บทที่ 10

การทดสอบสมมติฐาน (Test Hypothesis)

10.1 **สมมติฐานทางสถิติ** (Statistical hypothesis) **นั้นหมายถึงคำกล่าวหรือข้อเสนอที่เกี่ยวกับลักษณะของประชากรที่ต้องการจะศึกษาหรือสำรวจ โดยนำเอาข้อมูลที่ได้จากตัวอย่างที่สุ่มมาจากประชากรมาใช้ในการพิสูจน์คำกล่าว หรือข้อเสนอนั้น ๆ**

วิธีการทดสอบสมมติฐานทางสถิตินั้นทำได้โดยการที่เราตั้งสมมติฐานขึ้นมาแล้วพยายามหาข้อเท็จจริงที่เรารวบรวมได้จากตัวอย่างมาเป็นเครื่องมือช่วยในการตัดสินใจว่าจะยอมรับหรือปฏิเสธสมมติฐานที่ตั้งขึ้น

ดังนั้นในการทดสอบสมมติฐานเราจะมีสมมติฐานอยู่ 2 สมมติฐานด้วยกัน คือ

1. **สมมติฐานหลัก** (Null hypothesis) ใช้สัญลักษณ์แทนด้วย H_0 เป็นสมมติฐานที่กำหนดขึ้นในลักษณะที่เราต้องการจะไม่ยอมรับ ซึ่งมักจะเป็นข้อความที่ตรงข้ามกับคำกล่าวอ้างของผู้ที่ทำการทดสอบหรือผู้กล่าวข้อความ

2. **สมมติฐานรอง** (Alternative hypothesis) ใช้สัญลักษณ์แทนด้วย H_a เป็นสมมติฐานที่มักจะเป็นข้อความตามคำกล่าวอ้างของผู้ที่ทำการทดสอบ หรือผู้กล่าวข้อความ

ตัวอย่างที่ 1

ในการผลิตของอย่างหนึ่งจะมีของเสียเป็นสัดส่วน 20% หลังจากปรับปรุงกระบวนการผลิตแล้วได้สุ่มตัวอย่างมาจำนวนหนึ่งแล้วนับจำนวนของเสีย ต้องการจะทดสอบดูว่าการปรับปรุงกระบวนการผลิตจะลดจำนวนของเสียลงหรือไม่

ดังนั้น สมมติฐานที่เราตั้งขึ้น คือ

H_o : $\pi = 0.20$

H_a : $\pi < 0.20$

ตัวอย่างที่ 2

โรงพิมพ์แห่งหนึ่งเชื่อว่าแท่นพิมพ์ขนาดใหญ่มีอายุการใช้งานเฉลี่ยเครื่องละ 13,000 และ $\sigma = 2{,}000$ ชั่วโมง จากตัวอย่างสุ่มที่สุ่มมาจำนวนหนึ่งจะสรุปได้ไหมว่าอายุการใช้งานน้อยกว่า 13,000 ชั่วโมง

$H_o : \mu = 13{,}000$ ชั่วโมง

$H_a : \mu < 13{,}000$ ชั่วโมง

ตัวอย่างที่ 3

ธนาคารแห่งหนึ่งมีลูกหนี้ 8,000 ราย เมื่อสุ่มมา 30% พบว่าเป็นข้าราชการ 37% แต่จากรายงานของธนาคารเมื่อ 5 ปีก่อนมีลูกหนี้ที่เป็นข้าราชการ 32% จงทดสอบว่าอัตราการกู้ยืมของข้าราชการเปลี่ยนแปลงไปจากเดิมหรือไม่

$H_o : \pi = 0.32$ (ไม่เปลี่ยนแปลงไปจากเดิม)

$H_a : \pi \neq 0.32$ (เปลี่ยนแปลงไปจากเดิม)

ตัวอย่างที่ 4

พ่อค้าขายลำไยตั้งเกณฑ์ในการรับซื้อลำไยจากชาวสวนว่าจะรับซื้อลำไย ถ้ามีลำไยเสียปนอยู่ไม่เกิน 10% ดังนั้น ในการทดสอบสมมติฐานเราจะตั้งสมมติฐานว่า

$H_o : \pi = 0.10$

$H_a : \pi > 0.10$

ตัวอย่างที่ 5

โรงงานผลิตยาได้ทดสอบคุณภาพของยาชนิดที่ 1 และชนิดที่ 2 โดยให้คนไข้กินยาชนิดที่ 1 และคนไข้อีกกลุ่มหนึ่งให้กินยาชนิดที่ 2 แล้วดูจำนวนผู้ที่หายจากโรคจากคนไข้ทั้ง 2 กลุ่ม ต้องการทดสอบดูว่ายาชนิดที่ 2 จะให้ผลในการรักษาโรคได้ดีกว่ายาชนิดที่ 1 หรือไม่

$H_o : \pi_1 = \pi_2$ (ยาทั้ง 2 ชนิดให้ผลเหมือนกัน)

$H_a : \pi_2 > \pi_1$ (ยาชนิดที่ 2 ให้ผลดีกว่ายาชนิดที่ 1)

ตัวอย่างที่ 6

ในการสำรวจเพศของลูกฝาแฝดจำนวน 300 คู่ ปรากฏว่ามี 84 คู่ที่เป็นชายทั้ง 2 คน มี 126 คู่ที่เป็นชาย 1 คน และหญิง 1 คน และอีก 90 คู่ เป็นหญิงทั้ง 2 คน ต้องการทดสอบดูว่าจำนวนคู่แฝดจะอยู่ในอัตราส่วน ชช : ชญ : ญญ = 1 : 2 : 1 หรือไม่

H_o : อัตราส่วนของคู่แฝด ชช : ชญ : ญญ เป็น 1 : 2 : 1

H_a : อัตราส่วนของคู่แฝด ชช : ชญ : ญญ ไม่เป็น 1 : 2 : 1

ตัวอย่างที่ 7

จากอายุการใช้งาน (ปี) ของแบตเตอรี่รถยนต์ จากโรงงานแห่งหนึ่งจำนวน 40 ตัว ได้ค่าสังเกตต่าง ๆ ต้องการทดสอบดูว่าอายุการใช้งานของแบตเตอรี่รถยนต์ที่ผลิตจากโรงงานนี้มีการกระจายเป็นแบบปกติ (Normal distribution)

H_o : อายุการใช้งานของแบตเตอรี่รถยนต์จากโรงงานนี้มีการแจกแจงแบบปกติ

H_a : อายุการใช้งานของแบตเตอรี่รถยนต์จากโรงงานนี้ไม่มีการแจกแจงแบบปกติ

จากตัวอย่างทั้ง 7 ข้างต้นเราได้ข้อสังเกต ดังนี้

1. สมมติฐานที่ตั้งขึ้นจะต้องเกี่ยวกับลักษณะของประชากรหรือพารามิเตอร์ของประชากร เช่น ค่าเฉลี่ยของประชากร (μ) ความแปรปรวนของประชากร (σ^2) สัดส่วนของประชากร (π) เป็นต้น

นอกจากนี้สมมติฐานที่ตั้งขึ้นอาจจะเกี่ยวกับการแจกแจงของประชากร เช่น ตัวอย่างที่ 6 และ 7 เป็นต้น หรือทั้งการแจกแจงและพารามิเตอร์ของประชากร

2. การตั้งสมมติฐานหลัก (H_o) นั้น ข้อความหรือคำกล่าว H_o มักจะมีคำว่า "เท่ากับ" รวมอยู่ด้วยเสมอ ส่วนสมมติฐานรอง (H_a) ข้อความหรือคำกล่าวใน H_a มักจะเป็นคำกล่าวของผู้ที่ทำการทดสอบ หรือผู้กล่าวข้อความ นอกเสียจากว่าถ้าในคำกล่าวนั้นมีคำว่า "เท่ากับ" รวมอยู่ด้วย คำกล่าวนั้นเราจะนำไปตั้งในสมมติฐานหลัก (H_o) แทน

3. สำหรับสมมติฐานรอง (H_a) นั้นมี 2 แบบด้วยกัน คือ

ก. สมมติฐานรองทางเดียว (One sided alternatives)

ถ้าให้ θ เป็นพารามิเตอร์ของประชากร

θ_0 เป็นค่าที่ระบุไว้ของพารามิเตอร์ θ

H_a ที่ตั้งขึ้นอาจจะเป็นรูปใดรูปหนึ่งใน 2 แบบ ดังนี้

H_a : $\theta > \theta_0$ หรือ

H_a : $\theta < \theta_0$

ดังตัวอย่างที่ 1, 2, 4, 5 เป็นต้น

ข. สมมติฐานรองสองทาง (Two-sided alternatives)

ถ้าให้ θ เป็นพารามิเตอร์ของประชากร

θ_0 เป็นค่าที่ระบุไว้ของพารามิเตอร์ θ

H_a ที่ตั้งขึ้นจะมีลักษณะดังนี้ คือ

H_a : $\theta \neq \theta_0$ ซึ่งหมายความว่า θ มากกว่า θ_0 และ θ น้อยกว่า θ_0

ดังตัวอย่างที่ 3 เป็นต้น

10.2 ความผิดพลาดประเภทที่ 1 และความผิดพลาดประเภทที่ 2 (Type I error and Type II error)

ในการทดสอบสมมติฐานนั้น เราต้องอาศัยตัวอย่างที่สุ่มจากประชากรมาใช้ในการตัดสินใจว่าจะยอมรับ หรือปฏิเสธสมมติฐานหลัก (H_0) ที่เราตั้งขึ้นมา ดังนั้นการตัดสินใจอาจจะมีการผิดพลาดได้ ซึ่งความผิดพลาดนี้เราแบ่งได้เป็น 2 ประเภท คือ

1. **ความผิดพลาดประเภทที่ 1** (Type I error) **หรือความคลาดเคลื่อนแบบ 1** เป็นความผิดพลาดที่เกิดจากการที่เราปฏิเสธสมมติฐานหลัก (H_0) ทั้ง ๆ ที่สมมติฐานหลัก (H_0) นั้นเป็นจริง (ถูกต้อง) ซึ่งเป็นการตัดสินใจที่ผิดพลาด และความน่าจะเป็นหรือโอกาสที่จะเกิดความผิดพลาดประเภทนี้เราเรียกว่า การเสี่ยงแบบ 1 (Alpha risk) หรือระดับนัยสำคัญ (Level of significance) ซึ่งเรามักจะใช้สัญลักษณ์แทนด้วย α

$$\therefore \alpha = P[\text{ เกิดความผิดพลาดประเภทที่ 1 }]$$

$$= P[\text{ ปฏิเสธ } H_0 \mid H_0 \text{ จริง }]$$

2. ความผิดพลาดประเภทที่ 2 (Type II error) **หรือความคลาดเคลื่อนแบบ 2 เป็นความผิดพลาดที่เกิดจากการที่เรายอมรับสมมติฐานหลัก (H_0) ทั้ง ๆ ที่สมมติฐานหลัก (H_0) นั้นไม่เป็นจริง (ไม่ถูกต้อง) ซึ่งเป็นการตัดสินใจที่ผิดพลาด และความน่าจะเป็น**หรือโอกาสที่จะเกิดความผิดพลาดประเภทนี้เราเรียกว่า การเสี่ยงแบบ 2 (Beta risk) ซึ่งเขียนสัญลักษณ์แทนด้วย β

$\therefore \beta$ = P [เกิดความผิดพลาดประเภทที่ 2]

= P [ยอมรับ H_0 | H_0 ไม่จริง]

เราสามารถสรุปความผิดพลาดทั้ง 2 ชนิด ในการทดสอบสมมติฐานได้ดังตารางต่อไปนี้

ผลสรุปในการทดสอบ (การตัดสินใจ)	เหตุการณ์จริงที่เราไม่ทราบ	
	H_0 จริง	H_0 ไม่จริง
ยอมรับ H_0	✓	✗ เกิด Type II error
ปฏิเสธ H_0	✗ เกิด Type I error	✓

ในการทดสอบสมมติฐานนั้น เราต้องการให้โอกาสที่จะเกิดความผิดพลาดทั้ง 2 ชนิดนี้มีค่าน้อย ๆ แต่เนื่องจากเราไม่สามารถทำให้ α และ β มีค่าน้อยพร้อมกันได้ ในทางปฏิบัติเราจึงกำหนดค่า α ให้มีค่าน้อยแล้วพยายามหาวิธีการทดสอบที่ทำให้ β มีค่าน้อย ซึ่งโดยทั่วไปแล้วเรามักจะกำหนดให้ $\alpha = .05$ หรือ $\alpha = .01$

10.3 ลำดับขั้นตอนในการทดสอบสมมติฐาน

การทดสอบสมมติฐานไม่ว่าจะเป็นการทดสอบเกี่ยวกับพารามิเตอร์ตัวใดก็ตามจะมีลำดับขั้นตอนในการทดสอบเหมือนกันหมด ซึ่งมีทั้งหมด 6 ขั้นตอนด้วยกันดังนี้

1. **ตั้งสมมติฐาน** (Hypotheses Formulation) สมมติฐานที่จะตั้งขึ้นนั้นมีสมมติฐานหลัก (H_0) และสมมติฐานรอง (H_a) ซึ่งมี 3 แบบด้วยกันที่จะเป็นไปได้ ดังนี้

ถ้าให้ θ เป็นพารามิเตอร์ของประชากร

θ_0 เป็นค่าที่ระบุไว้ของพารามิเตอร์ θ

แบบที่ 1

$H_0 : \theta = \theta_0$ (รวมค่า $\theta < \theta_0$ ไว้ด้วย)

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

แบบที่ 2

$H_0 : \theta = \theta_0$ (รวมค่าที่ $\theta > \theta_0$ ไว้ด้วย)

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

แบบที่ 3

$H_0 : \theta = \theta_0$

$H_a : \theta \neq \theta_0$ (เป็นสมมติฐานรองแบบ 2 ทาง)

2. กำหนดระดับนัยสำคัญ (Level of significance หรือ α) และขนาดตัวอย่าง (Sample size หรือ n)

3. กำหนดตัวสถิติที่ใช้ทดสอบ (Test statistic) โดยที่ถ้า H_0 จริงแล้ว เราต้องทราบการแจกแจงของการสุ่มตัวอย่าง (Sampling distribution) ของตัวสถิตินั้น

4. กำหนดเขตวิกฤต หรือเขตปฏิเสธ H_0 (Critical region = CR.) โดยดูจาก

ก. สมมติฐานรองที่ตั้งขึ้นว่าเป็นแบบทางเดียวหรือ 2 ทางเพื่อจะได้ทราบทิศทางของเขตปฏิเสธ H_0 ว่าจะอยู่ทางด้านไหน

ข. ระดับนัยสำคัญ (α) เพื่อที่จะได้ทราบขนาดของเขตปฏิเสธ H_0

ค. การแจกแจงของตัวสถิติที่ใช้ทดสอบ เพื่อจะได้ทราบจุดแบ่งเขตปฏิเสธ H_0 และเขตยอมรับ H_0

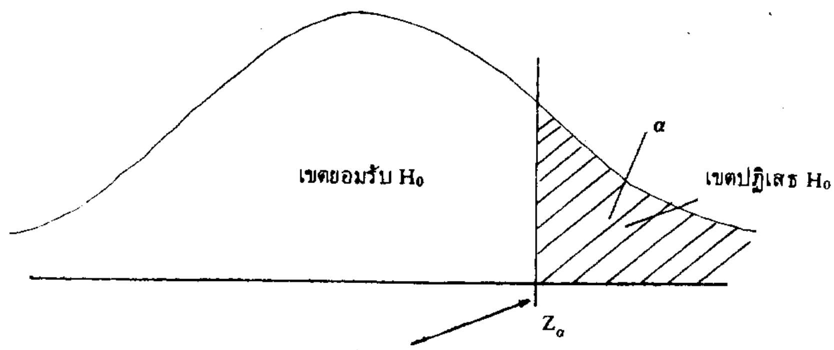
ซึ่งเขตวิกฤตมีแบบต่าง ๆ ตาม H_a ดังนี้

แบบที่ 1

$H_0 : \theta = \theta_0$

$H_a : \theta > \theta_0$

สมมติว่าการแจกแจงของตัวสถิติที่ใช้ทดสอบ คือ การแจกแจงแบบปกติ

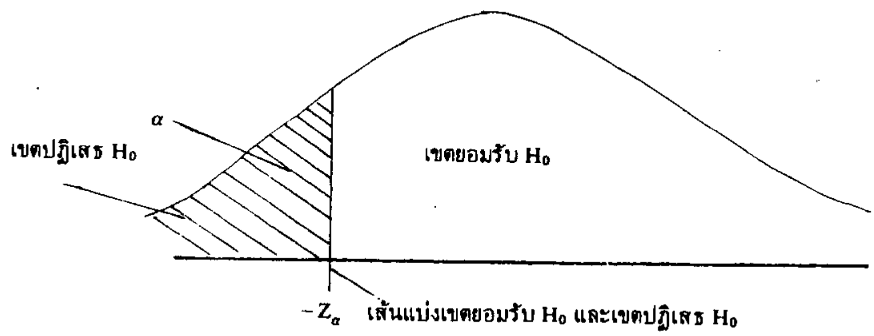


เส้นแบ่งเขตยอมรับ H_0 และเขตปฏิเสธ H_0 ซึ่งจะมีค่าเท่าใดนั้น
เราเปิดจากตาราง Z หาค่า Z_α ว่าจะมีค่าเท่าใด

แบบที่ 2

$H_0 : \theta = \theta_0$

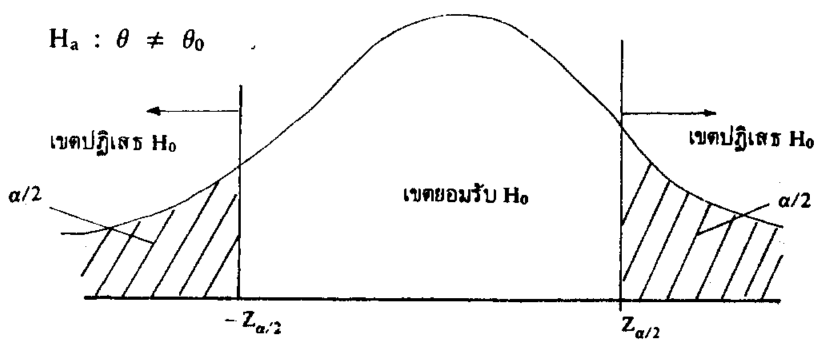
$H_a : \theta < \theta_0$



แบบที่ 3

$H_0 : \theta = \theta_0$

$H_a : \theta \neq \theta_0$



5. ทำการสุ่มตัวอย่างแล้วคำนวณค่าของตัวสถิติที่ใช้ทดสอบโดยใช้ข้อมูลจากตัวอย่างที่สุ่มได้

6. สรุปผล ถ้าค่าที่คำนวณได้ในขั้นที่ 5 ตกอยู่ในเขตวิกฤต (CR.) เราจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ $= \alpha$ แต่ถ้าค่าที่คำนวณได้ในขั้นที่ 5 ตกอยู่นอกเขตวิกฤต (CR.) เราจะยอมรับ H_0 ที่ระดับนัยสำคัญ $= \alpha$

สิ่งที่สำคัญที่สุดในการสรุปผลนั้น จะต้องสรุปผลให้สอดคล้องกับคำกล่าวของผู้ทำการทดสอบ หรือสมมติฐานนั้น ๆ ด้วยเสมอ ไม่ใช่บอกแต่เพียงว่าปฏิเสธ H_0 หรือยอมรับ H_0

10.4 การทดสอบเกี่ยวกับค่าเฉลี่ยของประชากร (μ) ที่มีการแจกแจงแบบปกติ

เป็นการทดสอบสมมติฐานที่กำหนดว่าค่าเฉลี่ยของประชากร (μ) มีค่าเท่ากับค่าใดค่าหนึ่งที่กำหนดไว้ (μ_0) แล้วนำข้อมูลที่รวบรวมได้ จากตัวอย่างมาคำนวณหาค่าสถิติ เพื่อใช้เป็นเครื่องมือในการตัดสินใจว่าจะยอมรับ หรือปฏิเสธสมมติฐานหลัก (H_0) ที่ตั้งไว้ ซึ่งเราแบ่งได้เป็น 2 กรณีด้วยกัน ดังนี้

กรณีที่ 1

การทดสอบค่าเฉลี่ยของประชากร (μ) เมื่อทราบว่าความแปรปรวนของประชากร (σ^2) หรือเมื่อไม่ทราบความแปรปรวนของประชากรแต่ใช้ตัวอย่างขนาดโต ($n \geqslant 30$) ขั้นตอนในการทดสอบมีดังนี้

1. **ตั้งสมมติฐาน**

$H_0 : \mu = \mu_0$

H_a : 1. $\mu > \mu_0$ หรือ

2. $\mu < \mu_0$ หรือ

3. $\mu \neq \mu_0$

2. กำหนด α (นิยมใช้ $\alpha = .05$ หรือ .01) และขนาดตัวอย่าง n

3. ตัวสถิติที่ใช้ในการทดสอบ คือ

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \quad \text{(ในกรณีที่ทราบ } \sigma^2\text{)}$$

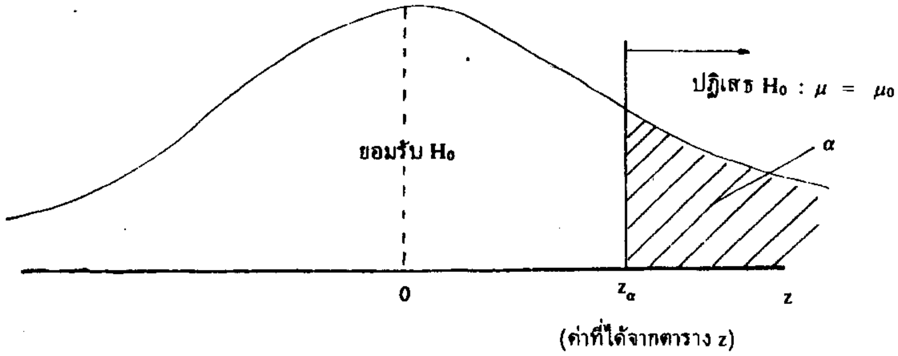
หรือ

$$Z = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$ (ในกรณีที่ไม่ทราบ σ^2 แต่ $n \geq 30$)

4. เขตวิกฤต (CR.) โดยดูจาก H_a แบบต่าง ๆ กันทั้ง 3 แบบ ดังนี้

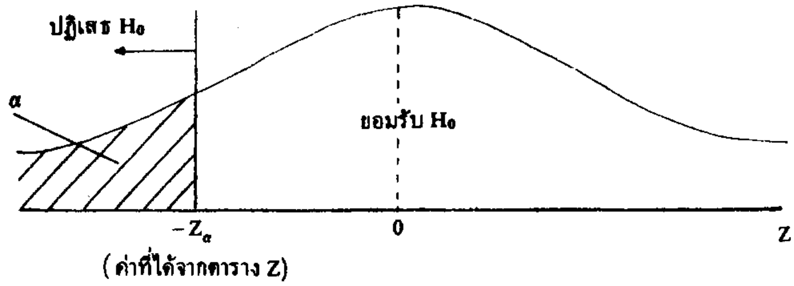
1. ถ้า $H_a : \mu > \mu_0$

CR. : ปฏิเสธ H_0 ถ้า $Z_c > Z_\alpha$ ซึ่งเขียนรูปแสดงได้ ดังนี้



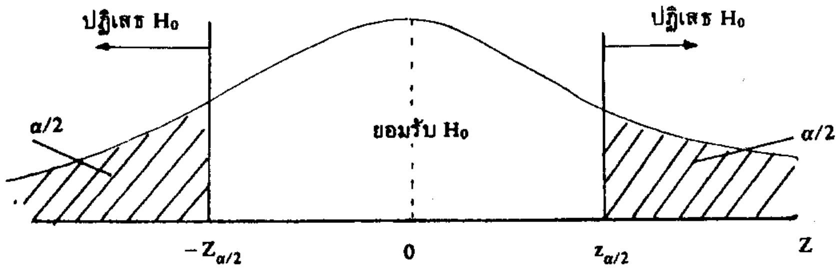
2. ถ้า $H_a : \mu < \mu_0$

CR : ปฏิเสธ H_0 ถ้า $Z_c < -Z_\alpha$ ดังรูป



3. ถ้า $H_a : \mu \neq \mu_0$

CR : ปฏิเสธ H_0 ถ้า $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ ดังรูป



5. จากตัวอย่างที่สุ่มมาขนาด n คำนวณหาค่า Z_c จาก

$$Z_c = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \quad (\text{กรณีทราบ } \sigma^2)$$

หรือ $$Z_c = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \quad (\text{กรณีไม่ทราบ } \sigma^2 \text{ แต่ } n \geqslant 30)$$

เมื่อ $$S^2 = \frac{1}{n-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n} \right]$$

6. สรุปผล

1. ถ้า Z_c ตกอยู่ใน CR. เราจะปฏิเสธ H_0 นั่นคือ ยอมรับ H_a
2. ถ้า Z_c ตกอยู่นอก CR. เราจะไม่ปฏิเสธ H_0 นั่นคือ ยอมรับ H_0

ตัวอย่างที่ 10.4.1

โรงงานผลิตของเด็กเล่นแห่งหนึ่ง ทราบว่าคนงานคนหนึ่งสามารถผลิตของเล่นได้โดยเฉลี่ยชั่วโมงละ 15 ชิ้น และค่าเบี่ยงเบนมาตรฐาน 5 ชิ้น มีผู้เสนอให้ใช้วิธีการใหม่ซึ่งจะช่วยทำให้คนงานสามารถผลิตของเล่นเด็กได้มากขึ้น เมื่อได้ทดลองวิธีใหม่ในเวลา 100 ชั่วโมงดู ปรากฏว่าคนงานผลิตของเล่นเด็กได้โดยเฉลี่ยชั่วโมงละ 16 ชิ้น ถ้าให้ $\alpha = .01$ มีเหตุผลเพียงพอที่จะเชื่อว่าวิธีการใหม่จะช่วยให้คนงานผลิตของเล่นได้มากขึ้นหรือไม่

1. $H_0 : \mu = 15$

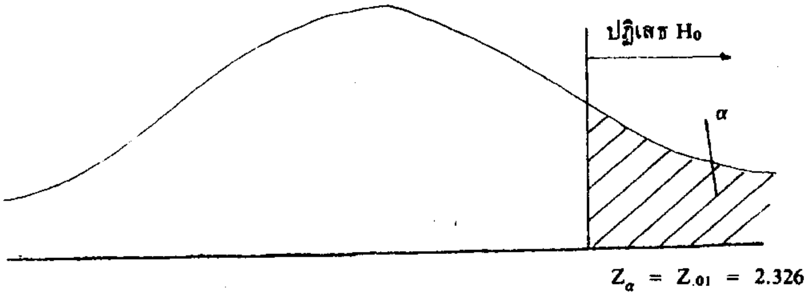
   $H_a : \mu > 15$
2. $\alpha = 0.01$ $n = 100$ ชั่วโมง
3. ตัวสถิติที่ใช้ทดสอบ คือ

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

เนื่องจากเราทราบความแปรปรวนของประชากร σ^2 ว่า

$$\sigma^2 = (5)^2 = 25$$

4. CR : ปฏิเสธ H_0 ถ้า $Z_c > Z_\alpha$



ในที่นี้โจทย์กำหนดให้ $\alpha = .01$

$\therefore$ CR : ปฏิเสธ H_0 ถ้า $Z_c > Z_{.01}$

(เปิดตาราง Z จะได้ $Z_{.01} = 2.326$)

นั่นคือ CR : ปฏิเสธ H_0 ถ้า $Z_c > 2.326$

5. คำนวณค่า Z_c จาก

$$Z_c = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$$

แทนค่า $\overline{X} = 16$

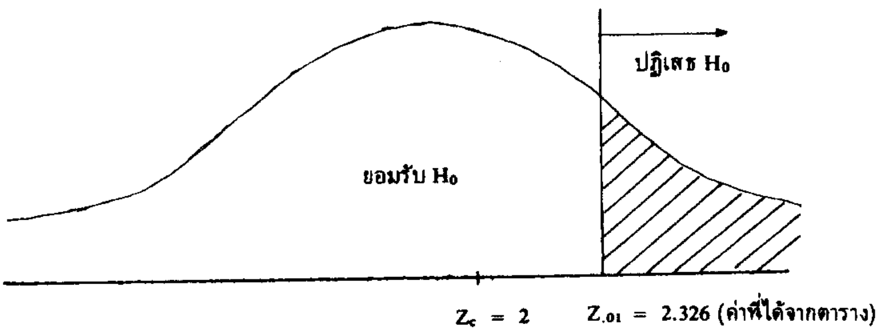
$\mu_0 = 15$

$\sigma = 5$

$n = 100$

$$\therefore \quad Z_c = \frac{16-15}{5/\sqrt{100}} = \frac{1}{\frac{5}{10}} = \frac{10}{5} = 2$$

6. สรุปผล จะเห็นได้ว่าค่า Z_c ตกอยู่นอก CR. ดังรูป



$\therefore$ เราจะไม่ปฏิเสธ H_0 นั่นคือเรายอมรับ H_0 ที่ว่า $\mu = 15$

แสดงว่าไม่มีเหตุผลเพียงพอที่จะเชื่อว่าวิธีการใหม่จะช่วยให้คนงานผลิตของเล่นได้มากขึ้น

ตัวอย่างที่ 10.4.2

บริษัทผลิตอุปกรณ์ทางด้านกีฬาตกปลาได้โฆษณาว่าสายเบ็ดที่เขาปรับปรุงใหม่ สามารถต้านทานน้ำหนักได้โดยเฉลี่ย 15 กิโลกรัม และมีส่วนเบี่ยงเบนมาตรฐาน 0.5 กิโลกรัม จงทดสอบที่ $\alpha = 0.01$ ว่าคำโฆษณานี้เป็นจริงหรือไม่ โดยสุ่มตัวอย่างสายเบ็ดมา 50 อัน เพื่อทดลองดู และพบว่ามีความต้านทานน้ำหนักโดยเฉลี่ยได้ 14.8 กิโลกรัม

1. $H_0 : \mu = 15$ กิโลกรัม

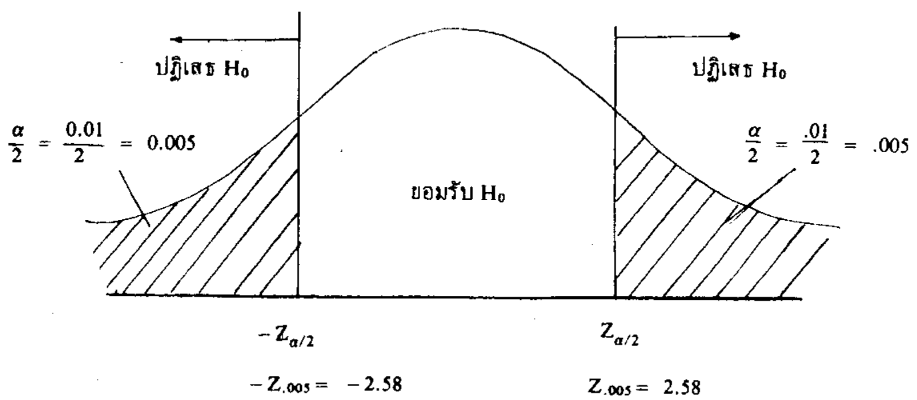
   $H_a : \mu \neq 15$ กิโลกรัม

2. $\alpha = 0.01, n = 50$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$

4.



CR : ปฏิเสธ H_0 ถ้า $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$

นั่นคือปฏิเสธ H_0 ถ้า $Z_c < -2.58$ หรือ $Z_c > 2.58$

5. $\overline{X} = 14.8, S = 0.5, \mu_0 = 15, n = 50$

จาก

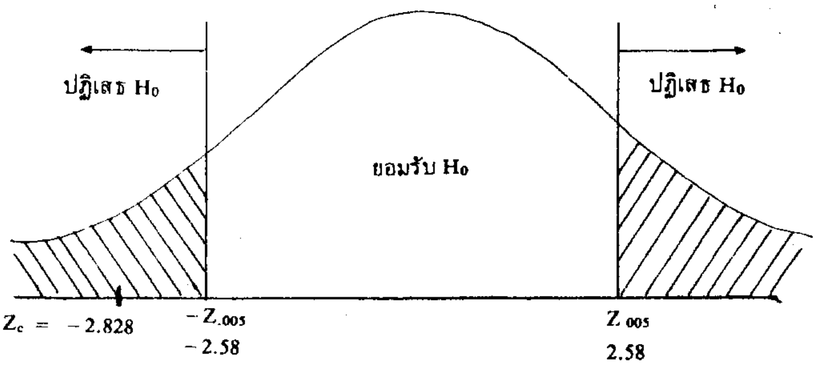
$$Z_c = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$

$$\therefore Z_c = \frac{14.8-15}{0.5/\sqrt{50}} = \frac{-0.2\times\sqrt{50}}{0.5}$$

$$= -0.4\times\sqrt{50}$$

$$= -2.828$$

6. สรุปผล



$Z_c < -2.58$

จะเห็นได้ว่า Z_c ที่คำนวณได้ตกอยู่ใน CR. $\therefore$ เราปฏิเสธ H_0 ที่ว่า $\mu = 15$ กิโลกรัม และยอมรับ $H_a : \mu \neq 15$ กิโลกรัม แสดงว่าการที่บริษัทโฆษณาว่าสายเบ็ดที่ปรับปรุงใหม่มีความต้านทานน้ำหนักโดยเฉลี่ยได้ 15 กิโลกรัมนั้นไม่เป็นจริง

กรณีที่ 2

การทดสอบค่าเฉลี่ยของประชากร (μ) เมื่อไม่ทราบค่าความแปรปรวนของประชากร (σ^2) และตัวอย่างมีขนาดเล็ก ($n < 30$) ขั้นตอนในการทดสอบมีดังนี้

1. ตั้งสมมติฐาน

$H_0 : \mu = \mu_0$

H_a : 1. $\mu > \mu_0$ หรือ

2. $\mu < \mu_0$ หรือ

3. $\mu \neq \mu_0$

2. กำหนด α และขนาดตัวอย่าง (n)

3. ตัวสถิติที่ใช้ทดสอบ คือ

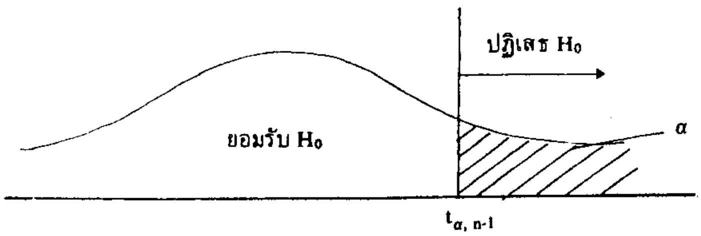
$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

df. = n − 1

4. เขตวิกฤต (CR)

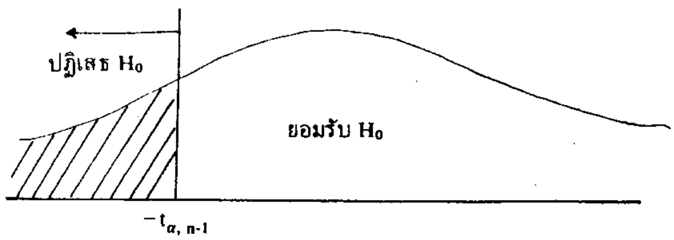
1. ถ้า $H_a : \mu > \mu_0$

CR : ปฏิเสธ H_0 ถ้า $t_c > t_{\alpha, n-1}$



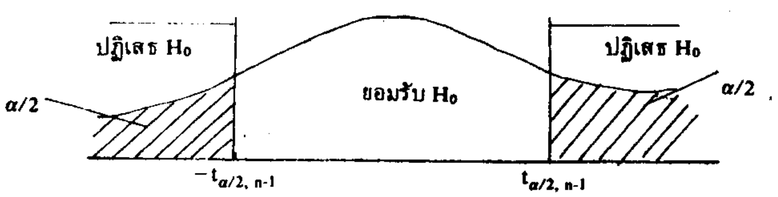
2. ถ้า $H_a : \mu < \mu_0$

CR : ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha, n-1}$



3. ถ้า $H_a : \mu \neq \mu_0$

CR : ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha/2, n-1}$ หรือ $t_c > t_{\alpha/2, n-1}$



5. จากตัวอย่างที่สุ่มมาขนาด n คำนวณหา t_c จากสูตร

$$t_c = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

เมื่อ $S^2 = \frac{1}{n-1}\left[\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}\right]$

6. สรุปผล

1. ถ้า t_c ตกอยู่ใน CR. เราจะปฏิเสธ H_0 นั่นคือยอมรับ H_a
2. ถ้า t_c ตกอยู่นอก CR. เราจะยอมรับ H_0

ตัวอย่างที่ 10.4.3

ในการลงทะเบียนเรียนของมหาวิทยาลัยแห่งหนึ่งพบว่าโดยเฉลี่ยแล้วนักศึกษาจะใช้เวลาคนละ 50 นาที เจ้าหน้าที่แผนกลงทะเบียนได้ปรับปรุงวิธีใหม่ โดยนำเครื่องคอมพิวเตอร์เข้ามาช่วย จากการสุ่มตัวอย่างนักศึกษามา 12 คน พบว่าการนำเครื่องคอมพิวเตอร์เข้ามาช่วยเขาจะใช้เวลาลงทะเบียน โดยเฉลี่ยคนละ 42 นาที และมีส่วนเบี่ยงเบนมาตรฐาน 11.9 นาที จงทดสอบว่าวิธีใหม่ทำให้การลงทะเบียนของนักศึกษาใช้เวลาน้อยกว่า 50 นาที ที่ $\alpha = 0.01$

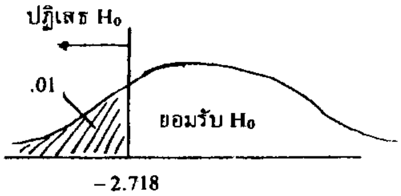
1. $H_0 : \mu = 50$ นาที

   $H_a : \mu < 50$ นาที

2. $\alpha = 0.01, n = 12$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}, \; d.f. = n - 1 = 11$$



4. CR : ปฏิเสธ H_0 ถ้า $t_c < -t_{.01,\,11}$

   เปิดตารางได้ $-t_{.01,\,11} = -2.718$

   $\therefore$ ปฏิเสธ H_0 ถ้า $t_c < -2.718$

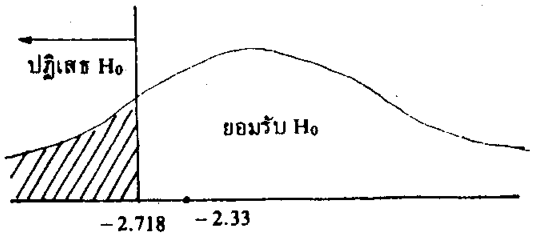
5. $\overline{X} = 42, n = 12, \mu_0 = 50, S = 11.9$

แทนค่าจะได้

$$t_c = \frac{42-50}{11.9/\sqrt{12}} = \frac{-8}{11.9}\times\sqrt{12}$$

$$= \frac{-8}{11.9}\times 3.46$$

$$= \frac{-27.68}{11.9} = -2.33$$



6. สรุปผล

$\because t_c > -2.718$

จะเห็นว่า t_c ตกอยู่นอก CR. ดังนั้นเราจึงยอมรับ H_0 ที่ว่า $\mu = 50$ นาที แสดงว่าการใช้วิธีใหม่ไม่ทำให้การลงทะเบียนของนักศึกษาใช้เวลาน้อยกว่า 50 นาที

ตัวอย่างที่ 10.4.4

ไก่พันธุ์หนึ่งมีอัตราการเพิ่มของน้ำหนักโดยเฉลี่ยเท่ากับ 65 กรัม ในระหว่างอายุ 3 เดือนแรก ได้มีการนำไก่พันธุ์นี้มาทดลองให้อาหารเฉพาะชนิดหนึ่ง 12 ตัว ตั้งแต่แรกเกิดจนกระทั่งอายุได้ 3 เดือนปรากฏว่าน้ำหนักเป็นดังนี้

54 65 55 62 58 64 60 59 62 67 61 62

จากข้อมูลที่ได้ เราพอจะเชื่อได้ไหมว่าอาหารที่ให้นั้นมีส่วนในการเปลี่ยนแปลงน้ำหนักของไก่พันธุ์นี้ ที่ $\alpha = 0.05$

1. $H_0 : \mu = 65$ กรัม

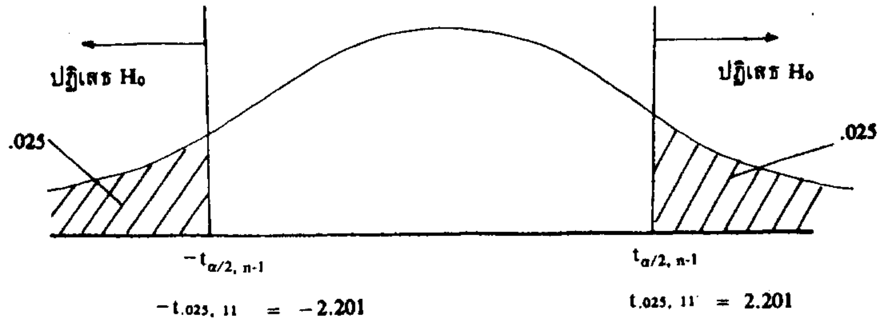
   $H_a : \mu \neq 65$ กรัม

2. $\alpha = 0.05, n = 12$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$t = \frac{\overline{X}-\mu}{S/\sqrt{n}}$$

4. CR : ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha/2,\, n-1}$ หรือ $t_c > t_{\alpha/2,\, n-1}$



$\therefore$ CR : ปฏิเสธ H_0 ถ้า $t_c < -2.201$ หรือ $t_c > 2.201$

5. $\overline{X} = \frac{\Sigma X}{n} = \frac{729}{12} = 60.75$

$$S^2 = \frac{1}{n-1}\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right]$$

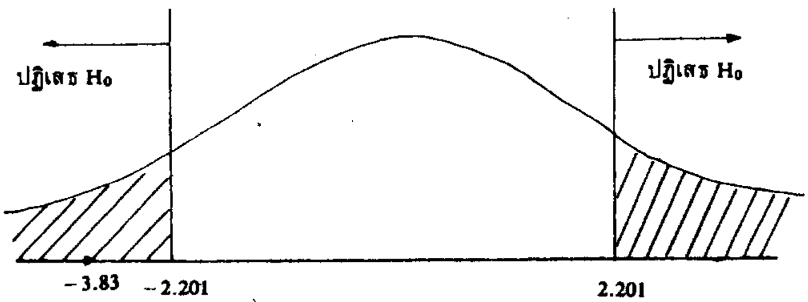
$$= \frac{1}{11}\left[44449 - \frac{(729)^2}{12}\right] = \frac{1}{11}\left[\frac{(44449)(12)-(729)^2}{12}\right]$$

$$= \frac{1}{11}\left[\frac{533388-531441}{12}\right] = \frac{1}{11}\times\frac{1947}{12} = 14.75$$

$$\therefore S = \sqrt{14.75} = 3.84$$

$$\therefore t_c = \frac{60.75-65}{3.84/\sqrt{12}} = \frac{-4.25\times 3.46}{3.84} = -3.83$$

6. สรุปผล



$\because t_c < -2.201$ ดังนั้น t_c ตกอยู่ใน CR. เราจึงปฏิเสธ H_0 ที่ว่า $\mu = 65$ กรัม แสดงว่าอาหารชนิดนั้นมีส่วนทำให้เกิดการเปลี่ยนแปลงน้ำหนักของไก่พันธุ์นี้

10.5 การทดสอบเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ย 2 ค่าจากสองประชากรที่มีการแจกแจงแบบปกติ

เมื่อต้องการเปรียบเทียบค่าเฉลี่ยของ 2 ประชากรแบบปกติว่ามีความแตกต่างกันหรือไม่ เราทำได้โดยการสุ่มตัวอย่างขนาด n_1 และ n_2 จากประชากร ทั้ง 2 ตามลำดับ ซึ่งในการทดสอบสมมติฐานนั้นเราแบ่งออกได้เป็น 2 กรณีด้วยกัน ดังนี้

กรณีที่ 1

เมื่อตัวอย่างที่สุ่มมาเป็นอิสระกัน เราแบ่งได้เป็น 3 ลักษณะด้วยกัน ดังนี้

ก. **เมื่อทราบความแปรปรวนของประชากรทั้ง 2** มีขั้นตอนในการทดสอบ ดังนี้

1. ตั้งสมมติฐาน มี 3 แบบด้วยกัน

$H_0 : \mu_1 - \mu_2 = D_0$

H_a : 1. $\mu_1 - \mu_2 > D_0$ หรือ

2. $\mu_1 - \mu_2 < D_0$ หรือ

3. $\mu_1 - \mu_2 \neq D_0$

เมื่อ D_0 = ค่าที่กำหนดให้ค่าใดค่าหนึ่ง (ซึ่งจะมีค่าเป็นเท่าใดก็ได้)

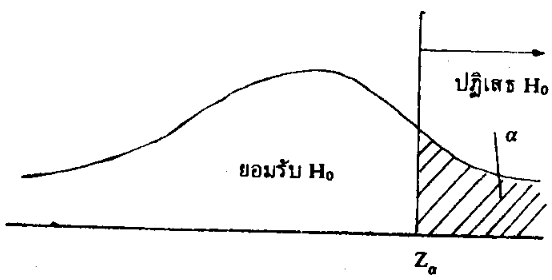
2. กำหนด α, n_1 และ n_2

3. ตัวสถิติที่ใช้ทดสอบ คือ

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

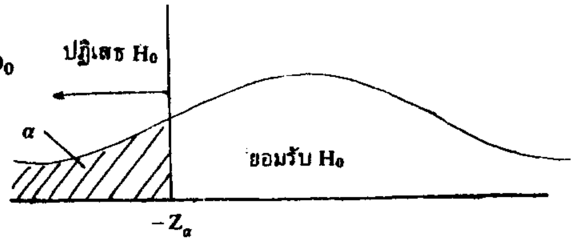
4. CR :

1. ถ้า $H_a : \mu_1 - \mu_2 > D_0$



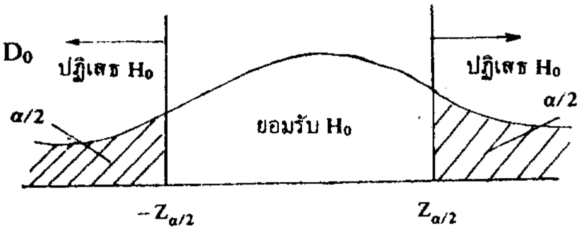
CR : ปฏิเสธ H_0 ถ้า $Z_c > Z_\alpha$

2. ถ้า $H_a : \mu_1 - \mu_2 < D_0$



CR : ปฏิเสธ H_0 ถ้า $Z_c < -Z_\alpha$

3. ถ้า $H_a : \mu_1 - \mu_2 \neq D_0$



CR : ปฏิเสธ H_0 ถ้า $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$

5. คำนวณหาค่า Z_c จาก

$$Z_c = \frac{(\bar{X}_1 - \bar{X}_2) - D_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

โดยที่ $D_0 = \mu_1 - \mu_2$

6. สรุปผล : 1. ถ้า Z_c ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า Z_c ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่างที่ 10.5.1

บริษัทผลิตรถยนต์ต้องการซื้อยางรถยนต์ ซึ่งมียางรถยนต์ให้เลือกตัดสินใจอยู่ 2 ชนิดด้วยกัน โดยบริษัทจะเลือกชนิดที่มีความทนทานมากกว่า จากการทดลองความทนทานของยางทั้ง 2 ชนิด เขาจึงสุ่มยางแต่ละชนิดมาอย่างละ 25 เส้น แล้วบันทึกระยะทางที่วิ่งได้จนเสื่อมคุณภาพได้ผลดังนี้

	ชนิดที่ 1	ชนิดที่ 2
ค่าเฉลี่ย	39,780 ไมล์	40,650 ไมล์
ส่วนเบี่ยงเบน	3,000 ไมล์	3,200 ไมล์

บริษัทจะเลือกซื้อยางชนิดใดถ้าจากประสบการณ์ทำให้ทราบว่ายางชนิดที่ 1 และชนิดที่ 2 มีส่วนเบี่ยงเบนเท่ากับ 2,400 และ 3,600 ไมล์ ตามลำดับที่ $\alpha = 0.05$

1. ตั้งสมมติฐาน

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 < 0$

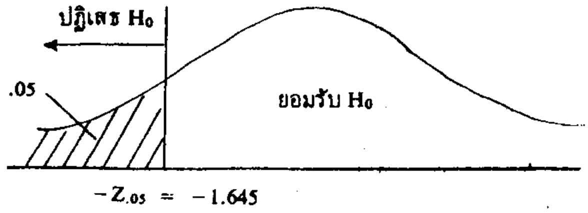
2. $\alpha = 0.05$, $n_1 = 25$, $n_2 = 25$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

4. CR : ปฏิเสธ H_0 ถ้า

$Z_c < -1.645$



5. $\overline{X}_1 = 39{,}780$, $\overline{X}_2 = 40{,}650$

$\sigma_1^2 = (2{,}400)^2$, $\sigma_2^2 = (3{,}600)^2$

$n_1 = 25$, $n_2 = 25$

หาค่า Z_c จาก

$$Z_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

ในที่นี้ $\mu_1 - \mu_2 = 0$

$$\therefore Z_c = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

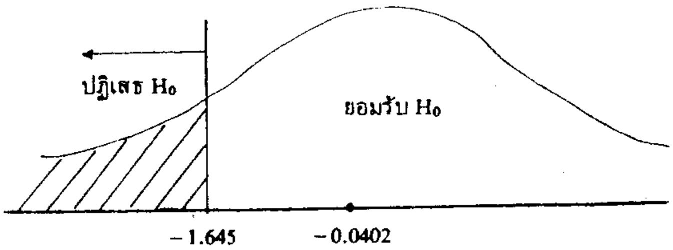
$$= \frac{(39{,}780 - 40{,}650)}{\sqrt{\frac{(2{,}400)^2}{25} + \frac{(3{,}600)^2}{25}}}$$

$$= \frac{-870}{\frac{1}{5}\sqrt{5{,}760{,}000 - 12{,}960{,}000}}$$

$$= \frac{-870}{5\sqrt{18{,}720{,}000}} = \frac{-870}{5 \times 4{,}326.6}$$

$$= \frac{-870}{21{,}633.307} = -0.0402$$

6. สรุปผล



$\therefore Z_c > -1.645$

$\therefore$ จะเห็นได้ว่า Z_c ตกอยู่นอก CR. ดังนั้นเราจึงยอมรับ H_0 ที่ว่า $\mu_1 - \mu_2 = 0$ หรือ $\mu_1 = \mu_2$ แสดงว่าบริษัทจะเลือกซื้อยางชนิดใดก็ได้

ข. เมื่อไม่ทราบความแปรปรวน แต่ทราบว่ามีความแปรปรวนเท่ากัน ($\sigma_1^2 = \sigma_2^2 = \sigma^2$) โดยที่ความแปรปรวนของประชากรเราประมาณได้จากความแปรปรวนของตัวอย่าง คือ จาก S_1^2 และ S_2^2 ดังนี้

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

เมื่อ S_p^2 คือ ความแปรปรวนร่วม

สำหรับขั้นตอนในการทดสอบมีดังนี้

1. ตั้งสมมติฐาน

$H_0 : \mu_1 - \mu_2 = D_0$

H_a : 1. $\mu_1 - \mu_2 > D_0$ หรือ

2. $\mu_1 - \mu_2 < D_0$ หรือ

3. $\mu_1 - \mu_2 \neq D_0$

2. กำหนด α และขนาดตัวอย่าง n_1, n_2

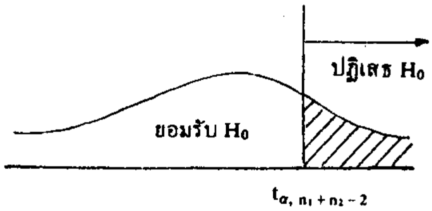
3. ตัวสถิติที่ใช้ทดสอบ คือ

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$d.f. = (n_1 - 1) - (n_2 - 1) = n_1 + n_2 - 2$$

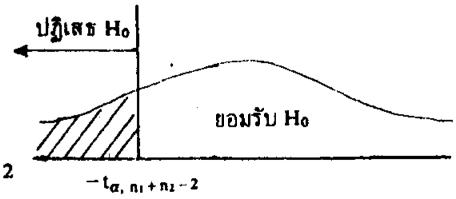
4. CR :

1. ถ้า $H_a : \mu_1 - \mu_2 > D_0$



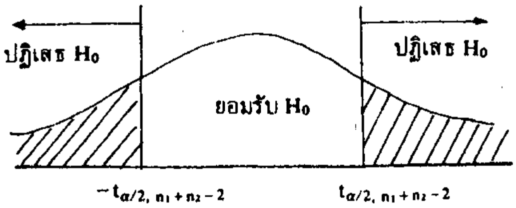
ปฏิเสธ H_0 ถ้า $t_c > t_{\alpha, n_1+n_2-2}$

2. ถ้า $H_a : \mu_1 - \mu_2 < D_0$



ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha, n_1+n_2-2}$

3. ถ้า $H_a : \mu_1 - \mu_2 \neq D_0$



ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha/2, n_1+n_2-2}$ หรือ $t_c > t_{\alpha/2, n_1+n_2-2}$

5. คำนวณค่า t_c จาก

$$t_c = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

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

6. สรุปผล : 1. ถ้า t_c ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า t_c ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่าง 10.5.2

ในการสอนวิชาคอมพิวเตอร์เพื่อเปรียบเทียบว่าการสอนแบบมีปฏิบัติการจะให้ผลดีกว่าการสอนแบบไม่มีปฏิบัติการหรือไม่ ปรากฏว่าในห้องที่ใช้สอนแบบมีปฏิบัติการมีนักศึกษาเข้าเรียน 11 คน และสอบได้คะแนนเฉลี่ย 85 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 4.7 และห้องที่สอนแบบไม่มีปฏิบัติการมีนักศึกษาเข้าเรียน 17 คน เมื่อใช้ข้อสอบเดียวกันกับกลุ่มแรก ได้คะแนนเฉลี่ย 79 คะแนน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 6.1 ถ้ากำหนดให้คะแนนของนักศึกษามีการแจกแจงใกล้โค้งปกติและมีความแปรปรวนเท่ากัน จงทดสอบว่าการสอนแบบมีปฏิบัติการจะช่วยเพิ่มคะแนนเฉลี่ยของนักศึกษาให้สูงขึ้นโดยใช้ $\alpha = .01$

1. ตั้งสมมติฐาน

ให้ μ_1 = ค่าเฉลี่ยของคะแนนจากการสอนแบบมีปฏิบัติการ

μ_+ = ค่าเฉลี่ยของคะแนนจากการสอนแบบไม่มีปฏิบัติการ

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 > 0$

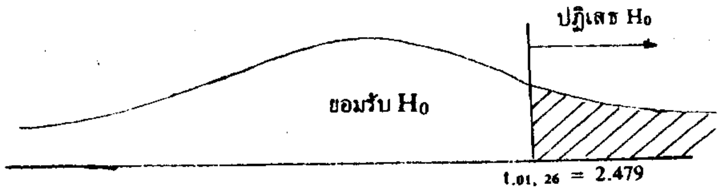
2. $\alpha = .01, n_1 = 11, n_2 = 17$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

d.f. $= n_1 + n_2 - 2 = 11 + 17 - 2 = 26$

4. CR : ปฏิเสธ H_0 ถ้า $t_c > 2.479$



5. $n_1 = 85$, $n_2 = 17$

$\overline{X}_1 = 85$, $\overline{X}_2 = 79$

$S_1 = 4.7$, $S_2 = 6.1$

หา S_p^2 จาก

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

$$= \frac{(10)\,(4.7)^2+(16)\,(6.1)^2}{11+17-2}$$

$$= \frac{(10)\,(22.09)+(16)\,(37.21)}{26}$$

$$= \frac{220.9+595.36}{26}$$

$$= \frac{816.26}{26} = 31.395$$

$$\therefore\ S_p = \sqrt{31.395} = 5.603$$

จาก $t_c = \dfrac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$

ในที่นี้ $\mu_1-\mu_2 = 0$

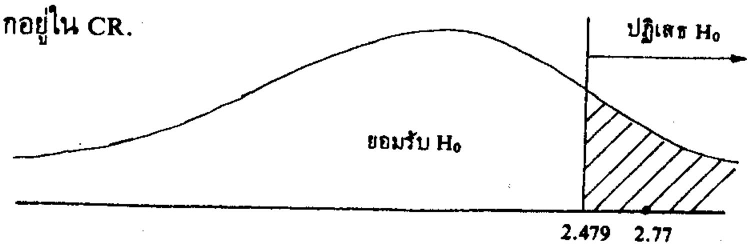
$$\therefore\ t_c = \frac{(85-79)-0}{5.603\sqrt{\dfrac{1}{11}+\dfrac{1}{17}}}$$

$$= \frac{6}{5.603\sqrt{0.15}} = \frac{6}{5.603\times .387}$$

$$= \frac{6}{2.168} = 2.77$$

6. สรุปผล

$\therefore$ t_c ตกอยู่ใน CR.



$\therefore$ เราปฏิเสธ H_0 ที่ว่า $\mu_1-\mu_2 = 0$ แสดงว่าการสอนแบบมีปฏิบัติการจะทำให้คะแนนเฉลี่ยของนักศึกษาสูงขึ้น

ค. ไม่ทราบความแปรปรวนของประชากร แต่ทราบว่ามีความแปรปรวนไม่เท่ากัน ($\sigma_1^2 \neq \sigma_2^2$) มีขั้นตอนในการทดสอบ ดังนี้

1. ตั้งสมมติฐาน

$H_0 : \mu_1 - \mu_2 = D_0$

H_a : 1. $\mu_1 - \mu_2 > D_0$ หรือ

2. $\mu_1 - \mu_2 < D_0$ หรือ

3. $\mu_1 - \mu_2 \neq D_0$

2. กำหนด α, n_1 และ n_2

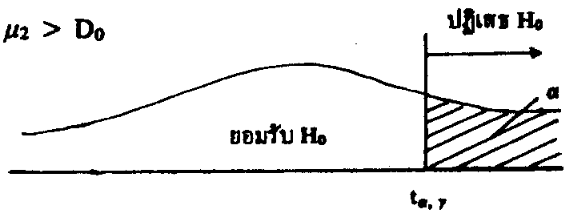
3. ตัวสถิติที่ใช้ทดสอบ คือ

$$T' = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

โดยที่ T′ จะมีการแจกแจงแบบ t ที่มี df = ν โดยที่

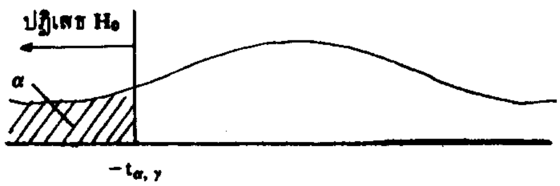
$$\nu = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_1 - 1)} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{(n_2 - 1)}}$$

4. CR : 1. ถ้า $H_a : \mu_1 - \mu_2 > D_0$



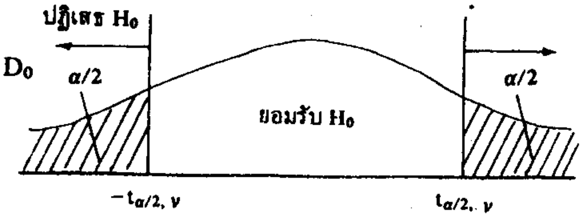
CR : จะปฏิเสธ H_0 ถ้า $t_c' > t_{\alpha, \nu}$

2. ถ้า $H_a : \mu_1 - \mu_2 < D_0$



CR : ปฏิเสธ H_0 ถ้า $t_c' < -t_{\alpha, \nu}$

3. ถ้า $H_a : \mu_1 - \mu_2 \neq D_0$



CR : ปฏิเสธ H_0 ถ้า $t_c' < -t_{\alpha/2,\,\nu}$ หรือ

$t_c' > t_{\alpha/2,\,\nu}$

5. คำนวณค่า t_c' จากสูตร

$$t_c' = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

6. สรุปผล 1. ถ้า t_c' ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า t_c' ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่างที่ 10.5.3

จากการเก็บสถิติตัวเลขเกี่ยวกับปริมาณน้ำฝนที่ตกในเดือนกันยายน บันทึกผลไว้ดังนี้ ในเขต ก. ปริมาณน้ำฝนที่ตกในช่วง 10 ปีที่ผ่านมามีจำนวนโดยเฉลี่ย 1.8 นิ้ว ส่วนเบี่ยงเบนมาตรฐาน 0.3 นิ้ว ส่วนใน เขต ข. มีปริมาณน้ำฝนที่ตกในช่วง 7 ปีที่ผ่านมามีจำนวนโดยเฉลี่ย 1.1 นิ้ว ส่วนเบี่ยงเบนมาตรฐาน 0.2 นิ้ว จากข้อมูลที่ได้จะถือว่าปริมาณน้ำฝนที่ตกใน เขต ก. มากกว่าปริมาณน้ำฝนที่ตกใน เขต ข. ได้หรือไม่ ที่ $\alpha = .01$ ถ้าถือว่าปริมาณน้ำฝนที่ตกในแต่ละเขตมีการกระจายแบบปกติที่มีความแปรปรวนไม่เท่ากัน

1. $H_0 : \mu_ก - \mu_ข = 0$

   $H_a : \mu_ก - \mu_ข > 0$

2. $\alpha = 0.01,\ n_1 = 10,\ n_2 = 7$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$T' = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

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

$$= \frac{\left[\frac{(.3)^2}{10}+\frac{(.2)^2}{7}\right]^2}{\frac{\left[\frac{(.3)^2}{10}\right]^2}{9}\ \frac{\left[\frac{(.2)^2}{7}\right]^2}{6}}$$

$$= \frac{\left(\frac{.09}{10}+\frac{.04}{7}\right)^2}{\frac{\left(\frac{.09}{10}\right)^2}{9}+\frac{\left(\frac{.04}{7}\right)^2}{6}}$$

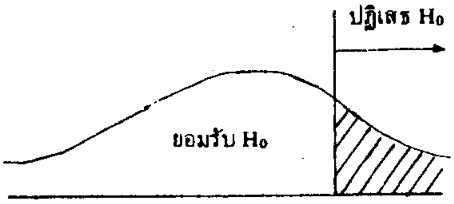
$$\doteq \frac{.000225}{.0000148}$$

$$= 15.203$$

$\therefore$ df $\approx$ 15

4. CR : ปฏิเสธ H_0 ถ้า $t_c' > t_{.01,\,15}$

คือ $t_c' > 2.602$



5. จาก $t_c' = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

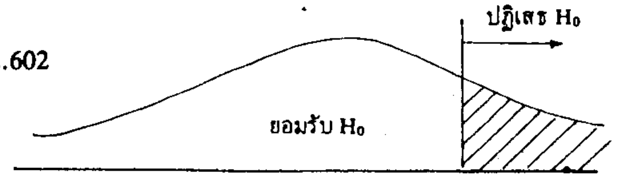
$$= \frac{(1.8-1.1)-0}{\sqrt{\frac{(.3)^2}{10}+\frac{(.2)^2}{7}}}$$

$$= \frac{0.7}{\sqrt{\frac{.09}{10}+\frac{.04}{7}}}$$

$$= \frac{0.7}{\sqrt{.009+.006}} = \frac{0.7}{\sqrt{.015}}$$

$$= \frac{0.7}{0.128} = 5.47$$

6. สรุปผล $\because t'_c > 2.602$



$\therefore t'_c$ ตกอยู่ใน CR. เราจึงปฏิเสธ H_0 ที่ว่า

$\mu_1 - \mu_2 = 0$ และยอมรับ H_a ที่ว่า

$\mu_1 - \mu_2 > 0$ แสดงว่าจากข้อมูลที่ได้เราถือว่าปริมาณน้ำฝนที่ตกใน เขต ก. มากกว่าปริมาณน้ำฝนที่ตกใน เขต ข.

กรณีที่ 2

เมื่อตัวอย่างที่สุ่มมาไม่เป็นอิสระกัน หรือกรณีที่ข้อมูลมีลักษณะจับกันเป็นคู่ ๆ (Dependent random sample or paired observations) มีขั้นตอนในการทดสอบ ดังนี้

1. ตั้งสมมติฐาน

$H_0 : \mu_1 - \mu_2 = D_0$

H_a : 1. $\mu_1 - \mu_2 > D_0$ หรือ

2. $\mu_1 - \mu_2 < D_0$ หรือ

3. $\mu_1 - \mu_2 \neq D_0$

2. กำหนด α, n

3. ตัวสถิติที่ใช้ทดสอบ คือ

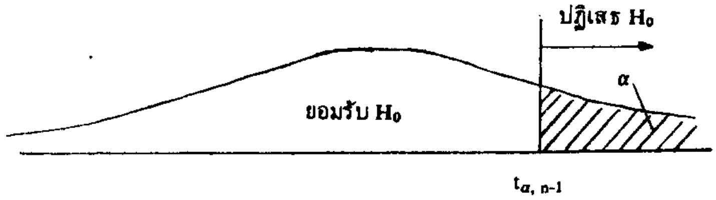
$$t = \frac{\bar{d} - D_0}{S_d/\sqrt{n}} \quad \text{มี } d.f = n-1$$

โดยที่ d_i = ผลต่างของหน่วยตัวอย่างคู่ที่ i ; i = 1, 2, ..., n

$$\bar{d} = \frac{\Sigma d_i}{n}$$

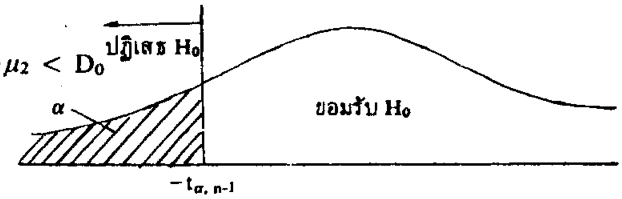
$$S_d^2 = \frac{1}{n-1}\left[\Sigma d_i^2 - \frac{(\Sigma d_i)^2}{n}\right]$$

4. CR : 1. ถ้า $H_a : \mu_1 - \mu_2 > D_0$



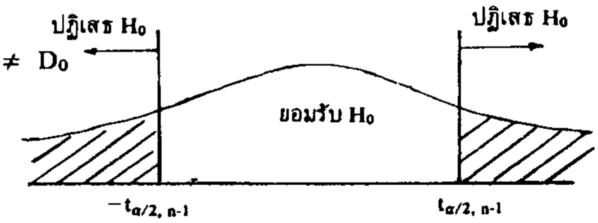
ปฏิเสธ H_0 ถ้า $t_c > t_{\alpha, n-1}$

2. ถ้า $H_a : \mu_1 - \mu_2 < D_0$



ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha, n-1}$

3. ถ้า $H_0 : \mu_1 - \mu_2 \neq D_0$



ปฏิเสธ H_0 ถ้า $t_c < -t_{\alpha/2, n-1}$ หรือ $t_c > t_{\alpha/2, n-1}$

5. คำนวณหาค่า t_c จาก

$$t_c = \frac{\bar{d} - D_0}{S_d/\sqrt{n}}$$

โดยที่ $\bar{d} = \frac{\Sigma d_i}{n}$

และ $S_d^2 = \frac{1}{n-1}\left[\Sigma d_i - \frac{(\Sigma d_i)^2}{n}\right]$

6. สรุปผล

1. ถ้า t_c ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า t_c ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่างที่ 10.5.4

สถานบริการลดความอ้วนโฆษณาว่า วิธีการลดความอ้วนแบบใหม่จะช่วยลดน้ำหนักได้ภายใน 2 อาทิตย์ ได้มีการทดลองใช้วิธีการแบบใหม่กับคน 7 คน โดยชั่งน้ำหนักก่อนทดลอง และหลังจากทดลองการใช้วิธีการลดความอ้วนแบบใหม่แล้ว 2 อาทิตย์ นำมาชั่งน้ำหนักได้ผลดังนี้ (น้ำหนักเป็นปอนด์)

คนที่	1	2	3	4	5	6	7
นน. ก่อนทดลอง	129	133	136	152	141	138	125
นน. หลังทดลอง	130	121	128	137	129	132	120

จะใช้ข้อมูลที่ได้จากตัวอย่างนี้สนับสนุนการโฆษณาได้หรือไม่ที่ $\alpha = .01$

1. ตั้งสมมติฐาน

$H_0 : \mu_1 - \mu_2 = 0$

$H_a : \mu_1 - \mu_2 > 0$

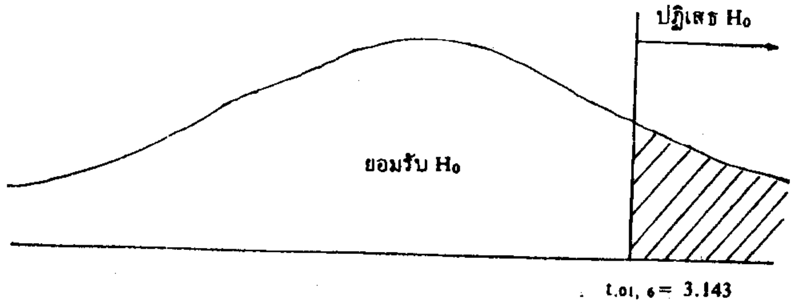
2. $\alpha = .01$, $n = 7$ คู่

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$t = \frac{\bar{d} - 0}{S_d/\sqrt{n}}$$

$df = n - 1 = 6$

4. CR : ปฏิเสธ H_0 ถ้า $t_c > 3.143$



5. คำนวณหาค่า d_i, $\bar{d}$, S_d^2 ได้ดังนี้

คนที่	1	2	3	4	5	6	7	รวม
นน. ก่อนทดลอง	129	133	136	152	141	138	125	
นน. หลังทดลอง	130	121	128	137	129	132	120	
d_i (ผลต่าง)	−1	12	8	15	12	6	5	57
d_i^2	1	144	64	225	144	36	25	639

$$\therefore \bar{d} = \frac{\Sigma d_i}{n} = \frac{57}{7} = 8.14$$

$$S_d^2 = \frac{1}{6}\left[639 - \frac{(57)^2}{7}\right]$$

$$= \frac{1}{6}\left[639 - \frac{3249}{7}\right] = \frac{1}{6}[639 - 464.14]$$

$$= \frac{1}{6}[174.86] = 29.14$$

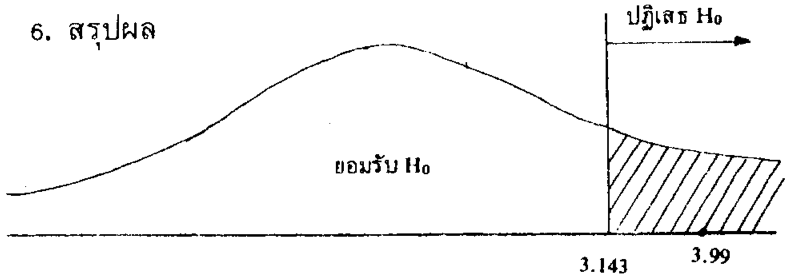
$$\therefore S_d = \sqrt{29.14} = 5.4$$

$$\therefore t_c = \frac{8.14}{5.4/\sqrt{7}} = \frac{8.14}{\frac{5.4}{2.65}}$$

$$= \frac{8.14 \times 2.65}{5.4} = \frac{21.57}{5.4}$$

$$= 3.99$$

6. สรุปผล



$\therefore t_c$ ตกอยู่ใน CR. ดังนั้น จึงปฏิเสธ H_0 และยอมรับ H_a ที่ว่า $\mu_1 - \mu_2 > 0$ แสดงว่าวิธีการแบบใหม่ช่วยในการลดน้ำหนักจริง

10.6 การทดสอบเกี่ยวกับสัดส่วนของประชากร

มีขั้นตอนในการทดสอบ ดังนี้

1. ตั้งสมมติฐาน

$H_0 : \pi = \pi_0$

H_a : 1. $\pi > \pi_0$ หรือ

2. $\pi < \pi_0$ หรือ

3. $\pi \neq \pi_0$

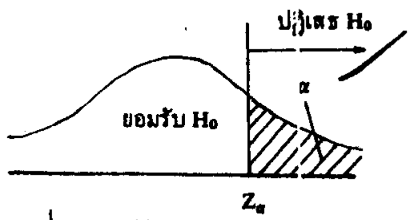
2. กำหนด α, และขนาดตัวอย่าง n

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0 (1 - \pi_0)}{n}}}$$

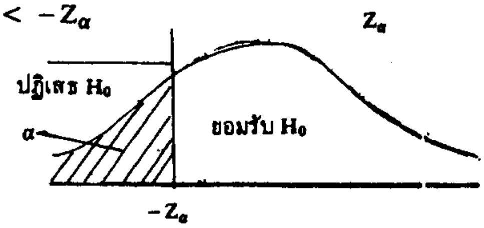
4. CR : 1. ถ้า $H_a : \pi > \pi_0$

จะปฏิเสธ H_0 ถ้า $Z_c > Z_\alpha$

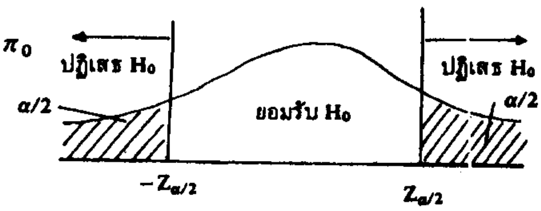


2. ถ้า $H_a : \pi < \pi_0$

จะปฏิเสธ H_0 ถ้า $Z_c < -Z_\alpha$



3. ถ้า $H_a : \pi \neq \pi_0$



จะปฏิเสธ H_0 ถ้า $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$

5. คำนวณหา Z_c จากสูตร

$$Z_c = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

6. สรุปผล : 1. ถ้า Z_c ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า Z_c ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่าง 10.6.1

บริษัทผู้ผลิตอ้างว่า กล้องถ่ายรูปที่ผลิตออกมาจะมีกล้องที่ชำรุดไม่เกิน 10% จากการสุ่มตัวอย่างกล้อง 200 อัน ปรากฏว่ามีกล้องที่ชำรุด 4 อัน จงทดสอบคำกล่าวอ้างนี้ที่ $\alpha = 0.01$

1. ตั้งสมมติฐาน

$H_0 : \pi = 0.01$

$H_a : \pi > 0.01$

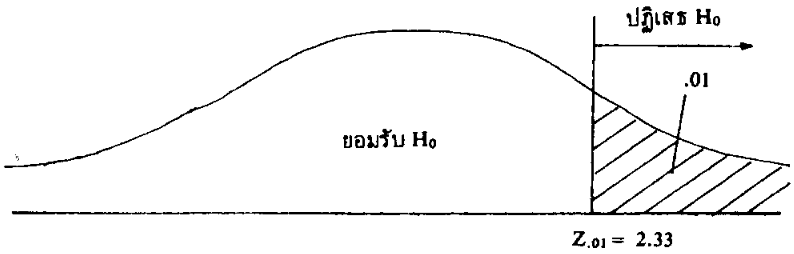
2. $\alpha = 0.01, n = 200$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

4. CR : ปฏิเสธ H_0 ถ้า $Z_c > Z_{.01}$

คือ $Z_c > 2.33$



5. $P = \frac{4}{200} = \frac{1}{50} = 0.02$

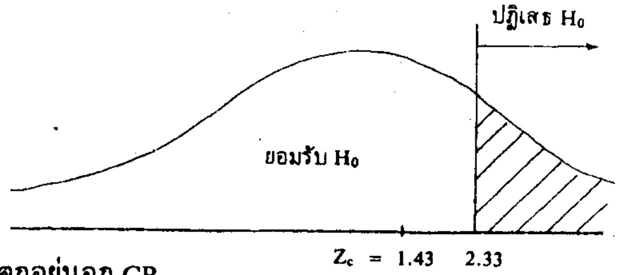
$\pi_0 = 0.01, \; n = 200$

$$\therefore Z = \frac{.02 - .01}{\sqrt{\frac{(.01)(.99)}{200}}}$$

$$= \frac{.02 - .01}{.007}$$

$$= \frac{.01}{.007} = 1.43$$

6. $Z_c < 2.33$



จะเห็นได้ว่า Z_c ตกอยู่นอก CR.

$\therefore$ เรายอมรับ H_0 ที่ว่า $\pi = 0.01$

แสดงว่า กล้องถ่ายรูปที่ผลิตจากบริษัทนี้จะมีของชำรุดไม่เกิน 1% จริง

ตัวอย่าง 10.6.2

ในการสำรวจพัสดุที่ส่งทางไปรษณีย์ 5,000 ชิ้น ปรากฏว่ามี 19 ชิ้นที่ไม่ถึงผู้รับปลายทาง จากผลการสำรวจนี้จะสรุปด้วย $\alpha = 0.05$ ได้หรือไม่ว่าอัตราพัสดุสูญหายได้เพิ่มสูงจากเดิมซึ่งมีไม่เกิน 0.3% หรือไม่

1. $H_0 : \pi = 0.003$

   $H_a : \pi > 0.003$

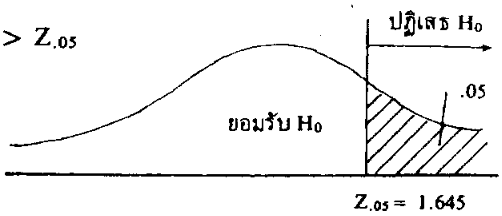
2. $\alpha = 0.05, \; n = 5,000$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0 (1 - \pi_0)}{n}}}$$

4. CR : จะปฏิเสธ H_0 ถ้า $Z_c > Z_{.05}$

คือ $Z_c > 1.645$



5. คำนวณหา Z_c จาก

$$Z_c = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

เมื่อ $P = \frac{19}{5,000} = 0.0038$

$\pi_0 = 0.003, \ n = 5,000$

$$\therefore \quad Z_c = \frac{0.0038 - 0.003}{\sqrt{\frac{0.003\,(1-0.003)}{5,000}}}$$

$$= \frac{0.0008}{\sqrt{\frac{(.003)\,(.997)}{5,000}}}$$

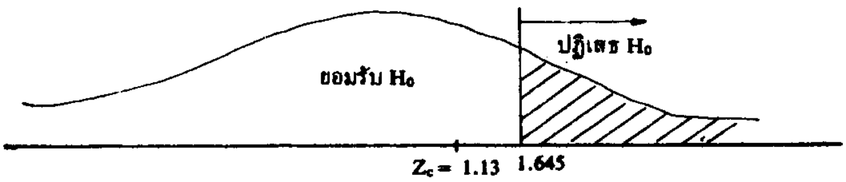
$$= \frac{.0008}{\sqrt{\frac{.00299}{5,000}}}$$

$$= \frac{.0008}{\sqrt{.0000005}}$$

$$\therefore Z_c = \frac{.0008}{.0007071} = 1.13$$

6. สรุปผล

จะเห็นว่า Z_c ตกอยู่นอก CR. ดังนั้นจึงยอมรับ H_0 ที่ว่า $\pi = 0.003$ แสดงว่าอัตราการสูญหายของพัสดุไม่ได้เพิ่มสูงจากเดิม



10.7 การทดสอบความแตกต่างระหว่างอัตราส่วน 2 ค่า

มีขั้นตอนในการทดสอบ ดังนี้

1. ตั้งสมมติฐาน

$H_0 : \pi_1 - \pi_2 = D_0$

H_a : 1. $\pi_1 - \pi_2 > D_0$ หรือ

2. $\pi_1 - \pi_2 < D_0$ หรือ

3. $\pi_1 - \pi_2 \neq D_0$

2. กำหนด α และ n_1, n_2

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{(P_1 - P_2)\,(\pi_1 - \pi_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}}$$

เมื่อ $p_1 = \frac{x_1}{n_1},\ q_1 = 1 - p_1$

$p_2 = \frac{x_2}{n_2},\ q_2 = 1 - p_2$

ในกรณีที่ D_0 มีค่า $= 0$ (คือ $\pi_1 = \pi_1$) **ตัวสถิติที่ใช้ทดสอบ คือ**

$$Z = \frac{(P_1 - P_2) - 0}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

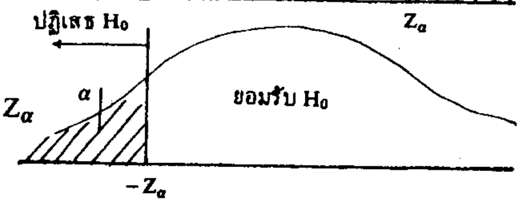
เมื่อ $P = \frac{x_1 + x_2}{n_1 + n_2}$, $Q = 1 - P$

4. CR : 1. ถ้า $H_a : \pi_1 - \pi_2 > D_0$

จะปฏิเสธ H_0 ถ้า $Z_c > Z_\alpha$

2. $H_a : \pi_1 - \pi_2 < D_0$

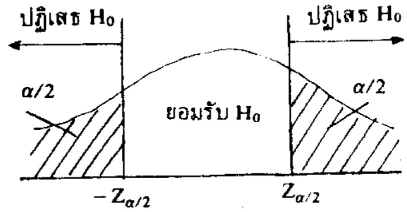
จะปฏิเสธ H_0 ถ้า $Z_c < -Z_\alpha$



3. $H_a : \pi_1 - \pi_2 \neq D_0$

จะปฏิเสธ H_0 ถ้า $Z_c < -Z_{\alpha/2}$ หรือ

$Z_c > Z_{\alpha/2}$



5. คำนวณค่า Z_c จาก

1. ถ้า $H_0 : \pi_1 - \pi_2 = D_0$

$$Z_c = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

เมื่อ $p_1 = \frac{x_1}{n_2}$, $q_1 = 1 - p_1$

$p_2 = \frac{x_2}{n_2}$, $q_2 = 1 - p_2$

2. ถ้า $H_0 : \pi_1 - \pi_2 = 0$

$$Z_c = \frac{(P_1 - P_2)}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

เมื่อ $P = \frac{X_1 + x_2}{n_1 + n_2}$, $Q = 1 - P$

6. สรุปผล 1. ถ้า Z_c ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า Z_c ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่างที่ 10.7.1

มีกล่องอยู่ 2 กล่อง คือ A และ B ในแต่ละกล่องมีลูกหินจำนวนเท่า ๆ กัน แต่สัดส่วนของลูกหินสีแดงและสีขาวเราไม่ทราบจากการสุ่มตัวอย่างลูกหิน 50 ลูก แบบแทนที่ จากแต่ละกล่องพบว่ามีลูกหินสีแดง 32 ลูก จากกล่อง A และลูกหินสีแดง 23 ลูก จากกล่อง B ให้ใช้ $\alpha = .01$ ทดสอบสมมติฐานว่ากล่อง A มีอัตราส่วนของลูกหินสีแดงมากกว่ากล่อง B

ให้ π_A = สัดส่วนของลูกหินสีแดงในกล่อง A

π_B = สัดส่วนของลูกหินสีแดงในกล่อง B

1. $H_0 : \pi_A = \pi_B$ หรือ $\pi_A - \pi_B = 0$

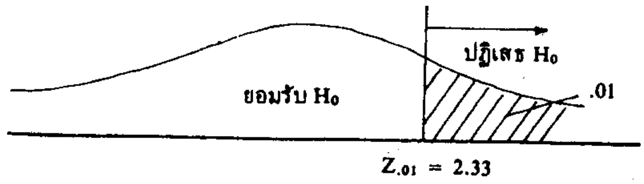
$H_a : \pi_A > \pi_B$ หรือ $\pi_A - \pi_B > 0$

2. $\alpha = 0.01,\ n_A = 50,\ n_B = 50$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{P_A - P_B}{\sqrt{PQ\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

4. CR : จะปฏิเสธ H_0 ถ้า $Z_c > Z_{.01}$ คือ $Z_c > 2.33$



5. คำนวณค่า Z_c จาก

$$Z_c = \frac{P_A - P_B}{\sqrt{PQ\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}}$$

$$P_A = \frac{32}{50},\ P_B = \frac{23}{50}$$

$$\therefore P = \frac{32+23}{50+50} = \frac{55}{100} = 0.55$$

$$\therefore Q = 1-0.55 = 0.45$$

แทนค่า

$$Z_c = \frac{\left(\frac{32}{50}-\frac{23}{50}\right)}{\sqrt{(.55)(.45)\left(\frac{1}{50}+\frac{1}{50}\right)}}$$

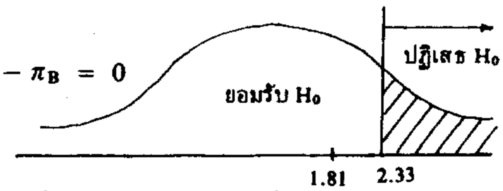
$$= \frac{\frac{32-23}{50}}{\sqrt{0.2475 \times \frac{100}{2,500}}}$$

$$= \frac{\frac{9}{50}}{\sqrt{0.2475 \times .04}} = \frac{0.18}{\sqrt{0.0099}}$$

$$= \frac{0.18}{.0995} = 1.81$$

6. สรุปผล $\because Z_c$ ตกอยู่นอก CR.

$\therefore$ เรายอมรับ H_0 ที่ว่า $\pi_A - \pi_B = 0$



แสดงว่า กล่อง A มีอัตราส่วนของลูกหินสีแดงไม่มากกว่ากล่อง B

ตัวอย่างที่ 10.7.2

ผู้จัดการผลิตเยลลี่ชนิดหนึ่งได้สุ่มตัวอย่างจากร้านขายขนมเด็ก 300 ร้าน พบว่ามี 53% ที่ขายเยลลี่ของบริษัทตนเอง ต่อมาทางบริษัทได้ใช้วิธีขายแบบใหม่และได้สุ่มตัวอย่างร้านขายขนมเด็กมี 400 ร้าน พบว่ามี 40% ที่ขายเยลลี่ของบริษัทตนเอง จงทดสอบที่ $\alpha = 0.05$ ว่า การใช้วิธีขายแบบเดิมทำให้การขายเยลลี่สูงกว่าแบบใหม่ 5%

ให้ π_1 = อัตราส่วนของร้านที่ขายเยลลี่ของบริษัทโดยใช้วิธีขายแบบเดิม

π_2 = อัตราส่วนของร้านที่ขายเยลลี่ของบริษัทโดยใช้วิธีขายแบบใหม่

1. $H_0 : \pi_1 - \pi_2 = 0.05$

   $H_a : \pi_1 - \pi_2 > 0.05$

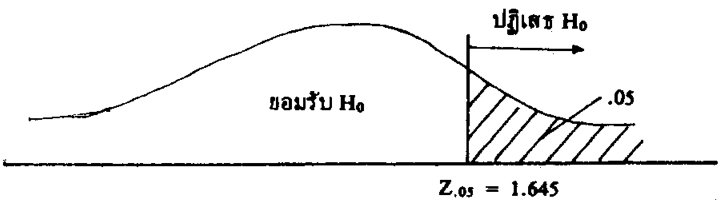
2. $\alpha = 0.05, n_1 = 300, n_2 = 400$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$Z = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}}$$

4. CR : จะปฏิเสธ H_0 ถ้า $Z_c > Z_{.05}$

   คือ $Z_c > 1.645$



5. คำนวณหา Z_c จาก

$$Z_c = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_1}}}$$

เมื่อ $p_1 = 0.53$, $p_2 = 0.40$

$q_1 = 0.47$, $q_2 = 0.60$

$n_1 = 300$, $n_2 = 400$

$$\therefore Z_c = \frac{(0.53 - 0.40) - (0.05)}{\sqrt{\frac{(.53)\ (.47)}{300} + \frac{(.40)\ (.60)}{400}}}$$

$$= \frac{.13 - .05}{\sqrt{\frac{.249}{300} + \frac{.24}{400}}}$$

$$= \frac{.08}{\sqrt{.00083 + .0006}}$$

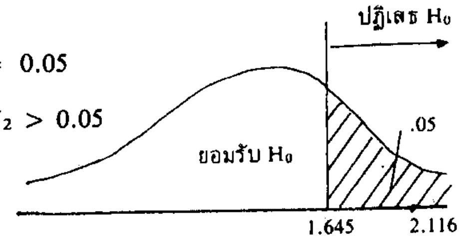
$$= \frac{.08}{\sqrt{.00143}} = \frac{.08}{.0378}$$

$$= 2.116$$

6. สรุปผล $\because Z_c$ ตกอยู่ใน CR.

$\therefore$ ปฏิเสธ $H_O : \pi_1 - \pi_2 = 0.05$

และยอมรับ $H_a : \pi_1 - \pi_2 > 0.05$



แสดงว่าการขายแบบเดิมทำให้การขายเยลลี่สูงกว่าแบบใหม่ 5% จริง

10.8 การทดสอบเกี่ยวกับความแปรปรวนของประชากรและส่วนเบี่ยงเบนมาตรฐานของประชากรที่มีการแจกแจงแบบปกติ

สุ่มตัวอย่างขนาด n จากประชากรที่มีการแจกแจงแบบปกติ ซึ่งมีความแปรปรวนเป็น σ^2 ขั้นตอนในการทดสอบมีดังนี้

1. ตั้งสมมติฐาน

สำหรับความแปรปรวน

$H_0 : \sigma^2 = \sigma_0^2$

H_a : 1. $\sigma^2 > \sigma_0^2$ หรือ

2. $\sigma^2 < \sigma_0^2$ หรือ

3. $\sigma^2 \neq \sigma_0^2$

สำหรับส่วนเบี่ยงเบนมาตรฐาน

$H_0 : \sigma = \sigma_0$

H_a : 1. $\sigma > \sigma_0$ หรือ

2. $\sigma < \sigma_0$ หรือ

3. $\sigma \neq \sigma_0$

2. กำหนด α และขนาดตัวอย่าง n

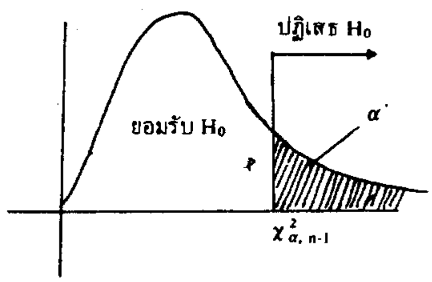
3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \frac{(n-1)\,S^2}{\sigma_0^2}$$

มี df = n − 1

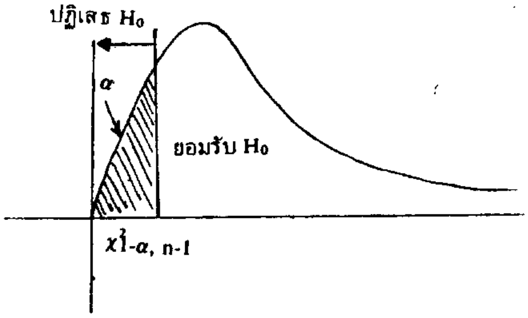
4. CR : 1. ถ้า $H_a : \sigma^2 > \sigma_0^2$ หรือ $\sigma > \sigma_0$

จะปฏิเสธ H_0 ถ้า $\chi_c^2 > \chi^2_{\alpha,\, n-1}$



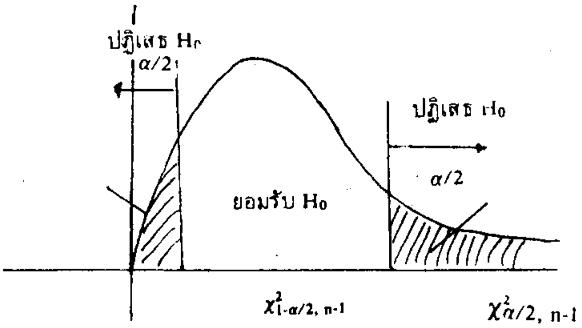
2. ถ้า $H_a : \sigma^2 < \sigma_0^2$ หรือ $\sigma < \sigma_0$

จะปฏิเสธ H_0 ถ้า $\chi_c^2 < \chi^2_{1-\alpha,\, n-1}$



3. ถ้า $H_a : \sigma^2 \neq \sigma_0^2$ หรือ $\sigma \neq \sigma_0$

จะปฏิเสธ H_0 ถ้า $\chi_c^2 < \chi^2_{1-\alpha/2,\, n-1}$ หรือ $\chi_c^2 > \chi^2_{\alpha/2,\, n-1}$



5. คำนวณหาค่า χ_c^2 จาก

$$\chi_c^2 = \frac{(n-1)\, S^2}{\sigma_0^2}$$

เมื่อ $S^2 = \frac{1}{n-1}\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right]$

6. สรุปผล

1. ถ้า χ_c^2 ตกอยู่ใน CR. จะปฏิเสธ H_0 และยอมรับ H_a
2. ถ้า χ_c^2 ตกอยู่นอก CR. จะยอมรับ H_0

ตัวอย่าง 10.8.1

บริษัทผลิตยางรถยนต์โฆษณาว่า อายุการใช้งานของยางรถยนต์มีความแปรปรวนเท่ากับ 0.81 ปี ถ้าสุ่มตัวอย่างยางรถยนต์มา 10 เส้นแล้วคำนวณส่วนเบี่ยงเบนมาตรฐานได้ 1.2 ปี จงทดสอบว่า $\sigma^2 > 0.81$ ปี โดยใช้ $\alpha = 0.05$

1. $H_0 : \sigma^2 = 0.81$ ปี

   $H_a : \sigma^2 > 0.81$ ปี

2. $\alpha = 0.05,\ n = 10$

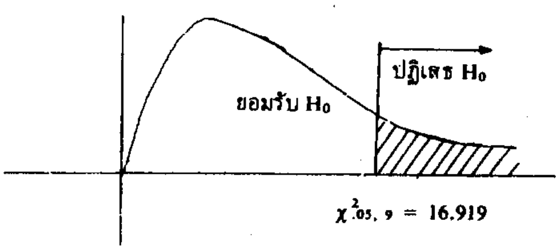
3. ตัวสมมติที่ใช้ทดสอบ คือ

$$\chi^2 = \frac{(n-1)\, S^2}{\sigma^2}$$

$df = n - 1 = 10 - 1 = 9$

4. CR : ปฏิเสธ H_0 ถ้า $\chi_c^2 > \chi^2_{.05,\,9}$

คือ $\chi_c^2 > 16.919$



5. คำนวณค่า χ_c^2 จาก

$$\chi_c^2 = \frac{(n-1)\,S^2}{\sigma_0^2}$$

$$= \frac{(9)\,(1.2)^2}{0.81}$$

$$= \frac{(9)\,(1.44)}{0.81} = \frac{12.36}{0.81}$$

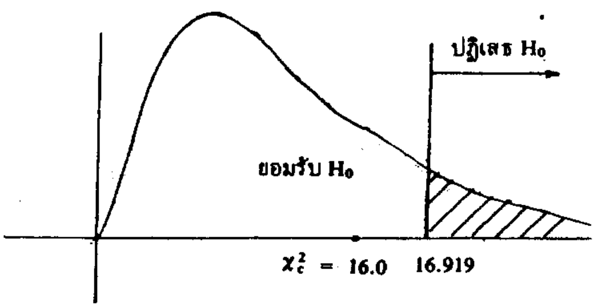
$$= 16.0$$

6. สรุปผล

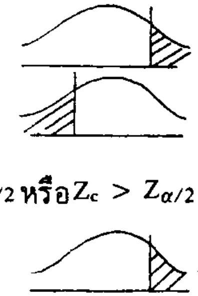
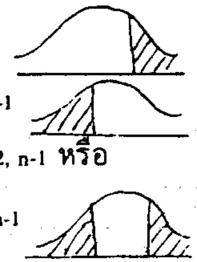
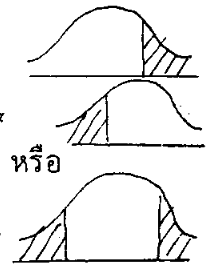
$\because$ χ_c^2 ตกอยู่นอก CR.

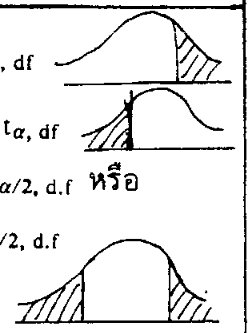
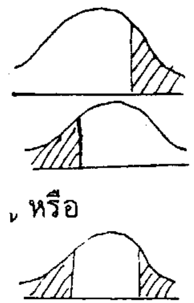
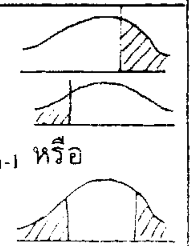
$\therefore$ ยอมรับ H_0 ที่ว่า $\sigma^2 = 0.81$

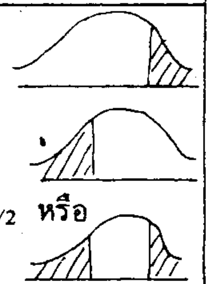
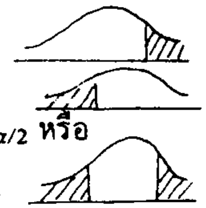
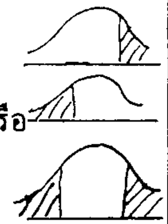
แสดงว่าอายุการใช้งานของยางรถยนต์ที่ผลิตจากบริษัทนี้ มีความแปรปรวนเท่ากับ 0.81 ปี

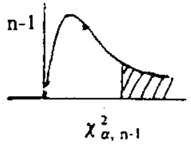
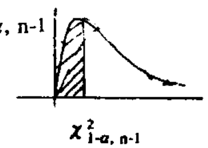
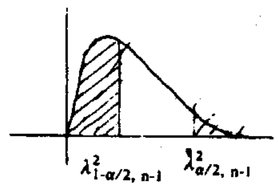


10.9 ตารางสรุปเกี่ยวกับการทดสอบสมมติฐานของค่าเฉลี่ยประชากรสัดส่วนประชากร และความแปรปรวนของประชากร

	H_0	เงื่อนไข	ตัวสถิติที่ใช้ทดสอบ	H_a	CR :
1	$\mu = m_0$	ทราบความแปรปรวน และประชากรมีการแจกแจงแบบปกติ หรือเมื่อไม่ทราบความแปรปรวนแต่ตัวอย่างมีขนาดโต $(n \geqslant 30)$	$Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{N}} \sim N(0, 1)$ (ทราบความแปรปรวน) $Z = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$ (ไม่ทราบ σ^2)	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$Z_c > Z_\alpha$ $Z_c < -Z_\alpha$ $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ 
2	$\mu = \mu_0$	ไม่ทราบความแปรปรวนและประชากรมีการแจกแจงแบบปกติ เมื่อ n ขนาดเล็ก $(n < 30)$	$T = \frac{\overline{X}-\mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$ $df = n-1$ $S^2 = \frac{1}{n-1}\left[\Sigma x^2 - \frac{(\Sigma x)^2}{n}\right]$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \neq \mu_0$	$\tau_c > t_{\alpha, n-1}$ $t_c < -t_{\alpha, n-1}$ $t_c < -t_{\alpha/2, n-1}$ หรือ $t^c > t_{\alpha/2, n-1}$ 
3.	$\mu_1-\mu_2 = D_0$	ทราบความแปรปรวนของประชากรทั้ง 2 และประชากรทั้ง 2 มีการแจกแจงแบบปกติ ตัวอย่างที่สุ่มมาขนาด n_1 และ n_2 เป็นอิสระต่อกัน	$Z = \frac{(\overline{X}_1-\overline{X}_2)-D_0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$	$\mu_1-\mu_2 > D_0$ $\mu_1-\mu_2 < D_0$ $\mu_1-\mu_2 \neq D_0$	$Z_c > Z_\alpha$ $Z_c < -Z_\alpha$ $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ 

	H_0	เงื่อนไข	ตัวสถิติที่ใช้ทดสอบ	H_a	CR.
4.	$\mu_1-\mu_2 = D_0$	ประชากรทั้ง 2 มีการแจกแจงแบบปกติไม่ทราบความแปรปรวนของประชากรทั้งสอง แต่ทราบว่าเท่ากัน ตัวอย่างที่สุ่มมาเป็นอิสระต่อกันและ n_1, n_2 มีขนาดเล็ก	$T = \dfrac{(\overline{X}_1-\overline{X}_2)-D_0}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$ $d.f = n_1+n_2-2$ $S_p^2 = \dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$\mu_1-\mu_2 > D_0$ $\mu_1-\mu_2 < D_0$ $\mu_1-\mu_2 \neq D_0$	$t_c > t_{\alpha, df}$ $t_c < -t_{\alpha, df}$ $t_c < -t_{\alpha/2, d.f}$ หรือ $t_c > t_{\alpha/2, d.f}$ 
5.	$\mu_1-\mu_2 = D_0$	ประชากรทั้ง 2 มีการแจกแจงแบบปกติ ไม่ทราบความแปรปรวนของประชากรทั้ง 2 แต่ทราบว่าไม่เท่ากัน ตัวอย่างที่สุ่มมาเป็นอิสระกัน	$T' = \dfrac{(\overline{X}_1-\overline{X}_2)-D_0}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_\nu$ $\nu = \dfrac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_1-1)}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{(n_2-1)}}$	$\mu_1-\mu_2 > D_0$ $\mu_1-\mu_2 < D_0$ $\mu_1-\mu_2 \neq D_0$	$t_c' > t_{\alpha, \nu}$ $t_c' < -t_{\alpha, \nu}$ $t_c' < -t_{\alpha/2, \nu}$ หรือ $t_c' > t_{\alpha/2, \nu}$ 
6.	$\mu_1-\mu_2 = D_0$	ข้อมูลมีลักษณะจับคู่กันหรือกรณีที่ตัวอย่างที่สุ่มมาไม่เป็นอิสระต่อกัน (N = จำนวนคู่)	$T = \dfrac{\overline{d}-D_0}{S_d/\sqrt{n}} \sim t_{n-1}$ $\overline{d} = \dfrac{\Sigma d_i}{n}$ $S_d^2 = \dfrac{1}{n-1}\left[\Sigma d_i^2 - \dfrac{(\Sigma d_i)^2}{n}\right]$	$\mu_1-\mu_2 > D_0$ $\mu_1-\mu_2 < D_0$ $\mu_1-\mu_2 \neq D_0$	$t_c > t_{\alpha, n-1}$ $t_c < -t_{\alpha, n-1}$ $t_c < -t_{\alpha/2, n-1}$ หรือ $t_c > t_{\alpha/2, n-1}$ 

	H_0	เงื่อนไข	ตัวสถิติที่ใช้ทดสอบ	H_a	CR.
7.	$\pi = \pi_0$		$Z = \dfrac{P-\pi}{\sqrt{\dfrac{\pi_0(1-\pi}{n}}} \sim N(0,1)$ $P = \dfrac{x}{n}$	$\pi > \pi_0$ $\pi < \pi_0$ $\pi \neq \pi_0$	$Z_c > Z_\alpha$ $Z_c < -Z_\alpha$ $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ 
8	$\pi_1 - \pi_2 = D_0$		$Z = \dfrac{(P_1-P_2)-D_0}{\sqrt{\dfrac{p_1q_1}{n_1}+\dfrac{p_2q_2}{n_2}}} \sim N(0,1)$ $p_1 = \dfrac{x_1}{n_1}$, $q_1 = 1-p_1$ $p_2 = \dfrac{x_2}{n_2}$, $q_2 = 1-p_2$	$\pi_1 - \pi_2 > D_0$ $\pi_1 - \pi_2 < D_0$ $\pi_1 - \pi_2 \neq D_0$	$Z_c > Z_\alpha$ $Z_c < Z_\alpha$ $Z_c < -Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ 
9.	$\pi_1 - \pi_2 = 0$ เมื่อ $D_0 = 0$		$Z = \dfrac{(P_1-P_2)}{\sqrt{PQ\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}} \sim N(0,1)$ $p_1 = \dfrac{x_1}{n_1}$, $P = \dfrac{x_1+x_2}{n_1+n_2}$ $p_2 = \dfrac{x_2}{n_2}$, $Q = 1-P$	$\pi_1 - \pi_2 > 0$ $\pi_1 - \pi_2 < 0$ $\pi_1 - \pi_2 \neq 0$	$Z_c > Z_\alpha$ $Z_c < -Z_\alpha$ $Z_c < Z_{\alpha/2}$ หรือ $Z_c > Z_{\alpha/2}$ 

	H_a	เงื่อนไข	ตัวสถิติที่ใช้ทดสอบ	H_a	CR.
10.	$\sigma^2 = \sigma_0^2$ หรือ $\sigma = \sigma_0$	ประชากรมีการแจกแจงแบบปกติ	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{n-1}$ $df = n-1$	$\sigma^2 > \sigma_0^2$ หรือ $\sigma > \sigma_0$	$\chi_c^2 > \chi^2_{\alpha, n-1}$ 
				$\sigma_2 < \sigma_0^2$ หรือ $\sigma < \sigma_0$	$\chi^2 < \chi^2_{1-\alpha, n-1}$ 
				$\sigma_2 \neq \sigma_0^2$ หรือ $\sigma \neq \sigma_0$	$\chi^2 < \chi^2_{1-\alpha/2, n-1}$ หรือ $\chi^2 > \chi^2_{\alpha/2, n-1}$ 

10.10 การวิเคราะห์ข้อมูลที่จัดจำแนกแล้ว (Analysis of Categorized Data)

Observed frequency และ Expected frequency

Observed frequency หมายถึง ความถี่ที่สังเกตหรือนับได้จากการทดลอง ซึ่งถ้าการทดลองมีเหตุการณ์ที่เกิดขึ้น k เหตุการณ์ ∴ observed frequency คือ O_1, O_2, ..., O_k ซึ่ง $\sum_{i=1}^{k} O_i = n$

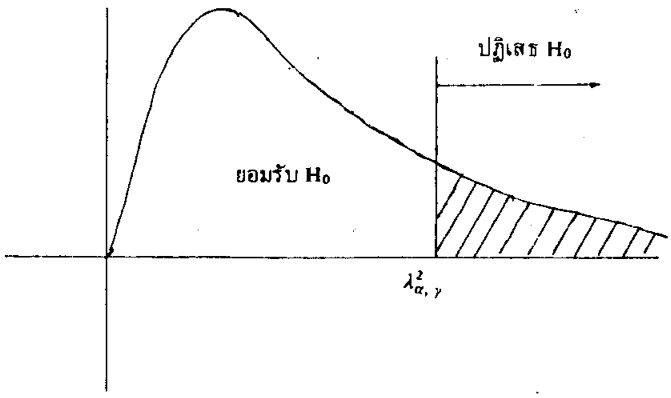
Expected frequency หมายถึงความถี่ที่คาดหมายว่าควรจะเป็นหรือควรจะเกิดขึ้นในการทดลอง หาได้โดยใช้หลักความน่าจะเป็น (probability) เราใช้สัญลักษณ์แทนด้วย E_i ∴ $\sum_{i=1}^{k} E_i = n$

10.10.1 การทดสอบ goodness of fit

เป็นการทดสอบว่าค่าความถี่ที่สังเกตหรือที่นับมาได้จากตัวอย่างที่สุ่มมา (observed frequency แทนด้วย O_i) และค่าความถี่ที่คาดหมายว่าควรจะเป็น (Expected frequency แทนด้วย E_i) นั้นมีค่าใกล้เคียงกันดีหรือไม่ ซึ่งการทดสอบขึ้นอยู่กับค่า χ^2 เมื่อ

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

ค่า O_i และ E_i (ที่คู่กัน) จะใกล้เคียงกันถ้าค่า χ^2_c มีค่าเล็กหรือถ้าค่า O_i และ E_i เท่ากันค่า χ^2_c จะมีค่าเป็นศูนย์ ดังนั้นในการทดสอบ goodness of fit ว่า H_0 : ข้อมูลที่ได้จากการสังเกตกับข้อมูลที่ได้จากการคาดหมายว่าควรจะเป็นนั้นใกล้เคียงกันดี หรือไม่เราจะปฏิเสธ H_0 เมื่อ χ^2_c มีค่ามากกว่า $\chi^2_{\alpha, \nu}$



โดยที่ ν (df) = จำนวนเทอมที่รวมกันเป็นค่า χ_c^2 − 1 − จำนวนพารามิเตอร์ที่ต้องประมาณค่า

นอกจากนี้ในการทดสอบ goodness of fit เรายังจะทดสอบว่าข้อมูลชุดที่เราสังเกตมาได้มีการแจกแจงใกล้เคียงการแจกแจงความน่าจะเป็น (Probability distribution) ที่เรารู้จักหรือไม่ เช่น มีการแจกแจงแบบปกติหรือมีการแจกแจงแบบทวินาม หรือไม่ เป็นต้น

ตัวอย่างที่ 10.10.1.1

จากการทดลองของเมนเดลเกี่ยวกับเมล็ดถั่ว ผลปรากฏดังนี้

ลักษณะของเมล็ดถั่ว	เมล็ดกลมและสีเหลือง	เมล็ดกลมและสีเขียว	เมล็ดย่นและสีเหลือง	เมล็ดย่นและสีเขียว
จำนวนเมล็ดถั่ว (O_i)	315	108	101	32

จากทฤษฎีของเขาเกี่ยวกับกรรมพันธุ์จำนวนที่ได้ ควรจะเป็นอัตราส่วน 9 : 3 : 3 : 1 จงทดสอบที่ α = 0.10 ว่า ข้อมูลที่สังเกตได้นั้นจะเป็นจริงตามทฤษฎีของเมนเดล หรือไม่

1. H_0 : อัตราส่วนของลักษณะเมล็ดถั่วจะเป็น 9 : 3 : 3 : 1

   H_a : อัตราส่วนของลักษณะเมล็ดถั่วไม่เป็น 9 : 3 : 3 : 1

2. $\alpha = 0.10$, $n = 315+108+101+32 = 556$

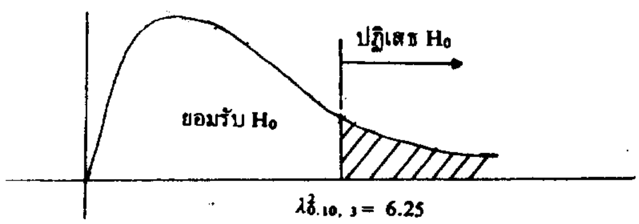
3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi_c^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

df = 4 − 1 = 3

4. CR : ปฏิเสธ H_0 ถ้า $\chi_c^2 > \chi^2_{.10,\,3}$

   คือ $\chi_c^2 > 6.25$



5. คำนวณค่า χ_c^2 จาก

$$\chi_c^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$$

ถ้า H_0 จริง

ได้ $P_i = \frac{1}{16}$

$\therefore P_1 = P$ [ได้เมล็ดกลมสีเหลือง] $= \frac{9}{16}$

$P_2 = P$ [ได้เมล็ดกลมสีเขียว] $= \frac{3}{16}$

$P_3 = P$ [ได้เมล็ดย่นสีเหลือง] $= \frac{3}{16}$

$P_4 = P$ [ได้เมล็ดย่นสีเขียว] $= \frac{1}{16}$

หา Expected frequency ได้จาก $E_i = n \times P_i$

$$\therefore E_1 = n \times P_1 = 556 \times \frac{9}{16} = 312.75$$

$$E_2 = n \times P_2 = 556 \times \frac{3}{16} = 104.25$$

$$E_3 = n \times P_3 = 556 \times \frac{3}{16} = 104.25$$

$$E_4 = n \times P_4 = 556 \times \frac{1}{16} = 34.75$$

ซึ่งเขียนสรุปลงในตารางได้ ดังนี้

	เมล็ดกลมสีเหลือง	เมล็ดกลมสีเขียว	เมล็ดย่นสีเหลือง	เมล็ดย่นสีเขียว
O_i	315	108	101	32
E_i	312.75	104.25	104.25	34.75

$$\therefore \chi_c^2 = \frac{(315-312.75)^2}{312.75} + \frac{(108-104.25)^2}{104.25} + \frac{(101-104.25)^2}{104.25} + \frac{(32-34.75)^2}{34.75}$$

$$= \frac{(2.25)^2}{312.75} + \frac{(3.75)^2}{104.25} + \frac{(-3.25)^2}{104.25} + \frac{(-2.75)^2}{34.75}$$

$$= \frac{5.0625}{312.75} + \frac{14.0625}{104.25} + \frac{10.5625}{104.25} + \frac{7.5625}{34.75}$$

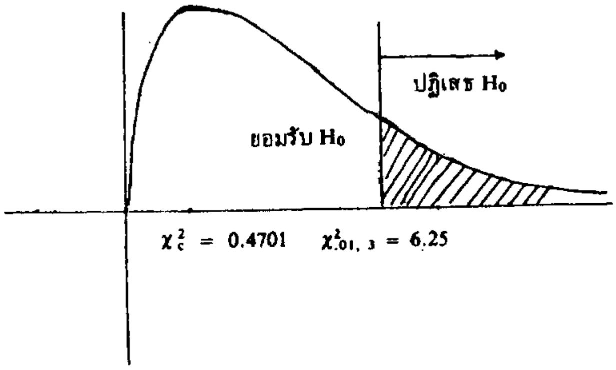
$$= 0.0162 + 0.1349 + 0.1013 + 0.2177$$

$$= 0.4701$$

6. สรุปผล $\because \chi_c^2$ ตกอยู่นอก CR.

$\therefore$ ยอมรับ H_0 แสดงว่าทฤษฎีนี้เป็นจริง

คือ อัตราส่วนดังกล่าวเป็น 9 : 3 : 3 : 1 จริง



ตัวอย่างที่ 10.10.1.2

โยนลูกเต๋า 180 ครั้ง ได้ผลดังนี้

หน้าลูกเต๋า	1	2	3	4	5	6	รวม
ความถี่ (O_i)	20	35	25	35	32	33	180

จงทดสอบว่าลูกเต๋านี้เป็นลูกเต๋าที่สมดุลย์ ที่ $\alpha = 0.01$

1. H_0 : ลูกเต๋านี้เป็นลูกเต๋าที่สมดุลย์

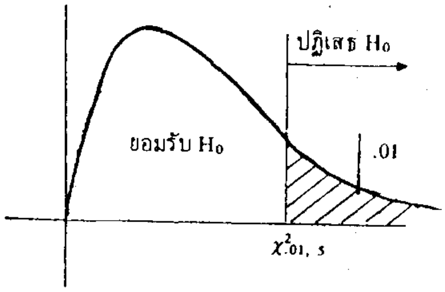
   H_a : ลูกเต๋านี้เป็นลูกเต๋าที่ไม่สมดุลย์

2. $\alpha = 0.01$, n = 180

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$$

$$df = 6 - 1 - 0 = 5$$



4. CR : ปฏิเสธ H_0 ถ้า $\chi^2_{.01,5}$

คือ $\chi^2_c > 15.086$

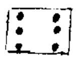
5. คำนวณค่า χ^2_c ต้องหาค่า E_i จาก

$$E_i = n \times p_i$$

ถ้า H_0 จริง ความน่าจะเป็นที่จะทอดลูกเต๋าได้แต่หน้า จะมีค่าเท่ากันคือ

$$P_i = \frac{1}{6}$$

$\therefore \quad E_i = n \times \frac{1}{6} = 180 \times \frac{1}{6} = 30$ ดังตาราง

ผลการทดลอง	ความถี่ที่คาดหมาย (E_i)
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$
	$\frac{1}{6} \times n = \frac{1}{6} \times 180 = 30$

ดังนั้นเราจะได้ตารางค่า O_i และ E_i ดังนี้

หน้าลูกเต๋า	1	2	3	4	5	6
O_i	20	35	25	35	32	33
E_i	30	30	30	30	30	30

จาก $\chi_c^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$

$$\therefore \quad \chi_c^2 = \frac{(20-30)^2}{30} + \frac{(35-30)^2}{30} + \frac{(25-30)^2}{30} + \frac{(35-30)^2}{30} + \frac{(32-30)^2}{30} + \frac{(33-30)^2}{30}$$

$$= \frac{(-10)^2}{30} + \frac{(5)^2}{30} + \frac{(-5)^2}{30} + \frac{(5)^2}{30} + \frac{(2)^2}{30} + \frac{(3)^2}{30}$$

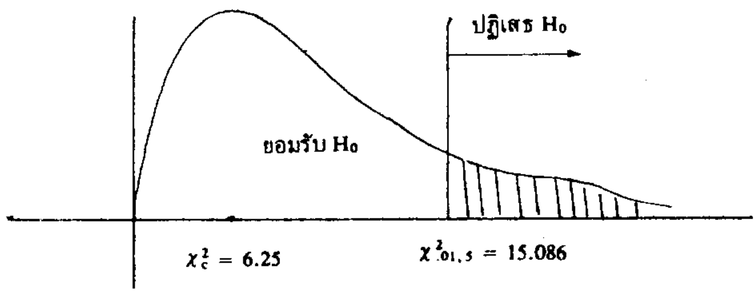
$$= \frac{100}{30} + \frac{25}{30} + \frac{25}{30} + \frac{25}{30} + \frac{4}{30} + \frac{9}{30}$$

$$= 3.33 + .83 + .83 + .83 + .13 + .30$$

$$= 6.25$$

6. สรุปผล $\therefore \ \chi_c^2 < 15.086$

$\therefore \ \chi_c^2$ ตกอยู่นอก CR. ดังนั้น เราจึงยอมรับ H_0 ที่ว่าลูกเต๋านี้เป็นลูกเต๋าที่สมดุลย์



ตัวอย่างที่ 10.10.1.3

โยนเหรียญ 5 อัน 1000 ครั้ง ได้ผลดังนี้

จำนวนหัว (x)	0	1	2	3	4	5	รวม
O_i	38	144	342	287	164	25	1000

จงทดสอบว่าเหรียญทั้ง 5 มีการแจกแจงของจำนวนหัวที่ได้เป็นแบบทวินาม (binomial distribution) ที่ $\alpha = 0.01$

ให้ X = จำนวนหัวที่ได้

$$P[X = x] = \binom{n}{x} p^x q^{n-x} ; x = 0, 1, \ldots, n$$

ในการหาค่า Expected frequency (E_i) เราหาได้จาก

$$E_i = NP_i \text{ เมื่อ } N = 1000$$

แต่เนื่องจากค่า P เราไม่ทราบค่า จึงต้องหาค่า P โดยคำนวณจากค่าเฉลี่ย (Mean) ดังนี้

X	f (คือ O_i)	fX
0	38	0
1	144	144
2	342	684
3	287	861
4	164	656
5	25	125
รวม	1000	2470

$$\text{Mean} = \frac{\Sigma fX}{N} = \frac{2470}{1000} = 2.470$$

ถ้าการแจกแจงของจำนวนหัวที่ได้จากการโยนเหรียญทั้ง 5 เป็น binomial

$(x = 0, 1, 2, 3, 4, 5; \therefore n = 5)$

$\text{Mean} = np$

$2.470 = 5 \times p$

$$\therefore p = \frac{2.470}{5} = 0.49 \approx 0.5$$

1. H_0 : ข้อมูลมีการแจกแจงแบบทวินาม (binomial distribution)

H_a : H_0 ไม่จริง

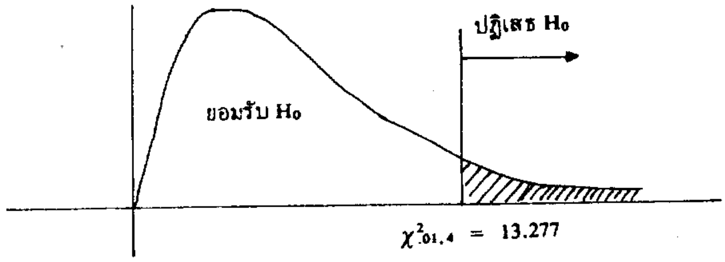
2. $\alpha = 0.01$

3. ตัวสถิติที่ใช้ทดสอบคือ $\chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$

df = จำนวนเทอมที่รวมกันเป็น $\chi_c^2 - 1 -$ พารามิเตอร์ที่ต้องการประมาณค่าพารามิเตอร์ที่ต้องประมาณค่าในที่นี้เท่ากับ 1 เพราะเราต้องประมาณค่าพารามิเตอร์ P

$\therefore \quad df = 6 - 1 - 1 = 4$

4. CR: ปฏิเสธ H_0 ถ้า $\chi_c^2 < \chi^2_{.01,4}$ คือ $\chi_c^2 < 13.277$



5. คำนวณค่า χ_c^2 โดยหา E_i ได้ดังนี้

X	O_i	$P\|X = x\| = \binom{n}{x} p^x q^{n-x}$ p = .05 (ดูจากตาราง binomial)	Expected frequeney $= N \times P\|X = x\|$ โดยที่ N = 1000
0	38	0.0332	33
1	144	0.1619	162
2	342	0.3162	316
3	287	0.3087	309
4	164	0.1507	151
5	25	0.0294	29
รวม	1000	1.0000	1000

จาก $\chi^2_c = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$

$$= \frac{(38-33)^2}{33} + \frac{(144-162)^2}{162} + \frac{(342-316)^2}{316} + \frac{(287-309)^2}{309} + \frac{(164-151)^2}{151} + \frac{(25-29)^2}{29}$$

$$= \frac{5^2}{33} + \frac{(-18)^2}{162} + \frac{(26)^2}{316} + \frac{(-22)^2}{309} + \frac{(13)^2}{151} + \frac{(-4)^2}{29}$$

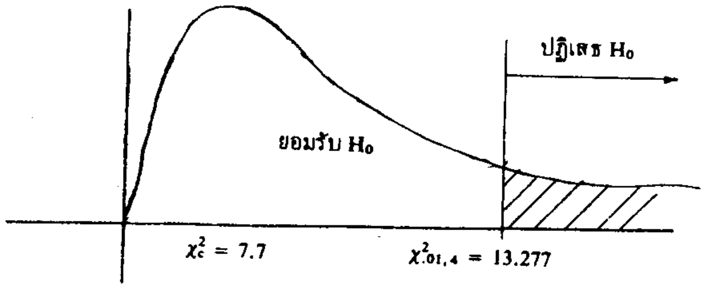
$$= \frac{25}{33} + \frac{324}{162} + \frac{676}{316} + \frac{484}{309} + \frac{169}{151} + \frac{16}{29}$$

$$= 0.7 + 2.0 + 2.1 + 1.2 + 1.1 + 0.6$$

$$= 7.7$$

6. สรุปผล $\therefore \quad \chi^2_c < 13.277$

$\therefore$ ค่า χ^2_c ตกอยู่นอก CR ดังนั้นเราจึงยอมรับ H_0 ที่ว่าข้อมูลที่ได้มามีการแจกแจงแบบทวินาม



ตัวอย่างที่ 10.10.1.4

น้ำหนักของนักศึกษาชายมหาวิทยาลัยรามคำแหง 100 คน เป็นดังนี้

น้ำหนัก (กก)	60 – 62	63 – 65	66 – 68	69 – 71	72 – 74	รวม
ความถี่ (O_i)	5	18	42	27	8	100

จงทดสอบว่าข้อมูลที่ได้มีการแจกแจงแบบปกติ (Normal distribution) ที่ $\alpha = 0.05$

ต้องหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของน้ำหนัก

ให้ X แทนน้ำหนักของนักศึกษาชาย

หาค่า Mean และ Standard deviation โดยวิธีลัด

X (ค่ากึ่งกลาง)	ความถี่ f (คือ O_i)	d	fd	fd^2
61	5	−2	−10	20
64	18	−1	−18	18
67	42	0	0	0
70	27	1	27	27
73	8	2	16	32
รวม	100		15	97

$$\overline{X} = a + i\,d$$

$$= 67 + \left(3 \times \frac{15}{100}\right)$$

$$= 67 + \frac{9}{20} = 67 + 0.45$$

$$= 67.45 \approx 67.5$$

$$S^2 = i^2 \left[\frac{\Sigma fd^2 - \frac{(\Sigma fd)^2}{n}}{n-1} \right]$$

$$\therefore \quad S = i\sqrt{\frac{\Sigma fd^2 - \frac{(\Sigma fd)^2}{n}}{n-1}}$$

$$= 3\sqrt{\frac{97 - \frac{(15)^2}{100}}{100-1}}$$

$$= 3\sqrt{\frac{97 - \frac{225}{100}}{99}}$$

$$= 3\sqrt{\frac{97-2.25}{99}}$$

$$= 3\sqrt{\frac{94.75}{99}}$$

$$= 3\sqrt{0.9571} = 3 \times .978$$

$$\therefore \quad S = 2.934 \approx 3$$

1. H_0 : ข้อมูลมีการแจกแจงแบบปกติ

   H_a : ข้อมูลไม่มีการแจกแจงแบบปกติ

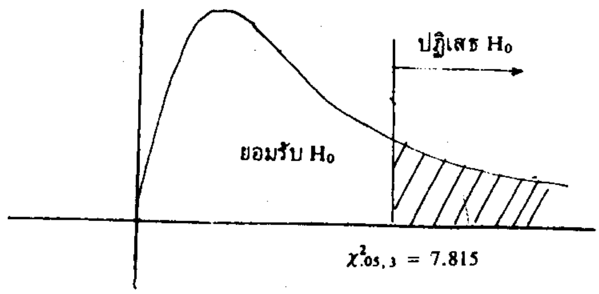
2. $\alpha = 0.05$

3. ตัวสถิติที่ใช้ทดสอบคือ

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$

$$df = 6-1-2 = 3$$

4. CR: ปฏิเสธ H_0 ถ้า $\chi_c^2 > \chi_{.05,3}^2$ คือ $\chi_c^2 > 7.815$



5. คำนวณค่า χ_c^2 ซึ่งจะต้องหาค่า E_i ก่อนโดยการแปลงค่า X ให้เป็นค่า Z โดยที่ $\bar{X} = 67.5$ และ $\sigma_X = 3$ และเปิดตาราง Z หาค่า P(X) ต่าง ๆ ได้ดังนี้

ค่า $X = 59.5$ ได้ค่า $Z = \frac{59.5-67.5}{3} = 2.66$

ค่า $X = 62.5$ ได้ค่า $Z = \frac{62.5-67.5}{3} = -1.66$

ทำโดยวิธีนี้จนครบจะได้ดังตาราง

น้ำหนัก	O_i	Classboundaries	P(X)	$E_i = 100 \times P(X)$
ต่ำกว่า 60	0 } 5	59.5	.0039	.39 } 4.85
60 – 62	5	62.5	.0446	4.46
63 – 65	18	65.5	.2061	20.61
66 – 68	42	68.5	.3747	37.47
69 – 71	27	71.5	.2789	27.89
72 – 74	8	74.5	.0819	8.19
มากกว่า 74	0	มากกว่า 74.5	.0099	0.99
			1.0000	1.00

$$\chi_c^2 = \frac{(5-4.85)^2}{4.85} + \frac{(18-20.61)^2}{20.61} + \frac{(4.2-37.47)^2}{37.47} + \frac{(27-27.89)^2}{27.89} + \frac{(8-8.19)^2}{8.19} + \frac{(0-.99)^2}{.99}$$

$$= \frac{(-.15)}{4.85} + \frac{(-2.61)^2}{20.61} + \frac{(4.33)^2}{37.47} + \frac{(-0.89)^2}{17.89} + \frac{(-.19)^2}{8.19} + \frac{(-.99)^2}{.99}$$

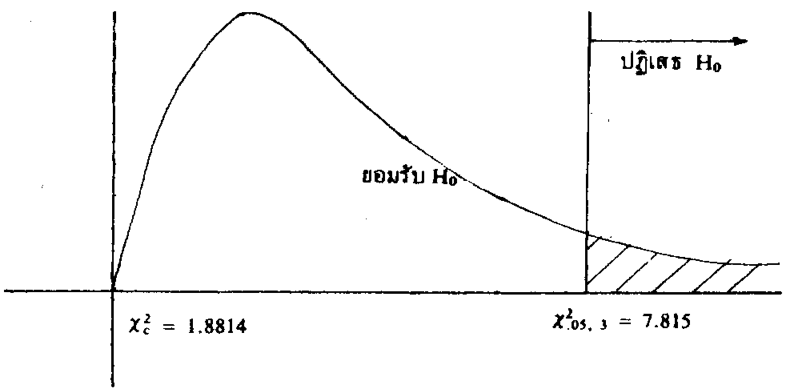
$$= \frac{0.0225}{3.85} + \frac{6.8121}{20.61} + \frac{19.6149}{37.47} + \frac{0.7921}{27.89} + \frac{0.0361}{8.19} + \frac{0.9801}{.99}$$

$$= .0046 + .3305 + .5235 + .0284 + .0044 + .99$$

$$= 1.8814$$

6. สรุปผล : χ_c^2 ตกอยู่นอก CR.

$\therefore$ เรายอมรับ H_0 ที่ว่าข้อมูลนี้มีการแจกแจงแบบปกติ



สรุปขั้นตอนในการคำนวณ

1. คำนวณหาค่าเฉลี่ยและความแปรปรวนของข้อมูลที่ได้
2. คำนวณความน่าจะเป็นของตัวแปร โดยแปลงค่า X ให้เป็นค่า Z และเปิดตารางหาความน่าจะเป็น
3. คำนวณจำนวนความถี่ที่คาดหมายตามทฤษฎีจาก $n \times P(x)$
4. หาค่า x_c^2

10.10.2 การทดสอบความเป็นอิสระต่อกันของลักษณะ 2 ลักษณะ

(Test for independence)

ในการศึกษาลักษณะ (characteristies) 2 ลักษณะในประชากร หรือกลุ่มที่เราสนใจว่าจะมีความสัมพันธ์กันหรือไม่นั้น เราจะจัดจำแนกข้อมูลในตัวอย่างขนาด n ออกเป็นกลุ่ม ๆ โดยใช้ 2 ลักษณะเป็นตัวจัดจำแนก ถ้าให้ลักษณะหนึ่งเป็น A โดย A แบ่งได้เป็น r ส่วนย่อย คือ $A_1, A_2, \ldots\ldots, A_r$ และอีกลักษณะหนึ่งเป็น B โดย B แบ่งได้เป็น C ส่วนย่อย คือ $B_1, B_2, \ldots\ldots B_c$

เมื่อทำการสุ่มตัวอย่างขนาด n เราจะได้ Contigency table ขนาด $r \times c$ ของค่าสังเกตหรือความถี่ (Oij) ดังนี้

	O_{ij}	ลักษณะ B				ผลรวมทาง row
		B_1	B_2	....	B_c	
ลักษณะ A	A_1	O_{11}	O_{12}	....	O_{1c}	R_1
	A_2	O_{21}	O_{22}	....	O_{2c}	R_2
	A_3	O_{31}	O_{32}	....	O_{3c}	R_3
	A_r	O_{r1}	O_{r2}	....	O_{rc}	R_r
ผลรวมทาง column		C_1	C_2	....	C_c	n

เมื่อ O_{ij} = ความถี่ที่มีลักษณะ A_i และ B_j

R_i = ผลรวมของ row ที่ i (คือความถี่ของ A_i นั่นเอง)

C_j = ผลรวมของ column ที่ j (คือความถี่ของ B_j นั่นเอง)

ในการทดสอบว่าลักษณะ A กับลักษณะ B เป็นอิสระต่อกัน หรือไม่นั้น เราต้องสร้างตาราง π_{ij} ดังนี้

π_{ij}		ลักษณะ B				Row Total
		B_1	B_2	....	B_c	
	A_1	π_{11}	π_{12}	....	π_{1c}	$\pi_{1\cdot}$
	A_2	π_{21}	Π_{22}	....	π_{2c}	$\pi_{2\cdot}$
ลักษณะ A	⋮	⋮	⋮	⋮	⋮	⋮
	A_r	π_{r1}	π_{r2}	....	π_{rc}	$\pi_{r\cdot}$
Column Total		$\pi_{.1}$	$\pi_{.2}$	....	$\pi_{.c}$	1

ถ้าให้ $\pi_{ij} = P(A_i \cap B_j)$

$\pi_{i.} = P(A_i)$

$\pi_{.j} = P(B_j)$

ดังนั้น ถ้าเราต้องการจะทดสอบว่าลักษณะ A และลักษณะ B เป็นอิสระกัน นั่นคือ เราจะต้องทดสอบว่า

$P(A_i \cap B_j) = P(A_i).P(B_j)$

หรือ $\pi_{ij} = \pi_{i.}\,\pi_{.j}$

จากตารางความถี่หรือค่าสังเกต ถ้าให้ $P_{i.}$ และ $P_{.j}$ เป็นตัวประมาณค่า $\pi_{i.}$ และ $\pi_{.j}$ ตามลำดับจะได้

$P_{i.} = \frac{R_i}{n}$ และ $P_{.j} = \frac{C_j}{n}$

ดังนั้น ถ้า $\pi_{ij} = \pi_{i.}\,\pi_{.j}$ จริงจะได้

$$\pi_{ij} = \frac{R_i}{n} \times \frac{C_j}{n}$$

และเราสามารถหา E_{ij} ได้จาก $n \times P_{ij}$

$$\therefore E_{ij} = n \times \frac{R_i}{n} \times \frac{C_j}{n} = \frac{R_i \times C_j}{n}$$

การหา df.

df. = จำนวน cells – 1 – จำนวนพารามิเตอร์ที่ต้องการประมาณค่า เนื่องจาก $\sum_{i=1}^{r} P_i = 1$ $\therefore$ เราประมาณค่า $P_{i.}$ เพียง $(r-1)$ ตัว ส่วนตัวสุดท้าย หาได้จากการลบผลบวกของ P_i ทั้ง $(r-1)$ ตัวออกจาก 1 ในทำนองเดียวกัน $\therefore \sum_{j=1}^{c} P_j = 1$ $\therefore$ เราประมาณค่า $P_{.j}$ เพียง $(c-1)$ ตัว

$\therefore$ จำนวนพารามิเตอร์ที่ต้องประมาณค่า $= (r-1)+(c-1)$

$= r+c-2$

และจำนวน cells $= r \times c$

$$\begin{aligned} \therefore df &= rc-1-(r+c-2) \\ &= rc-1-r-c+2 \\ &= rc-r-c+1 \\ &= r(c-1)-(c-1) \\ &= (r-1)(c-1) \end{aligned}$$

ขั้นตอนในการทดสอบความเป็นอิสระ

1. ตั้งสมมติฐาน

H_0 : ตัวแปรทั้ง 2 เป็นอิสระต่อกัน (หรือไม่มีความสัมพันธ์ต่อกัน)

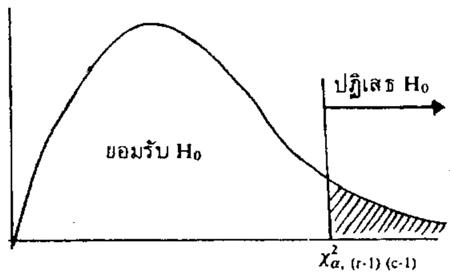
H_a : ตัวแปรทั้ง 2 ไม่เป็นอิสระต่อกัน (หรือมีความสัมพันธ์ต่อกัน)

2. กำหนด α และขนาดตัวอย่าง

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

4. CR : ปฏิเสธ H_0 ถ้า $\chi^2_{\alpha, (r-1)(c-1)}$



5. คำนวณหาค่า χ^2_c จาก

$$\chi^2_c = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

โดยที่ $E_{ij} = \frac{R_i \times C_j}{n}$

6. สรุปผล :

1. ถ้า χ^2_c ตกอยู่ใน CR. เราจะปฏิเสธ H_0 และยอมรับ H_a
2. ถ้า χ^2_c ตกอยู่นอก Cr. เราจะยอมรับ H_0

ตัวอย่างที่ 10.10.2.1

ต้องการจะทดสอบดูว่ามีความสัมพันธ์ระหว่างเพศกับการชอบสี หรือไม่ที่ $\alpha = 0.05$ สมมติว่าเลือกสุ่มตัวอย่างคนมา 100 คน ปรากฏผลดังนี้

	ชาย	หญิง	รวม
สีแดง	10	20	30
สีเขียว	20	10	30
สีเหลือง	30	10	40
รวม	60	40	100

1. ตั้งสมมติฐาน

H_0 : เพศและการชอบสีเป็นอิสระต่อกัน (หรือไม่มีความสัมพันธ์ต่อกัน)

H_a : เพศและการชอบสีไม่เป็นอิสระต่อกัน (หรือมีความสัมพันธ์กัน)

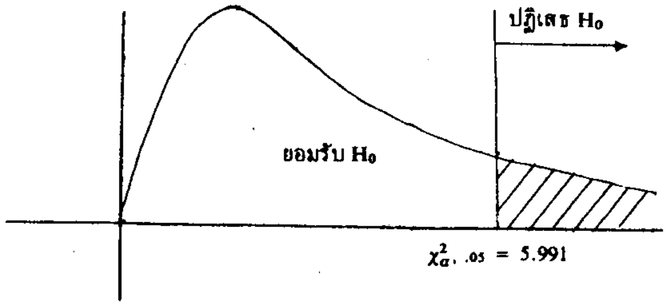
2. $\alpha = 0.05, n = 100$

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

4. CR : ปฏิเสธ H_0 ถ้า $\chi_c^2 > \chi^2_{(3-1)(2-1), .05}$

คือ $\chi_c^2 > 5.991$



5. คำนวณค่า χ_c^2 ซึ่งต้องหาค่า E_{ij} ก่อนจาก

$E_{ij} = \frac{R_i \times C_j}{n}$ ซึ่งจะได้ดังตารางต่อไปนี้

O_{ij}	10	20	30	20	10	10
E_{ij}	$\frac{30\times 60}{100} = 18$	$\frac{30\times 60}{100} = 18$	$\frac{40\times 60}{100} = 24$	$\frac{30\times 40}{100} = 12$	$\frac{30\times 40}{100} = 12$	$\frac{40\times 40}{100} = 16$

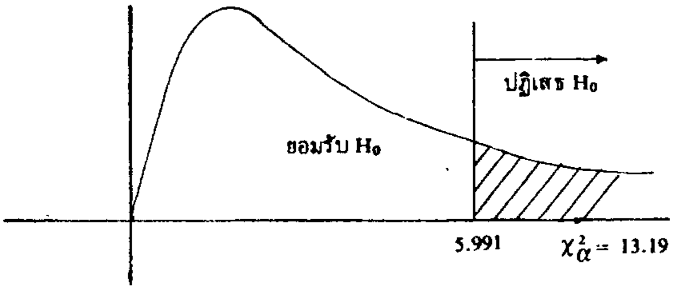
$$\therefore \chi_c^2 = \frac{(10-18)^2}{18}+\frac{(20-18)^2}{18}+\frac{(30-24)^2}{24}+\frac{(20-12)^2}{12}+\frac{(10-12)^2}{12}+\frac{(10-16)^2}{16}$$

$$= \frac{(-8)^2}{18}+\frac{(2)^2}{18}+\frac{(6)^2}{24}+\frac{(8)^2}{12}+\frac{(-2)^2}{12}+\frac{(-6)^2}{16}$$

$$= \frac{64}{18}+\frac{4}{18}+\frac{36}{24}+\frac{64}{12}+\frac{4}{12}+\frac{36}{16}$$

$= 3.56 + 0.22 + 1.50 + 5.33 + 0.33 + 2.25$

$= 13.19$



6. สรุปผล $\because \chi^2_c$ ตกอยู่ใน CR.

$\therefore$ เราปฏิเสธ H_0 ที่ว่าเพศและความชอบสีเป็นอิสระต่อกัน

ดังนั้น สรุปได้ว่าเพศและความชอบสีมีความสัมพันธ์กัน

10.10.3 การทดสอบความเป็นเอกภาพ (Test for Homogeniety)

เป็นการทดสอบดูว่าประชากรหรือกลุ่มต่าง ๆ เป็นเอกภาพ หรือเหมือนกันหรือไม่โดยการสุ่มตัวอย่างขนาด $R_1, R_2, \ldots\ldots, R_r$ จาก r ประชากรซึ่งคือ $A_1, A_2, \ldots\ldots, A_r$ แล้วจัดจำแนก R_i ตามลักษณะของ B ซึ่งมีอยู่ C ลักษณะย่อย คือ $B_1, B_2, \ldots\ldots, B_c$ เราจะได้ตาราง $r \times c$ ดังนี้

		ลักษณะ B				ผลรวมทาง
	O_{ij}	B_1	B_2	....	B_c	row
ประชากร	A_1	O_{11}	O_{12}	....	O_{1c}	R_1
	A_2	O_{21}	O_{22}	....	O_{2c}	R_2
	⋮	⋮	⋮		⋮	⋮
	A_r	O_{r1}	O_{r2}	....	O_{rc}	R_r
ผลรวมของ column		C_1	C_2	....	C_c	n

ขั้นตอนในการทดสอบ

1. ตั้งสมมติฐาน

H_0 : ประชากรต่าง ๆ เป็นเอกภาพกัน

H_a : ประชากรต่าง ๆ ไม่เป็นเอกภาพกัน

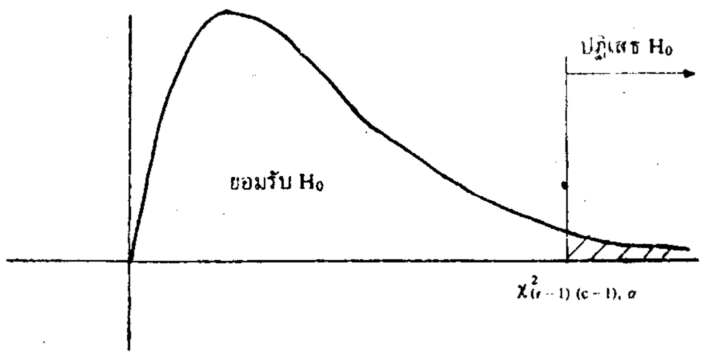
2. α, n

3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{\text{all } i} \sum_{\text{all } j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$$df = (r-1)(c-1)$$

4. CR : ปฏิเสธ H_0 ถ้า $\chi^2_c > \chi^2_{(r-1)(c-1),\alpha}$



5. คำนวณค่า χ^2_c จาก

$$\chi^2_c = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

โดยที่ $E_{ij} = \dfrac{R_i \times C_j}{n}$

6. สรุปผล 1. ถ้า χ^2_c ตกอยู่ใน CR. เราจะปฏิเสธ H_0 และยอมรับ H_a

2. ถ้า χ^2_c ตกอยู่นอก CR. เราจะยอมรับ H_0

ตัวอย่างที่ 10.10.3.1

ตัวแทนจำหน่ายน้ำยาปรับผ้านุ่มชนิดหนึ่งกล่าวว่าความนิยมน้ำยาปรับผ้านุ่มชนิดนี้ในแต่ละภาคมีพอ ๆ กัน เพื่อยืนยันคำกล่าวนี้ เขาได้สุ่มตัวอย่างมาจาก 4 ภาค ได้ผลดังนี้

ภาค	จำนวนคนที่นิยม	จำนวนคนที่ไม่นิยม	รวม
เหนือ	120	80	200
กลาง	200	50	250
ตะวันออกเฉียงเหนือ	200	100	300
ใต้	180	70	250
รวม	700	300	1,000

จงทดสอบที่ $\alpha = .01$ ว่าคำกล่าวของตัวแทนจำหน่ายน้ำยาปรับผ้านุ่มชนิดนี้ว่าเชื่อถือได้หรือไม่

1. **ตั้งสมมติฐาน**

H_0 : ความนิยมน้ำยาปรับผ้านุ่มชนิดนี้ในแต่ละภาคมีพอ ๆ กัน

H_a : H_0 ไม่จริง

2. $\alpha = 0.01$

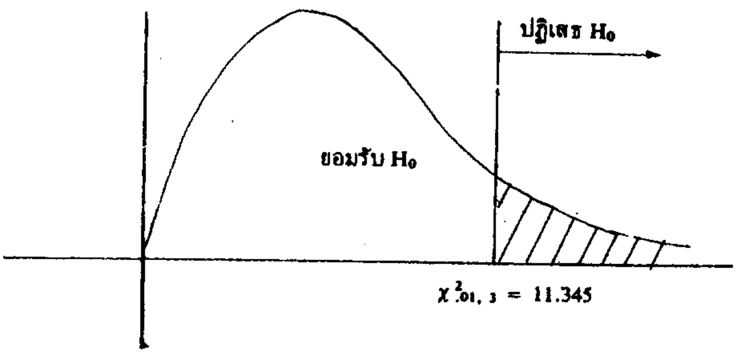
3. ตัวสถิติที่ใช้ทดสอบ คือ

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$df = (r-1)(c-1) = (4-1)(2-1) = 3 \times 1 = 3$

4. CR : ปฏิเสธ H_0 ถ้า $\chi^2_{.01,\,3}$

คือ $\chi^2_c > 11.345$



5. คำนวณหาค่า χ^2_c โดยหา E_{ij} ก่อนจาก $E_{ij} = \frac{R_i \times C_j}{n}$ ซึ่งจะได้ดังนี้

ตาราง E_{ij} สำหรับ O_{ij} ใน Colum n ที่ 1 (คนที่นิยม)

O_{ij}	120	200	200	180
E_{ij}	$\frac{700 \times 200}{1,000} = 140$	$\frac{700 \times 250}{1,000} = 175$	$\frac{700 \times 300}{1,000} = 210$	$\frac{700 \times 250}{1,000} = 175$

ตาราง E_{ij} สำหรับ O_{ij} ใน Column ที่ 2 (คนที่ไม่นิยม)

O_{ij}	80	50	100	70
E_{ij}	$\frac{300 \times 200}{1,000} = 60$	$\frac{300 \times 250}{1,000} = 75$	$\frac{300 \times 300}{1,000} = 90$	$\frac{300 \times 250}{1,000} = 75$

$$\therefore \chi_c^2 = \frac{(120-140)^2}{140} + \frac{(200-175)^+}{175} + \frac{(200-210)^2}{210} + \frac{(180-175)^2}{175}$$

$$+ \frac{(80-60)^2}{60} + \frac{(5075)^2}{75} + \frac{(100-90)^2}{90} + \frac{(70-75)^2}{75}$$

$$= \frac{(-20)^2}{140} + \frac{(25)^2}{175} + \frac{(-10)^2}{210} + \frac{(5)^2}{175} + \frac{(20)^2}{60} + \frac{(-25)^2}{75} + \frac{(10)^2}{90} + \frac{(-5)^2}{75}$$

$$= \frac{400}{140} + \frac{625}{175} + \frac{100}{210} + \frac{25}{175} + \frac{400}{.60} + \frac{625}{75} + \frac{100}{90} + \frac{25}{75}$$

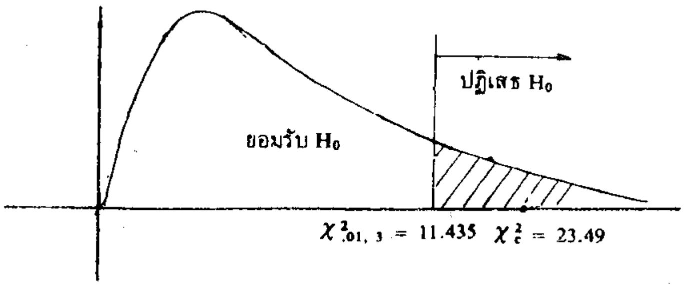
$$= 2.86 + 3.57 + 0.48 + 0.14 + 6.67 + 8.33 + 1.11 + 0.33$$

$$= 23.49$$

6. สรุปผล $\because \chi_c^2 > 11.345$ (ตกอยู่ใน CR.)

$\therefore$ เราปฏิเสธ H_0 และยอมรับ H_a

แสดงว่าความนิยมน้ำยาปรับผ้านุ่มชนิดนี้ในแต่ละภาคไม่เท่ากัน



แบบฝึกหัด

1. จงนิยามความหมายของค่าต่อไปนี้

   ก) สมมติฐาน

   ข) ระดับนัยสำคัญ

   ค. ความคลาดเคลื่อนแบบ 1

   ง. ความคลาดเคลื่อนแบบ 2

   จ. เขตวิกฤต

2. สมมติฐานหลัก และสมมติฐานรองคืออะไร

3. การทดสอบแบบทางเดียวและการทดสอบแบบสองทางคืออะไร จงอธิบาย พร้อมทั้งยกตัวอย่างประกอบ

4. จงบอกลำดับขั้นในการทดสอบสมมติฐาน

5. การหาเขตวิกฤตจะต้องอาศัยอะไรบ้าง

6. ตัวสถิติทดสอบคืออะไร

7. ความหมายของคำพูดที่ว่า "เราปฏิเสธสมมติฐานหลัก" คืออะไร

8. ถ้า H_o ถูกต้องและเรายอมรับ H_o หมายความถึงอะไร

9. จงพิจารณาสมมติฐานต่อไปนี้ว่าเป็นการทดสอบแบบทางเดียวหรือแบบสองทาง

   ก) $H_1 : \mu \neq 4.10$

   ข) $H_1 : \mu < 4.10$

   ค) $H_1 : \mu > 81$

   ง) $H_1 ; \mu > .66$

   จ) $H_1 : \mu \neq 1.90$

   ฉ) $H_1 : \mu < 3$

10. ในแต่ละหัวข้อต่อไปนี้จงบอกว่าข้อใดเป็นสมมติฐานทางเดียว ข้อใดเป็นสมมติฐานทางเดียว ข้อใดเป็นสมมติฐานสองทางพร้อมทั้งเขียนรูปโค้งปกติและแลเงาส่วนที่เป็นเขตวิกฤตด้วย

ก) $H_o : \mu = 10, \ H_1 : \mu \neq 10, \ \alpha = .02$

ข) $H_o : \mu = .037, \ H_1 : \mu > .037, \ \alpha = .05$

ค) $H_o : \mu = 3.2, \ H_1 : \mu < 3.2, \ \alpha = .01$

ง) $H_o : \mu = 17.45, \ H_1 : \mu > 17.45, \ \alpha = .05$

จ) $H_o : \mu_1 = \mu_2, \ H_1 : \mu_1 \neq \mu_2. \ \alpha = .02$

ฉ) $H_o : \pi = .05, \ H_1 : \pi \neq .05, \ \alpha = .01$

ช) $H_o : \pi_1 - \pi_2 = 0, \ H_1 \ \pi_1 - \pi_2 > 0, \ \alpha = .05$

11. จงตั้งสมมติฐานหลักและสมมติฐานรองจากข้อความต่อไปนี้

ก) ค่าใช้จ่ายในการซื้อรองเท้าต่อปี โดยเฉลี่ยของนักศึกษาหญิงในมหาวิทยาลัยแห่งหนึ่งมากกว่า 500 บาท

ข) นักศึกษาคณะรัฐศาสตร์ มหาวิทยาลัยรามคำแหง เห็นด้วยกับนโยบายของรัฐบาลในการแก้ปัญหาเศรษฐกิจมากกว่าครึ่งหนึ่ง

ค) อาหารของบริษัท ก. ให้น้ำหนักของหมูมากกว่าอาหารของบริษัท ข.

ง) ความนิยมของผงซักฟอกยี่ห้อบรีสมีพอ ๆ กันทุก ๆ ภาคของประเทศไทย

จ) จำนวนบุตรมีความสัมพันธ์กับระดับรายได้ของครอบครัว

ฉ) การสูบบุหรี่มีความสัมพันธ์กับโรคปอด

ช) บริษัทผลิตหลอดไฟแห่งหนึ่ง โฆษณาว่าหลอดไฟที่ผลิตได้จะมีอายุการใช้งานโดยเฉลี่ยอย่างน้อยที่สุด 800 ชั่วโมง

ซ) เจ้าหน้าที่แผนกลงทะเบียนของมหาวิทยาลัยแห่งหนึ่ง กล่าวว่า การลงทะเบียนเรียนภาคหนึ่งใช้เวลา โดยเฉลี่ยไม่น้อยกว่า 50 นาที

ฌ) บริษัทผลิตรถยนต์โฆษณาว่า 35% ของรถยนต์ทั้งหมดที่สร้างโดยบริษัท ตั้งแต่ปี 1954 ยังมีสภาพดีอยู่

12. จากตัวอย่างนักเรียนชั้นประถมปีที่ 1 ของโรงเรียนเทศบาลแห่งหนึ่งจำนวน 10 คน พบว่า ได้ค่าขนมไปโรงเรียนวันละ 3,5,7,6,4,5,6,4,3,7, บาท ตามลำดับ จงหาว่าข้อมูลนี้สนับสนุนสมมติฐานที่ว่านักเรียนโรงเรียนนี้ได้ค่าขนมเฉลี่ยวันละ 4 บาท หรือไม่ โดยใช้ $\alpha = 0.05$

13. ผลการสำรวจการขาดงานของลูกจ้าง 30 คน ในโรงงานอุตสาหกรรมแห่งหนึ่ง พบว่ามีความแปรปรวนเท่ากับ 90 แต่จากประสบการณ์เชื่อว่าความแปรปรวนควรจะเท่ากับ 102 อยากทราบว่าจะใช้ข้อเท็จจริงจากตัวอย่างสรุปได้ไหมว่า ความแปรปรวนของการขาดงานของประชากรเท่ากับ 102 ถ้าถือว่า การขาดงานของลูกจ้างมีการแจกแจงแบบปกติ โดยใช้ $\alpha = 0.01$

14. ในกระบวนการผลิตเครื่องโทรทัศน์ยี่ห้องหนึ่งจะผลิตเครื่องเสียออกมา 20% เมื่อปรับปรุงการผลิตใหม่แล้วสุ่มตัวอย่างออกมาทดสอบ 100 ชิ้น ปรากฏว่า พบของเสีย 16 ชิ้น จะเชื่อได้ไหมว่าการปรับปรุงการผลิตใหม่ทำให้ลดการผลิตของเครื่องเสียลงกำหนดให้ $\alpha = 0.05$

15. บริษัทผลิตรถยนต์โฆษณาว่า 35% ของรถยนต์ทั้งหมดที่สร้างโดยบริษัทตั้งแต่ปี 1954 ยังมีสภาพดีอยู่ ถ้าสุ่มตัวอย่างรถยนต์ดังกล่าวมา 800 คัน ปรากฏว่า ยังมีสภาพดีอยู่ ถึง 257 คัน อยากทราบว่าการโฆษณานี้เป็นที่ยอมรับ ณ.ระดับนัยสำคัญ 0.01 หรือไม่

16, จงเปิดตารางหาค่า

$t_{.015,9}$, $t_{.05,9}$, $z_{.01}$, $z_{.05}$, $z_{.99}$, $x^2_{.01,9}$, $x^2_{.99,9}$

17. ผู้ผลิตรถยนต์รุ่นหนึ่งอ้างว่ารถยนต์ที่เขาผลิตขึ้นมารุ่นนี้จะวิ่งได้ระทาง โดยเฉลี่ย 31 ไมล์ต่อน้ำมันธรรมดา 1 ถัง ถ้าสุ่มรถยนต์รุ่นนี้มี 9 คัน โดยที่แต่ละคันให้ทดลองวิ่ง โดยใช้น้ำมันธรรมดา 1 ถัง ปรากฏว่าวัดระยะทางโดยเฉลี่ยต่อน้ำมัน 1 ถัง ได้เท่ากับ 29.43 ไมล์ และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3 ไมล์ ท่านจะสรุปผลอย่างไรเกี่ยวกับคำกล่าวอ้างของผู้ผลิตรถยนต์รุ่นนี้ โดยใช้ $\alpha = .05$

18. สุ่มตัวอย่างเนื้อวัวที่บรรจุในถุงพลาสติกที่นำมาขายในห้างสรรพสินค้าแห่งหนึ่ง 36 ถุง มาตรวจสอบดูเปอร์เซนต์ของไขมัน ซึ่งที่ถุงบรรจุเขียนไว้ว่า "มีไขมันไม่เกิน 25%" จากการสุ่มตัวอย่าง ปรากฏว่า ไขมันโดยเฉลี่ยวัดได้ 0.265 และส่วนเบี่ยงเบนมาตรฐาน 0.030 ท่านจะสรุปผลอย่างไรถ้า $\alpha = .05$

19. สุ่มตัวอย่างนักศึกษามา 36 จากวิทยาลัยแห่งหนึ่งปรากฏว่า การลงทะเบียนภาคฤดูร้อน ใช้เวลาเฉลี่ยแล้ว 45 นาที และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 นาที จงทดสอบสมมติฐานที่ว่าการลงทะเบียนเรียนโดยเฉลี่ยไม่น้อยกว่า 50 นาที โดยใช้ $\alpha = .05$

20. ในการสอบวิชาสถิติวิชาหนึ่ง มีนักศึกษาหญิงเข้าสอบ 50 คน นักศึกษาชายเข้าสอบ 75 คน นักศึกษาหญิงทำคะแนนเฉลี่ยได้ 76 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 6 ส่วนนักศึกษาชายทำคะแนนเฉลี่ยได้ 82 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 8 จงสรุปว่านักศึกษาชายและนักศึกษาหญิงทำคะแนนเฉลี่ย แตกต่างกันหรือไม่ โดยใช้ $\alpha = 0.05$

21. สุ่มตัวอย่างหลอดไฟจากบริษัท ก. มา 80 หลอด หาอายุการใช้งานโดยเฉลี่ยของหลอดไฟได้เท่ากับ 1,258 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐาน 94 ชั่วโมง และสุ่มตัวอย่างหลอดไฟจากบริษัท ข. มา 60 หลอด หาอายุการใช้งานโดยเฉลี่ยของหลอดไฟได้เท่ากับ 1,029 ชั่วโมง ส่วนเบี่ยงเบนมาตรฐาน 68 ชั่วโมง เนื่องจากราคาหลอดไฟ ของบริษัท ก. แพงกว่า บริษัท ข. จึงตั้งเกณฑ์ไว้ว่า

เราจะตัดสินใจซื้อหลอดไฟของบริษัท ก. ถ้าหลอดไฟของบริษัท ก. มีอายุการใช้งาน โดยเฉลี่ยมากกว่าบริษัท ข. 200 ชั่วโมงขึ้นไป ท่านจะตัดสินใจว่าจะซื้อของบริษัทใด ถ้า $\alpha = .01$

22. ในการศึกษาเกี่ยวกับวิธีการสอบ 2 วิธี เพื่อเปรียบเทียบดูว่าวิธีที่ 1 มีนักเรียน 11 คน สอบได้คะแนนเฉลี่ย 85 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 4.7 และห้องที่ใช้การสอน วิธีที่ 2 มีนักเรียน 17 คน เมื่อสอบโดยใช้ข้อสอบชุดเดียวกับกลุ่มแรก ได้คะแนนเฉลี่ย 79 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 6.1 ถ้าให้คะแนนสอบของนักเรียนมีการแจกแจงใกล้โค้งปกติ และมีความแปรปรวนเท่ากัน จงทดสอบว่าการสอนวิธีที่ 1 จะช่วยทำให้คำแนนเสอบเฉลี่ยของนักเรียนสูงขึ้น โดยใช้ $\alpha = .01$

23. กำหนดข้อมูลจากตัวอย่างให้ดังนี้

จังหวัดอยุธยา	จังหวัดชัยนาท
$\bar{x}_1 = 23$ ปี	$\bar{x}_2 = 21$ ปี
$s_1 = 4$ ปี	$s_2 = 5$ ปี
$n_1 = 100$ ปี	$n_2 = 50$ ปี

มีความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่างทั้ง 2 อย่างมีนัยสำคัญหรือไม่ $\alpha = .01$

24. ในการเปรียบเทียบดูอายุการใช้งานโดยเฉลี่ยของแบตเตอรี่ 2 กลุ่ม โดยสุ่มตัวอย่างมากลุ่มละ 100 อัน ปรากฏว่ากลุ่มแรกได้ค่าเฉลี่ยของอายุการใช้งานเท่ากับ 47 ชั่วโมง และส่วนเบี่ยงเบนมาตรฐาน 4 ชั่วโมง กลุ่มที่ 2 ได้ค่าเฉลี่ย 48 ชั่วโมง และส่วนเบี่ยงเบนมาตรฐาน 3 ชั่วโมง อยากทราบว่ามีความแตกแต่างระหว่างค่าเฉลี่ยของตัวอย่างทั้ง 2 กลุ่มนี้หรือไม่ ถ้า $\alpha = 0.05$

25. ในการทดสอบความสามารถในการอ่านของนักเรียนชั้นประถมของโรงเรียน 2 โรงเรียน โดยโรงเรียนที่ 1 สุ่มนักเรียนมา 12 คน และโรงเรียนที่ 2 สุ่มนักเรียนมา 10 คน ได้ผลดังนี้

โรงเรียนที่ 1	โรงเรียนที่ 2
$\bar{x}_1 = 74$	$\bar{x}_2 = 70$
$s_1 = 8$	$s_2 = 10$

อยากทราบว่ามีความแตกต่างระหว่างค่าเฉลี่ยของตัวอย่างทั้ง 2 หรือไม่ ถ้า $\alpha = .05$

23. พนักงานขาย 2 คน ได้บันทึกสถิติการขาย เครื่องสำอางยี่ห้องหนึ่ง ปรากฏว่าผลการขาย 12 ครั้ง ของพนักงานคนแรกขายได้จำนวนเฉลี่ย 125 ชุด และจากการขาย 8 ครั้ง ของพนักงานคนที่ 2 ได้ 105 ชุด โดยเฉลี่ย จากประสบการณ์ที่ผ่านมาชี้ให้เห็นว่าผลการขายของพนักงานทั้ง 2 นี้ต่างก็มีความแปรปรวนเท่ากับ 270 จงทดสอบดูว่าการขายของพนักงานทั้ง 2 มีความแตกต่างกันหรือไม่โดยใช้ $\alpha = .05$

27. ผู้จัดการขายสินค้าอย่างหนึ่งตั้งเกณฑ์ในการรับของที่ส่งมาจากโรงงานไว้ว่า ถ้าสุ่มตัวอย่างมา 25 ชิ้นแล้ว ส่วนเบี่ยงเบนมาตรฐานของความยาวของตัวอย่างจะต้องไม่มากกว่า 0.01 มม. โรงงานได้ส่งสินค้ามาให้ เมื่อสุ่มตัวอย่างแล้ววัดความยาวดู ปรากฏว่า ได้ความแปรปรวนเป็น 0.0002 ผู้จัดการควรจะรับสินค้านั้นหรือไม่ ที่ $\alpha = .05$

28. การทดสอบ goodness of fit คืออะไร และใช้ตัวสถิติทดสอบตัวไหนในการทดสอบ goodness of fit

29. ในการศึกษาความสัมพันธ์ระหว่างการสูบบุหรี่และการเป็นโรคปอด โดยอาศัยตัวอย่าง 200 คน ได้ข้อมูลดังนี้

	เป็นโรคปอด	ไม่เป็นโรคปอด	รวม
สูบบุหรี่	40	80	120
ไม่สูบบุหรี่	10	70	80
รวม	50	150	200

จากข้อมูลดังกล่าว จะสรุปได้ไหมว่า การเป็นโรคขึ้นอยู่กับการสูบบุหรี่ ที่ $\alpha = 0.05$

30. สถานบริการลดความอ้วนแห่งหนึ่งโฆษณาว่า วิธีการลดความอ้วนแบบใหม่จะช่วยลดน้ำหนักได้ภายใน 2 อาทิตย์ ได้มีการทดลองใช้วิธีการแบบใหม่กับคน 7 คน โดยชั่งน้ำหนักก่อนและหลัง การใช้วิธีลดความอ้วนแบบใหม่ (2 อาทิตย์) ผลที่ได้มีดังนี้

คนที่	1	2	3	4	5	6	7
น้ำหนักก่อน	129	133	136	152	141	138	125
น้ำหนักหลัง	130	121	128	137	129	132	120

ข้อมูลจากตัวอย่างนี้สนับสนุนโฆษณาได้หรือไม่ที่ $\alpha = 0.01$

31. ผู้ผลิตสินค้าต้องการทราบว่าความนิยมสินค้าแบบต่าง ๆ มี ความสัมพันธ์กับเพศของลูกค้าหรือไม่ จากการสุ่มตัวอย่างผู้ซื้อสินค้าของบริษัทมา 1,000 รายได้ข้อมูลดังนี้

เพศ	แบบที่ซื้อ			รวม
	แบบที่ 1	แบบที่ 2	แบบที่ 3	
ชาย	100	100	200	400
หญิง	300	150	150	600
รวม	400	250	350	1,000

จงใช้ α = 0.05 ทดสอบว่าเพศและความนิยมไม่มีความสัมพันธ์กัน

32. บริษัทแห่งหนึ่งต้องการศึกษาดูว่าการแต่งงานเป็นสาเหตุที่ทำให้พนักงานหญิงของบริษัทขาดงานบ่อยหรือไม่ จึงสุ่มคนงานที่เคยขาดงานมา 400 คน ผลปรากฏดังนี้

	ขาดบ่อย	ขาดนาน ๆ ครั้ง	รวม
แต่งงาน	84	96	180
ไม่แต่งงาน	62	158	220
รวม	146	254	400

จงทดสอบที่ α = .05 ว่าสถานภาพสมรสกับการขาดงานเป็นอิสระต่อกันหรือไม่

33. ในการศึกษาเพื่อว่าอัตราการสูบบุหรี่ของนักศึกษารวมค่าแหงปีต่าง ๆ แตกต่างกันหรือไม่ ได้สุ่มนักศึกษาปีที่ 1 มา 50 คน ปีที่ 2 มา 40 คน ปีที่ 3 มา 60 คน ปีที่ 4 มา 50 คน จัดจำแนกตามนิสัยการสูบบุหรี่ได้ผลดังนี้

	นิสัยการสูบบุหรี่			รวม
	น้อยมาก	ปานกลาง	มาก	
ชั้นปีที่ 1	21	12	17	50
ชั้นปีที่ 2	13	8	19	40
ชั้นปีที่ 3	13	18	29	60
ชั้นปีที่ 4	3	22	25	50

จงทดสอบที่ α = 0.05 ว่าอัตรการสูบบุหรี่ของนักศึกษารามคำแหงทั้ง 4 ปี ไม่ต่างกัน

34. จากการเพาะเมล็ดพืชตัวอย่าง 400 เมล็ดพบว่ามี 346 เมล็ดที่งอก จงทดสอบว่าจำนวนเมล็ดพืชงอกมากกว่า 80% โดยใช้ $\alpha = .02$

35 จากการสุ่มตัวอย่างคน 900 คน ในจังหวัดลำปางซึ่งมีเพศชาย 467 คน เพศหญิง 433 คน พบว่าเพศชาย 8 คน เป็นตาบอดสี และเพศหญิง 2 คน เป็นตาบอดสี จากข้อมูลนี้จะสรุปได้ไหมว่าอัตราส่วนของการเป็นตาบอดสีของเพศหญิงน้อยกว่าเพชชาย 0.5% โดยใช้ $\alpha = 0.05$

36. จากข้อมูลต่อไปนี้

	รายได้ต่อเดือน		
	น้อยกว่า 5,000 บาท	5,000-10,000 บาท	มากกว่า 10,000 บาท
มีบ้านขอตนเอง	5%	35%	10%
เช่าบ้านอยู่	15%	25%	10%

สุ่มตัวอย่างครอบครัวมา 400 ครอบครัว จงทดสอบว่าการมีบ้านเป็นของตนเองขึ้นอยู่กับระดับรายได้ที่ $\alpha = .01$