## การพยากรณ์ข้าวนาปี

# มีขั้นตอนดังนี้

## 1. ข้อตกลงเบื้องต้น (Assumption)

 1.1 ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเนื้อที่เพาะปลูกปีที่พยากรณ์หรือเนื้อที่เพาะปลูก ปีที่ t ได้แก่

1) เนื้อที่เพาะปลูกในปีที่แล้ว หรือเนื้อที่เพาะปลูกปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า เกษตรกรจะ ตัดสินใจ ขยายหรือลดเนื้อที่เพาะปลูก ต้องคำนึงถึงเนื้อที่เพาะปลูกที่เคยทำมาในปีที่แล้ว

2) ราคาเฉลี่ย ราคาสูงสุด ราคาต่ำสุด ราคา ณ เดือนที่ผลผลิตออกมาก (เดือนพฤศจิกายน) ที่เกษตรกรขายได้ในปีที่แล้วหรือปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า เกษตรกรจะตัดสินใจในการขยายหรือลดเนื้อที่ เพาะปลูกขึ้นอยู่กับราคาดังกล่าว โดยคาดว่าหากราคาที่เกษตรกรขายได้ในปีที่แล้วมีแนวโน้มเพิ่มขึ้นก็จะจูงใจให้ เกษตรกรขยายเนื้อที่เพาะปลูกเพิ่มขึ้นในปีถัดไป และหากราคาที่เกษตรกรขายได้ในปีที่แล้วมีแนวโน้มลดลง เกษตรกรก็จะลดเนื้อที่เพาะปลูกลงในปีถัดไป

3) ราคาพืชแข่งขันปีที่แล้ว ซึ่งสามารถปลูกได้ในพื้นที่เดียวกัน ภายใต้ข้อตกลงเบื้องต้นว่า เช่น ถ้าราคาอ้อยโรงงานเมื่อปีที่แล้วมีแนวโน้มเพิ่มขึ้นหรือให้ผลตอบแทนที่ดีกว่า เกษตรกรก็จะปรับเปลี่ยนจากการ ปลูกข้าวไปปลูกอ้อยโรงงานแทน และหากราคาอ้อยโรงงานเมื่อปีที่แล้วมีแนวโน้มลดลงหรือผลตอบแทนน้อยกว่า เกษตรกรก็จะหันมาปลูกข้าวแทนอ้อยโรงงาน

4) ต้นทุนการผลิตปีที่แล้ว ภายใต้ข้อตกลงเบื้องต้นว่า หากต้นทุนการผลิตเมื่อปีที่แล้วมีแนวโน้ม ลดลง เกษตรกรจะตัดสินใจขยายเนื้อที่เพาะปลูกเพิ่มขึ้นเพื่อเพิ่มการผลิตและได้รับผลตอบแทนที่มากขึ้น และหาก ต้นทุนเมื่อปีที่แล้วมีแนวโน้มเพิ่มขึ้นและไม่คุ้มค่ากับผลตอบแทนที่ได้รับ เกษตรกรก็จะตัดสินใจลดเนื้อที่ปลูกข้าว และปรับเปลี่ยนไปปลูกพืชอื่นที่ให้ผลตอบแทนที่ดีกว่า

5) ปริ้มาณฝนตก และจำนวนวันฝนตก ภายใต้ข้อตกลงเบื้องต้นว่า หากปริมาณน้ำฝนหรือจำนวน วันที่ฝนตกในปีปัจจุบันเพียงพอต่อการเพาะปลูกเกษตรกรก็อาจจะขยายเนื้อที่เพาะปลูกเพิ่มขึ้น

6) ราคาน้ำมันดีเซล ภายใต้ข้อตกลงเบื้องต้นว่า หากราคาน้ำมันดีเซลต่ำลง เกษตรกรจะเพิ่ม ปริมาณการใช้เครื่องจักร เครื่องยนต์ในการเพาะปลูกมากขึ้น เช่น การเตรียมดิน ทำให้มีการขยายเนื้อที่เพาะปลูก มากขึ้น เนื่องจากมีความสะดวกและรวดเร็วขึ้น

7) แนวโน้มเวลา (Time Trend) เป็นตัวแปรอิสระที่กำหนดให้แทนการเปลี่ยนแปลงของเทคโนโลยี คือ เมื่อระยะเวลาผ่านไป มีผลให้เทคโนโลยีการผลิตเปลี่ยนแปลงไปด้วย ซึ่งอาจจะมีผลทำให้เนื้อที่เพาะปลูก เพิ่มขึ้น

8) ดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิต เป็นตัวแปรอิสระที่สะท้อนภาวะการเจริญเติบโตทาง เศรษฐกิจ โดยทั่วไปจะใส่ไว้ในแบบจำลองเพื่อถ่วงค่า (Deflate) ราคาอื่น เพราะขณะที่ราคาพืชผลการเกษตร สูงขึ้น อาจจะเป็นภาพลวงตาว่าสินค้านั้นมีราคาดีขึ้น หากค่าดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิตสูงขึ้นด้วย

1.2 ตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามผลผลิตต่อไร่ปีที่พยากรณ์หรือผลผลิตต่อไร่ปีที่ t ได้แก่ 1) ราคาเฉลี่ย ราคาสูงสุด ราคาต่ำสุด ราคา ณ เดือนที่ผลผลิตออกมาก (เดือนพฤศจิกายน) ที่เกษตรกรขายได้ในปีที่แล้วหรือปีที่ t-1 ภายใต้ข้อตกลงเบื้องต้นว่า ผลผลิตต่อไร่จะเพิ่มขึ้นหรือลดลงขึ้นอยู่กับการ ดูแลเอาใจใส่ของเกษตรกร หากราคาที่เกษตรกรขายได้ในปีที่ที่แล้วมีแนวโน้มเพิ่มขึ้นก็จะจูงใจให้เกษตรกรบำรุง ดูแลรักษาดีส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้น และหากราคาที่เกษตรกรขายได้ในปีแล้วมีแนวโน้มลดลง เกษตรกรก์จะ ดูแลเอาใจใส่ลดลงอาจส่งผลให้ผลผลิตต่อไร่ลดลง 2) ดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิต เป็นตัวแปรอิสระที่สะท้อนภาวะการเจริญเติบโตทางเศรษฐกิจ โดยทั่วไปจะใส่ไว้ในแบบจำลองเพื่อถ่วงค่า (Deflate) ราคาอื่น เพราะขณะที่ราคาพืชผลการเกษตรสูงขึ้น อาจจะ เป็นภาพลวงตาว่าสินค้านั้นมีราคาดีขึ้น หากค่าดัชนีราคาผู้บริโภค และดัชนีราคาผู้ผลิตสูงขึ้นด้วย

3) ต้นทุนการผลิตปีที่พยากรณ์ ภายใต้ข้อตกลงเบื้องต้นว่า หากต้นทุนการผลิต เช่น ราคาปุ๋ย ยากำจัด วัชพืช เมื่อปีที่พยากรณ์มีแนวโน้มราคาลดลง เกษตรกรสามารถซื้อปัจจัยดังกล่าว เพื่อไปบำรุงดูแลรักษาต้นข้าว ส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้น หรือในทางกลับกัน หากแนวโน้มราคาที่เกษตรกรขายได้เพิ่มขึ้น เกษตรกรก็อาจมี แรงจูงใจในการเพิ่มต้นทุนการผลิต เพื่อให้ได้ผลผลิตต่อไร่ที่เพิ่มขึ้น

4) ปริมาณฝนตก และจำนวนวันฝนตก ภายใต้ข้อตกลงเบื้องต้นว่า ปริมาณน้ำฝน จะมีผลต่อผลผลิตต่อไร่ หากปริมาณน้ำฝนในปีใดมีมากสม่ำเสมอ ผลผลิตต่อไร่น่าจะมีแนวโน้มเพิ่มมากขึ้น และในทางตรงกันข้าม เมื่อเกิด ภาวะฝนแล้งและฝนทิ้งช่วง จะทำให้ได้ผลผลิตลดลง

5) อัตราการใช้เมล็ดพันธุ์ และการใส่ปุ๋ยปีที่พยากรณ์ ภายใต้ข้อตกลงเบื้องต้นว่า หากมีการใช้เมล็ดพันธุ์ และการใส่ปุ๋ยเพิ่มขึ้น จะส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้นด้วย

6) ราคาน้ำมันดีเซล ภายใต้ข้อตกลงเบื้องต้นว่า ราคาน้ำมันดีเซลต่ำลง เกษตรกรจะเพิ่มปริมาณการใช้ เครื่องจักร เครื่องยนต์ ในการดูแลรักษามากขึ้น ส่งผลให้ผลผลิตต่อไร่เพิ่มขึ้นด้วย

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

### 2.กำหนดตัวแปรในแบบจำลอง

2.1 ตัวแปรตาม (Dependent Variable) เป็นตัวแปรที่ต้องการพยากรณ์ ประกอบด้วย

- Y1<sub>t</sub> คือ เนื้อที่เพาะปลูกข้าวนาปี ปีที่ t
- Y2<sub>t</sub> คือ ผลผลิตต่อไร่ข้าวนาปี ปีที่ t

**2.2 ตัวแปรอิสระ** (Independent Variable) เป็นตัวแปรที่มีสมมติฐานว่ามีอิทธิพลต่อการเปลี่ยนแปลง ของตัวแปรที่จะทำการพยากรณ์ 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 เดือนที่
	ผลผลิตออกมามาก ปีที่ t-1
Lag_X5 หรือ X5_1	คือ ราคาอ้อยโรงงาน ที่เกษตรกรขายได้ ปีที่ t-1
Cost <sub>t</sub>	คือ ต้นทุนการผลิตข้าวขาวดอกมะลิ 105 ภาคตะวันออกเฉียงเหนือ ปีที่ t
Seedt	คือ อัตราการใช้เมล็ดพันธุ์ข้าวต่อไร่ของข้าวนาปี ปีที่ t
Rate_Fer <sub>t</sub>	คือ อัตราการใส่ปุ๋ยต่อเนื้อที่เพาะปลูกต่อไร่ของข้าวนาปี ปีที่ t
Poil <sub>t</sub>	คือ ราคาน้ำมันดีเซล ปีที่ t
Trendt	คือ แนวโน้มเวลา ปีที่ t
RT5 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงพฤษภาคม ปีที่ t
RT6 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงมิถุนายน ปีที่ t
RT7 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงกรกฎาคม ปีที่ t
RT8 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงสิงหาคม ปีที่ t
RT9 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงกันยายน ปีที่ t
RT10 <sub>t</sub>	คือ ปริมาณน้ำฝนสะสมเดือนเมษายนถึงตุลาคม ปีที่ t
RD5 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงพฤษภาคม ปีที่ t
RD6 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงมิถุนายน ปีที่ t
RD7 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงกรกฎาคม ปีที่ t
RD8 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงสิงหาคม ปีที่ t
RD9 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงกันยายน ปีที่ t
RD10 <sub>t</sub>	คือ จำนวนวันฝนตกสะสมเดือนเมษายนถึงตุลาคม ปีที่ t

# 3. จัดเตรียมข้อมูลเพื่อการพยากรณ์

### 3.1 นำข้อมูลตัวแปรอิสระและตัวแปรตามที่ได้จากการสืบค้นมาจัดเรียงตามปีน้อยไปมาก ตัวอย่าง ที่ 3.1 การจัดเตรียมข้อมูล เพื่อพยากรณ์ผลผลิตข้าวนาปี ของจังหวัดอบุลราชธานี

rı	100		J.1	1119	UV16V	1909	106	าช่าย	6110	NOI	11996	MPIM		1919		ВЛ	υv	VI (	911	ยุบ	619	IUC	1112					
	А	В	С	D	Е	F	G	н	1	J	к	L	М	N	0	Ρ	Q	R	S	Т	U	х	Y	Z	AA	AB	AC	AD
1	year	Y1	X1	X2	Х3	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

🕅 gretl		-		$\times$	📓 gretl: open file				×
<u>File</u> <u>T</u> ools <u>D</u> ata <u>V</u> iew <u>A</u> dd <u>S</u> a	ample <u>V</u> ariable <u>M</u> odel <u>H</u> elp			•	🔪 🔪 83 ตู่มือพยากรถม์	giño Gretl			
<u>O</u> pen data	User file	Ctrl+O sirirut\[	Documents	\gretl	<u>P</u> laces	Name	▲ Size	Modified	*
Append data	Sample file			1	C Search	01-regression-analysis.pdf	3.5 MB	12:38	
Save data Ct	rl+S				Recently Used	07-time-series-forecasting.pdf	2.4 MB	12:41	
Save data <u>a</u> s	1. datatest.gdt				🛅 gretl	(g) 10-handbook.pdf	12.4 ME	3 09:51	
Export data	2. ตัวอย่างอุบลนาบ2.gat				irirut.OAE.000	gretl-2019b-64.exe	28.6 ME	3 09:52	
Send To	4 Run adt				Desktop	g Manual.pdf	6.1 MB	16:14	
New data set	rl+N 5 data monthtime sere50 61 vis				ap Local Disk (C:)	G Test1.gdt	9.8 kB	14:53	
	6 Run savat I-si sa visv				I DATA (D:)	🐏 การใช้งาน GRETL console.docx	15.1 kB	12:18	
Ciear data set					HP_RECOVERY (E:)	🕙 การพยากรณ์ ด้วย Gretl.doc	1.4 MB	08:11	=
Working directory					🥔 DVD RW Drive (F:)	🕙 การพยากรณ์ ด้วย Gretl ร่าง.doc	1.7 MB	08:11	
Script files	•				Info_CAI (\\cai-srv2)	🐏 ตัวอย่างอุบคนาปี.xls	312.3 ki	3 08:16	
Session files						🐏 ตัวอย่างอุบคนาปี 2.xls	309.2 ki	3 08:12	
Databases	<b>b</b>					6 ตัวอย่างอุบลนาปี ร่าง.gdt	4.2 kB	13:58	
Function packages	<b>b</b>					😤 แนวปฏิบัติที่ดีการ์โซโปรแกรม Gretl เพื่อการวิจัย.pdf	935.6 ki	3 12:40	
₹ Ctr	rl+Q				irirut (\\w_zone) (Y:)				
									Ŧ
							all files (*.*)		•
🖩 🌶 🏲 🗮 fx 💢 🏄	β 🕮 🖿						Cancel	<u>O</u> pen	
2) เลือก Sheet ที่	เราได้เตรียมข้อมลไว้ แ	เล้วกด C	)K ข้อ	อมล	ที่เตรียมไว้ได้นํ	าเข้าโปรแกรมแล้ว ดังรเ	]		

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

4 1 ป

📓 gretl: sp	readsheet im	port	8
	Start import	at:	
column:	1 * r	ow: 1	*
(A)			
	Sheet to imp	oort:	
data			~
data2			
data¥1			-
	<u>C</u> ancel	<u>о</u> к	

								ข		
🕅 gi	etl							-		$\times$
<u>F</u> ile	Tools	<u>D</u> ata	View	<u>A</u> dd	Sample	<u>V</u> ariable	Model	Help		-
Test1.	gdt						C:\User	s\sirirut\[	ocumen	ts\gret
ID # 4	Variab	ile nam	e ◀ De	scripti	/e label					4 ^
0	const									
1	Y1									
2	X1									
3	X2									
4	X3									
5	X4									
6	X5									
7	Cost									
8	Trend									
9	RT5									
10	RT6									
11	RT7									
12	RT8									
13	RT9									~
				Ann	ual: Full ra	nge 2542 -	2563			
	D 10		6.	<b>9</b> -8	1 B	0 D				
3131	💌 L		1	8 <b></b> 8	<u> </u>	H (C)				

3) หากต้องการแปลงข้อมูลให้อยู่ในรูป log ทำได้ดังนี้ คลิกขวาที่ตัวแปรที่จะทำการแปลงข้อมูล แล้วเลือก Add log ก็จะได้ตัวแปรใหม่ ดังรูป

gretl		_		×	1	gretl							_		$\times$
<u>File Tools Data View A</u> dd Sample Variable	<u>M</u> odel	<u>H</u> elp		6	<u>F</u> ile	<u>T</u> oo	ls <u>D</u> ata	a <u>V</u> iew	<u>A</u> dd	<u>S</u> ample	<u>V</u> ariable	<u>M</u> odel	<u>H</u> elp		
Test1.gdt *	C:\User	s\sirirut\	Docume	ents\gretl	Test	.gdt *						C:\User	s\sirirut\[	Documents	s∖gretl
ID #      Variable name      Descriptive label				• •	ID #	< V	ariable r	name 🖣	Descrip	otive label					• ^
0 const					9	- I	RT5								
= 1 Y1 U Direleventure					1	0 I	RT6								
22 V1 1 Display values					1	1 1	RT7								
					1	2 1	RT8								
						3 I	RT9								
4 X3      Frequency distribution						4 I	RT10								
Correlogram						5 1	RD5								
b X5     Periodogram						6 I	RD6								
Cost Edit attributes						/ 1	KD7								
Copy to clipboard						0 I 0 I									
10 PT6 Delete						9 I 0 I	2010								
						1 1									
						8 1	¥1		= log (	of Y1	_				~
Add percent change	25.62			Ŧ		-			A	unde Erell en	mme 2542	2562			
Add index values	2503					~		_	Ann		inge 2042 -	2305			
Dummify						2	<u>-</u>	fx		β					

### 4. ขั้นตอนพยากรณ์ผลผลิตข้าวหอมมะลินาปี

### 4.1 วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ

ตัวแปรทั้งหมดต้องนำมาศึกษาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ ว่ามีความสัมพันธ์กันหรือไม่ เพื่อ เป็นการคัดเลือกตัวแปรเข้าไปในแบบจำลอง โดยการทดสอบค่าสัมประสิทธิ์ของสหสัมพันธ์เพียร์สัน (Pearson Product Moment Correlation Coefficient) โดยพิจารณาจากการทดสอบสมมติฐานดังนี้

 $\mathbf{H}_{_{0}}$ : ตัวแปรตามและตัวแปรอิสระไม่มีความสัมพันธ์กัน (ho= 0)

 $\mathrm{H_{i}}$ : ตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กัน (ho 
eq 0)

ปฏิเสธสมมติฐานหลัก  $\mathbf{H}_0$  เมื่อ  $t \ge t_{\frac{\alpha}{2},n-2}$  และ  $t \le -t_{\frac{\alpha}{2},n-2}$  หรือ Significance ของสถิติทดสอบ  $\mathbf{t} < \alpha$ 

โดยจะใช้โปรแกรม Gretl ในการหา Correlation มีวิธีการดังนี้

 หลังจากที่น ำข้อมูลเข้าโปรแกรมแล้ว ไปที่ View เลือก Correlation matrix จะได้หน้าต่าง Correlation matrix แล้วเลือกตัวแปรอิสระที่จะทดสอบ ไปด้านขวามือแล้วกด OK ดังรูป





2) จะได้ Correlation ของตัวแปรตามกับตัวแปรอิสระ โดยดูจากแถวแรกของแต่ละกลุ่ม ซึ่งจะเป็น ความสัมพันธ์ระหว่าง Y1 กับตัวแปรอิสระต่างๆ

gretl: correlation matrix		🛐 greti: correlation matrix
<b>7</b> 8 6 9 X	0	
Correlation Coefficients, using the observations 2543 - 2562 5% critical value (two-tailed) = 0.4438 for n = 20	*	RD5         RD6         RD7         RD8         RD9           -0.1313         -0.0021         -0.0018         -0.1544         0.3682         Y1
Y1         Lag_Y1         Lag_X1         Lag_X2           1.0000         0.9503         0.6481         0.6092           1.0000         0.6308         0.5997           1.0000         0.9359         1.0000           1.0000         1.0000         1.9369	Lag_X3 0.6718 Y1 0.9745 Lag_Y1 0.8616 Lag_X2 1.0000 Lag_X3	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$
Lag_X4         Lag_X5         Cost         Trend           0.6882         0.7963         0.7593         0.9271           0.6662         0.7971         0.7507         0.9200           0.9546         0.7757         0.8908         0.7491           0.813         0.7292         0.8655         0.7161           0.9766         0.7819         0.8880         0.7670           1.0000         0.7814         0.8811         0.7391           1.0000         0.9030         0.8381         1.0000           0.8670         1.0000         1.0000         0.8670	RT5           -0.1418         Y1           -0.0111         Lag_Y1           -0.2758         Lag_X2           -0.2062         Lag_X3           -0.1201         Lag_X5           -0.1201         Lag_X5           -0.3209         Cost           -0.2277         Trend           1.0000         RTS	-0.1650 -0.0755 -0.0784 -0.1528 0.2999 Trend 0.4338 0.1923 0.0579 -0.3457 0.0065 RTS 0.2104 0.4054 0.1228 -0.2613 0.0032 RT6 -0.0933 0.2625 0.4833 -0.2002 -0.0562 RT7 -0.1030 0.1646 0.2969 0.5359 -0.1160 RT8 0.1063 0.0400 -0.5308 0.2596 0.4449 RT9 -0.1881 -0.1802 0.0292 0.3668 -0.4085 RT10 1.0000 0.1179 0.1262 -0.0755 0.0002 RD5 1.0000 0.0197 0.1932 0.4907 RD6 1.0000 0.0086 -0.0884 RD7 1.0000 0.099
RI6         RI7         RI8         RT9           0.0178         0.1915         -0.2482         -0.1538           -0.0341         0.2724         -0.3462         -0.0809           -0.2861         0.2635         -0.3367         0.0445	RT10 -0.0760 Y1 -0.2076 Lag_Y1 -0.1005 Teg V1	RD10 POIL -0.0148 0.5652 Y1 0.0215 0.5638 Lag_Y1

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 โดย จะนำตัวแปรตามและตัวแปรอิสระที่มีความสัมพันธ์กัน จากข้อ 4.1 มาสร้างแบบจำลอง ด้วยวิธีการวิเคราะห์การ ถดถอยเชิงพหุ (Multiple Regression Analysis) เพื่อจะทดสอบค่าสัมประสิทธิ์การถดถอยทุกตัวพร้อมกันของแต่ ละแบบจำลอง และตรวจสอบว่าตัวแปรอิสระทุกตัวสามารถอธิบายการผันแปรตัวแปรตามในแต่ละแบบจำลองได้ หรือไม่ 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

🦉 lag o	rder						×
Variable	-	I	ags		-	or	specific lags
default X1 X2	0	* *	to 0 to 1	•			
	1	• • •	to 1 to 1				
X5 Cost	1	• • •	to 1				
Frend POIL	0	* * *	to 0 to 0		• • •		
✓ Lags Y1	of d	epende	ent va to 1	riable ;			
<u>H</u> el	lp				<u>C</u> a	incel	<u>O</u> K

3) จะได้หน้าต่าง model ดังรูป

💐 gretl: m	odel 1					-		>	<
<u>F</u> ile <u>E</u> dit	<u>T</u> ests <u>S</u> ave <u>G</u>	<u>ì</u> raphs <u>A</u> nal	lysis L	aTeX				Ę	3
Model 1: Depender	: OLS, using nt variable:	observat Yl	tions	2543-2562	(T = 20)	_			
	coeff	icient	sto	l. error	t-ratio	F	-value		l
const	963675		3923	09	2.456	6	.0339	**	L
X1 1	-43	.5463	1	.02.588	-0.4245		.6802		L
X2_1	29	.3417		48.2760	0.6078		.5569		L
X3_1	-67	.2141		71.8227	-0.9358		.3714		L
X4 1	98	.4628		53.2760	1.848		.0943	*	L
X5 1	345	.440	4	04.437	0.8541		.4130		L
Cost	-128	.361	1	67.063	-0.7683		.4600		I
Trend	60031	.1	240	29.1	2.498		.0315	**	I
POIL	-11967	.5	139	10.8	-0.8603		.4098		I
¥1_1	0	.310558		0.273764	1.134		.2831		l
Mean dep	pendent var	210493	36 5	.D. depend	dent var	467	085.1		1
Sum squa	ared resid	2.12e+1	11 卢	E of reg	greesion	145	600 4		
R-square	ed	0.94878	38 <b>Z</b>	djusted R-	-squared	0.9	02698		
F(9, 10)	)	20.5852	29 E	-value(F)		0.0	00026		
Log-like	elihood	-259.233	33 <b>7</b>	kaike crit	terion	538	.4666		
Schwarz	criterion	548.423	39 F	lannan-Quir	nn	540	.4104		
rho		0.27044	40 I	)urbin-Wats	son	1.4	15730		
Excludir	ng the const	ant, p-va	alue v	as highest	t for varia	able	23 (X	1_1)	

**จากข้อมูล** Model 1 พบว่า แบบจำลองนี้ มีค่า

Adjusted R Square 0.902698

หมายความว่า แบบจ่ำลองนี้ตัวแปรอิสระ สามารถอธิบายตัวแปรตาม ได้ถึง 90.26%

F-test มีค่า Significance F 0.000026

Significant F = 0.000026 < lpha= 0.05 แสดงว่า ปฏิเสธ  $H_0$  ยอมรับ

 $H_1$ หมายความว่า แบบจำลองนี้ มีค่าพารามิเตอร์ของตัวแปรอิสระบางตัว ไม่

เท่ากับศูนย์ คือ แบบจำลองนี้สามารถอธิบายตัวแปรตามที่เราศึกษาได้

แต่จากการพิจารณาตัวแปรอิสระแต่ละตัวมี 2 ตัวเท่านั้น คือ X4\_1 และ Trend ที่สามารถอธิบายตัวแปรตามได้

ดังนั้นควรพิจารณา แบบจำลองอื่น โดยทดงองตัดตัวแปรที่มีค่า P-value สูง ออกทีละตัว หรือเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามสูงเข้าที ละตัว แล้วทอสอบ Run Regression จนตัวแปรอิสระที่อยู่ในแบบจำลองทุก ตัว มีค่า P-value น้อยกว่า 0.01 หรือ 0.05 หรือ 0.10 ซึ่งเป็นช่วงที่สามารถ ยอมรับความคลาดเคลื่อนได้ ที่ 1% หรือ5% หรือ 10%

4) ทำข้อ 1) ใหม่ โดยทดลองเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม และให้ค่า P-value น้อย กว่า 0.01 หรือ 0.05 หรือ 0.10 ซึ่งเป็นช่วงที่สามารถยอมรับความคลาดเคลื่อนได้ ที่ 1% หรือ5% หรือ 10% จะ ได้ model ดังนี้

				LTV			
	sts <u>s</u> ave <u>u</u>	rapris <u>A</u> na	iysis		(T = 0.0)		4
Dependent v	variable:	Y1	tior	18 2543-2562	(1 = 20)		
	coeff	icient	5	std. error	t-ratio	p-value	
const	605408		24	14955	2.472	0.0259	**
X3 1	-86	.8220		43.7217	-1.986	0.0656	*
X4_1	77	.7986		39.4046	1.974	0.0670	*
Trend	39466	. 8	1	6001.5	2.466	0.0262	**
¥1_1	0	.534514		0.180206	2.966	0.0096	***
Mean depend	dent var	21049	36	S.D. depende	ent var	467085.1	
Sum squared	d resid	2.57e+1	11	S.E. of reg	ression	130962.5	
R-squared		0.93793	36	Adjusted R-s	squared	0.921386	
F(4, 15)		56.671	55	P-value(F)		7.10e-09	
Log-likeli	nood	-261.15	53	Akaike crite	erion	532.3105	
Schwarz cr:	iterion	537.289	92	Hannan-Quinr	1	533.2824	-
rho		0.3516	69	Durbin's h		2.656397	

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< lpha= 0.05 แสดงว่า ปฏิเสธ  $H_0$ 

ยอมรับ  $H_1$ หมายความว่า แบบจำลองนี้ มีค่าพารามิเตอร์ของตัวแปร

อิสระบางตัว ไม่เท่ากับศูนย์ คือ แบบจำลองนี้สามารถอธิบายตัวแปรตามที่ เราศึกษาได้

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  $\mathbf{H}_{0}$ : ค่าความคลาดเคลื่อน  $\mathbf{e}_{i}$  และ  $\mathbf{e}_{j}$  เป็นอิสระกัน

 $\mathbf{H}_{i}$ : ค่าความคลาดเคลื่อน  $\mathbf{e}_{i}$  และ  $\mathbf{e}_{j}$  ไม่เป็นอิสระกัน

ยอมรับสมมติฐานหลัก  $\mathbf{H}_{_0}$  เมื่อค่า DW เข้าใกล้ 2 หรืออยู่ระหว่าง 1.5 - 2.5

ถ้าค่า DW < 1.5 แสดงว่าค่าความสัมพันธ์ของ e<sub>i</sub>และ e<sub>j</sub> อยู่ในทิศทางบวก และถ้า DW มีค่าเข้าใกล้ 0 แสดงว่า e<sub>i</sub>และ e<sub>j</sub> มีความสัมพันธ์กันมาก

ถ้าค่า DW > 2.5 แสดงว่าค่าความสัมพันธ์ของ e<sub>i</sub>และ e<sub>j</sub> อยู่ในทิศทางลบ และถ้า DW มีค่าเข้าใกล้ 4 แสดงว่า e<sub>i</sub>และ e<sub>j</sub> มีความสัมพันธ์กันมาก

### **ขั้นตอนที่ 2** สรุปผล

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

### 4.3.2 การตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระ เป็นอิสระกันหรือไม่

ซึ่งการตรวจสอบนี้จะตรวจสอบเฉพาะการวิเคราะห์การถดถอยที่มีรูปแบบเป็น linear โดยใช้สถิติ ทดสอบ Variance Inflation Factor (VIF)

**ขั้นตอนที่ 1** กำหนดสมมติฐาน

 $\mathbf{H}_{0}$ : ตัวแปรอิสระ  $\mathbf{X}_{i}$  เป็นอิสระกันกับตัวแปรอิสระอื่น

 $\mathbf{H}_{\mathrm{l}}$ : ตัวแปรอิสระ  $\mathbf{X}_{\mathrm{i}}$  ไม่เป็นอิสระกันกับตัวแปรอิสระอื่น

ยอมรับสมมติฐานหลัก  $H_0$ ถ้าหากค่า Tolerance ของตัวแปรอิสระ  $X_i$  เข้าใกล้ 1 แสดงว่าตัวแปรเป็น อิสระจากกัน แต่ถ้าค่าใกล้ศูนย์ แสดงว่าเกิดปัญหา Multicollinearity คือยอมรับสมมติฐานรอง  $H_1$ 

สำหรับค่า VIF ถ้ามีค่ามากแสดงว่า ตัวแปรอิสระ  $\mathbf{X}_{i}$ มีความสัมพันธ์กับตัวแปรอิสระอื่นมาก ถ้าค่า VIF มากกว่า 10 มีหลักฐานเกี่ยวกับความสัมพันธ์ที่มากพอ ตัวแปรอิสระ  $\mathbf{X}_{i}$ มีความสัมพันธ์กับตัวแปรอิสระอื่น

**ขั้นตอนที่ 2** ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก

Analysis>Collinearity จะได้หน้าต่าง Collinearity ดังรูป

🛐 gretl: model 1	I I I I I I I I I I I I I I I I I I I	
<u>File Edit Iests Save Graphs</u> <u>A</u> nalysis LaTeX		8
Model 1: 0LS, using obse       Display actual, fitted, residual         Dependent variable: Y1       Coefficien         coefficien       Confidence intervals for coefficients         const       605408         X3_1       -86.8220         Zefficient covariance matrix       Coefficient covariance matrix         X4_1       77.7986         Trend       39466.8         Y1_1       0.5345         Mean dependent var       2104936         Sum squared resid       2.57e+11         R-squared       0.937936         Adjusted R-squared       0.937936         F(4, 15)       56.67155	value       Variance Inflation Factors         Minimum possible value = 1.0         Values > 10.0 may indicate a collinearity problem         0259 **         .0656 *	E
Log-likelihood -261.1553 Akaike criterion 53 Schwarz criterion 537.2892 Hannan-Quinn 53 rho 0.351669 Durbin's h 2.	2,3105 3,2824 656397 656397 6,642 6,650 6,000 6,000 6,642 6,642 0,068 0,001 0,000	<b>.</b>

จากการทดสอบจะเห็นได้ว่า X3\_1 และ X4\_1 มีค่า VIF สูงมาก คือ 27.034 ลั 23.991 ตามลำดับ แสดงว่าเกิดปัญหาความสัมพันธ์เชิงเส้นของตัวแปรอิสระ (Multicolliearity) การแก้ไขคือ ตัดตัวแปร อิสระ X3\_1 หรือ X4\_1 ออก แล้วทำการสร้างแบบจำลองอีกครั้ง

### 4.3.3 การตรวจสอบว่าค่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่ หรือไม่

โดยใช้สถิติทดสอบของ Breusch-Pagan Test หรือ LM test

**ขั้นตอนที่ 1** กำหนดสมมติฐาน

 $\mathbf{H}_{0}$ : ค่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่/ Homoscedasticity

 $\mathbf{H}_1$ : ค่าความแปรปรวนของความคลาดเคลื่อนมีค่าไม่คงที่/ Heteroscedasticity ยอมรับสมมติฐานหลัก  $\mathbf{H}_0$  เมื่อ ค่า p-value มากกว่าหรือเท่ากับ  $\boldsymbol{\alpha}_{0.05}$  คือ ไม่เกิดปัญหา Heteroscedasticity ปฏิเสธสมมติฐานหลัก  $\mathbf{H}_0$  และยอมรับ  $\mathbf{H}_1$ เมื่อ ค่า p-value น้อยกว่า  $\boldsymbol{\alpha}_{0.05}$  คือ เกิดปัญหา Heteroscedasticity

ขั้นตอนที่ 2 ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก

Test> Heteroscedasticity>Breusch-Pagan จะเดหนาตาง LM test

🛐 gretl: m	odel 4		- 0	×	📓 gretl: LM test (heteroskedasticity) — 🗆	$\times$
<u>F</u> ile <u>E</u> dit	<u>Tests</u> <u>S</u> ave <u>G</u> raphs <u>A</u> nalysis <u>L</u> a	TeX		-		8
Model 4: Depender	Omit variables Add variables	43-2562 (T = 20)			Breusch-Pagan test for heteroskedasticity OLS, using observations 2543-2562 (T = 20) Dependent variable: scaled uhat^2	
	Sum of coefficients Linear restrictions	error t-ratio	o p-value		coefficient std.error t-ratio p-va	lue
const X3_1 X4_1 Trend	Non-linearity (squares) Non-linearity ( <u>l</u> ogs) <u>R</u> amsey's RESET	2.472 .7217 -1.986 .4046 1.974 .5 2.466	0.0259 0.0656 0.0670 0.0262	** * * **	const         -1.00994         2.54611         -0.3967         0.65           X3 1         0.000856316         0.000454453         1.884         0.07           X4_1         -0.000471818         0.000409580         -1.152         0.26           Trend         -0.185230         0.166233         -1.114         0.27	972 791 * 574 829
Y1_1 Mean der	Heteroskedasticity Normality of residual	White's test White's test (squares	6 ; only)	***	Y1_1 2.64533e-07 1.87309e-06 0.1412 0.88 Explained sum of squares = 15.7383	96
Sum square R-square F(4, 15) Log-like Schwarz rho	<u>C</u> how test <u>Autocorrelation</u> <u>Durbin-Watson p-value</u> A <u>R</u> CH <u>Q</u> LR test <u>C</u> USUM test	Breusch-Pagan Koenker alue(F) ike criterion nan-Quinn bin's h	5 7.10e-09 532.3105 533.2824 2.656397		<pre>Test statistic: LM = 7.869157, with p-value = P(Chi-square(4) &gt; 7.869157) = 0.096490</pre>	

จากการทดสอบ LM test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.096490 ซึ่ง มากกว่า  $lpha_{_{0.05}}$  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

 $\mathbf{H}_{0}$ : ค่าความคลาดเคลื่อนมีการแจกแจงปกติ

H<sub>1</sub>: ค่าความคลาดเคลื่อนไม่ได้มีการแจกแจงปกติ

ยอมรับสมมติฐานหลัก  $\mathbf{H}_{_0}$  เมื่อ ค่า p-value มากกว่าหรือเท่ากับ  $\pmb{lpha}_{_{0.05}}$  คือ มีการแจกแจงปกติ

ปฏิเสธสมมติฐานหลัก  $\mathbf{H}_{_0}$  และยอมรับ  $\mathbf{H}_{_1}$ เมื่อ ค่า p-value น้อยกว่า  $\pmb{lpha}_{_{0.05}}$  คือ ไม่ได้มีการแจกแจงปกติ

**ขั้นตอนที่ 2** ทำการทดสอบ โดยไปที่หน้าต่าง model ที่ต้องการทดสอบ แล้วเลือก

Test>Normality of residual จะได้หน้าต่าง residual dist. และ Graph ดังรูป

🎉 gretl: mo	del 4		- 🗆	$\times$
<u>F</u> ile <u>E</u> dit	<u>Tests</u> <u>Save</u> <u>G</u> raphs <u>A</u> nalysis <u>L</u> aTeX			
Model 4: Depender	Qmit variables 4 Add variables Sum of coefficients	3-2562 (T = 20) rror t-ratio	p-value	
const X3_1 X4_1 Trend	Non-linearity (squares) Non-linearity (logs) <u>R</u> amsey's RESET	2.472 7217 -1.986 4046 1.974 5 2.466	0.0259 0.0656 0.0670 0.0262	* * * * **
YI_I Mean der Sum squa R-square F(4, 15)	Heteroskedasticity Normality of residual Chow test Autocorrelation Durbin-Watson p-value	dependent var of regression sted R-squared lue (F)	467085.1 130962.5 0.921386 7.10e-09	***
Schwarz rho	A <u>R</u> CH n. QLR test b: <u>C</u> USUM test	an-Quinn in's h	532.3105 533.2824 2.656397	
🕅 gretl: residu	al dist.		- 0	$\times$
<b>7</b> 8 P				8
Frequency number of -1.084e+( -1.042e+( -2.005e+( 6.411e+( 1.483e+(	<pre>distribution for uhat1, obs 1-20 bins = 7, mean = 2.09548e-010, sd erval midpt frequency &lt; -1.884e+005 -2.305e+005 051.042e+005 -1.463e+005 052.005e+004 - 6.213e+004 04 - 6.411e+004 2.203e+004 04 - 1.483e+005 1.062e+005 05 - 2.324e+005 1.903e+005 &gt;= 2.324e+005 2.745e+005</pre>	= 130962 rel. cum. 1 5.00% 5.0 2 10.00% 15.0 7 35.00% 50.0 5 25.00% 75.0 3 15.00% 90.0 1 5.00% 100.0	)0% * )0% *** )0% ******** )0% ****** )0% * *	****
Test for r Chi-square	ull hypothesis of normal distribu (2) = 1.902 with p-value 0.38627	tion:		



จากการทดสอบ Normality test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.38627 ซึ่ง มากกว่า α<sub>0.05</sub> นั่นคือ ค่าความคลาดเคลื่อนมีการแจกแจงปกติ

### 4.3.5 การตรวจสอบว่าค่าเฉลี่ยของความคลาดเคลื่อนเป็นศูนย์ หรือไม่

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

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

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

#### 4.3.6 ตัวอย่างแบบจำลอง

gretl: model 35				—		$\times$
<u>F</u> ile <u>E</u> dit <u>T</u> ests <u>S</u> ave	<u>G</u> raphs <u>A</u> nalysis	<u>L</u> aTeX				e
Model 35: OLS, usin Dependent variable	ng observatio : Yl	ns 2542-2562	(T = 21)			
coe	fficient	std. error	t-ratio	p-value		
const	1.53936e+06	127935	12.03	4.83e-010	***	
Trend 8782	5.6	9152.88	9.595	1.68e-08	***	
POIL -1876	5.9	8017.32	-2.341	0.0310	**	
Mean dependent var	2075680	S.D. depende	ent var	474588.1		
Sum squared resid	4.6le+11	S.E. of reg	ression	160103.1		
R-squared	0.897575	Adjusted R-	squared	0.886194		
F(2, 18)	78.86875	P-value(F)		1.24e-09		
Log-likelihood	-279.8342	Akaike crite	erion	565.6683		
Schwarz criterion	568.8019	Hannan-Quin	n	566.3484		
rho	0.550292	Durbin-Wats	on	0.897830		

Model 31: OLS, using observations 2542-2562 (T = 21)           Dependent variable: Y1	p-value
coefficient         std. error         t-ratio           const         1.29254e+06         80535.1         16.05         1           Trend         71194.3         6413.80         11.10         3           Wean dependent var         2075680         S.D. dependent var         4	p-value
const 1.29254e+06 80535.1 16.05 1 Trend 71194.3 6413.80 11.10 9 Mean dependent var 2075680 S.D. dependent var 4 Sum squared resid 6 02e+11 S.F. of regression 12	67e-012
Trend         71194.3         6413.80         11.10         9           Mean dependent var         2075680         S.D. dependent var         47           Nm squared resid         6.02e+11         S.F. of regression         11	
Mean dependent var 2075680 S.D. dependent var 47 Sum squared resid 6 02e+11 S.F. of regression 17	9.55e-010
Sum squared resid 6 02e+11 S F of regression 1	74588.1
SAW SOUGTER TESTO SIDE SIDE OF TESTERSTON T	7975.6
R-squared 0.866399 Adjusted R-squared 0.	859367
F(1, 19) 123.2142 P-value(F) 9.	55e-10
Log-likelihood -282.6243 Akaike criterion 56	59.2485
Schwarz criterion 571.3376 Hannan-Quinn 56	59.7019
rho 0.652162 Durbin-Watson 0.	670137

11

🦉 gretl: mode	el 40				—		$\times$	gretl: model	42
<u>F</u> ile <u>E</u> dit <u>T</u> e	ests <u>S</u> ave <u>G</u>	araphs <u>A</u> nalysis	<u>L</u> aTeX				•	<u>F</u> ile <u>E</u> dit <u>T</u> e	sts <u>S</u> a
Model 40: Dependent	OLS, usin variable:	g observatio Yl	ons 2543-2562	(T = 20)				Model 42: 0 Dependent	OLS, varia
	coef	ficient	std. error	t-ratio	p-value				co
const	1	.45677e+06	133781	10.89	1.61e-08	***		const 1 X4 1	1
Trend	86648	.0	9292.29	9.325	1.24e-07	***		1 Trend	
POIL	-18644	.5	9070.36	-2.056	0.0577	*		1_POIL	-
X4_1	105	.901	41.8741	2.529	0.0231	**		_	
X3_1	-103	.273	47.6822	-2.166	0.0468	**		Mean depend	dent d res
Mean depen	dent var	2104936	S.D. depende	nt var	467085.1			R-squared	
Sum square	d resid	3.18e+11	S.E. of regr	ession	145707.1			F(3, 16)	
R-squared		0.923174	Adjusted R-s	quared	0.902687			Log-likeli	nooa
F(4, 15)		45.06169	P-value(F)		3.47e-08			schwarz cr.	Iteri
Log-likeli	.hood	-263.2890	Akaike crite	rion	536.5781			1110	
Schwarz cr	iterion	541.5567	Hannan-Quinn		537.5499			Log-likeli	hood
rho		0.287641	Durbin-Watso	n	1.328409				
								1	

🕅 gretl: model 42				-		×
<u>F</u> ile <u>E</u> dit <u>T</u> ests <u>S</u> a	ve <u>G</u> raphs <u>A</u> na	alysis <u>L</u> aTeX				
Model 42: OLS, Dependent varia	using observ ble: 1_Yl	ations 25	43-2562 (T = :	20)		
co	efficient	std. erro	r t-ratio	p-value		
const 1	2.8780	0.900955	14.29	1.57e-010	***	
1 X4 1	0.238834	0.118995	2.007	0.0619	*	
1 Trend	0.345223	0.0760595	4.539	0.0003	***	
1_POIL -	0.425497	0.166168	-2.561	0.0209	**	
Mean dependent	var 14.536	83 S.D.	dependent va	r 0.21880	9	
Sum squared res	id 0.1863	49 S.E.	of regression	n 0.10792	20	
R-squared	0.7951	47 Adju	sted R-square	d 0.75673	37	
F(3, 16)	20.701	60 P-va	lue(F)	9.36e-0	06	
Log-likelihood	18.379	89 Akai	ke criterion	-28.7597	78	
Schwarz criteri	on -24.776	85 Hann	an-Quinn	-27.9822	27	
rho	0.5188	40 Durb	in-Watson	0.85161	15	
Log-likelihood	for Y1 = -27	2.357				

### 4.4 การพยากรณ์ข้อมูล

### 4.4.1 สร้างแบบจำลองและตรวจสอบเงื่อนไขของการวิเคราะห์การถดถอยเชิงพหุ

🛐 gretl: model 42				-	
<u>F</u> ile <u>E</u> dit <u>T</u> ests	<u>S</u> ave <u>G</u> raphs <u>A</u> r	nalysis <u>L</u> aTe	(		e
Model 42: OLS, Dependent vari	using obser iable: 1_Yl	vations 2	543-2562 (T	= 20)	
c	coefficient	std. err	or t-rati	o p-value	
const	12.8780	0.900955	14.29	1.57e-010	***
1 X4 1	0.238834	0.118995	2.007	0.0619	*
1 Trend	0.345223	0.076059	5 4.539	0.0003	***
1_POIL	-0.425497	0.166168	-2.561	0.0209	**
Mean dependent	var 14.53	683 S.D	. dependent	var 0.2188	09
Sum squared re	esid 0.186	349 S.E	. of regres	sion 0.1079	20
R-squared	0.795	147 Adj	usted R-squ	ared 0.7567	37
F(3, 16)	20.70	160 P-v	alue(F)	9.36e-	06
Log-likelihood	1 18.37	989 Aka	ike criteri	on -28.759	78
Schwarz criter	ion -24.77	685 Han	nan-Quinn	-27.982	27
rho	0.518	840 Dur	bin-Watson	0.8516	15
Log-likelihood	i for Y1 = -2	72.357			

เป็นรูปแบบจำลอง Double log Model

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

0.7567

Adjusted R Square

F-test มีค่า Significance F 0.00000936

- 1. ค่า DW = 0.85 (มีปัญหา Autocorrelation)
- ค่า VIF ตัวแปรอิสระทุกตัว มีค่าต่ำกว่า 10 แสดงว่า ไม่ เกิดปัญหา Multicolliearity

 LM test ค่า p-value = 0.6792 ซึ่ง มากกว่า α<sub>0.05</sub> นั่น คือ ไม่เกิดปัญหา Heteroscedasticity

4. Normality test ค่า p-value เท่ากับ 0.6299 ซึ่ง มากกว่า  ${f lpha}_{0.05}$  นั่นคือ ค่า e มีการแจกแจงปกติ

4.2.2 จากการตรวจสอบเงื่อนไข พบว่าเกิดปัญหา Autocorrelation ดังนั้นต้องทำการแก้ปัญหาก่อนดังนี้

1) หน้าต่าง model ไปที่ Edit>Modify model

📓 gretl: model 1							x
<u>File</u> <u>E</u> dit <u>T</u> ests	<u>S</u> ave <u>G</u> raphs	<u>A</u> nalysis <u>I</u>	aTeX				8
Mode Copy	Ctr	H+C lions	2543-2562	(T = 20	)		*
Dept Modify m	nodel						
	coefficient	std.	error t-	ratio	p-value		
const	12.8780	0.900	955 14	.29	1.57e-010	***	
1_X4_1	0.238834	0.118	995 2	.007	0.0619	*	=
1_Trend	0.345223	0.076	0595 4	.539	0.0003	***	-
1_POIL	-0.425497	0.166	168 -2	.561	0.0209	**	
Mean dependent	t var 14.	53683	S.D. depen	dent var	0.21880	9	
Sum squared re	esid 0.1	86349	S.E. of re	gression	0.10792	20	
R-squared	0.7	95147 3	Adjusted R	-squared	0.75673	37	
F(3, 16)	20.	70160	P-value(F)		9.36e-0	06	
Log-likelihoo	d 18.	37989 1	Akaike cri	terion	-28.7597	78	
Schwarz crite:	rion -24.	77685 1	Hannan-Qui	nn	-27.9822	27	
rho	0.5	18840 1	Durbin-Wat	son	0.85161	15	
Log-likelihoo	d for Y1 =	-272.357					+

3) จะเห็นได้ว่าค่า Std. error ลดลง เมื่อเทียบกับ ยังไม่ได้แก้ไขรูปด้านบน แต่อ่า พารรณิตวร์ และ DW ผู้มหมืองเอิง

	แตคา พารามเตอร และ	DW 8	องเหมอ	อนเดม			
ſ	🛐 gretl: model 2						×
	<u>File Edit Tests Save Graphs</u>	<u>A</u> nalysis	<u>L</u> aTeX				8
	Model 2: OLS, using obse Dependent variable: 1_Y1 HAC standard errors, ban	rvation dwidth	ns 2543- 2 (Bart	-2562 (T = )	20) 1)		
	coefficient	std.	error	t-ratio	p-value		
	const 12.8780	0.88	9201	14.48	1.29e-010	***	
	1 X4 1 0.238834	0.10	6969	2.233	0.0402	**	
	1 Trend 0.345223	0.09	943519	3.659	0.0021	***	
	1_POIL -0.425497	0.12	6930	-3.352	0.0040	***	
	Maan dependent war 14		8 D (	Jonondont v	0 2199	0.0	
	Sum equared regid 0.1	96349	S.F. (	of regressi	ar 0.2100	20	
	R-squared 0.7	95147	Adjust	ted R-squar	ed 0.7567	37	
	F(3, 16) 41.	31911	P-valu	ie (F)	9.23e-	08	
	Log-likelihood 18.	37989	Akaike	e criterion	-28.759	78	
	Schwarz criterion -24.	77685	Hannar	n-Quinn	-27,982	27	
	rho 0.5	18840	Durbin	n-Watson	0.8516	15	
	Log-likelihood for Y1 =	-272.35	57				

2) ทำเครื่องหมายถูกหน้า Robust standard error แล้วกด OK



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...

🙀 gretl: model 2					×
<u>File Edit Tests Save Graphs</u>	<u>A</u> nalysis	<u>L</u> aTeX			8
Model 2: OLS, using obse	<u>D</u> ispla	y actual, fitted, residual			
Dependent variable: 1_Y1	<u>Forec</u>	asts			
HAC standard errors, bar	<u>C</u> onfi	dence intervals for coefficients	;		
coefficient	Confi	dence <u>e</u> llipse		ne	
	Coeff	icient covariance <u>m</u> atrix			
const 12.8780	<u>C</u> ollin	earity		010 ***	
1 X4 1 0.238834	Influe	ntial observations		**	
1_Trend 0.345223	ANO	/A		***	
1_POIL -0.425497	Boots	trap		***	
Mean dependent var 14.	53683	S.D. dependent var	0.2	18809	
Sum squared resid 0.1	86349	S.E. of regression	0.1	07920	
R-squared 0.7	95147	Adjusted R-squared	0.7	56737	
F(3, 16) 41.	31911	P-value(F)	9.2	3e-08	
Log-likelihood 18.	37989	Akaike criterion	-28.	75978	
Schwarz criterion -24.	77685	Hannan-Quinn	-27.	98227	
rho 0.5	18840	Durbin-Watson	0.8	51615	
Log-likelihood for Y1 =	-272.35	7			

2.3) จะได้หน้าต่างข้อมูลที่พยากรณ์

### 2.2) จะได้หน้าต่าง Forecast กด OK

Start End
Forecast range: 2563 + 2563 +
<ul> <li>automatic forecast (dynamic out of sample)</li> </ul>
O dynamic forecast
⊖ static forecast
O recursive k-step ahead forecasts: k = 1 ▲
Number of pre-forecast observations to graph
Show fitted values for pre-forecast range
Plot confidence interval using error bars
1 - α = 0.95 💂
<u>H</u> elp <u>C</u> ancel <u>O</u> K

2.4) ข้อมูลที่พยากรณ์ได้คือ

Log(Y1)<sub>2563</sub> = 14.836743

แต่เราพยากรณ์ค่า Y1 ดังนั้นต้องทำการถอดค่า Log ออก จะได้ Y1<sub>2563</sub> = 2,776,615

🔣 greti	: forecasts					
8	1					8
For	95% confidence	intervals,	t(16, 0.0	25) =	2.120	
	1_Y1	prediction	std.	error	95%	interval
2553	14.621362	14.583141				
2554	14.711100	14.610785				
2555	14.725579	14.655421				
2556	14.604007	14.672426				
2557	14.611680	14.680483				
2558	14.716056	14.735420				
2559	14.841385	14.765441				
2560	14.829168	14.677499				
2561	14.844650	14.739691				
2562	14.857947	14.856904	_			
2563		14.836743	0.117	591	14.587461	- 15.086025
			1			

## 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

🛐 gretl: model 7										
<u>F</u> ile <u>E</u> dit <u>T</u> ests	<u>S</u> ave <u>G</u> rapl	ns <u>A</u> nalysis	<u>L</u> aTeX			8				
Model 7: OLS, using observations 2542-2562 (T = 21) Dependent variable: Y2										
	coefficie	nt std.	error	t-ratio	p-value					
const	202.485	34.6	5781	5.839	1.57e-05	***				
Rate Fer	3.7430	7 1.7	75795	2.129	0.0473	**				
Trend	2.2298	9 0.5	534618	4.171	0.0006	***				
Mean depender	nt var 3	06.2857	s.D. d	ependent v	ar 22.079	973				
Sum squared 1	resid 2	908.034	S.E. o	f regressi	on 12.71(	052				
R-squared	0	.701749	Adjust	ed R-squar	ed 0.668	510				
F(2, 18)	2	1.17591	P-valu	e(F)	0.000	019				
Log-likelihoo	od -8	1.57016	Akaike	criterion	169.14	103				
Schwarz crite	erion 1	72.2739	Hannan	-Quinn	169.82	204				
rho	-0	.109645	Durbin	-Watson	2.128	521				

\_\_\_\_\_ 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 แสดงว่าค่าความสัมพันธ์ของค่าความ คลาดเคลื่อนเป็นอิสระกัน 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 หรือไม่Breusch-Pagan test for heteroskedasticityOLS, using observations 2542-2562 (T = 21)Dependent variable: scaled uhat^2coefficient std. error t-ratio p-value<math>const 5.13306 5.59326 0.9177 0.3709Rate_Fer -0.159225 0.283542 -0.5616 0.5813Trend -0.0691783 0.0862290 -0.8023 0.4329Explained sum of squares = 8.15056Test statistic: LM = 4.075278,<br/>with p-value = P(Chi-square(2) > 4.075278) = 0.130336
```

จากการทดสอบ LM test จะเห็นได้ว่า ค่า p-value เท่ากับ 0.130336 ซึ่ง มากกว่า  $lpha_{
m 0.05}$  นั่น คือ ไม่เกิดปัญหา Heteroscedasticity

#### 2.4 ตรวจสอบว่าความคลาดเคลื่อนมีการแจกแจงปกติ หรือไม่



ค่าพยากรณ์ที่ได้จากแบบจำลอง
 Y2 = 335.04 กิโลกรัมต่อไร่

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 **מ**<sub>0.10</sub> นั่น คือ ค่าความคลาดเคลื่อนมีการแจกแจงปกติ