

121-107

สถิติพื้นฐานเพื่อการ
วิเคราะห์ข้อมูลเบื้องต้น

บทที่ 4

การวิเคราะห์ข้อมูล



วัตถุประสงค์การเรียนรู้

เมื่อจบการเรียนรู้ในสัปดาห์นี้แล้ว นักศึกษาสามารถ

1. เข้าใจวิธีการวิเคราะห์ข้อมูลแต่ละวิธีโดยภาพรวม
2. เลือกใช้วิธีการวิเคราะห์ข้อมูลที่ถูกต้องเหมาะสมกับข้อมูลที่รวบรวมได้

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลขั้นต้น
หรือ

การวิเคราะห์ด้วยสถิติแบบบรรยาย

นำสถิติมาใช้ในการบรรยาย
หรืออธิบายลักษณะของ
ข้อมูล

การวิเคราะห์ข้อมูลขั้นสูง
หรือ

การวิเคราะห์ด้วยสถิติแบบอ้างอิง

นำสถิติมาใช้ในการหาข้อสรุป
เกี่ยวกับข้อมูลของประชากร
โดยอาศัยค่าของข้อมูลที่ได้จาก
การเลือกตัวอย่าง

การวิเคราะห์ข้อมูลด้วยสถิติแบบบรรยาย

การแจกแจงความถี่

บรรยายข้อมูล
ด้วยการนับ

การวัดการกระจาย

บรรยายความ
แตกต่างของข้อมูล

- พิสัย
- ค่าเบี่ยงเบนมาตรฐาน
- สัมประสิทธิ์
ความแปรผัน

โค้งความถี่

บรรยายข้อมูลด้วย
เส้นโค้งจากการ
แจกแจงความถี่

การวัดแนวโน้มเข้าสู่ส่วนกลาง

บรรยายข้อมูลด้วยตัวเลข
ที่เป็นตัวแทนของข้อมูล

- ค่าเฉลี่ย
- ฐานนิยม
- ค่าที่บอกตำแหน่ง

การหาค่ามาตรฐาน

การปรับข้อมูลให้อยู่
ในมาตรฐานเดียวกัน

การวิเคราะห์ข้อมูลด้วยสถิติแบบอ้างอิง

การประมาณค่า

นำค่าของข้อมูลที่ได้จากตัวอย่างไปประมาณค่าของประชากร

การทดสอบสมมติฐาน

นำค่าของข้อมูลที่ได้จากตัวอย่างไปทดสอบข้อสมมติหรือข้อคาดเดาเกี่ยวกับค่าพารามิเตอร์ของประชากร

การหาความสัมพันธ์

การศึกษารูปแบบของความสัมพันธ์ระหว่างข้อมูลตั้งแต่ 2 ชุด ขึ้นไป โดยทำการศึกษาจากข้อมูลที่ได้จากตัวอย่าง

การพยากรณ์

สร้างตัวแบบความสัมพันธ์จากข้อมูลตัวอย่างเพื่อการพยากรณ์

การประมาณค่า

การประมาณค่าแบบจุด

ประมาณค่าพารามิเตอร์
ของประชากรด้วยตัวเลขใด
ตัวเลขหนึ่ง โดยใช้ค่าสถิติที่
คำนวณจากข้อมูลตัวอย่าง

การประมาณค่าแบบช่วง

ประมาณค่าพารามิเตอร์ของ
ประชากรว่าจะอยู่ในช่วงใด
ช่วงหนึ่ง ด้วยหลักการ

ตัวประมาณค่าแบบจุด \pm ความคลาดเคลื่อน

การทดสอบสมมติฐาน (Hypothesis Testing)

กำหนดสมมติฐาน H_0 และ H_1



กำหนดเกณฑ์การตัดสินใจ α



เลือกและคำนวณค่าของ
ตัวสถิติทดสอบ



สรุปผล

ตัวสถิติทดสอบ

Z-test

t-test

F-test

χ^2 -test

ใช้ทดสอบเกี่ยวกับค่าเฉลี่ย หรือค่าสัดส่วนสำหรับข้อมูลหนึ่งกลุ่มหรือสองกลุ่มที่ทราบค่าเบี่ยงเบนมาตรฐานของข้อมูลทั้งหมด

ใช้แทน Z-test ในกรณีที่ไม่ทราบค่าเบี่ยงเบนมาตรฐานของข้อมูลทั้งหมด

ใช้ทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยสำหรับข้อมูลตั้งแต่ 2 กลุ่มขึ้นไป

ใช้ในการทดสอบเกี่ยวกับความเป็นอิสระกันระหว่างคุณลักษณะของข้อมูลโดยใช้จำนวนหรือความถี่ของข้อมูลมาเป็นตัวทดสอบ

การหาความสัมพันธ์ (Correlation)

การทดสอบความสัมพันธ์

ทดสอบว่าข้อมูล ตั้งแต่ 2 ชุด
มีความสัมพันธ์กันหรือไม่โดย
การทดสอบสมมติฐาน

การหาขนาดและทิศทาง ของความสัมพันธ์

ศึกษาว่าข้อมูล ตั้งแต่ 2 ชุดมี
ความสัมพันธ์กันมากน้อยเพียงใด
และสัมพันธ์ในทิศทางเดียวกัน
หรือทิศทางตรงกันข้าม

การพยากรณ์

การวิเคราะห์การถดถอย

การศึกษาคovariance ของตัวแปร 2 กลุ่ม กลุ่มหนึ่งเรียกว่าตัวแปรอิสระ อีกกลุ่มหนึ่งเรียกว่าตัวแปรตาม เพื่อประโยชน์ในการพยากรณ์ค่าของตัวแปรตามเมื่อทราบค่าของตัวแปรอิสระ

การวิเคราะห์อนุกรมเวลา

การศึกษาข้อมูลที่มีการเก็บรวบรวมแบบต่อเนื่องของช่วงเวลา โดยการสร้างความสัมพันธ์ระหว่างองค์ประกอบของข้อมูลอนุกรมเวลา เพื่อประโยชน์ในการพยากรณ์ข้อมูลอนุกรมเวลาในอนาคต

คำถามเพื่อทบทวน สัปดาห์ที่ 6

1. จากข้อความต่อไปนี้ ข้อใดสรุปถูกต้อง

ก) การวิเคราะห์ข้อมูลเป็นการจัดระบบระเบียบให้กับข้อมูล
ที่เก็บรวบรวมได้

ข) การวิเคราะห์ข้อมูลขั้นต้น เป็นการอธิบายลักษณะของ
ข้อมูลที่เก็บรวบรวมได้

1. ข้อ ก. ถูก ข้อ ข. ผิด
2. ข้อ ก. ผิด ข้อ ข. ถูก
3. ข้อ ก. และ ข้อ ข. ผิด
4. ข้อ ก. และ ข้อ ข. ถูก

คำถามเพื่อทบทวน สัปดาห์ที่ 6

2. ข้อใด **ไม่** จัดเป็นการวิเคราะห์ข้อมูลด้วยสถิติแบบบรรยาย
 1. การแจกแจงความถี่
 2. การวัดค่าการกระจาย
 3. การวัดแนวโน้มเข้าสู่ส่วนกลาง
 4. การประมาณค่า

คำถามเพื่อทบทวน สัปดาห์ที่ 6

จากข้อมูลต่อไปนี้ ตอบคำถามข้อ 3-4

ในการศึกษาพฤติกรรมการอ่านหนังสือของนักศึกษา มหาวิทยาลัยสยาม โดยทำการเก็บข้อมูลจากกลุ่มตัวอย่าง จำนวน 396 คน พบว่ากลุ่มตัวอย่างอ่านหนังสือโดยเฉลี่ย วันละ 15 นาที

คำถามเพื่อทบทวน สัปดาห์ที่ 6

3. ข้อใดกล่าวไม่ถูกต้อง

1. เวลาเฉลี่ยในการอ่านหนังสือของกลุ่มตัวอย่าง 15 นาทีต่อวัน
จัดเป็นค่าสถิติ
2. ประเมินได้ว่านักศึกษามหาวิทยาลัยสยามทั้งหมดอ่านหนังสือ
โดยเฉลี่ยวันละ 15 นาที
3. ตามหลักการสร้างช่วงการประมาณแล้วอาจจะประมาณได้ว่า
เวลาเฉลี่ยในการอ่านหนังสือของนักศึกษามหาวิทยาลัยสยาม
ทั้งหมดอยู่ในช่วง 10 ถึง 17 นาทีต่อวัน
4. ค่าพารามิเตอร์ที่สนใจในที่นี้คือเวลาเฉลี่ยในการอ่านหนังสือของ
นักศึกษามหาวิทยาลัยสยามทั้งหมด

คำถามเพื่อทบทวน สัปดาห์ที่ 6

4. ถ้ามีการคาดเดาว่านักศึกษามหาวิทยาลัยสยามทั้งหมดอ่านหนังสือโดยเฉลี่ยวันละ 12 นาที ข้อใดกล่าวไม่ถูกต้อง
1. ต้องการทดสอบว่าค่าพารามิเตอร์เท่ากับ 12 นาทีหรือไม่
 2. ตัวสถิติทดสอบอาจจะใช้ T-test เนื่องจากไม่ทราบค่าเบี่ยงเบนมาตรฐานของประชากร
 3. การกำหนดสมมติฐานทางสถิติจะต้องมี 2 ส่วนคือ H_0 และ H_1
 4. ในการทดสอบสมมติฐานไม่จำเป็นต้องใช้ข้อมูลเวลาเฉลี่ยในการอ่านหนังสือของกลุ่มตัวอย่าง

คำถามเพื่อทบทวน สัปดาห์ที่ 6

5. จากข้อความต่อไปนี้ ข้อใดสรุปถูกต้อง

- ก) การพยากรณ์ยอดขายร่วมในเดือนกรกฎาคมเมื่อมีการเก็บข้อมูล
ยอดขายร่วมรายเดือนตั้งแต่เดือนมกราคม ปี 2551 ถึงเดือนเมษายน
ปี 2553 จะต้องใช้ความรู้ในเรื่องการวิเคราะห์การถดถอย
- ข) การศึกษาว่ายอดขายขึ้นอยู่กับราคาขายหรือไม่ จะต้องใช้ความรู้ในเรื่อง
ของการทดสอบความสัมพันธ์
1. ข้อ ก. ถูก ข้อ ข. ผิด
 2. ข้อ ก. ผิด ข้อ ข. ถูก
 3. ข้อ ก. และ ข้อ ข. ผิด
 4. ข้อ ก. และ ข้อ ข. ถูก