
**ตัวแบบการวิเคราะห์ภาวะเจริญพันธุ์
ในช่วงภาวะเจริญพันธุ์ต่ำของประเทศไทย**
Fertility modeling in low fertility of Thai society

ธนานนท์ บัวทอง* และ วิรากรณ์ โพธิศิริ**
Thananon Buathong and Wiraporn Pothisiri



*นิสิตปริญญาเอก วิทยาลัยประชากรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ได้รับการสนับสนุนทุนการศึกษาจากทุนการศึกษาหลักสูตรดุษฎีบัณฑิต "100 ปี จุฬาลงกรณ์มหาวิทยาลัย"

**อาจารย์ประจำ วิทยาลัยประชากรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

Buathong, T. and Pothisiri, W (2017). 13 (2): 217-240

DOI: 10.14456/jssnu.2017.23

Copyright © 2017 by Journal of Social Sciences, Naresuan University: JSSNU

All rights reserved

บทคัดย่อ

ตัวแปรระดับภาวะเจริญพันธุ์ เป็นข้อมูลจำนวนนับที่ได้รับความสนใจวิจัยมาอย่างต่อเนื่อง เพื่อเป็นข้อมูลพื้นฐานในการกำหนดเป้าหมายของนโยบายประชากร แต่ตัวแปรจำนวนนับไม่ใช่ตัวแปรต่อเนื่อง ทำให้สมการถดถอยเชิงเส้นไม่ใช่ตัวแบบการวิเคราะห์ที่เหมาะสม สำหรับบทความนี้ได้สำรวจและนำเสนอวิธีการวิเคราะห์ข้อมูลที่เหมาะสมกับการกระจายตัวของตัวแปรระดับภาวะเจริญพันธุ์ของผู้หญิงไทยในปี 2552 และผลการวิจัยพบว่า 1) ค่าความแปรปรวนของตัวแปรจำนวนบุตรเกิดรอด และตัวแปรจำนวนบุตรที่ต้องการมีเพิ่ม ไม่แตกต่างไปจากค่าเฉลี่ยมากนัก แต่มีตัวอย่างรายงานว่า ไม่ต้องการมีบุตรเพิ่มเป็นจำนวนมาก ทำให้ตัวแปรจำนวนบุตรเกิดรอดควรเลือกใช้ตัวแบบการถดถอยปัวซอง ส่วนตัวแปรจำนวนบุตรที่ต้องการมีเพิ่มควรเลือกใช้ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ และ 2) ผลการทดสอบด้วยสถิติอนุมานยืนยันความเหมาะสมของตัวแบบการวิเคราะห์เช่นเดียวกับผลการวิเคราะห์ด้วยสถิติพรรณนา โดย 3) ผลการวิเคราะห์ความแตกต่างของภาวะเจริญพันธุ์ของประชากรด้วยตัวแบบการถดถอยที่เหมาะสมกับการกระจายตัวของตัวแปรตาม เปรียบเทียบผลจากการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้น (เป็นตัวแบบที่ถูกนำมาใช้ในการวิเคราะห์ภาวะเจริญพันธุ์อย่างต่อเนื่อง) พบว่า ระดับความมีนัยสำคัญทางสถิติและทิศทางของความสัมพันธ์มีความแตกต่างเล็กน้อย สำหรับข้อเสนอแนะทางวิชาการคือ ในกรณีงานวิจัยมีวัตถุประสงค์เพียงพิจารณาแนวโน้มความสัมพันธ์ระหว่างตัวแปรอธิบายกับตัวแปรตามที่เป็นจำนวนนับอาจเลือกนำเสนอผลการวิเคราะห์จากตัวแบบการถดถอยเชิงเส้นได้ในกรณีที่ได้ทดสอบความสอดคล้องของผลการวิเคราะห์เรียบร้อยแล้ว ส่วนในกรณีที่เป็นการวิจัยมีวัตถุประสงค์ต้องการสร้างสมการพยากรณ์ของตัวแปรตามที่เป็นจำนวนนับ หรือสนใจขนาดความแตกต่างระหว่างกลุ่มย่อยประชากร จำเป็นต้องนำเสนอผลการวิเคราะห์จากตัวแบบการถดถอยที่มีความเหมาะสมกับการกระจายตัวของข้อมูลตัวแปรตาม เช่น ตัวแบบการถดถอยปัวซอง ตัวแบบการถดถอยทวินามลบ ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ หรือตัวแบบการถดถอยทวินามลบที่มีผลกระทบจากศูนย์ เป็นต้น

คำสำคัญ: ตัวแบบการวิเคราะห์ ตัวแปรจำนวนนับ ภาวะเจริญพันธุ์ต่ำ ประเทศไทย

Abstract

Modeling of fertility has been widely studied in the literature to understand the past, current and future trend of population and to identify the suitable population policy and target population to sustain fertility level. The current studies show using Poisson regression models are relevant as fertility outcome rather than ordinary least square model. This paper concerns the suitable model to predict the number of children ever born and the desired number of additional children of ever-married women aged 18-49 in 2009 and compared the result from the linear regression model. The results showed that 1) the variance of number of children ever born and the desired number of additional children are similar to the mean, while most of sample reported "zero child-number intention"; 2) The Poisson regression model is relevant as number of children ever born, while Poisson zero-inflated regression is statistically appropriate for the modeling the desired number of additional children than Poisson regression in low fertility of Thailand; 3) compared with results from linear regression models show the statistically significant and the direction of relationships are slightly different with the appropriate model. Regarding an academics recommendations, it depends on an objective of research. If the research aims to analyses association between explanatory variables and count data variable, result of linear regression model in case concordance of analysis result was tested. Should the research aims to create a forecast equation of count data or focuses on magnitude of difference between the population subgroup. Analysis Result of regression model which is suitable for distribution of count data needs to be presented, namely Poisson regression model, Negative-binomial regression model, Zero-inflated Poisson regression model and Zero-inflated Negative-binomial regression model.



Keyword: Fertility modeling, Count data, Low fertility, Thailand

ที่มาและความสำคัญ
 ในสังคมที่ภาวะเจริญพันธุ์ของประชากรอยู่ในระดับต่ำกว่าระดับทดแทน¹ ความเข้าใจที่ถูกต้องเกี่ยวกับความแตกต่างของภาวะเจริญพันธุ์ระหว่างประชากรกลุ่มย่อยมีความสำคัญเป็นอย่างยิ่งในการกำหนดนโยบายส่งเสริมการมีบุตร รวมไปถึงการคาดประมาณจำนวนและโครงสร้างของประชากรในอนาคต แต่ทว่าปัญหาสำคัญประการหนึ่งในการวิเคราะห์ภาวะเจริญพันธุ์ของประชากร คือ การใช้ตัวแบบวิเคราะห์ที่อาจไม่เหมาะสมกับลักษณะของข้อมูล

ในงานวิจัยที่ผ่านมา ตัวแปรระดับภาวะเจริญพันธุ์ของประชากรที่สำคัญและนิยมใช้ในการศึกษาอย่างต่อเนื่อง คือ จำนวนบุตรเกิดรอด (Number of children ever born) และจำนวนบุตรที่ยังมีชีวิตอยู่ (Number of living children) ซึ่งการนับจำนวนดังกล่าวจะขึ้นกับช่วงเวลาที่ต้องการศึกษา เช่น ตั้งแต่เข้าสู่วัยเจริญพันธุ์จนเวลาปัจจุบันที่ศึกษา หรือหลังจากสิ้นสุดวัยเจริญพันธุ์ไปแล้ว นอกจากนี้ พบว่าในช่วงเวลาปัจจุบันที่สังคมมีภาวะเจริญพันธุ์ต่ำกว่าระดับทดแทน นักวิจัยได้ให้ความสำคัญกับการวิเคราะห์ขนาดครอบครัวที่พึงปรารถนา (Family size preference) ของผู้หญิง ซึ่งเป็นคำถามเกี่ยวกับจำนวนบุตรที่ผู้หญิงปรารถนาจะมีในตลอดช่วงวัยเจริญพันธุ์ ตัวแปรจำนวนบุตรเหล่านี้มีลักษณะเป็นจำนวนนับ (Count data) ค่าของข้อมูลจะเป็นจำนวนเต็ม (Integer) ไม่มีค่าติดลบและเกิดจากการนับจำนวน ไม่ใช่การเรียงลำดับจำนวน เนื่องด้วยตัวแปรจำนวนบุตรไม่ใช่ตัวแปรต่อเนื่อง (Continuous variable) ทำให้โดยทฤษฎีแล้ว การเลือกใช้วิธีวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้น (Linear regression model) ไม่มีความเหมาะสม (Pandey & Kaur, 2015)

นับจากปี ค.ศ. 1990 เป็นต้นมา ตัวแบบการถดถอยแบบไม่ใช่เส้นตรง (Non-linear regression model) ได้ถูกนำมาใช้ในการวิเคราะห์ตัวแปรจำนวนบุตร โดยในระยะแรกตัวแบบการถดถอยปัวซอง (Poisson regression model) เป็นตัวแบบในการวิเคราะห์ภาวะเจริญพันธุ์ของประชากร (Pandey & Kaur, 2015) เช่น งานวิจัยของ Fagbamigbe and Adebawale (2014) ในประเทศไนจีเรียปี 2550 หรืองานวิจัยในประเทศอินเดีย ในช่วงปี 2548-2549 ของ Pandey and Kaur (2015) เป็นต้น แต่ด้วยข้อจำกัดของตัวแบบการถดถอยปัวซองที่ไม่เหมาะสมในการวิเคราะห์ข้อมูลที่ตัวแปรตามที่มีค่าความแปรปรวนสูงกว่าค่าเฉลี่ย (Over-dispersion) และข้อมูลตัวแปรตามที่มีค่าเท่ากับ 0 เป็นจำนวนมาก (Excess of zeros) (Poston Jr & McKibben, 2003; Zeileis, Kleiber, & Jackman, 2008) ทำให้มีการพัฒนาวิธีการวิเคราะห์ข้อมูลภาวะเจริญพันธุ์ของประชากรอย่างต่อเนื่อง

ในปัจจุบัน มีนักวิจัยบางส่วนวิเคราะห์ข้อมูลด้วยตัวแบบการถดถอยที่มีผลกระทบจากศูนย์ (Zero-inflated regression model) เช่น Poston Jr and McKibben (2003) ที่เปรียบเทียบผลการวิเคราะห์จำนวนบุตรเกิดรอดของประชากรในสังคมที่มีภาวะเจริญพันธุ์ต่ำ² ของประเทศสหรัฐอเมริกาด้วยตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ (Zero-inflated Poisson regression) เปรียบเทียบกับตัวแบบการถดถอยทวินามลบที่มีผลกระทบจากศูนย์ (Zero-inflated negative-binomial regression) หรืองานวิจัยของ Testa, Bordone, Osiewalska, and Skirbekk (2016) ในการวิเคราะห์การวิเคราะห์จำนวนบุตรที่ต้องการ/ตั้งใจเพิ่มของประชากรในประเทศออสเตรีย ประเทศบัลแกเรีย ประเทศอิตาลี และประเทศนอร์เวย์ด้วยตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ เป็นต้น

ในขณะเดียวกัน มีนักวิจัยบางส่วนเลือกใช้การวิเคราะห์ด้วยตัวแบบการถดถอยอันดับ (Ordered logit regression model) เช่น การวิเคราะห์ความตั้งใจมีบุตรของประเทศในแถบยุโรปของ Testa (2014) เป็นต้น และตัวแบบการถดถอยพหุ (Multinomial logit regression model) ด้วยเช่นเดียวกัน เช่น การวิเคราะห์ความตั้งใจมีบุตรของประเทศในแถบยุโรปของ Berrington and Pattaro (2014) เป็นต้น

วัตถุประสงค์หลักของงานวิจัย คือ 1) เพื่อค้นหาตัวแบบการวิเคราะห์ข้อมูลที่มีความเหมาะสมกับการกระจายตัวของข้อมูลภาวะเจริญพันธุ์ของประชากรในช่วงภาวะเจริญพันธุ์ต่ำของประเทศไทย และ 2) เพื่อเปรียบเทียบความสอดคล้องต้องกันของผลการวิเคราะห์ที่ได้จากตัวแบบการวิเคราะห์ที่เหมาะสมกับตัวแบบตัวแบบการถดถอยเชิงเส้น ซึ่งเป็นตัวแบบการวิเคราะห์ที่ถูกใช้ในการวิเคราะห์ภาวะเจริญพันธุ์ของประชากรมาอย่างต่อเนื่อง โดยตัวแปรภาวะเจริญพันธุ์ของประชากรที่สนใจ คือ 1) จำนวนบุตรเกิดรอด และ 2) จำนวนบุตรที่ต้องการมีเพิ่ม (Desired number of additional children) ซึ่งเป็นตัวแปรภาวะเจริญพันธุ์ของประชากรที่มีความสำคัญในการทำความเข้าใจแนวโน้มของภาวะเจริญพันธุ์ของประชากร และความเข้าใจที่ถูกต้องยังช่วยให้สามารถกำหนดประชากรเป้าหมายส่งเสริมการมีบุตรได้อย่างถูกต้องอีกด้วย



การทบทวนวรรณกรรม

ในช่วงปี 2510-2530 นักประชากรศาสตร์ทำการวิเคราะห์ความแตกต่างของภาวะเจริญพันธุ์ของประชากรที่มีคุณลักษณะทางประชากรศาสตร์ เศรษฐกิจและสังคมแตกต่างกันด้วยตัวแบบการถดถอยเชิงเส้น (Ordinary least square regression model) (Pandey & Kaur, 2015) แต่โดยหลักการทางสถิติแล้ว ผลการวิเคราะห์ตัวแปรจำนวนนับด้วยตัวแบบการถดถอยเชิงเส้นมักมีปัญหาไม่มีประสิทธิภาพ ไม่คงเส้นคงวา และไม่เที่ยงตรง (Inefficient, inconsistent and biased estimates) เนื่องจากการกระจายตัวของข้อมูลมีความเบ้ขวาจากการที่ข้อมูลจำนวนนับไม่มีค่าติดลบ (Long & Freese, 2006) ทำให้ตัวแบบการถดถอยเชิงเส้นไม่ใช่ตัวแบบที่มีความเหมาะสมในการวิเคราะห์ข้อมูลภาวะเจริญพันธุ์ของประชากร โดยเฉพาะในสังคมที่ประชากรมีแนวโน้มมี

บุตรจำนวนน้อยประมาณ 1-2 คน และมีบางส่วนไม่ต้องการมีบุตร (Poston Jr & McKibben, 2003)

โดยหลักการทางสถิติแล้ว ตัวแบบการวิเคราะห์ที่มีความเหมาะสมกับตัวแปรจำนวนนับ คือ ตัวแบบการถดถอยปัวซอง ซึ่งจัดเป็นสมการถดถอยแบบไม่ใช่เส้นตรง (Nonlinear regression model) เช่นเดียวกับตัวแบบการถดถอยโลจิต (Logit regression model) เนื่องจากตัวแปรตามมีลักษณะเป็นข้อมูลไม่ต่อเนื่อง (Discrete variable) (Cameron & Trivedi, 2013; Long & Freese, 2006) และในปัจจุบันการวิเคราะห์ภาวะเจริญพันธุ์ของประชากรมีแนวโน้มเลือกใช้ตัวแบบการวิเคราะห์ในกลุ่มตัวแบบการถดถอยปัวซองมากขึ้น โดยตัวแบบที่เหมาะสมแตกต่างกันไปตามลักษณะการกระจายตัวของข้อมูล³

ตัวแบบการถดถอยปัวซอง จัดอยู่ในกลุ่มตัวแบบการถดถอยแบบ generalized linear พัฒนาโดย King (1989) และ Winkelmann and Zimmermann (1994) ที่มักใช้ในการวิเคราะห์จำนวนบุตรเกิดรอดของประชากร เช่น งานวิจัยของ Fagbamigbe and Adebawale (2014) ในประเทศไนจีเรียปี 2550 หรืองานวิจัยในประเทศอินเดียในช่วงปี 2548-2549 ของ Pandey and Kaur (2015) เป็นต้น ซึ่งเป็นตัวแบบที่มีความเหมาะสมกับข้อมูลมีลักษณะการกระจายตัวแบบปัวซอง (ไม่มีค่าติดลบ และมีความเบ้ขวา) และในปัจจุบันได้มีการพัฒนามีวิธีการทดสอบความเหมาะสม (Goodness of fit) ของตัวแบบการถดถอยปัวซองกับการกระจายตัวของตัวแปรจำนวนนับด้วยคำสั่ง estat gof (เป็นคำสั่งในโปรแกรม STATA) โดยในกรณีผลการทดสอบไม่มีนัยสำคัญทางสถิติ (มีค่ามากกว่า 0.05) หมายความว่า การกระจายตัวของตัวแปรตามเหมาะสมกับการวิเคราะห์ด้วยตัวแบบการถดถอยปัวซอง ส่วนกรณีที่มีนัยสำคัญทางสถิติ จำเป็นต้องทดสอบหาตัวแบบที่เหมาะสมต่อไป (Cameron & Trivedi, 2013) สำหรับสมการพยากรณ์จำนวนบุตรเกิดรอดด้วยตัวแบบการถดถอยปัวซอง คือ

$$\mu_i = \exp(a + X_{1i} b_1 + X_{2i} b_2 + \dots + X_{ki} b_k)$$

เมื่อ μ_i เป็นจำนวนบุตรเกิดรอดที่คาดหวังของผู้หญิงคนที่ i

X_{ki} เป็นคุณลักษณะของผู้หญิง

b_k เป็นสัมประสิทธิ์ถดถอยปัวซอง

ตัวแบบการถดถอยทวินามลบ (Negative-binomial regression model) จัดอยู่ในกลุ่มตัวแบบการถดถอยแบบ generalized linear เช่นเดียวกับตัวแบบการถดถอยปัวซอง เป็นตัวแบบที่ได้รับการพัฒนาจากข้อด้อยของตัวแบบการถดถอยปัวซองที่ไม่เหมาะสมกับข้อมูลที่มีค่าความแปรปรวนของตัวแปรตามสูงกว่าค่าเฉลี่ยเป็นอย่างมาก (Problem of over-dispersion) ซึ่งมีผลทำให้ค่าส่วนเบี่ยงเบนมาตรฐานที่ได้จากการวิเคราะห์มีความโน้มเอียงไปในทางน้อยกว่าค่าที่ควรจะเป็น และส่งผลให้มีแนวโน้มปฏิเสธสมมติฐานหลัก แต่ทิศทางของความสัมพันธ์ยังคงมีความถูกต้อง (Cameron & Trivedi, 2013; Long & Freese, 2006) หรืออาจกล่าวได้ว่า ผลการวิเคราะห์ด้วยสถิติอนุมานมีปัญหาเกี่ยวกับความน่าเชื่อถือของระดับความมีนัยสำคัญทางสถิติ สำหรับสมการถดถอยทวินามลบ อาจเรียกได้ว่า เป็นรูปแบบจำเพาะของสมการถดถอยปัวซอง เนื่องจากยังคงยึดสมการโครงสร้างเช่นเดียวกับสมการถดถอยปัวซอง แต่ได้เพิ่มค่าพารามิเตอร์ที่อนุญาตให้ค่าความแปรปรวนของตัวแปรตามสูงกว่าค่าเฉลี่ย (ϵ_i) เข้าไปในสมการ ทำให้สามารถวิเคราะห์ข้อมูลตัวแปรตามที่มีปัญหา ความแปรปรวนสูงกว่าค่าเฉลี่ยได้ สำหรับสมการพยากรณ์จำนวนบุตรเกิดรอดด้วยตัวแบบการถดถอยทวินามลบ คือ

$$\mu_i = \exp(a + X_{1i} b_1 + X_{2i} b_2 + \dots + X_{ki} b_k + \epsilon_i)$$

ในปัจจุบัน ได้มีการพัฒนามีวิธีการทดสอบความเหมาะสมของตัวแบบการถดถอยทวินามลบเปรียบเทียบกับตัวแบบการถดถอยปัวซอง เรียกว่า Likelihood-ratio test of $\alpha=0$ และในกรณีที่ผลการทดสอบมีนัยสำคัญทางสถิติ (มีค่าน้อยกว่า 0.05) หมายความว่า ข้อมูลมีปัญหาค่าความแปรปรวนสูงกว่าค่าเฉลี่ย และไม่เหมาะสมกับการวิเคราะห์ด้วยตัวแบบการถดถอยปัวซอง (Cameron & Trivedi, 2013) หรือการกระจายตัวของข้อมูลมีความเหมาะสมที่จะใช้ตัวแบบการวิเคราะห์การถดถอยทวินามลบ

ตัวแบบการถดถอยที่มีผลกระทบจากศูนย์ (Zero-inflated regression model) จัดเป็นตัวแบบการวิเคราะห์ส่วนขยายของตัวแบบการถดถอยปัวซอง และตัวแบบการถดถอยทวินามลบ ในกรณีข้อมูลที่มีการรายงานตัวแปรตามมีค่าเท่ากับ 0 เป็นจำนวนมาก (Problem of Excess zeros) และถูกนำมาใช้วิเคราะห์ภาวะเจริญพันธุ์ของประชากรในสังคมที่มีภาวะเจริญพันธุ์ต่ำ⁴ เช่น การวิเคราะห์จำนวนบุตรเกิดรอดของประชากรในประเทศสหรัฐอเมริกาของ Poston Jr and McKibben (2003) ที่วิเคราะห์ด้วยตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ (Zero-inflated Poisson regression) เปรียบเทียบกับตัวแบบการถดถอยทวินามลบที่มีผลกระทบจากศูนย์ (Zero-inflated negative-binomial regression) เป็นต้น และการวิเคราะห์จำนวนบุตรที่ต้องการ/ตั้งใจมีเพิ่มของประชากรในประเทศออสเตรเลีย ประเทศบัลแกเรีย ประเทศอิตาลี และประเทศนอร์เวย์ของ Testa et al. (2016) ที่วิเคราะห์ด้วยตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ เป็นต้น



ในกรณีที่มีการรายงานตัวแปรตามมีค่าเท่ากับ 0 เป็นจำนวนมาก ผลที่ได้จากการวิเคราะห์ด้วยตัวแบบการถดถอยปัวซอง และตัวแบบการถดถอยทวินาม มักมีปัญหาในการพยากรณ์สัดส่วนของตัวอย่างที่รายงานตัวแปรตามเท่ากับ 0 น้อยกว่าความเป็นจริง ทำให้ไม่ใช้ตัวแบบที่เหมาะสมในการวิเคราะห์ข้อมูล โดยในปัจจุบัน Cameron and Trivedi (2013) และ Long and Freese (2006) แนะนำวิธีการทดสอบทางสถิติ 2 วิธี คือ การทดสอบด้วยวิธี Vuong พัฒนาโดย Vuong (1989) และการทดสอบด้วยวิธี vce และหากผลการวิเคราะห์ที่มีนัยสำคัญทางสถิติ หมายความว่า ตัวแบบการถดถอยที่มีผลกระทบจากศูนย์เป็นตัวแบบการวิเคราะห์ที่มีความเหมาะสมกับการกระจายตัวของตัวแปรตามมากกว่าตัวแบบการถดถอยปัวซอง

ข้อสังเกตประการสำคัญเกี่ยวกับวิธีการทดสอบความเหมาะสมของตัวแบบการถดถอยทวินาม ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ หรือตัวแบบการถดถอยทวินามที่มีผลกระทบจากศูนย์ ล้วนเป็นการเปรียบเทียบผลการวิเคราะห์กับตัวแบบการถดถอยปัวซอง ซึ่งอาจกล่าวได้ว่า ตัวแบบการวิเคราะห์ตัวแปรจำนวนนับได้รับการพัฒนาบนพื้นฐานของตัวแบบการถดถอยปัวซอง

โดยสรุป ตัวแบบการวิเคราะห์ที่เหมาะสมกับตัวแปรภาวะเจริญพันธุ์ในกลุ่มตัวแบบการถดถอยแบบ generalized linear ที่เป็นไปได้มีอยู่ 4 วิธีการ คือ 1) ตัวแบบการถดถอยปัวซอง 2) ตัวแบบการถดถอยทวินาม 3) ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ และ 4) ตัวแบบการถดถอยทวินามที่มีผลกระทบจากศูนย์ ขึ้นอยู่กับการกระจายตัวของข้อมูลตัวแปรตาม

อย่างไรก็ดี ยังคงมีตัวแบบการวิเคราะห์อื่นๆ ที่ถูกนำมาใช้ในการวิเคราะห์ระดับภาวะเจริญพันธุ์ของประชากร ได้แก่ ตัวแบบการถดถอยอันดับ (Ordered logit regression model) เช่น การวิเคราะห์ความตั้งใจมีบุตรของประเทศในแถบยุโรปของ Testa (2014) เป็นต้น และตัวแบบการถดถอยพหุ (Multinomial logit regression model) เช่น การวิเคราะห์ความตั้งใจมีบุตรของประเทศในแถบยุโรปของ Berrington and Pattaro (2014) เป็นต้น โดยต่างเป็นตัวแบบที่เหมาะสมกับข้อมูลตัวแปรตามที่มีลักษณะเป็นกลุ่มรายการ แต่มีความแตกต่างเล็กน้อยในการเลือกใช้ กล่าวคือ ตัวแบบการถดถอยพหุมีความเหมาะสมในการวิเคราะห์ข้อมูลจำนวนบุตรเกิดรอดหรือจำนวนบุตรที่ต้องการมีเพิ่มในกรณีที่มีประชากรตัวอย่างรายงานว่า ไม่แน่ใจ หรือไม่ทราบ หรือแล้วแต่พระผู้เป็นเจ้าเป็นจำนวนมาก และผู้วิจัยต้องการเก็บข้อมูลไว้ในการวิเคราะห์ข้อมูล ซึ่งมีความน่าสนใจในทางวิชาการเป็นอย่างมาก โดยเฉพาะงานวิจัยเกี่ยวกับจำนวนบุตรที่ต้องการมีเพิ่ม

สำหรับตัวแบบการถดถอยอันดับ มักถูกนำมาใช้ในการวิเคราะห์จำนวนบุตรเกิดรอดหรือจำนวนบุตรที่ต้องการมีเพิ่มในกรณีที่ตัดข้อมูลประชากรตัวอย่างที่รายงานว่า ไม่แน่ใจ หรือไม่ทราบ หรือแล้วแต่พระผู้เป็นเจ้า ตัดออกจากการวิเคราะห์เรียบร้อยแล้ว โดยมีหลักคิดว่า จำนวนบุตรที่ประชากรตัวอย่างรายงานเป็นกลุ่มของคำตอบที่สามารถเรียงลำดับได้ (Ordinal scale) เช่น กลุ่มผู้หญิงที่ไม่มีบุตร กลุ่มผู้หญิงที่บุตร 1 คน และกลุ่มผู้หญิงที่บุตร 2 คน ตามลำดับ เป็นต้น

การเลือกใช้ตัวแบบการถดถอยอันดับ มีข้อดคลงเบื้องต้นสำคัญ คือ ตัวแปรตามต้องมีคุณสมบัติค่าสัมประสิทธิ์ของตัวแปรอิสระต้องคงที่ระหว่างตัวแปรตามในแต่ละ

อันดับชั้น และมีวิธีการพิสูจน์ด้วยสถิติ Omodel test of proportionality of odds พัฒนาโดย Wolfe and Gould (1998) และ Brant Test of Parallel Regression Assumption รวมทั้งวิธีการตรวจสอบทางสถิติในกลุ่มของ generalized ordinal logistic regression (gologit) พัฒนาโดย Fu (2002) และ Williams (2006) โดยมีวิธีการแปลผลคล้ายคลึงกัน กล่าวคือ หากผลการทดสอบมีนัยสำคัญทางสถิติ หมายความว่า “ตัวแปรตาม ไม่มีคุณสมบัติเหมาะสมที่จะใช้ตัวแบบการถดถอยอันดับในการวิเคราะห์ข้อมูล” (Long & Freese, 2006)

ข้อมูล ตัวแปร และวิธีวิเคราะห์

◆ ข้อมูล

การศึกษาค้างนี้ใช้ข้อมูลจากโครงการสำรวจอนามัยการเจริญพันธุ์ ปี พ.ศ. 2552 ซึ่งเป็นข้อมูลระดับประเทศที่จัดเก็บโดยสำนักงานสถิติแห่งชาติ มีวัตถุประสงค์เพื่อศึกษาสถานการณ์ด้านอนามัยการเจริญพันธุ์ของประชากรไทย และการเข้าถึงบริการสาธารณสุขที่เกี่ยวข้อง สำหรับการสุ่มเลือกประชากรตัวอย่างให้เป็นตัวแทนของประชากรไทยใช้วิธีสุ่มตัวอย่างแบบแบ่งชั้นสองขั้นตอน (Stratified Two-stage sampling Technique) มีประชากรเป้าหมาย คือ ผู้หญิงอายุ 15-49 ปี และผู้ชายอายุ 15-24 ปี โดยคำถามเกี่ยวกับภาวะเจริญพันธุ์ จะถามเฉพาะผู้หญิงที่เคยสมรสเท่านั้น

ในการศึกษาค้างนี้ ผู้วิจัยเลือกเฉพาะผู้หญิงที่เคยสมรส มีอายุ 18-49 ปี และตอบคำถามครบถ้วนในทุกตัวแปรที่ใช้ในการวิเคราะห์ข้อมูล การไม่นับรวมผู้หญิงอายุ 15-17 ปี ในการวิเคราะห์ เนื่องจากเกือบทั้งหมดของผู้หญิงอายุ 15-17 ปี ยังไม่เคยสมรส และช่วงอายุก่อนข้างน้อยอาจจะยังไม่ได้ตัดสินใจหรือคิดเกี่ยวกับเรื่องการมีบุตร ข้อมูลที่ได้จึงอาจมีความน่าเชื่อถือน้อยกว่าข้อมูลที่ไดจากกลุ่มตัวอย่างที่มีอายุมากกว่า อีกทั้งในกลุ่มตัวอย่างที่มีอายุก่อนข้างน้อยมักมีปัญหาเกี่ยวกับความน่าเชื่อถือของจำนวนบุตรที่ต้องการมีเพิ่มที่รายงาน อันเป็นประเด็นเกี่ยวเนื่องกับความไม่มั่นคงด้านอาชีพ/รายได้ หรือความไม่มั่นคงของชีวิตคู่/ชีวิตสมรส



หลังจากตรวจสอบความสอดคล้องภายในของคำตอบในแต่ละข้อคำถามที่ใช้พบว่า ข้อมูลมีความสอดคล้องภายในเป็นอย่างดี เนื่องด้วยเป็นข้อมูลการสำรวจระดับชาติที่จัดทำโดยผู้เชี่ยวชาญของสำนักงานสถิติแห่งชาติ และจัดตัดข้อมูลของตัวอย่างที่ “ไม่ตอบ” หรือตอบว่า “ไม่ทราบ หรือไม่แน่ใจ” ออกไปแล้ว (มีประมาณร้อยละ 1 ของตัวอย่าง) ได้จำนวนตัวอย่างที่ใช้ในการวิเคราะห์ 22,188 คน

◆ ตัวแปรตาม: ระดับภาวะเจริญพันธุ์ประชากรไทย

ตัวแปรตามของงานวิจัย คือ ระดับภาวะเจริญพันธุ์ของประชากรไทย ประกอบด้วย 1) จำนวนบุตรเกิดรอด และ 2) จำนวนบุตรที่ต้องการมีเพิ่มของผู้หญิงเคยสมรส

จำนวนบุตรเกิดรอด วัดจากข้อคำถามในแบบสอบถามเกี่ยวกับจำนวนบุตรที่สตรีเคยให้กำเนิด^๕ ทั้งหมดจากการสมรสทุกครั้ง (ไม่รวม บุตรเลี้ยงและบุตรบุญธรรม) จำนวน 2 ข้อ คือ 1) “.. มีบุตรเกิดรอดที่ปัจจุบันยังมีชีวิตอยู่กี่คน” และ 2) “.. มีบุตรเกิดรอดที่ปัจจุบันเสียชีวิตแล้ว

ก่อกัน” และจากข้อมูลพบว่าจำนวนบุตรเกิดรอดมีค่าอยู่ระหว่าง 0-13 คน เนื่องจากมีเพียงร้อยละ 6 ของกลุ่มตัวอย่างที่รายงานว่ามีบุตรเกิดรอดจำนวน 4-13 คน ผู้วิจัยได้รวมกลุ่มตัวอย่างดังกล่าวเข้าด้วยกัน และจัดรหัสคำตอบใหม่ ได้แก่ 0=ไม่มีบุตรเกิดรอด 1=มีบุตรเกิดรอด 1 คน 2=มีบุตรเกิดรอด 2 คน 3=มีบุตรเกิดรอด 3 คน และ 4 = มีบุตรเกิดรอด 4 คนหรือมากกว่า

จำนวนบุตรที่ต้องการมีเพิ่ม วัดจากข้อคำถามในแบบสอบถาม คือ “นอกเหนือจากบุตรที่มีอยู่ในขณะนี้ ท่านต้องการมีบุตรเพิ่มอีกกี่คน” ซึ่งเป็นการวัดขนาดครอบครัวที่ปรารถนาตามแนวคิดความตั้งใจมีบุตรเพิ่ม (Fertility intention) อันเป็นแนวคิดที่ให้ประชากรตัวอย่างพิจารณาจากจำนวนบุตรที่มีอยู่แล้วร่วมกับข้อจำกัดอื่นๆของประชากรตัวอย่าง (Testa, 2014) จากข้อมูลพบว่าจำนวนบุตรที่ต้องการมีเพิ่มมีค่าอยู่ระหว่าง 0-9 คน เนื่องจากมีเพียงร้อยละ 1 ของกลุ่มตัวอย่างที่รายงานว่าการมีบุตรเพิ่ม 3-9 คน ผู้วิจัยได้รวมกลุ่มตัวอย่างดังกล่าวเข้าด้วยกัน และจัดรหัสคำตอบใหม่ ได้แก่ 0=ไม่ต้องการมีบุตรเพิ่ม 1=ต้องการมีบุตรเพิ่มอีก 1 คน 2=ต้องการมีบุตรเพิ่มอีก 2 คนขึ้นไป

◆ ตัวแปรอธิบาย

สำหรับตัวแปรอธิบายความผันแปรของของระดับภาวะเจริญพันธุ์ ผู้วิจัยเลือกโดยยึดตามแนวทางของ Poston Jr and McKibben (2003) Testa (2014) และ Pandey and Kaur (2015) ประกอบด้วย ตัวแปรระดับบุคคล ได้แก่ กลุ่มอายุ ระดับการศึกษา สถานภาพสมรส ระยะเวลาที่สมรส เขตที่อยู่อาศัย การมีบุตรเสียชีวิต สถานภาพการทำงาน และตัวแปรระดับครัวเรือน ได้แก่ ความมั่งคั่งของครัวเรือน ทั้งนี้ ในการวิเคราะห์ความต้องการมีบุตรเพิ่ม จะเพิ่มตัวแปรจำนวนบุตรเกิดรอดเข้ามาในตัวแบบการวิเคราะห์เนื่องจากพบว่าความต้องการมีบุตรเพิ่มขึ้นอยู่กับว่า ผู้หญิงมีจำนวนบุตรได้ครบตามจำนวนที่ต้องการแล้วหรือยัง (Morgan, 1982) โดยแต่ละตัวแปรมีรายละเอียดของการนิยามและวิธีการสร้างตัวแปรดังนี้

กลุ่มอายุของตัวอย่าง เป็นตัวแปรระดับจัดอันดับจำนวน 7 กลุ่มรายการ คือ กลุ่มรายการช่วงอายุ 2 ปี (18-19 ปี) และอีก 6 กลุ่มรายการช่วงอายุ 5 ปี โดยให้กลุ่มอายุ 30-34 ปี เป็นกลุ่มอ้างอิง ส่วนตัวแปรระยะเวลาที่สมรส เป็นตัวแปรจัดอันดับที่สร้างจากข้อคำถามเกี่ยวกับอายุแรกสมรสของตัวอย่าง และจัดเป็น 4 กลุ่มรายการ คือ 0-4 ปี 5-9 ปี 10-14 ปี และ 15 ปีขึ้นไป โดยให้ 0-4 ปี เป็นกลุ่มอ้างอิง สำหรับตัวแปรสถานภาพสมรส เป็นตัวแปรระดับจัดกลุ่มจำนวน 2 กลุ่มรายการ คือ 1) กำลังสมรสหรืออยู่กินฉันท์สามีภรรยา และ 2) เป็นหม้าย/หย่าร้าง / แยกกันอยู่ โดยให้กลุ่มกำลังสมรสหรืออยู่กินฉันท์สามีภรรยาเป็นกลุ่มอ้างอิง ส่วนตัวแปรศาสนา เป็นตัวแปรระดับจัดกลุ่ม 3 กลุ่มรายการ คือ 1) พุทธ 2) อิสลาม และ 3) คริสต์ และอื่นๆ โดยให้ศาสนาพุทธเป็นกลุ่มอ้างอิง

ตัวแปรระดับการศึกษาของผู้หญิง เป็นตัวแปรระดับจัดอันดับจากข้อคำถามเกี่ยวกับระดับการศึกษาสูงสุดและจัดกลุ่มเป็น 4 กลุ่มรายการ คือ 1) การศึกษาระดับต่ำที่สุด (ไม่ได้รับการศึกษา และได้รับการศึกษาแต่ไม่สำเร็จการศึกษาระดับประถมศึกษา) 2) การศึกษาระดับต่ำ (สำเร็จการศึกษาระดับประถมศึกษา) 3) การศึกษาระดับกลาง (สูง

กว่าประถมศึกษา แต่ไม่เกินมัธยมปลาย) และ 4) การศึกษาระดับสูง (สูงกว่ามัธยมปลายขึ้นไป) โดยให้การศึกษาระดับต่ำเป็นกลุ่มอ้างอิง

ตัวแปรจำนวนบุตรเกิดรอดที่ยังมีชีวิต เป็นตัวแปรระดับจัดอันดับจากข้อคำถามที่ว่า “ท่านมีบุตรเกิดรอดที่ปัจจุบันยังมีชีวิตอยู่กี่คน (รวมบุตรที่อยู่ในครรภ์ขณะนี้และอยู่ที่อื่น) โดยกลุ่มตัวอย่างรายงานว่า จำนวนบุตรเกิดรอดที่ยังมีชีวิตอยู่ในช่วง 0-12 คน และผู้วิจัยได้รวมกลุ่มตัวอย่างที่รายงานว่า มีบุตรเกิดรอดที่ยังมีชีวิตจำนวน 5-12 คน (ประมาณร้อยละ 1 ของขนาดตัวอย่าง) เข้ากับกลุ่มตัวอย่างรายงานว่า มีบุตรเกิดรอดที่ยังมีชีวิตจำนวน 4 คน (ประมาณร้อยละ 3 ของขนาดตัวอย่าง) และจัดเป็น 5 กลุ่มรายการ คือ ไม่มี / มี 1 คน / มี 2 คน / มี 3 คน และมีตั้งแต่ 4 คนขึ้นไป โดยให้กลุ่มมีบุตรเกิดรอดที่ยังมีชีวิต 2 คน เป็นกลุ่มอ้างอิง

ตัวแปรการมีบุตรเสียชีวิต เป็นตัวแปรระดับจัดอันดับจากข้อคำถามที่ว่า “ท่านมีบุตรเกิดรอดที่ปัจจุบันเสียชีวิตแล้วกี่คน” โดยกลุ่มตัวอย่างรายงานว่า จำนวนบุตรเกิดรอดที่เสียชีวิตอยู่ในช่วง 0-6 คน และผู้วิจัยได้รวมกลุ่มตัวอย่างที่รายงานว่า มีบุตรเกิดรอดที่เสียชีวิตจำนวน 2-6 คน (ประมาณร้อยละ 2 ของขนาดตัวอย่าง) เข้ากับกลุ่มตัวอย่างรายงานว่า มีบุตรเกิดรอดที่เสียชีวิตจำนวน 1 คน (ประมาณร้อยละ 2 ของขนาดตัวอย่าง) และจัดเป็น 2 กลุ่มรายการ คือ ไม่มี และมีบุตรเกิดรอดที่เสียชีวิต โดยให้กลุ่มไม่มีบุตรที่เสียชีวิตเป็นกลุ่มอ้างอิง

เขตที่อยู่อาศัย เป็นตัวแปรระดับจัดกลุ่มสร้างจากข้อคำถามเกี่ยวกับภูมิภาค และข้อคำถามเกี่ยวกับเขตที่อยู่อาศัยในเขต/นอกเขตเทศบาล และผู้วิจัยจัดเป็น 3 กลุ่มรายการ คือ 1) เขตชนบท (อาศัยอยู่นอกเขตเทศบาลในต่างจังหวัด) 2) เขตเมือง (อาศัยอยู่ในเขตเทศบาลในต่างจังหวัด) และ 3) กรุงเทพมหานคร โดยให้เขตเมืองเป็นกลุ่มอ้างอิง

สถานภาพการทำงาน เป็นตัวแปรจัดกลุ่มจากข้อคำถามที่ว่า “ในระหว่าง 12 เดือนก่อนวันสัมภาษณ์ท่านทำงานอะไร” และผู้วิจัยได้จัดเป็น 3 กลุ่มรายการ คือ 1) ไม่ได้ทำงาน 2) ทำงานในระบบ (ทำงานเป็นข้าราชการ/ลูกจ้างรัฐบาล พนักงาน/ลูกจ้างรัฐวิสาหกิจ พนักงาน/ลูกจ้างเอกชน) และ 3) ทำงานนอกระบบ (ค้าขาย/ประกอบธุรกิจส่วนตัว เกษตรกร (ปลูกพืช เลี้ยงสัตว์ ประมง) รับจ้างทั่วไป/กรรมกร) เพื่อสะท้อนถึงความแตกต่างในความสามารถในการบริหารจัดการเวลาในการดูแลบุตรและทำงานตามสถานภาพการทำงาน

ตัวแปรระดับความมั่งคั่งของครัวเรือน เป็นตัวแปรระดับจัดอันดับ 5 กลุ่มรายการตามระดับความมั่งคั่งของครัวเรือน (ต่ำที่สุด-สูงที่สุด) โดยวิธีการสร้างยัฒวิธีการคำนวณค่าถ่วงน้ำหนักขององค์ประกอบในการวัดระดับความมั่งคั่งของครัวเรือนด้วยแนวคิดการถือครองสินทรัพย์ตามวิธีการของ Filmer and Pritchett (2001) และองค์ประกอบหลักที่นำมาใช้ในการคำนวณ คือ 1) ลักษณะของที่อยู่อาศัย 2) สถานภาพการครอบครองที่อยู่อาศัย และ 3) การเป็นเจ้าของทรัพย์สิน ได้แก่ เติงยนอนไม้หรือโลหะ เตาอบไมโครเวฟ กระจกต้มน้ำไฟฟ้า ตู้เย็น โทรทัศน์ เครื่องเล่นวิดีโอ / ซีดี เครื่องซักผ้า เครื่องปรับอากาศ เครื่องทำน้ำร้อน / น้ำอุ่น ในห้องน้ำ เครื่องคอมพิวเตอร์ โทรศัพท์ (รวมเครื่องต่อพ่วง PCT) โทรศัพท์เคลื่อนที่รถยนต์ส่วนบุคคล รถบรรทุกเล็ก / รถปิกอัพ / รถตู้ รถอีแต่น และรถอื่น ๆ ประเภทเดียวกัน รถจักรยานยนต์



◆ การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลแบ่งออกเป็น 3 ส่วนหลัก ประกอบด้วย 1) คุณลักษณะของประชากรตัวอย่าง เป็นการวิเคราะห์การกระจายตัวตัวอย่าง จำแนกตามตัวแปรอธิบาย ด้วยสถิติพรรณนา และ 2) การวิเคราะห์ตัวแบบที่เหมาะสมด้วยสถิติพรรณนาเปรียบเทียบความสอดคล้องต้องกันของผลการวิเคราะห์ตัวแบบที่เหมาะสมที่ได้จากการทดสอบทางสถิติ

การวิเคราะห์ตัวแบบที่เหมาะสมด้วยสถิติพรรณนาเป็นการ 1) เปรียบเทียบค่าเฉลี่ยกับความแปรปรวนของตัวแปรตาม เพื่อทดสอบว่า ข้อมูลมีปัญหา Over-dispersion หรือไม่ และ 2) พิจารณาจากร้อยละของตัวอย่างที่รายงานจำนวนบุตรเกิดรอดหรือจำนวนบุตรที่ต้องการ เท่ากับ 0 เพื่อทดสอบว่า ข้อมูลมีปัญหา Excess of zeros หรือไม่ สำหรับวิธีการทดสอบทางสถิติ ได้แสดงไว้ในการทบทวนวรรณกรรม

ตาราง 1: การวิเคราะห์ตัวแบบที่เป็นไปได้ในการวิเคราะห์ระดับภาวะเจริญพันธุ์ ด้วยสถิติพรรณนา
ที่มา: คำนวณโดยผู้วิจัย

ปัญหา Over-dispersion	ปัญหา Excess of zeros	ตัวแบบที่เหมาะสม
ไม่มี	ไม่มี	การถดถอยปัวซอง
มี	ไม่มี	การถดถอยทรีนามัล
ไม่มี	มี	การถดถอยปัวซองที่มีผลกระทบจากศูนย์
มี	มี	การถดถอยทรีนามัลที่มีผลกระทบจากศูนย์

ส่วนสุดท้าย การเปรียบเทียบความสอดคล้องต้องกันของผลการวิเคราะห์ที่ได้จากตัวแบบการวิเคราะห์ที่เหมาะสมกับตัวแบบตัวแบบการถดถอยเชิงเส้น ซึ่งเป็นตัวแบบการวิเคราะห์ที่ถูกใช้ในการวิเคราะห์ภาวะเจริญพันธุ์ของประชากรมาอย่างต่อเนื่อง โดยพิจารณาจากความมีนัยสำคัญทางสถิติ และทิศทางความสัมพันธ์ของตัวแปรอธิบายต่อระดับภาวะเจริญพันธุ์ของประชากรไทย (ค่าสัมประสิทธิ์) เพื่อชี้ให้เห็นว่า ผลการวิจัยที่ได้จากการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้น อาจนำไปสู่ข้อสรุปของผลการวิจัยที่ผิดพลาดอย่างไร ในการวิเคราะห์ระดับภาวะเจริญพันธุ์ของประชากรไทยในช่วงที่ประชากรมีภาวะเจริญพันธุ์ต่ำกว่าระดับทดแทน

ผลการวิจัย

◆ คุณลักษณะของประชากรตัวอย่าง

ในปี 2552 ผู้หญิงที่เคยสมรสอายุ 18-49 ปี มีจำนวนบุตรเกิดรอดเฉลี่ยประมาณ 1.76 คน (รายละเอียดดังตาราง 2) และมีเพียงร้อยละ 14 ของประชากรตัวอย่างที่ยังไม่มีบุตรเกิดรอด โดยสัดส่วนของผู้หญิงที่รายงานว่ามีบุตรเกิดรอด ลดลงตามกลุ่มอายุที่เพิ่มขึ้น เห็นได้ว่า ในสังคมไทยปี 2552 ประชากรไทยมีภาวะเจริญพันธุ์สมรสอยู่ในระดับต่ำกว่าระดับทดแทน แต่การมีคู่ครองแต่ไม่มีบุตร อาจจะยังไม่ใช้ปรากฏการณ์ที่พบเห็นได้โดยทั่วไปในสังคมไทย โดยเฉพาะอย่างยิ่งหลังอายุ 30 ปีขึ้นไป (สัดส่วนผู้หญิงรายงานว่ามีบุตรเกิดรอดต่ำกว่าร้อยละ 10)

สำหรับประเด็นความต้องการมีบุตรเพิ่ม พบว่า ส่วนใหญ่มีจำนวนบุตรเกิดรอดตามที่ต้องการแล้ว (รายละเอียดดังตาราง 3) โดยมีเพียงร้อยละ 19 ที่ต้องการมีบุตรเพิ่ม และเมื่อพิจารณาจำนวนที่ต้องการมีเพิ่ม จำแนกตามจำนวนบุตรเกิดรอดที่ยังมีชีวิตแล้ว เห็นได้ชัดว่า กลุ่มตัวอย่างที่ต้องการมีบุตรเพิ่มเป็นผู้หญิงที่ยังไม่มีบุตร หรือมีบุตรเพียง 1 คน โดยเกือบทั้งหมดของผู้หญิงที่มีบุตรเกิดรอดที่ชีวิตจำนวน 2 คน ขึ้นไป ไม่ต้องการมีบุตรเพิ่ม ซึ่งอาจสะท้อนถึง การมีบุตร 2 คน เป็นค่านิยมของผู้หญิงไทย

ตาราง 2: การกระจายตัวร้อยละ และค่าเฉลี่ยของจำนวนบุตรเกิดรอด จำแนกตามกลุ่มอายุ
ที่มา: คำนวณโดยผู้วิจัย

กลุ่มอายุ	การกระจายตัวร้อยละ					ค่าเฉลี่ย
	ไม่มีบุตร	บุตร 1 คน	บุตร 2 คน	บุตร 3 คน	4 คนขึ้นไป	
อายุ 18-19 ปี	46.13	47.78	5.93	0.16	-	0.67
อายุ 20-24 ปี	45.64	39.69	12.93	1.38	0.35	0.96
อายุ 25-29 ปี	17.03	46.37	29.34	5.64	1.61	1.25
อายุ 30-34 ปี	9.95	33.38	41.75	10.82	4.10	1.67
อายุ 35-39 ปี	6.64	23.75	46.87	16.16	6.57	1.95
อายุ 40-44 ปี	5.97	21.00	45.05	18.12	9.86	2.13
อายุ 45-49 ปี	5.34	16.49	45.35	20.81	12.01	2.33
จำนวนตัวอย่าง	13.96	28.99	37.89	12.93	6.23	1.76



ตาราง 3: การกระจายตัวของจำนวนบุตรที่ต้องการมีเพิ่ม จำแนกตามจำนวนบุตรเกิดรอดที่ยังมีชีวิต
ที่มา: คำนวณโดยผู้วิจัย

จำนวนบุตรที่ต้องการเพิ่ม	ตัวอย่างรวม	จำแนกตามจำนวนบุตรเกิดรอดที่ยังมีชีวิต			
		ไม่มีบุตร	มีบุตร 1 คน	มีบุตร 2 คน	มากกว่า 2 คน
0 คน	81.15	52.86	65.11	94.12	95.48
1 คน	14.54	24.03	30.79	4.77	1.97
2 คน	3.91	21.54	3.59	1.01	0.78
3 คนขึ้นไป	0.40	1.57	0.51	0.10	1.77
รวม	100.00	100.00	100.00	100.00	100.00

สำหรับข้อมูลด้านประชากรศาสตร์ ประชากรตัวอย่างเป็นผู้หญิงเคยสมรสอายุเฉลี่ยประมาณ 35 ปี เกือบทั้งหมดยังคงมีคู่ครอง (ประมาณร้อยละ 90) มีอายุเมื่อแรกสมรสโดยเฉลี่ยประมาณ 22 ปี มีจำนวนบุตรเกิดรอดที่ยังมีชีวิตโดยเฉลี่ยประมาณ 1.64 คน และมีเพียงร้อยละ 6 ที่มีบุตรเสียชีวิต

ข้อมูลด้านเศรษฐกิจและสังคม เกือบทั้งหมดนับถือศาสนาพุทธ (ประมาณร้อยละ 93 เปรียบเทียบกับร้อยละ 7 นับถือศาสนาอิสลาม) และประมาณครึ่งหนึ่งมีการศึกษาตั้งแต่มัธยมศึกษาขึ้นไป (ประมาณร้อยละ 51 เปรียบเทียบกับประมาณร้อยละ 15 มีการศึกษาต่ำกว่าประถมศึกษาและไม่ได้รับการศึกษา) โดยส่วนใหญ่ทำงานอยู่ในภาคนอกระบบ (ร้อยละ 56 เปรียบเทียบกับร้อยละ 18 ไม่ได้ทำงาน) และอาศัยอยู่ในเขตเมือง (ประมาณร้อยละ 54 เปรียบเทียบกับประมาณร้อยละ 5 อาศัยอยู่ในเขตกรุงเทพมหานคร)

เมื่อพิจารณาการกระจายตัวของประชากรตัวอย่าง เห็นได้ว่า ข้อมูลไม่ได้เบี่ยงเบนไปจากความเป็นจริงมากนัก และในแต่ละกลุ่มรายการย่อยมีจำนวนมากพอในการวิเคราะห์ด้วยสมการถดถอยอีกด้วย(ไม่มีกลุ่มรายการย่อยถูกตัดออกจากการวิเคราะห์ด้วยโปรแกรมทางสถิติ)

♦ ตัวแบบการวิเคราะห์ที่เหมาะสม: การพิจารณาด้วยสถิติพรรณนา และการทดสอบทางสถิติ

การพิจารณาตัวแบบการวิเคราะห์ที่เหมาะสมด้วยสถิติพรรณนา ข้อมูลในตารางที่ 4 ชี้ให้เห็นว่า ค่าเฉลี่ยของตัวแปรจำนวนบุตรเกิดรอด และจำนวนบุตรที่ต้องการมีเพิ่ม ไม่แตกต่างจากค่าความแปรปรวนมากนัก (ทั้งกรณีพิจารณาในภาพรวม หรือจำแนกตามตัวแปรอธิบาย) หมายความว่า การกระจายตัวของตัวแปรตามไม่มีปัญหา Over-dispersion ไม่จำเป็นต้องใช้ตัวแบบการถดถอยพหุนามในการวิเคราะห์ข้อมูล สำหรับประเด็นปัญหา Excess of zeros ข้อมูลในตารางที่ 2 ได้ชี้ให้เห็นว่า มีเพียงร้อยละ 14 ของตัวอย่างที่รายงานว่า ไม่มีบุตรเกิดรอด ซึ่งหมายความว่า ตัวแปรจำนวนบุตรเกิดรอดของผู้หญิงเคยสมรสไม่มีปัญหาการรายงานคำตอบเท่ากับ 0 เป็นจำนวนมาก ตรงกันข้ามกับตัวแปรจำนวนบุตรที่ต้องการมีเพิ่มที่ประมาณร้อยละ 81 ของตัวอย่างรายงานว่า ไม่ต้องการมีบุตรเพิ่ม ทำให้มีแนวโน้มควรเลือกใช้ตัวแบบการถดถอยที่มีผลกระทบจากศูนย์ในการวิเคราะห์ข้อมูล

ผลการพิจารณาตัวแบบการวิเคราะห์ที่เหมาะสมด้วยสถิติพรรณนา เห็นได้ว่า ตัวแปรจำนวนบุตรเกิดรอดควรเลือกใช้ตัวแบบการถดถอยปึงซองในการวิเคราะห์ข้อมูล ส่วนการกระจายตัวของตัวแปรจำนวนบุตรที่ต้องการมีเพิ่มมีความเหมาะสมกับตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์

ตาราง 4: ค่าเฉลี่ย และค่าความแปรปรวนของตัวแปรตาม จำแนกตามตัวแปรการศึกษาของผู้หญิง
ที่มา: คำนวณโดยผู้วิจัย

ระดับการศึกษา	จำนวนบุตรเกิดรอด		จำนวนบุตรที่ต้องการมีเพิ่ม	
	ค่าเฉลี่ย	ค่าความแปรปรวน	ค่าเฉลี่ย	ค่าความแปรปรวน
การศึกษาต่ำที่สุด	2.10	1.01	0.09	0.12
การศึกษาระดับต่ำ	1.94	0.89	0.16	0.18
การศึกษาระดับกลาง	1.45	0.80	0.31	0.32
การศึกษาระดับสูง	1.31	0.73	0.36	0.37
รวม	1.76	0.94	0.23	0.26

หมายเหตุ ผู้วิจัยได้ทดสอบกับทุกตัวแปรอธิบายเรียบร้อยแล้ว และผลการทดสอบไม่แตกต่างกันมากนัก

ตาราง 5: ผลการทดสอบความเหมาะสมของตัวแบบการวิเคราะห์ ด้วยวิธีการทางสถิติ
ที่มา: คำนวณโดยผู้วิจัย

การทดสอบ	จำนวนบุตรเกิดรอด	จำนวนบุตรที่ต้องการมีเพิ่ม
ตัวแบบการถดถอยอันดับ		
ค่า Pseudo R2	0.1422	0.2563
การทดสอบ proportionality of odds		
ค่า prob chi ² จากค่าสั่ง omodel	0.000	0.000
ค่า prob chi ² จากค่าสั่ง Brant	0.000	0.000
ค่า prob chi ² จากค่าสั่ง gologit	0.000	0.000
ข้อสรุป	ไม่ผ่านข้อตกลงเบื้องต้น	ไม่ผ่านข้อตกลงเบื้องต้น
ตัวแบบการถดถอยปัวซอง		
ค่า Pseudo R2	0.0661	0.2423
การสอบความเหมาะสมของตัวแบบ		
ค่า Prob chi2 จากค่าสั่ง estat gof	1.000	0.000
ข้อสรุป	มีความเหมาะสม	ต้องพิจารณาตัวแบบอื่น
ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์		
ค่า Pr z จากการทดสอบด้วยวิธี Vuong		0.000
ค่า Prob chi ² จากการทดสอบด้วยวิธี		0.000
vce		
สรุปตัวแบบที่เหมาะสม	ตัวแบบการถดถอยปัวซอง	ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์



การพิจารณาตัวแบบการวิเคราะห์ที่เหมาะสมด้วยการทดสอบทางสถิติ ข้อมูลในตารางที่ 5 ชี้ให้เห็นว่า ตัวแปรจำนวนบุตรเกิดรอด และตัวแปรจำนวนบุตรที่ต้องการมีเพิ่ม ไม่ผ่านการทดสอบข้อตกลงเบื้องต้นในการเลือกใช้ตัวแบบการถดถอยอันดับ โดยตัวแปรจำนวนบุตรเกิดรอดมีความเหมาะสมในการวิเคราะห์ด้วยตัวแบบการถดถอยปัวของ ส่วนการกระจายตัวของตัวแปรจำนวนบุตรที่ต้องการมีเพิ่ม ไม่มีความเหมาะสมกับตัวแบบการถดถอยปัวของ แต่ควรเลือกใช้ตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์ในการวิเคราะห์ข้อมูล เห็นได้ว่า ผลการวิเคราะห์ตัวแบบการวิเคราะห์ที่เหมาะสมจากการพิจารณาด้วยสถิติพรรณนา สอดคล้องกับการทดสอบทางสถิติ

◆ ความแตกต่างของผลการวิเคราะห์ระหว่างตัวแบบ

ตัวแบบการถดถอยเชิงเส้น กับตัวแบบการถดถอยปัวของ

เมื่อเปรียบเทียบผลการวิเคราะห์จำนวนบุตรเกิดรอดของผู้หญิงที่เคยสมรสอายุ 18-49 ปี ระหว่างตัวแบบการถดถอยเชิงเส้นกับตัวแบบการถดถอยปัวของ ซึ่งเป็นตัวแบบการวิเคราะห์ที่มีความเหมาะสมกับการกระจายของข้อมูลจำนวนบุตรเกิดรอด (รายละเอียดดังตารางที่ 6) พบว่า ผลการวิจัยมีความสอดคล้องต้องกันในกรณีที่พิจารณาจากความมีนัยสำคัญทางสถิติ และทิศทางความสัมพันธ์ระหว่างตัวแปรอธิบายกับจำนวนบุตรเกิดของผู้หญิง แต่ขนาดความสัมพันธ์ (Magnitude) จากตัวแบบการถดถอยเชิงเส้นสูงกว่าโดยเปรียบเทียบกับตัวแบบการถดถอยปัวของในเกือบทุกกลุ่มรายได้ของตัวแปรอธิบาย เช่น ตัวแปรการศึกษาระดับกลาง เป็นกลุ่มรายการที่มีจำนวนบุตรเกิดร่อน้อยกว่าการศึกษาระดับต่ำอย่างมีนัยสำคัญทางสถิติ โดยค่าสัมประสิทธิ์เชิงเส้นเท่ากับ -0.14 เปรียบเทียบกับค่าสัมประสิทธิ์ปัวของเท่ากับ -0.09 เป็นต้น

ความเบี่ยงของขนาดความสัมพันธ์ที่ได้จากตัวแบบการถดถอยเชิงเส้นมีผลต่องานวิจัยที่มีความสนใจความขนาดของความแตกต่างของจำนวนบุตรเกิดรอดระหว่างกลุ่มย่อยประชากร หรือต้องการค้นหาสมการที่ดีที่สุดในการพยากรณ์จำนวนบุตรเกิดรอดของผู้หญิง เนื่องจากนำไปสู่ข้อสรุปและข้อเสนอเชิงนโยบายที่เบี่ยงไปจากที่ควรเป็น กล่าวคือ ผลการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้นชี้ให้เห็นว่า ผู้หญิงที่มีการศึกษาระดับกลางและการศึกษาระดับสูงมีจำนวนบุตรเกิดร่อน้อยกว่าค่อนข้างมากโดยเปรียบเทียบผู้หญิงที่มีการศึกษาระดับต่ำ สมควรเป็นกำหนดให้เป็นประชากรเป้าหมายในการส่งเสริมการมีบุตร ซึ่งผลการวิเคราะห์ด้วยตัวแบบที่เหมาะสมกับการกระจายตัวของข้อมูลจำนวนบุตรเกิดรอดให้ผลแตกต่างไปเล็กน้อย โดยผู้หญิงที่มีการศึกษาระดับกลางและการศึกษาระดับสูงมีจำนวนบุตรเกิดร่อน้อยกว่าเล็กน้อยโดยเปรียบเทียบกับผู้หญิงที่มีการศึกษาระดับต่ำ

ตาราง 6: ผลการวิเคราะห์จำนวนบุตรเกิดรอดด้วยสถิติถดถอย
ที่มา: คำนวณโดยผู้วิจัย

ตัวแปรอธิบาย	ตัวแบบการถดถอยเชิงเส้น ค่าสัมประสิทธิ์	ตัวแบบการถดถอยปัวซอง	
		ค่าสัมประสิทธิ์	ค่า IRR
ระดับการศึกษา			
การศึกษาต่ำที่สุด	0.06 ***	0.02	1.02
การศึกษาระดับต่ำ (อ้างอิง)			
การศึกษาระดับกลาง	-0.11 ***	-0.06 ***	0.94 ***
การศึกษาระดับสูง	-0.14 ***	-0.09 ***	0.91 ***
กลุ่มอายุของผู้หญิง			
18-19 ปี	-0.21 ***	-0.19 ***	0.82 ***
20-24 ปี	-0.06 *	-0.02	0.97
25-29 ปี	-0.03	-0.01	0.99
30-34 ปี (อ้างอิง)			
35-39 ปี	-0.11 ***	-0.07 ***	0.93 ***
40-44 ปี	-0.19 ***	-0.10 ***	0.92 ***
45-49 ปี	-0.15 ***	-0.08 ***	0.92 ***
ระยะเวลาที่สมรส			
0-4 ปี (อ้างอิง)			
5-9 ปี	0.57 ***	0.05 ***	1.72 ***
10-14 ปี	0.91 ***	0.76 ***	2.15 ***
15 ปีขึ้นไป	1.34 ***	0.98 ***	2.68 ***
เขตที่อยู่อาศัย			
ในเขตเมือง (อ้างอิง)			
ในเขตชนบท	0.09 ***	0.05 ***	1.05 ***
กรุงเทพมหานคร	-0.12 ***	-0.13 ***	0.92 **
สถานภาพสมรส			
กำลังสมรส (อ้างอิง)			
หม้าย/หย่า/แยก	-0.21 ***	-0.12 ***	0.89 ***
ศาสนา			
พุทธ (อ้างอิง)			
อิสลาม	0.52 ***	0.28 ***	1.31 ***
อื่นๆ	0.21 ***	0.12 ***	1.13 ***
การมีบุตรเสียชีวิต	0.12 ***	0.06 **	1.06 **
สถานภาพการทำงาน			
ไม่ได้ทำงาน (อ้างอิง)			
ทำงานในระบบ	-0.25 ***	-0.17 ***	0.85 ***
ทำงานนอกระบบ	-0.11 ***	- 0.07 ***	0.93 ***
ระดับความมั่งคั่งของครัวเรือน			
ต่ำที่สุด	0.02	0.01	1.01
ต่ำ	0.01	0.01	1.01
ปานกลาง (อ้างอิง)			
สูง	- 0.03 *	-0.02	0.98
สูงที่สุด	0.01	0.01	1.01
ค่าคงที่	1.01 ***	-0.04	0.91

หมายเหตุ * มีนัยสำคัญที่ระดับ 0.05 ** มีนัยสำคัญที่ระดับ 0.01 *** มีนัยสำคัญที่ระดับ 0.001



ประเด็นที่น่าสนใจ คือ มีความแตกต่างของขนาดความมีนัยสำคัญทางสถิติระหว่างตัวแบบในบางกลุ่มรายของตัวแปรอธิบาย เช่น ในกรณีของตัวแปรระดับการศึกษาของผู้หญิง ผลการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้น แสดงให้เห็นว่า กลุ่มรายการระดับการศึกษาต่ำที่สุดมีจำนวนบุตรเกิดรอดสูงกว่าอย่างมีนัยสำคัญทางสถิติโดยเปรียบเทียบกับผู้หญิงที่มีการศึกษาต่ำ แต่ผลการวิเคราะห์ด้วยตัวแบบการถดถอยปัวของกลับ พบว่า จำนวนบุตรเกิดรอดไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติระหว่างผู้หญิงที่มีการศึกษาระดับต่ำที่สุดกับผู้หญิงที่มีการศึกษาระดับต่ำ หรือกรณีของตัวแปรระดับความมั่งคั่งของครัวเรือน ที่พบความแตกต่างของขนาดความมีนัยสำคัญทางสถิติระหว่างตัวแบบการถดถอยเชิงเส้นโดยเปรียบเทียบกับตัวแบบการถดถอยปัวของ ซึ่งเป็นตัวแบบการวิเคราะห์ที่มีความเหมาะสมกับการกระจายของข้อมูลจำนวนบุตรเกิดรอดด้วยเช่นกัน

ตัวแบบการถดถอยเชิงเส้น กับตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์

เมื่อเปรียบเทียบผลการวิเคราะห์จำนวนบุตรที่ต้องการมีเพิ่มของผู้หญิงที่เคยสมรสอายุ 18-49 ปี ระหว่างตัวแบบการถดถอยเชิงเส้น ตัวแบบการถดถอยปัวของ และตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์ ซึ่งเป็นตัวแบบการวิเคราะห์ที่มีความเหมาะสมกับการกระจายของข้อมูลจำนวนบุตรที่ต้องการมีเพิ่ม (รายละเอียดดังตารางที่ 7) พบว่า 1) ผลการวิจัยความสอดคล้องต้องกันในกรณีที่พิจารณาจากความมีนัยสำคัญทางสถิติ และทิศทางความสัมพันธ์ระหว่างตัวแปรอธิบายกับจำนวนบุตรที่ต้องการมีเพิ่ม แต่ 2) ขนาดความสัมพันธ์ (Magnitude) จากตัวแบบการถดถอยเชิงเส้นแตกต่างไปจากตัวแบบการถดถอยปัวของ และตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์ในเกือบทุกกลุ่มรายของตัวแปรอธิบาย อีกทั้ง ยังพบความแตกต่างของขนาดความมีนัยสำคัญทางสถิติระหว่างตัวแบบในบางกลุ่มรายของตัวแปรอธิบายเช่นเดียวกับผลการวิเคราะห์จำนวนบุตรเกิดรอด เช่น กลุ่มรายการจำนวนบุตรที่ยังมีชีวิต = 3 หรือกลุ่มอายุ 25-29 ปี หรือกลุ่มรายการระยะเวลาที่สมรส 0-4 ปี และ 5-9 ปี หรือกลุ่มรายการอาศัยอยู่ในเขตชนบท หรือกลุ่มรายการทำงานนอกระบบ เป็นต้น

ตาราง 7: ผลการวิเคราะห์จำนวนบุตรที่ต้องการมีเพิ่มขึ้นด้วยสถิติอนุमान
ที่มา: คำนวณโดยผู้วิจัย

ตัวแปรอธิบาย	โมเดล 1	โมเดล 2	โมเดล 3	
	ค่าสัมประสิทธิ์ เชิงเส้น	ค่าสัมประสิทธิ์ ปัวซอง	ค่าสัมประสิทธิ์ ปัวซอง	ค่าสัมประสิทธิ์ โลจิต
ระดับการศึกษา				
การศึกษาต่ำที่สุด	-0.01	-0.08	-0.12	-0.31
การศึกษาระดับต่ำ (อ้างอิง)				
การศึกษาระดับกลาง	-0.01	0.01	0.02	0.19
การศึกษาระดับสูง	0.06 ***	0.17 ***	0.15 **	-0.27
จำนวนบุตรเกิดรอดที่ยังมีชีวิต				
ไม่มี	0.52 ***	2.06 ***	0.85 ***	-1.82
1 คน	0.24 ***	1.47 ***	0.33 ***	- 0.32
2 คน (อ้างอิง)				
3 คน	-0.02	-0.40 ***	0.57 ***	1.64 ***
4 คนขึ้นไป	-0.07 ***	-0.69 ***	- 0.42	1.09
กลุ่มอายุ				
18-19 ปี	-0.07 **	-0.20 *	-0.18 *	-0.68
20-24 ปี	0.01	-0.06	-0.03	-0.16
25-29 ปี	0.04 ***	0.02	0.03	-0.09
30-34 ปี (อ้างอิง)				
35-39 ปี	-0.07 ***	-0.24 ***	-0.17 **	0.37 *
40-44 ปี	-0.12 ***	-0.63 ***	-0.53 ***	0.65 **
45-49 ปี	-0.13 ***	-1.02 ***	-0.80 ***	1.21 ***
ระยะเวลาที่สมรส				
0-4 ปี (อ้างอิง)				
5-9 ปี	-0.09 ***	-0.01	-0.08	-0.25
10-14 ปี	-0.14 ***	-0.08	-0.13 *	-0.21
15 ปีขึ้นไป	-0.15 ***	-0.40 ***	-0.38 ***	-0.14
เขตที่อยู่อาศัย				
ในเขตเมือง (อ้างอิง)				
ในเขตชนบท	0.12	-0.07 ***	0.07 *	0.13
กรุงเทพมหานคร	- 0.09 ***	-0.54 ***	-0.46 ***	-0.03
สถานภาพสมรส				
กำลังสมรส (อ้างอิง)				
หม้าย/หย่า/แยก	-0.16 ***	-1.56 ***	-1.69 ***	-1.04
ศาสนา				
พุทธ (อ้างอิง)				
อิสลาม	0.29 ***	0.92 ***	0.65 ***	-2.13 ***
อื่นๆ	0.08	0.33	0.36	0.01
การมีบุตรเสียชีวิต	0.01	0.05	0.06	-0.14
สถานภาพการทำงาน				
ไม่ได้ทำงาน (อ้างอิง)				
ทำงานในระบบ	-0.03 **	-0.12 **	-0.09 *	0.26
ทำงานนอกระบบ	0.02 *	0.06	0.07	-0.01



ตัวแปรอธิบาย	โมเดล 1	โมเดล 2	โมเดล 3	
	ค่าสัมประสิทธิ์ เชิงเส้น	ค่าสัมประสิทธิ์ ปัวซอง	ค่าสัมประสิทธิ์ ปัวซอง	ค่าสัมประสิทธิ์ โลจิต
ระดับความมั่งคั่งของครัวเรือน				
ต่ำที่สุด	0.01	0.05	0.01	-0.40
ต่ำ	-0.01	-0.03	-0.05	-0.24
ปานกลาง (อ้างอิง)				
สูง	-0.01	0.01	0.01	0.08
สูงที่สุด	-0.01	0.03	0.02	-0.08
ค่าคงที่	0.25 ***	-4.61 ***	-1.03 ***	1.03 ***

หมายเหตุ โมเดล 1 วิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้น

โมเดล 2 วิเคราะห์ด้วยตัวแบบการถดถอยปัวซอง

โมเดล 3 วิเคราะห์ด้วยตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์

* มีนัยสำคัญที่ระดับ 0.05 ** มีนัยสำคัญที่ระดับ 0.01 *** มีนัยสำคัญที่ระดับ 0.001

สรุปและอภิปรายผล

ตัวแปรระดับภาวะเจริญพันธุ์ของประชากร จัดเป็นตัวแปรไม่ต่อเนื่องที่มีลักษณะเป็นข้อมูลจำนวนนับ (ไม่มีค่าติดลบ) ทำให้การวิเคราะห์ความแตกต่างของจำนวนบุตรเกิดรอด หรือจำนวนบุตรที่ต้องการมีเพิ่มระหว่างกลุ่มประชากรย่อยด้วยตัวแบบการถดถอยเชิงเส้นมีความไม่เหมาะสม (Pandey & Kaur, 2015) โดยตัวแบบการวิเคราะห์ที่เหมาะสมจัดอยู่ในกลุ่ม generalized linear model เช่น ตัวแบบการถดถอยอันดับ หรือตัวแบบการถดถอยปัวซอง เป็นต้น ซึ่งขึ้นอยู่กับการกระจายตัวของข้อมูลตัวแปรตามเป็นสำคัญ

ประเด็นที่น่าสนใจเกี่ยวกับการกระจายตัวของตัวแปรระดับภาวะเจริญพันธุ์ของผู้หญิงอายุ 18-49 ปี ในปี 2552 คือ 1) ผู้หญิงไทยมีระดับภาวะเจริญพันธุ์สมรสในปัจจุบันอยู่ในระดับต่ำกว่าระดับทดแทน และความผันแปรของจำนวนบุตรเกิดรอดไม่สูงนัก (ตารางที่ 4 ค่าเฉลี่ยเท่ากับ 1.76 และค่าความแปรปรวนเท่ากับ 0.94) โดยมีเพียงประมาณร้อยละ 14 ไม่มีบุตรเกิดรอด และประมาณร้อยละ 19 มีจำนวนบุตรเกิดรอด 3 คนขึ้นไป (ดังแสดงตารางที่ 2) หมายความว่า ประชากรไทยมีแนวโน้มมีจำนวนบุตรเกิดรอดค่อนข้างน้อย และไม่มีความแตกต่างกันระหว่างประชากรกลุ่มย่อยมากนัก ทำให้การกระจายของข้อมูลจำนวนบุตรเกิดรอดในปี 2552 ไม่มีปัญหา Over-dispersion และไม่มีปัญหาการรายงานตัวแปรมีค่า เท่ากับ 0 เป็นจำนวนมาก จึงไม่จำเป็นต้องใช้ตัวแบบการถดถอยทวินาม หรือตัวแบบการถดถอยที่มีผลกระทบจากศูนย์ในการวิเคราะห์ข้อมูลแต่อย่างใด

สำหรับตัวแปรจำนวนบุตรที่ต้องการมีเพิ่ม เกือบทั้งหมดไม่ต้องการมีบุตรเพิ่ม (ประมาณร้อยละ 81) โดยมีเพียงร้อยละ 4 ที่ต้องการมีบุตรเพิ่มอีก 2 คนขึ้นไป (ดังแสดงตารางที่ 3) เห็นได้ว่า การกระจายตัวของจำนวนบุตรที่ตั้งใจมีเพิ่มไม่มีปัญหา Over-dispersion (ตารางที่ 4 ค่าเฉลี่ยเท่ากับ 0.23 และค่าความแปรปรวนเท่ากับ 0.26) เช่นเดียวกันกับตัวแปรจำนวนบุตรเกิดรอด หรืออาจกล่าวได้ว่า ผู้หญิงที่เคยสมรสอายุ 18-49 ปี มีภาวะเจริญพันธุ์ในปัจจุบันและความตั้งใจมีบุตรเพิ่มไม่แตกต่างกันมากนักระหว่างประชากรกลุ่มย่อย แต่อย่างไรก็ตามตัวแปรจำนวนบุตรที่ต้องการมีเพิ่มมีปัญหา

การรายงานตัวแปรที่มีค่า เท่ากับ 0 เป็นจำนวนมาก ทำให้ตัวแบบการถดถอยที่มีผลกระทบจากศูนย์เป็นตัวแบบที่เหมาะสม

ผลการวิเคราะห์ตัวแบบการวิเคราะห์ที่เหมาะสมจากการพิจารณาด้วยสถิติพรรณนา มีความสอดคล้องต้องกันกับการทดสอบทางสถิติ กล่าวคือ ตัวแปรจำนวนบุตรเกิดรอดควรเลือกใช้ตัวแบบการถดถอยปัวซองในการวิเคราะห์ข้อมูล ส่วนการกระจายตัวของตัวแปรจำนวนบุตรที่ต้องการมีเพิ่มมีความเหมาะสมกับตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ ซึ่งมีประเด็นที่น่าสนใจเกี่ยวกับลักษณะจำเพาะของการกระจายตัวของข้อมูลระดับภาวะเจริญพันธุ์ในปี 2552 กล่าวคือ 1) ค่าเฉลี่ย และค่าความแปรปรวนไม่ต่างกันมากนัก (ตารางที่ 4 ค่าเฉลี่ยเท่ากับ 1.76 เปรียบเทียบกับค่าความแปรปรวนเท่ากับ 0.94 และค่าเฉลี่ยเท่ากับ 0.23 เปรียบเทียบกับค่าความแปรปรวนเท่ากับ 0.26) ทำให้มีความชัดเจนว่า การกระจายของข้อมูลไม่มีปัญหา Over-dispersion แต่ในบางกรณีที่มีค่าความแปรปรวนของตัวแปรตามสูงกว่าค่าเฉลี่ย อาจทำให้การวิเคราะห์ตัวแบบที่เหมาะสมด้วยสถิติพรรณนาอาจนำไปสู่ข้อสรุปที่ไม่ถูกต้อง จำเป็นต้องใช้การทดสอบทางสถิติ และ 2) สัดส่วนของตัวอย่างที่รายงานว่า ไม่มีบุตรเกิดรอด หรือไม่ต้องการมีบุตรเพิ่ม (ประมาณร้อยละ 14 เปรียบเทียบกับร้อยละ 81) ค่อนข้างชัดเจนว่า ตัวแบบการถดถอยที่มีผลกระทบจากศูนย์หรือไม่ ทำให้ผลการวิเคราะห์ตัวแบบการวิเคราะห์ที่เหมาะสมจากการพิจารณาด้วยสถิติพรรณนามีความสอดคล้องต้องกันกับการทดสอบทางสถิติ

ตัวแบบการถดถอยอันดับ จัดอยู่ในกลุ่ม generalized linear model เช่นเดียวกับตัวแบบการถดถอยปัวซอง แต่ไม่ใช่ตัวแบบที่มีความเหมาะสมในการวิเคราะห์จำนวนบุตรเกิดรอด หรือจำนวนบุตรที่ต้องการมีเพิ่มของผู้หญิงที่เคยสมรสอายุ 18-49 ปี ในปี 2552 เนื่องจากไม่ผ่านข้อตกลงเบื้องต้นจากการทดสอบ proportionality of odds (ดังแสดงตารางที่ 5) ซึ่งเป็นปกติวิสัยของข้อมูลที่ได้จากการสำรวจ (Long & Freese, 2006)

เมื่อเปรียบเทียบผลการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงกับตัวแบบที่มีความเหมาะสมแล้ว พบว่า ส่วนใหญ่มีความสอดคล้องต้องกันของความมีนัยสำคัญทางสถิติ และทิศทางความสัมพันธ์ โดยมีบางกลุ่มรายการที่มีความแตกต่างของระดับความมีนัยสำคัญระหว่างตัวแบบในการวิเคราะห์ กล่าวคือ ผลการวิเคราะห์ด้วยตัวแบบการถดถอยเชิง มีความแตกต่างอย่างมีนัยสำคัญ แต่ผลการวิเคราะห์ด้วยตัวแบบที่เหมาะสมไม่มีนัยสำคัญ หรือในทางกลับกันอีกด้วย (ดังแสดงตารางที่ 6 และตารางที่ 7) ซึ่งอาจเป็นไปได้ว่า ในปี 2552 ประชากรตัวอย่างมีจำนวนบุตรเกิดรอดของประมาณ 1-2 คน และส่วนใหญ่ไม่ต้องการมีบุตรเพิ่ม โดยตัวอย่างที่ต้องการมีบุตรเพิ่มส่วนใหญ่ต้องการมีเพิ่มประมาณ 1-2 คน เห็นได้ว่า การกระจายตัวของตัวแปรตามมีลักษณะไม่ใช่ต่อแปรต่อเนื่องอย่างเห็นได้ชัด ทำให้ผลการวิเคราะห์ด้วยตัวแบบการถดถอยเชิงเส้นมีความเบี่ยงไปจากความเป็นจริง และอาจนำไปสู่การกำหนดประชากรเป้าหมายในการส่งเสริมการมีบุตรที่ไม่ถูกต้อง

ตัวแบบการถดถอยปัวซองที่มีผลกระทบจากศูนย์ เป็นตัวแบบที่มีจุดเด่นในการวิเคราะห์โอกาสของประชากรตัวอย่างจะไม่ต้องการมีบุตรเพิ่ม (พิจารณาได้จากค่า



สัมประสิทธิ์โลจิสต์) ไปพร้อมๆ กับการวิเคราะห์จำนวนบุตรที่ต้องการมีเพิ่ม ซึ่งช่วยให้เข้าใจความต้องการของประชากรตัวอย่างได้อย่างชัดเจนยิ่งขึ้น โดยการแปรผลจากตัวแบบการถดถอยเชิงเส้น และตัวแบบการถดถอยปัวของ เป็นการพิจารณาจำนวนบุตรเกิดที่เพิ่มขึ้นหรือลดลงจากการที่ตัวแปรอิสระที่เป็นตัวแปรต่อเนื่องเพิ่มหรือลดลง เช่น ค่าสัมประสิทธิ์เชิงเส้นหรือสัมประสิทธิ์ปัวของของอายุเต็มปี เท่ากับ -0.01 หมายความว่า อายุที่เพิ่มขึ้น 1 ปี ทำให้จำนวนบุตรที่ต้องการมีเพิ่มลดลง 0.01 คน หรือในกรณีที่เป็นตัวแปรกลุ่มรายการ เช่น ค่าสัมประสิทธิ์เชิงเส้นหรือสัมประสิทธิ์ปัวของของระดับการศึกษาสูง เท่ากับ 0.06 หมายความว่า โดยเปรียบเทียบกับประชากรตัวอย่างที่มีการศึกษาระดับต่ำแล้ว ประชากรตัวอย่างที่มีการศึกษาสูงมีจำนวนบุตรที่ต้องการมีเพิ่มสูงกว่าโดยเปรียบเทียบประมาณ 0.06 คน สำหรับการแปรผลค่าสัมประสิทธิ์โลจิสต์จากตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์ เป็นการพิจารณาโอกาสของประชากรตัวอย่างที่ไม่ต้องการมีบุตรเพิ่มเทียบกับประชากรตัวอย่างกลุ่มอ้างอิง เช่น กรณีมีบุตรเกิดรอดที่ยังมีชีวิตจำนวน 3 คน มีค่าสัมประสิทธิ์โลจิสต์ = 1.64 หมายความว่า โดยเปรียบเทียบกับประชากรตัวอย่างที่มีบุตรเกิดรอดที่ยังมีชีวิตจำนวน 2 คนแล้ว ประชากรตัวอย่างที่มีบุตรเกิดรอดที่ยังมีชีวิตจำนวน 3 คน มีสัดส่วนความเสี่ยงที่จะไม่ต้องการมีบุตรเพิ่มมากกว่าประมาณ 4.16 เท่า ($e^{1.64} - 1 = 5.16 - 1 = 4.16$) หรือกรณีนับถือศาสนาอิสลามมีค่าสัมประสิทธิ์โลจิสต์ = -2.13 หมายความว่า โดยเปรียบเทียบกับผู้หญิงที่นับถือศาสนาพุทธแล้ว ผู้หญิงที่นับถือศาสนาอิสลามมีสัดส่วนความเสี่ยงที่จะไม่ต้องการมีบุตรน้อยกว่าประมาณ 7.41 เท่า เป็นต้น

ข้อเสนอแนะในเชิงวิชาการ ในกรณีงานวิจัยมีวัตถุประสงค์เพียงพิจารณาแนวโน้มความสัมพันธ์ระหว่างตัวแปรอธิบายกับตัวแปรตามที่เป็นจำนวนนับ แต่ผลการทดสอบพบว่า ตัวแบบการถดถอยปัวของเป็นตัวแบบการวิเคราะห์ที่เหมาะสมกับการกระจายตัวของข้อมูล ผู้วิจัยอาจเลือกนำเสนอข้อมูลจากตัวแบบการถดถอยเชิงเส้นหรือตัวแบบการถดถอยอันดับ ซึ่งเป็นตัวแบบการวิเคราะห์ที่ได้รับความนิยมโดยทั่วไปย่อมสามารถทำได้ ในกรณีที่ได้ทดสอบความสอดคล้องของผลการวิเคราะห์เรียบร้อยแล้ว ส่วนในกรณีที่งานวิจัยมีวัตถุประสงค์ต้องการสร้างสมการพยากรณ์ของตัวแปรตามที่เป็นจำนวนนับ หรือสนใจขนาดความแตกต่างระหว่างกลุ่มย่อยประชากร จำเป็นต้องนำเสนอผลการวิเคราะห์จากตัวแบบการถดถอยที่มีความเหมาะสมกับข้อมูลตัวแปรตาม เช่น ตัวแบบการถดถอยปัวของ ตัวแบบการถดถอยทวินามลบ ตัวแบบการถดถอยปัวของที่มีผลกระทบจากศูนย์ หรือตัวแบบการถดถอยทวินามลบที่มีผลกระทบจากศูนย์ เป็นต้น สำหรับในกรณีที่ข้อมูลตัวแปรตามมีปัญหาการรายงานค่าเท่ากับ 0 เป็นจำนวนมาก และการรายงานค่า 0 มีความหมายในเชิงกำหนดนโยบาย เช่น การไม่ต้องการมีบุตรเพิ่ม เป็นต้น ควรนำเสนอด้วยตัวแบบการถดถอยที่มีผลกระทบจากศูนย์ยอมทำให้ได้ภาพความต้องการของประชากรชัดเจนยิ่งขึ้น

กิตติกรรมประกาศ

ผู้เขียน ขอขอบคุณสำนักงานสถิติแห่งชาติที่อนุญาตให้ใช้ข้อมูลดิบจากการสำรวจอนามัยการเจริญพันธุ์ ปี 2552 เป็นข้อมูลหลักในการวิเคราะห์ข้อมูล และขอขอบพระคุณ ผศ.ดร. วิราภรณ์ โพธิศิริ วิทยาลัยประชากรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ดร. รยา มุตตารักษ์ ดร. Maria Rita Teats และ ดร. Erich Stressing สังกัด Vienna Institute of Demography ประเทศออสเตรีย ที่ให้ข้อเสนอแนะในการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติ และประเด็นอันเป็นประโยชน์ต่อบทความ

บทความนี้เป็นส่วนหนึ่งของวิทยานิพนธ์ปริญญาดุษฎีบัณฑิต สาขาประชากรศาสตร์ วิทยาลัยประชากรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ผู้เขียนขอขอบพระคุณอาจารย์ที่ปรึกษา และคณะกรรมการผู้ทรงคุณวุฒิทุกท่านที่กรุณาให้ความเป็นอันเป็นประโยชน์ต่องานของศิษย์เสมอมา

Notes

¹ โดยนิยามแล้ว หมายถึง สังคมที่ผู้หญิงมีบุตรตลอดช่วงวัยเจริญพันธุ์โดยเฉลี่ยประมาณ 2.1 คน โดยหลังจากปี 2533 เป็นต้นมา ภาวะเจริญพันธุ์ของประเทศไทยต่ำกว่าระดับทดแทน และในปี 2553 อยู่ที่ระดับประมาณ 1.5-1.6 คน

² ประชากรมีแนวโน้มเลื่อนการสมรสออกไปทำให้ในบางกลุ่มอายุรายงานว่ามีจำนวนบุตรเกิดรอดเท่ากับ 0 เป็นจำนวนมาก ประกอบกับมีประชากรบางส่วนยินดีครองโสดตลอดช่วงชีวิต หรือยินดีที่จะมีคู่ครอง แต่ไม่มีบุตร หรือมีปัญหาบุตรยากอีกด้วย สำหรับประเด็นความต้องการมีบุตรเพิ่ม ประชากรรายงานว่า ไม่ต้องการมีบุตรเพิ่ม เนื่องจากส่วนใหญ่มีบุตรได้ครบจำนวนที่ต้องการประมาณ 1-2 คน เรียบร้อยแล้ว และบางส่วนยินดีที่จะมีคู่ครอง แต่ไม่มีบุตร

³ สำหรับฟังก์ชันของแต่ละตัวแบบการวิเคราะห์ ศึกษาโดยละเอียดเพิ่มเติมได้จาก Long & Freese (2006) Zeileis et al. (2008) และ Cameron & Trivedi (2013)

⁴ ประชากรมีแนวโน้มเลื่อนการสมรสออกไปทำให้ในบางกลุ่มอายุรายงานว่ามีจำนวนบุตรเกิดรอดเท่ากับ 0 เป็นจำนวนมาก ประกอบกับมีประชากรบางส่วนยินดีครองโสดตลอดช่วงชีวิต หรือยินดีที่จะมีคู่ครอง แต่ไม่มีบุตร หรือมีปัญหาบุตรยากอีกด้วย สำหรับประเด็นความต้องการมีบุตรเพิ่ม ประชากรรายงานว่า ไม่ต้องการมีบุตรเพิ่ม เนื่องจากส่วนใหญ่มีบุตรได้ครบจำนวนที่ต้องการประมาณ 1-2 คน เรียบร้อยแล้ว และบางส่วนยินดีที่จะมีคู่ครอง แต่ไม่มีบุตร

⁵ การศึกษานี้เลือกใช้จำนวนบุตรเกิดรอด เนื่องจากพบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติกับจำนวนบุตรที่ยังมีชีวิตอยู่ ทั้งนี้จะเป็นเพราะอัตราตายของทารกและมารดาของประเทศไทยอยู่ในระดับต่ำ (มีกลุ่มตัวอย่างเพียงร้อยละ 6 รายงานว่ามีบุตรเกิดรอดที่เสียชีวิต (Dead children)) ทำให้ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างจำนวนบุตรเกิดรอดกับจำนวนบุตรเกิดรอดที่ยังมีชีวิตในปัจจุบัน

References

- Berrington, A., & Pattaro, S. (2014). Educational differences in fertility desires, intentions and behaviour: A life course perspective. *Advances in life course research*, 21: 10-27.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53): Cambridge university press.



- Fagbamigbe, A. F., & Adebowale, A. S. (2014). Current and Predicted Fertility using Poisson Regression Model: Evidence from 2008 Nigerian Demographic Health Survey. *African journal of reproductive health*, 18 (1): 71-83.
- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure Data—Or tears: An application to educational enrollments in states of india*. *Demography*, 38 (1): 115-132.
- Fu, V. K. (1998). Estimating generalized ordered logit models. *Stata Technical Bulletin*, 8 (44): 27-30.
- King, G. (1989). Variance specification in event count models: From restrictive assumptions to a generalized estimator. *American Journal of Political Science*, 762-784.
- Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata*: Stata press.
- Morgan, S. P. (1982). Parity-specific fertility intentions and uncertainty: The United States, 1970 to 1976. *Demography*, 19 (3): 315-334.
- Pandey, R., & Kaur, C. (2015). Modelling fertility: an application of count regression models. *Chinese Journal of Population Resources and Environment*, 13 (4): 349-357.
- Poston Jr, D. L., & McKibben, S. L. (2003). Using zero-inflated count regression models to estimate the fertility of US women. *Journal of Modern Applied Statistical Methods*, 2 (2): 10.
- Testa, M. R. (2014). On the positive correlation between education and fertility intentions in Europe: Individual-and country-level evidence. *Advances in Life Course Research*, 21: 28-42.
- Testa, M. R., Bordone, V., Osiewalska, B., & Skirbekk, V. (2016). Are daughters' childbearing intentions related to their mothers' socio-economic status? *Demographic Research*, 35: 581-616.
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, 57: 307-333.
- Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal*, 6 (1): 58-82.
- Winkelmann, R., & Zimmermann, K. F. (1994). Count data models for demographic data. *Mathematical Population Studies*, 4 (3): 205-221.
- Wolfe, R., & Gould, W. (1998). An approximate likelihood-ratio test for ordinal response models. *Stata Technical Bulletin*, 7 (42): 22-27.
- Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, 27 (8): 1-25.