

บทที่ 1

บทนำ

1.1 ความสำคัญของปัญหา

การวิเคราะห์การถดถอยเป็นเทคนิคการพยากรณ์ ที่นิยมใช้กันอย่างกว้างขวางมากที่สุดในสาขาวิชาต่างๆ โดยใช้ตัวแปรอิสระ (Independent Variable) หรือตัวแปรพยากรณ์ (Predictor) อย่างน้อยหนึ่งตัวไปพยากรณ์ หรืออธิบายความผันแปรของตัวแปรตาม (Dependent Variable) ถ้าการวิเคราะห์การถดถอยประกอบด้วยตัวแปรอิสระหนึ่งตัว ซึ่งวัดเป็นค่าในเชิงปริมาณ เรียกว่า การวิเคราะห์การถดถอยอย่างง่าย (Simple regression analysis) ส่วนกรณีที่มีตัวแปรอิสระมากกว่าหนึ่งตัว จะเรียกว่า การวิเคราะห์การถดถอยแบบพหุ (Multiple regression analysis)

การวิเคราะห์การถดถอยส่วนหนึ่งจะเกี่ยวข้องกับการศึกษาว่า ความผันแปรของตัวแปรตามในรูปแบบการถดถอยต่างๆ เป็นผลเนื่องมาจากปัจจัยใดบ้าง ปัจจัยที่นำมาศึกษาจะแทนด้วยตัวแปรอิสระ ตัวแปรอิสระที่ใช้กันมากมีลักษณะตามชื่อของตัวแปร ได้แก่ ตัวแปรปัจจัย (factor variable) ตัวแปรเวลา (time variable) และตัวแปรดัมมี่ (dummy variable) โดยมีรายละเอียดดังนี้

- ตัวแปรปัจจัย จะแทนปัจจัยต่างๆ ที่มีอิทธิพลต่อตัวแปรตาม เช่น ยอดขายของบริษัทจะมากหรือน้อยจะขึ้นอยู่กับปัจจัยต่างๆ ได้แก่ ยอดโฆษณา ราคาสินค้าที่ขาย และราคาสินค้าของบริษัทคู่แข่ง ผลผลิตข้าวจะมากหรือน้อยขึ้นอยู่กับปัจจัยต่างๆ ได้แก่ ปริมาณน้ำฝน ปริมาณปุ๋ยที่ใช้ และคุณภาพของดิน จำนวนสินค้าที่มีรอยตำหนิที่ผลิตจากโรงงานแห่งหนึ่งจะมากหรือน้อยขึ้นอยู่กับปัจจัยต่างๆ ได้แก่ เครื่องจักรที่ใช้ผลิต ระดับความสามารถของผู้ควบคุมเครื่องจักร เงินเดือนของพนักงานในองค์กรแห่งหนึ่งจะมากหรือน้อยขึ้นอยู่กับ เพศ อายุ ระดับการศึกษาจำนวนวันลาในปีที่ผ่านมา เป็นต้น

- ตัวแปรเวลา เป็นกรณีที่ศึกษาการเคลื่อนไหวของอนุกรมเวลา ซึ่งส่วนใหญ่จะเป็นข้อมูลทางธุรกิจหรือทางเศรษฐกิจ มักจะเป็นการศึกษาแนวโน้มซึ่งเป็นการเคลื่อนไหวของค่าสังเกตตามเวลาในระยะยาว รูปแบบแนวโน้มที่สร้างขึ้นจะกำหนดตัวแปรเวลาเป็นตัวแปรอิสระที่มีค่าเป็นรหัส (coded value) แทนวัน เดือน ไตรมาส หรือปี ที่เก็บรวบรวมข้อมูลมา

- ตัวแปรคัมมี เป็นตัวแปรที่สร้างขึ้นเพื่อระบุกลุ่มหรือชุดที่ค่าสังเกตนั้นอยู่ จะกำหนดค่าของตัวแปรคัมมีเป็น 1 หรือ 0 ซึ่งตัวแปรคัมมีจะมีค่าเท่ากับ 1 เมื่อค่าสังเกตนั้นอยู่ในกลุ่มที่สนใจ และมีค่าเป็น 0 เมื่อค่าสังเกตนั้นไม่อยู่ในกลุ่มที่สนใจ

ในการเขียนสมการถดถอยเพื่อประมาณค่าตัวแปรตาม ส่วนใหญ่จะให้ความสำคัญกับตัวแปรเชิงปริมาณหรือตัวแปรปัจจัยแล้ว ตัวแปรเชิงคุณภาพหรือตัวแปรคัมมีก็มีความสำคัญไม่น้อยไปกว่ากัน แต่ในการวิเคราะห์การถดถอยโดยทั่วไปจะเน้นศึกษาเฉพาะตัวแปรปัจจัย เพราะสามารถวัดค่าออกมาได้ ทำให้ง่ายต่อการนำไปวิเคราะห์ และมองข้ามตัวแปรคัมมีซึ่งอาจมีอิทธิพลต่อตัวแปรตามเช่นกัน ส่วนตัวแปรคัมมีในการวิเคราะห์การถดถอยจะถือเป็นตัวแปรจัดประเภท (Categorical Variable) เพื่อใช้ในการระบุกลุ่มของข้อมูล

จากแนวคิดในการวิเคราะห์การถดถอย ผู้วิจัยจึงสนใจที่จะนำเอาตัวแปรคัมมีมาใช้ในสมการถดถอยในการประมาณคะแนนสอบปลายภาคของนักศึกษา โดยอาศัยในการเก็บรวบรวมข้อมูลจากคะแนนสอบกลางภาคและเกรดเฉลี่ยสะสม (Grade Point Average หรือ GPA) ของนักศึกษา โดยแบ่งข้อมูลออกเป็นกลุ่มตาม เพศ และสาขาวิชาที่นักศึกษากำลังศึกษา

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาตัวแปรคัมมีที่ใช้ในรูปแบบการถดถอย
2. เพื่อศึกษาความสามารถของตัวแปรคัมมีในการประมาณคะแนนสอบปลายภาคของนักศึกษา

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถใช้เป็นแนวทางในการศึกษาตัวแปรคัมมีในรูปแบบการถดถอย
2. ทำให้ทราบถึงความสำคัญของตัวแปรคัมมีในการประมาณคะแนนสอบปลายภาคของนักศึกษา
3. นำผลการวิจัยที่ได้มาถ่ายทอดองค์ความรู้ให้กับนักศึกษาได้

1.4 ขอบเขตของงานวิจัย

งานวิจัยครั้งนี้จะทำการศึกษาในการแบ่งกลุ่มข้อมูลโดยอาศัยตัวแปรคัมมีในรูปแบบการถดถอย เพื่อนำมาใช้ในการประมาณคะแนนสอบปลายภาคของนักศึกษา จากการเก็บรวบรวมข้อมูลได้จากนักศึกษา สาขาวิชาสถิติประยุกต์ และสาขาวิชาจุลชีววิทยาอุตสาหกรรมหรือจุลชีววิทยาในคณะวิทยาศาสตร์ และนักศึกษาสาขาวิชาสัตวศาสตร์ในคณะเทคโนโลยีการเกษตร เพื่อให้คะแนนสอบปลายภาคและคะแนน

สอบกลางภาคที่ใช้ในการวิเคราะห์ในแต่ละสาขาวิชามีความน่าเชื่อถือ จึงทำการหาคุณภาพของข้อสอบประกอบในการวิเคราะห์