

กราฟรูปกล่อง

ความหมาย

กราฟรูปกล่อง (Box Plot) หมายถึงการนำเสนอข้อมูลเชิงปริมาณในรูปแบบกราฟเพื่อแสดงลักษณะการแจกแจงตำแหน่งและการกระจายของข้อมูล โดยใช้รูปกล่องแสดงตำแหน่งข้อมูล ขอบซ้ายและขอบขวาของรูปกล่องเป็นค่าควอร์ไทล์ที่ 1 และควอร์ไทล์ที่ 3 ตามลำดับ เส้นตั้งฉากในรูปกล่องเป็นค่ามัธยฐาน (ควอร์ไทล์ที่ 2) ใช้หมวดแนว (whisker) ซึ่งเป็นเส้นตรงที่ลากจากรูปกล่องไปยังค่าน้อยที่สุดและค่ามากที่สุด หมวดแนวแสดงการกระจายของข้อมูลพร้อมทั้งเน้นส่วนหางของการแจกแจง

ประวัติความเป็นมา

ถ้าเราสุ่มข้อมูลจากการแจกแจงปกติ \bar{X} และ S จะเป็นค่าประมาณที่ดีของ μ และ σ ตามลำดับ แต่ถ้าเราสุ่มข้อมูลจากการแจกแจงที่เบี่ยงเบนจากการแจกแจงปกติโดยเบี่ยงไปทางที่ยาวขึ้น \bar{X} และ S จะเป็นค่าประมาณที่ไม่เหมาะสมของ μ และ σ ตามลำดับ ดังนั้นในปี ค.ศ.1977 จอห์น ดับบลิว ทูกี (John.W.Tukey) ได้นำเสนอข้อมูลเชิงปริมาณด้วยกราฟรูปกล่อง เพื่อแสดงการแจกแจงของข้อมูล โดยสร้างรูปกล่องและแสดงตำแหน่งของมัธยฐานว่ามีระยะห่างจากควอร์ไทล์ที่ 1 และที่ 3 อย่างไร เพื่อดูความสมมาตรของการแจกแจงพร้อมทั้งเน้นส่วนหางของการแจกแจงโดยสร้างหมวดแนวเพื่อชี้พิสัยด้านซ้ายและด้านขวาที่ตกนอกรูปกล่อง หมวดแนวจะมีความยาวมากกว่าพิสัยระหว่างควอร์ไทล์ (interquartile range) ได้ไม่เกิน 50% ของพิสัยระหว่างควอร์ไทล์ ทูกีเรียกขอบของรูปกล่อง

(ควอร์ไทล์ที่ 1 และควอร์ไทล์ที่ 3) ว่า Tukey's hinges และเรียกกราฟนี้ว่า box and whisker plot

วิธีการสร้างกราฟรูปกล่อง

ขั้นตอนการสร้างกราฟรูปกล่องมีดังต่อไปนี้

1. คำนวณมัธยฐานหรือควอร์ไทล์ที่ 2 (Q_2) ควอร์ไทล์ที่ 1 (Q_1) และควอร์ไทล์ที่ 3 (Q_3)
2. คำนวณพิสัยระหว่างควอร์ไทล์ (IQR) ซึ่งเท่ากับควอร์ไทล์ที่ 3 ลบด้วยควอร์ไทล์ที่ 1
3. คำนวณส่วนขยายของข้อมูลเพื่อแสดงตำแหน่งของค่าสังเกตผิดปกติซึ่งเป็นค่านอกกลุ่ม (outlier) ดังต่อไปนี้

รั้วชั้นใน (inner fence) ด้านซ้ายเท่ากับ

$$Q_1 - 1.5(IQR)$$

รั้วชั้นในด้านขวาเท่ากับ $Q_3 + 1.5(IQR)$

รั้วชั้นนอก(outer fence) ด้านซ้ายเท่ากับ

$$Q_1 - 3(IQR)$$

รั้วชั้นนอกด้านขวาเท่ากับ $Q_3 + 3(IQR)$

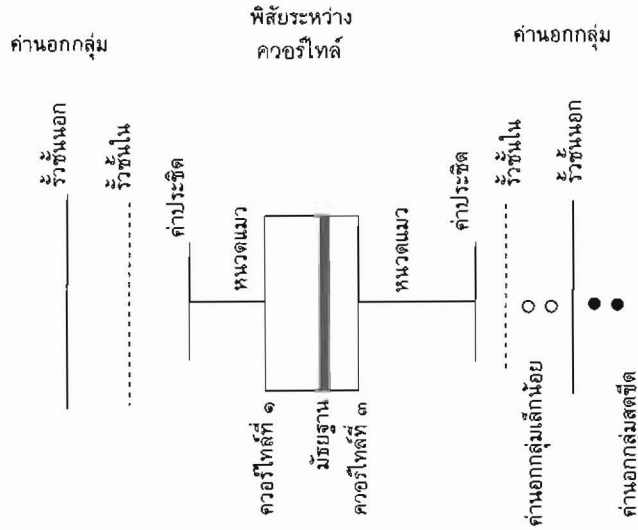
ค่าประชิด (adjacent values) คือ ค่าสุดขีดมากที่สุดและน้อยที่สุด (the most extreme values) ที่ตกอยู่ในรั้วชั้นใน

ค่านอกกลุ่มเล็กน้อย (mild outliers) เป็นค่าสังเกตที่ตกอยู่ระหว่างรั้วชั้นนอกและรั้วชั้นใน

ค่านอกกลุ่มสุดขีด (extreme outliers) เป็นค่าสังเกตที่ตกอยู่นอกรั้วชั้นนอก

4. สร้างรูปกล่องโดยสร้างสี่เหลี่ยมผืนผ้าความยาวสั้นสุดที่ Q_1 และ Q_3 ลากเส้นตั้งฉากภายในสี่เหลี่ยมผืนผ้าที่จุด Q_2 ลงจุดค่าประชิดในแนวกึ่งกลางของด้านกว้างของสี่เหลี่ยมผืนผ้า ลากเส้นตรงขนานกับแกนนอนไปยังค่าประชิด เรียกเส้น

นี้ว่า หนวดแมว ถ้ามีค่านอกกลุ่มเล็กน้อยจะลงจุดโดยใช้วงโปร่งและค่านอกกลุ่มสุดขีดจะลงจุดโดยใช้วงทึบ ดังรูปที่ 1

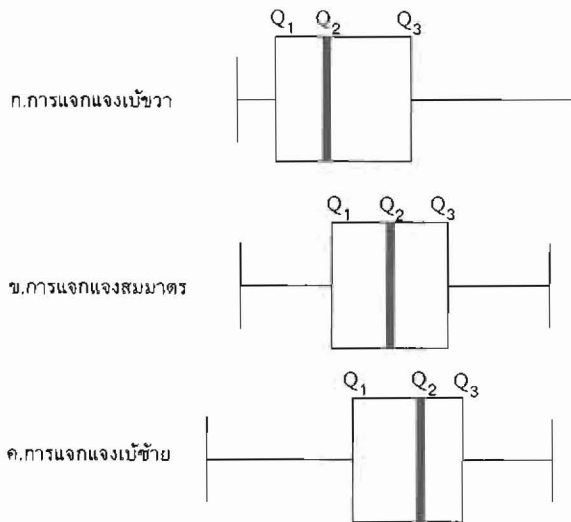


รูปที่ 1 ส่วนประกอบต่างๆ ของกราฟรูปกล่อง

การแปลผล

รูปกล่องประกอบด้วยจำนวนข้อมูลที่อยู่ส่วนกลางร้อยละ 50 ของข้อมูลทั้งหมด ในการบอกรูปร่างการแจกแจงจะพิจารณาจากระยะห่างระหว่างมัธยฐาน กับควอร์ไทล์ที่ 1 และควอร์ไทล์ที่ 3 สำหรับกรณีข้อมูลที่ไม่มีค่านอกกลุ่มดังนี้

1. ถ้ามัธยฐานเข้าใกล้ควอร์ไทล์ที่ 1 มากกว่าควอร์ไทล์ที่ 3 แสดงว่าข้อมูลมีการแจกแจงเบ้ขวา รูปที่ 2 (ก)
2. ถ้ามัธยฐานอยู่กึ่งกลางระหว่างควอร์ไทล์ที่ 1 และควอร์ไทล์ที่ 3 แสดงว่าข้อมูลมีการแจกแจงสมมาตร รูปที่ 2 (ข)
3. ถ้ามัธยฐานเข้าใกล้ควอร์ไทล์ที่ 3 มากกว่าควอร์ไทล์ที่ 1 แสดงว่าข้อมูลมีการแจกแจงเบ้ซ้าย รูปที่ 2 (ค)



รูปที่ 2 กราฟรูปกล่องแสดงการแจกแจงแบบต่างๆ

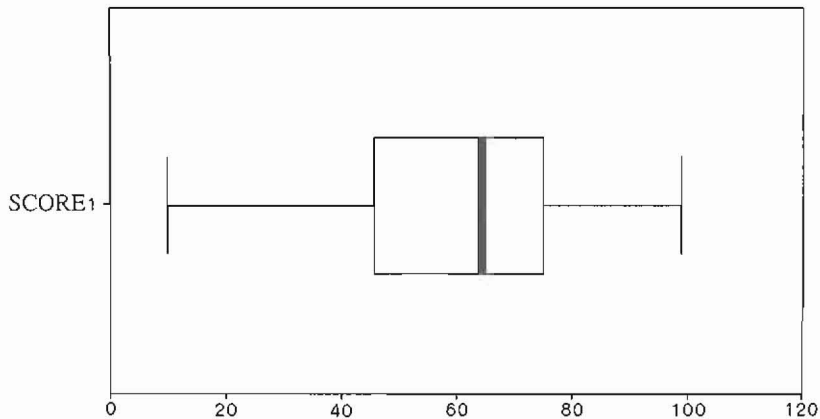
ตัวอย่างที่ 1 คะแนนสอบความถนัดของนักเรียน 50 คนเป็นดังนี้

34	56	78	89	62	23	10	79	99	10
25	56	46	67	87	95	56	75	89	60
68	87	75	70	45	34	23	55	67	55
72	45	65	44	65	64	35	80	67	65
87	67	46	34	89	54	67	61	66	63

จากข้อมูลคะแนนข้างต้นคำนวณค่าต่าง ๆ ได้ดังตารางต่อไปนี้

	ค่าสถิติ		ค่าสถิติ
ร้อยละนอกด้านขวา	162.85	ร้อยละนอกด้านซ้าย	-42
ร้อยละในด้านขวา	118.88	ร้อยละในด้านซ้าย	1.88
ปลายหมวดแฉกด้านขวา	99	ปลายหมวดแฉกด้านซ้าย	10
ค่าสูงสุด	99	ค่าต่ำสุด	10
ควอร์ไทล์ที่ 3	75	ควอร์ไทล์ที่ 1	45.75
มัธยฐาน	61.5	พิสัยระหว่างควอร์ไทล์	29.25

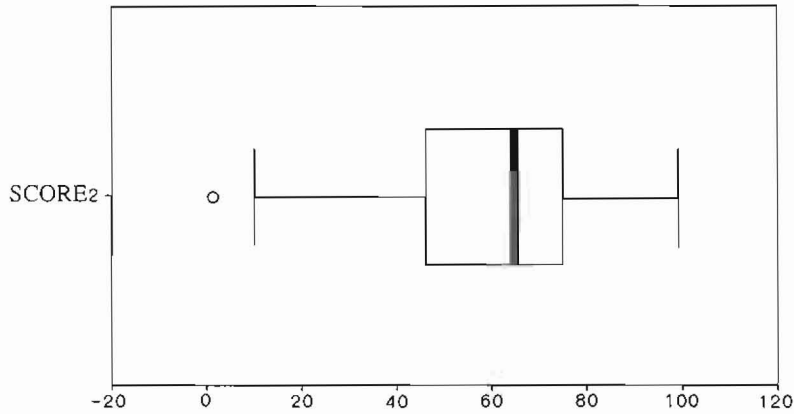
จากค่าสถิติดังกล่าวสามารถสร้างกราฟรูปกล่องได้ดังแสดงในรูปที่ 3



รูปที่ 3 กราฟรูปกล่องของคะแนนสอบความถนัดของนักเรียน 50 คน

จากรูปที่ 3 พบว่าการแจกแจงเบ้ซ้ายเนื่องจากมัธยฐานเข้าใกล้ควอร์ไทล์ที่ 3 มากกว่าควอร์ไทล์ที่ 1

ในกรณีที่มีค่านอกกลุ่มบอกรหัสผลตจะแสดงค่านอกกลุ่มโดยพิมพ์วงไปรงหรือวงทีบ ดังรูปที่ 4 แสดงกราฟรูปกล่อง จากข้อมูลตัวอย่างที่ 1 โดยเปลี่ยนข้อมูลค่าสุดท้ายในบรรทัดแรกจาก 10 เป็น 1 รูปที่ 4 แสดงค่านอกกลุ่มเล็กน้อย 1 ค่าคือ 1

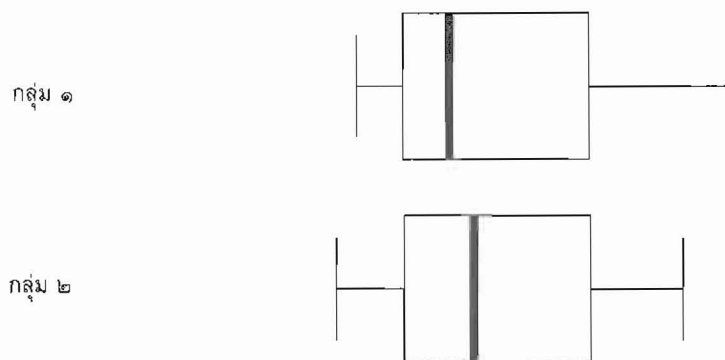


รูปที่ 4 กราฟรูปกล่องของคะแนนสอบความถนัดของนักเรียน 50 คน กรณีมีค่า outliers

ประโยชน์ของกราฟรูปกล่อง

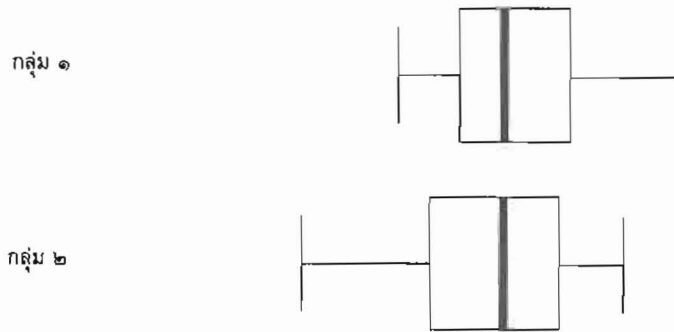
1. ใช้กราฟรูปกล่องแสดงสารสนเทศ (information) เบื้องต้นเกี่ยวกับการแจกแจงข้อมูลซึ่งช่วยให้มองเห็นภาพรวมของข้อมูลได้สะดวกและรวดเร็ว
2. ใช้กราฟรูปกล่องแสดงควอร์ไทล์และค่า outliers
3. ใช้กราฟรูปกล่องเปรียบเทียบการแจกแจงของตัวแปรตั้งแต่สองตัวขึ้นไปได้อย่างมีประสิทธิภาพ โดยสร้างกราฟรูปกล่องข้อมูลตัวอย่างที่สุ่มมาจากแต่ละประชากรแล้วเปรียบเทียบมัธยฐานและพิสัยระหว่างควอร์ไทล์ซึ่งวัดการกระจายของข้อมูลดังนี้

(1) ประชากรกลุ่มที่ 1 มีมัธยฐานน้อยกว่าประชากรกลุ่มที่ 2 ประชากรทั้ง 2 กลุ่มมีการแจกแจงเบ้ แต่มีการกระจายเท่ากันแสดงดังรูปที่ 5



รูปที่ 5 กราฟรูปกล่องแสดงการแจกแจงเบ้ของประชากร 2 กลุ่ม แต่มีการกระจายเท่ากัน

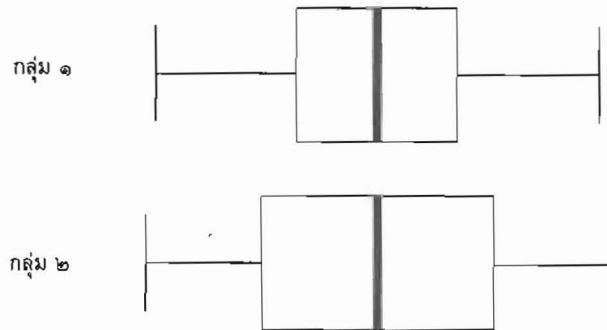
(2) ประชากรกลุ่มที่ 1 และกลุ่มที่ 2 มีมัธยฐานเท่ากัน ประชากรทั้งสองกลุ่มมีการแจกแจงเบ้แต่ ประชากรกลุ่มที่ 1 มีการกระจายแคบกว่าประชากรกลุ่มที่ 2 (แสดงดังรูปที่ 6)



รูปที่ 6

กราฟรูปกล่องแสดงการแจกแจงเบ้ของประชากร 2 กลุ่ม ที่มีมัธยฐานเท่ากันและมีการกระจายต่างกัน

(3) ประชากรกลุ่มที่ 1 และกลุ่มที่ 2 มีมัธยฐานเท่ากัน และมีการแจกแจงสมมาตรแต่ประชากรกลุ่มที่ 1 มีการกระจายแคบกว่าประชากรกลุ่มที่ 2 (แสดงดังรูปที่ 7)



รูปที่ 7

กราฟรูปกล่องแสดงการแจกแจงสมมาตรของประชากร 2 กลุ่ม ที่มีมัธยฐานเท่ากันแต่มีการกระจายต่างกัน

บรรณานุกรม

- Attwood,G.,Skipworth,G.and Dyer,G. **Statistics 1**. Oxford : Heinemann Publishers(Oxford) Ltd, 1994.
- Berenson,M.L.and Levine, D.M. **Statistics forBusiness and Economics**. 2nd ed. New Jersey : Prentice - Hall, 1993.
- Devore, J.L. **Probability and Statistics for Engineering and the Sciences**. California: Brooks/ Cole Publishing Company, 1982.
- Koopmans,L.M. **An Introduction to Contemporary Statistics**. Boston:Duxbury Press. 1981.
- ราชบัณฑิตยสถาน. **ศัพท์คณิตศาสตร์ ฉบับราชบัณฑิตยสถาน**. กรุงเทพฯ : โรงพิมพ์มหาวิทยาลัยราชภัฏวชิราวุฒวิทยาลัย, 2538.
- อรพินทร์ เจียรพงษ์. “การแจกแจงความถี่,” ใน **สารานุกรมศึกษาศาสตร์**. จัดพิมพ์เป็นที่ระลึกเนื่องในมหามงคลสมัยฉลองสิริราชสมบัติครบ 50 ปี พุทธศักราช 2539. หน้า 549-555 คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2539.