Multiple Regression Workshop#4 -Nonlinear relationship

หนึ่งในข้อสมมติฐานของการวิเคราะห์การถดถอยพหุคูณนั้น กำหนดให้ความสัมพันธ์ระหว่างตัวแปรตาม (dependent variable) กับตัวแปรอิสระ (independent variable(s)) หนึ่งหรือหลายตัวต้องเป็นความสัมพันธ์เชิงเส้นตรง (linear form)

ตารางด้านล่างแสดงข้อมูลของปริมาณผลผลิต (Q-Quantity) จำนวนแรงงานที่ใช้ (L-labor) และทุนในรูปเครื่องมือเครื่องจักรในการผลิต (K-Capital)

File Edit View Data Transform Analyze Graphs Utilities

1 : Q 240

	Q	L	K	var
1	240	1480	410	
2	400	1660	450	
3	110	1150	380	
4	530	1790	430	
5	590	1880	480	
6	470	1860	450	
7	450	1940	490	
8	160	1240	395	
9	290	1240	430	
10	490	1850	460	
11	350	1570	435	
12	550	1700	470	
13	560	2000	480	
14	430	1850	440	

หากความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรต้นอยู่ในรูปแบบเส้นตรง สมการถดถอยทวีคูณจะอยู่ในรูปแบบดังต่อไปนี้

$Q = b_O + b_1 L + b_2 K + e$

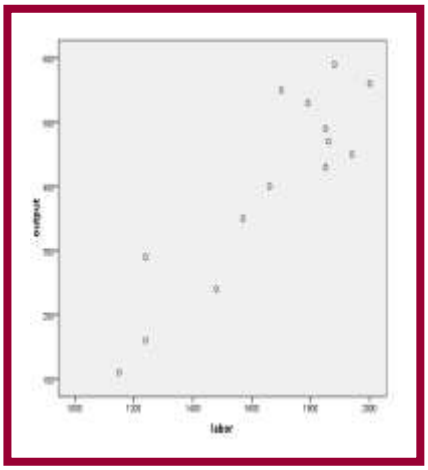
โดยที่ Q= ปริมาณผลผลิต

L= จำนวนหน่วยแรงงานที่ใช้ในการผลิต

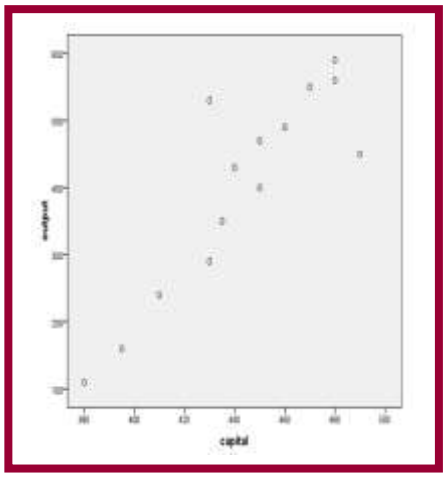
K = จำนวนหน่วยทรัพยากรทุนที่ใช้ในการผลิต

e = error term

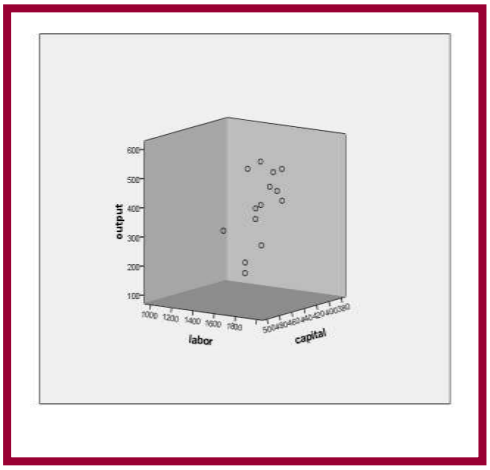
ภาพ A, B และ C ด้านล่างแสดงให้เห็นกราฟแสดงความสัมพันธ์ระหว่างผลผลิต (Q) กับแรงงาน (L) ในรูปสองมิติ กราฟแสดงความสัมพันธ์ระหว่างผลผลิต (Q) กับทรัพยากรทุน (K) ในรูปสองมิติและท้ายที่สุดกราฟแสดงความสัมพันธ์ระหว่างผลผลิต Q กับ L และ K ในรูปสามมิติ



ภาพ A



ภาพ B



ภาพ C

จะเห็นได้ชัดว่า ความสัมพันธ์ระหว่างตัวแปรตาม Q กับตัวแปรต้น L และ K ไม่ได้มีลักษณะเป็นความสัมพันธ์เชิงเส้นตรง ดังนั้นการหาความสัมพันธ์ระหว่างตัวแปรทั้งสามโดยใช้การวิเคราะห์การถดถอยพหุคูณอาจเป็นไปอย่างไม่มีประสิทธิผล เพราะละเมิดสมมติฐานข้อที่ 4 (ดู Statistics Talks # 41)

ผู้ที่มีพื้นฐานทางเศรษฐศาสตร์มาก่อนคงพอจะระลึกถึงฟังก์ชั่นในการผลิตที่มีชื่อว่า Cobb-Douglas Production Function ซึ่งมีรูปแบบดังนี้

$Q = b_O L^{b1} . K^{b2}$ ..... (1)

โดยที่Q = ผลผลิต

b_O = ค่าคงที่

L = แรงงานที่ใช้ในการผลิต

K = ทรัพยากรทุนที่ใช้ในการผลิต

b_1 = ความยืดหยุ่นของผลผลิตต่อแรงงาน (output elasticity of labor)

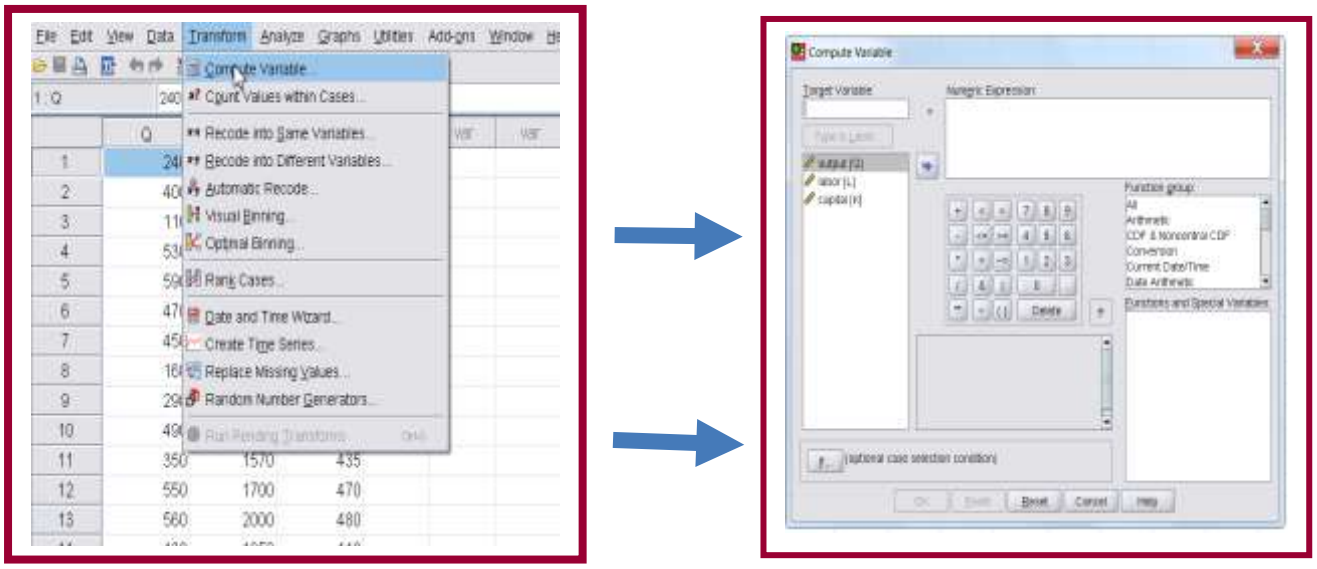
b_2 = ความยืดหยุ่นของผลผลิตต่อทรัพยากรทุน (output elasticity of capital)

เราจะทำการแปลงตัวแปรตามและตัวแปรอิสระตาม (1) ให้อยู่ในรูป natural log โดยใช้ log operation ทั้งสองข้าง เราจะได้สมการที่เป็นเส้นตรงดังนี้

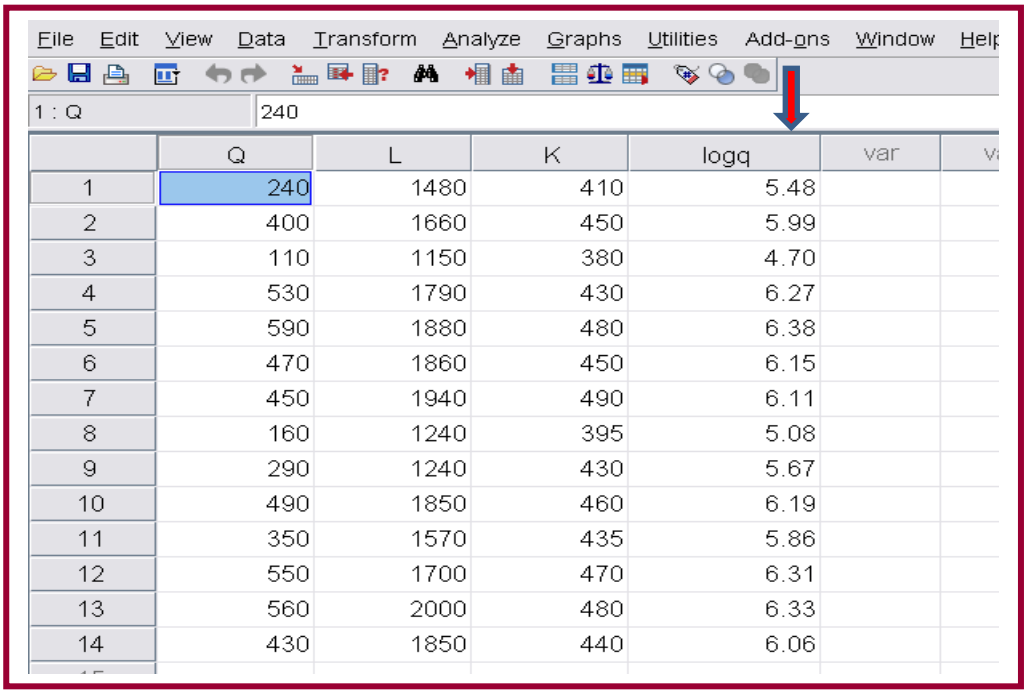
Log Q = log b_O + b_1 log L +b_2 log K ..... (2)

ขั้นตอนต่อไป เราจะทำการหาค่าของตัวแปรหลังจากมีการแปลงข้อมูลให้อยู่ในรูป ln

◉ คลิ๊กที่ menu transform> compute variable จะปรากฏ menu และ submenu ดังต่อไปนี้

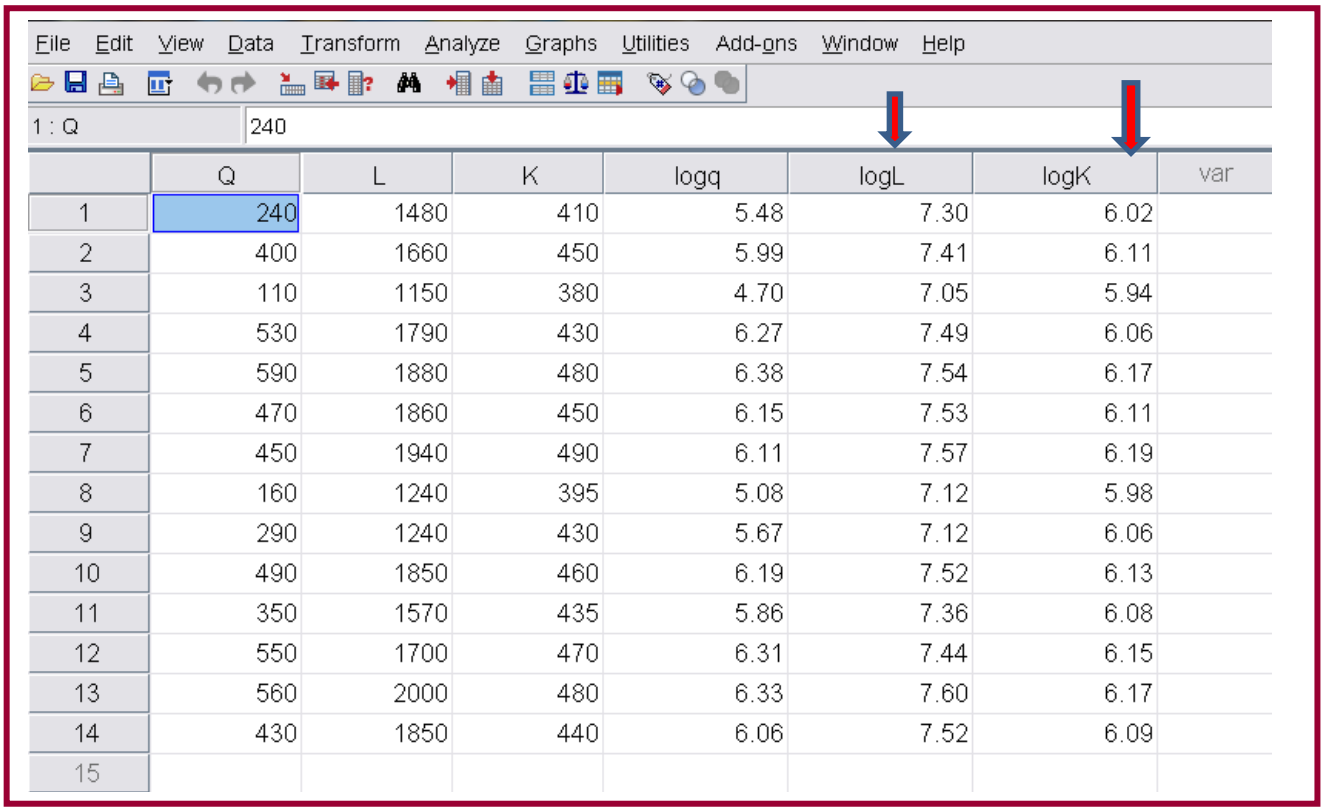


◉ ในกล่อง input variable กรอก logQ ส่วนในช่อง numeric expression ให้ใช้ arithmetic operation menu ย่อยด้านขวาทำการหาค่าln(Q) จากนั้นกด enter จะปรากฏตารางแสดงค่าของตัวแปร Q,L,K และ logQ แสดงได้ดังนี้



	Q	L	K	logq	var
1	240	1480	410	5.48	
2	400	1660	450	5.99	
3	110	1150	380	4.70	
4	530	1790	430	6.27	
5	590	1880	480	6.38	
6	470	1860	450	6.15	
7	450	1940	490	6.11	
8	160	1240	395	5.08	
9	290	1240	430	5.67	
10	490	1850	460	6.19	
11	350	1570	435	5.86	
12	550	1700	470	6.31	
13	560	2000	480	6.33	
14	430	1850	440	6.06	

◉ ขั้นตอนต่อไป เราจะใช้คำสั่ง transform แปลงค่า L และ K ที่เหลือให้อยู่ในรูปตัวแปร logL และ logK โดยใช้ arithmetic operation หาค่าของ ln(L) และ ln (K) ซึ่งผลที่ได้จะปรากฏแสดงได้ดังตารางด้านล่างนี้



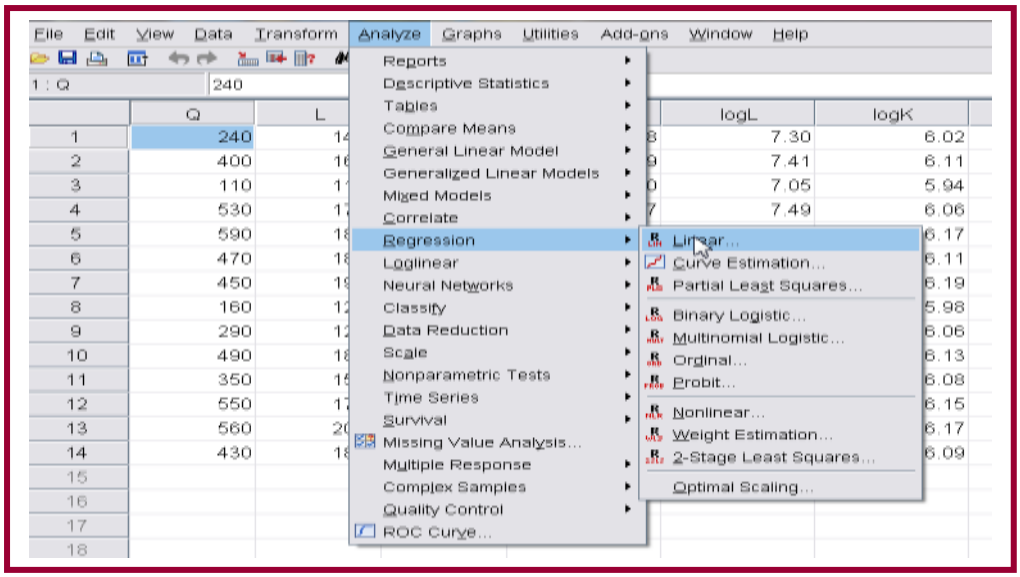
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : Q 240

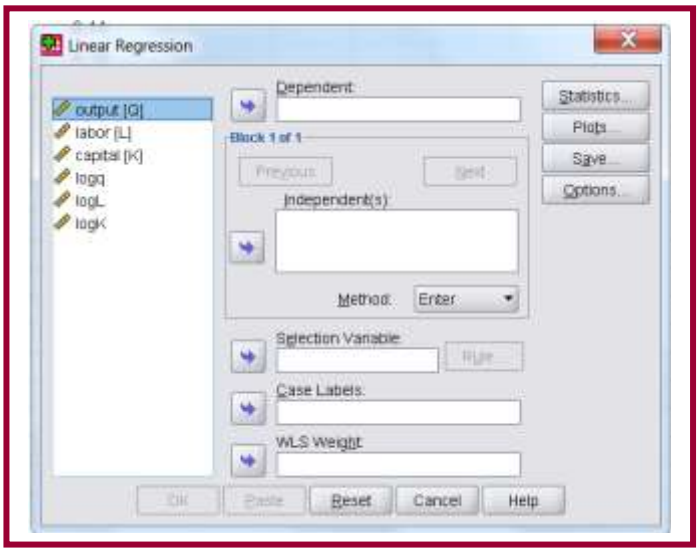
	Q	L	K	logq	logL	logK	var
1	240	1480	410	5.48	7.30	6.02	
2	400	1660	450	5.99	7.41	6.11	
3	110	1150	380	4.70	7.05	5.94	
4	530	1790	430	6.27	7.49	6.06	
5	590	1880	480	6.38	7.54	6.17	
6	470	1860	450	6.15	7.53	6.11	
7	450	1940	490	6.11	7.57	6.19	
8	160	1240	395	5.08	7.12	5.98	
9	290	1240	430	5.67	7.12	6.06	
10	490	1850	460	6.19	7.52	6.13	
11	350	1570	435	5.86	7.36	6.08	
12	550	1700	470	6.31	7.44	6.15	
13	560	2000	480	6.33	7.60	6.17	
14	430	1850	440	6.06	7.52	6.09	
15							

◉ เราจะทำการหาสมการถดถอยพหุคูณในรูป log form โดยใช้ menu ดังต่อไปนี้

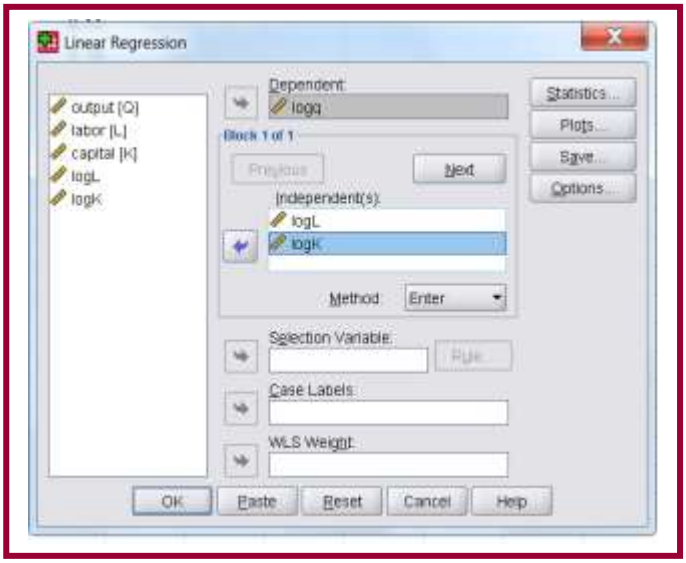
Analyze>regression>linear



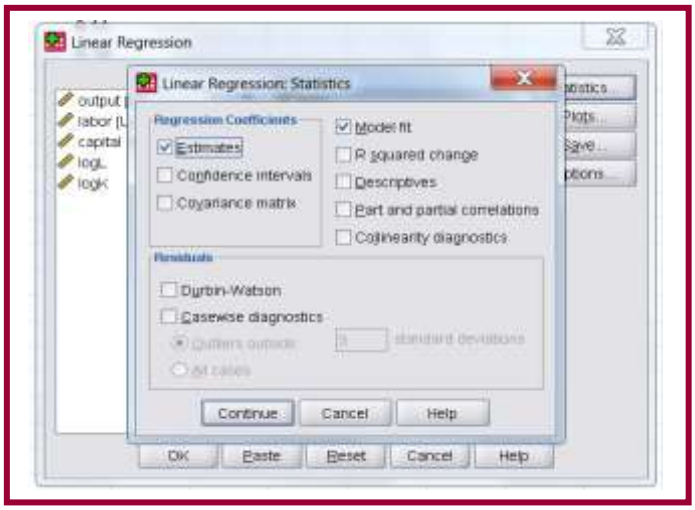
ภาพด้านล่างแสดง submenu-linear regression



ให้คลิ๊ก logq ไปอยู่ในช่อง dependent และคลิ๊ก logL logK ไปอยู่ในช่อง independent (s) ตามภาพดังต่อไปนี้



คลิ๊กที่ปุ่ม Statistics จะปรากฏ submenu แสดงได้ดังภาพถัดไป



คลิ๊กเลือก Estimates ในส่วนของ Regression Coefficients คลิ๊กเลือก Model fit เลือก Continue คลิ๊กปุ่ม OK

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.937[a]	.878	.856	.19128

a. Predictors: (Constant), logK, logL

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.898	2	1.449	39.597	.000[a]
	Residual	.402	11	.037		
	Total	3.300	13			

a. Predictors: (Constant), logK, logL

b. Dependent Variable: logq

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-23.231	5.236		-4.436	.001
	logL	1.430	.561	.519	2.550	.027
	logK	3.045	1.367	.453	2.227	.048

a. Dependent Variable: logq

จาก Computer output ที่ได้สามารถประมวลผลได้ดังนี้

1. สมการการถดถอยเชิงพหุคูณในรูป natural log form เป็นดังนี้

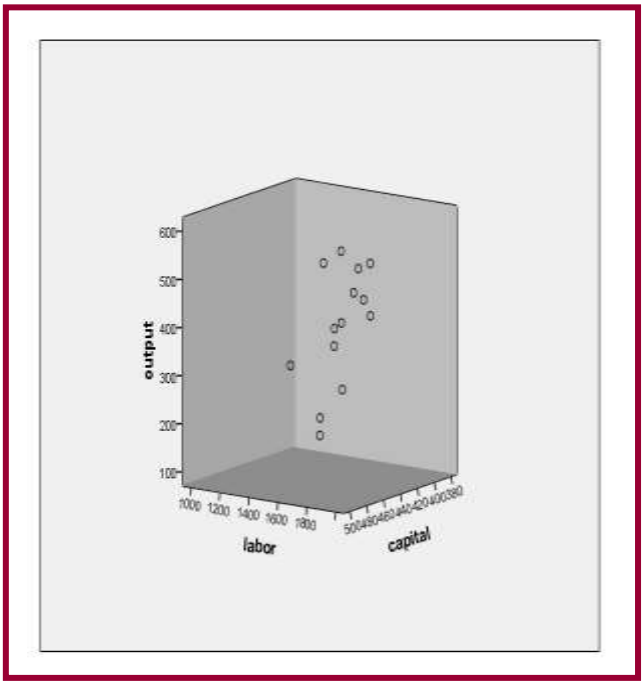
Log Q = -23.231 + 1.430log L + 3.045log K พร้อมค่า t-statisticดังนี้

(2.550) (2.227)

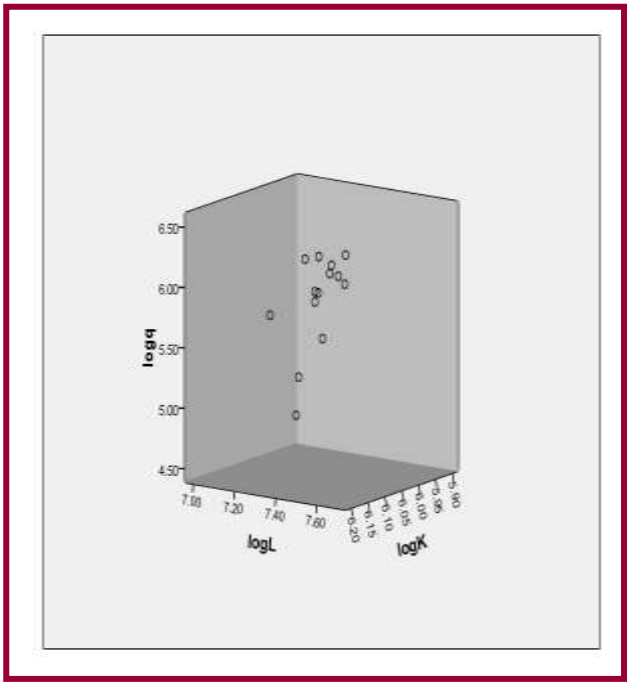
2.ผลจากตาราง ANOVA ชี้ให้เห็นว่า เราสามารถใช้ตัวแปรต้นสองตัวที่เป็น natural log form ในการกำหนดจำนวนผลผลิตในรูป log form โดยค่าสถิติ F มีนัยสำคัญ และ R^2 = 0.878

3.output elasticity ของแรงงาน (L) และทรัพยากรทุน (K) เท่ากับ 1.43 และ 3.045 ตามลำดับ ยิ่งไปกว่านั้นหากรวมสัมประสิทธิ์ทั้งสองตัว (1.43+3.045) จะเท่ากับ 4.475 ซึ่งมีค่ามากกว่า 1 นั่นหมายความว่า ในอุตสาหกรรมนี้มี increasing returns to scale นั่นก็คือหากเพิ่มแรงงานและทรัพยากรทุนเท่าๆ กันร้อยละหนึ่ง จะมีผลทำให้ผลผลิตเพิ่มขึ้นในอัตราร้อยละ 4.475

สุดท้ายเราลองมาพิจารณาเปรียบเทียบภาพสามมิติของ Q,L,K กับ Lnq, LnL, LnK ดูกัน



ภาพ 3 มิติ Q , L , K



ภาพ 3 มิติ Ln(q),Ln(L) ,Ln(K)

จะเห็นได้ว่าความสัมพันธ์ระหว่าง Ln (Q) ,Ln (L) และ Ln (K) มีลักษณะเป็นความสัมพันธ์ที่มีลักษณะเป็น linear function มากกว่า Q,L และ K

ผู้อ่านบางท่านอาจจะคิดขึ้นมาในใจว่า ถ้าเราจะดื้อดึงและฝืนสมมติฐานเกี่ยวกับความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรตามและตัวแปรต้นในแบบจำลองการถดถอยทวีคูณ เราก็จะได้ผลแสดงได้โดย computer output ดังต่อไปนี้

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.923[a]	.853	.826	63.323

a. Predictors: (Constant), capital, labor

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	255263.017	2	127631.508	31.829	.000[a]
	Residual	44108.412	11	4009.856		
	Total	299371.429	13			

a. Predictors: (Constant), capital, labor

b. Dependent Variable: output

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-931.640	306.490		-3.040	.011
	labor	.297	.120	.550	2.477	.031
	capital	1.897	1.034	.408	1.835	.094

a. Dependent Variable: output

สมการถดถอยที่ได้จะเป็น Q= -931.64 +.297L +1.897K แต่จะเห็นได้ว่า

1.ค่า R^2 ที่ได้มีค่าต่ำกว่าค่าที่มีการใช้ log transform

2.ผลการทดสอบสมมติฐานทางสถิติพบว่า regression coefficient หน้า K-capital มีค่าไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ (ค่า sig. ที่ได้มีค่ามากกว่า αซึ่งเท่ากับ .05) ดังนั้น หากตัดค่า K ออกจากสมการจะขัดแย้งกับทฤษฎี เพราะในทางเศรษฐศาสตร์การผลิตต้องใช้ทรัพยากรทั้งที่เป็นแรงงาน (L) และเครื่องจักร (K)

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 ฯลฯ ได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง(กำหนด DEBUT 1 เมษายน 2569)

* ผู้แต่งหนังสือ"การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน"ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม

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