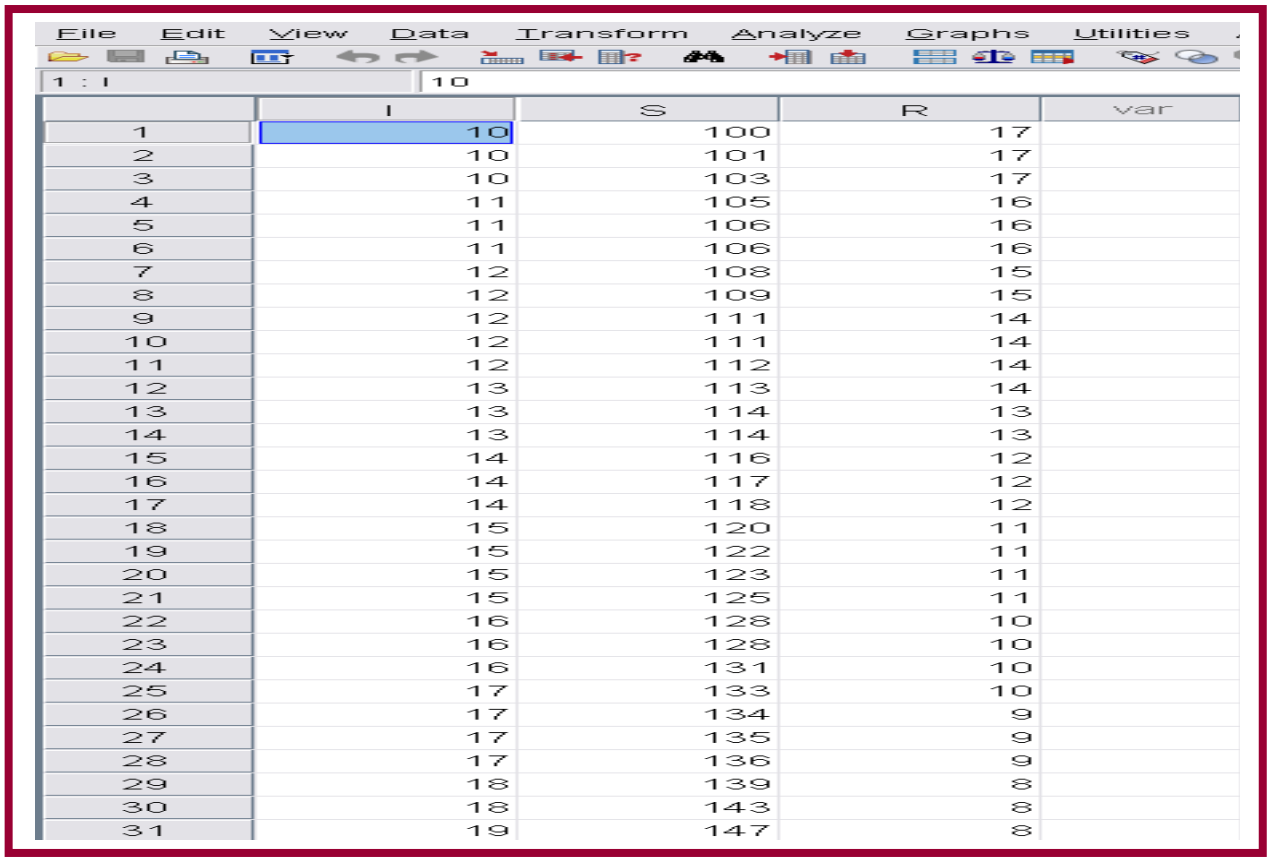
48

Multiple Regression Workshop#7-Detecting heteroscedasticity using Goldfeld-Quant Procedure

ตารางด้านล่างแสดงข้อมูลเกี่ยวกับสินค้าคงคลัง (I-inventories) ยอดขาย (S-sales) และอัตราดอกเบี้ยเงินกู้ (R-borrowing rate) โดยในที่นี้ตัวแปรตามได้แก่ สินค้าคงคลัง ส่วนตัวแปรอิสระได้แก่ยอดขายและอัตราดอกเบี้ยเงินกู้



File Edit View Data Transform Analyze Graphs Utilities

1 : I 10

	I	S	R	var
1	10	100	17	
2	10	101	17	
3	10	103	17	
4	11	105	16	
5	11	106	16	
6	11	106	16	
7	12	108	15	
8	12	109	15	
9	12	111	14	
10	12	111	14	
11	12	112	14	
12	13	113	14	
13	13	114	13	
14	13	114	13	
15	14	116	12	
16	14	117	12	
17	14	118	12	
18	15	120	11	
19	15	122	11	
20	15	123	11	
21	15	125	11	
22	16	128	10	
23	16	128	10	
24	16	131	10	
25	17	133	10	
26	17	134	9	
27	17	135	9	
28	17	136	9	
29	18	139	8	
30	18	143	8	
31	19	147	8	

เอกสารวิชาการนี้ต้องการจะแสดงตัวอย่างขั้นตอนในการตรวจสอบปัญหาของ heteroscedasticity ที่อาจเป็นต้นเหตุให้การวิเคราะห์สมการถดถอยทวีคูณไม่มีประสิทธิผล โดยขั้นตอนในการตรวจสอบนี้เรียกว่า Goldfeldt –Quandt Procedureในส่วนสมมติฐานทางสถิติที่ต้องการทดสอบมีดังนี้

H_O : $Var(\varepsilon_I)$ ของแบบจำลองการถดถอยมีค่าคงที่ หรืออีกนัยหนึ่ง error terms มีลักษณะ homoscedastic

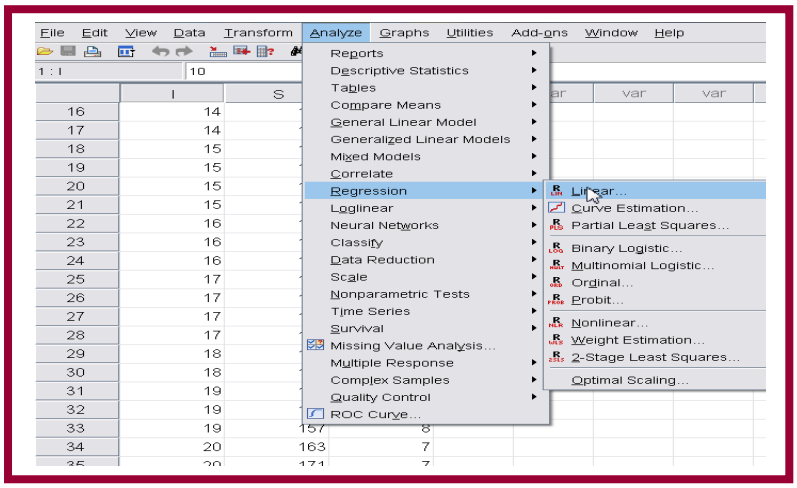
H_a : $Var(\varepsilon_I)$ ของแบบจำลองการถดถอยขึ้นอยู่กับตัวแปรอิสระ หรืออีกนัยหนึ่ง error terms มีลักษณะ heteroscedastic

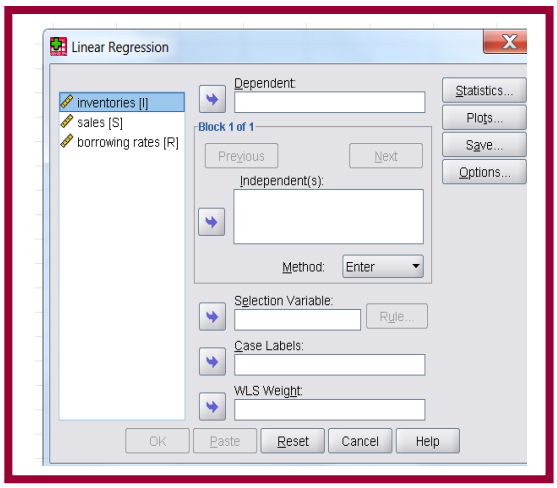
Goldfeldt-Quandt Procedure

1. จัดเรียงลำดับข้อมูลทั้งหมด โดยให้ตัวแปรอิสระมีค่าจากน้อยไปหามาก ในกรณีที่มีตัวแปรอิสระมากกว่าหนึ่งตัว ให้เลือกตัวแปรอิสระที่ท่านมีความกังวลว่าerror variance ของตัวแปรตามมีความสัมพันธ์กับความผันผวนของตัวแปรอิสระตัวใดตัวหนึ่งเป็นการเฉพาะ หรือจากผลการศึกษาเบื้องต้นจากการถดถอยอย่างง่ายที่มุ่งเน้นความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระหนึ่งๆ
2. แบ่งข้อมูลออกเป็นสองส่วน ส่วนแรกประมาณ 2 ใน 5 ส่วนที่สองประมาณ 2 ใน 5 โดยเหลือข้อมูลส่วนกลางไว้ราว 1 ใน 5
3. ทำการหาสมการถดถอยเส้นที่หนึ่งจากข้อมูลชุดแรก และหาสมการถดถอยเส้นที่สองจากข้อมูลชุดที่ 2
4. ทำการหารค่า Error Sum of Squares (ESS) จากข้อมูลชุดที่ 2 ด้วย Error Sum of Squares จากข้อมูลชุดที่ 1 จะได้ค่าของสถิติ F
5. เปรียบเทียบค่า F ที่ได้กับค่าวิกฤติF ที่ degree of freedom เท่ากับ $n_2 - k$ และ $n_1 - k$ โดยที่ n_1และ n_2 เป็นdegree of freedom ของข้อมูลชุดที่ 2 และข้อมูลชุดที่ 1 ส่วน k=จำนวนตัวแปรอิสระ
6. หากค่าสถิติ F ที่ได้ต่ำกว่า critical F มีความหมายว่า heteroscedasticity ไม่ได้เป็นปัญหาที่ต้องกังวล

จากข้อมูลซึ่งมีขนาดกลุ่มตัวอย่างเท่ากับ 35 เราจะทำการแบ่งข้อมูลออกเป็นสามส่วน โดยส่วนแรกจะประกอบด้วยกลุ่มตัวอย่าง 14รายแรก (#1-14)ส่วนที่สองจะประกอบด้วยกลุ่มตัวอย่าง 14 รายหลัง (#22-35) โดยเราจะงดเว้นไม่รวมกลุ่มตัวอย่างที่เหลืออยู่ตรงกลางระหว่างสองกลุ่มจำนวน 7 ราย (#15-21)

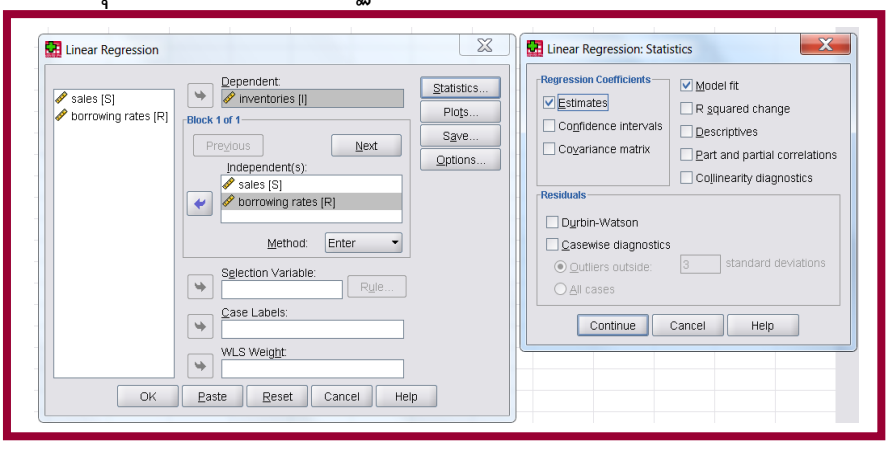
◉คลิ๊กAnalyze>regression>linear จะปรากฏsubmenu แสดงได้ดังภาพล่างขวามือ





◉คลิ๊กเลือก inventories ไปอยู่ในกล่องDependent คลิ๊กเลือก Sales และ Borrowing rates ไปอยู่ในช่อง Independent(s)

◉คลิ๊กที่ปุ่ม Statistics จะปรากฏ submenu แสดงได้ดังภาพล่างขวามือ



ในส่วน Regression Coefficients คลิ๊กเลือก Estimates คลิ๊กเลือก Model fit

ต่อไปนี้เป็น computer output ของข้อมูลชุดแรก (#1-14)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.970[a]	.941	.930	.287

a. Predictors: (Constant), borrowing rates, sales

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14.521	2	7.261	88.002	.000[a]
	Residual	.908	11	.083		
	Total	15.429	13			

a. Predictors: (Constant), borrowing rates, sales

b. Dependent Variable: inventories

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-2.230	13.081		-.170	.868
	sales	.158	.084	.683	1.896	.084
	borrowing rates	-.220	.273	-.291	-.808	.436

a. Dependent Variable: inventories

ในทำนองเดียวกัน computer output ของข้อมูลชุดที่สอง (#22-35) เป็นดังนี้

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.977[a]	.955	.947	.328

a. Predictors: (Constant), borrowing rates, sales

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	25.177	2	12.589	117.367	.000[a]
	Residual	1.180	11	.107		
	Total	26.357	13			

a. Predictors: (Constant), borrowing rates, sales

b. Dependent Variable: inventories

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.298	3.939		3.884	.003
	sales	.055	.016	.525	3.442	.006
	borrowing rates	-.626	.201	-.476	-3.123	.010

a. Dependent Variable: inventories

ผลจากการใช้โปรแกรมสำเร็จรูปพบว่าจากกลุ่มตัวอย่าง 14 รายแรก

Inventories=-2.23 +.158Sales -.22Borrowing rate R^2 = .94 ; RSS=.908

ส่วนผลจากกลุ่มตัวอย่าง 14 รายหลัง

Inventories= 15.298+.055sales-.626Borrowing rate R^2 = .96 ; RSS=1.118

G-Q's F-statistic $= \frac{1.118/(14-2)}{.908/(14-2)} = 1.23$

เนื่องจากสถิติ F ที่ได้ (1.23) มีค่าต่ำกว่า ค่าวิกฤติ F(12,12) ที่ α=0.05ซึ่งเท่ากับ 2.69

ดังนั้นยอมรับ H_O: Variance ของ error terms มีค่าคงที่

สรุปก็คือ Goldfeldt-Quant ไม่พบว่าข้อมูล error terms มีลักษณะเป็น heteroscedastic

Contribution this issue: ดร.ดนัย ปัตตพงศ์*

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

*ผู้แต่ง MBA's Made Easy(160+ issues) เอกสารวิชาการด้านศาสตร์การบริหารธุรกิจที่ช่วยให้ธุรกิจสามารถยืนหยัดและอยู่รอดได้ในภาวะที่โลกเปลี่ยนแปลงอยู่ตลอดเวลา

*ผู้พัฒนา FINALYSIS… a dedicated software สำหรับให้บริการนักธุรกิจที่ต้องการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์(บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง(กำหนด DEBUT 1 เมษายน 2569)

*ผู้แต่งหนังสือ"การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน"ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม

*Co-developer ซอฟแวร์en◉gex◉cel®สำหรับใช้ทดสอบ/เรียนรู้ศัพท์(ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ฯลฯอีกมาก(กำหนด DEBUT 1 เมษายน 2569)