บทที่ 5
การประยุกต์ใช้การแจกแจงที่มีลักษณะหางยาว

5.1 กล่าวนำ

เนื่องจากการแจกแจงหางยาวที่ทำการศึกษาในครั้งนี้ประกอบไปด้วย การแจกแจงลอกนอร์มอล การแจกแจงโคชี และการแจกแจงลาปลาซ ซึ่งแต่ละการแจงแจงมีความเหมาะสมในการนำไปประยุกต์ใช้กับงานด้านต่างๆแตกต่างกันไป เช่น การแจกแจงลอกนอร์มอลซึ่งมีการนำไปประยุกต์ใช้อย่างกว้างขวาง เช่น นำไปอธิบายการกระจายรายได้ อธิบายน้ำหนักตัว อธิบายจำนวนเงินผลตอบแทนที่บริษัทประกันภัยต้องจ่ายให้กับลูกค้า หรือแม้แต่อธิบายความเร็วลม เป็นต้น การแจกแจงโคชีนำไปใช้ในทางฟิสิกส์เกี่ยวกับทฤษฎีกลศาสตร์ และ สนามแม่เหล็กไฟฟ้า ส่วนการแจกแจงลาปลาซนำไปใช้ในการอธิบายอัตราการเจริญเติบโตของบริษัท จากความแตกต่างในด้านการนำไปประยุกต์ใช้ดังกล่าวแล้วข้างต้น ดังนั้นในที่นี้จึงจะได้นำเสนอตัวอย่างการประยุกต์ใช้ของแต่ละการแจกแจงดังต่อไปนี้

5.2 การประยุกต์ใช้การแจกแจงลอกนอร์มอล

การแจกแจงลอกนอร์มอลเป็นการแจกแจงที่ถูกนำไปประยุกต์ใช้กับงานต่างๆอย่างกว้างขวาง เช่น นำไปอธิบายความเร็วลม อธิบายน้ำหนักตัว อธิบายจำนวนเงินผลตอบแทนที่บริษัทประกันภัยต้องจ่ายให้กับลูกค้าในแต่ละปี เป็นต้น ซึ่งในเรื่องของจำนวนเงินผลตอบแทนที่บริษัทประกันภัยต้องจ่ายให้กับลูกค้าแต่ละปีนั้นพบว่า การแจกแจงแกมมามีความสามารถในการอธิบายได้ดี แต่ถ้าข้อมูลมีลักษณะของการกระจายตัวสูงและมีความเบ้มาก การแจกแจงลอกนอร์มอลจะมีความสามารถในการอธิบายได้ดีกว่า[1]

นอกจากการประยุกต์ที่กล่าวไปแล้วข้างต้น การแจกแจงลอกนอร์มอลยังถูกพบว่ามีลักษณะการนำไปประยุกต์ใช้สัมพันธ์กับการแจกแจงหางหนัก(Heavy Tailed Distribution) ซึ่งการแจกแจงหางหนักก็มีลักษณะของการนำไปประยุกต์ใช้ได้กว้างขวางเช่นกัน เช่น นำไปอธิบายการกระจายรายได้ การอธิบายการแจกแจงของขนาดแฟ้มข้อมูลในอินเทอร์เน็ต เป็นต้น ซึ่งจากความสัมพันธ์ของการแจกแจงทั้งสองจึงทำให้การแจกแจงลอกนอร์มอลถูกนำไปประยุกต์กับงานด้านต่างๆได้

[1] D. Papush, G. Patrik, F. Podgaits. "Approximations of the aggregate loss distribution" :176-186

เพิ่มมากขึ้น ทั้งนี้เนื่องจากมีผู้นำข้อมูลที่อธิบายได้ด้วยการแจกแจงหางหนักมาทดลองอธิบายด้วยการแจกแจงลอกนอร์มอล เพื่อหาการแจกแจงที่เหมาะสมที่สุด ซึ่งก็รวมถึงการกระจายรายได้ด้วย ซึ่ง Aitchison และ Brown ได้ทำการศึกษาและพบว่า การแจกแจงลอกนอร์มอลมีความเหมาะสมในการอธิบายรายได้ที่ต่ำๆ ส่วนการแจกแจงหางหนักมีความเหมาะสมในการอธิบายรายได้ที่ค่อนข้างสูง

จากการประยุกต์ใช้ที่กว้างขวางดังกล่าวของการแจกแจงลอกนอร์มอลนั้น ในที่นี้จะได้นำเสนอตัวอย่างการนำการแจกแจงลอกนอร์มอลไปอธิบายจำนวนเงินผลตอบแทนที่บริษัทประกันภัยต้องจ่ายให้กับลูกค้า และการอธิบายการกระจายรายได้ดังนี้

ตัวอย่าง 5.1 การแจกแจงความถี่ของจำนวนเงินผลตอบแทนจากการประกันภัยกับการแจกแจงลอกนอร์มอล

เป็นที่ทราบกันโดยทั่วไปว่า ธุรกิจประกันภัย หรือการทำประกันนั้น จะมีหลายรูปแบบที่แตกต่างกัน เช่น การประกันชีวิต การประกันสุขภาพ การประกันทรัพย์สิน เป็นต้น และธุรกิจประกันภัยเหล่านี้ในปัจจุบันเป็นธุรกิจที่สร้างรายได้และได้รับความสนใจจากประชาชนโดยทั่วไปเป็นอย่างดี โดยเฉพาะอย่างยิ่งในสหรัฐอเมริกา ดังนั้นบริษัทประกันภัยของสหรัฐอเมริกาจึงสนใจที่จะปรับปรุงธุรกิจดังกล่าวให้มีมาตรฐานที่ดีและสร้างผลประโยชน์และความพอใจให้กับทั้งลูกค้าและบริษัทเอง ซึ่งการปรับปรุงในครั้งนี้ทางบริษัทเลือกที่จะมุ่งไปที่การประกันทรัพย์สิน ทั้งนี้เนื่องจากเป็นที่สังเกตของผู้เชี่ยวชาญในปัจจุบันว่า ในแต่ละปีบริษัทต้องเสียเงินประกันทรัพย์สินให้กับลูกค้าเป็นจำนวนมาก โดยเฉพาะปีที่เกิดมหันตภัยทางธรรมชาติ เช่นในปี พ.ศ. 2542 บริษัทต้องเสียเงิน 28.6 พันล้านดอลล่าร์ โดย 85%ของเงินจำนวนนี้มีสาเหตุมาจากมหันตภัยทางธรรมชาติ และ 15% มีสาเหตุจากการกระทำของมนุษย์ และในปี พ.ศ. 2535 นั้น ต้องเสียเงินมากที่สุดถึง 32.4 พันล้านดอลล่าร์ ซึ่งมีสาเหตุหลักมาจากการเกิดมหันตภัยธรรมชาติเช่นกัน โดยการปรับปรุงครั้งนี้ ทางบริษัทได้ทำการปรับปรุงดัชนีผลตอบแทนที่ต้องจ่ายให้ลูกค้า ซึ่งดัชนีผลตอบแทนดังกล่าวนี้จะเป็นตัวที่บอกว่าทางบริษัทจะต้องจ่ายผลตอบแทนให้กับลูกค้าจำนวนเท่าไรในกรณีที่ลูกค้าประสบภัยทางทรัพย์สิน สำหรับการสร้างดัชนีดังกล่าวทางบริษัทได้อาศัยข้อมูลการจ่ายผลตอบแทนให้กับลูกค้าในแต่ละปีของบริษัทในอดีต ตั้งแต่ปี พ.ศ. 2493-12542 และทำการศึกษาการแจกแจงของข้อมูลก่อนที่จะใช้สถิติทำการวิเคราะห์ความสัมพันธ์ต่างๆของข้อมูลต่อไป ซึ่งในส่วนของการศึกษาการแจกแจงของข้อมูลนั้นพบว่าการแจกแจงแบบลอกนอร์มอลมีความสามารถในการอธิบายลักษณะ

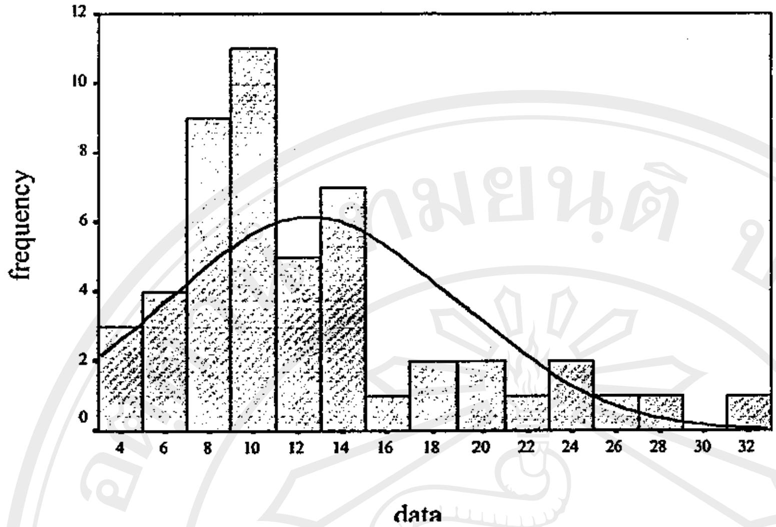
ของข้อมูลได้ดีกว่าการแจกแจงอื่นๆ[2] เนื่องจากข้อมูลมีการกระจายตัวและมีความเบ้มากดังกล่าวแล้วข้างต้น ซึ่งในที่นี้จะได้แสดงให้เห็นถึงลักษณะของข้อมูลดังนี้

ตาราง 5.1 แสดงข้อมูลจำนวนเงินผลตอบแทน (พันล้านดอลล่าร์) ที่บริษัทประกันจ่ายให้กับลูกค้าต่อปี ตั้งแต่ปีพ.ศ. 2493 - 2542

ปีที่	จำนวนเงิน	ปีที่	จำนวนเงิน	ปีที่	จำนวนเงิน	ปีที่	จำนวนเงิน	ปีที่	จำนวนเงิน
1	4.9	11	7.3	21	8.2	31	11.6	41	14
2	3.6	12	8.6	22	9.8	32	9.7	42	18.8
3	5.1	13	9.5	23	13.4	33	12.9	43	32.4
4	9.3	14	10.6	24	8.4	34	10.1	44	22.3
5	8.2	15	13.6	25	10.2	35	12.4	45	14.9
6	7.8	16	7.3	26	11.3	36	10.7	46	25.7
7	6.8	17	19.1	27	9.5	37	20.5	47	23.8
8	13.4	18	4.7	28	14.2	38	13.9	48	16.9
9	7.2	19	6.3	29	8.3	39	24.6	49	17.8
10	5.8	20	10.8	30	9.2	40	12.9	50	28.6

จากข้อมูลจำนวนเงินผลตอบแทนที่บริษัทประกันต้องจ่ายให้กับลูกค้าในแต่ละปีในตารางที่ 5.1 เมื่อนำมาเขียนเป็นกราฟแสดงการแจกแจงความถี่ของข้อมูลดังรูปที่ 5.1 พบว่ากราฟมีลักษณะเบ้ขวาคล้ายกับโค้งความถี่ของการแจกแจงลอกนอร์มอลดังนั้นจึงทำการตรวจสอบว่าข้อมูลชุดดังกล่าวมีความเหมาะสมที่จะอธิบายด้วยการแจกแจงลอกนอร์มอลหรือไม่

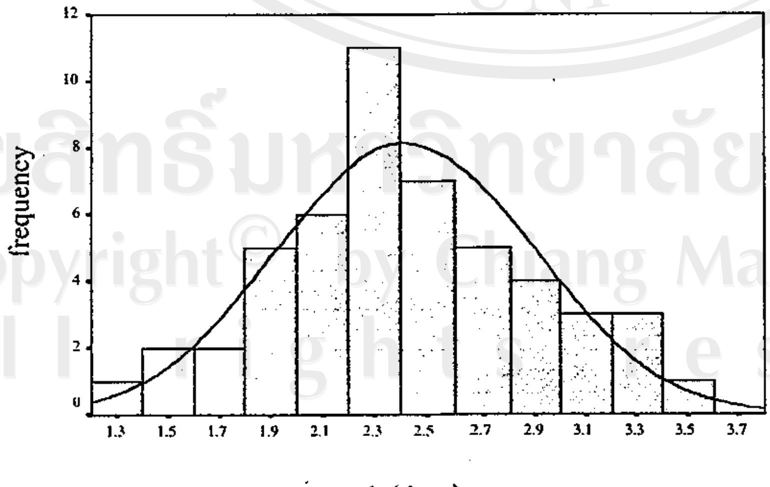
[2] K. Burnecki , G. Kukla and R. Wero, "Property insurance loss distributions", Physica A 287 . (2000) : 269-278



รูป 5.1 แสดงโค้งความถี่ของจำนวนเงินผลตอบแทนที่บริษัทประกันภัยต้องจ่ายให้กับลูกค้าต่อปี ตั้งแต่ปี พ.ศ. 2493 - 2542

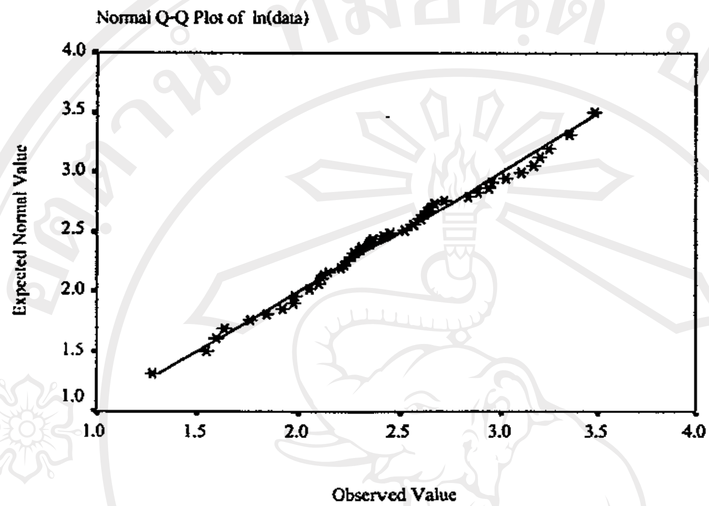
ในการแสดงว่าข้อมูลดังกล่าวมีความเหมาะสมกับการแจกแจงลอกนอร์มอลต้องแสดงว่า ลอกธรรมชาติ (natural log) ของข้อมูลมีการแจกแจงปกติ ซึ่งมีหลายวิธีในการตรวจสอบ และในที่นี้จะใช้การพิจารณาจากการแจกแจงความถี่ของข้อมูลและการเขียนกราฟช่วย ดังต่อไปนี้

ทำการแปลงข้อมูลโดยใช้ลอกธรรมชาติ (lnx) แล้วเขียนกราฟแสดงการแจกแจงความถี่ของข้อมูลที่เกิดจากการแปลงได้ดังต่อไปนี้



รูป 5.2 แสดงโค้งความถี่ของจำนวนเงินผลตอบแทนที่บริษัทประกันภัยจะต้องจ่ายให้กับลูกค้าต่อปี ตั้งแต่ปี พ.ศ. 2493 - 2542 ที่เกิดจากการแปลงโดยใช้ลอกธรรมชาติ (ln x)

จากโค้งความถี่ที่เกิดจากการแปลงพบว่า ลักษณะโค้งใกล้เคียงกับโค้งความถี่ของการแจกแจงแบบปกติ จึงตรวจสอบว่าข้อมูลที่เกิดจากการแปลงมีการแจกแจงแบบปกติหรือไม่ โดยการใช้ Normal Probability Plot ดังรูปข้างล่าง



รูป 5.3 แสดง Normal Probability Plot ของจำนวนเงินผลตอบแทนที่บริษัทประกันภัยจะต้องจ่ายให้กับลูกค้าต่อปี ตั้งแต่ปี พ.ศ. 2493 - 2542 ที่เกิดจากการแปลงโดยใช้ลอกธรรมชาติ($\ln x$)

จากกราฟสังเกตได้ว่าข้อมูลอยู่รอบๆเส้นตรงอย่างสุ่ม ดังนั้นสามารถกล่าวได้ว่า จำนวนเงินผลตอบแทนที่บริษัทประกันภัยจะต้องจ่ายให้กับลูกค้าต่อปี ตั้งแต่ปี พ.ศ. 2493-2542 ที่เกิดจากการแปลงมีการแจกแจงแบบปกติ

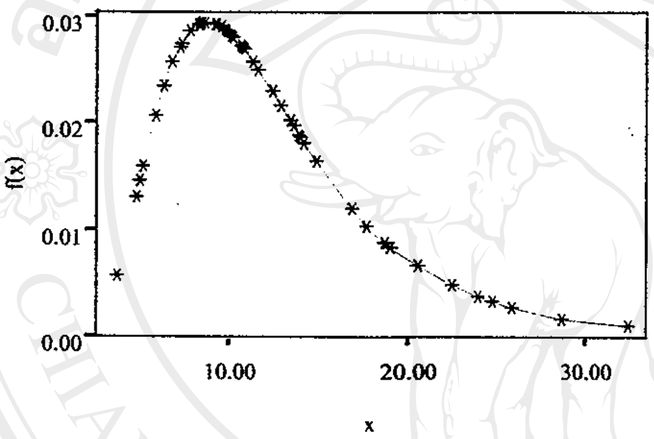
ดังนั้นถ้าให้ X คือผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้าต่อปีตั้งแต่ปี พ.ศ. 2493 - 2542 จะได้ว่า $\ln X \sim N(m,\sigma_m^2)$ นั่นคือ $X \sim LN\left(m,\sigma_m^2\right)$ การคำนวณหาค่าประมาณของพารามิเตอร์ m นั้นทำได้โดยการหาค่าเฉลี่ยของ $\ln x_i$ ซึ่งจะได้ค่าประมาณคือ $\hat{m} = 2.41$ และการหาค่าประมาณของพารามิเตอร์ σ_m^2 หาได้จากสูตร $\frac{1}{n}\sum_{i=1}^{n}\left(\ln x_i - \hat{m}\right)^2$ ซึ่งจะได้ $\hat{\sigma}_m^2 = 0.2393$ หรือเพื่อความสะดวกอาจจะใช้โปรแกรม SPSS ช่วยในการวิเคราะห์เพื่อหาค่าของตัวประมาณก็ได้

จากการที่ทราบค่าของตัวประมาณ ทำให้สามารถแสดงให้เห็นถึงลักษณะของฟังก์ชันความหนาแน่นน่าจะเป็นของข้อมูลได้โดยการแทนค่าของตัวประมาณคือ $\hat{m} = 2.41$ และ $\hat{\sigma}_m^2 = 0.2393$ ลงในฟังก์ชันความหนาแน่นน่าจะเป็น

จาก $$f(x) = \frac{1}{x\sigma_m\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - m}{\sigma_m}\right)^2}$$

$$= \frac{1}{x\sqrt{0.478\pi}} e^{-\frac{1}{0.478}(\ln x - 2.41)^2}$$

จากฟังก์ชันความหนาแน่นน่าจะเป็น เมื่อแทนค่าของ x แต่ละตัว แล้วเขียนกราฟของคู่อันดับ $(x, f(x))$ จะได้กราฟดังนี้



รูป 5.4 กราฟแสดงฟังก์ชันความหนาแน่นน่าจะเป็น ของการแจกแจงลอกนอร์มอลที่มี พารามิเตอร์ $m = 2.41$ และ $\sigma_m^2 = 0.239$

เนื่องจากเราทราบว่าข้อมูลดังกล่าวมีการแจกแจงลอกนอร์มอล และทราบค่าของตัวประมาณ ดังนั้นเราสามารถที่จะหาค่าเฉลี่ย ค่ามัธยฐาน ฐานนิยม และความแปรปรวนของข้อมูลได้ดังนี้

(1) ค่าเฉลี่ย ของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้า คือ

จาก $$\hat{\mu} = e^{\hat{m}+\frac{1}{2}\hat{\sigma}_m^2}$$

$$= e^{2.41+\frac{1}{2}(0.239)}$$

$$= 12.547$$

นั่นคือตั้งแต่ปี พ.ศ. 2493-2542 บริษัทต้องจ่ายเงินผลตอบแทนให้กับลูกค้าโดยเฉลี่ยปีละ 12.547 พันล้านดอลล่าร์

(2) ค่ามัธยฐาน ของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้า คือ

$$\text{จาก } Med(X) = e^{\hat{m}}$$

$$= e^{2.41}$$

$$= 11.134$$

นั่นคือ ค่ามัธยฐานของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้าต่อปีตั้งแต่ปี พ.ศ. 2493 - 2542 คือ 11.134 พันล้านดอลล่าร์

(3) ค่าฐานนิยม ของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้า คือ

$$\text{จาก } Mode(X) = e^{\hat{m}-\hat{\sigma}_m^2}$$

$$= e^{2.41-0.239}$$

$$= 8.767$$

นั่นคือ ค่าฐานนิยมของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้าต่อปีตั้งแต่ปี พ.ศ. 2493 - 2542 คือ 8.767 พันล้านดอลล่าร์

(4) ค่าความแปรปรวน ของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้า คือ

$$\text{จาก } \hat{\sigma}^2 = e^{2\hat{m}+\hat{\sigma}_m^2}\left(e^{\hat{\sigma}_m^2}-1\right)$$

$$= e^{2(2.41)+0.239}\left(e^{0.239}-1\right)$$

$$= (157.433)(0.270)$$

$$= 42.507$$

$$\hat{\sigma} = 6.486$$

นั่นคือ ส่วนเบี่ยงเบนมาตรฐานของจำนวนเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้าต่อปีตั้งแต่ปี พ.ศ. 2493-2542 เท่ากับ 6.486 พันล้านดอลล่าร์

นอกจากการหาค่าต่างๆภายในชุดข้อมูลดังที่แสดงให้เห็นแล้ว ยังสามารถที่ทราบลักษณะของข้อมูลโดยรวม โดยใช้สถิติอนุมานได้อีกด้วย เช่น ต้องการทราบว่าโดยเฉลี่ยแล้วบริษัทประกันทรัพย์สินจะต้องจ่ายเงินให้กับลูกค้าถึง 15 พันล้านดอลล่าร์หรือไม่ สามารถที่จะทราบได้ด้วยวิธีการดังนี้

สมมุติฐาน โดยเฉลี่ยแล้วปีหนึ่งบริษัทประกันทรัพย์สินจ่ายเงินผลตอบแทนให้กับลูกค้าไม่ถึง 15 พันล้านดอลล่าร์

สมมุติฐาน $H_0: \quad \mu \geq 15$

$H_1: \quad \mu < 15$

กำหนดอาณาเขตวิกฤตที่ระดับนัยสำคัญ $\alpha = 0.05$

สถิติที่ใช้ทดสอบ คือ $Z_{cal} = \dfrac{\hat{\mu} - \mu_0}{\sigma_m e^{m+\frac{1}{2}\sigma_m^2} \sqrt{1+\frac{1}{2}\sigma_m^2} \Big/ \sqrt{n}}$

$$= \frac{12.547 - 15}{(0.489)e^{2.41+\frac{1}{2}(0.239)} \sqrt{1+\frac{1}{2}(0.239)} \Big/ \sqrt{50}}$$

$$= \frac{-2.453}{(6.136)(0.149)}$$

$$= -2.683$$

เนื่องจากที่ระดับนัยสำคัญ $\alpha = 0.05$ ค่าของ Z ที่ได้จากการเปิดตารางมีค่าเท่ากับ 1.96 แสดงว่า $Z_{cal} < -Z_{0.975}$ ดังนั้นจึงปฏิเสธสมมุติฐาน H_0 นั่นคือ ในแต่ละปีบริษัทประกันภัยทรัพย์สินในสหรัฐอเมริกาจ่ายเงินผลตอบแทนให้ลูกค้าโดยเฉลี่ยน้อยกว่า 15 พันล้านดอลล่าร์

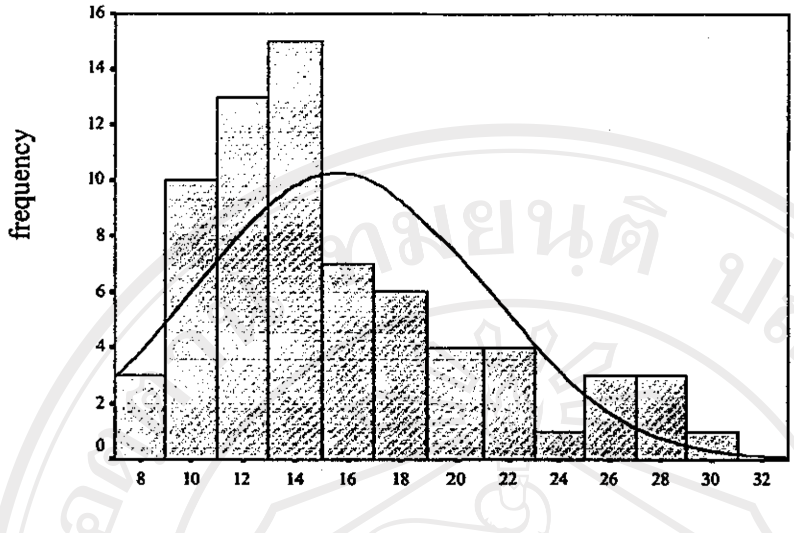
ตัวอย่าง 5.2 การกระจายรายได้กับการแจกแจงลอกนอร์มอล

ดังกล่าวข้างต้นแล้วว่า Aitchison และ Brown ได้ทำการศึกษาและพบว่า การแจกแจงลอกนอร์มอลมีความเหมาะสมในการอธิบายรายได้ที่ต่ำๆ ซึ่งในที่นี้จะได้แสดงตัวอย่างการนำการแจกแจงลอกนอร์มอลมาใช้ในการอธิบายเงินเดือนที่ได้รับต่อปี (ล้านฟรังค์) ของชาวฝรั่งเศสในปี พ.ศ. 2528 ดังต่อไปนี้

ตาราง 5.2 แสดงจำนวนเงินเดือนที่ได้รับต่อปี (ล้านฟรังค์) ของชาวฝรั่งเศสในปี พ.ศ. 2528 จำนวน 70 คน

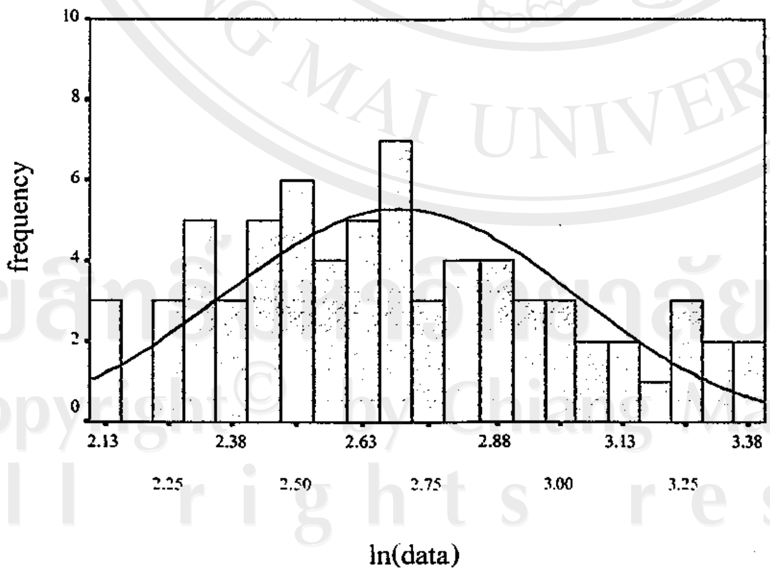
คนที่	รายได้	คนที่	รายได้	คนที่	รายได้	คนที่	รายได้
1	10.5	19	14.1	37	16.6	55	9.4
2	15.6	20	14.8	38	22.9	56	13.4
3	15.2	21	14.4	39	28.7	57	12.7
4	13.7	22	18.9	40	25.4	58	21.2
5	11.3	23	11	41	10.7	59	14.6
6	11.5	24	12.1	42	11.1	60	14.3
7	14.4	25	28.2	43	11.9	61	19.3
8	25.4	26	20.7	44	12.3	62	11.7
9	18.2	27	8.3	45	19.6	63	10.2
10	12.5	28	13.1	46	27.7	64	12.4
11	9.9	29	13.1	47	14.4	65	20.2
12	9.9	30	17.3	48	15.6	66	26.3
13	16.4	31	16.9	49	8.5	67	16.6
14	21.1	32	9.5	50	10	68	18.6
15	13	33	29.8	51	11.3	69	14.9
16	14	34	10.4	52	17.7	70	17.9
17	8.2	35	9.2	53	22.1		
18	12.3	36	23.5	54	14.2		

จากข้อมูลจำนวนเงินเดือนที่ได้รับต่อปีของชาวฝรั่งเศสในตารางที่ 5.2 เมื่อนำมาเขียนเป็นกราฟแสดงการแจกแจงความถี่ของข้อมูลดังรูปที่ 5.5 พบว่ากราฟมีลักษณะเบ้ขวาคล้ายกับโค้งความถี่ของการแจกแจงลอกนอร์มอลดังนั้นจึงทำการตรวจสอบว่าข้อมูลชุดดังกล่าวมีความเหมาะสมที่จะอธิบายด้วยการแจกแจงลอกนอร์มอลหรือไม่



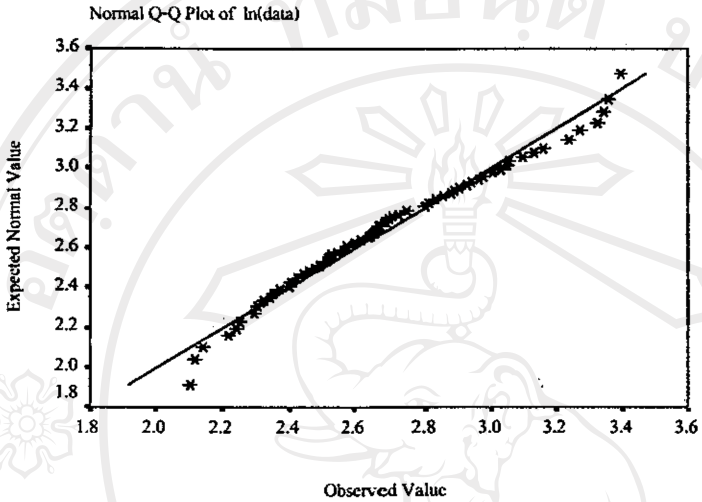
รูป 5.5 แสดงโค้งความถี่ของจำนวนเงินเดือนที่ได้รับต่อปีของชาวฝรั่งเศสในปี พ.ศ. 2528

ในการแสดงว่าข้อมูลจำนวนเงินเดือนต่อปีของชาวฝรั่งเศสมีความเหมาะสมกับการแจกแจงลอกนอร์มอลต้องแสดงว่าลอกธรรมชาติ (natural log) ของข้อมูลมีการแจกแจงแบบปกติ ดังนั้นจึงพิจารณาที่โค้งความถี่ของข้อมูลที่เกิดจากการแปลงโดยใช้ลอกธรรมชาติ (natural log) ซึ่งได้ผลดังนี้



รูป 5.6 แสดงโค้งความถี่ของข้อมูลรายได้ที่เกิดจากการแปลงโดยใช้ลอกธรรมชาติ (ln x)

จากโค้งความถี่ที่เกิดจากการแปลงพบว่า ลักษณะโค้งความถี่ใกล้เคียงกับโค้งความถี่ของการแจกแจงแบบปกติ จึงตรวจสอบว่าข้อมูลที่เกิดจากการแปลงมีการแจกแจงแบบปกติหรือไม่ โดยการใช้ Normal Probability Plot ดังรูปต่อไปนี้



รูป 5.7 แสดงNormal Probability Plot ของข้อมูลที่เกิดจากการแปลงโดยใช้ลอกธรรมชาติ(ln x)

จากกราฟสังเกตได้ว่าตำแหน่งข้อมูลอยู่รอบๆเส้นตรงอย่างสุ่ม ดังนั้นสามารถกล่าวได้ว่าข้อมูลจำนวนเงินเดือนที่ได้รับต่อปีของชาวฝรั่งเศสในปีพ.ศ. 2528 ที่เกิดจากการแปลงมีการแจกแจงแบบปกติ

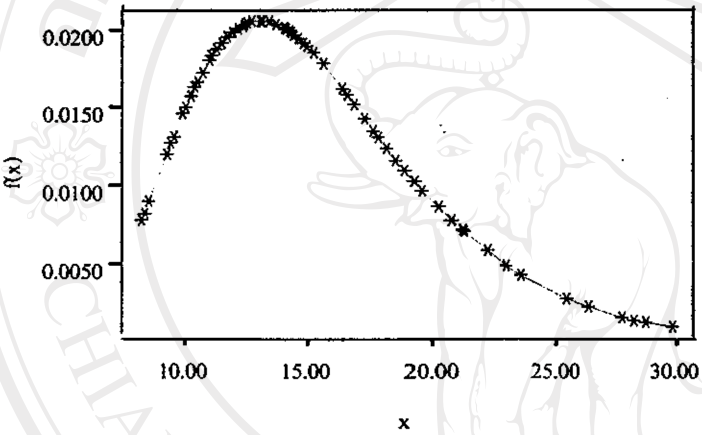
ดังนั้นถ้าให้ X คือจำนวนเงินเดือนที่ได้รับต่อปีของประชากรชาวฝรั่งเศสในปี พ.ศ. 2528 จะได้ว่า $\ln X \sim N(m,\sigma_m^2)$ นั่นคือ $X \sim LN\left(m,\sigma_m^2\right)$ การคำนวณหาค่าประมาณของพารามิเตอร์ m นั้นทำได้โดยการหาค่าเฉลี่ยของ $\ln x_i$ ซึ่งจะได้ค่าประมาณคือ $\hat{m} = 2.6782$ และการหาค่าประมาณของพารามิเตอร์ σ_m^2 หาได้จากสูตร $\frac{1}{n}\sum_{i=1}^{n}\left(\ln x_i - \hat{m}\right)^2$ ซึ่งจะได้ $\hat{\sigma}_m^2 = 0.1107$ หรือเพื่อความสะดวกอาจจะใช้โปรแกรม SPSS ช่วยในการวิเคราะห์เพื่อหาค่าของตัวประมาณก็ได้

จากการที่ทราบค่าของตัวประมาณ ทำให้สามารถแสดงให้เห็นถึงลักษณะของฟังก์ชันความหนาแน่นน่าจะเป็นของข้อมูลได้โดยการแทนค่าของตัวประมาณคือ $\hat{m} = 2.6782$ และ $\hat{\sigma}_m^2 = 0.1107$ ลงในฟังก์ชันความหนาแน่นน่าจะเป็น

จาก $$f(x) = \frac{1}{x\sigma_m\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x - m}{\sigma_m}\right)^2}$$

$$= \frac{1}{x\sqrt{0.2214\pi}} e^{-\frac{1}{0.2214}(\ln x - 2.6782)^2}$$

จากฟังก์ชันความหนาแน่นน่าจะเป็น เมื่อแทนค่าของ x แต่ละตัว แล้วเขียนกราฟของคู่อันดับ $(x, f(x))$ จะได้กราฟดังนี้



รูป 5.8 กราฟแสดงฟังก์ชันความหนาแน่นน่าจะเป็น ของการแจกแจงลอกนอร์มอลที่มี
พารามิเตอร์ $m = 2.8$ และ $\sigma_m^2 = 0.064$

เนื่องจากเราทราบว่าข้อมูลจำนวนเงินเดือนต่อปีของชาวฝรั่งเศสมีการแจกแจงลอกนอร์มอล และทราบค่าของตัวประมาณ ดังนั้นเราสามารถที่จะหาค่าเฉลี่ย ค่ามัธยฐาน ฐานนิยม และความแปรปรวนของข้อมูลได้ดังนี้

(1) ค่าเฉลี่ย ของจำนวนเงินเดือนต่อปีของชาวฝรั่งเศส คือ

จาก $$\hat{\mu} = e^{\hat{m}+\frac{1}{2}\hat{\sigma}_m^2}$$

$$= e^{2.6782+\frac{1}{2}(0.1107)}$$

$$= 15.3874$$

นั่นคือ ในปี พ.ศ. 2528 ชาวฝรั่งเศสมีเงินเดือนต่อปี โดยเฉลี่ยคนละประมาณ 15.3784 ล้านฟรังค์

(2) ค่ามัธยฐาน ของจำนวนเงินเดือนต่อปีของชาวฝรั่งเศส คือ

$$\text{จาก } Med(X) = e^{\hat{m}} = e^{2.6782} = 14.5589$$

นั่นคือ ในปี พ.ศ. 2528 ชาวฝรั่งเศสมีค่ามัธยฐานของเงินเดือนต่อปีคนละประมาณ 14.5589 ล้านฟรังค์

(3) ค่าฐานนิยม ของจำนวนเงินเดือนต่อปีของชาวฝรั่งเศส คือ

$$\text{จาก } Mode(X) = e^{\hat{m}-\hat{\sigma}_m^2} = e^{2.6782-0.1107} = 13.0332$$

นั่นคือในปี พ.ศ. 2528 ชาวฝรั่งเศสมีค่าฐานนิยมของเงินเดือนต่อปีคนละประมาณ 13.0332 ล้านฟรังค์

(4) ค่าความแปรปรวน ของจำนวนเงินเดือนต่อปีของชาวฝรั่งเศส คือ

$$\text{จาก } \hat{\sigma}^2 = e^{2\hat{m}+\hat{\sigma}_m^2}\left(e^{\hat{\sigma}_m^2}-1\right)$$

$$= e^{2(2.6782)+0.1107}\left(e^{0.1107}-1\right)$$

$$= (236.7726)(0.117)$$

$$= 27.7165$$

$$\hat{\sigma} = 5.2646$$

นั่นคือ ในปี พ.ศ. 2528 นั้นส่วนเบี่ยงเบนมาตรฐานของเงินเดือนต่อปีที่ชาวฝรั่งเศสได้รับเท่ากับ 5.2646 ล้านฟรังค์

5.3 การประยุกต์ใช้การแจกแจงโคชี

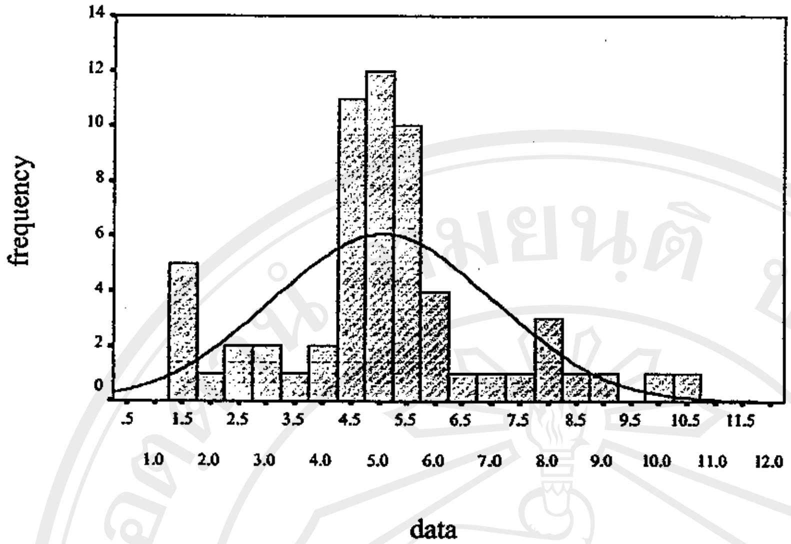
การแจกแจงโคชีเป็นการแจกแจงที่มีการนำไปประยุกต์ใช้กับงานด้านต่างๆค่อนข้างน้อย แต่อย่างไรก็ตาม การแจกแจงดังกล่าวก็ยังมีประโยชน์และถูกนำไปใช้บ้างในทางฟิสิกส์เกี่ยวกับทฤษฎีกลศาสตร์ และสนามแม่เหล็กไฟฟ้า และในที่นี้จะได้นำเสนอตัวอย่างการนำการแจกแจงโคชี : $C(a,b)$ โดยที่พารามิเตอร์ b เท่ากับหนึ่ง ไปอธิบายความหนาแน่นของสนามไฟฟ้าที่ถูกสร้างขึ้นบนพื้นผิวโลก

ตัวอย่าง 5.3 จากการศึกษาข้อมูลความหนาแน่นของสนามไฟฟ้าที่อยู่บนพื้นผิวโลกชุดหนึ่งเป็นดังนี้

ตาราง 5.3 แสดงความหนาแน่นของสนามไฟฟ้าบนพื้นผิวโลก (คูลอมบ์/ตร.เมตร)

พื้นที่	ความหนาแน่น	พื้นที่	ความหนาแน่น	พื้นที่	ความหนาแน่น	พื้นที่	ความหนาแน่น	พื้นที่	ความหนาแน่น
1	4.596	13	4.673	25	4.494	37	2.965	49	4.799
2	5.592	14	4.546	26	9.018	38	3.686	50	4.247
3	2.117	15	5.238	27	8.480	39	4.443	51	4.638
4	10.188	16	5.189	28	2.357	40	8.133	52	4.777
5	6.829	17	4.994	29	1.539	41	5.925	53	6.544
6	5.552	18	1.560	30	3.172	42	8.218	54	5.132
7	2.707	19	5.762	31	4.615	43	4.994	55	5.352
8	4.852	20	5.233	32	4.702	44	4.258	56	5.671
9	5.862	21	1.352	33	1.447	45	4.484	57	5.426
10	5.482	22	3.962	34	4.753	46	5.421	58	7.817
11	5.529	23	10.350	35	7.489	47	5.944	59	4.509
12	1.260	24	5.462	36	4.872	48	5.088	60	5.578

จากข้อมูลความหนาแน่นของสนามไฟฟ้าบนพื้นผิวโลกในตารางที่ 5.3 เมื่อนำมาเขียนเป็นกราฟแสดงการแจกแจงความถี่ของข้อมูลดังรูปที่ 5.9 พบว่ากราฟมีลักษณะคล้ายกับการแจกแจงโคชีซึ่งมีลักษณะสมมาตรและมีการกระจายตัวค่อนข้างสูงดังนั้นจึงทำการตรวจสอบว่าข้อมูลชุดดังกล่าวมีความเหมาะสมที่จะอธิบายด้วยการแจกแจงโคชีหรือไม่



รูป 5.9 แสดงโค้งความถี่ของข้อมูลความหนาแน่นสนามไฟฟ้าบนพื้นผิวโลก

ในการตรวจสอบว่าการแจกแจงโคชีมีความเหมาะสมในการอธิบายการกระจายของข้อมูลหรือไม่นั้น ในขั้นตอนแรกต้องทำการประมาณค่าพารามิเตอร์ของข้อมูล ซึ่งพารามิเตอร์ที่ต้องประมาณคือ พารามิเตอร์ a และขั้นตอนต่อไปคือ ตรวจสอบว่าข้อมูลมีการแจกโคชีจริงหรือไม่ โดยการเปรียบเทียบความถี่ของข้อมูลที่สังเกตได้กับความถี่ทางทฤษฎีว่าสอดคล้องกันหรือไม่ ถ้าความถี่ทั้งสองมีความสอดคล้องกันก็แสดงว่าข้อมูลดังกล่าวมีการแจกแจงโคชีจริง

ในการประมาณค่าพารามิเตอร์ a นั้นทำได้โดยใส่ค่าของข้อมูลความหนาแน่นของสนามไฟฟ้าที่มีอยู่ลงในโปรแกรม Newton Raphson เพื่อให้โปรแกรมช่วยคำนวณ ซึ่งจะได้ค่าของตัวประมาณคือ

$$\hat{a} = 5.005$$

จากค่าของตัวประมาณที่ได้ ทำให้ได้ฟังก์ชันความหนาแน่นน่าจะเป็นของข้อมูลความหนาแน่นของสนามไฟฟ้า คือ

$$f(x) = \frac{1}{\pi\left[1+(x-a)^2\right]}$$

$$= \frac{1}{\pi\left[1+(x-5.005)^2\right]}$$

และได้ฟังก์ชันการแจกแจงสะสมได้คือ

$$F(x) = \frac{1}{2}+\frac{1}{\pi}\tan^{-1}(x-a)$$

$$= \frac{1}{2}+\frac{1}{\pi}\tan^{-1}(x-5.005)$$

เมื่อทราบฟังก์ชันการแจกแจงสะสมจึงสามารถหาความน่าจะเป็นของความหนาแน่นสนามไฟฟ้าในช่วงต่างๆได้ เช่น ถ้าต้องการทราบความน่าจะเป็นที่ความหนาแน่นของสนามไฟฟ้าจะอยู่ในช่วง 0 ถึง 1 จะหาได้โดย

$$P(a\leq x\leq b) = F(b)-F(a)$$

ดังนั้น $P(0.0\leq x\leq 1.0) = \frac{1}{2}+\frac{1}{\pi}\tan^{-1}(1.0-5.005)-\frac{1}{2}-\frac{1}{\pi}\tan^{-1}(0-5.005)$

$$= \frac{1}{\pi}\left[\tan^{-1}(-4.005)-\tan^{-1}(-5.005)\right]$$

$$= 0.0252$$

สำหรับการหาความน่าจะเป็นในช่วงอื่นๆนั้นจะใช้วิธีการเช่นเดียวกันซึ่งหาได้ดังตารางต่อไปนี้

ตาราง 5.4 แสดงความถี่จากค่าสังเกตและความน่าจะเป็นที่ความหนาแน่นสนามไฟฟ้าจะอยู่ในช่วง a ถึง b หรือ $P(a \leq x \leq b)$

ความหนาแน่นสนามไฟฟ้า(x_i)	ความถี่จากค่าสังเกต	$F(a)$	$F(b)$	$P(a \leq x \leq b)$
น้อยกว่า 1.0	1	0.000	0.0879	0.0879
1.0-2.0	5	0.0879	0.1223	0.0344
2.0-3.0	4	0.1223	0.1775	0.0552
3.0-4.0	3	0.1775	0.2892	0.1117
4.0-5.0	19	0.2892	0.5484	0.2592
5.0-6.0	19	0.5484	0.8092	0.2608
6.0-7.0	3	0.8092	0.9221	0.1129
7.0-8.0	2	0.9221	0.9674	0.0453
8.0-9.0	3	0.9674	0.9879	0.0205
มากกว่า 9.0	1	0.9879	1.0000	0.0121
รวม	60			1.0000

จากความน่าจะเป็นในช่วงต่างๆของความหนาแน่นสนามไฟฟ้า สามารถนำมาหาความถี่ในทางทฤษฎีโดย นำความน่าจะเป็นที่หาได้คูณกับผลรวมความถี่จากค่าสังเกตทั้งหมด หรือความถี่ในทางทฤษฎีหาได้จาก $E_i = N\,P(a \leq x \leq b)$ เช่น ความถี่ในทางทฤษฎีของความหนาแน่นไฟฟ้าในช่วง 0 ถึง 1 เท่ากับ $0.0252 \times 60 = 1.512$ ส่วนความถี่ในช่วงอื่นๆสามารถหาได้ดังตารางที่ 5.5

จากตารางที่ 5.5 ซึ่งแสดงการเปรียบเทียบความถี่จากค่าสังเกตและความถี่ในทางทฤษฎี เราสามารถแสดงได้ด้วยกราฟดังรูปที่ 5.10

เมื่อพิจารณากราฟในรูป 5.10 ซึ่งแสดงการเปรียบเทียบความถี่จากค่าสังเกตและความถี่ในทางทฤษฎีของความหนาแน่นสนามไฟฟ้า พบว่าลักษณะของกราฟมีแนวโน้มไปในทิศทางเดียวกัน ดังนั้น กล่าวได้ว่าการแจกแจงโคชีมีความเหมาะสมในการอธิบายข้อมูลความหนาแน่นสนามไฟฟ้า

เมื่อทราบว่าข้อมูลความหนาแน่นสนามไฟฟ้ามีการแจกแจงโคชี และทราบค่าของตัวประมาณดังนั้นเราสามารถใช้สถิติโคชีช่วยในการตรวจสอบลักษณะของข้อมูลที่เราสนใจได้ดังนี้

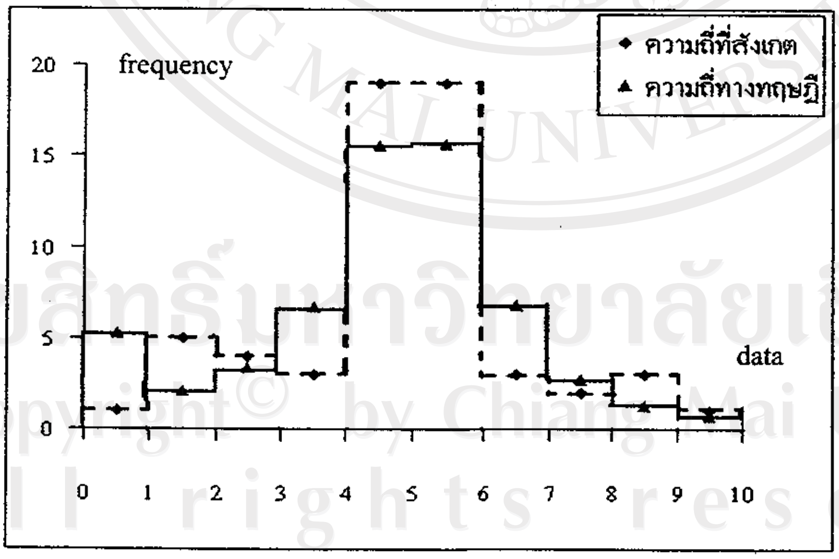
(1) ค่ามัธยฐานของความหนาแน่นสนามไฟฟ้า คือ

$$\text{จาก } Med(X) = \hat{a}$$
$$= 5.005$$

นั่นคือ ค่ามัธยฐานของความหนาแน่นสนามไฟฟ้าเท่ากับ 5.005

ตาราง 5.5 แสดงความถี่จากค่าสังเกตและความถี่ในทางทฤษฎีของความหนาแน่นสนามไฟฟ้า

ความหนาแน่นสนามไฟฟ้า (x_i)	ความถี่จากค่าสังเกต	ความถี่ในทางทฤษฎี
0.0 - 1.0	1	5.274
1.0 - 2.0	5	2.064
2.0 - 3.0	4	3.312
3.0 - 4.0	3	6.702
4.0 - 5.0	19	15.552
5.0 - 6.0	19	15.648
6.0 - 7.0	3	6.774
7.0 - 8.0	2	2.718
8.0 - 9.0	3	1.23
9.0 - 10.0	1	0.726
รวม	60	60



รูป 5.10 แสดงการเปรียบเทียบ ความถี่จากค่าสังเกตและความถี่ทางทฤษฎีของความหนาแน่นสนามไฟฟ้า

(2) ค่าฐานนิยมของความหนาแน่นสนามไฟฟ้า คือ

จาก Mode(X) $=$ $\hat{a}$

$=$ 5.005

นั่นคือ ค่าฐานนิยมของความหนาแน่นสนามไฟฟ้าเท่ากับ 5.005

5.4 การประยุกต์ใช้การแจกแจงลาปลาซ

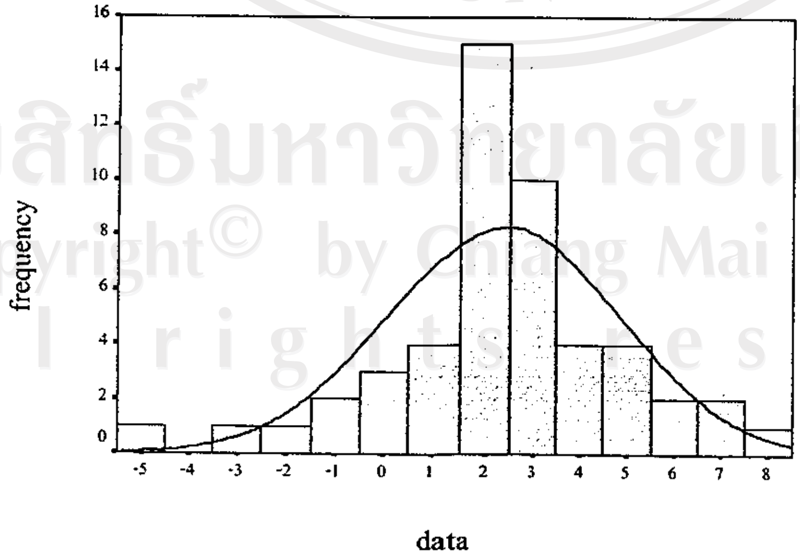
การแจกแจงลาปลาซถูกนำไปอธิบายเหตุการณ์ต่างๆในชีวิตประจำวันโดยอาศัยกระบวนการสโทแคสติก(stochastic)เป็นส่วนมากซึ่งเหตุการณ์ต่างๆที่กล่าวถึงนั้นได้รวมไปถึงในทางเศรษฐศาสตร์ ซึ่งมีบทบาทที่สำคัญในการดำรงอยู่ของมนุษย์ในปัจจุบันด้วย และในทางเศรษฐศาสตร์นี้การแจกแจงลาปลาซได้ถูกนำไปใช้ในการอธิบายการแจกแจงของอัตราการเจริญเติบโตของบริษัท ซึ่งจากการศึกษาของหลายท่านและในครั้งล่าสุดซึ่งศึกษาโดย Giulio Bottazzi และ Angelo Secchi แสดงให้เห็นว่า การแจกแจงลาปลาซมีความเหมาะสมในการอธิบายการเจริญเติบโตของบริษัท ซึ่งการศึกษาดังกล่าวนี้ได้ทำควบคู่ไปกับการศึกษาตัวแบบการเจริญเติบโต ทั้งนี้เพื่อประโยชน์ในการคาดคะเนอัตราการเจริญเติบโตในอนาคตและสามารถวางแผนการปฏิบัติงานของบริษัทได้อย่างถูกต้อง อย่างไรก็ตามการนำการแจกแจงลาปลาซไปอธิบายการเจริญเติบโตของบริษัทดังกล่าวอยู่บนเงื่อนไขที่ว่า จำนวนบริษัทและจำนวนโอกาสทางธุรกิจของบริษัทจะต้องมีอยู่มาก รวมทั้งบริษัทที่ศึกษานั้นจะต้องมีลักษณะที่ใกล้เคียงกัน คือ เป็นบริษัทที่มีอุตสาหกรรมการผลิตอย่างเดียวกันและมีขนาดของบริษัทใกล้เคียงกัน และในที่นี้จะได้นำเสนอตัวอย่างของการนำการแจกแจงลาปลาซไปประยุกต์ใช้กับการศึกษาอัตราการเจริญเติบโตของบริษัทที่ทำอุตสาหกรรมการผลิต

ตัวอย่าง 5.4 การศึกษาและวิเคราะห์ข้อมูลอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังจำนวน 50 บริษัทดังนี้

ตาราง 5.6 แสดง อัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังจำนวน 50 บริษัท

บริษัท	อัตราเจริญเติบโต	บริษัท	อัตราเจริญเติบโต	บริษัท	อัตราเจริญเติบโต	บริษัท	อัตราเจริญเติบโต	บริษัท	อัตราเจริญเติบโต
1	2.191	11	2.078	21	2.770	31	6.114	41	1.876
2	6.682	12	3.821	22	1.700	32	2.251	42	2.087
3	1.312	13	-1.019	23	0.581	33	4.841	43	2.012
4	2.712	14	-4.685	24	2.226	34	2.303	44	0.308
5	3.847	15	3.050	25	2.960	35	0.723	45	6.847
6	2.001	16	0.409	26	2.935	36	1.041	46	-1.905
7	4.258	17	6.258	27	4.301	37	2.737	47	3.200
8	2.872	18	2.250	28	5.234	38	1.647	48	7.594
9	2.133	19	2.375	29	-1.424	39	4.707	49	2.380
10	-2.665	20	5.210	30	2.687	40	3.225	50	-0.173

จากข้อมูลอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังจำนวน 50 บริษัทในตารางที่ 5.6 เมื่อนำมาเขียนเป็นกราฟแสดงการแจกแจงความถี่ของข้อมูลดังรูปที่ 5.11 พบว่ากราฟมีลักษณะที่น่าจะสามารถนำการแจกแจงลาปลาซมาช่วยในการอธิบายข้อมูลได้ดังนั้นจึงทำการตรวจสอบว่าข้อมูลชุดดังกล่าวมีความเหมาะสมที่จะอธิบายด้วยการแจกแจงลาปลาซดังที่คาดไว้หรือไม่



รูป 5.11 แสดงโค้งความถี่ของข้อมูลอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง

ในการตรวจสอบว่าการแจกแจงลาปลาซมีความเหมาะสมในการอธิบายข้อมูลชุดดังกล่าวหรือไม่นั้น ในขั้นตอนแรกต้องทำการประมาณค่าพารามิเตอร์ของข้อมูล ซึ่งพารามิเตอร์ที่ต้องประมาณคือ พารามิเตอร์ a และ b และขั้นตอนต่อไปคือ ตรวจสอบว่าข้อมูลมีการแจกลาปลาซจริงหรือไม่ โดยการเปรียบเทียบความถี่ของข้อมูลที่สังเกตได้กับความถี่ทางทฤษฎีว่าสอดคล้องกันหรือไม่ ถ้าความถี่ทั้งสองมีความสอดคล้องกันก็แสดงว่าข้อมูลดังกล่าวมีการแจกแจงลาปลาซจริง

สำหรับการประมาณค่าพารามิเตอร์ทำได้ดังนี้

เนื่องจาก $\hat{a} \quad = \quad \text{median}\left(x_i\right)$

ดังนั้น $\hat{a}$ ของอัตราการเจริญเติบโตของทั้ง 50 บริษัท จึงเกิดจาก การเรียงข้อมูลจากน้อยไปหามากแล้วนำค่าที่อยู่ในตำแหน่งที่ 25 และ 26 มาหาค่าเฉลี่ย

ซึ่งจะได้ $\hat{a} \quad = \quad 2.277$

และเนื่องจาก

$$\hat{b} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \hat{a}\right|$$

ดังนั้นสามารถหาค่า $\hat{b}$ โดยการแทนค่า ตัวแปรลงในสูตร ซึ่งจะได้

$$\hat{b} = \frac{1}{50}\sum_{i=1}^{n}\left|x_i - 2.277\right|$$

เมื่อแทนค่า x_i จะได้ค่า $\hat{b} = 1.688$

เนื่องจากในการแจกแจงลาปลาซซึ่งประกอบไปด้วย พารามิเตอร์ a และ b นั้น มีเงื่อนไขว่า $-\infty < a < \infty$ และ $b > 0$ ดังนั้น เมื่อเราทราบค่าของตัวประมาณคือ $\hat{a}$ และ $\hat{b}$ แล้ว เราจึงทำการทดสอบค่าพารามิเตอร์ว่าเป็นไปตามเงื่อนไขหรือไม่ ซึ่งจะเห็นได้ว่า พารามิเตอร์ที่ต้องทำการทดสอบคือ พารามิเตอร์ b ดังต่อไปนี้

ทดสอบสมมุติฐาน $b \le 0$

สมมุติฐาน

$H_0: \quad b \le 0$

$H_1: \quad b > 0$

กำหนดอาณาเขตวิกฤตที่ระดับนัยสำคัญ $\alpha = 0.05$

สถิติที่ใช้ทดสอบ คือ $Z_{cal} = \dfrac{\hat{b} - b_0}{b/\sqrt{n}} = \dfrac{1.688}{1.688/\sqrt{50}} = 7.071$

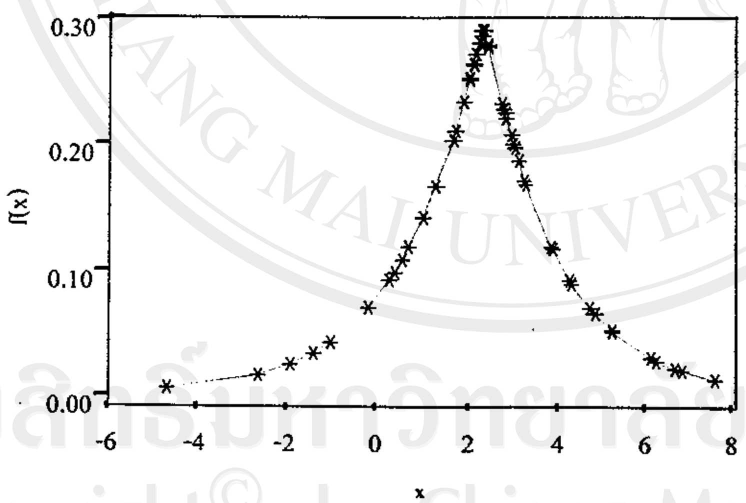
เนื่องจากที่ระดับนัยสำคัญ $\alpha = 0.05$ ค่าของ Z ที่ได้จากการเปิดตารางมีค่าเท่ากับ 1.96 แสดงว่า $Z_{cal} > Z_{0.975}$ ดังนั้นจึงปฏิเสธสมมุติฐานที่ว่า $b \leq 0$ นั่นคือ $b > 0$ ซึ่งตรงกับเงื่อนไขของการแจกแจงลาปลาซ

จากการที่ทราบค่าของตัวประมาณ ทำให้สามารถแสดงให้เห็นถึงลักษณะของฟังก์ชันความหนาแน่นน่าจะเป็นของข้อมูลชุดดังกล่าว โดยการแทนค่าของตัวประมาณคือ $\hat{a} = 2.277$ และ $\hat{b} = 1.688$ ลงในฟังก์ชันความหนาแน่นน่าจะเป็น ดังนี้

$$\text{จาก } f(x) = \frac{1}{2b} e^{-\frac{|x-a|}{b}}$$

$$= \frac{1}{3.376} e^{-\frac{|x-2.277|}{1.688}}$$

เมื่อแทนค่าของ x_i แต่ละตัว แล้วเขียนกราฟของคู่อันดับ $(x_i, f(x_i))$ จะได้กราฟดังนี้



รูป 5.12 กราฟแสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงลาปลาซที่มีพารามิเตอร์ $a = 2.277$ และ $b = 1.688$

สำหรับการหาความถี่ในทางทฤษฎีของข้อมูลอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังทั้ง 50 บริษัทนั้น สามารถหาได้ดังนี้

$$\text{จาก } f(x) = \frac{1}{3.376} e^{-\frac{|x-2.277|}{1.688}}$$

$$\text{และ } F(x) = \begin{cases} \frac{1}{2} e^{-\left(\frac{2.277-x}{1.688}\right)} & , x < 2.277 \\ 1-\frac{1}{2} e^{-\left(\frac{x-2.277}{1.688}\right)} & , x \geq 2.277 \end{cases}$$

ดังนั้น ถ้าต้องการทราบความน่าจะเป็นที่อัตราการเจริญเติบโตของบริษัทอยู่ในช่วง 1 ถึง 2.5 จะหาได้โดย

$$P(a \leq x \leq b) = F(b) - F(a)$$

$$\text{ดังนั้น } P(1 \leq x \leq 2.5) = 1-\frac{1}{2} e^{-\left(\frac{2.5-2.277}{1.688}\right)} - \frac{1}{2} e^{-\left(\frac{2.277-1}{1.688}\right)}$$

$$= 1 - 0.43812 - 0.23465$$

$$= 0.32723$$

สำหรับการหาความน่าจะเป็นในช่วงอื่นๆก็ใช้วิธีการเช่นเดียวกันซึ่งได้ดังตารางต่อไปนี้

ตาราง 5.7 แสดงความถี่จากค่าสังเกตและความน่าจะเป็นที่บริษัทจะมีอัตราการเจริญเติบโตอยู่ในช่วง a ถึง b หรือ $P(a \le x \le b)$

อัตราเจริญเติบโต (x_i)	ความถี่จากค่าสังเกต	$F(a)$	$F(b)$	$P(a \le x \le b)$
น้อยกว่า (-3.5)	1	0.00000	0.01632	0.01632
-3.5-(-2)	1	0.01632	0.03968	0.02336
-2-(-0.5)	3	0.03968	0.09649	0.05681
-0.5-1	5	0.09649	0.23465	0.13816
1-2.5	17	0.23465	0.56188	0.32723
2.5-4	12	0.56188	0.81983	0.25795
4-5.5	6	0.81983	0.92591	0.10608
5.5-7	4	0.92591	0.96953	0.04362
มากกว่า 7	1	0.96953	1.00000	0.03047
รวม	50			1.00000

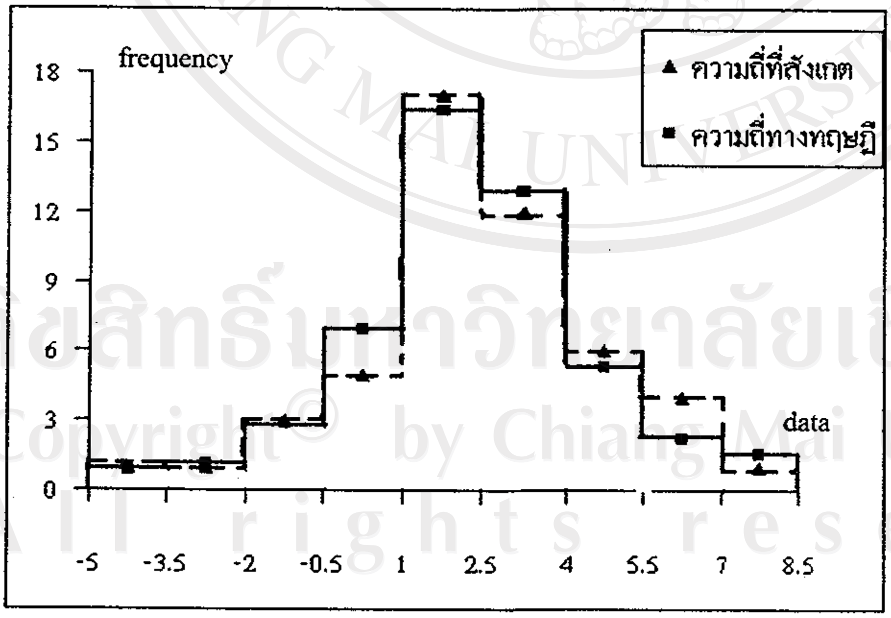
จากความน่าจะเป็นของอัตราการเจริญเติบโตในช่วงต่างๆของบริษัทผลิตเครื่องหนังที่หาได้ สามารถนำมาหาความถี่ในทางทฤษฎีโดย นำความน่าจะเป็นที่หาได้คูณกับผลรวมความถี่จากค่าสังเกตทั้งหมด เช่น ความถี่ในทางทฤษฎีที่บริษัทจะเจริญเติบโตอยู่ในช่วง 1 ถึง 2.5 เท่ากับ $50 \times 0.3273 = 16.3615$ ส่วนความถี่ในช่วงอื่นๆ สามารถหาได้ดังตารางที่ 5.8

จากตารางที่ 5.8 ซึ่งแสดงการเปรียบเทียบความถี่จากค่าสังเกตและความถี่ในทางทฤษฎี เราสามารถแสดงได้ด้วยกราฟดังรูปที่ 5.13

เมื่อพิจารณากราฟในรูปที่ 5.13 ซึ่งแสดงการเปรียบเทียบความถี่จากค่าสังเกตและความถี่ในทางทฤษฎีของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังจำนวน 50 บริษัท พบว่าลักษณะของกราฟมีแนวโน้มไปในทิศทางเดียวกัน ดังนั้น กล่าวได้ว่าการแจกแจงลาปลาซมีความเหมาะสมในการอธิบายข้อมูลชุดดังกล่าว

ตาราง 5.8 แสดงความถี่จากค่าสังเกตและความถี่ในทางทฤษฎีของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง

อัตราเจริญเติบโต (x_i)	ความถี่จากค่าสังเกต	ความถี่ในทางทฤษฎี
(-5.0) - (-3.5)	1	0.8160
(-3.5) - (-2.0)	1	1.1680
(-2.0) - (-0.5)	3	2.8405
(-0.5) - 1.0	5	6.9080
1.0 - 2.5	17	16.3615
2.5 - 4.0	12	12.8975
4.0 - 5.5	6	5.3040
5.5 - 7.0	4	2.1810
7.0 - 8.5	1	1.5235
รวม	50	50



รูป 5.13 แสดงการเปรียบเทียบ ความถี่จากค่าสังเกตและความถี่ทางทฤษฎีของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง

เมื่อทราบว่าข้อมูลอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังมีการแจกแจงลาปลาซ และทราบค่าของตัวประมาณ ดังนั้นเราสามารถที่จะหาค่าเฉลี่ย ค่ามัธยฐาน ฐานนิยม และความแปรปรวนของข้อมูลได้ดังนี้

(1) ค่าเฉลี่ย ของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง คือ

จาก $\hat{\mu} = \hat{a}$

$= 2.277$

นั่นคือบริษัทผลิตเครื่องหนังมีอัตราการเจริญเติบโตโดยเฉลี่ยเท่ากับร้อยละ 2.277

(2) ค่ามัธยฐาน ของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง คือ

จาก $\mathrm{Med}(X) = \hat{a}$

$= 2.277$

นั่นคือ ค่ามัธยฐานของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังเท่ากับร้อยละ 2.277

(3) ค่าฐานนิยม ของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง คือ

จาก $\mathrm{Mode}(X) = \hat{a}$

$= 2.277$

นั่นคือ ค่าฐานนิยมของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังเท่ากับร้อยละ 2.277

(4) ค่าความแปรปรวน ของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนัง คือ

จาก $\hat{\sigma}^2 = 2b^2$

$= 2\times(1.688)^2$

$= 5.6986$

$\hat{\sigma} = 2.387$

นั่นคือ ส่วนเบี่ยงเบนมาตรฐานของอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังเท่ากับร้อยละ 2.387

นอกจากการหาค่าต่างๆภายในชุดข้อมูลดังที่แสดงให้เห็นแล้ว ยังสามารถที่ทราบลักษณะของข้อมูลโดยรวม โดยใช้สถิติอนุมานได้อีกด้วย เช่น ต้องการทราบว่าอัตราการเจริญเติบโตของบริษัทผลิตเครื่องหนังโดยเฉลี่ยในช่วงเวลาดังกล่าวมากกว่าร้อยละ 3.5 หรือไม่ ทดสอบสมมุติฐานได้ ดังนี้

สมมุติฐาน ค่าเฉลี่ยของอัตราการเจริญเติบโตของบริษัทไม่มากกว่าร้อยละ 3.5

สมมุติฐาน $H_0: \quad a \leq 3.5$

$H_1: \quad a > 3.5$

กำหนดอาณาเขตวิกฤตที่ระดับนัยสำคัญ $\alpha = 0.05$

สถิติที่ใช้ทดสอบ คือ $Z_{cal} = \dfrac{\hat{a}-a_0}{\hat{b}/\sqrt{n}} = \dfrac{2.277-3}{1.688/\sqrt{50}} = -3.029$

เนื่องจากที่ระดับนัยสำคัญ $\alpha = 0.05$ ค่าของ Z ที่ได้จากการเปิดตารางมีค่าเท่ากับ 1.96 แสดงว่า $Z_{cal} < Z_{0.975}$ ดังนั้นจึงยอมรับสมมุติฐาน H_0 นั่นคือ อัตราการเจริญเติบโตโดยเฉลี่ยของบริษัทผลิตเครื่องหนังในช่วงเวลาดังกล่าวไม่มากกว่าร้อยละ 3.5