

บทที่ 3

วิธีดำเนินการวิจัย

แหล่งของข้อมูล (data source)

ข้อมูลที่ใช้ในการศึกษา เป็นข้อมูลที่ได้จากฟาร์มเอกชนแห่งหนึ่งในภาคตะวันออกเฉียงเหนือของประเทศไทย ซึ่งเก็บรวบรวมข้อมูลผลผลิตและข้อมูลพันธุ์ประวัติ ตั้งแต่ปี พ.ศ. 2533 ถึง ปี พ.ศ. 2545 เป็นสุกรพ่อแม่พันธุ์พันธุ์แท้ที่ทางฟาร์มได้สั่งซื้อสุกรพ่อแม่พันธุ์จากต่างประเทศ ได้แก่ พันธุ์ลาจัวท์ ยอร์กเชียร์ และแลนด์เรซ จากประเทศอังกฤษ พันธุ์ดูรอด จากประเทศอเมริกาและแคนาดา ข้อมูลที่นำมาวิเคราะห์เป็นบันทึกผลผลิตจริงของแม่สุกรภายในฟาร์ม ได้แก่ ลักษณะจำนวนลูกแรกเกิดทั้งหมด จำนวนลูกแรกเกิดมีชีวิต น้ำหนักแรกเกิด จำนวนลูกหย่านม และน้ำหนักลูกหย่านม ทั้งหมด 4 พันธุ์ดังกล่าวมาแล้วข้างต้น

โครงสร้างของข้อมูล (data structure)

ข้อมูลที่ใช้ในการศึกษา ประกอบด้วย 2 แฟ้มข้อมูล คือ

1. แฟ้มข้อมูลพันธุ์ประวัติ (pedigree file)

- หมายเลขประจำตัวสัตว์ (animal ID)
- หมายเลขพ่อพันธุ์สุกร
- หมายเลขแม่พันธุ์สุกร
- วัน เดือน ปี เกิดของตัวสัตว์

2. แฟ้มข้อมูลของลักษณะสมรรถภาพทางการสืบพันธุ์

- หมายเลขประจำตัวสุกร (animal ID : ID)
- พันธุ์ (breed : B)
- วัน-เดือน-ปี ที่สุกรให้ลูก (birth-date : BD)
- ฤดูกาลที่สุกรให้ลูก (season : S)
- ครั้งที่สุกรให้ลูก (parity :P)
- จำนวนลูกแรกเกิดทั้งหมด (total born : TB)
- จำนวนลูกแรกเกิดมีชีวิต (total born alive : BA)
- น้ำหนักแรกเกิด (birth weight : BW)

- จำนวนลูกหย่านม (number wean : NW)
- น้ำหนักลูกหย่านม (weaning weight : WW)

การจัดการและการเตรียมข้อมูล (data manipulation and editing)

1. การจัดการข้อมูลเบื้องต้น

เนื่องจากข้อมูลได้มาจากทะเบียนประวัติของฟาร์ม ซึ่งเก็บในโปรแกรม LOGIPORC เป็นแฟ้มข้อมูลลักษณะการให้ผลผลิตและแฟ้มข้อมูลพันธุ์ประวัติ แต่ในโปรแกรม LOGIPORC จะมีแฟ้มข้อมูลทางด้านเศรษฐกิจรวมอยู่ในส่วนการเงินของโปรแกรมดังกล่าวด้วย ดังนั้น จึงทำการดึงข้อมูลลักษณะการให้ผลผลิตและแฟ้มข้อมูลพันธุ์ประวัติออกมาไว้ในโปรแกรมหมูยิ้ม และทำการแปลงแฟ้มข้อมูลทั้งหมดให้เป็นแฟ้มตัวอักษร (text file) เพื่อสามารถจัดการกับข้อมูลเบื้องต้นได้ จำนวนข้อมูลลักษณะสมรรถภาพทางการสืบพันธุ์ได้แก่ จำนวนลูกแรกเกิดทั้งหมด (TB) จำนวนลูกแรกเกิดมีชีวิต (BA) น้ำหนักแรกเกิด (BW) จำนวนลูกหย่านม (NW) และน้ำหนักลูกหย่านม (WW) ข้อมูลที่เข้าทำการวิเคราะห์ทั้งสิ้น 2,838 ตัว ดังรายละเอียดในตารางที่ 3.1

ตารางที่ 3.1 จำนวนข้อมูลของลักษณะทางการสืบพันธุ์แยกตามพันธุ์สุกรที่ทำการศึกษา

รายการ	ลักษณะสมรรถภาพทางการสืบพันธุ์					รวม
	TB	BA	BW	NW	WW	
พันธุ์ลาร์จไวท์	2,295	2,294	2,289	2,221	2,209	11,308
พันธุ์แลนด์เรซ	2,173	2,165	2,160	2,114	2,104	10,716
พันธุ์ดูรอด	1,203	1,178	1,177	1,122	1,121	5,801
พันธุ์ออร์ดเชียร์	900	884	883	857	854	4,378
รวม	6,571	6,521	6,483	6,314	6,288	32,203

2. การจำแนกอิทธิพลปัจจัยคงที่ ประกอบด้วย

2.1 อิทธิพลของปีที่แม่สุกรคลอด จากข้อมูลที่ทำการศึกษาอยู่ในช่วงตั้งแต่ปี พ.ศ. 2533-2545

2.2 อิทธิพลของเดือน จำแนกตามเดือนที่ให้ผลผลิตทั้งหมด 12 เดือน

2.3 อิทธิพลของพันธุ์ เป็นสุกรพันธุ์แท้ 4 พันธุ์ คือ

- พันธุ์ลาร์จไวท์
- พันธุ์แลนด์เรซ
- พันธุ์ดูรอด

- พันธุ์ยอร์คเชียร์

2.4 อิทธิพลของลำดับครอก มีทั้ง 13 ลำดับครอก

3. การตรวจสอบการกระจายของข้อมูล

การศึกษาครั้งนี้ใช้วิธีวิเคราะห์ข้อมูลด้วยแบบหุ้ผสมเชิงเส้นตรง โดยใช้โมเดลของสัตว์แต่ละตัว (animal model) ทำการวิเคราะห์องค์ประกอบความแปรปรวนด้วยวิธี Restricted Maximum Likelihood (REML) และวิเคราะห์ค่าการผสมพันธุ์ด้วยวิธี Best Linear Unbiased Prediction (BLUP) (Henderson, 1984) การวิเคราะห์องค์ประกอบความแปรปรวนด้วยวิธี REML มีข้อกำหนด (assumption) ว่า ข้อมูลที่นำมาวิเคราะห์มีการกระจายเป็นแบบปกติ (normal distribution) ดังนั้น ข้อมูลลักษณะสมรรถภาพทางการสืบพันธุ์ ได้แก่ จำนวนลูกแรกเกิด มีชีวิต, น้ำหนักแรกเกิด, จำนวนลูกหย่านม และน้ำหนักลูกหย่านม ก่อนทำการวิเคราะห์จึงต้องทำการตรวจสอบการกระจายของข้อมูลด้วยชุดคำสั่ง PROC UNIVARIATE ในโปรแกรมสำเร็จรูป SAS 6.12 for Windows (Statistical Analysis System, 1998) พบว่า ข้อมูลที่สุ่มมาจากประชากรมีการกระจายโมเดลปกติที่ระดับความเชื่อมั่น 99 เปอร์เซ็นต์ทุกลักษณะ

การวิเคราะห์ข้อมูล

1. การวิเคราะห์เบื้องต้น

การวิเคราะห์เบื้องต้นเพื่อทดสอบปัจจัยคงที่ ที่คาดว่าจะมีอิทธิพลต่อลักษณะที่ศึกษา ด้วยคำสั่ง PROC MIXED โดยใช้โปรแกรมคอมพิวเตอร์สำเร็จรูป SAS 6.12 for Windows (Statistical Analysis System, 1998) ดังนี้

1.1 โมเดลสำหรับการทดสอบหาอิทธิพลที่มีผลกระทบต่อลักษณะที่ศึกษา

จากการตรวจเอกสารพบว่ามีปัจจัยคงที่ (fixed effects) คือ ปีที่แม่สุกรคลอด เดือนที่แม่สุกรคลอด พันธุ์ของสุกร และลำดับการคลอด สำหรับปัจจัยสุ่ม (random effects) คือ อิทธิพลเนื่องจากตัวสัตว์และค่าความคลาดเคลื่อน โดยทำการวิเคราะห์คราวละ 1 ลักษณะ โดยมีโมเดลในการวิเคราะห์รวมทุกพันธุ์และแยกตามพันธุ์ ดังนี้

(1) วิเคราะห์ข้อมูลรวมทุกพันธุ์

$$y_{ijklm} = \mu + B_i + P_j + YM_k + AGEF_l + An_m + e_{ijklm} \quad \dots(3.1)$$

(2) วิเคราะห์ข้อมูลแยกตามพันธุ์

$$y_{ijklm} = \mu + P_i + YM_j + AGEF_k + An_l + e_{ijklm} \quad \dots(3.2)$$

โดยที่

y_{ijklm} , y_{ijkl} = ลักษณะทางการสืบพันธุ์ที่ศึกษาที่ i ที่ได้รับอิทธิพลจากพันธุ์ที่ i ลำดับการคลอดที่ j ปี-เดือนที่ k และอิทธิพลอื่นๆ มีค่าอธิบายในสัญลักษณ์ตามโมเดลที่กล่าวมาทั้งหมด

โดยลักษณะที่ทำการวิเคราะห์คือ

1. จำนวนลูกเกิดทั้งหมด
2. จำนวนลูกเกิดมีชีวิต
3. น้ำหนักแรกเกิด
4. จำนวนลูกหย่านม
5. น้ำหนักลูกหย่านม

μ = ค่าเฉลี่ยของทุกลักษณะ

B_i = อิทธิพลคงที่เนื่องจากพันธุ์ที่ i

P_j = อิทธิพลคงที่เนื่องจากลำดับการคลอดที่ j

YM_k , YM_j = อิทธิพลคงที่เนื่องจากปี-เดือนที่ k หรือ j

$AGEF_i$, $AGEF_k$ = อิทธิพลคงที่เนื่องจากอายุคลอดลูกครั้งที่ i หรือ k

An_k , An_m = อิทธิพลสุ่มเนื่องจากตัวสัตว์ที่ m หรือ j

โดยให้ An_k , An_m , \sim NID $(0, A\sigma_a^2)$

A เป็นเมตริกซ์ความสัมพันธ์ระหว่างสัตว์

e_{ijklm} , e_{ijkl} = อิทธิพลสุ่มอื่นๆ ที่ค่าสังเกตได้รับ

โดยให้ $e \sim$ NID $(0, I\sigma_e^2)$

2. การเตรียมแฟ้มพันธุ์ประวัติและผลผลิต

ก่อนการประมาณค่าองค์ประกอบความแปรปรวนและความแปรปรวนร่วม (Variance and covariance component) ด้วยโปรแกรม REMLF90 ต้องมีการจัดการข้อมูลโดยการ RENUM ข้อมูลพันธุ์ประวัติและผลผลิตเพื่อให้มีการจัดลำดับ แบ่งกลุ่มตามปีและเดือนหรือมีการแบ่งเป็นฝูงและฤดูกาลทั้งนี้ขึ้นกับการทดสอบอิทธิพลที่ในขั้นตอนที่ผ่านมา รวมทั้งการให้สัญลักษณ์หมายเลขตัวสัตว์ที่เป็นตัวอักษรและตัวเลขเปลี่ยนเป็นตัวเลขใหม่ทั้งหมด การ RENUM พันธุ์ประวัติจะได้แฟ้มพันธุ์ประวัติทั้งหมด 2 แฟ้ม คือ แฟ้มพันธุ์ประวัติที่คิด

3.1.3 โมเดลที่มีอิทธิพลสุ่มแบบบวกสะสมของยีน อัตราเลือดชิด และอิทธิพลแบบข้ามของยีน

$$y = X\beta + Z_1b\Delta + Z_2u + Wf + Z_3p + e \quad \dots(\text{โมเดลที่ 3})$$

- y = เวกเตอร์ของค่าสังเกต
 β = เวกเตอร์ของปัจจัยคงที่
 b = เวกเตอร์ของสัมประสิทธิ์อัตราเลือดชิด
 Δ = ค่าคงที่ของอิทธิพลอัตราเลือดชิด
 u = เวกเตอร์แบบสุ่มของอิทธิพลแบบบวกสะสมของยีน
 f = เวกเตอร์แบบสุ่มของอิทธิพลแบบข้ามของยีน ของ parent subclass
 p = เวกเตอร์ของอิทธิพลของสิ่งแวดล้อมที่ถาวร
 e = เวกเตอร์แบบสุ่มของอิทธิพลที่เหลือ

$X, Z_{1,2,3}, W$ = เมตริกซ์ของเหตุการณ์ที่สัมพันธ์กับเวกเตอร์ที่ระบุไว้แล้ว

3.2 ประเมินค่าอัตราพันธุกรรม (heritability, h^2)

ประเมินค่าอัตราพันธุกรรมจากค่าองค์ประกอบความแปรปรวนที่หาได้จากโมเดลทั้ง 3 โมเดล สามารถคำนวณค่าอัตราพันธุกรรมเนื่องจากอิทธิพลแบบบวกสะสมของยีน ค่าอัตราพันธุกรรมเนื่องจากอิทธิพลแบบข้ามของยีน ได้จากสูตรการคำนวณดังนี้

3.2.1 ค่าอัตราพันธุกรรมเนื่องจากอิทธิพลแบบบวกสะสมของยีน (h^2)

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2} \quad \dots(\text{สำหรับโมเดลที่ 1,2})$$

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_{pe}^2 + \sigma_e^2} \quad \dots(\text{สำหรับโมเดลที่ 3})$$

3.2.2 ค่าอัตราพันธุกรรมเนื่องจากอิทธิพลแบบข้ามของยีน (d^2)

$$d^2 = \frac{\sigma_d^2}{\sigma_a^2 + \sigma_d^2 + \sigma_{pe}^2 + \sigma_e^2} \quad \dots(\text{สำหรับโมเดลที่ 3})$$

โดยที่

$$\begin{aligned}\sigma_a^2 &= \text{ความแปรปรวนเนื่องจากอิทธิพลแบบบวกสะสมของยีน} \\ \sigma_d^2 &= \text{ความแปรปรวนเนื่องจากอิทธิพลแบบข่มของยีน} \\ \sigma_{pe}^2 &= \text{ความแปรปรวนเนื่องจากอิทธิพลของสิ่งแวดล้อมถาวร} \\ \sigma_e^2 &= \text{ความแปรปรวนเนื่องจากอิทธิพลสุ่มอื่นๆ ที่ค่าสังเกตได้รับ}\end{aligned}$$

3.3 ประเมินค่าการผสมพันธุ์ (breeding value, BV)

จากค่าองค์ประกอบความแปรปรวนของลักษณะที่ทำการศึกษา ที่ประมาณได้ของแต่ละโมเดลจะนำไปใช้ในการประเมินค่าทางพันธุกรรมโดยวิธี Best Linear Unbiased Predictions (BLUP) ด้วยโปรแกรมสำเร็จรูป BLUPF90 PC (Duangjinda and Misztal, 2001) ด้วยวิธี BLUP (Henderson, 1984) ทำการประเมินค่าการผสมพันธุ์โดยทำการวิเคราะห์ที่ละลักษณะ โปรแกรมจะใช้ค่าความแปรปรวนที่คำนวณได้จากโปรแกรม REMLF90 ซึ่งจะเป็นค่าที่ใช้ในสมการ Mixed Model Equation (MME) ทั้งนี้โปรแกรมจะหยุดการคำนวณเมื่อผลเฉลยในแต่ละรอบของการคำนวณที่ติดกันมีค่าต่างกันไม่เกิน 0.0000000001 หรือ 1×10^{-10}

3.4 ความสัมพันธ์ระหว่างค่าการผสมพันธุ์จาก 3 โมเดล

ค่าการผสมพันธุ์ของสุกรแต่ละตัวทั้ง 5 ลักษณะ ที่ได้จากการวิเคราะห์ 3 โมเดล นำมาหาความสัมพันธ์กับค่าการผสมพันธุ์ โดยใช้วิธี Spearman rank correlation coefficient เพื่อพิจารณาการจัดลำดับแม่พันธุ์จากแต่ละโมเดลว่ามีความแตกต่างกันหรือไม่

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย