ทฤษฎีตัดสินใจทางสถิติ กับปัญหาการถดถอย

Statistical Decision Theory approach to Regression Problems

ปัญหาการถดถอยเกี่ยวข้องกับวิธีที่จะสืบหาความสัมพันธ์ระหว่างตัวแปรที่สังเกตได้ ตัวแปรตัวหนึ่งเป็นแบบสุ่ม ซึ่งปกติเราเรียกว่า ตัวแปรตาม และสมมติว่าตัวแปรเชิงสุ่มนี้เป็นฟังก์ชันของตัวแปรหลายตัวที่ไม่เป็นแบบสุ่มซึ่งสังเกตได้ ตัวแปรเหล่านี้เรียกว่าตัวแปรอิสระ ความสัมพันธ์ที่ง่ายสุดระหว่างตัวแปรตาม y และตัวแปรอิสระ $x_1, x_2, .., x_m$ นั้นเป็นแบบเชิงเส้น ซึ่งกำหนดไว้ด้วย

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \epsilon$$

การสืบสวนหรือการศึกษาเกี่ยวกับความสัมพันธ์นั้นจะเป็นการประมาณค่า และทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ $\beta_0, \beta_1, \beta_2, \ldots, \beta_m$ เราจะเริ่มการวิเคราะห์สำหรับกรณีของปัญหาถดถอยเชิงเส้นแบบง่าย ซึ่งมีตัวแปรอิสระเพียงตัวเดียว ต่อไปเราจึงทำการวิเคราะห์ปัญหาการถดถอยที่มีตัวแปรอิสระหลายตัว กรณีหลังเรียกว่าปัญหาการถดถอยเชิงเส้นแบบซ้อน

1. <u>ตัวแบบถดถอยเชิงเส้นแบบง่าย</u> (Simple Linear Regression Model)

ตัวแบบถดถอยเชิงเส้นแบบง่ายจะมีรูปแบบเป็น

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \ , \quad i = 1, 2, .., n$$

ในเมื่อ y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม, x_i เป็นค่าสังเกตที่ i ของตัวแปรอิสระ, ϵ_i เป็นค่าที่สังเกตไม่ได้ของตัวแปรคลาดเคลื่อน, และ β_0, β_1 เป็นพารามิเตอร์ถดถอย

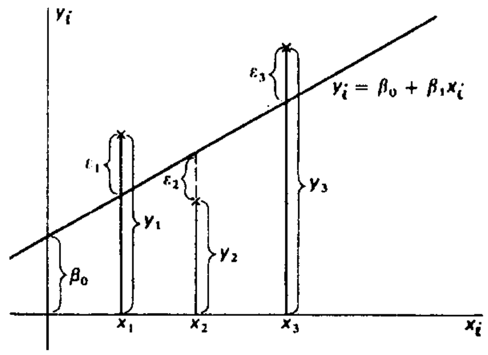
เราจะเห็นได้ว่าตัวแบบข้างบนนี้เป็นความสัมพันธ์แบบฟังก์ชันที่เป็นเชิงเส้นในพารามิเตอร์ β_0 และ β_1 เท่านั้น ตัวแปร x_i อาจจะปรากฏในแบบฟอร์มที่ไม่เป็นเชิงเส้นก็ได้ ดังเช่นความสัมพันธ์ที่กำหนดไว้เป็น

$$y_i = \beta_0 + \beta_1 e^{x_i} + \epsilon_i$$

ความสัมพันธ์นี้ก็เป็นเชิงเส้นอยู่ถึงแม้ว่าจะเป็นฟังก์ชันที่ไม่เป็นเชิงเส้นของ x_i เพราะว่าพารามิเตอร์ถดถอยยังปรากฏในรูปแบบที่เป็นเชิงเส้น ด้วยเหตุผลเดียวกัน ตัวแปรอิสระ

อาจจะปรากฏอยู่ในรูปแบบที่ไม่เป็นเชิงเส้นใด ๆ ได้ แต่ถ้าพารามิเตอร์ถดถอยมีความสัมพันธ์เป็นเชิงเส้น แล้วเราก็เรียกว่า ตัวแบบถดถอยเชิงเส้น อันนี้ก็เป็นจริงสำหรับตัวแบบถดถอยเชิงเส้นแบบซ้อน ที่มีตัวแปรอิสระหลายตัว

ปัญหาตัดสินใจหลายปัญหาสามารถแทนได้ด้วยตัวแบบถดถอยเชิงเส้น สมมติเราสนใจที่จะประมาณความสัมพันธ์แบบฟังก์ชันที่อธิบายถึงความผันแปรในค่าใช้จ่ายบริโภคด้วยความผันแปรของรายได้ (ตัวแปรอิสระ) ของ 100 ครัวเรือนที่เป็นแบบเดียวกันในชุมนุมชนเล็ก ๆ แห่งหนึ่ง กำหนด y_i เป็นค่าสังเกตตัวที่ i ของค่าใช้จ่ายบริโภค และ x_i เป็นค่าสังเกตตัวที่ i ของรายได้สำหรับแต่ละครัวเรือน เราไม่หวังว่าทั้ง 100 ครัวเรือนจะมีแบบแผนบริโภคเหมือนกันสำหรับรายได้ที่กำหนดไว้ เพราะบางคนใช้มาก บางคนใช้น้อย ดังนั้นเราจึงหวังว่ากลุ่ม (clusters) ของจุดที่แทนค่าใช้จ่ายบริโภคนั้นจะรอบค่าที่แทนรายได้ที่กำหนดไว้ ซึ่งหมายความว่าความสัมพันธ์ระหว่าง y_i และ x_i ไม่พอดีทีเดียว และเราก็จะถือว่าเป็นเพราะตัวแปรเชิงสุ่ม ϵ_i เพราะฉะนั้นความสัมพันธ์เชิงเส้นอาจจะกำหนดไว้ดังที่กล่าวมาแล้ว เรามีเหตุผลอื่นที่รวมตัวแปรเคลื่อนคลาด ϵ_i ไว้ในตัวแบบถดถอยเชิงเส้นด้วย คือมันเกือบจะเป็นไปไม่ได้ และทางปฏิบัติก็ทำไม่ได้ที่จะรวมตัวแปรที่เกี่ยวข้องทั้งหมดเพื่ออธิบายความผันแปรในตัวแปรตาม y_i ดังนั้นจึงเป็นไปได้ที่จะไม่รวมตัวแปรบางตัวหรืออาจจะยากที่จะหาข้อมูลของตัวแปรบางตัว ยิ่งกว่านั้นก็มีสิ่งที่ไม่สามารถคาดได้ (แบบสุ่ม) ในการตอบสนองของมนุษย์เพื่อที่จะอธิบายความผันแปรใน y_i ดังนั้นจึงแทนสิ่งต่าง ๆ นั้นด้วยตัวคลาดเคลื่อนเชิงสุ่ม ϵ_i และแปลความหมายของ ϵ_i ในแง่ของตัวแปรที่อธิบายความผันแปรใน y_i ซึ่งอธิบายไม่ได้ด้วย x_i สถานการณ์เช่นนี้แสดงได้เป็น



ตัวแบบถดถอยเชิงเส้นแบบง่ายมีข้อกำหนดดังนี้ (1) ค่า x_i $(i=1,2,..,n)$ จะเป็นจำนวนจำกัดในตัวอย่างที่ซ้ำ ๆ กัน (2) ตัวแปรคลาดเคลื่อนเชิงสุ่ม ϵ_i มีการแจกแจงปกติที่มีค่าเฉลี่ย 0 และความแปรปรวน σ^2 ร่วมกัน และ ϵ_i ต่าง ๆ นั้นเป็นอิสระกัน (3) ตัวแปรทุกตัวจะวัดได้โดยไม่มีความเคลื่อนคลาด

ข้อกำหนด 1 กล่าวว่า x_i ไม่เป็นตัวแปรเชิงสุ่ม อันนี้ไม่เคร่งครัดนัก และกล่าวได้ว่า x_i อาจจะมีการแจกแจงน่าจะเป็นบางอย่างที่ไม่เกี่ยวกับ β_0, β_1, และ σ^2 นั่นคือเป็นอิสระกับการแจกแจงของ ϵ_i

ข้อกำหนด 2 หมายความว่าตัวแปร $y_1, y_2, .., y_n$ เป็นอิสระ และแจกแจงปกติที่มีค่าเฉลี่ย $\beta_0 + \beta_1 x_i$ และความแปรปรวน σ^2 ร่วมกัน ฟังก์ชันน่าจะเป็น (Likelihood Function) สำหรับ β_0, β_1, และ σ^2 แสดงได้ด้วย

$$p(y/\beta_0, \beta_1, \sigma^2, x) = (1/2\pi\sigma^2)^{n/2} e^{-\frac{1}{2\sigma^2}\sum(y_i-\beta_0-\beta_1 x_i)^2}$$
$$\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum(y_i-\beta_0-\beta_1 x_i)^2}$$

ในเมื่อ y และ x เป็นเวคเตอร์แนวตั้ง (Column Vector) ของ y_i และ x_i ต่าง ๆ

ค่าประมาณของ β_0 และ β_1 หาได้จากผลรวมกำลังสอง L มีค่าน้อยที่สุด

$$L = \sum(y_i - \beta_0 - \beta_1 x_i)^2$$

วิธีนี้เรียกว่า หลักของกำลังสองน้อยสุด (Principle of Least Squares) แล้วเราจะได้ค่าประมาณเป็น b_0 และ b_1 ดังนี้

$$b_0 = \bar{y} - b_1\bar{x}$$
$$b_1 = \sum(y_i - \bar{y})(x_i - \bar{x}) / \sum(x_i - \bar{x})^2$$

สำหรับค่าประมาณของ σ^2 จะได้เป็น s^2

$$s^2 = \sum e_i^2/(n-2)$$

เราจะเห็นว่าค่าประมาณของ β_0 และ β_1 เป็นแบบเชิงเส้นกับ $y_1, y_2, .., y_n$ และไม่เอียงเฉ กับทั้งยังมีความแปรวนน้อยระหว่างกลุ่มของตัวประมาณที่ไม่เอียงเฉเชิงเส้น ผลอันนี้เป็นทฤษฎีเกาส์–มาร์คอฟ (Gauss-Markov Theorem) สำหรับ s^2 นั้นก็ไม่เอียงเฉด้วย

ค่าประมาณของ β_0 และ β_1 นั้นสามารถพัฒนาได้จากการทำให้ฟังก์ชั่นน่าจะเป็น $p(y/\beta_1, \beta_2, \sigma^2, x)$ มากที่สุด นั่นคือถ้าเราหาอนุพันธ์ของฟังก์ชั่นน่าจะเป็นโดยเทียบกับ β_0, β_1, และ σ^2 และให้เป็น 0 ซึ่งเราจะได้ค่าประมาณ $\hat{b}_0$, $\hat{b}_1$, และ $\hat{s}^2$ ดังนี้

$$\hat{b}_0 = \bar{y} - \hat{b}_1\bar{x}$$

$$\hat{b}_1 = \sum(y_i - \bar{y})(x_i - \bar{x}) / \sum(x_i - \bar{x})^2$$

$$\hat{s}^2 = (1/n)\sum(y_i - \hat{b}_0 - \hat{b}_1 x_i)^2$$

ค่าประมาณเหล่านี้เรียกว่า ค่าประมาณน่าจะเป็นสูงสุด (Maximum Likelihood Estimate) $\hat{b}_0$ และ $\hat{b}_1$ ที่ได้นี้จะเป็นเช่นเดียวกับค่าประมาณแบบกำลังสองน้อยสุดที่หามาแล้วนั่นเอง แต่ $\hat{s}^2$ ต่างจาก s^2 ซึ่งหมายความว่า $\hat{s}^2$ จะเป็นค่าประมาณที่เอียงเฉของ σ^2

2. <u>การอนุมานเชิงสถิติเกี่ยวกับ β_0, β_1, และ σ^2</u>

เนื่องจาก b_0 เป็นฟังก์ชั่นเชิงเส้นใน $y_1, y_2, .., y_n$ เราจึงได้ว่า b_0 เป็นตัวแปรเชิงสุ่มแบบปกติที่มีค่าเฉลี่ย และความแปรปรวน ดังนี้

$$E(b_0) = \beta_0$$

$$V(b_0) = \sigma^2 \frac{\sum x_i^2/n}{\sum(x_i-\bar{x})^2}$$

ถ้าทราบค่าของ σ^2 แล้วอัตราส่วน Z

$$Z = (b_0 - \beta_0)/\sqrt{V(b_0)}$$

จะมีการแจกแจงปกติมาตรฐาน N(0, 1)

โดยทั่วไปไม่ทราบ σ^2 จึงประมาณด้วย s^2 ดังนั้นค่าประมาณของ $V(b_0)$ จะเป็น $s^2_{b_0}$

$$s^2_{b_0} = s^2 \frac{\sum x_i^2/n}{\sum(x_i-\bar{x})^2}$$

อัตราส่วนของ T

$$T = (b_0 - \beta_0)/s_{b_0}$$

จะมีการแจกแจง t ที่มีองศาความเป็นอิสระ $n-2$

เพราะฉะนั้นถ้าเราทดสอบสมมติฐานเกี่ยวกับค่าของ β_0 เราก็ใช้อัตราส่วน หรือตัวสถิติทดสอบ T หรือ Z แล้วแต่ว่าทราบ σ^2 หรือไม่ทราบ

ในทำนองเดียวกันเราสามารถแสดงได้ว่า

$$E(b_1) = \beta_1$$
$$V(b_1) = \sigma^2/\sum(X_i - \bar{X})^2$$

และ b_1 เป็นตัวแปรเชิงสุ่มแบบปกติด้วย ดังนั้นในการทดสอบสมมติฐานเกี่ยวกับค่าของ β_1 หรือทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระว่ามีจริงหรือไม่นั้น เราใช้ตัวสถิติทดสอบ Z หรือ T

$$Z = (b_1 - \beta_1^*)/(\sigma/\sqrt{\sum(X_i - \bar{X})^2}) \quad \text{ถ้าทราบ } \sigma^2$$
$$T = (b_1 - \beta_1^*)/(s/\sqrt{\sum(X_i - \bar{X})^2}) \quad \text{ถ้าไม่ทราบ } \sigma^2$$

ในการทดสอบสมมติฐานเกี่ยวกับ σ^2 นั้นเราใช้ตัวสถิติทดสอบ U

$$U = (n-2)s^2/\sigma_0^2$$

ซึ่งมีการแจกแจงไคสแควร์ ที่มีองศาความเป็นอิสระ $n - 2$

แบบทดสอบที่กล่าวมานั้นใช้สำหรับทดสอบพารามิเตอร์ตัวใดตัวหนึ่งเท่านั้น ถ้าเราสนใจทดสอบสมมติฐาน

$$H_0: \beta_0 = \beta_0^* ; \quad \beta_1 = \beta_1^*$$

ซึ่งเป็นสมมติฐานเกี่ยวกับทั้ง β_0 และ β_1 สมมติฐานแบบนี้เรียกว่า สมมติฐานเชิงเส้น (Linear Hypothesis) ตัวสถิติทดสอบของสมมติฐานนี้จะเป็น

$$F = (Q/2)/(s^2/\sigma^2)$$

ซึ่งมีการแจกแจงเอฟ ที่มีองศาความเป็นอิสระ 2 และ $n - 2$ ในเมื่อ Q กำหนดไว้ดังนี้

$$Q = (1/\sigma^2)\{n(b_0 - \beta_0)^2 + 2n\bar{x}(b_0 - \beta_0)(b_1 - \beta_1) + \sum X_i^2(b_1 - \beta_1)^2\}$$

ตัวสถิติทดสอบ T นี้เราสามารถใช้พิจารณาช่วงเชื่อมั่น $100(1-\alpha)\%$ สำหรับพารามิเตอร์ถดถอย β_0 และ β_1 ได้เป็น

$$\beta_0 = b_0 \pm t_{\alpha/2}^{(n-2)} s_{b_0}$$
$$\beta_1 = b_1 \pm t_{\alpha/2}^{(n-2)} s_{b_1}$$

3. <u>การแยกความผันแปรทั้งหมด</u> (Dedomposition of Total Variation)

เราได้กล่าวมาแล้วว่า จุดประสงค์ที่สำคัญสำหรับการประมาณความสัมพันธ์แบบฟังก์ชันระหว่าง y และ x เพื่ออธิบายความผันแปรใน y อันเนื่องมาจากการเปลี่ยนแปลงใน x อย่างไรก็ตามความผันแปรทั้งหมดใน y อาจจะไม่เนื่องมาจากความผันแปรใน x เพราะว่าอาจจะมีอิทธิพลบางอย่างที่ไม่ใช่ x ซึ่งสามารถอธิบายความผันแปรใน y ในตัวแบบถดถอยแบบง่ายนั้นจะสมมติว่าอิทธิพลอื่น ๆ ทั้งหมดเป็นตัวคลาดเคลื่อนเชิงสุ่ม ปัญหาที่เราสนใจก็คือ มีความผันแปรใน y เท่าไรที่เนื่องมาจากความผันแปรใน x และมีเท่าไรที่เนื่องมาจากความผันแปรของตัวคลาดเคลื่อนเชิงสุ่ม ปัญหานี้พิจารณาได้จากการแยกความผันแปรทั้งหมดของตัวแปรตามดังนี้

$$\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 + \sum e_i^2 \;;\; \bar{y} = \bar{\hat{y}}$$

เราจะเห็นได้ว่า ความผันแปรในค่า y รอบค่าเฉลี่ย $\bar{y}$ ซึ่งเรียกว่า ผลรวมกำลังสองทั้งหมด (Total Sum of Squares, SST) นั้นเนื่องมาจาก ความผันแปรในค่า $\hat{y}$ รอบค่าเฉลี่ย $\bar{y}$ ซึ่งเป็นอิทธิพลของค่า x และเรียกว่า ผลรวมกำลังสองเนื่องจากการถดถอย (Sum of Squares due to Regression, SSR) กับความผันแปรที่อธิบายไม่ได้ในค่า y รอบเส้นกำลังสองน้อยสุด ซึ่งเรียกว่า ตัวเศษหรือผลรวมกำลังสองของความคลาดเคลื่อน (Residual or Error Sum of Squares, SSE) ดังนั้นเราจึงได้

$$SST = SSR + SSE$$

จากการแยกความผันแปรของค่า y เราจะได้มาตราวัดที่เรียกว่า สัมประสิทธิ์ของการตัดสินใจ (Coefficient of Determination, R^2) ซึ่งเป็นสัดส่วนของความผันแปรใน y ที่เนื่องมาจากความผันแปรใน x นั่นคือ

$$\begin{aligned} R^2 &= 1 - SSE/SST \\ &= 1 - \sum e_i^2 / \sum(y_i - \bar{y})^2 \end{aligned}$$

จากสมการนี้เราจะเห็นว่าค่าสูงสุดของ R^2 เป็น 1 และจะเกิดขึ้นก็ต่อเมื่อ $\sum e_i^2 = 0$ ในกรณีเช่นนี้จะมีการปรับที่ดีที่สุด (Exact fit) ของเส้นกำลังสองน้อยสุด ในทำนองเดียวกัน

ค่าต่ำสุดของ R^2 จะเป็น 0 ซึ่งจะเกิดขึ้นเมื่อ $\sum e_i^2 = \sum(y_i - \bar{y})^2$ หรือ $b_1 = 0$ นั่นเอง เงื่อนไขนี้ได้จากความจริงที่ว่า

$$\begin{aligned} SSR &= \sum(\hat{y}_i - \bar{y})^2 \\ &= \sum(b_0 + b_1 x_i - \bar{y})^2 \\ &= b_1^2 \sum(x_i - \bar{x})^2 \end{aligned}$$

ดังนั้น R^2 จะเป็นศูนย์ต่อเมื่อไม่มีอิทธิพลเชิงเส้น (linear influence) ของ x ต่อ y ในการทดสอบสมมติฐานที่ว่า ไม่มีอิทธิพลเชิงเส้น นั่นคือ $H_0: \beta_1 = 0$ และ $H_a: \beta_1 \neq 0$ เราสามารถใช้ตัวสถิติทดสอบ $T = (b_1 - \beta_1^*)/S_{b_1}$ หรือใช้วิธีการอื่นที่ชื่อว่า การวิเคราะห์ความแปรปรวน (Analysis of Variance, ANOVA) ซึ่งใช้ตัวสถิติ

$$F = MSR/MSE = (SSR/1)/(SSE/(n-2))$$

ซึ่งมีการแจกแจงแบบเอฟ ที่มีองศาความเป็นอิสระ 1 และ $(n-2)$

4. <u>การทำนาย</u> (Prediction)

สมมติว่าเรามีจุดประสงค์จะทำนายค่าของ y สำหรับบางค่าของ x ที่ระบุไว้ ซึ่งให้เป็น x_0 ถ้าพารามิเตอร์ถดถอย β_0 และ β_1 ทราบมาก่อนทดลอง แล้วเราใช้ $E(y_0)$ เป็นตัวทำนาย (predictor) ของ y ณ x_0 เพราะว่า

$$E(y_0) = \beta_0 + \beta_1 x_0$$

ตามปกติเราไม่ทราบค่าของ β_0 และ β_1 เราจึงต้องประมาณโดยอาศัยค่าสังเกตจากตัวอย่าง สมมติว่าเราใช้ค่าประมาณแบบกำลังสองน้อยสุด b_0 และ b_1 ของ β_0 และ β_1 แล้วค่าประมาณแบบจุด $\hat{y}_0$ สำหรับค่า x_0 จะเป็น

$$\hat{y}_0 = b_0 + b_1 x_0$$

ผลต่างระหว่างค่าพยากรณ์ที่แท้จริง (actual forecast) y_0 และค่าพยากรณ์ที่ประมาณได้ (estimated forecast) $\hat{y}_0$ นี้เรียกว่า ความคลาดเคลื่อนของการพยากรณ์ (Forecast Error) ซึ่งกำหนดไว้ว่า

$$y_0 - \hat{y}_0 = (\beta_0 + \beta_1 x_0 + \epsilon_0) - (b_0 - b_1 x_0)$$

เราจะได้ว่า
$$E(y_0 - \hat{y}_0) = 0$$
$$V(y_0 - \hat{y}_0) = E(y_0 - E(y_0))^2 + E\{E(y_0) - \hat{y}_0\}^2$$
หรือ $$\sigma_f^2 = \sigma^2 + \sigma_0^2$$

ดังนั้นความแปรปรวนของการพยากรณ์แยกออกได้เป็น 2 ส่วนคือ ความแปรปรวนของตัวคลาดเคลื่อน กับความแปรปรวนของตัวทำนาย y_0 สำหรับ σ_0^2 จะแยกออกได้เป็น
$$\sigma_0^2 = \sigma^2(1/n + (x_0 - \bar{x})^2/\sum(x_i - \bar{x})^2)$$
ดังนั้นความแปรปรวนของการพยากรณ์จะเป็น
$$\sigma_f^2 = \sigma^2\left(1 + 1/n + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)$$
ซึ่งเราจะเห็นได้ว่า σ_f^2 จะน้อยลง ถ้าตัวอย่างขนาดโต หรือการกระจัดกระจายของค่า x รอบค่าเฉลี่ยมีค่ามาก หรือระยะทางระหว่าง x_0 กับ $\bar{x}$ มีค่าน้อย จากผลอันนี้หมายความว่าการพยากรณ์จะดีมากขึ้น ถ้าค่าของ x ใกล้ค่า $\bar{x}$ มาก ๆ

สำหรับปัญหาถดถอยส่วนมากนั้น σ^2 จะไม่ทราบ จึงต้องประมาณ ซึ่งเราจะได้ค่าประมาณไม่เอียงเฉของ σ^2 ดังนี้
$$s_f^2 = s^2(1 + 1/n + (x_0 - \bar{x})^2/\sum(x_i - \bar{x})^2)$$
เราสามารถหาตัวสถิติทดสอบ T ซึ่งกำหนดไว้ว่า
$$(\hat{y}_0 - E(y_0))/s_0$$
สำหรับทดสอบสมมติฐาน $H_0: E(y_0) = y_0^*$ ในเมื่อ $E(y_0) = \beta_0 + \beta_1 x_0$ ส่วน s_0^2 นั้นเป็นค่าประมาณของความแปรปรวนของตัวทำนาย y_0 ซึ่งกำหนดไว้ว่า
$$s_0^2 = s^2(1/n + (x_0 - \bar{x})^2/\sum(x_i - \bar{x})^2)$$
ตัวสถิติทดสอบ T นี้ ถ้าสมมติฐาน H_0 เป็นจริง จะมีการแจกแจง t ที่มีองศาความเป็นอิสระ $(n-2)$

ช่วงเชื่อมั่น $100(1-\alpha)$ % ของตัวทำนาย $E(y_0)$ หรือ $E(y_0/x_0)$ จะเป็น
$$E(y_0) = (b_0 - b_1 x_0) \pm t_{\alpha/2}^{(n-2)} s_0$$
บางครั้งเราสนใจที่จะหาช่วงเชื่อมั่น $100(1-\alpha)$ % สำหรับ y_0 ใด ๆ ที่เกี่ยวพันกับค่าเฉพาะ x_0 ของ x ซึ่งจะได้เป็น
$$(b_0 + b_1 x_0) \pm t_{\alpha/2}^{(n-2)} s\sqrt{(1 + 1/n + (x_0 - \bar{x})^2/\sum(x_i - \bar{x})^2)}$$

5. <u>การวิเคราะห์ตัดสินใจกับตัวแบบถดถอยเชิงเส้นแบบง่าย</u>

เราจะพิจารณาการประมาณค่า และการอนุมานของตัวแบบถดถอยเชิงเส้นแบบง่ายจากแง่ของทฤษฎีตัดสินใจ วิธีการแบบนี้ก็คือการสมมติฟังก์ชันหนาแน่นน่าจะเป็นก่อนทดลองให้แก่พารามิเตอร์ถดถอย และความแปรปรวนของตัวคลาดเคลื่อนเชิงสุ่ม และแล้วจึงปรับปรุงฟังก์ชันน่าจะเป็นก่อนทดลองเหล่านี้ โดยอาศัยข้อมูลข่าวสารตัวอย่าง $y_1, y_2, \ldots, y_n$ และ $x_1, x_2, \ldots, x_n$ สำหรับตัวแบบถดถอยนั้นจะเป็นเชิงเส้นแบบง่ายดังที่กล่าวมาแล้ว และมีข้อกำหนด 1 – 3 กำกับอยู่ ฟังก์ชันน่าจะเป็น $f(y_1, y_2, \ldots, y_n / \beta_0, \beta_1, \sigma^2, x_1, x_2, \ldots, x_n)$ จะได้เป็น

$$\begin{aligned} & f(y_1, y_2, \ldots, y_n / \beta_0, \beta_1, \sigma^2, x_1, x_2, \ldots, x_n) \\ &= (1/2\pi\sigma^2)^{n/2} e^{-(1/2\sigma^2)\sum(y_i - \beta_0 - \beta_1 x_i)^2} \\ &\propto (1/\sigma^2)^{n/2} e^{-(1/2\sigma^2)\sum(y_i - \beta_0 - \beta_1 x_i)^2} \end{aligned}$$

ฟังก์ชันหนาแน่นน่าจะเป็นก่อนทดลองของพารามิเตอร์ถดถอย β_0, β_1, และความแปรปรวน σ^2 สมมติว่าเป็นดังนี้

$$\begin{aligned} p(\beta_0) &\propto k_1, \quad -\infty < \beta_0 < \infty, \quad k_1 \text{ เป็นค่าคงที่} \\ p(\beta_1) &\propto k_2, \quad -\infty < \beta_1 < \infty, \quad k_2 \text{ เป็นค่าคงที่} \\ p(\sigma^2) &\propto 1/\sigma^2, \quad \sigma > 0 \end{aligned}$$

ถ้าเราสมมติว่าฟังก์ชันหนาแน่นน่าจะเป็นก่อนทดลองของ β_0, β_1, และ σ^2 เป็นอิสระกันแล้ว เราจะได้ฟังก์ชันหนาแน่นน่าจะเป็นร่วมของพารามิเตอร์เหล่านี้เป็น

$$p(\beta_0)\,p(\beta_1)\,p(\sigma^2) \propto k_1 k_2/\sigma^2$$

โดยการใช้สมการนี้ ฟังก์ชันน่าจะเป็น และทฤษฎีเบยส์ เราก็จะสามารถหาฟังก์ชันหนาแน่นน่าจะเป็นที่ปรับปรุงร่วมของ β_0, β_1, และ σ^2 ได้ ซึ่งจะแทนด้วย $p(\beta_0, \beta_1, \sigma^2 / y_1, y_2, \ldots, y_n; x_1, x_2, \ldots, x_n)$ ดังนั้นเราได้

$$\begin{aligned} & p(\beta_0, \beta_1, \sigma^2 / y_1, y_2, \ldots, y_n; x_1, x_2, \ldots, x_n) \\ &\propto (1/\sigma^2)^{(n+2)/2} e^{-(1/2\sigma^2)\sum(y_i - \beta_0 - \beta_1 x_i)^2} \end{aligned}$$

กำหนด $\bar{y} = \sum y_i/n$, $\bar{x} = \sum x_i/n$

$$b_0 = \bar{y} - b_1\bar{x} \ , \ b_1 = \sum(y_i - \bar{y})(x_i - \bar{x}) / \sum(x_i - \bar{x})^2$$

$$s^2 = \sum(y_i - b_0 - b_1x_i)^2/(n-2)$$

พิจารณา

$$\begin{aligned}\sum(y_i - \beta_0 - \beta_1x_i)^2 &= \sum(y_i - \beta_0 - \beta_1x_i - b_0 - b_1x_i + b_0 + b_1x_i)^2 \\ &= \sum\{(y_i - b_0 - b_1x_i) - (\beta_0 - b_0) - (\beta_1 - b_1)x_i\}^2 \\ &= \sum(y_i - b_0 - b_1x_i)^2 + n(\beta_0 - b_0)^2 + (\beta_1 - b_1)^2\sum x_i^2 \\ &\quad + 2(\beta_0 - b_0)(\beta_1 - b_1)\sum x_i \\ &= (n-2)s^2 + n(\beta_0 - b_0)^2 + (\beta_1 - b_1)\sum x_i^2 \\ &\quad + 2(\beta_0 - b_0)(\beta_1 - b_1)\sum x_i\end{aligned}$$

เพราะฉะนั้นเราจะได้ฟังก์ชั่นหนาแน่นน่าจะเป็นร่วมที่ปรับปรุง ดังนี้

$$\begin{aligned}&p(\beta_0, \beta_1, \sigma^2 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) \\ &\quad \propto (1/\sigma^2)^{(n+2)/2} \exp\{-(1/2\sigma^2)[(n-2)s^2 + n(\beta_0 - b_0)^2 \\ &\quad\quad + (\beta_1 - b_1)^2 \sum x_i^2 + 2(\beta_0 - b_0)(\beta_1 - b_1)\sum x_i]\}\end{aligned}$$

ผลที่ได้มานี้ กล่าวได้ดังทฤษฎีต่อไปนี้

<u>ทฤษฎี 1</u> ให้ $y_i = \beta_0 + \beta_1x_i + \epsilon_i$ เป็นตัวแบบถดถอยเชิงเส้นแบบง่าย ที่สอดคล้องกับสมมติฐาน หรือข้อกำหนด 1–3 สมมติฟังก์ชั่นหนาแน่นน่าจะเป็นก่อนทดลองของ $\beta_0, \beta_1,$ และ σ^2 เป็นอิสระกัน และสมมติว่าเป็นแบบยูนิฟอร์ม ดังนี้

$$\begin{aligned}p(\beta_0) &\propto k_1 \ , \ -\infty < \beta_0 < \infty, \ k_1 \text{ คงที่} \\ p(\beta_1) &\propto k_2 \ , \ -\infty < \beta_1 < \infty, \ k_2 \text{ คงที่} \\ p(\sigma^2) &\propto 1/\sigma^2, \quad \sigma > 0\end{aligned}$$

แล้วฟังก์ชั่นหนาแน่นน่าจะเป็นร่วมที่ปรับปรุง กำหนดไว้ดังสมการข้างบนนั้น ในเมื่อ $\bar{y}, \bar{x}$, $b_0, b_1,$ และ s^2 นิยามไว้ดังข้างบนนี้

ฟังก์ชั่นหนาแน่นน่าจะเป็นร่วมที่ปรับปรุงได้นี้ เป็นฟังก์ชั่นพื้นฐานสำหรับทำการอนุมานเกี่ยวกับ $\beta_0, \beta_1, \sigma^2$ ตัวอย่างเช่น เราสามารถหาฟังก์ชั่นร่วมที่ปรับปรุงของ β_0 และ β_1 ได้เป็นดังนี้

$$
\begin{aligned}
p(\beta_0, \beta_1 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) \\
&\propto \int_0^\infty p(\beta_0, \beta_1, \sigma^2 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n)\, d\sigma^2 \\
&\propto \left[(n-2)s^2 + n(\beta_0 - b_0)^2 + (\beta_1 - b_1)^2 \sum x_i^2 \right. \\
&\qquad \left. + 2(\beta_0 - b_0)(\beta_1 - b_1)\sum x_i\right]^{-n/2}
\end{aligned}
$$

ในทำนองเดียวกันเราจะได้ฟังก์ชั่นที่ปรับปรุงของ β_0 ดังนี้

$$
\begin{aligned}
p(\beta_0 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) \\
&\propto \left[(n-2) + (\beta_0 - b_0)^2 \frac{\sum (x_i - \bar{x})^2}{s^2 \sum x_i^2 / n}\right]^{-(n-1)/2}
\end{aligned}
$$

เมื่อกำหนด $T_0 = \dfrac{(\beta_0 - b_0)\sqrt{\sum (x_i - \bar{x})^2}}{\sqrt{s^2 \sum x_i^2 / n}}$ แล้วเราจะได้

$$
\begin{aligned}
p(t_0 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) &\propto \left[n - 2 + t_0^2\right]^{-(n-1)/2} \\
&\propto \left[1 + t_0^2/(n-2)\right]^{-(n-1)/2}
\end{aligned}
$$

ฟังก์ชั่นนี้จะเป็นแบบฟอร์มของการแจกแจง t ที่มีองศาความเป็นอิสระ $n - 2$

<u>ทฤษฎี 2</u> สำหรับตัวแบบถดถอยเชิงเส้นแบบง่าย และสำหรับฟังก์ชั่นหนาแน่นน่าจะเป็นก่อนทดลองของ β_0, β_1, และ σ^2 ดังที่กล่าวไว้ในทฤษฎี 1 การแจกแจงน่าจะเป็นที่ปรับปรุงของอัตราส่วน T_0 ที่กล่าวไว้ข้างบนนี้ เมื่อกำหนดข้อมูลข่าวสารตัวอย่าง $y_1, y_2, \ldots, y_n$ และ $x_1, x_2, \ldots, x_n$ ให้นี้ จะมีการแจกแจง t ที่มีองศาความเป็นอิสระ $n - 2$

จากฟังก์ชั่นร่วมของ β_0 และ β_0 นั้นเราสามารถหาฟังก์ชั่นของ β_1 ได้ดังนี้

$$
\begin{aligned}
p(\beta_1 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) \\
&\propto \left[(n-2) + \frac{\sum (x_i - \bar{x})^2}{s^2}(\beta_1 - b_1)^2\right]^{-(n-1)/2} \\
p(t_1 / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) &\propto \left[(n-2) + t_1^2\right]^{-(n-1)/2} \\
&\propto \left[1 + t_1^2/(n-2)\right]^{-(n-1)/2}
\end{aligned}
$$

ในเมื่อ $T_1 = \dfrac{\beta_1 - b_1}{\sqrt{s^2 / \sum (x_i - \bar{x})^2}}$

ดังนั้นการแจกแจงน่าจะเป็นที่ปรับปรุงของอัตราส่วน T_1 ที่กำหนดไว้ข้างบน จะมีการแจกแจงปกติที่มีองศาความเป็นอิสระ $n - 2$

ทฤษฎี 3 ข้อกำหนดเช่นเดียวกับทฤษฎี 1 การแจกแจงน่าจะเป็นที่ปรับปรุงของอัตราส่วน T_1 โดยกำหนด $y_1, y_2, .., y_n, x_1, x_2, .., x_n$ ให้นั้น จะมีการแจกแจง t ที่มีองศาความเป็นอิสระ $n - 2$

ในทำนองเดียวกัน จากฟังก์ชั่นร่วมที่ปรับปรุงของ β_0, β_1, และ σ^2 เราสามารถหาฟังก์ชั่นที่ปรับปรุงของ σ^2 ได้เป็น

$$p(\sigma^2 / y_1, y_2, .., y_n, x_1, x_2, .., x_n) \propto (1/\sigma^2)^{n/2} e^{-\frac{(n-2)s^2}{2\sigma^2}}$$

ถ้าเรากำหนด $U = (n - 2)s^2/\sigma^2$ แล้วเราจะได้

$$p(u / y_1, y_2, .., y_n, x_1, x_2, .., x_n) \propto u^{n/2} e^{-u/2}$$

นั่นคือ การแจกแจงน่าจะเป็นที่ปรับปรุงของ U เป็นแบบไคสแควร์ ที่มีองศาความเป็นอิสระ $(n - 2)$ ดังนั้นเราจะได้ทฤษฎีต่อไปนี้

ทฤษฎี 4 ข้อกำหนดเช่นเดียวกับทฤษฎี 1 แล้วการแจกแจงน่าจะเป็นที่ปรับปรุงของ $U = (n - 2)s^2/\sigma^2$ หลังจากสังเกต $y_1, y_2, .., y_n, x_1, x_2, .., x_n$ นั้นจะเป็นการแจกแจงไคสแควร์ ที่มีองศาความเป็นอิสระ $(n - 2)$

6. การอ้างอิงเกี่ยวกับ β_0, β_1, และ σ^2

การแจกแจงน่าจะเป็นที่ปรับปรุงของตัวสถิติ T_0, T_1, และ U ที่กล่าวมานั้นจะใช้สำหรับทำการอ้างอิงเกี่ยวกับ β_0, β_1, และ σ^2 แบบทดสอบ หรือตัวสถิตินี้จะใช้ได้เมื่อความแปรปรวนของตัวคลาดเคลื่อนเชิงสุ่มไม่ทราบ

ถ้าเราต้องการทดสอบสมมติฐานว่า ไม่มีอิทธิพลเชิงเส้นของ x ใน y นั่นคือ

$$H_0: \beta = 0 \quad \text{กับ} \quad H_a: \beta \neq 0$$

เราใช้ตัวสถิติทดสอบ T_1

เมื่อเราให้ $F = T_1^2$ เราจะได้ว่า F มีการแจกแจงแบบเอฟ ที่มีองศาความเป็นอิสระ $1, n - 2$ ดังนั้นเราจึงใช้ตัวสถิติทดสอบ

$$F = \frac{b_1^2}{s^2/\sum(x_i - \bar{x})^2}$$

ทดสอบสมมติฐานข้างบนนี้ได้

ช่วงเชื่อมั่น 100(1 $-\alpha$) % ของ β_1 พิจารณาได้จากขีดจำกัดของ β_1 ดังนี้

$$\beta_1 = b_1 \pm t_{\alpha/2}^{(n-2)} s/\sqrt{\sum(x_i-\bar{x})^2}$$

ในทำนองเดียวกัน ในการทดสอบสมมติฐาน

$$H_0: \beta_0 = \beta_0^* \quad \text{กับ} \quad H_a: \beta_0 \neq \beta_0^*$$

เราใช้ตัวสถิติทดสอบ T_0

ช่วงเชื่อมั่น 100(1$-\alpha$) % ของ β_0 กำหนดได้จาก

$$\beta_0 = b_0 \pm t_{\alpha/2}^{(n-2)} s\sqrt{\frac{\sum x_i^2/n}{\sum(x_i-\bar{x})^2}}$$

ในการทดสอบสมมติฐาน

$$H_0: \sigma^2 = \sigma^{2*} \quad \text{กับ} \quad H_a: \sigma^2 = \sigma^{2*}$$

เราใช้ตัวสถิติทดสอบ U

ในทำนองเดียวกัน การวิเคราะห์แบบทดสอบของสมมติฐานเกี่ยวกับ β_0 และ β_1 เมื่อทราบ σ^2 นั้นก็ทำได้โดยตรง คือใช้ตัวสถิติทดสอบ

$$Z = \frac{\beta_0 - b_0}{\sigma\sqrt{\sum x_i^2/n}}\sqrt{\sum(x_i-\bar{x})^2}$$

$$\text{และ} \quad Z = \frac{\beta_1 - b_1}{\sigma/\sqrt{\sum(x_i-\bar{x})^2}}$$

ต่อไปเราจะหาการแจกแจงน่าจะเป็นที่ปรับปรุงของค่าคาดหวังเงื่อนไข $E(y_0/x_0)$ เมื่อ x_0 เป็นค่าเฉพาะของตัวแปรอิสระ เราทราบว่า

$$E(y_0/x_0) = \eta_0 = \beta_0 + \beta_1 x_0$$

ถ้าไม่ทราบค่าของ σ^2 และถ้าการแจกแจงน่าจะเป็นที่ปรับปรุงของ η_0 กำหนด $y_1, y_2, \ldots, y_n$ และ $x_1, x_2, \ldots, x_n$ ให้ หรือ $p(\eta_0/y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n)$ แล้วเราจะสามารถแสดงได้ว่า

$$p(\eta_0/y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n) \propto \left[(n-2) + \frac{n\sum(x_i-\bar{x})^2}{s^2\sum(x_i-x_0)^2}(\eta_0-\hat{\eta}_0)^2\right]^{-(n-1)/2}$$

เมื่อ $\hat{\eta}_0 = b_0 - b_1 x_0$

ถ้าเรากำหนดตัวแปรใหม่เป็น $T = \dfrac{(\eta_0 - \hat{\eta}_0)\sqrt{n\sum(x_i-\bar{x})^2}}{s\sqrt{\sum(x_i - x_0)^2}}$

$$= \frac{\eta_0 - \hat{\eta}_0}{s\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}}}$$

ในเมื่อ $\sum(x_i - x_0)^2 = \sum(x_i - \bar{x})^2 + n(x_0 - \bar{x})^2$

เพราะฉะนั้น $p(t / y_1, y_2, \ldots, y_n, x_1, x_2, \ldots, x_n)$

$$\propto \left[(n-2) + t^2\right]^{-(n-1)/2}$$

$$\propto \left[1 + t^2/(n-2)\right]^{-(n-1)/2}$$

ดังนั้นอัตราส่วน T นี้มีการแจกแจง t ที่มีองศาความเป็นอิสระ $(n-2)$ เพราะฉะนั้นในการทดสอบสมมติฐาน

$$H_0: \eta_0 = \eta_0^* \quad \text{กับ} \quad H_a: \eta_0 \neq \eta_0^*$$

เราจึงใช้ตัวสถิติทดสอบ T นั้น

ช่วงเชื่อมั่น $100(1-\alpha)$ % ของ η_0 พิจารณาได้จาก

$$\eta_0 = \hat{\eta}_0 + t_{\alpha/2}^{(n-2)}\, s\sqrt{\frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$$

จะสังเกตเห็นว่า ความยาวของช่วงเชื่อมั่นจะกว้าง ถ้า x_0 บ่ายเบนมากจากค่าเฉลี่ย $\bar{x}$ ของค่าตัวอย่าง $x_1, x_2, \ldots, x_n$

ในทำนองเดียวกันถ้าเราประสงค์จะสร้างช่วงเชื่อมั่นสำหรับค่า y_0 ค่าหนึ่ง ๆ ที่เกี่ยวพันกับค่า x_0 ที่ระบุไว้ เราก็พิจารณาจากตัวสถิติ

$$T = \frac{y_0 - (b_0 + b_1x_0)}{s\sqrt{1 + 1/n + (x_0-\bar{x})^2/\sum(x_i-\bar{x})^2}}$$

ซึ่งมีการแจกแจงแบบ t ซึ่งมีองศาความเป็นอิสระ $(n-2)$ และเราจะได้ช่วงเชื่อมั่น $100(1-\alpha)$ % ของ y_0 ดังนี้

$$y_0 = (b_0 + b_1x_0) \pm t_{\alpha/2}^{(n-2)}\sqrt{s^2\left[1+\frac{1}{n} + \frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right]}$$

7. <u>การวิเคราะห์ตัดสินใจกับตัวแบบถดถอยเชิงเส้นทั่วไป</u>

ตัวแบบถดถอยเชิงเส้นทั่วไป กำหนดไว้ดังนี้

$$y_t = \beta_1 + \beta_2 x_{2t} + \cdots\cdots + \beta_k x_{kt} + \varepsilon_t ; \quad t = 1,2,..,n$$

ในเมื่อ y_t เป็นค่าสังเกตที่ t ของตัวแปรตาม x_{it} เป็นค่าสังเกตที่ t ในตัวแปรอิสระ ε_t เป็นค่าที่สังเกตไม่ได้ของตัวคลาดเคลื่อนเชิงสุ่ม และ $\beta_1, \beta_2, .., \beta_k$ เป็นพารามิเตอร์ถดถอย

ตัวแบบถดถอยเชิงเส้นทั่วไป หรือตัวแบบถดถอยเชิงพหุ สามารถแสดงในรูปแมทริคซ์ได้เป็น

$$Y = X\beta + \varepsilon$$

หรือ

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{21} & \cdots & x_{k1} \\ 1 & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & & \vdots \\ 1 & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

ตัวแบบถดถอยนี้มีข้อกำหนดกำกับไว้ ดังนี้

<u>ข้อกำหนด 1</u> สมาชิกของ ε เป็นอิสระกัน และแจกแจงแบบปกติ ที่มีค่าเฉลี่ย 0 และความแปรปรวนเท่ากันหมด σ^2 นั่นคือ $E(\varepsilon) = 0$ และ $V(\varepsilon) = E(\varepsilon'\varepsilon) = \sigma^2 I_n$ ในเมื่อ 0 เป็นเวคเตอร์ที่มีมิติเป็น $n \times 1$ และ I_n เป็นแมทริคซ์เอกลักษณ์ที่มีมิติเป็น $n \times n$

<u>ข้อกำหนด 2</u> แมทริคซ์ X ไม่เป็นแบบสุ่ม (Nonstochastic) ข้อกำหนดนี้ละเลยได้ เพื่อที่จะกล่าวว่า X เป็นตัวแปรเชิงสุ่ม แต่ตัวแปรอิสระ X_i $(i=1,2,..,n)$ แจกแจงเป็นอิสระกับ ε ที่มีการแจกแจงไม่เกี่ยวข้องกับพารามิเตอร์ $\beta_1, \beta_2, .., \beta_k$ และ σ^2

<u>ข้อกำหนด 3</u> ค่าสังเกตทั้งหมดวัดได้โดยไม่มีความคลาดเคลื่อน

<u>ข้อกำหนด 4</u> แมทริคซ์ X มีอันดับ (Rank) $k < n$

ค่าประมาณแบบกำลังสองน้อยสุดของ $\beta_1, \beta_2, .., \beta_k$ และ σ^2 พิจารณาได้จาก

$$\sum e_i^2 = \sum(y_i - \beta_1 x_{1i} - \beta_2 x_{2i} - \dots - \beta_k x_{ki})^2$$

หรือ $\varepsilon'\varepsilon = (y - x\beta)'(y - x\beta)$

โดยการทำให้มันมีค่าน้อยที่สุด เมื่อเทียบกับ β เราสามารถแสดงว่าค่าประมาณจะเป็น

$$\hat{\beta} = (x'x)^{-1}(y - x\beta)$$

และสามารถแสดงได้ว่า

$$E(\hat{\beta}) = \beta$$

$$E(\hat{\beta}) = \sigma^2(x'x)^{-1}$$

สำหรับ $\hat{\beta}$ นั้นจะเป็นตัวประมาณค่าไม่เอียงเฉเชิงเส้นที่ดีที่สุด (BLUE, Best Linear Unbiased Estimators) ค่าประมาณนั้นอาจได้จากวิธีการประมาณแบบน่าจะเป็นมากสุดอีกวิธีหนึ่ง

ต่อไปเราจะขยายผลข้างบนนี้ในแบบวิเคราะห์การตัดสินใจ ฟังก์ชั่นน่าจะเป็นของ β และ σ^2 กำหนดว่าข้อมูลข่าวสารจากตัวอย่าง y และ x สามารถเขียนได้เป็น

$$p(y/\beta, \sigma^2, x) \propto (1/\sigma^2)^{n/2} e^{-\frac{1}{2\sigma^2}(y - x\beta)'(y - x\beta)}$$

เราสมมติว่าฟังก์ชั่นก่อนทดลองของ β และ σ^2 เป็นอิสระกัน และแจกแจงแบบยูนิฟอร์ม ดังนั้น ฟังก์ชั่นก่อนทดลองร่วมกันของ β และ σ^2 สามารถแสดงได้เป็น

$$p(\beta, \sigma^2) \propto k/\sigma^2, \quad -\infty < \beta < \infty, \quad \sigma^2 > 0$$

เพราะฉะนั้นฟังก์ชั่นที่ปรับปรุงร่วมของ β และ σ^2 โดยกำหนด y และ x จะเป็น

$$p(\beta, \sigma^2 / y, x) \propto (1/\sigma^2)^{(n+1)/2} e^{-\frac{1}{2\sigma^2}(y - x\beta)'(y - x\beta)}$$

แต่ $(y - x\beta)'(y - x\beta) = [(y - x\hat{\beta}) - x(\beta - \hat{\beta})]'[(y - x\hat{\beta}) - x(\beta - \hat{\beta})]$

$$= (y - x\hat{\beta})'(y - x\hat{\beta}) + (\beta - \hat{\beta})'x'x(\beta - \hat{\beta})$$

ในเมื่อ $\hat{\beta} = (x'x)^{-1}x'y$

ดังนั้น $p(\beta, \sigma^2 / y, x) \propto (1/\sigma^2)^{(n+1)/2} exp\{-(1/2\sigma^2)[(y - x\hat{\beta})'(y - x\hat{\beta}) + (\beta - \hat{\beta})'x'x(\beta - \hat{\beta})]\}$

ให้ $S^2 = (y - x\hat{\beta})'(y - x\hat{\beta})/(n - k)$

ซึ่งเราจะได้ $$p(\beta,\sigma^2/y,x) \propto (1/\sigma^2)^{(n+1)/2} \exp[-(1/2\sigma^2)\{(\beta-\hat{\beta}) + (\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\}]$$

สมการนี้เป็นฟังก์ชั่นหนาแน่นน่าจะเป็นเบื้องต้นที่จะทำการอ้างอิงเกี่ยวกับตัวใดตัวหนึ่ง หรือเซทย่อยใด ๆ ของสมาชิก β และ σ^2 จากฟังก์ชั่นนี้เราจะได้

$$p(\beta/y,x) \propto \{(n-k)s^2 + (\beta-\hat{\beta})'X'X(\beta-\hat{\beta})\}^{-(n/2)}$$

ซึ่งเป็นรูปฟอร์ของการแจกแจงหลายตัวแปร t (Multivariate t Distribution) ในการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ถดถอยตัวใดตัวหนึ่ง β_t นั้นเราสามารถแสดงได้ว่า อัตราส่วน T

$$T = \frac{\beta_t - \hat{\beta}_t}{\sqrt{s^2 m^{tt}}}, \qquad t = 1,2,..,k$$

ในเมื่อ m^{tt} เป็นสมาชิกที่ (t,t) ของ $(X'X)^{-1}$ และ $\hat{\beta}_t$ เป็นสมาชิกที่ t ของ $(X'X)^{-1}X'Y$ T มีการแจกแจง t ที่มีองศาความเป็นอิสระ $(n-k)$

ตัวสถิติทดสอบ หรืออัตราส่วน T นี้นอกจากจะใช้ทดสอบสมมติฐาน แล้วยังใช้สร้างช่วงเชื่อมั่น $100(1-\alpha)$ % สำหรับ β_t ได้เป็น

$$\beta_t = \hat{\beta}_t \pm t_{\alpha/2}^{(n-k)} s\sqrt{m^{tt}}$$