บทที่5
การประมาณค่าและการทดสอบสมมติฐาน
(Estimation and Testing Hypothesis)

5.1 คำนำ

การประมาณค่า และ การทดสอบสมมติฐาน เป็นสถิติในส่วนที่เรียกว่า สถิติเชิงอ้างอิง คือ มีการศึกษาจากตัวอย่างเพื่อนำสรุปภาพรวมของประชากร โดยทั่วไปการศึกษาจากทุกหน่วยของประชากรที่ให้ข้อมูล ไม่สามารถทำได้ เนื่องจากข้อจำกัดของเวลา และ ค่าใช้จ่าย ทำให้ไม่สามารถศึกษาทั้งประชากรได้ ในบางครั้ง หากสามารถรวบรวมข้อมูลจากทุกหน่วยของประชากรได้ ข้อสนเทศที่ได้ไม่สามารถนำมาใช้ประโยชน์ ให้ทันต่อเหตุการณ์ได้ เช่น การศึกษาความต้องการเทคโนโลยี เป็นต้น

การศึกษาจากตัวอย่าง หรือ บางส่วนของประชากร ทำให้ได้ข้อสนเทศรวดเร็ว ประหยัดเวลา และ ค่าใช้จ่าย ดังนั้นการเลือกตัวอย่างเป็นสิ่งที่จำเป็นและสำคัญ สำหรับสถิติในส่วนของ การอ้างอิง สำหรับการเลือกตัวอย่างจะเลือกโดยใช้ความน่าจะเป็น โดยกำหนดให้ทุกหน่วยที่ให้ข้อมูลในประชากร มีโอกาสถูกเลือกเท่าๆกัน จากนั้นนำหน่วยตัวอย่างที่เลือกได้ หาค่าสถิติเพื่อ ทำการประมาณค่าของประชากร หรือ ทำการทดสอบสมมติฐาน

การประมาณค่าและการทดสอบสมมติฐาน จะต้องอาศัยการแจกแจงของตัวอย่างที่ได้ศึกษาจากบทที่3 เพื่อบอกว่า การประมาณค่าและการทดสอบสมมติฐานมีความถูกต้องเท่าใด

5.2 การประมาณค่า

การประมาณค่า คือ การประมาณสิ่งที่เราสนใจในประชากร โดยที่เราไม่ทราบว่าค่าในประชากร(ค่าพารามิเตอร์) มีค่าเป็นเท่าใด การประมาณสามารถประมาณโดยใช้ค่าเพียงค่าเดียวในการประมาณซึ่งเรียกว่าการประมาณแบบจุด หรือประมาณแบบช่วงซึ่งได้จากการสร้างช่วงมาหนึ่งช่วง การประมาณเป็นการประมาณค่าพารามิเตอร์ เช่น ค่าเฉลี่ย ค่าสัดส่วน และค่าส่วนเบี่ยงเบนมาตรฐาน เป็นต้น

5.2.1 ชนิดของการประมาณค่า

การประมาณค่าสำหรับค่าพารามิเตอร์ มี 2 ชนิดคือ

- ***การประมาณค่าแบบจุด***

การประมาณค่าแบบจุด คือ การใช้ค่าจากตัวอย่าง เพียงค่าเดียว ประมาณ ค่าพารามิเตอร์ของประชากรเช่น

ใช้ค่าเฉลี่ยตัวอย่างแทนด้วย $\overline{X}$ ประมาณ ค่าเฉลี่ยของประชากร แทนด้วย μ

ใช้ค่าสัดส่วนตัวอย่างแทนด้วย p ประมาณ ค่าสัดส่วนของประชากร แทนด้วย π

ใช้ค่าแปรปรวนตัวอย่างแทนด้วย s^2ประมาณค่าแปรปรวนของประชากรแทนด้วย σ^2

ตัวอย่างที่ 1

สุ่มตัวอย่างนักศึกษา 10 คนเพื่อทำการประมาณค่าใช้จ่าย เฉลี่ยต่อวันของนักศึกษาคณะบัญชีได้ ข้อมูลดังนี้

120 145 185 100 95 65 120 100 80 90

จากข้อมูลดังกล่าว โจทย์ต้องการประมาค่าเฉลี่ยคือ μ

สามารถใช้ค่าเฉลี่ยตัวอย่างคือ $\overline{X}$ ทำการประมาณค่า μ

$$n = 10 \quad \overline{X} = \frac{\sum_{i=1}^{10} X_i}{n} = \frac{120 + 145 + ... + 90}{10} = 110$$

$\therefore$ ด้วยการประมาณค่าแบบจุด สามารถประมาณค่าเฉลี่ยของประชากรเกี่ยวกับค่าใช้จ่ายเฉลี่ยต่อวันของนักศึกษาคณะบัญชีเป็น 110 บาท

ตัวอย่างที่2

ฝ่ายวางแผนของมหาวิทยาลัยต้องการ ประมาณจำนวนนักศึกษาที่มาจากต่างจังหวัดว่ามีร้อยละเท่าใด จึงทำการสุ่มตัวอย่างนักศึกษาจำนวน 50 คน มีจำนวนนักศึกษาที่มาจากต่างจังหวัด 15 คน จงประมาณค่าสัดส่วนของนักศึกษาที่มาจากต่างจังหวัดแบบจุด

จากโจทย์ ต้องการประมาณค่า π

สามารถใช้ค่าสัดส่วนตัวอย่างคือ p ทำการประมาณค่า π $n = 50$

ถ้าให้ X แทนนักศึกษาที่มาจากต่างจังหวัด

$$X = \begin{cases} 1, & \text{นักศึกษาที่มาจากต่างจังหวัด} \\ 0, & \text{นักศึกษาที่มาจากกรุงเทพ} \end{cases}$$

$$p = \overline{X} = \frac{\sum_{i=1}^{50} X_i}{n} = \frac{15}{50} = 0.3$$

$\therefore$ ด้วยการประมาณค่าแบบจุด สามารถประมาณค่าสัดส่วนของประชากรเกี่ยวกับค่าสัดส่วนของนักศึกษาที่มาจากต่างจังหวัดมีค่าเป็น 0.3

ในการประมาณโดยใช้ค่าเพียงค่าเดียวทำการประมาณนั้น หากค่าจากตัวอย่างแตกต่างจากประชากร การประมาณแบบจุดโดยค่าสถิตินั้น จะทำการประมาณค่าพารามิเตอร์ได้ไม่ถูกต้อง เช่น

สมมติค่าใช้จ่ายเฉลี่ยต่อวันของนักศึกษามหาวิทยาลัยหอการค้าไทยมีค่า 120 บาท (μ =120) สุ่มตัวอย่างนักศึกษา 10 คนได้ค่าใช้จ่ายเฉลี่ยต่อวัน 110 บาท($\overline{X}$=110)การนำค่า 110 บาทไปประมาณค่า 120 จะพบว่าค่าประมาณที่ได้ทำการประมาณไม่ถูกต้อง และโอกาสของการประมาณด้วยวิธีนี้ ถูกต้องยากเนื่องจากใช้ค่าเพียงค่าเดียว นักสถิติได้พยายามค้นคิดเพื่อให้การประมาณมีความถูกต้อง จึงได้คิดการสร้างช่วงเชื่อมั่นเพื่อทำการประมาณแบบช่วงสำหรับค่าพารามิเตอร์ ซึ่งการประมาณแบบช่วงจะมีความถูกต้องในการประมาณมากกว่าการประมาณแบบจุด เพราะการประมาณแบบช่วง สามารถครอบคลุม ค่าพารามิเตอร์ ที่ต้องการได้มากกว่าประมาณแบบจุด โดยช่วงของการประมาณ สร้างโดยอาศัยการแจกแจงของค่าที่ได้จากตัวอย่างสุ่มในบทที่3 ซึ่งสามารถบอกได้ว่าการประมาณในแต่ละครั้งมีความถูกต้องร้อยละเท่าใด

5.2.2 การประมาณแบบช่วง

การประมาณแบบช่วงคือการสร้างช่วงของตัวเลขช่วงหนึ่ง โดยอาศัยค่าที่ได้จากตัวอย่าง ช่วงเชื่อมั่นของการประมาณ ประกอบด้วย

<u>ขีดจำกัดของการประมาณ</u> สร้างโดยอาศัยการแจกแจงของค่าสถิติที่ได้จากตัวอย่าง เช่น การประมาณค่าเฉลี่ยของประชากร เราสามารถสร้างโดยพิจารณาจาก

การแจกแจงของค่าเฉลี่ยตัวอย่างสุ่ม ซึ่งได้กล่าวในบทที่3 ดังนั้นในการประมาณค่าพารามิเตอร์ จะต้องอาศัยการแจกแจงค่าสถิติของตัวอย่างเสมอ การประมาณค่าจะบอกถึงขีดจำกัดของการประมาณ ซึ่งประกอบด้วยขีดจำกัดล่างและขีดจำกัดบน

<u>ระดับความเชื่อมั่น</u> การสร้างช่วงเชื่อมั่นต้องบอกว่าช่วงที่สร้างมีโอกาสของความถูกต้องร้อยละเท่าใด กำหนดโดยใช้โอกาสของความถูกต้องเป็น 1- α และมีโอกาสผิดพลาดเป็น α ความผิดพลาด α เรียกว่าระดับนัยสำคัญ

ขั้นตอนของการประมาณค่า

- กำหนดพารามิเตอร์ที่ต้องการประมาณว่าเป็นค่าใด และจะใช้ค่าสถิติอะไรทำการประมาณ
- กำหนดโอกาสของการผิดพลาดหรือ ระดับนัยสำคัญ α
- พิจารณาการแจกแจงค่าสถิติว่ามีการแจกแจงแบบใด ในบทการแจกแจงค่าที่ได้จากตัวอย่าง และ เขียนช่วงเชื่อมั่น โดยอาศัยข้อมูลจากตัวอย่าง

5.2.3 ตารางสรุปการแจกแจงของค่าสถิติและช่วงของการประมาณค่า

การประมาณค่าเฉลี่ย สามารถประมาณแยกได้เป็นกรณีต่าง ๆดังนี้

กำหนดช่วงความเชื่อมั่นเป็น (1- α)100%

ค่าพารามิเตอร์	ค่าสถิติ	กรณีและการแจกแจงค่าสถิติ	ขีดจำกัดของการประมาณ
μ	$\overline{X}$	1. สุ่มตัวอย่างจากประชากรที่มีการแจกแจงปกติ หรือ 2. สุ่มตัวอย่างจากประชากรที่ไม่ทราบการแจกแจงแต่มีขนาดตัวอย่างใหญ่ (n$\geq$30) **การแจกแจง** $Z = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$	$\mu = \overline{X} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$
		3. การประมาณค่าเฉลี่ยกรณีไม่ทราบส่วนเบี่ยงเบนมาตรฐานของประชากร (σ) **การแจกแจง** $t = \dfrac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$	$\mu = \overline{X} \pm t_{\frac{\alpha}{2}, n-1} \dfrac{S}{\sqrt{n}}$
π	p	**การแจกแจง** $Z = \dfrac{p - \mu_p}{\sigma_p}$	$\pi = p \pm z_{\frac{\alpha}{2}} \sqrt{\dfrac{p(1-p)}{n}}$

ตัวอย่างที่3

ค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษาคณะบัญชีมีการแจกแจงปกติ ด้วยค่าส่วนเบี่ยงเบนมาตรฐาน 1600 บาท สุ่มตัวอย่างจำนวนนักศึกษา 1000 คน จากมหาวิทยาลัยแห่งหนึ่ง ค่าใช้จ่ายโดยเฉลี่ยต่อเดือน 3000 บาท จงประมาณค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษากลุ่มนี้ด้วยความเชื่อมั่น 95 %

วิธีทำ โจทย์ต้องการประมาณค่าเฉลี่ย

กำหนดให้ μ แทนค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษากลุ่มนี้

$\overline{X}$ แทนค่าใช้จ่ายเฉลี่ยต่อเดือนของนักศึกษาในตัวอย่าง

n= 1000 $\overline{X}$= 3000 σ =1600 α = 0.05 $z_{0.025}$= 1.96

ขีดจำกัดของการประมาณสำหรับค่า μ $= \overline{X} \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$

$= 3000 \pm 1.96 \dfrac{1600}{\sqrt{1000}}$

= (2900.83,3099.17)

$\therefore$ 95% ช่วงเชื่อมั่นของการประมาณสำหรับค่า μ คือ

$$2900.83 < \mu < 3099.17$$

ค่าใช้จ่ายโดยเฉลี่ยของนักศึกษากลุ่มนี้มีค่าอยู่ระหว่าง 2900.83 บาท และ 3099.17 บาท ด้วยความเชื่อมั่น 95 %

ตัวอย่างที่4

เวลาที่นักศึกษาใช้ในการรอสัมภาษณ์บุคคลสำคัญ มีค่าเบี่ยงเบนมาตรฐาน 9 นาที สุ่มตัวอย่างนักศึกษาที่ ไปสัมภาษณ์บุคคลสำคัญจำนวน 36 คน พบว่าโดยเฉลี่ยต้องรอ 30 นาที จงประมาณเวลาที่ไปรอการสัมภาษณ์โดยเฉลี่ยทั้งหมด α =0.1

วิธีทำ โจทย์ต้องการประมาณค่าเฉลี่ย

กำหนดให้ μ แทนเวลาที่ไปรอการสัมภาษณ์โดยเฉลี่ยทั้งหมด

$\overline{X}$ แทนเวลาที่ไปรอการสัมภาษณ์โดยเฉลี่ยในตัวอย่าง

n= 36 $\overline{X}$ = 30 σ =9 α = 0.1 $z_{0.05}$= 1.645

ขีดจำกัดของการประมาณสำหรับค่า μ $= \overline{X} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$

$= 30 \pm 1.645\dfrac{9}{\sqrt{36}}$

= (27.5325 ,32.4675)

$\therefore$ 90% ช่วงเชื่อมั่นของการประมาณสำหรับค่า μ คือ

$$27.5325 < \mu < 32.4675$$

เวลาที่ใช้ในการรอสัมภาษณ์โดยเฉลี่ยอยู่ระหว่าง 27.5325 นาที และ 32.4675 นาที ด้วยความเชื่อมั่น 90 %

ตัวอย่างที่5

ตำรวจจราจรต้องการประมาณ จำนวนอุบัติเหตุของรถจักรยานยนต์ที่เกิดขึ้น เพื่อเตือนให้ผู้ขับขี่สวมหมวกนิรภัย จากจำนวนอุบัติเหตุที่พบในท้องที่ต่อวันได้จำนวนอุบัตุเหตุที่เกิดดังนี้

10	9	10	11	16	15	8	6	18	17
4	12	15	14	15	9	7	8	16	14

จงประมาณช่วงแห่งความเชื่อมั่น 95 % ของจำนวนอุบัติเหตุโดยเฉลี่ยต่อวันในท้องที่นี้ ถ้าจำนวนครั้งของการเกิดอุบัติเหตุต่อวันมีการแจกแจงปกติ

วิธีทำกำหนดให้ μ แทน จำนวนอุบัติเหตุจากรถจักรยานยนต์เฉลี่ยต่อวันในท้องที่นี้ทั้งหมด

$\overline{X}$ แทน จำนวนอุบัติเหตุรถจักรยานยนต์เฉลี่ยต่อวันในท้องที่นี้ในตัวอย่าง

จากโจทย์ ข้อมูลนี้ไม่ทราบค่าส่วนเบี่ยงเบนมาตรฐานของประชากร ทำการประมาณค่าโดยใช้ค่าจากตัวอย่าง

วิธีทำ กำหนดให้ μ แทน จำนวนอุบัติเหตุโดยเฉลี่ยต่อวัน ในท้องที่นี้ในประชากร

$\overline{X}$ แทน จำนวนอุบัติเหตุโดยเฉลี่ยต่อวัน ในท้องที่นี้ในตัวอย่าง

$n = 6$ $\overline{X}=11.7$ $S = 4.0401$

$\alpha = 0.05$ $df = 20-1=19$ $t_{0.025,19} = 2.093$

ขีดจำกัดของการประมาณสำหรับ μ คือ $= \overline{X} \pm t_{\frac{\alpha}{2},n-1} \dfrac{S}{\sqrt{n}}$

$= 11.7 \pm 2.093 \left(\dfrac{4.0401}{\sqrt{20}}\right)$

$= (9.809, 13.591)$

$\therefore$ จำนวนอุบัติเหตุโดยเฉลี่ยต่อวันในท้องที่นี้มีค่าอยู่ระหว่าง 9.809 ครั้ง และ 13.591 ครั้ง ด้วยความเชื่อมั่น 95 %

<u>ตัวอย่างที่6</u>

จงประมาณค่าเฉลี่ยแบบจุดและแบบช่วง ด้วยความเชื่อมั่น 90 %ของอายุการใช้งานเฉลี่ยของล้อเครื่องบิน บริษัทผลิตล้อเครื่องบินได้สุ่มตัวอย่างทำการตรวจสอบล้อเครื่องบินที่ส่งมาเปลี่ยนได้ระยะทางที่วิ่ง (หน่วยเป็นพันไมล์)ดังนี้

46.7	47.2	49.1	56.5	56.8	59.2	59.9	63.2
63.3	63.4	63.7	64.1	67.1	67.7	93.3	118.5

วิธีทำ กำหนดให้ μ แทน อายุการใช้งานของล้อเครื่องบิน(พันไมล์) ทั้งหมด

$\overline{X}$ แทน อายุการใช้งานของล้อเครื่องบิน(พันไมล์) ในตัวอย่าง

$n = 16$ $\sum_{i=1}^{n} X_i^2 = 72350.71$ $\sum_{i=1}^{n} X_i = 1039.7$

$\overline{X} = 64.98125$ $s = 17.869$

$\alpha = 0.1$ $df = 16-1=15$ $t_{0.05,15} = 1.753$

ค่าประมาณแบบจุด = 64.98125

ขีดจำกัดของการประมาณสำหรับ μ คือ $= \overline{X} \pm t_{\frac{\alpha}{2},n-1} \dfrac{S}{\sqrt{n}}$

$= 64.98125 \pm 1.753 \left(\dfrac{17.869}{\sqrt{16}}\right)$

$= (57.150, 72.812)$

$\therefore$ 90 % ช่วงเชื่อมั่นของการประมาณสำหรับค่า μ คือ

$57.150 < \mu < 72.812$

$\therefore$ อายุการใช้งานของล้อเครื่องบินโดยเฉลี่ยมีค่าอยู่ระหว่าง 57.150 พันไมล์ และ 72.812 พันไมล์ ด้วยความเชื่อมั่น 90 %

ตัวอย่างที่7

ผู้วิจัยต้องการทราบว่านักศึกษาเพศชายมีสัดส่วนเท่าใดในมหาวิทยาลัยเอกชน จึงสุ่มตัวอย่างนักศึกษาในมหาวิทยาลัยเอกชนจำนวน 423 คน มีนักศึกษาชาย 196 คน จงประมาณสัดส่วนของนักศึกษาเพศชายในมหาวิทยาลัยเอกชนด้วยความเชื่อมั่น 95%

วิธีทำ กำหนด p แทนสัดส่วนนักศึกษาเพศชายในมหาวิทยาลัยเอกชนในตัวอย่าง

π แทนสัดส่วนนักศึกษาเพศชายในมหาวิทยาลัยเอกชนในประชากร

$$n = 423 \quad p = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = 196/423 = 0.463$$

$$\alpha = 0.05 \quad z_{0.025} = 1.96$$

ขีดจำกัดของการประมาณสำหรับค่า π คือ $p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$

$$= 0.463 \pm 1.96\sqrt{\frac{0.463(1-0.463)}{423}}$$

$$= (0.415, 0.511)$$

$\therefore$ 95% ช่วงเชื่อมั่นของการประมาณสำหรับค่า π คือ $0.415 < \pi < 0.511$

สัดส่วนของนักศึกษาเพศชายในมหาวิทยาลัยเอกชนมีค่าระหว่าง 0.415 และ 0.511 ด้วยความเชื่อมั่น 95 %

ตัวอย่างที่8

ผู้ผลิตต้องการทราบว่าการผลิตโทรทัศน์มีสัดส่วนของโทรทัศน์ที่ไม่ได้คุณภาพเป็นจำนวนเท่าใด จึงสุ่มตัวอย่างการผลิตโทรทัศน์จำนวน 50 เครื่องมีโทรทัศน์ที่ไม่ได้คุณภาพจำนวน 7 เครื่อง จงประมาณสัดส่วนของโทรทัศน์ที่ไม่ได้คุณภาพด้วยความเชื่อมั่น 95%

วิธีทำ

กำหนด p แทนสัดส่วนโทรทัศน์ที่ไม่ได้คุณภาพในตัวอย่าง

π แทนสัดส่วนโทรทัศน์ที่ไม่ได้คุณภาพในประชากร

$$n = 50 \quad p = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = 7/50 = 0.14 \quad \alpha = 0.05 \; z_{0.025} = 1.96$$

ขีดจำกัดของการประมาณสำหรับค่า π คือ $p \pm z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$

$$= 0.14 \pm 1.96\sqrt{\frac{0.14(1-0.14)}{50}}$$

$$= (0.044, 0.236)$$

$\therefore$ 95% ช่วงเชื่อมั่นของการประมาณสำหรับค่า π คือ $0.044 < \pi < 0.236$

สัดส่วนของโทรทัศน์ที่ไม่ได้คุณภาพมีค่าระหว่าง 0.044 และ 0.236 ด้วยความเชื่อมั่น 95 %

5.3 แบบทดสอบ

แบบทดสอบที่1

จากการศึกษา ดัชนีราคาผู้บริโภคหมวดข้าว มีการแจกแจงปกติด้วยค่าส่วนเบี่ยงเบนมาตรฐาน 15 ข้อมูลในตารางต่อไปนี้เป็นดัชนีราคาผู้บริโภคหมวดข้าวปี 25433-2546

ปี	2533	2534	2535	2536	2537	2538	2539	2540	2541	2542	2543	2544	2545	2546
ข้าว	45.4	47.5	49.7	49.2	51.3	53.3	62.2	80.2	100	85.5	87.5	85.6	78.6	86.3

ที่มา: ธนาคารแห่งประเทศไทย

จงประมาณดัชนีราคาผู้บริโภคหมวดข้าวด้วยความเชื่อมั่น 95%

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

แบบทดสอบที่2

ข้อมูลต่อไปนี้เป็นจำนวนการต่อโทรศัพท์ไปต่างประเทศ(ล้านครั้ง)

2541	2542	2543	2544	2545	2546
92.49	94.63	96.44	105.31	115.00	111.76

ที่มา: บริษัท ไปรษณีย์ไทย จำกัด

ถ้าจำนวนการต่อโทรศัพท์ไปต่างประเทศ(ล้านครั้ง) มีการแจกแจงปกติ จงประมาณช่วงเชื่อมั่น95%ของจำนวนการต่อโทรศัพท์ไปต่างประเทศ(ล้านครั้ง)

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

…………………………………………………………………………………………

<u>แบบทดสอบที่3</u>

บริษัทไมโครซอฟท์ได้ออกเกมเพลย์สเตชั่นรุ่นใหม่ เอ็กซ์360 และทางบริษัทคาดว่าจะได้สัดส่วนตลาดมากกว่าเดิม ในปัจจุบันบริษัทโซนี่เป็นผู้นำตลาดด้านเกมเพลย์สเตชั่น จากการสอบถาม บ้านที่ซื้อเกมเพลส์สเตชั่น 100 ครัวเรือน มีครัวเรือนที่สนใจซื้อเกมเพลย์สเตชั่นจำนวน 48 ครัวเรือน จงประมาณสัดส่วนของลูกค้าที่สนใจซื้อเกมเพลย์สเตชั่นรุ่นใหม่ด้วยความเชื่อมั่น 95%

....

....

....

....

....

....

....

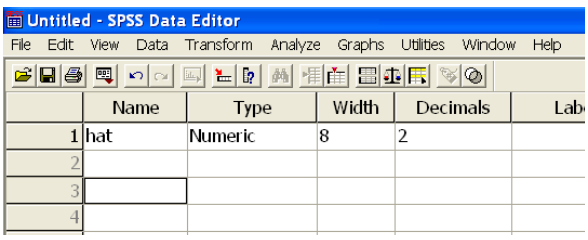
....

5.4 การใช้โปรแกรม SPSS ในการวิเคราะห์การประมาณค่า

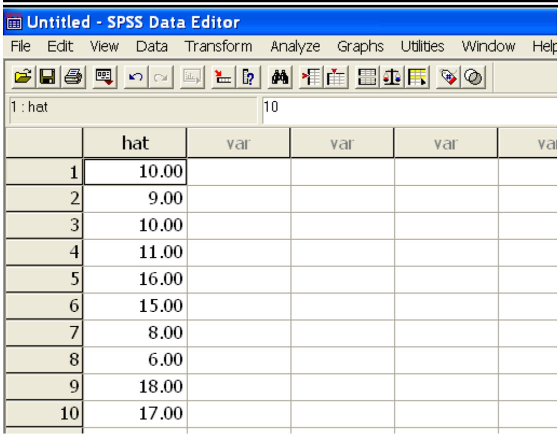
การประมาณค่าโดยใช้โปรแกรม SPSS ในการวิเคราะห์เพื่อหาการประมาณค่า ข้อมูลที่โจทย์กำหนดต้องเป็นข้อมูลดิบ ไม่ใช่ข้อมูลที่ผ่านการคำนวณค่าสถิติเช่นค่าเฉลี่ย หรือ ส่วนเบี่ยงเบนมาตรฐานเหมือนในตัวอย่างที่1-4 สำหรับในตัวอย่างที่5โจทย์ให้ข้อมูลดิบสามารถวิเคราะห์โดยใช้โปรแกรมได้

ขั้นตอนการวิเคราะห์มีดังนี้

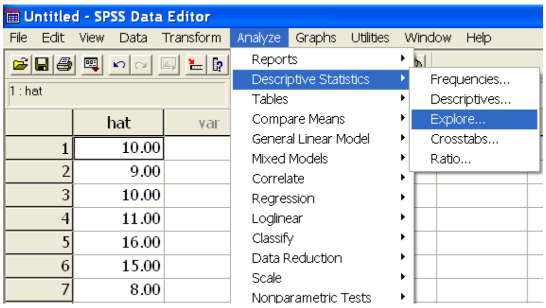
1. เมื่อเปิดโปรแกรม จะมีหน้าจอให้ 2 หน้าจอคือหน้าจอการใส่ข้อมูล Data View และหน้าจอกำหนดชื่อตัวแปร Variable View
2. ตั้งชื่อตัวแปรดังรูป



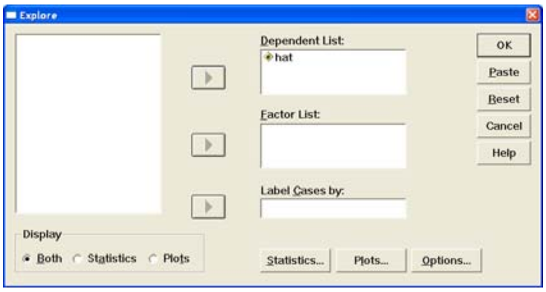
3. ป้อนข้อมูลโดยกลับไปที่หน้าจอ Data View จะได้หน้าจอ



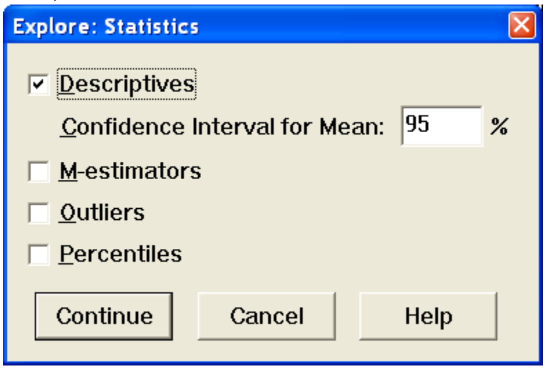
เลือกเมนู Analyze เลือก Explore จะได้หน้าจอ



เลือกตัวแปร hat ลงใช่อง Dependent List



กดปุ่ม Statistics เลือก Descriptive



กดปุ่มContinue จะได้หน้าจอเดิมและกดปุ่ม OK จะได้ผลวิเคราะห์ดังนี้

Explore

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
HAT	20	100.0%	0	.0%	20	100.0%

Descriptives

			Statistic	Std. Error
HAT	Mean		11.7000	.90350
	95% Confidence Interval for Mean	Lower Bound	9.8089	
		Upper Bound	13.5911	
	5% Trimmed Mean		11.7778	
	Median		11.5000	
	Variance		16.326	
	Std. Deviation		4.04058	
	Minimum		4.00	
	Maximum		18.00	
	Range		14.00	
	Interquartile Range		6.7500	
	Skewness		-.183	.512
	Kurtosis		-1.096	.992

ช่วงเชื่อมั่น

5.5 การทดสอบสมมติฐาน

สมมติฐาน คือ ข้อความหรือข้อกำหนดที่เกี่ยวข้องกับลักษณะประชากรที่เราสนใจ ซึ่งอาจจะเกิดจากความเชื่อของบุคคลใดบุคคลหนึ่งหรือบุคคลทั่วๆไปเช่น

- การจัดงาน Money Expo'2005 คาดว่าผู้เข้าชมงานจะสนใจทำสินเชื่อบัตรเครดิตมากที่สุด
- เวลาที่ใช้ในการรอสัมภาษณ์บุคคลสำคัญโดยเฉลี่ย 30 นาที
- นักศึกษาคณะบัญชีเลือกเรียนมหาวิทยาลัยหอการค้าไทย เพราะชื่อเสียงของสถาบัน เป็นต้น

นอกจากนี้สมมติฐานอาจจะเป็นสมมติฐานที่เกิดจากทางทฤษฎีต่างๆหรือเกิดจากการทดลอง เนื่องจากสมมติฐานเป็นข้อที่เกิดจากความเชื่อ การทดลอง หรือทางทฤษฎีดังนั้นเพื่อเป็นการพิสูจน์ว่าความเชื่อถูกต้องหรือไม่ เราจำเป็นต้องทำการเก็บข้อมูลเพื่อนำมาสนับสนุนความเชื่อดังกล่าว ในบางครั้งการเก็บข้อมูลไม่สามารถเก็บได้จากทุกหน่วยของประชากร ดังนั้นเราต้องอาศัยการเลือกตัวอย่างสุ่ม

การเลือกตัวอย่างที่ดีจะได้ค่าสถิติที่มีค่าใกล้เคียงกับค่าของประชากร หากตัวอย่างที่เก็บได้ไม่ดี ผลที่ได้ค่าแตกต่างจากประชากร และตัวอย่างนั้นจะไม่เป็นตัวแทนที่ดีของประชากรที่ต้องการศึกษา ดังนั้น ขั้นตอนการเลือกตัวอย่างจึงเป็นขั้นตอนที่สำคัญในการศึกษาสถิติเชิงอ้างอิง เมื่อเราได้ตัวอย่างต้องทำการทดสอบว่า ตัวอย่างที่ได้สนับสนุนความเชื่อของเราหรือไม่ เราสามารถทำได้โดย การทดสอบสมมติฐาน ถ้าตัวอย่างสนับสนุนความเชื่อของเรา สมมติฐานที่ตั้งไว้จะเป็นจริง หากตัวอย่างที่ได้ไม่สนับสนุนความเชื่อนั้น สมมติฐานของการทดสอบจะไม่เป็นจริง

5.5.1 การตั้งสมมติฐานทางสถิติ

สมมติฐานที่กล่าวข้างต้นเป็นสมมติฐานที่ต้องการศึกษา สำหรับการทดสอบเราต้องนำสมมติฐานมาเขียนให้เป็นสมมติฐานทางสถิติ การตั้งสมมติฐานทางสถิติเพื่อการทดสอบจะต้องประกอบด้วยสมมติฐาน 2 ชนิด คือ

- สมมติฐานหลัก (Null Hypothesis) ใช้สัญลักษณ์ H_0 เป็นสมมติฐานที่แสดงค่า พารามิเตอร์ ที่จะต้องมีเครื่องหมายเท่ากับรวมอยู่ในค่านั้นเสมอ เช่น

H_0 : เวลาที่ใช้ในการรอสัมภาษณ์บุคคลสำคัญเฉลี่ย 30 นาที

เขียนแทนด้วยประโยคสัญลักษณ์ $\mu = 30$

H_0 : ค่าใช้จ่ายในการเดินทางเฉลี่ยต่อวันของนักศึกษาชั้นปีที่ 1 ไม่เกิน 100 บาท

หรือเขียนแทนด้วยประโยคสัญลักษณ์ $\mu \leq 100$

H_0 : สัดส่วนของนักศึกษาหญิงในมหาวิทยาลัยหอการค้าไทยมีอย่างน้อยร้อยละ75

เขียนแทนด้วยประโยคสัญลักษณ์ $\pi \geq 0.75$

- สมมติฐานรอง (Alternative Hypothesis)

ใช้สัญลักษณ์ H_a หรือ H_1 เป็นสมมติฐานที่แสดงค่าพารามิเตอร์ที่ตรงกันข้ามกับค่าพารามิเตอร์ในสมมติฐานหลัก และไม่มีค่าพารามิเตอร์ซ้ำกันกับค่าพารามิเตอร์ที่อยู่ในสมมติฐานหลัก เช่น

H_1 : เวลาที่ใช้ในการรอสัมภาษณ์บุคคลสำคัญเฉลี่ยไม่เกิน 30 นาที หรือ

เขียนแทนด้วยประโยคสัญลักษณ์ $\mu \neq 30$

H_1 : ค่าใช้จ่ายในการเดินทางเฉลี่ยต่อวันของนักศึกษาชั้นปีที่ 1 เกินกว่า 100 บาท

หรือ เขียนแทนด้วยประโยคสัญลักษณ์ $\mu > 100$

H_1 : สัดส่วนของเพศหญิงในมหาวิทยาลัยหอการค้าไทยมีน้อยกว่า 75%

หรือ เขียนแทนด้วยประโยคสัญลักษณ์ $\pi < 0.75$

ตัวอย่างที่9

จงตั้งสมมติฐานเพื่อการทดสอบจากข้อมูลต่อไปนี้ในแต่ละกรณี

ก. บริษัทฟิล์มสีฟูจี คาดว่ายอดขายฟิล์มสีเฉลี่ยเป็น 300 ล้านม้วน

H_0 : $\mu = 300,000,000$

ข. นักศึกษาชายของมหาวิทยาลัยหอการค้าไทยขับรถมาเรียนมากกว่า 15%

H_0 : $\pi \geq 0.15$

ผลการทดสอบสมมติฐาน มี 2 อย่าง คือ ปฏิเสธสมมติฐานหลัก(ปฏิเสธ H_0)กับไม่ปฏิเสธสมมติฐานหลัก(ยอมรับ H_0)

5.5.2 การกำหนดระดับนัยสำคัญ

ในการทดสอบสมมติฐานเราต้องกำหนดว่า ในการทดสอบครั้งหนึ่งๆเราสามารถบอกได้ว่าถูกต้องเท่าใด และผิดพลาดเท่าใด เช่นกำหนดระดับนัยสำคัญ ของการทดสอบสมมติฐานเป็น 0.05 หมายถึงในการทดสอบร้อยครั้ง สามารถสรุปถูกต้อง 95 ครั้ง ผิดพลาด 5ครั้ง ดังนั้นในการทดสอบเราจะพยายามกำหนดค่าระดับนัยสำคัญให้มีค่าน้อยๆ แต่ไม่สามารถกำหนดเป็น 0 ได้เนื่องจากในการทดสอบสมมติฐานมีความผิดพลาดเกิดขึ้น 2 ประเภทคือ

- ความผิดพลาดชนิดที่ 1 (Type I error)

ถ้าสมมติฐานหลัก เป็นจริงแล้วได้ ผลสรุปจากการทดสอบสมมติฐานว่าปฏิเสธสมมติฐานหลักเรียกว่าเกิดความผิดพลาดชนิดที่ 1 (Type I error) และโอกาสที่จะเกิดความผิดพลาดชนิดที่ 1 เรียกว่า α ในที่นี้ α คือค่าระดับนัยสำคัญ

ดังนั้น $\boxed{\alpha = P(\text{ ปฏิเสธ } H_0 \ / \ H_0 \text{ เป็นจริง})}$

- ความผิดพลาดชนิดที่ 2 (Type II error)

ถ้าสมมติฐานหลักไม่จริง แต่ในการทดสอบสมมติฐานผลสรุปจากการทดสอบสมมติฐานว่าไม่ปฏิเสธสมมติฐานหลัก เรียกว่าเกิดความผิดพลาดชนิดที่ 2 (Type II error) และโอกาสที่จะเกิดความผิดพลาดชนิดที่ 2 เรียกว่า β

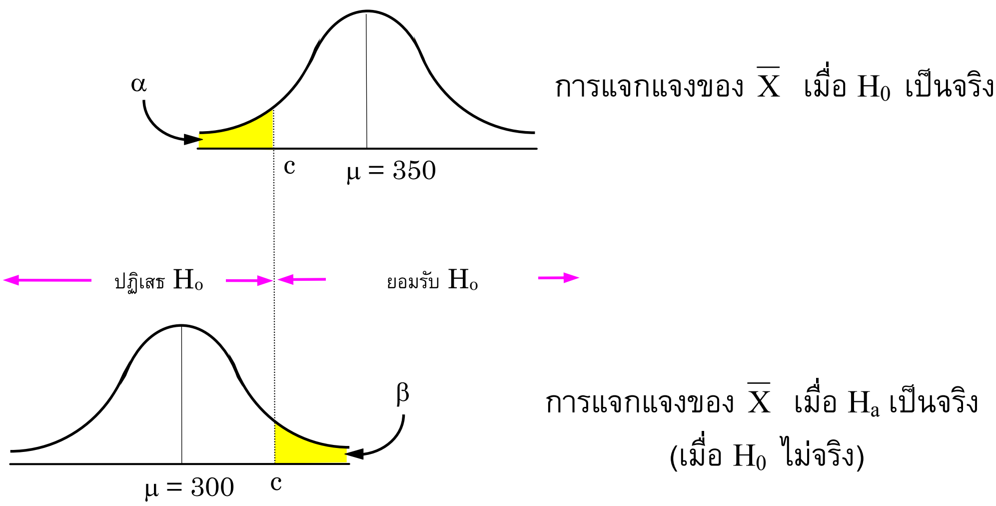
ดังนั้น $\boxed{\beta = P(\text{ ไม่ปฏิเสธ } H_0 \ / \ H_0 \text{ ไม่จริง})}$

เราเรียก 1- β ว่า อำนาจการทดสอบ

ในการทดสอบสมมติฐานครั้งหนึ่งๆเราต้องการให้ค่า ความผิดพลาดทั้งสองประเภทมีค่าน้อยๆ แต่เนื่องจากความสัมพันธ์ของค่าทั้งสองมีทิศทางตรงข้ามหมายถึงเมื่อกำหนด α มีค่าน้อยๆค่า β จะมีค่าเพิ่มขึ้น และหากกำหนดค่า β มีค่าน้อยๆ ค่า α มีค่าเพิ่มขึ้น ดังนั้นในการทดสอบสมมติฐานเราไม่สามาตรกำหนด $\alpha = 0$ ได้เนื่องจาก เมื่อกำหนดเป็น $\alpha = 0$ ค่า β จะมีค่าเพิ่มมากขึ้นจะทำให้การทดสอบนั้นมีอำนาจการทดสอบลดลง สมมติเราต้องการทดสอบ

$H_0 \ : \ \mu = 350$

$H_1 \ : \ \mu < 350$



ตามรูป ถ้ากำหนด c เป็นค่าวิกฤต

$\alpha = P(\bar{X} < c)$ เมื่อ $\mu = 350$ (H_0 เป็นจริง)

$\beta = P(\bar{X} > c)$ เมื่อ $\mu = 300$ (H_1 เป็นจริง)

ถ้า c เลื่อนไปทางขวา จะทำให้ α เพิ่มขึ้น และ β ลดลง

แต่ถ้า c เลื่อนไปทางซ้าย จะทำให้ α ลดลง และ β เพิ่มขึ้น

ในการทดสอบโดยทั่วๆไปเราจึงทำการกำหนดจากค่า α ซึ่งเรียกว่า ระดับนัยสำคัญ หรือ 1- α เป็นโอกาสที่จะไม่ปฏิเสธ H_0 เมื่อ H_0 เป็นจริง เมื่อคูณด้วย 100 จะเรียกว่าระดับความเชื่อมั่น โดยปกติมักกำหนด α = 0.01,0.05,0.10

5.5.3 ค่าวิกฤตและบริเวณวิกฤต

การทดสอบสมมติฐานเราต้องสร้างเกณฑ์เพื่อใช้ในการกำหนดบริเวณสำหรับการยอมรับและปฏิเสธ สมมติฐานหลัก เราใช้ค่าวิกฤตเป็นตัวแบ่ง

ค่าวิกฤต(Critical Value) คือค่าที่แบ่งบริเวณที่ใช้ทดสอบออกสองส่วนคือบริเวณการปฏิเสธสมมติฐานหลักกับบริเวณการยอมรับสมมติฐานหลัก

5.5.4 การทดสอบแบบด้านเดียวและการทดสอบแบบสองด้าน

ให้ θ เป็นพารามิเตอร์ และ k เป็นค่าคงที่ใดๆ

θ หมายถึง ค่าเฉลี่ย (μ) ค่าสัดส่วน(π) และ ค่าแปรปรวน(σ^2)

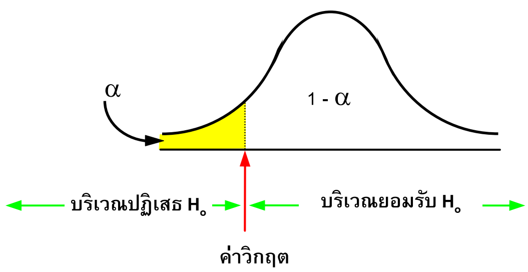
- การทดสอบแบบด้านเดียว

การทดสอบแบบด้านเดียวเกิดขึ้นเมื่อทำการทดสอบสมมติฐานการทดสอบสนใจด้านมากหรือด้านน้อยต่อไปนี้

ก. $H_0 : \theta \geq k$ หรือ $\theta = k$

$H_1 : \theta < k$

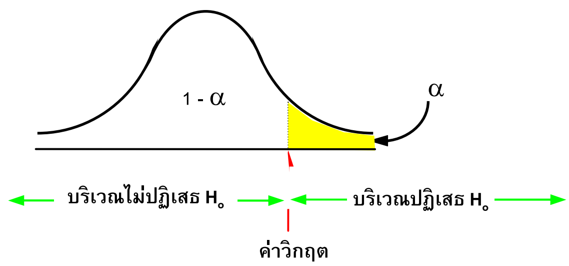
บริเวณวิกฤตหรือบริเวณที่จะปฏิเสธ H_0 จะอยู่ปลายหางด้านซ้ายมือ ซึ่งเรียกว่าเป็นการทดสอบด้านเดียวด้านน้อยดังรูป



ข. $H_0 : \theta \leq k$ หรือ $\theta = k$

$H_1 : \theta > k$

บริเวณวิกฤตหรือบริเวณที่จะปฏิเสธ H_0 จะอยู่ปลายหางด้านขวามือ ซึ่งเรียกว่าเป็น



การทดสอบด้านเดียวด้านมาก(Upper-tailed test) ดังรูป

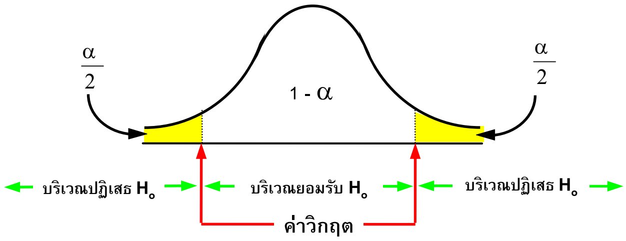
- การทดสอบแบบสองด้าน (two-tailed test)

การทดสอบแบบสองด้าน เป็นการทดสอบสมมติฐานที่สนใจทั้งสองด้านพร้อมกันสมมติฐานคือ

$H_0 \;:\; \theta = k$

$H_1 \;:\; \theta \neq k$

บริเวณวิกฤตหรือบริเวณที่จะปฏิเสธ H_0 จะอยู่ปลายหางทั้ง 2 ด้านดังรูป



ในการสรุปผลการทดสอบเมื่อคำนวณค่าสถิติที่เก็บจากตัวอย่างให้นำค่าที่ได้มาเปรียบเทียบกับค่าวิกฤตในรูป หากค่าที่ได้ตกบริเวณปฏิเสธ ตัดสินใจปฏิเสธ H_0 หากตกบริเวณยอมรับตัดสินใจยอมรับ H_0

5.6 ขั้นตอนการทดสอบสมมติฐาน

1. ตั้งสมมติฐานเราต้องพิจารณาว่าพารามิเตอร์ที่ต้องการทดสอบคือค่าใดเช่นถ้าเป็น ค่าเฉลี่ยแทนด้วย μ ค่าสัดส่วนแทนด้วยπ และ ค่าแปรปรวนแทนด้วยσ^2จากนั้นถามตัวเองว่า ข้อความที่ต้องการทดสอบมีความหมายบ่งบอกการเท่ากันหรือไม่ หากมีความหมายบ่งบอกการเท่ากัน ข้อความที่ต้องการทดสอบต้องตั้งที่ H_0 ส่วนที่เหลือจะเป็นข้อความที่ H_1 หากข้อความนั้น ไม่มีความหมายบ่งบอกการเท่ากันให้ตั้งที่ H_1 ส่วนที่เหลือจะเป็นข้อความที่ H_0

2. กำหนดระดับนัยสำคัญ (α) โดยทั่วไปกำหนดเป็น 0.01,0.05 และ 0.10

3. กำหนดสถิติที่ใช้ทดสอบ การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยและสัดส่วนของประชากรเดียวตัวสถิติที่ใช้ทดสอบคือ Z ,T,χ^2 และ F ซึ่งการที่จะกำหนดตัวสถิตินี้ จะต้องสอดคล้องกับพารามิเตอร์ที่สมมติฐานตั้งขึ้นเพื่อทำการทดสอบและลักษณะของประชากรที่เก็บรวบรวมข้อมูลมา สำหรับตัวสถิติทดสอบจะพูดรายละเอียดต่อไป

เมื่อกำหนดสถิติที่ใช้ทดสอบได้แล้วให้แทนค่าต่าง ๆจากข้อมูลที่รวบรวมได้ลงในสูตรของตัวสถิติที่ใช้ทดสอบพร้อมกับคำนวณค่า

4. ดูบริเวณการยอมรับและปฏิเสธสมมติฐาน โดยใช้ค่าวิกฤต ค่าวิกฤตที่กำหนดขึ้นมาจะต้องขึ้นกับตัวสถิติที่ใช้ทดสอบ เช่น ตัวสถิติที่ใช้ทดสอบคือ Z ค่าวิกฤตจะได้จากการเปิดตาราง Z และถ้าตัวสถิติที่ใช้ทดสอบคือ t ค่าวิกฤตก็จะได้จากการเปิดตาราง t ค่าวิกฤตจะมีค่าเดียวหรือ 2 ค่าก็ขึ้นอยู่กับการทดสอบว่าเป็นแบบด้านเดียวหรือแบบ 2 ด้าน ซึ่งจะทำให้บริเวณวิกฤตมีด้านเดียวหรือ 2 ด้าน

5. ตัดสินใจและสรุปผล ถ้าค่าสถิติที่คำนวณได้จากตัวอย่างตกอยู่ในบริเวณปฏิเสธจะตัดสินใจว่า ปฏิเสธ H_0 แต่ถ้าค่าสถิติคำนวณได้จากตัวอย่างตกอยู่ในบริเวณยอมรับตัดสินใจว่ายอมรับ H_0

ในการพิจารณานอกเหนือจากการหาค่าวิกฤต โดยใช้ค่าจากตาราง ซึ่งอาจจะเป็นการแจกแจงของ Z ,t,χ^2, และการแจกแจงของ F การสรุปผลสามารถใช้ค่า p-value ได้

เนื่องจากปัจจุบันเราวิเคราะห์โดยใช้เคร ื่องคอมพิวเตอร์ และโปรแกรมจะให้ค่า p-value มาให้ ทำให้สะดวกในการเปรียบเทียบกับค่า ระดับนัยสำคัญ α ได้ ค่า p-value ได้จากการนำค่าสถิติที่ได้จากตัวอย่างไปคำนวณหาค่าความน่าจะเป็น

หากค่าจากตัวอย่างให้ค่าความน่าจะเป็นน้อยกว่าค่า α แสดงว่าค่าจากตัวอย่างยังอยู่ในบริเวณของการปฏิเสธจึงตัดสินใจปฏิเสธ H_0

หากค่าจากตัวอย่างให้ค่าความน่าจะเป็นมากกว่าค่า α แสดงว่าค่าจากตัวอย่างอยู่ในบริเวณของการยอมรับจะตัดสินใจยอมรับ H_0

เกณฑ์การตัดสินใจ ถ้า p- value $< \alpha$ ผลสรุปคือ ปฏิเสธ H_0

ถ้า p- value $\geq \alpha$ ผลสรุปคือ ไม่ปฏิเสธ H_0

5.6.1 การทดสอบสมมติฐานประชากรเดียว

พารามิเตอร์	กรณี	ค่าสถิติทดสอบ
μ	• สุ่มตัวอย่างสุ่มมาจากการแจกแจงแบบปกติและทราบค่าความแปรปรวนของประชากร	$Z = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$
	• ตัวอย่าง สุ่มมาจากประชากรที่มีการแจกแจงแบบใด ๆที่ไม่ทราบค่าความแปรปรวนของประชากรแต่ขนาดตัวอย่างโตพอ(n≥30)	$Z = \dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ โดย ที่$S^2 = \dfrac{\Sigma\left(X_i-\overline{X}\right)^2}{n-1}$
	• กรณีที่ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ ไม่ทราบค่าความแปรปรวนของประชากรและขนาดตัวอย่างที่สุ่มน้อยกว่า 30	$t = \dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}$ โดยที่ df = n-1
π	• กรณีสุ่มตัวอย่างขนาดใหญ่	$Z = \dfrac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$

ตัวอย่างที่ 10

บริษัทผลิตกล้องถ่ายรูปได้กล่าวว่า กล้องถ่ายรูปของเขามีอายุการใช้งานเฉลี่ย35000 ชั่วโมง จากการสุ่มตัวอย่างกล้องถ่ายรูป 10 เครื่อง พบว่ามีอายุการใช้งานเฉลี่ย 29500 ชั่วโมง หากทราบว่าอายุการใช้งานมีการแจกแจงปกติด้วยค่าส่วนเบี่ยงเบนมาตรฐาน 5000 ชั่งโมง จงทดสอบสมมติฐานว่าคำกล่าวอ้างของบริษัทเป็นจริงหรือไม่ ที่ระดับนัยสำคัญ 0.05

กำหนดให้

μ แทน อายุการใช้งานเฉลี่ยของกล้องถ่ายรูปบริษัทนี้

$\overline{X}$ แทน อายุการใช้งานเฉลี่ยของกล้องถ่ายรูปบริษัทนี้ในตัวอย่าง

วิธีทำ

1. ตั้งสมมติฐาน ในที่นี้ต้องการทดสอบค่าเฉลี่ยดังนั้นสมมติฐานการทดสอบมีดังนี้

$H_0: \mu = 35000$

$H_1: \mu \neq 35000$

2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$
3. กำหนดตัวสถิติทดสอบและคำนวณค่า

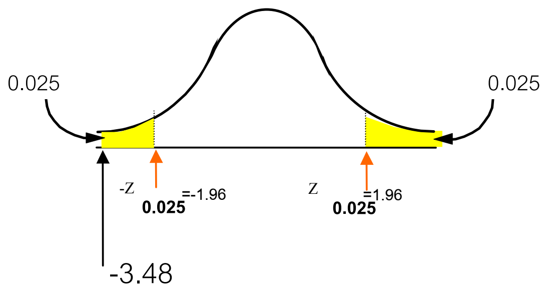
ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ตัวสถิติที่ใช้ในการทดสอบคือ

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

แทนค่า $\mu = 35000$ $\sigma = 5000$ $n = 10$ $\overline{X} = 29500$

$$Z = \frac{29500 - 35000}{\frac{5000}{\sqrt{10}}} = -3.48$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานสองด้าน $\alpha/2 = 0.025$



5. ตัดสินใจปฏิเสธ H_0 สรุปว่าอายุการใช้งานของกล้องถ่ายรูปไม่เป็น 35000 ชั่วโมง ที่ระดับนัยสำคัญ 0.05

ตัวอย่างที่ 11

จากใบสมัครเรียนของมหาวิทยาลัยแห่งหนึ่งกล่าวว่าเพื่อประสิทธิภาพของการเรียนในสถาบัน ทางมหาวิทยาลัยได้จัดชั้นเรียนขนาดเล็กโดยแต่ละห้องมีจำนวนนักศึกษาไม่เกิน 35 คน สุ่มตัวอย่างห้องเรียนจำนวน 64 ห้องพบว่ามีจำนวนนักศึกษาเรียนโดยเฉลี่ย 37 คนและส่วนเบี่ยงเบนมาตรฐาน 6 คน ท่านคิดว่าในใบสมัครเรียนที่กล่าวว่าเพื่อประสิทธิภาพของการเรียนนั้นเป็นจริงหรือไม่ที่ระดับนัยสำคัญ 0.05

กำหนดให้

μ แทน จำนวนนักศึกษาในชั้นเรียนของมหาวิทยาลัยเอกชนแห่งนี้

$\overline{X}$ แทน จำนวนนักศึกษาในชั้นเรียนของมหาวิทยาลัยเอกชนแห่งนี้ในตัวอย่าง

วิธีทำ

1. ตั้งสมมติฐาน ในที่นี้ต้องการทดสอบค่าเฉลี่ยดังนั้นสมมติฐานการทดสอบมีดังนี้

   $H_0: \mu \leq 35$

   $H_1: \mu > 35$

2. กำหนดระดับนัยสำคัญ

   $\alpha = 0.05$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

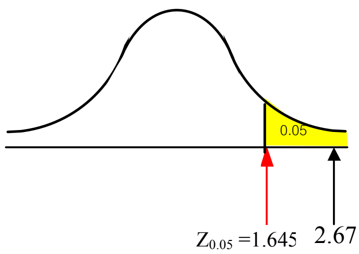
   ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ไม่ทราบค่าแปรปรวนและขนาดตัวอย่างใหญ่ ตัวสถิติที่ใช้ในการทดสอบคือ

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

แทนค่า $\mu = 35$ $S = 6$ $n = 64$ $\overline{X} = 37$

$$Z = \frac{37 - 35}{\frac{6}{\sqrt{64}}} = 2.67$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานแบบด้านเดียว ด้านมาก



5. ตัดสินใจปฏิเสธ H_0

สรุปว่าทางมหาวิทยาลัยไม่ได้จัดห้องเรียนตามที่ได้กล่าวไว้ ที่ระดับนัยสำคัญ 0.05

ตัวอย่างที่12

ผู้จัดการคิดว่าการลดน้ำหนักของลูกค้าจะต้องลดลงอย่างน้อย45 ปอนด์ ภายใน 6 เดือน สุ่มตัวอย่างลูกค้าที่ใช้บริการการลดน้ำหนัก 28 คน น้ำหนักโดยเฉลี่ยลดลง 35 ปอนด์ ส่วนเบี่ยงเบนมาตรฐาน 20 ปอนด์ ที่ระดับนัยสำคัญ 0.01 จงทดสอบว่าความเชื่อของผู้จัดการเกินความเป็นจริงหรือไม่ ถ้านำหนักที่ลดของลูกค้ามีการแจกแจงแบบปกติ

μ แทน น้ำหนักที่ลดลงโดยเฉลี่ยของลูกค้าที่ใช้บริการ

$\overline{X}$ แทน น้ำหนักที่ลดลงโดยเฉลี่ยของลูกค้าที่ใช้บริการของตัวอย่าง

วิธีทำ

1. ตั้งสมมติฐาน ในที่นี้ต้องการทดสอบค่าเฉลี่ยดังนั้นสมมติฐานการทดสอบมีดังนี้

   $H_0: \mu \geq 45$

   $H_1: \mu < 45$

2. กำหนดระดับนัยสำคัญ

   $\alpha = 0.01$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

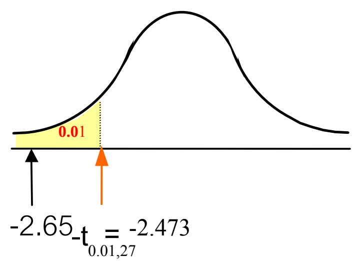
   ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ไม่ทราบค่าแปรปรวน และขนาดตัวอย่างเล็ก ตัวสถิติที่ใช้ในการทดสอบคือ

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

แทนค่า $\mu = 45$ $S = 20$ $n = 28$ $\overline{X} = 35$

$$t = \frac{35 - 45}{\frac{20}{\sqrt{28}}} = -2.65$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานแบบด้านเดียว ด้านน้อย



5. ตัดสินใจปฏิเสธ H_0 สรุปว่าคำกล่าวอ้างของผู้จัดการเกินความเป็นจริงที่ระดับนัยสำคัญ 0.01

5.6.2 การทดสอบสมมติฐานประชากรเดียวโดยใช้โปรแกรมSPSS

การป้อนข้อมูลเหมือนกับการประมาณค่า

ตัวอย่างที่13

ข้อมูลต่อไปนี้เป็นค่าใช้จ่ายแต่ละวันของนักศึกษากลุ่มหนึ่ง

145 228 150 130 142 265 160 114

จงทดสอบสมมติฐานว่าค่าใช้จ่ายนักศึกษาแต่ละวันมีค่าเป็น 150 บาทหรือไม่ ที่ระดับนัยสำคัญ 0.05 ผลการวิเคราะห์โดยใช้โปรแกรมได้ดังนี้

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
COST	8	166.7500	52.02952	18.39521

One-Sample Test

	Test Value = 150					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
COST	.911	7	.393	16.7500	-26.7478	60.2478

ขั้นตอนการทดสอบ

1. ตั้งสมมติฐาน H_0 : μ = 150

   H_1 : $\mu \neq 150$
2. กำหนดระดับนัยสำคัญ α = 0.05
3. ผลการทดสอบได้ค่า p-value = 0.393 มากกว่าค่า α = 0.05

   ตัดสินใจยอมรับสมมติฐาน H_0

สรุปว่า ค่าเฉลี่ยของประชากรมีค่าเท่ากับ 150 ที่ α = 0.05

5.7 การทดสอบค่าสัดส่วนของประชากร

การทดสอบสมมติฐานค่าสัดส่วนประชากร ใช้สำหรับทดสอบจำนวนสิ่งที่เราสนใจในประชากรว่าเป็นสัดส่วนเท่าใดของประชากรทั้งหมดแทนด้วย π เช่น

- สัดส่วนของนักศึกษาชายในมหาวิทยาลัยหอการค้าไทยที่ขับรถมาเรียน มีมากกว่าร้อยละ 25
- สัดส่วนของผู้เป็นสมาชิก IBC เคเบิ้ลทีวีเป็น 70 %
- นักศึกษาคณะบัญชีมีทักษะในการใช้โปรแกรม Microsoft Excel เป็นร้อยละ 45
- แชมพูสระผมที่นักศึกษาหญิงนิยมใช้คือ ซันซิล ร้อยละ 60

ผู้ทำการทดสอบจะต้องเก็บรวบรวมข้อมูลตัวอย่างจากประชากรที่สนใจศึกษาแล้ว คำนวณสัดส่วนของลักษณะที่สนใจในตัวอย่าง p เพื่อที่จะนำมาทดสอบกับค่าสัดส่วนของประชากรที่กำหนดไว้ในสมมติฐาน

ตัวอย่างที่14

เชื่อว่าร้อยละ70% ของผู้ชมที่เป็นสตรีเห็นด้วยกับ กรรมการตัดสินการประกวดนางงามที่ให้ สาวงามจากประเทศไทยได้รับรางวัลชนะเลิศประเภทแต่งกายประจำชาติ ในปีที่ประเทศไทยเป็นเจ้าภาพการประกวดนางงามจักรวาลปี 2005 จากการสุ่มตัวอย่างสตรีจำนวน 500 คน มีคนเห็นด้วย 300 คน จงทดสอบสมมติฐานดังกล่าวว่าเป็นจริงหรือไม่ ที่ระดับนัยสำคัญ 0.05

วิธีทำ

กำหนด p แทน สัดส่วนของผู้หญิงที่เห็นด้วยกับกรรมการตัดสินในตัวอย่าง

π แทน สัดส่วนของผู้หญิงที่เห็นด้วยกับกรรมการตัดสินในประชากร

1. ตั้งสมมติฐาน

$H_0: \quad \pi = 0.7$

$H_1: \quad \pi \neq 0.7$

2. กำหนดระดับนัยสำคัญ

$\alpha = 0.05$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

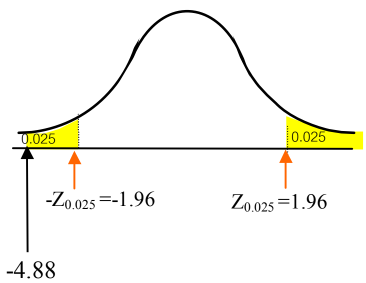
ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ตัวสถิติที่ใช้ในการทดสอบคือ

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

แทนค่า $p = 300/500 = 0.6$

$$Z = \frac{0.6 - 0.7}{\sqrt{\frac{0.7(1-0.7)}{500}}} = -4.88$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานสองด้าน $\alpha/2 = 0.025$



5. ตัดสินใจปฏิเสธ H_0 สรุปว่าสัดส่วนสัดส่วนของผู้หญิงที่เห็นด้วยกับกรรมการตัดสินไม่เป็น70% ที่ระดับนัยสำคัญ 0.05

ตัวอย่างที่15

ผู้จัดการต้องการทราบว่าสัดส่วนของผู้บริโภคแชมพูยี่ห้อ แฟซ่าอย่างมากเป็น 35 % หรือไม่ จึงทำการสุ่มตัวอย่างลูกค้าจำนวน 100 คน มีผู้ใช้แชมพูยี่ห้อนี้ 40 คน จงทดสอบสมมติฐานดังกล่าวเป็นจริงหรือไม่ ที่ระดับนัยสำคัญ 0.01

กำหนด p แทน สัดส่วนของผู้บริโภคแชมพูยี่ห้อ แฟซ่าในตัวอย่าง

π แทน สัดส่วนของของผู้บริโภคแชมพูยี่ห้อ แฟซ่าในประชากร

วิธีทำ

1. ตั้งสมมติฐาน

$H_0: \pi = 0.35$

$H_1: \pi > 0.35$

2. กำหนดระดับนัยสำคัญ

$\alpha = 0.01$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

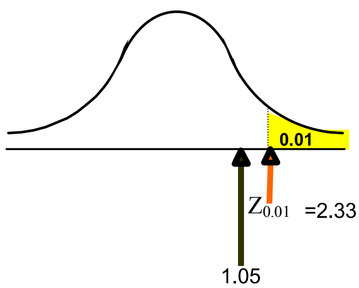
ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ตัวสถิติที่ใช้ในการทดสอบคือ

$$Z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

แทนค่า p = 40/100= 0.4

$$Z = \frac{0.4 - 0.35}{\sqrt{\frac{0.35(1-0.35)}{100}}} = 1.05$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานด้านเดียวด้านมาก



5. ตัดสินใจยอมรับH_0

สรุปว่าสัดส่วนของผู้ใช้แชมพูยี่ห้อแฟซ่าอย่างมากเป็น 35 %ที่ระดับนัยสำคัญ 0.01

5.8 การทดสอบค่าผลต่างของประชากร 2 กลุ่ม

ในการทดสอบสามารถแบ่งได้เป็น 2 กรณีคือ

5.8.1 ประชากรเป็นอิสระต่อกัน

ค่าสถิติ	กรณี	การแจกแจง	หมายเหตุ
$\bar{x}_1 - \bar{x}_2$	1.ประชากรมีการแจกแจงปกติ	$Z = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$	
	2.ประชากรมีการแจกแจงปกติ แต่ไม่ทราบค่าแปรปรวน 2.1 $\sigma_1^2 \neq \sigma_2^2$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \sim t_{(\nu)}$	$\nu = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$
	2.2 $\sigma_1^2 = \sigma_2^2$	$t = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_p^2}{n_1} + \dfrac{s_p^2}{n_2}}} \sim t_{(n_1+n_2-2)}$	$S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$
	3.ขนาดตัวอย่างทั้งสอง มีค่าตั้งแต่30	$Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \sim N(0,1)$	
p_1-p_2	ขนาดตัวอย่างใหญ่	$Z = \dfrac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\dfrac{\pi_1(1-\pi_1)}{n_1} + \dfrac{\pi_2(1-\pi_2)}{n_2}}} \sim N(0,1)$	ถ้าสมมติฐานตั้งว่า H_0 : $\pi_1 - \pi_2 = 0$ หมายถึง $\pi_1 = \pi_2$ และถ้าไม่ทราบค่าπ_1และπ_2ประมาณค่าค่าπ_1และπ_2 ด้วย p_1 และ p_2 p_1 และ p_2 ประมาณด้วย p $= \dfrac{X_1 + X_2}{n_1 + n_2} = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ ค่าสถิติที่ใช้คือ $Z = \dfrac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\dfrac{p(1-p)}{n_1} + \dfrac{p(1-p)}{n_2}}}$

ขั้นตอนในการทดสอบเหมือนกับการทดสอบประชากรเดียว สำหรับตัวสถิติทดสอบใช้การแจกแจงดังในตาราง

ตัวอย่างที่16

ฝ่ายวิจัยต้องการทดสอบว่า เวลานอนของนักศึกษาชายและหญิงคณะบัญชี แตกต่างกันหรือไม่ โดยการแจกแจงของเวลานอนมีการแจกแจงปกติ สุ่มตัวอย่างนักศึกษาคณะบัญชี ดังนี้

ชาย			หญิง		
7.0	6.0	8.0	8.0	6.0	7.0
5.5	4.0		7.0	7.0	8.0
9.0	7.0		8.5	8.0	7.5
6.5	6.0		6.0	7.0	7.0
6.0	7.0		6.0	6.0	8.5
7.0	7.5		8.0	7.5	9.5
7.5	7.0		6.5	8.0	
8.0	8.0		7.0	8.0	
7.0	8.0		7.0	6.0	
6.0	4.5		6.0	8.5	
6.0	7.0		7.5	8.0	
8.0	6.0		8.0	7.5	
7.0	8.0		9.0	6.0	
6.0	5.0		6.5	7.5	
10.0	6.5		8.0	6.0	
7.0	9.0		7.0	7.5	
6.0	7.0		7.0	8.0	
6.5	5.5		8.0	9.0	
7.0	6.5		5.5	5.0	
7.0	6.0		8.0	8.0	
5.0	8.5		9.0	7.0	

กำหนดระดับนัยสำคัญ 0.05 จงทดสอบสมมติฐานดังกล่าว

T-Test

Group Statistics

	GENDER	N	Mean	Std. Deviation	Std. Error Mean
TIMESLEE	male	43	6.8140	1.21510	.18530
	female	48	7.3438	1.01665	.14674

Independent Samples Test

			TIMESLEE Equal variances assumed	TIMESLEE Equal variances not assumed
Levene's Test for Equality of Variances	F		.434	
	Sig.		.512	
t-test for Equality of Means	t		-2.264	-2.241
	df		89	82.279
	Sig. (2-tailed)		.026	.028
	Mean Difference		-.5298	-.5298
	Std. Error Difference		.23406	.23637
	95% Confidence Interval of the Difference	Lower	-.99487	-.99998
		Upper	-.06472	-.05961

กำหนดให้

μ_1 แทนเวลานอนเฉลี่ยของนักศึกษาชายคณะบัญชี

μ_2 แทนเวลานอนเฉลี่ยของนักศึกษาหญิงคณะบัญชี

ขั้นตอนการทดสอบสมมติฐาน

1. ตั้งสมมติฐาน

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

2. กำหนดระดับนัยสำคัญ

$\alpha = 0.05$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

3.1 $\sigma_1^2 \neq \sigma_2^2$

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{(\nu)}$$

ได้ค่า p-value = 0.028 น้อยกว่า α

3.2 $\sigma_1^2 = \sigma_2^2$

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \sim t_{(n_1 + n_2 - 2)}$$

ได้ค่า p-value = 0.026 น้อยกว่า α

ตัดสินใจ ปฏิเสธ สมมติฐาน H_0

สรุปว่าเวลานอนหลับโดยเฉลี่ยของนักศึกษาคณะบัญชีเพศชายและหญิงแตกต่างกัน

แบบทดสอบ4

ฝ่ายวิจัยตลาดต้องการทดสอบคะแนนความนิยมของการเลือกใช้คอมพิวเตอร์แตกต่างกันหรือไม่ จึงสุ่มตัวอย่างผู้ใช้คอมพิวเตอร์ สองยี่ห้อคือ IBM กับ HPได้ข้อมูลคะแนนนิยมจากแบบสอบถามดังนี้

ไอบีเอ็ม 81 86 73 77 90 91 75 62 98 74

คอมแพ็ค 89 55 59 64 37 58 35 57 65 68 42 71 69 49 67

กำหนดระดับนัยสำคัญ 0.05

T-Test

Group Statistics

	COMPUTER	N	Mean	Std. Deviation	Std. Error Mean
MARK	IBM	10	80.7000	10.64633	3.36667
	HP	15	59.0000	14.19255	3.66450

Independent Samples Test

			MARK	
			Equal variances assumed	Equal variances not assumed
Levene's Test for Equality of Variances	F		.459	
	Sig.		.505	
t-test for Equality of Means	T		4.114	4.361
	Df		23	22.582
	Sig. (2-tailed)		.000	.000
	Mean Difference		21.7000	21.7000
	Std. Error Difference		5.27511	4.97625
	95% Confidence Interval of the Difference	Lower	10.78760	11.39529
		Upper	32.61240	32.00471

กำหนดให้

μ_1 แทน...

μ_2 แทน..

ขั้นตอนการทดสอบสมมติฐาน

1. ตั้งสมมติฐาน H_0 : ………………..

   H_1 : ………………..

2. กำหนดระดับนัยสำคัญ ....

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

3.1 $\sigma_1^2 \neq \sigma_2^2$ $$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{(\nu)}$$

ได้ค่า p-value = …………………………………..

3.2 $\sigma_1^2 = \sigma_2^2$ $$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}} \sim t_{(n_1 + n_2 - 2)}$$

ได้ค่า p-value = ……………………………………

ตัดสินใจ...................................... H_0

สรุปว่า ....

ตัวอย่างที่19

บริษัทต้องการทดสอบประสิทธิภาพการใช้เครื่องมือชนิด Aและ B โดยทราบว่าค่าแปรปรวนของประชากรไม่เท่ากัน ทำการสุ่มตัวอย่างดังนี้

n_1= 16 $\bar{x}_1$ = 0.41 S_1^2 =0.0025

n_2= 12 $\bar{x}_2$ = 0.38 S_2^2 =0.0004

จงทดสอบสมมติฐานว่า ประสิทธิภาพของเครื่องมือทั้งสองยี่ห้อแตกต่างกันหรือไม่ที่ระดับนัยสำคัญ 0.05

วิธีทำ

กำหนดให้ μ_1 แทนประสิทธิภาพของเครื่องมือยี่ห้อA

μ_2 แทนประสิทธิภาพของเครื่องมือยี่ห้อ **B**

1. ตั้งสมมติฐาน

H_0 : $\mu_1 - \mu_2 = 0$

H_1: $\mu_1 - \mu_2 \neq 0$

2. กำหนดระดับนัยสำคัญ

α = 0.05

3. กำหนดตัวสถิติทดสอบและคำนวณค่า ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ไม่ทราบค่าแปรปรวนและไม่เท่ากันขนาดตัวอย่างเล็ก ตัวสถิติที่ใช้ในการทดสอบคือ

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{(\nu)}$$

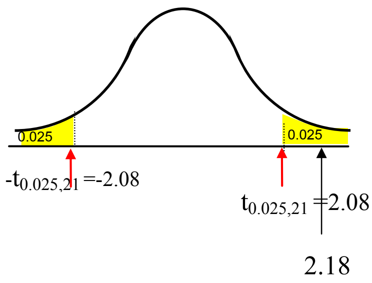
มีการแจกแจงแบบทีด้วยองศาแห่งความเป็นอิสระ ν โดยที่

$$\nu = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}} = 20.8 \cong 21$$

แทนค่า

$$t = \frac{0.41-0.38}{\sqrt{\frac{0.0025}{16}+\frac{0.0004}{12}}} = 2.18$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานแบบสองด้าน



5.ตัดสินใจปฏิเสธ H_0 สรุปว่าประสิทธิภาพการใช้เครื่องมือยี่ห้อ A และ B แตกต่างกันที่ระดับนัยสำคัญ 0.05

5.8.2 ประชากรไม่เป็นอิสระต่อกัน

ในกรณีที่ประชากร 2 กลุ่มไม่เป็นอิสระต่อกัน เช่นทดสอบการทำงานของคู่แฝด การทดลองยาลดความอ้วน ประสิทธิภาพการใช้งานก่อนและหลังการอบรมการใช้ จะสังเกตุเห็นว่าประชากรเป็นประชากร ไม่เป็นอิสระต่อกัน เราเรียกข้อมูลลักษณะนี้ว่าเป็นข้อมูลคู่ (pair data) ลักษณะข้อมูลเป็นดังนี้

เก็บข้อมูลจากประชากรที่มีการแจกแจงปกติมาเป็นคู่ n คู่

$(X_1, Y_1), (X_2, Y_2),\ldots,(X_n, Y_n)$หาค่าผลต่าง $D_1, D_2,\ldots, D_n$ หาค่า $\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$ และ

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}\left(D_i-\overline{D}\right)^2}{n-1}} = S_D = \sqrt{\frac{\sum_{i=1}^{n} D_i^2 - n\overline{D}^2}{n-1}}$$

ได้ว่า

ตัวสถิติ $t = \frac{\overline{D}-\mu_D}{\frac{S_D}{\sqrt{n}}}$ จะมีการแจกแจงแบบทีด้วยองศาแห่งความเป็นอิสระ n-1

หากจะประมาณค่าตัวสถิติ μ_Dสามารถทำได้เหมือนตัวสถิติอื่นๆในบทที่ 7 โดยเขียนขีดจำกัดการประมาณของ μ_D คือ $\overline{D} \pm t_{\frac{\alpha}{2},n-1} \frac{S_D}{\sqrt{n}}$

$\therefore$ (1-α) 100 % ช่วงของการประมาณค่า μ_D คือ

$$\overline{D} - t_{\frac{\alpha}{2},n-1} \frac{S_D}{\sqrt{n}} < \mu_D < \overline{D} + t_{\frac{\alpha}{2},n-1} \frac{S_D}{\sqrt{n}}$$

ถ้าข้อมูลที่สุ่มมาไม่ได้มาจากประชากรที่มีการแจกแจงปกติ และขนาดตัวอย่างมีค่าตั้งแต่ 30 สามารถใช้การแจกแจงปกติประมาณได้

<u>**ตัวอย่างที่18**</u>

ทดสอบความสามารถของนักศึกษาสาขาเลขานุการในการพิมพ์ โดยใช้เครื่องคอมพิวเตอร์ก่อนและหลังการอบรมได้เวลาในการใช้พิมพ์ดังนี้

คนที่	เวลาก่อนอบรม	เวลาหลังอบรม	ผลต่าง
1	55	50	5
2	46	42	4
3	78	70	8
4	61	63	-2
5	52	58	-6
6	45	35	10
7	47	46	1
8	57	52	5
9	71	60	11
10	58	49	9

ถ้าข้อมูลมีการแจกแจงปกติ ใช้ระดับนัยสำคัญ 0.05

18.1ทดสอบว่านักศึกษาสามารถพิมพ์ได้ดีขึ้นหรือไม่หลังการอบรม

18.2 จงหาช่วงการประมาณสำหรับผลต่างที่ได้

กำหนดผลการวิเคราะห์โดยใช้โปรแกรม SPSS ดังนี้

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BEFORE	57.0000	10	10.79094	3.41240
	AFTER	52.5000	10	10.43765	3.30067

Paired Samples Test

			Pair 1 BEFORE - AFTER
Paired Differences	Mean		4.5000
	Std. Deviation		5.48229
	Std. Error Mean		1.73365
	95% Confidence Interval of the Difference	Lower	.5782
		Upper	8.4218
T			2.596
Df			9
Sig. (2-tailed)			.029

วิธีทำ

18.1 ทดสอบว่านักศึกษาสามารถพิมพ์ได้ดีขึ้นหรือไม่หลังการอบรม

กำหนดให้ μ_D แทนผลต่างของเวลาที่ใช้ในการพิมพ์ก่อนและหลังการอบรม

ขั้นตอนการทดสอบมีดังนี้

1.ตั้งสมมติฐาน

$H_0 : \mu_D = 0$

$H_1 : \mu_D > 0$

2. กำหนดระดับนัยสำคัญ $\alpha = 0.05$

3. กำหนดตัวสถิติทดสอบและคำนวณค่า

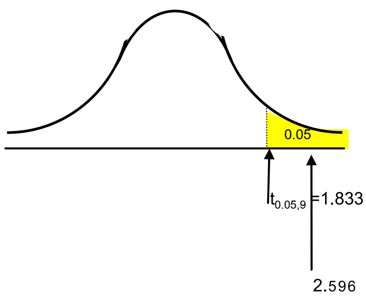
ตัวสถิติที่ใช้ในการทดสอบคือ

$$t = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}}$$

$$t = \frac{4.5 - 0}{\frac{5.48}{\sqrt{10}}} \qquad = 2.596$$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0 ตั้งสมมติฐานด้านเดียว

$\alpha = 0.05$ df = 10-1=9



5. ตัดสินใจปฏิเสธ H_0 สรุปว่าการอบรมมีผลทำให้การพิมพ์ดีขึ้น ที่ระดับนัยสำคัญ 0.05

ผลการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ค่า p-value = 0.029 <0.05
ตัดสินใจปฏิเสธ H_0 สรุปว่าการอบรมมีผลทำให้การพิมพ์ดีขึ้น ที่ระดับนัยสำคัญ 0.05

18.2 จงหาช่วงการประมาณสำหรับผลต่างที่ได้

μ_D = (0.5782,8.4218)

ผลต่างของเวลาที่ใช้ในการพิมพ์ก่อนและหลังการอบรมมีค่าอยู่ระหว่าง 0.5782 และ 8.4218 ด้วยความเชื่อมั่น 95 %

5.9 การทดสอบผลต่างค่าสัดส่วนของประชากร

เมื่อต้องการทดสอบว่า สัดส่วนของประชากร 2 กลุ่ม แตกต่างกันหรือไม่ เช่น สัดส่วนของคนที่ชอบใช้บัตรเครดิต ในอาชีพแม่บ้านกับพ่อบ้านแตกต่างกันหรือไม่ เราสามารถศึกษาความแตกต่างของสัดส่วนของประชากร 2 กลุ่ม (π_1-π_2) ได้เช่นเดียวกับความแตกต่างของค่าเฉลี่ยประชากร 2 กลุ่ม

ให้ p_1 และ p_2 แทนสัดส่วนของสิ่งที่สนใจศึกษาจากตัวอย่างขนาด n_1 และ n_2 ซึ่งสุ่มมาจากประชากรที่ 1 และ 2 ตามลำดับ

ถ้า π_1 และ π_2 เป็นสัดส่วนของประชากรที่ 1 และที่ 2 ตามลำดับ
ขั้นตอนในการทดสอบสมมติฐาน เหมือนการทดสอบค่าเฉลี่ยสองประชากร

ตัวอย่างที่19

จากการศึกษาอัตราการเสี่ยงชีวิตของยา 2 ชนิดคือ Aspirin และ Placebo สำหรับบุคคลผู้เป็นโรคหัวใจได้ผลตามตารางดังนี้

ผล	Aspirin	Placebo
ตาย	104	189
ขนาด	11037	11034

จงทดสอบว่าสัดส่วนการเสี่ยงของผู้เป็นโรคหัวใจที่ใช้ยา Aspirin น้อยกว่า Placebo ที่ระดับนัยสำคัญ 0.05

วิธีทำ

กำหนด π_1 เป็นสัดส่วนการตายเนื่องจากใช้ยา Aspirin

กำหนด π_2 เป็นสัดส่วนการตายเนื่องจากใช้ยา Placebo

1.ตั้งสมมติฐาน

$H_0 \quad : \pi_1 - \pi_2 = 0$

$H_1 \quad : \pi_1 - \pi_2 < 0$

2 . กำหนดระดับนัยสำคัญ $\alpha = 0.05$

3 . กำหนดตัวสถิติทดสอบ และคำนวณค่าตัวสถิติทดสอบนั้น

$p_1 = \dfrac{104}{11037} = 0.0094$ และ $p_2 = \dfrac{189}{11034} = 0.0171$

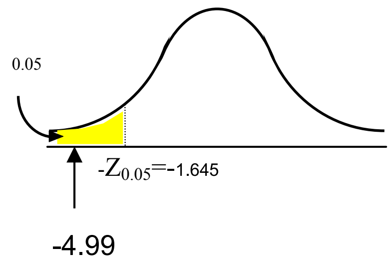
ใช้ $p = \dfrac{104 + 189}{11037 + 11034} = 0.0133$

ตัวสถิติทดสอบ

$$Z = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}} = \frac{(0.0094 - 0.0171) - 0}{\sqrt{\frac{0.0133(1-0.0133)}{11037} + \frac{0.0133(1-0.0133)}{11037}}}$$

$= -4.99$

4. ดูบริเวณการยอมรับและปฏิเสธ H_0



5. ตัดสินใจ ปฏิเสธ H_0

สรุปว่าสัดส่วนการเสี่ยงของผู้เป็นโรคหัวใจที่ใช้ยา Aspirin น้อยกว่า Placebo ที่ระดับนัยสำคัญ 0.05