

Multicollinearity

เมื่อใดก็ตามที่นักวิจัยใช้การวิเคราะห์การถดถอยพหุคูณ (multiple regression analysis) อาจเกิดกรณี multicollinearity ที่เกิดจากการที่ตัวแปรอิสระมีความสัมพันธ์ระหว่างกัน ซึ่งเป็นผลทำให้เกิดประเด็นทางสถิติหลายประการ ดังนี้

1. หากมีการเพิ่มหรือลดตัวแปรอิสระ หรือเปลี่ยนแปลงขนาดของกลุ่มตัวอย่าง ค่าสัมประสิทธิ์หน้าตัวแปรแต่ละตัวจะผันผวนมาก
2. ค่าสัมประสิทธิ์หน้าตัวแปรแต่ละตัวไม่มีนัยสำคัญ (จากการทดสอบด้วยค่าสถิติ t) แต่การทดสอบค่าสัมประสิทธิ์ทั้งหมดแตกต่างจากศูนย์อย่างมีนัยสำคัญ (จากการทดสอบด้วยค่าสถิติ F)
3. ค่าสัมประสิทธิ์หน้าตัวแปรแต่ละตัวในการวิเคราะห์การถดถอยพหุคูณ (multiple regression analysis) ไม่มีนัยสำคัญ แต่การวิเคราะห์การถดถอยอย่างง่าย (simple regression analysis) กลับพบว่า ค่าสัมประสิทธิ์หน้าตัวแปรแต่ละตัวแตกต่างจากศูนย์อย่างมีนัยสำคัญ
4. เรายังสามารถใช้สมการที่ได้จากการวิเคราะห์การถดถอยพหุคูณในการพยากรณ์ค่าของตัวแปรตามได้ แต่การใช้ตัวแปรอิสระใด ๆ ในการพยากรณ์ผลกระทบที่มีต่อตัวแปรตามจะไม่น่าเชื่อถือ

เพื่อให้เห็นลักษณะปัญหาของ multicollinearity เราลองพิจารณาตัวอย่างของข้อมูลต่อไปนี้โดยกำหนดให้ Y=ยอดนำเข้า X1= ผลิตภัณฑ์มวลรวมประชาชาติ(GDP) X2= อัตราเงินเฟ้อ

ปี	Y	X1	X2	ปี	Y	X1	X2
1	0.054	0.079	-0.1593	9	0.0841	0.0503	0.0527
2	0.0656	0.0505	-0.2683	10	0.1540	0.0621	-0.1500
3	0.1475	0.0780	0.4801	11	0.0578	0.0432	0.0251
4	0.0686	0.0750	0.1348	12	0.0918	0.0600	-0.1119
5	0.0455	0.0627	-0.0218	13	0.0949	0.0623	-0.0131
6	0.0827	0.0464	0.1612	14	0.0555	0.0585	-0.3613
7	-0.0157	0.0399	-0.2511	15	0.1593	0.0652	0.2579
8	0.0753	0.0640	-0.2611				

ผลจากการวิเคราะห์การถดถอยที่คอมพิวเตอร์ได้ computer output ดังนี้

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.650 ^a	.423	.327	.0374421

a. Predictors: (Constant), inflation, gross domestic product
b. Dependent Variable: growth rate of import

นั่นก็คือ ตัวแปร X1 และ X2 ใช้อธิบายความผันผวนของตัวแปรตามได้ถึงร้อยละ 42.3

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.012	2	.006	4.398	.037 ^a
	Residual	.017	12	.001		
	Total	.029	14			

a. Predictors: (Constant), inflation, gross domestic product
b. Dependent Variable: growth rate of import

F-test ระบุว่า ค่าสัมประสิทธิ์หน้าตัวแปรอิสระทุกตัวแตกต่างจากศูนย์ (0) อย่างมีนัยสำคัญ

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.002	.058		.029	.977
	gross domestic product	1.393	.958	.344	1.453	.172
	Inflation	.088	.047	.438	1.850	.089

a. Dependent Variable: growth rate of import

ผลจาก t-test ระบุว่า ค่าสัมประสิทธิ์หน้าตัวแปรอิสระแต่ละตัวมีค่าใกล้ศูนย์ (0) และไม่มีนัยสำคัญ นี่คือน่าสนใจที่ขัดแย้ง (paradox) ซึ่งจะเกิดขึ้นเมื่อใดก็ตามที่ตัวแปรอิสระมี multicollinearity

ผลจาก computer output เราได้สมการการถดถอยที่คอมพิวเตอร์ที่สามารถนำไปใช้พยากรณ์ค่าของ Y ถ้าเราทราบค่า X1 และ X2 ดังนี้

$$Y = 0.002 + 1.393X_1 + 0.088X_2$$

เรามาดูการพิจารณา computer output ที่ได้จากการใช้ตัวแปรอิสระตัวใดตัวหนึ่งโดด ๆ

ใช้ X1 พยากรณ์ Y

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.508 ^a	.258	.201	.0407808

a. Predictors: (Constant), gross domestic product
b. Dependent Variable: growth rate of import

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.041	.058		-.697	.498
	gross domestic product	2.059	.967	.508	2.128	.053

a. Dependent Variable: growth rate of import

นั่นก็คือ หากใช้ X1 ในการพยากรณ์ Y X1สามารถใช้อธิบายความผันผวนของ Y ได้ร้อยละ 25.8

และค่าสัมประสิทธิ์ของ X1 เกือบมีนัยสำคัญทางสถิติที่ระดับ 0.05 ยิ่งไปกว่านั้นสมการที่ใช้ X1 ในการ

พยากรณ์ค่าของ Y คือ $Y = -0.041 + 2.059X1$

ใช้ X2 พยากรณ์ Y

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.567 ^a	.321	.269	.0390107

a. Predictors: (Constant), inflation
b. Dependent Variable: growth rate of import

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.085	.010		8.356	.000
	inflation	.114	.046	.567	2.481	.028

a. Dependent Variable: growth rate of import

นั่นก็คือ หากใช้ X2 ในการพยากรณ์ Y X2สามารถใช้อธิบายความผันผวนของ Y ได้ร้อยละ 32.1

และค่าสัมประสิทธิ์ของ X2 มีนัยสำคัญทางสถิติที่ระดับ 0.05 และสมการที่ใช้ X2 ในการพยากรณ์ค่าของ Y คือ

$Y = .085 + .114X2$

เมื่อใดถึงควรกังวลเกี่ยวกับปัญหา multicollinearity?

เมื่อใดก็ตามที่เราพบกรณีอย่างใดอย่างหนึ่งดังต่อไปนี้

1. เมื่อค่า tolerance มีค่าอยู่ในระดับต่ำ(0.20,0.10 หรือต่ำกว่า)

2.เมื่อค่า Variance inflation factor(VIF) มีค่าอยู่ในระดับสูง (5, 10 หรือมากกว่า)

ตามตัวอย่างข้างต้น เราสามารถดูค่าของ tolerance หรือ VIF ได้โดยระบุคำสั่งเพิ่มเติม ซึ่งทำให้ได้ computer output แสดงได้ดังตารางต่อไปนี้

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	.002	.058		.029	.977		
	gross domestic product	1.393	.958	.344	1.453	.172	.859	1.164
	inflation	.088	.047	.438	1.850	.089	.859	1.164

a. Dependent Variable: growth rate of import

Contribution this issue: ดร. ดนัย ปัตตพงษ์

อยากเรียนรู้การนำสถิติข้างต้นนี้ไปใช้ในการวิจัยระดับสารนิพนธ์ (independent study) วิทยานิพนธ์ (thesis) ดุษฎีนิพนธ์(dissertation) ปรึกษาได้ที่ dpattaphongse@gmail.com / dpattaphongse@yahoo.com

- * ผู้แต่ง MBA's Made Easy (160+ issues) เอกสารวิชาการด้านศาสตร์การบริหารธุรกิจที่ช่วยให้ธุรกิจสามารถยืนหยัดและอยู่รอดได้ในภาวะที่โลกเปลี่ยนแปลงอยู่ตลอดเวลา
- * ผู้พัฒนา FINALYSIS... a dedicated software สำหรับช่วยในการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์ (บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯ ได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง(กำหนด debut 1 เมษายน 2569)
- * ผู้แต่งหนังสือ"การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน"ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม
- * Co-developer ซอฟต์แวร์ en@gex@cel[®] สำหรับใช้ทดสอบ/เรียนรู้ศัพท์(ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาในระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ ฯลฯ อีกมาก (กำหนด debut 1 เมษายน 2569)