

9-1-1983

## การพิจารณาขนาดตัวอย่างและเทคนิคการเลือกตัวอย่างในการทำวิจัยทางการแพทย์ (ตอนที่ 1 )

เต็มศรี ชำนิจารกิจ

Follow this and additional works at: <https://digital.car.chula.ac.th/clmjjournal>



Part of the [Medicine and Health Sciences Commons](#)

---

### Recommended Citation

ชำนิจารกิจ, เต็มศรี (1983) "การพิจารณาขนาดตัวอย่างและเทคนิคการเลือกตัวอย่างในการทำวิจัยทางการแพทย์ (ตอนที่ 1 )," *Chulalongkorn Medical Journal*: Vol. 27: Iss. 5, Article 1.

Available at: <https://digital.car.chula.ac.th/clmjjournal/vol27/iss5/1>

This Editorial is brought to you for free and open access by the Chulalongkorn Journal Online (CUJO) at Chula Digital Collections. It has been accepted for inclusion in Chulalongkorn Medical Journal by an authorized editor of Chula Digital Collections. For more information, please contact [ChulaDC@car.chula.ac.th](mailto:ChulaDC@car.chula.ac.th).

บทบรรณาธิการ

# การพิจารณาขนาดตัวอย่างและเทคนิคการเลือกตัวอย่าง ในการทำวิจัยทางการแพทย์ ตอนที่ 1

เต็มศรี ชำนิจารกิจ\*

บทนำ

ปัจจุบันนี้ยังมีสิ่งที่ต้องการค้นหาคำตอบอีกมากมาย โดยที่การค้นหาคำตอบเหล่านั้น จำเป็นจะต้องใช้วิธีการและเทคนิคต่าง ๆ ที่ถูกต้องตามหลักวิทยาศาสตร์ ซึ่งก็คือการทำวิจัยนั่นเอง จะเห็นได้ว่าก่อนจะเริ่มทำวิจัยก็จะต้องมีคำถามตัวเองเสียก่อนว่า ทำไมจะต้องทำวิจัย จะทำอะไร จะทำกับใคร จะทำอย่างไร ทำแล้วจะได้คำตอบอะไร รู้คำตอบแล้วจะทำอย่างไรต่อไป ฯลฯ คำตอบของคำถามเหล่านี้ก็คือขั้นตอนต่าง ๆ ของการทำวิจัยนั่นเอง ปัญหาหนึ่งที่ผู้วิจัยจะประสบหลังจากเลือกประชากรที่ศึกษาได้แล้ว ก็คือการพิจารณาขนาดตัวอย่าง และเทคนิคการเลือกตัวอย่างที่ศึกษา ทั้งนี้เนื่องจากว่าผู้ทำวิจัยไม่สามารถทำการศึกษาจากประชากรทั้งหมดได้ และก็จำเป็นต้องทำเช่นนั้นด้วย ดังนั้นจึงต้องมีวิธีการที่จะเลือกตัวอย่างให้มีคุณสมบัติเป็นผู้แทนของประชากรนั้นมาศึกษา และเพื่อเป็นการลดค่าใช้จ่าย ลดเวลาและลดจำนวนบุคลากรที่จะทำวิจัยโดยการ คำนวณหาขนาดตัวอย่างที่จะศึกษาอย่างมีหลักเกณฑ์ ซึ่งสามารถนำผลมาวิเคราะห์และแปลผลกลับไปเป็นของประชากรนั้น ๆ ได้

\* ภาควิชาเวชศาสตร์ป้องกันและสังคม คณะแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

## I. การพิจารณาขนาดตัวอย่าง (Sample size determination)

ในการพิจารณาขนาดตัวอย่างที่จะศึกษานั้น จะต้องพิจารณาถึงลักษณะข้อมูลและองค์ประกอบต่าง ๆ ที่เกี่ยวข้องด้วย เช่น ค่าเฉลี่ยหรือค่าส่วนเบี่ยงเบนมาตรฐาน หรืออัตราส่วน การเปลี่ยนแปลงต่าง ๆ ว่าเปลี่ยนจากเดิมมากน้อยเท่าใด ข้อมูลเหล่านี้อาจจะได้จากรายงานที่ผู้อื่นเคยศึกษามาก่อนแล้วก็ได้ หรือได้จากการถามผู้รู้หรือผู้ที่เคยมีประสบการณ์มาก่อน นอกจากนี้ขนาดของตัวอย่างยังขึ้นอยู่กับวัตถุประสงค์ของการศึกษา และชนิดของการออกแบบการศึกษานั้น ๆ ด้วย ดังนั้นวิธีคำนวณหาขนาดตัวอย่างจึงมีหลายวิธี ประการสุดท้ายขึ้นกับว่าผู้ศึกษานั้นมีข้อมูลพื้นฐานอะไรบ้างเกี่ยวกับเรื่องที่จะศึกษา หากไม่มีแนวทางอะไรมาก่อนเลย คือ ไม่ทราบอะไรเลยและไม่ไปถามผู้รู้หรือผู้ที่มีประสบการณ์มาก่อนเลยด้วย ในกรณีเช่นนี้จะไม่สามารถคำนวณหาขนาดตัวอย่างโดยวิธีใด ๆ ได้เลย

การพิจารณาขนาดตัวอย่างที่จะกล่าวในที่นี้ จะพิจารณาตามแบบการศึกษาที่ศึกษากันบ่อยเพื่อผู้อ่านจะได้นำไปใช้ประโยชน์ในงานที่ต้องการศึกษาต่อไป ดังนี้คือ

### 1. การพิจารณาขนาดตัวอย่างในการศึกษาเชิงพรรณนา (Descriptive Studies)

การคำนวณขนาดตัวอย่างในการศึกษาเชิงพรรณนา มีวิธีคำนวณหลายวิธี โดยขึ้นอยู่กับชนิดของข้อมูลที่เก็บดังนี้

1.1 การคำนวณขนาดตัวอย่างของข้อมูลที่ได้จากการนับ (Quantal Data) ถ้าข้อมูลที่เก็บนั้นเป็นการสังเกตคว่ามีอาการหรือไม่มีอาการหรือคุณลักษณะสำคัญต่าง ๆ ของโรคเพื่อหาอัตราการเกิดโรค หรือสาเหตุป่วย (Morbidity Data) ข้อมูลเช่นนี้เรียกว่า *Quantal Data* และจะต้องกำหนดความแม่นยำ (precision) ของการศึกษานั้น โดยอาศัยการดูค่า Standard error ของตัวอย่างซึ่งคือความคลาดเคลื่อนของตัวอย่าง (sampling error) นั้นเอง

ยกตัวอย่าง เช่น จากการสำรวจความชุกของพยาธิปากขอในชุมชนหนึ่งพบว่ามีความชุกของโรค (Prevalence rate) = 40 %

กำหนด  $p = \% \text{ คนเป็นโรค} = 40 \%$

$q = 1 - p$  หรือ  $= 100 - 40 = 60\%$  คนไม่เป็นโรค ก่อนอื่นต้องหาค่าความคลาดเคลื่อนของตัวอย่างก่อน (sampling error) โดย

สูตร

$$\text{Sampling error} = \sqrt{\frac{p \cdot q}{n}} \quad n = \text{จำนวนตัวอย่าง}$$

(หรือ Standard error)

$$= \sqrt{\frac{40 \times 60}{100}} = 5$$

$$95\% \text{ confidence limit} = 40 \pm 2 \cdot \text{Standard error}$$

$$= 40 \pm 10 = 30 - 50\%$$

สรุป 95% ของตัวอย่างที่ศึกษาจะพบอัตราการเป็นโรคพยาธิปากขอ 30 - 50% นั่นคือ อัตราความชุกของพยาธิปากขอในประชากรนั้นมี 30 - 50% ถ้าเพิ่มจำนวน n มากขึ้น ค่า Sampling error จะลดลง ฉะนั้นจาก Sampling error ที่คำนวณได้ = 5 นั้นเอง ที่นำมาใช้ในการคำนวณหาขนาดตัวอย่างต่อไป

ในกรณีที่ไม่มีข้อมูลอะไรมาก่อนตั้งเช่นข้างบนก็อาจใช้ค่า sampling error ที่ผู้อื่นเคยทำมาก่อน แล้วตั้ง precision ว่า จะยอมให้มีความผิดพลาด หรือมี sampling error 5 หรือ 10% ก็ได้

โจทย์ตัวอย่างที่ 1 ถ้าต้องการสำรวจความชุกของโรค ๆ หนึ่ง ในจังหวัดหนึ่งโดยพอทราบมาบ้างว่าพบโรคนี้นราว 10% จึงคำนวณขนาดตัวอย่างที่ใช้การสำรวจ โดยตั้ง precision ว่า จะยอมให้มีความผิดพลาดของโอกาสที่จะพบโรคได้ = 5% หรือ = .05 =  $\lambda$

$$p = .1 - q = .9$$

$$\text{สูตร} \quad \sqrt{\frac{pq}{n}} = \lambda p$$

$$\text{หรือ} \quad \frac{pq}{n} = \lambda^2 p^2$$

$$n = \frac{pq}{\lambda^2 p^2} = \frac{q}{\lambda^2 \cdot p}$$

$$= \frac{.9}{(.05)^2 \times .1} = 3600 \text{ ราย}$$

∴ จำนวนตัวอย่าง = 3600

ถ้าในจังหวัดนั้นพบโรคนี้อีกขึ้น จำนวนขนาดตัวอย่างก็ใช้น้อยลง เช่น ถ้าพบโรคในจังหวัดนั้นประมาณ 30 %

กำหนด  $\lambda = .05$  หรือ 5 %

$$p = .3 \quad q = 1 - .3 = .7$$

$$n = \frac{q}{\lambda^2 \cdot p} = \frac{.7}{(.05)^2 \times .3} = 933$$

∴ จำนวนตัวอย่าง 933

โจทย์ตัวอย่างที่ 2 จากการศึกษาของแพทย์คนหนึ่งทำการตรวจ pap smear จากปากมดลูกในสตรีตั้งครรภ์ พบว่าในสตรีมีครรภ์ 1000 ราย พบมี pap smear ให้ผลบวก 3 ราย ถ้าต้องการทำการศึกษานี้ในสตรีมีครรภ์ในโรงพยาบาลต่าง ๆ ในเขต กทม. จะต้องใช้ขนาดตัวอย่างสักเท่าไร

$$p = \frac{3}{1000} = .003$$

$$q = 1 - .003 = .997$$

กำหนด  $\lambda = .05$  คือ กำหนดให้มีความผิดพลาดของโอกาสที่จะพบ pap smear มีผลบวกได้ 5 %

$$\sqrt{\frac{pq}{n}} = \lambda p$$

$$n = \frac{q}{\lambda^2 \cdot p}$$

$$n = \frac{.997}{(.05)^2 \times .003} = 133,000$$

หมายเหตุ ที่ต้องใช้ขนาดตัวอย่างมากเพราะโอกาสที่พบมะเร็งปากมดลูกในสตรีตั้งครรภ์นั้น พบน้อยมาก ถ้าใช้ขนาดตัวอย่างที่จะศึกษาน้อยอาจไม่พบเลยก็ได้ ดังนั้นจึงต้องใช้ขนาดตัวอย่างที่มากพอจึงจะพบ pap smear ที่ให้ผลบวก และผลที่ได้ก็สามารถสรุปเป็นผลที่พบในประชากรที่ศึกษานั้นได้ ปัญหาที่มีอยู่ว่าผู้ทำการศึกษาคือทำได้หรือไม่ ผลที่ได้จะคุ้มกับค่าใช้จ่ายและเวลาที่ทำการศึกษาหรือไม่นั่นเอง

## 1.2 การคำนวณขนาดตัวอย่างของข้อมูลที่ได้จากการวัด (Quantitative Data)

คือข้อมูลที่เก็บเป็นข้อมูลที่วัดจากการวัด ซึ่งมีค่าเป็นตัวเลขต่อเนื่อง (Continuous data) โดยทั่วไปแล้วความแตกต่างระหว่างตัวอย่างอาจเกิดจากสาเหตุดังต่อไปนี้

ก. ความแตกต่างระหว่างแต่ละคน

ข. ความแตกต่างในคนเดียวกันแต่วัดในเวลาต่างกัน (เช่น diurnal Variation)

ในกรณีนี้ Precision คือ การกำหนดความแม่นยำของค่าที่วัดได้จากตัวอย่าง โดยอาศัยการวัดความคลาดเคลื่อนของค่าเฉลี่ยของตัวอย่างนั่นเอง (Standard error of the sample means) ซึ่งขึ้นอยู่กับความคลาดเคลื่อนจากการวัดนั้น ๆ

ถ้าค่าที่วัดได้จากตัวอย่างมีค่าที่แตกต่างกันมาก คือ มีพิสัย (range) กว้างก็ยิ่งต้องการจำนวนตัวอย่างมากขึ้น

สูตร 
$$S.E. \text{ (Standard error of the sample means)} = \frac{S.D.}{\sqrt{n}} \text{ หรือ}$$
$$\therefore n = \frac{S.D.^2}{S.E.^2}$$

ในการคำนวณหาขนาดตัวอย่างโดยใช้วิธีนี้จะต้องตั้ง precision ก่อนว่าจะให้มี S.D. เท่าใดหรือจะให้ S.E. เท่าใด ถ้าไม่มีตัวเลขที่เคยทำการศึกษามาก่อน ก็อาจจะใช้ข้อมูลของคนอื่นที่เคยทำมาก่อนแล้วก็ได้

สูตรนี้อาจคิดเป็นสัดส่วนก็ได้ โดยใช้สูตร ดังนี้

$$S.E. = \sqrt{\frac{pq}{n}} \quad \text{หรือ} \quad n = \frac{p \cdot q}{S.E.^2}$$

โดยตั้ง precision ว่าจะให้ S.E. เท่าใด

โจทย์ตัวอย่างที่ 1 ถ้าต้องการศึกษาหาระดับฮีโมโกลบินในสตรีตั้งครรภ์ปกติในโรงพยาบาลต่าง ๆ ในกรุงเทพมหานคร โดยเท่าที่เคยทราบว่ามีความเข้มข้นเฉลี่ย 11.5 gm.% และมีค่าเบี่ยงเบนมาตรฐาน 2 จงคำนวณหาขนาดตัวอย่าง

กำหนด S.E. ของค่าเฉลี่ยของระดับฮีโมโกลบิน = .05 หรือ 5%

สูตร 
$$S.E. = \frac{S.D.}{\sqrt{n}}$$
$$.05 = \frac{2}{\sqrt{n}}$$
$$n = \frac{2^2}{(.05)^2} = 1600$$
$$\therefore \text{จำนวนตัวอย่าง} = 1600$$

โจทย์ตัวอย่างที่ 2 จากการทำ pilot study ในการทดสอบผลการตรวจทางห้องปฏิบัติการชนิดหนึ่งมีค่า S.D. = 5 จะต้องใช้ขนาดตัวอย่างเท่าใดในการศึกษานี้

กำหนด precision ให้ S.E. = 10%

$$\begin{aligned} \text{สูตร} \quad \text{S.E.} &= \frac{\text{S.D.}}{\sqrt{n}} \\ .1 &= \frac{5}{\sqrt{n}} \\ n &= \frac{5^2}{(.1)^2} = 2500 \end{aligned}$$

ถ้ากำหนดให้ S.E. เป็น 25%

$$\therefore n = \frac{5^2}{(.25)^2} = 400$$

ในกรณีเช่นนี้ แสดงว่าจำนวนตัวอย่างไม่เหมาะสมเพราะว่า S.E. สูงมาก

จะสังเกตเห็นว่าถ้ากำหนดให้ S.E. น้อยขนาดตัวอย่างจะยิ่งมากขึ้น ถ้าไม่ทราบค่าอะไรเลยก็ให้ลองศึกษาในตัวอย่างจำนวนหนึ่งแล้วคำนวณหาค่า S.E. ถ้ามีค่าน้อยก็แสดงว่าจำนวนตัวอย่างนั้นเหมาะสมแล้ว แต่ถ้าค่า S.E. มีค่าสูงก็แสดงว่าจำนวนตัวอย่างนั้นอาจมีขนาดน้อยไปควรต้องเพิ่มตัวอย่างต่อไปอีก จนกว่าจะได้ค่า S.E. ตามที่ใช้อย่างทั่ว ๆ ไป

### 1.3 การคำนวณขนาดตัวอย่างโดยคิดจากสัดส่วน

ก. การคำนวณขนาดตัวอย่างโดยคิดจากสัดส่วนและทราบสัดส่วนของเหตุการณ์ที่เกิดขึ้น

กำหนดให้  $p$  = สัดส่วนของเหตุการณ์ในตัวอย่าง

$P$  = สัดส่วนของเหตุการณ์ในประชากร

และกำหนดให้  $p$  เป็นค่าประมาณของ  $P$

$\alpha$  เป็นค่าความน่าจะเป็นที่ค่าประมาณ  $p$  ที่ ผิดพลาดไปจาก  $P$

$$\text{นั่นคือ } \Pr(|p - P|) = \alpha$$

d เป็นความแตกต่างระหว่าง  $p$  และ  $P$

จากสูตร  $\sigma_p = \sqrt{\frac{pq}{n}}$  ,  $\sigma_p$  = ความคลาดเคลื่อนของโอกาสที่จะเกิดเหตุการณ์

$$q = 1 - p$$

$$\begin{aligned} \text{สูตร} \quad Z &= \frac{p-P}{\sqrt{\frac{pq}{n}}} \\ Z &= \frac{d}{\sqrt{\frac{pq}{n}}} \\ \therefore n &= \frac{Z^2 pq}{d^2} \end{aligned}$$

**โจทย์ตัวอย่าง** การสำรวจในชนบทเกี่ยวกับผู้ที่เป็นโรคระเพาะอักษะ เมื่อ 5 ปี ก่อนพบว่า สัดส่วนผู้เป็นโรคระเพาะอาหารอักษะไม่มากไปกว่า 0.35 และช่วงความเชื่อมั่น 95% กำหนดความแตกต่างระหว่างค่าประมาณของสัดส่วนจากตัวอย่างกับสัดส่วนของประชากรให้เท่ากับ 0.05 ( $d = 0.05$ ) จงคำนวณขนาดตัวอย่างที่จะทำการสำรวจโรคระเพาะอาหารอักษะในชุมชนชนบทแห่งหนึ่ง

$$\text{จากสูตร} \quad n = \frac{Z^2 pq}{d^2}$$

$$\text{ค่า } Z = 1.96, \quad p = 0.35, \quad q = 0.65, \quad d = 0.05$$

$$n = \frac{(1.96)^2 (0.35) (0.65)}{(0.05)^2} = 349.6$$

$\therefore$  ขนาดตัวอย่างที่จะทำการสำรวจ 350 คน

หมายเหตุ : ค่า Z (Two-tailed Test) ที่ระดับนัยสำคัญ .05 มีค่า = 1.96

(จากการตารางแสดงค่า Z)

ข. การคำนวณขนาดตัวอย่างโดยคิดจากสัดส่วน โดยทราบค่าความคลาดเคลื่อน (S.E)

**โจทย์ตัวอย่าง** จากการศึกษาเกี่ยวกับยาฉีดคุมกำเนิดและการตั้งครรภ์หลังหยุดใช้ยาแล้วในประเทศอินเดีย พบว่าหลังหยุดฉีดยาแล้ว 3 เดือน สตรีเหล่านั้นจะตั้งครรภ์ได้ 80% ถ้าต้องการทำการวิจัยเรื่องนี้ จะต้องใช้ขนาดตัวอย่างเท่าไร

$$\text{โอกาสที่สตรีหยุดฉีดยาแล้ว 3 เดือน} \quad \text{ตั้งครรภ์ได้} = p$$

$$= .8$$

$$q = .2$$



$$\begin{aligned} \text{สูตร} \quad \text{S.E.} &= \sqrt{\frac{p \cdot q}{n}} \\ \text{กำหนด S.E.} &= 5\% \\ \therefore .05 &= \sqrt{\frac{.8 \times .2}{n}} \\ n &= \frac{.8 \times .2}{(.05)^2} = 64 \\ \therefore \text{จำนวนตัวอย่างที่ศึกษา} &= 64 \end{aligned}$$

1.4 การกำหนดขนาดตัวอย่าง ในกรณีที่ทราบค่าเบี่ยงเบนมาตรฐานและค่าเฉลี่ยของข้อมูลที่ต้องการศึกษานั้นมาก่อน ซึ่งอาจได้จากการศึกษานำร่อง (pilot study) หรือผลการศึกษาของผู้คนมาก่อนก็สามารถนำมาใช้กำหนดขนาดตัวอย่างได้ ตามตัวอย่างต่อไปนี้

โจทย์ตัวอย่าง ต้องการศึกษาน้ำหนักของเด็กอายุต่าง ๆ ตั้งแต่แรกเกิดจนถึงอายุ 18 ปี เพื่อจะได้ทราบค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน สำหรับใช้เปรียบเทียบกับเด็กที่มีปัญหาต้านน้ำหนัก และส่วนสูงจากงานวิจัยที่เคยได้ทำมาแล้วพบว่า เด็กอายุระหว่าง 11-12 ปี มีน้ำหนักเฉลี่ย 29 ก.ก. ส่วนเบี่ยงเบนมาตรฐาน 4 ก.ก. จะต้องใช้จำนวนตัวอย่างเด็กกลุ่มนี้เท่าไร

$$\begin{aligned} \text{C.V. } (\bar{x}) &= \frac{s}{\sqrt{n} \cdot \bar{x}} & \text{C.V.} &= \text{ค่าสัมประสิทธิ์แห่งการกระจาย (Coefficient of variation)} \\ n &= \frac{s^2}{\{\bar{x} (\text{C.V.})\}^2} \end{aligned}$$

$$\text{ถ้ากำหนดให้ C.V.} = .01, n = \frac{16}{\{29 (.01)\}^2} = \frac{16}{.084} = 190$$

$$\text{ถ้ากำหนดให้ C.V.} = .02, n = 48$$

$$\text{ถ้ากำหนดให้ C.V.} = .05, n = 8$$

และในเด็กอายุ 13-14 ปี พบว่ามีน้ำหนักเฉลี่ย 35 ก.ก. S.D. = 7 ก.ก. จะต้องใช้จำนวนตัวอย่างเด็กกลุ่มนี้เท่าไร

$$\text{ถ้าให้ C.V.} = .02, n = 100$$

$$\text{C.V.} = .03, n = 44$$

$$\text{C.V.} = .05, n = 16$$

จะสังเกตเห็นว่า ถ้ากำหนดให้ C.V สูงค่า n จะลดลงและถ้ากำหนด C.V. ค่าค่า n จะต้องใช้มากขึ้น

## 2. การพิจารณาขนาดตัวอย่างในการศึกษาเชิงวิเคราะห์ (Analytical Studies)

### 2.1 การคำนวณขนาดตัวอย่างในการศึกษาชนิดย้อนหลัง (Case-Control Study)

ก่อนที่จะคำนวณขนาดตัวอย่างจำเป็นต้องทราบข้อมูลพื้นฐานต่าง ๆ เสียก่อน ดังต่อไปนี้คือ

1. คนปกติของประชากรเป้าหมาย จะได้รับองค์ประกอบมากน้อยเพียงใด (Relative frequency) โดยกำหนดให้  $= P_0$

2. ตั้งสมมติฐาน และกำหนดให้การเสี่ยงต่อการเกิดโรคหลังได้รับองค์ประกอบแล้ว  $= R$

3. ตั้งระดับความเชื่อมั่น  $= \alpha$  (Two-sided)

4. กำหนดความคลาดเคลื่อนชนิดที่ 2 (Type II error)  $= \beta$

สูตร I  $n = [Z_\alpha \sqrt{2\bar{p}\bar{q}} + Z_\beta \sqrt{p_1 q_1 + p_0 q_0}]^2 / (p_1 - p_0)^2$  (two-sided)

โดย  $p_1 = p_0 \cdot R / [1 + p_0 (R - 1)]$

$\bar{p} = \frac{1}{2} (p_1 + p_0)$ ,  $\bar{q} = 1 - \bar{p}$

$q_1 = 1 - p_1$        $q_0 = 1 - p_0$

สูตร II (Two-sided)

$n = 2 \bar{p}\bar{q} (Z_\alpha + Z_\beta)^2 / (p_1 - p_0)^2$

โดยกำหนด  $\alpha = .05$  (two-sided),  $\beta = .10$

$\therefore Z_\alpha = 1.96$ ,  $Z_\beta = 1.28$

ตัวอย่าง การศึกษาความสัมพันธ์ของโรคหัวใจในทารกแรกเกิด (Congenital Heart disease) ต่อการที่มารดาได้รับฮอร์โมนขณะตั้งครรภ์ ปรากฏว่ามีมารดาที่มีบุตรปกติ (กลุ่มควบคุม) มีประวัติได้รับฮอร์โมนขณะตั้งครรภ์ 20 %

ให้  $p_0$  สัดส่วนที่มารดาของเด็กปกติได้รับฮอร์โมนขณะตั้งครรภ์ .20

Relative risk ที่ตั้งไว้  $= 3$ ,  $\alpha = .05$  (two-sided),  $\beta = .10$

$$p_1 = p_0 \cdot R / [1 + p_0 (R - 1)]$$

$$\therefore p_1 = .20 \times 3 / [1 + .20 (3 - 1)] = .4286$$

$$\bar{p} = \frac{1}{2} (p_1 + p_0) = \frac{1}{2} (.4286 + .20) = .3143$$

$$\begin{aligned} \text{สูตร I } n &= [Z\alpha\sqrt{2\bar{q}\bar{p}} + Z\beta\sqrt{p_1q_1 + p_0q_0}]^2 / (p_1 - p_0)^2 \\ n &= \frac{[1.96\sqrt{2(.3143 \times .6857)} + 1.28\sqrt{.4286 \times .5714 + (.2 \times .8)}]^2}{(.4286 - .20)^2} \\ &= \frac{[1.287 + .8145]^2}{.0523} = 84 \end{aligned}$$

หรือคิดค่า  $n$  จากสูตร II

$$\begin{aligned} n &= 2\bar{q}\bar{p}(Z\alpha + Z\beta)^2 / (p_1 - p_0)^2 \\ &= 2(.3143 \times .6857)(1.96 + 1.28)^2 / (.4286 - .20)^2 \\ &= \frac{4.525}{.0523} = 87 \end{aligned}$$

ขนาดของตัวอย่างจะขึ้นอยู่กับค่า Relative risk ถ้าค่า Relative risk สูงขนาดตัวอย่างจะน้อยลง ดังแสดงในตารางข้างล่าง ซึ่งคำนวณขนาดตัวอย่างในการศึกษาอันหลังเกี่ยวกับโรคหัวใจเด็กแรกเกิด (Congenital Heart disease) กับการศึกษาที่ได้รับฮอร์โมนในระยะตั้งครรภ์ ดังนี้

Relative risk	ขนาดตัวอย่างของกลุ่ม
2	188
3	73
4	45
5	34
7	24
10	18

หมายเหตุ การคำนวณตัวอย่างข้างบนใช้สูตร I และกำหนด  $p_0 = .3$ ,  $\alpha = .05$  (two-sided) และ  $\beta = .10$

## 2.2 การคำนวณขนาดตัวอย่างในการศึกษาไปข้างหน้า (Cohort Study)

ก่อนที่จะคำนวณขนาดตัวอย่างในการศึกษานี้จำเป็นต้องทราบข้อมูลพื้นฐานต่าง ๆ เกี่ยวกับอุบัติการณ์ของโรคหรือเหตุการณ์ที่เป็นผลจากการได้รับองค์ประกอบนั้นก่อน โดยอาจจะใช้ตัวเลขจากผู้อื่นเคยทำการศึกษารวบรวมมาก่อนแล้วก็ได้ ขนาดตัวอย่างที่จะใช้ในการศึกษานี้จะมากน้อยขึ้นอยู่กับอุบัติการณ์ของโรค เหตุการณ์ที่พึงเกิดในกลุ่มที่ไม่ได้รับองค์ประกอบและ Relative risk ของโรคนั้นเป็นสำคัญ โดยจะแตกต่างไปจากขนาดตัวอย่างในการศึกษาชนิดย้อนหลัง (Case-Control studies) ซึ่งการคำนวณขนาดตัวอย่างในการศึกษาชนิดย้อนหลังนั้นจะขึ้นอยู่กับ การได้รับองค์ประกอบในกลุ่มคนที่เป็นโรคและไม่เป็นโรคว่ามีประวัติมากน้อยเพียงใด (Prevalence of Exposure to the Factor) รวมทั้งขึ้นอยู่กับ Relative risk ของโรคนั้นด้วย เป็นสำคัญ

ข้อมูลพื้นฐานที่จำเป็นต้องทราบก่อนที่จะคำนวณขนาดตัวอย่าง คือ

1. โอกาสที่จะเกิดโรคในกลุ่มที่ไม่ได้รับองค์ประกอบ (Non-Exposed)

โดยกำหนดให้ =  $p_1$

2. ค่าสัมประสิทธิ์และกำหนดให้การเสี่ยงต่อการเกิดโรค (Relative risk) =  $R$

3. ค่าระดับความเชื่อมั่น =  $\alpha$  (Two-sided)

4. กำหนดความคลาดเคลื่อนชนิดที่ 2 (Type II error) =  $\beta$

ให้  $n$  = ขนาดของตัวอย่างของแต่ละกลุ่ม

$$\text{สูตร I } n_{\text{(Two-sided)}} = \frac{\{Z_\alpha \sqrt{2 \bar{p} \bar{q}} + Z_\beta \sqrt{p_1 [1 + R - p_1 (1 + R^2)]}\}^2}{[p_1 (1 - R)]^2}$$

โดย  $\bar{p} = \frac{1}{2} p_1 (1 + R)$ ,  $\bar{q} = 1 - \bar{p}$

$Z_\alpha$  และ  $Z_\beta$  เป็นค่าความคลาดเคลื่อนชนิดที่ I และ II ( $\alpha$  error และ  $\beta$  error) ซึ่งจะได้จากตารางแสดงค่า  $Z_\alpha$  และ  $Z_\beta$

หมายเหตุ สูตรที่ I นี้จะเน้นความสำคัญให้เห็นว่าขนาดตัวอย่างขึ้นอยู่กับ  $p_1$  และ  $R$  ถ้าอุบัติการณ์ยังพบน้อยขนาดตัวอย่างยังต้องใช้มากขึ้น

$$\text{หรือใช้สูตร II } n_{\text{(Two-sided)}} = \frac{\{Z_\alpha \sqrt{2 \bar{p} \bar{q}} + Z_\beta \sqrt{p_1 (1 - p_1 + p_2 (1 - p_2))}\}^2}{(p_1 - p_2)^2}$$

โดยค่า  $\bar{p} = \frac{1}{2} (p_1 - p_2)$ ,  $\bar{q} = 1 - \bar{p}$

แทนค่า  $p_2$  จากสูตร  $p_2 = p_1 R$  ซึ่งขนาดของตัวอย่างจะมากหรือน้อยจะขึ้นอยู่กับค่า  $P_1$  และ  $R$

ตารางแสดงค่าความคลาดเคลื่อนชนิดที่ I และ II ( $Z_\alpha$  &  $Z_\beta$ )

$\alpha$ ( $/\beta$ )	One-sided test	Two-sided test
	$Z_\alpha$ ( $/ Z_\beta$ )	$Z_\alpha$
0.001	3.09	3.29
0.005	2.58	2.81
0.01	2.33	2.58
0.025	1.96	2.24
0.05	1.64	1.96
0.10	1.28	1.64
0.20	0.84	1.28
0.30	0.52	1.04

โจทย์ตัวอย่าง การศึกษาเพื่อหาความสัมพันธ์ของโรคหัวใจพิการแต่กำเนิด (Congenital heart disease - CHD) ต่อการที่มารดาได้รับยาเม็ดคุมกำเนิด (OC<sub>s</sub>) ก่อนหรือหลัง Conception 3 เดือน

$$p_1 = \frac{\text{อุบัติการณ์ของ CHD ในกลุ่มไม่ได้รับองค์ประกอบ}}{\text{8 ราย/1000 เด็กเกิดมีชีวิต}} = .008$$

กำหนด  $\alpha = .05$   $\beta = .10$

$$\text{จากสูตร I n} = \frac{\left\{ Z_\alpha \sqrt{2 \bar{p} \bar{q}} + Z_\beta \sqrt{p_1 [1 + R - p_1 (1 + R^2)]} \right\}^2}{p_1 (1 - R)^2}$$

$$\bar{p} = \frac{1}{2} p_1 (1 + R), \bar{q} = 1 - \bar{p}$$

กำหนด  $R = 2, 3, 4, 5, 7$  และ  $10$

ก. ถ้า Relative risk = 2,  $\bar{p} = \frac{1}{2} (0.008) (1 + 2) = 0.012$

$\therefore \bar{q} = 1 - 0.012 = 0.988$

$$\begin{aligned} \text{แทนค่าจากสูตร } I_n &= \frac{\left\{ 1.96 \sqrt{2(0.012)(0.988)} + 1.28 \sqrt{0.008 [1 + 2 - 0.008(1+4)]} \right\}^2}{[0.008(1-2)]^2} \\ &= \frac{0.248786}{(-0.008)^2} \\ &= 3887 \end{aligned}$$

เมื่อกำหนดขนาดตัวอย่างตามค่า Relative risk ดังกล่าวจะได้ผลตามตารางต่อไปนี้  
ตารางแสดงจำนวนตัวอย่างตามค่า Relative risk ต่าง ๆ กัน

Relative risk	ขนาดตัวอย่าง
2	3,887
3	1,289
4	712
5	478
7	280
10	168

### 3. การพิจารณาขนาดตัวอย่างในการวิจัยเชิงทดลอง (Experimental studies)

การพิจารณาขนาดตัวอย่างในการศึกษาชนิดนี้ จะต้องต้องมีข้อมูลที่ทราบมาก่อนแล้วบ้าง เช่นกันกับการวิจัยแบบอื่น ถ้าการวิจัยเชิงทดลองนี้เป็นการทดลองในกลุ่มเดียว ก็อาจเลือกใช้วิธีใดวิธีหนึ่งของการกำหนดขนาดตัวอย่างในการศึกษาเชิงพรรณนาดังกล่าวแล้วในตอนต้นได้ แต่ถ้การทดลองนั้นเป็นการทดลองใน 2 กลุ่ม ก็จะต้องมีวิธีการกำหนดขนาดตัวอย่างดังนี้

#### 3.1 การกำหนดขนาดตัวอย่างในการทดลอง 2 กลุ่มที่ทราบค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน

โดยที่การวิจัยชนิดนี้เป็นกรวิจัยที่ต้องการทดสอบความแตกต่างของค่าที่ได้จากประชากรสองกลุ่ม ดังนั้นการกำหนดตัวอย่างเพื่อทดสอบความแตกต่างของประชากรทั้งสองกลุ่มนี้ จึงต้องกล่าวถึงการทดสอบสมมติฐานเพื่อดูความแตกต่างของค่าเฉลี่ย 2 ชุด และความคลาดเคลื่อน (error) ทั้ง 2 ชนิดนั้นเสียก่อนดังนี้ คือ

- Type I error ( $\alpha$ ) — ความคลาดเคลื่อนที่ไม่ยอมรับทั้งที่สมมติฐานเป็นจริง
- Type II error ( $\beta$ ) — ความคลาดเคลื่อนที่ยอมรับทั้งที่สมมติฐานไม่เป็นจริง

## สมมติฐาน

	เป็นจริง	ไม่เป็นจริง
ยอมรับ (Accept)	$(1-\alpha)$	Type II error ( $\beta$ )
ไม่ยอมรับ (Reject)	Type I error ( $\alpha$ )	$(1-\beta)$

จากการทดสอบสมมติฐาน  $H_0 : \mu_1 = \mu_2$

$H_A : \mu_1 \neq \mu_2$

ใช้ Z test ในการทดสอบสมมติฐาน

$$Z_\alpha = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}, \quad Z_\beta = \frac{D - (\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

$$S_p^2 = \text{pooled Variance} = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}$$

D = ความแตกต่างของค่าเฉลี่ยที่ต้องการทดสอบ

$$Z_\alpha + Z_\beta = \frac{D}{\sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}}$$

$$\frac{n_1 n_2}{n_1 + n_2} = \frac{(Z_\alpha + Z_\beta)^2 S_p^2}{D^2}$$

ถ้าจำนวนตัวอย่างในกลุ่มที่ 1 เท่ากับจำนวนตัวอย่างในกลุ่มที่ 2 หรือ  $n_1 = n_2$

$$\therefore n = \frac{(Z_\alpha + Z_\beta)^2 S_p^2}{D^2}$$

ถ้าเป็นการทดลองในประชากรชุดเดียว แต่ต้องการดูความแตกต่างก่อนและหลังการทดลอง (paired test) ก็ใช้คำนวณขนาดตัวอย่างดังนี้

$$n = \frac{(Z_\alpha + Z_\beta)^2 S_p^2}{D^2}$$

ในทางชีววิทยาและการแพทย์มักจะกำหนดค่า  $\alpha = .01$  หรือ  $= .05$  ทั้งนี้ขึ้นอยู่กับความสำคัญของการทดลองนั้น ๆ ถ้าเป็นการทดลองที่ทำในคนแล้วไม่ควรกำหนดให้ค่า  $\alpha$  สูงกว่า .05 หากเป็นการทดลองในสัตว์หรือค้ำนธุรกิจก็อาจจะยอมให้  $\alpha = .01$  หรือ  $= .02$  ได้ส่วนค่า  $\beta$  จะมากกว่าหรือเท่ากับ  $\alpha$  แต่จะไม่น้อยกว่าค่า  $\alpha$

### 3.2 การคำนวณขนาดตัวอย่างในการทดลองสองกลุ่มที่ทราบค่าสัดส่วนของเหตุการณ์ที่เกิดขึ้น

ถ้ามีผลงานจากการทำวิจัยนำร่อง (Pilot study) หรือจากงานวิจัยของผู้อื่นที่บอกค่าเป็นสัดส่วนของเหตุการณ์ที่เกิดขึ้นก็สามารถ จะนำมาคำนวณหาขนาดของตัวอย่างต่อไปได้

$$\text{สูตรที่ 1 } n = \frac{\left\{ Z_\alpha \sqrt{2p_1(1-p_1)} + Z_\beta \sqrt{P_1(1-p_1) + P_2(1-p_2)} \right\}^2}{D^2}$$

$P_1$  = สัดส่วนที่เกิดเหตุการณ์ในตัวอย่างที่ 1

$P_2$  = สัดส่วนที่เกิดเหตุการณ์ในตัวอย่างที่ 2

$D = P_1 - P_2$  ความแตกต่างของสัดส่วนของเหตุการณ์ทั้งสอง

$$\text{สูตรที่ 2 } n = \frac{(Z_\alpha + Z_\beta)^2 2\bar{p}(1-\bar{p})}{D^2}$$

$$\bar{p} = \frac{1}{2} (P_1 + P_2)$$

ตัวอย่าง ในการเปรียบเทียบผลการใช้ยาชนิดที่ 1 รักษาโรคพยาธิใบไม้ในตับพบว่ามีโอกาสหายได้ 60% และใช้ยาชนิดที่ 2 รักษาพบว่าหายได้ 80% อยากทราบว่าจะต้องใช้ตัวอย่างกลุ่มละเท่าไรในการวิจัยเพื่อศึกษาประสิทธิภาพของยาทั้ง 2 ชนิดนี้ในการรักษาพยาธิใบไม้ในตับของผู้ป่วยนี้

$$P_1 = 0.6, P_2 = 0.8, \bar{p} = \frac{1}{2} (0.6 + 0.8) = 0.7$$

$$D = P_2 - P_1 = 0.2$$

กำหนดให้  $\alpha = .05$  (one-sided)  $Z_\alpha = 1.64$

$$\beta = .10 \quad Z_\beta = 1.28$$

$$n = \frac{(Z_\alpha + Z_\beta)^2 2\bar{p}(1-\bar{p})}{D^2}$$



$$= \frac{(1.64 + 1.28)^2 2(0.7 \times 0.3)}{(0.2)^2} = 89.2$$

∴ จะต้องตัวอย่างกลุ่มละ 89 คน

จะสังเกตได้ว่า ถ้าความแตกต่างของคุณภาพอย่างยิ่งมาก (D)

จำนวนตัวอย่างที่จะใช้ในงานวิจัยก็ยิ่งน้อยลงไป เช่น ถ้าความแตกต่างของยาสองชนิดมีมากขึ้นจาก 20% เป็น 30% จะได้

$$n = \frac{(1.64 + 1.28)^2 2(0.7)(0.3)}{(0.3)^2}$$

$$= 40$$

∴ จะใช้ตัวอย่างกลุ่มละ 40 คน

หมายเหตุ สำหรับงานวิจัยที่ใช้เวลานานเป็นปี ๆ เวลาที่คิดขนาดตัวอย่างจำเป็นจะต้องนึกถึงจำนวนผู้ที่หายไปหรือขาดการติดต่อ หรือตายไประหว่างการทดลอง (lost to follow up) ด้วย ควรจะประมาณไว้เลยว่าจะหายไประหว่างการทดลองสักเท่าไร แล้วนำจำนวนนี้ไปบวกกับขนาดตัวอย่างที่คำนวณได้อีกด้วย

## II. เทคนิคการเลือกตัวอย่าง (SAMPLING TECHNIQUES)

การสุ่มตัวอย่างเพื่อทำการศึกษานั้น จุดประสงค์ก็เพื่อเลือกตัวอย่างโดยไม่มีอคติลำเอียง ทุก ๆ คนหรือหน่วยที่จะเลือกตัวอย่างมีโอกาสที่จะถูกเลือกมาศึกษาเท่า ๆ กัน ตัวอย่างที่เลือกได้ก็ย่อมมีคุณสมบัติเป็นตัวแทนของประชากรที่เลือกตัวอย่าง (population sampled) นั้นได้ ผลงานที่ได้ก็สามารถสรุปเป็นผลของประชากรที่ต้องศึกษาได้ นอกจากนั้นยังช่วยประหยัดค่าใช้จ่ายลดระยะเวลาศึกษาให้สั้นลง และประหยัดกำลังคนด้วย เมื่อกำหนดวิธีเลือกตัวอย่างและจำนวนตัวอย่างที่จะศึกษาแล้ว ก็จะต้องปฏิบัติตามที่ได้กำหนดไว้ วิธีที่จะเลือกตัวอย่างให้ถูกต้องและมีคุณสมบัติเป็นตัวแทนที่ดีของประชากรที่ศึกษานั้นมีหลายวิธีด้วยกัน ซึ่งผู้ศึกษาจะต้องทราบถึงรายละเอียดของวิธีการเลือกตัวอย่างชนิดต่าง ๆ เสียก่อน แล้วนำมาพิจารณาเพื่อหาวิธีที่เหมาะสมซึ่งสามารถทำได้ต่อไป

ประเภทของการสุ่มตัวอย่าง

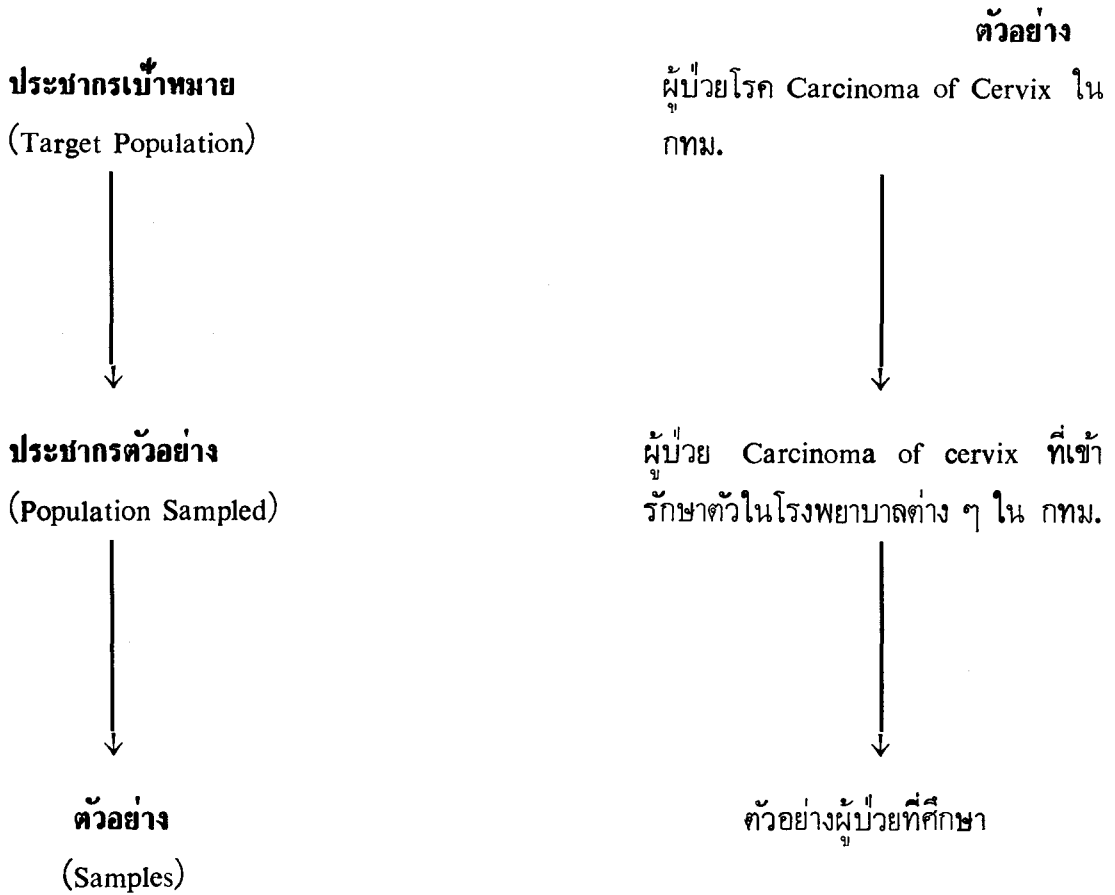
การสุ่มตัวอย่างแบ่งเป็น 2 ประเภท ด้วยกัน คือ

1. การสุ่มตัวอย่างโดยใช้หลักของความน่าจะเป็น (Probability) มาช่วยในการเลือกตัวอย่าง (Random Sampling) ตัวอย่างที่เลือกได้โดยวิธีนี้เรียกว่า "Probability Samples" หรือ "Random Samples" การสุ่มตัวอย่างชนิดนี้สามารถทราบถึงโอกาส (Probability) ของแต่ละคนหรือหน่วยที่จะถูกเลือกเข้ามาเป็นตัวอย่างการศึกษา

2. การเลือกตัวอย่างโดยไม่ต้องอาศัยหลักความน่าจะเป็นมาช่วย ตัวอย่างที่เลือกได้โดยวิธีนี้เรียกว่า "Non - Probability Sample" ได้แก่การเลือกตัวอย่างตามความสะดวกหรือจากอาสาสมัคร ตัวอย่างที่เลือกมาศึกษาโดยวิธีนี้จะไม่คุณสมบัติเป็นผู้แทนของประชากรที่จะศึกษาในการปฏิบัติจริงอาจจำเป็นต้องใช้วิธีนี้เพราะมีความสะดวก และสามารถทำได้ การปฏิบัติงานก็จะสิ้นเปลืองค่าใช้จ่ายน้อย และใช้ในกรณีที่ไม่สามารถเก็บตัวอย่างตามชนิดแรกได้

ถ้าตัวอย่างในกรณีนี้มีจำนวนมาก ก็อาจจะสุ่มเลือกตัวอย่างมาศึกษาอีกทีโดยใช้หลักความน่าจะเป็นตามวิธีแรกได้

## ประชากร-ประชากรตัวอย่าง และตัวอย่าง



### 1. การสุ่มตัวอย่างโดยใช้หลักความน่าจะเป็น

#### (PROBABILITY RANDOM SAMPLING)

1.1 การสุ่มเลือกตัวอย่าง โดยวิธี "Simple Random Sampling" วิธีนี้แต่ละคนหรือหน่วยในประชากรที่จะเลือกตัวอย่าง (population sampled) มีโอกาสจะถูกเลือกมาเป็นตัวอย่างเท่า ๆ กัน และจำเป็นต้องทราบจำนวนทั้งหมดที่จะถูกเลือกเสียก่อนจากนั้น จึงสุ่มเลือกตัวอย่างมาโดยใช้ตัวเลขในตารางสุ่มตัวอย่าง (random number table) ซึ่ง Fisher ได้เป็นผู้สร้างตารางนี้จากหลักความน่าจะเป็น (probability) และเป็นตารางแสดงตัวเลขที่อาจจะเลือกใช้ตามแนวใดก็ได้ ถ้าเลขที่ได้จากตารางตรงกับเลขที่ใดของผู้ถูกเลือกก็จะถูกเลือกเป็นตัวอย่างที่จะศึกษา

ได้ นอกจากนี้อาจจะสุ่มเลือกโดยวิธีการใส่เลขที่ทุก ๆ คน หรือหน่วยที่จะถูกเลือกแล้วจับสลาก ก็ได้รับการสุ่มตัวอย่างชนิดนี้ส่วนมากเป็นการสุ่มแบบไม่เอาผู้ที่สุ่มได้แล้วกลับไปแทนที่เดิมเพื่อเลือก อีก หมายถึง ผู้ที่ถูกเลือกแล้วไม่มีโอกาสถูกเลือกซ้ำอีกเลย (sampling without replacement) หมายเหตุ ผู้ที่จะสุ่มเลือกตัวอย่างโดยใช้ตารางเลขสุ่มนี้ ควรต้องศึกษาและทดลองสุ่มเลือกโดยใช้ตารางสุ่มเลขเสียก่อน ตัวอย่างเช่น ต้องการสุ่มเลือกตัวอย่างจากชุมชนหนึ่งซึ่งมี 400 ครัวเรือน และต้องการสัมภาษณ์หัวหน้าครัวเรือนในชุมชนนั้น 20% (80 ครัวเรือน) โดยใช้ตัวเลขจากตารางสุ่มประการแรกต้องใส่เลขที่ครัวเรือนเหล่านั้นตั้งแต่เลขที่ 1 ถึงเลขที่ 400 การเลือกใช้ตัวเลขจากตารางสุ่มดังแสดงตารางที่ 1 โดยอาจเริ่มที่แถวใดหรือเลขใดก็ได้แล้วเลือกไปตามแถวที่ละแถวตามแนวขนานหรือเลือกเป็นเส้นทะแยงก็ได้ จากนั้นกำหนดว่าจะใช้ตัวเลขแรก ๆ หรือท้าย ๆ ของเลขในตารางสุ่มเสียก่อน ในกรณีเช่นนี้ เนื่องจากมีจำนวนครัวเรือนทั้งหมด 400 ครัวเรือน และกำหนดจะใช้เลขจากท้าย 3 หลักที่มีค่าไม่เกิน 400 จากนั้นดูค่าตัวเลขจากตารางสุ่มไปทุกตัวตามแนวอนทกแถวหรือแนวทะแยงก็ได้ เมื่อได้ค่าเลขที่เท่าใดก็จะนำไปใช้ให้ตรงกับเลขที่ของครัวเรือนที่ได้กำหนดไว้แต่ก่อนเลือกซึ่งเลขที่ 3 หลักค้ำหลังของตัวเลขจากตารางสุ่มนี้ อาจจะมีค่าแตกต่างกันไป นอกจากนี้ควรกำหนดการไม่เลือกซ้ำไว้ด้วยเลย คือ ถ้าได้เลขที่ใดจากตาราง และกำหนดตามเลขที่ครัวเรือนแล้วก็ควรจะทำเครื่องหมายไว้ว่าเลือกแล้ว หากได้เลขสุ่มจากตารางซ้ำกันอีกก็ผ่านไปเลยสุ่มเลือกต่อไปจนกว่าจะได้ครบ 80 ครัวเรือน จำนวนครัวเรือน 80 ที่เลือกเป็นตัวอย่างเพื่อสัมภาษณ์นี้จะเป็นการสุ่มเลือกที่ยุติธรรมที่สุดโดยทุก ๆ ครัวเรือนมีโอกาสถูกเลือกเท่า ๆ กัน (ไม่มีอคติในการเลือกเลย) เมื่อเลือกตัวอย่างครัวเรือนแล้วก็ต้องปฏิบัติตามการสัมภาษณ์ตามที่ได้เลือกไว้ทุกครัวเรือน หากไม่ได้สัมภาษณ์ก็ครัวเรือน ที่กล่าวนี้เป็น การสุ่มตัวอย่าง โดยใช้ตารางเลขสุ่มซึ่งจะนำไปใช้ในการศึกษาแบบอื่น ๆ ได้

การสุ่มเลือกตัวอย่างชนิดนี้ในชุมชนใหญ่ จะมีปัญหาในด้านปฏิบัติงานและมีความยุ่งยากในการหาบ้านให้พบ หากสามารถปฏิบัติได้ก็จะเป็นการสุ่มตัวอย่างที่ถูกต้องที่สุด และเป็นตัวแทนของครัวเรือนในชุมชนโดยสมบูรณ์

1.2 การสุ่มเลือกตัวอย่างโดยเลือกตามกำหนดลำดับเลขที่ (Systematic Random Sampling) วิธีนี้เป็นวิธีที่นิยมใช้กันมาก เพราะง่ายสะดวกและรวดเร็ว จำนวนตัวอย่างที่ได้สามารถเลือกได้ตามสัดส่วนของประชากร เช่น ต้องการเลือกตัวอย่างนักเรียน 50 คน จากนักเรียนทั้งหมด

1000 คน ก็สามารถเลือกได้โดยหาช่วงของการสุ่มเลือกตัวอย่างก่อน (Sampling interval) ซึ่งคือ  $1000/50 = 20$  ดังนั้นจะสามารถเลือกตัวอย่างนักเรียน 1 คน ในทุกๆ 20 คน ก่อนเริ่มเลือกประชากรแรกนักเรียนทั้งหมด 1000 คน ต้องมีเลขที่ทุกคนแล้วจับสลากเลือกตัวอย่างแรกจากลำดับ 1-20 สมมติว่าจับได้เลขที่ 14 ฉะนั้นตัวอย่างที่จะเลือกตัวไปก็คือเลือกทุกๆ ช่วง 20 คือ เลขที่ 34, 54, 74, 94, 114, 134 ..... ฯลฯ จนครบ 50 ตัวอย่าง

1.3 การสุ่มเลือกตัวอย่างโดยแบ่งเป็นกลุ่มตามลักษณะต่างๆ เสียก่อน (Stratified Sampling) วิธีนี้ใช้แบ่งกลุ่มประชากรที่จะเลือกตัวอย่างออกเป็นกลุ่มหรือพวกตามลักษณะที่ต้องการ โดยให้แต่ละกลุ่มประกอบด้วยกลุ่มที่มีความคล้ายคลึงกันมากที่สุด (homogeneity within Stratum) ทั้งนี้เพื่อจุดประสงค์ในการเพิ่มประสิทธิภาพการสุ่มเลือกให้ถูกต้องยิ่งขึ้น และช่วยลดการกระจายของผลการวิจัยในแต่ละกลุ่ม เช่น แบ่งประชากรที่จะเลือกตัวอย่างตามกลุ่มอายุต่างๆ เสียก่อนแล้วจึงสุ่มตัวอย่างจากแต่ละกลุ่มอายุโดยวิธี 1.1 หรือ 1.2 ก็ได้ วิธีนี้จะช่วยลดจำนวนตัวอย่างที่จะต้องเลือกเป็นอย่างดี เพราะการจัดกลุ่มตามอายุ จะช่วยกำจัดการกระจายของตัวอย่างให้น้อยลงไปชั้นหนึ่งแล้ว เมื่อสุ่มเลือกตัวอย่างโดยวิธี 1.1 หรือ 1.2 จากแต่ละกลุ่มอายุจึงไม่จำเป็นต้องใช้จำนวนตัวอย่างมากนัก

1.4 การสุ่มเลือกตัวอย่างตามเขตสถานที่ [Cluster (area) Sampling] การสุ่มตัวอย่างชนิดนี้เหมาะในกรณีที่ต้องการสุ่มเลือกตัวอย่างจากประชากรที่ใหญ่มากเช่นต้องการสุ่มเลือกตัวอย่างเพื่อศึกษาทัศนคติ และการปฏิบัติตัวเกี่ยวกับการใช้บริการทางการแพทย์ของรัฐฯ ของประชากรในกรุงเทพมหานคร ก่อนสุ่มเลือกตัวอย่างก็ต้องแบ่งประชากรตามเขตพื้นที่เป็นกลุ่มใหญ่ๆ (Cluster) เสียก่อน จากนั้นจึงสุ่มเลือกจากกลุ่มนั้นๆ อีกที การสุ่มเลือกชนิดนี้อาจทำได้หลายขั้นตอนตามความต้องการของผู้ที่จะศึกษาเองได้ดังนี้

ก. สุ่มจากกลุ่มโดยแบ่งชั้นเดียว (Single Stage Cluster Sampling) การสุ่มเลือกชนิดนี้คงต้องหาข้อมูลก่อนว่า กทม. แบ่งเป็นกี่อำเภอ แล้วสุ่มเลือกจากอำเภอทั้งหมดของ กทม. โดยวิธี 1.1 หรือ 1.2 ก็ได้ตามจำนวนอำเภอที่ต้องการ ถ้าเลือกได้อำเภอใดก็ทำการศึกษาคูครัวเรือนในอำเภอนั้น

ข. สุ่มจากกลุ่มโดยแบ่ง 2 ชั้น (Two Stage Cluster Sampling) เช่น เริ่มสุ่มจากอำเภอก่อนเมื่อได้อำเภอแล้วก็สุ่มเลือกจากจำนวนตำบลในอำเภอนั้นๆ อีกทีตามวิธี 1.1 หรือ 1.2 ถ้าเลือกได้ตำบลใดก็ทำการศึกษาทั้งหมดในตำบลที่เลือกได้

ค. สุ่มจากกลุ่มโดยแบ่งหลายชั้น (Multi Stage Sampling) ถ้าประชากรที่ศึกษามีขนาดใหญ่มาก เช่น ประเทศไทยก็อาจจะสุ่มจังหวัด-อำเภอ-ตำบล-หมู่บ้าน จากทุกๆ ภาคของประเทศได้ สำหรับการสุ่มเลือกแต่ละชั้นของกลุ่ม (Cluster) ก็ใช้วิธี 1.1 หรือ 1.2 ดังกล่าวแล้วข้างต้นนั่นเอง

การสุ่มเลือกตัวอย่างชนิดนี้ถ้าประชากรที่จะศึกษามีขนาดใหญ่มากก็จะช่วยให้ทุนค่าใช้จ่ายและสะดวกในการปฏิบัติงานมาก สิ่งที่สำคัญที่สุด ก็คือผลที่ได้สามารถสรุปเป็นผลของประชากรเป้าหมายได้

## 2. การเลือกตัวอย่างโดยไม่ใช้หลักความน่าจะเป็น (Non-Probability Sampling)

การศึกษาบางอย่างก็จำเป็นต้องทำการศึกษาจากตัวอย่างที่ได้จากชนิดนี้ และจะเป็นการเก็บตัวอย่างจากหลายวิธี ดังนี้

ก. ตัวอย่างที่ได้จากความบังเอิญ (Accidental Sampling) คือ รวบรวมตัวอย่างเท่าที่จะหาได้ เช่น ต้องการสัมภาษณ์สตรีที่เดินอยู่แถวย่านการค้า และสัมภาษณ์คนใดก็ได้ที่ผ่านตามจำนวนที่ต้องการ

ข. ตัวอย่างที่ได้จากการกำหนดจำนวน (Quota Sampling) คล้ายกับชนิดที่เลือกตัวอย่างโดยแยกประเภท หรือลักษณะของประชากรก่อน (stratified sampling) เหมาะกับในกรณีที่ต้องการศึกษาในพวกที่มีลักษณะต่างกัน เช่น ต้องการศึกษาเกี่ยวกับรายละเอียดของโรคๆ หนึ่ง ในชายและหญิง (ผู้ใหญ่) ซึ่งต้องจำกัดจำนวนลงไปกว่าจะต้องการเลือกตัวอย่างจากผู้ป่วยชายและหญิงอย่างละกี่คน เช่น ศึกษาโดยกำหนดทำการศึกษาผู้ป่วยโรค myocardial infarction ชาย 100 ราย และหญิง 100 ราย เป็นต้น

ค. การเลือกตัวอย่างโดยกำหนดไว้ (Purposive Sampling) การเลือกตัวอย่างวิธีนี้ผู้ทำการศึกษาจะเป็นผู้กำหนดลักษณะของตัวอย่างที่จะทำการศึกษาเอง เช่น ต้องการศึกษาเกี่ยวกับประสิทธิภาพของยาชนิดหนึ่ง ผู้ศึกษาก็จะกำหนดว่าจะทำการศึกษาในรายที่มีลักษณะเฉพาะใดๆ ไว้ก่อน แล้วจึงเลือกตัวตามลักษณะที่กำหนดไว้

ง. การเลือกตัวอย่างตามความสะดวก (Convenient Sampling) การเลือกตัวอย่างชนิดนี้ต้องอาศัยความสนใจของผู้ที่จะเป็นตัวอย่างศึกษาเป็นใหญ่ เช่น ต้องการศึกษาดังผลของยาชนิด

หนึ่งจากตัวอย่างซึ่งเป็นอาสาสมัคร ซึ่งจะมีความสะดวกในการเลือกตัวอย่างและสะดวกในการศึกษาด้วย

หมายเหตุ วิธีการละเอียดเกี่ยวกับการสุ่มเลือกตัวอย่าง (Randomized Allocation) ในวิจัยเพื่อประเมินผลการรักษาในคลินิก (Clinical Trials) นั้น จะนำเผยแพร่ในตอนที่ 2 ต่อไป

ตารางที่ 1 แสดง Random Number

10480	15011	01536	02011	81647	91646
22368	46573	25595	85393	30995	89198
24130	48360	22527	97265	76395	64809
42167	93093	06243	61680	07856	16376
37570	39975	81837	16656	06121	91782
77921	06907	11008	42751	27756	53498
99562	72905	56420	69994	98872	31016
96301	91977	05463	07972	18876	20922
89579	14342	63661	10281	17453	18103
85475	36857	53342	53988	53060	59533
28918	69578	88231	33267	70997	79936
63553	40961	48235	03427	49626	69445
09429	93969	52636	92737	88947	33488
10365	61129	85729	85689	48237	52267
07119	97336	71048	08178	77233	13916
51085	12765	51821	51259	77452	16308
02368	21382	52404	60268	89368	19885
01011	54092	33362	94904	31273	04146
52162	53916	46369	58586	23216	14513
07056	97628	33787	09998	42698	06691

แหล่งที่มา : Table of Fisher and Yates (1963)

## อ้างอิง

1. Armitage P. Statistical methods in medical research. Oxford : Blackwell Scientific Publications, 1971.
2. Bailey NTJ. Statistical methods in biology. New York : Halsted Press, 1981.
3. Schor S. Fundamental of biostatistics. New York : Putnum, 1968.
4. Snedecor GW, Cochran WG. Statistical methods. 6<sup>th</sup> ed. Ames, Iowa : Iowa State University press, 1967.
5. Cochran WG. Sampling techniques. 2<sup>nd</sup> ed. New York : Wiley, 1963.
6. Lilienfeld AM. Foundations of Epidemiology. 2<sup>nd</sup> ed. Oxford : Oxford University Press, 1980.
7. Mausner JS, Bahn AK. Epidemiology : An introductory text. Philadelphia : Saunders, 1974.
8. Morton RF, Hebel JR. A study guide to epidemiology & biostatistics. Baltimore : University park press, 1979.
9. Schlesselman JJ. Case-control studies. 1982.
10. Schlesselman JJ. Sample size requirement in cohort and case-control studies of diseases. Am J Hyg 1974; 99 : 381-84