บทที่ 4

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง

ในบทนี้จะกล่าวถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ซึ่งมีหลายชนิดในหัวข้อต่อไปนี้ จะให้ตัวอย่างเพื่อนำไปสู่นิยาม ให้นิยาม ตรวจสอบคุณสมบัติของการเป็นฟังก์ชันของความน่าจะเป็น หา $E(X)$ และหา $V(X)$

4.1 การแจกแจงแบบเสมอต้นเสมอปลาย (Uniform distribution)

พิจารณาผลจากการโยนลูกเต๋าแบบเที่ยงตรง 1 ลูก (นั่นคือโอกาสของการเกิดแต้มแต่ละแต้มจะเท่ากัน) จะได้แซมเปิลสเปซเป็น $S = \{1,2,3,4,5,6\}$

ให้ X เป็นเลขแต้มที่ลูกเต๋าหงาย ดังนั้น $X(s) \varepsilon \{1,2,3,4,5,6\}$

ถ้าให้ $f(x) = P(X = x)$ จะได้

$$f(1) = P(X = 1) = \frac{1}{6}$$

$$f(2) = P(X = 2) = \frac{1}{6}$$

$$\vdots$$

$$f(5) = P(X = 5) = \frac{1}{6}$$

$$f(6) = P(X = 6) = \frac{1}{6}$$

จะเห็นว่า เมื่อ $x \varepsilon X$ ค่า $f(x)$ เท่ากันหมดคือ $\frac{1}{6}$ ซึ่งเป็นค่าคงที่

พิจารณาผลการดึงไพ่จากไพ่สำรับหนึ่ง ซึ่งประกอบไปด้วยหน้าที่มีหมายเลข 3,4,5,...., 9,10 ต้องการหยิบไพ่ 1 ใบจากไพ่เหล่านี้ ให้ X เป็นตัวเลขที่ปรากฏบนไพ่ที่หยิบได้ ดังนั้นแซมเปิลสเปซเป็น $S = \{3,4,5,6,\ldots,10\}$

$$X = 3,4,5,\ldots, 10$$

ถ้าให้ $f(x) = P(X = x)$ จะได้

$$f(3) = \frac{1}{8}$$
$$f(4) = \frac{1}{8}$$
$$\vdots$$
$$f(10) = \frac{1}{8}$$

จะเห็นว่าค่า $f(x)$ เมื่อ $x \varepsilon X$ เท่ากันหมดคือ $\frac{1}{8}$ ซึ่งเป็นค่าคงที่

นิยาม 4.1 ถ้าตัวแปรสุ่ม X มี $R_X = \{x_1, x_2, \ldots, x_n\}$ และฟังก์ชันความน่าจะเป็น f ของ X กำหนดได้โดย $f(x) = \frac{1}{n}$, สำหรับทุก ๆ $x \varepsilon R_X$ เรียก X ว่ามีการแจกแจงแบบเสมอต้นเสมอปลาย (Uniform distribution) มี n เป็นพารามิเตอร์

ทฤษฎี 4.1.1 ให้ f เป็นฟังก์ชันค่าจริงที่มีเซตย่อย S ของ R เป็นโดเมน มีสมาชิก n ตัว ซึ่ง $f(s) = \frac{1}{n}$ สำหรับทุก ๆ $s \varepsilon S$ ดังนั้นจะมีตัวแปรสุ่ม X ที่มีการแจกแจงแบบเสมอต้นเสมอปลายที่มี S เป็นแซมเปิลสเปซ และมี f เป็นฟังก์ชันความน่าจะเป็นของ X

พิสูจน์ ให้ $X(s) = s$ สำหรับทุก ๆ s ใน S ดังนั้น X เป็นตัวแปรสุ่ม (บน S เมื่อเทียบกับเพาเวอร์เซตของ S) ให้ P เป็นความน่าจะเป็นบนเพาเวอร์เซตของ S โดยที่ $P(\{s\}) = \frac{1}{n}$ จะเห็นว่า

1. $f(x) \geq 0$ สำหรับทุก ๆ $x \varepsilon R_X$

2. $\sum_{x \varepsilon R_X} f(x) = 1$

3. $f(x) = P(X = x)$ สำหรับทุก ๆ $x \varepsilon R_X$

<u>ตัวอย่าง 4.1</u> จงหาการแจกแจงความน่าจะเป็นแบบเสมอต้นเสมอปลายของการเลือกคณะกรรมการ 3 คน จากสมาชิกสภาผู้แทน 6 คนแบบสุ่ม

<u>วิธีทำ</u> จำนวนคณะกรรมการ 3 คนที่เลือกจากสมาชิกสภาผู้แทน 6 คน

$$= \binom{6}{3} = \frac{6!}{3!\ 3!}$$

$$= 20$$

แต่ละคณะมีความน่าจะเป็นที่จะได้รับเลือกเท่ากัน

ฉะนั้นการแจกแจงความน่าจะเป็นของคณะกรรมการเหล่านี้เป็นแบบเสมอต้นเสมอปลาย ดังนั้น

$$f(x) = \frac{1}{20} \quad , \quad x = 1,2,\ldots,20$$

นั่นคือ ถ้าคณะที่ 15 ถูกเลือกจะมีความน่าจะเป็นเท่ากับ $p(X = 15) = \frac{1}{20}$

<u>ทฤษฎี 4.1.2</u> ให้ X เป็นตัวแปรสุ่ม ซึ่งมี $R_X = \{x_1, x_2, \ldots, x_n\}$ และ X มีการแจกแจงแบบเสมอต้นเสมอปลาย ให้ f เป็นฟังก์ชัน ความน่าจะเป็นของ X ดังนั้น $E(X) = \frac{1}{n} \sum_{i=1}^{n} x_i$ และ $V(X) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2$

<u>พิสูจน์</u>

$$E(X) = \sum_{x \in R_X} x\, f(x) = \sum_{i=1}^{n} x_i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$E(X^2) = \sum_{x \in R_X} x^2 f(x) = \frac{1}{n} \sum_{i=1}^{n} x_i^2$$

$$V(X) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2$$

4.2 การแจกแจงแบบทวินาม (Binomial distribution)

การทดลองแต่ละครั้งที่ผลของการทดลองมีโอกาสเป็นไปได้สองอย่างเท่านั้นคือ เหตุการณ์ที่ต้องการ (success) และเหตุการณ์ที่ไม่ต้องการ (failure) เช่น การโยนเหรียญ 1 อัน 1 ครั้ง ผลที่เกิดขึ้นเป็นไปได้สองอย่าง คือการเกิดหัวและเกิดก้อย ถ้าให้หัวเป็นเหตุการณ์ที่ต้องการ ก้อยเป็นเหตุการณ์ที่ไม่ต้องการ เรียกการกระทำในแต่ละครั้งที่เกิดเหตุการณ์สองอย่างแบบนี้ว่า "การทดลองเบอร์นูลี" (Bernoulli trails) ถ้าทำการทดลองซ้ำ ๆ กัน n ครั้ง เรียกว่า "การทดลองแบบทวินาม" (Binomial experiment)

<u>นิยาม 4.2</u> ให้ n เป็นจำนวนเต็มบวก X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มี $R_X = \{0,1,2,\ldots, n\}$ และให้ฟังก์ชัน $f(x) = \binom{n}{x}p^x(1-p)^{n-x}$ สำหรับ $x \in R_X$ โดยที่ $0 \leq p \leq 1$, $q = 1 - p$ เป็นฟังก์ชันความน่าจะเป็นของ X แล้วเรียกตัวแปรสุ่ม X ว่า มีการแจกแจงทวินาม (Binomial distribution) ที่มี n, p เป็นพารามิเตอร์

<u>หมายเหตุ</u> 1. $\binom{n}{x} p^x q^{n-x}$ เป็นเทอมในการกระจายทวินามของ $(p+q)^n$ โดยที่ $q = 1 - p$

2. ให้ $X_1, \ldots, X_n$ เป็นตัวแปรสุ่มที่มี $R_{X_i} = \{0,1\}$ และ $P(X_i = 0) = p$, $P(X_i = 1) = q$ สำหรับทุก ๆ $i = 1,\ldots, n$, $p, q \geq 0$ และ $p + q = 1$

ให้ $X = X_1 + \ldots + X_n$

ดังนั้น X จะเป็นตัวแปรสุ่มที่มี $R_X = \{0,1,\ldots, n\}$

ถ้า $X_1, \ldots, X_n$ เป็นอิสระต่อกัน ดังนั้น X จะมีการแจกแจงทวินามที่มี n และ p เป็นพารามิเตอร์

3. จะเห็นว่า f ในนิยาม 4.2 มีคุณสมบัติสอดคล้องกับสองข้อแรกของนิยาม 3.2.1 ของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มทั้งนี้เพราะ

3.1 $f(x) \geq 0$ สำหรับทุกค่า $x \in R_X$

3.2 $\sum_{x \in R_X} f(x) = 1$ ซึ่งมีการพิสูจน์ดังนี้

$$\sum_{x=0}^{n} f(x) = f(0) + f(1) + \dots + f(n)$$

$$= \binom{n}{0} p^0 q^{n-0} + \binom{n}{1} p q^{n-1} + \dots + \binom{n}{n} p^n q^{n-n}$$

$$= \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x}$$

$$= (p+q)^n = 1^n = 1$$

<u>ตัวอย่าง 4.2</u> ข้อสอบวิชาภาษาอังกฤษเป็นข้อสอบแบบปรนัยมี 10 ข้อ แต่ละข้อมีคำตอบให้เลือก 4 คำตอบ ใน 4 คำตอบนี้มีคำตอบที่ถูกเพียงข้อเดียว จงหาความน่าจะเป็นที่นักเรียนจะทำข้อสอบถูก 3 ข้อใน 10 ข้อนี้

<u>วิธีทำ</u> p = โอกาสที่จะตอบถูกใน 1 ข้อ $= \frac{1}{4}$

$$q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}$$

ดังนั้นความน่าจะเป็นที่นักเรียนจะทำข้อสอบถูก 3 ข้อใน 10 ข้อคือ

$$\binom{n}{x} p^x q^{n-x} = \binom{10}{3} \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^7$$

$$= \frac{10!}{7!\,3!} \cdot \frac{1}{4^3} \cdot \frac{3^7}{4^7}$$

$$= 0.25$$

ทฤษฎี 4.2 ถ้า x เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง และมีการแจกแจงแบบทวินาม ที่มี n และ p เป็นพารามิเตอร์ แล้ว E(X) = np และ V(X) = npq

พิสูจน์

$$E(X) = \sum_{i=0}^{n} x_i f(x_i) = \sum_{x=0}^{n} x\binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} x \frac{n!}{(n-x)!x!} p^x q^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\ (n-1)-(x-1)\ !} \cdot p^{x-1} q^{(n-1)-(x-1)}$$

$$= np \sum_{x=0}^{n-1} \frac{(n-1)!}{x!\left[(n-1)-x\right]!} \cdot p^x q^{(n-1)-x}$$

$$= np(p+q)^{n-1}$$

$$E(X) = np$$

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$E(X^2) = \sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} x^2\binom{n}{x} p^x q^{n-x}$$

พิจารณา $\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{n(n-1)!}{x(x-1)!(n-x)!} = \frac{n}{x}\binom{n-1}{x-1}$

$$E(X^2) = \sum_{x=1}^{n} x^2 \frac{n}{x}\binom{n-1}{x-1} p^x q^{n-x} = np \sum_{x=1}^{n} x\binom{n-1}{x-1} p^{x-1} q^{(n-1)-(x-1)}$$

ให้ m = n−1 , y = x−1 แทนในสมการ เมื่อ x มีค่าตั้งแต่ 1 ถึง n จะได้ว่า y มีค่าตั้งแต่ 0 ถึง n−1 คือ 0 ถึง m

$$E(X^2) = np \sum_{y=0}^{m} (y+1)\binom{m}{y} p^y q^{m-y}$$

$$E(X^2) = np\left[\sum_{y=0}^{m} y\binom{m}{y} p^y q^{m-y} + \sum_{y=0}^{m} \binom{m}{y} p^y q^{m-y}\right]$$

พจน์แรกอยู่ในรูป $E(Y)$ พจน์หลัง $= 1$

$$= np\left[mp + 1\right]$$

$$= np\left[(n-1)p + 1\right]$$

$$V(X) = E(X^2) - \left[E(X)\right]^2 = n^2p^2 - np^2 + np - n^2p^2$$

$$= np(1 - p) = npq$$

4.3 <u>การแจกแจงปัวส์ซอง</u> (Poisson distribution)

ในการทดลองใด ๆ หรือการสังเกตเหตุการณ์ที่เกิดขึ้นในปรากฏการณ์ธรรมชาติ เช่น จำนวนอนุภาคที่แพร่กระจายออกจากสารกัมมันตภาพรังสี หรือสังเกตจำนวนรถยนต์ชนกันบริเวณสี่แยกแห่งหนึ่ง จะมีการแจกแจงแบบหนึ่ง ซึ่งประยุกต์เข้ากับสถานการณ์ทำนองดังกล่าวนี้ได้อย่างดี การแจกแจงนั้นคือการแจกแจงปัวส์ซอง ดังปรากฏในนิยามต่อไปนี้

<u>นิยาม 4.3</u> ให้ λ เป็นจำนวนจริงบวก x เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มี

$f(x) = \dfrac{e^{\lambda}\lambda^x}{x!}$ $(x = 0,1,2,\ldots)$ เป็นฟังก์ชันความน่าจะเป็นของ x

จะเรียก x ว่ามี<u>การแจกแจงแบบปัวส์ซอง</u>ที่มี λ เป็นพารามิเตอร์

<u>หมายเหตุ</u> f ในนิยาม 4.3 สอดคล้องคุณสมบัติสองข้อแรกของนิยาม 3.2.1 ของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ซึ่งพิสูจน์ได้ดังนี้

1. $f(x) \geq 0$

   $f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2,\ldots$ และ $\lambda > 0$

   ดังนั้น $\lambda^x > 0$, $e^{-\lambda} > 0$ จะได้ $f(x) \geq 0$

2. พิสูจน์ว่า $\sum_{x=0}^{\infty} f(x) = 1$

$$\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x}}{x!}$$

$$= e^{-\lambda}\left[1 + \lambda + \frac{\lambda^{2}}{2!} + \frac{\lambda^{3}}{3!} + \cdots\right]$$

$$= e^{-\lambda} e^{\lambda} = 1$$

ตัวอย่าง 4.3 บริเวณสี่แยกแห่งหนึ่งใน อ. เมืองของจังหวัดหนึ่ง มักมีอุบัติเหตุเสมอ โดยเฉลี่ย 1 วันจะมีรถชนกัน 2 คัน ถ้าให้การแจกแจงแบบปัวส์ซองเหมาะสมกับสถานการณ์นี้ จงหาความน่าจะเป็นที่จะมีรถชนกัน 4 คันใน 1 วัน

วิธีทำ ให้ x เป็นจำนวนรถที่ชนกันใน 1 วัน $X = 0,1,2,\ldots$

โดยเฉลี่ย 1 วันจะมีรถชนกัน 2 คัน

ดังนั้น $\lambda = E(X) = 2$

หา $P(X = 4) = \frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{e^{-2}\,2^{4}}{4!} = 0.09022$

ฉะนั้นความน่าจะเป็นที่จะมีรถชนกัน 4 คันใน 1 วันเท่ากับ 0.0922

ทฤษฎี 4.3 ถ้า x เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีการแจกแจงแบบปัวส์ซองที่มี λ เป็นพารามิเตอร์แล้ว $E(X) = \lambda$ และ $V(X) = \lambda$

พิสูจน์

$$E(X) = \sum_{x=0}^{\infty} x\, f(x) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= 0 + \frac{e^{-\lambda}\lambda}{1!} + \frac{e^{-\lambda}\lambda^{2}}{2!} + \cdots$$

$$E(X) = \lambda e^{-\lambda}\left[1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots\right]$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$= \lambda$$

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 f(x)$$

$$= \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= 0 + \sum_{x=1}^{\infty} x^2 \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda^x}{(x-1)!}$$

ให้ $y = x - 1$, $x = y + 1$

$$E(X^2) = \sum_{y=0}^{\infty} \frac{(y+1)e^{-\lambda}\lambda^{y+1}}{y!}$$

$$= \lambda \sum_{y=0}^{\infty} (y+1) \frac{e^{-\lambda}\lambda^y}{y!}$$

$$= \lambda \sum_{y=0}^{\infty} \frac{ye^{-\lambda}\lambda^y}{y!} + \lambda \sum_{y=0}^{\infty} \frac{e^{-\lambda}\lambda^y}{y!}$$

$$= \lambda E(Y) + \lambda E(1)$$

$$E(X^2) = \lambda.\lambda + \lambda.1$$

$$= \lambda^2 + \lambda$$

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$= \lambda^2 + \lambda - \lambda^2 = \lambda$$

4.4 การแจกแจงแบบไฮเปอร์จีออเมตริก (Hypergeometric distribution)

ในการทดลองใด ๆ เช่นการดึงไพ่ 5 ใบจากไพ่ทั้งสำรับ 52 ใบ ต้องการได้ไพ่สีแดง 3 ใบ จาก 5 ใบที่ดึงออกมา ฉะนั้นที่เหลือ 2 ใบจะเป็นไพ่สีดำ ซึ่งสีแดง 3 ใบนี้จะดึงจากสีแดงทั้งหมด 26 ใบ และสีดำ 2 ใบจะดึงจากสีดำทั้งหมด 26 ใบ ถ้าให้ไพ่สีแดง 3 ใบที่ดึงได้เป็นเหตุการณ์ที่ต้องการ (success) ฉะนั้นไพ่สีดำ 2 ใบที่เหลือก็จะเป็นเหตุการณ์ที่ไม่ต้องการ (failure) ใน 5 ใบที่ดึงออกมา ลักษณะการทดลองเช่นนี้เรียกว่า การทดลองแบบไฮเปอร์จีออเมตริก

นิยาม 4.4 ให้ m, n และ k เป็นจำนวนเต็มบวกที่ k น้อยกว่าหรือเท่ากับ m และ n ให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มี $R_X = \{0, 1, \ldots, k\}$ และมี $f(x) = \dfrac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$ สำหรับ $x \in R_X$ เป็นฟังก์ชันความน่าจะเป็นของ X เรียกว่า X มีการแจกแจงแบบไฮเปอร์จีออเมตริก ที่มี m, n และ k เป็นพารามิเตอร์

หมายเหตุ f ในนิยาม 4.4 สอดคล้องกับคุณสมบัติสองข้อแรกของนิยาม 3.2.1 ของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ซึ่งพิสูจน์ได้ดังนี้

1. $f(x) = \dfrac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$

$\binom{m}{x}$, $\binom{n}{k-x}$ และ $\binom{m+n}{k} > 0$ สำหรับ $x = 0, 1, 2, \ldots, k$

ดังนั้น $f(x) > 0$

2. $\sum\limits_{x=0}^{k} f(x) = \sum\limits_{x=0}^{k} \dfrac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$

$$= \frac{1}{\binom{m+n}{k}} \sum_{x=0}^{k} \binom{m}{x}\binom{n}{k-x}$$

พิจารณา $(1+a)^m . (1+a)^n = (1+a)^{m+n}$

จากทฤษฎีของการกระจายทวินามเราได้

$$\left[\binom{m}{0} + \binom{m}{1}a + \ldots + \binom{m}{k}a^k + \ldots + \binom{m}{m}a^m\right]\left[\binom{n}{0}+\binom{n}{1}a + \ldots + \binom{n}{k}a^k + \ldots + \binom{n}{n}a^n\right]$$

$$= \binom{m+n}{0} + \binom{m+n}{1}a + \ldots + \binom{m+n}{k}a^k + \ldots + \binom{m+n}{m+n}a^{m+n}$$

เทียบสัมประสิทธิ์ของ a^k จะได้

$$\binom{m}{0}\binom{n}{k} + \binom{m}{1}\binom{n}{k-1} + \binom{m}{2}\binom{n}{k-2} + \ldots + \binom{m}{x}\binom{n}{k-x} + \ldots + \binom{m}{k}\binom{n}{0} = \binom{m+n}{k}$$

$$\sum_{x=0}^{k} \binom{m}{x}\binom{n}{k-x} = \binom{m+n}{k}$$

ดังนั้น $\dfrac{1}{\binom{m+n}{k}} \sum\limits_{x=0}^{k} \binom{m}{x}\binom{n}{k-x} = 1$ นั่นคือ $\sum\limits_{x=0}^{k} f(x) = 1$

<u>ตัวอย่าง 4.4</u> ดึงไพ่ 5 ใบออกจากไพ่สำรับหนึ่ง 52 ใบ จงหาความน่าจะเป็นที่จะได้ไพ่ 3 ใบที่เป็นโพแดง

<u>วิธีทำ</u> $k = 5$, $m+n = 52$, $m = 13$ = ไพ่โพแดง $n = 39$ = ไพ่อื่น ๆ ที่ไม่ใช่ไพ่แดง x เป็นจำนวนไพ่โพแดงที่หยิบได้

ต้องการหา $P(X = 3) = f(x = 3) = \dfrac{\binom{13}{3}\binom{39}{2}}{\binom{52}{5}}$

$$= \frac{\dfrac{13!}{10!\ 3!} \cdot \dfrac{39!}{37!\ 2!}}{\dfrac{52!}{5!\ 47!}}$$

$$= 0.0815$$

ทฤษฎี 4.4 ถ้า x เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีการแจกแจงแบบไฮเปอร์จี-ออเมตริกที่มี m,n และ k เป็นพารามิเตอร์ และ $k \le \min(m,n)$ แล้ว

$$E(X) = \frac{km}{m+n} \quad \text{และ} \quad V(X) = \frac{kmn(m+n-k)}{(m+n)^2(m+n-1)}$$

พิสูจน์

$$E(X) = \sum_{x=0}^{k} x\, f(x)$$

$$= \sum_{x=0}^{k} x \frac{\binom{m}{x}\binom{n}{k-x}}{\binom{m+n}{k}}$$

$$= \frac{1}{\binom{m+n}{k}} \left[1\binom{m}{1}\binom{n}{k-1} + 2\binom{m}{2}\binom{n}{k-2} + \ldots + k\binom{m}{k}\binom{n}{0} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \left[m\binom{m-1}{0}\binom{n}{k-1} + m\binom{m-1}{1}\binom{n}{k-2} + \ldots + m\binom{m-1}{k-1}\binom{n}{0} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \left[m\{\binom{m-1}{0}\binom{n}{k-1} + \binom{m-1}{1}\binom{n}{k-2} + \ldots + \binom{m-1}{k-1}\binom{n}{0}\} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \left[m \sum_{y=0}^{k-1} \binom{m-1}{y}\binom{n}{(k-1)-y} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \cdot m\binom{m+n-1}{k-1}$$

$$= \frac{k!(m+n-k)!}{(m+n)!} \; \frac{m(m+n-1)!}{(k-1)!(m+n-k)!}$$

$$E(X) = \frac{km}{m+n}$$

$$V(X) = E(X(X-1)) + E(X) - \left[E(X)\right]^2$$

$$E(X(X-1)) = \sum_{x=0}^{k} x(x-1)f(x)$$

$$= 0 + 0 + 2.1\, f(2) + 3.2.f(3) + \ldots + k(k-1)\,f(k)$$

$$= \frac{1}{\binom{m+n}{k}} \left[2.1\binom{m}{2}\binom{n}{k-2} + 3.2\binom{m}{3}\binom{n}{k-3} + \ldots + k(k-1)\binom{m}{k}\binom{n}{0} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \left[\frac{2.1.m(m-1)}{2.1}\binom{m-2}{0}\binom{n}{k-2} + \frac{3.2.m(m-1)}{3.2}\binom{m-2}{1}\binom{n}{k-3} \right.$$

$$\left. + \ldots + k(k-1)\,\frac{m(m-1)}{k(k-1)}\binom{m-2}{k-2}\binom{n}{0} \right]$$

$$= \frac{1}{\binom{m+n}{k}} \left[m(m-1)\left\{ \binom{m-2}{0}\binom{n}{k-2} + \binom{m-2}{1}\binom{n}{k-3} + \ldots + \right. \right.$$

$$\left. \left. \binom{m-2}{k-2}\binom{n}{0} \right\} \right]$$

$$= \frac{m(m-1)\binom{m+n-2}{k-2}}{\binom{m+n}{k}}$$

$$= \frac{m(m-1)k!\,(m+n-k)!}{(m+n)!} \cdot \frac{(m+n-2)!}{(k-2)!(m+n-k)!}$$

$$E(X(X-1)) = \frac{m(m-1)k(k-1)}{(m+n)(m+n-1)}$$

$$V(X) = \frac{m(m-1)k(k-1)}{(m+n)(m+n-1)} + \frac{km}{m+n} - \frac{k^2m^2}{(m+n)^2}$$

$$= \frac{kmn(m+n-k)}{(m+n)^2(m+n-1)}$$

4.5 การแจกแจงเรขาคณิต (Geometric Distribution)

ในการทดลองใด ๆ ที่การกระทำแต่ละครั้งเกิดเหตุการณ์สองอย่างคือ เหตุการณ์ที่ต้องการ (success) และเหตุการณ์ที่ไม่ต้องการ (failure) ทำการทดลองซ้ำ ๆ จนกว่าจะได้เหตุการณ์ที่ต้องการ (success) เป็นครั้งแรก จึงยุติการทดลอง

นิยาม 4.5 ให้ p เป็นจำนวนจริงที่ $0 < p < 1$ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มี $R_X = \{1,2,...\}$ มี $f(x) = p(1-p)^{x-1}$ สำหรับ $x \in R_X$ เป็นฟังก์ชันความน่าจะเป็นของ X เรียก X ว่า การแจกแจงเรขาคณิต มี p เป็นพารามิเตอร์

หมายเหตุ f ในนิยาม 4.5 สอดคล้องคุณสมบัติสองข้อแรกของนิยาม 3.2.1 ของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องซึ่งพิสูจน์ได้ดังนี้

1. $f(x) \geq 0$

$$f(x) = p(1-p)^{x-1}$$

$0 < p < 1$ ดังนั้น $1 - p > 0$

$$(1-p)^{x-1} \geq 0$$

ฉะนั้น $f(x) \geq 0$

2.
$$\begin{aligned}\sum_{x=1}^{\infty} f(x) &= f(1) + f(2) + \dots \\ &= p + p(1-p) + p(1-p)^2 + \dots \\ &= p + pq + pq^2 + \dots, \quad q = 1-p \\ &= p(1 + q + q^2 + \dots) \\ &= p\left(\frac{1}{1-q}\right) = p\cdot\frac{1}{p} = 1\end{aligned}$$

ตัวอย่าง 4.5 จงหาความน่าจะเป็นที่ว่า โยนเหรียญเที่ยงตรง 1 อันแล้วจะได้หัวเป็นครั้งแรกในการโยนครั้งที่ 4

วิธีทำ x เป็นจำนวนครั้งที่โยนเหรียญ

$p =$ ความน่าจะเป็นที่จะได้หัว เท่ากับ $\frac{1}{2}$

$$p(X = 4) = p(1-p)^{x-1}$$

$$= \frac{1}{2}(1-\frac{1}{2})^3$$

$$= 0.0625$$

ทฤษฎี 4.5 ถ้า x เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีการแจกแจงเรขาคณิตที่มี p เป็นพารามิเตอร์แล้ว $E(X) = \frac{1}{p}$ และ $V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$

พิสูจน์

$$E(X) = \sum_{x=1}^{\infty} x\, f(x)$$

$$= p + 2p(1-p) + 3p(1-p)^2 + 4p(1-p)^3 + \ldots$$

$$= p + 2pq + 3pq^2 + \ldots, \quad q = 1-p$$

$$= p(1 + 2q + 3q^2 + 4q^3 + \ldots)$$

$$= p \cdot \frac{1}{(1-q)^2}$$

$$= \frac{p}{p^2}$$

$$= \frac{1}{p}$$

$$E(X) = \frac{1}{p}$$

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$E(X^2) = \sum_{x=1}^{\infty} x^2 f(x)$$

$$= 1f(1) + 4f(2) + 9f(3) + \ldots$$

$$= p + 4pq + 9pq^2 + \ldots , \quad q = 1-p$$

$$= 1 - q + 4q - 4q^2 + 9q^2 - 9q^3 + \ldots$$

$$= 1 + 3q + 5q^2 + 7q^3 + \ldots$$

$$= 1 + (2+1)q + (3+2)q^2 + (4+3)q^3 + (5+4)q^4 + \ldots$$

$$= (1 + 2q + 3q^2 + 4q^3 + \ldots) + (q + 2q^2 + 3q^3 + \ldots)$$

$$= \frac{1}{(1-q)^2} + \frac{q}{(1-q)^2}$$

$$E(X^2) = \frac{1+q}{(1-q)^2}$$

$$V(X) = \frac{1+q}{(1-q)^2} - \frac{1}{p^2}$$

$$V(X) = \frac{q}{p^2}$$

4.6 <u>การแจกแจงทวินามนิเสธ</u> (Negative Binomial Distribution)

การทดลองที่การกระทำแต่ละครั้งเกิดเหตุการณ์สองอย่างคือเหตุการณ์ที่ต้องการ (success) และเหตุการณ์ที่ไม่ต้องการ (failure) และความน่าจะเป็นที่จะได้เหตุการณ์ที่ต้องการมีค่าเป็น p ทำการทดลองซ้ำ ๆ จนกว่าจะได้เหตุการณ์ที่ต้องการ k ครั้งจึงยุติการทดลอง

<u>นิยาม 4.6</u> ให้ k เป็นจำนวนเต็มบวก p เป็นจำนวนจริงที่ $0 < p < 1$ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มี $R_X = \{k, k+1, \ldots\}$ และมี $f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$ สำหรับ $x \in R_X$ เป็นฟังก์ชันความน่าจะเป็นของ X เรียกว่า X <u>มีการแจกแจงทวินามนิเสธ</u> ที่มี k,p เป็นพารามิเตอร์

<u>หมายเหตุ</u> f ในนิยาม 4.6.1 สอดคล้องคุณสมบัติสองข้อแรกของนิยาม 3.2.1 ของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ซึ่งพิสูจน์ได้ดังนี้

1. $f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$

$\binom{x-1}{k-1} > 0$, p^k และ $(1-p)^{x-k} > 0$ ดังนั้นจะได้ $f(x) > 0$

2. $$\sum_{x=k}^{\infty} f(x) = \sum_{x=k}^{\infty} \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

$$= \sum_{x=k} \binom{x-1}{k-1} p^k q^{x-k} \quad \text{เมื่อ} \quad q = 1-p$$

$$= p^k \Sigma \binom{x-1}{k-1} q^{x-k}$$

$$= p^k (1-q)^{-k} = 1$$

<u>นิยาม 4.6.2</u> ถ้า n เป็นจำนวนเต็มบวก ให้ $\binom{-n}{i} = \frac{(-n)(-n-1)(-n-2)\ldots}{i!(-n-i)!}$

สำหรับจำนวนเต็มบวก i ที่ $i \leq n$

วิธีพิสูจน์ว่า $\sum_{x=k}^{\infty} \binom{x-1}{k-1} q^{x-k} = (1-q)^{-k}$

จาก $(a+b)^{-n} = \sum_{i=0}^{\infty} \binom{-n}{i} a^{-n-i} b^i$

$$(1-q)^{-k} = \sum_{n=0}^{\infty} \binom{-k}{n} (-q)^n$$

$$= 1 + \binom{-k}{1}(-q) + \binom{-k}{2}(-q)^2 + \binom{-k}{3}(-q)^3 + \ldots$$

$$= 1 + kq + \frac{(-k)(-k-1)q^2}{2!} + \frac{(-k)(-k-1)(-k-2)(-q)^3}{3!} + \ldots$$

$$= 1 + kq + \frac{k(k+1)q^2}{2!} + \frac{k(k+1)(k+2)q^3}{3!} + \ldots$$

และ $\quad \Sigma\binom{x-1}{k-1} q^{x-k} = 1 + \binom{k}{k-1} q + \binom{k+1}{k-1} q^2 + \binom{k+2}{k-1} q^3 + \ldots$

$$= 1 + \frac{k!}{(k-1)!\ 1!} q + \frac{(k+1)!}{(k-1)!\ 2!} q^2 + \frac{(k+2)!}{(k-1)!\ 3!} q^3 + \ldots$$

ฉะนั้น $\quad \sum_{x=k}^{\infty} \binom{x-1}{k-1} q^{x-k} = (1-q)^{-k}$

<u>ตัวอย่าง 4.6</u> ทดลองโยนเหรียญ 1 อันไปเรื่อย ๆ จงหาความน่าจะเป็นที่เหรียญจะขึ้นหัวเป็นครั้งที่ 4 ในการโยนครั้งที่ 7

<u>วิธีทำ</u> ให้ x เป็นจำนวนครั้งที่ทดลอง

$x = 7$

$k = 4$

$p =$ ความน่าจะเป็นที่จะได้หัว $= \frac{1}{2}$

$$P(X = 7) = \binom{x-1}{k-1} p^k q^{x-k}$$

$$= \binom{7-1}{4-1} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^3$$

$$= \binom{6}{3} \left(\frac{1}{2}\right)^7$$

$$= \frac{5}{32}$$

<u>ทฤษฎี 4.6</u> ถ้า X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีการแจกแจงทวินามนิเสธที่มี k และ p เป็นพารามิเตอร์ โดยที่ $0 < p < 1$ แล้ว $E(X) = \frac{k}{p}$ และ $V(X) = \frac{kq}{p^2}$

<u>พิสูจน์</u> ให้ Z_1 คือจำนวนครั้งที่ทำการทดลอง จนเกิดเหตุการณ์ A เป็นครั้งแรก

Z_2 คือจำนวนครั้งที่ทำการทดลองหลังเกิดเหตุการณ์ A เป็นครั้งแรกถึงเกิด A เป็นครั้งที่สอง

.
.
.

Z_k คือจำนวนครั้งที่ทำการทดลองหลังเกิดเหตุการณ์ A เป็นครั้งที่ $k-1$ ถึงเกิด A ครั้งที่ k

จะได้ Z_i เป็นอิสระต่อกัน และแต่ละ Z_i เป็นตัวแปรสุ่มที่มีการแจกแจงแบบเรขาคณิต

ดังนั้น $X = Z_1 + Z_2 + \ldots + Z_k$

$$E(X) = E(Z_1) + E(Z_2) + \ldots + E(Z_k)$$

$$= \frac{1}{p} + \frac{1}{p} + \ldots + \frac{1}{p}$$

$$= \frac{k}{p}$$

$$V(X) = V(Z_1) + V(Z_2) + \ldots + V(Z_k)$$

$$= \frac{q}{p^2} + \frac{q}{p^2} + \ldots + \frac{q}{p^2}$$

$$= \frac{kq}{p^2}$$

4.7 การแจกแจงเอนกนาม (Multinomial distribution)

ในการทดลองใด ๆ ผลที่เกิดขึ้นในการกระทำแต่ละครั้งมีเพียงสองอย่าง คือเหตุการณ์ที่ต้องการ (success) และเหตุการณ์ที่ไม่ต้องการ (failure) ถ้าทำการทดลองซ้ำ ๆ กัน n ครั้ง เรียกว่าการทดลองแบบทวินาม เช่นการดึงไพ่ออกจากสำรับ 52 ใบ โดยใส่กลับคืน (with replacement) ถ้าเราสนใจผลที่เกิดขึ้นเพียงสองอย่างว่าจะเป็นไพ่สีแดงหรือสีดำ เป็นการทดลองแบบทวินาม ถ้าสนใจผลที่เกิดขึ้นมากกว่าสองอย่างเช่น จะเป็นไพ่โพแดง โพดำ ดอกจิก หรือข้าวหลามตัด เป็นการทดลองแบบเอนกนาม (Multinomial experiment)

สมมติ แซมเปิลสเปซถูกแบ่งออกเป็นเหตุการณ์ที่ไม่เกิดร่วมกัน k เหตุการณ์ $E_1, E_2, \ldots, E_k$ และแต่ละเหตุการณ์มีความน่าจะเป็น $p_1, p_2, \ldots, p_k$ ตามลำดับ $(p_1 + p_2 + \ldots + p_k = 1)$ ดังนั้น

<u>ทฤษฎี 4.7</u> ในการทดลอง n ครั้ง ความน่าจะเป็นที่ E_1 เกิดขึ้น x_1 ครั้ง E_2 เกิดขึ้น x_2 ครั้ง,...., E_k เกิดขึ้น x_k ครั้ง มีค่าเป็น

$$\frac{n!}{x_1!\, x_2! \ldots x_k!}\, p_1^{x_1} \cdot p_2^{x_2} \cdot \ldots\, p_k^{x_k}$$

โดยที่ $x_1 + x_2 + \ldots + x_k = n$ และเรียกการแจกแจงความน่าจะเป็นดังกล่าวว่าการแจกแจงแบบเอนกนาม (Multinomial distribution)

<u>ตัวอย่าง 4.7</u> ในการทอดลูกเต๋า 1 ลูก 8 ครั้ง จงหาความน่าจะเป็นที่จะได้หน้า 5 และหน้า 6 อย่างละ 2 ครั้ง และหน้าอื่น ๆ อย่างละครั้ง

<u>วิธีทำ</u> $n = 8$

เหตุการณ์ที่เกิดขึ้นคือเลขหน้า 1,2,..., 5,6

ความน่าจะเป็นของแต่ละเหตุการณ์เป็น $\frac{1}{6}$

ดังนั้นความน่าจะเป็นที่จะได้หน้า 5 และ 6 อย่างละ 2 ครั้ง และหน้าอื่น ๆ อย่างละครั้ง มีค่าเป็น

$$\binom{8}{2,2,1,1,1,1}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{8!}{2!\ 2!\ 1!\ 1!\ 1!\ 1!}\left(\frac{1}{6}\right)^8$$

$$= 0.006$$
