การพยากรณ์ข้าวนาปี

มีขั้นตอนดังนี้

1. ข้อตกลงเบื้องต้น (Assumption)

1.1 ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเนื้อที่เพาะปลูกปีที่พยากรณ์หรือเนื้อที่เพาะปลูกปีที่ t ได้แก่

1) เนื้อที่เพาะปลูกในปีที่แล้ว หรือเนื้อที่เพาะปลูกปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า เกษตรกรจะตัดสินใจ ขยายหรือลดเนื้อที่เพาะปลูก ต้องคำนึงถึงเนื้อที่เพาะปลูกที่เคยทำมาในปีที่แล้ว

2) ราคาเฉลี่ย ราคาสูงสุด ราคาต่ำสุด ราคา ณ เดือนที่ผลผลิตออกมาก (เดือนพฤศจิกายน) ที่เกษตรกรขายได้ในปีที่แล้วหรือปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า เกษตรกรจะตัดสินใจในการขยายหรือลดเนื้อที่เพาะปลูกขึ้นอยู่กับราคาดังกล่าว โดยคาดว่าหากราคาที่เกษตรกรขายได้ในปีที่แล้วมีแนวโน้มเพิ่มขึ้นก็จะจูงใจให้เกษตรกรขยายเนื้อที่เพาะปลูกเพิ่มขึ้นในปีถัดไป และหากราคาที่เกษตรกรขายได้ในปีที่แล้วมีแนวโน้มลดลง เกษตรกรก็จะลดเนื้อที่เพาะปลูกลงในปีถัดไป

3) ราคาพืชแข่งขันปีที่แล้ว ซึ่งสามารถปลูกได้ในพื้นที่เดียวกัน ภายใต้ข้อตกลงเบื้องต้นว่า เช่น ถ้าราคาอ้อยโรงงานเมื่อปีที่แล้วมีแนวโน้มเพิ่มขึ้นหรือให้ผลตอบแทนที่ดีกว่า เกษตรกรก็จะปรับเปลี่ยนจากการปลูกข้าวไปปลูกอ้อยโรงงานแทน และหากราคาอ้อยโรงงานเมื่อปีที่แล้วมีแนวโน้มลดลงหรือผลตอบแทนน้อยกว่า เกษตรกรก็จะหันมาปลูกข้าวแทนอ้อยโรงงาน

4) ต้นทุนการผลิตปีที่แล้ว ภายใต้ข้อตกลงเบื้องต้นว่า หากต้นทุนการผลิตเมื่อปีที่แล้วมีแนวโน้มลดลง เกษตรกรจะตัดสินใจขยายเนื้อที่เพาะปลูกเพิ่มขึ้นเพื่อเพิ่มการผลิตและได้รับผลตอบแทนที่มากขึ้น และหากต้นทุนเมื่อปีที่แล้วมีแนวโน้มเพิ่มขึ้นและไม่คุ้มค่ากับผลตอบแทนที่ได้รับ เกษตรกรก็จะตัดสินใจลดเนื้อที่ปลูกข้าวและปรับเปลี่ยนไปปลูกพืชอื่นที่ให้ผลตอบแทนที่ดีกว่า

5) ปริมาณฝนตก และจำนวนวันฝนตก ภายใต้ข้อตกลงเบื้องต้นว่า หากปริมาณน้ำฝนหรือจำนวนวันที่ฝนตกในปีปัจจุบันเพียงพอต่อการเพาะปลูกเกษตรกรก็อาจจะขยายเนื้อที่เพาะปลูกเพิ่มขึ้น

6) ราคาน้ำมันดีเซล ภายใต้ข้อตกลงเบื้องต้นว่า หากราคาน้ำมันดีเซลต่ำลง เกษตรกรจะเพิ่มปริมาณการใช้เครื่องจักร เครื่องยนต์ในการเพาะปลูกมากขึ้น เช่น การเตรียมดิน ทำให้มีการขยายเนื้อที่เพาะปลูกมากขึ้น เนื่องจากมีความสะดวกและรวดเร็วขึ้น

7) แนวโน้มเวลา (Time Trend) เป็นตัวแปรอิสระที่กำหนดให้แทนการเปลี่ยนแปลงของเทคโนโลยี คือ เมื่อระยะเวลาผ่านไป มีผลให้เทคโนโลยีการผลิตเปลี่ยนแปลงไปด้วย ซึ่งอาจจะมีผลทำให้เนื้อที่เพาะปลูกเพิ่มขึ้น

8) ดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิต เป็นตัวแปรอิสระที่สะท้อนภาวะการเจริญเติบโตทางเศรษฐกิจ โดยทั่วไปจะใส่ไว้ในแบบจำลองเพื่อถ่วงค่า (Deflate) ราคาอื่น เพราะขณะที่ราคาพืชผลการเกษตรสูงขึ้น อาจจะเป็นภาพลวงตาว่าสินค้านั้นมีราคาดีขึ้น หากค่าดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิตสูงขึ้นด้วย

1.2 ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามผลผลิตต่อไร่ปีที่พยากรณ์หรือผลผลิตต่อไร่ปีที่ t ได้แก่

1) ราคาเฉลี่ย ราคาสูงสุด ราคาต่ำสุด ราคา ณ เดือนที่ผลผลิตออกมาก (เดือนพฤศจิกายน) ที่เกษตรกรขายได้ในปีที่แล้วหรือปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า ผลผลิตต่อไร่จะเพิ่มขึ้นหรือลดลงขึ้นอยู่กับการดูแลเอาใจใส่ของเกษตรกร หากราคาที่เกษตรกรขายได้ในปีที่ที่แล้วมีแนวโน้มเพิ่มขึ้นก็จะจูงใจให้เกษตรกรบำรุงดูแลรักษาดีส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้น และหากราคาที่เกษตรกรขายได้ในปีแล้วมีแนวโน้มลดลง เกษตรกรก็จะดูแลเอาใจใส่ลดลงอาจส่งผลให้ผลผลิตต่อไร่ลดลง

2) ดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิต เป็นตัวแปรอิสระที่สะท้อนภาวะการเจริญเติบโตทางเศรษฐกิจ โดยทั่วไปจะใส่ไว้ในแบบจำลองเพื่อถ่วงค่า (Deflate) ราคาอื่น เพราะขณะที่ราคาพืชผลการเกษตรสูงขึ้น อาจจะเป็นภาพลวงตาว่าสินค้านั้นมีราคาดีขึ้น หากค่าดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิตสูงขึ้นด้วย

3) ต้นทุนการผลิตปีที่พยากรณ์ ภายใต้ข้อตกลงเบื้องต้นว่า หากต้นทุนการผลิต เช่น ราคาปุ๋ย ยากำจัดวัชพืช เมื่อปีที่พยากรณ์มีแนวโน้มราคาลดลง เกษตรกรสามารถซื้อปัจจัยดังกล่าว เพื่อไปบำรุงดูแลรักษาต้นข้าว ส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้น หรือในทางกลับกัน หากแนวโน้มราคาที่เกษตรกรขายได้เพิ่มขึ้น เกษตรกรก็อาจมีแรงจูงใจในการเพิ่มต้นทุนการผลิต เพื่อให้ได้ผลผลิตต่อไร่ที่เพิ่มขึ้น

4) ปริมาณฝนตก และจำนวนวันฝนตก ภายใต้ข้อตกลงเบื้องต้นว่า ปริมาณน้ำฝน จะมีผลต่อผลผลิตต่อไร่ หากปริมาณน้ำฝนในปีใดมีมากสม่ำเสมอ ผลผลิตต่อไร่น่าจะมีแนวโน้มเพิ่มมากขึ้น และในทางตรงกันข้าม เมื่อเกิดภาวะฝนแล้งและฝนทิ้งช่วง จะทำให้ได้ผลผลิตลดลง

5) อัตราการใช้เมล็ดพันธุ์ และการใส่ปุ๋ยปีที่พยากรณ์ ภายใต้ข้อตกลงเบื้องต้นว่า หากมีการใช้เมล็ดพันธุ์และการใส่ปุ๋ยเพิ่มขึ้น จะส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้นด้วย

6) ราคาน้ำมันดีเซล ภายใต้ข้อตกลงเบื้องต้นว่า ราคาน้ำมันดีเซลต่ำลง เกษตรกรจะเพิ่มปริมาณการใช้เครื่องจักร เครื่องยนต์ ในการดูแลรักษามากขึ้น ส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้นด้วย

7) แนวโน้มเวลา (Time Trend) เป็นตัวแปรอิสระที่กำหนดให้แทนการเปลี่ยนแปลงของเทคโนโลยี คือ เมื่อระยะเวลาผ่านไป มีผลให้เทคโนโลยีการผลิตเปลี่ยนแปลงไปด้วย ซึ่งจะมีผลทำให้ผลผลิตต่อไร่เพิ่มขึ้น

2.กำหนดตัวแปรในแบบจำลอง

2.1 ตัวแปรตาม (Dependent Variable) เป็นตัวแปรที่ต้องการพยากรณ์ ประกอบด้วย

$Y1_t$ คือ เนื้อที่เพาะปลูกข้าวนาปี ปีที่ t

$Y2_t$ คือ ผลผลิตต่อไร่ข้าวนาปี ปีที่ t

2.2 ตัวแปรอิสระ (Independent Variable) เป็นตัวแปรที่มีสมมติฐานว่ามีอิทธิพลต่อการเปลี่ยนแปลงของตัวแปรที่จะทำการพยากรณ์ ซึ่งในการศึกษาครั้งนี้มีตัวแปรอิสระ ดังนี้

Lag_Y1 หรือ Y1_1 คือ เนื้อที่เพาะปลูกข้าวนาปี ปีที่ t-1

Lag_X1 หรือ X1_1 คือ ราคาข้าวเปลือกเจ้าหอมมะลิ 105 ที่เกษตรกรขายได้ ณ ความชื้น 15% ปีที่ t-1

Lag_X2 หรือ X2_1 คือ ราคาสูงสุดข้าวเปลือกเจ้าหอมมะลิ 105 ที่เกษตรกรขายได้ ณ ความชื้น 15% ปีที่ t-1

Lag_X3 หรือ X3_1 คือ ราคาต่ำสุดของข้าวเปลือกเจ้าหอมมะลิ 105 ที่เกษตรกรขายได้

ณ ความชื้น 15% ปีที่ t-1

Lag_X4 หรือ X4_1 คือ ราคาข้าวเปลือกเจ้าหอมมะลิ 105 ที่เกษตรกรขายได้ ณ ความชื้น 15% ณ เดือนที่ผลผลิตออกมามาก ปีที่ t-1

Lag_X5 หรือ X5_1 คือ ราคาอ้อยโรงงาน ที่เกษตรกรขายได้ ปีที่ t-1

$Cost_t$ คือ ต้นทุนการผลิตข้าวขาวดอกมะลิ 105 ภาคตะวันออกเฉียงเหนือ ปีที่ t

$Seed_t$ คือ อัตราการใช้เมล็ดพันธุ์ข้าวต่อไร่ของข้าวนาปี ปีที่ t

$Rate_Fer_t$ คือ อัตราการใส่ปุ๋ยต่อเนื้อที่เพาะปลูกต่อไร่ของข้าวนาปี ปีที่ t

$Poil_t$ คือ ราคาน้ำมันดีเซล ปีที่ t

$Trend_t$ คือ แนวโน้มเวลา ปีที่ t

$RT5_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงพฤษภาคม ปีที่ t

$RT6_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงมิถุนายน ปีที่ t

$RT7_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงกรกฎาคม ปีที่ t

$RT8_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงสิงหาคม ปีที่ t

$RT9_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงกันยายน ปีที่ t

$RT10_t$ คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงตุลาคม ปีที่ t

$RD5_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงพฤษภาคม ปีที่ t

$RD6_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงมิถุนายน ปีที่ t

$RD7_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงกรกฎาคม ปีที่ t

$RD8_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงสิงหาคม ปีที่ t

$RD9_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงกันยายน ปีที่ t

$RD10_t$ คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงตุลาคม ปีที่ t

3. จัดเตรียมข้อมูลเพื่อการพยากรณ์

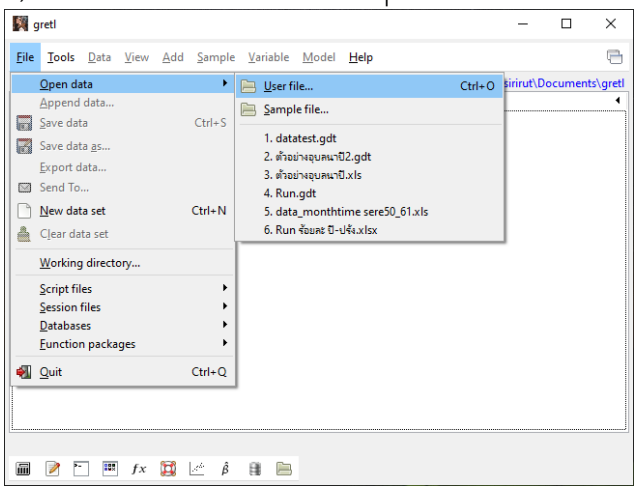
3.1 นำข้อมูลตัวแปรอิสระและตัวแปรตามที่ได้จากการสืบค้นมาจัดเรียงตามปีน้อยไปมาก

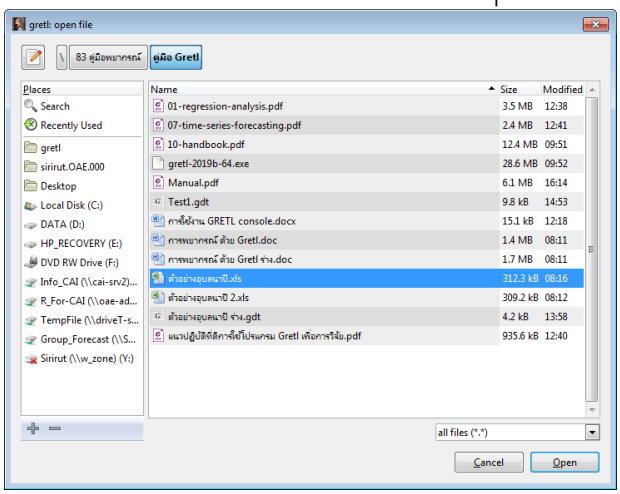
ตัวอย่าง ที่ 3.1 การจัดเตรียมข้อมูล เพื่อพยากรณ์ผลผลิตข้าวนาปี ของจังหวัดอุบลราชธานี

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	AA	AB	AC	AD
1	year	Y1	X1	X2	X3	X4	X5	Cost	Trend	RT5	RT6	RT7	RT8	RT9	RT10	RD5	RD6	RD7	RD8	RD9	RD10	POIL	Lag_Y1	Lag_X1	Lag_X2	Lag_X3	Lag_X4	Lag_X5
2	2542	1,490,565	7,205	8,700	6,700	6,800	488	1,468.13	1	175.76	102.36	141.71	70.74	165.06	56.60	21	20	20	22	21	14	8.96						
3	2543	1,590,342	6,521	7,125	4,700	7,125	421	1,697.84	2	326.60	235.58	247.88	168.52	157.66	76.47	25	25	28	25	25	20	12.93	1,490,565	7,205	8,700	6,700	6,800	488
4	2544	1,738,499	5,164	6,480	4,813	4,825	485	1,722.13	3	135.59	279.97	172.27	329.67	308.81	195.67	29	25	28	28	21	23	13.43	1,590,342	6,521	7,125	4,700	7,125	421
5	2545	1,684,993	6,795	9,240	5,804	6,188	433	1,747.58	4	139.94	191.50	331.98	288.75	385.40	93.69	30	27	31	30	29	21	13.12	1,738,499	5,164	6,480	4,813	4,825	485
6	2546	1,642,722	8,947	9,633	7,275	7,600	452	1,858.71	5	212.60	176.63	144.89	274.19	500.15	34.18	23	22	24	27	29	17	14.03	1,684,993	6,795	9,240	5,804	6,188	433
7	2547	1,646,172	7,863	8,060	7,600	8,060	359	1,885.11	6	260.35	138.53	299.00	238.99	170.25	11.30	28	29	30	30	25	5	14.59	1,642,722	8,947	9,633	7,275	7,600	452
8	2548	1,611,879	7,906	8,450	7,795	7,795	508	2,328.20	7	138.74	241.66	249.81	268.68	229.93	30.37	26	28	29	27	26	18	20.01	1,646,172	7,863	8,060	7,600	8,060	359
9	2549	1,662,816	8,225	8,970	7,990	8,100	747	2,482.52	8	63.16	115.93	342.53	374.86	198.03	181.31	18	21	28	31	20	17	25.56	1,611,879	7,906	8,450	7,795	7,795	508
10	2550	1,687,796	12,505	18,167	9,015	9,015	692	2,582.05	9	276.84	164.39	178.92	396.17	250.67	197.34	25	25	25	30	24	20	25.66	1,662,816	8,225	8,970	7,990	8,100	747
11	2551	1,721,866	14,108	15,900	12,580	13,450	622	3,534.60	10	33.56	143.40	131.96	274.22	355.90	124.07	31	25	26	30	29	28	31.16	1,687,796	12,505	18,167	9,015	9,015	692
12	2552	1,881,719	14,722	15,300	13,620	13,648	726	3,561.45	11	144.33	173.69	403.14	175.31	360.99	48.43	20	27	29	29	27	21	24.77	1,721,866	14,108	15,900	12,580	13,450	622
13	2553	2,238,602	13,715	14,880	13,013	14,380	984	3,727.29	12	-	111.25	217.20	360.97	251.13	157.19	0	27	27	30	29	17	28.59	1,881,719	14,722	15,300	13,620	13,648	726
14	2554	2,448,780	15,139	16,500	14,060	16,500	967	3,908.14	13	191.72	203.62	356.83	425.98	355.37	172.78	26	30	29	31	29	19	29.44	2,238,602	13,715	14,880	13,013	14,380	984
15	2555	2,484,492	15,691	15,958	15,094	15,625	1,044	4,250.47	14	209.30	109.12	204.04	220.74	265.21	59.56	24	27	29	29	28	11	30.41	2,448,780	15,139	16,500	14,060	16,500	967
16	2556	2,200,086	14,299	14,512	13,749	14,512	1,022	4,490.17	15	199.04	113.84	325.44	63.92	481.52	34.71	23	26	20	27	28	28	29.97	2,484,492	15,691	15,958	15,094	15,625	1,044
17	2557	2,217,033	13,390	14,033	12,151	12,175	933	4,570.65	16	109.28	398.92	480.90	170.56	209.98	87.90	20	28	30	26	25	18	29.73	2,200,086	14,299	14,512	13,749	14,512	1,022
18	2558	2,460,946	10,074	11,274	9,300	11,274	919	3,939.39	17	84.96	119.48	234.72	144.22	195.47	108.55	21	13	25	26	22	18	24.87	2,217,033	13,390	14,033	12,151	12,175	933
19	2559	2,789,533	9,241	12,825	8,583	8,583	798	3,774.95	18	97.02	348.08	161.02	180.48	330.49	64.56	25	28	23	30	30	22	23.25	2,460,946	10,074	11,274	9,300	11,274	919
20	2560	2,755,660	14,369	15,800	12,375	12,375	1,066	3,731.80	19	234.35	221.88	379.90	248.21	139.21	75.81	26	28	30	23	29	19	25.63	2,789,533	9,241	12,825	8,583	8,583	798
21	2561	2,798,656	16,333	16,660	16,180	16,660	799	3,789.27	20	134.58	190.00	312.55	240.97	190.88	79.31	26	27	30	29	27	17	28.35	2,755,660	14,369	15,800	12,375	12,375	1,066
22	2562	2,836,118	14,002	14,658	13,566	14,658	618	3,798.96	21	174.75	77.05	291.58	223.14	278.02	88.88	24	21	29	28	27.4	19	26.46	2,798,656	16,333	16,660	16,180	16,660	799
23	2563							3,817.97	22.00	143.50	189.32	296.42	240.20	295.09	88.04	22	26	27	28	26.4	17.4	26.81	2,836,118	14,002	14,658	13,566	14,658	618

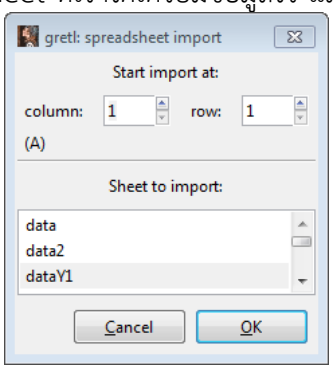
3.2 นำข้อมูลเข้าโปรแกรม Gretl

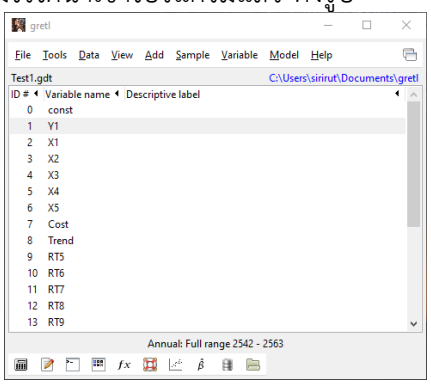
1) เปิดโปรแกรมแล้วเลือก File>Open data>User file จากนั้นค้นหาไฟล์ที่จะใช้ เลือก ไฟล์ แล้ว กด Open



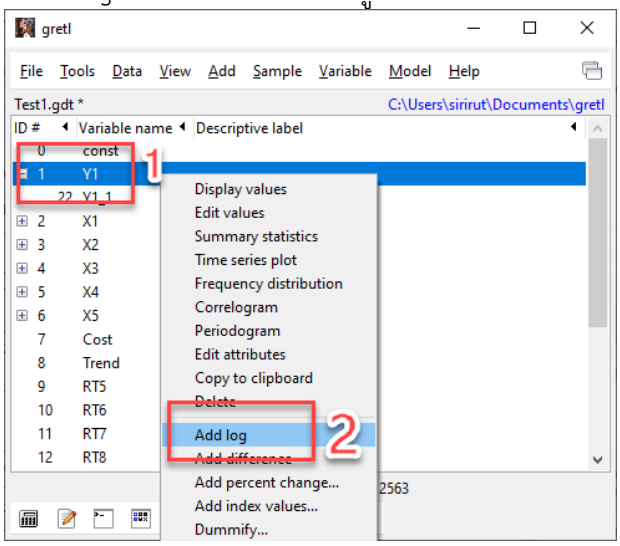


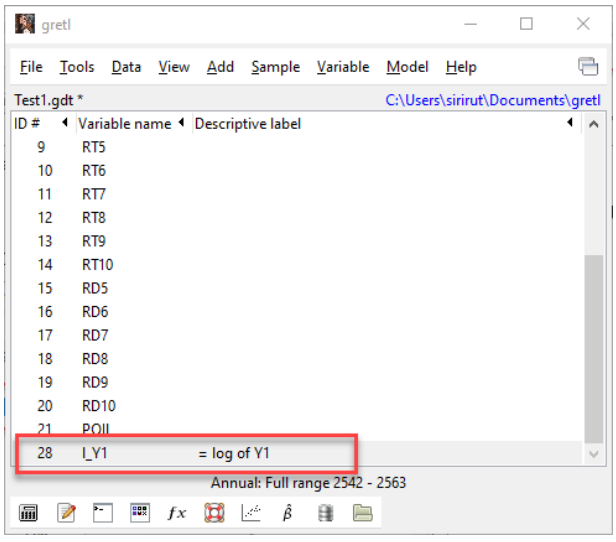
2) เลือก Sheet ที่เราได้เตรียมข้อมูลไว้ แล้วกด OK ข้อมูลที่เตรียมไว้ได้นำเข้าโปรแกรมแล้ว ดังรูป





3) หากต้องการแปลงข้อมูลให้อยู่ในรูป log ทำได้ดังนี้ คลิกขวาที่ตัวแปรที่จะทำการแปลงข้อมูล แล้วเลือก Add log ก็จะได้ตัวแปรใหม่ ดังรูป





4. ขั้นตอนพยากรณ์ผลผลิตข้าวหอมมะลินาปี

4.1 วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ

ตัวแปรทั้งหมดต้องนำมาศึกษาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ว่ามีความสัมพันธ์กันหรือไม่ เพื่อเป็นการคัดเลือกตัวแปรเข้าไปในแบบจำลอง โดยการทดสอบค่าสัมประสิทธิ์ของสหสัมพันธ์เพียร์สัน (Pearson Product Moment Correlation Coefficient) โดยพิจารณาจากการทดสอบสมมติฐานดังนี้

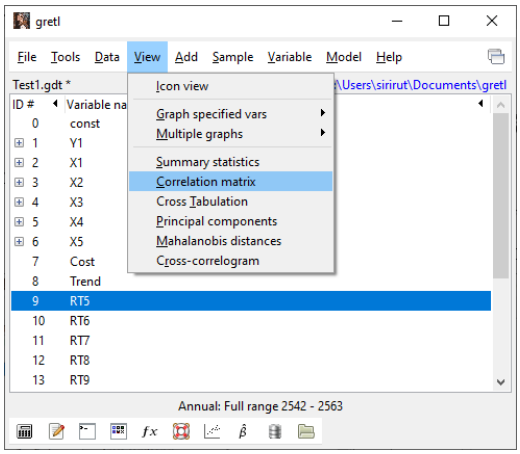
H_0 : ตัวแปรตามและตัวแปรอิสระไม่มีความสัมพันธ์กัน ($\rho = 0$)

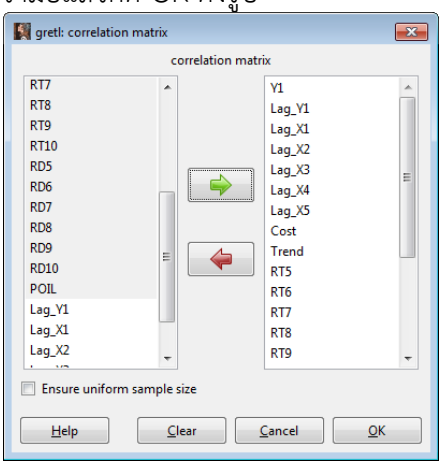
H_1: ตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กัน ($\rho \neq 0$)

ปฏิเสธสมมติฐานหลัก H_0 เมื่อ $t \geq t_{\frac{\alpha}{2},n-2}$ และ $t \leq -t_{\frac{\alpha}{2},n-2}$ หรือ Significance ของสถิติทดสอบ $t < \alpha$

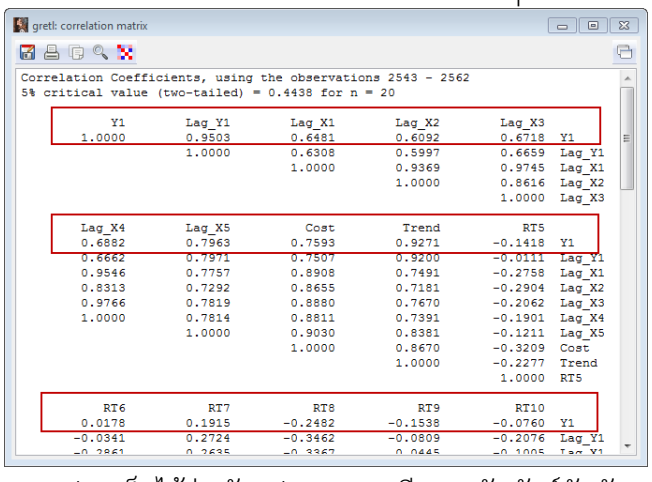
โดยจะใช้โปรแกรม Gretl ในการหา Correlation มีวิธีการดังนี้

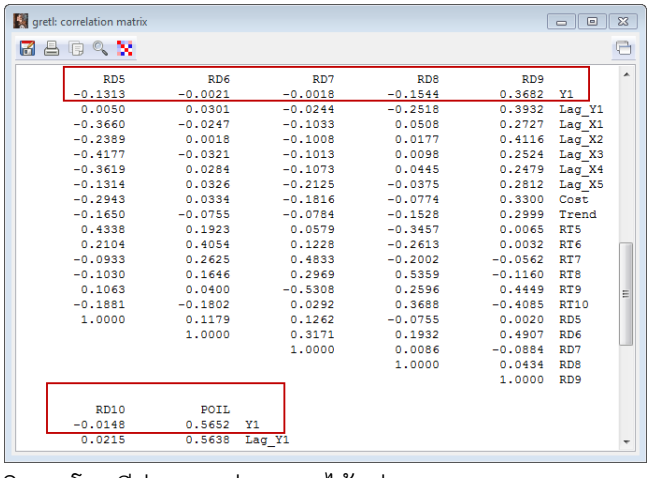
1) หลังจากที่นำข้อมูลเข้าโปรแกรมแล้ว ไปที่ View เลือก Correlation matrix จะได้หน้าต่าง Correlation matrix แล้วเลือกตัวแปรอิสระที่จะทดสอบ ไปด้านขวามือแล้วกด OK ดังรูป





2) จะได้ Correlation ของตัวแปรตามกับตัวแปรอิสระ โดยดูจากแถวแรกของแต่ละกลุ่ม ซึ่งจะเป็นความสัมพันธ์ระหว่าง Y1 กับตัวแปรอิสระต่างๆ





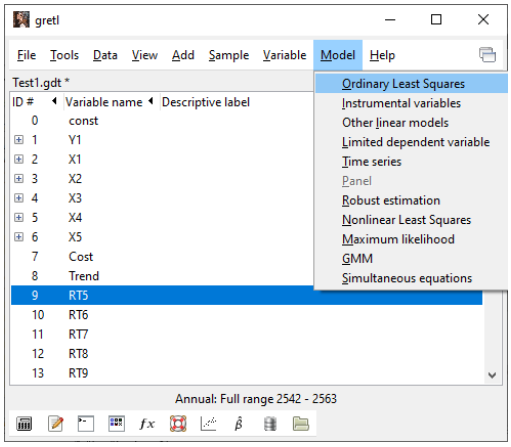
จากรูปจะเห็นได้ว่า ตัวแปรตาม Y1 มีความสัมพันธ์กับตัวแปรอิสระ โดยมีค่ามากกว่า 0.05 ได้แก่

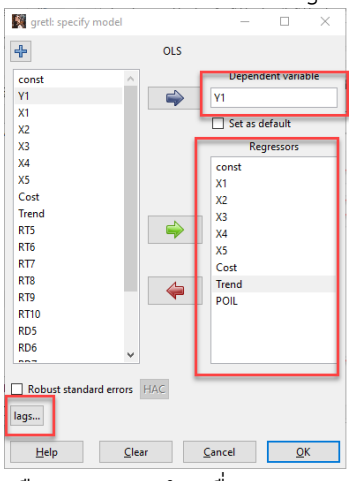
Lag_Y1 Lag_X1 Lag_X2 Lag_X3 Lag_X4 Lag_X5 Cost Trend POIL

4.2 การสร้างแบบจำลองการพยากรณ์

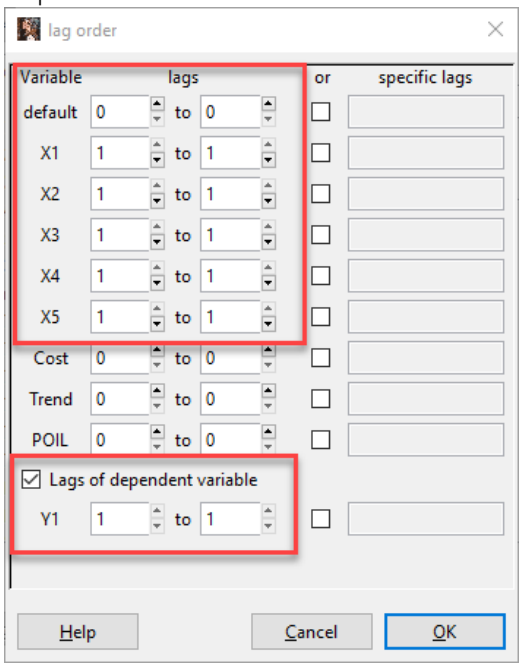
ในตัวอย่างนี้จะเป็นการสร้างแบบจำลองการพยากรณ์ เนื้อที่เพาะปลูกข้าวหอมมะลิ จังหวัดอุบลราชธานี โดยจะนำตัวแปรตามและตัวแปรอิสระที่มีความสัมพันธ์กัน จากข้อ 4.1 มาสร้างแบบจำลอง ด้วยวิธีการวิเคราะห์การถดถอยเชิงพหุ (Multiple Regression Analysis) เพื่อจะทดสอบค่าสัมประสิทธิ์การถดถอยทุกตัวพร้อมกันของแต่ละแบบจำลอง และตรวจสอบว่าตัวแปรอิสระทุกตัวสามารถอธิบายการผันแปรตัวแปรตามในแต่ละแบบจำลองได้หรือไม่ โดยมีวิธีการดังนี้

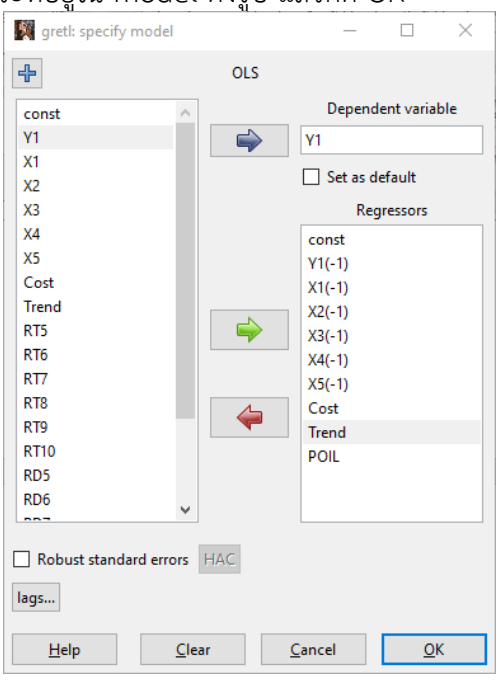
1) จากโปรแกรม Gretl ไปที่ model> Ordinary Least Square จะได้หน้าต่าง specify model จากนั้นเลือกตัวแปรตาม และตัวแปรอิสระ ที่มีความสัมพันธ์ใส่ช่องด้านขวามือ จากนั้นกด lags



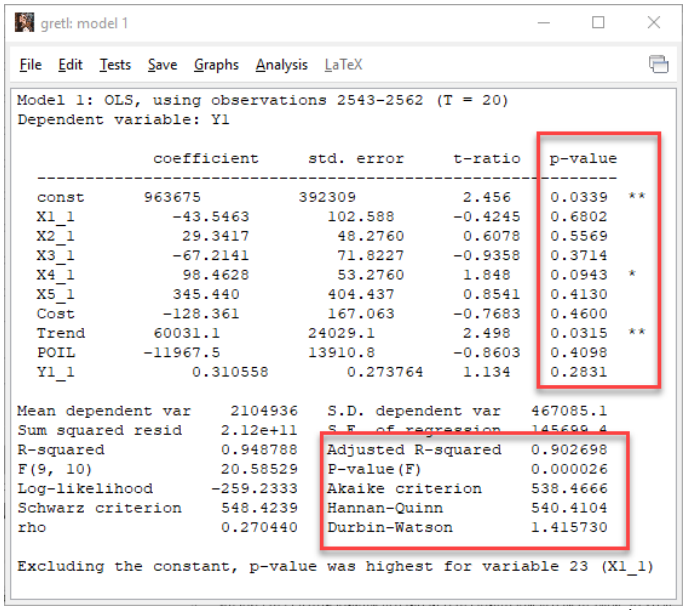


2) จะได้หน้าต่าง lag order โดยตัวแปรราคา X1 – X5 เลือก 1 และทำเครื่องหมายถูกตรง lags of dependent variable Y1 แล้วกด OK จะได้ตัวแปรอิสระที่อยู่ใน model ดังรูป แล้วกด OK





3) จะได้หน้าต่าง model ดังรูป



จากข้อมูล Model 1 พบว่า แบบจำลองนี้ มีค่า

Adjusted R Square 0.902698

หมายความว่า แบบจำลองนี้ตัวแปรอิสระ สามารถอธิบายตัวแปรตาม ได้ถึง 90.26%

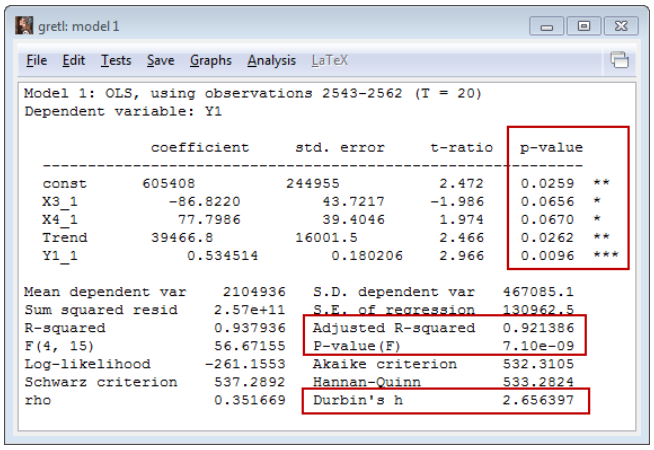
F-test มีค่า Significance F 0.000026

Significant F = 0.000026 < α= 0.05 แสดงว่า ปฏิเสธ H_0 ยอมรับ H_1 หมายความว่า แบบจำลองนี้ มีค่าพารามิเตอร์ของตัวแปรอิสระบางตัว ไม่เท่ากับศูนย์ คือ แบบจำลองนี้สามารถอธิบายตัวแปรตามที่เราศึกษาได้

แต่จากการพิจารณาตัวแปรอิสระแต่ละตัวมี 2 ตัวเท่านั้น คือ X4_1 และ Trend ที่สามารถอธิบายตัวแปรตามได้

ดังนั้นควรพิจารณา แบบจำลองอื่น โดยทดองตัดตัวแปรที่มีค่า P-value สูง ออกทีละตัว หรือเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามสูงเข้าทีละตัว แล้วทอสอบ Run Regression จนตัวแปรอิสระที่อยู่ในแบบจำลองทุกตัว มีค่า P-value น้อยกว่า 0.01 หรือ 0.05 หรือ 0.10 ซึ่งเป็นช่วงที่สามารถยอมรับความคลาดเคลื่อนได้ ที่ 1% หรือ5% หรือ 10%

4) ทำข้อ 1) ใหม่ โดยทดลองเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม และให้ค่า P-value น้อยกว่า 0.01 หรือ 0.05 หรือ 0.10 ซึ่งเป็นช่วงที่สามารถยอมรับความคลาดเคลื่อนได้ ที่ 1% หรือ5% หรือ 10% จะได้ model ดังนี้



จากข้อมูล Model 3 พบว่า แบบจำลองนี้ มีค่า

Adjusted R Square 0.921386

หมายความว่า แบบจำลองนี้ตัวแปรอิสระ สามารถอธิบายตัวแปรตาม ได้ถึง 92.13%

F-test มีค่า Significance F 0.0000000071

Significant F = 0.0000000071< α= 0.05 แสดงว่า ปฏิเสธ H_0 ยอมรับ H_1 หมายความว่า แบบจำลองนี้ มีค่าพารามิเตอร์ของตัวแปรอิสระบางตัว ไม่เท่ากับศูนย์ คือ แบบจำลองนี้สามารถอธิบายตัวแปรตามที่เราศึกษาได้

และเมื่อพิจารณาตัวแปรอิสระแต่ละตัว มีค่า P-value น้อยกว่า 0.01 หรือ 0.05 หรือ 0.10 ซึ่งเป็นช่วงที่สามารถยอมรับความคลาดเคลื่อนได้ ที่ 1% หรือ5% หรือ 10% ซึ่งทุกตัวสมารถอธิบายตัวแปรตามได้

มีค่า Durbin-Watson เท่ากับ 2.656

สมการของแบบจำลองนี้คือ

เป็นรูปแบบจำลอง Linear Model $Y1 = 605,408 + 0.53(Y1_1) - 86.82(X3_1) + 77.79(X4_1) + 39,466.8(Trend)$

4.3 การตรวจสอบเงื่อนไขของการวิเคราะห์การถดถอยเชิงพหุ

4.3.1 การตรวจสอบว่าค่าความคลาดเคลื่อนเป็นอิสระกัน หรือไม่

โดยใช้สถิติทดสอบ Durbin-Watson

ขั้นตอนที่ 1 กำหนดสมมติฐาน

H_0 : ค่าความคลาดเคลื่อน e_i และ e_j เป็นอิสระกัน

H_1 : ค่าความคลาดเคลื่อน e_i และ e_j ไม่เป็นอิสระกัน

ยอมรับสมมติฐานหลัก H_0 เมื่อค่า DW เข้าใกล้ 2 หรืออยู่ระหว่าง 1.5 - 2.5

ถ้าค่า DW < 1.5 แสดงว่าค่าความสัมพันธ์ของ e_i และ e_j อยู่ในทิศทางบวก และถ้า DW มีค่าเข้าใกล้ 0 แสดงว่า e_i และ e_j มีความสัมพันธ์กันมาก

ถ้าค่า DW > 2.5 แสดงว่าค่าความสัมพันธ์ของ e_i และ e_j อยู่ในทิศทางลบ และถ้า DW มีค่าเข้าใกล้ 4 แสดงว่า e_i และ e_j มีความสัมพันธ์กันมาก

ขั้นตอนที่ 2 สรุปผล

จาก Model 3 พบว่าให้ค่า Durbin-Watson เท่ากับ 2.656 นั่นคือ ค่า DW= 2.656 > 2.5 แสดงว่าค่าความสัมพันธ์ของค่าความคลาดเคลื่อนมีความสัมพันธ์กันมาก ซึ่งทำให้เกิดปัญหาสหสัมพันธ์ของตัวคลาดเคลื่อน (Autocorrelation) การแก้ไขคือ การแก้ไขคือ เพิ่มตัวแปรอิสระ เปลี่ยนรูปแบบฟังก์ชัน หรือแก้ด้วยวิธี GLS, Cochrane-Orcutt, การใส่ลำดับของ AR(p), HAC เป็นต้น

4.3.2 การตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระ เป็นอิสระกันหรือไม่

ซึ่งการตรวจสอบนี้จะตรวจสอบเฉพาะการวิเคราะห์การถดถอยที่มีรูปแบบเป็น linear โดยใช้สถิติทดสอบ Variance Inflation Factor (VIF)

ขั้นตอนที่ 1 กำหนดสมมติฐาน

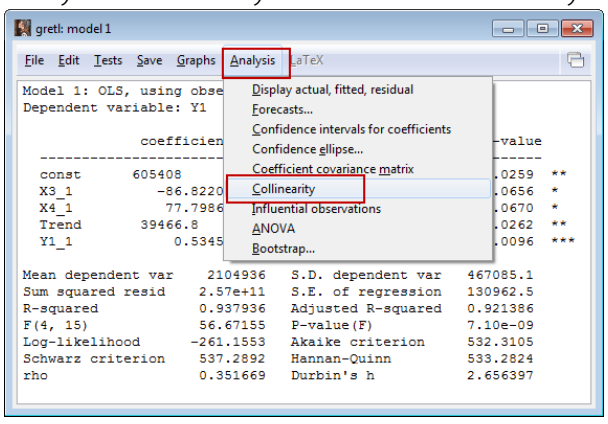
H_0 : ตัวแปรอิสระ X_i เป็นอิสระกันกับตัวแปรอิสระอื่น

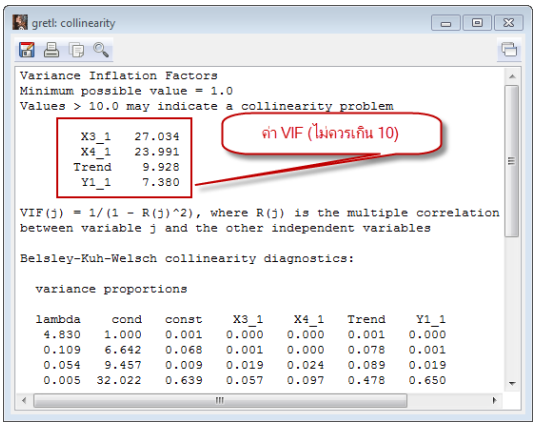
H_1 : ตัวแปรอิสระ X_i ไม่เป็นอิสระกันกับตัวแปรอิสระอื่น

ยอมรับสมมติฐานหลัก H_0 ถ้าหากค่า Tolerance ของตัวแปรอิสระ X_i เข้าใกล้ 1 แสดงว่าตัวแปรเป็นอิสระจากกัน แต่ถ้าค่าใกล้ศูนย์ แสดงว่าเกิดปัญหา Multicollinearity คือยอมรับสมมติฐานรอง H_1

สำหรับค่า VIF ถ้ามีค่ามากแสดงว่า ตัวแปรอิสระ X_i มีความสัมพันธ์กับตัวแปรอิสระอื่นมาก ถ้าค่า VIF มากกว่า 10 มีหลักฐานเกี่ยวกับความสัมพันธ์ที่มากพอ ตัวแปรอิสระ X_i มีความสัมพันธ์กับตัวแปรอิสระอื่น

ขั้นตอนที่ 2 ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก Analysis>Collinearity จะได้หน้าต่าง Collinearity ดังรูป





จากการทดสอบจะเห็นได้ว่า X3_1 และ X4_1 มีค่า VIF สูงมาก คือ 27.034 ลั 23.991 ตามลำดับ แสดงว่าเกิดปัญหาความสัมพันธ์เชิงเส้นของตัวแปรอิสระ (Multicolliearity) การแก้ไขคือ ตัดตัวแปรอิสระ X3_1 หรือ X4_1 ออก แล้วทำการสร้างแบบจำลองอีกครั้ง

4.3.3 การตรวจสอบว่าค่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่ หรือไม่

โดยใช้สถิติทดสอบของ Breusch-Pagan Test หรือ LM test

ขั้นตอนที่ 1 กำหนดสมมติฐาน

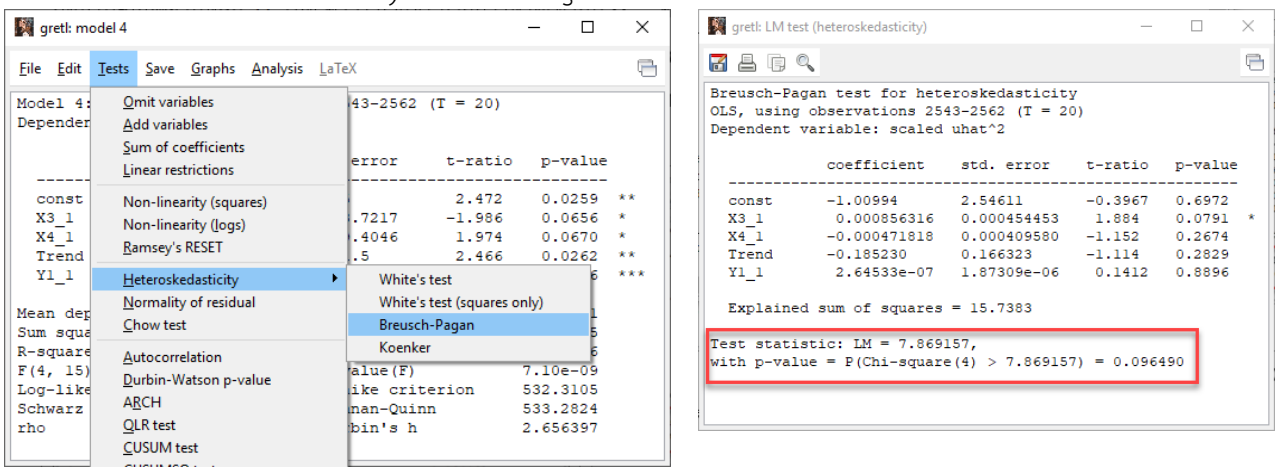
H_0 : ค่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่/ Homoscedasticity

H_1: ค่าความแปรปรวนของความคลาดเคลื่อนมีค่าไม่คงที่/ Heteroscedasticity

ยอมรับสมมติฐานหลัก H_0 เมื่อ ค่า p-value มากกว่าหรือเท่ากับ $\alpha_{0.05}$ คือ ไม่เกิดปัญหา Heteroscedasticity

ปฏิเสธสมมติฐานหลัก H_0 และยอมรับ H_1 เมื่อ ค่า p-value น้อยกว่า $\alpha_{0.05}$ คือ เกิดปัญหา Heteroscedasticity

ขั้นตอนที่ 2 ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก Test> Heteroscedasticity>Breusch-Pagan จะได้หน้าต่าง LM test



จากการทดสอบ LM test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.096490 ซึ่ง มากกว่า $\alpha_{0.05}$ นั่นคือ ไม่เกิดปัญหา Heteroscedasticity

หากเกิดปัญหา Heteroscedasticity แก้ไขโดย เพิ่มตัวแปรอิสระ หรือแก้ด้วยวิธี Weighted least squares (WLS) เป็นต้น

4.3.4 การตรวจสอบว่าความคลาดเคลื่อน e มีการแจกแจงปกติ หรือไม่

โดยใช้สถิติทดสอบ Normality test

ขั้นตอนที่ 1 กำหนดสมมติฐาน

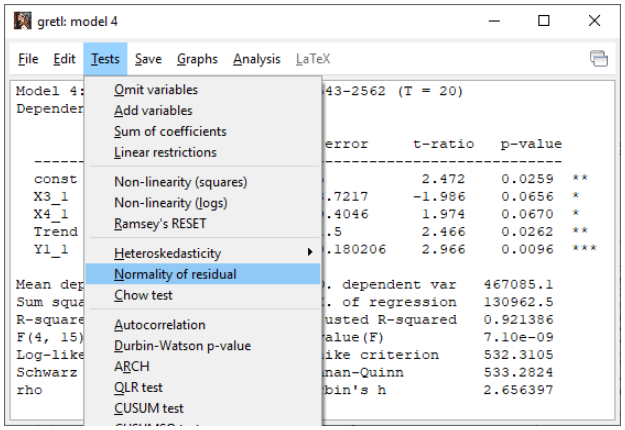
H_0 : ค่าความคลาดเคลื่อนมีการแจกแจงปกติ

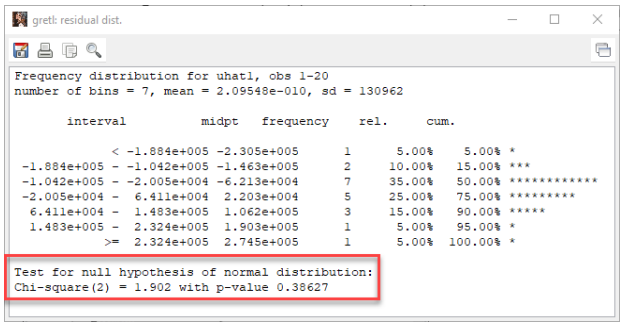
H_1: ค่าความคลาดเคลื่อนไม่ได้มีการแจกแจงปกติ

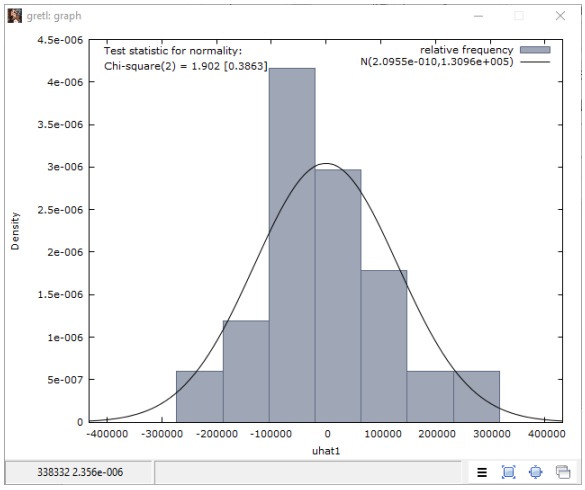
ยอมรับสมมติฐานหลัก H_0 เมื่อ ค่า p-value มากกว่าหรือเท่ากับ $\alpha_{0.05}$ คือ มีการแจกแจงปกติ

ปฏิเสธสมมติฐานหลัก H_0 และยอมรับ H_1 เมื่อ ค่า p-value น้อยกว่า $\alpha_{0.05}$ คือ ไม่ได้มีการแจกแจงปกติ

ขั้นตอนที่ 2 ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก Test>Normality of residual จะได้หน้าต่าง residual dist. และ Graph ดังรูป







จากการทดสอบ Normality test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.38627 ซึ่ง มากกว่า $\alpha_{0.05}$ นั่นคือ ค่าความคลาดเคลื่อนมีการแจกแจงปกติ

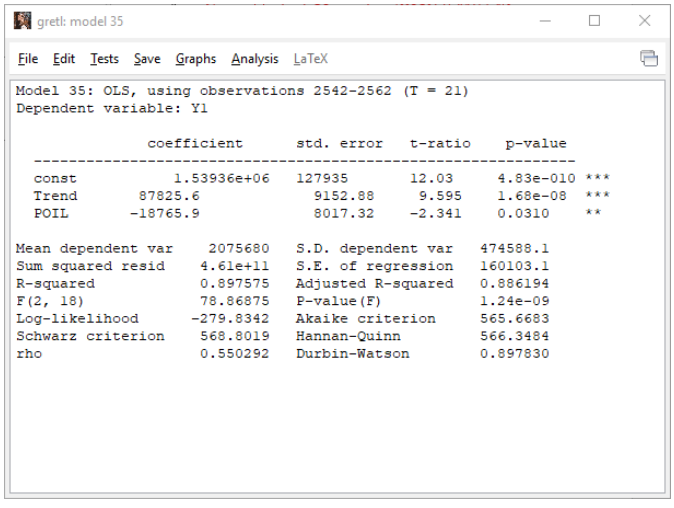
4.3.5 การตรวจสอบว่าค่าเฉลี่ยของความคลาดเคลื่อนเป็นศูนย์ หรือไม่

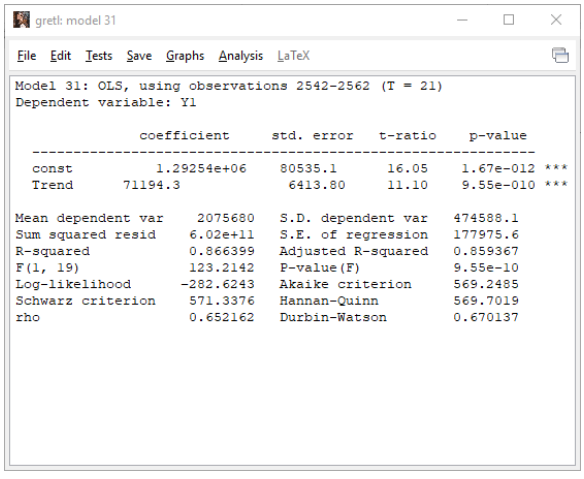
การประมาณค่าพารามิเตอร์ $\beta_1, \beta_2, ..., \beta_k$ โดยให้ผลบวกกำลังสองของค่าความคลาดเคลื่อนมีค่าต่ำที่สุด จะทำให้ $\sum e_i = 0$

ค่าเฉลี่ยของความคลาดเคลื่อน = $E(e) = \frac{\sum e_i}{n} = \frac{0}{n} = 0$

ดังนั้น เงื่อนไขข้อนี้ไม่จำเป็นต้องตรวจสอบ เนื่องจากเมื่อใช้วิธีกำลังสองน้อยที่สุด (Least Square) ในการประมาณค่าพารามิเตอร์ $\beta_1, \beta_2, ..., \beta_k$ จะทำให้ $E(e) = 0$ เสมอ

4.3.6 ตัวอย่างแบบจำลอง





gretl: model 40

File Edit Tests Save Graphs Analysis LaTeX

```
Model 40: OLS, using observations 2543-2562 (T = 20)
Dependent variable: Y1

              coefficient      std. error    t-ratio   p-value
  ------------------------------------------------------------
  const          1.45677e+06   133781        10.89     1.61e-08 ***
  Trend          86648.0         9292.29      9.325    1.24e-07 ***
  POIL          -18644.5         9070.36     -2.056    0.0577   *
  X4_1             105.901         41.8741    2.529    0.0231   **
  X3_1            -103.273         47.6822   -2.166    0.0468   **

Mean dependent var    2104936   S.D. dependent var   467085.1
Sum squared resid     3.18e+11  S.E. of regression   145707.1
R-squared             0.923174  Adjusted R-squared   0.902687
F(4, 15)              45.06169  P-value(F)           3.47e-08
Log-likelihood       -263.2890  Akaike criterion     536.5781
Schwarz criterion     541.5567  Hannan-Quinn         537.5499
rho                   0.287641  Durbin-Watson        1.328409
```

gretl: model 42

File Edit Tests Save Graphs Analysis LaTeX

```
Model 42: OLS, using observations 2543-2562 (T = 20)
Dependent variable: l_Y1

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       12.8780      0.900955     14.29     1.57e-010 ***
  l_X4_1       0.238834    0.118995      2.007    0.0619    *
  l_Trend      0.345223    0.0760595     4.539    0.0003    ***
  l_POIL      -0.425497    0.166168     -2.561    0.0209    **

Mean dependent var   14.53683   S.D. dependent var   0.218809
Sum squared resid    0.186349   S.E. of regression   0.107920
R-squared            0.795147   Adjusted R-squared   0.756737
F(3, 16)             20.70160   P-value(F)           9.36e-06
Log-likelihood       18.37989   Akaike criterion    -28.75978
Schwarz criterion   -24.77685   Hannan-Quinn        -27.98227
rho                  0.518840   Durbin-Watson        0.851615

Log-likelihood for Y1 = -272.357
```

4.4 การพยากรณ์ข้อมูล

4.4.1 สร้างแบบจำลองและตรวจสอบเงื่อนไขของการวิเคราะห์การถดถอยเชิงพหุ

gretl: model 42

File Edit Tests Save Graphs Analysis LaTeX

```
Model 42: OLS, using observations 2543-2562 (T = 20)
Dependent variable: l_Y1

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       12.8780      0.900955     14.29     1.57e-010 ***
  l_X4_1       0.238834    0.118995      2.007    0.0619    *
  l_Trend      0.345223    0.0760595     4.539    0.0003    ***
  l_POIL      -0.425497    0.166168     -2.561    0.0209    **

Mean dependent var   14.53683   S.D. dependent var   0.218809
Sum squared resid    0.186349   S.E. of regression   0.107920
R-squared            0.795147   Adjusted R-squared   0.756737
F(3, 16)             20.70160   P-value(F)           9.36e-06
Log-likelihood       18.37989   Akaike criterion    -28.75978
Schwarz criterion   -24.77685   Hannan-Quinn        -27.98227
rho                  0.518840   Durbin-Watson        0.851615

Log-likelihood for Y1 = -272.357
```

เป็นรูปแบบจำลอง Double log Model

$$Log(Y1) = 12.88 + 0.24Log(X4_1) + 0.35Log(Trend) - 0.43Log(Poil)$$

Adjusted R Square 0.7567

F-test มีค่า Significance F 0.00000936

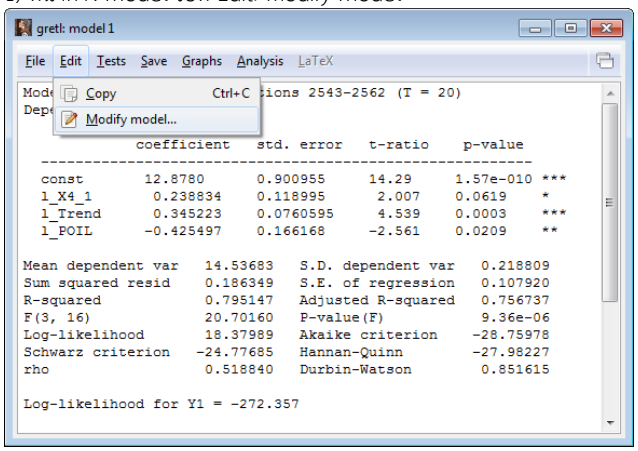
1. ค่า DW = 0.85 (มีปัญหา Autocorrelation)
2. ค่า VIF ตัวแปรอิสระทุกตัว มีค่าต่ำกว่า 10 แสดงว่า ไม่เกิดปัญหา Multicollieearity

```
l_X4_1    3.100
l_Trend   4.205
l_POIL    4.564
```

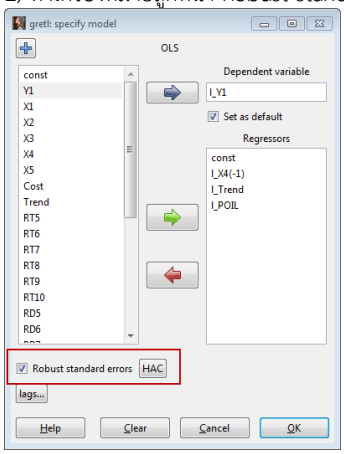
3. LM test ค่า p-value = 0.6792 ซึ่ง มากกว่า $\alpha_{0.05}$ นั่นคือ ไม่เกิดปัญหา Heteroscedasticity
4. Normality test ค่า p-value เท่ากับ 0.6299 ซึ่ง มากกว่า $\alpha_{0.05}$ นั่นคือ ค่า e มีการแจกแจงปกติ

4.2.2 จากการตรวจสอบเงื่อนไข พบว่าเกิดปัญหา Autocorrelation ดังนั้นต้องทำการแก้ปัญหาก่อนดังนี้

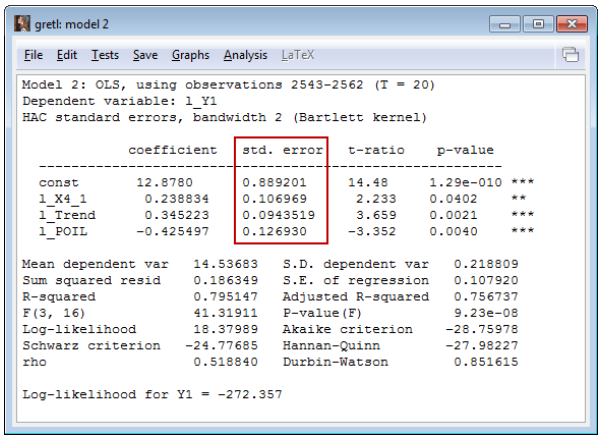
1) หน้าต่าง model ไปที่ Edit>Modify model



2) ทำเครื่องหมายถูกหน้า Robust standard error แล้วกด OK



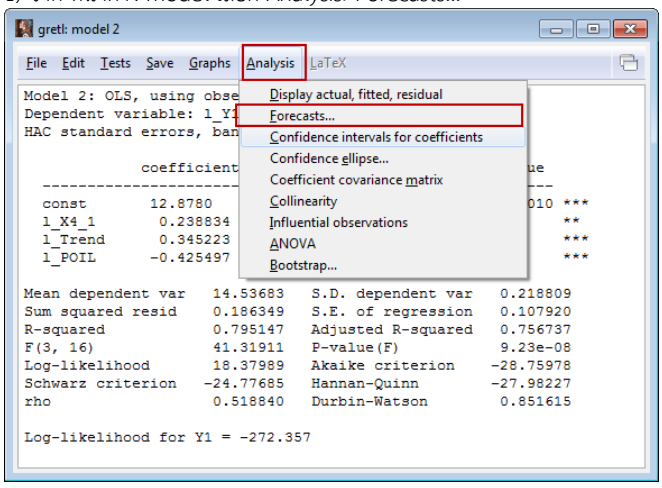
3) จะเห็นได้ว่าค่า Std. error ลดลง เมื่อเทียบกับ ยังไม่ได้แก้ไขรูปด้านบน แต่ค่า พารามิเตอร์ และ DW ยังเหมือนเดิม



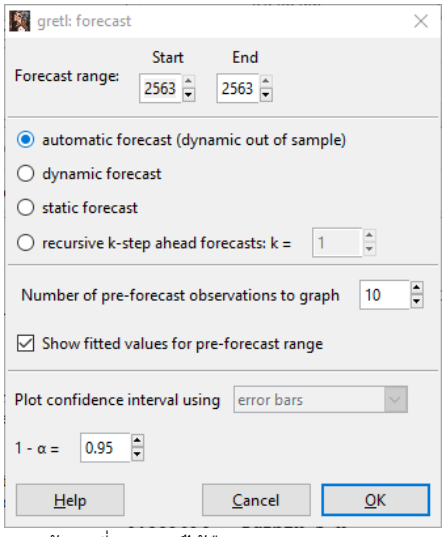
สามารถดูวิธีแก้ไขปัญหา Autocorrelation หรือแบบอื่นๆ ได้ที่
https://cj007blog.files.wordpress.com/2020/04/01-regression-analysis.pdf
จัดทำโดย ผู้ช่วยศาสตราจารย์ ดร.เฉลิมพล จตุพร

4.2.3 พยากรณ์ข้อมูล

1) จาก หน้าต่าง model เลือก Analysis>Forecasts...



2.2) จะได้หน้าต่าง Forecast กด OK



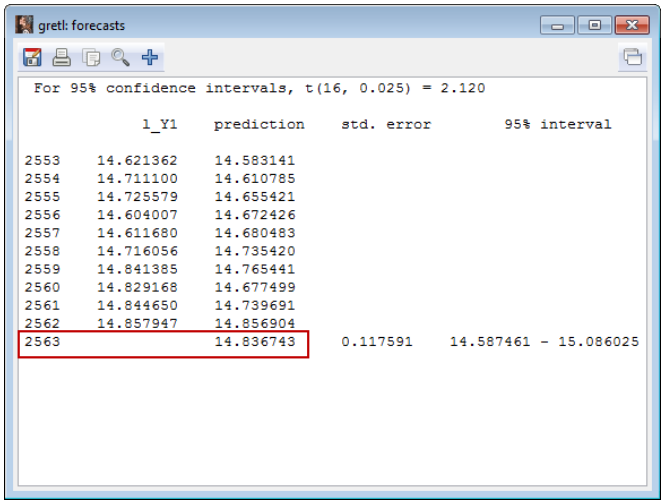
2.3) จะได้หน้าต่างข้อมูลที่พยากรณ์

2.4) ข้อมูลที่พยากรณ์ได้คือ

$Log(Y1)_{2563}$ = 14.836743

แต่เราพยากรณ์ค่า Y1 ดังนั้นต้องทำการถอดค่า Log ออก จะได้

$Y1_{2563}$ = 2,776,615



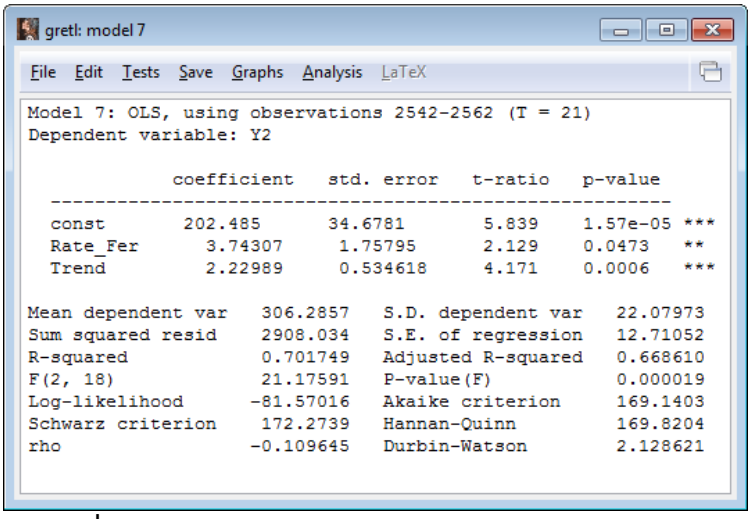
การพยากรณ์ ผลผลิตต่อไร่ข้าวหอมมะลินาปี (Y2) จังหวัดอุบลราชธานี

1. **ตรวจสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ**

   ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม คือ

ตัวแปรอิสระ	ค่า Correlation
Cost	0.6513
Seed	0.7516
Rate_Fer	0.7129
Trend	0.7690
Lag_X1	0.5429
Lag_X2	0.5240
Lag_X3	0.5210
Lag_X4	0.5593

2. **สร้างแบบจำลอง**



สมการที่ได้คือ Y2 = 202.485+3.743(Rate_Fer)+2.229(Trend)

Adj R^2 = 0.6686

2.1 ตรวจสอบว่าค่าความคลาดเคลื่อนเป็นอิสระกัน หรือไม่

Durbin-Watson เท่ากับ 2.12 นั่นคือ ค่า DW= 2.12 < 2.5 แสดงว่าค่าความสัมพันธ์ของค่าความคลาดเคลื่อนเป็นอิสระกัน ดังนั้นไม่เกิดปัญหา Autocorrelation

2.2 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระ เป็นอิสระกันหรือไม่

```
Variance Inflation Factors
Minimum possible value = 1.0
Values > 10.0 may indicate a collinearity problem

    Rate_Fer    1.362
       Trend    1.362
```

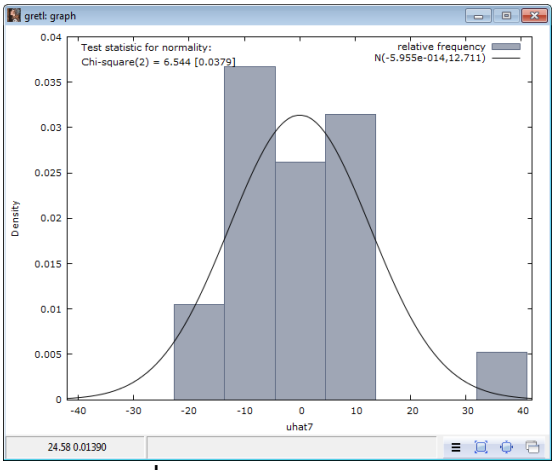
จากการทดสอบจะเห็นตัวแปรอิสระ Rate_Fer Trend มีค่า VIF คือ 1.362 1.362 ตามลำดับ มีค่าไม่เกิน 10 แสดงว่า ไม่เกิดปัญหา Multicolliearity

2.3 ตรวจสอบว่าค่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่ หรือไม่

```
Breusch-Pagan test for heteroskedasticity
OLS, using observations 2542-2562 (T = 21)
Dependent variable: scaled uhat^2

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       5.13306      5.59326      0.9177    0.3709
  Rate_Fer   -0.159225     0.283542    -0.5616    0.5813
  Trend      -0.0691783    0.0862290   -0.8023    0.4329

  Explained sum of squares = 8.15056

Test statistic: LM = 4.075278,
with p-value = P(Chi-square(2) > 4.075278) = 0.130336
```

จากการทดสอบ LM test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.130336 ซึ่ง มากกว่า $\alpha_{0.05}$ นั่นคือ ไม่เกิดปัญหา Heteroscedasticity

2.4 ตรวจสอบว่าความคลาดเคลื่อนมีการแจกแจงปกติ หรือไม่




```
Frequency distribution for uhat7, obs 1-21
number of bins = 7, mean = -5.95502e-014, sd = 12.7105

       interval          midpt    frequency     rel.      cum.

          < -13.545    -18.089          2      9.52%     9.52% ***
-13.545 - -4.4559    -9.0004          7     33.33%    42.86% ***********
-4.4559 -  4.6331   0.088623          5     23.81%    66.67% ********
 4.6331 -  13.722     9.1776          6     28.57%    95.24% **********
 13.722 -  22.811     18.267          0      0.00%    95.24%
 22.811 -  31.900     27.356          0      0.00%    95.24%
        >=  31.900     36.445          1      4.76%   100.00% *

Test for null hypothesis of normal distribution:
Chi-square(2) = 6.544 with p-value 0.03794
```

จากการทดสอบ Normality test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.03794 ซึ่ง มากกว่า $\alpha_{0.10}$ นั่นคือ ค่าความคลาดเคลื่อนมีการแจกแจงปกติ

3. ค่าพยากรณ์ที่ได้จากแบบจำลอง

Y2 = 335.04 กิโลกรัมต่อไร่