

ความหมายของค่าสถิติของการวิเคราะห์การถดถอยด้วยวิธี OLS

ความหมายของค่าสถิติจากการวิเคราะห์ความถดถอย

ผลการวิเคราะห์ความถดถอยนั้นจะแสดงค่าทางสถิติต่างๆ ดังในตารางที่ 4.7 ค่าต่างๆ ดังกล่าว มีความสำคัญต่อการตีความผลของการวิเคราะห์ถดถอย ค่าต่างๆ ที่สำคัญมีดังนี้

1. ค่าความคาดเคลื่อนมาตรฐานของสัมประสิทธิ์ (Standard errors of coefficient: s_b)

ค่าค่าความคาดเคลื่อนมาตรฐานของสัมประสิทธิ์จะได้มาพร้อมกับการประมาณค่าสัมประสิทธิ์ของสมการถดถอยซึ่งค่านี้จะใช้ในการกำหนดช่วงความเชื่อมั่นและใช้ในการหาค่าสถิติ t เพื่อทดสอบว่าค่าสัมประสิทธิ์นั้นมีค่าเท่ากับศูนย์หรือไม่

2. ค่าสถิติ t (t-statistics)

ค่าสถิติ t คำนวณจาก $t = b / S_b$ ใช้สำหรับการทดสอบสมมติฐานว่าค่าสัมประสิทธิ์นั้นมีค่าเท่ากับศูนย์หรือไม่ ถ้าค่าสถิติที่คำนวณได้มากกว่าค่าวิกฤตแล้ว แสดงว่าปฏิเสธสมมติฐานหลัก (H_0) นั่นคือ ค่าสัมประสิทธิ์นั้นไม่เท่ากับศูนย์

3. ค่านัยสำคัญ (Significance values)

ค่านัยสำคัญ อาจจะเรียกว่าเป็นค่า Prob-values หรือ p-value เนื่องจากค่านี้แสดงถึงความน่าจะเป็นของค่าสถิติ t ที่จะเกิดขึ้นถ้าไม่มีความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ดังนั้นถ้าค่าสถิติ t ที่คำนวณได้มากกว่าค่าวิกฤตที่ระดับนัยสำคัญ ร้อยละ 5 ค่า p-value ต้องน้อยกว่า 0.05 (สมมติว่าเป็นการทดสอบสองหาง หรือ two-tail test)

4. ค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณ (Standard error of the estimate: s_e)

ค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณ หมายถึง ค่าเบี่ยงเบนมาตรฐานของตัวแปรตาม ดังสมการที่ 4.11 k คือ จำนวนของตัวแปรอิสระ ค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณนี้มีประโยชน์ในการวัดค่าความถูกต้องแม่นยำของค่า \hat{Y} หรือ goodness of fit ของสมการถดถอยนั่นเอง ค่าดังกล่าวจะเท่ากับศูนย์ถ้ามีความสัมพันธ์กันอย่างสมบูรณ์ ดังนั้นค่านี้จึงเป็นค่าที่ผกผันกับ R^2

$$s_e = \sqrt{\frac{\sum (Y - \hat{Y})^2}{n - k - 1}} \quad 4.11$$

5. ค่าสัมประสิทธิ์ตัวกำหนด (Coefficient of determination: R^2)

ค่าสัมประสิทธิ์ตัวกำหนดเป็นค่าที่ชี้ให้เห็นว่าตัวแปรอิสระนั้นสามารถอธิบายความแปรปรวนของตัวแปรตามได้มากน้อยเพียงใด ดังสมการที่ 4.12 ตัวอย่างเช่น $R^2 = 0.98$ หมายความว่า ตัวแปรอิสระสามารถอธิบายความแปรปรวนของตัวแปรตามได้ร้อยละ 98 ส่วนอีกร้อยละ 2 นั้นเกิดจากปัจจัยอย่างอื่น หรือ ความแปรปรวนทั้งหมดของตัวแปรตามลดลงร้อยละ 98 อันเป็นผลมาจากตัวแปรอิสระที่นำมาพิจารณา

$$R^2 = \frac{\sum (\hat{Y} - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad 4.12$$

6. ค่าสัมประสิทธิ์ตัวกำหนดที่ปรับค่า (Adjusted Coefficient of determination: adj. R^2)

ความหมายเช่นเดียวกับข้อ 5 แต่เป็นการปรับค่าเพื่อขจัดปัญหาเกี่ยวกับการเพิ่มตัวแปรอิสระเข้าไปในสมการถดถอย เพราะโดยทั่วไปถ้าตัวแปรอิสระที่ไม่สัมพันธ์กับตัวแปรตามนั้นเมื่อมีมากขึ้นในสมการจะทำให้ค่า R^2 เพิ่มสูงขึ้น การขจัดปัญหาดังกล่าวจึงปรับโดยค่าองศาอิสระ (degree of freedom) ดังสมการที่ 4.12

$$adj. R^2 = 1 - \frac{(1 - R^2)(n - k)}{(n - k - 1)} \quad 4.13$$

7. ค่าสถิติ F (F-statistics)

ค่าสถิติ F เป็นค่าที่ใช้ทดสอบสัมประสิทธิ์ของตัวแปรอิสระเป็นกลุ่มหรือโดยรวมว่ามีความสัมพันธ์กับตัวแปรตามหรือไม่ กรณีสมการเป็นแบบถดถอยอย่างง่ายซึ่งมีตัวแปรอิสระตัวเดียวค่าสถิติ F ก็คือค่าสถิติ t นั่นเอง

อย่างไรก็ตามการจะนำสมการถดถอยมาใช้ในการอธิบายและพยากรณ์นั้นต้องอยู่ภายใต้ข้อสมมติของการวิเคราะห์ถดถอย นั่นคือต้องไม่มีการละเมิดในข้อสมมติของวิธี OLS (Ordinary Least Square) ซึ่งปัญหาการละเมิดข้อสมมติและปัญหาอื่นที่จะเกิดกับการวิเคราะห์การถดถอยนั้นจะได้กล่าวต่อไป