

# อภิธานศัพท์สถิติสำหรับการอ่านบทความวิจัย (ตอนที่ 2)

อรุณ จิรวัฒน์กุล\*

จิราพร เขียวอยู่\*

แก้วใจ เทพสุธรรมรัตน์\*\*

เจตต์นภิศ ระยัยกุล\*

เชษฐา งามจรัส\*

มาลินี เหล่าไพบูลย์\*

ยุพา ถาวรพิทักษ์\*

พงษ์เดช สารการ\*

พอใจ พัทธินิตย์ธรรม\*

ศิริพร คำสะอาด\*

\*ภาควิชาชีวสถิติและประชากรศาสตร์ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น

\*\*หน่วยระบาดวิทยาคลินิก คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น

## Likelihood ratio

อัตราส่วนความน่าจะเป็น คำนวณจากผลหารของความน่าจะเป็น 2 จำนวน เช่น Likelihood ratio ของการทดสอบที่ให้ผลบวก จะเท่ากับอัตราส่วนของความน่าจะเป็นที่การทดสอบให้ผลบวกในกลุ่มคนที่เป็นโรค กับความน่าจะเป็นที่การทดสอบให้ผลบวกในกลุ่มคนที่ไม่เป็นโรค

## Log - linear model

เป็นตัวแทนที่ใช้ในการวิเคราะห์หาความสัมพันธ์ระหว่างตัวแปรตามที่เป็นตัวแปรกลุ่ม (categorical variable) และตัวแปรอิสระทั้งหมดหรือบางส่วนเป็นตัวแปรกลุ่ม

## Logistic regression

เป็นตัวแทนถดถอยที่มีตัวแปรผลวัดออกได้เป็นสองลักษณะ (dichotomous) และจากตัวแปรที่มีตัวแปรผลเป็นการเกิดโรค ตัวแปรอิสระเป็นปัจจัยต่าง ๆ จะสามารถคำนวณค่า OR จากค่าสัมประสิทธิ์ของปัจจัยในสมการ ค่า OR ที่ได้นี้เป็นค่า OR ที่มีผลมา

จากปัจจัยดังกล่าวอย่างเดี่ยว โดยจะปรับอิทธิพลจากปัจจัยของตัวแปรอื่นให้เท่า ๆ กันทั้งสองกลุ่ม จึงนิยมใช้ logistic regression ในการปรับอิทธิพลของตัวแปรกวน

## Logrank test

การทดสอบนี้ใช้ในการเปรียบเทียบอัตราการรอดเหตุการณ์ของสองกลุ่ม โดยใช้อัตราการเกิดเหตุการณ์ตามช่วงเวลาประมาณค่าจำนวนที่คาดว่าจะพบในแต่ละช่วงเวลาเปรียบเทียบกับค่าที่ได้จากการสังเกต ถ้าจำนวนที่คาดหวังและค่าที่สังเกตได้แตกต่างกัน แสดงว่าอัตราการรอดเหตุการณ์ไม่แตกต่างกัน

## Mann-Whitney U test

เป็นสถิติใช้ในการทดสอบความแตกต่างของค่ามัธยฐานของข้อมูลอันดับ การทดสอบไม่ต้องมีข้อตกลงเบื้องต้นของการแจกแจงของข้อมูล ในกรณีที่เป็นตัวแปรต่อเนื่องมีขนาดตัวอย่างเล็ก หรือการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้น จะแปลงค่าข้อมูลให้เป็นค่าอันดับแล้ว นำไปทดสอบด้วย Mann-Whitney U test

**Mantel Haenszel Odds Ratio ( $OR_{MH}$ )**

เป็นวิธีการวิเคราะห์ที่ใช้ในการคำนวณค่า  $OR$  รวมของทุกกลุ่มจาก ค่า  $OR$  ที่มาจากกลุ่มต่าง ๆ ในตัวแปรเดียวกัน หรือค่า  $OR$  ที่นำมาจากหลายการศึกษาเข้าด้วยกัน

**McNemar test**

การทดสอบสมมติฐานด้วยสถิติ McNemar ใช้ในการทดสอบความสัมพันธ์ระหว่างตัวแปรกลุ่มสองตัวที่มาจากการศึกษาแบบไม่อิสระต่อกัน เช่น ได้จากการศึกษาที่จับคู่ matched หรือ paired เพื่อลดอิทธิพลของปัจจัยกวน โดยข้อมูลแต่ละตัวแปรจะแบ่งออกเป็นสองลักษณะ ผลการศึกษา อยู่ในรูปตาราง  $2 \times 2$  ค่าสถิติ McNemar จะมีการแจกแจงแบบ Chi square ที่มีชั้นความเป็นอิสระ (degree of freedom) เท่ากับ 1 ดังนั้น เมื่อนำการแจกแจงดังกล่าวมาคำนวณค่า p-value สถิตินี้จึงถูกเรียกว่า McNemar Chi square

**Meta - analysis**

การวิเคราะห์เมตาดา เป็นวิธีการทางสถิติที่ใช้รวมผลลัพธ์จากหลาย ๆ รายงานวิจัยที่ทำการศึกษาในเรื่องเดียวกันเข้าด้วยกันเพื่อประมาณค่ารวมของผลลัพธ์ดังกล่าว เช่น การใช้วิธีของ Mantel-Haenszel รวม  $OR$  ของแต่ละการศึกษาให้ได้เป็น  $OR_{MH}$  รวมของทุกการศึกษา

**Mortality rate**

อัตราการตายของโรค คำนวณจากจำนวนการตายจากโรคที่กำหนดหารด้วยจำนวนคนที่มีโอกาสเป็นโรคเดียวกัน ในระยะเวลาที่กำหนด

**Multiple comparison**

ในการวิเคราะห์เพื่อดูความแตกต่างค่าเฉลี่ยของประชากรมากกว่า 2 กลุ่ม เมื่อพบว่ามีย่อยหนึ่งกลุ่มที่แตกต่างกัน จะทำการเปรียบเทียบดูว่ามีคู่ใดที่

แตกต่างกันบ้าง เรียกการเปรียบเทียบดังกล่าวว่า multiple comparison ในการดำเนินการเปรียบเทียบโดยใช้ t-test ค่า  $\alpha$  error จะเพิ่มขึ้นตามจำนวนครั้งของการเปรียบเทียบ จึงได้พัฒนาวิธีการทางสถิติสำหรับวิเคราะห์ multiple comparison ขึ้นหลายวิธี เช่น LSD, Bonferroni, HSD, Newnan - Keuls

**Multivariable analysis**

เป็นวิธีการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระมากกว่าหนึ่งตัว กับตัวแปรตามหนึ่งตัว เช่น หาค่าความสัมพันธ์ระหว่างตัวแปรตาม ความดันโลหิตกับตัวแปรอิสระอายุ เพศ และน้ำหนัก

**Multivariate analysis**

การวิเคราะห์ตัวแปรพหุ เป็นวิธีการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระหนึ่งตัวหรือมากกว่า กับตัวแปรตามมากกว่าหนึ่งตัว เช่น ความดันโลหิตสูงมีการแบ่ง 2 เกณฑ์ (A, B) ตัวแปรตามจะมี 2 ตัว  $BP_A$ ,  $BP_B$  คือ โดยนำไปหาความสัมพันธ์กับตัวแปรอิสระ อายุ เพศ และน้ำหนักตัว

**Negative predictive value**

ร้อยละของผู้ไม่เป็นโรคจากผู้ที่มีผลการทดสอบให้ผลลบทั้งหมด ซึ่งค่านี้อาจแปรตามความชุกของโรค

**Non-parametric test**

เป็นวิธีการทดสอบทางสถิติที่ไม่ต้องมีข้อตกลงเบื้องต้นของลักษณะการแจกแจงของข้อมูล ใช้ในการทดสอบสมมติฐานของข้อมูลอันดับ และการทดสอบสมมติฐานของตัวแปรต่อเนื่อง ในกลุ่มนี้จะใช้ในกรณีขนาดตัวอย่างเล็ก ๆ หรือข้อมูลมีการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้น หรือไม่สามารถแปลงค่าให้เป็นไปตามข้อตกลงเบื้องต้น โดยค่าข้อมูลจะถูกปรับเปลี่ยนเป็นค่าตำแหน่ง (rank) แล้วคำนวณค่าสถิติจากค่าตำแหน่ง

### Normal distribution

การแจกแจงปกติ คือ การแจกแจงของข้อมูลต่อเนื่องที่มีลักษณะสมมาตรรอบค่าเฉลี่ย มีรูปร่างของการแจกแจงคล้ายระฆังคว่ำ วิธีการทางสถิติส่วนใหญ่มีข้อตกลงเบื้องต้นว่า ข้อมูลมีการแจกแจงแบบปกติ ในการทดสอบสมมติฐานหรือการประมาณค่าพารามิเตอร์ของประชากรที่ใช้สถิติที่มีข้อตกลงเบื้องต้นดังกล่าว จึงควรมีการตรวจสอบว่าข้อมูลมีการแจกแจงปกติหรือไม่

### Number Needed to Treat (NNT)

คือ จำนวนผู้ป่วยที่แพทย์ต้องรักษาเพื่อจะช่วยให้ผู้ป่วยรักษาได้ผลดีเพิ่มขึ้นหนึ่งราย ค่าสถิตินี้ใช้พิจารณาขนาดของประโยชน์ที่ได้รับจากการรักษาเมื่อเทียบวิธีการรักษาที่สนใจกับวิธีการรักษาในกลุ่มควบคุม

### Odds

คือ อัตราส่วนของการได้รับปัจจัยเสี่ยงเทียบกับการไม่ได้รับของกลุ่มที่เป็นโรคหรือกลุ่มที่ไม่เป็นโรค ใช้แสดงว่าแต่ละกลุ่มมีผู้ได้รับปัจจัยเสี่ยงเป็นกี่เท่าของผู้ไม่ได้รับปัจจัยเสี่ยง

### Odds Ratio (OR)

คือ อัตราส่วนของค่า Odds ของกลุ่มที่เป็นโรคเทียบกับ Odds ของกลุ่มที่ไม่เป็นโรค ใช้แสดงว่ากลุ่มที่เป็นโรคได้รับปัจจัยเสี่ยงเป็นกี่เท่าของกลุ่มที่ไม่เป็นโรค

### One-sided test

การตั้งสมมติฐานทางสถิติที่มีสมมติฐานเลือก (alternative hypothesis) เป็นด้านน้อยกว่าหรือมากกว่าเพียงด้านเดียว

### Ordinal data

ข้อมูลอันดับ คือ ข้อมูลที่แบ่งได้เป็นกลุ่ม ๆ ด้วยชื่อ และสามารถจัดอันดับได้ แต่ไม่สามารถบอกขนาด

ความต่างระหว่างกลุ่มได้ เช่น ความพึงพอใจของผู้มารับบริการ พอใจน้อยต้องแก้ไข พอใจ พอใจมาก เป็นต้น

### Paired t-test

เป็นสถิติที่ใช้ในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของประชากรสองกลุ่มที่มีความสัมพันธ์กัน ความสัมพันธ์เกิดจากจับคู่ matched หรือ paired หรือ วัดผลการศึกษาจากหน่วยศึกษาหน่วยเดียวกันสองช่วงเวลา เช่น การศึกษาเปรียบเทียบค่าเฉลี่ยของผลต่างความดันโลหิตของผู้ป่วยก่อนและหลังที่ได้รับยา

### Parameter

พารามิเตอร์ หมายถึง ค่าสรุปเพื่อใช้บรรยายหรือแสดงลักษณะของประชากร เช่น ค่าเฉลี่ยของประชากร ( $\mu$ ) ค่าส่วนเบี่ยงเบนมาตรฐาน ( $\sigma$ ) ซึ่งจะต้องคำนวณมาจากทุกหน่วยที่อยู่ในประชากร ในทางปฏิบัติ จะใช้การประมาณค่าพารามิเตอร์จากค่าสถิติของตัวอย่าง

### Per protocol analysis

การวิเคราะห์ผลการทดลองทางคลินิกที่นำเฉพาะหน่วยศึกษาที่สามารถรักษาได้ตามแผนการรักษา และติดตามผลได้ครบตามที่ได้กำหนดไว้ในโครงร่างการวิจัย มาทำการวิเคราะห์สรุปผล

### Person year

เป็นวิธีการนับเวลาของหน่วยศึกษาที่เข้ามาในโครงการศึกษาไม่พร้อมกัน ณ เวลาประเมินผล หรือสิ้นสุดโครงการ เช่น เมื่อสิ้นสุดโครงการ หน่วยศึกษาคคนที่ 1 เข้ามาในการศึกษาได้ 2 ปี จะนับเป็น 2 person year คนที่ 2 อยู่ 1 ปี 6 เดือน จะนับเป็น 1.5 person year

### Planned analysis

เพื่อป้องกันอคติในการวิเคราะห์ผลการทดลองทาง

คลินิก ICH/GCP ได้กำหนดแนวปฏิบัติไว้ว่า การวิเคราะห์ผลการทดลองโดยเฉพาะคำถามหลักจะต้องมีการวางแผนการวิเคราะห์และเขียนไว้ในโครงร่างวิจัยในการวิเคราะห์ตามแผนดังกล่าวเรียกว่า planned analysis

### Point estimate

การประมาณแบบค่าเดียว ทำโดยนำค่าสถิติที่ได้จากตัวอย่างที่ศึกษาไปเป็นค่าพารามิเตอร์ของประชากร เช่น ประมาณค่าค่าเฉลี่ยน้ำหนักของนักศึกษา ทำการสุ่มตัวอย่างนักศึกษาจำนวน 50 คน คำนวนค่าเฉลี่ยได้ 58 กิโลกรัม จะใช้ค่าเฉลี่ยจากตัวอย่างที่ได้เป็นค่าประมาณของนักศึกษา (ทั้งหมด)

### Poisson distribution

เป็นการแจกแจงทางสถิติของจำนวนเหตุการณ์ที่เกิดขึ้นในช่วงเวลา (สั้น ๆ) การแจกแจงนี้นำไปใช้คำนวณความน่าจะเป็นของโรคที่พบได้น้อย ๆ

### Positive predictive value

ร้อยละของผู้ที่เป็นโรคจากผู้ที่มีผลการทดสอบเป็นบวกทั้งหมด ซึ่งค่านี้จะแปรตามความชุกของโรค

### Post hoc analysis

การวิเคราะห์ข้อมูล การทดลองทางคลินิกที่ไม่ได้กำหนดไว้ในแผนการวิเคราะห์ของโครงร่างวิจัย โดยปรกติวิธีการวิเคราะห์แบบนี้จะพิจารณาจากลักษณะของข้อมูลที่ได้จากการศึกษา จึงทำให้มีโอกาสเกิดอคติในการวิเคราะห์และสรุปผลได้มาก

### Post-test odds

เป็นค่า Odds ของการเป็นโรครายหลังที่ทราบผลการตรวจวินิจฉัยในกรณีที่ผลการตรวจวินิจฉัยเป็นบวก จะเป็นค่า Odds ของผู้ที่เป็นโรค และถ้าผลตรวจวินิจฉัยเป็นลบ จะเป็นค่า Odds ของผู้ที่ไม่เป็นโรค

### Post-test probability

โอกาสเป็นโรคของผู้ป่วย ภายหลังจากที่ทราบผลการตรวจวินิจฉัยแล้ว

### Power

อำนาจการทดสอบ (power หรือ statistical power) เป็นความน่าจะเป็นที่ค่าสถิติที่ใช้ทดสอบให้ผลการทดสอบว่าต่างเมื่อความจริงต่าง โดยที่ค่าอำนาจการทดสอบมีความสัมพันธ์โดยตรงกับขนาดตัวอย่าง ถ้าต้องการให้มีอำนาจการทดสอบมากก็ให้เพิ่มขนาดตัวอย่างให้ใหญ่ โดยทั่วไปนิยมให้การทดสอบมีอำนาจการทดสอบไม่ต่ำกว่าร้อยละ 80 ดังนั้น ในการคำนวณขนาดตัวอย่างจึงนิยมกำหนดค่าอำนาจการทดสอบร้อยละ 80

### Pre-test odds

ค่า Odds ของการเป็นโรคของผู้ที่เข้ามารับการตรวจวินิจฉัย Pre-Test Odds = Prevalence / (1-Prevalence)

### Pre-test probability

โอกาสเป็นโรคของผู้ป่วยที่เข้ามารับการตรวจวินิจฉัย หรือเรียกว่าความชุกของโรค (prevalence)

### Prevalence rate

อัตราความชุก เป็นอัตราแสดงความชุกของโรค คำนวนจากจำนวนผู้ป่วยด้วยโรคที่กำหนดทั้งหมดหารด้วยจำนวนคนทั้งหมดที่มีโอกาสป่วยเป็นโรคนี้ในเวลาเดียวกัน

### Proportion

ค่าสัดส่วนของลักษณะที่สนใจของ dichotomous data โดยคำนวณจากจำนวนหน่วยที่มีลักษณะที่สนใจหารด้วย จำนวนหน่วยศึกษาทั้งหมด เช่น สัดส่วนของเพศชาย จะเท่ากับจำนวนเพศชาย หารด้วยจำนวน

หน่วยตัวอย่างทั้งหมด ตัวแปรผลของการศึกษาทางการแพทย์และสาธารณสุขส่วนใหญ่วัดออกมาเป็น dichotomous data ดังนั้น จึงใช้ค่าสัดส่วนในการบรรยายลักษณะในตัวอย่างและสรุปผลในประชากร

### p-value

คือ ค่าความน่าจะเป็นที่ใช้ในการปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักเป็นจริง

### Random-effect factors

ดูรายละเอียดใน Fixed-effect factor

### Random-effect model

ดูรายละเอียดใน Fixed-effect model

### Regression analysis

การวิเคราะห์ด้วยตัวแบบถดถอย เป็นวิธีการทางสถิติใช้ในการสร้างตัวแบบสำหรับแสดงความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระหนึ่งหรือหลายตัว ตัวแบบที่สร้างขึ้นจะใช้สำหรับทำนายค่าตัวแปรตามจากตัวแปรอิสระที่กำหนด และยังสามารถใช้ประเมินอิทธิพลของตัวแปรอิสระแต่ละตัว ต่อตัวแปรตาม

### Relative Risk Reduction (RRR)

เป็นร้อยละของการลดลงของ risk ในกลุ่ม treatment เมื่อเทียบกับกลุ่มควบคุม โดยมีสูตรคำนวณค่า ดังนี้  $RRR = [(CER - EER)/CER] \times 100$  ความหมายของ EER และ CER ดูจาก Absolute risk reduction

### Repeated measures

การวัดผลการทดลองทางคลินิกในบางกรณีจะมีการวัดผลซ้ำในหน่วยศึกษาหน่วยเดียวกัน ตามช่วงเวลาต่าง ๆ เช่น ในกรณี pre and post measure หรือ วัดปริมาณยาในกระแสเลือดทุก 20 นาที การวัดซ้ำใน

คน ๆ เดียวกัน ทำให้ข้อมูลที่ได้จากการวัดแต่ละครั้งไม่เป็นอิสระต่อกัน ในการวิเคราะห์จะใช้สถิติต่างกับที่ ไม่มีการวัดซ้ำ

### Risk Ratio [หรือ Relative Risk (RR)]

อัตราส่วนความเสี่ยงระหว่างกลุ่มทดลอง กับกลุ่มควบคุม หรือระหว่างกลุ่มที่ได้รับกับกลุ่มที่ไม่ได้รับ ปัจจัยเสี่ยง ความหมายของค่า RR ในกรณีหาปัจจัยเสี่ยง ถ้าค่า RR เท่ากับ 1 แสดงว่า ไม่เป็น risk ถ้าค่า RR น้อยกว่า 1 แสดงว่าเป็น protective effect ถ้าค่า RR มากกว่า 1 แสดงว่าเป็น risk effect

### ROC curve (receiver operating characteristic)

เป็นกราฟที่แสดงสมรรถนะของการตรวจวินิจฉัย โดยมีแกน Y เป็นค่าความไว (sensitivity) ส่วนแกน X เป็นผลบวกลวง (false-positive rate) การทดสอบที่มีความไวสูงและผลบวกลวงต่ำ จะเป็นการทดสอบที่ดี ซึ่งลักษณะโค้งจะอยู่ชิดมุมซ้ายด้านบน จะใช้ ROC curve ช่วยพิจารณาเลือกจุดตัดสินใจของการวินิจฉัยที่ให้ค่าความไวสูงและมีขนาดของผลบวกลวงที่ยอมรับได้

### Sensitivity analysis

เป็นวิธีการวิเคราะห์เพื่อจะดูว่าถ้าเงื่อนไขต่าง ๆ เปลี่ยนแปลงจากค่าที่พบจากการศึกษา ผลการวิเคราะห์ที่ได้จะมีการเปลี่ยนแปลงอย่างไร

### Sensitivity

ความไวเป็นคุณสมบัติของการทดสอบที่แสดงว่า ในกลุ่มที่มีผลการทดสอบเป็นบวกมีโอกาสเป็นโรคเท่าไร

### Shapiro-Wilk test (W)

สถิติที่ใช้ทดสอบการแจกแจงของตัวแปรที่จะนำมาทดสอบสมมติฐานหรือหาช่วงเชื่อมั่นว่าประชากรจะมีการแจกแจงแบบปกติหรือไม่

**Skewness**

ความเบ้ คือ การที่ข้อมูลมีการแจกแจงอยู่ด้านใดด้านหนึ่งของโค้งความถี่มากกว่าอีกด้านหนึ่ง วัดความเบ้โดยค่าสถิติ “สัมประสิทธิ์ความเบ้”

**Spearman rank correlation**

สถิติที่ใช้หาค่าสหสัมพันธ์ระหว่างตัวแปรอันดับค่าสัมประสิทธิ์สหสัมพันธ์ของ Spearman มีค่าระหว่าง -1 และ +1 (0 ถึง 1 หรือ 0 ถึง -1) เช่นเดียวกับ Pearson's correlation coefficient ในกรณีตัวแปรต่อเนื่องที่มีขนาดตัวอย่างเล็ก หรือมีลักษณะการแจกแจงไม่เป็นไปตามข้อตกลงเบื้องต้น การหาความสัมพันธ์จะแปลงค่าข้อมูลเป็นค่าอันดับ และคำนวณค่าสหสัมพันธ์ด้วยวิธีของ Spearman

**Specificity**

ความจำเพาะเป็นคุณสมบัติของการทดสอบที่แสดงว่าในกลุ่มที่มีผลการทดสอบเป็นลบ มีโอกาสไม่เป็นโรคเท่าไร

**Standard deviation**

ส่วนเบี่ยงเบนมาตรฐานใช้วัดการกระจายของข้อมูล ซึ่งเป็นค่าที่บอกว่าโดยเฉลี่ยแล้วค่าของข้อมูลแต่ละตัวห่างจากค่าเฉลี่ยของชุดข้อมูลนั้นเท่าไร ถ้าตัวแปรที่ศึกษามีส่วนเบี่ยงเบนมาตรฐานใหญ่ แสดงว่าข้อมูลของแต่ละหน่วยศึกษามีความแตกต่างกันมาก

**Standard error**

ความคลาดเคลื่อนมาตรฐาน คือ ค่าส่วนเบี่ยงเบนมาตรฐานของการแจกแจงค่าสถิติของตัวอย่าง โดยคำนวณจากค่าเบี่ยงเบนระหว่างค่าเฉลี่ยตัวอย่างแต่ละตัวกับค่าเฉลี่ยประชากร ค่าความคลาดเคลื่อนมาตรฐานจะใช้ในการอนุมานค่าพารามิเตอร์ของประชากร

**Standardized mortality rate**

อัตราการตายปรับฐานใช้สำหรับการเปรียบเทียบการตายระหว่างพื้นที่ เนื่องจากอัตราการตายของแต่ละพื้นที่มีโครงสร้างอายุของประชากรต่างกัน อัตราการตายจึงไม่สามารถเปรียบเทียบกันได้โดยตรง จะต้องนำอัตราการตายตามกลุ่มอายุของทั้งสองพื้นที่มาคำนวณจำนวนการตายจากประชากรฐานเดียวกัน อัตราการตายที่ปรับฐานนี้เรียกว่า standardized mortality rate

**Statistical model**

ตัวแบบทางสถิติ เป็นตัวแบบที่สร้างขึ้นเพื่อการทำนายผล โดยมีตัวแปรตามหนึ่งตัว และตัวแปรอิสระหนึ่งหรือหลายตัว ลักษณะของตัวแบบจะแตกต่างกันไปตามประเภทข้อมูล และความเป็นอิสระของตัวแปรที่อยู่ในตัวแบบ เช่น Regression model, Logistic regression model, Log-linear model, GEE, Cox proportional Hazard model

**Statistical power**

ดูรายละเอียดใน Power

**Statistical significance**

ในการทดสอบสมมติฐาน ถ้าค่าสถิติที่ใช้ในการทดสอบให้ค่า p value น้อยกว่าค่าที่กำหนด จะสรุปว่าการทดสอบมีความแตกต่างอย่างมีนัยสำคัญ ซึ่งหมายถึง ในประชากรจะมีความแตกต่างกันด้วย เรียกลักษณะที่พบนี้ว่ามี statistical significance ในทางตรงกันข้าม ถ้า p value มากกว่าค่าที่กำหนด แสดงว่าแตกต่างอย่างไม่มีนัยสำคัญ สรุปว่า statistical non-significance

**Sub-group analysis**

เป็นการวิเคราะห์เพื่อดูผลการทดลองของแต่ละ

กลุ่มย่อย เช่น ตามความหนักเบาของอาการ เพศ อายุ เป็นต้น ในการวิเคราะห์ในกลุ่มย่อยมักจะมีปัญหาขนาดตัวอย่างไม่มากพอที่จะบอกความต่าง ตามปกติ จะทำการวิเคราะห์ในกลุ่มย่อย เพื่อสามารถแสดง Interaction ของตัวแปรที่ใช้แบ่งกลุ่ม

### Survival analysis

การวิเคราะห์ความน่าจะเป็นของการปลอดเหตุการณ์ที่สนใจตามช่วงเวลาตัวอย่างเหตุการณ์ที่สนใจ เช่น การตาย การกลับมาเป็นซ้ำ

### Transformation of data

การแปลงหน่วยข้อมูลเพื่อลดขนาดของความแปรปรวน เช่น แปลงค่าคะแนน X เป็นค่า  $\log x$  การแปลงค่าข้อมูลเพื่อให้ข้อมูลมีการแจกแจงเป็นไปตามข้อตกลงเบื้องต้นของการใช้สถิติเพื่ออนุมานผล ในกรณีการประมาณค่าเมื่อได้ช่วงเชื่อมั่นของข้อมูลแปลงแล้ว ก็สามารถแปลงค่าช่วงเชื่อมั่นกลับมาเป็นหน่วยวัดเดิม สำหรับจากการทดสอบสมมติฐาน จะใช้ p value ที่คำนวณจากข้อมูลแปลง สรุปผลการทดสอบ

### t-test

การทดสอบสมมติฐานด้วยสถิติ t จะใช้ในการเปรียบเทียบค่าเฉลี่ยประชากรของข้อมูลต่อเนื่อง จะคำนวณค่า p-value ของการทดสอบด้วยการแจกแจง t

การใช้สถิติ t ในการทดสอบจะต้องมีข้อตกลงเบื้องต้นว่าตัวแปรที่จะทดสอบประชากรต้องมีการแจกแจงปกติ การทดสอบด้วยสถิติ t ในกรณีประชากรกลุ่มเดียวจะเป็นการเปรียบเทียบความต่างของค่าเฉลี่ยกับค่าคงที่ ๆ กำหนด ในการเปรียบเทียบค่าเฉลี่ยของประชากรสองกลุ่มจะมีสองกรณี คือ ประชากรสองกลุ่มเป็นอิสระต่อกันจะมีวิธีการคำนวณค่า t แบบ Independent t-test ถ้าประชากรสองกลุ่มไม่เป็นอิสระต่อกันจะคำนวณค่า t แบบ Paired t-test

### Unit of analysis

หน่วยวิเคราะห์โดยปกติจะเป็นหน่วยเดียวกันกับหน่วยศึกษา เช่น ผู้ป่วย บุคคล แต่ในบางครั้งอาจเป็นกลุ่มของหน่วยศึกษา เช่น โรงเรียน หมู่บ้าน หรือเป็นบางส่วนของคน เช่น พัน ตา

### Weighted mean

การหาค่าเฉลี่ยในกรณีที่ให้ข้อมูลแต่ละกลุ่มมีความสำคัญไม่เท่ากัน โดยจะคิดค่าน้ำหนักให้ตามความสำคัญของกลุ่ม ในการคำนวณค่าเฉลี่ยจะนำน้ำหนักมาคิดร่วมด้วย เช่น การคิดค่าคะแนนเฉลี่ยสะสม (GPA.) จะใช้หน่วยกิตของแต่ละวิชาเป็นน้ำหนักคูณกับหน่วยกิตที่ได้แล้วจึงหาค่าเฉลี่ยโดยหารด้วยหน่วยกิตทั้งหมด