

**ข้อ 1 บรรทัดที่ 1 : ความหมายของข้อมูลตัวแปรแต่ละประเภท พร้อมตัวอย่าง 5 ลักษณะ**

Continuous หมายถึง ข้อมูลที่มีค่าต่างๆ ทุกค่าต่อเนื่องกัน โดยแสดงได้ทั้งเศษส่วนหรือตัวเลขที่เป็นจำนวนเต็ม เช่น อายุ ส่วนสูง น้ำหนัก รายได้ หนี้สิน โดยการพรรณนาลักษณะของข้อมูล เป็นค่าสรุปเพียงค่าเดียว ใช้กรณีตัวแปรต่อเนื่อง ประกอบด้วย ค่าเฉลี่ย (mean) ค่ามัธยฐาน (median) ค่าฐานนิยม (mode)

Categorical หมายถึง ข้อมูลที่มีค่าต่างๆ แยกจากกันโดยเด็ดขาด เช่น เพศ อาชีพ หมู่เลือด สิทธิการรักษา โรคประจำตัว นำเสนอเป็น การวัดจำนวนนับ หรือความถี่ หรือสรุปลักษณะข้อมูล โดยใช้ ความถี่ ร้อยละ หรือค่าสัดส่วน

Count หมายถึง ข้อมูลจำนวนนับ ในช่วงเวลาหรือขอบเขตที่สนใจศึกษา ข้อมูลจะอยู่ในลักษณะ อัตรา ต่อ สิ่งที่สนใจ จะเป็นการแจกแจงแบบ Poisson หรือ การทำนายจำนวนของเหตุการณ์ในเวลาที่ระบุ เช่น จำนวนผู้มารับบริการใน รพ. แต่ละเดือน จำนวนป่วยในรอบหนึ่งปี จำนวนป่วยโรคเบาหวานในที่มารับการรักษาในแต่ละเดือน จำนวนอุบัติเหตุที่เกิดขึ้นในรอบหนึ่งปี จำนวนคนตายในพื้นที่ในรอบหนึ่งปี

Survival หมายถึง การวิเคราะห์การอยู่รอดหรือการวิเคราะห์ Time to event คือการนำข้อมูลในช่วงเวลาที่มีอยู่มาใช้ นำไปสู่การแก้ปัญหา การวิเคราะห์การอยู่รอด หรือ Cox Proportional Hazard หรือ Cox Regression ใช้ในการศึกษาความสัมพันธ์ระหว่างตัวแปรตามที่สนใจ กับตัวแปรอิสระต่างๆ ตั้งแต่หนึ่งตัวขึ้นไปโดยเป็นข้อมูลที่มีช่วงเวลาเข้ามาเกี่ยวข้องและมีลักษณะเป็นข้อมูลไม่สมบูรณ์ ข้อมูลที่มีลักษณะเป็น Censored data ซึ่งหมายถึง การมีข้อมูลที่ไม่ทราบค่าสังเกตที่แท้จริงว่าเกิดเหตุการณ์ใดขึ้น ก่อนสิ้นสุดการติดตาม เช่น ผลการรักษาผู้ป่วยเบาวานในระยะเวลาหนึ่งปี ผลการบำบัดผู้ติดสารเสพติดในระยะเวลาหนึ่งปี

**ข้อ 2 บรรทัดที่** **2 : หมายถึงอะไรพร้อมยกตัวอย่างคำถามวิจัย พร้อมยกตัวอย่างที่ใช้วิเคราะห์ เพื่อให้ได้ตามบรรทัดที่ 2**

**2.1 Mean** หมายถึง ค่าเฉลี่ยซึ่งเกิดจากข้อมูลของผลรวมทั้งหมดหารด้วยจำนวนรายการของข้อมูล

**คำถามวิจัย** เช่น จงหาน้ำหนักโดยเฉลี่ยรายได้ของหัวหน้าครอบครัว stat village

* + สถิติที่ใช้ในการวิเคราะห์ ได้แก่ ค่าเฉลี่ย ความถี่ ร้อยละ
  + Mean (SD)
  + Median (Min:Max)
  + Correlation coefficient: r and ICC

**2.2 median** หมายถึงมัธยฐาน (median) คือ ค่ากลางของข้อมูลที่อยู่กึ่งกลางของข้อมูลทั้งหมดหลังจากเรียงลำดับข้อมูลจากน้อยไปมากหรือจากมากไปน้อย

ตัวอย่าง จงหาค่ามัธยฐานของข้อมูล 3 , 7 19, 25, 12, 18 , 10

วิธีทำ เรียงข้อมูลจากน้อยไปมากได้ 3 , 7, 10, 12, 18, 19, 25

ข้อมูลมีทั้งหมด 7 ตัวเรียงข้อมูลแล้วตัวเลขที่อยู่ตรงกลางคือตัวเลขตำแหน่งที่ 4 ∴ตัวเลขตำแหน่งที่ 4 คือ 12 เป็นมัธยฐาน

**คำถามวิจัย** เช่น จงหาค่า median ของน้ำหนักหัวหน้าครอบครัวในหมู่บ้าน Stat Village

สถิติที่ใช้ในการวิเคราะห์ ได้แก่ ค่ามัธยฐาน

**2.3 Proportion** หมายถึงการเปรียบเทียบของสิ่งย่อย ในสิ่งที่ใหญ่กว่านั่นคือส่วนที่เป็นตัวตั้ง จะต้องอยู่เป็นส่วนหนึ่งของส่วนที่เป็นตัวหาร โดยมากแล้วการเปรียบเทียบ Proportion จะอยู่ในรูปของ{จำนวนหรือคนที่มีลักษณะบางอย่างที่เราสนใจ\overจำนวนหรือคนทั้งหมด}\times10^n$$  
โดย n จะเป็นเท่าไหร่ก็แล้วแต่ว่าเราต้องการเป็นแบบไหน (เช่นต่อร้อยคน ต่อแสนคน) โดยมากเรามักจะนิยม % ดังนั้น n จึงเป็น 2 (คูณ $10^2$ หรือ 100 นั่นเอง)  
จะสังเกตว่าการวัดแบบ Proportion นี้ก็เป็น Ratio แบบหนึ่งด้วย คือมีการเปรียบเทียบกับของสองสิ่งซึ่งอาจจะไม่เกี่ยวข้องกัน (แต่ในที่นี้ Proportion เกี่ยวข้องกัน) แต่ไม่ใช่ว่าทุก Ratio จะเป็น Proportion (เช่น Ratio ผู้ชาย:ผู้หญิง ไม่เป็น Proportion, แต่ % ของผู้ชายเป็น Proportion)  
เช่นเราทราบ Proportion ผู้ชาย = 45%  
ดังนั้น Proportion ผู้หญิง = 100% - 45% = 55%

\*Magnitude of effect สามารถใช้เป็น

* Difference ->เอาสองสัดส่วนมาลบกัน มักย่อเป็น **RD** = Risk difference
* Ratio -> เอาสองสัดส่วนมาหารกัน มักย่อเป็น **RR** = Risk Ratio หรือ Relative risk แต่บางกรณี หา Risk ไม่ได้ เช่นกรณี Case-control study จึงไปใช้ **OR** (Odds ratio) แทน

**คำถามวิจัย** เช่น อัตราการป่วยของเพศชายแตกต่างจากเพศหญิงหรือไม่

สถิติที่ใช้ในการวิเคราะห์ ได้แก่ OR RR , 95 % CI และ p - value

**2.4 rate per “ space”** หมายถึง อัตราเทียบกับระยะเวลา

**2.5 median survival** หมายถึง ค่ามัธยฐานของระยะปลอดเหตุการณ์

**2.6 RISK of events at** หมายถึง ปัจจัยที่ส่งผลกระทบต่อตัวแปรตามในช่วงเวลาใดเวลาหนึ่ง

**ข้อ 3.บรรทัดที่ 3 : หมายถึงอะไร พร้อมยกตัวอย่างคำถามวิจัย และสถิติที่ใช้ในการวิเคราะห์ เพื่อให้ได้ตาม**

3.1 Linear Reg

**- Simple linear regression** หมายถึง เป็นเทคนิคการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร 2 ตัว ซึ่งประกอบด้วยตัวแปรตาม 1 ตัว กับ ตัวแปรอิสระ 1 ตัว ที่มีความสัมพันธ์อยู่ในรูปเชิงเส้น โดยตัวแปรทั้ง 2 เป็นตัวแปรเชิงปริมาณ

**คำถามงานวิจัย** ได้แก่ อัตราป่วยชายต่างจากหญิงหรือไม่

สถิติที่ใช้ในการวิเคราะห์ t- test Mean difference

* **Multiple linear regression** หมายถึง เป็นเทคนิคการวิเคราะห์ความสัมพันธ์ของตัวแปรหลายตัว ซึ่งประกอบด้วยตัวแปรตาม 1 ตัว ซึ่งเป็นตัวแปรเชิงปริมาณ สเกลแบบช่วง หรือสเกลอัตราส่วน และ ตัวแปรต้น/ตัวแปรอิสระหลายตัว ซึ่งเป็นตัวแปรเชิงปริมาณ หรือตัวแปรเชิงกลุ่ม/เชิงคุณภาพ

**คำถามงานวิจัย** ได้แก่ การป่วยต่างกันหรือไม่ระหว่างกลุ่มสวัสดิการ

สถิติที่ใช้ในการวิเคราะห์ Mean difference binreg v8 i.v9 , rd/\*Bivariate analysis\*/

binreg v8 i.v9 age bmi, rd /\*Multivariable analysis\*/

**3.2 Logitis Reg.** หมายถึง เป็นการวิเคราะห์ถดถอยที่ตัวแปรตามเป็นตัวแปรเชิงกลุ่ม (categorical data)

**คำถามวิจัย** กลุ่มสวัสดิการ สัมพันธ์ กับ น้ำหนัก หรือไม่

สถิติที่ใช้ ในการวิเคราะห์ ได้แก่

**3.3 Possion Reg.** หมายถึง เป็นการศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปร ซึ่งมีความสัมพันธ์เป็นตัวแปรแบบการถดถอยปัวซง (Poisson regression modal) เพื่อใช้ในการทำนายหรือพยากรณ์ค่าของตัวแปรที่ศึกษา

**คำถามวิจัย** ได้แก่ จำนวนครั้งของการเจ็บป่วยของหัวหน้าครอบครัวในหมู่บ้าน stat village ในปีที่ผ่านมา

สถิติที่ใช้ในการวิเคราะห์ ได้แก่ Incidence rate ratio (IRR)

**3.4 cox Reg.** หมายถึง จะต้องมีตัวแปรตามเป็นความเสี่ยงต่อการเกิดเหตุการณ์ ณ เวลา t