การแจกแจงทางสถิติ
รูปแบบต่างๆ

Normal Probability Distribution

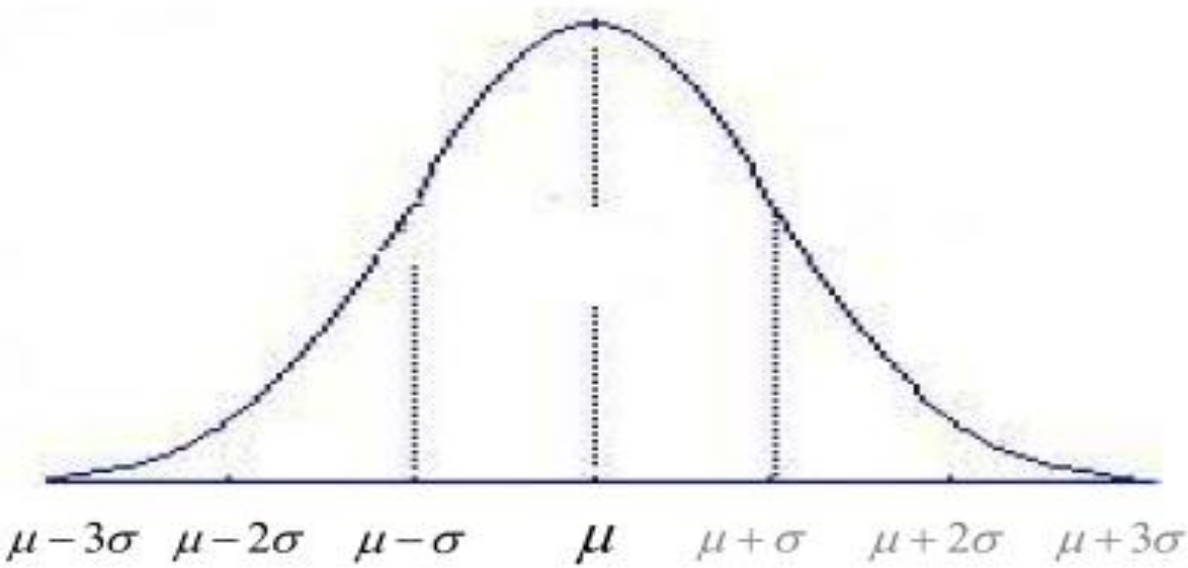
-การแจกแจงความน่าจะเป็นแบบปกติ มีการใช้กันมากที่สุด เนื่อง จากข้อมูลส่วนมากมีการแจกแจงแบบปกติหรือใกล้เคียงแบบปกติ

$$f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad , \quad -\infty \leq x \leq \infty$$

โดยที่ $-\infty \leq \mu \leq \infty$ และ $\sigma > 0$

$$E(x) = \mu$$

$$V(x) = \sigma^2$$



คุณสมบัติ

1. กราฟของฟังก์ชั่นเป็นรูประฆังคว่ำ

2. สมการ $x = \mu$ เป็นแกนสมมาตร

3. มีจุดสูงสุดอยู่จุดเดียว (unimodel) ที่ $x = \mu$ ค่าสูงสุด คือ $\dfrac{1}{\sigma\sqrt{2\pi}}$

4. ค่า Mean = Median = Mode

5. $\lim\limits_{x \to \infty} f(x) = \lim\limits_{x \to -\infty} f(x) = 0$

6. พื้นที่ใต้กราฟรวมกันทั้งหมด มีค่าเท่ากับ 1 นั่นคือ $\int_{-\infty}^{\infty} f(x)dx = 1$

การแจกแจงแบบปกติมาตรฐาน (Standard Normal Distribution)

-ก็คือการแจกแจงแบบปกติที่ $\mu = 0$ และ $\sigma^2 = 1$

-มี z เป็นตัวแปรสุ่ม (standard random variable) ที่แปลงค่ามาจากตัวแปรสุ่มแบบปกติ x โดยที่

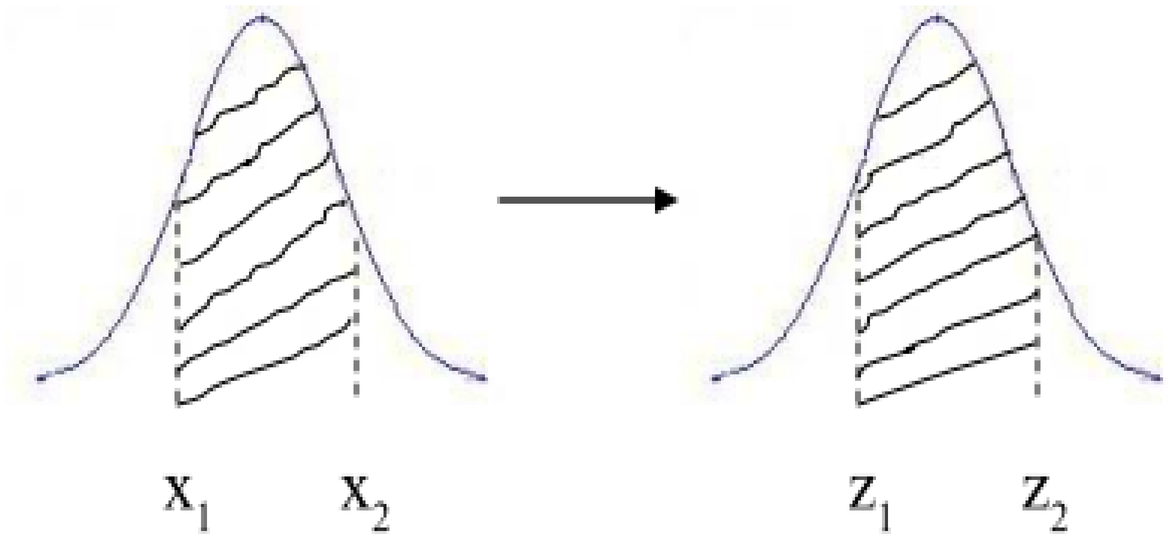
$$\boxed{z = \frac{x - \mu}{\sigma}}$$

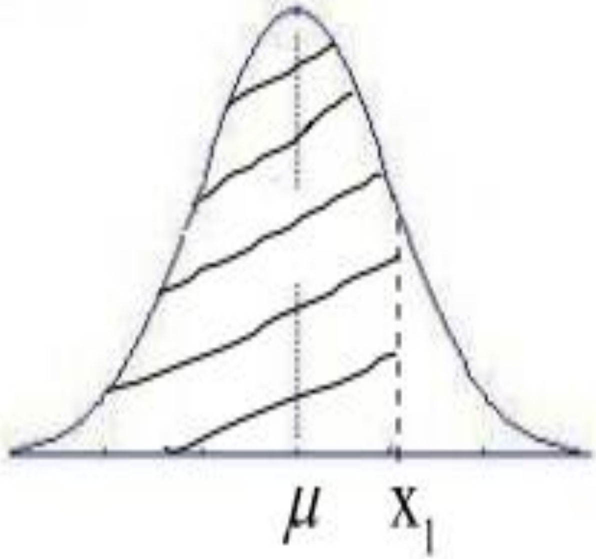
2

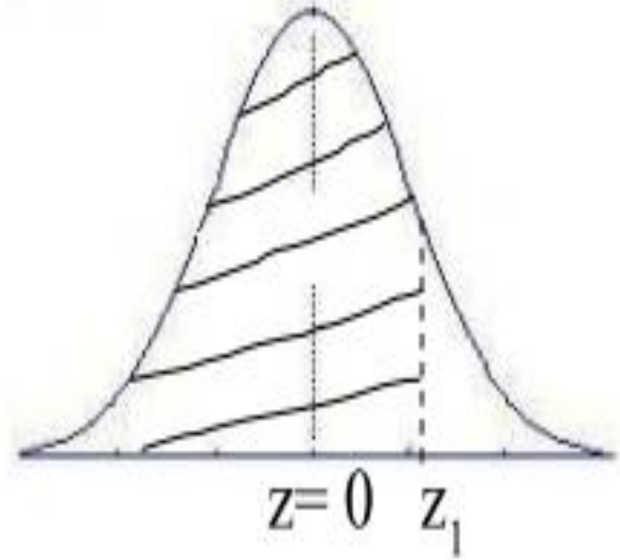
เมื่อ $z \sim N(0, 1)$ และ $f(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$, $-\infty \leq z \leq \infty$

-ในการหา $P(x_1 < x < x_2)$ ก็ปรับค่าเป็น $P(z_1 < z < z_2)$

โดยที่ $z_1 = \frac{x_1 - \mu}{\sigma}$ และ $z_2 = \frac{x_2 - \mu}{\sigma}$



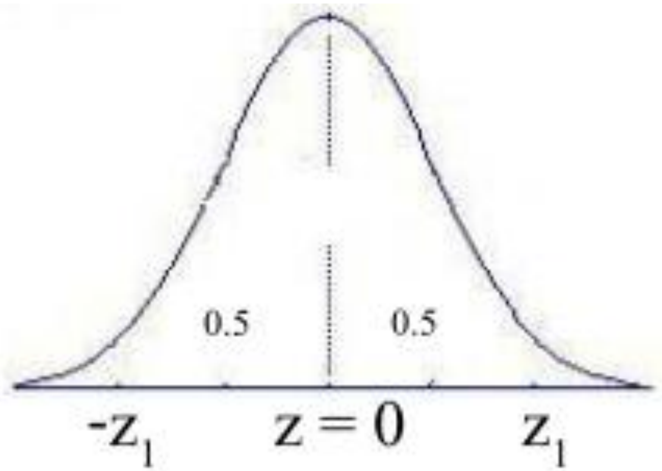




$P(x < x_1) = 1 - P(x > x_1)$

$P(z < z_1) = P(z < \frac{x_1 - \mu}{\sigma})$

-พื้นที่ใต้กราฟทั้งหมดมีค่าเท่ากับ 1 ดังนั้น ครึ่งหนึ่งของพื้นที่ใต้กราฟ จะเท่ากับ 0.5



-ถ้า $P(z < z_1) = 0.8$ จะได้ $P(z > z_1) = 1 - 0.8 = 0.2$ และ

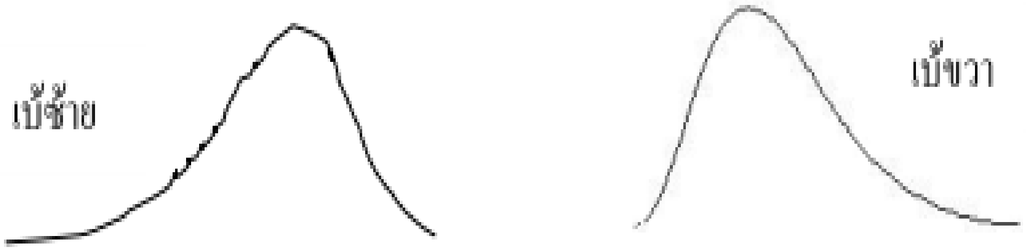
$P(z < -z_1) = 0.2$

$$P(-z_1 < z < z_1) = (0.8 - 0.5)(2) = 0.6$$

หรือ $= 1 - (0.2)(2) = 0.6$

3. Gamma Distribution

-เป็นการแจกแจงที่กราฟไม่มีแกนสมมาตร แต่มีความเบ้



-ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบแกมมา ฟังก์ชั่นแกมมา

ของ X จะมีพารามิเตอร์ 2 ตัว คือ α และ β ฟังก์ชั่นของ X ที่ได้ คือ

$$f(x,\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} e^{\frac{-x}{\beta}} x^{\alpha-1} \quad , \quad x > 0$$

$$= \quad 0 \quad , \quad \text{else}$$

โดยที่ $\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx$

$$E(x) = \alpha\beta$$

$$V(x) = \alpha\beta^2$$

เมื่อ $\alpha = 1$ จะได้ Exponential Probability Function

$$f(x,1,\beta) = \frac{1}{\beta} e^{\frac{-x}{\beta}} \quad , \quad x > 0$$

$$= \quad 0 \quad , \quad \text{else}$$

$$E(x) = \beta$$

$$V(x) = \beta^2$$

Chi-Square Distribution, (χ^2)

-คือ Gamma Distribution ที่ $\alpha = \frac{\upsilon}{2}$ และ $\beta = 2$ โดยที่ $\upsilon \in I^+$

- υ เป็น องศาแห่งความเป็นอิสระ (degree of freedom = df)

$$f(x, \upsilon) = \frac{1}{2^{\frac{\upsilon}{2}}\Gamma(\frac{\upsilon}{2})} e^{\frac{-x}{2}} x^{\frac{\upsilon}{2}-1} \quad , \quad x > 0$$

$$= 0 \quad , \quad \text{else}$$

$$E(x) = \upsilon$$

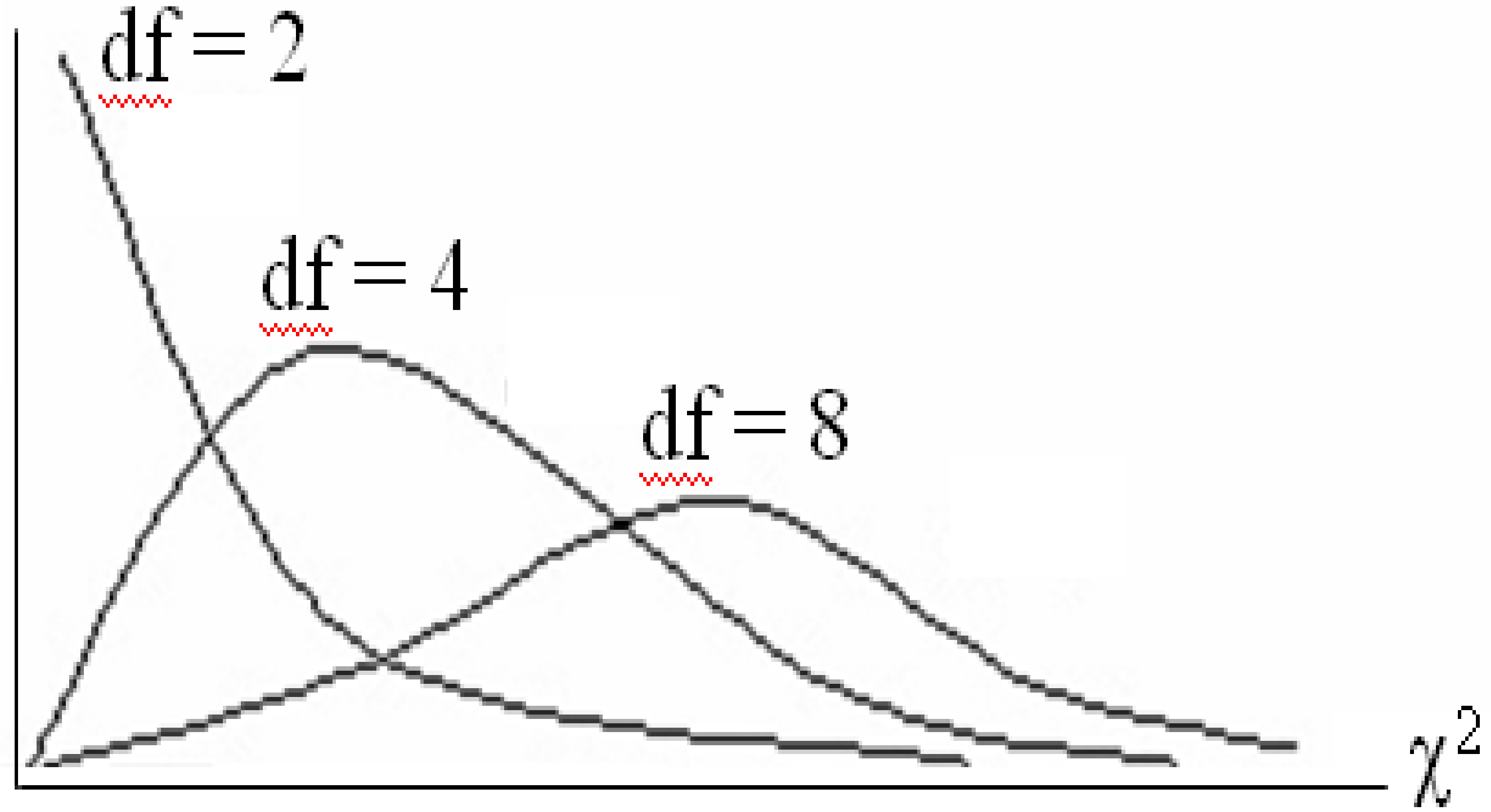
$$V(x) = 2\upsilon$$

-ถ้าตัวแปรสุ่ม $z_1, z_2, \ldots, z_n$ เป็นอิสระต่อกันและมีการแจกแจงแบบปกติมาตรฐาน (standard normal distribution) N(0,1) จะได้

$$z_1^2 + z_2^2 + \ldots + z_n^2 \sim \chi^2_{(n)} \quad (\mathrm{df} = \mathrm{n})$$

χ^2 อ่านว่า ไคสแควร์

การแจกแจงแบบไคสแควร์



คุณสมบัติ

1. $\chi^2 \geq 0$

2. พื้นที่ใต้กราฟรวมกันทั้งหมดมีค่าเท่ากับ 1

3. กราฟมีลักษณะเบ้ขวา แตกแต่งกันตามค่าของ df

4. $P(\chi^2 = x) = 0$

-เราสามารถหาค่าของตัวแปรสุ่มไคสแควร์ χ^2 ได้จากตารางการแจกแจงแบบไคสแควร์ เมื่อเรารู้ค่าความน่าจะเป็น α และค่า df

เช่น $\alpha = 0.1$ และ $df = 5$ จากตารางจะได้ $\chi^2_{0.1,5} = 9.236$

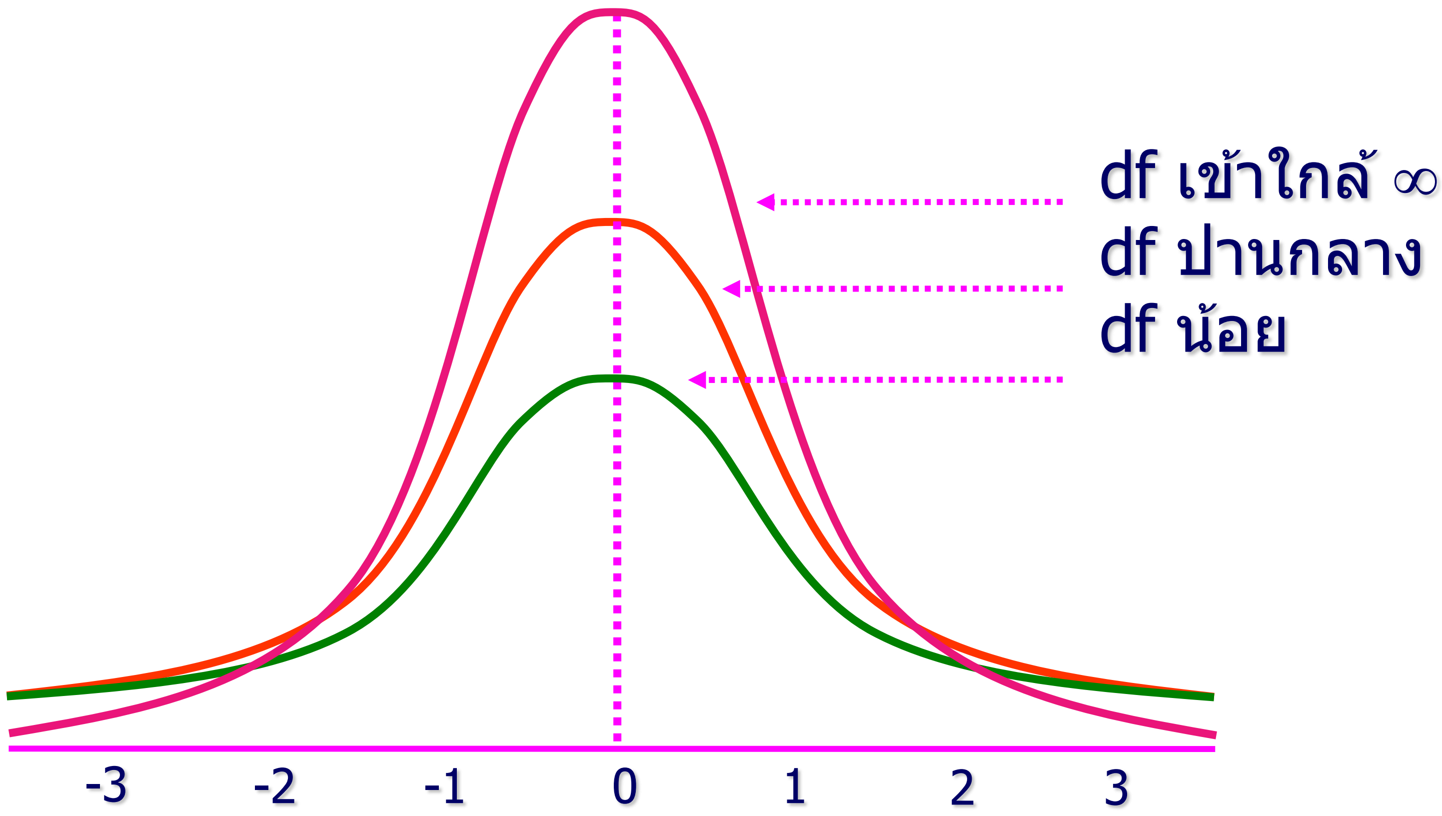
ตัวอย่าง

$$P(\chi^2 \le \chi^2_{0.05,20}) = 0.95 \qquad \chi^2_{0.05,20} \text{ มีค่าเท่าไร}$$

นั่นคือ $P(\chi^2 \ge \chi^2_{0.05,20}) = 0.05$

จากตาราง ที่ $df = 20$ และ $\alpha = 0.05$ จะได้ $\chi^2_{0.05,20} = 31.41$

Student-t Probability Distribution



Student-t Probability Distribution

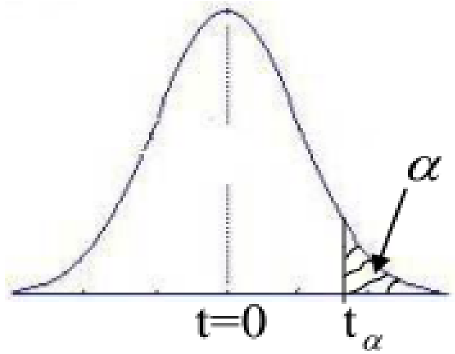
-ถ้า $z \sim N(0,1)$ และ $V \sim \chi^2_{(n)}$ ให้ T เป็นตัวแปรสุ่มที่มีค่าดังนี้

$$T = \frac{z}{\sqrt{V/n}}$$

T จะมีการแจกแจงแบบ Student-t ที่มี degree of freedom = n

และฟังก์ชั่นคือ

$$f(t\,,\,n) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi\,n}}(1+\frac{t^2}{n})^{-(\frac{n+1}{2})} \quad , \quad -\infty < t < \infty$$



-กราฟของฟังก์ชั่นมีลักษณะ

เหมือนกับกราฟ normal

<u>คุณสมบัติ</u>

1. กราฟของฟังก์ชั่นเป็นรูประฆังคว่ำ

2. สมการ $t = 0$ เป็นแกนสมมาตร

3. มีจุดสูงสุดอยู่จุดเดียว (unimodel) ที่ $t = 0$

5. $\lim_{t \to \infty} f(t) = \lim_{t \to -\infty} f(t) = 0$

6. พื้นที่ใต้กราฟรวมกันทั้งหมด มีค่าเท่ากับ 1

-เมื่อทราบค่า α และค่า df ก็สามารถหาค่า t ได้จากตาราง

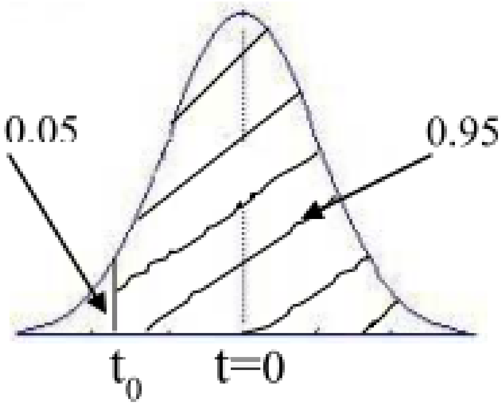
เช่น $\alpha = 0.05$ และ df = 23 จากตารางจะได้ $t_{0.05,23} = 1.714$

<u>ตัวอย่าง</u>

จงหาค่า t_0 ที่ทำให้ เมื่อ df = 23

ก. $P(t > t_0) = 0.95$ ข. $P(-t_0 \leq t \leq t_0) = 0.9$

ก. $P(t > t_0) = 0.95$



จะได้ว่า $t_0 = -t_{0.05,\,23}$

จากตารางที่ $\alpha = 0.05$ และ df = 23

จะได้ $t_{0.05,\,23} = 1.714$ ดังนั้น $t_0 = -1.714$

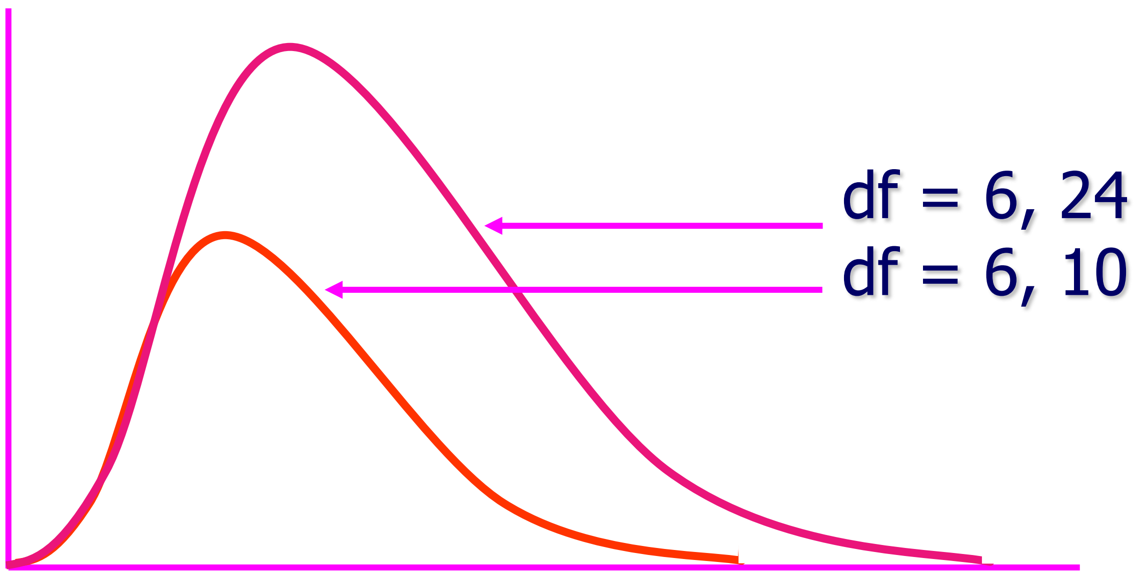
F Probability Distribution

เป็นการแจกแจงของตัวแปร f ที่เกิดจากอัตราส่วนของตัวแปรไคสแควร์ 2 ตัวที่เป็นอิสระต่อกัน

สูตร

$$F_{n1,n2} = \frac{\chi^2 / n_1}{\chi^2 / n_2}$$

F Probability Distribution



F Probability Distribution

-ถ้า U และ V เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยที่ U ~ $\chi^2_{(n)}$ และ V ~ $\chi^2_{(m)}$ จะได้ว่า

$$F = \frac{U/n}{V/m}$$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบ F มี df = n , m

-ฟังก์ชั่นการแจกแจงความน่าจะเป็นของ F คือ

$$G(f, n, m) = \frac{\Gamma(\frac{n+m}{2})\left(\frac{n}{m}\right)^{\frac{n}{2}} f^{(\frac{n}{2}-1)}}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})\left(1+\frac{nf}{m}\right)^{\frac{n+m}{2}}} \quad , \quad f > 0$$

$$= \quad 0 \quad , \quad \text{else}$$

คุณสมบัติ

1. กราฟของฟังก์ชั่นมีลักษณะเบ้ขวา

2. พื้นที่ใต้กราฟรวมกันทั้งหมด มีค่าเท่ากับ 1

3. $F \geq 0$

4. $P(F = f_0) = 0$

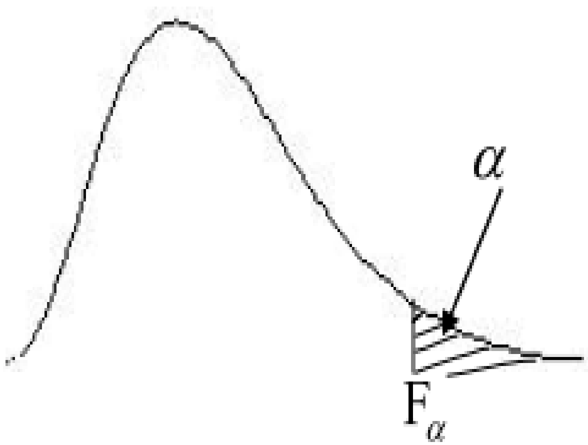
คุณสมบัติ

1. กราฟของฟังก์ชั่นมีลักษณะเบ้ขวา

2. พื้นที่ใต้กราฟรวมกันทั้งหมด มีค่าเท่ากับ 1

3. $F \geq 0$

4. $P(F = f_0) = 0$



-เมื่อทราบค่า α และค่า df (n, m) ก็

สามารถหาค่า $F_{\alpha\ n,m}$ ได้จากตาราง

เช่น $\alpha = 0.05$ และ df = 2, 3 จากตาราง

จะได้ค่า $F_{0.05\ 2,3} = 9.55$

-สำหรับค่า α บางค่า เราไม่สามารถหาค่า $F_{\alpha\ n,m}$ ได้โดยตรงจาก

ตาราง แต่เราอาจหาค่า $F_{\alpha\ n,m}$ ได้ โดยใช้คุณสมบัติ ดังต่อไปนี้

$$\boxed{f_{1-\alpha\ (n,m)} = \frac{1}{f_{\alpha\ (m,n)}}}$$

ตัวอย่างการทดสอบสมมติฐานและรูปแบบการแจกแจงต่างๆ

ตัวอย่าง

ผู้จัดการคิดว่าการลดนำหนักของลูกค้าจะต้องลดลงอย่างน้อย45 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

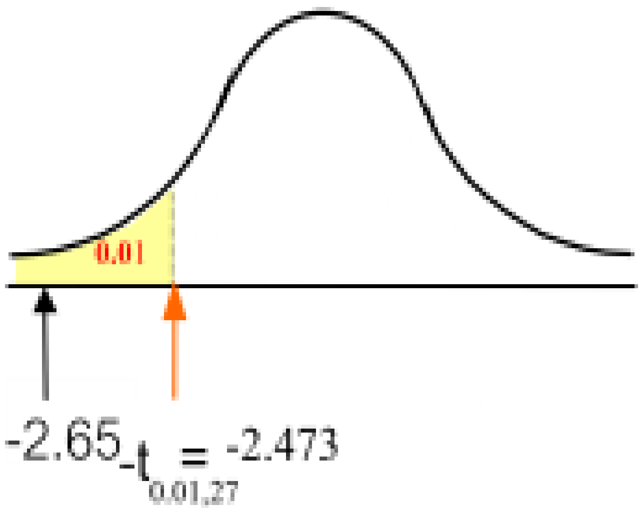
ตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ ไม่ทราบค่าแปรปรวน และขนาดตัวอย่างเล็ก ตัวสถิติที่ใช้ในการทดสอบคือ

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

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

ตัวอย่าง

จงหาค่า f_0 ที่ทำให้ $P(F < f_0) = 0.99$ โดยที่ df = 4,5

$P(F < f_0) = 0.99$ นั่นคือ $P(F > f_0)$

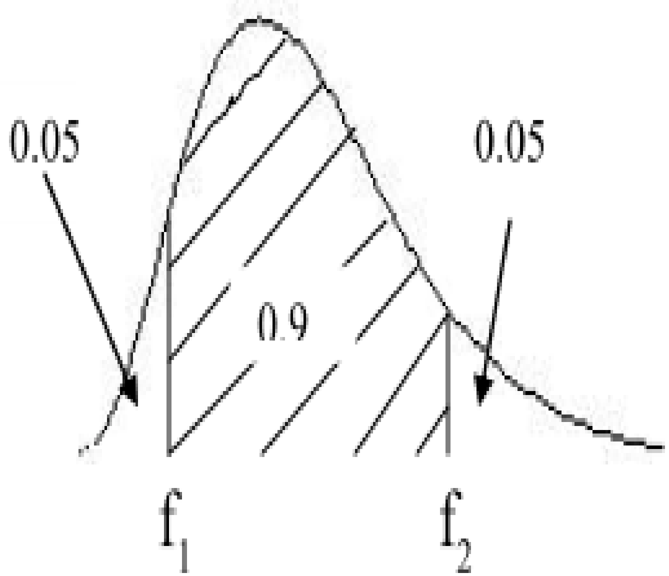
$= 0.01$

จะได้ว่า $f_0 = f_{0.01\ 4,5}$ จากตารางจะได้ $f_{0.01\ 4,5} = 11.39$

ตัวอย่าง

จงหาค่า f_1 และ f_2 ที่ทำให้ $P(f_1 < F < f_2) = 0.90$ และ $P(F \le f_1) = 0.05$

ที่ df เท่ากับ 9 และ 15



จะได้ $P(F>f_2) = 0.05$ และ $P(F>f_1) = 0.95$

ดังนั้น $f_2 = f_{0.05\ 9,15} = 2.59$

และ $f_1 = f_{0.95\ 9,15}$

เนื่องจากไม่สามารถหาค่า $f_{0.95\ 9,15}$ จากตารางได้ ดังนั้นจะใช้

คุณสมบัติ $$f_{1-\alpha\ (n,m)} = \frac{1}{f_{\alpha\ (m,n)}}$$

$$f_{0.95\ 9,15} = \frac{1}{f_{0.05\ 15,9}} = \frac{1}{3.01} = 0.332$$

ตัวอย่าง

ผู้จัดการของบริษัทชุมชนโอทอปแห่งหนึ่งต้องการทราบว่าสัดส่วนของผู้บริโภคแชมพูประคำดีควายอย่างมากเป็น 35 %หรือไม่ จึงทำการสุ่มตัวอย่างลูกค้าจำนวน 100 คน มีผู้ใช้แชมพูยี่ห้อนี้ 40 คน จงทดสอบสมมติฐานดังกล่าวเป็นจริงหรือไม่ ที่ระดับนัยสำคัญ 0.01