

น่าจะได้ ถ้านักศึกษาสายสังคมศาสตร์จะรู้จัก Probability Function

สมเพลิน เกษมรัตน์สันติ

บทคัดย่อ

ความเข้าใจในเรื่องรูปแบบการแจกแจงข้อมูลในวิชาสถิติ ถือเป็นกุญแจสำคัญคอกหนึ่งในการสร้างความเข้าใจวิชาสถิติเบื้องต้น อย่างไรก็ตาม ตัวรูปแบบเองก็มีเรื่องราวทางคณิตศาสตร์เข้ามาเกี่ยวข้องพอสมควร เช่น ฟังก์ชันความน่าจะเป็น (probability functions) การหาอนุพันธ์ (differentiation) การอินทิเกรต (Integration) เป็นต้น จึงทำให้เกิดอุปสรรคในการเข้าใจ โดยเฉพาะอย่างยิ่ง ถ้าเป็นนักศึกษาสายสังคมศาสตร์ ซึ่งไม่ถนัดวิชาคณิตศาสตร์ บทความนี้ จึงพยายามเขียนรูปแบบการกระจายข้อมูลแบบเบื้องต้น 2 รูปแบบ ซึ่งใช้กันมาก คือ รูปแบบการแจกแจงข้อมูลแบบโค้งปกติ (Normal distribution) และรูปแบบการแจกแจงข้อมูลแบบทวินาม (Binomial distribution) โดยอธิบายในแนวทางๆ ซึ่งเพียงพอแก่ความเข้าใจของนักศึกษาสายสังคมศาสตร์ และสามารถนำไปใช้ได้ ตลอดจนหลักการการยืมรูปแบบการแจกแจงข้อมูลแบบโค้งปกติมาใช้ในกรณีที่ ข้อมูลดั้งเดิมไม่ได้กระจายเป็นโค้งปกติ

Abstraction

The understanding of the probability distribution theory in Statistics is an important key to be into the course of Introduction to Statistics . However , the theory itself always involves , more or less , mathematical concepts , for example , the probability function , differentiation, integration, etc. This is quite a barrier for students in social science to understand since they have not been so keen in Mathematics . So , this writing is trying to explain two basic distributions namely Normal distribution theory and Binomial distribution theory in the easy manner so that they could understand and be able to use them correctly . The explanation also includes the use of the Normal distribution theory borrowed by the data originally Binomial distributed .

น่าจะดี ถ้านักศึกษาสายสังคมศาสตร์จะรู้จัก Probability function

สมเพณิณ เกษมรัตนสันติ

1. บทนำ

การศึกษาวิชาสถิติแม้จะเป็นเพียงสถิติเบื้องต้น แต่ก็ใช่ว่าจะง่ายสำหรับนักศึกษาสายสังคมศาสตร์ โดยเฉพาะอย่างยิ่ง ถ้านักศึกษาผู้นั้นสำเร็จการศึกษาชั้นมัธยมศึกษาสายศิลป์ภาษา ซึ่งมักเป็นผู้ที่ไม่ถนัดวิชาคณิตศาสตร์ แต่ถ้าเป็นสายศิลป์คำนวณก็ยังไม่พอสบายใจกับคณิตศาสตร์อยู่บ้าง อย่างไรก็ตาม แม้นักศึกษาเหล่านี้พยายามเลือกเรียนทางสายสังคมศาสตร์เพื่อหลีกเลี่ยงการเรียนวิชาคำนวณ และเพื่อจะได้เรียนในวิชาที่ตนถนัดและชอบ กระนั้นก็ดี คณะต่าง ๆ ทางสังคมศาสตร์หลายคณะจะมีวิชาบังคับที่เกี่ยวข้องกับวิชาสถิติอย่างน้อยหนึ่งวิชาที่นักศึกษาต้องเรียน ทั้งนี้เพราะวิชาสถิติจะให้ประโยชน์อย่างมากในการวิเคราะห์ข้อมูลต่าง ๆ ซึ่งนักศึกษาจำเป็นต้องใช้ในชีวิตทำงานในอนาคต แต่การวิเคราะห์ข้อมูลนี้มักต้องเกี่ยวข้องกับหลักเกณฑ์ทางคณิตศาสตร์และแนวคิดทางสถิติ ซึ่งบางเรื่องค่อนข้างจะเป็นมโนทัศน์และมีความแตกต่างอย่างสิ้นเชิงจากทางสังคมศาสตร์ จึงเป็นความลำบากใจทั้งผู้สอนและผู้เรียน ที่ต่างฝ่ายต่างต้องการให้บรรลุวัตถุประสงค์ของการเรียนวิชานี้

แต่อันที่จริง วิชาอะไรก็ตามถ้าผู้ศึกษาได้มีโอกาสเรียนรู้อย่างละเอียดถี่ถ้วนพอสมควร ผู้ศึกษาน่าจะเกิดความเข้าใจและนำไปใช้ได้ถูกต้อง ดีกว่าที่จะเรียนรู้อย่างผิวเผิน นักศึกษาสายสังคมศาสตร์ก็เช่นกัน ถ้าไม่ถูกจำกัดให้เรียนสถิติแต่เพียงพอประมาณ การรับรู้ที่มากขึ้นก็น่าจะทำให้สามารถใช้สถิติได้อย่างมีประสิทธิภาพและไม่ลำบากใจ เช่น สูตรมีมากมายจะใช้สูตรไหนดี ใช้เมื่อไร และอย่างไร บทความนี้เน้นการอธิบาย Probability function¹ ให้ง่ายแก่การเข้าใจของนักศึกษาสายสังคมศาสตร์โดยเฉพาะเพื่อจะได้นำไปใช้ได้ถูกต้อง

¹ Probability function มีลักษณะเป็นฟังก์ชันทางคณิตศาสตร์ฟังก์ชันหนึ่ง แต่แตกต่างจากฟังก์ชันทั่วไปตรงที่มีคุณสมบัติเรื่องค่าความน่าจะเป็น (probability) เพิ่มขึ้นใน function ความน่าจะเป็นหรือโอกาสที่จะเกิดมักคิดเป็นตัวเลขมีค่าที่ต่ำสุดเป็น 0 และค่าที่สูงที่สุดเป็น 1 ถ้าเหตุการณ์ใดมีค่าความน่าจะเป็นใกล้ 1 จึงหมายถึงเหตุการณ์นั้นมีโอกาสเกิดมาก โดยปกติความน่าจะเป็นจึงมีค่าเป็นทศนิยม ลองพิจารณาตัวอย่างต่อไปนี้

ตารางทางซ้ายมือเป็นตารางบันทึกคะแนนในการสอบของบุคคล 50 คน เราเรียกตารางนี้ว่า frequency distribution ส่วนตารางทางขวามือเป็นตารางที่เรียกว่า probability distribution



คะแนน	ความถี่
50 - 59	10
60 - 69	14
70 - 79	15
80 - 89	8
90 - 99	3
รวม	50

Frequency distribution

คะแนน	ความน่าจะเป็น (p)
50 - 59	0.20
60 - 69	0.28
70 - 79	0.30
80 - 89	0.16
90 - 99	0.06
รวม	1.00

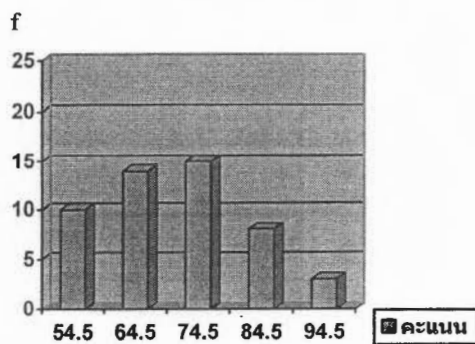
Probability distribution

ถ้าถามว่าความน่าจะเป็นในตารางขวามือหาได้อย่างไร เราสมมุติว่าบุคคลแต่ละคนได้มีโอกาสถูกสังเกตโดยทดเทียมกัน และสมมุติให้ช่วงของคะแนนแต่ละชั้นถูกเรียกว่า A ดังนั้นในการหาค่าความน่าจะเป็นจึงใช้ฟังก์ชันนี้

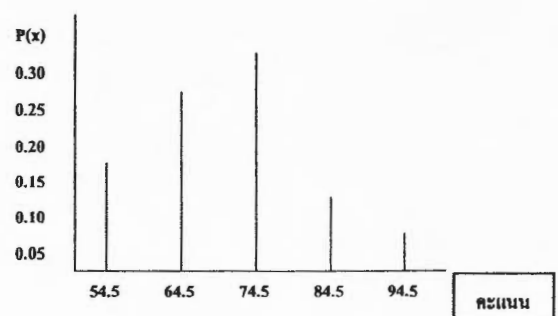
$$p(A) = \frac{\text{จำนวนความถี่ของ } A}{\text{จำนวนความถี่รวม}}$$

ฟังก์ชันดังกล่าวนี้ถูกเรียกว่า Probability function อย่างไรก็ตาม Probability function นี้เป็นชนิดไม่จับซ้อนและไม่มีกฎเกณฑ์ทางคณิตศาสตร์ (Mathematical rules) มากมาย บทความนี้ตั้งใจจะเสนอ Probability function ที่เป็นทฤษฎีการแจกแจงข้อมูลซึ่งมีกฎเกณฑ์ทางคณิตศาสตร์ที่ยุ่งยากมากขึ้นและเกี่ยวข้องกับการหาค่าความน่าจะเป็นของตัวแปร นั่นคือ probability function ของการแจกแจงแบบโค้งปกติ (Normal distribution) และของการแจกแจงแบบทวินาม (Binomial distribution)

เรามักเขียนกราฟแสดงตัวแปรแบบไม่ต่อเนื่องได้ 2 รูปแบบ คือรูปแบบที่เป็นแท่ง ๆ (histogram) และรูปแบบที่เป็น probability mass function ข้อมูลจากตารางข้างบนเขียนเป็นกราฟได้ดังนี้



ภาพที่ 1 รูปแบบที่เป็น histogram

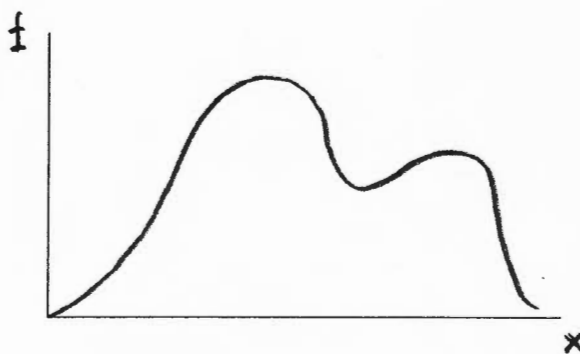


ภาพที่ 2 รูปแบบ probability mass function

การใช้สถิติในทางปฏิบัติ ก็คือการนำเอาข้อมูลดิบที่เก็บมาเองจากสนามหรือที่บันทึกไว้แล้ว ในเอกสารมาวิเคราะห์ทางสถิติ ข้อมูลที่เก็บมาอาจได้มาเป็นตัวเลขโดยตรง เช่น ในการขายสินค้า ในร้าน วันที่ 1 ขายได้ 56.50 บาท วันที่ 2 ขายได้ 60 บาท วันที่ 3 ขายได้ 72.75 บาท เป็นต้น หรือเป็นข้อมูลที่ต้องตีความหมายของข้อความให้เป็นตัวเลข เช่น ถ้ามีผู้ตอบ เห็นด้วยอย่างยิ่ง ก็แปลงเป็น 5 คะแนน เห็นด้วย ให้ 4 คะแนน เป็นต้น บางครั้งอาจมีคำตอบเรื่องเพศว่า เป็นหญิง 10 คน เป็นชาย 8 คน ก็เอาเลข 10 หรือเลข 8 มาเป็นข้อมูลดิบ เหล่านี้เป็นต้น ตัวอย่างที่ยกมาทั้งหมดนี้ เราสามารถแยกข้อมูลออกเป็นสองชนิดใหญ่ๆ ได้ ดังนี้

1. ตัวแปรชนิดต่อเนื่อง (continuous variable)

ถ้ามีความสนใจจะศึกษาเรื่องความสูง หรือน้ำหนัก ฯลฯ ซึ่งอาจเรียกความสูงหรือน้ำหนัก ว่าตัวแปรและใช้อักษร X แทนตัวแปรดังกล่าว และสมมุติว่ามีตัวเลข 2 ตัวซึ่งเป็นตัวเลขจริงที่มีค่าเป็นบวก (real number)² หรืออาจเป็นค่า 0 ก็ได้ ตัวอย่างเช่น 5 กับ 32 ถ้าตัวแปร X นี้มีค่าอยู่ในช่วง 5 ถึง 32 ($5 \leq X \leq 32$) คือเป็นตัวเลขใด ๆ ก็ได้ที่อยู่ในช่วงดังกล่าว เช่น 6 หรือ 9.118 หรือ 18.4557... หรือ 31.023689... (เป็นตัวเลขที่ไม่สิ้นสุด) เราจะเรียกตัวแปร X นี้ว่าเป็นตัวแปรต่อเนื่องซึ่งมีค่าทุก ๆ ค่าอยู่ในช่วงดังกล่าว ลักษณะนี้ทำให้ไม่มีช่องโหว่ระหว่างค่าของตัวแปร X เพราะ X มีค่าเป็นทศนิยมได้ไม่จำกัดจำนวน นั่นคือค่าของ X จากตัวหนึ่งไปยังอีกตัวหนึ่งจะเกิดความหนาที่บ (density) ตัวแปรชนิดนี้จะถูกเรียกว่าตัวแปรชนิดต่อเนื่อง บางครั้งนักวิชาการจะกล่าวถึงตัวแปรชนิดนี้อย่างง่าย ๆ ว่า เป็นตัวแปรที่สามารถนำมา ชั่ง ตวง หรือวัดได้ นั่นคือวัดเป็นค่าออกมาได้และมีทศนิยมได้ และสามารถเขียนกราฟที่มีความต่อเนื่อง ไม่มีช่วงว่างบนเส้นกราฟได้ ดังนี้



² ตัวเลขที่ไม่เป็น real number เช่น $\sqrt{-1}$, $4\sqrt{-1}$, $3\sqrt{-5}$ เป็นต้น

2. ตัวแปรชนิดไม่ต่อเนื่อง (discrete variable)

สมมุติว่ามีหนังสือสถิติเล่มหนึ่ง มีเลขหน้าซึ่งระบุไว้ตั้งแต่หน้า 1 – 1019 หน้า ดังนั้นค่าของตัวแปร X เช่นการเปิดหน้าหนังสือเล่มนี้เพื่อดูเนื้อหาของหนังสือ จึงเป็นตัวเลขที่มีขอบเขต (finite) คือเริ่มตั้งแต่หน้า 1 ถึง 1019 ซึ่งอาจจะไม่เปิดหน้าเรียงกันก็ได้ เช่น เปิดหน้า 1 ไปหน้า 2 ข้ามไปหน้า 5 หน้า 6 หน้า 7 ... อย่างไรก็ตามในการเปิดแต่ละครั้ง จะไม่ปรากฏตัวเลข 34.6 หรือ 1.1103 หรือตัวเลขอื่นที่มีทศนิยม ขึ้นเลย

อีกตัวอย่างหนึ่ง สมมุติว่ามีการดึงไฟจากสํารับหนึ่ง โดยดึงแล้วใส่คืนทุกครั้งพร้อมทั้งสับไฟก่อนการดึงครั้งต่อไป ให้ตัวแปร X คือการดึงไฟออกจากสํารับแล้วให้ได้ queen โพธิ์แดง ดังนั้น X จึงเป็นตัวเลขของครั้งที่ดึงไฟแล้วได้ queen โพธิ์แดง จึงเป็นไปได้ว่า X คือครั้งแรกแล้วได้ หรือครั้งที่ 3 แล้วได้ หรือครั้งที่ 21 แล้วได้ หรือครั้งที่ 302 แล้วได้ หรือครั้งที่ 5,000 แล้วได้ หรือครั้งที่ 4 ล้านแล้วได้ หรือไม่ได้เลย ตัวอย่างนี้แสดงให้เห็นว่า X เป็นตัวเลขที่ไม่มีขอบเขต แต่สามารถนับจำนวนครั้งได้ (countably infinite number of values) ทั้ง 2 ตัวอย่างนี้แสดงให้เห็นว่าค่าของตัวแปรชนิดนี้จะเป็นตัวเลขจำนวนนับและไม่มีทศนิยม กล่าวได้อย่างหลวม ๆ ว่าค่าของตัวเลขมีช่องว่างระหว่างตัวเลข ไม่ติดกันเป็นฟีด จึงมีลักษณะเป็น mass ไม่เป็น density จึงเรียกตัวแปรชนิดนี้ว่าเป็นตัวแปรไม่ต่อเนื่อง

ไม่ว่าจะเป็นข้อมูลแบบต่อเนื่องก็ดีหรือแบบไม่ต่อเนื่องก็ดี ข้อมูลมีความกระจัดกระจายพอสมควร กล่าวคือ เป็นตัวเลขมากบ้าง น้อยบ้าง แตกต่างกันไป

2. ทฤษฎีการแจกแจงความน่าจะเป็น (Probability Distribution Theory)

ในทางสถิติ นักคณิตศาสตร์ได้คิดสร้างทฤษฎีการแจกแจงข้อมูลไว้หลายรูปแบบด้วยกัน เพื่อให้ได้ครอบคลุมบรรดาข้อมูลดิบทั้งหลายที่มีอยู่ในโลกนี้เท่าที่เป็นไปได้ โดยแบ่งรูปแบบการแจกแจงข้อมูลออกเป็น 2 ชนิดใหญ่ ๆ คือ รูปแบบการแจกแจงข้อมูลแบบต่อเนื่อง และรูปแบบการแจกแจงข้อมูลแบบไม่ต่อเนื่อง รูปแบบเหล่านี้ถูกเรียกว่า ทฤษฎีการแจกแจงความน่าจะเป็น (Probability distribution theory) และทุก ๆ ทฤษฎีจะมี Probability function กำกับอยู่ (ถ้าเป็นฟังก์ชันพีชคณิตธรรมดาจะไม่มีคุณสมบัติค่าความน่าจะเป็นติดตัวอยู่กับฟังก์ชัน)

ตัวอย่าง Probability distribution theory ซึ่งกำกับข้อมูลแบบต่อเนื่อง (เช่น น้ำหนัก รายได้ ความสูง ฯลฯ) ได้แก่ Normal distribution, T – distribution, X^2 –distribution ฯลฯ เป็นต้น

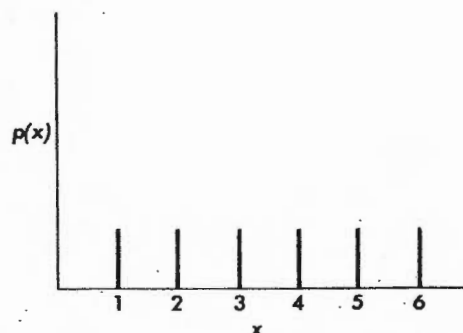
ตัวอย่าง Probability distribution theory ชนิดกำกับข้อมูลแบบไม่ต่อเนื่อง (เช่น เพศ ศาสนา อาชีพ ฯลฯ) ได้แก่ Binomial distribution, Multinomial distribution, Poisson distribution ฯลฯ เป็นต้น

เมื่อมีทฤษฎีรูปแบบการแจกแจงข้อมูลเหล่านี้ให้ใช้อย่างมากมาย ก็หมายถึงว่า เมื่อมีข้อมูลดิบ ไม่ว่าจะเป็นข้อมูลทางสังคมศาสตร์หรือข้อมูลทางวิทยาศาสตร์ ซึ่งได้แยกตามประเภทข้อมูลแล้ว คือ ข้อมูลแบบต่อเนื่องก็ดี แบบไม่ต่อเนื่องก็ดี ก่อนที่จะนำมาวิเคราะห์ทางสถิติได้ ต้องดูก่อนว่า ข้อมูลที่มีอยู่นั้นสามารถอยู่ภายใต้การกำกับของทฤษฎีการแจกแจงข้อมูลชนิดใด ทั้งนี้ เพื่อจะได้ใช้ คุณสมบัติของทฤษฎีการแจกแจงข้อมูลนั้น ๆ มาวิเคราะห์ข้อมูลที่เรา มีอยู่ เช่น ใช้สูตรอะไร ตารางสถิติตารางใด ถ้าพบว่าข้อมูลดิบที่ได้มาสามารถอยู่ภายใต้กำกับของทฤษฎีใดทฤษฎีหนึ่งเรียบร้อยแล้ว เราจึงเรียกตัวแปรของข้อมูลชุดนี้ว่าเป็น random variable³ นั่นคือ ถ้าตัวแปรที่เป็นเพียง variable ใดๆ จะยังไม่สามารถนำมาวิเคราะห์ทางสถิติได้ ตัวแปรจะต้องเป็น random variable เสียก่อนจึงจะทำการวิเคราะห์ทางสถิติได้

³เนื่องจากตัวแปรแบ่งออกเป็นแบบต่อเนื่อง (continuous) และแบบไม่ต่อเนื่อง (discrete) จึงแยกอธิบาย random variable ออกเป็น discrete random variable และ continuous random variable

Discrete Random Variable

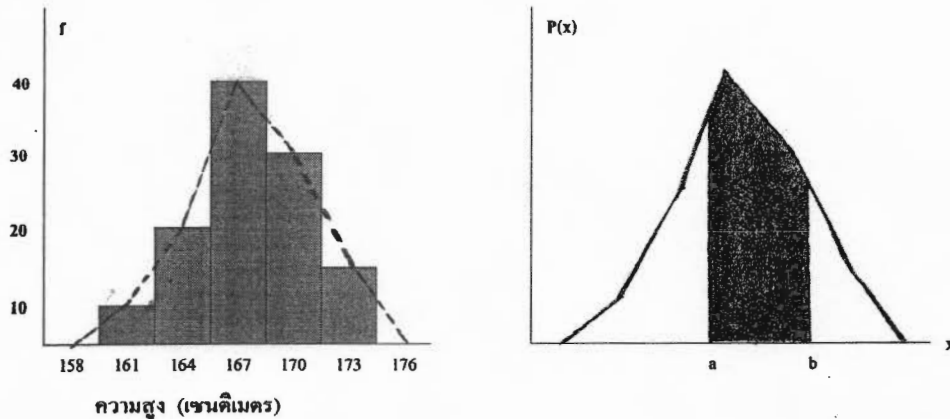
ถ้าตัวแปร X มีชุดของค่าที่เป็นตัวเลขไม่ต่อเนื่อง เช่น ลูกเต๋านึงลูกมีแต้มของหน้าต่างๆ ในแต่ละหน้าเป็นตัวเลขที่ไม่ต่อเนื่อง (discrete values) คือหน้า 1, 2, 3, 4, 5 และ 6 แต้ม ถ้าให้ลูกเต๋าลูกนี้เป็นลูกเต๋ามีความเที่ยงตรง ไม่ถูกถ่วงน้ำหนักให้หน้าไหนหน้าหนึ่งปรากฏอยู่เสมอ ดังนั้น หน้าทุกหน้าจะมีโอกาสเกิดเท่า ๆ กันในการโยนแต่ละครั้ง แต่ละหน้าของลูกเต๋าจึงมีค่าความน่าจะเป็นเท่า ๆ กันคือ $1/6$, $1/6$, $1/6$, $1/6$, $1/6$ และ $1/6$ โดยมีผลรวมของค่าความน่าจะเป็นทั้งหมดเป็น 1 เราจึงเรียกตัวแปรแบบนี้ว่า discrete random variable และ function $p(x)$ ซึ่งมีค่าเป็น $1/6$ ทั้งหมด ถูกเรียกว่า probability mass function ของตัวแปร X นั่นคือ เราเน้นความสำคัญของค่าความน่าจะเป็นที่มีประจำตัวเหตุการณ์ที่จะเกิดทุกเหตุการณ์ คือการเกิดหน้าที่มีแต้มต่าง ๆ และเมื่อรวมกันเป็น 1 เราจึงเรียกตัวแปรนี้ว่า random variable หรือ chance variable หรือ stochastic variable



Probability mass function for a discrete random variable, where $p(x) = 1/6$, for $x = 1, 2, \dots, 6$

Continuous Random Variable

แนวคิดของ discrete random variable สามารถขยายมาสู่ตัวแปร X ที่มีชุดของตัวเลขแบบต่อเนื่องได้ เส้นต่อเนื่องที่เรียกว่า relative frequency polygon (เส้นประในภาพซ้ายมือที่โยงจุดกึ่งกลางของแต่ละ histogram) ถูกลากต่อกันเป็นเส้นโค้งที่มีความต่อเนื่องที่แสดงในภาพขวา



โดยเส้นโค้งภาพขวานี้มีสมการคือ $Y = p(x)$ และมีคุณสมบัติเพิ่มเติมคือพื้นที่ใต้เส้นโค้งทั้งหมดจะเท่ากับ 1 โดยที่พื้นที่ใต้เส้นโค้งระหว่าง $X = a$ และ $X = b$ (พื้นที่แรเงา) ให้ค่าความน่าจะเป็นของ X ซึ่งอยู่ระหว่างค่า a และ b ย่อมมีค่าไม่ถึง 1 เขียนเป็นสัญลักษณ์ได้คือ $\Pr(a \leq X \leq b)$

เราเรียก $p(x)$ ว่าเป็น probability density function หรือเรียกสั้น ๆ ว่า density function และตัวแปร X ถูกเรียกว่า continuous random variable จะเห็นว่าตัวแปร X จะเป็น continuous random variable ได้ก็ต่อเมื่อมี probability function กำกับอยู่ และผลรวมของค่า probability function = 1 ค่าความน่าจะเป็นของตัวแปร X ซึ่งเป็นตัวแปรที่มีความต่อเนื่องจึงต้องหาเป็นช่วง เช่น หาค่าความน่าจะเป็นของ X ที่มีค่าตั้งแต่ a ถึง b ซึ่งย่อมมีค่าไม่ถึง 1 การหาค่าความน่าจะเป็นของตัวแปร X ที่ค่าใดค่าหนึ่งจึง $= 0$

ความคล้ายกันกับ discrete random variable คือ ตัวแปร X ของทั้งสองชนิดจะมี probability function กำกับอยู่และผลรวมของค่าความน่าจะเป็นของชุดตัวเลขของตัวแปร $X = 1$

ความแตกต่างกันคือ discrete random variable จะสามารถหาค่าความน่าจะเป็นของแต่ละตัวของชุดข้อมูลได้ เช่น หน้า 3 แด้มของลูกเต๋ามีค่าความน่าจะเป็น เป็น $1/6$ เป็นต้น แต่ continuous random variable ต้องหาค่าความน่าจะเป็นของตัวแปร X เป็นช่วง ๆ

เรื่องนี้อุปมาได้ว่า สมมุติมีประชาชนคนหนึ่งต้องการมีความรู้จนได้รับใบปริญญาบัตร ก็มีความจำเป็นสำหรับประชาชนคนหนึ่งที่ต้องหาสังกัดซึ่งเป็นมหาวิทยาลัยหรือวิทยาลัยเพื่อเข้าศึกษา เมื่อหาได้และสามารถเข้าเป็นนักศึกษาของมหาวิทยาลัยแห่งหนึ่งได้แล้ว ประชาชนคนนี้หรือเรียกใหม่ได้ว่า นักศึกษาผู้นี้ (เทียบได้กับ random variable) ก็ต้องทำตามระเบียบข้อบังคับว่าด้วยการศึกษาเพื่อให้ได้ใบปริญญาบัตร เมื่อศึกษาจนครบตามระเบียบ เขาก็จะได้รับใบปริญญาบัตรในที่สุด และทราบใดที่เขา

คนนี้ ตราบใดที่ยังไม่ทราบประเภทของข้อมูลและชนิดของ Probability function ที่สามารถกำกับข้อมูลชุดนี้ได้ ก็ยังไม่สามารถจะวิเคราะห์ทางสถิติได้ เพราะไม่ทราบจะวิเคราะห์อย่างไร นั่นคือต้องให้ข้อมูลชุดนี้มี Probability function ที่เหมาะสมกำกับข้อมูลชุดนี้ให้ได้ก่อน ซึ่งเรียกอย่างง่าย ๆ ว่าข้อมูลชุดนี้มีการแจกแจงเป็นแบบไหน การแจกแจงของข้อมูลทุกชนิดจะมี Probability function กำกับอยู่ เมื่อหาพบแล้วจึงทำการวิเคราะห์ทางสถิติต่อไปตามคุณสมบัติของแต่ละ Probability function

3. การแจกแจงแบบโค้งปกติ (Normal Distribution)

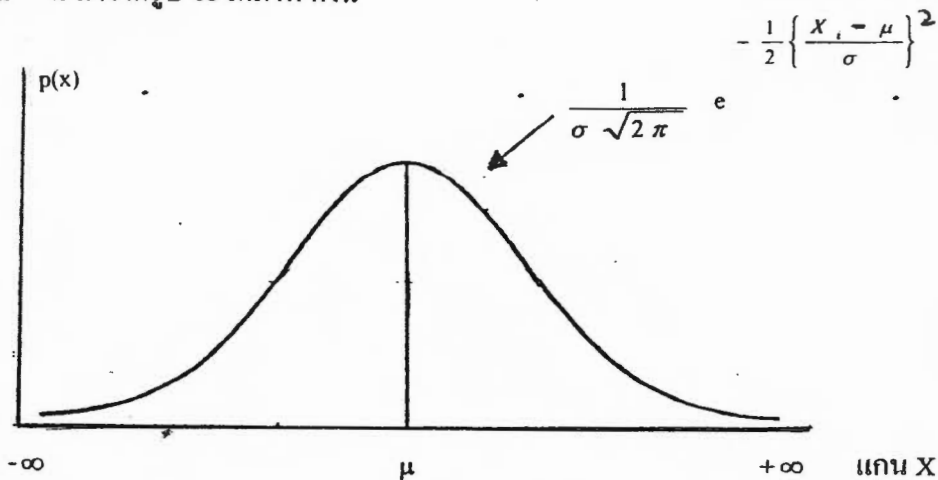
ลองมาเริ่มรู้จักกับการแจกแจงข้อมูลแบบที่รู้จักกันอย่างติดปากว่า “ โค้งปกติ ” ก่อน เพราะเป็นทฤษฎีการแจกแจงข้อมูลที่เป็นแบบพื้นฐานที่สุด โดยทั่ว ๆ ไป ในการทำวิจัย ผู้วิจัยมักจะเก็บตัวอย่าง เช่น ประมาณ 100 คน หรือมากกว่าขึ้นไปให้มาตอบแบบสอบถาม ในแบบสอบถามมักจะถามเกี่ยวกับเรื่องรายได้ อาชีพ หรือ เพศ เป็นต้น เรื่องรายได้มักจะให้ผู้ตอบกรอกตัวเลขเอง แต่เพศนั้น ผู้ตอบอาจจะกากบาทว่าเป็นหญิงหรือเป็นชาย ตัวแปรทั้งสองดังกล่าว พอจะแยกประเภทได้ว่า รายได้เป็นตัวแปรแบบต่อเนื่อง (continuous variable) ส่วนเพศเป็นตัวแปรแบบไม่ต่อเนื่อง (discrete variable) ตัวอย่างที่เก็บมา 100 คน เป็นอย่างน้อย ถือเป็นตัวอย่างที่มีขนาดใหญ่พอสมควร จึงพออนุมานได้ว่าตัวแปรรายได้น่าจะแจกแจงเป็นโค้งปกติได้ (Normal distribution) ซึ่งเป็นไปตามกฎของตัวอย่างขนาดใหญ่ ที่กล่าวว่าถ้าตัวอย่างมีขนาดใหญ่หรือใหญ่ขึ้น ใหญ่ขึ้น การกระจายของข้อมูลก็จะกลายเป็นเส้นโค้งปกติมากขึ้นมากขึ้น ดังนั้น จึงควรนำเรื่องการแจกแจงแบบโค้งปกติมาพิจารณาว่ามีคุณสมบัติอะไรบ้าง ที่นักวิจัยสามารถนำไปใช้ได้

การแจกแจงแบบโค้งปกติจะมี Probability function ประจำตัวอยู่ และเนื่องจากเป็น Probability function ของตัวแปรแบบต่อเนื่อง นักคณิตสถิติจึงเรียก function นี้อย่างเต็มยศว่า Normal probability density function ซึ่งตามคำหลังเขียนเป็นตัวย่อว่า p.d.f. คำว่า density จะทำให้ทราบว่า เป็นตัวแปรชนิด continuous คือ มีความต่อเนื่องระหว่างค่าของตัวแปร ถ้าเปรียบเทียบกับพื้นที่ก็จะทำให้เกิดพื้นที่แน่นอนที่บิ และสมการนี้จะทำให้เกิดเส้นโค้งแบบระฆังคว่ำ หางสองข้างลาดลงไปที่เท่าๆ กัน โดยไม่มีที่สิ้นสุด จะมียอดสูงเพียงยอดเดียว มีกฎเกณฑ์ทางคณิตศาสตร์เป็น function ดังนี้

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left\{ \frac{x - \mu}{\sigma} \right\}^2}$$

ค่า $\pi = 3.14159$, ค่า $e = 2.71828$

มี μ และ σ เป็นค่าพารามิเตอร์ (คือ ค่าสถิติของกลุ่มประชากร ประชากร จะหมายถึง เป็นคนก็ได้ เป็นสัตว์ก็ได้ เป็นสิ่งของก็ได้ แล้วแต่ว่านักวิจัยกำลังทำเรื่องเกี่ยวกับอะไรอยู่) ถ้านำเอา function นี้ มาวาดรูป จะได้ภาพ ดังนี้



หางของเส้นโค้งทั้งสองข้างจะไปบรรจบพื้นราบที่ $-\infty$ และ $+\infty$ (ซึ่งอยู่ไกลมาก การวาดภาพจึงต้องเปิดช่องว่างไว้ทั้งสองข้าง) และเนื่องจากเป็น probability function จึงกำหนดคุณสมบัติว่าพื้นที่ใต้เส้นโค้งทั้งหมดคือตั้งแต่ $-\infty$ วิ่งไปจนถึง $+\infty$ มีค่าเป็น 1 เสมอ ดังนั้น ถ้าตัดส่วนใดส่วนหนึ่งของภาพออกมา พื้นที่จะมีไม่ถึงค่า 1 คงเป็นทศนิยมเท่านั้น และ function ยังกำหนดให้ภาพนี้มีความโค้งหรือความสูงเป็นตัวเลขคือ 3 มีความเบ้ เป็น 0 คือ ไม่เบ้เลย หางสองข้างจะลาดออกไปเท่าๆ กัน เราอาจจะสงสัยว่านักคณิตศาสตร์ทำอะไรจึงแปลง function นี้ให้เป็นรูปกราฟดังที่วาดไว้แล้วได้ เพื่อเป็นการทบทวนความรู้เดิมเกี่ยวกับการวาดกราฟตามสมการพีชคณิต ให้ดูภาคผนวก ก. ประกอบ แต่สำหรับ Normal p.d.f. นี้ เป็นสมการที่ยากกว่าพีชคณิตมัธยมที่เคยเรียนมา จึงน่าจะทิ้งประเด็นความสงสัยในเรื่องได้ภาพนี้มาได้ได้อย่างไร โดยไว้วางใจในความถูกต้องและแม่นยำของนักคณิตศาสตร์ในการแปลงสมการมาเป็นภาพ เพราะตรงนี้ไม่ใช่ประเด็นสำคัญที่จะกล่าวถึง สิ่งที่น่าสนใจมากกว่ารูปกราฟก็คือคุณสมบัติอื่นๆซึ่งนำมาใช้ในการวิเคราะห์ข้อมูล นักคณิตศาสตร์กำหนดว่า

1. ถ้าข้อมูลกระจายเป็นโค้งปกติแล้ว การหาค่าเฉลี่ยให้ใช้สูตร

$$\mu = \frac{\sum_{i=1}^n X_i}{N}$$

2. การหาค่าเบี่ยงเบนมาตรฐานให้ใช้สูตร

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{N}}$$

3. โดยปกติตัวแปรที่อยู่บนเส้นแกน X ของกราฟ อาจเป็นการกระจายของคะแนนดิบก็ได้ แต่คะแนนดิบอาจมีค่าได้กว้างขวางมาก เช่น ค่าที่เป็นตัวเลขน้อยๆ ดังเช่นน้ำหนักของคนเป็นกิโลกรัม ซึ่งน้ำหนักสูงสุดคงไม่เกินตัวเลข 3 หลัก หรือค่าที่เป็นตัวเลขมาก ๆ เช่น รายได้ ซึ่งมีตั้งแต่หลักร้อย จนถึงหลักแสนก็ได้ ฉะนั้น ในกราฟของเส้นโค้งปกติ แกน X จึงอาจรวบรวมตัวเลขที่กว้างขวางหลายแบบหลายชนิดเช่นนี้ให้เป็นตัวเลขที่กะทัดรัดขึ้น นั่นคือการเปลี่ยนคะแนนดิบทั้งหลายให้เป็นคะแนน z-score (คือ คะแนนมาตรฐาน โดยเริ่มตั้งแต่เลข 0, 1, 2, ไปเรื่อย ๆ และเป็นคะแนนที่ไม่มีหน่วย เพื่อให้ข้อมูลทั้งหลายอยู่ในมาตรฐานเดียวกัน คะแนนมาตรฐานนี้สามารถมีทศนิยมได้ เช่น 0.23, 1.56,) การเปลี่ยนคะแนนดิบเป็นคะแนนมาตรฐานก็ทำได้ไม่ยาก เพียงนำเอาข้อมูลดิบแต่ละตัวมาเข้าสู่สูตร

$$Z = \frac{X_i - \bar{X}}{S.D.} \quad \text{สำหรับข้อมูลของกลุ่มตัวอย่าง}$$

$$Z = \frac{X_i - \mu}{\sigma} \quad \text{สำหรับข้อมูลของกลุ่มประชากร}$$

ค่า $\bar{X}$ และ S.D. (ค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง) ให้คำนวณจาก

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$$S.D. = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}$$

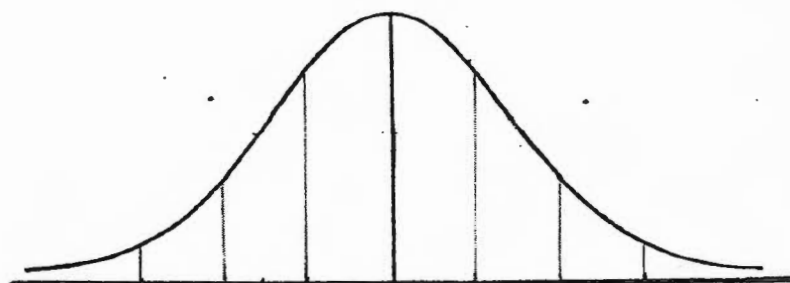
(ถ้าใช้ค่า S.D. เป็นค่าประมาณของ σ ประชากร สูตรจะเปลี่ยนไปเป็นหารด้วย n-1)

จึงได้ค่า z ออกมาเป็นคะแนน z-score ซึ่งเล็กกะทัดรัด โดยทั่ว ๆ ไปค่า z-score จะวิ่งอยู่ระหว่าง $-\infty$ ถึง $+\infty$ แต่ข้อมูลดิบต่างๆที่ถูกแปลงเป็น z-score แล้วมักจะอยู่ในช่วงระหว่าง -3 ถึง +3 z-score ไม่ว่าข้อมูลที่เก็บมาจะมีขนาดตัวอย่างใหญ่โตมากมายเพียงใด เช่น สมมุติว่ามีข้อมูลชุดหนึ่งเก็บมา 500 คน (อาจเป็นค่าน้ำหนักหรือความสูงหรือรายได้ ซึ่งเป็นตัวแปรแบบต่อเนื่อง) เมื่อเอาข้อมูลเหล่านี้มาคำนวณหาค่าเฉลี่ย ($\bar{X}$) และหาค่าเบี่ยงเบนมาตรฐาน (S.D.) ดังในสูตรข้างต้น ตัวอย่างเช่น สมมุติได้ค่าความสูงของคนกลุ่มนี้ ดังนี้

$$\bar{X} = 162 \text{ เซนติเมตร} \quad S.D. = 3 \text{ เซนติเมตร}$$

นั่นคือค่า S.D. ซึ่งเป็นค่าวัดการกระจายเฉลี่ยว่าข้อมูลอยู่ห่างจาก $\bar{X}$ เป็น 3 เซนติเมตร นั้น จะถูกยกให้เป็น 1 z-score เสมอ (แม้ค่า S.D. จะมีค่าเป็นตัวเลขใด ๆ ก็ตาม เช่น อาจเป็นเรื่องรายได้ = 2500

บาท หรือระยะห่าง = 150 กิโลเมตร ก็จะถูกเทียบให้ เท่ากับ 1 z-score (เสมอ) ซึ่งสามารถนำเสนอด้วยภาพได้ดังนี้



คะแนนดิบ	153	156	159	162	165	168	171	เซนติเมตร
คะแนน z-score	-3	-2	-1	0	+1	+2	+3	

นักคณิตสถิติได้คำนวณพื้นที่ใต้เส้นโค้งปกติออกมาดังนี้

พื้นที่ที่อยู่ระหว่าง -1 ถึง +1 z-score มีค่า probability = 0.6826

พื้นที่ที่อยู่ระหว่าง -2 ถึง +2 z-score มีค่า probability = 0.9544

พื้นที่ที่อยู่ระหว่าง -3 ถึง +3 z-score มีค่า probability = 0.9974

ซึ่งถ้าแปลงค่า probability เหล่านี้เป็นเปอร์เซ็นต์จะมีความหมาย ดังนี้

1. คนที่สูงระหว่าง 159 - 165 เซนติเมตร มี 68.26 เปอร์เซ็นต์ของจำนวนตัวอย่าง ดังนั้น ถ้าตัวอย่างมีขนาด 500 จะมีคนที่สูงระหว่าง 159 - 165 นิ้ว อยู่จำนวน 500 $(68.26 \div 100) = 341.3$ คน
2. คนที่สูงระหว่าง 156 - 168 เซนติเมตร มี 95.44 เปอร์เซ็นต์ของจำนวนตัวอย่าง คิดเป็น จำนวนคนได้ 500 $(95.44 \div 100) = 477.2$ คน
3. คนที่สูงระหว่าง 153 - 171 เซนติเมตร มี 99.74 เปอร์เซ็นต์ของจำนวนตัวอย่าง คิดเป็น จำนวนคนได้ 500 $(99.74 \div 100) = 498.7$ คน

นั่นคือ ไม่ว่าตัวอย่างที่นำมาศึกษาจะมีเป็น 1000 คนหรือใหญ่กว่านี้ก็ตาม 99.74 เปอร์เซ็นต์ ของจำนวนคนในตัวอย่างจะตกอยู่ในพื้นที่ -3 ถึง +3 z-score จึงเป็นเหตุผลที่เรามักจะเห็นภาพเส้น โค้งปกติมักจะวาดไว้เพียง 3 z-score ซ้ายและขวาเท่านั้น เหตุผลที่ค่า z-score ข้างซ้ายของเส้นโค้งปกติ เป็นค่าลบ ก็คือ ค่าของตัวแปรที่อยู่ทางซ้ายซ้ายของเส้นโค้งต้องเป็นค่าที่ลดลง จึงเอาค่าของ S.D. ไป ลบออกจากค่าเฉลี่ยเสมอ

จากภาพ ทำให้ทราบว่าถ้าเราเอาค่า $\bar{X}$ บวกกับค่า S.D. จะเป็นค่าคะแนนดิบที่อยู่ตรง 1 z-score ถ้ายกกับ 2 เท่าของค่า S.D. จะมีค่าเป็นคะแนนดิบที่อยู่ตรง 2 z-score หรือ $\bar{X}$ ลบกับค่า S.D. จะเป็น ค่าคะแนนดิบที่อยู่ตรง -1 z-score และถ้าลบกับ 2 เท่าของค่า S.D. จะเป็นค่าคะแนนดิบที่อยู่ตรง -2 z-score เป็นดังนี้ไปเรื่อยๆ

สมมุติว่าถ้ามีคะแนนดิบเป็น 63 จะคิดเป็นค่า z-score ได้อย่างไร นั่นคือ นักศึกษาต้องเอาคะแนนดิบ 63 ไปใส่ในสูตร z-score นั่นเอง

$$Z = \frac{63 - 62}{3} = +0.33$$

ปัญหาต่อไปก็คือ นักคณิตสถิติกำหนดตัวเลข 0.6826 หรือ 0.9544 หรือ 0.9974 ได้อย่างไร คำตอบที่สั้นที่สุดก็คือ คำนวณมาจาก function ที่กำหนดความโค้งของเส้นโค้งปกตินั่นเอง นั่นคือ สมการ Normal probability density function เป็นดังนี้

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left\{ \frac{X - \mu}{\sigma} \right\}^2}$$

ซึ่งคำนวณได้ดังนี้

ในการหาค่า probability ของสมการ Normal p.d.f. นี้ต้องอาศัยสูตร 2 สูตร คือ

1. อาศัยหลักของ Taylor's series

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

2. อาศัยหลักการ integration ซึ่งมีเครื่องหมายคล้ายตัว s แต่เขียนยืดยาว มีความหมายว่า “รวม” แต่การรวมแบบ integrate จะรวมแบบพื้นที่ทุกหย่อมหยั่วซึ่งจะไม่มีช่องว่างให้เหลืออยู่เลย จึงเหมาะกับการคิดหาค่า probability จากจุดหนึ่งไปอีกจุดหนึ่งโดยระยะห่างของทั้งสองจุดจะไม่มีช่องว่างใดแทรกอยู่เลย ซึ่งเหมาะกับตัวแปรแบบ continuous คือ มีค่าทศนิยมที่ติดต่อกันได้เป็นแพ มีสูตรที่ต้องใช้ดังนี้

$$\int f(x)^n dx = \frac{x^{n+1}}{n+1} + c \dots \quad (\underline{x^n \cdot x^1 = x^{n+1}})$$

สมการ Normal เขียนใหม่ได้ ดังนี้

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2}z^2} \dots \dots \dots (\text{เปลี่ยนกำลังของค่า } e \text{ ให้เป็น}$$

ค่า z-score)

$$\begin{aligned} \text{ซึ่ง} \quad e^{-\frac{1}{2}z^2} &= 1 + \frac{\left(-\frac{1}{2}z^2\right)}{1!} + \frac{\left(-\frac{1}{2}z^2\right)^2}{2!} + \frac{\left(-\frac{1}{2}z^2\right)^3}{3!} + \dots \\ &= 1 - \frac{z^2}{2(1!)} + \frac{z^4}{4(2!)} - \frac{z^6}{8(3!)} + \dots \end{aligned}$$

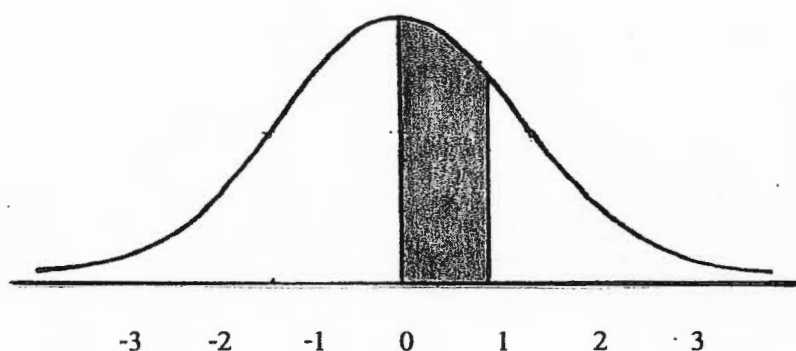
$$= 1 - \frac{z^2}{2} + \frac{z^4}{8} - \frac{z^6}{48} + \dots$$

ดังนั้น ถ้าหาค่า probability ของ z-score ที่เริ่มตั้งแต่ค่า 0 ถึง 1 z-score จะคำนวณได้ ดังนี้

$$\begin{aligned} & \int_{z=0}^{z=1} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \\ &= \frac{1}{\sqrt{2\pi}} \left[z - \frac{z^3}{6} + \frac{z^5}{8(5)} - \frac{z^7}{48(7)} + \dots + c \right] \Bigg|_{z=0}^{z=1} \\ &= \frac{1}{\sqrt{2\pi}} \left[z - \frac{z^3}{6} + \frac{z^5}{40} - \frac{z^7}{336} + \dots + c \right] \Bigg|_{z=0}^{z=1} \\ &= \frac{1}{\sqrt{2\pi}} \left[1 - \frac{1^3}{6} + \frac{1^5}{40} - \frac{1^7}{336} + \dots + c \right] - [0 + \dots + c] \\ &= \frac{1}{\sqrt{2\pi}} (1 - 0.1667 + 0.0250 - 0.00298 + \dots) \\ &= 0.3413 \end{aligned}$$

ข้อความทางคณิตศาสตร์ที่กล่าวมาทั้งหมด ก็เป็นเรื่องที่นักศึกษาสายสังคมศาสตร์ไม่ชอบอ่านเอาเลย ตรงนี้ถึงนักศึกษาอ่านไม่เข้าใจก็ไม่เป็นไร ขอแค่เพียงให้สรุปได้ดังนี้ก็พอแล้ว

1. ให้อ่านตัวเลขเพียงผ่านๆ ไป และจดจำเฉพาะผลลัพธ์ คือ ตัวเลข 0.3413 โดยยอมรับว่าผลลัพธ์นี้เป็นตัวเลขที่นักคณิตศาสตร์คิดออกมาอย่างถูกต้องแล้ว เป็นตัวเลขที่คิดมาจาก Normal p.d.f. ในช่วง z-score ตั้งแต่ 0-1 จึงเป็นค่า probability ของช่วงพื้นที่ดังกล่าว ดังรูป

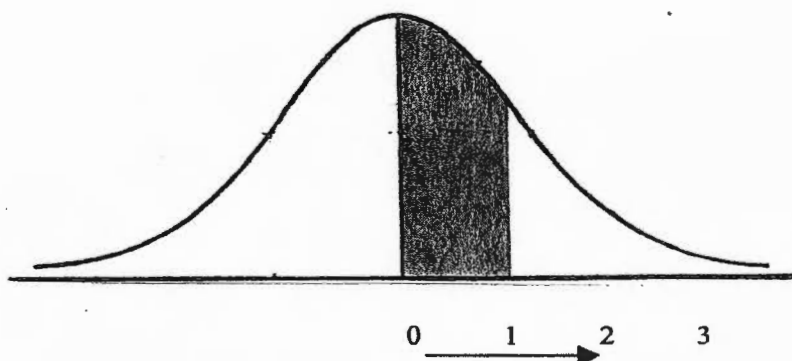


2. เนื่องจากความลาดของเส้นโค้งปกติทั้งสองข้างจากจุดที่มี $z\text{-score} = 0$ มีความลาดชันเท่ากัน ดังนั้น probability ของช่วง -1 ถึง 0 จึง $= 0.3413$ เหมือนกัน ถ้าเราเอาสองข้างมาบวกกันจะได้ 0.6826 จึงตอบคำถามที่ตั้งเอาไว้ว่านักคณิตศาสตร์คิดค่า 0.6826 มาได้อย่างไร ค่าอื่นๆ คือ 0.9544 และ 0.9974 ก็หาได้เช่นกัน

3. ถ้าเราเปิดตารางพื้นที่ใต้เส้นโค้งปกติ (ภาคผนวก ข) จะเห็นว่าตารางนี้ประกอบด้วยค่า $z\text{-score}$ ตั้งแต่ 0 ถึง 3.9 ใน column แรก และ 0 ถึง 9 ในแถวบนแรกซึ่งเป็นค่าทศนิยมตำแหน่งที่ 2 ต่อจาก column แรก ถ้าดูจากค่า $z\text{-score}$ ใน column แรก โดยวิ่งลงมาที่ตัวเลข 1.0 และมองถัดไปอีก 1 column คือ column ที่ 0 จะเห็นค่า 0.3413 ซึ่งเท่ากับที่คำนวณให้ดูแล้ว

4. ตารางนี้หาค่า probability ให้กับทุกๆ ค่าของ $z\text{-score}$ หหมดแล้ว ตั้งแต่ค่า 0 ถึง 3.9 ผู้ใช้จึงเลือกใช้เอาโดยอ่านจากตาราง ไม่ต้องคิดคำนวณเองเหมือนที่ทำได้เป็นตัวอย่าง เนื่องจากตารางนี้เป็นตารางของค่า $z\text{-score}$ กับค่า probability ถ้าทราบค่า $z\text{-score}$ ก็ไปอ่านค่า probability ได้ หรือถ้าทราบค่า probability ก็อ่านย้อนกลับมาเป็นค่า $z\text{-score}$ ได้

5. ตารางนี้เป็นการอ่านค่า probability ตั้งแต่ค่า $z\text{-score} = 0$ ออกไปทางหางขวาของรูปจนถึงค่า $z\text{-score}$ ที่เราต้องการ ซึ่งสังเกตได้จากค่า $z\text{-score} = 0$ มีค่า probability $= 0$ แปลว่ายังไม่มีพื้นที่เลย จึงเป็นตารางที่บอกพื้นที่ที่วิ่งจากจุดกึ่งกลางออกไปทางปลายหางขวา ดูภาคผนวก ข



ยังมีตารางชนิดอื่นที่อ่านพื้นที่จากปลายหางขวา ($+\infty$) วิ่งเข้าหา $z\text{-score} = 0$ ถ้าอ่านจากตารางนี้ ที่ $z\text{-score} = 0$ จะได้ค่า probability $= 0.5000$ และถ้าอ่านค่า $z\text{-score} = 1$ จะได้ค่า probability $= 0.1587$ ถ้าดูจากรูปก็คือพื้นที่นอกแฉาด้านขวา ซึ่งถ้าต้องการพื้นที่ที่แฉา ให้เอาค่า 0.1587 ไปลบออกจาก 0.5000 ก็ได้ค่า 0.3413 ซึ่งเท่ากับพื้นที่ที่แฉาข้างบน ดูภาคผนวก ค

และสุดท้ายก็ยังมีตารางอิทธิพลหนึ่งที่อ่านพื้นที่ตั้งแต่หางซ้ายโดยเริ่มที่ -3.5 ไปจนถึงปลายหางขวาคือค่า +3.5 เป็นพื้นที่สะสมไปเรื่อย แต่ไม่ค่อยเป็นที่นิยมใช้มากนัก ดูภาคผนวก ง1 และ ง2 ฉะนั้น เมื่อหีบตารางพื้นที่ใต้เส้นโค้งปกติออกมาใช้ จึงต้องตรวจสอบชนิดของตารางให้เรียบร้อยก่อนว่าเป็นแบบไหน จึงจะใช้ได้อย่างถูกต้อง

6. การหาค่า probability ไม่ได้หาได้จาก probability function แบบง่าย ๆ ที่ปรากฏใน footnote ที่ 1 เท่านั้น (จำนวนความถี่ของแต่ละอันดับการขึ้น ทารด้วย จำนวนความถี่ทั้งหมด) แต่หาจาก Probability function ที่คิดจากกฎเกณฑ์ทางคณิตศาสตร์ที่ยากขึ้นดังเช่น normal probability function ดังตัวอย่างที่กล่าวมาทั้งหมด ซึ่งก็อาศัยตารางพื้นที่ใต้เส้นโค้งปกติในการหาค่าความน่าจะเป็นโดยไม่ต้องคำนวณเอง

7. ข้อมูลดิบใด ๆ ที่ไม่ได้เป็นควมแปรปรวนต่อเนื่อง และไม่ได้กระจายเป็นโค้งปกติจะใช้สูตรเหล่านี้ไม่ได้

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

$$S.D. = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}$$

$$Z = \frac{X_i - \bar{X}}{S.D.}$$

และจะหาค่า probability โดยใช้ตาราง “พื้นที่ใต้เส้นโค้งปกติ” ก็ไม่ได้เช่นกัน

8. จากการอ่านตารางพื้นที่ใต้เส้นโค้งปกติ สมมุติว่าดูที่ z-score = 2.15 ได้ค่า probability = 0.4838 แสดงว่าเป็นค่า probability ของ z-score ตั้งแต่ 0 ถึง 2.15 จึงเป็นค่า probability ที่มาจากช่วงของข้อมูลที่อยู่ตรง 0 z-score ถึงข้อมูลที่อยู่ตรง 2.15 z-score ไม่ใช่เป็นค่า probability เฉพาะของข้อมูลที่อยู่ตรงจุด 2.15 z-score ดังนั้นการหาค่า probability ของข้อมูลแบบ continuous ณ จุดใดจุดหนึ่งจึงหาไม่ได้ ซึ่งก็คือ 0 นั่นเอง ต้องหาเป็นช่วงๆจึงจะหาได้ ซึ่งจะเห็นจากตารางว่าค่า probability ที่ z-score = 0.0 จะเท่ากับ 0.0000 (ตารางภาคผนวก ข)

มาถึงจุดนี้อาจมีความสงสัยว่าข้อความเบื้องต้นทั้งหมดจะทราบไปทำอะไร เป้าหมายใหญ่ของวิชาสถิติข้อมต้องนำไปใช้กับการวางแผนหรือสร้างนโยบายได้ ซึ่งกว่าจะถึงขั้นนั้นได้ก็ยังคงต้องทราบหลักการทางสถิติต่อไปอีก แต่ความรู้ในขั้นนี้ก็สามารถแก้โจทย์หรือทำโจทย์ทางสถิติได้ในระดับหนึ่งเพื่อความเข้าใจมากยิ่งขึ้น น่าจะลองศึกษา probability function ซึ่งเป็นของควมแปรปรวนไม่ต่อเนื่อง

(discrete variable) อีกหนึ่ง function จากนั้นจึงค่อยดูตัวอย่างโจทย์เปรียบเทียบกับกัน และวิธีแก้โจทย์ต่อไป

ข้อสรุปที่สำคัญของการแจกแจงแบบโค้งปกติ

จริงๆแล้ว การแจกแจงข้อมูลแบบโค้งปกติเป็น function ทางคณิตศาสตร์ที่เป็นทฤษฎีขั้นบริสุทธิ์ (pure theory) ไม่ได้เกิดจากข้อเท็จจริงที่ได้มาจากการสังเกตการณ์ในธรรมชาติ อย่างไรก็ตาม การแจกแจงแบบโค้งปกติก็มีประโยชน์อย่างมากมาในสถิติแบบอ้างอิง ซึ่งมีเหตุผลอย่างน้อย 4 ข้อในการอธิบายเรื่องนี้

1. ประชากรชุดหนึ่งอาจถูกยอมรับว่าแจกแจงเป็นโค้งปกติก็เพราะเป็นสิ่งที่รับรู้หรือสมมุติว่าเป็นจริง ซึ่งจริง ๆ แล้วอาจเป็นการแจกแจงแบบอื่น แต่คิดกันว่าการแจกแจงนั้นเป็นโค้งปกติ ตัวอย่างเช่น ความสูงของคนกลุ่มใหญ่กลุ่มหนึ่งก่อตัวเป็นรูปแบบของโค้งปกติโดยประมาณ การวัด IQ หรือน้ำหนักของคนก็เช่นกัน มีความเข้าใจกันว่าแจกแจงเป็นโค้งปกติ แม้ความคิดเช่นนี้โดดเด่นในศตวรรษที่ 19 แต่กฎเกณฑ์ทางคณิตศาสตร์ของโค้งปกติไม่ได้เกิดจากกฎที่มาจากธรรมชาติ (nature's rule) การแจกแจงหลายชนิดที่มีความคล้ายคลึงกับการแจกแจงแบบโค้งปกติเกิดขึ้นได้ในหลายสาขาวิชา นอกเหนือจากทางชีววิทยาดังกล่าวแล้วข้างต้น

2. อาจเป็นเรื่องของความสะดวก (บนพื้นฐานทางคณิตศาสตร์) ที่สมมุติว่าประชากรแจกแจงเป็นโค้งปกติ เพราะ function ของโค้งปกติมีคุณสมบัติทางคณิตศาสตร์ที่สำคัญ ๆ ซึ่งทฤษฎีการแจกแจงชนิดอื่นๆ ไม่มี การสมมุติว่าข้อมูลแจกแจงเป็นโค้งปกติทำให้ได้รับผลทางคณิตศาสตร์มากมายในการพัฒนาวิธีอื่น ๆ ได้ ปัญหาทางสถิติอันหลากหลายสามารถหาคำตอบได้ถ้าประชากรแจกแจงเป็นโค้งปกติ โค้งปกติยังเป็นแม่บทของการแจกแจงชนิดอื่น ๆ ในวิชาสถิติ และแม้แต่การประยุกต์ใช้วิธีทางสถิติต่าง ๆ ก็สามารถพัฒนาขึ้นมาได้โดยใช้ทฤษฎีของโค้งปกติ แม้ว่าสมมุติฐานของโค้งปกติไม่จริง

3. การแจกแจงแบบโค้งปกติให้บริการในด้านการประมาณค่าความน่าจะเป็นที่ดีกับทฤษฎีการแจกแจงข้อมูลชนิดอื่นๆ ได้ ถ้าการหาค่าความน่าจะเป็นของทฤษฎีอื่น ๆ มีความยุ่งยากและทำงานลำบาก ตัวอย่างเช่น การแจกแจงแบบโค้งปกติให้ค่าความน่าจะเป็นโดยประมาณที่ดีแก่ข้อมูลที่แจกแจงแบบ Binomial (ข้อมูลเป็นแบบไม่ต่อเนื่อง) ภายได้เงื่อนไขบางประการ (เช่นให้ n มีขนาดใหญ่เป็นต้น)

4. ข้อนี้เป็นเรื่องความเกี่ยวเนื่องอย่างใกล้ชิดระหว่างขนาดของกลุ่มตัวอย่าง (n) กับเรื่องราวของ Sampling distribution⁴ ที่วิ่งเข้าใกล้รูปแบบของโค้งปกติ กล่าวคือ Sampling distribution ที่กลุ่มตัวอย่างมีขนาดใหญ่จะมีการแจกแจงแบบโค้งปกติ แม้ประชากรของกลุ่มตัวอย่างจะไม่แจกแจงเป็นโค้งปกติก็ตาม กฎนี้ถูกเรียกว่า Central Limit Theorem กฎข้อนี้เป็นข้อสำคัญและเป็นหลักการที่เป็นประโยชน์มากซึ่งได้มาจากทฤษฎีของการแจกแจงแบบโค้งปกติ ซึ่งนำไปทดสอบสมมุติฐานกับการประมาณค่าแบบช่วงที่ใช้กันอย่างแพร่หลาย

⁴ Sampling distribution หมายถึงการกระจายของค่าสถิติ เช่น ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน ฯลฯ ของกลุ่มตัวอย่างที่เป็นไปได้ทั้งหมด ไม่ว่าจะสุ่มแบบใดหรือไม่สุ่มก่อนสุ่มครั้งต่อไปก็ตาม ตัวอย่างเช่น สมมุติมีประชากรหนึ่งประชากร ประกอบด้วยข้อมูล 5 ตัว คือ 2.2, 3.8, 4, 5.4 และ 6.6 ถ้าสุ่มตัวอย่างมาขนาด 2 โดยใช้วิธีสุ่มแล้วไม่ใส่คืน จะมีจำนวนตัวอย่างทั้งหมด 10 ตัวอย่างดังนี้

2.2, 3.8 มีค่าเฉลี่ย = 3	2.2, 4 มีค่าเฉลี่ย = 3.1	2.2, 5.4 มีค่าเฉลี่ย = 3.8
2.2, 6.6 มีค่าเฉลี่ย = 4.4	3.8, 4 มีค่าเฉลี่ย = 3.9	3.8, 5.4 มีค่าเฉลี่ย = 4.6
3.8, 6.6 มีค่าเฉลี่ย = 5.2	4, 5.4 มีค่าเฉลี่ย = 4.7	4, 6.6 มีค่าเฉลี่ย = 5.3
5.4, 6.6 มีค่าเฉลี่ย = 6		

จากนี้ให้นับค่าเฉลี่ยทั้งหมดที่ได้สืบความเขียนบนเส้นตรงที่เป็นแกน X ในกรณีนี้พบว่าไม่มีค่าเฉลี่ยตัวใดที่มีค่าซ้ำกันเลย จึงยังไม่มีเกิดเส้นโค้งที่เรียกว่าเส้นโค้งปกติ ถ้าเพิ่มขนาดตัวอย่างให้ใหญ่ขึ้นและถ้าใช้วิธีสุ่มแล้วใส่คืนก็จะทำให้มีจำนวนตัวอย่างมากขึ้น ซึ่งจะทำให้เกิดค่าเฉลี่ยของตัวอย่างมีค่าซ้ำๆ กันหลายตัว ค่าเฉลี่ยทั้งหมดจะก่อตัวเป็นเส้นโค้งปกติ เราเรียกเส้นกราฟนี้ใหม่ว่า Sampling distribution

4. การแจกแจงแบบทวินาม (Binomial Distribution)

ในการสร้างแบบสอบถาม ผู้วิจัยคงต้องมีคำถามเกี่ยวกับคุณสมบัติของผู้ตอบบ้าง เช่น เพศ เป็นชายหรือเป็นหญิง ทำงานหรือยัง ภูมิลำเนาเดิมอยู่จังหวัดอะไร เป็นต้น คำตอบนั้นผู้ตอบต้องเลือกตอบเอาตามคุณสมบัติของตัวเอง เช่น เป็นชาย ทำงานแล้ว อยู่จังหวัดลำปาง คำตอบเหล่านี้จะมีลักษณะเป็นตัวแปรแบบไม่ต่อเนื่อง เป็นชายก็เป็นชาย เป็นหญิงก็เป็นหญิง ถ้าหากมีจำนวนชายอยู่ 3 คน และหญิงอยู่ 5 คน ตัวเลขที่นำมาใช้จึงเป็นจำนวนนับคือเลข 3 หรือเลข 5 ไม่สามารถวัดเป็นค่าที่มีทศนิยมได้ โดยทั่วๆ ไป นักคณิตศาสตร์จึงคิด probability function ที่เป็นพื้นฐานที่สุดหรือเป็นเบื้องต้นที่สุด ของตัวแปรแบบไม่ต่อเนื่อง เรียกว่า Binomial probability mass function

Binomial p.m.f. มี function ดังนี้

$$p(X; n, p) = {}^n C_x p^x q^{(n-x)}$$

(มี n และ p เป็นค่าสถิติของประชากร หรือเรียกว่า parameter)

Binomial p.m.f. จะเป็น function ซึ่งใช้แทนการแจกแจงข้อมูลใดๆที่เป็นแบบ discrete variable และมีเหตุการณ์ 2 อย่างประกอบกัน เช่น คนในโลกนี้มี 2 เพศ คือชายกับหญิง ถ้าจะวิเคราะห์ข้อมูลโดยสนใจเพศชายเป็นเกณฑ์ ก็ใช้ p แทนเพศชาย ดังนั้น q ก็ถูกใช้แทนเพศหญิง (p จะถูกเรียกว่าเป็นเหตุการณ์ที่เป็น success และ q จะถูกเรียกว่าเป็นเหตุการณ์ที่เป็น failure ตามปกติ $p + q = 1$ เสมอ) หรือถ้าข้อมูลที่เก็บมามี 3 อาชีพ เช่น บางคนเป็นครู บางคนเป็นหมอ และบางคนเป็น

ทนายความ หากการทำวิจัยสนใจอาชีพหมอเป็นด้วยนั้น ก็จะใช้ p แทนอาชีพหมอ และ q แทนอาชีพอื่น ๆ (คือครูและทนายความ) เป็นต้น

$${}^nC_x \text{ เป็นสูตรของ combination คือ } = \frac{n!}{X!(n-X)!}$$

เช่น ถ้ากลุ่มตัวอย่าง (n) มี 5 คน และสุ่มมาเพียง (X) 2 คน โดยสุ่มแล้วไม่ใส่คืน จะคำนวณออกมาได้ดังนี้

$$= \frac{5!}{2!(5-2)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)(3 \times 2 \times 1)} = 10$$

(5! อ่านว่า ห้า factorial มีความหมายว่าให้เอาตัวเลข 5 คูณตัวเลขที่ลดลงจาก 5 หนึ่งค่า คือ 4 คูณตัวเลขที่ลดลงจาก 4 หนึ่งค่าคือ 3 คูณแบบนี้ไปเรื่อยจนกระทั่งถึงเลข 1)

หมายความว่า จำนวนรูปแบบที่คน 5 คนจะถูกสุ่มออกมาทีละ 2 คน จะได้รูปแบบที่ไม่ซ้ำกันเลย มี 10 รูปแบบ เช่นสมมุติให้คนทั้ง 5 คน ชื่อ ก, ข, ค, ง และ จ สุ่มทีละ 2 คนได้คู่ต่างๆ ดังนี้ กข กค กง กจ ขค ขง ขจ คง คจ และ งจ

p และ q เป็นค่า probability เบื้องแรกซึ่งผู้ทำวิจัยต้องทราบเสียก่อน เช่นโอกาสเกิดของเด็กหญิงในโลกนี้ = 0.5 ของชายก็ = 0.5 เช่นกัน โอกาสเกิดหน้า 1 แด้มของลูกเต๋าเพียงครั้ง = 1/6 ดังนั้น โอกาสเกิดหน้าอื่นๆ ของลูกเต๋านี้ จะ = 5/6 และ $p+q$ ต้อง = 1 เสมอ

คำถามมักจะเกิดขึ้นว่า จะทราบค่า probability เบื้องแรกได้อย่างไร คำตอบก็คือ ส่วนใหญ่จะเกิดจากการสุ่มตัวอย่างซ้ำแล้วซ้ำอีกหลาย ๆ ครั้ง แต่แต่ละครั้งจะนำตัวเลขมาหาค่าความน่าจะเป็น แล้วหาค่าเฉลี่ยของความน่าจะเป็นอีกครั้งหนึ่ง (Relative frequency definition of probability) เช่น โอกาสเกิดหัวหรือก้อยของสตาจ 1 อัน จะเท่ากัน (ต่างก็เท่ากับ 0.5) โอกาสเกิดของเด็กหญิงและเด็กชายจะเท่ากันคือ 0.5 บางครั้งก็ดูจากบันทึกเก่าๆ จะทำให้พอทราบได้ บางครั้งถ้าทำวิจัยซ้ำๆ กันหลายๆ ครั้ง ได้ตัวเลขตรงกันตลอดมา ก็ทำให้ทราบได้เช่นกัน ที่ง่ายที่สุดก็เป็นการอ่านมาจากโจทย์ หรือโจทย์บอกให้นั่นเอง เมื่อแทนค่าต่างๆ ลงใน probability function ตัวเลขที่ได้มาจึงเป็นค่า probability ตัวใหม่ซึ่งเราต้องการทราบ ตัวอย่างเช่น ในครอบครัวที่มีบุตร 2 คน โดยปกติจะมีรูปแบบที่บุตร 2 คนจะเกิดเป็นเพศชายหรือหญิงได้ดังนี้ (ชาย, ชาย) (ชาย, หญิง) (หญิง, ชาย) (หญิง, หญิง) ดังนั้น ถ้าความสนใจของผู้วิจัยอยากทราบ probability การเกิดของหญิงในครอบครัวที่มีบุตร 2 คน ตามรูปแบบการกระจายตามบรรทัดบนจะเป็นอย่างไร จึงต้องใช้ Binomial p.m.f. หาค่า probability ดังกล่าวได้ ดังนี้

จำนวนบุตรหญิง	${}^nC_x p^x q^{(n-x)}$	ค่าความน่าจะเป็น	คนที่ 1	คนที่ 2
0	${}^2C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^2$	$\frac{1}{4}$	ชาย	ชาย

			ชาย	หญิง
1	${}^2C_1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^1$	$\frac{2}{4}$	หญิง	ชาย
2	${}^2C_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^0$	$\frac{1}{4}$	หญิง	หญิง

ค่า $\frac{1}{4}$ ที่ได้จาก ${}^2C_0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^2$ คำนวณได้ดังนี้

$$= \frac{2!}{0!(2-0)!}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^2 \dots\dots\dots 0! = 1, \left(\frac{1}{2}\right)^0 = 1$$

$$= \frac{2!}{2!}(1)\left(\frac{1}{4}\right) = \frac{1}{4} \text{ เทอมอื่นๆ ก็หาค่าได้ในทำนองเดียวกัน}$$

ค่า probability ตัวใหม่ซึ่งคำนวณได้จาก Binomial p.m.f. จะตีความหมายดังนี้

ค่า $1/4$ ของครอบครัวที่มีบุตรหญิง 0 คน หมายความว่าโอกาสที่ครอบครัวๆหนึ่งมีบุตร 2 คน และได้บุตรทั้งสองเป็นชายหมดจะมีเพียง $1/4$ จากจำนวนเต็มคือ 1

ค่า $2/4$ ก็หมายความว่า โอกาสที่บุตรทั้งสองจะเป็นชายหนึ่งคนเป็นหญิงหนึ่งคนเป็น $2/4$ จากจำนวนเต็ม 1 ซึ่งก็คือ บุตรที่เป็นชายหนึ่งและหญิงหนึ่ง มีโอกาสเป็นไปได้มากกว่า ชายทั้งคู่ หรือ หญิงทั้งคู่

จะสังเกตว่า ค่า probability ที่เกิดขึ้นทั้งหมดรวมกัน ($1/4 + 2/4 + 1/4$) ย่อมเท่ากับ 1 เสมอ ซึ่งก็คล้ายกับการแจกแจงแบบ Normal คือ พื้นที่ใต้เส้นโค้งรวมกันทั้งหมด = 1 แต่ข้อมูลแบบ Binomial เป็นตัวแปรแบบไม่ต่อเนื่อง (ไม่มีทศนิยม) การรวมกันของ probability ของเหตุการณ์ทั้งหมดจึงใช้เครื่องหมาย summation (Σ) และ probability function ก็คำนวณได้ง่ายกว่าของการกระจายแบบโค้งปกติ อย่างไรก็ตาม ข้อแตกต่างที่เห็นได้ชัดเจนจะเป็นดังนี้

1. การหาค่าเฉลี่ยของข้อมูลแบบ Binomial ใช้สูตร

$$\bar{X} = np$$

และค่าเบี่ยงเบนใช้สูตร

$$S.D. = \sqrt{npq}$$

n คือ ขนาดตัวอย่าง

p คือ ค่า probability เบื้องแรกซึ่งหาไว้แล้ว

q คือ $1-p$

2. ไม่มีการย่อขนาดของข้อมูลดิบให้กลายเป็น z-score ไม่ว่า n จะใหญ่โตเพียงไร ดังนั้น z-score จึงเป็นการรวบรวมข้อมูลให้เล็กลง ทำให้เหลือการกระจายห่างออกไปจากเส้นตรงกลางของโค้งปกติ (ซึ่งมี z-score = 0) เป็น 3 z-score ทางซ้าย (ซึ่งปกติจะมีค่าเป็นลบ เพราะค่าลบจะทำให้คะแนนดิบมีค่าลดลง) และ 3 z-score ทางขวา ดังนั้นค่า z-score จึงเป็นเรื่องของเส้นโค้งปกติแต่อย่างเดียว ไม่เกี่ยวกับการแจกแจงแบบ Binomial (รวมทั้งการแจกแจงแบบอื่นๆด้วย)

3. ด้วยเหตุผลของข้อ 2 ถ้าลองเปิดตารางค่า probability ของ Binomial p.m.f. จะพบว่านักสถิติได้คำนวณค่า probability ไว้โดยกำหนด n มีค่าตั้งแต่ 1-20 ซึ่งต้องใช้หน้ากระดาษถึง 5 หน้า (ดูภาคผนวก จ1-จ5) ถ้า n ใหญ่นี้อีกก็จะใช้หน้ากระดาษเพิ่มขึ้น ซึ่งก็ไม่มีใครคำนวณไว้ ถ้า n มีขนาดใหญ่กว่า 20 ผู้วิจัยจะต้องคำนวณเอง หรือมีฉะนั้นก็ให้หาค่า probability โดยอาศัยตารางหาค่า probability ของโค้งปกติแทน ซึ่งเรียกว่าเป็น Binomial approximation to Normal ซึ่งสะดวกกว่ามาก แต่ต้องทำตามเงื่อนไขที่กำหนดไว้บางอย่าง (จะกล่าวต่อและยกตัวอย่างให้เห็นในภายหลัง)

เพื่อให้เกิดความชัดเจนมากยิ่งขึ้น ให้พิจารณาตัวอย่างโจทย์สองข้อจากหนังสือสถิติเบื้องต้นเปรียบเทียบกัน รวมทั้งการแก้โจทย์และตอบคำถามที่โจทย์ถาม

1. ชายคนหนึ่งต้องเดินทางจากบ้านที่อยู่นอกเมืองไปทำงานในเมืองโดยการเดินทางเฉลี่ย 24 นาที มีค่าเบี่ยงเบนมาตรฐาน = 3.8 นาที ให้การเดินทางเป็นการแจกแจงแบบโค้งปกติ

ก. จงหาความน่าจะเป็นที่การเดินทางจากบ้านจะใช้เวลาอย่างน้อยครึ่งชั่วโมง

ข. ถ้าที่ทำงานเริ่มงานเวลา 9 โมงเช้า และเขาออกจากบ้านเวลา 8.45 น. จงหาความน่าจะเป็น (probability) ที่เขาจะไปทำงานสาย

2. สมมุติ 15 % ของต้นไม้ในป่าดิบชื้นมีใบไม้ได้รับความเสียหาย เนื่องจากมลภาวะอากาศ ถ้าถ่มต้นไม้มา 5 ต้น จงหา

ก. ความน่าจะเป็นที่ใบไม้ได้รับความเสียหาย 3 ต้น

ข. ค่าเฉลี่ยของต้นไม้ที่มีใบเสียหาย

ถ้าพิจารณาโจทย์ทั้งสองข้อจะสังเกตเห็นข้อแตกต่างกันดังนี้

ในข้อที่หนึ่ง ตัวแปรเป็นเรื่องของเวลานับเป็นชั่วโมงและนาทีได้ ซึ่งเป็นตัวแปรแบบต่อเนื่องและมีข้อความในโจทย์ระบุชัดเจนว่า “ ให้การเดินทางแจกแจงเป็นโค้งปกติ ” คุณสมบัติทั้งสองอย่างของข้อมูลชุดนี้จึงถูกกำกับโดย Normal p.d.f. ดังนั้นการคำนวณหาความน่าจะเป็นตามที่โจทย์ถามจึงอาศัยตารางพื้นที่ใต้เส้นโค้งปกติ ซึ่งไม่ต้องคำนวณเองจาก probability function และอาศัยสูตร z-score ซึ่งเป็นคุณสมบัติของ Normal p.d.f. นี้ได้

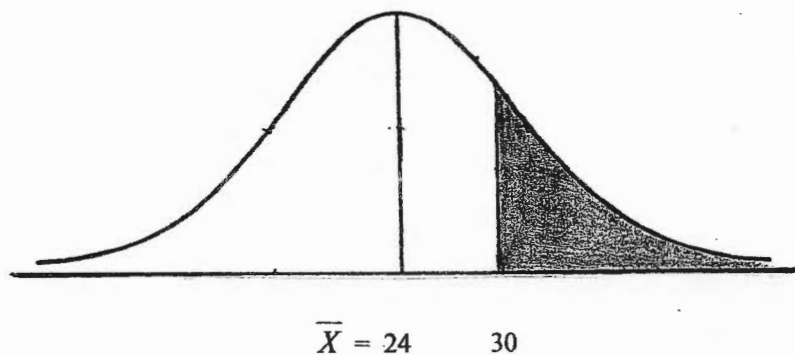
ส่วนในข้อสอง เป็นเรื่องของการสุ่มเอาต้นไม้มาพินิจพิจารณา ซึ่งต้องสุ่มมาเป็นคั่นๆ ไม่สามารถสุ่มมาเพียงครั้งคั่นหรือ $\frac{1}{4}$ ของคั่นได้ ข้อมูลแบบนี้จึงเป็นข้อมูลไม่ต่อเนื่อง (discrete variable) ทั้งยังไม่มีข้อความใดๆ ที่ระบุว่าเป็นการแจกแจงแบบโค้งปกติ แต่กลับมีข้อความกล่าวว่า 15% ของต้นไม้ในป่าดิบชื้นมิได้ได้รับความเสียหายจากมลภาวะอากาศ ย่อมแสดงว่ายังมีต้นไม้อีกจำนวน 85% ไม่ได้รับความเสียหาย เรื่องนี้จึงเป็นเหตุการณ์สองอย่างประกอบกัน ลักษณะข้อมูลแบบนี้จึงเข้าข่ายการแจกแจงแบบ Binomial และสามารถอยู่ภายใต้กำกับของ Binomial p.m.f. ได้ การหาค่าความน่าจะเป็นจึงต้องใช้ probability function ของ Binomial และการหาค่าเฉลี่ย ก็ต้องใช้สูตรของ Binomial เช่นกัน

นั่นคือ ในการเรียนสถิติขั้นพื้นฐาน ผู้เรียนยังเรียนเนื้อหาวิชาไม่มากนัก จึงยังไม่อาจทดสอบข้อมูลด้วยตนเองว่าข้อมูลที่ได้นั้นมีรูปแบบการแจกแจงเป็นแบบใด มี probability function ชนิดใดเป็นตัวกำกับ โจทย์จึงมักจะบอกข้อเท็จจริงตรงนี้ให้ ถ้าเราจับข้อเท็จจริงตรงนี้จากโจทย์ได้ ก็จะเป็นการง่ายในการแก้โจทย์เป็นอย่างมาก การแก้โจทย์หรือตอบคำถามโจทย์ทำได้ดังนี้

ข้อ 1. ค่าเฉลี่ยในการเดินทาง ($\bar{X}$) = 24 นาที

ค่าเบี่ยงเบนมาตรฐานของการเดินทาง (S.D.) = 3.8 นาที

ก. ให้หาความน่าจะเป็นที่การเดินทางจากบ้านจะใช้เวลาอย่างน้อย 30 นาที ถ้าวาดรูปจะได้ดังนี้



สิ่งที่โจทย์ถามคือพื้นที่หรือค่าความน่าจะเป็นในส่วนพื้นที่แรเงา

การวิเคราะห์ : การใช้ตารางพื้นที่ใต้เส้นโค้งปกติหาค่าความน่าจะเป็น จำเป็นต้องอ่านค่า z-score ใน column แรก เข้าไปใน column ต่อๆ ไป ซึ่งสุดท้ายแล้วพบว่าทศนิยมในตำแหน่งที่สองของค่า z-score จะปรากฏเป็นตัวเลขอะไร ก็ให้ไปที่ column นั้นซึ่งที่หัว column มีเลขตรงกัน แล้วอ่านค่าความน่าจะเป็นออกมา ดังนั้น จึงต้องแปลงค่าต่างๆ ที่โจทย์กำหนดให้ เป็นค่า z-score โดยทำ ดังนี้

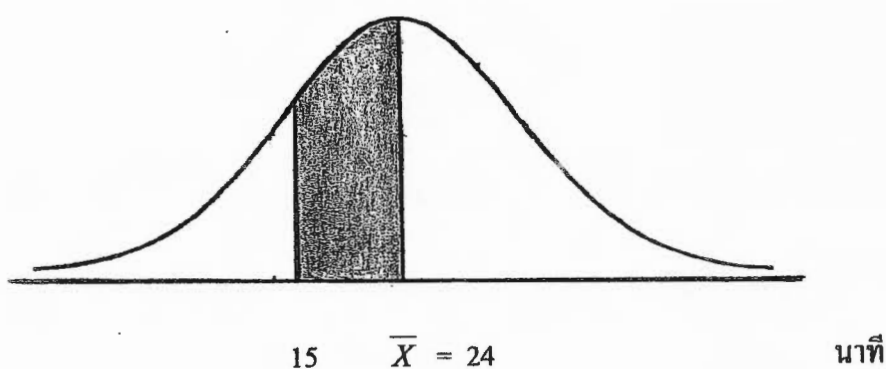
$$\begin{aligned}
 Z &= \frac{X_i - \bar{X}}{S.D.} \\
 &= \frac{30 - 24}{3.8} = 1.58
 \end{aligned}$$

จากนี้ จึงนำค่า z-score ไปเปิดตารางพื้นที่ใต้เส้นโค้งปกติ

กรณีที่ใช้ตารางในภาคผนวกหน้า ข ให้อ่านค่า z-score = 1.5 ใน column แรก และวิ่งเข้าไปที่ column ที่ 10 ซึ่งมีหัว column เขียนไว้ว่าเป็นเลข 8 ซึ่งเป็นทศนิยมตำแหน่งที่สอง จึงอ่านค่าความน่าจะเป็นได้ 0.4429 ดังนี้ถ้าดูในรูป ก็จะเป็นค่าที่ X_i อยู่ระหว่าง 24 ถึง 30 นาที เพราะตารางนี้อ่านค่าความน่าจะเป็นที่เริ่มตรงกึ่งกลางของเส้นโค้งตรงจุดที่ค่า z-score = 0 ออกไปทางหางขวาคือค่า $+\infty$ ดังนั้น คำตอบของข้อ ก. นี้จึงต้องเอาค่า 0.4429 ไปลบออกจาก 0.5000 ซึ่งเป็นพื้นที่ครึ่งหนึ่งของเส้นโค้งปกติ จะได้ = 0.0571 นั่นคือ ความน่าจะเป็นที่การเดินทางจากบ้านจะใช้เวลาอย่างน้อย 30 นาที เป็น 0.0571

ในกรณีที่ใช้ตารางในภาคผนวกหน้า ค ถ้าอ่านค่า z-score ดังที่คำนวณออกมาได้ (1.58) จะได้คำตอบโดยตรงคือ 0.0571 เพราะตารางนี้บอกค่าความน่าจะเป็นหรือก็คือค่าพื้นที่ใต้เส้นโค้งปกติจากทางสุดหางขวา ($+\infty$) เข้ามาสู่ค่า z-score (1.58) ที่เรากำลังเปิดดู จึงได้คำตอบพอดี

ข้อ ข. หมายความว่า เรามีเวลาเดินทางได้เพียง 15 นาที ก็ถึงเวลาเช้างานแล้ว (9.00 น.) ซึ่งจริงๆ แล้ว ณ เวลา 9.00 น. นั้น ชายผู้นี้ยังกำลังเดินทางอยู่ เพราะโดยเฉลี่ยเขาต้องใช้เวลาเดินทางถึง 24 นาที และเมื่อเป็นค่าเฉลี่ยของทุกๆ วันจึงตีความได้ว่า ถ้ามีเวลาเดินทางถึง 24 นาที เขาจะไม่ไปทำงานสาย ดังนั้น ค่าความน่าจะเป็นของการไปทำงานสายจึงอยู่ระหว่าง 15 ถึง 24 นาที ดังรูป



จึงหาค่าความน่าจะเป็นได้ดังนี้

$$\begin{aligned} Z &= \frac{15 - 24}{3.8} \\ &= -0.6842 \quad = -0.68 \end{aligned}$$

คิดเป็นค่าความน่าจะเป็น (โดยเปิดตารางภาคผนวก ข.) ได้ = 0.2518

ในข้อ 2. เขาสุ่มต้นไม้มา 5 ต้น ให้หาความน่าจะเป็นที่จะพบต้นไม้ที่ได้รับความเสียหาย 3 ต้น จากการวิเคราะห์แล้วทราบว่า ข้อ 2 นี้ คืออยู่ภายใต้กำกับของ Binomial p.m.f. จึงใช้สูตรในการหาความน่าจะเป็นดังนี้ คือ

$$p(x) = {}^n C_x p^x q^{(n-x)}$$

p คือค่าความน่าจะเป็นเบื้องต้น ซึ่งโจทย์บอกว่า 15% ของต้นไม้ในป่าดิบชื้นมีใบไม้ที่ได้รับความเสียหาย จึงมีค่าเป็น 0.15

$$q = 1 - 0.15 = 0.85 \dots\dots\dots (p + q = 1)$$

$$n = 5 \text{ ต้น}$$

$$x = 3 \text{ ต้นที่โจทย์ถาม}$$

จึงคำนวณหาความน่าจะเป็นได้ดังนี้

$$\begin{aligned} p(x) &= {}^5 C_3 (0.15)^3 (0.85)^{5-3} \\ &= \frac{5!}{3!(5-3)!} (0.15)^3 (0.85)^2 \\ &= \frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1)(2 \times 1)} (0.003375) (0.7225) \\ &= 0.0244 \end{aligned}$$

ซึ่งถ้าไม่คำนวณเองตามสูตร ก็สามารถเปิดตารางของ Binomial ในภาคผนวกหน้า จ1 เมื่อ $n=5$ หรือ x ในสูตรข้างบน $= 3$ $p = 0.15$ จะได้ตัวเลขเป็น 0.0244 ซึ่งตรงกับที่ได้คำนวณไว้พอดี และคำถามข้อ ข. ซึ่งให้หาค่าเฉลี่ยของต้นไม้ที่มีใบเสียหาย ให้ใช้สูตร

$$\begin{aligned} \bar{X} &= np \\ &= 5(0.15) = 0.75 \end{aligned}$$

ซึ่งคือคำตอบค่าเฉลี่ยที่โจทย์ถาม

จะเห็นว่าการแก้โจทย์ทั้งสองข้อ ต้องอาศัยความรู้เรื่อง probability function ของข้อมูลมา เป็นตัวช่วย และสามารถคำนวณหาคำตอบได้ตามที่ต้องการ

ในกรณีที่ ถ้าคำนวณด้วยอย่าง (n) ของการแจกแจงแบบ Binomial p.m.f. ใหญ่มากๆ เช่น เป็น 100 การคำนวณก็ทำได้ลำบากขึ้น จะใช้ตาราง Binomial ก็ไม่มีให้ใช้เพราะ n มีค่าใหญ่กว่า 20 ในเรื่องนี้ นักคณิตสถิติได้คิดค้นกันมานานแล้วพบว่า ถ้าข้อมูลใดมีขนาดตัวอย่างที่ใหญ่ การกระจายของข้อมูลก็จะมีลักษณะใกล้เคียงกับโค้งปกติ ดังนั้น ก็อาจขอใช้คุณสมบัติของเส้นโค้งปกติไปใช้ได้

โดยต้องทำความเข้าใจที่กำหนด และมีการปรับข้อมูลเล็กน้อย ลักษณะข้อมูลแบบ Binomial ถ้าจะขอ
ยืมคุณสมบัติของเส้นโค้งปกติไปใช้ จึงเรียกว่า Binomial approximation to Normal ซึ่งมีข้อกำหนดดังนี้

1. เมื่อขนาดตัวอย่างใหญ่ ซึ่งก็ไม่ได้กำหนดความใหญ่ที่แน่นอน จึงแนะนำให้ดูค่า np หรือ nq ถ้าคูณกันแล้วค่าก็มากกว่า 5 โดยที่ค่า p และค่า q ต่างก็ไม่ใกล้เคียงกับ 0 การกระจายแบบ Binomial ก็จะใกล้เคียงกับเส้นโค้งปกติ โดยมีสูตร standard score (z-score) เป็น ดังนี้

$$z = \frac{X - np}{\sqrt{npq}} \dots\dots(x \text{ คือจำนวน } p)$$

2. ให้เอาค่า 0.5 บวกหรือลบออกจากค่า X (แล้วแต่กรณี ดูย่อหน้าสุดท้ายในหน้านี) เพื่อให้ข้อมูลแบบขาดช่วงมีความต่อเนื่องมากขึ้น (เรียก Yate's correction for continuity) เพราะคุณสมบัติของเส้นโค้งปกติ กำหนดให้ข้อมูลที่น่ามาใช้ต้องเป็นข้อมูลแบบต่อเนื่อง จากนั้น จึงนำค่าต่างๆ มาเข้าสู่สูตร z-score ที่กล่าวแล้ว และนำค่า z-score ไปเปิดตารางพื้นที่ใต้เส้นโค้งปกติ ก็จะได้ค่า probability ที่ใกล้เคียงและประหยัดเวลามากกว่าการคิดคำนวณจาก Binomial probability function ตัวอย่างเช่น

ประชากรประกอบด้วยนักศึกษาชั้นปีหนึ่ง 1/6 ส่วน นอกนั้นเป็นนักศึกษาชั้นปีอื่นๆ ในการ
สุ่มจับฉลากชื่อนักศึกษามา 120 ครั้ง ให้หาความน่าจะเป็นที่จะได้นักศึกษาชั้นปีหนึ่ง 18 ครั้งหรือน้อยกว่า (สุ่มแล้วใส่คืน)

โจทย์ข้อนี้บอกถึงลักษณะของประชากรประกอบด้วย 2 ลักษณะ คือ นักศึกษาชั้นปีหนึ่ง มี probability เบื้องแรก = 1/6 และนักศึกษาชั้นปีอื่นๆ มี probability เบื้องแรก = 5/6 การหาค่า probability จึงควรอาศัย probability function ของ Binomial ซึ่งก็คือ ${}^nC_x p^x q^{n-x}$

$$\text{ในที่นี้ } p = 1/6 \quad q = 5/6$$

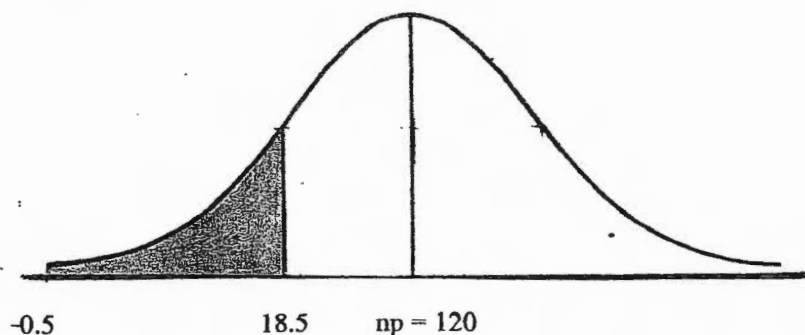
$$n = 120 \quad x = 18, 17, 16, \dots, 0 \text{ คน}$$

ดังนั้นจึงหาค่า probability ได้ ดังนี้

$${}^{120}C_{18} (1/6)^{18} (5/6)^{102} + {}^{120}C_{17} (1/6)^{17} (5/6)^{103} + \dots\dots\dots {}^{120}C_0 (1/6)^0 (5/6)^{120}$$

จะเห็นว่า ค่าต่างๆ จะคำนวณได้ลำบากและเสียเวลามาก และเมื่อพิจารณาจาก $np = 120 (1/6) = 20$ หรือ $nq = 120 (5/6) = 100$ ซึ่งเป็นตัวเลขที่ใหญ่กว่า 5 จึงทำให้ทราบว่าโจทย์ข้อนี้สามารถใช้ Binomial approximation to Normal ได้

จากตัวเลข 0 คน ถึงตัวเลข 18 คน ซึ่งเป็นค่า X จำเป็นต้องใช้ตัวเลข (-0.5) ถึง (18.5) แทน (ในที่นี้ให้เอาตัวเลข 0 คน ลบ 0.5 จึงได้ค่า -0.5 และเอาตัวเลข 18 คน บวก 0.5 จึงได้ค่า 18.5 ทั้งนี้เพื่อให้ค่าทั้งสองผ่านจุดที่โจทย์ถามคือ 0 และ 18 มิฉะนั้นจะได้คำตอบที่ไม่ถูกต้อง) ซึ่งการบวก หรือลบออกนั้น ก็เป็นการทำให้ข้อมูลแบบไม่ต่อเนื่อง มีความใกล้เคียงกับข้อมูลแบบต่อเนื่อง โดยใช้วิธีของ Yate's correction for continuity พื้นที่แฉกจากภาพข้างล่างแสดงถึงค่า probability ที่โจทย์ถาม



$$z = \frac{(-0.5) - 20}{\sqrt{120(1/6)(5/6)}} = -5.02$$

$$z = \frac{(18.5) - 20}{\sqrt{120(1/6)(5/6)}} = -0.37$$

ค่า z-score (-5.02) นำไปเปิดตารางพื้นที่ใต้เส้นโค้งปกติภาคผนวก ข. ได้ค่า 0.5000

ค่า z-score (-0.37) นำไปเปิดตารางพื้นที่ใต้เส้นโค้งปกติภาคผนวก ข. ได้ค่า 0.1443

$$\begin{aligned} \text{ดังนั้น ค่าความน่าจะเป็นจึงมีค่า} &= 0.5000 - 0.1443 \\ &= 0.3557 \end{aligned}$$

ซึ่งคำนวณได้ง่ายกว่าการใช้ Binomial p.d.f. และทำได้รวดเร็วกว่ามาก การคำนวณเช่นนี้มีข้อสังเกตดังนี้

1. โจทย์กำหนดข้อมูลให้เป็นแบบ Binomial p.d.f. แต่ตามค่า probability เป็นแบบช่วง จึงสามารถนำเอาคุณสมบัติของการแจกแจงแบบโค้งปกติมาใช้แทนได้ เช่น สูตร z-score ที่ปรับให้รับกับ Binomial และตารางพื้นที่ใต้เส้นโค้งปกติ เป็นต้น ถ้าโจทย์ถาม ณ จุดใดจุดหนึ่ง เช่น ให้หาค่า probability ที่สุ่มแล้วได้นักศึกษา 15 คน ในกรณีนี้ก็ต้องคำนวณจาก Binomial p.d.f. (${}^nC_x p^x q^{n-x}$) แม้จะต้องเสียเวลาก็ตาม บางครั้งอาจมีการคำนวณจากค่า 14.5 - 15.5 อยู่บ้าง แต่ก็ไม่ใช่ที่นิยมมากนัก และทำให้ทราบว่าการหาค่า probability ของข้อมูลแบบต่อเนื่อง ณ จุดใดจุดหนึ่งไม่สามารถหาได้

2. เป็นเรื่องที่ทำให้ทราบว่ามีการยืมคุณสมบัติของ probability function อื่นๆมาใช้ได้ แต่ต้องทำความเข้าใจอย่างเคร่งครัด และต้องมีสูตรที่นักสถิติคิดไว้แล้วเพื่อให้เรานำมาใช้ มิฉะนั้นก็ไม่สามารถยืมกันได้ (เรื่องนี้จะสามารถเห็นได้ชัดขึ้นในเรื่องการทดสอบ χ^2 test of independence ซึ่งข้อมูลที่แจกแจงแบบ Chi-square (χ^2) ต้องเป็นข้อมูลแบบต่อเนื่อง หากเรานำเอาข้อมูลแบบไม่ต่อเนื่อง เช่น Multinomial p.d.f. มาขอยืมคุณสมบัติของ Chi-square ไปใช้ ก็ต้องทำความเข้าใจอย่างเคร่งครัดเหมือนกัน)

เรื่องแบบนี้ทำให้เกิดฐานคติ (assumptions) ต่างๆในวิชาสถิติ ซึ่งผู้ใช้งานจำเป็นต้องปฏิบัติตามอย่างเคร่งครัด เพื่อให้คำนวณได้ตัวเลขที่ถูกต้อง และจะได้นำไปใช้กำหนดแผนงานหรือวางนโยบายได้อย่างมีประสิทธิภาพ

5. สรุป

โดยสรุป ความเข้าใจในเรื่อง probability function ของการแจกแจงข้อมูลจึงน่าจะเป็นเรื่องสำคัญที่แม้จะเป็นนักศึกษาสายสังคมศาสตร์ก็ควรจะได้รับทราบ เพราะจะทำให้เข้าใจสถิติมากขึ้น จะได้วิเคราะห์โจทย์ได้ถูกและแก้โจทย์เป็น ทำให้ทราบว่าตารางสถิติแต่ละตารางมีที่มาอย่างไร และเป็นพื้นฐานในการเรียนรู้สถิติขั้นสูงต่อไป เช่น การทดสอบสมมติฐาน การวิเคราะห์ความแปรปรวน เป็นต้น และสามารถนำผลการทดสอบไปใช้ปฏิบัติในการวางแผนหรืออื่นๆ ต่อไป

บรรณานุกรม

ธีระพร วีระถาวร , ความน่าจะเป็นกับการประยุกต์ . ต.ค. 2539 น้าอักษรการพิมพ์ กรุงเทพฯ

วินัส พิษวนิชย์ และ ชินนะพงษ์ บำรุงทรัพย์, ทฤษฎีความน่าจะเป็น : พื้นฐานและการประยุกต์
(Basic Probability Theory and Application) สำนักพิมพ์ประกายพริ้ง เม.ย. 2547 กรุงเทพฯ

รัชกมล กบิลจิตต์, ความน่าจะเป็น : ทฤษฎีและการประยุกต์ (Probability Theory and Application)
โรงพิมพ์มหาวิทยาลัยศิลปากร นครปฐม 2545

Freund, John E., **Mathematical Statistics**. Prentice-Hall Inc. Englewood Cliffs, N.J. 1962.

Hays, William L., **Statistics for the Social Sciences**. 2nd edition, Holt, Rinehart and
Winston, Inc. 1973.

Hoyt, John P., **A Brief Introduction to Probability Theory**. International Textbook Co. 1967.

Hogg, Robert V. and Craig, Allen T., **Introduction to Mathematical Statistics**. Copyright,
the Macmillan Co. 1971.

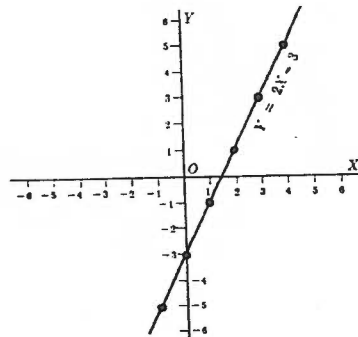
Spiegel, Murrey R. , **Theory and Problems of Statistics**. Schaum's Outline Series, McGraw-
Hill. 1961.

ภาคผนวก ก.

เพื่อเป็นการทบทวนความจำ จึงได้ยกกรณีตัวอย่างการเขียนกราฟจากสมการพีชคณิตธรรมดา (ไม่ใช่ probability function) มาให้ดู จำนวน 2 กราฟ ซึ่งไม่มีคุณสมบัติเรื่องค่าความน่าจะเป็น สำหรับ function ของการแจกแจงแบบโค้งปกติจึงสามารถเขียนเป็นกราฟได้เช่นกัน และยังมีคุณสมบัติเรื่องค่าความน่าจะเป็นเข้ามาเกี่ยวข้องด้วย

1. $Y = 2X - 3$ หาค่า X และ Y ได้ดังนี้

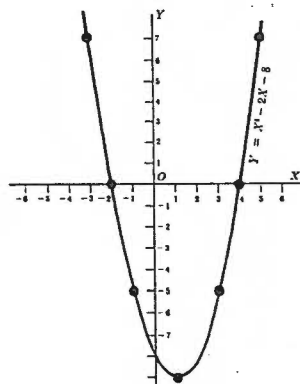
X โดยสมมุติ	-1	0	1	2
Y	-5	-3	-1	1



สมการนี้ ให้กราฟเป็นรูปเส้นตรง

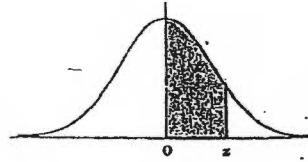
2. $Y = X^2 - 2X - 8$ หาค่า X และ Y ได้ดังนี้

X โดยสมมุติ	-2	-1	0	1
Y	0	-5	-8	-9



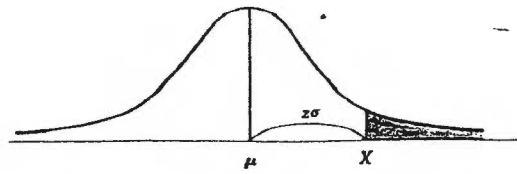
สมการนี้ให้กราฟเป็นรูปเส้นโค้ง parabola

**AREAS
under the
STANDARD
NORMAL CURVE
from 0 to z**

[illegible]

ภาคผนวก ก.

Ordinates of the Normal Curve



Example

$$Z = \frac{X - \mu}{\sigma}$$

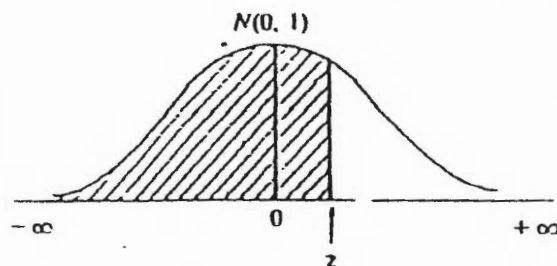
$$P[Z > 1] = .1587$$

$$P[Z > 1.96] = .0250$$

Normal Deviate z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

ภาคผนวก ๓1.

Cumulative distribution function for the standard normal distribution(SND)



For the SND, the table gives the cdf value at z , that is,

$$\int_{-\infty}^z N(0, 1) dz$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017
-3.4	0.00034	0.00033	0.00031	0.00030	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024
-3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00040	0.00039	0.00038	0.00036	0.00035
-3.2	0.00069	0.00066	0.00064	0.00062	0.00060	0.00058	0.00056	0.00054	0.00052	0.00050
-3.1	0.00097	0.00094	0.00090	0.00087	0.00085	0.00082	0.00079	0.00076	0.00074	0.00071
-3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.00114	0.00111	0.00107	0.00104	0.00100
-2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1057	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Table 31 is reprinted, with permission, from B. L. Grant and R. S. Leavenworth, *Statistical Quality Control*, McGraw-Hill Book Company, New York, 1972.

ภาคผนวก ง2.

Cumulative distribution function for the standard normal distribution(SND)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
+0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
+0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
+0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
+0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6870
+0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
+0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
+0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
+0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
+0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
+1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
+1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
+1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
+1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
+1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
+1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
+1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
+1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
+1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
+1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
+2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
+2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
+2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
+2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
+2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
+2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
+2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
+2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
+2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
+2.9	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
+3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
+3.1	0.99903	0.99906	0.99910	0.99913	0.99915	0.99918	0.99921	0.99924	0.99926	0.99929
+3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
+3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
+3.4	0.99966	0.99967	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
+3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983

ภาคผนวก จ1.

Table
BINOMIAL PROBABILITIES $(^N_r)p^r q^{N-r}$

N	r	p									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	.0500	.1000	.1500	.2000	.2500	.3000	.3500	.4000	.4500	.5000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000
	2	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625
	1	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500
	2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750
	3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1562
	2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
	3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
	5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2760	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094
	3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039

Table : (continued)

		p									
N	r	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0277	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
	1	.3293	.3835	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0054
	2	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	.0001	.0025	.0132	.0388	.0803	.1231	.1830	.2207	.2360	.2256
	6	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611
	8	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806
	9	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
	7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002

ภาคผนวก จ3.

Table . (continued)

N	r	p									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	.0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095
	8	.0000	.0000	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873
	10	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056
	3	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222
	4	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095
	8	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222
	10	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.3658	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.1348	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417
	5	.0006	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10	.0000	.0000	.0000	.0001	.0007	.0030	.0096	.0245	.0515	.0916
	11	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018

ภาคผนวก จ4.

Table (continued)

N	r	p									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
16	3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278
	5	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001
	2	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472
	6	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0001	.0014	.0084	.0279	.0644	.1143	.1606	.1883	.1855
	9	.0000	.0000	.0003	.0021	.0093	.0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001
	2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117
	5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0002	.0052	.0316	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855
	10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
	11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214
	12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117

ภาคผนวก จ5.

Table (continued)

N	r	p									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
15	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0062	.0018
	4	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
16	16	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	17	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	18	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	19	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
17	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000
	2	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002
	3	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011
	4	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046
	5	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148
18	16	.0003	.0089	.0454	.1091	.1686	.1916	.1712	.1244	.0746	.0370
	17	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739
	18	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201
	19	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602
	10	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762
	11	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602
	12	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201
	13	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739
	14	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148
19	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	1	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	2	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	4	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

This table is reproduced by permission from R. S. Burington and D. C. May, *Handbook of Probability and Statistics with Tables*. McGraw-Hill Book Company, (ed. 2), 1970.