

บทที่ 2

กรอบแนวคิดทางทฤษฎี แบบจำลองทางเศรษฐมิติ และงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่ใช้เป็นแนวทางในการศึกษา

2.1.1 Random Utility Model

ตัวแปรของข้อมูลทางเลือกเชิงคุณภาพกำหนดโดย random utility model จากสมมติฐานของ Nakosteen-Zimmer ถ้าให้ U^a และ U^b แทนตัวแปรของความพอใจ (utility) ที่มี 2 ทางเลือก เช่น สมมติให้ U^a เป็นความพอใจในการลงทะเบียนเข้าศึกษา และ U^b เป็นความพอใจที่จะไม่ลงทะเบียนเข้าศึกษา ตัวแปรซึ่งเป็นตัวแปรตามนี้จะมีค่าระหว่าง 2 ค่า คือ 1 และ 0 ซึ่งถ้าตัวแปรตามมีค่าเท่ากับ 1 ก็จะแสดงให้เห็นว่าความพอใจในการลงทะเบียนเข้าศึกษามีมากกว่าความพอใจที่จะไม่ลงทะเบียนเข้าศึกษา และถ้าตัวแปรตามมีค่าเท่ากับ 0 ก็จะแสดงให้เห็นว่าความพอใจในการไม่ลงทะเบียนเข้าศึกษามีมากกว่าความพอใจที่จะลงทะเบียนเข้าศึกษา

ดังนั้น ค่าตัวแปรตามที่เท่ากับ 1 แสดงว่า $U^a > U^b$ และมีค่าตัวแปรตามเป็น 0 ถ้า $U^a < U^b$ โดยมีรูปแบบสมการดังนี้

$$U^a = \beta'_a x + e_a \text{ และ } U^b = \beta'_b x + e_b$$

สมมติให้ $Y=1$ แสดงความพอใจในการลงทะเบียนเข้าศึกษา จะได้

$$\begin{aligned} \text{Prob}[Y=1|x] &= \text{Prob}[U^a > U^b] \\ &= \text{Prob}[\beta'_a x + e_a - \beta'_b x + e_b > 0 | x] \\ &= \text{Prob}[(\beta'_a - \beta'_b)'x + e_a - e_b > 0 | x] \\ &= \text{Prob}[\beta'x + e > 0 | x] \end{aligned}$$

2.1.2 Discrete Choice Model

เนื่องจากการศึกษาครั้งนี้ เกี่ยวข้องกับปัจจัยที่มีอิทธิพลต่อการลงทะเบียนเข้าศึกษา นักศึกษามีการตัดสินใจในทางเลือกเชิงคุณภาพ 2 ทางเลือก (binary choice) คือการลงทะเบียน หรือ การไม่ลงทะเบียนเข้าศึกษา ตัวแปรตามในการศึกษาครั้งนี้จึงมีลักษณะเป็นตัวแปรหุ่นที่มีค่าเป็น 1 หรือ 0 เมื่อตัวแปรตามมีลักษณะดังกล่าวจึงไม่สามารถใช้วิธีการวิเคราะห์ความถดถอยแบบทั่วไปในการวิเคราะห์อิทธิพลของตัวแปรต้นที่ส่งผลต่อตัวแปรตามได้ โดย กัลยา วานิชย์บัญชา (2548)

ได้กล่าวถึงเหตุผลที่ไม่สามารถใช้การวิเคราะห์ความถดถอยทั่วไปในการวิเคราะห์เมื่อสถานการณ์ของตัวแปรตามมีลักษณะดังกล่าวไว้ดังนี้

ประการแรกตัวแปรตามที่ศึกษาจะมีได้เพียง 2 ค่า ทำให้ค่าประมาณของ Y ซึ่งเป็นโอกาสที่เหตุการณ์ที่สนใจจะเกิด ซึ่งมีค่าระหว่าง 0 ถึง 1 ถ้าใช้สมการความถดถอยเชิงเส้นปกติ คือ $\hat{Y} = a + \beta x$ ซึ่งจะทำให้ค่า \hat{Y} ที่ได้อาจจะไม่ได้อยู่ในระหว่าง 0 ถึง 1 หรืออาจมีค่าน้อยกว่า 0 หรือมากกว่า 1 ดังนั้นถ้าใช้สมการความถดถอยเชิงเส้นปกติจะทำให้ค่า \hat{Y} ที่ได้ ออกนอกช่วง 0 ถึง 1 จึงไม่สามารถใช้สมการความถดถอยเชิงเส้นแบบปกติในการประมาณค่าตัวแปรตามได้

ประการที่สอง ในการวิเคราะห์ความถดถอยทั่วไปมีเงื่อนไขว่าค่าคลาดเคลื่อนต้องมีการแจกแจงแบบปกติ แต่เมื่อ Y มีได้เพียง 2 ค่า คือ 0 กับ 1 ค่าคลาดเคลื่อนมีได้เพียง 2 ค่าด้วย ซึ่งเป็นไปไม่ได้ที่ค่าคลาดเคลื่อนจะมีการแจกแจงแบบปกติ จึงไม่สามารถใช้สมการความถดถอยเชิงเส้นแบบปกติในการประมาณค่าตัวแปรตามได้

ประการที่สาม จากเงื่อนไขของการวิเคราะห์ความถดถอย คือค่าแปรปรวนของค่าคลาดเคลื่อน หรือ $V(e)$ ต้องคงที่ทุกค่าของ X แต่เมื่อ Y มีค่าได้เพียง 2 ค่าทำให้เงื่อนไขว่า $V(e)$ คงที่ไม่เป็นจริงจึงไม่สามารถใช้สมการความถดถอยเชิงเส้นแบบปกติได้

ทรงศักดิ์ ศรีบุญจิตต์ (2548) ได้กล่าวเพิ่มเติมว่าสาเหตุที่สำคัญที่ทำให้ไม่สามารถใช้สมการความถดถอยเชิงเส้นแบบปกติในการประมาณค่าตัวแปรตามเมื่อตัวแปรตามที่มีค่าได้เพียง 2 ค่า คือสมการความถดถอยเชิงเส้นแบบปกติจะประมาณค่าความชันด้วยวิธีกำลังสองน้อยที่สุดสามัญ (ordinary least squares) ซึ่งจะทำให้ค่าความชันที่ประมาณได้ อาจสูงหรือต่ำกว่าความเป็นจริงเมื่อค่าสังเกต (x) มีค่าสุดโต่ง หรือ เกาะกลุ่มกันอยู่ตรงกลางจำนวนมากเกินไป

จากเหตุผลที่กล่าวมาข้างต้นจึงต้องใช้แบบจำลองที่เหมาะสมสำหรับการวิเคราะห์ โดย ชัยวุฒิ ชัยพันธ์ (2542) กล่าวว่าแบบจำลองที่เหมาะสมและเป็นที่ยอมรับกันทั่วไปได้แก่แบบจำลองโพรบิต (Probit Model) และแบบจำลองโลจิท (Logit Model) หรือการวิเคราะห์ความถดถอยแบบโพรบิต และการวิเคราะห์ความถดถอยแบบโลจิสติก โดยที่ความแตกต่างของการวิเคราะห์ทั้งสองอยู่ที่การแจกแจงของตัวแปรตาม หากตัวแปรตามมีการแจกแจงแบบปกติ (normal distribution) ก็จะใช้การวิเคราะห์ความถดถอยแบบโพรบิตหรือแบบจำลองโพรบิต แต่หากตัวแปรตามมีการแจกแจงแบบโลจิสติกก็จะใช้การวิเคราะห์ความถดถอยแบบโลจิสติก หรือแบบจำลองโลจิท

การวิจัยครั้งนี้ได้ตั้งสมมติฐานว่าตัวแปรตามได้มีการแจกแจงแบบโลจิสติก ดังนั้นในที่นี้จะกล่าวถึงเฉพาะการวิเคราะห์ความถดถอยแบบโลจิสติก รายละเอียดของแบบจำลองโลจิทมีดังต่อไปนี้

2.2 แบบจำลองทางเศรษฐมิติ

2.2.1 Logit Model

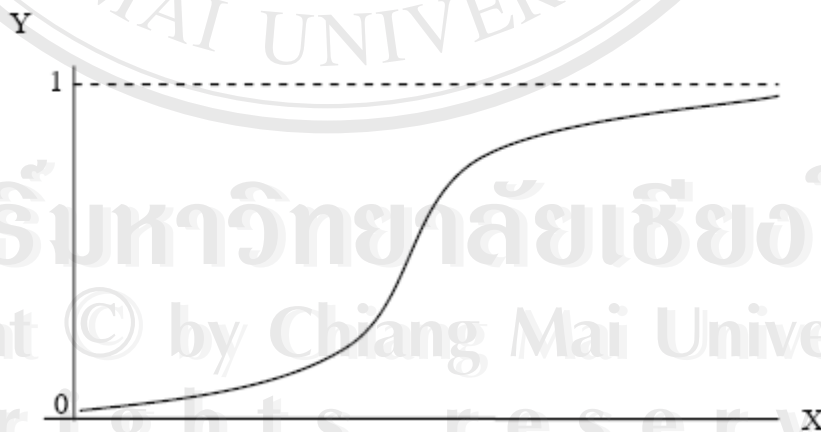
การวิเคราะห์ความถดถอยแบบโลจิสติก มีวัตถุประสงค์และแนวคิดเหมือนกับการวิเคราะห์ความถดถอยแบบปกติ คือเพื่อศึกษาความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระ และนำสมการถดถอยที่ได้ไปประมาณหรือพยากรณ์ค่าตัวแปรตาม เมื่อกำหนดค่าตัวแปรอิสระ

กัลยา วานิชย์บัญชา (2548) กล่าวถึงขนาดตัวอย่างที่จะนำมาใช้ในการวิเคราะห์การถดถอยแบบโลจิสติกว่า จะต้องใช้ขนาดตัวอย่างในจำนวนที่มากกว่า การวิเคราะห์ความถดถอยเชิงเส้นแบบปกติทั่วไป โดยจะต้องใช้จำนวนตัวอย่างเท่ากับ $n \geq 30p$ โดยที่ p เป็นจำนวนตัวแปรอิสระ และการวิเคราะห์การถดถอยแบบโลจิสติกมีเงื่อนไขที่สำคัญอยู่ 5 ข้อดังนี้

1. ตัวแปรอิสระ X อาจเป็นข้อมูลชนิด dichotomous (มีได้ 2 ค่า) หรือเป็นสเกลอันดับ (interval scale) และสเกลอัตราส่วน (ratio scale) ก็ได้
2. ค่าคาดหวังของค่าคลาดเคลื่อนต้องเป็นศูนย์ หรือ $E(e) = 0$
3. ค่าคลาดเคลื่อน e_i และ ค่าคลาดเคลื่อน e_j เป็นอิสระกัน
4. ค่าคลาดเคลื่อน e_i และตัวแปรอิสระ X_i เป็นอิสระกัน
5. ตัวแปรอิสระไม่ควรมีความสัมพันธ์กัน หรือไม่ควรเกิดปัญหา multicollinearity

กราฟของโลจิสติกแสดงได้ดังภาพที่ 1

ภาพที่ 1 กราฟของโลจิสติก



โดย ทรงศักดิ์ ศรีบุญจิตต์ (2548) ได้อธิบายถึงการแจกแจงแบบโลจิสติก (logistic distribution) ดังนี้

จากการแจกแจงแบบโลจิสติก (logistic distribution)

$$\text{Prob}(y=1) = \frac{e^{\beta'X}}{1+e^{\beta'X}} = \Lambda(\beta'X) \quad (1)$$

โดยที่ $\Lambda(\cdot)$ คือ ฟังก์ชันการแจกแจงสะสมแบบโลจิสติก (logistic cumulative distribution function)

จากแบบจำลองความน่าจะเป็น (probability model)

$$E[Y|X] = 0[1 - F(\beta'X)] + 1[F(\beta'X)] \quad (2)$$

เราจะได้ว่า

$$\begin{aligned} \frac{\partial E[Y|X]}{\partial X} &= \left\{ \frac{dF(\beta'X)}{d(\beta'X)} \right\} \beta \\ &= f(\beta'X) \beta \end{aligned} \quad (3)$$

โดยที่ $f(\cdot)$ คือ ฟังก์ชันความหนาแน่น (density function) ซึ่งคล้อยกับฟังก์ชันการแจกแจงสะสม (cumulative distribution) $F(\cdot)$ สำหรับการแจกแจงปกติ (normal distribution) เราจะได้ว่า

$$\frac{\partial E[Y|X]}{\partial X} = \phi(\beta'X) \beta$$

โดยที่ $\phi(\cdot)$ คือ ฟังก์ชันความหนาแน่น ปกติมาตรฐาน (standard normal density function) สำหรับการแจกแจงแบบโลจิสติก (logistic distribution)

$$\begin{aligned} \frac{d\Lambda[\beta'X]}{d(\beta'X)} &= \frac{e^{\beta'X}}{(1+e^{\beta'X})^2} \\ &= \Lambda(\beta'X)[1 - \Lambda(\beta'X)] \end{aligned} \quad (4)$$

เพราะฉะนั้นในแบบจำลองโลจิท (logit model) จะได้ว่า

$$\frac{\partial E[Y|X]}{\partial X} = \Lambda(\beta' X)[1 - \Lambda(\beta' X)]\beta \quad (5)$$

สำหรับตัวประมาณค่า นั้น ในกรณีของแบบจำลองโลจิท (ซึ่งแตกต่างจากกรณีของแบบจำลองพหุคูณ)

$$\beta = \sum_i (Y_i - \Lambda_i)^2 X_i X_i' \quad (6)$$

ซึ่งเป็นการคำนวณเมทริกซ์ความแปรปรวนร่วมเกี่ยวเชิงเส้นกำกับ (asymptotic covariance matrix) วิธีหนึ่ง

จาก

$$\hat{f} = \hat{\Lambda}(1 - \hat{\Lambda})$$

จะได้

$$\frac{d\hat{f}}{dz} = (1 - 2\hat{\Lambda}) \left(\frac{d\hat{\Lambda}}{dz} \right) = (1 - 2\hat{\Lambda}) \hat{\Lambda}(1 - \hat{\Lambda})$$

เมื่อจัดพจน์ (terms) ต่างๆ เข้าด้วยกันจะได้

$$\text{Asy. var} [\hat{\gamma}] = [\Lambda(1 - \Lambda)]^2 [\mathbf{I} + (1 - 2\Lambda)\beta X'] \nu [\mathbf{I} + (1 - 2\Lambda)X\beta']$$

(Greene, 2540 : 885 อ้างจาก ทรงศักดิ์ ศรีบุญจิตต์, 2548)

ดังนั้นในแบบจำลองนี้จึงใช้รูปสมการ

$$\text{Prob}(Y = 1) = \frac{1}{1 + e^{-\beta X}}$$

หรือสามารถเขียนให้อยู่ในรูป logit model ได้ดังนี้

$$\log \left[\frac{\text{Prob}(Y = 1)}{1 - \text{Prob}(Y = 1)} \right] = \alpha + \beta X_i$$

2.3 งานวิจัยที่เกี่ยวข้อง

2.3.1 สาเหตุของการสละสิทธิ์

2.3.1.1 อันดับการเลือกของคณะวิชาที่สอบ

อันดับการเลือกต่างๆ ของคณะวิชาที่สอบได้ จะมีผลต่อการสละสิทธิ์มากขึ้น โดยผู้สละสิทธิ์ส่วนใหญ่สอบได้คณะในอันดับการเลือกต่างๆ ซึ่งจากการศึกษาของ ประยูร ศรีประสาธน์ (2526) ที่ศึกษาองค์ประกอบที่คาดว่าจะมีความสัมพันธ์กับการสละสิทธิ์การศึกษาในระดับอุดมศึกษาของรัฐ ปีการศึกษา 2522-2524 พบว่า ผู้สละสิทธิ์ส่วนใหญ่สอบได้ในคณะอันดับการเลือกต่างๆ โดยอันดับการเลือก 3 อันดับแรก จะอยู่ในกลุ่มสาขาวิชาเดียวกันค่อนข้างมาก ส่วนอันดับการเลือก 3 อันดับสุดท้าย จะมีลักษณะการเลือกคณะวิชาแบบกระจัดกระจาย และจากการศึกษาของ นิสากร สุทธิแป้น (2547) ศึกษาสาเหตุการสละสิทธิ์ของผู้มีสิทธิ์เข้าศึกษาในมหาวิทยาลัยเชียงใหม่ระดับปริญญาตรีปีการศึกษา 2547 พบว่าตัวแปรที่มีผลต่อการสละสิทธิ์ ได้แก่ คณะหรือสาขาวิชาที่สอบได้ ไม่นัดหรือไม่เหมาะสมกับตนเองมีจำนวนมากที่สุด รองลงมา คือ คณะหรือสาขาวิชาที่สอบได้ เมื่อเรียนจบแล้วหางานทำยาก

2.3.1.2 การขาดแคลนทุนทรัพย์

การขาดแคลนทุนทรัพย์ในการศึกษาเป็นปัจจัยหนึ่งที่มีผลต่อการสละสิทธิ์เข้าศึกษาของนักเรียน เนื่องจากไม่มีทุนทรัพย์ จากการศึกษาของ บัญชร แก้วส่อง (2525) ศึกษาสาเหตุการสละสิทธิ์ของนักศึกษามหาวิทยาลัยขอนแก่น ปีการศึกษา 2522-2525 พบว่า สาเหตุที่สำคัญที่สุดที่ทำให้นักศึกษาสละสิทธิ์ คือ สาเหตุจากการขาดแคลนทุนทรัพย์ ผู้ปกครองไม่สนับสนุนให้เรียน ต้องการให้นักศึกษาประกอบอาชีพช่วยบิดามารดา หรือบางส่วนก็ต้องศึกษาในสถาบันอาชีพ เพื่อช่วยผู้ปกครองประกอบอาชีพ

2.3.1.3 สภาพครอบครัว

สภาพครอบครัวมีผลต่อการสละสิทธิ์ของนักเรียน กล่าวคือบุคคลที่นักศึกษาเคารพนับถือ บิดา มารดา และคนรอบข้างจะมีผลต่อการตัดสินใจของนักเรียนในการตัดสินใจสละสิทธิ์ในสถาบันอุดมศึกษา จากการศึกษาของบัญญัติ แก้วส่อง (2525) ที่ศึกษาสาเหตุการสละสิทธิ์ของนักศึกษามหาวิทยาลัยขอนแก่น ปีการศึกษา 2522-2525 พบว่า การเลือกเรียนในสถาบันใกล้บ้านช่วยให้ค่าใช้จ่ายลดลงเป็นอันมาก หากมหาวิทยาลัยที่อยู่ไกลบ้าน ไม่มีผู้ปกครองดูแล นักเรียนจึงไม่กล้ามาศึกษาต่อจึงเลือกสถาบันใกล้บ้าน และจากการศึกษาของ นิสากร สุทธิแป้น (2547) ที่ศึกษาสาเหตุการสละสิทธิ์ของผู้มีสิทธิ์เข้าศึกษาในมหาวิทยาลัยเชียงใหม่ระดับปริญญาตรีปีการศึกษา 2547 พบว่า ตัวนักศึกษาเองมีผลต่อการตัดสินใจมากที่สุด รองลงมาคือ บิดามารดา หรือผู้ปกครอง บุคคลที่สนิทหรือให้ความเคารพนับถือ

2.3.2 อัตราการสละสิทธิ์

2.3.2.1 แยกตามอันดับการเลือก

อันดับการเลือกของคณะที่สอบได้อันดับท้าย ๆ จะมีอัตราการสละสิทธิ์มาก โดยจากการศึกษาของ สุจินต์ สารงาม (2531) ถึงเหตุผลการสละสิทธิ์ของผู้สอบผ่านการสอบโควต้า เพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปีการศึกษา 2530 พบว่า อันดับการเลือกอันดับ 1 คิดเป็นร้อยละ 17.25 อันดับ 2 คิดเป็นร้อยละ 30.35 และอันดับ 3 ร้อยละ 52.40

การศึกษาการสละสิทธิ์เข้าศึกษาต่อสถาบันอุดมศึกษาของรัฐสังกัดทบวงมหาวิทยาลัยปีการศึกษา 2525-2532 ของฝ่ายวิจัยสถาบันและสารสนเทศ กองแผนงาน สำนักงานปลัดทบวงมหาวิทยาลัย (2532) พบว่าอัตราร้อยละของการสละสิทธิ์ในอันดับการเลือกอันดับ 1 ถึง 3 ไม่เกินร้อยละ 20 และผู้ที่ผ่านข้อเขียนได้ในอันดับ 6 จะสละสิทธิ์สูงถึงร้อยละ 32-39

การศึกษาการสละสิทธิ์ของผู้สอบผ่านข้อเขียนการสอบโควต้าเพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปี 2543 ของนิสากร สุทธิแป้น (2544) พบว่า การสละสิทธิ์ในคณะที่มีอันดับการเลือกอันดับ 3 มากที่สุด รองลงมาได้แก่อันดับ 2 คิดเป็นร้อยละ 53.35 และ 30.64 ตามลำดับ

2.3.2.2 แยกตามคณะที่สอบได้

คณะที่มีอัตราการสละสิทธิ์มากที่สุดได้แก่ คณะกลุ่มสาขาวิทยาศาสตร์และเทคโนโลยี เป็นส่วนมาก ได้แก่คณะวิทยาศาสตร์ เกษตรศาสตร์ และวิศวกรรมศาสตร์ เป็นต้น โดยจากการศึกษาของ สุจินต์ สารงาม (2531) ถึงเหตุผลการสละสิทธิ์ของผู้สอบผ่านการสอบโควต้า เพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปีการศึกษา 2530 พบว่า การสละสิทธิ์ของนักเรียนที่สอบได้ในคณะต่าง ๆ เมื่อเทียบกับจำนวนผู้สอบผ่านข้อเขียน พบว่าคณะที่มีอัตราการสละสิทธิ์สูงสุด 3 อันดับแรก คือ คณะวิทยาศาสตร์ คณะเกษตรศาสตร์ และคณะศึกษาศาสตร์

การศึกษาการสละสิทธิ์เข้าศึกษาต่อสถาบันอุดมศึกษาของรัฐสังกัดทบวงมหาวิทยาลัย ปีการศึกษา 2525-2532 ของฝ่ายวิจัยสถาบันและสารสนเทศ กองแผนงาน สำนักงานปลัดทบวงมหาวิทยาลัย (2532) พบว่า อัตราการสละสิทธิ์มากที่สุด 3 อันดับแรก คือ คณะวิทยาศาสตร์ ร้อยละ 30 คณะพยาบาลศาสตร์ ร้อยละ 27 คณะศึกษาศาสตร์ ร้อยละ 25 ส่วนคณะที่อัตราการสละสิทธิ์น้อยที่สุด 3 อันดับแรกได้แก่ คณะแพทยศาสตร์ ร้อยละ 3 คณะจิตรศิลป์ร้อยละ 5 คณะทันตแพทยศาสตร์ ร้อยละ 8

การศึกษาการสละสิทธิ์ของผู้สอบผ่านข้อเขียนการสอบโควต้าเพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปี 2543 ของนิสากร สุทธิแป้น (2544) พบว่า คณะที่มีผู้สละสิทธิ์มากที่สุดได้แก่คณะวิทยาศาสตร์ รองลงมาได้แก่คณะวิศวกรรมศาสตร์ ส่วนคณะที่มีผู้สละสิทธิ์น้อยที่สุดได้แก่คณะจิตรศิลป์ รองลงมาได้แก่ คณะแพทยศาสตร์

2.3.2.3 แยกตามเขตพื้นที่จังหวัด และอำเภอ

เขตพื้นที่จังหวัดและอำเภอ มีผลต่ออัตราการสละสิทธิ์เข้าศึกษา กล่าวคือ หากเขตพื้นที่ภูมิลำเนาอยู่ใกล้กับสถาบันการศึกษาจะมีอัตราการสละสิทธิ์มาก

การศึกษาของ สุจินต์ สารงาม (2531) ถึงเหตุผลการสละสิทธิ์ของผู้สอบผ่านการสอบโควต้า เพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปีการศึกษา 2530 พบว่า จังหวัดที่มีอัตราการสละสิทธิ์สูงสุด 5 อันดับแรกได้แก่ จังหวัดแม่ฮ่องสอน กำแพงเพชร ลำพูน พะเยา เพชรบูรณ์ และพิจิตร

การศึกษาการสละสิทธิ์ของผู้สอบผ่านข้อเขียนการสอบโควต้าเพื่อเข้าศึกษาในมหาวิทยาลัยเชียงใหม่ ปี 2543 ของนิสสาร สุทธิแป้น (2544) พบว่า ผู้สละสิทธิ์ทั้งภาคเหนือตอนล่างและตอนบนมีผู้สละสิทธิ์อยู่ในอำเภอเมืองมากกว่าอำเภอรอบนอก และผู้สละสิทธิ์ในอำเภอเมืองของจังหวัดภาคเหนือตอนล่างมีจำนวนสูงกว่าจังหวัดในภาคเหนือตอนบน