

**ข้อสอบปลายภาค**  
**วิชา Research Methodology in Finance : MGMG 522**  
**ภาคเรียนที่ 3 ปี 2547**

ผู้สอน ดร.ชาญ สรณาคมน์

(คะแนนสอบปลายภาค = 30% ของคะแนนรวม)  
**(ข้อสอบมีทั้งหมด 4 ข้อ ให้ทำทุกข้อ แต่ละข้อมีคะแนนเท่ากัน)**

**คำสั่ง**

1. อนุญาตให้นิสิตนำตำรา เอกสารอื่น ๆ และเครื่องคิดเลขเข้าห้องสอบได้ แต่ไม่อนุญาตให้หยิบยืมสิ่งต่าง ๆ เหล่านั้นระหว่างกันในช่วงการสอบ
2. ห้ามใช้ฟังก์ชันเครื่องคิดเลขในเครื่องโทรศัพท์เคลื่อนที่
3. ไม่อนุญาตให้ใช้เครื่องมือสื่อสารทุกชนิดระหว่างการสอบ (ห้ามใช้ PDA และ Notebook ทุกชนิดด้วย)
4. อ่านโจทย์ให้ดี แล้วตอบให้ตรงประเด็นคำถาม
5. พยายามอย่าปล่อยคำตอบให้ว่าง
6. ตรวจคำตอบให้ดีก่อนส่งกระดาษคำตอบ
7. ห้ามออกนอกห้องสอบระหว่างการสอบไม่ว่ากรณีใด ๆ ทั้งสิ้น ยกเว้นแต่จะได้ส่งคืนกระดาษคำตอบและสละสิทธิ์ที่จะทำข้อสอบต่อ

ข้อที่ 1. ปัญหา Measurement Error ในตัวแปรต้น (กล่าวคือ ค่าของตัวแปรต้นที่เก็บมาได้ไม่ใช่ค่าที่แท้จริงแต่มีความคลาดเคลื่อนของข้อมูล) มีความคล้ายคลึงกับปัญหาในกรณีรัน Ordinary Least Squares (ไม่ใช่ Two-Stage Least Squares) กับสมการโครงสร้าง (Structural Equations) คือ สัมประสิทธิ์ที่ได้จากทั้งสองกรณีจะ biased จงอธิบายโดยละเอียดว่า bias ในค่าสัมประสิทธิ์เกิดขึ้นได้เพราะอะไร?

#### แนวทางการตอบ

ปัญหา Measurement Error ในตัวแปรต้นและปัญหาในกรณีรัน OLS กับสมการโครงสร้างก่อให้เกิด bias ในค่า coefficient estimate เพราะ error term จะ correlate กับตัวแปรต้น ซึ่งจะ violate classical assumption #3

ควรเขียนแสดงสมการโครงสร้างและสมการที่ตัวแปรต้นมีปัญหาความคลาดเคลื่อนในการวัดประกอบด้วย พร้อมทั้งอธิบายประกอบโดยชี้ให้อาจารย์เห็นว่าทุกครั้งที่ค่า error term เปลี่ยนไป ตัวแปรต้นก็เปลี่ยนตามไปด้วยเสมอ แสดงให้เห็นถึง correlation ระหว่างตัวแปรต้นกับ error term

ข้อที่ 2. จงตอบคำถามต่อไปนี้

- ตัวสถิติ Durbin-Watson  $h$  ใช้ทดสอบอะไร?
- เมื่อไรเราจึงจะใช้ Durbin-Watson  $h$  test?
- แล้วทำไมเราจะไม่ใช้ Durbin-Watson  $d$  test ไม่ได้?
- Null hypothesis ของ Durbin-Watson  $h$  test คืออะไร? แล้วถ้าเราปฏิเสธ null hypothesis เราจะพอสสรุปได้ว่าอย่างไร?

#### แนวทางการตอบ

- ตัวสถิติ Durbin-Watson  $h$  ใช้ทดสอบ first-order serial correlation หรือ first-order autocorrelation ในสมการ regression ที่มีตัวแปรตาม ณ เวลาที่แล้ว ( $Y_{t-1}$ ) เป็นหนึ่งในตัวแปรต้นในสมการ regression
- เราจะไม่ใช้ Durbin-Watson  $h$  test เมื่อเราต้องการทดสอบหา first-order autocorrelation กรณีที่สมการ regression ของเรามีตัวแปร  $Y$  ณ เวลา  $t$  ( $Y_t$ ) อยู่ทางซ้ายมือ และมีตัวแปร  $Y$  ณ เวลา  $t-1$  ( $Y_{t-1}$ ) กับตัวแปรต้นตัวอื่น ๆ อยู่ทางขวามือของสมการ regression
- เราใช้ Durbin-Watson  $d$  test ไม่ได้เพราะถ้าเมื่อใดก็ตามเรามีตัวแปรล่าของตัวแปรตาม ( $Y_{t-1}$ ) อยู่ทางขวามือของสมการ regression ค่า Durbin-Watson  $d$  ที่ได้จาก output ของโปรแกรมสถิติจะมีค่าเข้าใกล้ 2 ซึ่งเราอาจจะสรุปผิดพลาดว่าไม่มีปัญหา first-order autocorrelation ในสมการ regression ของเรา เราจึง

จำเป็นต้องปรับค่า Durbin-Watson  $d$  ให้เป็นค่า Durbin-Watson  $h$  แล้วทดสอบสมมุติฐานด้วย Durbin-Watson  $h$  ว่าปัญหา autocorrelation ยังคงมีอยู่ในสมการ regression ของเราหรือไม่

- d.) Null hypothesis ของ Durbin-Watson  $h$  test คือไม่มีปัญหา first-order autocorrelation ในสมการ regression ของเราที่มีตัวแปรล่าของตัวแปรตามเป็นหนึ่งในตัวแปรต้น และถ้าเราปฏิเสธ null hypothesis เราจะพอสรุปได้ว่าโมเดลของเรามีปัญหา autocorrelation ในระดับ first-order

ข้อที่ 3. ในการเปิดหลักสูตรปริญญาเอกสาขาบริหารธุรกิจของมหาวิทยาลัยแห่งหนึ่งในประเทศไทย มีผู้สมัครเข้าศึกษาต่อ 30 คน มีผู้ได้เข้าศึกษาต่อ 8 คน นี่เป็นผลที่ได้จากการรันสมการ Linear Probability Model (ใช้ OLS แบบธรรมดา)

Dependent Variable: Y Method: Least Squares Date: 05/20/05 Time: 19:58 Sample: 1 30 Included observations: 30				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.299683	0.821531	-0.364786	0.7182
X1	-0.065599	0.246916	-0.265675	0.7926
X2	0.295805	0.126828	2.332332	0.0277
X3	0.010653	0.002831	3.762738	0.0009
R-squared	0.488580	Mean dependent var	0.266667	
Adjusted R-squared	0.429569	S.D. dependent var	0.449776	
S.E. of regression	0.339702	Akaike info criterion	0.802070	
Sum squared resid	3.000333	Schwarz criterion	0.988896	
Log likelihood	-8.031046	F-statistic	8.279598	
Durbin-Watson stat	0.664140	Prob(F-statistic)	0.000496	

โดยที่ตัวแปรต่าง ๆ ข้างต้นมีความหมายดังนี้

- Y = 1 ถ้าผู้สมัครได้เข้าศึกษาต่อ, = 0 หากผู้สมัครถูกปฏิเสธ
- X1 = คะแนน GPA ตอนจบปริญญาโทของผู้สมัคร (บน scale จาก 0.00 ถึง 4.00)
- X2 = 1 ถ้าผู้สมัครจบปริญญาตรีสายวิทยาศาสตร์ (วิศวกรรมศาสตร์ แพทยศาสตร์ เป็นต้น), = 0 ถ้าจบปริญญาตรีสายอื่น ๆ ที่ไม่ใช่สายวิทยาศาสตร์
- X3 = คะแนนสอบข้อเขียนตามแนว GMAT ของทางมหาวิทยาลัยแห่งนี้ (บน scale จาก 0 ถึง 100)

จงตอบคำถามต่าง ๆ ต่อไปนี้

- A.) จงประเมินสมการ regression ดังกล่าว ว่าในภาพรวมโมเดลนี้ดีหรือไม่อย่างไร?
- B.) จงประเมินว่าตัวแปรต้นแต่ละตัวมีนัยสำคัญหรือไม่อย่างไร? สอดคล้องกับความเป็นจริงหรือทฤษฎีหรือไม่อย่างไร?
- C.) จงพยากรณ์ว่าถ้าผู้สมัครเข้าศึกษาต่อคนหนึ่งได้คะแนนสอบข้อเขียน 40 คะแนน จบปริญญาตรีคณะอักษรศาสตร์ด้วยเกรดเฉลี่ย 2.00 จบปริญญาโทสาขาการตลาดด้วยเกรดเฉลี่ย 3.00 จะมีโอกาสเข้าศึกษาต่อปริญญาเอกมากน้อยแค่ไหน?

#### แนวทางการตอบ

- A.)  $\text{Prob}(F\text{-statistic}) < 0.05$  แสดงว่า โดยรวม ตัวแปรต้นทั้งหมดประกอบกันอธิบายตัวแปรตามได้ ค่า Adjusted R-squared บอกว่า ตัวแปรต้นทั้งหมดประกอบกันอธิบายตัวแปรตามได้ประมาณ 43 เปอร์เซ็นต์ ซึ่งจัดอยู่ในเกณฑ์พอใช้ (หนึ่ง ต้องพึงระลึกเสมอว่า ค่า R-squared ที่ได้จาก OLS ที่มีตัวแปรตามเป็นตัวแปรหุ่นที่มีค่าได้ 2 ค่าจะต่ำกว่าความเป็นจริงเสมอและมีค่ามากที่สุดไม่เกิน 0.7)
- B.) ตัวแปรต้น  $X_3$  และ  $X_2$  มีนัยสำคัญทางสถิติที่ระดับ 0.05 ซึ่งค่อนข้างสอดคล้องกับความเป็นจริงที่ว่าคะแนนสอบเข้าที่สูงขึ้นจะเพิ่มโอกาสในการเข้าศึกษาต่อ และผู้สมัครที่จบสายวิทยาศาสตร์ก็น่าจะมีโอกาสเข้าศึกษาต่อมากกว่าผู้จบสายอื่น ๆ แต่ตัวแปรต้น  $X_1$  มีสัมประสิทธิ์เป็นลบ ซึ่งแปลว่า GPA ตอนจบปริญญาโทของผู้สมัครที่เพิ่มขึ้นจะทำให้โอกาสเข้าศึกษาต่อนั้นน้อยลง ซึ่งค่อนข้างค้านกับความเป็นจริง แต่ว่าความสัมพันธ์เชิงลบดังกล่าวไม่มีนัยสำคัญ ณ ระดับ 0.05
- C.) แทนค่า  $X_1 = 3, X_2 = 0, X_3 = 40$ , จะได้  $Y = -0.07036$  โอกาสความน่าจะเป็นในการเข้าศึกษาต่อของผู้สมัครรายนี้มีค่าน้อยกว่าศูนย์ แต่นี้เป็นปัญหาของ Linear Probability Model เพราะเรารู้อยู่แล้วว่าความน่าจะเป็นของการเกิดเหตุการณ์ใดก็แล้วแต่ไม่มีทางน้อยกว่าศูนย์ (prob. = 0 คือ เหตุการณ์นั้นไม่มีทางเกิดขึ้นได้) และไม่มีทางเกินหนึ่ง (prob. = 1 คือ เหตุการณ์ดังกล่าวเกิดขึ้นแน่นอนร้อยเปอร์เซ็นต์) ดังนั้นเราพอจะสรุปได้ว่า ผู้สมัครคนนั้นมีโอกาสเข้าศึกษาต่อ 0 เปอร์เซ็นต์ คือ ไม่มีโอกาสเลยนั่นเอง

ข้อที่ 4. ต่อเนื่องจากคำถามข้อที่ 3 ถ้านำข้อมูลชุดเดิมไปรัน regression โดยใช้แบบจำลอง Probit แทนจะได้ผลลัพธ์ดังนี้

Dependent Variable: Y				
Method: ML - Binary Probit (Quadratic hill climbing)				
Date: 05/20/05 Time: 19:58				
Sample: 1 30				
Included observations: 30				
Convergence achieved after 8 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-9.978405	10.60789	-0.940659	0.3469
X1	-2.360045	3.083064	-0.765487	0.4440
X2	2.177406	1.441501	1.510513	0.1309
X3	0.213027	0.101414	2.100564	0.0357
Mean dependent var	0.266667	S.D. dependent var		0.449776
S.E. of regression	0.222396	Akaike info criterion		0.539549
Sum squared resid	1.285960	Schwarz criterion		0.726375
Log likelihood	-4.093235	Hannan-Quinn criter.		0.599316
Restr. log likelihood	-17.39746	Avg. log likelihood		-0.136441
LR statistic (3 df)	26.60844	McFadden R-squared		0.764722
Probability(LR stat)	7.11E-06			
Obs with Dep=0	22	Total obs		30
Obs with Dep=1	8			

จงตอบคำถามต่าง ๆ ต่อไปนี้

- A.) จงเปรียบเทียบโมเดล Probit นี้กับ Linear Probability Model ก่อนหน้านี้โดยพิจารณาระดับนัยสำคัญของตัวแปรต้นเป็นหลัก (ไม่ต้องเปรียบเทียบในภาพรวม โมเดลทั้งสองมีนัยสำคัญทั้งคู่)
- B.) จงพยากรณ์อีกครั้งภายใต้ Probit Model ว่าผู้สมัครคนเดิมจะมีโอกาสเข้าศึกษาต่อมากน้อยแค่ไหน?

แนวทางการตอบ

- A.) ตัวแปรต้น X1 ยังคงไม่มีนัยสำคัญเหมือนเดิมและความสัมพันธ์กับตัวแปรตามก็ยังคงเป็นเชิงลบเช่นเดิม ตัวแปรต้น X3 ยังคงมีความสัมพันธ์เชิงบวกกับตัวแปรตามและมีนัยสำคัญเช่นเดิม แต่ตัวแปรต้น X2 ถึงแม้จะมีความสัมพันธ์เชิงบวกกับตัวแปรตามดังความคาดหมาย แต่ระดับนัยสำคัญลดลงจนอยู่ในระดับที่ไม่มีนัยสำคัญทางสถิติ (ณ ระดับ 0.05)
- B.) แทนค่า  $X1 = 3$ ,  $X2 = 0$ ,  $X3 = 40$ , จะได้ค่า -8.53746 ซึ่งตัวเลขดังกล่าวเป็นค่า Z-score ต้องนำไปเปิดตารางหาพื้นที่ใต้โค้งของการแจกแจงปกติมาตรฐาน (Standard Normal) ซึ่งค่า Z-score ที่น้อยขนาดนี้จะได้พื้นที่ใต้โค้งประมาณน้อยกว่า 0.001 หรือ น้อยกว่า 0.1 เปอร์เซนต์ (พื้นที่ใต้โค้งเมื่อ  $Z = 3.09$  คือ 0.001)