องศาความเป็นอิสระ

อาวีพร ปานทอง[*]

องศาความเป็นอิสระ หรือ ระดับขั้นความเสรี มีชื่อทางภาษาอังกฤษว่า degree of freedom เขียนสัญลักษณ์แทนด้วย ν ซึ่งถูกนำไปใช้ในหลายแขนงวิชา อาทิเช่น ฟิสิกส์ วิศวกรรม และสถิติ เป็นต้น ในที่นี้ผู้เขียนขออธิบายเกี่ยวกับองศาความเป็นอิสระในแขนงวิชาสถิติ ที่ถูกนำไปใช้ในเรื่องของการทำงานวิจัยมากมาย ซึ่งยังมีหลายท่านไม่เข้าใจเกี่ยวกับคำคำนี้

ในทางสถิติ องศาความเป็นอิสระ (degree of freedom) คือ จำนวนของข้อมูลในการนำมาคำนวณค่า โดยสามารถแทนค่าของตัวแปรหรือข้อมูลนั้นได้อย่างเป็นอิสระ ถ้าตัวแปรตัวนั้น มี n ค่า จะสามารถแทนค่าเป็นจำนวนใดๆ ได้อย่างอิสระเพียง n-1 ค่า และจะมีอยู่ 1 ค่า ที่ไม่มีความเป็นอิสระในการแทนค่า ถ้าหากเราต้องการจะหาค่าเฉลี่ยจากการที่ไปเก็บข้อมูลจากกลุ่มตัวอย่างมา n ตัวอย่าง แล้วนำมาหาค่าเฉลี่ย เราจะใช้ข้อมูลจากกลุ่มตัวอย่างจำนวน n ตัว ของเราไปใช้หาค่าเฉลี่ยของประชากร ซึ่งเราไม่รู้ค่าแต่เรารู้ว่ามันต้องมีค่าอยู่ค่าหนึ่ง ความที่เรารู้ว่ามันมีค่าหนึ่งนี่แหละทำให้ต้องมีการลบหนึ่ง เพราะว่า ถ้าเราเชื่อว่าค่าเฉลี่ยของกลุ่มตัวอย่างของเราต้องเท่ากับค่าเฉลี่ยของประชากร มันเลยมีเงื่อนไขทำให้ค่าของตัวเอย่างแต่ละตัวไม่สามารถเป็นไปตามใจ หรืออาจจะบอกได้ว่าเป็นค่าใดก็ได้ตามสบายจำนวน n-1 ตัว ส่วนตัวสุดท้ายจะถูก lock ให้คำตอบมีเพียงคำตอบเดียว เพื่อให้ค่าเฉลี่ยของข้อมูลชุดนี้ถูกต้อง ทำให้ความอิสระของตัวแปรหายไปหนึ่งตัว

ยกตัวอย่างเช่น เราต้องการหาค่าของข้อมูลชุดหนึ่งโดยข้อมูลชุดนี้มี 10 จำนวน และมีค่าเฉลี่ยเท่ากับ 5 เนื่องจากเราไม่ทราบค่าที่แท้จริงของข้อมูลชุดนี้ว่าในแต่ละข้อมูลมีค่าเป็นอะไรบ้าง ทราบแต่เพียงว่ามีค่าเฉลี่ยเท่ากับ 5 ดังนั้นเราจึงมีการประมาณค่า (Estimate) ของข้อมูลชุดนี้ว่าแต่ละตัวมีค่าเป็นอะไรได้บ้าง โดยมีเงื่อนไขที่ว่าค่าเฉลี่ยต้องมีค่าเท่ากับ 5 จำนวนข้อมูลที่มีอิสระในการแทนค่าจะมีเพียง 9 จำนวน เท่านั้น ส่วนจำนวนที่ 10 จะไม่มีอิสระในการแทนค่า ต้องขึ้นอยู่กับอีก 9 จำนวนที่แทนค่าไปแล้ว

[*]อาจารย์ประจำสาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์

ถ้าให้ จำนวนที่ 1 แทนค่าอย่างอิสระเป็น 6
จำนวนที่ 2 แทนค่าอย่างอิสระเป็น 2
จำนวนที่ 3 แทนค่าอย่างอิสระเป็น 3
จำนวนที่ 4 แทนค่าอย่างอิสระเป็น 10
จำนวนที่ 5 แทนค่าอย่างอิสระเป็น 5
จำนวนที่ 6 แทนค่าอย่างอิสระเป็น 2
จำนวนที่ 7 แทนค่าอย่างอิสระเป็น 7
จำนวนที่ 8 แทนค่าอย่างอิสระเป็น 2
จำนวนที่ 9 แทนค่าอย่างอิสระเป็น 4
จำนวนที่ 10 ไม่มีสิทธิ์แทนค่า ต้องเป็น 9 เท่านั้น จึงจะทำให้ค่าเฉลี่ยของข้อมูลชุดนี้เท่ากับ 5

$$\bar{x} = \frac{6+2+3+10+5+2+7+2+4+9}{10}$$

จากตัวอย่างข้างต้น อาจจะแทนค่าตัวแปรต่าง ๆ เป็นจำนวนอื่นๆ ได้มากมาย แต่ตัวสุดท้ายจะถูกกำหนดให้มีค่าเป็นไปตามเงื่อนไขของค่าเฉลี่ยของข้อมูลชุดนี้ หรือสมการ a = b + c ถึงสมการนี้จะมีตัวแปรอยู่ 3 ตัวแต่ว่า ถ้าเรารู้ค่าของตัวแปรอยู่ 2 ตัวเราก็จะแก้สมการหาค่าของตัวแปรตัวที่ 3 ได้เลย นั่นก็คือ ค่าของตัวแปรตัวที่ 3 ถูกกำหนดไว้ ทำให้ไม่เป็นอิสระในการกำหนดค่าของตัวแปรแบบตามใจตนเอง จึงทำให้ค่าความเป็นอิสระถูกลดลงไปหนึ่งค่านั้นเอง

Degree of freedom parameters in probability distributions

สถิติที่นำองศาความเป็นอิสระมาใช้ ได้แก่ การแจกแจงไคกำลังสอง (Chi-square Distribution) การแจกแจงที (t- Distribution) การแจกแจงเอฟ (F-Distribution) การวิเคราะห์ความแปรปรวน (Analysis of Variance) การพยากรณ์ (Regression) เป็นต้น

การแจกแจงไคกำลังสอง (Chi-square Distribution)

การแจกแจงไคกำลังสอง (Chi-square Distribution) เริ่มเมื่อปี ค.ศ. 1900 คาร์ล เพียร์สัน (Karl Pearson) ได้เผยแพร่การคิดค้นเรื่องของการแจกแจงไคกำลังสองสำหรับการทดสอบภาวะสารูปสนิทดี (Goodness of Fit Test) รูปร่างการแจกแจงไคกำลังสองขึ้นอยู่กับองศาความเป็นอิสระ คือ ถ้าตัวแปรสุ่ม x มีการแจกแจงปกติโดยมีค่าเฉลี่ยเท่ากับ μ และส่วนเบี่ยงเบนมาตรฐาน σ เมื่อสุ่มตัวอย่าง

จากประชากรมา n ตัวอย่าง จะได้ค่าของข้อมูล คือ $x_1, x_2, x_3, x_4, x_5, \ldots, x_n$ ซึ่งนำมาแปลงข้อมูลที่ได้เป็นค่ามาตรฐาน Z แล้วยกกำลังสองของค่ามาตรฐานจะได้ค่าไคกำลังสอง เขียนแทนด้วย χ^2 ที่มีองศาความเป็นอิสระเท่ากับหนึ่ง ดังนี้

ข้อมูล	Z_n	$(Z_n)^2$	χ^2	ν
x_1	$Z_1=\frac{x_1-\mu}{\sigma}$	$Z_1^2=\frac{(x_1-\mu)^2}{\sigma^2}$	$\chi^2_{1(1)}$	1
x_2	$Z_2=\frac{x_2-\mu}{\sigma}$	$Z_2^2=\frac{(x_2-\mu)^2}{\sigma^2}$	$\chi^2_{2(1)}$	1
x_3	$Z_3=\frac{x_3-\mu}{\sigma}$	$Z_3^2=\frac{(x_3-\mu)^2}{\sigma^2}$	$\chi^2_{3(1)}$	1
x_4	$Z_4=\frac{x_4-\mu}{\sigma}$	$Z_4^2=\frac{(x_4-\mu)^2}{\sigma^2}$	$\chi^2_{4(1)}$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$Z_n=\frac{x_n-\mu}{\sigma}$	$Z_n^2=\frac{(x_n-\mu)^2}{\sigma^2}$	$\chi^2_{n(1)}$	1

เมื่อนำค่ามาตรฐานกำลังสองมาบวกกัน ผลลัพธ์ที่ได้เท่ากับค่าไคกำลังสองซึ่งมีองศาความเป็นอิสระเท่ากับจำนวนของค่ามาตรฐานกำลังสองมาบวกกัน

ผลบวกค่า Z^2	χ^2	ν
$Z_1^2+Z_2^2=\sum_{i=1}^{2}\frac{(x_i-\mu)^2}{\sigma^2}$	$=\chi^2_{(2)}$	2
$Z_1^2+Z_2^2+Z_3^2=\sum_{i=1}^{3}\frac{(x_i-\mu)^2}{\sigma^2}$	$=\chi^2_{(3)}$	3
$\vdots$	$\vdots$	$\vdots$
$Z_1^2+Z_2^2+Z_3^2+\ldots+Z_{n-1}^2=\sum_{i=1}^{n-1}\frac{(x_i-\mu)^2}{\sigma^2}$	$=\chi^2_{(n-1)}$	n-1

จาก $\chi^2_{(n-1)}=\sum_{i=1}^{n-1}\frac{(x_i-\mu)^2}{\sigma^2}$

แทนค่า μ ด้วย $\bar{x}$ จะได้

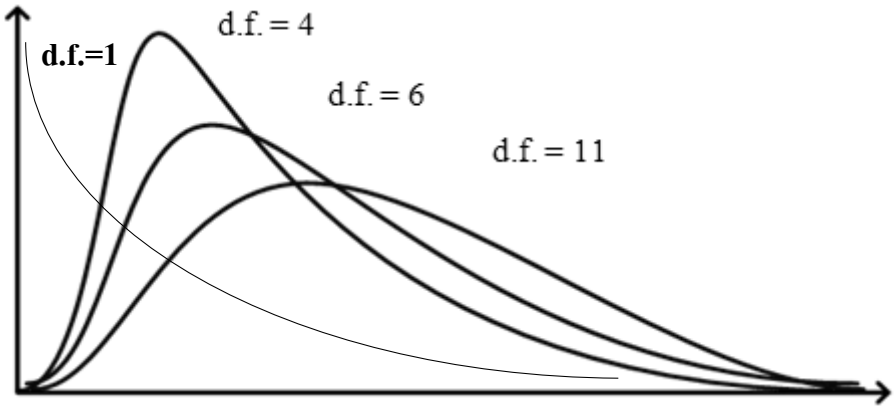
$$\chi^2_{(n-1)}=\sum_{i=1}^{n-1}\frac{(x_i-\bar{x})^2}{\sigma^2}$$

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

เมื่อทราบว่า $(n-1)\frac{s^2}{\sigma^2}$ มีการแจกแจงไคกำลังสอง โดยมีองศาความเป็นอิสระเท่ากับ n-1 นำมาใช้เป็นประโยชน์ในเรื่องของการอนุมานทางสถิติ ซึ่งเกี่ยวข้องกับการประมาณค่าและการทดสอบความแปรปรวนของประชากร

คุณสมบัติของการแจกแจงไคกำลังสอง

1. มีลักษณะโค้งไม่สมมาตรและมียอดเดียว โดย χ^2 มีค่า 0 ถึง ∞
2. เส้นโค้งมีลักษณะแตกต่างกันตามองศาความเป็นอิสระ
    - ถ้า องศาความเป็นอิสระ มีค่าเท่ากับ 1 หรือ 2 โค้งของการแจกแจงไคกำลังสองจะมีลักษณะคล้ายรูปตัว J กลับข้าง
    - ถ้า องศาความเป็นอิสระ มีค่ามากกว่าหรือเท่ากับ 3 โค้งของการแจกแจงไคกำลังสองจะมีลักษณะเบ้ขวา และเมื่อองศาความเป็นอิสระ มีค่ามากขึ้นโค้งของการแจกแจงจะมีลักษณะใกล้เคียงโค้งปกติ



แผนภาพแสดงลักษณะของเส้นโค้งการแจกแจงไคกำลังสอง

ค่าเฉลี่ยเลขคณิตของการแจกแจงไคกำลังสอง เท่ากับ $E\left[\chi_\nu^2\right]=\nu$

ค่าความแปรปรวนของการแจกแจงไคกำลังสอง เท่ากับ $Var\left[\chi_\nu^2\right]=2\nu$

การแจกแจงที (t- Distribution)

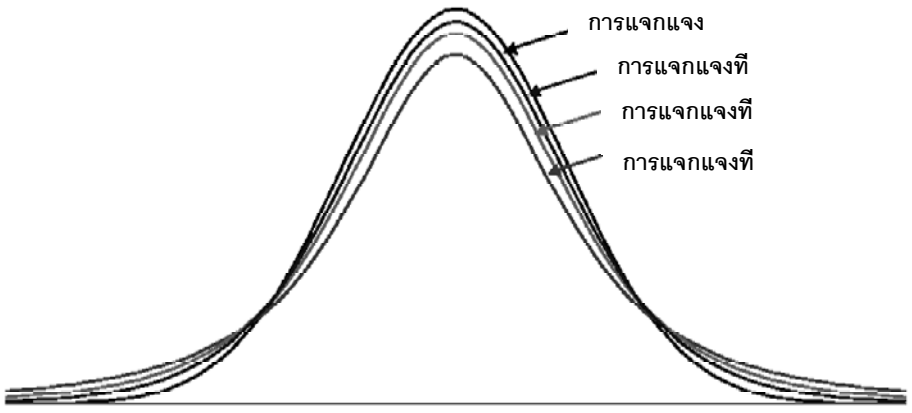
ผู้ค้นพบคือ วิลเลียม ซีลี กอสเซต (William Sealy Gosset) และออกเผยแพร่โดยใช้ชื่อว่า Student ดังนั้นการแจกแจงที จึงเรียกอีกชื่อหนึ่งว่า การแจกแจงสติวเดนท์ที (Student's t- Distribution) ซึ่งมีลักษณะการแจกแจงคล้ายคลึงการแจกแจงปกติ เมื่อตัวอย่างที่สุ่มมา มีขนาดใหญ่มาก ค่าส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง (s) จะมีค่าใกล้เคียงกับค่าส่วนเบี่ยงเบนมาตรฐานของประชากร (σ) การแจกแจงทีจะเหมือนการแจกแจงปกติ แต่หากกรณีที่สุ่มกลุ่มตัวอย่างมามีขนาดน้อย ($n < 30$) จะทำให้การแจกแจงที แตกต่างไปจากการแจกแจงปกติ โดยการแจกแจงทีจะมีความโด่งของโค้งน้อยกว่าการแจกแจงปกติ พื้นที่ทั้งหมดภายใต้โค้งของการแจกแจงทีเท่ากับหนึ่งเหมือนการแจกแจงปกติ การแจกแจงทีเป็นกลุ่มของการแจกแจง ซึ่งรูปโค้งของการแจกแจงทีจะขึ้นอยู่กับขนาดของกลุ่มตัวอย่าง (Sample) หรือองศาความเป็นอิสระ

การแจกแจงที่มีความสัมพันธ์การแจกแจงปกติและการแจกแจงไคกำลังสอง ดังนี้

$$t_{(\nu)} = \frac{Z}{\sqrt{\frac{\chi^2_{(\nu)}}{\nu}}}$$

ค่าเฉลี่ยเลขคณิตของการแจกแจงไคกำลังสอง เท่ากับ $E\left[t_{(\nu)}\right]=0$

ค่าความแปรปรวนของการแจกแจงไคกำลังสอง เท่ากับ $Var\left[t_{(\nu)}\right]=\frac{\nu}{\nu-2}$ เมื่อ $\nu>2$



แผนภาพแสดงลักษณะของเส้นโค้งการแจกแจงที

การแจกแจงเอฟ (F- Distribution)

การแจกแจงเอฟ เป็นการตั้งชื่อเพื่อเป็นเกียรติแก่ เซอร์โรนัลด์ เอ. ฟิชเชอร์ (Sir Ronald A. Fisher) การแจกแจงเอฟมีหน้าที่ในการศึกษาถึงอัตราส่วนความแปรปรวนของตัวอย่างสุ่ม ซึ่งได้มาจากประชากรที่มีการแจกแจงปกติ 2 กลุ่มที่เป็นอิสระต่อกัน ถ้า x_1 และ x_2 เป็นตัวแปรสุ่มอิสระต่อกัน และมีการแจกแจงไคกำลังสอง หรือการแจกแจงปกติ ด้วยองศาความเป็นอิสระ ν_1 และ ν_2 ตามลำดับ ดังนั้น การแจกแจงเอฟเป็นอัตราส่วนของการแจกแจงไคกำลังสองที่หารด้วยองศาความเป็นอิสระ จะได้ว่า

$$F=\frac{\left[\frac{x_1}{\nu_1}\right]}{\left[\frac{x_2}{\nu_2}\right]}$$

ถ้าสุ่มตัวอย่างขนาด n_1 และ n_2 จากประชากรที่มีการแจกแจงปกติ โดยมีค่าเฉลี่ยเท่ากับ μ_1, μ_2 และความแปรปรวนเท่ากับ σ_1, σ_2 ตามลำดับ จากการแจกแจงไคกำลังสองที่มีองศาความเป็นอิสระเท่ากับ n-1 จะได้ว่า

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

จะได้

$$F=\frac{\frac{\chi^2_{n_1-1}}{(n_1-1)}}{\frac{\chi^2_{n_2-1}}{(n_2-1)}}=\frac{\left(\frac{(n_1-1)\frac{S_1^2}{\sigma_1^2}}{(n_1-1)}\right)}{\left(\frac{(n_2-1)\frac{S_2^2}{\sigma_2^2}}{(n_2-1)}\right)}=\frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}=\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}$$

เป็นการแจกแจงเอฟที่มีองศาความเป็นอิสระ $\nu_1 = n_1\text{-}1$ และ $\nu_2 = n_2\text{-}1$ ตามลำดับ

หากกรกณีที่ $\sigma_1^2=\sigma_2^2$ จะได้ว่า

$$F=\frac{S_1^2}{S_2^2}$$

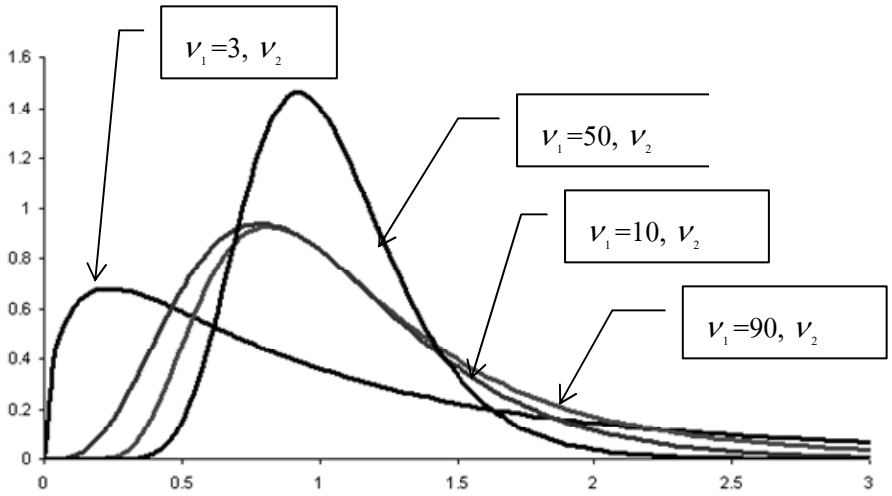
ค่าเฉลี่ยเลขคณิตของการแจกแจงเอฟ เท่ากับ $E\left[F_{(\nu_1,\nu_2)}\right]=\frac{\nu_2}{\nu_2-2}$ เมื่อ $\nu_2>2$

ค่าความแปรปรวนของการแจกแจงเอฟ เท่ากับ $Var\left[F_{(\nu_1,\nu_2)}\right]=\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$

เมื่อ $\nu_2 > 4$

คุณสมบัติของการแจกแจงเอฟ

1. มีลักษณะโค้งไม่สมมาตรและมียอดเดียว โดย F มีค่าตั้งแต่ 0 ถึง ∞
2. เส้นโค้งมีลักษณะแตกต่างกันตามองศาความเป็นอิสระ
   - ถ้า ν มีค่าเท่ากับ 1 หรือ 2 โค้งมีลักษณะคล้ายรูปตัว J กลับข้าง
   - ถ้า องศาความเป็นอิสระ มีค่ามากกว่าหรือเท่ากับ 3 โค้งของการแจกแจงเอฟจะมีลักษณะเบ้ขวา และเมื่อองศาความเป็นอิสระ มีค่ามากขึ้นโค้งของการแจกแจงจะมีลักษณะใกล้เคียงโค้งปกติ



แผนภาพแสดงลักษณะของเส้นโค้งการแจกแจงเอฟ

****

เอกสารอ้างอิง

ศรีวรรณ ฤกษ์ภูริทัต. “การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม” **หลักสถิติ** ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยนเรศวร. 1(1): 156-161.

Good, IJ (1973). "What Are Degrees of Freedom?". **The American Statistician**. 27 (5): 227–228.

Wikipedia the free encyclopedia. (2009, August) “Degrees of Freedom (http://en.wikipedia.org /wiki/Degrees_of_freedom_(statistics))” **Degrees of freedom (statistics)**.

Walker, H. M. (1940, April). "Degrees of Freedom". **Journal of Educational Psychology.** 31 (4): 253–269. doi:10.1037/h0054588.

****