วารสารวิชาการสาธารณสุข
ปีที่ 33 ฉบับเพิ่มเติม 1 กันยายน - ตุลาคม 2567

Journal of Health Science of Thailand
Vol. 33 Supplement 1, September - October 2024

นิพนธ์ต้นฉบับ *Original article*

ความแม่นยำในการประมาณค่าระหว่างสมการ ZINB for Count กับ ZINB for Rate ในงานวิจัยทางด้านทันตสาธารณสุข

นิตยา พวงราช ปร.ด. (วิทยาการระบาดและชีวสถิติ)
โรงพยาบาลคูเมือง อำเภอคูเมือง จังหวัดบุรีรัมย์
ติดต่อผู้เขียน: นิตยา พวงราช Email: nidarach@hotmail.com

วันรับ: 5 ต.ค. 2566
วันแก้ไข: 6 ก.พ. 2567
วันตอบรับ: 16 ก.พ. 2567

บทคัดย่อ การวิเคราะห์ข้อมูลด้วยสมการถดถอยพหุปัวส์ซอง จำแนกเป็น 2 รูปแบบ คือ Poisson for Count และ Poisson for Rate โรคฟันผุมีปัจจัยทางด้านตัวฟัน อาหาร เชื้อโรค และระยะเวลาที่กรดสัมผัสกับตัวฟัน เป็นปัจจัยเสี่ยงต่อการเกิดโรคดังกล่าว การวิเคราะห์ข้อมูลที่ดีควรคำนึงถึงผลกระทบของระยะเวลาที่สัมผัสปัจจัยเสี่ยงต่อการสรุปผล หากผู้วิจัยไม่มีข้อมูลระยะเวลาสัมผัสปัจจัยเสี่ยง ตัวแปรที่สามารถนำมาวิเคราะห์ข้อมูลแทนได้ เช่น อายุ หรือจำนวนที่สามารถระบุความสัมพันธ์กับระยะเวลาที่สัมผัสปัจจัยเสี่ยง การศึกษานี้ใช้ข้อมูลจากการศึกษาเรื่อง “ความสัมพันธ์ระหว่างเงินค่าขนมกับปริมาณฟันผุ ถอน อุด (DMFT) ในเด็กอายุ 12 ปี ที่ศึกษาในโรงเรียนประถมศึกษา อำเภอคูเมือง จังหวัดบุรีรัมย์” ตัวแปรตามพบ over-dispersion และมี zero outcome ร้อยละ 11.37 ผู้วิจัยใช้สมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (zero inflated negative binomial regression; ZINB) ซึ่งเป็นสถิติตัวหนึ่งในตระกูลปัวส์ซอง เป็นเครื่องมือในการแสวงหาคำตอบ จากการศึกษาที่ผ่านมาพบว่ามีการนำสมการ ZINB มาใช้วิเคราะห์การเกิดโรคฟันผุในเด็ก แต่ไม่มีการระบุว่าเป็นการวิเคราะห์แบบ Count หรือแบบ Rate ดังนั้น การศึกษานี้มีวัตถุประสงค์เพื่อประเมินความแม่นยำในการประมาณค่าของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate) และเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของวิธีการทางสถิติดังกล่าว ผลการศึกษาพบว่า การประมาณค่าพารามิเตอร์ของสมการ ZINB เมื่อพิจารณาความกระชับในการประมาณค่า สมการ ZINB for Rate ให้ค่าช่วงความเชื่อมั่น 95% CI แคบกว่าสมการ ZINB for Count นอกจากนี้ สมการ ZINB for Rate ยังประมาณค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในโมเดลการเกิดโรคฟันผุต่ำกว่าสมการ ZINB for Count จุดเด่นของสมการ ZINB คือ สมการมีรูปแบบโมเดลเป็นแบบ Mixed Effect Model ซึ่งในสมการมีการประมาณค่าพารามิเตอร์แบบ Fixed Effect และ Random Effect จึงช่วยทำให้ความคลาดเคลื่อนในการประมาณค่าลดลง

คำสำคัญ: **ปัวส์ซอง; สมการ ZINB; การประมาณค่า; ความคลาดเคลื่อน**

บทนำ

การวิเคราะห์ข้อมูลด้วยสมการถดถอยพหุปัวส์ซอง จำแนกเป็น 2 รูปแบบ คือ Poisson for Count ใช้ประมาณค่าหรือหาความสัมพันธ์กรณีที่ตัวแปรตามเป็นจำนวนนับ และ Poisson for Rate ใช้ประมาณค่าหรือหาความสัมพันธ์กรณีที่ตัวแปรตามเป็นอัตรา ซึ่งคำนวณจากจำนวนการเกิดเหตุการณ์ทั้งหมดหารด้วยระยะเวลาที่เพิ่มขึ้นหรือปัจจัยอื่นที่บ่งชี้ถึงขอบเขตที่กำหนด (index of size)[1,2] เป้าหมายของการวิเคราะห์ถดถอยพหุปัวส์ซอง คือ หาความสัมพันธ์ระหว่างตัวแปรตามที่มีลักษณะของข้อมูลเป็นจำนวนนับ (count data) กับตัวแปรอิสระในโมเดลตั้งแต่ 2 ตัวแปรขึ้นไป โดยควบคุมอิทธิพลของปัจจัยอื่นไปพร้อม ๆกัน ทั้งนี้ตัวแปรอิสระสามารถเป็นข้อมูลเชิงคุณภาพหรือข้อมูลเชิงปริมาณก็ได้ อย่างไรก็ตามมีการกล่าวถึง อัตรา (rate) หมายถึง ตัวแปรตามที่มีลักษณะเป็นจำนวนนับ เก็บรวบรวมภายใต้ช่วงระยะเวลา หรือพื้นที่ที่กำหนด[3] ส่วนอัตราอุบัติการณ์ (incidence rate) หมายถึง จำนวนผู้ป่วยใหม่ที่เกิดขึ้นต่อหน่วยประชากรที่เฝ้าระวังในช่วงระยะเวลาที่กำหนด[4] เป็นตัวบ่งชี้ถึงความรวดเร็วของการเกิดเหตุการณ์หรือโรคในระยะเวลาที่ศึกษาและเป็นการระบุถึงความรุนแรงของการแพร่กระจายโรคในแต่ละพื้นที่ [5] ทั้งนี้ความรุนแรงของโรคมีการพิจารณาจากหลากหลายตัวชี้วัด เช่น ความรุนแรงของโรคฟันผุพิจารณาจากปริมาณฟันถาวรผุ (Decayed) + ปริมาณฟันถาวรที่ถูกถอน (Missing) + ปริมาณฟันถาวรที่ได้รับการอุด (Filled) มีหน่วยนับเป็นซี่ต่อคน (DMFT) โดยมีปัจจัยทางด้านตัวฟัน อาหาร เชื้อโรค และระยะเวลาที่กรดสัมผัสกับตัวฟัน เป็นปัจจัยเสี่ยงต่อการเกิดโรคฟันผุ[6] อย่างไรก็ตาม Long and Freese [7] กล่าวถึง ระยะเวลาสัมผัสปัจจัยเสี่ยง (exposure times) เป็นเรื่องยากที่จะระบุว่ากลุ่มตัวอย่างที่ศึกษาสัมผัสปัจจัยเสี่ยงเท่ากัน แม้ว่าระยะเวลาในการศึกษาจะเท่ากัน (same period of time) ดังนั้น การวิเคราะห์ข้อมูลจึงมีการคำนึงถึงผลกระทบของระยะเวลาที่สัมผัสปัจจัยเสี่ยงต่อการสรุปผล-หากผู้วิจัยไม่มีข้อมูลระยะเวลาสัมผัสปัจจัยเสี่ยง ตัวแปรที่สามารถนำมาวิเคราะห์ข้อมูลแทนได้ เช่น อายุ หรือจำนวนที่สามารถระบุความสัมพันธ์กับระยะเวลาที่สัมผัสปัจจัยเสี่ยง[7] อย่างไรก็ตาม สมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (Zero Inflated Negative Binomial Regression; ZINB) เป็นสถิติอีกตัวหนึ่งในตระกูลปัวส์ซอง ซึ่งจากการทบทวนวรรณกรรมพบว่ามีการนำสมการ ZINB มาใช้วิเคราะห์การเกิดโรคฟันผุในเด็ก แต่ไม่มีการระบุว่าเป็นการวิเคราะห์แบบ Count หรือแบบ Rate[8-11] ทั้งนี้เนื่องจากวิธีการประมาณค่าที่แตกต่างกันจะทำให้ค่าความคลาดเคลื่อนแตกต่างกันและผลลัพธ์ข้อสรุปที่ได้จากการวิเคราะห์ข้อมูลย่อมแตกต่างกันด้วย

สมการถดถอยทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) เป็นสมการที่มีรูปแบบการแจกแจงแบบผสม (Mixed Distribution) ระหว่างการแจกแจงทวินาม (binomial distribution) กับการแจกแจงแกมม่า (gamma distribution) ซึ่งเป็นโมเดลผสม (mixed model) ของโมเดลการถดถอยโลจิสติกส์ (logistic model) หรือโมเดลการถดถอยโปรบิท (probit model) กับโมเดลการถดถอยทวินามแบบลบ (Negative Binomial Regression Model) สำหรับรูปแบบสมการมีดังนี้

1) ฟังก์ชั่นความน่าจะเป็นของสมการถดถอยพหุปัวส์ซองกรณีตัวแปรตามมีค่าศูนย์

$$\Pr(y_i = 0 \mid x_i, z_i) = \psi_i + (1-\psi_i)\exp(-\mu_i)$$

$$\Pr(y_i > 0 \mid x_i, z_i) = (1-\psi_i)\frac{\exp(-\mu_i)\mu_i^{y_i}}{y_i!}$$

2) ฟังก์ชั่นความน่าจะเป็นของสมการถดถอยพหุทวินามแบบลบ

$$\Pr[y_i \mid x_i, \alpha] = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i}$$

α (dispersion parameter) ≥ 0, $y_i = 0, 1, 2, 3\ldots$

ฟังก์ชั่นความน่าจะเป็นของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) เขียนเป็นสมการได้ ดังนี้

3) $$\Pr(y_i = 0 \mid x_i, z_i) = \psi_i + (1-\psi_i)\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}$$

4) $\Pr(y_i>0|x_i,z_i)=(1-\psi_i)\frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i}$

เมื่อกำหนดให้

$y_i = 0$, with probability ψ_i

y_i = Poisson-Gamma (μ_i), with probability $1-\psi_i$

$E(y_i|x_i, z_i) = (1-\psi_i)\mu_i$

$V(y_i|x_i, z_i) = \mu_i(1-\psi_i)[1+\mu_i(\psi_i+\alpha)]$

$\mu_i = \exp(\beta_i' x_i)$

$$\psi_i = \frac{\exp(z_i'\gamma)}{1+\exp(z_i'\gamma)}$$

β_i คือ ค่าสัมประสิทธิ์การถดถอยพหุจาก Poisson - Gamma Component

γ_i คือ ค่าสัมประสิทธิ์การถดถอยพหุจาก Logistic Component

x_i และ z_i คือ ตัวแปรอิสระที่มีอยู่ในโมเดล

การศึกษานี้มีวัตถุประสงค์เพื่อประเมินความแม่นยำในการประมาณค่าของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate) และเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของวิธีการทางสถิติดังกล่าว

วิธีการศึกษา

การศึกษานี้เป็นส่วนหนึ่งจากการศึกษา เรื่องความสัมพันธ์ระหว่างเงินค่าขนมกับปริมาณฟันผุ ถอน อุด (DMFT) ในเด็กอายุ 12 ปี ที่ศึกษาในโรงเรียนประถมศึกษา อำเภอคูเมือง จังหวัดบุรีรัมย์ ซึ่งผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์จากมหาวิทยาลัยขอนแก่น เป็นการศึกษาเชิงวิเคราะห์แบบภาคตัดขวาง (analytical cross sectional study) กลุ่มตัวอย่าง 299 คน ได้จากการประมาณการด้วยวิธีของ Signorini[12] ซึ่งปรับขนาดตัวอย่างตามการวิเคราะห์ถดถอยพหุปัวส์ซอง และค่า item non response rate ใช้วิธีการสุ่มตัวอย่างแบบมีระบบ (systematic random sampling) เก็บรวบรวมข้อมูลโดยใช้แบบสัมภาษณ์ แบบบันทึกเงินค่าขนม แบบบันทึกการตรวจฟัน ผู้วิจัยปรับมาตรฐานการตรวจฟันกับทันตแพทย์

การศึกษาปริมาณฟันผุ ถอน อุด (DMFT) กำหนดให้กลุ่มตัวอย่างแต่ละคนมีระยะเวลาในการศึกษาเท่ากัน โดยนับระยะเวลาตั้งแต่ฟันถาวรซี่แรกขึ้นมาในช่องปาก เมื่ออายุ 6 ปี และฟันถาวรขึ้นครบหรือใกล้เคียง 28 ซี่ เมื่ออายุประมาณ 12 ปี แม้ว่ากลุ่มตัวอย่างจะมีระยะเวลาในการศึกษาเท่ากัน แต่เมื่อพิจารณาอายุของฟันที่ขึ้นในช่องปากเด็กแต่ละคน พบว่า ฟันแต่ละซี่ที่ขึ้นมาในช่องปากมีอายุการขึ้นต่างกัน ซึ่งมีผลต่อจำนวนซี่ฟันในช่องปาก รวมถึงระยะเวลาในการสัมผัสปัจจัยเสี่ยงของฟันแต่ละซี่ในเด็กแต่ละคนย่อมแตกต่างกัน ดังนั้น เพื่อเปรียบเทียบให้เห็นความแตกต่างในการประมาณค่า ผู้วิจัยจึงกำหนดให้จำนวนซี่ฟันที่มีทั้งหมดในช่องปากของเด็กแต่ละคนเป็นพื้นที่ที่ศึกษา (index of size) โดยกำหนดให้เป็นตัวแปร Offset และนำเข้าสู่การวิเคราะห์ข้อมูลในโมเดล ZINB for Rate ในขณะที่โมเดล ZINB for Count ไม่ได้นำตัวแปรจำนวนซี่ฟันทั้งหมดในช่องปากเข้าสู่การวิเคราะห์ข้อมูล โดยค่าอัตราส่วนอัตราอุบัติการณ์ (incident rate ratio: IRR) สำหรับโมเดล ZINB for Rate ของแต่ละตัวแปรคำนวณมาจาก (ปริมาณฟันผุ ถอน อุด หารด้วยจำนวนซี่ฟันที่มีทั้งหมดในช่องปากของเด็กแต่ละคน ในกลุ่มที่สัมผัสปัจจัยเสี่ยง)/(ปริมาณฟันผุ ถอน อุด หารด้วยจำนวนซี่ฟันที่มีทั้งหมดในช่องปากของเด็กแต่ละคน ในกลุ่มที่ไม่ได้สัมผัสปัจจัยเสี่ยง)

การศึกษาความสัมพันธ์ระหว่างเงินค่าขนมกับปริมาณฟันผุ ถอน อุด (DMFT) เก็บรวบรวมข้อมูลในปี 2556 พบ over-dispersion และมี zero outcome ร้อยละ 11.37 ผู้วิจัยใช้สมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) เป็นเครื่องมือในการแสวงหา

คำตอบ วิเคราะห์ข้อมูลโมเดล ZINB for Count และ ZINB for Rate ด้วยโปรแกรม STATA version 14

ผลการศึกษา

กลุ่มตัวอย่างเด็กอายุ 12 ปี ปราศจากโรคฟันผุ 34 คน ร้อยละ 11.37 (95%CI: 8.0-15.53) ค่าเฉลี่ยของปริมาณฟันผุ ถอน อุด (DMFT) มีค่าเท่ากับ 4.08 ซี่ต่อคน (ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 3.06 ซี่ต่อคน) โดยพบ ค่าเฉลี่ยของปริมาณฟันถาวรผุ (Decayed) เท่ากับ 2.74 ซี่ต่อคน (ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 2.67 ซี่ต่อคน) ค่าเฉลี่ยของฟันถาวรที่ถูกถอนจากโรคฟันผุ (Missing) เท่ากับ 0.16 ซี่ต่อคน (ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 0.52 ซี่ต่อคน) ค่าเฉลี่ยของฟันถาวรที่ได้รับการอุด (Filled) เท่ากับ 1.18 ซี่ต่อคน (ส่วนเบี่ยงเบนมาตรฐาน เท่ากับ 1.52 ซี่ต่อคน) รายละเอียดนำเสนอในตารางที่ 1

การประมาณค่าพารามิเตอร์ของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate) เมื่อพิจารณาความกระชับในการประมาณค่า พบว่า สมการ ZINB for Rate ให้ค่าช่วงความเชื่อมั่น 95% CI แคบกว่าสมการ ZINB for Count นอกจากนี้สมการ ZINB for Rate ยังประมาณค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในโมเดลต่ำกว่าสมการ ZINB for Count เมื่อพิจารณาความแม่นยำในการประมาณค่าเป็นรายตัวแปรโดยเฉพาะตัวแปรการแปรงฟัน ซึ่งผู้วิจัยกำหนดให้การแปรงทุกวันเป็นปัจจัยป้องกัน (Protective Factor) พบว่า สมการ ZINB for Rate สามารถประมาณค่า IRR_{Adj} = 1.37 (95%CI: 0.98-1.90) ออกมาในทิศทางตรงกันข้ามกับสมการ ZINB for Count ซึ่งให้ค่าเป็นปัจจัยเสี่ยง (Risk Factor) IRR_{Adj} = 1.51 (95%CI: 1.09-2.09) และสมการ ZINB for Rate สามารถประมาณค่าตัวแปรปฏิสัมพันธ์ (interaction term) ระหว่างความถี่ในการกิน*การแปรงฟัน ได้แม่นยำกว่า โดยเมื่อเด็กอายุ 12 ปี รับประทานขนมบ่อยครั้งและแปรงฟันบ่อยครั้งโอกาสในการเกิดโรคฟันผุเกิดขึ้นได้, IRR_{Adj} = 1.35 (95%CI: 0.89-2.07) แต่จะต่ำกว่าเมื่อเทียบกับการประมาณค่าในสมการ ZINB for Count, IRR_{Adj} = 1.51 (95%CI: 0.99-2.30) เช่นเดียวกับความถี่ในการแปรงฟันซึ่งในโมเดลนี้ตัวแปรดังกล่าวถูกกำหนดให้เป็นตัวแปรที่ทำให้การเกิดโรคฟันผุลดลง (inflated variable) กล่าวคือ เมื่อเด็กแปรงฟันเพิ่มขึ้นจาก 2 ครั้งต่อวัน เป็น 3 ครั้งต่อวัน โอกาสในการเกิดโรคฟันผุจะลดลงแต่การศึกษานี้ไม่พบการมีนัยสำคัญทางสถิติ รายละเอียดนำเสนอในตารางที่ 2 และตารางที่ 3 อย่างไรก็ตามเมื่อพิจารณาความเหมาะสมของโมเดลในภาพรวม พบว่า สมการ ZINB for Rate สามารถประมาณค่า Akaike's information criterion (AIC) และ Bayesian information criterion (BIC) ได้ดีกว่า สมการ ZINB for Count โดยพบ AIC เท่ากับ 1,431.56 และ BIC เท่ากับ 1,498.16 สำหรับสมการ ZINB for Rate และ AIC เท่ากับ 1,446.92 และ BIC เท่ากับ 1,513.53 สำหรับสมการ ZINB for Count

ตารางที่ 1 ปริมาณฟันผุ ถอน อุด (DMFT) ของเด็กอายุ 12 ปี ที่ศึกษาในโรงเรียนประถมศึกษา อำเภอคูเมือง จังหวัดบุรีรัมย์

ความรุนแรงของโรคฟันผุ (n=299)	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน	ค่ามัธยฐาน	ค่าต่ำสุด	ค่าสูงสุด
จำนวนซี่ฟันทั้งหมดในช่องปาก	26.54	2.48	28	17	28
ฟันผุ (Decayed)	2.74	2.67	2	0	14
ฟันที่ถูกถอน (Missing)	0.16	0.52	0	0	4
ฟันที่ได้รับการอุด (Filled)	1.18	1.52	1	0	8
Decayed, Missing, Filled teeth	4.08	3.06	4	0	17

ตารางที่ 2 การประมาณค่าพารามิเตอร์ของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate)

ตัวแปรอิสระ	ZINB for Count			ZINB for Rate		
	IRRAdj	95%CI	p-value	IRRAdj	95%CI	p-value
Count Model (DMFT >0)						
รายจ่าย (80 บาท/สัปดาห์)	1.60	1.21-2.12	0.001	1.56	1.19-2.06	0.001
ความถี่การกิน (10 ครั้ง/สัปดาห์)	1.17	0.90-1.52	0.24	1.15	0.89-1.48	0.27
การแปรงฟัน (แปรงทุกวัน)	1.51	1.09-2.09	0.01	1.37	0.98-1.90	0.06
รายจ่าย*ความถี่ในการกิน	0.53	0.36-0.77	0.001	0.55	0.37-0.79	0.002
ความถี่ในการกิน*การแปรงฟัน	1.51	0.99-2.30	0.06	1.35	0.89-2.07	0.15
อนามัยช่องปาก (OHIS >2)	1.12	0.95-1.33	0.16	1.13	0.96-1.34	0.13
กลุ่มชาติพันธุ์						
- ไทยโคราช	1.30	0.93-1.83	0.13	1.28	0.92-1.78	0.14
- ไทยอีสาน	1.34	1.04-1.75	0.03	1.34	1.04-1.72	0.02
ผู้เลี้ยงดูเรียนจบต่ำกว่ามัธยมศึกษา	1.31	1.05-1.62	0.01	1.29	1.05-1.59	0.01
อาชีพของผู้เลี้ยงดู						
- รับจ้าง/เกษตรกรรม	0.95	0.74-1.22	0.70	0.97	0.76-1.23	0.81
- รับราชการ/ค้าขาย	1.31	0.89-1.93	0.17	1.36	0.93-1.97	0.11
สมาชิกในครอบครัวสูบบุหรี่ (คน)						
- 1	0.98	0.83-1.16	0.85	0.98	0.83-1.15	0.80
- 2-3	1.28	1.01-1.63	0.04	1.27	1.01-1.61	0.03
Logit Model (DMFT=0)						
ความถี่ในการแปรงฟัน (ครั้ง)						
- 2	1.43	-1.65-4.53	0.36	1.17	-1.47-3.82	0.38
- 3	0.96	-2.31-4.23	0.56	0.85	-1.97-3.67	0.55

ตารางที่ 3 ค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate)

ตัวแปรอิสระ	Standard Error	
	ZINB for Count	ZINB for Rate
Count Model (DMFT >0)		
รายจ่าย (80 บาท/สัปดาห์)	0.23039	0.21924
ความถี่การกิน (10 ครั้ง/สัปดาห์)	0.15636	0.14989
การแปรงฟัน (แปรงทุกวัน)	0.24967	0.22992
รายจ่าย*ความถี่ในการกิน	0.10365	0.10471
ความถี่ในการกิน*การแปรงฟัน	0.32563	0.29256
อนามัยช่องปาก (OHIS >2)	0.09667	0.09483

ตารางที่ 3 ค่าความคลาดเคลื่อนมาตรฐานของพารามิเตอร์ในสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate) (ต่อ)

ตัวแปรอิสระ	Standard Error	
	ZINB for Count	ZINB for Rate
กลุ่มชาติพันธุ์		
- ไทยโคราช	0.22455	0.21640
- ไทยอีสาน	0.17959	0.17358
ผู้เลี้ยงดูเรียนจบต่ำกว่ามัธยมศึกษา	0.14247	0.13735
อาชีพของผู้เลี้ยงดู		
- รับจ้าง/เกษตรกรรม	0.11944	0.11873
- รับราชการ/ค้าขาย	0.25761	0.25896
สมาชิกในครอบครัวสูบบุหรี่ (คน)		
- 1	0.08516	0.08298
- 2-3	0.15579	0.15034
Logit Model (DMFT=0)		
ความถี่ในการแปรงฟัน (ครั้ง)		
- 2	1.57808	1.35193
- 3	1.66925	1.44154

วิจารณ์

งานวิจัยทางด้านทันตสาธารณสุข เช่น การศึกษาของ Solinas และคณะ [11] ประยุกต์ใช้สถิติ Poisson, Negative Binomial (NB), Zero Inflated Poisson (ZIP) และ Zero Inflated Negative Binomial (ZINB) ถ้าหาก Solinas และคณะ [11] วิเคราะห์โมเดลแบบ Count อาจมีความเหมาะสมเพราะเด็กอายุ 4 ปี มีฟันน้ำนมขึ้นครบ 20 ซี่ และเป็นกลุ่มอายุที่เด็กยังไม่มีฟันแท้ขึ้น ดังนั้น จำนวนซี่ฟันน้ำนม หรือจำนวนด้านของฟันน้ำนม รวมถึง Offset ของเด็กอายุ 4 ปีจึงเท่ากัน แต่แตกต่างจากการศึกษาปัจจัยเสี่ยงต่อการเกิดโรคฟันผุ (DMFT) ของเด็กอายุ 12 ปี ซึ่งเมื่อพิจารณาอายุของฟันที่ขึ้นในช่องปากเด็กแต่ละคน พบว่า ฟันแท้แต่ละซี่ที่ขึ้นมาในช่องปากมีอายุการขึ้นต่างกัน ซึ่งมีผลต่อจำนวนซี่ฟันในช่องปาก รวมถึงระยะเวลาในการสัมผัสปัจจัยเสี่ยงของฟันแท้แต่ละซี่ในเด็กแต่ละคนย่อมแตกต่างกัน โดยเฉพาะการศึกษานี้พบค่าเฉลี่ยจำนวนฟันทั้งหมดในช่องปากของเด็กแต่ละคนเท่ากับ 26.54 ซี่ต่อคน และพบค่าต่ำสุด เท่ากับ 17 ซี่ต่อคน รวมถึงพบค่าสูงสุด เท่ากับ 28 ซี่ต่อคน ดังนั้น ในการศึกษาปัจจัยเสี่ยงต่อการเกิดโรคฟันผุ (DMFT) ของเด็กอายุ 12 ปี หากผู้วิจัยกำหนดให้จำนวนซี่ฟันที่มีทั้งหมดในช่องปากของเด็กแต่ละคนเป็นพื้นที่ที่ศึกษา (Index of Size) หรือ Offset ในโมเดลแบบอัตรา (Rate) จะทำให้การวิเคราะห์ข้อมูลแบบอัตรา (Rate) มีความแม่นยำกว่าการวิเคราะห์ข้อมูลแบบจำนวนนับ (Count) รวมถึง เมื่อพิจารณาความเหมาะสมของโมเดลในภาพรวม พบว่า สมการ ZINB for Rate สามารถประมาณค่า AIC และ BIC ได้ดีกว่า สมการ ZINB for Count

จุดเด่นของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) คือ สมการมีรูปแบบโมเดล

เป็นแบบผสม (Mixed Effect Model) ซึ่งในสมการมีการประมาณค่าพารามิเตอร์แบบคงที่ (Fixed effect) และแบบสุ่ม (Random effect) ทั้งนี้การประมาณค่าพารามิเตอร์แบบสุ่ม (Random effect) คำนึงถึงความแตกต่างระหว่างกลุ่ม (Heterogeneity between group) และความแตกต่างระหว่างคน หรือความแตกต่างของตัวอย่างที่อยู่ภายในกลุ่มเดียวกัน (heterogeneity of subjects within group) จึงช่วยทำให้ความคลาดเคลื่อนในการประมาณค่าลดลง[7] ข้อจำกัดของการศึกษานี้ เนื่องการศึกษานี้มีวัตถุประสงค์เพื่อประเมินความแม่นยำในการประมาณค่าของสมการถดถอยพหุทวินามแบบลบกรณีตัวแปรตามมีค่าศูนย์ (ZINB) ระหว่างสมการแบบจำนวนนับ (Count) กับสมการแบบอัตรา (Rate) และเปรียบเทียบค่าความคลาดเคลื่อนมาตรฐานในการประมาณค่าของวิธีการทางสถิติดังกล่าว ซึ่งข้อมูลและตัวแปรได้จากการศึกษาเรื่องความสัมพันธ์ระหว่างเงินค่าขนมกับปริมาณฟันผุ ถอน อุด (DMFT) ในเด็กอายุ 12 ปี ที่ศึกษาในโรงเรียนประถมศึกษา อำเภอคูเมือง จังหวัดบุรีรัมย์ ซึ่งผู้วิจัยได้คัดเลือกตัวแปรและ Fitted โมเดลเพื่อหาความเหมาะสมของโมเดลเรียบร้อยแล้ว ดังนั้นตัวแปรต่าง ๆ จึงอยู่ในบริบทของอำเภอคูเมือง จังหวัดบุรีรัมย์ ซึ่งโมเดลมีความเหมาะสมสำหรับความตรงภายนอก (goodness of fitted via external validity) ร้อยละ 73

กิตติกรรมประกาศ

ขอขอบคุณสำนักงานกองทุนสนับสนุนการสร้างเสริมสุขภาพ (สสส.) สำนักงานพัฒนานโยบายสุขภาพระหว่างประเทศ (IHPP) แผนยุทธศาสตร์การพัฒนาดัชนีประเมินภาระโรคและสุขภาพของประชากรไทย (BOD) ที่สนับสนุนทุนในการทำวิจัย และขอขอบคุณผู้ที่เกี่ยวข้องทุกท่าน ที่ให้ความร่วมมือและสนับสนุนการดำเนินงานวิจัยนี้จนสำเร็จลุล่วงด้วยดี

เอกสารอ้างอิง

1. Agresti A. An introduction to categorical data analysis. New York: Wiley; 2007.
2. Agresti A. An introduction to categorical data analysis. New York: Wiley; 1996.
3. Hilbe JM. Negative binomial regression. Cambridge: Cambridge University press; 2011.
4. ไพบูลย์ โล่ห์สุนทร. ระบาดวิทยา. กรุงเทพมหานคร: จุฬาลงกรณ์มหาวิทยาลัย; 2547.
5. กรมควบคุมโรค. อัตราอุบัติการณ์ [อินเทอร์เน็ต]. [สืบค้นเมื่อ 12 ก.ค. 2556]. แหล่งข้อมูล: http://lms.kmddc.go.th/e_learning/e-data/32/S1_003/3_26-6.swf
6. Seow WK. Biological mechanisms of early childhood caries. Community Dent Oral Epidemiol 1998; 26(Suppl. 1): 8-27.
7. Long JS., Freese J. Regression models for categorical dependent variables using STATA. USA: Stata Press; 2001.
8. Böhning D, Dietz E, Schlattmann P, Mendonça L, Kirchner U. The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. Journal of the Royal Statistical Society Series A: Statistics in Society 1999;162(2):195-209.
9. Lewsey JD, Thomson WM. The utility of the zero-inflated Poisson and zero-inflated negative binomial models: A case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. Community Dentistry and Oral Epidemiology 2004;32(3):183-9.
10. Tramini P, Molinari N, Tentscher M, Demattei C, Schulte AG. Association between caries experience and body mass index in 12-year-old French children. Caries Research 2009;43(6):468-73.
11. Solinas G, Campus G, Maida C, Sotgiu G, Cagetti MG, Lesaffre E, et al. What statistical method should be used

to evaluate risk factors associated with dmfs index? Evidence from the National Pathfinder Survey of 4-year-old Italian children. Community Dentistry and Oral Epidemiology 2009;37(6):539-46.

12. Signorini DF. Sample size for Poisson regression. Biometrika 1991;78(2):446-50.

The Precision of Estimation between ZINB for Count and ZINB for Rate in Dental Health Research

Nittaya Phuangrach, Ph.D. (Epidemiology and Biostatistics)

Khumuang Hospital, Khumuang District, Buriram Province, Thailand

Journal of Health Science of Thailand 2024;33(Suppl 1):86-S93.

Corresponding author: Nittaya Phuangrach, Email: nidarach@hotmail.com

Abstract: Poisson regression is divided into two types which are Poisson for count and Poisson for rate; when considering dental caries, susceptible tooth and host, substrate, and bacteria including exposure time are the risk factors. Researcher should consider the impact of the exposure time before conclusion. However, if the researcher does not have this information, age or number that can identify the association with exposure time can be put instead of this variable. This study is a part of the study on "the Association Between Daily Pocket Money and DMFT among children aged 12 years in a primary school at Khumuang District, Buriram Province" which data had over-dispersion and 11.37% of zero outcomes, the researcher considered the total number of teeth in the mouth as an offset and using the zero-inflated negative binomial regression (ZINB) which is a part of Poison statistics for the research question. As a result, the previous study researchers used ZINB to analyze the data but they did not identify that the analysis was by count or rate then the objective of this study was to assess the precision of estimation between ZINB for count and ZINB for rate and compared the standard error between the two methods in dental health research. The results found that the parameters in the ZINB for rate model had more precisions than the parameters in the ZINB for count model. Moreover, ZINB for rate could estimate standard errors less than the ZINB for count. The prominent feature of the zero-inflated negative binomial regression (ZINB) is that it is a "mixed effect model", which estimates the parameters in both fixed effect and random effect and thus reduces the standard error of estimation.

Keywords: Poisson regression; ZINB; estimation; standard error