

สมการทำนายน้ำหนักมีชีวิตของสุกรลูกผสม โดยใช้การประมาณค่าวงนัยทั่วไป เทียบกับตัวแบบเชิงเส้น

Predicted Equation in Live -Weight of Crossbred Swine Using Generalized Estimating Equation and Linear Mixed Effect Model

อัญชลี ทองกำเหนิด^{1/} และ ลีลิ อินศรีสว่าง^{2/}
Unchalee Tonggumnead^{1/} and Lily Insrisawang^{2/}

Abstract: The data were collected 4 times during 64 – 70 weeks of age from 173 heads of crossbred swine (Landracc x Large White x Duroc Jersey). Dependent variable is (weight, kilogram) while 6 independent variables are live-weight heart girth (L1, centimeter), body length (L2, centimeter), age (age, week), shoulder width (W1, centimeter), hump width (W2, centimeter) and belly width (W3, centimeter) respectively. Constructing predicted equation on average live - weight of crossbred swine using generalized estimating equation (GEE) was compared with linear mixed effect model. The results showed that in the GEE model, the intercept, heart girth(L1), body length (L2) and hump width (W2) are highly related to live- weight at 0.01 significant level when mean deviance, Pearson chi-square/df and residual plot were considered using empirical standard error estimate and AR(1) correlation structure. The most appropriate predicted equation found is $Y = -191.73 + 0.4292L1 + 2.8302L2 + 0.1684W2$. In linear mixed effect model with random intercept, the estimated parameters using residual maximum likelihood (REML) gave the results close to the GEE method. Equation of the mixed effect is $Y_i = -191.16 + 0.4311L1 + 2.8271L2 + 0.1702W2 + \alpha_{0i}$. Since the random variance, $Var(\alpha_{0i})$, in each crossbred swine was not difference, thus this linear predicted equation may also be used to predict the live-weight of an individual swine.

Keywords: Generalized Estimating Equation, Linear Mixed Effect Model, live weight, crossbred swine

^{1/} ภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคล จ.ปทุมธานี 12110

^{1/} Department of Applied Statistic, Faculty of Science, Rajamangala University of Technology Thanyaburi, Pathum Thani 102110

^{2/} ผู้ช่วยศาสตราจารย์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพฯ 10900

^{2/} Assistant Professor, Department of Statistic, Faculty of Science, Kasetsart University, Bangkok 10900

บทคัดย่อ: เก็บข้อมูลจากสุกรลูกผสม (แลนด์เรซ x ลาร์จไวท์ x ดุริโอกเจอร์ซี) จำนวน 173 ตัว โดยเก็บ 4 ครั้งห่างกันทุก 2 สัปดาห์ ตั้งแต่สุกรมีอายุ 64 ถึง 70 สัปดาห์ ตัวแปรตามได้แก่น้ำหนักมีชีวิตของสุกร (live-weight) ตัวแปรอิสระจำนวน 6 ตัวแปร ได้แก่ ความยาวรอบอก (L1, เซนติเมตร) , ความยาวลำตัว (L2, เซนติเมตร), อายุ (age, สัปดาห์), ความกว้างไหล่ (W1, เซนติเมตร) , ความกว้างสะโพก (W2, เซนติเมตร) และ ความกว้างเอว (W3, เซนติเมตร) สร้างตัวแบบสมการถดถอยเพื่อใช้ทำนายน้ำหนักมีชีวิตของสุกร ใช้สมการการประมาณค่าวงนัยทั่วไป (generalized estimating equations) เปรียบเทียบกับตัวแบบเชิงเส้นผสม (linear mixed effect model) ผลการศึกษาพบว่า กรณีใช้สมการการประมาณค่าวงนัยทั่วไป ค่าคงที่ (intercept) และตัวแปรอิสระ (L1) (L2) และ (W2) ส่งผลต่อค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกรลูกผสมอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 โดยตัวแบบที่เหมาะสมเมื่อพิจารณาจากค่า mean deviance , Pearson chi-square/df และ residual plot เฉพาะกรณี empirical standard error estimates และกำหนดโครงสร้างของความสัมพันธ์เป็นแบบ AR(1) สมการ คือ $Y = -191.73 + 0.4292L1 + 2.8302L2 + 0.1684W2$ กรณีใช้ตัวแบบเชิงเส้นผสมที่มีอิทธิพลสุ่มได้แก่ random intercept และประมาณค่าพารามิเตอร์โดยใช้วิธี residual maximum likelihood (REML) พบว่าให้ผลที่ใกล้เคียงกับการประมาณค่าวงนัยทั่วไป โดยตัวแบบการถดถอยในส่วนของอิทธิพลคงที่อยู่ในรูปแบบ $Y_i = -191.16 + 0.4311L1 + 2.8271L2 + 0.1702W2 + \alpha_{oi}$ และเนื่องจากความแปรปรวนสุ่ม $Var(\alpha_{oi})$ ของสุกรลูกผสมแต่ละตัวไม่แตกต่างกัน จึงสามารถใช้สมการนี้ทำนายน้ำหนักของสุกรแต่ละตัวได้ด้วย

คำสำคัญ: สมการทำนายน้ำหนักมีชีวิตของสุกร การประมาณค่าวงนัยทั่วไป ตัวแบบเชิงเส้นผสม

คำนำ

ธุรกิจการผลิตสุกรในประเทศไทยมีเครือข่ายที่ครอบคลุมทั้งด้านการผลิต โรงงานอาหารสัตว์ พันธุ์สัตว์ โรงฆ่า โรงงานตัดแต่งและแปรรูปเนื้อสุกร ปัจจุบันประเทศไทยมีจำนวนสุกรประมาณ 12 ล้านตัว เป็นการผลิตเพื่อตอบสนองการบริโภคภายในประเทศประมาณกว่าร้อยละ 75 ซึ่งปริมาณการบริโภคเนื้อสุกรของคนไทยโดยเฉลี่ยคนละ 13 กิโลกรัมต่อปี รูปแบบของตลาดสุกรนั้นมีทั้ง 1) สุกรมีชีวิต 2) เนื้อสุกรแช่เย็นและแช่แข็ง 3) เนื้อสุกรสุก และ 4) ผลิตภัณฑ์สุกร ตามปกติการจำหน่ายสุกรมีชีวิตจะมีน้ำหนักระหว่าง 80 - 125 กิโลกรัม (อยู่ในช่วงอายุ 70 - 74 สัปดาห์) เพราะเป็นช่วงที่เหมาะสมแก่การจำหน่ายและได้ราคาดี (กรมปศุสัตว์, 2541: 15) ดังนั้นน้ำหนักของสุกรจึงเป็นสิ่งสำคัญที่จะต้องพิจารณา ในทางปฏิบัติการชั่งน้ำหนักมักทำเฉพาะเวลาจำหน่าย แต่การจัดการอื่นๆ ต้องอาศัยน้ำหนักประกอบด้วย ซึ่งการชั่งน้ำหนักโดยตรงยังขาดความสะดวกและมีข้อจำกัด โดยเฉพาะอย่างยิ่งในเกษตรกรรายย่อย เพราะการซื้อเครื่องชั่งที่มีราคาแพงย่อมไม่คุ้มทุน จึงมีการใช้สายวัดน้ำหนักสำเร็จรูปที่นำเข้าจาก

ต่างประเทศและค่อนข้างมีราคาแพง ประกอบกับสายวัดน้ำหนักดังกล่าว อาจไม่มีความเหมาะสมในการประมาณน้ำหนักของสุกรในประเทศไทย จึงควรทำการศึกษาวิจัยในเรื่องการสร้างตัวแบบหรือสมการเพื่อทำนายน้ำหนักมีชีวิตของสุกร การศึกษาวิจัยในครั้งนี้มีวัตถุประสงค์เพื่อสร้างสมการทำนายน้ำหนักมีชีวิตของสุกรลูกผสมโดยใช้สมการการประมาณค่าวงนัยทั่วไป (generalized estimating equation) เทียบกับตัวแบบเชิงเส้นผสม (linear mixed effect model) ซึ่งการประมาณค่าวงนัยทั่วไป เป็นวิธีการหนึ่งในการอนุมานเกี่ยวกับค่าสัมประสิทธิ์ β ภายใต้ marginal model เมื่อค่าวัดของตัวแปรตาม Y เป็นข้อมูลที่มีการวัดซ้ำหรือข้อมูลระยะยาว (longitudinal data) การประมาณการถดถอยสามารถทำได้โดยหาคำตอบของสมการที่อยู่ในรูป

$$U_i(\beta, \alpha) = \sum_{i=1}^k \left(\frac{\partial \mu_i}{\partial \alpha} \right)^T \text{COV}^{-1}(Y_i, \beta, \alpha) [y_i - \mu_i(\beta)] = 0$$

เมื่อ $\mu_i(\beta) = E(Y_i)$ คือ marginal mean ของ Y_i โดยที่ Y_i เป็น $n_i \times 1$ เวกเตอร์ และเมตริกซ์ความแปรปรวนร่วม $\text{COV}(Y_i)$ ขึ้นอยู่กับ β และ α ซึ่งเป็นพารามิเตอร์ที่แสดงคุณสมบัติภายในกลุ่ม (within-cluster) ส่วนตัวแบบเชิงเส้นผสมเป็นอีกทางเลือกหนึ่งใน

การจัดการข้อมูล เมื่อค่าวัดของตัวแปรตาม Y ที่มีการวัดซ้ำ โดยตัวแปรตาม Y ต้องเป็นข้อมูลต่อเนื่องเพื่อใช้ศึกษาความสัมพันธ์เชิงเส้นระหว่างค่าเฉลี่ยของตัวแปรตาม Y กับตัวแปรอิสระ X ตั้งแต่ 1 ตัวขึ้นไปที่ประกอบด้วยปัจจัยอิทธิพลคงที่ (fixed effect) และปัจจัยอิทธิพลสุ่ม (random effect) โดยปัจจัยอิทธิพลสุ่มจะอธิบายลักษณะเฉพาะของแต่ละหน่วยศึกษา (Shoukri and Chaudhary, 2007)

อุปกรณ์และวิธีการ

ข้อมูลที่ใช้ในการศึกษา ได้แก่ น้ำหนักมีชีวิต (live weight, กิโลกรัม) ซึ่งกำหนดให้เป็นตัวแปรตาม ส่วนตัวแปรอิสระ ที่มีอิทธิพลต่อตัวแปรตาม ได้แก่ ความยาวรอบอก (L1, เซนติเมตร) ความยาวลำตัว (L2, เซนติเมตร) อายุ (age, สัปดาห์) ความกว้างไหล่ (W1, เซนติเมตร) ความกว้างสะโพก (W2, เซนติเมตร) ความกว้างเอว (W3, เซนติเมตร) ทำการศึกษาในสุกรลูกผสม (แลนด์เรซ x ลาร์จไวท์ x ดูริคเจอร์ซี) จำนวน 173 ตัว เลี้ยงที่สถานีวิจัยทับกวาง มหาวิทยาลัยเกษตรศาสตร์ อำเภอบึงสามพัน จังหวัดสระบุรี เก็บข้อมูลต่อเนื่อง 4 ครั้ง ห่างกันทุก 2 สัปดาห์ ตั้งแต่สุกรมีอายุ 64 จนถึง 70 สัปดาห์ ตัวอย่างของข้อมูลที่เก็บได้แสดงในตารางที่ 1

การวิเคราะห์ข้อมูล

สมการการประมาณค่าพารามิเตอร์ทั่วไป (generalized estimating equation)

1. ทำการตรวจสอบการแจกแจงของตัวแปรตาม (live-weight) และกำหนด link function แล้วทดสอบการเป็นอิสระกันของตัวแปรอิสระทั้ง 6 ตัวแปร

2. คัดเลือกตัวแบบที่เหมาะสมโดยพิจารณาตัวแบบในรูปแบบ univariate ของตัวแปรอิสระแต่ละตัวแปรจากสถิติทดสอบ Z ที่มีค่า P - value ต่ำกว่า 0.25 และคัดเลือกตัวแปรโดยวิธี backward elimination ซึ่งพิจารณาจากสถิติทดสอบ Z ที่มีค่า P-value อยู่ระหว่าง 0.15 – 0.20 (Hosmer and Lemeshow , 2002) รวมทั้ง

พิจารณาจากค่า 1) mean deviance 2) mean scaled deviance 3) mean Pearson chi-square และ 4) mean scaled Pearson X²

3. กำหนดรูปแบบ correlation structure โดยใช้รูปแบบดังต่อไปนี้ 1) Independent 2) Exchangeable 3) First-order autoregressive 4) Unstructured และ 5) Toeplitz

4. ใช้คำสั่ง Proc GENMOD ใน SAS โปรแกรม โดยพิจารณาการแจกแจง (distribution) ของตัวแปรตาม Y เพื่อการกำหนด linked function

5. เปรียบเทียบตัวแบบที่ได้จากการกำหนดรูปแบบ correlation structure ที่แตกต่างกัน โดยพิจารณาถึงความแตกต่างระหว่างค่า parameter estimate, standard error ตลอดจน correlation structure ที่เหมาะสม

6. ตรวจสอบความเหมาะสมของตัวแบบ โดยพิจารณาจากความคลาดเคลื่อน (residual plot) ซึ่งสร้างขึ้นโดยการ plot residual กับ $\hat{Y}$ และตรวจสอบการแจกแจงของความคลาดเคลื่อน ถ้าความคลาดเคลื่อนมีการเคลื่อนไหวที่เป็นอิสระต่อกันโดยกระจายรอบ ๆ ค่า 0 และมีการแจกแจงแบบปกติแสดงว่าตัวแบบมีความเหมาะสม

ตัวแบบเชิงเส้นผสม (linear mixed model)

1. คัดเลือกตัวแบบที่เหมาะสมโดยพิจารณาตัวแบบในรูปแบบ univariate ของตัวแปรอิสระแต่ละตัวแปรจากสถิติทดสอบ t ที่มีค่า P - value ต่ำกว่า 0.25 (Hosmer and Lemeshow , 2002) และคัดเลือกตัวแปรโดยใช้ backward elimination ที่พิจารณาจากค่า -2 log likelihood

2. กำหนดตัวแบบเชิงเส้นผสมที่ประกอบไปด้วยอิทธิพลคงที่ (fixed effect) และอิทธิพลสุ่ม (random effect) โดยกำหนดอิทธิพลสุ่มคือ random intercept ซึ่งตัวแบบอยู่ในรูป $Y_i = X_i\beta + \alpha_{0i} + \varepsilon_i$

3. ใช้คำสั่ง Proc MIXED ใน SAS เพื่อสร้างตัวแบบ/สมการทำนายน้ำหนักมีชีวิตของสุกรพันธุ์ผสม (แลนด์เรซ x ลาร์จไวท์ x ดูริคเจอร์ซี)

4. พิจารณาความเหมาะสมของตัวแบบจากเกณฑ์ -2 log likelihood

Table 1 Examples of data being collected 4 times from 173 heads of crossbred swine (Large White X Landrace X Duroc Jersey) during 64 to 70 weeks of age

NO.	Time	Weight	L1	L2	Age	W1	W2	W3
1	1	78.13	91.40	79.40	64	33.25	33.00	27.32
	2	80.48	92.80	80.00	66	33.80	33.20	27.87
	3	83.00	95.00	81.00	68	34.00	33.60	28.07
	4	84.15	96.00	81.50	70	34.25	34.00	28.32
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
173	1	109.60	110.13	87.70	64	35.80	36.30	29.87
	2	110.90	112.48	88.40	66	36.40	36.50	30.47
	3	111.00	115.00	89.20	68	37.00	36.70	31.07
	4	114.20	118.15	89.80	70	37.70	37.00	31.77
Minimum		70.13	86.90	77.30	64.00	17.87	23.80	24.50
Maximum		127.15	119.30	92.20	70.00	33.07	39.00	38.00
Mean		101.83	104.92	85.53	66.99	29.56	35.49	34.98
Variance		105.28	33.35	7.21	5.01	2.44	2.43	1.36

Note : Live-weight (weight : kilogram), heart girth (L1 : centimeter) , body length (L2 : centimeter) , age(age : week) , shoulder width (W1 : centimeter) , hump width (W2 : centimeter) and belly width (W3: centimeter)

ผลการวิเคราะห์ข้อมูล

ส่วนที่ 1 ผลการตรวจสอบการแจกแจงของตัวแปรตาม (weight) และการตรวจสอบความเป็นอิสระกันของตัวแปรอิสระ

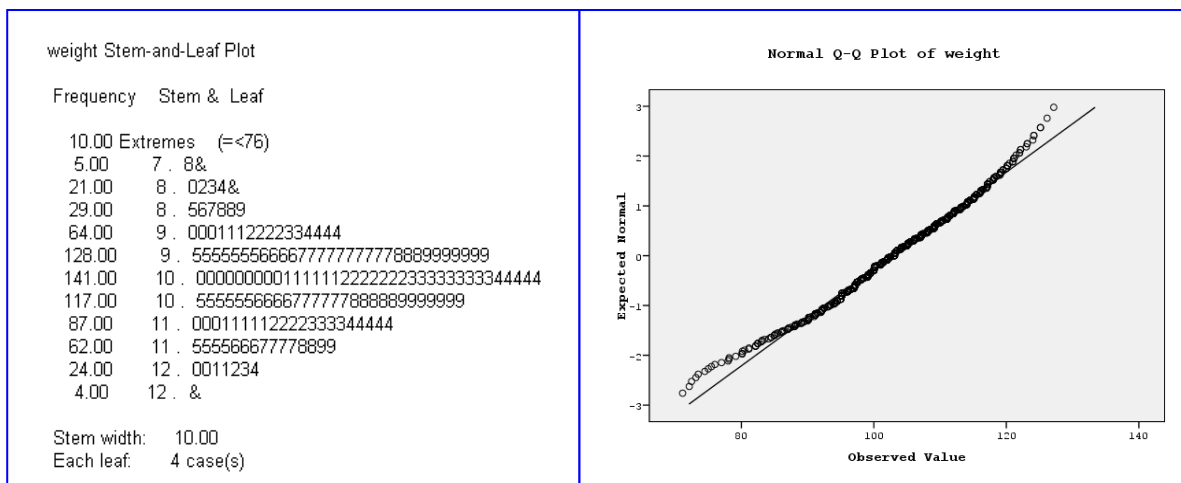
จากภาพที่ 1 ผลการตรวจสอบการแจกแจงของตัวแปรตาม (weight) ในเบื้องต้นโดยพิจารณาจากกราฟ ได้แก่ a) stem & leaf b) normal probability plot และ c) box plot พบว่าการแจกแจงของ ตัวแปรตามเป็นแบบปกติ และจากสถิติทดสอบ Kolmogorov – Smirnov ได้ค่า P- value เท่ากับ 0.013 แสดงว่าตัวแปรตาม มีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.01 ในการศึกษา

ครั้งนี้จึงกำหนด link function เป็น identity link (Myers et al., 2002)

ส่วนการตรวจสอบความเป็นอิสระกันของตัวแปรอิสระพิจารณาจากค่า VIF(variance inflation factor) ของตัวแปรอิสระ X_i โดย VIF ของตัวแปร

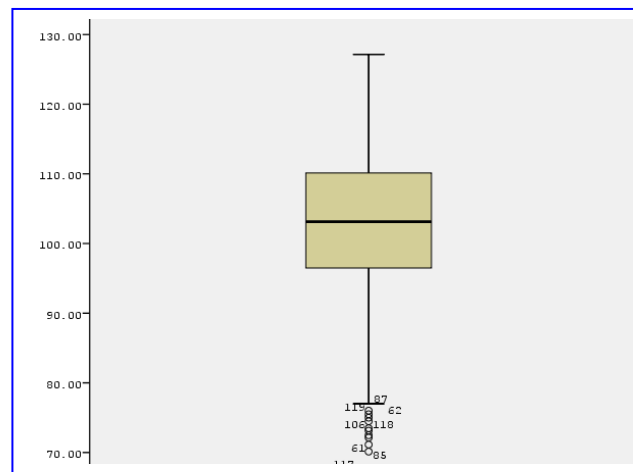
$$X_i = VIF_i = \frac{1}{1 - R_i^2}$$

ถ้า VIF_i มีค่ามากแสดงว่าตัวแปรอิสระ X_i มีความสัมพันธ์กับตัวแปรอิสระอื่น ๆ มาก หรือเกิดปัญหา multicollinearity ระหว่างตัวแปรอิสระ โดยถ้าค่า VIF มีค่ามากกว่า 10 ขึ้นไปจะถือว่าเกิดปัญหา multicollinearity ระหว่างตัวแปรอิสระ (Kutner et al. , 2005)



a) Stem & Leaf Plot

b) Normal Probability Plot



c) Box Plot

Figure 1 Examination of dependent variable (weight)

ผลการตรวจสอบการเป็นอิสระกันของตัวแปรอิสระทั้ง 6 ตัวแปรได้แก่ ความยาวรอบอก (L1) ความยาวลำตัว (L2) อายุ (age) ความกว้างไหล่ (W1) ความกว้างสะโพก (W2) และ ความกว้างเอว (W3) มีค่าอยู่ระหว่าง 1.091 – 3.435 รายละเอียดเกี่ยวกับค่า VIF ของตัวแปรแต่ละตัวแสดงดังตารางที่ 2 เนื่องจากค่าดังกล่าว VIF ไม่เกิน 10 แสดงว่าไม่เกิดปัญหา multicollinearity ระหว่างตัวแปรอิสระ

ส่วนที่ 2 ผลการวิเคราะห์ข้อมูลโดยใช้วิธีสมการการประมาณค่าวงนัยทั่วไป

การศึกษาครั้งนี้เลือกที่จะนำเสนอตัวแบบการถดถอยที่มีโครงสร้างความสัมพันธ์เป็น AR(1) correlation structure เนื่องจากเมื่อพิจารณาจาก covariance matrix (model-based) และ covariance matrix (empirical) พบว่าโครงสร้างความสัมพันธ์ AR (1) เป็นโครงสร้างที่ covariance matrix มีค่าใกล้เคียงกันมากที่สุด นอกจากนี้

การที่เลือก AR(1) correlation structure เนื่องจาก มีรายงานการใช้รูปแบบ first - order autoregressive ว่าเป็นรูปแบบที่เหมาะสมกับข้อมูลเกี่ยวกับระบบสืบพันธุ์และการเจริญเติบโตของสิ่งมีชีวิต (วีระศักดิ์ และคณะ 2549 อ้างถึงใน Rajala-Schultz and Razer 2003) ผลการวิเคราะห์ข้อมูลหลังจากคัดเลือกตัวแปรด้วยวิธี backward elimination แสดงดังตารางที่ 3

จากตารางที่ 4 เมื่อพิจารณาเฉพาะกรณี empirical standard error estimates จากสถิติทดสอบ Z-test เพื่อตรวจสอบตัวแปรอิสระที่ส่งผลต่อค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกร พบว่าค่าคงที่ (intercept) , ความยาวรอบอก (L1) , ความยาวลำตัว (L2) และความกว้างสะโพก (W2) ให้ค่า P- value น้อยกว่า 0.0001 ทุกกรณี แสดงว่าค่าคงที่ (intercept) และตัวแปรอิสระ ส่งผลต่อ

ค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกร อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ดังนั้นตัวแบบอยู่ในรูปของ

$Y_i = -191.73 + 0.4292L1 + 2.8302L2 + 0.1684W2$ นอกจากนั้นจากตารางที่ 3 ค่า mean deviance มีค่าเท่ากับ 0.9877 และค่า Pearson chi - square/df มีค่าเท่ากับ 0.9887 ซึ่งใกล้เคียงกับ 1 แสดงว่าตัวแบบดังกล่าวเป็นตัวแบบที่เหมาะสม

ส่วนที่ 3 ตรวจสอบความเหมาะสมของตัวแบบ

โดยพิจารณาจากความคลาดเคลื่อน (residual plot)

การพิจารณาความเหมาะสมของตัวแบบ $Y_i = -191.73 + 0.4292L1 + 2.8302L2 + 0.1684W2$ โดยพิจารณาจาก residual plot จากภาพที่ 2 d) เป็นการ plot residual กับ $\hat{Y}$ ซึ่งแสดงให้เห็นว่าความคลาดเคลื่อนมีการเคลื่อนไหวที่เป็นอิสระต่อกันโดย

Table 2 Checking of multicollinearity

Variable	VIF value
L1	2.675
L2	2.567
Age	1.091
W1	1.874
W2	3.256
W3	3.435

Table 3 Parameters estimated using first – order autoregressive correlation

Correlation structure	Parameter	Estimate	Empirical			Model-based		
			Standard Error	Z	P-value	Standard Error	Z	P-value
First-order autoregressive correlation	Intercept	-191.73	5.9376	-32.20	<0.0001**	2.7215	-70.25	<0.0001**
	L1	0.4292	0.0762	5.63	<0.0001**	0.0332	12.94	<0.0001**
	L2	2.8302	0.1681	16.84	<0.0001**	0.0720	39.31	<0.0001**
	W2	0.1684	0.0480	4.13	<0.0001**	0.0395	4.26	<0.0001**
Mean Deviance			0.9877		Pearson chi – square/df		0.9887	

** significant (P-value < 0.01)

กระจายรอบ ๆ ค่า 0 ภาพที่ 2 e) เป็นการตรวจสอบการแจกแจงของความคลาดเคลื่อน เมื่อพิจารณาเบื้องต้นจากกราฟ พบว่าการแจกแจงของความคลาดเคลื่อนมีการแจกแจงแบบปกติ และจากสถิติทดสอบ Kolmogorov – Smirnov ได้ค่า P- value เท่ากับ 0.014 ดังนั้นความคลาดเคลื่อนมีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.01

ผลการตรวจสอบความเหมาะสมของตัวแบบโดยพิจารณาจากค่าวัดความเหมาะสมของตัวแบบด้วย mean deviance , Pearson chi – square/df และ residual plot พบว่าตัวแบบ

$$Y_i = -191.73 + 0.4292L1 + 2.8302L2 + 0.1684W2 \quad (1)$$
 เป็นตัวแบบที่เหมาะสม จากสมการที่ (1) สามารถอธิบายได้ดังนี้

เมื่อความยาวรอบอกของสุกร (L1) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 0.4292 กิโลกรัม เมื่อความยาวลำตัวของสุกร (L2) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 2.8302 กิโลกรัม และเมื่อความกว้างสะโพกของสุกร (W2) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 0.1684 กิโลกรัม

ส่วนที่ 4 ผลการวิเคราะห์ข้อมูลโดยใช้ตัวแบบเชิงเส้นผสม (linear mixed effect model)

ผลการวิเคราะห์ข้อมูลโดยใช้ ตัวแบบเชิงเส้นผสม โดยประมาณค่าพารามิเตอร์ด้วยวิธี residual maximum likelihood (REML) ซึ่งมีโครงสร้างความสัมพันธ์เป็น AR(1) correlation หลังจากใช้วิธี backward elimination ในการคัดเลือกตัวแปร แสดงดังตารางที่ 4

จากสถิติทดสอบ Z-test เพื่อตรวจสอบตัวแปรอิสระที่ส่งผลต่อค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกร พบว่าค่าคงที่ (intercept) , ความยาวรอบอก (L1) , ความยาวลำตัว (L2) และความกว้างสะโพก (W2) ให้ค่า P- value น้อยกว่า 0.0001 ทุกกรณี แสดงว่าค่าคงที่ และตัวแปรอิสระทุกตัว ส่งผลต่อค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกร อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 และตัวแบบอยู่ในรูปของ

$$Y_i = -191.16 + 0.4311L1 + 2.8271L2 + 0.1702W2 + \alpha_{oi} \quad (2)$$
 ซึ่ง สามารถอธิบายได้ดังนี้

เมื่อความยาวรอบอกของสุกร (L1) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 0.4311 กิโลกรัม ในขณะที่ความยาวลำตัวของสุกร (L2) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 2.8271 กิโลกรัม และ เมื่อความกว้างสะโพกของสุกร (W2) เพิ่มขึ้น 1 เซนติเมตร จะทำให้น้ำหนักมีชีวิตของสุกร เพิ่มขึ้น 0.1702 กิโลกรัม

นอกจากนี้จากสถิติทดสอบ Z – test เพื่อตรวจสอบความแปรปรวนสุ่มของสุกรแต่ละตัว มีค่า P- value เท่ากับ 0.0501 แสดงว่าความแปรปรวนสุ่ม $Var(\alpha_{oi})$ ของสุกรแต่ละตัวไม่แตกต่างกันที่ระดับนัยสำคัญ 0.05

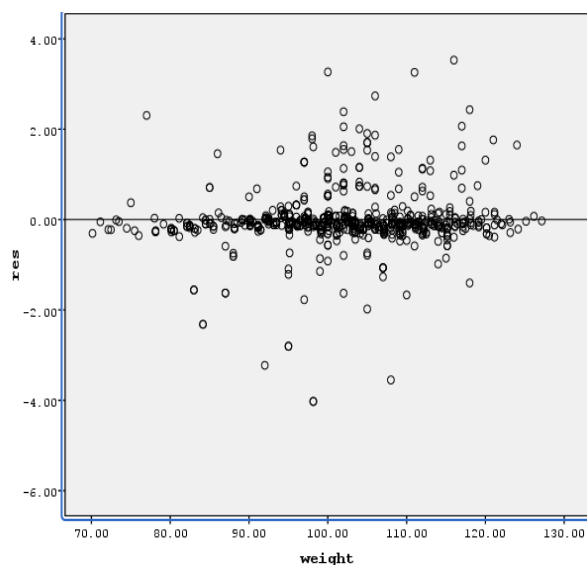
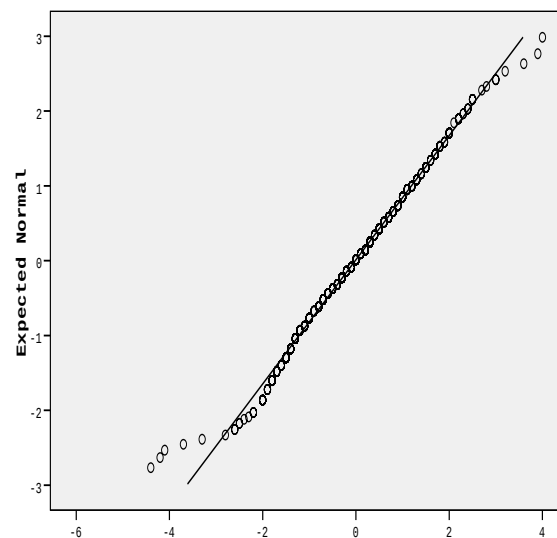
อภิปรายและสรุปผลการทดลอง

ในการศึกษานี้พบว่าตัวแปรอิสระที่ส่งผลต่อค่าเฉลี่ยของน้ำหนักมีชีวิตของสุกรลูกผสม อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 มีทั้งหมด 3 ตัวแปรได้แก่ ความยาวรอบอก (L1) ความยาวลำตัว (L2) และความกว้างสะโพก (W2) ซึ่งค่าประมาณพารามิเตอร์โดยใช้สมการการประมาณค่าพารามิเตอร์ทั่วไป กรณี empirical standard error estimates ที่มีโครงสร้างความสัมพันธ์เป็น AR(1) ให้ค่าประมาณพารามิเตอร์ที่ใกล้เคียงกันมากกับการใช้ตัวแบบเชิงเส้นผสม จึงสามารถนำไปใช้ประโยชน์ได้ทั้ง 2 วิธีขึ้นอยู่กับวัตถุประสงค์ในการใช้งาน เช่นหากต้องการสร้างสมการน้ำหนักมีชีวิตในภาพรวมของประชากรสามารถเลือกใช้ได้ทั้ง 2 วิธี แต่ถ้าหากต้องการสร้างสมการน้ำหนักมีชีวิตในหน่วยย่อยของสุกรแต่ละตัว จะต้องเลือกใช้วิธีที่ 2 คือ ตัวแบบเชิงเส้นผสม

สำหรับข้อมูลที่ใช้การศึกษานี้เมื่อวิเคราะห์โดยใช้ตัวแบบเชิงเส้นผสม (linear mixed effect model) พบว่าความแปรปรวนสุ่ม $Var(\alpha_{oi})$ ของสุกรลูกผสม แต่ละตัวไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ดังนั้นจึงไม่จำเป็นต้องสร้างสมการทำนายน้ำหนักของสุกรแต่ละตัวเนื่องจากจะให้ค่าประมาณพารามิเตอร์เช่นเดียวกันในสุกรแต่ละตัว โดยสมการทำนายจะอยู่ในรูปของ $Y_i = -191.16 + 0.4311L1 + 2.8271L2 + 0.1702W2 + \alpha_{oi}$ ซึ่งอาจจะแตกต่างกันเพียงเล็กน้อยในเทอมของค่าคงที่

(intercept) สอดคล้องกับการศึกษาของ รักปัญญา (2546) ที่เก็บข้อมูลจากสุกรลูกผสม จำนวน 450 ตัว ในพื้นที่ จังหวัดปทุมธานี โดยวัดค่าของตัวแปรตามและตัวแปรอิสระเพียง 1 ครั้ง (ไม่มีการวัดซ้ำ) พบว่าตัวแปรอิสระ ได้แก่ ความยาวรอบอก, ความยาวลำตัว, อายุ ส่งผลต่อน้ำหนักมีชีวิตของสุกรอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 และได้สมการทำนายน้ำหนักมีชีวิตของสุกรแบบการถดถอยเชิงเส้นตรง (linear regression) คือ น้ำหนักมีชีวิต = $-80.68 + 1.09(\text{ความยาวรอบอก}) + 0.23(\text{ความยาว}$

ลำตัว) + 2.03 (อายุ) อย่างไรก็ตามในการวิจัยในครั้งนี้ใช้ข้อมูลที่มีการวัดซ้ำหรือข้อมูลระยะยาว (longitudinal data) ซึ่งการใช้ตัวแบบการถดถอยเชิงเส้นตรง จะวิเคราะห์ข้อมูลได้ไม่ดีพอ เนื่องจากมีอิทธิพลสุ่มจากการวัดซ้ำสุกรแต่ละตัวเข้ามาเกี่ยวข้อง ดังนั้นจึงควรใช้สมการการประมาณค่าวงนัยทั่วไป สำหรับการใช้ตัวแบบเชิงเส้นผสม ทั้ง 2 วิธี ให้ผลไม่แตกต่างกันจึงสามารถนำไปใช้ประโยชน์ในการทำนายน้ำหนักมีชีวิตของสุกรลูกผสมได้

d) Residual with $\hat{Y}$ plot

e) Normal probability plot

Figure 2 Analysis for appropriate model using residual plot

Table 4 Parameter estimating using Residual Maximum Likelihood (REML)

Parameter	Estimate	Standard Error	Z	P-value
Intercept	-191.16	2.7358	-69.88	<0.0001**
L1	0.4311	0.03335	12.92	<0.0001**
L2	2.8271	0.07237	39.06	<0.0001**
W2	0.1702	0.03974	4.28	<0.0001**
$\text{Var}(\alpha_{oi})$	0.02589	0.01579	1.64	0.0501*
-2 Log Likelihood	615.8			

** significant (P-value < 0.01) * significant (P-value < 0.05)

เอกสารอ้างอิง

- กรมปศุสัตว์. 2541. ข้อมูลเศรษฐกิจการปศุสัตว์ ประจำปี 2541. หน้า 2- 3. ใน : รายงานข้อมูลเศรษฐกิจการปศุสัตว์ 2541 . กรมปศุสัตว์.
- วีรศักดิ์ ปัญญาพรวิทยา, ขวัญชัย เครือสุคนธ์, อ้อมฤทัย ทอگی และ ทวีศาสตร์ ตันกิตยานนท์. 2549. การประเมินประสิทธิภาพการสืบพันธุ์จากโปรแกรมสุขภาพโคนมระดับฝูงการศึกษาเปรียบเทียบ 3 ปี. *เชียงใหม่สัตวแพทยสาร*. 4(2) : 107 – 115.
- รักปัญญ์ ธนอมวงศ์วัฒนะ. 2546. สมการทำนายน้ำหนักมีชีวิตของสุกรลูกผสม. *วารสารเกษตร*. 19(2) : 160 – 161.
- Hosmer, D.W. and Lemeshow, S. 2002. *Applied Logistic Regression* . 2nd ed. New York : John Wiley & Sons. 345 pp.
- Kutner, M. K., Nachtsheim C.J., Neter J. and William L.I. 2005. *Applied Linear Statistical Models*. 5th ed. Mcgraw.Hill companies. 1396 pp.
- Shoukri, M. M. and Chaudhary, M. A. 2007. *Analysis of Correlated Data with SAS and R*. Taylor & Francis Group,LLc. 295 pp.
- Myers, R. H., Montgomery, D. and Vining, G. G. 2002. *Generalized Linear Models*. New York : John Wiley & Sons. 342 pp.