บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในการเปรียบเทียบวิธีการสร้างตัวเลขสุ่มเมื่อนำไปประมาณค่าพารามิเตอร์สำหรับการแจกแจงของตัวแปรสุ่มแบบต่อเนื่องและการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่อง ในที่นี้ได้เรียบเรียงข้อมูลที่เกี่ยวข้องกับงานวิจัยตามลำดับ ดังนี้

1. วิธีการสร้างตัวเลขสุ่ม 5 วิธี
2. ตัวแปรสุ่มแบบต่อเนื่องและตัวแปรสุ่มแบบไม่ต่อเนื่อง
3. ค่าคาดหวังและความแปรปรวนของตัวแปรสุ่มแบบต่อเนื่องและตัวแปรสุ่มแบบไม่ต่อเนื่อง
4. วิธีการสร้างค่าตัวแปรสุ่ม
5. ฟังก์ชันการแจกแจงของตัวแปรสุ่มแบบต่อเนื่องและฟังก์ชันการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องและการสร้างค่าตัวแปรสุ่มของแต่ละการแจกแจง
6. วิธีการประมาณค่า 2 วิธี คือ

6.1 การประมาณค่าแบบสภาวะสูงสุด (Maximum Likelihood Estimation)

6.2 วิธีกำลังสองน้อยที่สุด (Method of Least Square)

1. วิธีการสร้างตัวเลขสุ่ม 5 วิธี

ในที่นี้ศึกษาวิธีการสร้างตัวเลขสุ่มที่มีผู้คิดค้นขึ้น 5 วิธี คือ

1. วิธีตัดกลางกำลังสอง (Midsquare Method)
2. วิธีตัดกลางของผลคูณ (Midproduct Technique)
3. วิธีตัวคูณคงที่ (Constant Multiplier Technique)
4. วิธี Additive Congruential (Additive Congruential Method)
5. วิธี Linear Congruential (Linear Congruential Method)

1. วิธีตัดกลางกำลังสอง (Midsquare Method)

โดย Von-Neuman & Metropolis วิธีนี้อาศัยตัวเลขเริ่มต้น (Seed) มายกกำลังสองและตัวตัวเลขส่วนหัวและส่วนท้ายออกด้วยจำนวนหลักเท่า ๆกัน ตัวเลขที่อยู่ตรงกลางจะนำมาใช้เป็นตัวเลขสุ่มโดยทำให้ตัวเลขนั้นเป็นเลขทศนิยม ทำเช่นนี้ต่อเนื่องไปเรื่อย ๆจนได้ตัวเลขสุ่มตามจำนวนที่ต้องการ

ตัวอย่าง ให้ X0 เป็นตัวเลขเริ่มต้น เท่ากับ 3579

$X_0^2 = (3579)^2 = 12809241 = X_1 = 8092$

$R_1 = 0.8092$

$X_1^2 = (8092)^2 = 65480464 = X_2 = 4804$

$R_2 = 0.4804$

.

.

2. วิธีตัดกลางของผลคูณ (Midproduct Technique)

วิธีการนี้คล้ายกับวิธีตัดกลางกำลังสอง แต่ตัวเลขเริ่มต้นมี 2 ตัว คือ X_0' เป็นตัวเลขเริ่มต้นตัวแรก X_0 เป็นตัวเลขเริ่มต้นตัวที่สอง โดย X_0' , X_0 เป็นเลขขนาด d หลัก นำ X_0' และ X_0 คูณกันแล้วนำตัวเลขตรงกลางขนาด d หลักมาเป็น X_1 ทำเป็นเลขจุดทศนิยมก็จะกลายเป็น R_1 นำ X_0 คูณกับ X_1 แล้วนำตัวเลขตรงกลาง d หลักมาเป็น X_2 แล้วสร้าง R_2 ต่อไปทำเช่นนี้ไปเรื่อย ๆจนได้ตัวเลขสุ่มตามจำนวนที่ต้องการ

ตัวอย่าง $X_0' = 3579$ $X_0 = 9753$

$U_1 = X_0' X_0 = (3579)(9753) = 34905987 = X_1 = 9059$

$R_1 = 0.9059$

$U_2 = X_0 X_1 = (9753)(9059) = 88352427 = X_2 = 3524$

$R_2 = 0.3524$

.

.

.

3. วิธีตัวคูณคงที่ (Constant Multiplier Technique)

วิธีตัวคูณคงที่ ต่างจากวิธีตัวกลางของผลคูณ โดยวิธีการนี้มีค่าคงที่ 1 ตัวคือ k นำมาคูณกับ ตัวเลขเริ่มต้น (X_0) ซึ่ง X_0 มีขนาด d หลัก แล้วนำเอาตัวเลขตรงกลางขนาด d หลักมาเป็น X_1 แล้วนำมาสร้าง R_1 จากนั้นนำ k มาคูณกับ X_1 ซึ่งมีขนาด d หลัก แล้วนำเอาตัวเลขตรงกลางขนาด d หลักมาเป็น X_2 แล้วนำมาสร้าง R_2 และทำเช่นนี้ไปเรื่อย ๆ จนได้ตัวเลขสุ่มตามจำนวนที่ต้องการ

ตัวอย่าง $k = 1357$, $X_0 = 3579$

$V_1 = k X_0$

$V_1 = (1357)(3579) = 04856703 = X_1 = 8567$

$R_1 = 0.8567$

$V_2 = k X_1$

$V_2 = (1357)(8567) = 11625419 = X_2 = 6254$

$R_2 = 0.6254$

.

.

4. วิธี **Additive Congruential (Additive Congruential Method)**

วิธีนี้กำหนดตัวเลขจำนวนเต็ม $X_1, X_2, \ldots, X_n$ ตามลำดับ นำตัวเลขเหล่านี้มาสร้าง $X_{n+1}, X_{n+2}, \ldots$ จากนั้นนำตัวเลขที่สร้างมานี้มาสร้าง $R_1, R_2, \ldots$ ต่อไป โดยสูตรการสร้างซึ่งมีการทำ Modulo m ซึ่งจะทำให้ค่าที่ได้เป็นเลขเศษที่มีค่าน้อยกว่า m ซึ่งการคำนวณค่าตามสูตรต่อไปนี้

$$X_i = (X_{i-1} + X_{i-n}) \bmod m$$

$$R_{i-n} = \frac{X_i}{m}$$

ตัวอย่าง $X_1 = 65$, $X_2 = 24$, $X_3 = 39$, $X_4 = 88$, $X_5 = 75$,
$n = 5$, $m = 100$

$X_i = (X_{i-1} + X_{i-n}) \bmod m$

$X_6 = (X_5 + X_1) \bmod 100 = (75 + 65) \bmod 100$
$= 140 \bmod 100 = 40$

$R_{6-5} = R_1 = 40/100 = 0.40$

$X_7 = (X_6 + X_2) \bmod 100 = (40 + 24) \bmod 100$
$= 64 \bmod 100 = 64$

$R_{7-5} = R_2 = 0.64$

.

.

.

5. วิธี **Linear Congruential (Linear Congruential Method)**

วิธีนี้ได้พัฒนา โดย Lehmer ในปี ค.ศ.1951 และเป็นวิธีที่นิยมใช้กันในปัจจุบัน วิธีนี้มีสูตรการคำนวณหาดังนี้

$$X_{i+1} = (a X_i + c) \bmod m \qquad ; i = 0,1,2,...$$

$X_1, X_2, X_3, ..., X_i, X_{i+1}$ เป็นเลขจำนวนเต็มอยู่ระหว่าง 0 กับ m -1 เมื่อ

X_0 เป็นตัวเลขเริ่มต้น

a เป็นค่าคงที่ที่ใช้ในการคูณ

c เป็นค่าที่เพิ่มขึ้น

m เป็นตัว Modulus

ดังนั้นจะได้ตัวเลขสุ่ม

$$R_{i+1} = \frac{X_{i+1}}{m}$$

ตัวอย่าง a = 65539 c = 435 m = 10000 และ X_0 = 3579

$X_{i+1} = (a X_i + c) \bmod m$

X_1 = (65539*3579 + 435) mod 10000

X_1 = 234564516 mod 10000 = 4516

R_1 = 4516/10000 = 0.4516

X_2 = (65539*4516 + 435) mod 10000

X_2 = 4559

R_2 = 4559/10000 = 0.4559

.

.

.

โดยตัวเลขสุ่มที่ได้ต้องมีคุณสมบัติ 2 ข้อ ดังนี้

1. มีความสม่ำเสมอ (Uniformity) ทดสอบด้วยการทดสอบการแจกแจง (Distribution Test) ด้วยไคสแควร์ (Chi-Square Test) เปรียบเทียบการแจกแจงของตัวเลขที่ผลิตขึ้นกับการแจกแจงสม่ำเสมอ U(0,1)

2. มีความเป็นอิสระต่อกัน (Independence) ทดสอบด้วยการทดสอบรัน (Run Test)

2. ตัวแปรสุ่มแบบต่อเนื่องและตัวแปรสุ่มแบบไม่ต่อเนื่อง

ตัวแปรสุ่ม(Random Variable)

ตัวแปรสุ่มจำแนกได้ 2 ประเภท คือ

2.1 ตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable)

2.2 ตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Random Variable)

2.1 ตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable)

<u>นิยาม</u>

ตัวแปรสุ่มแบบต่อเนื่อง (Continuous Random Variable)

ให้ตัวแปรสุ่ม X เป็นตัวแปรสุ่มเราไม่อาจจะนับจำนวนค่าที่เป็นไปได้ของตัวแปรนั้น ค่าของตัวแปรสุ่มประเภทนี้จะกำหนดเป็นช่วง ๆ เช่น

$$X = \{x \mid 0 < x < 10\}$$

<u>นิยาม</u>

ตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีค่าอยู่ในช่วง (a,b) โดย $a < b$ f(x) เรียกว่าฟังก์ชันความน่าจะเป็น ถ้า f(x) มีคุณสมบัติดังนี้

1. $f(x) \geq 0$
2. $\int_a^b f(x)dx = 1$
3. ถ้า c และ d เป็นค่าในช่วง (a,b) และ $c < d$

$$P(c < X < d) = \int_c^d f(x)dx$$

<u>นิยาม</u>

ถ้าตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีค่าอยู่ในช่วง (a,b) F(x) จะเรียกว่าฟังก์ชันการแจกแจงสะสม ถ้า F(x) กำหนดไว้ดังนี้

$$F(x) = P(X \leq x) = \int_a^x f(x)dx$$

คุณสมบัติของฟังก์ชันการแจกแจงสะสม

1. $0 \leq F(x) \leq 1$
2. ถ้า $x_1 < x_2$ แล้ว $F(x_1) \leq F(x_2)$
3. $f(x) = \frac{dF(x)}{dx}$
4. $P(a < X \leq b)$
   $P(a \leq X \leq b)$

$P(a \le X < b) \quad = F(b) - F(a)$

$P(a < X < b)$

5. $P(X=a)=0$
6. ฟังก์ชันการแจกแจงสะสมเป็นฟังก์ชันที่ขึ้นทางขวาอย่างเดียว

2.2 ตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Random Variable)

<u>นิยาม</u>

ตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Random Variable)

ให้ตัวแปรสุ่ม X เป็นตัวแปรสุ่ม ถ้าค่าที่เป็นไปได้ของ X มีจำนวนจำกัด (Finite)

นับจำนวนได้เรียก X ว่า ตัวแปรสุ่มไม่ต่อเนื่อง (Discrete Random Variable) กล่าวคือ ตัวแปรสุ่ม $X = \{x/x=x_1, x_2, x_3, ...\}$

<u>นิยาม</u>

ตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีค่าเป็นจำนวนนับได้และมีค่าเป็น $x_1, x_2, x_3, ...$ ฟังก์ชัน p เรียกว่า ฟังก์ชันความน่าจะเป็นของ X

ถ้า $p(x_i)$ กำหนดไว้ดังนี้

$$p(x_i) = P(X=x_i) \; ; \; i = 1,2,3,...$$

โดยที่ $p(x_i)$ ต้องมีคุณสมบัติ ดังนี้

1. $0 \le p(x_i) \le 1$
2. $\sum_{x_i} p(x_i) = 1$

<u>นิยาม</u> ให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ฟังก์ชัน F(x) เรียกว่าฟังก์ชันการแจกแจงสะสมของ X ถ้า F(x) กำหนดไว้ดังนี้

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i) \quad \text{ซึ่ง } x_i \le x$$

คุณสมบัติของฟังก์ชันการแจกแจงสะสม

1. $0 \le F(x) \le 1$
2. ถ้า $x_1 < x_2$ แล้ว $F(x_1) \le F(x_2)$
3. ถ้า $x_1 < x_2 < \ldots < x_{i-1} < x_i$ แล้ว $P(X = x_i) = F(x_i) - F(x_{i-1})$
4. $P(c < X \le d) = F(d) - F(c)$
   $P(c \le X \le d) = F(d) - F(c) + P(X = c)$
   $P(c \le x < d) = F(d) - F(c) + P(X = c) - P(X = d)$
   $P(c < x < d) = F(d) - F(c) - P(X = d)$
5. ฟังก์ชันการแจกแจงสะสมเป็นฟังก์ชันที่ขึ้นทางขวาแบบขั้นบันได

3. ค่าคาดหวังและความแปรปรวน (Expected Value and Variance)

<u>นิยาม</u> ให้ X เป็นตัวแปรสุ่ม E(X) เรียกว่า ค่าคาดหวังของ X ถ้า

$$E(X) = \sum_{\forall x} x f(x) = \mu$$ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง

$$E(X) = \int_a^b x f(x)\,dx$$ X เป็นตัวแปรสุ่มแบบต่อเนื่อง

มีค่า $a < x < b$

EPU

<u>นิยาม</u> ให้ X เป็นตัวแปรสุ่ม E(X) = μ เป็นค่าคาดหวังของ X เราเรียก V(X)=σ^2 ว่า

ความแปรปรวนของ X ถ้า

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

$$E(X^2) = \sum_{\forall x} x^2 f(x)$$ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง

$$E(X^2) = \int_a^b x^2 f(x)dx$$ X เป็นตัวแปรสุ่มแบบต่อเนื่อง

มีค่า $a < x < b$

$$S(X) = \sqrt{V(X)}$$ S(X) เป็นส่วนเบี่ยงเบนมาตรฐานของ X

4. วิธีการสร้างค่าตัวแปรสุ่ม

ในการศึกษานี้ใช้ 2 วิธีของการสร้างค่าตัวแปรสุ่มตามการแจกแจงที่ต้องการทดสอบ คือ

1. วิธีการแปลงผกผัน (The Inverse Transform Method) ใช้สำหรับการสร้างค่าตัวแปรสุ่มของการแจกแจงแบบเอ๊กซ์โปเนนเชียล(Exponential Distribution) การแจกแจงแบบพัวซอง(Poisson Distribution) และการแจกแจงแบบทวินาม (Binomial Distribution)
2. วิธีการปฏิเสธ (The Rejection Method) ใช้สำหรับสร้างค่าตัวแปรสุ่มของการแจกแจงแบบปกติ(Normal Distribution)

1. วิธีการแปลงผกผัน (The Inverse Transform Method)

วิธีนี้สะดวกเนื่องจากถ้าเราทราบรูปแบบฟังก์ชันการแจกแจงสะสม(Cumulative Distribution Function) ของตัวแปรสุ่มนั้นแล้วนำฟังก์ชันการแจกแจงสะสมนี้มาเท่ากับตัวเลขสุ่มที่มีการแจกแจงแบบสม่ำเสมอ (0,1) นั้นคือ ฟังก์ชันการแจกแจงสะสมต้องมีการแจกแจงแบบสม่ำเสมอ (0,1) เช่นกัน

<u>**ทฤษฎี**</u> ฟังก์ชันการแจกแจงสะสม(Cumulative Distribution Function) ของตัวแปรสุ่ม X คือ F(x) จะมีการแจกแจงแบบสม่ำเสมอ (0,1)

เมื่อ F(X) มีการแจกแจงแบบสม่ำเสมอ และ R (Random Number) ก็มีการแจกแจงแบบสม่ำเสมอ จึงสามารถใช้วิธีการแปลงผกผันหาค่าตัวแปรสุ่ม ได้ดังนี้

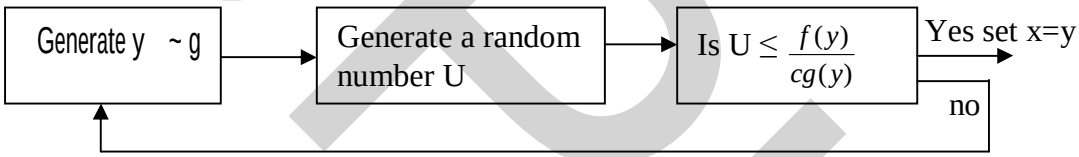
$$F(X) = R$$
$$X = F^{-1}(R)$$

2. วิธีการปฏิเสธ (The Rejection Method)

วิธีการในการสร้างค่าตัวแปรสุ่มที่มีฟังก์ชันความหนาแน่น g(x)เราสามารถใช้ในการสร้างจากฟังก์ชั่น f(x) โดยการสร้าง y จากฟังก์ชัน g หลังจากนั้นสร้างค่าความน่าจะเป็นด้วยสัดส่วน $\frac{f(y)}{g(y)}$ โดยให้ c เป็นค่าคงที่

$$\frac{f(y)}{g(y)} \le c \quad \text{สำหรับทุกค่าของ } y$$

ดังแผนภาพการสร้างค่าตัวแปรสุ่ม x ที่มีฟังก์ชันหนาแน่น f ด้วยวิธีการปฏิเสธ ดังนี้



มีขั้นตอนการทำงาน ดังนี้

ขั้นตอนที่ 1 สร้าง y ที่มีความหนาแน่น g

ขั้นตอนที่ 2 สร้างตัวเลขสุ่ม U

ขั้นตอนที่ 3 ถ้า $U \le \frac{f(y)}{cg(y)}$, ให้ x = y ถ้าไม่ใช่กลับไปที่ขั้นตอนที่ 1

5. ฟังก์ชันการแจกแจงของตัวแปรสุ่มแบบต่อเนื่องและการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องและการสร้างค่าตัวแปรสุ่มของแต่ละการแจกแจง

ในการศึกษานี้ศึกษาถึงการแจกแจงของตัวแปรสุ่มแบบต่อเนื่อง 2 ประเภท คือ

1. การแจกแจงแบบเอ๊กซ์โปเนนเชียล(Exponential Distribution)
2. การแจกแจงแบบปกติ(Normal Distribution)

ส่วนการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องศึกษาจาก 2 ประเภทของการแจกแจง คือ

1. การแจกแจงแบบพัวซอง(Poisson Distribution)
2. การแจกแจงแบบทวินาม(Binomial Distribution)

สำหรับการสร้างค่าตัวแปรสุ่มในการศึกษาครั้งนี้ใช้วิธีการแปลงผกผัน(The Inverse Transform Method) สำหรับการแจกแจงแบบเอ๊กซ์โปเนนเชียล การแจกแจงแบบพัวซอง และการแจกแจงแบบทวินาม ส่วนการแจกแจงแบบปกติใช้วิธีการปฏิเสธ(The Rejection Method) ในการสร้างค่าตัวแปรสุ่ม

การแจกแจงแบบเอ๊กซ์โปเนนเชียล(Exponential Distribution)

มีฟังก์ชันหนาแน่นความน่าจะเป็น(Probability density function) ดังนี้

$$f(x) = \begin{cases} \frac{1}{\beta} \cdot e^{-\frac{x}{\beta}} & \text{ถ้า } x \geq 0 \\ 0 & \text{อื่น ๆ} \end{cases}$$

$E(X) = \beta \quad Var(X) = \beta^2$

เมื่อใช้วิธีการแปลงผกผันในการสร้างค่าตัวแปรสุ่มหาค่าตัวแปรสุ่มดังนี้

$$F(x) = R$$
$$x = F^{-1}(R)$$

$$F(x) = P(X \leq x)$$
$$= \int_0^x f(x)dx = \int_0^x \frac{1}{\beta} e^{-\frac{x}{\beta}} dx$$
$$F(x) = 1 - e^{-\frac{x}{\beta}}$$
$$F(x) = R$$
$$1 - e^{-\frac{x}{\beta}} = R$$
$$e^{-\frac{x}{\beta}} = R$$
$$x = -\beta * \ln(R)$$

และมีขั้นตอนการทำงาน (Algorithm) ดังนี้

ขั้นตอนที่ 1 สร้างตัวเลขสุ่ม U

ขั้นตอนที่ 2 $x = -\beta * \ln(R)$

ขั้นตอนที่ 3 Return

การแจกแจงแบบปกติ (Normal Distribution)

ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ มีค่าเฉลี่ย (Mean) เท่ากับ μ และความแปรปรวน (Variance) เท่ากับ σ^2 โดยมีฟังก์ชันหนาแน่นความน่าจะเป็น(Probability density function) ดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}.e^{-(x-\mu)^2/2\sigma^2} \quad , -\infty < x < \infty$$

$$E(X) = \mu \qquad Var(X) = \sigma^2$$

สำหรับการแจกแจงแบบปกติในการศึกษานี้ใช้วิธีการปฏิเสธ(The Rejection Method)ในการสร้างค่าตัวแปรสุ่ม ดังนี้

การสร้างค่าตัวแปรสุ่มที่มีการแจกแจงแบบปกติ (Generating a Normal Random Variable)

การสร้างตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน Z~N(0,1) นั้น เมื่อค่าสัมบูรณ์ของ Z มีฟังก์ชันหนาแน่นความน่าจะเป็น (Probability density function)

$$f(x)= \frac{2}{\sqrt{2\pi}}.e^{-\frac{x^2}{2}} \qquad ; \quad 0 < x < \infty$$

เมื่อให้ g คือ ฟังก์ชันหนาแน่นน่าจะเป็นแบบเอ๊กซ์โปเนนเชียลมีค่าเฉลี่ยเท่ากับ 1 นั่นคือ

$$g(x) = e^{-x} \qquad ; \quad 0 < x < \infty$$

ดังนั้น

$$\frac{f(x)}{g(x)} = \sqrt{2/\pi e^{x-\frac{x^2}{2}}}$$

ค่าสูงสุดของ $\frac{f(y)}{g(y)}$ เกิดขึ้นเมื่อ x ทำให้ $x-\frac{x^2}{2}$ มีค่าสูงสุด เมื่อ x=1

ดังนั้น

$$c = \text{Max}\ \frac{f(y)}{g(y)} = \frac{f(1)}{g(1)} = \sqrt{2e/\pi}$$

เพราะว่า $\frac{f(y)}{cg(y)} = \exp\{x-\frac{x^2}{2}-\frac{1}{2}\}$

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

ดังนั้นสามารถสร้างค่าตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน Normal(0,1) ตามขั้นตอน ดังนี้

ขั้นตอนที่ 1 สร้าง y ที่มีการแจกแจงแบบเอ๊กซ์โปเนนเชียลมีค่าเฉลี่ยเท่ากับ 1

ขั้นตอนที่ 2 สร้างตัวเลขสุ่ม U

ขั้นตอนที่ 3 ถ้า $U \le \exp\{-{(y-1)^2}/{2}\}$,ให้ x=y ถ้าไม่ใช่กลับไปที่ขั้นตอนที่ 1

ดังนั้นเราสร้างตัวแปรสุ่ม x ที่มีฟังก์ชันหนาแน่นดังสมการ

$$f(x)= \frac{2}{\sqrt{2\pi}}.e^{-\frac{x^2}{2}} \quad ; \quad 0 < x < \infty$$

ที่มีการแจกแจงแบบปกติมาตรฐาน ดังนั้นจะได้ $z \sim N(0,1)$ โดยให้ $z = x$ หรือ $-x$ และในขั้นตอนที่ 3 ค่าของ x=y ถ้า $U \le \exp\{-{(y-1)^2}/{2}\}$ ซึ่งมีค่าเท่ากับ $\log U \ge {(y-1)^2}/{2}$ ดังนั้นสามารถเขียนขั้นตอนการทำงานใหม่ได้ดังนี้

ขั้นตอนที่ 1 สร้างตัวแปรสุ่มที่มีการแจกแจงแบบเอ๊กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1 y_1 ,y_2

ขั้นตอนที่ 2 ถ้า $y_2 \ge {(y-1)^2}/{2}$, ให้ $x = y_1$ ถ้าไม่ใช่กลับไปที่ขั้นตอนที่ 1

ถ้าขณะนี้ได้ $x= y_1$ และเรารู้ว่า $y_2 \ge {(y-1)^2}/{2}$ ซึ่งมากกว่าอยู่เท่าใดนั้น คำตอบก็คือให้ y_2 มีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1 และให้มีค่ามากกว่าค่าใด ๆ จำนวนที่ y_2 มากกว่า ${(y-1)^2}/{2}$ ก็จะมีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1 นั่นคือ เมื่อเราผ่านในขั้นตอนที่ 2 ไม่เพียงแต่เราจะได้ x (ค่าสัมบูรณ์ของการแจกแจงแบบปกติมาตรฐาน) แต่จาก

$y_2 - {(y-1)^2}/{2}$ เราสามารถสร้างค่าตัวแปรสุ่มแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1 ได้

ดังนั้นโดยสรุปขั้นตอนการสร้างค่าตัวแปรที่มีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1 และการสร้างค่าตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน ดังนี้

ขั้นตอนที่ 1 สร้าง y_1 ที่มีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1

ขั้นตอนที่ 2 สร้าง y_2 ที่มีการแจกแจงแบบเอ็กซ์โปเนนเชียลที่มีค่าเฉลี่ยเท่ากับ 1

ขั้นตอนที่ 3 ถ้า $y_2 - {(y-1)^2}/{2} > 0$, ให้ $y = y_2 - {(y-1)^2}/{2}$ และไปที่ขั้นตอนที่ 4 ถ้าไม่ใช่กลับไปที่ขั้นตอนที่ 1

ขั้นตอนที่ 4 สร้างตัวเลขสุ่ม U และกำหนด

$$Z = \frac{y_1 - \mu}{\sigma} \quad ; \quad U \le \frac{1}{2}$$

$$\frac{-y_1 - \mu}{\sigma} \quad ; \quad U > \frac{1}{2}$$

การแจกแจงแบบพัวซอง(Poisson Distribution)

มีฟังก์ชันหนาแน่นความน่าจะเป็น(Probability density function) ดังนี้

$$p_i = P\{X=i\} = e^{-\lambda}\frac{\lambda^i}{i!} \quad ; i = 0,1,...$$

$$E(X) = Var(X) = \lambda$$

เมื่อใช้วิธีการแปลงผกผัน (Inverse Transform) ในการสร้างค่าตัวแปรสุ่มแบบพัวซองที่มีค่าเฉลี่ยเท่ากับ λ โดยใช้สมการดังนี้

$$p_{i+1} = \frac{\lambda}{i+1}p_i \quad , \quad i \ge 0$$

โดยมีขั้นตอนการทำงาน (Algorithm) ดังนี้

ขั้นตอนที่ 1 สร้างตัวเลขสุ่ม U

ขั้นตอนที่ 2 ให้ $i = 0$, $p = e^{-\lambda}$, $F = p$

ขั้นตอนที่ 3 ถ้า $U < F$ ให้ $X = i$ และหยุด

ขั้นตอนที่ 4 $p = \frac{\lambda}{i+1}p_i$, $F = F + p$, $i = i + 1$

ขั้นตอนที่ 5 กลับไปที่ขั้นตอนที่ 3

จาก 5 ขั้นตอนในการสร้างค่าตัวแปรสุ่มที่มีการแจกแจงแบบพัวซองนั้น $p = p_i$ เป็นค่าความน่าจะเป็นที่ $X = i$ และ $F = F(i)$ เป็นความน่าจะเป็นที่ X น้อยกว่าหรือเท่ากับ i!

การแจกแจงแบบทวินาม(Binomial Distribution)

เมื่อต้องการสร้างค่าตัวแปรสุ่มที่มีการแจกแจงแบบทวินามที่มีพารามิเตอร์ n,p โดยที่มีฟังก์ชันหนาแน่นความน่าจะเป็น(Probability density function) ดังนี้

$$P\{X=i\} = \frac{n!}{i!(n-i)!} p^{i}(1-p)^{n-i} \ , \ i = 0,1,....,n$$

เมื่อใช้วิธีการแปลงผกผันโดยการกระทำซ้ำ(Loop) เมื่อ pr=P{X=i} เป็นค่าความน่าจะเป็นที่ X เท่ากับ i และ F = F(i) เป็นค่าความน่าจะเป็นที่ X น้อยกว่าหรือเท่ากับ i โดยการทำซ้ำ(Loop) ตามสมการและขั้นตอนการทำงาน (Algorithm) ดังนี้

$$P\{X=i+1\}=\frac{n-i}{i+1}.\frac{p}{1-p}.P\{X=i\}$$

ขั้นตอนที่ 1 สร้างตัวเลขสุ่ม U

ขั้นตอนที่ 2 $c = \frac{p}{(1-p)}$, i=0 , pr = $(1-p)^{n}$, F = pr

ขั้นตอนที่ 3 ถ้า U < F , ให้ X = i และหยุด

ขั้นตอนที่ 4 pr = [c(n-i)/(i+1)]pr , F = F + pr, i = i+1

ขั้นตอนที่ 5 กลับไปที่ขั้นตอนที่ 3

6. วิธีการประมาณค่า

ในการศึกษาวิจัยนี้ใช้วิธีการประมาณค่า 2 วิธี คือ การประมาณค่าแบบสภาวะสูงสุด (Maximum Likelihood Estimation) และการประมาณค่าด้วยวิธีกำลังสองน้อยที่สุด(Method of Least Square) ซึ่งในแต่ละวิธีมีทฤษฎีและวิธีการดังนี้

6.1 การประมาณค่าแบบสภาวะสูงสุด(Maximum Likelihood Estimation)

การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุด (เป็นวิธีที่ใช้กันแพร่หลายมากที่สุด มีแนวความคิดมานานตั้งแต่คริสต์ศตวรรษที่ 18 คาร์ล ฟรีดริค เกาส์ (Karl Friedrich Gauss,1777-1855) และแดเนียล เบอร์นูลลี่ (Daniel Bernoulli) ได้ใช้วิธีการนี้มาแล้วต่อมาในต้นศตวรรษที่ 20 โรนัลด์ ไอล์เมอร์ ฟิเชอร์ (Ronald Aylmer Fisher 1890-1962) ได้ทำการศึกษาคุณสมบัติของวิธีการนี้ ทำให้มีผู้ใช้กันกว้างขวางขึ้น และถือว่าวิธีการนี้เป็นผลงานของฟิเชอร์ โดยเขาได้นำเสนอผลงานเกี่ยวกับวิธีการนี้ในปี ค.ศ.1912 พร้อมทั้งมีการปรับปรุงแก้ไขส่วนที่เกี่ยวข้องให้เหมาะสมขึ้นอีกด้วย นักสถิติคนอื่น ๆก็มีส่วนทำให้วิธีการนี้เป็นที่นิยมแพร่หลายยิ่งขึ้นด้วย

นิยาม ให้ $x_1,\ldots,x_n$ เป็นตัวอย่างสุ่มจากประชากรที่มีฟังก์ชันความหนาแน่น $f(x;\theta)$, $\theta \in \Omega$ ฟังก์ชันภาวะน่าจะเป็น(Likelihood Function) ของตัวอย่างสุ่ม คือ ฟังก์ชันความหนาแน่นร่วม $L = L(\theta;x_1,x_2,\ldots,x_n)=L(\theta)$ ของตัวอย่างสุ่มนั้นที่ถือว่าเป็นฟังก์ชันของพารามิเตอร์ θ นั่นคือ

$$L(\theta;x_1,x_2,\ldots,x_n) = \prod_{i=1}^{n} f(x_i;\theta)$$

นิยาม ค่าของพารามิเตอร์ θ ในเทอมของค่าสังเกตของตัวอย่างสุ่ม $x_1,x_2,\ldots,x_n$ ที่ทำให้ฟังก์ชันภาวะน่าจะเป็นมีค่าสูงสุด เรียกว่า ตัวประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimator-MLE) ของ θ นั่นคือค่าของ θ คือ $\hat{\theta}=\hat{\theta}(x_1,\ldots,x_n)$ เป็น MLE ของ θ ก็ต่อเมื่อ $L(\hat{\theta})=L(\hat{\theta}(x_1,\ldots,x_n))$ มีค่าสูงสุด

วิธีการหาตัวประมาณแบบภาวะน่าจะเป็นสูงสุด

เป็นวิธีการหาค่าของพารามิเตอร์ θ ที่ทำให้ฟังก์ชัน $L(\theta)$ สูงสุด ในการนี้มีข้อควรสังเกตดังต่อไปนี้

1. เป้าหมายในการหา MLE ของ θ คือ การหาค่า θ เรียกว่า $\hat{\theta}=\hat{\theta}(x_1,\ldots,x_n)$ ที่ทำให้

   $L(\hat{\theta};x_1,\ldots,x_n) \geq L(\theta;x_1,x_2,\ldots,x_n)$

   เมื่อ $\theta \in \Omega$ และ $X_1=x_1,\ldots,x_n=x_n$

2. ถ้าฟังก์ชันภาวะน่าจะเป็น $L(\theta)$ เป็นฟังก์ชันที่หาอนุพันธ์ได้(Differentiable function) เมื่อเทียบกับ θ อาจใช้อนุพันธ์หา MLE ของ θ ได้ เมื่อเรนจ์ของ $f(x;\theta)$ ไม่ขึ้นอยู่กับ θ และ θ อยู่ในช่วงจำนวนจริงช่วงหนึ่งในกรณีดังกล่าว $\hat{\theta}$ คือ รากของสมการ

   $$\frac{\partial L}{\partial \theta} = 0$$

   เงื่อนไขพอเพียง (Sufficient Condition) ที่ $\hat{\theta}$ ทำให้ $L(\hat{\theta}) \geq L(\theta), \theta \in \Omega$ คือ

   $$\frac{\partial^2 L}{\partial \theta^2} < 0$$

   เมื่อ $\theta = \hat{\theta}$

3. การใช้อนุพันธ์หา MLE ในหลายกรณีใช้ $\ln L$ จะสะดวกกว่าที่จะใช้ L

   ควรสังเกตว่า

   $$\frac{\partial \ln L}{\partial \theta} = \frac{1}{L}.\frac{\partial L}{\partial \theta}$$

   และ $L > 0$ ดังนั้นเมื่อให้

   $$\frac{\partial \ln L}{\partial \theta} = 0$$

   เราจะได้ $\frac{\partial L}{\partial \theta} = 0$ ดังนั้น นอกจากนั้น เมื่อ

   $\frac{\partial^2 \ln L}{\partial \theta^2} < 0$ ก็จะทำให้ $\frac{\partial^2 L}{\partial \theta^2} < 0$ ด้วย

นิยาม สมการที่ใช้หา MLE คือ $\frac{\partial L}{\partial \theta} = 0$ หรือ $\frac{\partial \ln L}{\partial \theta} = 0$ เราเรียกว่า สมการภาวะน่าจะเป็น (Likelihood equation)

การหาตัวประมาณค่าพารามิเตอร์ของการแจกแจงแบบทวินาม (Binomial Distribution)

ให้ X_i ~ Bernulli เมื่อ i = 1,...,n

L(p ; X) = $p^{x_1}(1-p)^{1-x_2}.p^{x_2}(1-p)^{1-x_2}.....p^{x_n}(1-p)^{1-x_n}$

L(p ; X_i) = $p^{\sum x_i}(1-p)^{n-\sum x_i}$

ln L = $\sum X_i \ln p + (n-\sum X_i).\ln(1-p)$

$$\frac{\partial \ln L}{\partial p} = \frac{\sum X_i}{p} + \frac{n-\sum X_i}{1-p}(-1) = 0$$

$$\frac{\sum X_i}{p} - \frac{n-\sum X_i}{1-p} = 0$$

$$\frac{\sum X_i}{p} = \frac{n-\sum X_i}{1-p}$$

$$\sum X_i - p\sum X_i = np - p\sum X_i$$

$$\overline{X} = \frac{\sum X_i}{n} = p$$

ดังนั้น ตัวประมาณค่าพารามิเตอร์ p ของการแจกแจงแบบทวินาม คือ $\overline{X}$

การหาตัวประมาณค่าพารามิเตอร์ของการแจกแจงแบบพัวซอง (Poisson Distribution)

ให้ X_i ~ Poisson เมื่อ i = 1,...,n

โดย X_i มีฟังก์ชันการแจกแจง ดังนี้

f($x_i, \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; x = 0,1,2,...,$

$$L = f(x_1,\lambda).f(x_2,\lambda).f(x_3,\lambda)...f(x_n,\lambda)$$

$$= \frac{e^{-\lambda}.\lambda^{x_1}}{x_1!} \times \frac{e^{-\lambda}.\lambda^{x_2}}{x_2!} \times ... \times \frac{e^{-\lambda}.\lambda^{x_n}}{x_n!}$$

$$= \frac{e^{-n\lambda}.\lambda^{\sum x_i}}{\prod x_i!}$$

$$\ln L = -n\lambda + \sum x_i \ln\lambda + \ln\prod x_i!$$

$$\frac{\partial \ln L}{\partial \lambda} = -n + \frac{\sum x_i}{\lambda} = 0$$

$$\frac{\sum x_i}{\lambda} = n$$

$$\lambda = \frac{\sum x_i}{n}$$

ดังนั้นตัวประมาณค่าพารามิเตอร์ λ ของการแจกแจงแบบพัวซอง คือ $\overline{X}$

การหาตัวประมาณค่าพารามิเตอร์ของการแจกแจงแบบเอ๊กซ์โปเนนเชียล (Exponential Distribution)

ให้ $X_i \sim$ Exponential เมื่อ i = 1,...,n

โดย X_i มีฟังก์ชันการแจกแจง ดังนี้

$$f(x_i, \beta) = \frac{1}{\beta}.e^{-\frac{x}{\beta}}; x_i > 0,$$

$$L = f(x_1,\beta).f(x_2,\beta).f(x_3,\beta)...f(x_n,\beta)$$

$$= \frac{1}{\beta}.e^{-\frac{x_1}{\beta}} \times \frac{1}{\beta}.e^{-\frac{x_2}{\beta}} \times \frac{1}{\beta}.e^{-\frac{x_3}{\beta}} \times ... \times \frac{1}{\beta}.e^{-\frac{x_n}{\beta}}$$

$$= \frac{1}{\beta^n}.e^{-\frac{1}{\beta}\sum x_i}$$

$$\ln L = -n\ln\beta - \frac{1}{\beta}\sum x_i$$

$$\frac{\partial \ln L}{\partial \beta} = -\frac{n}{\beta} + \frac{\sum x_i}{\beta^2} = 0$$

$$\frac{\sum x_i}{\beta^2} = \frac{n}{\beta}$$

$$\frac{\sum x_i}{n} = \frac{\beta^2}{\beta}$$

$$\overline{X} = \frac{\sum x_i}{n} = \beta$$

ดังนั้นตัวประมาณค่าพารามิเตอร์ β ของการแจกแจงแบบเอ๊กซ์โปเนนเชียล คือ $\overline{X}$

การหาตัวประมาณค่าพารามิเตอร์ของการแจกแจงแบบปกติ (Normal Distribution)

ให้ $X_i \sim$ Normal(μ, σ^2) เมื่อ i = 1,...,n

โดย X_i มีฟังก์ชันการแจกแจง ดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}}.e^{-(x-\mu)^2/2\sigma^2} \quad , -\infty < x < \infty$$

DPU

$$L(\mu,\sigma^2;x_1,x_2,...,x_n)=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}.e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\ln L(\mu,\sigma^2)=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

ดังนั้น $\frac{\partial \ln L(\mu,\sigma^2)}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)=0$

และ $\frac{\partial \ln L(\mu,\sigma^2)}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2=0$

จะได้ $\mu=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{X}$ และ $\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

นั่นคือ MLE ของ μ คือ $\hat{\mu}=\frac{1}{n}\sum_{i=1}^{n}x_i=\overline{X}$

และ MLE ของ σ^2 คือ $\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2=S^2$

ในที่นี้ $\hat{\sigma}^2$ ขึ้นอยู่กับ μ จะต้องประมาณ μ ก่อนจึงจะประมาณ σ^2 ได้ ในกรณีเช่นนี้มักเรียก μ ว่า พารามิเตอร์รบกวน (Nuisance parameter)

6.2 การประมาณค่าพารามิเตอร์ด้วยวิธีกำลังสองน้อยที่สุด (Method of Least Square)

แนวความคิดและวิธีการหาตัวประมาณแบบกำลังสองน้อยที่สุด วิธีนี้เป็นวิธีการสำคัญ มีรากฐานมาจากทฤษฎีการประมาณเชิงเส้น(Theory of Linear estimation) โดยไม่จำเป็นต้องทราบรูปการแจกแจงความน่าจะเป็นแต่อาศัยผลต่างระหว่างค่าสังเกตและค่าคาดหมายเป็นสำคัญ วิธีการนี้คิดขึ้นโดยคาร์ล ฟรีดริค เกาส์ (Karl Friedrich Gauss, 1777-1855) และอังเดร อังเดรเยวิช มาร์โคฟ (Andrei Andreevich Markov, 1856-1922)

นิยาม ให้ $x_1,...,x_n$ เป็นตัวอย่างสุ่มจากประชากรหนึ่งที่มีค่าคาดหมาย E(x) ค่าของพารามิเตอร์ในเทอมของค่าสังเกตที่ทำให้ผลบวกของกำลังที่สองของผลต่างระหว่างค่าสังเกตกับค่าคาดหมายต่ำสุดจะเรียกว่า ตัวประมาณแบบกำลังสองน้อยที่สุด (Least Squares estimator) ของพารามิเตอร์ที่เกี่ยวข้อง

นั่นคือ ตัวประมาณแบบกำลังสองน้อยที่สุดของพารามิเตอร์ θ (อาจเป็นเวคเตอร์)ที่ปรากฏอยู่ในค่าคาดหมาย E(x) ได้แก่ ค่าของ θ ในเทอมของค่าสังเกตที่ทำให้

$Q=\sum_{i=1}^{n}(x_i-E(x_i))^2$ ต่ำที่สุด

มักหาตัวประมาณดังกล่าว โดยการใช้อนุพันธ์กล่าวคือ ใช้

$$\frac{\partial Q}{\partial \theta} = 0$$

และเรียกสมการนี้ว่า สมการปกติ (Normal equation(s))

คุณสมบัติของตัวประมาณแบบกำลังสองน้อยที่สุด

ตัวประมาณแบบกำลังสองน้อยที่สุดมีคุณสมบัติที่สำคัญหลายประการ โดยเฉพาะในตัวแบบเชิงเส้น ดังต่อไปนี้

1. เมื่อ $\underline{Y} = X\underline{\theta} + \underline{\in}$ ที่ $E(\underline{\in}) = \underline{0}$ ตัวประมาณแบบกำลังสองน้อยที่สุดของ $\underline{\theta}$ เป็นตัวประมาณไม่เอนเอียง

   พิสูจน์ จาก $\hat{\underline{\theta}} = (X'X)^{-1}X'\underline{Y} = (X'X)^{-1}X'(X\underline{\theta} + \underline{\in})$

   $$= \underline{\theta} + (X'X)^{-1}X'\underline{\in}$$

   ดังนั้น $E(\hat{\theta}) = \underline{\theta}$ เนื่องจาก $E(\underline{\in}) = \underline{0}$

2. เมื่อ $\underline{Y} = X\underline{\theta} + \underline{\in}, E(\underline{\in}) = \underline{0}, V(\underline{\in}) = \sigma^2 I$ เมตริกซ์ของความแปรปรวนและความแปรปรวนร่วมของตัวประมาณแบบกำลังสองน้อยที่สุดของ $\underline{\theta}$ คือ $(X'X)^{-1}.\sigma^2$

   พิสูจน์ $V(\hat{\underline{\theta}}) = E[(\hat{\underline{\theta}} - \underline{\theta})(\hat{\underline{\theta}} - \underline{\theta})']$

   $$= E[((X'X)^{-1}X'\underline{\in})(\underline{\in}'X(X'X)^{-1})]$$

   $$= (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1}$$

3. ในตัวแบบเชิงเส้น $\underline{Y} = X\underline{\theta} + \underline{\in}$ ที่ $E(\underline{\in}) = \underline{0}, E(\underline{\in}\underline{\in}') = \sigma^2 I$

   ตัวประมาณแบบกำลังสองน้อยที่สุด $\hat{\underline{\theta}}$ ของ $\underline{\theta}$ เป็นตัวประมาณที่มีความแปรปรวนต่ำสุดในบรรดาตัวประมาณที่ไม่เอนเอียงของ $\underline{\theta}$

   คุณสมบัตินี้ คือ ทฤษฎีของเกาส์และมาร์โคฟ

ทฤษฎีบทที่ 1 (เกาส์และมาร์โคฟ) เมื่อ

$$\underline{Y} = X\underline{\theta} + \in \;,\; E(\in) = 0 \;,\; E(\underline{\in}\,\underline{\in}') = \sigma^2 I \;,\; E(\underline{Y}) = X\underline{\theta}$$

ตัวประมาณแบบกำลังสองน้อยที่สุด $\hat{\underline{\theta}} = (X'X)^{-1}X'\underline{Y}$ เป็นตัวประมาณที่เป็นฟังก์ชันเชิงเส้นไม่เอนเอียงและดีที่สุดของ $\underline{\theta}$

พิสูจน์ ให้ $\hat{\underline{\theta}}^* = [(X'X)^{-1}X' + A]\underline{Y}$ เป็นฟังก์ชันเชิงเส้นของ $\underline{Y}$ ที่ใช้ประมาณ $\underline{\theta}$ จะหาเมตริกซ์ A ที่ทำให้ $\hat{\underline{\theta}}^*$ เป็นตัวประมาณไม่เอนเอียงของ $\underline{\theta}$ ที่มีส่วนประกอบแต่ละตัวมีความแปรปรวนต่ำที่สุดในบรรดาตัวประมาณไม่เอนเอียงของส่วนประกอบของ $\underline{\theta}$

จาก $E(\hat{\underline{\theta}}^*) = \underline{\theta}$ จะได้

$$\underline{\theta} = [(X'X)'X' + A]E(\underline{Y}) = [(X'X)^{-1}X' + A]X\underline{\theta} = \underline{\theta} + AX\underline{\theta}$$

ดังนั้น $AX\underline{\theta} = \underline{\theta}$ หรือ $AX = 0$

$$V(\underline{\hat{\theta}^*}) = E[(\underline{\hat{\theta}^*} - \underline{\theta})(\underline{\hat{\theta}^*} - \underline{\theta})']$$
$$= E[((X'X)'X' + A)\underline{Y} - \underline{\theta}][((X'X)^{-1}X' + A)Y - \underline{\theta}]'$$

$$=E[(X'X)^{-1}X'\underline{\in}\underline{\in}'X(X'X)^{-1} + A\underline{\in}\underline{\in}'A' + (X'X)^{-1}X'\underline{\in}\underline{\in}'A' + A\underline{\in}\underline{\in}X(X'X)^{-1}]$$
$$=[(X'X)^{-1} + AA']\sigma^2$$

ให้ $AA' = B = (b_{ij})$ เมตริกซ์บวกกำหนดแน่(Positive definite matrix) จึงมี $b_{ij} \geq 0$ ทุกค่า i=1,...,p

ต้องการหา B ที่มีสมาชิกบนเส้นทแยงมุมหลักมีค่าน้อยที่สุด เพื่อทำให้ $V(\hat{\theta}^*_i)$ มีสมาชิกบนเส้นทแยงมุมหลักต่ำที่สุด เมื่อ $b_{ii} = \sum_{j=1}^{n} a^2_{ij} = 0$ ทุกค่า i =1,...,p

แต่เมื่อ $A = (a_{ij})$ จะได้ $b_{ii} = \sum_{j=1}^{n} a^2_{ij} = 0$

ดังนั้น $a_{ij} = 0$ ทุกค่า i และ j , i=1,2,...,p ; j = 1,2,...,n

นั่นคือ A = 0 และ $\underline{\hat{\theta}^*} = (X'X)^{-1}X'Y = \underline{\hat{\theta}}$ ซึ่งเป็นตัวประมาณแบบกำลังสองต่ำสุดของ $\underline{\theta}$

4. ตัวประมาณไม่เอนเอียงและมีความแปรปรวนต่ำสุดของฟังก์ชันเชิงเส้นใด ๆของพารามิเตอร์ $\theta_1,...,\theta_p$ คือ ฟังก์ชันเดียวกันของตัวประมาณแบบกำลังสองน้อยที่สุดของ $\theta_1,...,\theta_p$

ทฤษฎีบท ในตัวแบบเชิงเส้น $\underline{Y} = X\underline{\theta} + \underline{\in}$ ที่ $E(\underline{\in}) = \underline{0}$

$E(\underline{\in}\underline{\in}') = \sigma^2 I$, $E(\underline{Y}) = X\underline{\theta}$ ให้ t เป็นเวคเตอร์ขนาด p x 1 ที่มีส่วนประกอบคงที่ตัวประมาณเชิงเส้นไม่เอนเอียง และมีความแปรปรวนต่ำที่สุดของ $\underline{t}'\underline{\hat{\theta}}$ เมื่อ $\underline{\hat{\theta}} = (X'X)^{-1}X'\underline{Y}$

พิสูจน์ ให้ $T\underline{Y}$ เป็นฟังก์ชันเชิงเส้นของ $\underline{Y}$ ที่ใช้ประมาณ $\underline{t}'\underline{\hat{\theta}}$ จะหา T ที่ทำให้ $T\underline{Y}$ เป็นตัวประมาณไม่เอนเอียงและมีความแปรปรวนต่ำที่สุดของ $\underline{t}'\underline{\hat{\theta}}$

จาก $E(T\underline{Y}) = \underline{t}'\underline{\hat{\theta}}$ จะได้ $TX\underline{\theta} = \underline{t}'\underline{\hat{\theta}}$ ดังนั้น $TX = \underline{t}'$

$$V(T\underline{Y}) = E[(T\underline{Y} - \underline{t}'\underline{\theta})(T\underline{Y} - \underline{t}'\underline{\theta})']$$
$$T\underline{Y} - \underline{t}'\underline{\theta} = T(X\underline{\theta} + \underline{\in}) - \underline{t}'\underline{\theta} = TX\underline{\theta} + T\in - \underline{t}'\underline{\theta} = T\in$$

จึงได้

$$V(T\underline{Y}) = E[T\underline{\in}\underline{\in}'T'] = TT'\sigma^2$$
$$TT' = [TX(X'X)^{-1}X'][TX(X'X)^{-1}X']' + [T - TX(X'X)^{-1}X'][T - TX(X'X)^{-1}X']'$$
$$= [\underline{t}'(X'X)^{-1}X'][\underline{t}'(X'X)^{-1}X']' + [T - \underline{t}'(X'X)^{-1}X'][T - \underline{t}'(X'X)^{-1}X']'$$

ขวามือของสมการนี้เป็นผลบวกของเมตริกซ์ในรูป A A' ซึ่งแต่ละเทอมมีสมาชิกบนเส้นทแยงมุมหลักไม่ติดลบ เฉพาะเทอมหลังเท่านั้นที่เป็นฟังก์ชันของ T
ดังนั้น $T\,T'$ มีสมาชิกบนเส้นทแยงมุมหลักน้อยที่สุด เมื่อสมาชิกบนเส้นทแยงมุมหลักของเทอมหลังเป็น 0 หมด ซึ่งจะเป็นเช่นนั้นได้เมื่อ T = $\underline{t}'(X'X)^{-1}X'$
ดังนั้น ตัวประมาณไม่เอนเอียงและมีความแปรปรวนต่ำสุดของ $\underline{t}'\underline{\theta}$ ได้แก่ T$\underline{Y}$ = $\underline{t}'(X'X)^{-1}X'\underline{Y} = \underline{t}'\hat{\underline{\theta}}$

การประมาณความแปรปรวนในตัวแบบเชิงเส้น

ทฤษฎี ในตัวแบบเชิงเส้น $\underline{Y}$ = X$\underline{\theta} + \underline{\in}$ ที่ E($\underline{\in}$)=$\underline{0}$ E($\underline{\in}\underline{\in}'$) $= \sigma^2 I$ ตัวประมาณที่ไม่เอนเอียงของ σ^2 คือ $\frac{1}{n-p}(\underline{Y} - X\hat{\underline{\theta}})(\underline{Y} - X\hat{\underline{\theta}})$ เมื่อ $\underline{Y}$ เป็นเวคเตอร์ขนาด n x 1 และ x เป็นเมตริกซ์ขนาด p x 1 , $\underline{\in}$ เป็นเวคเตอร์ขนาด n x 1 และ $\hat{\underline{\theta}} = (X'X)^{-1}X'\underline{Y}$

พิสูจน์

$\underline{Y} - X\hat{\underline{\theta}}$ =(X$\underline{\theta} + \underline{\in}$)

I-X($X'X)^{-1}X'$ เป็นเมตริกซ์สมมาตรและไอเดมโพเทนต์ (Symmetric idempotent matrix)

ดังนั้น

$$(\underline{Y} - X\hat{\underline{\theta}})'(\underline{Y} - X\hat{\underline{\theta}}) = \underline{\in}'[\text{I-X}(X'X)^{-1}X']\underline{\in}$$

$$E[(\underline{Y} - X\hat{\underline{\theta}})'(\underline{Y} - X\hat{\underline{\theta}})] = \sigma^2 tr[I - X(X'X)^{-1}X']$$

$$= \sigma^2[trI - tr(X(X'X)^{-1}X')]$$

$$= \sigma^2[n - tr.Ip] = (\text{n-p})\sigma^2$$

ดังนั้น $\frac{1}{n-p}(\underline{Y} - X\underline{\theta})'(\underline{Y} - X\underline{\theta})$ เป็นตัวประมาณไม่เอนเอียงของ σ^2

ตัวประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) และ

วิธีกำลังสองน้อยที่สุด (Method of Least Square) ของแต่ละการแจกแจงที่ใช้สำหรับการศึกษา ดังนี้

การแจกแจง	พารามิเตอร์	ตัวประมาณค่า
การแจกแจงแบบพัวซอง	λ	$\hat{\lambda} = \overline{X}$

(Poisson Distribution)		
การแจกแจงแบบทวินาม (Binomial Distribution)	p	$\hat{p} = \overline{X}$
การแจกแจงแบบเอ๊กซ์โปเนนเชียล (Exponential Distribution)	β	$\hat{\beta} = \overline{X}$
การแจกแจงแบบปกติ (Normal Distribution)	μ และ σ^2	$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = S^2$

DPU