \sum_j (1-\hat{W}_j/W)^2 / (n_j-1)$$

ตัวอย่าง ในการศึกษาเพื่อเปรียบเทียบกลุ่ม 4 กลุ่ม เกี่ยวกับตัวแปรทางจิตวิทยาว่ามีค่าเฉลี่ยระหว่างกลุ่มแตกต่างกันหรือไม่ ได้ข้อมูลสรุปจากตัวอย่างสุ่ม ดังนี้

กลุ่ม	n_i	$\bar{X}_{\cdot j}$	S_i^2	$\hat{w}_i$	$\hat{w}_i \bar{X}_{\cdot i}$	$\bar{X}_i - \bar{X}_0$	$\frac{1}{n_i - 1}(1-\hat{w}_i/w)^2$
1	6	12.17	13.37	.4488	5.4619	−4.56	.0357
2	7	17.43	43.29	.1617	2.8184	.70	.1045
3	6	28.00	71.60	.0838	2.3464	11.27	.1592
4	5	28.80	60.70	.0824	2.3731	12.07	.1998
				.7767	12.9998		.4992

$$\begin{array}{l} Ho : \mu_j = \mu;\ \forall_j \\ Ha : \mu_j \neq \mu;\ \exists_j \end{array} \quad j = 1,2,3,4$$

เราได้ $\bar{X}_o = 12.9998/(.7767) = 16.73$ ดังนั้นตัวสถิติ F_o จะมีค่าเป็น

$$F_o = \frac{1}{4-1}(.4488(-4.56)^2 + .1617(.70)^2 + .0838(11.27)^2 + .0824(12.07)^2)$$

$$= 10.68$$

สำหรับ $A = .4992,\ \nu_0 = (4^2-1)/3(.4992) \approx 10$ เราจะได้ $F_{.05}^{(3,10)} = 3.71$ จึงสรุปได้ว่ากลุ่มทั้งหมดให้ค่าเฉลี่ยไม่เท่ากันหมด ($F_o > 3.71$)

เมื่อสมมติฐานหลัก H_o ได้รับการปฏิเสธ เราสามารถสร้างช่วงมั่น $100(1-\alpha)\%$ สำหรับ $\mu_i - \mu_j,\ i<j$ ได้เป็น

$$\mu_i - \mu_j = (\bar{X}_i - \bar{X}_{oj}) \pm S_o\sqrt{S_i^2/n_i + S_j^2/n_j}$$

ในเมื่อ $S_o^2 = (k-1)F_\alpha^{(k-1,\ \nu_o)}$ และ $i,j = 1,2,\ldots,k\ ;\ i<j$

ข. **กรณีตัวอย่างไม่เป็นอิสระกัน** ในการทดสอบสมมติฐานเกี่ยวกับการเท่ากันของหลายค่าเฉลี่ยประชากร เมื่อตัวอย่างมีความสัมพันธ์กัน หรือไม่เป็นอิสระกันนั้นก็เป็นกรณีทั่วไปของการเปรียบเทียบค่าเฉลี่ยของสองประชากรโดยใช้ตัวอย่างที่มีความสัมพันธ์กัน การที่จะเปรียบเทียบหลายค่าเฉลี่ยประชากรนี้จะต้องกำหนดว่าความแปรปรวนในแต่ละประชากรนั้นเท่ากันหมด สัมประสิทธิ์สหพันธ์ระหว่างสองประชากรใด ๆ จะต้องเท่ากัน และประชากรต่าง ๆ นั้นต้องมีการแจกแจงแบบปกติ

ตัวสถิติทดสอบสำหรับสมมติฐานหลักนั้นจะอาศัยวิธีการแยกความผันแปร (Variation) ของค่าสังเกตจากตัวอย่างทั้งหมดออกตามแหล่งที่ก่อให้เกิดความผันแปร วิธีการนี้ได้ชื่อว่า การวิเคราะห์ความแปรปรวน Analysis of Variance) ซึ่งจะได้กล่าวในบทต่อไปอย่างละเอียดอีก ดังนั้นตัวสถิติทดสอบจึงกำหนดไว้ดังนี้

$$F = SSA/(k-1)/SSE/(n-1)(k-1) = MSA/MSE$$

ในเมื่อ SSA เป็นความผันแปรเนื่องจากกลุ่มตัวอย่าง, SSE เป็นความผันแปรที่อธิบายไม่ได้, MSA เป็นค่าเฉลี่ยของ SSA ต่อองศาความเป็นอิสระ, และ MSE ก็เป็นค่าเฉลี่ยของ SSE ต่อองศาความเป็นอิสระเช่นกัน, k เป็นกลุ่มตัวอย่าง, n เป็นจำนวนผู้รับการทดลองหรือบลอค (Block)

ตัวสถิติทดสอบ F นี้ภายใต้สมมติฐานหลักจะมีการแจกแจงแบบเอฟ ด้วยองศาความเป็นอิสระ $k-1$ และ $(n-1)(k-1)$ การคำนวณตัวสถิติ F นี้จะอาศัยการแยกความผันแปรทั้งหมด (SST) ออกเป็นความผันแปรระหว่างกลุ่ม (SSA), ความผันแปรระหว่างบลอคหรือผู้รับการทดลอง (SSS), และความผันแปรที่เหลือซึ่งอธิบายไม่ได้ (SSE) ความผันแปรดังกล่าวหาได้จากสูตรต่อไปนี้

(1) $C = (\sum X_{ij})^2/nk$

(2) $SST = \sum X_{ij}^2 - C$

(3) $SSA = \sum X_{.j}^2/n - C$

(4) $SSS = \sum X_{i.}^2/k - C$

(5) $SSE = SST - SSA - SSS$

ในเมื่อ $X_{.j}$ เป็นผลรวมของค่าสังเกตในกลุ่ม j ($j = 1,2,\ldots,k$), $x_{i.}$ เป็นผลรวมของค่าสังเกตในบลอค i ($i = 1,2,\ldots,$), และ X_{ij} เป็นค่าสังเกตของกลุ่ม j ในบลอค i

ตัวอย่าง ในการทดลองเพื่อเปรียบเทียบค่าเฉลี่ยของตัวแปรทางจิตวิทยาในกลุ่มหรือกรรมวิธีทดลอง 5 แบบ โดยการแบ่งหน่วยทดลองหรือตัวอย่างออกเป็น 8 พวก (บลอค) ซึ่งแต่ละพวกจะมีหน่วยทดลองคล้ายคลึงกัน เมื่อการทดลองสิ้นสุดลงได้ค่าสังเกตุดังนี้

กรรมวิธี	1	2	3	4	5	ผลรวม (X_i)
บลอค 1	0	3	2	3	4	12
2	6	8	9	9	9	41
3	9	10	10	12	11	51
4	5	6	6	6	7	30
5	6	8	5	11	11	41
6	3	3	6	7	10	29
7	4	5	6	7	6	28
8	3	6	9	8	10	36

ผลรวม $(X_{.j})$ 36 48 53 63 68 268

ค่าเฉลี่ย $(\bar{X}_{.j})$ 4.50 6.00 6.625 7.875 8.50 6.70

สำหรับตัวสถิติทดสอบ F คำนวณได้ดังนี้

$$H_o : \mu_1 = \mu_2 = = \mu_5 = \mu$$

$$H_a : \mu_j \neq \mu; \quad \exists_j \quad (j=1,2,\ldots,5)$$

(1) $C = (\sum X_{ij})^2/nk = (268)^2/8(5) = 1795.60$

(2) $SST= \sum X_{ij}^2 - C = (0^2+6^2+9^2+\ldots+10^2+6^2+10^2) - C$

$= 2112-1795.60 = 316.40$

(3) $SSA= \sum X_{.j}^2/n - C = (36^2+48^2+ \ldots+68^2)/8-C$

$= 1875.25 - 1795.60 = 79.65$

(4) $SSS= \sum X_{i.}^2/k-C = (12^2+41^2+\ldots+36^2)/5-C$

$= 1985.60-1795.60 = 190.00$

(5) $SSE = SST-SSA-SSS = 316.40 - 79.69 - 190.00$

$= 46.75$

ดังนั้น $F = 79.65/(5-1)/46.75/(8-1)(5-1) = 11.92$

เมื่อ α = .05 เราได้ $F_{.05}^{(4,28)} = 2.71$ โดยที่ $k-1 = 4$ และ $(n-1)(k-1) = 28$ จึงปฏิเสธ H_o ($F > 2.71$) นั่นคือกรรมวิธีทดลอง 5 แบบ ให้ค่าเฉลี่ยของตัวแปรทางเศรษฐศาสตร์ไม่เท่ากันหมด

เมื่อสมมติฐานหลัก H_o ได้รับการปฏิเสธ และเราต้องการทราบว่ากลุ่มหรือกรรมวิธีไหนที่ให้ค่าเฉลี่ยแตกต่างกันบ้าง เราทำได้โดยอาศัยช่วงเชื่อมั่น $100(1-\alpha)\%$ ของ $\mu_i - \mu_j$ $(i<j)$ ดังนี้

$$\mu_i - \mu_j = (\bar{X}_i - \bar{X}_j) \pm \sqrt{S^2}\sqrt{2MSE/n}$$

ในเมื่อ $S^2 = (k-1)F_\alpha^{(\nu_1, \nu_2)}$, $\nu_1 = k-1$, $\nu_2 = (n-1)(k-1)$

$i,j = 1,2,\ldots k \qquad i < j$

ถ้าต้องการหาสัมประสิทธิ์สหสัมพันธ์ระหว่าง 2 กลุ่มใด ๆ ที่ถือว่าเท่ากันนั้น เราจะได้ค่าประมาณ r จากสมการต่อไปนี้

$$MSE = MSW(1-r)$$

ในเมื่อ MSW = SSW/k(n−1) โดยที่ SSW เป็นความผันแปรที่ไม่เกี่ยวกับกลุ่มหรือกรรมวิธี ซึ่งหาได้จาก SSW = SST − SSA

จากตัวอย่างเราได้ SSW = 316.40 − 79.65 = 236.75

MSW = 236.75/5(8 − 1) = 6.76

MSE = 46.75/(8 − 1)(5 − 1) = 1.67

ดังนั้น 1.67 = 6.76(1 − λ)

λ = 1 − 1.67/6.76 = 0.76

6.4.2 **การทดสอบเกี่ยวกับความแปรปรวน** (Tests of Variances)

กรณีประชากรเดียวเราจะทดสอบสมมติฐานเกี่ยวกับค่าของแปรปรวนที่ระบุไว้ ส่วนกรณีสองประชากรขึ้นไปเราจะทำการทดสอบความแตกต่างของความแปรปรวนในเทอมของอัตราส่วนเพื่อสรุปผลเกี่ยวกับการเท่ากัน

6.4.2.1 ทดสอบค่าของความแปรปรวนที่ระบุไว้ (σ_0^2) ถ้าประชากรมีการแจกแจงแบบปกติที่มีความแปรปรวน σ^2 ซึ่งเราไม่ทราบค่า แต่คาดว่าควรจะเป็น σ_0^2 แล้วเราต้องการทดสอบการคาดหวังนั้น นั่นคือทดสอบสมมติฐานแบบหนึ่งแบบใดดังต่อไปนี้

(1) Ho: $\sigma^2 \geq \sigma_0^2$; Ha : $\sigma^2 < \sigma_0^2$

(2) Ho: $\sigma^2 \leq \sigma_0^2$; Ha : $\sigma^2 > \sigma_0^2$

(3) Ho: $\sigma^2 = \sigma_0^2$; Ha : $\sigma^2 \neq \sigma_0^2$

ในการทดสอบสมมติฐานหลัก H_0 ก็จะอาศัยตัวอย่างสุ่มขนาด n จากประชากรแบบปกติที่สนใจ ตัวสถิติทดสอบก็จะอาศัยความแปรปรวนตัวอย่าง S^2 นั่นคือตัวสถิติทดสอบจะเป็น

$$X^2 = (n-1)S^2/\sigma_0^2$$

ซึ่งมีการแจกแจงแบบไคสแควร์ ด้วยองศาความเป็นอิสระ ν = n − 1 ถ้าสมมติฐานหลักเป็นจริง

เกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α จะกำหนดไว้ตามสมมติฐานของ Ha ต่าง ๆ ดังนี้

- (1) สำหรับ H_a : $\sigma^2 < \sigma_0^2$ จะปฏิเสธ Ho ถ้า $X^2 < X^{2(n-1)}_{1-\alpha}$
- (2) สำหรับ H_a : $\sigma^2 > \sigma_0^2$ จะปฏิเสธ Ho ถ้า $X^2 > X^{2(n-1)}_{\alpha}$
- (3) สำหรับ H_a : $\sigma^2 \neq \sigma_0^2$ จะปฏิเสธ Ho ถ้า $X^2 < X^{2(n-1)}_{1-\alpha/2}$ หรือ $X^2 > X^{2(n-1)}_{\alpha/2}$

ในเมื่อ $X^{2(n-1)}_{\alpha}$ และ $X^{2(n-1)}_{1-\alpha}$ เป็นค่าคงที่จากตารางไคสแควร์ที่มีองศาความเป็นอิสระ n − 1

และทำให้เกิดพื้นที่ทางขวามือเท่ากับ α และ $1-\alpha$ ตามลำดับ

ในกรณีตัวอย่างขนาดโต ค่าวิกฤตของตารางไคสแควร์ไม่ได้ทำไว้สำหรับองศาความเป็นอิสระมาก ๆ จึงต้องอาศัยการประมาณค่าด้วยการแจกแจงปกติมาตรฐาน ดังนี้

$$\chi^{2(\nu)}_{\alpha} = (1/2)(Z_{\alpha} + \sqrt{2\nu-1})^2$$

หรือจะใช้ตัวสถิติทดสอบ Z

$$Z = \sqrt{2\chi^2} - \sqrt{2\nu - 1}$$

ซึ่งมีการแจกแจงปกติมาตรฐาน N(0,1)

ตัวอย่าง จากการศึกษาความผันแปรของผลการสอบวิชาสถิติที่ผ่านมาได้ส่วนเบี่ยงเบนมาตรฐาน 12.50 แต่ภาคเรียนใหม่นี้อาจารย์ผู้สอนคาดว่าจะมีความผันแปรน้อยกว่าเพราะมีอุปกรณ์ช่วยสอนมากขึ้น จากผลการสอบของภาคเรียนใหม่นั้นปรากฏว่านักศึกษา 30 คน ที่เป็นตัวอย่างมีส่วนเบี่ยงเบนมาตรฐานของคะแนนสถิติเป็น 10.55

ความผันแปรน่าจะน้อยลงหรือไม่ ?

$$H_o : \sigma^2 \geq 12.50 \quad ; \quad H_a : \sigma^2 < 12.50$$

จากข้อมูลเรามี $n = 30$, $S^2 = 10.55$, $\sigma_o^2 = 12.50$ จึงได้ค่าของตัวสถิติทดสอบเป็น

$$\chi^2 = (n-1)\,S^2/\,\sigma_0^2 = (30-1)\,(10.55)/12.50 = 24.476$$

เมื่อ $\alpha = .01$, $\nu = 30-1 = 29$ เราได้ค่าจากตารางไคสแควร์เป็น $\chi^{2(29)}_{.01} = 14.2565$ จึงสรุปได้ว่า "ปฏิเสธ Ho ไม่ได้" นั่นคือความแปรปรวนในคะแนนสอบวิชาสถิติจะไม่น้อยกว่าที่ระบุไว้

6.4.2.2 ทดสอบการเท่ากันของสองความแปรปรวนประชากร เมื่อตัวอย่างเป็นอิสระกัน เมื่อสองประชากรแบบปกติที่มีค่าเฉลี่ย μ_1, μ_2 และความแปรปรวน σ_1^2, σ_2^2 ซึ่งไม่ทราบค่านั้น ถ้าเราต้องการทดสอบสมมติฐานเกี่ยวกับการเท่ากันของสองความแปรปรวนแล้วสมมติฐานที่จะทดสอบดังนี้

(1) $H_o : \sigma_1^2 \leq \sigma_2^2$ หรือ $H_o : \sigma_1^2/\sigma_2^2 \leq 1$

$H_a : \sigma_1^2 > \sigma_2^2$ หรือ $H_a : \sigma_1^2/\sigma_2^2 > 1$

(2) $H_o : \sigma_1^2 \geq \sigma_2^2$; $H_a : \sigma_1^2 < \sigma_2^2$

(3) $H_o : \sigma_1^2 = \sigma_2^2$; $H_a : \sigma_1^2 \neq \sigma_2^2$

ในการทดสอบเราจะอาศัยตัวอย่างสุ่มขนาด n_1 และ n_2 จากประชากรทั้งสองนั้น แล้วก็ทำการเปรียบเทียบความแปรปรวนในรูปอัตราส่วน นั่นคือตัวสถิติทดสอบสำหรับสมมติฐานหลักนี้จะเป็น

$$F = S_1^2/S_2^2 \quad ; \quad S_1^2 > S_2^2$$

เมื่อสมมติฐานหลัก Ho เป็นจริง ตัวสถิติ F จะมีการแจกแจงแบบเอฟ ด้วยองศาความเป็นอิสระ n_1-1, n_2-1 ดังนั้นเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α กำหนดไว้ตามสมมติฐานรองดังนี้

เมื่อสมมติฐานหลัก Ho เป็นจริง ตัวสถิติ F จะมีการแจกแจงแบบเอฟ ด้วยองศาความเป็นอิสระ n_1-1, n_2-1 ดังนั้นเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α กำหนดไว้ตามสมมติฐานรองดังนี้

(1) Ha : $\sigma_1^2 > \sigma_2^2$ จะปฏิเสธ Ho ถ้า $F > F_{\alpha}^{(\nu_1,\nu_2)}$
(2) Ha : $\sigma_1^2 < \sigma_2^2$ จะปฏิเสธ Ho ถ้า $F < F_{1-\alpha}^{(\nu_1,\nu_2)}$
(3) Ha : $\sigma_1^2 \neq \sigma_2^2$ จะปฏิเสธ Ho ถ้า $F > F_{\alpha/2}^{(\nu_1,\nu_2)}$ หรือ $F < F_{1-\alpha/2}^{(\nu_1,\nu_2)}$

ในเมื่อ $\nu_1 = n_1-1$, $\nu_2 = n_2-1$ กับ $F_{1-\alpha}^{(\nu_1,\nu_2)}$ และ $F_{\alpha}^{(\nu_1,\nu_2)}$ เป็นค่าคงที่จากตารางเอฟซึ่งสอดคล้องกับ ν_1, ν_2 และทำให้เกิดพื้นที่ทางขวามือเป็น $1-\alpha$, α ตามลำดับ

เนื่องจาก $F_{1-\alpha}^{(\nu_1,\nu_2)}$ ในตารางเอฟมักไม่นิยมทำไว้ ดังนั้นในการทดสอบทางด้านซ้าย ($Ha: \sigma_1^2 < \sigma_2^2$) จึงดัดแปลงเป็นทดสอบด้านขวา ($Ha: \sigma_2^2 > \sigma_1^2$) หรือถ้าต้องการทดสอบด้านซ้าย เราต้องหาค่าวิกฤต $F_{1-\alpha}^{(\nu_1,\nu_2)}$ ทางด้านซ้าย ด้วยความสัมพันธ์ที่ว่า

$$F_{1-\alpha}^{(\nu_1,\nu_2)} = 1/F_{\alpha}^{(\nu_2,\nu_1)}$$

ตัวอย่าง ในการศึกษาเพื่อเปรียบเทียบความผันแปรของผลการเรียน (G.P.A.) **ของนักศึกษาชายและหญิงปี** 4 คณะศึกษาศาสตร์ ได้ข้อมูลสรุปดังนี้

$$n_1 = 5,\ n_2 = 12,\ s_1^2 = 27,\ s_2^2 = 15$$

มีเหตุผลเพียงพอหรือไม่ที่จะสรุปว่าความแปรปรวนทั้งสองนี้แตกต่างกัน ?

$$Ho: \sigma_1^2 = \sigma_2^2;\ Ha: \sigma_1^2 \neq \sigma_2^2$$

เกณฑ์ตัดสินใจ ปฏิเสธ Ho **ณ ระดับนัยสำคัญ** .05 **ถ้า** $F > F_{\alpha/2}^{(\nu_1,\nu_2)} = 2.74$; $\nu_1 = 5-1 = 4$, $\nu_2 = 12-1 = 11$

ค่าของตัวสถิติ F ไม่อยู่ในเขตปฏิเสธ จึงยอมรับ Ho นั่นคือ ไม่มีเหตุผลเพียงพอที่จะสรุปว่าความผันแปรในผลการเรียนของชายและหญิงนั้นแตกต่างกัน

เมื่อยอมรับสมมติฐานหลัก $Ho: \sigma_1^2 = \sigma_2^2 = \sigma^2$ เราจึงประมาณค่า σ^2 ได้เป็น S_p^2

$$S_p^2 = \{(n_1-1)S_1^2 + (n_2-1)S_2^2\}/(n_1+n_2-2)$$

ในบางครั้งเราต้องการทดสอบสมมติฐานหลัก Ho ที่ว่าอัตราส่วนของความแปรปรวนประชากรมีค่าหนึ่ง (k_o) หรือความแปรปรวนของประชากรหนึ่งมีค่าเป็นกี่เท่าของอีกประชากรหนึ่ง นั่นคือสมมติฐานหลักจะเป็น

$$Ho : \sigma_1^2 / \sigma_2^2 = k_o \quad \text{หรือ} \quad Ha : \sigma_1^2 = k_o \sigma_2^2$$

ในการทดสอบจึงต้องดัดแปลงตัวสถิติทดสอบเป็น

$$F = (S_1^2/S_2^2)/(\sigma_1^2/\sigma_2^2) = (S_1^2/S_2^2)/k_o$$

ซึ่งมีการแจกแจงแบบเอฟ ด้วยองศาความเป็นอิสระเช่นเดิม

6.4.2.3 ทดสอบการเท่ากันของสองความแปรปรวนเมื่อตัวอย่างมีสหสัมพันธ์กัน ในการทดสอบสมมติฐานเกี่ยวกับการเท่ากันของสองความแปรปรวน หรือ

$$Ho : \sigma_1^2 = \sigma_2^2 ; \quad Ha : \sigma_1^2 \neq \sigma_2^2$$

นั้นเมื่อตัวอย่างทั้งสองที่มีขนาด n มีความสัมพันธ์กัน ตัวสถิติทดสอบจะเป็น

$$T = r / \sqrt{(1-r^2)/(n-2)}$$

ในเมื่อ r เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่าง S_i และ D_i โดยที่ $S_i = X_{1i} - X_{2i}$ และ

$D_i = X_{1j} = X_{2i} \qquad (i=1,2,\ldots n)$

ตัวสถิติทดสอบ T จะมีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ n-2 ถ้าสมมติฐานหลักเป็นจริง

ตัวอย่าง จากการใช้ตัวอย่างขนาด 16 เพื่อที่จะทดสอบสมมติฐานว่า ความแปรปรวนของคะแนนสอบในวิชาคณิตศาสตร์ และสถิติศาสตร์ ของนักศึกษาจิตวิทยาจะไม่แตกต่างกัน คะแนนที่ได้จากตัวอย่างเป็นดังนี้

นักศึกษา	คณิตศาสตร์	สถิติศาสตร์	นักศึกษา	คณิตศาสตร์	สถิติศาสตร์
1	5.0	4.9	9	5.3	5.2
2	4.8	5.0	10	5.3	5.5
3	4.3	4.3	11	5.3	5.5
4	5.1	5.3	12	5.9	5.9
5	4.1	4.1	13	6.5	6.8
6	4.0	4.0	14	6.6	6.6

7	7.1	6.9	15	5.2	4.8
8	5.9	6.3	16	5.2	6.3

$$Ho : \sigma_1^2 = \sigma_2^2 \quad ; \qquad Ha : \sigma_1^2 \neq \sigma_2^2$$

จากข้อมูล เราได้ $\Sigma S_i = 176.6$, $\Sigma S_i^2 = 200\ 1.16$, $\Sigma S_i D_i = -13.64$, $\Sigma D_i = -1.2$, $\Sigma D_i^2 = 0.48$ ดังนั้น r จะเป็น

$$r = \{n \Sigma S_i D_i - (\Sigma S_i)(\Sigma D_i)\} / \sqrt{(\Sigma S_i^2 - (\Sigma S_i)^2)(n\Sigma D_i^2 - (\Sigma D_i)^2)}$$

$$= \{16(-13.64) - 176.6(-1.2)\} / \sqrt{(12(2001.16) - (176.6)^2)\ (12(0.48) - (-1.2)^2)}$$

$$= 0.0878$$

ดังนั้น $\quad T \quad = \quad -0.0878/\sqrt{(1-(-0.0878)^2)/(16-2)}$

$$= \quad -0.330$$

สำหรับ $\alpha = 0.05$ เราได้ $t_{.025}^{(16-2)} = 2.145$ จึงปฏิเสธ Ho ไม่ได้ นั่นคือความแปรปรวนในวิชาทั้งสองของนักศึกษาจะไม่แตกต่างกัน

ตัวสถิติทดสอบสำหรับสมมติฐานหลักดังกล่าว เราสามารถใช้ตัวสถิติ T ดังต่อไปนี้ได้

$$T = (S_2^2 - S_1^2) / (2S_1S_2\sqrt{(1-r^2)/(n-2)})$$

ในเมื่อ r เป็นสัมประสิทธิ์สหพันธ์ระหว่าง X_{1i} กับ X_{2i} ($i = 1,2,\ldots,n$) และตัวสถิตินี้มีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ n-2

ตัวอย่าง ในการศึกษาเพื่อทดสอบการเท่ากันของสองความแปรปรวนโดยที่ตัวอย่างทั้งสองมีสหสัมพันธ์กัน และมีขนาด 54 เราได้ข้อมูลสรุปดังนี้

$$S_1 = 3.75, \quad S_2 = 12.28, \quad r = .65$$

สมมติฐาน

$$Ho : \sigma_1^2 = \sigma_2^2 \quad ; \qquad Ha : \sigma_1^2 \neq \sigma_2^2$$

$$T = \{(12.28)^2 - (3.75)^2\} / 2(3.75)(12.28)\sqrt{(1-(.65)^2)/(54-2)}$$

$$= \quad 14.10$$

เมื่อ $\alpha = .05$ เราได้ $t_{.025}^{(54-2)} = 2.00$ จึงปฏิเสธ Ho นั่นคือความแปรปรวนแตกต่างกัน

บางครั้งเราต้องการทดสอบความแตกต่างของส่วนเบี่ยงเบนมาตราฐานที่มีสหสัมพันธ์กันโดยใช้ตัวอย่างขนาดโต ($n \rightarrow \infty$) เราใช้ตัวสถิติทดสอบ

$$T = (S_1 - S_2 - \Delta)/S_{(s_1 - s_2)}$$

ในเมื่อ $\triangle$ เป็นความแตกต่างของส่วนเบี่ยงเบนมาตราฐานที่ระบุไว้ และ $S_{(s_1-s_2)}$ หาได้จาก

$$S^2_{(s_1-s_2)} = S^2_{s_1} + S^2_{s_2} - 2r^2 S_{s_1} S_{s_2}$$

โดยที่ $S_{s_i} = 0.71 \; S_i/\sqrt{n} \quad ; \quad i = 1,2$

ตัวสถิติทดสอบ T นี้จะมีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ n-1 ถ้าสมมติฐานหลัก Ho เป็นจริง

ตัวอย่าง ในการศึกษาผลการเรียนวิชาสถิติ 1 และสถิติ 2 ของนักศึกษาศึกษาศาสตร์โดยอาศัยตัวอย่างขนาด 64 ปรากฏว่าได้ส่วนเบี่ยงเบนมาตร ฐานในวิชาสถิติทั้งสอง เป็น 6.00 และ 5.00 และสัมประสิทธิ์สหสัมพันธ์ระหว่างวิชาทั้งสองเป็น 0.60

มีเหตุผลเพียงพอที่จะสรุปว่าส่วนเบี่ยงเบนมาตร ฐานวิชาสถิติทั้งสองไม่แตกต่างกันหรือไม่ ?

$$Ho : \sigma_1 = \sigma_2; \qquad Ha : \sigma_1 \neq \sigma_2$$

$$S_{s_1} = 0.71(6.00)/\sqrt{64} = 0.53, \qquad S_{s_2} = 0.71(5.00)/\sqrt{64} = 0.44$$

$$S_{(s_1-s_2)} = \sqrt{(.53)^2+(.44)^2-2(.60)(.53)(.44)} = 0.55$$

$$T = \{(6.00-5.00)-0\}/0.55 = 1.82$$

เมื่อ α = .05 เราได้ $t^{(64-1)}_{.025}$ = 2.00 จึงสรุปได้ว่า ส่วนเบี่ยงเบนมาตร ฐานของวิชาสถิติทั้งสองจะไม่แตกต่างกัน

6.4.2.4 ทดสอบการเท่ากันของความแปรปรวนหลายประชากร ในกรณีหลายประชากรแบบปกติ เมื่อเราต้องการทดสอบการเท่ากันของความแปรปรวน นั่นคือสมมติฐานที่ทดสอบจะเป็น

$$Ho : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$$

หรือ $Ho : \sigma_j^2 = \sigma^2 \quad ; \quad \forall j$

$$Ha : \sigma_j^2 \neq \sigma^2 \quad ; \quad \exists j \qquad j = 1,2,\ldots k$$

เราจะอาศัยตัวอย่างสุ่มที่เป็นอิสระกันขนาด $n_1, n_2, \ldots, n_k$ ตามลำดับจากประชากรแบบปกติที่สนใจเหล่านั้น ตัวสถิติทดสอบสำหรับสมมติฐานหลัก จะเป็น

ก. ตัวสถิติฮาร์ทเลย์ (Hartley's Test statistic) ซึ่งใช้ขนาดตัวอย่างสุ่มเท่า ๆ กัน และกำหนดไว้ว่า

$$F_{max} = S^2_{max} / S^2_{min}$$

ในเมื่อ S^2_{max} และ S^2_{min} เป็นค่าความแปรปรวนที่มากที่สุดและน้อยที่สุดจากตัวอย่าง

เกณฑ์ตัดสินใจสำหรับตัวสถิติ F_{max} นี้จะอาศัยตารางพิเศษที่สอดคล้องกับกลุ่มตัวอย่าง (k) และขนาดตัวอย่างลบด้วย 1 หรือ n-1 เมื่อขนาดตัวอย่างไม่เท่ากันก็ให้เอาขนาดตัวอย่างของกลุ่มที่มากที่สุด หรือเฉลี่ยขนาดตัวอย่างแบบฮาร์โมนิค ซึ่งจะให้เป็น n_o

ข. ตัวสถิติคอคราน (Cochran's Test statistic) ตัวสถิตินี้จะใช้ได้ดีกว่าตัวสถิติฮาร์ทเลย์ เพราะใช้ข้อเท็จจริงจากข้อมูลที่รวบรวมได้มากกว่า นั่นคือตัวสถิติทดสอบกำหนดไว้เป็น

$$C = S^2_{max} \Big| \sum_j^k S_j^2$$

ตัวสถิติ C นี้จะใช้กับตัวอย่างที่มีขนาดเท่ากันเช่นเดียวกับตัวสถิติ F_{max}

เกณฑ์ตัดสินใจสำหรับตัวสถิติ C นี้อาศัยตารางพิเศษเช่นเดียวกับตัวสถิติฮาร์ทเลย์

ค. ตัวสถิติบาร์ทเลทท์ (Bartlett's Test statistic) ตัวสถิตินี้นิยมใช้กันแพร่หลาย แต่การคำนวณยุ่งยากกว่าตัวสถิติทั้งสองที่กล่าวมาแล้ว สำหรับขนาดตัวอย่างก็ไม่จำเป็นต้องเท่ากัน แต่ก็ไม่ควรน้อยกว่า 3 ซึ่งส่วนมากควรจะมากกว่า 5 ตัวสถิติบาร์ทเลทท์ กำหนดไว้ดังนี้

ในเมื่อ

$$B = (1/a)(\nu \ln(\sum \nu_j S_j^2/\nu) - \sum \nu_j \ln S_j^2)$$

$$=(2.3026/a)\nu(\log \sum \nu_j S_j^2/\nu) - \sum \nu_j \log S_j^2)$$

$$a = 1+1/\{3(k-1)\}(\sum 1/\nu_j - 1/\nu);\ \nu_j = n_j - 1;\ \nu = \sum \nu_j = n-k$$

เมื่อตัวอย่างขนาดโต และสมมติฐานหลักเป็นจริง ตัวสถิติ B นี้จะมีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ k-1 ดังนั้นเกณฑ์ตัดสินใจสำหรับสมมติฐานหลักจึงกำหนดไว้ว่า "ปฏิเสธ Ho ถ้า B > $\chi^{2(k-1)}_{\alpha}$

ตัวอย่าง จากการศึกษารายได้ต่อวันของกลุ่มกรรมกร 4 ประเภท โดยอาศัยตัวอย่าง ได้ความแปรปรวนของรายได้ ดังตารางต่อไปนี้

กลุ่ม	n_j	S_j^2	$\nu_j S_j^2$	$\nu_j \log S_j^2$	$1/\nu$
1	9	84.2	673.6	15.40248	0.125
2	21	63.8	1276.0	36.09650	0.050
3	6	88.6	443.0	9.73715	0.2
4	11	72.1	721.0	18.57940	0.10

47 308.7 3113.6 79.81543 0.475

กรรมกรทั้ง 4 กลุ่ม มีความแปรปรวนในรายได้ต่อวันแตกต่างกันหรือไม่ ?

$$Ho : \sigma_j^2 = \sigma^2 ; \quad \forall_j$$
$$Ha : \sigma_j^2 \neq \sigma^2 ; \quad \exists_j \qquad j = 1,2,3,4$$

เมื่อใช้ตัวสถิติทดสอบ B เราจะได้ค่าเป็น

$$B = 2.3026(43(1.5979)-79.81543)/\{1+(1/3(4-1))(0.475-1/43)\} = 0.341$$

ในเมื่อ $\nu = 47-4 = 43$, $\log \sum \nu_j S_j^2/\nu = \log(3113.6/43) = 1.85979$

สำหรับ $\alpha = .05$ เราได้ $X^{2(4-1)}_{.05} = 7.81$ จึงปฏิเสธ Ho ไม่ได้ เพราะ $B < 7.81$ แสดงว่าทั้ง 4 กลุ่ม มีความแปรปรวนไม่แตกต่างกัน

เมื่อใช้ตัวสถิติ C เราได้ค่าเป็น

$$C = 88./308.7 = 0.2871$$

สำหรับ $\alpha = .05$ เราได้ค่าวิกฤต $C_{\alpha}^{(k,n_o-1)} = C_{.05}^{(4,20)} = .4366$ จึงปฏิเสธ Ho ไม่ได้เช่นเดียวกัน ($C < .4366$)

เมื่อใช้ตัวสถิติ F_{max} เราได้ค่าเป็น

$$F_{max} = 88.6/63.8 = 1.3888$$

สำหรับ $\alpha = .05$ เราได้ค่าวิกฤต $H^{(k,n_o-1)} = H_{.05}^{(4,20)} = 3.29$ จึงปฏิเสธไม่ได้ เพราะ $F_{max} < 3.29$

เนื่องจากตัวอย่างไม่เท่ากัน ตัวสถิติ F_{max} และ C จึงใช้ขนาดตัวอย่างที่มากที่สุดคือ $n_o = 21$

ในการทดสอบสมมติฐานเกี่ยวกับการเท่ากันของหลายความแปรปรวนนั้น เมื่อยอมรับสมมติฐานหลัก Ho : $\sigma_j^2 = \sigma^2, \forall_j$ ($j = 1,2,\ldots,k$) เราสามารถประมาณ σ^2 แบบช่วงได้โดยอาศัยตัวประมาณค่า S_p^2 นั่นคือช่วงเชื่อมั่น $100(1-\alpha)$ % สำหรับ σ^2 จะเป็น

$$(n-k)S_p^2/X^{2(n-k)}_{\alpha/2} \leq \sigma^2 \leq (n-k)S_p^2/X^{2(n-k)}_{1-\alpha/2}$$

ในเมื่อ $S_p^2 = \sum(n_j-1)S_j^2/(n-k)$; $n = n_1+n_2+\ldots+n_k$

แต่ถ้าสมมติฐาน Ho ได้รับการปฏิเสธ เราก็สามารถที่จะหาว่าความแปรปรวนคู่ใด

ที่แตกต่างกันบ้าง โดยอาศัยช่วงเชื่อมั่นสำหรับอัตราส่วนคู่ใด ๆ นั้น นั่นคือช่วงเชื่อมั่น $100(1-\alpha)$ % สำหรับ σ_i^2/σ_j^2 จะเป็น

$$(S_i^2/S_j^2)/F_{\alpha_0/2}^{(\nu_i,\nu_j)} \leq \sigma_i^2/\sigma_j^2 \leq (S_i^2/S_j^2)/F_{1-\alpha_0/2}^{(\nu_i,\nu_j)}$$

ในเมื่อ $i,j = 1,2,\ldots k(i < j)$; $\alpha_0 = \alpha/\binom{k}{2}$; $\nu_i = n_i-1$ $\nu_j = n_j-1$

กรณีที่ความแปรปรวนต่าง ๆ นั้นเท่ากัน และถ้าเราต้องการทดสอบว่าค่าที่เท่ากันนั้นจะเท่ากับค่าที่ระบุไว้ (σ_o^2 หรือไม่ เราก็ทำการทดสอบได้โดยมีสมมติฐานหลักเป็น

$$Ho: \sigma^2 = \sigma_o^2$$

และตัวสถิติทดสอบเป็น

$$X^2 = (n-k)S_p^2/\sigma_o^2$$

ซึ่งมีการแจกแจงแบบไคสแควร์ ด้วยองศาความเป็นอิสระ n-k

6.4.3 การทดสอบเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์ (Test of Correlation Coefficient)

ในกรณีประชากรแบบปกติชนิดสองตัวแปรนั้น ถ้าสนใจประชากรเดียวเราก็ทำการทดสอบค่าที่ระบุไว้ของสัมประสิทธิ์สหสัมพันธ์ (ρ_o) แต่ถ้ามีประชากรที่สนใจตั้งแต่สองประชากรขึ้นไป เราจะทำการทดสอบการเท่ากันของสัมประสิทธิ์สหสัมพันธ์

6.4.3.1 ทดสอบสัมประสิทธิ์สหสัมพันธ์ที่ระบุไว้ (ρ_o) ถ้าประชากรแบบปกติชนิดสองตัวแปรมีสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรทั้งสองเป็น ρ ซึ่งเราไม่ทราบค่า และเมื่อเรามีสมมติฐานเกี่ยวกับ ρ ว่าเป็น ρ_o นั่นคือเราทดสอบสมมติฐานแบบใดแบบหนึ่งดังนี้

(1) Ho : $\rho \geq \rho_o$; Ha : $\rho < \rho_o$

(2) Ho : $\rho \leq \rho_o$; Ha $\rho > \rho_o$

(3) Ho : $\rho = \rho_o$; Ha : $\rho \neq \rho_o$

ในการทดสอบก็อาศัยตัวอย่างสุ่มจากประชากรที่สนใจนั้น โดยมีตัวสถิติทดสอบเป็น

$$T = (Z - Z_o)/\sqrt{(1/(n-3))} = (Z-Z_o)\sqrt{n-3}$$

ในเมื่อ $Z = \tan h^{-1} r = (1/2)\ln\{(1+r)/(1-r)\}$ และ $Z_o = \tan h^{-1} \rho_o = (1/2)\ln((1+\rho_o)/(1-\rho_o))$

ค่า Z และ Z_o พิจารณาได้จากตารางแปลงค่าที่ชื่อว่า Fisher Z Transformation Table

ตัวสถิติทดสอบ T จะมีการแจกแจงแบบปกติมาตรฐาน ถ้าสมมติฐานหลักเป็นจริง ดังนั้นเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α จึงกำหนดได้ดังนี้

(1) สำหรับ Ha : $\rho < \rho_0$ จะปฏิเสธ Ho ถ้า $T < -Z_{\alpha}$

(2) สำหรับ Ha : $\rho > \rho_0$ จะปฏิเสธ Ho ถ้า $T > Z_{\alpha}$

(3) สำหรับ Ha : $\rho \neq \rho_0$ จะปฏิเสธ Ho ถ้า $T < -Z_{\alpha/2}$ หรือ $T > Z_{\alpha/2}$

ในกรณีที่ $\rho_0 = 0$ หรือทดสอบสมมติฐานหลัก Ho : $\rho = 0$ แล้วตัวสถิติทดสอบจะเป็น

$$T_1 = r \Big/ \sqrt{(1-r^2)/(n-2)}$$

ซึ่งมีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ n-2 เกณฑ์ตัดสินใจจึงอาศัยค่าวิกฤตจากตาราง ที่ ช่วยกำหนด

บางครั้งเราต้องการประมาณค่า ρ แบบช่วง เราพิจารณาได้จากช่วง เชื่อมั่นของ E (Z) นั่นคือ

$$E(Z) = Z \pm Z_{\alpha/2}\sigma_Z \quad ; \qquad \sigma_Z = 1/\sqrt{n-3}$$

แล้วแปลงค่า E(Z) กลับไปเป็น ρ ค่าเดิมอีก โดยอาศัยตารางแปลงค่า และเราจะได้ช่วงเชื่อมั่น $100(1-\alpha)\%$ ของ ρ

ตัวอย่าง (1) ในการทดสอบคำกล่าวที่ว่า "สัมประสิทธิ์สหสัมพันธ์ระหว่างระดับสติปัญญากับผลการเรียนเศรษฐศาสตร์จะมีค่าประมาณ 0.80" นั้นได้อาศัยตัวอย่างขนาด 40 ของนักศึกษากลุ่มที่สนใจนั้น ปรากฏว่าได้สัมประสิทธิ์สหสัมพันธ์ตัวอย่างเป็น 0.8895

ผลสรุปของการทดสอบจะเป็นอย่างไร ?

$$Ho : \rho = 0.80 \ ; \qquad Ha : \rho \neq 0.80$$

จากตารางแปลงค่า เราได้ $Z = 1.4195$ เมื่อ $r = 0.8895$ และ $Z_0 = 1.0990$ เมื่อ $\rho_0 = 0.80$ ดังนั้น สถิติทดสอบ T จึงมีค่าเป็น

$$T = (1.4195-1.0990)/\sqrt{40-3} = 1.9495$$

สำหรับ $\alpha = .05$ เราได้ $Z_{.025} = 1.96$ จึงปฏิเสธ Ho ไม่ได้ ($Z < 1.96$) นั่นคือสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรทั้งสองไม่แตกต่างจาก 0.80

(2) ในการทดสอบสมมติฐานที่ว่า "ตัวแปร X กับตัวแปร Y มีความสัมพันธ์กันทางบวก ($\rho > 0$)" ได้อาศัยตัวอย่างสุ่มขนาด 27 ปรากฏว่าได้ค่า r เป็น 0.60

สมมติฐานที่กล่าวนั้นน่าจะเชื่อถือได้หรือไม่?

$$Ho : \rho = 0; \ Ha : \rho > 0$$

สำหรับ $n = 27, r = 0.60$ เราได้ค่าของตัวสถิติทดสอบ T_1 เป็น

$$T_1 = 0.60 / \sqrt{(1-0.60^2)/(27-2)} = 3.75$$

เมื่อ α = .05 เราได้ $t_{.05}^{(25)}$ = 1.708 จึงปฏิเสธ Ho นั่นคือตัวแปรทั้งสองมีความสัมพันธ์ทางบวก

(3) จากการศึกษาความสัมพันธ์ระหว่างตัวแปรสองตัวของประชากรแบบปกติชนิดสองตัวแปรโดยใช้ขนาดตัวอย่างเป็น 22 ปรากฏว่าได้สัมประสิทธิ์สหสัมพันธ์ตัวอย่างเป็น .76 จงประมาณสัมประสิทธิ์สหสัมพันธ์ ระหว่างตัวแปรทั้งสอง โดยใช้ระดับความเชื่อมั่น 0.95

จากตารางแปลงค่า เมื่อ r = .76 เราได้ Z = .99621 ดังนั้นช่วงเชื่อมั่น 95% สำหรับ E(Z) จึงเป็น

$$E(Z) = .99621 \pm 1.96\,(1/\sqrt{22-3})$$
$$= .99621 \pm .4479 = .54831,\ 1.44411$$

เมื่อแปลงช่วงของ E(Z) เราจึงได้ช่วงเชื่อมั่น 95% สำหรับ ρ เป็น

$$\rho = 0.495,\ 0.895$$

6.4.3.2 ทดสอบสัมประสิทธิ์สหสัมพันธ์ที่ไม่เกี่ยวพันกัน ในกรณีที่สนใจสัมประสิทธิ์สหสัมพันธ์ของสองประชากรที่เป็นอิสระกันเพื่อต้องการเปรียบเทียบความแตกต่าง หรือทดสอบสมมติฐานหลัก

$$Ho: \rho_1 = \rho_2$$

เราก็ใช้ตัวอย่างสุ่มที่เป็นอิสระกันจากสองประชากรนั้นด้วยขนาด n_1 และ n_2 และกำหนดตัวสถิติทดสอบไว้เป็น

$$T = (Z_1 - Z_2)/\sqrt{1/(n_1-3) + 1/(n_2-3)}$$

ในเมื่อ $Z_1 = (1/2)\ln\{(1+r_1)/(1-r_1)\}$ และ $Z_2 = (1/2)\ln\{(1+r_2)/(1-r_2)\}$

ตัวสถิติ T นี้จะมีการแจกแจงแบบปกติมาตรฐาน ถ้าสมมิฐานหลักเป็นจริง ดังนั้นเกณฑ์ตัดสินใจจึงอาศัยการแจกแจงแบบปกติมาตรฐานนี้

ในกรณีที่เรายอมรับสมมติฐานหลัก Ho : $\rho_1 = \rho_2$ เราก็สามารถหาค่าประมาณที่ดีของ $\rho_1 = \rho_2 = \rho$ ได้จากค่าเฉลี่ยของ Z_1 และ Z_2 ดังนี้

$$Z = \{(n_1-3)Z_1 + (n_2-3)Z_2\}/\{(n_1-3)+(n_2-3)\}$$

แล้วแปลงค่า Z กลับเป็น r อีก ซึ่งค่า r นี้จะเป็นค่าประมาณที่ดีของ ρ นั่นเอง

ถ้าเราต้องการประมาณผลต่างของ $\rho_1 - \rho_2$ ในเทอมของ $E(Z_1) - E(Z_2)$ ก็จะได้ช่วงเชื่อมั่น 100(1-α) % สำหรับ $E(Z_1) - E(Z_2)$ ดังนี้

$$E(Z_1)-E(Z_2) = (Z_1-Z_2) \pm Z_{\alpha/2}\sqrt{1/(n_1-3)+1/(n_2-3)}$$

ตัวอย่าง ในการศึกษาถึงความสัมพันธ์ระหว่างความถนัดทางการเรียนและผลการเรียนทางเศรษฐศาสตร์ เบื้องต้นของนักศึกษาชายหญิง โดยอาศัยตัวอย่างของนักศึกษาชายและหญิงกลุ่มละ 53 ราย ได้สัมประสิทธิ์สหสัมพันธ์เป็น $r_1 = .32$ และ $r_2 = .56$ ตามลำดับ

ความแตกต่างระหว่างสัมประสิทธิ์สหสัมพันธ์จะมีจริงหรือไม่?

$$Ho : \rho_1 = \rho_2 \; ; \qquad Ha : \rho_1 \neq \rho_2$$

สำหรับ $r_1 = .32$ และ $r_2 = .56$ เมื่อแปลงเป็นค่า Z ได้เป็น $Z_1 = .33165$ และ $Z_2 = .63283$ ดังนั้นตัวสถิติทดสอบจะมีค่าเป็น

$$T = (.33165-.63283)/\sqrt{1/(53-3)-1/(53-3)} = -1.56$$

เมื่อ $\alpha = .05$ ค่าวิกฤตเป็น $\pm Z_{.025} = \pm 1.96$ จึงปฏิเสธ Ho ไม่ได้ นั่นคือความแตกต่างระหว่างสัมประสิทธิ์สหสัมพันธ์จะไม่มี

6.4.3.3 ทดสอบการเท่ากันของสัมประสิทธิสหสัมพันธ์ที่เกี่ยวข้องกัน ในบางครั้งสองสัมประสิทธิ์สหสัมพันธ์ ρ_1 และ ρ_2 มีความสัมพันธ์กัน เช่นสัมประสิทธิ์ ρ_1 และ ρ_2 ที่พิจารณาจากประชากรเดียวกัน ในเมื่อ ρ_1 เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X_1 กับ X_2 และ ρ_2 เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X_1 และ X_3 เมื่ออยากทราบว่าสัมประสิทธิ์สหสัมพันธ์ทั้งสองนั้นแตกต่างกันหรือไม่ เราก็ทำการทดสอบสมมติฐานหลัก

$$Ho : \rho_{12} = \rho_{13}$$

โดยอาศัยตัวอย่างขนาด n จากประชากรที่สนใจนั้น และกำหนดตัวสถิติทดสอบที่ Hotelling เสนอไว้คือ

$$T = (r_{12}-r_{13})\sqrt{(n-3)(1+r_{23})/2(1-r_{23}^2-r_{12}^2-r_{13}^2+2r_{23}r_{12}r_{13})}$$

ในเมื่อ r_{12}, r_{13}, และ r_{23} เป็นสัมประสิทธิ์สหสัมพันธ์ตัวอย่างระหว่างตัวแปร 1 กับ 2, 1 กับ 3, และ 2 กับ 3 ตามลำดับ

ตัวสถิติ T นี้มีการแจกแจงแบบที ด้วยองศาความเป็นอิสระ n-3 และตัวสถิติ $F = T^2$ จะมีการแจกแจงแบบเอฟ ด้วยองศาความเป็นอิสระ 1, n-3

ตัวอย่าง ในการศึกษาถึงคะแนนความถนัดทางการเรียน คะแนนวิชาสถิติ และคะแนนวิชาเศรษฐศาสตร์ เพื่อทดสอบว่าความสัมพันธ์ระหว่างคะแนนความถนัดทางการเรียนกับคะแนนวิชาเศรษฐศาสตร์ จะมากกว่าความสัมพันธ์ระหว่างคะแนนความถนัดทางการเรียนกับคะแนนวิชาสถิติ ได้อาศัยตัวอย่างของนักศึกษาเศรษฐศาสตร์ 200 ราย ปรากฏว่าได้ $r_{12} = .55$, $r_{13} = .45$, และ $r_{23} = .60$

มีเหตุผลเพียงพอที่จะสรุปตามสมมติฐานนั้นหรือไม่?

$$Ho : \rho_{12} = \rho_{13} \quad ; \quad Ha : \rho_{12} > \rho_{13}$$

$$T = (.55-.45)\sqrt{\{(200-3)(1-.60)\}/2(1-.60^2-.55^2-.45^2+2(.60)(.55)(.45))}$$
$$= 1.91$$

เมื่อ α = .05 เราได้ $t_{.05}^{(200-3)}$ = 1.645 จึงสรุปได้ตามสมมติฐานที่ระบุไว้

6.4.3.4 ทดสอบการเท่ากันของสัมประสิทธิ์สหสัมพันธ์จากหลาย ๆ ประชากร ถ้าเราสนใจประชากรแบบปกติชนิดสองตัวแปรต่าง ๆ ที่เป็นอิสระกัน k ประชากร โดยมี $\rho_1, \rho_2, \ldots, \rho_k$ เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างสองตัวแปร และต้องการทดสอบสมมติฐานที่ว่า

$$Ho : \rho_1 = \rho_2 = \ldots\ldots = \rho_k = \rho$$

หรือ $$Ho : \rho_j = \rho \; ; \forall_j \; j = 1, 2, \ldots, k$$

$$Ha : \rho_j \neq \rho \; ; \exists_j \; j = 1, 2, \ldots k$$

เราก็อาศัยตัวอย่างสุ่มจากประชากรต่าง ๆ และใช้ตัวสถิติทดสอบ X^2 ดังนี้

$$X^2 = \Sigma(n_j-3)(Z_j-Z_o)^2$$

ในเมื่อ $Z_j = (1/2)\ln\{(1+r_j)/(1-r_j)\}; j = 1, 2, \ldots\ldots,$ และ $Z_o = (\sum(n_j-3)Z_j)/(\Sigma(n_j-3))$

ตัวสถิติทดสอบ X^2 นี้จะมีการแจกแจงแบบไคสแควร์ ด้วยองศาความเป็นอิสระ k-1 ถ้าสมมติฐานหลัก Ho เป็นจริง

ถ้าสมมติฐานหลักได้รับการยอมรับ แล้วค่าประมาณที่ดีของ ρ คำนวณได้จากการแปลงค่า Z_o เป็น r ในเมื่อ Z_o เป็นค่าเฉลี่ยของ Z_j และกำหนดไว้ดังที่กล่าวมาแล้ว นั่นคือ

$$Z_o = \{\Sigma(n_j-3)Z_j\}/\{\Sigma(n_j-3)\}$$

ถ้าปฏิเสธ Ho และต้องการทราบว่าประชากรคู่ใดบ้างมีสัมประสิทธิ์สหสัมพันธ์เท่ากัน เราก็อาศัยช่วงเชื่อมั่น 100(1-α) % สำหรับ $E(Z_i) - E(Z_j)$ ดังนี้

$$E(Z_i)-E(Z_j) = (Z_i-Z_j) \pm \sqrt{X_{\alpha}^{2(k-1)}}\sqrt{1/\{(n_i-3)+1/(n_j-3)\}}$$

ในเมื่อ $i, j = 1, 2, \ldots k \; ; \; i < j$

เมื่อสนใจทดสอบความแผกกัน Ψ เราก็อาศัยช่วงเชื่อมั่น 100(1-α) % สำหรับดังนี้

$$\Psi = \hat{\Psi} \pm \sqrt{X_{\alpha}^{2(k-1)}}$$

ในเมื่อ $\hat{\psi} = a_1 Z_1 + a_2 Z_2 + \ldots + a_k Z_k$ และ $S^2_{\hat{\psi}} = \Sigma a_j^2/(n_j-3)$

ตัวอย่าง ในการศึกษาความสัมพันธ์ระหว่างตัวแปรทางจิตวิทยาสองตัวของกลุ่มกรรมการ 5 กลุ่ม โดยอาศัยตัวอย่างสุ่ม ได้ข้อมูลมาดังนี้

กลุ่ม	1	2	3	4	5
n_j	58	68	113	37	91
r_j	.66	.70	.68	.92	.44
Z_j	.793	.867	.829	1.589	.472
$S^2 Z_j = 1/(n_j-3)$	.0182	.0154	.0091	.0294	.0114

มีเหตุผลเพียงพอหรือไม่ที่จะสรุปว่า ตัวอย่างจาก 5 กลุ่ม นั้นให้สัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรทั้งสองไม่เท่ากันหมด

$$\text{Ho}: \rho_j = \rho \; ; \quad \forall j$$
$$\text{Ha}: \rho_j \neq \rho \; ; \quad \exists j \qquad j=1,2,\ldots k$$

เมื่อ $Z_o = \{55(.793) + 65(.867) + \ldots + 88(.472)\}/(55+65+\ldots+88) = 0.814$

เราได้

$$X^2 = 55(.793-.814)^2 + 65(.867-.814)^2 + \ldots\ldots + 88(.472-.814)^2$$
$$= 30.94$$

เนื่องจากค่าวิกฤตสำหรับ $\alpha = .05$ **เป็น** $X^2_{.05}(5-1) = 9.49$ **จึงปฏิเสธ** Ho **และ สรุป** ได้ว่า สัมประสิทธิ์สหสัมพันธ์ของสองตัวแปรจากกลุ่มกรรมกรทั้ง 5 กลุ่ม ไม่เท่ากันหมด

เมื่อเราต้องการทราบว่าคู่ไหนแตกต่างกันบ้าง เราก็อาศัยช่วงเชื่อมั่น 95% สำหรับ $E(Z_j)$ **ดังนี้**

$E(Z_1)-E(Z_2) = (.793-.867) \pm \sqrt{9.49}\sqrt{.0182+.0154}$
$= -.64, .49$ **ไม่มีนัยสำคัญ**

$E(Z_1)-E(Z_3) = (.793-.829) + \sqrt{9.49}\sqrt{.0182+.0091}$
$= -.54, .47$ **ไม่มีนัยสำคัญ**

$E(Z_1)-E(Z_4) = -1.47, -.12$ **มีนัยสำคัญ**

$E(Z_1)-E(Z_5) = -.21, .85$ **ไม่มีนัยสำคัญ**

$E(Z_2)-E(Z_3) = -.44, .52$ **ไม่มีนัยสำคัญ**

$E(Z_2)-E(Z_4)$ = -1.37, -.07 มีนัยสำคัญ

$E(Z_2)-E(Z_5)$ = -.11, .90 ไม่มีนัยสำคัญ

$E(Z_3)-E(Z_4)$ = -1.36, - 16 มีนัยสำคัญ

$E(Z_3)-E(Z_5)$ = -.08, .80 ไม่มีนัยสำคัญ

$E(Z_4)-E(Z_5)$ = .50, 1.74 มีนัยสำคัญ

เราจะเห็นได้ว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรทั้งสองของกลุ่มกรรมกรต่าง ๆ จะแตกต่างจากกลุ่ม 4

ถ้าเราจะสนใจจะเปรียบเทียบกรรมกรกลุ่ม 1, 2, 3 ซึ่งถือว่าเป็นกลุ่มหนึ่งนั้นกับกลุ่มกรรมกร 5 เราจะได้ความแผกกันเป็น

$$\psi = \frac{1}{3}(Z_1 + Z_2 + Z_3) - Z_5$$

ซึ่งประมาณค่าได้เป็น $\hat{\psi} = (1/3)(.793 + .867 + .829) - .472 = .358$

และมีความแปรปรวนเป็น $S^2_{\hat{\psi}} = (1/9)(S^2_{\hat{\psi}_1} + S^2_{\hat{\psi}_2} + S^2_{\hat{\psi}_3}) + S^2_{\hat{\psi}_5}$

$$= (1/9)(.0182 + .0154 + .0091) + .0114 = .0161$$

ดังนั้นช่วงเชื่อมั่น 95% สำหรับความแผกกัน ψ จะเป็น

$$\psi = \hat{\psi} \pm \sqrt{\chi^2_{.05}(k-1)}\, S_{\hat{\psi}}$$

$$= .358 \mp \sqrt{9.49}\sqrt{.0161} = -.032,\ .748$$

เนื่องจากช่วงเชื่อมั่นนี้รวม 0 ไว้ด้วย จึงเห็นได้ว่าความแผกกันที่สนใจนั้นไม่มีนัยสำคัญ

6.5 การทดสอบพารามิเตอร์ของประชากรทวินาม (Tests of Parameters of Binomial Population)

ประชากรทวินามจะมีหน่วยในประชากรแบ่งออกเป็น 2 ประเภท ซึ่งจะให้เป็น S และ F หรือ 1 และ 0 และสัดส่วนของหน่วยที่เป็น S หรือ 1 จะให้เป็น π ซึ่งจะเป็นพารามิเตอร์ประชากร และค่าของ π จะอยู่ระหว่าง 0 กับ 1 ดังนั้นการทดสอบสมมติฐานเกี่ยวกับสัดส่วนของประชากรทวินามที่สนใจก็คือการทดสอบสัดส่วนที่ระบุไว้ ทดสอบความแตกต่างของสองสัดส่วน และทดสอบการเท่ากันของสัดส่วนในหลาย ๆ ประชากร

6.5.1 ทดสอบสัดส่วนหรือเปอร์เซ็นต์ที่ระบุไว้ (π_o) สมมติฐานเกี่ยวกับสัดส่วนประชากรทวินามที่ระบุไว้เป็น π_o นั้นจะเป็นแบบใดแบบหนึ่ง ดังนี้

$$(1)\ Ho: \pi = \pi_o;\ Ha: \pi \neq \pi_o$$

$$(2)\ Ho: \pi \geq \pi_o;\ Ha: \pi < \pi_o$$

$$(3)\ Ho: \pi \leq \pi_o;\ Ha: \pi > \pi_o$$

ในการทดสอบสมมติฐานเหล่านี้ก็ต้องอาศัยตัวอย่างขนาดหนึ่ง (n) จากประชากรที่สนใจ ให้ X เป็นจำนวนหน่วยในตัวอย่างที่มีลักษณะตามที่สนใจ หรือจำนวนหน่วยที่เป็น S แล้ว X จะเป็นตัวสถิติทดสอบสำหรับสมมติฐานหลักนั้น

ถ้าสมมติฐานหลักเป็นจริง แล้วตัวสถิติ X จะมีการแจกแจงทวินามที่มีพารามิเตอร์ n และ π_o ดังนั้นเกณฑ์ตัดสินใจจึงขึ้นอยู่กับการแจกแจงทวินามนี้ และกำหนดตามสมมติฐานรอง Ha ดังนี้

(1) สำหรับ $Ha: \pi < \pi_o$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $p < \alpha$ ในเมื่อ p กำหนดไว้ว่า

$$p = P(X \leq x / Ho \text{ จริง}) = \sum_{X \leq x} \binom{n}{x} \pi_o^{x} (1-\pi_o)^{n-x}$$

(2) สำหรับ $Ha: \pi > \pi_o$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $p < \alpha$ ในเมื่อ p กำหนดไว้ว่า

$$p = P(X \geq x / Ho \text{ จริง}) = \sum_{X \geq x} \binom{n}{x} \pi_o^{x} (1-\pi_o)^{n-x}$$

(3) สำหรับ $Ha: \pi = \pi_o$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $p < \alpha/2$ ในเมื่อ p กำหนดไว้ว่า

เมื่อ $x < n\pi_o$ แล้ว $p = P(X \leq x / Ho \text{ จริง})$

เมื่อ $x > n\pi_o$ แล้ว $p = P(X \geq x / Ho \text{ จริง})$

เนื่องจากการใช้ตารางทวินามไม่ค่อยสะดวก เราสามารถหลีกเลี่ยงได้โดยอาศัยการแจกแจงแบบเอฟ และวิธีนี้จะใช้ได้กับทุกขนาดตัวอย่าง n เกณฑ์ตัดสินใจจะกำหนดไว้ดังนี้

(1) สำหรับ $Ha: \pi < \pi_o$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $Z > F_\alpha(\nu_1, \nu_2)$ ในเมื่อ $\nu_1 = 2(x+1), \nu_2 = 2(n-x)$ และ Z กำหนดไว้ว่า

$$Z = \{(n-x)/(x+1)\}\{\pi_0/(1-\pi_0)\}$$

(2) สำหรับ Ha : $\pi > \pi_0$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $Y > F_{\alpha}(\nu_1, \nu_2)$ ในเมื่อ $\nu_1 = 2(n+1-x); \nu_2 = 2x$ และ Y กำหนดไว้ว่า

$$Y = \{x/(n+1-x)\}(1-\pi_0)/\pi_0$$

(3) สำหรับ Ha : $\pi \neq \pi_0$ จะปฏิเสธ Ho ณ ระดับนัยสำคัญ α ถ้า $Y > F_{\alpha/2}(\nu_1, \nu_2)$ หรือถ้า $Z > F_{\alpha/2}(\nu_3, \nu_4)$ ในเมื่อ $\nu_1, \nu_2, \nu_3, \nu_4$ กำหนดไว้ว่า

$$\nu_1 = 2(n+1-x), \quad \nu_2 = 2x, \quad \nu_3 = 2(x+1), \quad \nu_4 = 2(n-x)$$
$$Y = \{x/(n+1-x)\}(1-\pi_0)/\pi_0, \quad Z = \{(n-x)/(x+1)\}\pi_0/(1-\pi_0)$$

สำหรับตัวอย่างขนาดโต (n 20) เราอาศัยการแจกแจงต่อไปนี้ช่วยประมาณการแจกแจงทวินาม

ก. เมื่อ π_0 มีค่าใกล้ 0 หรือ 1 ใช้การแจกแจงปัวซอง (Poisson) ช่วยคำนวณหาค่า p

ข. เมื่อ π_0 มีค่าไม่ใกล้ 0 หรือ 1 นั่นคือ π_0 ใกล้ ๆ 0.5 จะใช้การแจกแจงปกติมาตรฐาน N(0,1) ดังนั้นตัวสถิติทดสอบจะเป็น

$$Z = (X - n\pi_0)/\sqrt{n\pi_0(1-\pi_0)} = (P-\pi_0)/\sqrt{\pi_0(1-\pi_0)/n} \; ; \; P = X/n$$

และเกณฑ์ตัดสินใจ ณ ระดับนัยสำคัญ α จะกำหนดไว้ดังนี้

(1) สำหรับ Ha : $\pi < \pi_0$ จะปฏิเสธ Ho ถ้า $Z < -Z_{\alpha}$

(2) สำหรับ Ha : $\pi > \pi_0$ จะปฏิเสธ Ho ถ้า $Z > Z_{\alpha}$

(3) สำหรับ Ha : $\pi \neq \pi_0$ จะปฏิเสธ Ho ถ้า $Z < -Z_{\alpha/2}$ หรือ $Z > Z_{\alpha/2}$

ในเมื่อ Z_{α} และ $Z_{\alpha/2}$ เป็นค่าคงที่จากตารางปกติมาตรฐานที่ทำให้เกิดพื้นที่ทางขวามือเป็น α และ $\alpha/2$ ตามลำดับ

ตัวอย่าง นักโบราณคดีได้ขุดพบโครงกระดูก ของมนุษย์ที่บ้านเชียง 15 ราย ปรากฏว่าเป็นเพศชาย 4 ราย มีเหตุผลเพียงพอที่เชื่อได้ไหมว่าในสมัยโบราณนั้นจำนวนผู้หญิงมากกว่าผู้ชาย?

สมมติฐาน Ho : $\pi = 0.5$, Ha : $\pi > 0.5$ เมื่อ π เป็นสัดส่วนของจำนวนผู้หญิง

ระดับนัยสำคัญ $\alpha = .05$ และขนาดตัวอย่าง $n = 15$

ตัวสถิติทดสอบ คือตัวสถิติ X ซึ่งเป็นจำนวนเพศหญิงในตัวอย่าง เกณฑ์ตัดสินใจก็คือ ปฏิเสธ Ho ณ ระดับนัยสำคัญ 0.05 ถ้า $p = P(x \geq x/\text{Ho เป็นจริง}) < 0.05$

จากตัวอย่างเราได้ X = 11 ดังนั้น $p = P(X>11/\pi = 0.5)$

$$p = \sum_{x=11}^{15} \binom{15}{x} (0.5)^x (0.5)^{15-x} = 0.059$$

การสรุปผล เนื่องจาก $p = 0.059 > \alpha = 0.05$ จึงปฏิเสธ Ho ไม่ได้ นั่นคือ ในสมัยโบราณจำนวนผู้หญิงกับผู้ชายไม่น่าจะแตกต่างกัน

เมื่ออาศัยการแจกแจงแบบเอฟ เราจะได้ว่า $\nu_1 = 2(15+1-11)=10, \nu_2 = 2(11) = 22$ และ $F_{.05}^{(10,22)} = 2.30$

เนื่องจาก $Y = \frac{11}{15+1-11} \cdot \frac{1-0.5}{0.5} = 2.20 < 2.30$ จึงปฏิเสธ Ho ไม่ได้ จะเห็นได้ว่าผลสรุปนั้นเหมือนกับการใช้การแจกแจงทวินาม

ตัวอย่าง สุ่มอาจารย์มา 100 คน และสอบถามทัศนคติเกี่ยวกับการสอนนักศึกษาโดยใช้วิทยุปรากฏว่ามีอาจารย์ 16 คน เห็นว่าการสอนโดยใช้วิทยุจะไม่ได้ผล

จงทดสอบสมมติฐานที่ว่ามีอาจารย์น้อยกว่า 20% ที่คิดว่าการสอนโดยใช้วิทยุจะไม่ได้ผล

สมมติฐาน $Ho : \pi = 0.20$; $Ha : \pi < 0.20$

ระดับนัยสำคัญ $\alpha = .05$ **ขนาดตัวอย่าง** $n = 100$

ตัวสถิติทดสอบ $Z = (P - \pi_0)/\sqrt{\pi_0(1-\pi_0)/n}$ เกณฑ์ตัดสินใจ ปฏิเสธ Ho เมื่อ $Z < Z_{.05} = -1.645$ **หรือ** $Z > Z_{.05} = 1.645$

คำนวณค่าตัวสถิติทดสอบ $P = 16/100 = 0.16$

$$Z = (0.16-0.20)/\sqrt{0.20(0.80)/100} = -1.00$$

สรุปผล เนื่องจาก $Z = -1.00$ ตกอยู่ในเขตยอมรับ Ho จึงปฏิเสธ Ho ไม่ได้ นั่นคือคำกล่าวที่ว่า "มีอาจารย์น้อยกว่า 20 % ที่คิดว่าการสอนโดยใช้วิทยุจะไม่ได้ผล" นั้นไม่น่าจะเป็นจริง

เมื่ออาศัยการแจกแจงแบบเอฟ เราจะได้ว่า $\nu_1 = 2(16+1) = 34, \nu_2 = 2(100-16) = 168$ ดังนั้น $F_{05}^{(34,168)} = 1.46$

เนื่องจาก $Z = \frac{100-16}{16+1} \times \frac{0.2}{1-0.2} = 1.235 < 1.46$ จึงปฏิเสธ Ho ไม่ได้จะเห็นได้ว่าผลสรุปจะเป็นเช่นเดียวกับการใช้การแจกแจงปกติมาตรฐาน

6.5.2 ทดสอบความแตกต่างของสองสัดส่วนหรือเปอร์เซนต์ที่เป็นอิสระกัน ถ้าประชากรทวินามทั้งสองมีสัดส่วนของหน่วยที่สนใจเป็น π_1 และ π_2 เมื่อเราต้องการทดสอบความแตก ต่างของมัน นั่นคือทดสอบสมมติฐานที่ว่า

(1) $Ho: \pi_1 \geq \pi_2$; $Ha: \pi_1 < \pi_2$

(2) $Ha: \pi_1 \leq \pi_2$; $Ha: \pi_1 > \pi_2$

(3) $Ha: \pi_1 = \pi_2$; $Ha: \pi_1 \neq \pi_2$

ในการทดสอบสมมติฐานเหล่านี้ก็อาศัยตัวอย่างสุ่มที่เป็นอิสระจากประชากรทั้งสองโดยมีขนาดตัวอย่างเป็น n_1 และ n_2 ตามลำดับ ให้ X_1 และ X_2 เป็นจำนวนหน่วยที่เป็น S แล้วตัวสถิติทดสอบจะกำหนดได้ดังนี้

เมื่อตัวอย่างขนาดเล็ก เราสามารถสร้างตารางจรณ์ขนาด 2 x 2 และใช้ตัวสถิติไคสแควร์ หรืออาจจะใช้การแจกแจงไฮเปอร์จีโอเมทริค ดังนี้ "ถ้า X_1 และ X_2 มีการแจกแจงทวินามที่มีพารามิเตอร์ n_1, π_1 และ n_2, π_2 และ X_1 เป็นอิสระกับ X_2 เมื่อ $\pi_1 = \pi_2$ แล้วการแจกแจงเงื่อนไขของ X_1 โดยกำหนด $X_1 + X_2$ จะเป็นแบบไฮเปอร์จีโอเมทริค มีพารามิเตอร์ $N = n_1 + n_2$, $n = x_1 + x_2$, $r = n_1$" นั่นคือ

$$f(x_1 \mid (x_1+x_2)) = \binom{n_1}{x_1}\binom{n_1+n_2-n_1}{x_1+x_2-x_1} \Big/ \binom{n_1+n_2}{x_1+x_2}$$

$$= \binom{n_1}{x_1}\binom{n_2}{x_2} \Big/ \binom{n_1+n_2}{x_1+x_2}$$

เกณฑ์ตัดสินใจของการทดสอบ $Ho: \pi_1 = \pi_2$ ณ ระดับนัยสำคัญ α จะกำหนดไว้ตามสมมติฐานรอง Ha ดังนี้

(1) สำหรับ $Ha: \pi_1 < \pi_2$ จะปฏิเสธ Ho ถ้า $p < \alpha$ ในเมื่อ p กำหนดว่า

$$p = p(x_1 \leq x_1 / (x_1 + x_2))$$

(2) สำหรับ $Ha: \pi_1 > \pi_2$ จะปฏิเสธ Ho ถ้า $p < \alpha$ ในเมื่อ p กำหนดไว้ว่า

$$p = p(x_1 \geq x_1 / (x_1 + x_2))$$

(3) สำหรับ $Ha: \pi_1 \neq \pi_2$ จะปฏิเสธ Ho ถ้า $p_1 < \alpha/2$ หรือ $p_2 < \alpha/2$ ในเมื่อ

$$p_1 = P(x_1 \leq x_1 / (x_1 + x_2)) \text{ และ } p_2 = P(x_1 \geq x_1 / (x_1 + x_2))$$

กรณีตัวอย่างขนาดโต เราใช้ตัวสถิติทดสอบ

$$Z = p_1 - p_2 / \sqrt{pQ(1/n_1 + 1/n_2)} \quad ; \quad p = (x_1+x_2)/(n_1+n_2), \quad Q = 1-p$$

ซึ่งมีการแจกแจงปกติมาตรฐาน

ถ้าสมมติฐานหลักกำหนดไว้เป็น

$$Ho: \pi_1 - \pi_2 = \Delta$$