

ฟังก์ชันที่ใช้ในโปรแกรม R

[Function in R]

ฟังก์ชันที่มีอยู่ในโปรแกรม R Package เบื้องต้นสามารถจำแนกได้ 5 กลุ่มใหญ่ๆดังต่อไปนี้

- ☐ Numeric Functions ☐ Statistical Functions ☐ Character Functions
☐ Useful Functions ☐ Graphical Functions

ในที่นี้จะกล่าวถึงบางฟังก์ชันที่จำเป็นเท่านั้น

2.1 Graphical Functions

เป็นฟังก์ชันประเภทที่ใช้งานเกี่ยวกับการสร้างกราฟ ซึ่งมีฟังก์ชันมาตรฐานที่ควรรู้ดังต่อไปนี้

ฟังก์ชัน	วัตถุประสงค์การใช้งาน
<code>plot()</code>	ใช้สร้างกราฟแบบ Scatterplot ของ 2 ตัวแปร
<code>hist()</code>	ใช้สร้างกราฟแบบ แบบแท่งต่อเนื่อง ของ 1 ตัวแปร
<code>barplot()</code>	ใช้สร้างกราฟแบบ แท่งแบบแยกแท่ง ของ 1 ตัวแปร
<code>boxplot()</code>	ใช้สร้างกราฟแบบ กล่อง ของ 1 ตัวแปร
<code>dotplot()</code>	ใช้สร้างกราฟแบบ จุด ของ 1 ตัวแปร
<code>piechart()</code>	ใช้สร้างกราฟแบบ วงกลม ของ 1 ตัวแปร

ผู้ใช้สามารถระบุ ขนาดตัวอักษร สี ชนิดของเส้น และอื่นๆ นอกจากผู้ใช้อย่างยังสามารถสร้างกราฟหลายรูปแบบมาผสมกันได้ รายละเอียดของการสร้างกราฟ สามารถทำได้ โดยใช้ฟังก์ชัน `help` และระบุชื่อฟังก์ชันที่ต้องการ ตัวอย่างการใช้คำสั่ง `help(barplot)` จะปรากฏวินโดวส์ แสดงข่าวสารของการใช้ ฟังก์ชัน

รายละเอียดต่างๆผู้ใช้อย่างสามารถดูตัวอย่างการใช้ที่มีมาให้เมื่อ ร้องขอด้วยคำสั่ง `help()`

2.2 Statistical Functions

เป็นฟังก์ชันที่ใช้สำหรับการวิเคราะห์ข้อมูล สามารถจัดเป็นประเภทต่างๆได้ 3 ประเภทดังนี้

- ☐ Statistical Standard Functions
- ☐ Statistical Probability Functions
- ☐ Statistical Standard Method Functions

ในที่นี้จะกล่าวถึงเฉพาะ Statistical Standard Functions

มีดังนี้

① Statistical Standard Functions

เป็นฟังก์ชันที่สามารถเรียกใช้งานโดยพื้นฐานทั่วไปโดยประกอบไปด้วยฟังก์ชันต่อไปนี้

ฟังก์ชัน	วัตถุประสงค์การใช้งาน
<code>mean(x, trim=0, na.rm=FALSE, ...)</code>	ใช้ หาค่าเฉลี่ย ของ object x โดยไม่นำค่า Missing ออก ดังนั้นถ้าข้อมูลมี Missing โปรแกรม จะให้ผลลัพธ์เป็น NA ดังตัวอย่างต่อไปนี้ $x = c(2, 3, NA, 4)$; <code>mean(x)</code> จะได้ผลคือ NA <code>mean(x, na.rm = TRUE)</code> จะได้ผลคือ 3
<code>sd(x, na.rm=FALSE)</code>	ใช้ หาค่าส่วนเบี่ยงเบนมาตรฐาน (s ไม่ใช่ σ) ของ object x โดยถือว่า Object เป็นข้อมูล ตัวอย่าง n และไม่นำค่า Missing ออก
<code>var(x, na.rm=FALSE)</code>	ใช้ หาค่าความแปรปรวน (s^2 ไม่ใช่ σ^2) ของ object x โดยถือว่า Object เป็นข้อมูล ตัวอย่าง n และไม่นำค่า Missing ออก
<code>COV(x, na.rm=FALSE)</code>	ใช้ หาค่าความแปรปรวนร่วม ของ object x และ y โดยถือว่า Object เป็นข้อมูล ตัวอย่าง n และไม่นำค่า Missing ออก
<code>median(x, na.rm=FALSE)</code>	ใช้ หาค่ามัธยฐาน ของ object x โดยไม่นำค่า Missing ออก
<code>range(x, na.rm=FALSE)</code>	ใช้ หาค่าพิสัย ของ object x โดยไม่นำค่า Missing ออก

② Statistical Standard Method Functions

เป็นฟังก์ชันที่ใช้เกี่ยวกับการวิเคราะห์ข้อมูลสถิติ ซึ่งจำแนกได้ 3 กลุ่มใหญ่ๆ ดังนี้คือ

- ☐ การวิเคราะห์ข้อมูล สำหรับข้อมูล แบบ ต่อเนื่อง (Continuous Response)
- ☐ การวิเคราะห์ข้อมูล สำหรับข้อมูล แบบ ไม่ต่อเนื่อง (Discrete Response)
- ☐ การวิเคราะห์ข้อมูล แบบนอนพารามेटริก (Nonparametric)

ฟังก์ชันที่จะกล่าวถึงต่อไปนี้เป็นฟังก์ชัน เพียงบางฟังก์ชัน ที่อยู่ในชุด Package stats

① ฟังก์ชันสำหรับการวิเคราะห์ข้อมูลแบบต่อเนื่อง (Continuous Response)

โดยส่วนใหญ่ฟังก์ชันในกลุ่มนี้จะใช้กับข้อมูลที่สามารถนำมาวิเคราะห์แบบ Parametric

ฟังก์ชัน	วัตถุประสงค์การใช้งาน
<code>t.test()</code>	ใช้ ทดสอบค่าเฉลี่ยสำหรับ 1 หรือ 2 กลุ่มตัวอย่าง ด้วยค่าสถิติ t
<code>oneway.test()</code> หรือ <code>aov()</code>	ใช้ ทดสอบค่าเฉลี่ย สำหรับ หลายกลุ่มตัวอย่าง ด้วยค่าสถิติ F โดยการวิเคราะห์ความแปรปรวน ซึ่งสามารถใช้ วิเคราะห์ได้ทั้งแบบ ทางเดียว และ แบบ หลายทาง
<code>var.test()</code>	ใช้ ทดสอบความแปรปรวนสำหรับ 2 กลุ่มตัวอย่าง ด้วยค่าสถิติ F
<code>bartlett.test()</code>	ใช้ ทดสอบความแปรปรวน หลายกลุ่มตัวอย่าง ด้วยค่าสถิติ Bartlett
<code>levene.test()</code>	ใช้ ทดสอบความแปรปรวน หลายกลุ่มตัวอย่าง ด้วยค่าสถิติ Levene
<code>cor.test()</code>	ใช้ ทดสอบความสัมพันธ์ด้วยค่าสถิติ t แปรปรวน
<code>levene.test()</code>	ใช้ ทดสอบความแปรปรวน หลายกลุ่มตัวอย่าง ด้วยค่าสถิติ Levene
<code>lm()</code>	ใช้ ในการวิเคราะห์การถดถอย ซึ่งสามารถใช้ได้ทั้ง Simple Regression และ Multiple Regression พร้อมค่า สถิติต่างๆ ที่ใช้ในการทดสอบ ซึ่งจะมีฟังก์ชันที่สามารถใช้ในการวิเคราะห์ร่วมกันอีกหลายฟังก์ชัน เช่น <code>coef()</code> <code>fitted()</code> <code>residuals()</code> <code>dwtest()</code> และอื่นๆ

นอกจากนี้ยังมีฟังก์ชันอื่นๆที่สามารถนำมาใช้วิเคราะห์ได้โดยไม่ต้องเรียก Package เพิ่มเติม

② ฟังก์ชันสำหรับการวิเคราะห์ข้อมูลแบบไม่ต่อเนื่อง (Discrete Response)

โดยส่วนใหญ่ฟังก์ชันในกลุ่มนี้จะใช้กับข้อมูลแบบแจกแจงนับในรูปของ สัดส่วน หรือ ร้อยละ

ฟังก์ชัน	วัตถุประสงค์การใช้งาน
<code>binom.test()</code>	ใช้ ทดสอบค่าสัดส่วน 1 กลุ่ม ด้วยค่าสถิติ Binomial และ Sign Test
<code>prop.test()</code>	ใช้ ทดสอบค่าสัดส่วน หลายกลุ่ม ด้วยค่าสถิติ Chi-Square
<code>fisher.test()</code>	ใช้ ทดสอบค่าสัดส่วนสำหรับกลุ่มตัวอย่างย่อย ด้วยค่าสถิติ Fisher
<code>chisq.test()</code>	ใช้ ทดสอบที่เกี่ยวข้องกับ ค่าสถิติ Chi-Square เช่น ทดสอบเกี่ยวกับ Independence หรือ Goodness of fit

นอกจากนี้ยังมีฟังก์ชันอื่นๆที่สามารถนำมาใช้วิเคราะห์ได้โดยไม่ต้องเรียก Package เพิ่มเติม

③ ฟังก์ชันสำหรับการวิเคราะห์ข้อมูลแบบนอนพารามेटริก (Nonparametric)

โดยส่วนใหญ่ฟังก์ชันในกลุ่มนี้จะใช้กับข้อมูลที่ไม่สามารถใช้การทดสอบแบบ พารามेटริกได้

ฟังก์ชัน	วัตถุประสงค์การใช้งาน
<code>wilcox.test()</code>	ใช้ ทดสอบสำหรับ 1 กลุ่ม / 2 กลุ่มตัวอย่าง ด้วยค่าสถิติ Wilcoxon
<code>kruskal.test()</code>	ใช้ ทดสอบสำหรับ หลายกลุ่มตัวอย่าง ด้วยค่าสถิติ Kruskal-Wallis
<code>friedman.test()</code>	ใช้ ทดสอบทดสอบสำหรับ หลายกลุ่มตัวอย่าง แบบ Two-Ways Analysis of Variance ด้วยค่าสถิติ Friedman
<code>cor.test(..., method = "kendall")</code> หรือ <code>"spearman")</code>	ใช้ ทดสอบที่เกี่ยวข้องกับความสัมพันธ์แบบ Nonparametric ซึ่งใช้ฟังก์ชันเดียวกันกับ การหาความสัมพันธ์แบบ Parametric ที่ใช้ค่าของ Pearson แต่ด้วยวิธีของ Nonparametric จะใช้วิธีของ Spearman และ Kendall แต่ใช้ ฟังก์ชัน <code>cor.test()</code> เหมือนกัน