



สาขาวิชาวิทยาศาสตร์สุขภาพ
มหาวิทยาลัยสุโขทัยธรรมศาสตร์

การสอนเสริมครั้งที่ 2

เอกสารโสตทัศนศึกษา

50103

สถิติและการวิจัย
สำหรับ
วิทยาศาสตร์สุขภาพ

Statistics and Research in Health Sciences

สงวนลิขสิทธิ์

เอกสารสไลด์ทัศนศึกษา สถิติและการวิจัยสำหรับวิทยาศาสตร์สุขภาพ การสอนเสริมครั้งที่ 2
จัดทำขึ้นเพื่อเป็นบริการแก่นักศึกษาในการสอนเสริม

จัดทำต้นฉบับ : คณะกรรมการกลุ่มผลิตชุดวิชา

บรรณาธิการ/ออกแบบ : หน่วยผลิตสื่อสอนเสริม ศูนย์สไลด์ทัศนศึกษา
สำนักเทคโนโลยีการศึกษา

จัดพิมพ์ : สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช

พิมพ์ที่ : โรงพิมพ์มหาวิทยาลัยสุโขทัยธรรมมาธิราช

พิมพ์ครั้งที่ 28 ภาค 2/2557 (ปรับปรุง)

ไสตท์ศน์ # 2.6.1 ลักษณะและแหล่งที่มาของข้อมูลเพื่อการวิจัย

ลักษณะของข้อมูล

1. ข้อมูลเชิงคุณภาพ
2. ข้อมูลเชิงปริมาณ

ระดับข้อมูล

1. ข้อมูลระดับนามบัญญัติ (nominal scale)
2. ข้อมูลระดับอันดับ (ordinal scale)
3. ข้อมูลระดับช่วงชั้นหรืออันตรภาค (interval scale)
4. ข้อมูลระดับอัตราส่วน (ratio scale)

ลักษณะของข้อมูลที่ตี

1. ความถูกต้อง
2. ความทันสมัย
3. ความสมบูรณ์ครบถ้วน
4. ความชัดเจน
5. ความสอดคล้อง
6. เปรียบเทียบกับข้อมูลอื่นได้

แหล่งที่มาของข้อมูล

1. ข้อมูลปฐมภูมิ
2. ข้อมูลทุติยภูมิ

ไสตท์ศน์ # 2.6.2 วิธีการเก็บรวบรวมข้อมูล

วิธีการเก็บรวบรวมข้อมูล

1. การเก็บรวบรวมข้อมูลจากเอกสาร
2. วิธีการบันทึกข้อมูลจากการวัดหรือนับ
3. การใช้แบบสอบถาม
4. การทดลอง
5. การสัมภาษณ์
6. การทดสอบ
7. การสังเกต

สไลด์ทัศน์ # 2.6.3 เทคนิคการเก็บรวบรวมข้อมูล

เทคนิคในการเก็บรวบรวมข้อมูล

1. การวางแผนการเก็บข้อมูล
2. วิธีการเก็บรวบรวมข้อมูล
3. เครื่องมือที่ใช้เพื่อเก็บรวบรวมข้อมูล

เทคนิคของการเก็บข้อมูลด้วยวิธีสังเกต

1. ศึกษาวัตถุประสงค์ของการวิจัย
2. สร้างแบบแผนการเก็บตัวอย่าง
3. อบรมผู้ที่จะดำเนินการเก็บรวบรวมข้อมูลให้เข้าใจถึงวัตถุประสงค์ของงานวิจัยและขั้นตอนการสังเกตข้อมูล
4. เก็บข้อมูลอย่างเป็นระบบ ตามวัตถุประสงค์ของงานวิจัย
5. ตรวจสอบข้อมูลและประมวลผลเบื้องต้น

เทคนิคของการเก็บข้อมูลด้วยวิธีแจกแบบสอบถาม

1. การแจกแบบสอบถามโดยตรง
2. การส่งแบบสอบถามให้ผู้ตอบทางไปรษณีย์

เทคนิคของการเก็บข้อมูลด้วยการสัมภาษณ์

1. การสัมภาษณ์จากผู้ให้คำตอบโดยตรง
2. การสัมภาษณ์ทางโทรศัพท์

สไลด์ทัศน์ # 2.7.1 การเตรียมและการตรวจสอบแก้ไขข้อมูล

1. การเตรียมข้อมูล (data preparation)

มีขั้นตอนดังนี้

1. การตรวจสอบแก้ไขข้อมูลเบื้องต้น
2. การให้รหัสข้อมูล
3. การบันทึกข้อมูล
4. การตรวจสอบแก้ไขรหัสข้อมูลครั้งสุดท้าย

2. การตรวจสอบแก้ไขข้อมูลเบื้องต้น

1. ความครบถ้วนสมบูรณ์
2. ความสอดคล้องของคำตอบ
3. ความสอดคล้องกับวัตถุประสงค์

โลตทัศน์ # 2.7.2 ตารางจำลอง

ตารางจำลอง (dummy table)

ช่วยเป็นกรอบในการวิเคราะห์ รวมทั้งการแปลผล ทำให้การศึกษามีข้อผิดพลาดน้อย ประหยัดค่าใช้จ่ายในการวิเคราะห์ด้วยคอมพิวเตอร์

องค์ประกอบของตาราง

1. หมายเลขของตาราง เป็นการระบุลำดับที่ของตาราง กำหนดให้อยู่ซ้ายมือ
2. ชื่อของตาราง เป็นข้อความที่อยู่ถัดจากหมายเลขของตารางและอยู่แถวเดียวกัน เป็นข้อความสั้นๆ เพื่อให้ทราบข้อมูลในตารางเป็นเรื่องอะไร ที่ไหน และเมื่อใด
3. หัวเรื่องแถว เป็นส่วนของตารางที่อธิบายข้อมูลที่อยู่ในแนวนอน
4. หัวเรื่องสดมภ์ เป็นส่วนของตารางที่อธิบายข้อมูลที่อยู่ในแนวตั้ง
5. ตัวเรื่อง เป็นส่วนของตารางที่เป็นตัวเลข และตัวเลขนี้จะจำแนกตามรายละเอียด ของหัวเรื่องแถว และหัวเรื่องสดมภ์

ตัวอย่าง

หมายเลขของตาราง

ตารางที่ 1 จำนวนและร้อยละของการทำงานเพื่อค่าจ้าง หลังจากสมรสของสตรีที่เคยสมรสแล้ว

จำแนกตามระดับการศึกษา → ชื่อของตาราง

ระดับการศึกษา	สตรีที่เคยสมรสแล้ว		← หัวเรื่องสดมภ์
	จำนวน	ร้อยละ	
ประถมศึกษา	30	30.0	← ตัวเรื่อง
มัธยมศึกษา	30	30.0	
อนุปริญญา	20	20.0	
ปริญญาตรีหรือสูงกว่า	20	20.0	
รวม	100	100.0	

← หัวเรื่องแถว

ถ้าบอกสถานที่และระยะเวลาด้วยยิ่งดี แต่ถ้าเป็นที่เข้าใจกันในเล่มเดียวกันก็ไม่ต้องเขียน

ชนิดของตาราง

1. ตารางอย่างง่าย (univariate tabulation) เป็นตารางที่มีตัวแปรเดียว ในหัวตารางที่เป็นแถวตั้งจะมีเพียงจำนวน และร้อยละ มักจะนำเสนอคุณลักษณะของตัวอย่าง ไม่ได้แสดงความสัมพันธ์ระหว่างตัวแปร

ตารางที่ 2 จำนวนและร้อยละของผู้ตอบแบบสอบถามเรื่องการวางแผนครอบครัว จำแนกตามเพศและอายุ

ตัวแปร	จำนวน	ร้อยละ
เพศ		
ชาย	300	37.5
หญิง	500	62.5
รวม	800	100.0
กลุ่มอายุ (ปี)		
20-29	200	25.0
30-49	250	31.3
50 ขึ้นไป	350	43.7
รวม	800	100.0

2. ตาราง 2 ตัวแปร (bivariate tabulation) เป็นตารางที่มี 2 ตัวแปร ในหัวตารางที่เป็นแถวนอนและแถวตั้งจะเป็นตัวแปรคนละพวก เช่น ตัวแปรอิสระ หรือตัวแปรตาม เป็นต้น มักเป็นการวิเคราะห์แบบตารางไขว้การที่จะจัดตัวแปรตามหรือตัวแปรต้นเป็นแถวตั้งหรือแถวนอนแล้วแต่ความเหมาะสมของหน้ากระดาษ

ตาราง 3 จำนวนของสิ่งตัวอย่างจำแนกตามการเจ็บป่วยด้วยโรคพยาธิใบไม้ตับและประวัติการกินปลาดิบหรือดิบๆ สุกๆ

ประวัติการกินปลาดิบหรือดิบๆ สุกๆ	การเจ็บป่วยด้วยโรคพยาธิใบไม้ตับ		รวม จำนวน	รวม ร้อยละ
	ป่วย	ไม่ป่วย		
เคยกิน	80	20	100	100.0
ไม่เคยกิน	0	100	100	100.0
รวม	80	120	200	100.0

การคำนวณร้อยละใช้ตัวแปรต้นเป็นฐาน 100 ในการคำนวณ แต่แปลผลให้แปลผลตามตัวแปรตาม

ไสตท์ศน์ # 2.7.3 ชนิดของตาราง (ต่อ)

3. ตารางที่มีตัวแปรมากกว่า 2 ตัวแปร (multivariate tabulation) เป็นตารางที่มีตัวแปรมากกว่า 2 ตัว ในหัวตารางที่เป็นแนวนอนอาจแสดงตัวแปร 2 ตัว หรือมากกว่า หรือในแถวตั้งอาจแสดงตัวแปร 2 ตัว หรือมากกว่าก็ได้

ตาราง 4 ความสัมพันธ์ระหว่างอาชีพ ระดับการศึกษา และความพึงพอใจต่อการรับบริการในโรงพยาบาลของผู้มารับบริการ

การศึกษา	เกษตรกร		ข้าราชการ		รวม
	พอใจ	ไม่พอใจ	พอใจ	ไม่พอใจ	
ต่ำกว่าปริญญาตรี	80	20	60	40	200
ปริญญาตรีหรือสูงกว่า	20	80	40	60	200
รวม	100	100	100	100	400

ไสตท์ศน์ # 2.7.4 การบันทึกข้อมูลด้วยคอมพิวเตอร์

1. ความสำคัญของการบันทึกข้อมูลด้วยคอมพิวเตอร์

การใช้คอมพิวเตอร์ในการวิเคราะห์ข้อมูล จำเป็นต้องบันทึกข้อมูลด้วยคอมพิวเตอร์เช่นเดียวกันเพื่อคอมพิวเตอร์จะได้อ่านข้อมูลได้ วิเคราะห์ได้ และส่งข้อมูลผ่านออกมาได้ การที่บันทึกข้อมูลด้วยคอมพิวเตอร์สามารถเรียกขึ้นมาดูบนจอได้ โดยใช้โปรแกรมที่ใช้บันทึกข้อมูล หรือโปรแกรมที่จะวิเคราะห์ข้อมูล สามารถบันทึกข้อมูลเพิ่มเติมเป็นระยะได้ ถ้าหากพบข้อผิดพลาดก็สามารถแก้ไขใหม่ได้ตลอดเวลา หรืออาจมีการแทรกตัวแปรที่ไปเก็บข้อมูลเพิ่มเติม หรือแทรกชุดของข้อมูลก็ได้ ซึ่งถือว่าการบันทึกข้อมูลด้วยคอมพิวเตอร์อำนวยความสะดวกให้แก่กวิจัยเป็นอย่างดี

2. สื่อที่ใช้ในการบันทึกข้อมูลด้วยคอมพิวเตอร์

เนื่องจากคอมพิวเตอร์ไม่สามารถอ่านข้อมูลจากเอกสารต่างๆ ได้โดยตรง จึงต้องมีสื่อที่ให้เครื่องคอมพิวเตอร์อ่านได้ สื่อสำหรับคอมพิวเตอร์ได้แก่

- แผ่นซีดี
- คอมแพคแฟลช
- การ์ดมัลติมีเดียและซีเคียวร์ติจิตอล
- สื่ออื่นๆ เช่น removable disk

ไสตท์ศน์ # 2.7.4 การบันทึกข้อมูลด้วยคอมพิวเตอร์ (ต่อ)

3. วิธีการบันทึกข้อมูล

หลังจากรวบรวมข้อมูล ผ่านกระบวนการตรวจสอบข้อมูลเบื้องต้น ได้มีการสร้างคู่มือลงรหัส และลงรหัสในแบบฟอร์มที่สร้างขึ้นแล้ว และทำความเข้าใจกับคอมพิวเตอร์ อุปกรณ์หรือสื่อต่างๆ ที่ใช้กับคอมพิวเตอร์ ทราบโปรแกรมคอมพิวเตอร์ที่จะเลือกใช้พร้อมกับมีโปรแกรมอื่นๆ ในเครื่องคอมพิวเตอร์ ก็สามารถที่จะบันทึกหรือคีย์ข้อมูลเข้าคอมพิวเตอร์ได้ ทั้งนี้ในการบันทึกข้อมูลสำหรับงานวิจัยเรื่องหนึ่งควรใช้โปรแกรมคอมพิวเตอร์ในการบันทึกเพียงโปรแกรมเดียว ไม่ควรใช้หลายโปรแกรม หรือแบ่งกันบันทึกข้อมูลก็ควรต้องเป็นโปรแกรมเดียวกัน

4. การตรวจสอบข้อมูลขั้นสุดท้าย (final editing)

ประโยชน์จากการใช้คอมพิวเตอร์ช่วยในการตรวจสอบข้อมูลขั้นสุดท้าย ก่อนทำการวิเคราะห์ คือ

1. ตรวจรหัสข้อมูลจากแผ่นพิมพ์
2. ตรวจรหัสข้อมูลโดยการพิมพ์ซ้ำ
3. ตรวจรหัสข้อมูลโดยดูตัวเลขที่เป็นไปไม่ได้

การตรวจสอบแก้ไขข้อมูลครั้งสุดท้ายนี้เป็นสิ่งจำเป็น เพราะถ้าข้อมูลไม่ชัดเจนมีข้อผิดพลาดเกิดขึ้น จะทำให้การวิเคราะห์ข้อมูลสะดุดเป็นระยะ ต้องเสียเวลาไปค้นหาข้อมูลที่ถูกต้อง หรือเมื่อวิเคราะห์ข้อมูลไปแล้วเพิ่งมาพบข้อผิดพลาดหรือความเป็นไปไม่ได้ ก็ต้องย้อนกลับมาวิเคราะห์ข้อมูลใหม่อีกครั้ง ทำให้เสียเวลา งบประมาณ และความรู้สึก

ไสตท์ศน์ # 2.8.1 การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูล คือการนำข้อมูลที่รวบรวมมาได้ นำมาสู่วิธีการทางสถิติ แบ่งออกเป็น 2 ส่วน คือ

1. สถิติพรรณนา (descriptive statistics)
2. การอนุมานเชิงสถิติ (inference statistics)

นอกจากนี้การวิเคราะห์ข้อมูลยังต้องมีความรู้เกี่ยวกับประเภทของข้อมูล แบ่งเป็น 2 ประเภท คือ

1. ข้อมูลคุณภาพ (qualitative data)
2. ข้อมูลปริมาณ (quantitative data)

ข้อมูลทั้ง 2 ประเภท ยังมีระดับการวัดตัวแปรหรือข้อมูลเป็น 4 ระดับ คือ

1. ระดับนามมาตรา (nominal scale)
2. ระดับอันดับมาตรา (ordinal scale)
3. ระดับช่วงมาตรา (interval scale)
4. ระดับอัตราส่วนมาตรา (ratio scale)

ไต่ทัศน์ # 2.8.2 การแจกแจงความถี่ และหลักการสร้างตารางแจกแจงความถี่แบบมีอันตรภาคชั้น

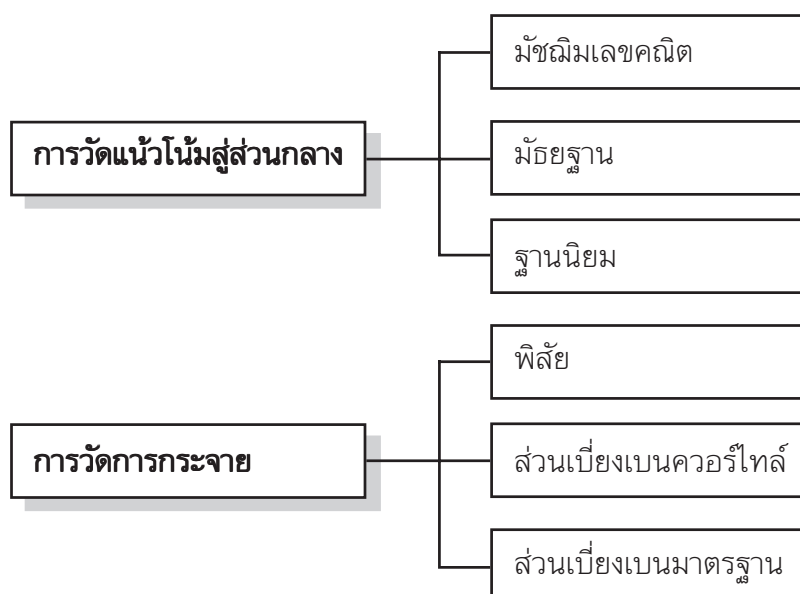
การแจกแจงความถี่

1. การแจกแจงแบบเรียงตัวเลข
2. การจัดเรียงเป็นหมวดหมู่แล้วแจกแจงความถี่
 - ตารางแจกแจงความถี่แบบธรรมดา
 - ตารางแจกแจงความถี่แบบมีอันตรภาคชั้น

หลักการสร้างตารางแจกแจงความถี่แบบมีอันตรภาคชั้น

1. หาค่าสูงสุดของข้อมูล
 2. หาค่าต่ำสุดของข้อมูล
 3. หาพิสัย พิสัย = ค่าสูงสุด - ค่าต่ำสุด
 4. กำหนดจำนวนชั้น
 5. หาอันตรภาคชั้น = $\frac{\text{พิสัยจำนวน}}{\text{จำนวนชั้น}}$
 6. การจัดเรียงชั้น อาจเริ่มเรียงให้ชั้นที่มีค่าน้อยอยู่ชั้นบนก่อน หรืออาจเรียงให้ชั้นที่มีค่ามากอยู่ชั้นบนก่อนก็ได้ แต่การเรียงให้ชั้นค่าน้อยอยู่ชั้นบนก่อน แล้วชั้นค่ามากเรียงชั้นล่างลงไปตามลำดับ จะทำให้ดูง่ายกว่า
 7. ในตารางที่เรียงค่าน้อยไว้ชั้นบน ค่าของขีดล่างในชั้นนี้จะต้องมีค่าน้อยกว่าหรือเท่ากับค่าต่ำสุดของข้อมูล และค่าของขีดจำกัดบนของชั้นสูงสุด จะต้องมีความมากกว่าหรือเท่ากับค่าสูงสุดของข้อมูล
 8. แจกแจงความถี่ของค่าของข้อมูลที่มีค่าอยู่ในชั้นต่าง ๆ ด้วยเครื่องหมาย tally mark แล้วรวมลงเป็นค่าความถี่
- (ดูตัวอย่างในเอกสารการสอน หน้า 11 ดังตารางที่ 8.3)

ไต่ทัศน์ # 2.8.3 การวัดแนวโน้มสู่ส่วนกลาง และการจัดการกระจาย



มัชฌิมเลขคณิต

1.1 ข้อมูลที่ไม่ได้แจกแจงความถี่

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x}{n}$$

1.2 ข้อมูลที่แจกแจงความถี่

$$\bar{x} = \frac{\sum fx}{n}$$

ถ้าข้อมูล คือ 2, 3, 5, 6, 4	มีมัชฌิมเลขคณิต = 4
7, 8, 10, 11, 9	มีมัชฌิมเลขคณิต = ...
1, 2, 4, 5, 3	มีมัชฌิมเลขคณิต = ...
$\frac{3}{4}, \frac{5}{4}, \frac{7}{4}, \frac{9}{4}, \frac{8}{4}$	มีมัชฌิมเลขคณิต = ...

มัชฐาน

คือ ค่าที่ตำแหน่งอยู่ตรงกลางของข้อมูลทั้งหมด หาได้โดยการนำข้อมูลทั้งหมดมาเรียงกันจากน้อยไปหามาก หรือจากมากไปหาน้อย แล้วนับจำนวน ถ้าจำนวนได้จำนวนคี่ ค่าของข้อมูลที่มีตำแหน่งอยู่ตรงกลาง คือ ค่ามัชฐาน แต่ถ้าจำนวนได้จำนวนคู่ ค่าของข้อมูลที่มีตำแหน่งอยู่ตรงกลาง รวมกันแล้วหารด้วย 2 คือ ค่ามัชฐาน

มัชฐาน ค่าของข้อมูลตำแหน่งกลาง

1. ข้อมูลที่ไม่ได้แจกแจงความถี่

เช่น มัชฐานของข้อมูล 2 3 2 6 8 คือ 5

$$\text{มัชฐานของข้อมูล } 1 \ 4 \ 5 \ 7 \ 8 \text{ คือ } \frac{5 + 7}{2} = 6$$

2. ข้อมูลที่แจกแจงความถี่

$$\text{มัชฐาน} = L + i \left[\frac{\frac{N}{2} - C}{f} \right]$$

(ดูตัวอย่างในเอกสารการสอน หน้า 18 ดังตารางที่ 8.7)

ไต่ทัศน์ # 2.8.4 มัชฌิมเลขคณิต มัธยฐาน และฐานนิยม (ต่อ)

ฐานนิยม

คือ ค่าของข้อมูลที่มีจำนวนมากที่สุด หรือมีความถี่สูงสุด ข้อมูลชุดหนึ่ง ๆ อาจมีฐานนิยมได้หลายค่า ถ้าข้อมูลชุดนั้นมีค่าของข้อมูลที่มีความถี่สูงสุดเท่ากันหลายค่า บางชุดของข้อมูลอาจไม่มีฐานนิยมถ้าข้อมูลชุดนั้นมีความถี่เท่ากัน วิธีการหาค่าฐานนิยมทำได้โดยการเรียงข้อมูลตามลำดับจากมากไปหาน้อยหรือจากน้อยไปหามาก ค่าของข้อมูลที่มีความถี่สูงสุดคือฐานนิยม

ฐานนิยม ค่าที่มีจำนวนซ้ำมากที่สุด

1. ข้อมูลที่ไม่ได้แจกแจงความถี่

เช่น ฐานนิยมของข้อมูล 3 8 5 6 4 5 คือ 5

2. ข้อมูลที่แจกแจงความถี่

$$\text{ฐานนิยม} = L + i \left[\frac{F_2}{F_1 + F_2} \right]$$

(ดูตัวอย่างในเอกสารการสอน หน้า 20 ดังตารางที่ 8.8)

ไต่ทัศน์ # 2.8.5 พิสัย ส่วนเบี่ยงเบนมาตรฐาน และสัมประสิทธิ์แห่งการกระจาย (the coefficient of dispersion)

การวัดการกระจาย

เป็นการพิจารณาว่าค่าของข้อมูลทั้งหมดมีการกระจายออกจากค่ากลางมากน้อยเพียงใด การวัดการกระจายที่นิยมใช้กันมากในงานวิจัยทั่ว ๆ ไป คือ พิสัยและส่วนเบี่ยงเบนมาตรฐาน

พิสัย

พิสัย = ค่ามากที่สุด - ค่าน้อยที่สุด

เช่น ข้อมูล 6 4 8 9 7 7 5 6 4 8

พิสัย = 9 - 4 = 5

ส่วนเบี่ยงเบนมาตรฐาน

คือ ค่าที่ใช้วัดความแตกต่างระหว่างค่าแต่ละค่าของข้อมูลนั้นกับมัชฌิมเลขคณิตซึ่งเป็นตัวแทนของข้อมูลทั้งหมด

1. การหาส่วนเบี่ยงเบนมาตรฐานสำหรับข้อมูลที่ไม่ได้แจกแจงความถี่

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

โสตทัศน # 2.8.5 (ต่อ)

ตัวอย่าง สุ่มตัวอย่างสอบถามเกี่ยวกับการให้คะแนนความสะอาดของห้องทำงานแห่งหนึ่งคะแนนเต็มเป็น 15 มีผู้ตอบแบบสอบถาม 6 คน ให้คะแนนดังนี้ 2, 5, 7, 8, 9, 11

หาค่าเฉลี่ยของข้อมูล

$$\bar{x} = \frac{\sum x}{n} = \frac{2 + 5 + 7 + 8 + 9 + 11}{6} = \frac{42}{6} = 7$$

x	(x - \bar{x})	(x - \bar{x}) ²
2	-5	25
5	-2	4
7	0	0
8	1	1
9	2	4
11	4	16
$\sum x = 42$	-	50

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} = \sqrt{\frac{50}{6 - 1}} = \sqrt{10}$$

$$= 3.16$$

2. การหาส่วนเบี่ยงเบนมาตรฐานสำหรับข้อมูลที่มีการแจกแจงความถี่

$$s = \sqrt{\frac{\sum f(x - \bar{x})^2}{n - 1}}$$

โสตทัศน # 2.8.5 (ต่อ)

สัมประสิทธิ์แห่งการกระจาย

สัมประสิทธิ์แห่งการกระจายเป็นวิธีการที่นำมาใช้ในการเปรียบเทียบการกระจายของข้อมูลต่างชนิดกัน ซึ่งข้อมูลต่างชุดกันที่มีค่าการกระจายเท่ากัน อาจมีขนาดของข้อมูลต่างกันมากก็ได้

$$\text{สัมประสิทธิ์แห่งการกระจาย} = \frac{S}{\bar{X}} \times 100$$

ค่านี้อาจเรียกได้อีกชื่อหนึ่งว่า “สัมประสิทธิ์ของความแปรปรวน” (coefficient of variation, C.V.)

ห้องเรียน	มัธยมเลขคณิต	ส่วนเบี่ยงเบนมาตรฐาน
ห้อง ก	20	5
ห้อง ข	24	3

จะได้ $CV_g = \dots\dots\dots$

จะได้ $CV_b = \dots\dots\dots$

ดังนั้น $\dots\dots\dots$ มีการกระจายมากกว่า ห้อง $\dots\dots\dots$

โสตทัศน # 2.8.6 การเปรียบเทียบค่าเฉลี่ยและการเปรียบเทียบค่าสัดส่วน

การเปรียบเทียบค่าเฉลี่ย

1. การเปรียบเทียบค่าเฉลี่ยของประชากร 1 กลุ่ม
 - 1.1 ทราบค่าความแปรปรวนของประชากร (σ^2 known)
 - 1.2 ไม่ทราบค่าความแปรปรวนของประชากร (σ^2 unknown)
2. การเปรียบเทียบความแตกต่างของค่าเฉลี่ยของประชากร 2 กลุ่ม
 - 2.1 ประชากร 2 กลุ่มเป็นอิสระต่อกัน
 - 2.2.1 ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2 known)
 - 2.2.2 ไม่ทราบค่าความแปรปรวนของประชากร (σ_1^2, σ_2^2 unknown)
 - 2.2 ประชากร 2 กลุ่มไม่เป็นอิสระต่อกัน

ไสตท์ศน์ # 2.8.6 การเปรียบเทียบค่าเฉลี่ยและการเปรียบเทียบค่าสัดส่วน (ต่อ)

การเปรียบเทียบค่าสัดส่วน

1. การเปรียบเทียบค่าสัดส่วนของประชากร 1 กลุ่ม
2. การเปรียบเทียบความแตกต่างของค่าสัดส่วนของประชากร 1 กลุ่ม
 - 2.1 ประชากร 2 กลุ่มที่เป็นอิสระต่อกัน
 - 2.2 ประชากร 2 กลุ่มที่ไม่เป็นอิสระต่อกัน
3. การเปรียบเทียบความแตกต่างของค่าสัดส่วนของประชากร 2 กลุ่ม
 - 3.1 ประชากรมากกว่า 2 กลุ่มที่เป็นอิสระต่อกัน
 - 3.2 ประชากรมากกว่า 2 กลุ่มที่ไม่เป็นอิสระต่อกัน

ไสตท์ศน์ # 2.8.7 การหาความสัมพันธ์ระหว่างตัวแปร การหาค่าตัวทำนาย

การหาความสัมพันธ์ระหว่างตัวแปร

1. การหาความสัมพันธ์ของข้อมูลเชิงคุณภาพ
 - 1.1 ตารางการณ์จร $r \times c$
 - 1.2 ตารางการณ์จร 2×2
 - 1.3 การทดสอบความแม่นยำตรงของพิชเซอร์ สำหรับตาราง 2×2
2. การหาความสัมพันธ์ของข้อมูลเชิงปริมาณ

การหาค่าตัวทำนาย

ในการหาค่าตัวทำนายนี้หมายความว่า ถ้าตัวแปรหนึ่ง ๆ มีความสัมพันธ์กับตัวแปรอีกตัวหนึ่ง หรือตัวแปรอิสระมีผลต่อตัวแปรตาม ตัวแปรอิสระ (X) จะสามารถทำนายการเกิดตัวแปรตาม (Y) ได้หรือไม่ มีความแม่นยำในการทำนายได้เพียงใด บางกรณีเรียกว่า การหาตัวพยากรณ์ (prediting forecasting) เป็นสถิติขั้นสูง (advanced statistics) ซึ่งต้องคำนึงถึงข้อตกลงเบื้องต้นก่อนนำไปใช้

1. การวิเคราะห์ถดถอย (regression analysis)
2. การวิเคราะห์จำแนกพหุ (multiple classification analysis)
3. การวิเคราะห์จำแนกประเภท (discriminant analysis)

ไสตท์ศน์ # 2.8.8 การนำเสนอข้อมูล

การนำเสนอข้อมูล

1. การนำเสนอข้อมูลในรูปบทความ
2. การนำเสนอข้อมูลในรูปบทความกึ่งตาราง
3. การนำเสนอข้อมูลในรูปตาราง
 - 3.1 ตารางแจกแจงทางเดียว หรือตารางเดียว
 - 3.2 ตารางแจกแจงสองทาง หรือตารางไขว้
 - 3.3 ตารางแจกแจงสามทาง
4. การนำเสนอข้อมูลในรูปกราฟและแผนภูมิ
 - 4.1 การนำเสนอข้อมูลแบบรูปกราฟเส้น
 - 4.2 การนำเสนอข้อมูลแบบแผนภูมิแท่ง
 - 4.3 การนำเสนอข้อมูลแบบแผนภูมิวงกลม
 - 4.4 การนำเสนอข้อมูลแบบรูปภาพ
 - 4.5 การนำเสนอข้อมูลแบบแผนภูมิแผนที่
 - 4.6 การนำเสนอข้อมูลแบบแผนภาพกระจาย หรือแผนภาพสหลัมพันธ์
 - 4.7 การนำเสนอข้อมูลแบบฮีสโตแกรม
 - 4.8 การนำเสนอข้อมูลแบบรูปหลายเหลี่ยมแห่งความถี่
 - 4.9 การนำเสนอข้อมูลแบบแผนภูมิพีระมิต

ไสตท์ศน์ # 2.9.1 เซต

เซต

หมายถึง กลุ่มของสิ่งของต่าง ๆ ซึ่งมีคุณสมบัติที่แน่ชัด โดยที่สามารถบอกสมาชิกของเซตนั้นได้ การดำเนินการของเซต ได้แก่ ยูเนียน อินเตอร์เซกชัน และคอมพลีเมนต์ เป็นการนำหลาย ๆ เซตมาดำเนินการให้เกิดเซตใหม่

$A \cup B$

คือ เซตที่มีสมาชิกที่เป็นสมาชิกของเซต A หรือสมาชิกของเซต B หรือของทั้งสองเซต

เขียนอธิบายลักษณะของสมาชิกได้ว่า $A \cup B = \{x \mid x \text{ เป็นสมาชิกของ } A \text{ หรือ } x \text{ เป็นสมาชิกของ } B\}$

$A \cap B$

คือ เซตที่มีสมาชิกที่เป็นสมาชิกของเซต A และเซต B

เขียนอธิบายลักษณะของสมาชิกได้ว่า $A \cap B = \{x \mid x \text{ เป็นสมาชิกของ } A \text{ และ } x \text{ เป็นสมาชิกของ } B\}$

A^c , A' หรือ \bar{A}

คอมพลีเมนต์ของเซต A ใดๆ เมื่อเทียบกับเอกภพสหลัมพันธ์ U คือเซตของทุกสมาชิกที่มีอยู่ใน U แต่ไม่อยู่ใน A

เขียนอธิบายลักษณะของสมาชิกได้ว่า $A^c = \{x \mid x \text{ เป็นสมาชิกของ } U \text{ และ } x \text{ ไม่เป็นสมาชิกของ } A\}$

ไสตท์ศน์ # 2.9.2 การเรียงสับเปลี่ยนและการจัดหมู่**การเรียงสับเปลี่ยนและการจัดหมู่**

เป็นเรื่องราวของจำนวนวิธีการจัดสิ่งของที่แตกต่างกัน โดยดูอันดับสิ่งของเรียกว่าการเรียงสับเปลี่ยน และไม่ดูอันดับสิ่งของเรียกว่าการจัดหมู่

การเรียงสับเปลี่ยนของสิ่งของ n สิ่ง ซึ่งแตกต่างกันทั้งหมด โดยจัดอันดับทีละ r สิ่ง ($r < n$) มีค่าคือ

$${}^n P_r = \frac{n!}{(n-r)!}$$

การจัดหมู่ของสิ่งของที่แตกต่างกัน n สิ่ง โดยที่มีหมู่ละ r สิ่ง ($r \leq n$) มีค่าคือ

$${}^n C_r = \frac{n!}{r!(n-r)!}$$

ไสตท์ศน์ # 2.9.3 การทดลองเชิงสุ่ม ปริภูมิตัวอย่าง และเหตุการณ์**การทดลองเชิงสุ่ม**

หมายถึง การทดลองที่ไม่สามารถพยากรณ์ผลลัพธ์ได้อย่างถูกต้องแน่นอน เนื่องจากผลลัพธ์อาจเกิดขึ้นได้หลายอย่าง

ปริภูมิตัวอย่าง

หมายถึง เซตที่มีสมาชิกเป็นผลลัพธ์ที่อาจจะเป็นไปได้ทั้งหมดของการทดลองเชิงสุ่ม

เหตุการณ์

หมายถึง เซตย่อยของปริภูมิตัวอย่าง

ไสตท์ศน์ # 2.9.4 ความน่าจะเป็น และการแจกแจงความน่าจะเป็น**ความน่าจะเป็น**

หมายถึง ค่าความถี่สัมพัทธ์ที่จะบอกให้ทราบว่าเหตุการณ์นั้นๆ มีโอกาสเกิดขึ้นได้มากน้อยเพียงใด

คุณสมบัติของความน่าจะเป็น

1. ความน่าจะเป็นของเหตุการณ์ E ใดๆ มีค่าตั้งแต่ 0 ถึง 1 เสมอ นั่นคือ $0 \leq P(E) \leq 1$
2. ความน่าจะเป็นของปริภูมิตัวอย่าง S มีค่าเท่ากับ 1 นั่นคือ $P(S) = 1$
3. ความน่าจะเป็นของเหตุการณ์ที่เป็นเซตว่างเท่ากับ 0 นั่นคือ $P(\emptyset) = 0$

ไสตท์ศน์ # 2.9.4 ความน่าจะเป็น และการแจกแจงความน่าจะเป็น (ต่อ)

4. ถ้า $E_1, E_2, E_3, \dots, E_n$ เป็นเหตุการณ์ที่ไม่มีส่วนร่วม จะได้ว่า

$$P(E_1 \cup E_2 \cup E_3 \cup \dots \cup E_n) = P(E_1) + P(E_2) + P(E_3) + \dots + P(E_n)$$

5. ถ้า E_1 และ E_2 เป็นเหตุการณ์ใดๆ ที่เป็นเซตย่อยของปริภูมิตัวอย่าง S แล้ว

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

6. เหตุการณ์ E_1 และ E_2 เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ก็ต่อเมื่อ

$$P(E_1 \cap E_2) = P(E_1) P(E_2)$$

การแจกแจงความน่าจะเป็น

หมายถึง การแจกแจงความถี่ของตัวแปรสุ่มชนิดไม่ต่อเนื่อง และชนิดต่อเนื่อง ด้วยค่าความน่าจะเป็น การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่อง ได้แก่ การแจกแจงทวินาม และการแจกแจงปัวซอง การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง ได้แก่ การแจกแจงปกติ

ไสตท์ศน์ # 2.9.5 การแจกแจงทวินาม

การแจกแจงทวินาม

คือ การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่อง ที่ผลการทดลองมี 2 อย่าง คือ สำเร็จและล้มเหลว

ความน่าจะเป็นของการแจกแจงทวินาม

$$f(x) = {}^n C_x p^x (1-p)^{n-x}; x = 0, 1, 2, 3, \dots, n$$

โดยที่ n คือ จำนวนครั้งของการทดลอง

x คือ จำนวนครั้งของความสำเร็จ

${}^n C_x$ คือ จำนวนรูปแบบของผลการทดลองที่เป็นไปได้ทั้งหมด

$$\text{ค่า } {}^n C_x = \frac{n!}{x!(n-x)!}$$

p คือ ความน่าจะเป็นของความสำเร็จในแต่ละครั้งของการทดลอง

$1-p$ คือ ความน่าจะเป็นของความล้มเหลวในแต่ละครั้งของการทดลอง (หรือให้ $q = 1-p$)

การแจกแจงทวินามมีค่าเฉลี่ย = np และ ค่าความแปรปรวน = npq

ไสตท์ศน์ # 2.9.6 การแจกแจงปัวซอง

การแจกแจงปัวซอง

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดไม่ต่อเนื่อง ที่ค่าเฉลี่ยของความสำเร็จเกิดขึ้นในช่วงเวลาหรือขอบเขตที่กำหนด

ความน่าจะเป็นของการแจกแจงปัวซอง

$$f(x) = \frac{e^{-\mu} \mu^x}{x!}; x = 0, 1, 2, 3, \dots, \infty$$

โดยที่ x คือ จำนวนของความสำเร็จในการทดลองแบบปัวซอง

μ คือ ค่าเฉลี่ยของความสำเร็จที่เกิดขึ้นในช่วงเวลาหรือในขอบเขตที่กำหนด

e คือ ค่าคงที่ = 2.718281828...

การแจกแจงปัวซองมีค่าเฉลี่ย = μp และ ค่าความแปรปรวน = μp

ไสตท์ศน์ # 2.9.7 การแจกแจงปกติ

การแจกแจงปกติ

คือ การแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่อง ที่โค้งการแจกแจงเป็นรูปประฆังคว่ำสมมาตร

ความน่าจะเป็นของการแจกแจงปกติ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}}; -\infty < x < \infty$$

โดยที่ μ คือ มัชฌิมเลขคณิต หรือเรียกว่าค่าเฉลี่ย

σ คือ ส่วนเบี่ยงเบนมาตรฐาน

π คือ ค่าคงที่ = $\frac{22}{7} = 3.141592654...$

e คือ ค่าคงที่ = 2.718281828...

การแจกแจงปกติมีค่าเฉลี่ย = μ และ ค่าความแปรปรวน = σ^2

ไสตท์ศน์ # 2.9.7 การแจกแจงปกติ (ต่อ)

ความน่าจะเป็นของการแจกแจงปกติมาตรฐาน

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}; -\infty < z < \infty$$

โดยที่ $z = \frac{x - \mu}{\sigma}$

π คือ ค่าคงที่ = $\frac{22}{7} = 3.141592654\dots$

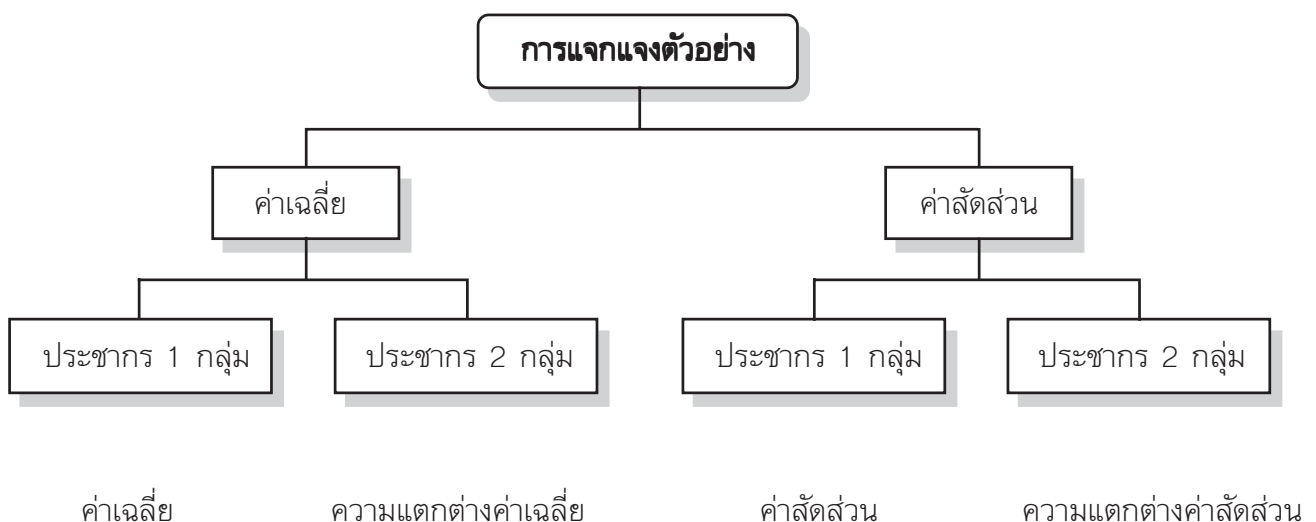
e คือ ค่าคงที่ = 2.718281828...

การแจกแจงปกติมาตรฐานมีการแจกแจงปกติด้วยค่าเฉลี่ย $\mu = 0$ และ ค่าความแปรปรวน $\sigma^2 = 1$ หรือเขียนด้วยสัญลักษณ์ว่า $z \sim N(0, 1)$

ไสตท์ศน์ # 2.9.8 การแจกแจงตัวอย่าง

การแจกแจงตัวอย่าง

คือการแจกแจงความน่าจะเป็นของค่าสถิติที่ได้จากตัวอย่าง คุณสมบัติของการแจกแจงค่าเฉลี่ยและค่าสัดส่วนของตัวอย่าง ได้นำไปใช้ในการอนุมานเชิงสถิติ



ไสถทัศน์ # 2.9.9

การแจกแจงความน่าจะเป็นของค่าเฉลี่ย \bar{x} ที่ได้จากการสุ่มตัวอย่างจากประชากรที่มีการแจกแจงปกติ ด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวน เท่ากับ σ^2 จะได้รับการแจกแจงค่าเฉลี่ยของตัวอย่างมีค่าเฉลี่ย $\mu_{\bar{x}}$ และความแปรปรวน $\sigma_{\bar{x}}^2$ ดังนี้

1. ค่าเฉลี่ย $\mu_{\bar{x}} = \mu$
2. ความแปรปรวน $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$
3. ความคลาดเคลื่อนมาตรฐาน $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
4. มีโค้งการแจกแจงปกติ เมื่อประชากรมีการแจกแจงปกติหรือไม่มีการแจกแจงปกติก็ตาม แต่สุ่มตัวอย่างมาขนาดใหญ่พอ (จากทฤษฎีซีตจังก์กลาง)

การแจกแจงความแตกต่างค่าเฉลี่ยของตัวอย่างที่มีการแจกแจงปกติ มีค่าเฉลี่ย $\mu_{(\bar{x}_1 - \bar{x}_2)}$ และความแปรปรวน $\sigma_{(\bar{x}_1 - \bar{x}_2)}^2$ ดังนี้

1. ค่าเฉลี่ย $\mu_{(\bar{x}_1 - \bar{x}_2)} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2$
2. ความแปรปรวน $\sigma_{(\bar{x}_1 - \bar{x}_2)}^2 = \sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$
3. ความคลาดเคลื่อนมาตรฐาน $\sigma_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
4. การแจกแจงนี้มีการแจกแจงปกติ ถ้าการแจกแจงค่าเฉลี่ยของตัวอย่างแต่ละชุดนั้นมีการแจกแจงปกติ

การแจกแจงค่าเฉลี่ยของตัวอย่าง การแจกแจงของ p มีค่าเฉลี่ย μ_p ความแปรปรวน σ_p^2 ดังนี้

1. ค่าเฉลี่ย $\mu_p = \pi$
2. ความแปรปรวน $\sigma_p^2 = \frac{\pi(1-\pi)}{n}$
3. ความคลาดเคลื่อนมาตรฐาน $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$
4. เป็นการแจกแจงทวินาม แต่เมื่อ n มีขนาดมาก การแจกแจงนี้สามารถประมาณค่าได้จากการแจกแจงปกติ

ไสตท์ศน์ # 2.9.9 (ต่อ)

การแจกแจงความแตกต่างค่าเฉลี่ยของตัวอย่าง มีคุณสมบัติของการแจกแจงของ $p_1 - p_2$ มีค่าเฉลี่ย $\mu_{(p_1 - p_2)}$ และความแปรปรวน $\sigma^2_{(p_1 - p_2)}$ ดังนี้

$$1. \text{ ค่าเฉลี่ย } \mu_{(p_1 - p_2)} = \pi_1 - \pi_2$$

$$2. \text{ ความแปรปรวน } \sigma^2_{(p_1 - p_2)} = \frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}$$

$$3. \text{ ความคลาดเคลื่อนมาตรฐาน } \sigma_{(p_1 - p_2)} = \sqrt{\frac{\pi_1(1 - \pi_1)}{n_1} + \frac{\pi_2(1 - \pi_2)}{n_2}}$$

ไสตท์ศน์ # 2.10.1 ประเภทของการประมาณค่า

ประเภทของการประมาณค่า

1. **การประมาณค่าเป็นจุด** คือการนำเอาค่าสถิติที่ได้จากข้อมูลตัวอย่างไปใช้ประมาณค่าพารามิเตอร์เลย การประมาณค่าพารามิเตอร์หนึ่งๆ โดยกล่าวว่ามีความเท่ากับค่าเพียงค่าเดียวเท่านั้น ก็จะมีโอกาสหรือความน่าจะเป็นที่ผิดพลาดได้ง่าย

2. **การประมาณค่าเป็นช่วง** คือการนำเอาค่าสถิติที่ได้จากข้อมูลของตัวอย่างไปใช้ประมาณค่าพารามิเตอร์ โดยการกำหนดค่าประมาณเป็นช่วงว่า ค่าพารามิเตอร์ที่ต้องการทราบจะมีค่าอยู่ในช่วงนี้ ด้วยความเชื่อมั่นว่าจะถูกต้องเป็นร้อยละหรือเปอร์เซ็นต์เท่าใด และเรียกช่วงนั้นว่าช่วงความเชื่อมั่น (confidence interval)

ไสตท์ศน์ # 2.10.2 คุณสมบัติของตัวประมาณค่าเป็นจุดที่ดี

คุณสมบัติของตัวประมาณค่าเป็นจุดที่ดี

1. เป็นตัวประมาณค่าที่ไม่อคติ (unbiased estimator)
2. มีความพ้องและสม่ำเสมอหรือมีความคงที่ (consistency)
3. มีความแปรปรวนน้อยที่สุด (minimum variance)
4. มีความเพียงพอ (sufficiency)

ขั้นตอนของการทดสอบสมมติฐาน

1. การตั้งสมมติฐานทางสถิติ
 - 1.1 สมมติฐานว่าง
 - 1.2 สมมติฐานทางเลือก

2. การตั้งสถิติที่ใช้ทดสอบ
 - 2.1 z-test
 - 2.2 t-test
 - 2.3 χ^2 -test
 - 2.4 F-test

3. การกำหนดความผิดพลาดในการทดสอบ
 - 3.1 ความผิดพลาดประเภท 1
 - 3.2 ความผิดพลาดประเภท 2

4. ระดับนัยสำคัญ

5. บริเวณวิกฤต

6. การทดสอบสถิติ
 - 6.1 การทดสอบสองข้าง
 - 6.2 การทดสอบข้างเดียว

7. ค่าพี (p-value)

8. สรุปผลการทดสอบ

ไฮททัศน์ 2.10.4 # การประมาณค่าเฉลี่ยและการประมาณค่าสัดส่วน

การประมาณค่าเฉลี่ย

ตัวประมาณค่าเฉลี่ยของประชากร (μ) คือค่าเฉลี่ยเลขคณิตของตัวอย่าง (x) มีสูตรในการประมาณค่า ดังนี้

$$(1 - \alpha) 100\% \text{ ขีดจำกัดความเชื่อมั่นของ } \mu = x \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

การประมาณค่าสัดส่วน

ตัวประมาณค่าสัดส่วนของประชากร (P) คือค่าสัดส่วนของตัวอย่าง (p) มีสูตรในการประมาณค่าดังนี้

$$(1 - \alpha) 100\% \text{ ขีดจำกัดความเชื่อมั่นของ } P = p \pm z_{\alpha/2} \sqrt{p(1-p)/n}$$

ยกตัวอย่างการประมาณค่าเฉลี่ย และการประมาณค่าสัดส่วน

เซ็ง เครียด ทำไง?

เคยไหม? นอนไม่หลับ เหนื่อยใจ ปวดเนื้อปวดตัวกับที่ไม่ได้ใช้แรง
ทำใช้...คุณกำลังเข้าสู่ภาวะความเครียด!

จะรู้เท่ากับความเครียด...



กดดัน ปวดศีรษะ

หมกมุ่นครุ่นคิด นำไปสู่ยาเสพติด
โรคกระเพาะ **กลัว** วิตกกังวล
ไม่สนใจสิ่งรอบตัว ความดันโลหิตสูง
เสียวความมั่นใจ ปวดหลัง โรคหัวใจ
บ่นไปหลับ ทำร้ายผู้อื่น
ต้นเหตุโรคจิตโรคประสาท
หกลดเลือดอุดตัน
โรคอ้วน

ไม่เอา
ไม่(อยาก)
เครียด...

-  คิดในทางบวก
-  ดนตรี ฟังเพลง
-  ทำสมาธิ
-  ออกกำลังกาย
-  บริการเพื่อน ครอบครัว หรือจิตแพทย์

ความเครียดศัตรูตัวร้ายของทุกคน
ปราบมันให้ได้เพื่ออยู่บนโลกที่ดูวุ่นวายใบนี้อย่างมีความสุข

Source: www.pharmacy.mahidol.ac.th

Infographic MOVE

