

# ข้อพึงระวังในการวิเคราะห์ความสัมพันธ์ของตัวแปร โดยวิธีวิเคราะห์การถดถอยในการศึกษาตลาดทุน

ดร.นนทวรรณ ยมจินดา\*

## บทคัดย่อ

บทความนี้นำเสนอข้อพึงระวังและตัวอย่างปัญหาทางสถิติในการทำการศึกษาความสัมพันธ์ระหว่างตัวแปรที่รวบรวมได้จากฐานข้อมูลตลาดทุน โดยใช้วิธีการทางสถิติ คือ การวิเคราะห์การถดถอย (Regression Analysis) ผู้วิจัยควรระวังและคำนึงถึงสมมติฐานทางสถิติที่ใช้ในการวิเคราะห์ข้อมูล รวมถึงการแปลความหมายของผลที่ได้รับจากการวิเคราะห์โดยใช้วิธีการทางสถิติที่คลาดเคลื่อน

**คำสำคัญ:** การวิเคราะห์การถดถอย การวิเคราะห์ความสัมพันธ์ ตลาดทุน

## ABSTRACT

This paper provides concerns and examples on statistical issues on using multiple linear regression methods in capital market studies. Concerns on violating statistical assumptions of market data, potential problems on stating hypotheses and regression model, as well as on concluding the results of study are also discussed in this paper.

**Keywords:** Regression Analysis, Correlation Analysis, Capital Market

\* ผู้ช่วยศาสตราจารย์ประจำ ภาควิชาการบัญชี คณะพาณิชยศาสตร์และการบัญชี มหาวิทยาลัยธรรมศาสตร์

## บทนำ

การวิเคราะห์ความสัมพันธ์โดยใช้สมการถดถอย (Regression Analysis) เป็นวิธีการทางสถิติที่มีการใช้กันอย่างแพร่หลายในงานวิจัยทางบัญชี เพื่อหาความสัมพันธ์ระหว่างตัวแปรตาม กับตัวแปรต้น โดยเฉพาะในงานวิจัยที่ใช้ฐานข้อมูลจากตลาดทุน การศึกษาความสัมพันธ์ของข้อมูลโดยใช้วิธีการวิเคราะห์ความสัมพันธ์นี้ช่วยให้ผู้วิจัยทราบว่าค่าของตัวแปรตามจะเปลี่ยนแปลงไปอย่างไร เมื่อตัวแปรต้นตัวใดตัวหนึ่งมีค่าเปลี่ยนแปลงไป ในขณะที่ตัวแปรต้นตัวอื่น ๆ ไม่เปลี่ยนแปลง วิธีการวิเคราะห์ความสัมพันธ์ของตัวแปรที่นิยมใช้ในงานวิจัยทางบัญชีมีหลายวิธี เช่น การวิเคราะห์การถดถอยเชิงเส้นแบบพหุ (Multiple Linear Regression: MLR) การวิเคราะห์ความสัมพันธ์แบบวิธีกำลังสองน้อยที่สุด (Ordinary Least Square: OLS) หรือการวิเคราะห์การถดถอยโลจิสติกส์ (Logistic Model Regression) เป็นต้น

## ข้อสมมติฐานทางสถิติ

ในการวิเคราะห์ความสัมพันธ์โดยวิธีการวิเคราะห์การถดถอยนี้ต้องเป็นไปตามข้อสมมติฐานและเงื่อนไขทางสถิติหลายข้อซึ่งอาจเป็นข้อจำกัดที่ทำให้การศึกษาข้อมูลจากตลาดทุนอาจไม่เป็นไปตามหลักทางสถิติ หรืออาจทำให้ผลการศึกษาไม่ถูกต้องตรงกับข้อเท็จจริง ทั้งนี้ข้อสมมติขั้นพื้นฐานสำหรับการวิเคราะห์การถดถอยเชิงเส้นแบบพหุ หรือการถดถอยเชิงเส้นแบบหลายตัวแปร (Multiple Linear Regression: MLR) ได้แก่

1) ตัวแปรที่ศึกษามีความสัมพันธ์แบบเส้นตรง (Linear in Parameters) คือ ตัวแปรตาม (Dependent Variable: Y) ที่ต้องการศึกษา มีความสัมพันธ์แบบเส้นตรงกับตัวแปรอิสระ (Independent Variable: X) ทุกตัว ซึ่งทำให้สมการความสัมพันธ์ของประชากรที่ศึกษา สามารถเขียนให้อยู่ในรูปของสมการ

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + u$$

โดยค่าสัมประสิทธิ์  $\beta_0, \beta_1, \beta_2, \dots, \beta_k$  คือ ค่าสัมประสิทธิ์ความสัมพันธ์ที่ได้จากการคำนวณ และ  $u$  คือ ค่าความคลาดเคลื่อนของข้อมูล

2) การเลือกตัวอย่างโดยวิธีการสุ่ม (Random Sampling) หมายถึง ตัวอย่างของข้อมูลที่ใช้ศึกษาต้องได้มาโดยวิธีการสุ่ม นั่นคือ โอกาสที่จะได้รับเลือกมาเป็นกลุ่มตัวอย่างที่ทำการศึกษานั้นเท่ากัน จึงจะทำให้ตัวอย่างดังกล่าวเป็นตัวแทนของประชากร และสามารถนำผลการศึกษาไปใช้ในการทำนายความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้

3) ไม่มีตัวแปรอิสระที่สัมพันธ์กันแบบสมบูรณ์ (Perfect Collinearity) หมายถึง ตัวแปรอิสระแต่ละตัว จะต้องไม่เป็นค่าคงที่ หรือมีความสัมพันธ์กับตัวแปรอิสระอื่นสูงมากจนสามารถถือได้ว่าตัวแปรนั้นมีความสัมพันธ์กันเป็นแบบเส้นตรงกับตัวแปรอิสระอื่น

4) ค่าเฉลี่ยเลขคณิตของความคลาดเคลื่อนมีค่าเป็นศูนย์ (Zero Conditional Mean) หรือ ค่าคาดหวัง (expected value) สำหรับค่าความคลาดเคลื่อนของสมการถดถอยเชิงเส้นแบบพหุมีค่าเท่ากับศูนย์ หรือเขียนให้อยู่ในรูปสมการได้ ดังนี้  $E(u) = 0$  ซึ่งจะทำให้ค่าคาดหวังความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามได้โดยอยู่ในรูปของสมการ

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

5) ค่าความแปรปรวน (variance) ของความคลาดเคลื่อนของสมการ (error term)  $u$  เป็นค่าคงที่ (Homoskedasticity) ไม่แปรไปตามค่าของตัวแปรอิสระ

## ปัญหาทางสถิติที่อาจเกิดขึ้น

การวิเคราะห์ข้อมูลที่ได้รับจากฐานข้อมูลในตลาดทุน นอกจากผู้วิจัยจะต้องระมัดระวังเกี่ยวกับการที่ข้อมูลที่ใช้ในการศึกษาไม่เป็นไปตามข้อสมมติขั้นพื้นฐานสำหรับการใช้สมการถดถอยเชิงเส้นแบบพหุ ในการศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ จนทำให้ผลการศึกษา

ความคลาดเคลื่อนไปจากข้อเท็จจริงแล้วยังมีปัญหาเชิงสถิติในด้านอื่น ๆ อีก ดังนี้

### 1. ปัญหาที่เกิดจากการเลือกตัวอย่างที่ใช้ศึกษา (Population and Samples)

เนื่องจากข้อมูลทางบัญชี เป็นผลจากการบันทึกรายการบัญชีตามสภาพแวดล้อมทางเศรษฐกิจและสังคมของแต่ละกิจการ และตลาดทุนแต่ละแห่ง จึงทำให้ตัวเลขทางบัญชี และข้อมูลอื่น ๆ ที่ได้จากตลาดทุน เช่น ราคาและปริมาณการซื้อขายหลักทรัพย์ หรือข้อมูลข่าวสารต่าง ๆ ที่เกี่ยวข้องกับการศึกษา เกิดจากผลการดำเนินงานเชิงเศรษฐกิจที่เกิดขึ้นจริงและมีลักษณะเฉพาะสำหรับกิจการ ไม่ใช่ข้อมูลที่ได้จากการใช้วิธีสุ่มคัดเลือก (random sampling) ดังนั้น การใช้วิธีคัดเลือกกลุ่มตัวอย่างเพื่อให้เป็นตัวแทนของประชากร จึงอาจเป็นวิธีการที่ไม่เหมาะสมสำหรับการศึกษาข้อมูลในตลาดทุน ข้อมูลที่จะนำมาใช้ในงานวิจัย จึงควรเป็นประชากรทั้งหมดที่ต้องการศึกษา โดยผ่านกระบวนการคัดกรอง เพื่อให้ได้ประชากรที่เกี่ยวข้องกับเรื่องที่จะศึกษาเท่านั้น

ตัวอย่างการคัดกรองประชากรที่ใช้ในการศึกษาที่เห็นได้ชัดเจน ได้แก่ การเลือกประชากรที่มีข้อมูลที่คงการศึกษาอย่างครบถ้วน หรือการเลือกช่วงเวลาที่เกี่ยวข้องกับหัวข้อที่ทำการศึกษา เช่น ในการศึกษาผลของการเปลี่ยนแปลงมาตรฐานการบัญชีต่อตัวแปรอื่น ๆ ผู้วิจัยไม่จำเป็นต้องใช้ข้อมูลที่มีทั้งหมดตลอดระยะเวลาที่ตลาดทุนที่ทำการศึกษานั้น แต่ควรตัดตอนช่วงเวลาการศึกษาเฉพาะช่วงที่เกิดการเปลี่ยนแปลงมาตรฐาน (ไอซี) เพื่อมิให้ประชากรที่นำมามีผลกระทบจากปัจจัยอื่น เช่น สภาวะทางการเมืองและเศรษฐกิจ ปะปนลงด้วย ดังงานวิจัยของ Barth, Landsman, and Lang (2008) ซึ่งศึกษาความสัมพันธ์ระหว่างการใช้มาตรฐานการบัญชีระหว่างประเทศ (IAS) กับคุณภาพของข้อมูลทางบัญชี โดยเลือกประชากรที่ศึกษาจาก 21 ประเทศ แม้ว่าจะคณะกรรมการมาตรฐานการบัญชีระหว่างประเทศ (International Accounting Standard

Committee: IASC) จะได้เริ่มประกาศใช้มาตรฐานการบัญชีระหว่างประเทศมาตั้งแต่ปี ค.ศ. 1975 แต่องค์กรกำกับดูแลหลักทรัพย์ระหว่างประเทศ (International Organization of Security Commission: IOSCO) ได้แนะนำให้คณะกรรมการกำกับหลักทรัพย์ในประเทศต่าง ๆ ทั่วโลกอนุญาตให้บริษัทจดทะเบียนใช้มาตรฐานการบัญชีระหว่างประเทศ (IAS) สำหรับหลักทรัพย์ซื้อขายในตลาดต่างประเทศได้ตั้งแต่ปี ค.ศ. 2000 และภายในปี ค.ศ. 2005 ก็ได้มีการบังคับใช้มาตรฐานการบัญชีและการรายงาน (International Financial Reporting Standards: IFRS) กับบริษัทจดทะเบียนในประเทศต่าง ๆ ทั่วภาคพื้นยุโรปและในประเทศต่าง ๆ อย่งทั่วถึง ดังนั้น คณะผู้วิจัยจึงเลือกประชากรที่ศึกษาในช่วงปี ค.ศ. 1994-2003 เพื่อให้ประชากรที่ศึกษาอยู่ในช่วงเวลาที่มีการเปลี่ยนแปลงการบังคับใช้มาตรฐานการบัญชีระหว่างประเทศ (ปี ค.ศ. 2000) เท่าๆกันและเพื่อหลีกเลี่ยงจากผลกระทบด้านเศรษฐกิจสังคม และปัจจัยอื่น ๆ ที่ไม่เกี่ยวข้องกับเรื่องที่ต้องการศึกษาโดยตรง

ตัวอย่างการคัดกรองประชากรอีกตัวอย่างหนึ่ง

ที่เห็นได้ชัดเจนคือ งานของ Doyle, Ge และ McVay (2007) ซึ่งศึกษาข้อบ่งชี้ถึงความอ่อนแอของระบบการควบคุมภายในสำหรับการรายงานทางการเงินของกิจการที่จดทะเบียนในตลาดหลักทรัพย์ในประเทศสหรัฐอเมริกา ซึ่งต้องปฏิบัติตามกฎหมาย Sarbanes-Oxley Act ที่มีผลบังคับใช้ตั้งแต่ปี ค.ศ. 2002 ในงานวิจัยนี้คณะผู้วิจัยได้ตั้งสมมติฐานเกี่ยวกับปัจจัยที่อาจส่งผลให้ผู้บริหารกิจการต้องเปิดเผยข้อมูลเกี่ยวกับความอ่อนแอของระบบการควบคุมภายในด้านต่างๆ ของกิจการ โดยทำการทดสอบสมมติฐานว่ามีตัวแปรอิสระ 11 ตัวแปร ซึ่งเป็นตัวแทนของข้อบ่งชี้ในด้านต่างๆ เช่น ขนาด ความซับซ้อนของกิจการ ความเสี่ยงด้านการเงินและความเสี่ยงด้านการรายงาน เป็นต้น คณะผู้วิจัยจึงต้องรวบรวมข้อมูลของกิจการที่เปิดเผยมูลด้านการควบคุมภายในตามกฎหมาย โดยเริ่มตั้งแต่ปี ค.ศ. 2003 (เพื่อให้มั่นใจว่าบริษัทต่างๆ ที่เป็นประชากรที่ศึกษาได้เริ่มปฏิบัติ

ตามกฎหมายโดยครบถ้วนแล้ว) จนถึงปีล่าสุดที่ทำการศึกษาคือปี ค.ศ. 2005 แม้ว่าฐานข้อมูลตลาดทุนของประเทศสหรัฐอเมริกาจะมีข้อมูลทางการเงินของบริษัทที่จดทะเบียนในตลาดหลักทรัพย์ต่าง ๆ เริ่มตั้งแต่ปี ค.ศ. 1958 ก็ตาม เนื่องจากข้อมูลที่เกิดขึ้นก่อนหน้านั้น ถือว่าไม่เกี่ยวข้องกับเรื่องที่ทำการศึกษา นอกจากนี้ยังกำหนดให้กิจการที่นำมาศึกษาจะต้องมีข้อมูลทางการเงินที่นำมาใช้คำนวณค่าตัวแปรอิสระที่เกี่ยวข้อง ซึ่งได้มาจากฐานข้อมูลอื่น (ได้แก่ EDGAR และ Compustat) อย่างครบถ้วนเท่านั้น จึงทำให้มีจำนวนค่าสังเกต (Observations) ที่สามารถนำมาใช้ประชากรที่นำมาใช้ศึกษาได้ทั้งสิ้น 5,826 ค่าสังเกตเท่านั้น โดยแบ่งเป็นกิจการที่เป็นกลุ่มที่ใช้ทำการทดสอบจำนวน 779 ค่าสังเกต และกลุ่มควบคุมจำนวน 5,047 ค่าสังเกต

## 2. ปัญหาจากการมีตัวแปรอิสระที่มีความสัมพันธ์กันเอง (Correlated Variables)

ปัญหาจากการมีตัวแปรอิสระที่มีความสัมพันธ์กันเอง คือ การที่ตัวแปรอิสระตัวหนึ่ง มีความสัมพันธ์กับตัวแปรอื่นอื่นสูงมาก<sup>1</sup> จึงทำให้ตัวแปรอิสระมีค่าเปลี่ยนแปลงไปตามตัวแปรอิสระอื่น เป็นผลให้ไม่สามารถหาค่าที่แท้จริงของตัวแปรตามได้ และทำให้ตัวแบบ (Model) ที่ได้จากการศึกษาจะได้ผลลัพธ์ที่ไม่มีเสถียรภาพ เกิดปัญหาความเชื่อถือของตัวแบบ (Model Reliability) คือไม่สามารถนำ Model ที่ได้ไปใช้พยากรณ์ค่าตัวแปรตามได้<sup>2</sup>

นอกจากนี้ หากตัวแปรอิสระเป็นค่าคงที่ หมายถึงตัวแปรอิสระนั้นไม่ได้มีความสัมพันธ์กับตัวแปรตามแต่อย่างใดและไม่ควรนำมารวมในสมการแสดงความสัมพันธ์ระหว่างข้อมูลที่ต้องการศึกษาเนื่องจากตัวแปรอิสระดังกล่าว

ไม่สามารถช่วยทำให้อธิบายการเปลี่ยนแปลงค่าของตัวแปรตามได้เพิ่มขึ้นแต่อย่างใด ดังนั้น ผู้วิจัยจึงควรใช้วิธีการถดถอยและความเชี่ยวชาญในการคัดเลือกตัวแปรอิสระที่ใช้ในสมการถดถอยเชิงเส้นเพื่อการอธิบายความสัมพันธ์ของข้อมูลที่ศึกษา โดยให้แน่ใจว่าไม่มีตัวแปรอิสระที่มีความสัมพันธ์กันเองสูงเกินไป และในขณะเดียวกัน ตัวแปรอิสระที่เลือกมาใช้ขึ้นนี้ ก็เป็นตัวแปรที่มีอยู่จริงจะเป็นในการใช้อธิบายตัวแปรตามได้ โดยหากเลือกตัวแปรที่จะนำมาใช้ในตัวแบบสมการถดถอยนี้ จะอยู่ในขั้นตอนของการออกแบบงานวิจัย ผู้วิจัยอาจต้องใช้วิธีการลองผิด-ลองถูก (Trial and Error) เพื่อคัดสรรหาตัวแปรอิสระที่เหมาะสม เพื่อใช้อธิบายการเปลี่ยนแปลงของตัวแปรตามได้อย่างสมเหตุสมผล และให้ตัวแบบสมการถดถอยมีรูปแบบที่กระชับและไม่เยิ่นเย้อจนเกินไปจนจำเป็น

## 3. การมีตัวแปรในสมการตัวแบบไม่ครบถ้วน (Omitted Variables)

การมีตัวแปรอิสระที่สำคัญในสมการถดถอยไม่ครบถ้วน เกิดจากการที่ผู้วิจัยไม่ได้ใส่ตัวแปรอิสระที่สำคัญที่ใช้ในการอธิบายการเปลี่ยนแปลงของตัวแปรตามลงในสมการถดถอย ซึ่งอาจเกิดจากการที่ผู้วิจัยไม่ได้นึกถึงตัวแปรอิสระนั้น หรือไม่คิดว่าตัวแปรอิสระนั้นจะมีความเกี่ยวข้องในการใช้อธิบายการเปลี่ยนแปลงค่าของตัวแปรตาม หรือไม่สามารถหาวิธีที่เหมาะสมที่ใช้จะวัดค่าตัวแปรอิสระ หรือตัวแปรอิสระนั้น ๆ ไม่สามารถวัดค่าเป็นตัวเลขได้

การละตัวแปรอิสระที่สำคัญออกไปจากตัวแบบ แสดงความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามนั้น ส่งผลให้ค่าสัมประสิทธิ์ที่คำนวณได้จากสมการถดถอยเป็นค่าที่

<sup>1</sup> ดูได้จากค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) ระหว่างตัวแปรสองตัว โดยตามหลักทั่วไป จะถือว่า ตัวแปรมีความสัมพันธ์กันสูงเมื่อค่าสัมประสิทธิ์สหสัมพันธ์เกิน 0.8

<sup>2</sup> อาจสังเกตว่าการมีตัวแปรอิสระที่มีความสัมพันธ์กันเองสูงรวมอยู่ด้วยหรือไม่ ได้หลายวิธี เช่น มีตัวแปรบางตัวที่มีค่า t-statistic ต่ำ แต่ค่าการถดถอยโดยรวมมีค่า R<sup>2</sup> สูงมาก หรือการเปลี่ยนจำนวนตัวอย่างเป็นเพียงเล็กน้อยทำให้ค่าสัมประสิทธิ์ความสัมพันธ์ของตัวแปรอิสระในสมการถดถอยมีค่าเปลี่ยนแปลงไปอย่างมาก เป็นต้น

ไม่เที่ยงตรง (Biased) ดังตัวอย่างต่อไปนี้

สมมติให้ตัวแบบที่ใช้อธิบายความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม คือ

$$\text{สมการ (1) : } Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

แต่ผู้วิจัยใช้สมการในการประมาณค่าตัวแปรตามดังนี้

$$\text{สมการ (2) : } Y = \gamma_0 + \gamma_1 X_1 + v$$

นั่นคือ ผู้วิจัยได้ละตัวแปร  $X_2$  ออกไปจากสมการ โดยจะตั้งใจหรือไม่ตั้งใจก็ตาม จึงทำให้สัมประสิทธิ์ความสัมพันธ์ระหว่างตัวแปรอิสระ  $X_1$  กับตัวแปรตาม  $Y$  ที่คำนวณได้ มีค่าเท่ากับ  $\gamma_1$  แทนที่จะเท่ากับ  $\beta_1$  ซึ่งทำให้ค่าสัมประสิทธิ์ความสัมพันธ์ที่คำนวณได้ ( $\gamma_1$ ) นี้ มีค่าคลาดเคลื่อนจากค่าสัมประสิทธิ์ที่แท้จริง ( $\beta_1$ ) เท่ากับ  $\beta_2 + \delta$  โดยที่  $\delta$  คือ ค่าสัมประสิทธิ์ความแปรปรวนระหว่างตัวแปร  $X_1$  และ  $X_2$  ( $cov_{x_1, x_2}$ ) นั่นเอง

#### 4. การสรุปผลการทดสอบสมมติฐานผิดพลาด

##### (False conclusion)

ผู้วิจัยจะต้องการทดสอบสมมติฐานว่า โยเบกจิ ในการทดสอบความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามนั้น ผู้วิจัยมักมีเป้าหมายในการทดสอบว่าตัวแปรตาม ( $Y$ ) มีความสัมพันธ์กับตัวแปรอิสระ ( $X$ ) นั่นคือ  $\beta_i \neq 0$  แต่ในการทดสอบสมมติฐานผู้วิจัยจะต้องใช้ความระมัดระวังในการตั้งสมมติฐาน โดยตั้งสมมติฐานว่า

$$H_0: \beta_i = 0 \quad \text{และ}$$

$$H_a: \beta_i \neq 0$$

เนื่องจาก หาก  $\beta_i = 0$  แสดงว่าตัวแปรตาม ( $Y$ ) ไม่ได้มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรอิสระ ( $X$ ) และหากผลการทดสอบสมมติฐานสรุปว่าไม่สามารถปฏิเสธสมมติฐาน  $H_0$  (Fail to Reject  $H_0$ ) จะต้องรายงานผลทางสถิติว่าตัวแปรต้นและตัวแปรตามไม่มีความสัมพันธ์กัน

อย่างมีนัยสำคัญทางสถิติ ทั้งนี้ในการวิจัยทางบัญชีโดยทั่วไป จะมีการตั้งระดับนัยสำคัญ (Significant Level:  $\alpha$ ) เท่ากับ 0.05 ดังนั้น หากมีการปฏิเสธสมมติฐานศูนย์ จะสรุปได้ว่าค่าสัมประสิทธิ์ความสัมพันธ์  $\beta_i$  มีค่าไม่เท่ากับศูนย์ (ค่าตัวแปรตามมีความสัมพันธ์กับตัวแปรต้น) ระดับความเชื่อมั่นร้อยละ 95 ( $1 - \alpha$ )

ในทางปฏิบัติ โปรแกรมคอมพิวเตอร์ทางสถิติมักรายงานค่า p-value ซึ่งคือค่าความน่าจะเป็นที่จะปฏิเสธสมมติฐานศูนย์เมื่อสมมติฐานนั้นเป็นความจริง ดังนั้น หากค่า p-value มีค่าน้อยกว่าค่าที่กำหนดไว้สำหรับระดับนัยสำคัญ ( $\alpha$ ) ก็จะแสดงว่าโอกาสที่จะปฏิเสธสมมติฐานศูนย์เมื่อสมมติฐานนั้นเป็นจริง (Type I Error) ก็จะน้อยลง และความเชื่อมั่นว่าสมมติฐานทางเลือก (Alternative Hypothesis:  $H_a$ ) เป็นจริงจะอยู่ในระดับสูง

นอกจากนี้ผู้วิจัยส่วนใหญ่ไม่ได้ให้ความสำคัญกับการควบคุมทางสถิติ เพื่อป้องกันปัญหาการไม่ปฏิเสธสมมติฐานหลัก ( $H_0$ ) ในขณะที่สมมติฐานหลักนั้นไม่เป็นจริง (Type II Error) นั่นคือ ไม่ได้ปฏิเสธว่าตัวแปรตามไม่ได้มีความสัมพันธ์กับตัวแปรอิสระ ในขณะที่ความจริงแล้วตัวแปรทั้งสองนั้นมีความสัมพันธ์กันจริง (False Negative) ทั้งนี้ ผู้วิจัยควรใช้ความระมัดระวังในการตั้งสมมติฐานเกี่ยวกับความสัมพันธ์ของตัวแปร และให้ความสำคัญและหาทางป้องกันข้อผิดพลาดในการสรุปผลทางสถิติทั้งสองแบบโดยเท่าเทียมกัน

#### ปัญหาเกี่ยวกับการรายงานผลการศึกษา

การรายงานผลการศึกษาด้วยค่าทางทดสอบทางสถิติที่มีนัยสำคัญอาจเป็นสิ่งที่นักวิจัยปรารถนา เนื่องจากค่าสัมประสิทธิ์ที่มีนัยสำคัญทางสถิตินั้นใช้เป็นหลักฐานในการตอบข้อสมมติฐานงานวิจัยตั้งขึ้น อย่างไรก็ตาม สิ่งที่มีความสำคัญมากกว่าความมีนัยสำคัญทางสถิติ คือการพบค่าสัมประสิทธิ์ที่ไม่มีนัยสำคัญทางสถิติ อาจมีความสำคัญและเป็นจุดกำเนิดงานวิจัยต่อเนื่องในอนาคต ซึ่งทำให้งานวิจัยดังกล่าวมีคุณค่ามากกว่างานวิจัยที่รายงาน

ค่าสัมประสิทธิ์ที่มีนัยสำคัญทางสถิติที่ไม่สามารถนำไปใช้ประโยชน์ต่อไปในเชิงเศรษฐกิจต่อไปได้ (Ziliak and McCloskey:2004)

นอกจากความมีนัยสำคัญเชิงเศรษฐกิจนี้แล้ว ความสามารถในการแปลความหมายของตัวเลขต่าง ๆ ที่แสดงในตารางแสดงผลของงานวิจัยเป็นเรื่องราวที่ทำให้ผู้อ่านงานวิจัยเกิดความเข้าใจและตระหนักถึงความสำคัญของข้อค้นพบจากเรื่องที่ทำการศึกษา คำอธิบายผลการวิจัยโดยเน้นถึงผลกระทบเชิงเศรษฐกิจที่มีน้ำหนัก และมีทฤษฎีหรือผลจากงานวิจัยอื่นที่เกี่ยวข้องรองรับอย่างสมเหตุสมผลย่อมเป็นสิ่งสำคัญที่ทำให้งานวิจัยน่าเชื่อถือ ทำให้ผู้อ่านสามารถนำผลงานวิจัยไปใช้ประโยชน์ต่อไปได้ ส่งผลให้งานวิจัยมีคุณค่ายิ่งขึ้น ดังตัวอย่างงานวิจัยของ Ogneva และคณะ (2007) และ Ashbaugh-Skaife และคณะ (2009) ซึ่งศึกษาความสัมพันธ์ระหว่างการรายงานข้อบกพร่องในการควบคุมภายในกับต้นทุนในการลงทุน (Cost of Equity) และได้ผลที่ไม่สามารถใช้เป็นข้อสรุปได้ Gordon and Wilford (2012) จึงได้ทำการศึกษาในเรื่องที่คล้ายคลึงกันอีกครั้งหนึ่งโดยการใช้ข้อมูลการเปิดเผยข้อบกพร่องในการควบคุมภายในแบบต่อเนื่อง และได้ผลการศึกษาที่เป็นไปในทิศทางเดียวกับงานของของ Ashbaugh-Skaife และคณะ (2009) ในกรณีนี้ คณะผู้วิจัยยังได้กล่าวถึงข้อจำกัดและข้อเสนอแนะอื่น ๆ เช่น การใช้ระบบการควบคุมภายในโดยใช้เทคโนโลยี (Technology-based Internal Control Monitoring System) ซึ่งเป็นประเด็นที่ไม่เคยมีการกล่าวถึงในงานวิจัยก่อนหน้านี้ไว้ได้อย่างน่าสนใจและสามารถใช้เป็นแนวทางในการวิจัยเพิ่มเติมในอนาคตสำหรับผู้สนใจและมีความสามารถด้านเทคโนโลยีอีกด้วย

## สรุป

ในการวิเคราะห์ความสัมพันธ์ของตัวแปรตามในงานวิจัยทางบัญชี ซึ่งใช้ข้อมูลทางการเงินจากตลาดทุนหรือจากแหล่งข้อมูลอื่นที่เกี่ยวข้อง ผู้วิจัยควรระวังในความระมัดระวังในการใช้วิธีวิจัยที่เหมาะสม รวมถึงการแปลผลข้อมูลให้ถูกต้องตามหลักสถิติ ตรรก และความมีนัยสำคัญเชิงเศรษฐกิจ ทั้งนี้ เนื่องจากวิธีวิเคราะห์ข้อมูลโดยใช้สมการถดถอยเชิงเส้นแบบหลายตัวแปร เป็นวิธีการที่นิยมนั้นมีข้อจำกัดเกี่ยวกับสมมติฐานในการจัดเก็บข้อมูลและคุณสมบัติของข้อมูลที่ใช้อย่างมาก ในขณะที่ข้อมูลทางการเงินที่ได้รับจากตลาดทุนมีลักษณะเฉพาะที่อาจไม่เป็นตามข้อสมมติฐานดังกล่าว

ในการออกแบบสมการถดถอยเชิงเส้นที่ใช้อธิบายความสัมพันธ์ของตัวแปรอิสระและตัวแปรตามนั้น ผู้วิจัยยังต้องระวังความระมัดระวังตั้งแต่ขั้นตอนการตั้งสมมติฐาน การเลือกทรงประชากรและกลุ่มตัวอย่าง หรือการคัดเลือกตัวแปรที่เกี่ยวข้อง เพื่อลดปัญหาอันจะทำให้ลดความน่าเชื่อถือของสมการถดถอย หรือทำให้สมการมีความซับซ้อนเกินความจำเป็นจนทำให้ใช้อธิบายการเปลี่ยนแปลงของตัวแปรตามได้ยากขึ้น

นอกจากข้อควรระวังในการวิเคราะห์เชิงสถิติแล้ว ในการสรุปผลการศึกษาผู้วิจัยควรตระหนักถึงทั้งนัยสำคัญเชิงสถิติและนัยสำคัญเชิงเศรษฐกิจ รวมถึงการอธิบายผลการศึกษาโดยให้มีความสมเหตุสมผล มีน้ำหนักรองรับหรือสามารถอ้างอิงถึงทฤษฎี หรือเปรียบเทียบกับงานวิจัยอื่นที่เกี่ยวข้องได้ อันจะทำให้ผลการศึกษาเป็นที่น่าเชื่อถือและมีคุณค่าต่อการนำไปใช้อ้างอิงต่อไปได้

## เอกสารอ้างอิง

- Ashbaugh-Skaife, H., D. W. Collins, W.R. Kinney, and R. Lafond (2009). The effect of SOX internal control deficiencies on firm risk and cost of equity. *Journal of Accounting Research*, 47(1), 1–43.
- Barth, M. E., Landsman, W. R., & Lang, M. H. (2008). International accounting standards and accounting quality. *Journal of accounting research*, 46(3), 467–498.
- Doyle, J., Ge, W., and McVay, S. (2007). Determinants of weakness in internal control over financial reporting. *Journal of Accounting & Economics*, 44(2007). 193–223.
- Gordon, L., A. and A.L. Wilford (2012). An Analysis of Multiple Consecutive Years of Material Weaknesses in Internal Control. *The Accounting Review*. 87(6), 2027–2060.
- Kutner, M. H., Neter, J., Nachtsheim. C. J. (2004). *Applied Linear Regression Models*. Fourth edition. New York: McGraw-Hill.
- Ogneva, M., K. R. Subramanyam, and K. Raghunandan (2007). Internal control weakness and cost of equity: Evidence from SOX section 404 disclosures. *The Accounting Review*, 82(5), 1255–1297.
- Snedecor, George W. and Cochran, William G. (1989). *Statistical Methods*. Eighth Edition. Iowa State University Press.
- Wooldridge, Jeffrey M. (2009). *Introductory Econometrics: A Modern Approach*. 4e. Mason, OH: South-Western CENGAGE Learning.
- Ziliak, S. T., and D. McCloskey (2004). Size matters: The standard error of regressions in the American Economic Review. *The Journal of Socio-Economics*, 33(5), 527–546.