



เทคนิคการใช้สถิติเพื่อการวิจัย



รองศาสตราจารย์ชูศรี วงศ์รัตนะ

เทคนิคการใช้สถิติเพื่อการวิจัย

รองศาสตราจารย์ชูศรี วงศ์รัตน์

ค.บ. (เกียรตินิยม) ค.ม. (การวิจัยทางการศึกษา)

เทคนิคการใช้สถิติเพื่อการวิจัย

จัดทำโดย ชุติ วังศรีคันะ

พิมพ์ครั้งที่ 15/2564

จำนวน 1,000 เล่ม

ราคา 300 บาท

สงวนลิขสิทธิ์ตามพระราชบัญญัติลิขสิทธิ์ ห้ามลอกเลียนแบบหรือคัดลอกส่วนใดส่วนหนึ่งของหนังสือเล่มนี้
นอกจากจะได้รับอนุญาตเป็นลายลักษณ์อักษรจากผู้เขียน

ISBN 978-616-429-977-1

ผู้จัดจำหน่าย ศูนย์หนังสือจุฬาลงกรณ์มหาวิทยาลัย

ถนนพญาไท เขตปทุมวัน กรุงเทพฯ 10330

สาขา

ศาลาพระเกี้ยว โทร. 0 2218 7000-3 โทรสาร 0 2255 4441

สยามสแควร์ โทร. 0 2218 9881-2 โทรสาร 0 2254 9495

ม.นเรศวร จ.พิษณุโลก โทร. 0 5526 0162-4 โทรสาร 0 5526 0165

ม.เทคโนโลยีสุรนารี จ.นครราชสีมา โทร. 0 4421 6131-4 โทรสาร 0 4421 6135

ม.บูรพา จ.ชลบุรี โทร. 0 3839 4855-9 โทรสาร 0 3839 3239

โรงเรียนนายร้อย จปร. จ.นครนายก โทร. 0 3739 3023 โทรสาร 0 3739 3023

ม.พะเยา จ.พะเยา โทร. 0 5446 6799-800 โทรสาร 0 5446 6798

จัตุรัสจามจุรี (CHAMCHURI SQUARE) ชั้น 4 โทร. 0 2160 5301-2

โทรสาร 0 2160 5304

หัวหมาก โทร. 0 2374 1378 โทรสาร 0 2374 1377

ม.เทคโนโลยีราชมงคลอีสาน จ.นครราชสีมา โทร. 0 4492 2662-3

โทรสาร 0 4492 2664 มือถือ 08 6392 7785

Call Center (จัดส่งทั่วประเทศ) โทร. 0 2255 4433 <http://www.chulabook.com>

และเครือข่ายฯ

ร้านค้าติดต่อ แผนกขายส่ง สาขาหัวหมาก โทร. 0 2374 1374-5 โทรสาร 0 2374 1377

พิมพ์ที่ อมรการพิมพ์

157, 159, 161 ซอยเพชรเกษม 48 แยก 4 ถนนเพชรเกษม แขวงบางด้วน

เขตภาษีเจริญ กรุงเทพฯ 10160 โทร. 0 2869 5747, 0 2467 1110

นายอมร เจียรสุพัฒน์ ผู้พิมพ์ผู้โฆษณา

คำนำ

หนังสือเทคนิคการใช้สถิติเพื่อการวิจัย ได้รับความสนใจอย่างกว้างขวางจากนิสิตนักศึกษาทั้งระดับปริญญาตรีและระดับบัณฑิตศึกษา ตลอดจนครูอาจารย์และนักวิจัย โดยใช้หลักการว่า ไม่เน้นทฤษฎีทางสถิติ แต่เน้นให้ผู้อ่านมีความรู้ความเข้าใจเกี่ยวกับสถิติแต่ละชนิด เพื่อให้สามารถนำไปใช้ตอบโจทย์การวิจัยได้ โดยนำเสนอสาระที่สำคัญ คือ

1. นำเสนอสถิติแต่ละชนิด โดยเริ่มด้วยการอธิบายความคิดรวบยอด หลักการใช้ ตัวอย่างงานวิจัย ลักษณะงานวิจัย สูตรคำนวณ การสรุปผล และการแปลความหมาย แล้วตามด้วยตัวอย่างการคำนวณ

2. เมื่อจบสถิติแต่ละชนิด จะมีการนำเสนอตัวอย่างที่หลากหลายจากงานวิจัยสาขาวิชาต่างๆ แล้วตามด้วยตัวอย่างรายชื่องานวิจัยที่ใช้สถิติแต่ละชนิด เพื่อให้ผู้อ่านได้ศึกษาค้นคว้าเพิ่มเติม

ในการพิมพ์ครั้งนี้ ได้มีการปรับปรุงและเพิ่มเติมเนื้อหา ตลอดจนมีการจัดลำดับบทต่างๆ ใหม่ เพื่อให้มีความชัดเจนและสมบูรณ์ยิ่งขึ้น ดังนี้

1. เพิ่มบทใหม่ ชื่อบทว่า สหสัมพันธ์และการถดถอย โดยปรับปรุงเพิ่มเติมจากบทที่ 14 (สหสัมพันธ์) และบทที่ 15 (สหสัมพันธ์พหุคูณ) เดิม เพื่อให้มีเนื้อหาครอบคลุมทั้งสหสัมพันธ์อย่างง่าย สหสัมพันธ์พหุคูณ การถดถอยอย่างง่าย และการถดถอยพหุคูณ

2. เพิ่มบทใหม่ ชื่อบทว่า สถิติวิเคราะห์ตัวแปรพหุนาม โดยในบทนี้ กล่าวถึงแนวคิด หลักการใช้ และตัวอย่างงานวิจัย

3. สำหรับสถิติการวิเคราะห์ความแปรปรวน ซึ่งอยู่ในบทที่ 12 เดิม มีการปรับปรุงเพิ่มเติม ดังนี้

- 3.1 แยกเนื้อหาการวิเคราะห์ความแปรปรวนในบทที่ 12 เดิมออกเป็น 2 บท คือ บทที่ว่าด้วยการวิเคราะห์ความแปรปรวนแบบมีหนึ่งตัวประกอบ และบทที่ว่าด้วยการวิเคราะห์ความแปรปรวนแบบมีสองตัวประกอบ โดยมีการเปลี่ยนแปลงและเพิ่มเติมตัวอย่างงานวิจัยต่างๆ ในทั้ง 2 บท

- 3.2 ในหัวข้อการวิเคราะห์ความแปรปรวนแบบมีหนึ่งตัวประกอบแบบวัดซ้ำ ได้เพิ่มตัวอย่างงานวิจัยเชิงทดลองที่ใช้สถิติชนิดนี้

4. เพิ่มเดิมตัวอย่างงานวิจัยที่เป็นปัจจุบันและหลากหลายมากขึ้น เพื่อให้ครอบคลุมสถิติประเภทต่างๆ

ตัวอย่างงานวิจัยต่างๆ ที่เสนอในแต่ละบท ผู้เขียนได้คัดเลือกมาโดยใช้เกณฑ์ 2 ข้อ คือ (1) เป็นงานวิจัยที่ใช้ระเบียบวิธีวิจัยที่เหมาะสม ชัดเจน (2) เป็นงานวิจัยที่ศึกษาตัวแปรที่น่าสนใจ และมีความสำคัญต่อการจัดการศึกษา เช่น ความคิดสร้างสรรค์ทางคณิตศาสตร์ ทักษะการแสวงหาความรู้ด้วยตนเอง เป็นต้น

ผู้เขียนหวังเป็นอย่างยิ่งว่า หนังสือเล่มนี้จะช่วยให้นิสิต นักศึกษา ทั้งระดับปริญญาตรี ระดับบัณฑิตศึกษา ตลอดจนนักวิจัย มีแนวทางในการเลือกใช้สถิติเพื่อการวิจัยได้อย่างถูกต้องและเหมาะสมยิ่งขึ้น

ความดีของหนังสือเล่มนี้ ขอมอบแด่คุณพ่อ คุณแม่ บุรพาจารย์ทุกท่าน และขอขอบคุณเจ้าของผลงานทุกท่านที่ได้ระบุไว้ในแต่ละบทและในบรรณานุกรม

รศ.ชูศรี วงศ์รัตน์
พฤษภาคม 2564

สารบัญ

	หน้า
บทที่ 1 ความเข้าใจเบื้องต้นเกี่ยวกับ “สถิติ”	1
1.1 ความหมายของคำว่า “สถิติ”	1
1.2 คำจำกัดความของคำสำคัญที่ใช้ในวิชาสถิติ	2
1.3 ตัวแปร	5
1.4 สเกลของการวัด	7
1.5 ความรู้พื้นฐานทางคณิตศาสตร์	12
แบบฝึกหัดที่ 1	15
บทที่ 2 การแจกแจงความถี่	17
2.1 ความหมายของการแจกแจงความถี่	17
2.2 วิธีการแจกแจงความถี่	22
2.3 คำศัพท์ที่เกี่ยวข้องกับการแจกแจงความถี่	27
2.4 การแจกแจงความถี่ด้วยกราฟแท่งหรือฮิสโทแกรม	29
2.5 การแจกแจงความถี่ด้วยรูปหลายเหลี่ยมแห่งความถี่	30
2.6 การแจกแจงความถี่สะสม	31
2.7 โค้งความถี่สะสม	31
2.8 ลักษณะโค้งต่างๆ ที่เกิดจากการแจกแจงความถี่	32
แบบฝึกหัดที่ 2	34
บทที่ 3 การวัดแนวโน้มเข้าสู่ส่วนกลาง	36
3.1 ตัวกลางเลขคณิต หรือค่าเฉลี่ย	37
3.2 มัธยฐาน	45
3.3 ฐานนิยม	48
3.4 การเลือกใช้วิธีการวัดแนวโน้มเข้าสู่ส่วนกลาง	49
3.5 ความสัมพันธ์ระหว่างค่าเฉลี่ย มัธยฐาน และฐานนิยม	50
3.6 ข้อสังเกตในการใช้วิธีการวัดแนวโน้มเข้าสู่ส่วนกลางด้วยวิธีต่างๆ	51
3.7 ลักษณะของงานวิจัยที่มีการคำนวณหาค่าเฉลี่ย	53
3.8 ตัวอย่างงานวิจัยที่หาค่าเฉลี่ย	55
3.9 ตัวอย่างรายชื่องานวิจัยที่มีการคำนวณหาค่าเฉลี่ย	59
แบบฝึกหัดที่ 3	61
บทที่ 4 การวัดการกระจาย	63
4.1 พิสัย	67
4.2 ความเบี่ยงเบนควอไทล์	68

4.3 ความเบี่ยงเบนเฉลี่ย	70
4.4 ความแปรปรวนและความเบี่ยงเบนมาตรฐาน	72
4.5 การแปลความหมายค่าความเบี่ยงเบนมาตรฐานโดยประยุกต์ใช้กับ การแจกแจงปกติ	90
4.6 ตัวอย่างรายชื่องานวิจัยที่มีการคำนวณค่าสถิติพื้นฐาน	91
แบบฝึกหัดที่ 4	94
บทที่ 5 การแจกแจงปกติ	96
5.1 โถงปกติ	97
5.2 คุณสมบัติที่สำคัญของโถงปกติ	101
5.3 การหาพื้นที่ใต้โถง	102
แบบฝึกหัดที่ 5	111
บทที่ 6 ทฤษฎีการเลือกตัวอย่าง	113
6.1 การแจกแจงของกลุ่มตัวอย่าง	113
6.2 การแจกแจงของค่าเฉลี่ย	115
6.3 การแจกแจงของผลต่างของค่าเฉลี่ย	117
6.4 การแจกแจงแบบท	118
6.5 การแจกแจงแบบโค-สแควร์	119
6.6 การแจกแจงแบบเอฟ	120
แบบฝึกหัดที่ 6	123
บทที่ 7 การทดสอบสมมติฐาน	124
7.1 ความหมายของสมมติฐาน	125
7.2 ประเภทของสมมติฐาน	125
7.3 ขั้นตอนการทดสอบสมมติฐาน	130
7.4 ชนิดของความคลาดเคลื่อน	131
7.5 ระดับนัยสำคัญ	133
7.6 การทดสอบสมมติฐานแบบมีทิศทาง และแบบไม่มีทิศทาง	135
แบบฝึกหัดที่ 7	139
บทที่ 8 การทดสอบเกี่ยวกับค่าเฉลี่ยและสัดส่วน	140
8.1 การทดสอบเกี่ยวกับค่าเฉลี่ยในหนึ่งตัวอย่าง	140
8.1.1 สูตรการทดสอบค่าซี (Z-test) และการทดสอบค่าที (t-test)	141
8.1.2 ขั้นตอนในการทดสอบ	143
8.1.3 วิธีเปิดตารางหาค่า Z	144
8.1.4 การเปรียบเทียบค่า Z ที่คำนวณได้กับค่า Z ที่เปิดจากตาราง	145

8.1.5	วิธีเปิดตารางหาค่า t	148
8.1.6	การเปรียบเทียบค่า t ที่คำนวณได้กับค่า t ที่ได้จากตาราง	148
8.1.7	การแปลความหมายผลการทดสอบ	149
8.1.8	ตัวอย่างการคำนวณ	151
8.1.9	ลักษณะงานวิจัยที่มีการใช้การทดสอบค่าเฉลี่ยในหนึ่งตัวอย่าง	153
8.1.10	ตัวอย่างงานวิจัยที่วิเคราะห์โดยใช้ t -test for one sample	154
8.2	การทดสอบความแตกต่างระหว่างค่าเฉลี่ยสองค่าที่ได้จากกลุ่มตัวอย่าง สองกลุ่มที่เป็นอิสระจากกัน	157
8.2.1	ลักษณะกลุ่มตัวอย่าง 2 กลุ่ม ที่เป็นอิสระจากกัน	158
8.2.2	ตัวอย่างงานวิจัยที่ใช้การทดสอบค่าเฉลี่ยและการทดสอบค่าที	159
8.2.3	หลักเกณฑ์การพิจารณางานวิจัยที่ต้องทำการทดสอบโดยใช้การทดสอบ ค่าเฉลี่ยหรือการทดสอบค่าที	161
8.2.4	ลักษณะงานวิจัยที่ใช้การทดสอบค่าเฉลี่ยหรือการทดสอบค่าที	161
8.2.5	ข้อดกลงเบื้องต้น	162
8.2.6	ลักษณะข้อมูลที่ใช้การทดสอบค่าเฉลี่ยหรือการทดสอบค่าที	163
8.2.7	ค่าต่างๆ ที่เกี่ยวข้องกับการคำนวณ	164
8.2.8	สูตรการทดสอบค่าเฉลี่ย (Z-test) และการทดสอบค่าที (t-test)	164
8.2.9	กรณีตัวอย่าง (Case study) งานวิจัยที่ใช้ t -test for independent Samples	167
8.2.10	สรุปการใช้สูตรการทดสอบค่าเฉลี่ย (Z-test) และการทดสอบค่าที (t-test)	168
8.2.11	แนวทางการตัดสินใจเลือกใช้ t -test ในทางปฏิบัติ	170
8.2.12	การทดสอบความแปรปรวน	170
8.2.13	การตั้งสมมติฐานทางสถิติ	171
8.2.14	ขั้นตอนการทดสอบ	172
8.2.15	วิธีเปิดตารางหาค่า Z	174
8.2.16	การเปรียบเทียบค่า Z ที่คำนวณได้กับค่า Z ที่เปิดจากตาราง	174
8.2.17	วิธีเปิดตารางหาค่า t	174
8.2.18	การเปรียบเทียบค่า t ที่คำนวณได้กับค่า t ที่เปิดจากตาราง	174
8.2.19	การแปลความหมายผลการทดสอบ	174
8.2.20	ตัวอย่างการคำนวณ	176
8.2.21	ตัวอย่างงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้การทดสอบค่าที (t-test for independent samples)	186

8.2.22 การนำเสนอตารางแสดงผลการวิเคราะห์ข้อมูลโดยใช้ t-test for independent samples	189
8.2.23 ตัวอย่างการใช้สถิติ t-test for independent samples เมื่อวิเคราะห์ด้วยโปรแกรมสำเร็จรูปโดยใช้คำสั่ง Independent sample T-test	189
8.2.24 ตัวอย่างรายชื่องานวิจัยที่ใช้ t-test for independent samples	191
8.3 การทดสอบความแตกต่างระหว่างค่าเฉลี่ย 2 ค่า ที่ได้จากกลุ่มตัวอย่างสองกลุ่มที่ไม่เป็นอิสระจากกันหรือกลุ่มตัวอย่างสัมพันธ์กัน	193
8.3.1 ลักษณะของกลุ่มตัวอย่าง	193
8.3.2 ลักษณะงานวิจัยที่ทดสอบโดยใช้ t-test for dependent samples	194
8.3.3 สูตรการทดสอบ	196
8.3.4 ลักษณะของข้อมูลเมื่อบรรจุในตาราง	196
8.3.5 ขั้นตอนการทดสอบค่าที่	197
8.3.6 การแปลความหมาย	198
8.3.7 ตัวอย่างการคำนวณ	198
8.3.8 ตัวอย่างการใช้สถิติ t-test for dependent samples เมื่อวิเคราะห์ด้วยโปรแกรมสำเร็จรูปโดยใช้คำสั่ง Paired-Sample T-test	202
8.3.9 ตัวอย่างงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้การทดสอบค่าที่ (เมื่อกลุ่มตัวอย่างเป็นแบบ Dependent samples)	204
8.3.10 ตัวอย่างรายชื่องานวิจัยที่ใช้ t-test for dependent samples	209
8.4 การทดสอบความแตกต่างระหว่างสัดส่วน	211
8.4.1 ตัวอย่างงานวิจัยที่ใช้การทดสอบความแตกต่างระหว่างสัดส่วน	211
8.4.2 สูตรการทดสอบ	211
8.4.3 ขั้นตอนการทดสอบ	212
8.4.4 ตัวอย่างการคำนวณ	213
แบบฝึกหัดที่ 8	215
บทที่ 9 การวิเคราะห์ความแปรปรวนแบบมีหนึ่งตัวประกอบ	219
9.1 ความเข้าใจพื้นฐานบางประการเกี่ยวกับการวิเคราะห์ความแปรปรวน	220
9.2 ข้อตกลงเบื้องต้นในการวิเคราะห์ความแปรปรวน	221
9.3 ผลบวกของคะแนนเบี่ยงเบนยกกำลังสอง	221
9.4 การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ	222
ก. ลักษณะงานวิจัยที่ใช้การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ	223
ข. ลักษณะข้อมูลที่ใช้ในการวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ	225
ค. สูตรการทดสอบ	228

ง. ขั้นตอนการทดสอบ	230
จ. การสรุปผลและแปลความหมาย	230
ฉ. ขั้นตอนการคำนวณค่า F	231
ช. ตัวอย่างการคำนวณ	234
ซ. การทดสอบความเท่ากันของความแปรปรวนหลายๆ ตัว	237
ณ. การเปรียบเทียบพหุคูณ	240
ญ. การวิเคราะห์ด้วยโปรแกรมสำเร็จรูปโดยใช้คำสั่ง One-Way ANOVA	251
ฎ. ตัวอย่างงานวิจัยที่ใช้การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ	255
9.5 การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ แบบวัดซ้ำ	267
ก. ลักษณะงานวิจัยที่ใช้การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ แบบวัดซ้ำ	267
ข. ลักษณะข้อมูลที่ใช้การวิเคราะห์ความแปรปรวนแบบมี 1 ตัวประกอบ แบบวัดซ้ำ	267
ค. สูตรการทดสอบ	268
ง. ตัวอย่างการคำนวณ	269
จ. ตัวอย่างงานวิจัยที่วิเคราะห์โดยใช้ One-Way analysis of variance with repeated measures	273
9.6 ตัวอย่างรายงานวิจัยที่มีการวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ ความแปรปรวนแบบมีหนึ่งตัวประกอบ	284
แบบฝึกหัดที่ 9	285
บทที่ 10 การวิเคราะห์ความแปรปรวนแบบมีสองตัวประกอบ	288
10.1 แบบแผนการทดลองแบบแฟกทอเรียล	289
10.2 ตัวอย่างวิธีการดำเนินการวิจัยของงานวิจัยที่วิเคราะห์ด้วยการวิเคราะห์ ความแปรปรวนแบบมี 2 ตัวประกอบ	291
10.3 ข้อค้นพบที่ได้จากแบบแผนการทดลองแบบแฟกทอเรียล	292
10.4 ตัวอย่างการวิเคราะห์ผลหลักและผลของปฏิสัมพันธ์	294
10.5 ตัวอย่างของแบบแผนการทดลองแบบแฟกทอเรียลที่มีกราฟแสดง ทั้งกรณีที่มีปฏิสัมพันธ์และไม่มีปฏิสัมพันธ์	298
10.6 การเขียนกราฟแสดงปฏิสัมพันธ์	302
10.7 ลักษณะตารางบรรจุข้อมูลที่ใช้การวิเคราะห์ความแปรปรวนแบบมี 2 ตัวประกอบ	304
10.8 ขั้นตอนการทดสอบ	305
10.9 สูตรการทดสอบ	307

10.10	การสรุปผลและแปลความหมาย	309
10.11	ขั้นตอนการคำนวณค่า F	310
10.12	ตัวอย่างการคำนวณ	313
10.13	ตัวอย่างงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ความแปรปรวน แบบมี 2 ตัวประกอบ	318
10.14	ตัวอย่างงานวิจัยที่มีการวิเคราะห์ข้อมูลโดยใช้การวิเคราะห์ ความแปรปรวนแบบมี 2 ตัวประกอบ	328
	แบบฝึกหัดที่ 10	332
บทที่ 11	การวิเคราะห์ความแปรปรวนร่วม	334
11.1	ลักษณะการวิจัยที่ต้องใช้การวิเคราะห์ความแปรปรวนร่วม	334
11.2	ลักษณะข้อมูลที่ใช้การวิเคราะห์ความแปรปรวนร่วม	336
11.3	ขั้นตอนในการวิเคราะห์ความแปรปรวนร่วม	336
11.4	ตัวอย่างการคำนวณ	338
11.5	ตัวอย่างงานวิจัยที่มีการวิเคราะห์ด้วยสถิติการวิเคราะห์ ความแปรปรวนร่วม	346
11.6	ตัวอย่างงานวิจัยที่วิเคราะห์ด้วยสถิติการวิเคราะห์ ความแปรปรวนร่วม	350
	แบบฝึกหัดที่ 11	352
บทที่ 12	สหสัมพันธ์และการวิเคราะห์ถดถอย	354
12.1	ความหมายของคำว่า “สหสัมพันธ์”	354
12.2	การเลือกวิธีวัดความสัมพันธ์ระหว่างตัวแปร	355
12.3	ลักษณะที่สำคัญของสหสัมพันธ์	356
12.4	ลักษณะของความสัมพันธ์ระหว่างข้อมูล	357
12.5	สหสัมพันธ์อย่างง่ายและสหสัมพันธ์พหุคูณ	359
12.6	การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่าย	359
	ก. สูตรการคำนวณและตัวอย่างการคำนวณ	360
	ข. การแปลความหมายค่าสัมประสิทธิ์สหสัมพันธ์	362
	ค. การทดสอบนัยสำคัญของค่า r	364
	ง. ตัวอย่างงานวิจัยที่คำนวณค่าสัมประสิทธิ์สหสัมพันธ์โดยใช้ สหสัมพันธ์แบบเพียร์สัน	366
12.7	การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ	367
	ก. ตัวอย่างงานวิจัยที่ใช้สหสัมพันธ์พหุคูณ	368
	ข. สูตรการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ	368

ค. การทดสอบนัยสำคัญของ R	370
12.8 การวิเคราะห์ถดถอยอย่างง่ายและการวิเคราะห์ถดถอยพหุคูณ	371
ก. การวิเคราะห์ถดถอยอย่างง่าย	374
ข. การวิเคราะห์ถดถอยพหุคูณ	375
ค. ตัวอย่างการวิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูปโดยใช้คำสั่ง Linear Regression	376
ง. ตัวอย่างงานวิจัยที่ใช้สถิติการวิเคราะห์ถดถอยพหุคูณ	379
12.9 ตัวอย่างรายงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้สหสัมพันธ์พหุคูณและ การวิเคราะห์ถดถอยพหุคูณ	385
แบบฝึกหัดที่ 12	387
บทที่ 13 สถิติวิเคราะห์ตัวแปรพหุนาม	393
13.1 แนวคิดการใช้สถิติวิเคราะห์ตัวแปรพหุนาม	393
13.2 หลักการใช้สถิติวิเคราะห์ตัวแปรพหุนาม	393
13.3 ตัวอย่างงานวิจัย	394
13.3.1 ตัวอย่างงานวิจัยที่ใช้สถิติการวิเคราะห์ความแปรปรวนตัวแปร พหุนามแบบสองทาง	394
13.3.2 ตัวอย่างงานวิจัยที่ใช้สถิติการวิเคราะห์ถดถอยพหุคูณตัวแปร พหุนาม	401
13.3.3 ตัวอย่างงานวิจัยที่ใช้สถิติการวิเคราะห์สหสัมพันธ์คาโนนิกอล	406
13.3.4 ตัวอย่างงานวิจัยที่ใช้สถิติการวิเคราะห์จำแนก	407
บทที่ 14 การทดสอบไค-สแควร์	409
14.1 การทดสอบข้อมูลที่มี 1 มิติ	411
ก. คำที่เกี่ยวข้องกับการทดสอบ	411
ข. สูตรการคำนวณ	415
ค. ขั้นตอนการทดสอบ	415
ง. การเปิดตารางค่า X^2	416
จ. การแปลความหมาย	417
ฉ. ลักษณะข้อมูลเมื่อบรรจุในตาราง	417
ช. ตัวอย่างการคำนวณ	418
14.2 การทดสอบความเป็นอิสระ	420
ก. ความหมายของการทดสอบความเป็นอิสระและลักษณะข้อมูล	420
ข. ลักษณะตารางบรรจุข้อมูล	421
ค. สูตรการคำนวณ	422

ง. ขั้นตอนการคำนวณ	422
จ. การสรุปผลและการแปลความหมาย	423
ฉ. ตัวอย่างการคำนวณ	424
ช. ตัวอย่างการใช้สถิติ X^2 -test เพื่อทดสอบความสัมพันธ์เมื่อวิเคราะห์ ด้วยโปรแกรมสำเร็จรูปโดยใช้คำสั่ง Crosstab	427
ซ. ตัวอย่างงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้การทดสอบไค-สแควร์	429
ณ. Contingency table แบบ 2×2	430
ญ. ค่าแก้ของเยตส์	432
ฎ. การหาระดับความสัมพันธ์จากค่าไค-สแควร์	433
14.3 ตัวอย่างงานวิจัยที่มีการวิเคราะห์ข้อมูลโดยใช้การทดสอบ ไค-สแควร์	434
แบบฝึกหัดที่ 14	435
บทที่ 15 สถิตินอนพาราเมตริก	436
15.1 การทดสอบเครื่องหมาย	436
15.2 การทดสอบของวิลค็อกซัน	441
15.3 การทดสอบของแมน-วิทนี	443
15.4 การทดสอบของครัสคาล-วอลลิส	449
15.5 ตัวอย่างงานวิจัยที่วิเคราะห์ข้อมูลโดยใช้ Wilcoxon matched-pairs signed-ranks Test	451
15.6 ตัวอย่างงานวิจัยที่มีการวิเคราะห์ข้อมูลโดยใช้สถิตินอนพารา เมตริก	453
บรรณานุกรม	456
ตารางสถิติ	463

บทที่ 1

ความเข้าใจเบื้องต้นเกี่ยวกับ “สถิติ”

1.1 ความหมายของคำว่า “สถิติ”

คำว่า “สถิติ”(Statistics) มีความหมายอยู่ 2 ประการคือ

1. หมายถึง ตัวเลขสถิติ ซึ่งเป็นตัวเลขที่แสดงถึงข้อเท็จจริงของสิ่งต่างๆ ซึ่งอาจเป็นบุคคล สิ่งของ หรือปรากฏการณ์ต่างๆ ฯลฯ คำว่า สถิติ ในความหมายนี้จะต้องประกอบด้วยตัวเลขหลายๆ ตัว ตัวอย่างเช่น ตัวเลขแสดงจำนวนผู้สมัครเรียนต่อระดับปริญญาโท สาขาวิชาต่างๆ ของมหาวิทยาลัยศรีนครินทรวิโรฒ ตั้งแต่ปี 2540 ถึงปัจจุบันเป็นตัวเลขสถิติ

2. หมายถึง “สถิติศาสตร์” วิชาสถิติศาสตร์เป็นแขนงหนึ่งของระเบียบวิธีการทางวิทยาศาสตร์ซึ่งประกอบด้วย การเก็บรวบรวมข้อมูล การจัดระบบและการนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการตีความหมายข้อมูล ข้อมูลที่เก็บรวบรวมมาอาจจะได้จากการสำรวจหรือการทดลอง จุดมุ่งหมายที่สำคัญของวิชาสถิติก็คือ เพื่อบรรยายและลงสรุปเกี่ยวกับคุณลักษณะของประชากร

วิชาสถิติศาสตร์แบ่งออกเป็น 2 สาขาใหญ่ๆ คือ

1. สถิติเชิงบรรยาย (Descriptive statistics) เป็นสถิติที่บรรยายให้เห็นคุณลักษณะของสิ่งที่ต้องการศึกษาจากกลุ่มใดกลุ่มหนึ่งโดยเฉพาะ ซึ่งอาจจะเป็นกลุ่มใหญ่หรือกลุ่มเล็กก็ได้ ผลของการศึกษาไม่สามารถนำไปอ้างอิงถึงกลุ่มอื่นได้ เช่น ครูหาน้ำหนักเฉลี่ยของนักเรียนชายห้อง ม.1/1 ได้ 45 กิโลกรัม ครูคนนี้จะสรุปอ้างอิงว่านักเรียนชาย ม.1 ห้องอื่นๆ มีน้ำหนักเฉลี่ย 45 กิโลกรัมด้วยไม่ได้

2. สถิติเชิงอ้างอิง (Inferential statistics) เป็นสถิติที่ศึกษากับกลุ่มตัวอย่าง แล้วสรุปผลที่ศึกษาได้จากกลุ่มตัวอย่างนั้นอ้างอิงไปถึงกลุ่มประชากรด้วยโดยอาศัยทฤษฎีความน่าจะเป็น สถิติแบบนี้สำคัญอยู่ที่กลุ่มตัวอย่างจะต้องเป็นตัวแทนที่ดีของกลุ่มประชากร จึงจะทำให้การสรุปอ้างอิงจากกลุ่มตัวอย่างไปถึงกลุ่มประชากรถูกต้องและเชื่อถือได้ เช่น โรงเรียนแห่งหนึ่งมีนักเรียนชั้น ม.1 จำนวน 300 คน ครูคนหนึ่งต้องการทราบน้ำหนักเฉลี่ยของนักเรียนชั้น ม.1 ทั้งหมดถ้าครูคนนี้ใช้สถิติเชิงอ้างอิงก็ไม่

จำเป็นต้องชั่งน้ำหนักของนักเรียนชั้น ม.1 ทุกคน เพียงแต่สุ่มกลุ่มตัวอย่างมาจำนวนหนึ่ง เช่น สุ่มมา 100 คน ชั่งน้ำหนักนักเรียน 100 คนนี้แล้วหาค่าเฉลี่ย สมมติว่าได้ 40 กิโลกรัม ครูคนนี้จะสรุปผลอ้างอิงไปถึงนักเรียนชั้น ม.1 ทั้ง 300 คนได้โดยใช้สถิติเชิงอ้างอิง

สถิติเชิงอ้างอิงยังแบ่งเป็น 2 ประเภทย่อยคือ

2.1 สถิติประมาณ (Estimation statistics) เป็นสถิติที่ใช้ประมาณค่าพารามิเตอร์ เช่น ประมาณค่า μ จาก $\bar{X}$ เป็นต้น

2.2 สถิติทดสอบ (Test statistics) เป็นสถิติเกี่ยวกับการทดสอบสมมุติฐาน

1.2 คำจำกัดความของคำสำคัญที่ใช้ในวิชาสถิติ

ในการศึกษาสถิติเชิงอ้างอิง มีคำที่ต้องทำความเข้าใจอยู่หลายคำ ได้แก่

1.2.1 ตัวแปร (Variable)

ตัวแปร คือ คุณลักษณะใดๆ ที่มีความแตกต่างกันระหว่างบุคคล (เช่น ตัวแปร เพศ, อายุ, ...) ระหว่างสิ่งของ (เช่น ตัวแปร สีของรถ, ชนิดของรถ, ...) หรือระหว่างสภาพแวดล้อม (เช่น สภาพทางภูมิอากาศ, สภาพโรงเรียน,...)

รายละเอียดจะกล่าวอีกครั้งในหัวข้อ 1.3

1.2.2 ข้อมูล (Data)

ข้อมูล คือ ข้อเท็จจริงเกี่ยวกับตัวแปรที่สนใจจะศึกษา ซึ่งรวบรวมได้จากแต่ละหน่วยของกลุ่มที่ต้องการศึกษา อาจเป็นข้อเท็จจริงที่อยู่ในรูปของตัวเลข เช่น จำนวน ปริมาณ ระยะทาง เป็นต้น ซึ่งมีชื่อเรียกเฉพาะว่า **ข้อมูลเชิงประมาณ (Quantitative data)** หรืออาจเป็นข้อเท็จจริงที่ไม่อยู่ในรูปของตัวเลข เช่น ชื่อ เพศ ที่อยู่ สถานภาพสมรส ระดับการศึกษา เป็นต้น มีชื่อเรียกเฉพาะว่า **ข้อมูลเชิงคุณภาพ (Qualitative data)** หรือเรียกว่าข้อมูลเชิงคุณลักษณะ

ตัวอย่างเช่น

ถ้าเราต้องการศึกษาผลสัมฤทธิ์ทางการเรียนกลุ่มสาระภาษาไทยของนักเรียนชั้นประถมศึกษาปีที่ 6/1 ตัวแปรที่สนใจจะศึกษาคือ ผลสัมฤทธิ์ทางการเรียนกลุ่มสาระภาษาไทย ข้อมูลก็คือคะแนนผลการสอบกลุ่มสาระภาษาไทยของนักเรียนชั้น

ประถมศึกษาปีที่ 6/1 แต่ละคน เป็นข้อมูลเชิงปริมาณ ซึ่งเป็นตัวเลข ถ้ามีนักเรียนชั้นประถมศึกษาปีที่ 6/1 อยู่ 40 คน ก็จะได้ข้อมูลทั้งหมด 40 ตัว

คำว่า data เป็นพหูพจน์ใช้แทนข้อมูลที่มีมากกว่า 1 จำนวน ถ้าเป็นเอกพจน์ใช้ datum

ข้อมูลที่เป็นตัวเลข (Numerical data) มี 2 ชนิด

1) ข้อมูลที่เป็นความถี่ (frequency) ซึ่งเป็นข้อมูลที่ได้จากการแจกแจง หรือจัดประเภท เช่น ในโรงเรียนแห่งหนึ่งมีครูชาย 50 คน ครูหญิง 70 คน เป็นต้น เรียกอีกชื่อหนึ่งว่าข้อมูลจัดประเภท (categorical data)

2) ข้อมูลที่เป็นค่าสเกลต่างๆ (scale values) ซึ่งเป็นข้อมูลที่ได้จากการวัด (Measurement) เช่น คะแนนสอบวิชาต่างๆ ซึ่งวัดได้จากแบบทดสอบ คะแนนที่ได้จากแบบวัดความเครียด น้ำหนักของนักเรียนที่วัดได้จากเครื่องชั่ง เป็นต้น

ระเบียบวิธีทางสถิติทั้งหลายที่ใช้ จะเกี่ยวข้องกับข้อมูล 2 ชนิดดังกล่าวเสมอ

1.2.3 ประชากร (Population)

ประชากรหมายถึง กลุ่มทั้งหมดของคน สัตว์หรือสิ่งของ ซึ่งมีคุณลักษณะร่วมกันตรงกับที่ผู้วิจัยสนใจจะศึกษา เช่น ถ้าต้องการศึกษาความคิดเห็นของนิสิตมหาวิทยาลัยเกษตรศาสตร์ที่มีต่อการปกครองแบบประชาธิปไตย นิสิตทุกคนในมหาวิทยาลัยเกษตรศาสตร์ก็เป็นประชากร

1.2.4 กลุ่มตัวอย่าง (Sample)

กลุ่มตัวอย่างหมายถึง บางส่วนของกลุ่มประชากร

1.2.5 ค่าพารามิเตอร์ (Parameter)

ค่าพารามิเตอร์หมายถึง ค่าที่ใช้บอกคุณลักษณะใดๆ ของกลุ่มประชากร เช่น รายได้เฉลี่ยต่อวันของเด็กชายพวงมาลัยตามสี่แยกต่างๆ ในกรุงเทพมหานคร อายุเฉลี่ยของนิสิตระดับปริญญาโทในมหาวิทยาลัยศรีนครินทรวิโรฒ เป็นต้น

สัญลักษณ์ที่ใช้แทนค่าพารามิเตอร์ที่พบบ่อยคือ μ (มิว) ซึ่งแทนค่าเฉลี่ยของประชากร และ σ (ซิกมา) ซึ่งแทนค่าความเบี่ยงเบนมาตรฐานของประชากร

1.2.6 ค่าสถิติ (Statistic)

ค่าสถิติหมายถึง ค่าที่คำนวณได้จากข้อมูลที่เก็บจากกลุ่มตัวอย่าง ค่าสถิตินี้จะใช้ประมาณค่าพารามิเตอร์ โดย $\bar{X}$ แทนค่าเฉลี่ยของกลุ่มตัวอย่าง และ S หรือ SD แทนค่าความเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

สมมติว่าในกรุงเทพฯมีเด็กขายพวงมาลัยทั้งหมด 1,000 คน (กลุ่มประชากร) ถ้าสอบถามรายได้ที่เด็กเหล่านี้ได้รับใน 1 วัน โดยถามทั้ง 1,000 คนแล้วนำรายได้ของทั้ง 1,000 คน มาหาค่าเฉลี่ย ค่าเฉลี่ยที่คำนวณได้คือค่าพารามิเตอร์ (μ) แต่ถ้าสุ่มเด็กขายพวงมาลัยมาจากสี่แยกต่างๆ 300 คน (กลุ่มตัวอย่าง) สอบถามรายได้ที่เด็กเหล่านี้ได้รับใน 1 วัน แล้วคำนวณหาค่าเฉลี่ย ค่าเฉลี่ยที่คำนวณได้คือค่าสถิติ ($\bar{X}$) ซึ่งค่าสถิติตัวนี้จะนำไปใช้ประมาณค่าพารามิเตอร์ได้โดยใช้สถิติเชิงอ้างอิง

ข้อสังเกต ถ้าเป็นวิชาสถิติ ใช้คำว่า Statistics (อยู่ในรูปพหูพจน์)

ตาราง 1.1 จุดมุ่งหมายของวิชาสถิติและสัญลักษณ์ที่ใช้

ก เพื่อบรรยายคุณลักษณะของกลุ่มตัวอย่าง	ข เพื่ออ้างอิงคุณลักษณะของกลุ่มประชากร
1. การวัดแนวโน้มเข้าสู่ส่วนกลาง เช่น - ค่าเฉลี่ย ($\bar{X}$) - มัชฐาน (m) - สัดส่วน (p)	1. การวัดแนวโน้มเข้าสู่ส่วนกลาง เช่น - ค่าเฉลี่ย (μ) - มัชฐาน (Md) - สัดส่วน (P)
2. การวัดการกระจาย - ความแปรปรวน (S^2) - ความเบี่ยงเบนมาตรฐาน (S)	2. การวัดการกระจาย - ความแปรปรวน (σ^2) - ความเบี่ยงเบนมาตรฐาน (σ)
3. ความเบ้ (SK)	3. ความเบ้ (α_3)
4. ความโค้ง (K)	4. ความโค้ง (α_4)
5. สหสัมพันธ์ (r)	5. สหสัมพันธ์ (p)

หมายเหตุ สัญลักษณ์ในวงเล็บในช่อง ก คือค่าสถิติ

สัญลักษณ์ในวงเล็บในช่อง ข คือค่าพารามิเตอร์

ข้อสังเกต ค่าสถิติแต่ละตัวที่ได้จากข้อมูลที่เก็บจากกลุ่มตัวอย่าง จะมีพารามิเตอร์ที่สอดคล้องกันเสมอ เช่น $\bar{X}$ สอดคล้องกับ μ , S สอดคล้องกับ σ เป็นต้น

1.3 ตัวแปร (Variable)

1.3.1 ความหมายของตัวแปร

คุณลักษณะใดๆ ของบุคคล สิ่งของ หรือสภาพแวดล้อม ซึ่งมีค่าได้ต่างๆ กันไม่ว่าจะเป็นค่าที่อยู่ในรูปของปริมาณหรือคุณภาพก็ตามเรียกว่าตัวแปร เช่น ความสูงของนักเรียนกลุ่มหนึ่งเป็นตัวแปร เพราะความสูงของนักเรียนแต่ละคนในกลุ่มนี้มีค่าได้ต่างๆ กันออกไป บางคนสูง 145 เซนติเมตร บางคนสูง 171 เซนติเมตร บางคนสูง 153 เซนติเมตร ค่าในตัวอย่างนี้อยู่ในรูปของปริมาณ เพศ ก็เป็นตัวแปร ซึ่งมี 2 ค่าคือ เพศชายกับเพศหญิง ค่าในที่นี้อยู่ในรูปของคุณลักษณะ เป็นต้น

ในการทำวิจัยจะมีตัวแปรที่ศึกษาอยู่ 2 ชนิด คือ ตัวแปรตาม (Dependent variable) และ ตัวแปรอิสระหรือตัวแปรต้น (Independent variable)

ตัวแปรตาม คือ ตัวแปรที่ผู้วิจัยต้องการศึกษาจากกลุ่มตัวอย่าง เช่น ความพึงพอใจต่อการบริการของพนักงานธนาคาร เจตคติต่อวิชาภาษาไทยของนักเรียน เป็นต้น

ตัวแปรอิสระ หรือ ตัวแปรต้น คือ ตัวแปรที่ผู้วิจัยต้องการศึกษาว่าเป็นสาเหตุของตัวแปรตามหรือไม่

1.3.2 ชนิดของตัวแปร

ชนิดของตัวแปรที่เกี่ยวข้องกับการเลือกใช้สถิติ มีดังนี้

ก) ตัวแปรเชิงปริมาณและตัวแปรเชิงคุณภาพ (Quantitative and Qualitative variable)

ตัวแปรเชิงปริมาณ คือ ตัวแปรที่มีค่าต่างๆ ซึ่งอยู่ในรูปของจำนวนหรือขนาด โดยที่จำนวนหรือขนาดเหล่านี้ใช้บอกปริมาณความมากน้อยได้ ตัวอย่างของตัวแปรเชิงปริมาณ ได้แก่ อายุ ความสูง รายได้ ขนาดของประชากร ขนาดของครอบครัว ระยะเวลาที่นักโทษถูกจำคุก อัตราการเกิด และคุณลักษณะอื่นๆ ที่อยู่ในรูปของตัวเลข คนแต่ละคนจะมีอายุต่างๆ กันไป บางคนมีอายุมาก บางคนมีอายุน้อย เมืองแต่ละเมืองก็มีขนาดต่างๆ กันไป บางเมืองมีขนาดใหญ่ บางเมืองมีขนาดเล็ก ตัวแปรต่างๆ เหล่านี้สามารถวัดได้ และแต่ละค่าที่วัดออกมาได้เหล่านี้เรียกว่า “Variate” เขตของ Variate เช่น ความสูงของทหารแต่ละคนในกองทัพนั่ง หรือจำนวนคนในแต่ละครอบครัวซึ่งมีทั้งหมด

1,000 ครอบครัว สามารถจัดเรียงขนาดจากเล็กที่สุดไปถึงใหญ่ที่สุดได้ ค่าที่วัดได้นี้ถือว่าเป็นค่าสเกล (Scale values) การจัดเรียงข้อมูลดังกล่าวนี้เป็นลักษณะที่สำคัญที่ช่วยบอกความแตกต่างจากตัวแปรเชิงคุณภาพได้

ตัวแปรเชิงคุณภาพ คือ ตัวแปรที่มีค่าได้ต่างๆ กัน แต่ค่าดังกล่าว ไม่ได้อยู่ในรูปของจำนวนหรือขนาด ส่วนใหญ่จะอยู่ในรูปของคุณภาพหรือชนิด คุณภาพเหล่านี้เรียกว่า “คุณลักษณะ”(Attribute) เช่น “เพศ” เป็นคุณลักษณะที่มีค่าได้ 2 ชนิดคือ “ชาย” กับ “หญิง” “สถานภาพสมรส” ก็เป็นคุณลักษณะที่มีลักษณะได้ต่างๆ กัน เช่น เป็นโสด แต่งงานแล้ว เป็นหม้าย หย่าร้างกัน เป็นต้น “เชื้อชาติ” ของคนมีได้ต่างๆ กัน เช่น ไทย จีน พม่า อินเดีย เป็นต้น คุณลักษณะ (Attribute) ที่จัดเป็นตัวแปรเชิงคุณภาพนี้ไม่ถือว่าเป็นค่าสเกลเพราะไม่สามารถจัดเรียงลำดับตามขนาดได้ เพศ เชื้อชาติ สถานภาพสมรสไม่สามารถพิจารณาได้ว่าค่าใด “สูงกว่า” “ใหญ่กว่า” หรือ “มากกว่า” ค่าใดได้

ในการใช้สถิติวิเคราะห์ข้อมูลจะต้องพิจารณาว่า ข้อมูลที่จะนำมาวิเคราะห์นั้นเป็นข้อมูลเชิงปริมาณ (Quantitative data) หรือข้อมูลเชิงคุณภาพ (Qualitative data) แล้วเลือกใช้สถิติให้เหมาะสมกับชนิดของข้อมูล

ข) ตัวแปรที่มีค่าต่อเนื่องและมีค่าไม่ต่อเนื่อง (Continuous and Discrete variable)

ในการวิเคราะห์ข้อมูลเชิงปริมาณ (Quantitative data) ผู้วิเคราะห์จะต้องทราบความแตกต่างระหว่างตัวแปรที่มีค่าต่อเนื่อง (Continuous variable) กับตัวแปรที่มีค่าไม่ต่อเนื่อง (Discrete variable)

ตัวแปรที่มีค่าต่อเนื่อง คือ ตัวแปรที่มีค่าใดๆ ก็ได้ในพิสัยหนึ่งๆ ที่กำหนดให้ ค่าที่อยู่ในพิสัยนี้มีจำนวนมากมายนับไม่ถ้วน

ตัวแปรที่มีค่าไม่ต่อเนื่อง คือ ตัวแปรที่ไม่อาจมีค่าได้ทุกค่าในพิสัยหนึ่งที่กำหนดให้ ค่าที่อยู่ในพิสัยนี้ไม่ต่อเนื่องกันและนับจำนวนได้ว่ามีกี่ค่า

ตัวอย่างตัวแปรที่มีค่าต่อเนื่อง เช่น “น้ำหนัก” ระหว่างน้ำหนัก 50 กิโลกรัม กับ 51 กิโลกรัม จะมีค่าน้ำหนักต่างๆ ที่เป็นไปได้นับไม่ถ้วน เช่น 50.01 กิโลกรัม 50.53 กิโลกรัม 50.99 กิโลกรัม ฯลฯ “เวลา” “อายุ” ก็เป็นตัวอย่างของตัวแปรที่มีค่าต่อเนื่อง

ตัวอย่างตัวแปรที่มีค่าไม่ต่อเนื่อง เช่น จำนวนหนังสือประเภทต่างๆ ในห้องสมุด จำนวนบุตรในครอบครัว ฯลฯ

ตัวแปรและชนิดของตัวแปร แสดงเป็นแผนภาพได้ดังนี้



1.4 สเกลของการวัด (Scales of measurement)

สเกลของการวัด มีความสำคัญยิ่งต่อการพิจารณาเลือกใช้สถิติ

“การวัด” (measurement) คือการกำหนดตัวเลขหรือสัญลักษณ์อย่างมีระบบให้กับสิ่งของหรือเหตุการณ์เพื่อแทนปริมาณหรือคุณภาพของสิ่งที่ต้องการวัด ค่าที่วัดได้แบ่งได้เป็น 4 ชนิด เรียงลำดับจากมาตราวัดสูงสุด ไปหามาตราวัดต่ำสุด ดังนี้

1. มาตราอัตราส่วน (Ratio scale)
2. มาตราอันตรภาค (Interval scale)
3. มาตราเรียงอันดับ (Ordinal scale)
4. มาตรานามบัญญัติ (Nominal scale)

1.4.1 มาตราอัตราส่วน (Ratio scale)

มาตราของการวัดใดที่มีคุณสมบัติครบ 3 ประการคือ

- 1) เป็นตัวเลขที่แสดงถึงปริมาณความมากน้อย (Magnitude)
- 2) ความแตกต่างของแต่ละหน่วยเท่ากัน (Equal interval) และ
- 3) มีศูนย์แท้ (Absolute zero)

เรียกว่า อัตราส่วน (Ratio scale) ข้อมูลที่จัดอยู่ในมาตรานี้ เช่น “ความสูง” พ่อสูง 160 เซนติเมตร ลูกชายสูง 161 เซนติเมตร ตัวเลข 160 และ 161 เซนติเมตร แสดงถึงปริมาณความสูง และแสดงว่าลูกชายสูงกว่าพ่อ 1 เซนติเมตร ถ้าลูกสาวสูง 159 เซนติเมตร แสดงว่าพ่อสูงกว่าลูกสาว 1 เซนติเมตร หนึ่งเซนติเมตรตัวแรกกับหนึ่งเซนติเมตรตัวหลังห่างเป็นระยะเท่ากัน (Equal interval) และถ้าเรากล่าวว่าสูง 0 เซนติเมตรก็แปลว่าไม่มีความสูงเลย 0 ตัวนี้เรียกว่า ศูนย์แท้ (Absolute zero) ตัวอย่างอื่นๆของข้อมูลที่จัดอยู่ในมาตราอัตราส่วน เช่น น้ำหนัก ความเร็ว ฯลฯ ข้อมูลในมาตรานี้เป็นข้อมูลชนิด Scale values ข้อมูลที่อยู่ในมาตราอัตราส่วนมักเป็นข้อมูลทางฟิสิกส์ ส่วนข้อมูลทางการศึกษา ทางจิตวิทยา ทางสังคมวิทยามักจะไม่อยู่ในมาตรานี้

1.4.2 มาตราอันตรภาค (Interval scale)

มาตราของการวัดที่มีคุณสมบัติเพียง 2 ประการ คือ

- 1) เป็นตัวเลขที่แสดงถึงปริมาณความมากน้อย (Magnitude) และ
- 2) ความแตกต่างของแต่ละหน่วยเท่ากัน (Equal interval)

เรียกว่ามาตราอันตรภาค (Interval scale) ตัวอย่างเช่น การวัดอุณหภูมิ เป็นองศาเซลเซียส วันนี้อุณหภูมิ 33 องศาเซลเซียส เมื่อวานอุณหภูมิ 35 องศาเซลเซียส กล่าวได้ว่าอุณหภูมิวันนี้ต่ำกว่าเมื่อวาน (Magnitude) โดยอุณหภูมิ 33 กับ 35 องศาเซลเซียสต่างกัน 2 องศาเซลเซียส เช่นเดียวกับอุณหภูมิ 30 กับ 32 องศาเซลเซียส (Equal interval) ถ้าวัดอุณหภูมิได้ 0 องศาเซลเซียส ไม่ได้แปลว่าไม่มีความร้อนเลย (ไม่ใช่ศูนย์แท้) ถ้าวันนี้อุณหภูมิ 30 องศาเซลเซียส เมื่อวานอุณหภูมิ 15 องศาเซลเซียส จะบอกว่าอุณหภูมิสูงกว่าเมื่อวาน 2 เท่าไม่ได้ กล่าวคือถ้าข้อมูลที่วัดได้ไม่มีศูนย์แท้จะหาอัตราส่วนไม่ได้ คะแนนที่ได้จากการตอบแบบทดสอบต่างๆ คะแนนที่ได้จากการตอบแบบสอบถามที่เป็นมาตราส่วนประมาณค่า (Rating scale) หรือคะแนนที่ได้จากเครื่องมือการวิจัยอื่นๆ ก็ตาม จัดเป็นข้อมูลในมาตรานี้ (Levin, 1983 : 8) ซึ่งเป็นข้อมูลชนิด Scale values

1.4.3 มาตราเรียงอันดับ (Ordinal scale)

มาตราของการวัดที่มีคุณสมบัติเพียง 1 ประการ คือ

เป็นตัวเลขที่แสดงถึงปริมาณความมากน้อย (Magnitude) เรียกว่า มาตราเรียงอันดับ (Ordinal scale) การวัดในมาตรานี้ ความแตกต่างระหว่างแต่ละหน่วยไม่เท่ากัน (ไม่มี Equal interval) และไม่มีศูนย์แท้ (ไม่มี Absolute zero)

ข้อมูลที่อยู่ในมาตรานี้เป็นตัวเลข ตัวเลขที่ต่างกันแสดงให้เห็นว่าสิ่งนั้นมีคุณสมบัติที่ต่างกัน แต่ตัวเลขนั้นไม่สามารถบอกให้ทราบว่า สิ่งต่างๆเหล่านั้นมีคุณสมบัติแตกต่างกันเป็นปริมาณเท่าใด และตัวเลขที่ต่างกัน 1 หน่วยไม่ได้แทนปริมาณที่เท่ากัน เหมือนข้อมูลในมาตราอันดับหรืออัตราส่วน

ตัวอย่างที่ 1 การประกวดพานดอกไม้ มีการจัดลำดับจากพานที่สวยงามที่สุดเป็นอันดับที่ 1 อันดับที่ 2 และอันดับที่ 3 ความแตกต่างของความสวยงามระหว่างพานอันดับที่ 1 กับอันดับที่ 2 และระหว่างอันดับที่ 2 กับอันดับที่ 3 ไม่เท่ากัน และบอกไม่ได้ว่ามีความสวยงามต่างกันอยู่มากน้อยเท่าใด ข้อมูลในระดับนี้ไม่สามารถนำมาบวก ลบ คูณ หาร กัน

ตัวอย่างที่ 2 การจัดแถวนักเรียนในห้องตามลำดับความสูง คนที่สูงที่สุดจัดเป็นอันดับ 1 สูงรองลงมาจัดเป็นอันดับ 2 จัดเช่นนี้ไปเรื่อยๆ อันดับ 1 กับอันดับ 2 ต่างกัน 1 อันดับ เช่นเดียวกับอันดับ 2 กับอันดับ 3 ซึ่งความสูงที่ได้จากการวัดเป็นเซนติเมตรของ 3 คนนี้อาจเป็นดังนี้

ความแตกต่าง			
อันดับ	ความสูง (เซนติเมตร)	อันดับ	เซนติเมตร
1	78	1	6
2	72	1	1
3	71		

การจัดอันดับนี้จะเห็นว่าระหว่างอันดับ 1 กับอันดับ 2 ต่างกัน 1 หน่วย เช่นเดียวกับระหว่างอันดับ 2 กับอันดับ 3 แต่ถ้ามองในแง่ของความสูงเป็นเซนติเมตรจะเห็นว่า อันดับ 1 กับอันดับ 2 ต่างกัน 6 เซนติเมตร แต่อันดับ 2 กับอันดับ 3 ต่างกัน 1

เซนติเมตร ลักษณะเช่นนี้เรียกว่า ไม่มี Equal interval กล่าวคือ ต่างกัน 1 อันดับเหมือนกันแต่ไม่ได้แทนปริมาณที่เท่ากัน

1.4.4 มาตรฐานนามบัญญัติ (Nominal scale)

มาตรฐานนามบัญญัติเป็นมาตรการวัดที่กำหนดสัญลักษณ์ขึ้นเพื่อเรียกชื่อหรือเพื่อจำแนก หรือจัดประเภทสิ่งของตามลักษณะ เช่น จำแนกคนออกเป็น 2 เพศคือ เพศหญิง-เพศชาย จำแนกคนที่นับถือศาสนาต่างๆ เป็นนับถือศาสนาพุทธ-คริสต์-อิสลาม-ฮินดู จำแนกระดับการศึกษาของครูเป็นระดับปริญญาตรี-สูงกว่าปริญญาตรี เป็นต้น ในการวิจัยข้อมูลในระดับนี้มักจะใช้เป็นตัวแปรอิสระ (Independent variable) นอกจากนี้อาจจะมีการกำหนดตัวเลขให้กับสิ่งต่างๆเพื่อใช้ในการสื่อความหมาย โดยที่ตัวเลขดังกล่าวไม่มีความหมายในเชิงปริมาณ เช่น ห้องหมายเลข 927 สื่อความหมายว่าเป็นห้องๆหนึ่ง ห้องหมายเลข 323 สื่อความหมายว่าเป็นอีกห้องหนึ่ง เป็นต้น ห้องที่มีหมายเลขมากกว่าไม่ได้หมายความว่าจะจุคนได้มากกว่า

กล่าวโดยสรุปก็คือ สัญลักษณ์ที่กำหนดขึ้นนั้นอาจเป็นตัวเลข ตัวอักษร ชื่อ หรือสัญลักษณ์ใดๆก็ได้ แต่ตัวเลขดังกล่าวไม่ได้มีคุณสมบัติเหมือนใน 3 มาตรฐานแรกที่กล่าวมา ดังนั้นจึงไม่สามารถนำตัวเลขต่างๆ มาบวก ลบ คูณ หารกัน ลักษณะที่กระทำได้คือการนับเพื่อดูความถี่หรือการเกิดซ้ำๆกันของข้อมูลแต่ละประเภท

เช่น จากการสำรวจสถานภาพสมรสของครูในโรงเรียน ก. พบว่าเป็นครูที่มีสถานภาพสมรสโสด 45 คนแต่งงานแล้ว 15 คน และเป็นหม้าย 2 คน สถานภาพสมรสของครูซึ่งจำแนกออกเป็น 3 ประเภทนั้นจัดเป็นการวัดในมาตรฐานนามบัญญัติ จากการนับจำนวนครูจำแนกตามสถานภาพสมรสพบว่าได้ตัวเลขมา 3 ตัวคือ 45, 15, และ 2 ตัวเลข 3 ตัวนี้เรียกว่าเป็นความถี่ที่ได้จากการนับ ซึ่งเป็นตัวเลขที่จะนำไปใช้ในการวิเคราะห์ข้อมูลต่อไป ดังนั้นข้อมูลที่อยู่ในมาตรานี้จะเป็นชนิด frequency

ตัวอย่างการวัดในมาตราต่างๆ

มาตราการวัด	ตัวอย่างการวัด
Ratio scale	ความสูง, น้ำหนัก, เวลา, รายได้, ความดันโลหิต ฯลฯ
Interval scale	วันในปฏิทิน, ระดับอุณหภูมิ, ผลสัมฤทธิ์ทางการเรียน, เจตคติต่ออาชีพ ฯลฯ
Ordinal scale	ยศทางทหาร, ผลการประกวดนางงาม ฯลฯ
Nominal scale	สีตา, เชื้อชาติ, สถานภาพสมรส, ภูมิภาค, หมายเลขนักฟุตบอล ฯลฯ

สรุป มาตราวัดแต่ละชนิดเหมาะสำหรับสถิติที่ใช้ในการวิเคราะห์ข้อมูลแต่ละชนิด ถ้าเราไม่สามารถจัดว่าข้อมูลที่รวบรวมได้อยู่ในมาตราวัดใด ก็อาจจะเลือกใช้เทคนิคการวิเคราะห์ข้อมูลที่ไม่เหมาะสม ซึ่งทำให้การสรุปผลผิดพลาด (Mueller, 1970: 16)

ตาราง 1.2 ระดับข้อมูลและสถิติที่เหมาะสมกับระดับข้อมูลนั้นๆ

ระดับข้อมูล	ตัวอย่างสถิติต่างๆ ที่เหมาะสม	การทดสอบนัยสำคัญที่เหมาะสม
นามบัญญัติ	ฐานนิยม การแจกแจงความถี่ Contingency coefficient	การทดสอบน็อนพารามตริก (Nonparametric statistical test)
เรียงอันดับ	มัธยฐาน เปอร์เซ็นต์ไทล์ Rank correlation	การทดสอบน็อนพารามตริก (Nonparametric statistical test)
อันตรภาค	ค่าเฉลี่ย ความเบี่ยงเบนมาตรฐาน Pearson product-moment correlation	การทดสอบพารามตริก (Parametric statistical test)
อัตราส่วน	ใช้สถิติได้ทุกตัว	ใช้ได้ทุกชนิด

1.5 ความรู้พื้นฐานทางคณิตศาสตร์

1.5.1 เครื่องหมายหรือสัญลักษณ์ทางคณิตศาสตร์

ที่ใช้กันเสมอมีอยู่ 4 อย่าง คือ Σ (อ่านว่า ซิกมา), X , N , $\sqrt{\quad}$ ซึ่งจัดเป็น 3 ประเภท คือ

ก. **คำนามทางคณิตศาสตร์** (Mathematical nouns) เป็นสัญลักษณ์ที่กำหนดขึ้นเพื่อใช้แทนปริมาณ ที่สำคัญมี 2 ตัว คือ X กับ N

X เป็นสัญลักษณ์ใช้แทนค่าของคะแนนหรือปริมาณ เช่น คะแนนผลการสอบวิชาสถิติ ค่าความสูงของนักเรียนอายุ 10 ขวบ เป็นต้น ในการวิเคราะห์ข้อมูล ถ้ามีข้อมูลชุดเดียว จะใช้สัญลักษณ์ X แทนดังตาราง 1.3 ถ้ามีข้อมูล 2 ชุดก็ใช้สัญลักษณ์ 2 ตัว คือ X และ Y แทนดังตาราง 1.4

ตาราง 1.3

คนที่	X
1	10
2	25
3	13
4	20
5	9

ตาราง 1.4

คนที่	X	Y
1	20	7
2	17	8
3	30	10
4	25	11
5	20	9

N เป็นสัญลักษณ์ใช้แทนจำนวนคะแนนหรือจำนวนปริมาณ เช่น ถ้ามีคะแนนผลการสอบวิชาสถิติอยู่ 10 จำนวน เขียนเป็นสัญลักษณ์ได้ $N = 10$ วัดส่วนสูงของนักเรียน 30 คน จะได้ $N = 30$ เป็นต้น

ข. **คำคุณศัพท์ทางคณิตศาสตร์** (Mathematical adjectives) ใช้เมื่อต้องการขยายคำนามทางคณิตศาสตร์ โดยจะเขียนตัวเลขหรืออักษรกำกับไว้ข้างล่างเฉียงไปทางขวา เช่น ถ้ามีคะแนนชุดหนึ่ง N จำนวน จะใช้สัญลักษณ์ X_i แทนค่าของคะแนนต่างๆ N ค่า ตั้งแต่ค่าที่ 1 ถึงค่าที่ N เขียนกระจายได้ดังนี้: $X_1, X_2, \dots, X_N$

ค. คำกริยาทางคณิตศาสตร์ (Mathematical verbs) ใช้เมื่อต้องการบอกให้ทราบว่า ต้องทำอะไรบ้าง คำกริยาทางคณิตศาสตร์ที่สำคัญ ได้แก่

1) Σ บอกให้ทราบว่า จะต้องเอาคะแนนหรือปริมาณต่างๆ มารวมกัน

2) $\sqrt{\quad}$ บอกให้ทราบว่า ต้องหารากที่สองที่เป็นจำนวนบวกของคะแนนหรือปริมาณต่างๆ

1.5.2 การใช้เครื่องหมายผลรวมในวิชาสถิติ

สัญลักษณ์ที่ใช้คือ $\sum_{i=1}^N X_i$ ซึ่ง i ได้ Σ และ N เหนือ Σ แสดงให้เห็นว่า จะต้องรวมค่า X ทุกค่า จากค่าที่ 1 จนถึงค่าที่ N โดยที่ X_1 แทนคะแนนของหน่วยที่ 1, X_2 แทนคะแนนของหน่วยที่ 2, X_i แทนคะแนนของหน่วยที่ i และ N แทนจำนวนหน่วยทั้งหมด ดังนั้นจะได้ว่า

$$X_1 + X_2 + \dots + X_i + \dots + X_N = \sum_{i=1}^N X_i$$

ตัวอย่างเช่น ถ้าเราต้องการรวมค่าของตัวแปร X_i ซึ่งมีค่าเป็น 4, 5, 6, 10, 15

สัญลักษณ์ $\sum_{i=1}^5 X_i$ ก็จะหมายความว่าให้รวมทุกๆ ค่าของตัวแปร X_i โดยเริ่มจากค่าที่หนึ่งคือ 4 ไปจนครบทั้ง 5 ค่า ดังนี้

$$\begin{aligned} \sum_{i=1}^5 X_i &= X_1 + X_2 + X_3 + X_4 + X_5 \\ &= 4 + 5 + 6 + 10 + 15 \\ &= 40 \end{aligned}$$

ถ้าข้อมูลอยู่ในรูปของเมตริกซ์ (Matrix) ดังตาราง 1.5 ต้องใช้

เครื่องหมายผลรวม 2 ตัว คือ $\sum_{i=1}^r \sum_{j=1}^c X_{ij}$ เมื่อ i แทนแถว (row) และ j

แทนคอลัมน์ (Column) โดยที่แถวมีพิสัยจาก 1 ถึง r และคอลัมน์มีพิสัยจาก 1 ถึง c

หมายเหตุ ไม่จำเป็นต้องใช้ i แทนแถว และ j แทนคอลัมน์เสมอไป อาจใช้ j แทนแถว และ k แทนคอลัมน์ก็ได้ และเขียนสัญลักษณ์ผลรวมเป็น

$$\sum_{j=1}^r \sum_{k=1}^c X_{jk} \quad \text{ทั้งนี้ขึ้นอยู่กับผู้เขียน}$$

ตาราง 1.5 Matrix

		คอลัมน์		
		1	2	3
แถว	1	10	2	8
	2	15	16	2

เครื่องหมายผลรวม 2 ตัว หมายความว่าต้องรวมค่าต่างๆ ใน Matrix ทั้งทางคอลัมน์ และทางแถว ดังนี้

$$\sum_{i=1}^2 \sum_{j=1}^3 X_{ij} = (10 + 2 + 8) + (15 + 16 + 2) = 53$$

ถ้าต้องการกำหนดค่าเฉพาะค่าใดค่าหนึ่ง เช่น X_{21} ก็หมายความว่า เป็นคะแนนในแถวที่ 2 คอลัมน์ที่ 1 ซึ่งมีค่าเท่ากับ 15 ($X_{21} = 15$)

ถ้าต้องการรวมเฉพาะแถวที่หนึ่ง เขียนเป็น $\sum_{j=1}^3 X_{1j} = (10 + 2 + 8) = 20$

ถ้าต้องการรวมเฉพาะคอลัมน์ที่หนึ่ง เขียนเป็น $\sum_{i=1}^2 X_{i1} = (10 + 15) = 25$

โดยทั่วไปเราอาจเขียน $\sum X$ แทน $\sum_{i=1}^N X_i$



แบบฝึกหัดที่ 1

1. จงเปรียบเทียบความหมายของคำว่า “สถิติ” ที่ท่านค้นได้จากพจนานุกรมกับที่กล่าวไว้ในบทที่ 1
2. จงพิจารณาว่าข้อใดเป็นตัวแปรและข้อใดเป็นค่าคงที่
 - ก. จำนวนวันในเดือนสิงหาคม
 - ข. อายุของนิสิตปีที่ 1 คณะศึกษาศาสตร์ มศว
 - ค. เวลาที่ใช้ในการทำรายงานเรื่องต่างๆ
 - ง. คะแนนที่ได้จากการสอบแบบทดสอบแบบเลือกตอบ ซึ่งมีจำนวน 100 ข้อ
 - จ. จำนวนเงินที่ใช้ในการซื้อหนังสือในรอบ 1 ปีของนิสิต 10 คน
3. จงยกตัวอย่าง “ประชากร” มา 3 ตัวอย่าง
4. จงพิจารณาว่าข้อความที่กำหนดให้เป็นค่าสถิติ ข้อมูล หรือเป็นการอ้างสรุปจากข้อมูล
 - ก. กลุ่มตัวอย่างจำนวน 250 คนในจังหวัดราชบุรีมีรายได้เฉลี่ย 70,000 บาทต่อปี
 - ข. สัดส่วนของนางสาวจารุรัตน์ คือ 36 – 25 – 35
 - ค. อัตราการเกิดในประเทศสหรัฐอเมริกาลดลง 5%จากปีก่อน
5. คำตอบที่ได้จากการตอบคำถามต่อไปนี้เป็นข้อมูลที่อยู่ในมาตราใด
 - ก. ท่านสูงเท่าใด
 - ข. ท่านหนักเท่าใด
 - ค. ท่านประกอบอาชีพอะไร
 - ง. ผลการเรียนของท่านเป็นลำดับที่เท่าใดของห้อง
 - จ. ท่านชอบอ่านข่าวประเภทใด
6. จงยกตัวอย่างข้อมูลที่อยู่ในมาตรานามบัญญัติ เรียงอันดับ อันตรภาค และอัตราส่วน มาตราละ 2 อย่าง

7. บริษัทวิจัยแห่งหนึ่ง ได้ทำการสำรวจข้อมูลเกี่ยวกับการใช้ไฟฟ้าของครอบครัวต่างๆ ในหมู่บ้านแห่งหนึ่งดังนี้

- ก. อายุของหัวหน้าครอบครัว
- ข. อาชีพของหัวหน้าครอบครัว
- ค. การใช้เครื่องปรับอากาศ (ใช้กับไม่ใช้)
- ง. จำนวนชั่วโมงโดยเฉลี่ยที่ใช้ไฟฟ้าในบ้าน
- จ. รายได้ของครอบครัว
- ฉ. จัดลำดับปริมาณการใช้ไฟฟ้าจากเดือนมกราคมถึงเดือนพฤษภาคม

7.1 จงพิจารณาว่าตัวแปรในข้อ ก-ฉ ที่บริษัทวิจัยต้องการสำรวจ ตัวแปรใดเป็นตัวแปรเชิงปริมาณและตัวแปรใดเป็นตัวแปรเชิงคุณภาพ

7.2 จงพิจารณาว่าตัวแปรแต่ละตัว เมื่อเก็บข้อมูลแล้วจะเป็นข้อมูลในระดับใด

บทที่ 2

การแจกแจงความถี่ (Frequency distribution)

2.1 ความหมายของการแจกแจงความถี่

หลังจากที่กำหนดตัวแปรที่ต้องการศึกษาและมีการเก็บรวบรวมข้อมูลเกี่ยวกับตัวแปรที่ต้องการศึกษาแล้ว ข้อมูลที่เก็บได้เรียกว่าข้อมูลดิบ (Raw data) เช่นต้องการศึกษาตัวแปร “ผลสัมฤทธิ์ทางการเรียนวิชาสถิติ” ก็จะสร้างเครื่องมือคือแบบทดสอบวัดผลสัมฤทธิ์ทางการเรียนวิชาสถิติ นำไปสอบกับกลุ่มนิสิตนักศึกษาที่ต้องการวัด แล้วตรวจให้คะแนน คะแนนที่ได้เรียกว่าข้อมูลดิบ (Raw data) หรือคะแนนดิบ (Raw score) ซึ่งตัวข้อมูลดิบเองยังไม่มี ความหมายอะไรมากนัก ปัญหา ก็คือจะทำข้อมูลดิบเหล่านี้ให้มีความหมายได้อย่างไร โดยทั่วไปเราต้องการที่จะบรรยายคุณลักษณะของคะแนนทั้งกลุ่มมากกว่าที่จะสนใจคะแนนแต่ละตัว เช่น ต้องการทราบพิสัยของคะแนนชุดนั้น ต้องการทราบว่าคะแนนส่วนใหญ่เกาะกลุ่มกันอย่างไร คะแนนชุดนั้นมีการกระจายเป็นอย่างไ เป็นต้น วิธีง่ายๆ วิธีหนึ่งที่จะทำให้ข้อมูลดิบมีความหมายก็คือทำการแจกแจงความถี่

การแจกแจงความถี่ คือ การนำข้อมูลที่รวบรวมได้มาจัดใหม่ให้เป็นระเบียบ เป็นหมวดหมู่ เรียงจากค่ามากไปหาค่าน้อย (หรือเรียงจากค่าน้อยไปหาค่ามาก) เพื่อแสดงให้เห็นว่าข้อมูลแต่ละค่า (เมื่อเป็นการแจกแจงความถี่แบบไม่จัดเป็นกลุ่ม) หรือข้อมูลแต่ละกลุ่ม (เมื่อเป็นการแจกแจงแบบจัดเป็นกลุ่ม) เกิดขึ้นซ้ำๆ กันกี่ครั้ง (หรือที่เรียกว่ามีความถี่มากน้อยใด) ซึ่งเป็นการย่อข้อมูลเพื่อให้แปลความหมายได้มากขึ้น โดยจะต้องสร้างตารางแจกแจงความถี่ขึ้น

ตารางแจกแจงความถี่จะเป็นตารางซึ่งประกอบด้วยส่วนประกอบที่สำคัญ 3 ส่วน คือ ส่วนที่หนึ่งเป็นคะแนน แทนด้วย X (คำว่า “คะแนน” ในที่นี้หมายถึงข้อมูลที่รวบรวมได้ไม่จำเป็นต้องเป็นคะแนนสอบเสมอไป) ส่วนที่สองเป็นจัตรอยคะแนน (Tally) และส่วนที่สามเป็นค่าของความถี่แทนด้วย f ซึ่งนับได้จากจำนวนจัตรอยคะแนน

ตัวอย่างที่ 2.1 สมมติว่ามีการทดสอบย่อยวิชาประวัติศาสตร์กับนิสิต 10 คน โดยใช้แบบทดสอบแบบอัตนัย คะแนนเต็ม 10 คะแนน ผลการสอบปรากฏใน ตอน ก. ของตาราง 2.1 วิธีแรกที่จะทำให้เห็นภาพคะแนนเหล่านี้ซึ่งมีจำนวนน้อยให้ชัดเจนขึ้นก็คือเรียงลำดับคะแนนจากมากไปหาน้อย ตามที่ปรากฏใน ตอน ข. ของตาราง 2.1 จะสังเกตเห็นว่า 10 คะแนน มี 1 จำนวน 9 คะแนน มี 2 จำนวน 8 และ 7 คะแนนมีอย่างละ 3 จำนวน และ 6 คะแนนมี 1 จำนวน ถ้าใช้สัญลักษณ์ X แทนค่าของคะแนน แล้วขีดรอยคะแนน (tally) ว่าคะแนนแต่ละตัวเกิดขึ้นกี่ครั้ง การทำเช่นนี้ เรียกว่า การแจกแจงความถี่ดังปรากฏใน ตอน ค. ของตาราง 2.1 ซึ่งช่อง f จะชี้ให้เห็นว่าแต่ละค่าของตัวแปร X มีกี่จำนวน

ตาราง 2.1 พัฒนาการของการแจกแจงความถี่

ตอน ก. คะแนนสอบวิชาประวัติศาสตร์ของนิสิต 10 คน

8	8
10	7
9	6
7	8
9	7

ตอน ข. คะแนนสอบวิชาประวัติศาสตร์เรียงจากมากไปน้อย

10	8
9	7
9	7
8	7
8	6

ตอน ค. การแจกแจงความถี่คะแนนสอบวิชาประวัติศาสตร์

X	Tally	F
10	/	1
9	//	2
8	///	3
7	///	3
6	/	1
		N = 10

ตัวอย่างที่ 2.2 เป็นการแจกแจงความถี่คำตอบของกลุ่มตัวอย่างจำนวน 80 คนที่เลือกตอบจากคำถาม 1 ข้อ ของแบบสอบถามวัดความคิดเห็นแบบมาตราส่วนประมาณค่า ซึ่งมีคำตอบให้เลือกตอบ 5 คำตอบ และแต่ละคำตอบให้คะแนนดังนี้

เห็นด้วยน้อยที่สุด	ให้	1	คะแนน
เห็นด้วยน้อย	ให้	2	คะแนน
เห็นด้วยปานกลาง	ให้	3	คะแนน
เห็นด้วยมาก	ให้	4	คะแนน
เห็นด้วยมากที่สุด	ให้	5	คะแนน

กลุ่มตัวอย่างแต่ละคนจะเลือกเพียงคำตอบเดียวที่ตรงกับความคิดเห็นของเขาเองมากที่สุด แล้วนำคำตอบมาทำการแจกแจงความถี่ดังปรากฏในตาราง 2.2

ตาราง 2.2 การแจกแจงความถี่คำตอบเกี่ยวกับความคิดเห็นของกลุ่มตัวอย่าง

คำตอบ	X	f
เห็นด้วยน้อยที่สุด	1	6
เห็นด้วยน้อย	2	25
เห็นด้วยปานกลาง	3	10
เห็นด้วยมาก	4	30
เห็นด้วยมากที่สุด	5	9
		N = 80

ทั้ง 2 ตัวอย่างที่กล่าวมาคะแนนมีค่าน้อย (ตัวอย่างที่ 2.1 คะแนนมีค่าจาก 6 – 10 ตัวอย่างที่ 2.2 มีค่าจาก 1 – 5) และตัวอย่างที่ 2.1 มี N น้อย (N = 10) แต่ตัวอย่างที่ 2.2 มี N มาก (N = 80)

การแจกแจงความถี่จะมีประโยชน์มากถ้ามีคะแนนจำนวนมาก และคะแนนที่รวบรวมได้มีการกระจายมาก ดังตัวอย่างที่ 2.3

ตัวอย่างที่ 2.3 จากการสอบความรู้พื้นฐานทางคณิตศาสตร์ของนักเรียน 112 คน ได้คะแนนดังตาราง 2.3 จงสร้างตารางแจกแจงความถี่

ตาราง 2.3 คะแนนความรู้พื้นฐานทางคณิตศาสตร์ของนักเรียน 112 คน

79	51	67	50	78	71	77	75
62	89	83	73	80	67	74	63
88	48	60	71	79	79	47	55
89	63	55	93	71	81	72	68
41	81	46	50	61	72	86	66
59	50	90	75	61	82	73	57
75	98	53	79	80	64	67	51
70	37	42	72	74	78	91	69
67	73	79	67	85	74	70	62
91	73	77	36	77	45	39	59
53	67	85	74	77	78	73	61
71	43	42	96	83	83	84	67
70	92	59	86	53	71	49	68
32	67	67	71	71	59	80	66

จากตาราง 2.3 ซึ่งเป็นตารางแสดงคะแนนสอบของนักเรียนทั้ง 112 คน เราไม่สามารถแปลความหมายของคะแนนชุดนี้ได้ทันที ต้องสร้างตารางแจกแจงความถี่ก่อนจึงจะแปลความหมายได้ชัดเจน

ถ้าสร้างตารางแจกแจงความถี่ตามตัวอย่างที่ 2.1 และตัวอย่างที่ 2.2 ซึ่งเรียกว่าการแจกแจงความถี่แบบไม่จัดกลุ่ม (Ungrouped data) จะต้องแจกแจงว่าคะแนนแต่ละค่ามีกี่จำนวนซึ่งไม่เหมาะที่จะใช้ในกรณีที่มีคะแนนมากเช่นนี้ ควรจัดกลุ่มคะแนนก่อนแล้วแจกแจงว่าแต่ละกลุ่มคะแนนมีคะแนนกี่จำนวน ในตาราง 2.4 จะเห็นว่ามีคะแนนเพียง 2 จำนวนระหว่างช่วงคะแนน 95 ถึง 99, มีคะแนน 5 จำนวนระหว่างช่วงคะแนน 90 ถึง 94,... ซึ่งเขียนสรุปได้ดังตาราง 2.5

ข้อแตกต่างระหว่างการแจกแจงความถี่ในตัวอย่างที่ 2.3 กับตัวอย่างที่ 2.1 และ 2.2 ที่กล่าวมาแล้วข้างต้นก็คือ แทนที่จะดูว่าคะแนนแต่ละค่ามีกี่จำนวนก็จะดูว่าคะแนนแต่ละกลุ่ม เช่น 30 – 34, 35 – 39, ... มีกี่จำนวน ลักษณะนี้เรียกว่าการแจกแจงความถี่แบบจัดเป็นกลุ่ม (Grouped data)

ตาราง 2.4 คะแนนความรู้พื้นฐานทางคณิตศาสตร์ของนักเรียน 112 คน เรียงจากมากไปหาน้อย

-	74, 74, 74, 74	49
98	73, 73, 73, 73, 73	48
-	72, 72, 72	47
96	71, 71, 71, 71, 71, 71, 71	46
-	70, 70, 70	45
-	69	-
93	68, 68	43
92	67, 67, 67, 67, 67, 67, 67, 67, 67	42, 42
91, 91	66, 66	41
90	-	-
89, 89	64	39
88	63, 63	-
-	62, 62	37
86, 86	61, 61, 61	36
85, 85	60	-
84	59, 59, 59, 59	-
83, 83, 83	-	-
82	57	32
81, 81	-	-
80, 80, 80, 80	55, 55	-
79, 79, 79, 79, 79	-	-
78, 78, 78	53, 53, 53	-
77, 77, 77, 77	-	-
-	51, 51	-
75, 75, 75	50, 50, 50	-

ตาราง 2.5 ตารางแจกแจงความถี่คะแนนความรู้พื้นฐานทางคณิตศาสตร์ของนักเรียน
112 คน

อันตรภาคชั้น (Class interval)	ขีดรอยคะแนน (Tally)	ความถี่ (f)
95 – 99	//	2
90 – 94	////	5
85 – 89	//// //	7
80 – 84	//// ////	10
75 – 79	//// //// ////	15
70 – 74	//// //// //// //// //	22
65 – 69	//// //// ///	14
60 – 64	//// ///	9
55 – 59	//// //	7
50 – 54	//// ///	8
45 – 49	////	5
40 – 44	///	4
35 – 39	///	3
30 – 34	/	1
		N = 112

สรุปแล้วการแจกแจงความถี่มี 2 วิธี ดังจะกล่าวในหัวข้อ 2.2

2.2 วิธีการแจกแจงความถี่มี 2 ชนิดคือ

1. แจกแจงความถี่แบบไม่จัดกลุ่ม (Ungrouped data) วิธีนี้จะเรียงลำดับคะแนนจากค่ามากไปหาค่าน้อย หรือจากค่าน้อยไปหาค่ามาก แล้วหาว่าคะแนนแต่ละคะแนนนั้นมีกี่จำนวน วิธีนี้ช่วงห่างระหว่างแต่ละคะแนนเป็นหนึ่งหน่วยเท่ากันตลอด

ตัวอย่างที่ 2.4 จากการสอบวัดเชาว์ปัญญา (I.O.) ของนักเรียนชั้นมัธยมศึกษาตอนปลาย 110 คนปรากฏผล ดังนี้

154	131	122	100	113	119	121	128	112	93
133	119	115	117	110	104	125	85	120	135
116	103	103	121	109	147	103	113	107	98
128	93	90	105	118	134	89	143	108	142
85	108	108	136	115	119	110	80	111	127
100	100	114	123	126	117	122	102	100	106
105	111	127	108	106	91	123	132	97	110
150	130	87	89	108	137	124	96	111	101
118	104	127	94	115	101	125	129	131	110
97	135	108	139	133	107	115	83	109	116
110	113	112	82	114	112	113	142	145	123

จงแจกแจงความถี่แบบไม่จัดกลุ่ม

- วิธีทำ**
- 1) พิจารณาจากค่าคะแนนสูงสุดและคะแนนต่ำสุด คะแนนสูงสุดคือ 154 คะแนนต่ำสุดคือ 80 คะแนน
 - 2) สร้างตารางบรรจุคะแนนจากค่าสูงสุดไปหาค่าต่ำสุด แล้วหาความถี่

ตาราง 2.6 ตารางแจกแจงความถี่คะแนนเข้าวัดปัญญา (I.Q.) ของนักเรียน 110 คน

X	ขีดรอยคะแนน	f	X	ขีดรอยคะแนน	f	X	ขีดรอยคะแนน	f
154	/	1	129	/	1	104	//	2
153	-	0	128	//	2	103	///	3
152	-	0	127	///	3	102	/	1
151	-	0	126	/	1	101	//	2
150	/	1	125	//	2	100	////	4
149	-	0	124	/	1	99	-	0
148	-	0	123	///	3	98	/	1
147	/	1	122	//	2	97	//	2
146	-	0	121	//	2	96	/	1
145	/	1	120	/	1	95	-	0
144	-	0	119	///	3	94	/	1
143	/	1	118	//	2	93	//	2
142	//	2	117	//	2	92	-	0
141	-	0	116	//	2	91	/	1
140	-	0	115	////	4	90	/	1
139	/	1	114	//	2	89	//	2
138	-	0	113	////	4	88	-	0
137	/	1	112	///	3	87	/	1
136	/	1	111	///	3	86	-	0
135	//	2	110	////	5	85	//	2
134	/	1	109	//	2	84	-	0
133	//	2	108	////	6	83	/	1
132	/	1	107	//	2	82	/	1
131	//	2	106	//	2	81	-	0
130	/	1	105	//	2	80	/	1

การแจกแจงความถี่แบบไม่จัดกลุ่มนี้ ถ้าคะแนนสูงสุดและคะแนนต่ำสุดมีค่าไม่แตกต่างกันมากนัก จะทำให้เห็นภาพการกระจายของข้อมูลและแปลความหมายได้ชัดเจน แต่ถ้าคะแนนสูงสุดและต่ำสุดมีค่าแตกต่างกันมากควรแจกแจงความถี่แบบจัดเป็นกลุ่ม

2. การแจกแจงความถี่แบบจัดกลุ่ม (Grouped data)

วิธีนี้คล้ายคลึงกับการแจกแจงความถี่แบบไม่จัดเป็นกลุ่มตรงที่จะต้องเรียงคะแนนจากมากไปหาน้อย แต่แตกต่างกันตรงที่ว่าแต่ละชั้นคะแนนจะประกอบด้วยกลุ่มของคะแนนแทนที่จะเป็นคะแนนเพียงหนึ่งตัว

การแจกแจงความถี่แบบจัดเป็นกลุ่ม มีขั้นตอน ดังนี้

ขั้นที่ 1 หาพิสัยของคะแนนพิสัย = คะแนนสูงสุด - คะแนนต่ำสุด

ขั้นที่ 2 ประมาณจำนวนชั้นคะแนน ซึ่งไม่ควรต่ำกว่า 10 ชั้น และไม่เกิน 20 ชั้น ทั้งนี้ให้พิจารณาจากพิสัย ถ้าพิสัยมีค่ามากก็ควรกำหนดหลายชั้น ถ้าพิสัยมีค่าน้อยก็ไม่จำเป็นต้องมีจำนวนชั้นคะแนนมาก

ขั้นที่ 3 หาว่าในแต่ละชั้นคะแนนมีคะแนนกี่จำนวน นั่นคือหาอันตรภาคชั้น (Class interval) โดยใช้สูตร

$$\text{อันตรภาคชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}}$$

ขั้นที่ 4 เขียนขีดจำกัดชั้นของคะแนนแต่ละชั้นลงในช่องคะแนน โดยเริ่มจากคะแนนสูงสุดหรือต่ำสุดก็ได้

ขั้นที่ 5 จีดรอยคะแนน (Tally) โดยเอาคะแนนที่รวบรวมได้แต่ละคะแนนไปขีดลงในชั้นคะแนนแต่ละชั้น

ขั้นที่ 6 นับจำนวนรอยคะแนนแล้วรวมใส่ในช่องความถี่

ตัวอย่างที่ 2.5 จากโจทย์ในตัวอย่างที่ 2.4 จงสร้างตารางแจกแจงความถี่แบบจัดเป็นกลุ่ม

ขั้นที่ 1 หาพิสัย

$$\text{พิสัย} = 154 - 80 = 74$$

ขั้นที่ 2 ประมาณจำนวนชั้นคะแนน

ในตัวอย่างนี้สมมติว่าต้องการ 15 ชั้น

ขั้นที่ 3 หาอันตรภาคชั้น

$$\begin{aligned}\text{อันตรภาคชั้น} &= \frac{\text{พิสัย}}{\text{จำนวนชั้น}} \\ &= \frac{74}{15} = 4.9 \approx 5\end{aligned}$$

ขั้นที่ 4 เขียนขีดจำกัดชั้นของคะแนน ในตัวอย่างนี้จะเริ่มจากคะแนนสูงสุด ให้คะแนนสูงสุดเป็นขีดจำกัดบนของชั้นของคะแนน
สูงสุด ขีดจำกัดบนของชั้นคะแนนถัดลงไปก็จะห่างจากชั้นบนเท่ากับอันตรภาคชั้น (5 คะแนน)

ขั้นที่ 5 ขีดรอยคะแนน

ขั้นที่ 6 นับจำนวนรอยคะแนนแล้วรวมใส่ในช่องความถี่

ตาราง 2.7 ตารางแจกแจงความถี่แบบจัดเป็นกลุ่ม

(N = 110)

ชั้นคะแนน (X)	ขีดรอยคะแนน (Tally)	ความถี่ (f)
150 – 154	//	2
145 - 149	//	2
140 - 144	///	3
135 - 139	////	5
130 - 134	//// //	7
125 - 129	//// ///	9
120 - 124	//// ///	9
115 - 119	//// //// //	13
110 - 114	//// //// //// //	17
105 – 109	//// //// ///	14
100 - 104	//// //// //	12
95 – 99	////	4
90 – 94	////	5
85 - 89	////	5
80 – 84	///	3

2.3 คำศัพท์ที่เกี่ยวข้องกับการแจกแจงความถี่

1. อินตรภาคชั้น (i) คือ ความกว้างของชั้นคะแนนแต่ละชั้น
2. ขีดจำกัดบน (Upper limit) คือ คะแนนสูงสุดของคะแนนแต่ละชั้น
ขีดจำกัดล่าง (Lower limit) คือ คะแนนต่ำสุดของคะแนนแต่ละชั้น
3. จุดกลาง (Mid point) คือ คะแนนที่อยู่กึ่งกลางระหว่างขีดจำกัดบนและ

ขีดจำกัดล่าง หาได้จากสูตร

$$\text{จุดกลาง} = \frac{[\text{ขีดจำกัดบน} + \text{ขีดจำกัดล่าง}]}{2}$$

4. ขีดจำกัดชั้นที่แท้จริง (Exact limits หรือ Real limits) คือ คะแนนที่แต่ละชั้นคะแนนมีค่าต่อเนื่องกัน

ในกรณีที่ขีดจำกัดบนและขีดจำกัดล่างของแต่ละชั้นคะแนนเป็นจำนวนเต็ม จะหาขีดจำกัดชั้นที่แท้จริงได้โดยเอา 0.5 ลบออกจากขีดจำกัดล่าง และเอา 0.5 บวกกับขีดจำกัดบน เช่น ชั้นคะแนน 80 – 84 เขียนเป็นขีดจำกัดที่แท้จริง ได้ 79.5 – 84.5

ในกรณีที่ขีดจำกัดบนและขีดจำกัดล่างของแต่ละชั้นคะแนนเป็นทศนิยม 1 ตำแหน่ง จะหาขีดจำกัดชั้นที่แท้จริงได้โดยเอา 0.05 ลบออกจากขีดจำกัดล่าง และเอา 0.05 บวกกับขีดจำกัดบน เช่น ชั้นคะแนนเป็น 70.5 – 74.5 เขียนเป็นขีดจำกัดชั้นที่แท้จริง ได้ 70.45 – 74.55 ถ้าขีดจำกัดบนและขีดจำกัดล่างของแต่ละชั้นคะแนนเป็นทศนิยม 2 ตำแหน่ง จะหาขีดจำกัดชั้นที่แท้จริงได้โดยเอา 0.005 ลบออกจากขีดจำกัดล่าง และเอา 0.005 บวกกับขีดจำกัดบน เช่น คะแนนเป็น 4.05 – 9.05 เขียนเป็นขีดจำกัดชั้นที่แท้จริงได้ 4.045 – 9.055 เป็นต้น

การกำหนดจำนวนชั้นคะแนนของตารางแจกแจงความถี่แบบจัดเป็นกลุ่ม

จุดมุ่งหมายของการสร้างตารางแจกแจงความถี่ก็คือต้องการสรุปย่อข้อมูล (Summarization of data) เพื่อจะได้เห็นภาพข้อมูลทั้งกลุ่มได้อย่างถูกต้อง ดังนั้น จำนวนชั้นคะแนนจะต้องเหมาะสม ในตาราง 2.5 มีชั้นคะแนน 14 ชั้น ถ้านับจากคะแนนต่ำสุดชั้นคะแนนชั้นที่หนึ่งคือ 30 – 34, ชั้นที่สองคือ 35 – 39, ... ชั้นคะแนนที่สิบสี่คือ 95 – 99 ถ้าใช้จำนวนชั้นคะแนนมากเกินไปหรือน้อยเกินไปจะทำให้สรุปย่อข้อมูลทั้งกลุ่มได้ไม่เหมาะสม

ได้มีผู้แนะนำว่าจำนวนชั้นคะแนนที่เหมาะสมคือ 10 – 20 ชั้น (Mc.Call. 1970: 26) แต่ทั้งนี้ไม่ใช่กฎเกณฑ์ตายตัว ถ้าจำนวนคะแนนน้อย (N อยู่ในช่วง 20 – 50) จำนวนชั้นคะแนนที่ใช้ก็จะน้อยกว่ากรณีที่จำนวนคะแนนมากถึง 100 – 200 จำนวน ถ้ามีคะแนนเพียง 20 จำนวน อาจจะมีจำนวนชั้นคะแนนเพียง 4 หรือ 5 ชั้น แต่ถ้ามีคะแนนถึง 200 จำนวน ก็ต้องมีชั้นคะแนนเกิน 20 ชั้น

ที่กล่าวมาในข้อ 2.2 และข้อ 2.3 เป็นการแจกแจงความถี่ของข้อมูลเชิงปริมาณ (Quantitative data) สำหรับข้อมูลเชิงคุณภาพ (Qualitative data) ก็สามารถนำมาแจกแจงความถี่ได้โดยแจกแจงความถี่จำแนกตามประเภทหรือตามคุณลักษณะ ดังตัวอย่างที่ 2.6

ตัวอย่างที่ 2.6 จากการสอบถามนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนสาธิต

มหาวิทยาลัยศรีนครินทรวิโรฒ (ฝ่ายมัธยม) จำนวน 20 คน ว่าสนใจเรียน
คณะอะไรมากที่สุด ได้คำตอบดังนี้

บัญชี	สถาปัตย์	วิสวะ	บัญชี	บัญชี	วิสวะ
แพทย์	แพทย์	แพทย์	สถาปัตย์	วิสวะ	วิสวะ
สถาปัตย์	เศรษฐศาสตร์	วิสวะ	แพทย์	แพทย์	แพทย์
บัญชี	บัญชี				

ทำการแจกแจงความถี่จำแนกตามคณะดังนี้

คณะ	ขีดรอยคะแนน	ความถี่
บัญชี	////	5
สถาปัตย์	///	3
วิสวะ	////	5
แพทย์	//// /	6
เศรษฐศาสตร์	/	1

2.4 การแจกแจงความถี่ด้วยกราฟแท่งหรือฮิสโทแกรม

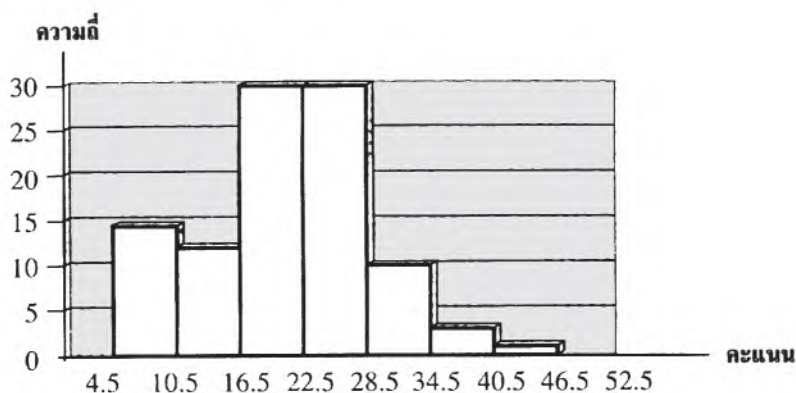
ฮิสโทแกรม (Histogram) หมายถึง แผนภูมิแท่งรูปสี่เหลี่ยมผืนผ้าชุดหนึ่ง ซึ่งมีความสูงของแต่ละแท่งแทนจำนวนความถี่ของคะแนนแต่ละชั้น ความกว้างของแต่ละแท่งแทนระยะระหว่างขีดจำกัดชั้นที่แท้จริง

วิธีสร้างฮิสโทแกรม

1. ลากแกนนอนและแกนตั้ง
2. แบ่งสเกลตามแกนตั้งให้สูงสุดเท่าจำนวนความถี่ของชั้นคะแนนที่มีความถี่มากที่สุด
3. แบ่งสเกลตามแนวนอนโดยใช้ขีดจำกัดชั้นที่แท้จริง ให้เท่ากับจำนวนชั้นของคะแนน
4. สร้างแท่งสี่เหลี่ยมผืนผ้าบนแต่ละสเกลที่แบ่งได้ในข้อ 3. โดยที่แต่ละแท่งสูงเท่าความถี่ของคะแนนแต่ละชั้น

ตัวอย่าง 2.7 จงสร้างฮิสโทแกรมจากตารางแจกแจงความถี่ที่กำหนดให้

ชั้นคะแนน	ขีดรอยคะแนน	ความถี่
41 - 46	/	1
35 - 40	///	3
29 - 34	//// //	10
23 - 28	//// // // // // //	30
17 - 22	//// // // // // //	30
11 - 16	//// // //	12
5 - 10	//// // //	14
		N = 110



รูปที่ 2.1 การแจกแจงความถี่ด้วยฮิสโทแกรม

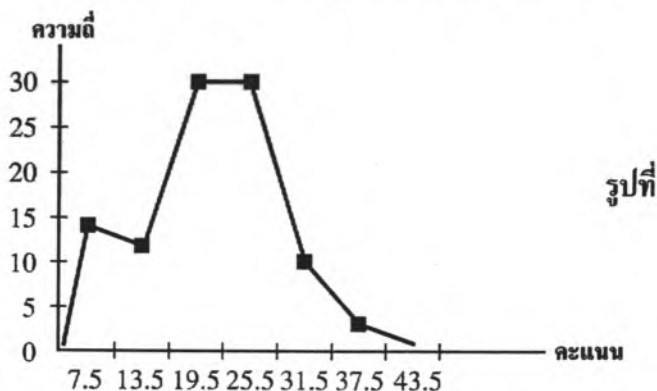
2.5 การแจกแจงความถี่ด้วยรูปหลายเหลี่ยมแห่งความถี่

รูปหลายเหลี่ยมแห่งความถี่ (*Frequency polygon*) คือ แผนภูมิเส้นที่แสดงความถี่ของคะแนนแต่ละชั้น

วิธีสร้างรูปหลายเหลี่ยมแห่งความถี่ทำได้ 2 วิธี

วิธีที่ 1 สร้างฮิสโทแกรม แล้วลากเส้นโยงจุดกลางของแต่ละแท่งสี่เหลี่ยมผืนผ้าของฮิสโทแกรม

วิธีที่ 2 แบ่งสเกลตามแนวนอน โดยใช้จุดกึ่งกลางของคะแนนแต่ละชั้น แล้วลากเส้นตรงโยงทุกจุดที่ประกอบด้วยความถี่และจุดกลางของคะแนนชั้นนั้นๆ ดังแสดงในรูปที่ 2.2



รูปที่ 2.2 การแจกแจงความถี่ด้วยรูปหลายเหลี่ยมแห่งความถี่

2.6 การแจกแจงความถี่สะสม (Cumulative frequency distribution)

ในกรณีที่ต้องการทราบจำนวนคะแนนที่มากกว่าหรือน้อยกว่าคะแนนที่กำหนดให้ จะทราบได้โดยใช้ตารางแจกแจงความถี่สะสม เขียนย่อๆ ว่า cf หรือ cuf ตารางบางเล่มใช้สัญลักษณ์ F

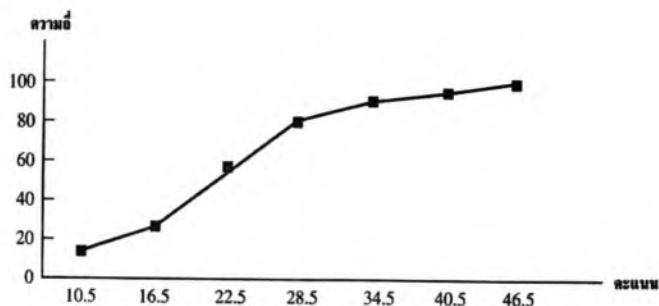
การแจกแจงความถี่สะสม ทำได้ 2 ลักษณะคือ หาความถี่สะสมจากคะแนนมากไปหาคะแนนน้อย และหาความถี่สะสมจากคะแนนน้อยไปหาคะแนนมาก

ตาราง 2.8 ตารางแจกแจงความถี่สะสม

ชั้นคะแนน	ความถี่	ความถี่สะสมจากคะแนน น้อยไปมาก	ความถี่สะสมจากคะแนน มากไปน้อย
41 - 46	1	100	1
35 - 40	3	99	4
29 - 34	10	96	14
23 - 28	30	86	44
17 - 22	30	56	74
11 - 16	12	26	86
5 - 10	14	14	100

2.7 โคง้ความถี่สะสม (Cumulative frequency polygon)

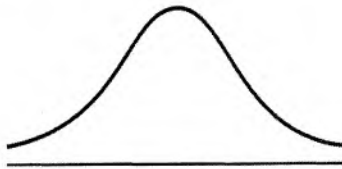
โคง้ความถี่สะสมได้จากความสัมพันธ์ระหว่างคะแนนกับความถี่สะสม ซึ่งสร้างได้โดยลากแกนตั้งให้สูงเท่าจำนวนข้อมูลทั้งหมดและแกนนอนแทนขีดจำกัดชั้นที่แท้จริง



รูปที่ 2.3 โคง้ความถี่สะสม

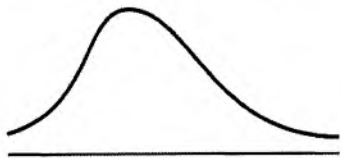
2.8 ลักษณะโค้งต่างๆ ที่เกิดจากการแจกแจงความถี่

1. โค้งปกติ (Normal curve)



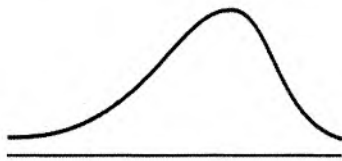
โค้งแบบนี้คนที่ได้คะแนนสูงและ
คะแนนต่ำมีจำนวนพอๆกัน คน
ส่วนมากได้คะแนนปานกลาง
ลักษณะ โค้งคล้ายระฆังคว่ำ

2. โค้งเบ้ทางบวก (Positively skewed or Positive skewness)



โค้งแบบนี้แสดงให้เห็นว่า
คนส่วนใหญ่ได้คะแนนน้อยและ
มีคนส่วนน้อยที่ได้คะแนนมาก

3. โค้งเบ้ทางลบ (Negatively skewed or Negative skewness)



โค้งแบบนี้แสดงให้เห็นว่า คน
ส่วนน้อยได้คะแนนน้อยและ
มีคนส่วนใหญ่ได้คะแนนมาก

4. โค้งรูปตัวยู (U – shaped)



โค้งแบบนี้แสดงให้เห็นว่า
คนส่วนมากได้คะแนนน้อยและ
คะแนนมาก คนส่วนน้อยได้คะแนน
ปานกลาง

ข้อสังเกต

- (1) ในเรื่องของการแจกแจงความถี่ จะเห็นว่ากล่าวถึงคำว่า “คะแนน” อยู่เสมอ ซึ่งคำว่า “คะแนน” นี้ไม่ได้หมายความว่าเฉพาะผลการสอบ แต่หมายถึงข้อมูล (data) ที่ผู้วิจัยรวบรวมมาได้ และต้องการนำข้อมูลนั้นมาแจกแจงความถี่
- (2) ทางซ้ายของแกนนอนมีค่าน้อย และมากขึ้นเรื่อยๆ ไปทางขวาของแกนนอน



แบบฝึกหัดที่ 2

1. จงหาขีดจำกัดชั้นที่แท้จริง จุดกึ่งกลางและอันตรภาคชั้นของชั้นคะแนนต่อไปนี้
 ก. 8 - 12 ข. 6 - 7 ค. 0 - 2 ง. 5 - 14
 จ. $(-2) - (-8)$ ฉ. $2.5 - 3.5$ ช. $(-3) - (+3)$
2. จากการสอบวิชาสถิติของนักเรียนกลุ่มหนึ่งปรากฏผลดังตารางข้างล่าง

63	76	98	82	87	84	78	74
68	75	52	86	77	87	70	67
77	79	92	92	81	77	40	66
94	79	81	82	81	78	70	60
88	46	81	77	83	76	70	61

กำหนด $i = 5$

- ก. จงสร้างตารางแจกแจงความถี่
 - ข. จงหาขีดจำกัดชั้นที่แท้จริงและจุดกึ่งกลาง
 - ค. จงหาความถี่สะสม
 - ง. จงสร้างฮิสโทแกรม
 - จ. จงสร้างโค้งความถี่สะสม
3. ใช้ข้อมูลในข้อ 2 สร้างตารางแจกแจงความถี่โดยใช้ i ต่างๆ กัน ดังนี้
 ก. $i = 1$ ข. $i = 3$ ค. $i = 10$ ง. $i = 20$

จงอภิปรายถึงข้อดีข้อเสียของการใช้ค่า i ต่างๆ กัน

4. จงยกตัวอย่างข้อมูลที่มีลักษณะการแจกแจงต่อไปนี้
 ก. โค้งปกติ
 ข. โค้งรูปตัวยู
 ค. โค้งเบ้ทางบวก
 ง. โค้งเบ้ทางลบ

5. จากการทดสอบความสามารถในการอ่านของนักเรียนชั้นประถมปีที่ 4 จำนวน 2 ห้อง
ปรากฏผลดังนี้

ชั้นคะแนน	ความถี่ของห้อง ก.	ความถี่ของห้อง ข.
50 - 52	5	2
47 - 49	12	3
44 - 46	18	5
41 - 43	19	8
38 - 40	26	12
35 - 37	19	24
32 - 34	13	35
29 - 31	9	22
26 - 28	6	11
23 - 25	4	9
20 - 22	3	6
17 - 19	1	4
14 - 16	2	2

ก. จงสร้างรูปหลายเหลี่ยมแห่งความถี่ทั้ง 2 รูปบนแกนเดียวกัน

ข. จงอธิบายและเปรียบเทียบการแจกแจงความถี่ของทั้ง 2 ห้อง



บทที่ 3

การวัดแนวโน้มเข้าสู่ส่วนกลาง

(Measures of central tendency)

ในบทที่ 2 ได้กล่าวถึงการแจกแจงความถี่ ซึ่งเป็นวิธีหนึ่งที่จะช่วยให้ข้อมูลดิบที่รวบรวมได้มีความหมายมากขึ้น โดยต้องมีการนำเสนอข้อมูลด้วย การวัดแนวโน้มเข้าสู่ส่วนกลางเป็นอีกวิธีหนึ่งที่ช่วยบรรยายลักษณะข้อมูลแต่ละชุดที่รวบรวมได้โดยใช้ตัวเลขเพียงตัวเดียวหรือใช้ค่าใดค่าหนึ่ง การวัดแนวโน้มเข้าสู่ส่วนกลางเป็นระเบียบวิธีสถิติที่ใช้หาค่าค่าหนึ่ง ซึ่งเป็นตัวแทนของข้อมูลทั้งชุด เพื่อสะดวกในการจดจำและสรุปเรื่องราวที่สำคัญของข้อมูลชุดนั้น ค่าที่หาได้จะช่วยให้ทราบลักษณะของข้อมูลทั้งชุดว่าเป็นอย่างไร เช่น ช่วยให้ครูทราบระดับผลการเรียนของนักเรียนทั้งชั้น ทราบอายุเฉลี่ยของนักเรียนทั้งชั้น เป็นต้น

การวัดแนวโน้มเข้าสู่ส่วนกลาง มีวิธีหาได้หลายวิธีคือ

1. ตัวกลางเลขคณิต (Arithmetic mean) หรือค่าเฉลี่ย
2. มัธยฐาน (Median)
3. ฐานนิยม (Mode)
4. ตัวกลางเรขาคณิต (Geometric mean)
5. ตัวกลางฮาร์โมนิก (Harmonic mean)

แต่ละวิธีมีทั้งข้อดีและข้อเสีย และความถูกต้องเหมาะสมในการนำไปใช้ไม่เหมือนกันขึ้นอยู่กับลักษณะของข้อมูล และวัตถุประสงค์ของผู้ใช้ข้อมูลชนิดนั้นๆ ซึ่งจะให้ออกข้อสังเกตไว้ในท้ายบทนี้ สำหรับในการวิจัยทางการศึกษา วิธีที่ใช้มากที่สุดคือ ตัวกลางเลขคณิต ซึ่งจะกล่าวอย่างละเอียด ส่วนวิธีมัธยฐานและฐานนิยมมีโอกาสนำมาใช้บ้าง ก็จะกล่าวไว้พอสังเขป ส่วนตัวกลางเรขาคณิตและตัวกลางฮาร์โมนิกจะไม่กล่าวถึงเพราะมีโอกาสใช้น้อยมากในการวิจัยทางการศึกษา