

สถิติธุรกิจ และ การวิเคราะห์ข้อมูล ด้วยโปรแกรม SPSS

เหมาะสำหรับนักศึกษาคณะบริหารธุรกิจ
รายวิชาสถิติธุรกิจ สาขาวิชาการบัญชี
การเงินและการธนาคาร การจัดการธุรกิจ
การจัดการโรงแรม การตลาด
การจัดการธุรกิจระหว่างประเทศ
และระบบสารสนเทศ
เพื่อการจัดการ

พศ.ดร. พิษณาตา พันพา

สาขาวิชาวิทยาการวิจัยทางการศึกษา (GPA 3.92)
จุฬาลงกรณ์มหาวิทยาลัย

inspiration starts here

สถิติธุรกิจและการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS

โดย ผศ.ดร. พิษณุดา พันผา

สงวนลิขสิทธิ์ตามกฎหมาย โดย ผศ.ดร. พิษณุดา พันผา © พ.ศ. 2566

ห้ามคัดลอก ลอกเลียน ดัดแปลง ทำซ้ำ จัดพิมพ์ หรือกระทำการอื่นใด โดยวิธีการใดๆ ในรูปแบบใดๆ
ไม่ว่าส่วนหนึ่งส่วนใดของหนังสือเล่มนี้ เพื่อเผยแพร่ในสื่อทุกประเภท หรือเพื่อวัตถุประสงค์ใดๆ
นอกจากจะได้รับอนุญาต

ข้อมูลทางบรรณานุกรมของหอสมุดแห่งชาติ

พิษณุดา พันผา.

สถิติธุรกิจและการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS. -- กรุงเทพฯ : ซีเอ็ดดูเคชั่น, 2566.
352 หน้า.

1. ธุรกิจ -- การประมวลผลข้อมูล. 2. ธุรกิจศึกษา.
 3. สถิติการค้า. I. ชื่อเรื่อง.
- 650.0285

Barcode (e-book) 9786160847488

inspiration starts here

ผลิตและจัดจำหน่ายโดย

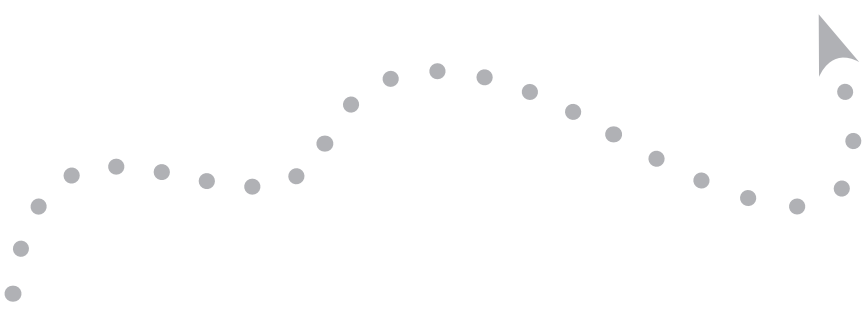


บริษัท ซีเอ็ดดูเคชั่น จำกัด (มหาชน)
SE-EDUCATION PUBLIC COMPANY LIMITED

เลขที่ 1858/87-90 ถนนเทพรัตน แขวงบางนาใต้ เขตบางนา กรุงเทพฯ 10260
โทรศัพท์ 0-2826-8000

หากมีคำแนะนำหรือติชม สามารถติดต่อได้ที่ comment@se-ed.com

คำนำ



จากการที่ผู้เขียนได้สอนวิชา สถิติธุรกิจและการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS และ วิชาสถิติธุรกิจและวิเคราะห์ปริมาณ 1 มาหลายปี นิสิต นักศึกษาที่เรียนก็มีหลากหลายสาขาวิชา เช่น บัญชี การเงินและการธนาคาร การจัดการธุรกิจ การจัดการการโรงแรม การตลาด การจัดการธุรกิจระหว่างประเทศ และระบบสารสนเทศเพื่อการจัดการ ซึ่งจากประสบการณ์ที่ผ่านมา นิสิต นักศึกษาในสาขาต่าง ๆ บางกลุ่มยังไม่มีพื้นฐานทางด้านสถิติ หรือบางกลุ่มมีพื้นฐานทางด้านสถิติมาบ้าง หรือบางกลุ่มก็มีพื้นฐานความรู้ทางสถิติมาพอสมควร หนังสือเล่มนี้จึงเขียนขึ้นเพื่อให้นักศึกษาที่เรียนในศาสตร์บริหารธุรกิจใช้ประกอบการเรียนวิชาสถิติธุรกิจ ซึ่งมีวัตถุประสงค์รายวิชาเพื่อให้ผู้เรียนสามารถนำความรู้ทางสถิติที่ได้จากการเรียนวิชานี้ไปเป็นพื้นฐานในการทำวิจัยทางธุรกิจ วิจัยทางสังคมศาสตร์ หรือเรียนต่อในระดับปริญญาโท และผู้เรียนสามารถนำความรู้สถิติธุรกิจไปประยุกต์ใช้ในงาน ในการประกอบอาชีพเพื่อประกอบการตัดสินใจในทางธุรกิจหรือทางสังคมศาสตร์ได้

เนื้อหาในหนังสือเล่มนี้ ประกอบด้วย ความหมายและความสำคัญของสถิติ สถิติเชิงบรรยาย ความน่าจะเป็น สถิติเชิงอ้างอิง ได้แก่ t-test, z-test, One-Way ANOVA, Two-Way ANOVA สถิติแบบนอนพาราเมตริกแบบไคกำลังสอง (Chi-square) การวิเคราะห์สหสัมพันธ์และการวิเคราะห์ถดถอยเชิงเส้นอย่างง่ายและพหุคูณ การวิเคราะห์อนุกรมเวลา เลขดัชนี และการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS for Windows พร้อมด้วยตัวอย่างงานวิจัยหลายเรื่องที่ใช้สถิติดังกล่าว ซึ่งการวิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูปทางสถิติมีความสำคัญมากในการทำวิจัยเพราะสามารถวิเคราะห์ข้อมูลที่สลับซับซ้อนและมีจำนวนมากได้ และสามารถวิเคราะห์ข้อมูลได้ถูกต้อง แม่นยำ รวดเร็ว และประหยัดเวลาในการวิเคราะห์

ผู้เขียนขอกราบขอบพระคุณ รศ.ดร. กัลยา วาณิชบัญชา ภาควิชาสถิติ คณะพาณิชยศาสตร์ และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย รศ.ดร. สุชาดา บวรกิตติวงศ์ สาขาสถิติทางการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย และ รศ.ดร. สุพรรณิ อึ้งปัญญ์ตวงศ์ ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ที่ช่วยพิจารณาและเสนอแนะให้หนังสือเล่มนี้สมบูรณ์มากยิ่งขึ้น ผู้เขียนหวังว่าหนังสือเล่มนี้จะมีประโยชน์แก่นิสิต นักศึกษา นักวิจัย และผู้สนใจทั่วไป หนังสือเล่มนี้มีข้อบกพร่องประการใด ถ้าท่านผู้อ่านมีข้อเสนอแนะใด กรุณาส่งกลับมายังผู้เขียน ผู้เขียนขอน้อมรับข้อเสนอแนะ เพื่อการปรับปรุงแก้ไขให้ดียิ่งขึ้น และขอขอบพระคุณเป็นอย่างยิ่ง

ผศ.ดร. พิษณุดา พันผา

สาขาการจัดการธุรกิจ

คณะบริหารศาสตร์ มหาวิทยาลัยอุบลราชธานี

SE-ED

inspiration starts here

สารบัญ

	หน้า
บทที่ 1 ความรู้เบื้องต้นทางสถิติ	9
1.1 การใช้สถิติเพื่อการตัดสินใจทางธุรกิจ	12
1.2 ประเภทของสถิติ	13
1.3 วงจรการคำนวณในทางสถิติ	14
1.4 มาตรวัดของข้อมูล (Measurement Scales of Data)	15
1.5 สถิติเชิงบรรยาย (Descriptive Statistics)	17
1.6 ขั้นตอนการสร้างตารางแจกแจงความถี่	19
1.7 การวัดแนวโน้มเข้าสู่ส่วนกลาง (Measures of Central Tendency)	23
1.8 การวัดการกระจายของข้อมูล (Measure of Distribution)	27
1.9 การวัดรูปร่างของข้อมูล (Measure of Shape)	29
1.10 การวัดตำแหน่งของข้อมูล (Measure of Position)	32
1.11 การทบทวนคณิตศาสตร์ระดับมัธยมศึกษา	36
1.12 การวิเคราะห์ค่าสถิติเบื้องต้นด้วยโปรแกรม SPSS	39
1.13 สรุปเนื้อหาบทที่ 1	45
1.14 สรุปสูตรบทที่ 1	46
แบบฝึกหัดบทที่ 1	47
บทที่ 2 ตัวแปรสุ่มและการแจกแจงความน่าจะเป็น	49
2.1 ความหมายและประเภทของตัวแปรสุ่ม	50
2.2 การแจกแจงความน่าจะเป็น (Probability Distribution)	51
2.3 เงื่อนไขของความน่าจะเป็น	51
2.4 ค่าคาดหวังหรือค่าเฉลี่ยของตัวแปรสุ่ม (Expected Value; $E(X)$)	52
2.5 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Probability Distributions)	59
2.6 การแจกแจงความน่าจะเป็นแบบต่อเนื่อง	70
2.7 สรุปเนื้อหาบทที่ 2	83
2.8 สรุปสูตรบทที่ 2	83
แบบฝึกหัดบทที่ 2	84
บทที่ 3 การสุ่มตัวอย่าง	91
3.1 การสุ่มตัวอย่าง	92
3.2 ประเภทของการสุ่มตัวอย่าง	93

3.3	ขนาดกลุ่มตัวอย่าง (Sample Size)	100
3.4	สรุปเนื้อหาบทที่ 3	112
3.5	สรุปสูตรบทที่ 3	113
	แบบฝึกหัดบทที่ 3	115

บทที่ 4 การประมาณค่าพารามิเตอร์ 116

4.1	นิยามของการประมาณค่า (Estimation)	117
4.2	ชนิดของการประมาณค่า	117
4.3	สถิติทดสอบซี (Z-test)	119
4.4	สถิติทดสอบที (t-test)	122
4.5	การประมาณค่าแบบจุด (Point Estimation)	124
4.6	การประมาณค่าแบบช่วง (Interval Estimation)	124
4.7	การประมาณค่าผลต่างระหว่างค่าเฉลี่ยของประชากร 2 กลุ่ม ($\mu_1 - \mu_2$)	128
4.8	การประมาณค่าสัดส่วนประชากร	133
4.9	สรุปเนื้อหาบทที่ 4	137
4.10	สรุปสูตรบทที่ 4	138
	แบบฝึกหัดบทที่ 4	140

บทที่ 5 การทดสอบสมมติฐาน 142

5.1	ความหมายการทดสอบสมมติฐานและประเภทของสมมติฐาน	143
5.2	ขั้นตอนการทดสอบสมมติฐาน	145
5.3	การทดสอบสมมติฐานค่าเฉลี่ยประชากร 1 กลุ่ม	147
5.4	การทดสอบสมมติฐานผลต่างค่าเฉลี่ยประชากร 2 กลุ่ม ($\mu_1 - \mu_2$)	147
5.5	ประชากร 1 กลุ่ม แต่วัดซ้ำ 2 ครั้ง หรือประชากร 2 กลุ่มที่สัมพันธ์กัน (Paired Sample)	148
5.6	การทดสอบค่าสัดส่วนของประชากร (P)	149
5.7	การวิเคราะห์การประมาณค่าและทดสอบสมมติฐานค่าเฉลี่ยในโปรแกรม SPSS	157
5.8	ตัวอย่างงานวิจัยที่ใช้ t-test แบบเปรียบเทียบผลต่างค่าเฉลี่ยของประชากร 2 กลุ่ม ($\mu_1 - \mu_2$)	160
5.9	สรุปเนื้อหาบทที่ 5	163
5.10	สรุปสูตรบทที่ 5	164
	แบบฝึกหัดบทที่ 5	166

บทที่ 6 การวิเคราะห์ความแปรปรวน 168

6.1	การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA)	169
6.2	การวิเคราะห์ความแปรปรวนแบบจำแนก 2 ทาง (Two-way ANOVA)	186
6.3	สรุปเนื้อหาบทที่ 6	197
6.4	สรุปสูตรบทที่ 6	198
	แบบฝึกหัดบทที่ 6	202

บทที่ 7 การทดสอบไคกำลังสอง 205

7.1	ไคกำลังสองหรือไคสแควร์	206
7.2	ข้อตกลงในการทดสอบไคกำลังสอง (χ^2 -test)	208
7.3	การทดสอบไคกำลังสอง (χ^2) แบบลักษณะเดียว	209
7.4	การทดสอบไคกำลังสองแบบตัวแปร 2 ตัว	212

7.5	การแปลความหมายความสัมพันธ์ของสถิติ Phi (ϕ) และ Cramer's V	214
7.6	การวิเคราะห์ไคสแควร์ (χ^2) ด้วยโปรแกรม SPSS	219
7.7	สรุปเนื้อหาบทที่ 7	221
7.8	สรุปสูตรบทที่ 7	222
	แบบฝึกหัดบทที่ 7	223

บทที่ 8 การวิเคราะห์สหสัมพันธ์และการวิเคราะห์ถดถอย 225

8.1	การวิเคราะห์สหสัมพันธ์	226
8.2	สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's Correlation Coefficient)	227
8.3	ทิศทางของความสัมพันธ์ (Direction of the Relationship)	227
8.4	ข้อตกลงเบื้องต้นของการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์	228
8.5	ข้อควรระวังในการวิเคราะห์สหสัมพันธ์	229
8.6	ข้อสังเกตในการแปลความหมายของค่าสัมประสิทธิ์สหสัมพันธ์	229
8.7	การแปลความหมายค่าสัมประสิทธิ์สหสัมพันธ์	229
8.8	สัมประสิทธิ์การอธิบาย (Coefficient of Determination)	229
8.9	การทดสอบความสัมพันธ์กันเชิงเส้นตรงระหว่างตัวแปร X กับ Y (Hypothesis Testing)	230
8.10	การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Regression Analysis)	236
8.11	โมเดลการถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Model)	236
8.12	ข้อตกลงเบื้องต้นในการวิเคราะห์การถดถอย	238
8.13	การหาสมการถดถอยเชิงเส้นอย่างง่ายถ้าตัวแปรทำนายคือเวลา	241
8.14	การใช้โปรแกรมคอมพิวเตอร์ช่วยในการวิเคราะห์ข้อมูล	244
8.15	ขั้นตอนการวิเคราะห์สหสัมพันธ์ด้วยโปรแกรม SPSS	245
8.16	ขั้นตอนการวิเคราะห์ถดถอยเชิงเส้นอย่างง่ายด้วยโปรแกรม SPSS	246
8.17	วิธีการคัดเลือกตัวแปรในโปรแกรม SPSS	247
8.18	การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis)	250
8.19	สมการพยากรณ์การถดถอยพหุคูณในรูปแบบคะแนนดิบ	250
8.20	สมการพยากรณ์การถดถอยพหุคูณในรูปแบบมาตรฐาน	251
8.21	ความสัมพันธ์ระหว่างสัมประสิทธิ์คะแนนดิบ (b) กับสัมประสิทธิ์คะแนนมาตรฐาน (β)	251
8.22	ข้อตกลงเบื้องต้นของการวิเคราะห์ถดถอยพหุคูณ	252
8.23	ตัวอย่างงานวิจัยการวิเคราะห์ถดถอยพหุคูณ	253
8.24	สรุปเนื้อหาบทที่ 8	261
8.25	สรุปสูตรบทที่ 8	262
	แบบฝึกหัดบทที่ 8	263

บทที่ 9 การวิเคราะห์ข้อมูลอนุกรมเวลา 265

9.1	ความหมายของการวิเคราะห์อนุกรมเวลา	266
9.2	ความหมายพยากรณ์และประโยชน์การพยากรณ์ด้านต่าง ๆ	266
9.3	การพยากรณ์เชิงปริมาณ (Quantitative Forecasting)	268
9.4	การพยากรณ์เชิงคุณภาพ (Qualitative Forecasting)	268
9.5	ลักษณะข้อมูลที่จะนำมาวิเคราะห์อนุกรมเวลา	269

9.6 องค์ประกอบของอนุกรมเวลา	269
9.7 สูตรการคำนวณโมเดลอนุกรมเวลา	271
9.8 วิธีอัตราส่วนระหว่างข้อมูลอนุกรมเวลากับค่าแนวโน้ม	274
9.9 วิธีอัตราส่วนระหว่างข้อมูลอนุกรมเวลากับค่าเฉลี่ยเคลื่อนที่	275
9.10 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยโปรแกรม SPSS	279
9.11 การวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธี Seasonal Decomposition	280
9.12 สรุปเนื้อหาบทที่ 9	284
9.13 สรุปสูตรบทที่ 9	285
แบบฝึกหัดบทที่ 9	286

บทที่ 10 เลขดัชนี 290

10.1 ความหมายของเลขดัชนี	291
10.2 ประโยชน์ของเลขดัชนี	291
10.3 ประเภทของเลขดัชนี (Type of Index Number)	292
10.4 สัญลักษณ์ในการคำนวณเลขดัชนี	292
10.5 การคำนวณเลขดัชนีอย่างง่าย	293
10.6 สูตรในการคำนวณเลขดัชนีราคา	295
10.7 เลขดัชนีราคาผู้บริโภค (Consumer Price Index)	302
10.8 ประโยชน์ของดัชนีราคาผู้บริโภค	303
10.9 ข้อจำกัดของดัชนีราคาผู้บริโภค	304
10.10 การเปลี่ยนพื้นฐานของเลขดัชนี	305
10.11 สรุปเนื้อหาบทที่ 10	307
10.12 สรุปสูตรบทที่ 10	308
แบบฝึกหัดบทที่ 10	309

บทที่ 11 การใช้โปรแกรมคอมพิวเตอร์สำหรับการวิเคราะห์ข้อมูล 312

11.1 ความรู้เบื้องต้นเกี่ยวกับ SPSS	313
11.2 หน้าต่าง SPSS เมื่อเปิดโปรแกรม	313
11.3 เมนูคำสั่งหลัก	317
11.4 การสร้างแฟ้มข้อมูล	320
11.5 การคำนวณค่าสถิติพื้นฐาน	321
11.6 การแจกแจงความถี่แบบ 2 ทาง (Crosstabs)	323
11.7 การคำนวณค่าสหสัมพันธ์ (Correlation)	323
11.8 คำสั่งที่ใช้ในการประมาณค่าและการทดสอบสมมติฐานของผลต่างค่าเฉลี่ยของ 2 กลุ่ม ($\mu_1 - \mu_2$)	323
11.9 คำสั่งที่ใช้ในการประมาณค่าและการทดสอบสมมติฐานของค่าเฉลี่ยของการวัดซ้ำ 2 ครั้ง	324
11.10 การวิเคราะห์หัตถด้อยเชิงเส้นอย่างง่ายและการวิเคราะห์หัตถด้อยเชิงเส้นพหุคูณ	324
11.11 การจัดการผลการวิเคราะห์และการออกจากโปรแกรม	325

บรรณานุกรม 327

ภาคผนวก 331

ความรู้เบื้องต้นทางสถิติ

(Introduction to Statistics)

1

วัตถุประสงค์การเรียนรู้

เมื่อนักศึกษาเรียนจบบทนี้แล้ว

1. มีความรู้ ความเข้าใจความหมายของสถิติ ประเภทของสถิติ สถิติธุรกิจและมาตรวัดข้อมูล
2. สามารถคำนวณหาค่าสถิติเชิงบรรยาย ได้แก่ การวัดแนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจายของข้อมูล การวัดรูปร่างของข้อมูล การหาตำแหน่งของข้อมูล
3. มีความรู้ ความเข้าใจ และสามารถคำนวณคณิตศาสตร์

Inspiration starts here

สถิติมีประโยชน์ต่อทุกสาขาวิชาและทุกหน่วยงาน เพราะทุกสาขาวิชาและหน่วยงานต่าง ๆ ทั้งภาครัฐบาลและภาคเอกชนจะต้องมีการศึกษาข้อมูล เก็บรวบรวมข้อมูลเพื่อสรุปผลไปใช้ในการวางแผนและการตัดสินใจ เพื่อใช้ในการดำเนินงานและใช้เป็นข้อมูลสำหรับการตัดสินใจแก้ปัญหาต่าง ๆ อยู่ตลอดเวลา ซึ่งประโยชน์สำหรับงานในด้านต่าง ๆ พอสังเขปเป็นดังนี้

ด้านการพัฒนาประเทศ การพัฒนาประเทศไม่ว่าจะเป็นทางด้านเศรษฐกิจ หรือด้านสังคม รัฐบาลต้องทราบสถิติข้อมูลด้านต่าง ๆ เช่น ด้านประชากร ด้านเศรษฐกิจหรือสังคมในปัจจุบัน ตัวอย่าง เช่น สถิติการนำเข้า-ส่งออกสินค้า สถิติของราคาน้ำมันในตลาดโลก สถิติการศึกษา สถิติแรงงาน สถิติการเกษตร สถิติการเกิดการตาย ปริมาณนักท่องเที่ยวในแต่ละปี เป็นต้น เพื่อให้สามารถวางแผนหรือกำหนดนโยบายระยะสั้นและระยะยาวในการพัฒนาประเทศให้ถูกต้องและดียิ่งขึ้น

ด้านธุรกิจ ในการทำธุรกิจ นักธุรกิจจะต้องมีการวางแผนในการดำเนินธุรกิจ เช่น การวางแผนการลงทุน การผลิต การจำหน่าย การบริหารจัดการ การจัดส่งสินค้า การสต็อกสินค้า เป็นต้น เมื่อเกิดปัญหาเกี่ยวกับธุรกิจ เช่น ปัญหาด้านการผลิต ปัญหาด้านการตลาด ปัญหาเกี่ยวกับแรงงาน และปัญหาคุณภาพสินค้า เป็นต้น นักธุรกิจจะต้องทำการตัดสินใจเพื่อแก้ปัญหาหรือดำเนินการตามแผนการธุรกิจ ซึ่งการวางแผนที่ดีและการตัดสินใจที่ถูกต้องนั้นย่อมอาศัยข้อมูลสนับสนุนที่ถูกต้องส่งผลให้การดำเนินทางธุรกิจนั้นราบรื่นและมีโอกาสประสบความสำเร็จสูง ดังนั้นการวางแผนและการตัดสินใจทางธุรกิจที่ดีจำเป็นต้องอาศัยข้อมูลทางสถิติมาช่วยประกอบการตัดสินใจในการทำธุรกิจ

ด้านการเกษตร ในการทำการเกษตร ผลผลิตคือเป้าหมายที่เกษตรกรต้องการทำให้มากที่สุด ในระยะเวลาที่ต้องการ การวางแผนการดำเนินการ เช่น เพาะปลูก ไถพรวนดิน ช่วงเวลาใส่ปุ๋ย ให้น้ำ เก็บเกี่ยวผลผลิต จัดจำหน่าย เกษตรกรมีความจำเป็นต้องใช้ข้อมูลทางสถิติมาช่วยในการประกอบการตัดสินใจ เช่น ข้อมูลปริมาณน้ำ ข้อมูลราคาปุ๋ย ราคาน้ำมัน พยากรณ์อากาศ พยากรณ์ราคาพืชผลการเกษตร เป็นต้น ดังนั้นในการทำการเกษตรจึงต้องอาศัยข้อมูลจากสถิติต่างๆ ที่เกี่ยวข้องเพื่อวางแผนในการทำการเกษตรและเพื่อให้ได้ผลผลิตตามเป้าหมายที่คาดหวัง

ด้านการศึกษา ในด้านการศึกษา สถิติถูกนำมาใช้เพื่อการพัฒนาคุณภาพการศึกษามากมาย เช่น สถิติการสอบเข้าเรียน การออกกลางคันของนักศึกษา สถิติการกู้ยืมเงินของนักศึกษา สถิติการใช้ห้องสมุด หรือนำสถิติมาใช้ในการบวนการเรียนการสอน การทำวิจัยในชั้นเรียน เป็นต้น

นอกจากด้านที่กล่าวมาแล้วนั้นในงานด้านอื่นๆ เช่น ด้านการแพทย์ ด้านวิศวกรรม ด้านวิทยาศาสตร์ ด้านการพยาบาล ด้านการกีฬา ด้านการขนส่ง ด้านการคมนาคม ด้านการท่องเที่ยว เป็นต้น ก็ได้นำสถิติมาใช้อย่างกว้างขวาง โดยเฉพาะนักวิจัยและนักวิชาการในสาขาวิชาชีพต่างๆ ที่นำสถิติมาอธิบายถึงผลการวิจัย ช่วยทดสอบเครื่องมือเก็บข้อมูลให้มีคุณภาพ การวิเคราะห์ข้อมูลก็ใช้สถิติมาสรุปอ้างอิง การนำเสนอข้อมูล ข้อมูลที่ได้จากการวิเคราะห์ทางสถิติก็สามารถนำไปใช้ในการวางแผนและการตัดสินใจ ดังนั้นจึงสามารถกล่าวได้ว่าสถิติได้เข้ามามีบทบาทในทุกสาขาวิชาชีพเกือบทุกด้าน

สถิติในทางธุรกิจนั้น ถูกนำมาใช้เป็นเครื่องมือในการวิเคราะห์และวิจัยเพื่อใช้ประกอบการตัดสินใจและวางแผนในการแก้ไขปัญหาต่างๆ ทางธุรกิจ โดยเฉพาะการหาแนวโน้ม การทำนาย การเปรียบเทียบ การหาความสัมพันธ์ หรือการหาความน่าจะเป็น การหาโอกาสในการลงทุนที่จะประสบความสำเร็จในธุรกิจนั้นๆ การที่นักลงทุน ผู้ประกอบการ หรือองค์กรต่างๆ มีข้อมูลทางสถิติเพื่อประกอบการตัดสินใจและวางแผนเกี่ยวกับการประกอบธุรกิจที่เหมาะสม ก็จะทำให้การดำเนินงานของธุรกิจผิดพลาดน้อย ลงทุนถูกต้อง และดำเนินธุรกิจได้อย่างมีประสิทธิภาพมากยิ่งขึ้น ข้อมูลทางสถิติเป็นสิ่งที่ต้องการและใช้กันอย่างกว้างขวาง รวมถึงแพร่หลายในวงการธุรกิจ เอกชน ธุรกิจทุกขนาดต้องอาศัยข้อมูลในการวางแผนด้านต่างๆ อย่างรอบคอบ เพื่อให้การดำเนินธุรกิจมีโอกาสที่จะประสบความสำเร็จ การวางแผนด้านการผลิต การตลาด การโฆษณา การกำหนดราคาสินค้าหรือบริการให้

เหมาะสมกับกำลังซื้อและสภาวะการแข่งขัน จะต้องอาศัยการศึกษาและการวิเคราะห์ข้อมูลต่าง ๆ ด้วยสถิติ เพื่อนำผลการวิเคราะห์ข้อมูลที่ได้มาใช้ให้เป็นประโยชน์ในการวางแผน และการตัดสินใจในด้านต่าง ๆ เพื่อเพิ่มความสำเร็จในการดำเนินการของธุรกิจต่อไป

เมื่อกล่าวถึง “สถิติ (Statistics)” หลายคนคงนึกถึงตัวเลข ข้อมูล จำนวนต่าง ๆ ที่แสดงในรูปแบบ กราฟ แผนภูมิ หรือตาราง เช่น จำนวนคนไทยในแต่ละจังหวัด จำนวนนักธุรกิจในจังหวัดอุบลราชธานี ปริมาณการส่งออก-นำเข้าสินค้า ซึ่งสถิตินั้นได้มีนักวิจัยและนักวิชาการได้ให้ความหมายไว้ในแนวทางเดียวกัน ดังนี้

กัลยา วานิชย์บัญชา (2551) กล่าวว่า **สถิติ** หมายถึงศาสตร์ที่ว่าด้วยการเก็บรวบรวมข้อมูล การนำเสนอข้อมูล และการวิเคราะห์ข้อมูล

อวยพร เรืองตระกูล (2552) ได้ให้ความหมายของสถิติว่า หมายถึงตัวเลขที่ใช้อธิบายหรือบรรยายลักษณะของข้อมูลจากตัวอย่างที่ศึกษา และอีกความหมายหนึ่งหมายถึงศาสตร์แขนงหนึ่งซึ่งประกอบด้วยกระบวนการ 4 ขั้นตอนคือ

- (1) การเก็บรวบรวมข้อมูล
- (2) การนำเสนอข้อมูล
- (3) การวิเคราะห์ข้อมูล
- (4) การแปลความหมายและสรุปผลข้อมูล

ลินด์ เอ. (Lind A.), แมนสัน อาร์. ดี. และ มาร์ชัล ดับเบิลยู. ดี. (Manson R. D. & Marchal W. D. 2000) ได้ให้ความหมายของวิชาสถิติว่า คือวิชาหรือศาสตร์ในการรวบรวมข้อมูล การจัดการข้อมูล การนำเสนอ การวิเคราะห์ข้อมูล และการแปลความหมายข้อมูลเพื่อช่วยในการตัดสินใจให้มีประสิทธิภาพมากยิ่งขึ้น

ดังนั้นผู้เขียนจึงสรุปความหมายได้ว่า **สถิติ** หมายถึงศาสตร์ที่ศึกษาเกี่ยวกับการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การแปลความหมาย การสรุปผล และการนำเสนอข้อมูล หรืออีกความหมายหนึ่งคือ ตัวเลขหรือจำนวนที่ใช้อธิบาย บรรยายลักษณะข้อมูลจากตัวอย่างที่เราสนใจศึกษาที่ได้มาจากกระบวนการสุ่มจากประชากร ข้อมูลที่เป็นตัวเลข เช่น รายได้ ค่าใช้จ่าย น้ำหนัก ปริมาณการส่งออก ปริมาณการนำเข้าสินค้า ราคาสินค้า ระดับความพึงพอใจของลูกค้า เป็นต้น

ส่วน “**สถิติทางธุรกิจ**” หมายถึงศาสตร์ที่ศึกษาเกี่ยวกับการเก็บรวบรวมข้อมูล ตัวเลข หรือจำนวนทางด้านธุรกิจ เช่น กำไร ขาดทุน รายรับ รายจ่าย ยอดการลงทุน ราคาสินค้า ต้นทุนราคา เป็นต้น แล้วนำข้อมูลมาวิเคราะห์ แปลความหมาย สรุปผล และนำเสนอข้อมูล เพื่อบรรยายลักษณะข้อมูลจากตัวอย่างที่สนใจศึกษาซึ่งข้อมูลที่สนใจศึกษาจะเป็นข้อมูลทางธุรกิจ ผลการวิเคราะห์ข้อมูลที่ได้จะเป็นประโยชน์ต่อการดำเนินการทางธุรกิจในลักษณะต่าง ๆ เช่น การวางแผนการลงทุน การขยายการลงทุน หรือการตัดสินใจซื้อสินค้า เป็นต้น

1.1 การใช้สถิติเพื่อการตัดสินใจทางธุรกิจ

ผู้เขียนประยุกต์มาจากแนวคิดของลำไผ่ ตระกูลสันติ (2559) ดังนี้

1. ด้านการบัญชี องค์กรต่าง ๆ ต้องใช้ข้อมูลทางบัญชี ในการวางแผนและควบคุม การดำเนินการ การตัดสินใจลงทุนและการขยายธุรกิจ ข้อมูลต้นทุนที่ใช้ในการลงทุน อาจได้มาจาก งบประมาณรายรับ-รายจ่าย เช่น ข้อมูลรายรับ-รายจ่ายทางการเงินใน 3 ปีที่ผ่านมา โครงการที่จะลงทุนอีก 5 ปีข้างหน้า นโยบายในการบริหารงาน ต้นทุนในแต่ละปี ซึ่งการที่จะดำเนินโครงการต่าง ๆ ในองค์กรนั้น ก็ต้องใช้ข้อมูลทางบัญชีที่ทำเป็นสถิติไว้ เพื่อประกอบการวางแผนการตัดสินใจ การลงทุนโครงการในอนาคต การวางแผนค่าใช้จ่ายก่อนที่จะมีการใช้จ่ายจริง

2. ด้านการเงิน ในธุรกิจนั้นข้อมูลด้านการเงินที่สำคัญ เช่น อัตราดอกเบี้ย จำนวนเงินฝาก การกู้เงิน การออม การลงทุนกับกองทุนต่าง ๆ หรือการซื้อหุ้นนั้น จะเป็นประโยชน์ต่อการตัดสินใจ ในด้านการเงิน เช่น การพิจารณากลุ่มลูกค้าเป้าหมาย การให้กู้เงินลงทุน จำนวนเงินที่สมควรให้การประเมินผลโครงการ หรืออัตราเสี่ยงในการลงทุน เป็นต้น

3. ด้านการผลิต จะมีส่วนสัมพันธ์กับคนและเครื่องจักรเป็นส่วนใหญ่ ปัจจุบันการควบคุมคุณภาพของสินค้า มีความสำคัญในการดำเนินธุรกิจอย่างมาก วิธีการในการควบคุมคุณภาพสินค้า มี 2 วิธีคือ การควบคุมกรรมวิธีระหว่างการผลิต และการตรวจสอบสินค้าที่ผลิตแล้ว วิธีการทางสถิติ จะมีบทบาทในการควบคุมคุณภาพทั้งในระหว่างการผลิตและการตรวจสอบสินค้าที่ผลิตแล้ว

4. ด้านการตลาด การวางแผนเกี่ยวกับผลิตภัณฑ์ สร้างภาพพจน์ ตราสินค้า แนะนำผลิตภัณฑ์ใหม่ กำหนดตลาดเป้าหมาย การกำหนดตราสินค้า ตำแหน่งผลิตภัณฑ์ และการหาส่วนแบ่งทางการตลาด ผู้บริหารขององค์กรจะต้องวิเคราะห์สถานการณ์ ทั้งการวิเคราะห์ปัจจัยภายใน และปัจจัยภายนอกองค์กร และการวิเคราะห์ความเสี่ยงขององค์กร

5. ด้านการบริหาร การบริหารเป็นหัวใจสำคัญของการดำเนินธุรกิจ เป็นการกำหนดเป้าหมายขององค์กรและกำหนดนโยบายสำคัญต่าง ๆ ในธุรกิจ เช่น กำหนดการแบ่งงานให้หน่วยงานต่าง ๆ ในการรับผิดชอบ การวางแผนบุคลากร การกำหนดแผนงานในการดำเนินงาน เป็นต้น ดังนั้นข้อมูลในระดับนี้ส่วนใหญ่มาจากผลสรุปเกี่ยวกับการดำเนินงานของฝ่ายต่าง ๆ ซึ่งส่วนใหญ่เป็นข้อมูลที่น่าสนใจในรูปสถิติ ตัวเลข แผนภูมิ กราฟการเปรียบเทียบ ตาราง เพื่อง่ายต่อการทำความเข้าใจ และตีความหมายได้อย่างชัดเจน เป็นรูปธรรมมากกว่าการนำเสนอเป็นข้อความ

6. ด้านการจัดการทรัพยากรบุคคล พนักงานถือเป็นส่วนสำคัญที่ทำให้การดำเนินงานขององค์กรประสบความสำเร็จ องค์กรจึงจำเป็นต้องมีข้อมูลเกี่ยวกับพนักงานเพื่อจัดการทรัพยากรบุคคลให้เหมาะสม ทั้งด้านการวางแผนงานบุคคล การขยายงาน การเพิ่ม-ลดพนักงานหรือการโยกย้ายงาน เป็นต้น

1.2 ประเภทของสถิติ

สถิติแบ่งออกเป็น 2 ประเภท

1.2.1 สถิติเชิงบรรยายหรือเชิงพรรณนา (Descriptive Statistics)

เป็นสถิติที่ใช้ในการนำเสนอลักษณะข้อมูลเบื้องต้น ใช้ในการศึกษาข้อเท็จจริงจากข้อมูลที่รวบรวมมาได้ เพื่อให้ทราบรายละเอียดเกี่ยวกับลักษณะของข้อมูลกลุ่มนั้นโดยไม่ได้สรุปอ้างอิงผลการศึกษาไปยังกลุ่มข้อมูลกลุ่มอื่น หรือไม่ได้สรุปอ้างอิงไปยังประชากรที่ศึกษา การบรรยายลักษณะของกลุ่มข้อมูล ได้แก่ การแจกแจงความถี่ การวัดตำแหน่ง การเปรียบเทียบ การวัดแนวโน้มเข้าสู่ส่วนกลาง และการวัดการกระจายของข้อมูล เป็นต้น

1.2.2 สถิติเชิงอ้างอิงหรือเชิงอนุมาน (Inferential Statistics)

สถิติเชิงอ้างอิงหรือเชิงอนุมานเป็นสถิติที่ว่าด้วยการสรุปอ้างอิงข้อมูลเพียงบางส่วนจากตัวอย่างไปยังประชากร คือเป็นการนำค่าที่ได้จากการวิเคราะห์ข้อมูลจากตัวอย่างที่เป็นตัวแทนประชากร สรุปอ้างอิงไปยังประชากรโดยใช้ทฤษฎีความน่าจะเป็น การประมาณค่า และการทดสอบสมมติฐาน ถ้าหากแบ่งกลุ่มสถิติอ้างอิงตามลักษณะพารามิเตอร์ จะแบ่งได้ 2 กลุ่ม (นงลักษณ์ วิรัชชัย, 2555) ได้แก่

1. **กลุ่มสถิติพารามетริก (Parametric Statistics)** เป็นกลุ่มสถิติวิเคราะห์ที่มีข้อกำหนดหรือข้อตกลงเบื้องต้นเกี่ยวกับค่าพารามิเตอร์และการแจกแจงของสถิติทดสอบ เช่น t-test Z-test F-test เป็นต้น

2. **กลุ่มสถิตินอนพารามетริก (Non-parametric Statistics)** หรือสถิติที่ไม่มีเงื่อนไขการแจกแจง ซึ่งเป็นสถิติวิเคราะห์ที่ไม่มีข้อกำหนดเกี่ยวกับค่าพารามิเตอร์ เช่น χ^2 -test, Kolmogorov-smirnov Test, Mann-whitney U Test, Kruskal-wallis One-way ANOVA เป็นต้น

ประชากร (Population) คือสมาชิกทุกหน่วยในเรื่องที่สนใจศึกษา ซึ่งอาจจะเป็นบุคคล กลุ่มบุคคล องค์กร สัตว์หรือสิ่งของก็ได้ เช่น เราสนใจผู้ประกอบการ SMEs ทุกคนในจังหวัดอุบลราชธานี ลูกค้าที่เข้ามาซื้อสินค้าในตลาดนัดคนเดินทุกคน นักท่องเที่ยวในภาคใต้ทุกคน พนักงานบริษัทเอกชนทั้งหมดทุกคนในกรุงเทพฯ เจ้าของโรงสีทุกคนในประเทศไทย หรือถ้าเป็นสัตว์ สิ่งของ เช่น ไก่ทุกตัวในฟาร์ม ทุเรียนทุกลูกในสวนในจังหวัดนนทบุรี รถยนต์ทุกคันในจังหวัดศรีสะเกษ เป็นต้น สัญลักษณ์ที่ใช้แทนจำนวนประชากรคือ N ประชากรแบ่งออกเป็น 2 ประเภทคือ

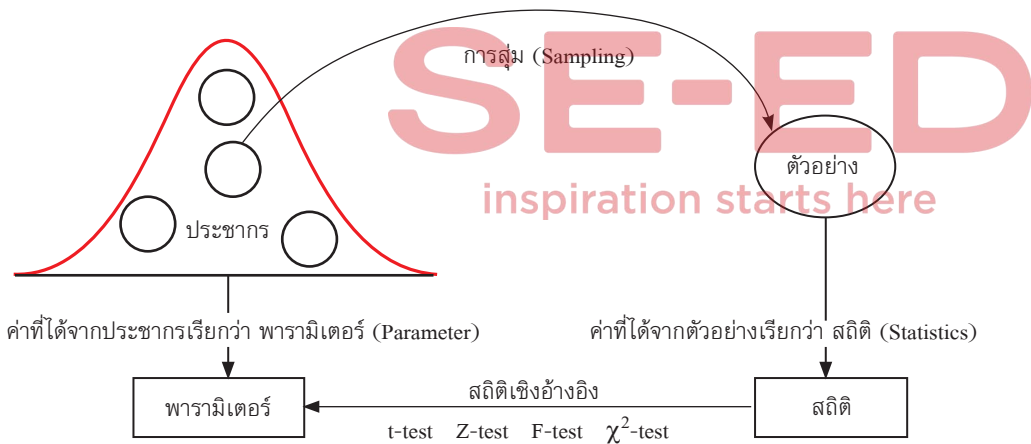
1. **ประชากรจำนวนจำกัด (Finite Population)** คือประชากรที่เราทราบจำนวนที่แน่นอนว่ามีอยู่เท่าไร เช่น นักธุรกิจขนาดกลางและขนาดย่อม (SMEs) ในจังหวัดอุบลราชธานีทั้งหมด 55,742 หน่วยธุรกิจ (ที่มา : ข้อมูลจาก <http://www.sme.go.th/> ของสำนักงานส่งเสริมวิสาหกิจขนาดกลางและขนาดย่อม, 2559) โดยสามารถแบ่งเป็นหมวดหมู่ได้ดังนี้ ธุรกิจขนาดเล็ก 55,630 ราย ธุรกิจขนาดกลาง 112 ราย ธุรกิจขนาดใหญ่ 56 ราย ทั้งหมดจำนวน 55,742 ราย เป็นต้น

2. ประชากรจำนวนไม่จำกัด (Non-finite Population) คือประชากรที่เราไม่ทราบจำนวนที่แน่นอน เช่น จำนวนลูกค้าที่ซื้อสินค้าผ่านอินเทอร์เน็ตในประเทศไทย จำนวนเมล็ดข้าวที่ผลิตได้ต่อ 1 ไร่ ปริมาณน้ำในแม่น้ำโขง เป็นต้น

ตัวอย่าง (Sample) คือส่วนหนึ่งของประชากรที่ได้มาจากการสุ่มจากประชากรทั้งหมดที่สนใจศึกษา หรือสมาชิกเพียงบางส่วน of ประชากรทั้งหมดที่เราสนใจศึกษาที่เป็นตัวแทนที่ดีของประชากร สัญลักษณ์ที่ใช้แทนจำนวนตัวอย่างคือ n

1.3 วงจรการคำนวณในทางสถิติ

วงจรของการคำนวณในทางสถิติหรือวิจัยจะเป็นกระบวนการดังนี้ สิ่งที่เราสนใจศึกษาทั้งหมดจะเรียกว่า **ประชากร** ค่าที่วิเคราะห์ได้จากประชากรเรียกว่า **พารามิเตอร์ (Parameter)** ซึ่งเป็นค่าที่ใช้บอกคุณลักษณะบางสิ่งที่เราสนใจศึกษาในระดับประชากร แต่เนื่องจากประชากรโดยส่วนใหญ่มีขนาดใหญ่มาก ทำให้นักวิจัยหรือนักสถิติไม่สามารถนำประชากรมาศึกษาทั้งหมดได้จึงทำการสุ่ม (Sampling) สิ่งที่เราสนใจศึกษาบางส่วนจากประชากรออกมา ซึ่งเราเรียกว่า **ตัวอย่าง** ค่าที่วิเคราะห์ได้จากตัวอย่างเรียกว่า **สถิติ (Statistics)** ซึ่งเป็นค่าที่ใช้บอกคุณลักษณะบางสิ่งที่เราสนใจศึกษาในระดับตัวอย่าง จากนั้นนักสถิติหรือนักวิจัยจึงนำค่าสถิติที่ได้จากตัวอย่างอ้างอิงกลับไปยังประชากรที่เป็นค่าพารามิเตอร์ โดยใช้สถิติเชิงอ้างอิง ได้แก่ t-test Z-test F-test χ^2 -test เป็นต้น วงจรของการคำนวณในทางสถิติสามารถแสดงได้ดังแสดงในรูปที่ 1.1



รูปที่ 1.1 วงจรของการคำนวณในทางสถิติ

สัญลักษณ์ค่าพารามิเตอร์และค่าสถิติที่สำคัญที่ใช้ในการคำนวณในบทต่าง ๆ ดังแสดงในตารางที่ 1.1 ดังนี้

ตารางที่ 1.1 สัญลักษณ์ค่าพารามิเตอร์และค่าสถิติที่สำคัญที่ใช้ในการคำนวณในบทต่างๆ

ประชากร (ค่าพารามิเตอร์)	ตัวอย่าง (ค่าสถิติ)
ค่าเฉลี่ย μ (Mu)	ค่าเฉลี่ย $\bar{x}$
จำนวนประชากร N	จำนวนตัวอย่าง n
ค่าสัดส่วน P	ค่าสัดส่วน p หรือ $\hat{p}$
ความแปรปรวน (Variance) σ^2	ความแปรปรวน (Variance) $s^2, S.D.^2, S^2$
ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation; SD) σ (Sigma)	ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation; SD) s, S.D., S
ความสัมพันธ์ ρ (Rho)	ความสัมพันธ์ r, R

1.4 มาตรฐานวัดของข้อมูล (Measurement Scales of Data)

ในชีวิตประจำวันของเรานั้นจะมีข้อมูลที่เป็นข้อเท็จจริงอยู่มากมายหรือข้อมูลที่เป็นค่าของตัวแปรที่เราเก็บรวบรวมได้จากหน่วยผู้ให้ข้อมูลต่างๆ โดย **ตัวแปรจะหมายถึงข้อมูลต่างๆ ที่เป็นสิ่งที่เราสนใจศึกษาและสามารถแปรค่าได้** เช่น ตัวแปรอายุ เป็นสิ่งที่เราสนใจศึกษา เมื่อเราได้หน่วยผู้ให้ข้อมูลมาจำนวน 10 คน เราก็จะได้ข้อมูลจากตัวแปรอายุนี้อีก 10 ค่า เป็นต้น ข้อมูลหรือตัวแปรที่เราสนใจศึกษามีความละเอียดมากน้อยแตกต่างกัน **ข้อมูลบางอย่างเป็นข้อมูลเพียงจัดกลุ่มหรือจัดประเภทเท่านั้น** บางข้อมูลเป็นตัวเลขที่มีความต่อเนื่อง และมีความละเอียดบวก ลบ คูณ หาร และเปรียบเทียบความมากน้อยกันได้ ดังนั้นเพื่อเป็นการวิเคราะห์ข้อมูลและเลือกใช้สถิติวิเคราะห์ข้อมูลให้ถูกต้อง ข้อมูลโดยทั่วๆ ไป จึงแบ่งออกเป็น 4 มาตรฐาน ดังนี้ (อวยพร เรื่องตระกูล, 2552; กัลยา วานิชย์บัญชา, 2551)

inspiration starts here

1.4.1 มาตรฐานบัญญัติ (Nominal Scale)

เป็นมาตรฐานวัดระดับต่ำสุดใช้กับข้อมูลเชิงคุณภาพที่สามารถจัดกลุ่ม หรือแบ่งกลุ่มพวกเดียวกัน แบ่งพวก (Classification) ข้อมูลที่มีสมบัติต่างกันจะแยกกลุ่มออกจากกันได้อย่างชัดเจน แต่ไม่สามารถบอกได้ว่า กลุ่มไหนดีกว่ากลุ่มไหน หรือกลุ่มใดมาก่อน การจัดกลุ่มให้เห็นเพียงความแตกต่างของสิ่งต่างๆ เท่านั้น

- ตัวอย่างเช่น
- เพศ แบ่งเป็นหญิงและชาย
 - เชื้อชาติ แบ่งเป็นไทย ลาว พม่า จีน ญี่ปุ่น อเมริกัน และอื่นๆ
 - สาขาอาชีพ แบ่งเป็นรับราชการ นักธุรกิจ เกษตรกร รัฐวิสาหกิจ เป็นต้น
 - สี แบ่งเป็นขาว เขียว น้ำเงิน แดง เงิน ทอง เหลือง และอื่นๆ
 - กลุ่มเลือด แบ่งเป็น A B AB และ O
 - ศาสนา แบ่งเป็นพุทธ คริสต์ อิสลาม ฮินดู และอื่นๆ

นอกจากนี้รหัสไปรษณีย์ บ้านเลขที่ เบอร์โทรศัพท์ สถานภาพสมรส หมายเลขหลังเลือนักกีฬา เป็นต้น ก็จัดเป็นนามบัญญัติด้วยเช่นกัน

ลักษณะข้อมูลในมาตรวัดนามบัญญัตินี้ เป็นข้อมูลที่ไม่สามารถบอกปริมาณมากน้อยหรือจัดระดับสูงต่ำระหว่างกลุ่มได้ ไม่สามารถนำข้อมูลมาบวก ลบ คูณ หาร ระหว่างกลุ่มได้ แม้จะได้ให้ค่าเป็นตัวเลข เช่น เลขที่บ้าน เบอร์โทรศัพท์ เลขไปรษณีย์แต่ละประเทศ เป็นต้น ตัวเลขดังกล่าวเป็นเพียงการกำหนดสัญลักษณ์เท่านั้น ไม่มีความหมายในเชิงปริมาณ ถ้ามีค่าสังเกตเพียง 2 ค่าหรือ 2 กลุ่ม เรียกตัวแปรเหล่านี้ว่า ตัวแปรทวินาม (Binomial Variable หรือ Dichotomous Variable) เช่น เพศ (ชาย-หญิง) ผลการสอบ (สอบได้-สอบตก) ชายของ (ชายได้-ชายไม่ได้) ตัดสินใจเลือก (เลือก-ไม่เลือก) เป็นต้น

1.4.2 มาตรเรียงอันดับ (Ordinal Scale)

เป็นมาตรวัดอันดับที่สูงกว่าหรือเพิ่มจากมาตรวัดนามบัญญัติ เป็นมาตรวัดที่สามารถจัดกลุ่มข้อมูลออกเป็นกลุ่มๆ ได้ สามารถจัดลำดับของข้อมูลว่ามาก ปานกลาง หรือน้อยได้ และบอกได้ว่ากลุ่มไหนดีกว่ากลุ่มไหน กลุ่มไหนมาก่อนมาหลัง แต่บอกความแตกต่างระหว่างกลุ่มว่าต่างกัน ปริมาณเท่าใดไม่ได้ เช่น

- ระดับการศึกษา แบ่งเป็นประถมศึกษา มัธยมศึกษา ระดับอนุปริญญา ปริญญาตรี ปริญญาโท และปริญญาเอก
- ยศทางทหาร แบ่งเป็นพลทหาร นายสิบ นายร้อย นายพัน และนายพล
- ระดับความคิดเห็น เช่น เห็นด้วยอย่างยิ่ง เห็นด้วย เฉย ๆ ไม่ค่อยเห็นด้วย หรือไม่เห็นด้วยอย่างยิ่ง
- ลำดับผลการประกวด แบ่งเป็นลำดับที่ 1 ที่ 2 ที่ 3 ... เป็นต้น

ลักษณะของข้อมูลในมาตรวัดเรียงอันดับนี้ ไม่สามารถบอกปริมาณช่วงห่างระหว่างลำดับได้ ข้อมูลที่ได้ไม่สามารถนำมาบวก ลบ คูณ หรือหารกันได้

เราเรียกข้อมูลในมาตรนามบัญญัติและมาตรวัดเรียงอันดับหรือมาตรวัด 1 และ 2 ว่าข้อมูลจัดกลุ่มหรือข้อมูลจัดประเภท (Discrete Data/Categorical Data) สถิติที่ใช้ในการวิเคราะห์ ได้แก่ ร้อยละ (Percentage) สัดส่วน (Proportion) ความถี่ (Frequency) เป็นต้น

1.4.3 มาตรอันตรภาค (Interval Scale)

เป็นมาตรวัดสูงกว่ามาตรนามบัญญัติและเรียงอันดับ เป็นมาตรวัดที่สามารถกำหนดช่วงเท่าได้ แต่ไม่มีศูนย์แท้มีแต่ศูนย์เทียม (Arbitrary Zero or Relative Zero : ศูนย์เทียมในที่นี้ไม่ได้แปลว่าไม่มี แต่เป็นศูนย์ที่กำหนดขึ้นว่าตรงนั้นคือศูนย์) เช่น คะแนนสอบ (ถ้าเราสอบได้ศูนย์คะแนนในวิชาสถิติธุรกิจ ไม่ได้หมายความว่านักศึกษาไม่มีความรู้เลย แต่หมายความว่าความรู้นักศึกษามีอยู่นั้น

ยังมีไม่เพียงพอที่จะสอบได้คะแนนวิชาสถิติธุรกิจ) เวลา ระดับ ความสามารถทางอารมณ์ (Emotion Quotient; EQ) ความสามารถทางจริยธรรม (Morality Quotient; MQ) ความสามารถทางสติปัญญา (Intelligence Quotient; IQ) ความสามารถในการเปลี่ยนแปลงอาชีพหรือการทำงาน (Protean-Career Quotient; PQ) ความสามารถทางเทคโนโลยี (Technology Quotient; TQ) ความสามารถทางวัฒนธรรม (Culture Quotient; CQ) ความสามารถในการเป็นผู้ประกอบการที่ดี (Entrepreneurial-Spirit Quotient; ES-Q) อุดมภูมิ ระดับความดัง ระดับความเข้มของแสง เป็นต้น ข้อมูลในมาตรวัดนี้สามารถบอกความแตกต่างระหว่างช่วงของตัวแปรได้ โดยแต่ละหน่วยมีระยะห่างเท่า ๆ กัน

1.4.4 มาตรอัตราส่วน (Ratio Scale)

เป็นมาตรวัดที่มีความละเอียดสูงที่สุดคือ วัดได้ทั้งขนาดและทิศทาง มีระยะห่างเท่ากัน ช่วงเท่ากัน และมีศูนย์แท้ (Absolute Zero) ศูนย์แปลว่าไม่มี สามารถเปรียบเทียบหรือคำนวณข้อมูลได้ เช่น รายได้ ค่าใช้จ่าย อายุ น้ำหนัก ส่วนสูง ปริมาณสินค้า เป็นต้น ข้อมูลในมาตรวัดนี้สามารถวัดออกมาเป็นตัวเลขเชิงปริมาณ บอกความแตกต่างของแต่ละช่วงได้ แต่ละหน่วยการวัดมีระยะห่างเท่า ๆ กัน มีศูนย์แท้ ค่าที่ได้สามารถนำมาบวก ลบ คูณ หาร ยกกำลัง และถอดรากได้ จึงสามารถนำมาวิเคราะห์โดยใช้สถิติได้ทุกตัว

เราเรียกข้อมูลในมาตรอันดับและมาตรอัตราส่วนหรือมาตรวัดที่ 3 และ 4 ว่าข้อมูลแบบต่อเนื่อง (Continuous Data) สถิติที่ใช้ในการวิเคราะห์ ได้แก่ บวก ลบ คูณ หาร ยกกำลัง ถอดรากได้ ค่าเฉลี่ย ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน สถิติเชิงอ้างอิง การวิเคราะห์ถดถอย การวิเคราะห์สหสัมพันธ์ เป็นต้น คือข้อมูล 2 มาตรวัดนี้สามารถวิเคราะห์ “สถิติพาราเมตริกได้ทุกตัว” สามารถศึกษาสถิติพาราเมตริกและสถิติอนพาราเมตริกเพิ่มเติมได้ในบทที่ 7

1.5 สถิติเชิงบรรยาย (Descriptive Statistics)

สถิติเชิงบรรยายสามารถนำเสนอข้อมูลได้หลายวิธี เช่น การแจกแจงความถี่ ร้อยละ สัดส่วน ในรูปแบบภูมิแท่ง แผนภูมิรูปภาพ กราฟ หรือตารางต่าง ๆ นอกจากนี้ยังสามารถนำเสนอค่าสถิติในลักษณะการวัดแนวโน้มเข้าสู่ส่วนกลาง ได้แก่ ค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) ค่าฐานนิยม (Mode) นำเสนอในรูปแบบสัดส่วน (Proportion) เปอร์เซ็นต์หรือร้อยละ (Percentage) การแจกแจงความถี่ (Frequency) ค่าต่ำสุด (Minimum) ค่าสูงสุด (Maximum) นำเสนอในรูปแบบการกระจายของข้อมูล เช่น พิสัย ความแปรปรวน (Variance; $V(x)$) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation; SD) สัมประสิทธิ์การกระจาย (Coefficient Variation; CV) ส่วนเบี่ยงเบนเฉลี่ย ส่วนเบี่ยงเบนควอร์ไทล์ การวัดรูปร่างของข้อมูล (Measure of Shape) เช่น ค่าความเบ้ (Sk) ค่าความโด่ง (Ku) และการหาตำแหน่ง ข้อมูล เช่น เปอร์เซ็นต์ไทล์ (Percentiles) ควอร์ไทล์ (Quartiles) เดไซล์ (Deciles) เป็นต้น โดยแต่ละวิธีอธิบายได้ดังนี้

1.5.1 การแจกแจงความถี่

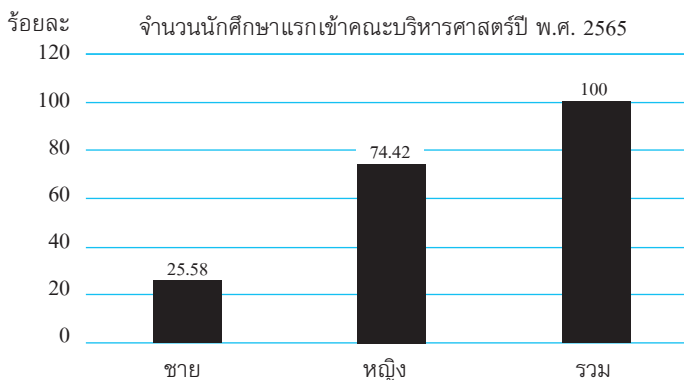
การแจกแจงความถี่เป็นการนำข้อมูลที่เป็นค่าของตัวแปรที่เราสนใจศึกษา มาจัดเรียงตามลำดับจากมากไปน้อย และแบ่งเป็นช่วงเท่าๆ กัน จำนวนข้อมูลในแต่ละช่วงเรียกว่า **ความถี่** ในกรณีที่มีความแตกต่างระหว่างคะแนนสูงสุดกับคะแนนต่ำสุดไม่มาก ไม่จำเป็นต้องแบ่งช่วงเป็นกลุ่ม การแจกแจงความถี่มีจุดมุ่งหมายเพื่อให้ทราบภาพรวมของการแจกแจงข้อมูลทั้งหมดอย่างเป็นระบบ การจัดระบบและนำเสนอข้อมูลในเบื้องต้น สามารถนำเสนอข้อมูลในรูปของตารางและแผนภูมิ ในที่นี้ขอนำเสนอเป็น 2 ส่วนคือ (1) การแจกแจงความถี่และร้อยละของข้อมูลที่น่าสนใจที่เป็นไปได้ทั้งหมด (2) การแจกแจงความถี่และร้อยละในแต่ละช่วงของข้อมูลที่น่าสนใจ

1. การแจกแจงความถี่และร้อยละของข้อมูลที่น่าสนใจที่เป็นไปได้ทั้งหมด การแจกแจงความถี่แบบนี้ใช้กับข้อมูลที่มีจำนวนลักษณะที่เป็นไปได้ทั้งหมดไม่มากนัก เช่น จำแนกนักศึกษาตามเพศชาย-หญิง จำแนกตามระดับการศึกษา จำแนกตามอาชีพหลัก เป็นต้น ดังตัวอย่างต่อไปนี้

ตัวอย่างจำนวนนักศึกษาที่จำแนกตามเพศ ที่ได้รับคัดเลือกเข้าศึกษาในคณะบริหารศาสตรมหาวิทยาลัยอุบลราชธานีในปีการศึกษา 2565 (ข้อมูลจากงานทะเบียน ม. อุบลราชธานี, 2565) โดยแจกแจงความถี่ ได้ดังนี้

เพศ	จำนวนนักศึกษา (คน)	ร้อยละ
ชาย	233	25.58
หญิง	678	74.42
รวม	911	100.00

เมื่อทำเป็นกราฟแท่งเป็นดังนี้



2. การแจกแจงความถี่และร้อยละในแต่ละช่วงของข้อมูลที่สนใจ การแจกแจงความถี่แบบนี้ใช้กับข้อมูลที่มีจำนวนลักษณะที่เป็นไปได้ทั้งหมดจำนวนมาก เช่น ศึกษารายได้ของคนไทยทั้งหมด หรืออายุของคนไทยทั้งหมด เป็นต้น ดังนั้นในการแจกแจงความถี่จึงควรแบ่งข้อมูลทั้งหมดออกเป็นช่วง ๆ ที่ต่อเนื่องกัน โดยแต่ละช่วงประกอบด้วยข้อมูลหลายๆ ค่า ทำให้ลดจำนวนค่าที่เป็นไปได้ทั้งหมดลง ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 1.1 สุ่มตัวอย่างพนักงานบริษัทวัสดุก่อสร้างในจังหวัดแห่งหนึ่งในภาคตะวันออกเฉียงเหนือมา 200 คน เพื่อสอบถามค่าใช้จ่ายรายวัน แล้วจัดเป็นช่วง ๆ ได้ 5 ช่วงชั้น ดังนี้

ค่าใช้จ่ายรายวัน (บาท)	จำนวนพนักงาน	ร้อยละ
0 – 100	10	5.00
101 – 150	36	18.00
151 – 200	84	42.00
201 – 250	54	27.00
251 – 300	16	8.00
รวม	200	100.00

การสร้างตารางแจกแจงความถี่สำหรับค่าในแต่ละช่วงของลักษณะที่สนใจมีขั้นตอนในการสร้างดังนี้

1.6 ขั้นตอนการสร้างตารางแจกแจงความถี่

ขั้นตอนการสร้างตารางแจกแจงความถี่ มีดังนี้

1. หาค่าพิสัยของข้อมูล (Range) ได้จาก พิสัย (Range; R) = ค่าสูงสุด – ค่าต่ำสุด
2. กำหนดจำนวนชั้น (k) $k = 3.322 + \log N$
3. คำนวณหาความกว้างของชั้น (Class Interval; CI)

$$CI = \text{ความกว้างของชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}} = \frac{R}{K}$$
4. คำนวณหาขีดจำกัด (Class Limit)
 ขีดจำกัดล่างของชั้นแรก = ค่าต่ำสุด $\frac{(CI \times k) - R}{2}$
 หรือใช้ค่าต่ำสุดเป็นขีดจำกัดบนของชั้นต่ำสุดของการแจกแจง
5. คำนวณหาจุดกึ่งกลางของแต่ละชั้น (Mid Point)
 จุดกึ่งกลางชั้น = (ขีดจำกัดบน + ขีดจำกัดล่าง)/2
6. คำนวณหาขีดจำกัดชั้นที่แท้จริง (Class Boundaries)
 ขีดจำกัดชั้นที่แท้จริง = (ขีดจำกัดบนของชั้น + ขีดจำกัดล่างของชั้นถัดไป)/2
7. นับจำนวนค่าของข้อมูล (ความถี่) ในแต่ละชั้น

เมื่อได้จำนวนแล้วสามารถหาความถี่สะสม ความถี่สัมพัทธ์ และร้อยละโดยการหาความถี่สะสม จะเริ่มหาผลบวกของความถี่ที่เริ่มจากชั้นแรกบวกไปเรื่อย ๆ เมื่อถึงชั้นนั้น ๆ การหาความถี่สัมพัทธ์ หรือสัดส่วน (Proportion) ของชั้นใดก็นำความถี่ของชั้นนั้นหารด้วยความถี่ทั้งหมดและคูณด้วยร้อยละ เรียกว่า เปอร์เซนต์หรือร้อยละ (Percentage)

การหาความถี่สัมพัทธ์หรือสัดส่วน $P(E) = \frac{\text{ความถี่ของชั้นนั้น}}{\text{ความถี่ทั้งหมด}} = \frac{n}{N}$ เมื่อคูณสัดส่วนด้วยร้อยละ ($\times 100$) จะเรียกว่า เปอร์เซนต์หรือร้อยละ

ตัวอย่างที่ 1.2

จากข้อมูลเป็นจำนวนคนที่มาใช้บริการธนาคารพาณิชย์ไทยแห่งหนึ่งใน อำเภวารินชำราบ จังหวัดอุบลราชธานี จำนวน 20 วัน เป็นดังนี้ 84 61 80 65 72 80 85 64 90 79 76 65 73 70 85 75 62 89 62 93 คน จงสร้างตารางแจกแจงความถี่ที่เหมาะสม พร้อมทั้งหาร้อยขีด ความถี่ ความถี่สัมพัทธ์ และร้อยละความถี่สัมพัทธ์ จากข้อมูลดังกล่าว

วิธีทำ

1. หาพิสัย (R) = Max – Min = 93 – 62 = 31

2. กำหนดจำนวนอันตรภาคชั้น ในที่นี้ $N = 20$

$$\begin{aligned}\text{สูตรหาจำนวนชั้น } k &= 1 + 3.322 \log N \\ &= 1 + 3.322 \log_{10} 20 \\ &= 1 + 3.322 (1.3010) = 5.3219 \approx 5\end{aligned}$$

3. กำหนดขนาดหรือความกว้างของอันตรภาคชั้น (Class Interval; CI)

$$CI = \text{ความกว้างของชั้น} = \frac{\text{พิสัย}}{\text{จำนวนชั้น}} = \frac{31}{5} = 6.2 \approx 6$$

4. จัดช่วงชั้นของจำนวนลูกค้าในแต่ละวันและแจกแจงความถี่ในแต่ละชั้น

อันตรภาคชั้น	รอยขีด	ความถี่	ขอบล่าง	ขอบบน	ขีดจำกัดล่าง	ขีดจำกัดบน	จุดกึ่งกลางชั้น
61 – 67		6	60.5	67.5	61	67	64
68 – 74		3	67.5	74.5	68	74	71
75 – 81		5	74.5	81.5	75	81	78
82 – 88		3	81.5	88.5	82	88	85
89 – 95		3	88.5	95.5	89	95	92
		N = 20					

อันตรภาคชั้น	รอยขีด	ความถี่	ความถี่สะสม	ความถี่สัมพัทธ์	ร้อยละความถี่สัมพัทธ์
61 – 67		6	6	0.3	0.3
68 – 74		3	9	0.15	0.45
75 – 81		5	14	0.25	0.70
82 – 88		3	17	0.15	0.85
89 – 95		3	20	0.15	1.00
รวม		20		1.00	

คำศัพท์ที่สำคัญ

1. อันตรภาคชั้น (Class Interval) คือแต่ละชั้นของข้อมูล เช่น จากตารางมี 5 อันตรภาคชั้น จำนวนอันตรภาคชั้นโดยทั่วไปนิยมใช้ระหว่าง 5–20 ชั้น ถ้าจำนวนอันตรภาคชั้นน้อยเกินไป รายละเอียดของข้อมูลจะสูญเสียไป แต่ถ้าจำนวนอันตรภาคชั้นมากเกินไป ข้อมูลจะขาดความกระชับ
2. รอยขีดคือ ค่าจากการสังเกตในแต่ละอันตรภาคชั้น (ไม่มีก็ได้)
3. ความถี่ (Frequency) ใช้สัญลักษณ์ f คือจำนวนครั้งที่ค่าจากการสังเกตปรากฏในข้อมูล (ผลรวมของรอยขีด) ถ้าไม่มีข้อมูล ห้ามใส่เลข “0” ลงไป ให้ใช้เครื่องหมาย – แทน
4. ขอบล่าง (Lower Boundary) คือค่ากึ่งกลางระหว่างค่าที่น้อยที่สุดของชั้นนั้นกับค่าที่มากที่สุดของชั้นรองลงไป เช่น ขอบล่างของอันตรภาคชั้น 75–81 คือ $\frac{75 + 74}{2} = 74.5$
5. ขอบบน (Upper Boundary) คือค่ากึ่งกลางระหว่างค่าที่มากที่สุดของชั้นนั้นกับค่าที่น้อยที่สุดของชั้นถัดไป เช่น ขอบบนของอันตรภาคชั้น 75–81 คือ $\frac{81 + 82}{2} = 81.5$

หมายเหตุ : เพื่อความสะดวกและรวดเร็ว การหาขอบบนและขอบล่าง

- ถ้าข้อมูลเป็นจำนวนเต็ม เช่น อันตรภาคชั้น 68–74 ขอบล่างจะลดลงจากค่าต่ำสุด 0.5 คือ 67.5 ขอบบนจะเพิ่มจากค่าสูงสุด 0.5 คือ 74.5
- ถ้าข้อมูลเป็นทศนิยม เช่น อันตรภาคชั้น 2.0–4.9 ขอบล่างและขอบบนจะลดและเพิ่ม 0.05 จะได้ ขอบล่างคือ 1.95 ขอบบนคือ 4.95

6. ขีดจำกัดบน (Upper Limit) และขีดจำกัดล่าง (Lower Limit) คือค่าที่สูงที่สุดและค่าต่ำที่สุดของอันตรภาคชั้นนั้นๆ เช่น อันตรภาคชั้น 68–74 ขีดจำกัดล่างคือ 68 และขีดจำกัดบนคือ 74 (* ในการคำนวณส่วนมาก จะใช้ขอบบนและขอบล่างเป็นส่วนใหญ่)

7. ความกว้างของอันตรภาคชั้น (CI) คือผลต่างของขอบบนกับขอบล่างของอันตรภาคชั้นนั้น เช่น อันตรภาคชั้น 82–88 คือ $88.5 - 81.5 = 7$ ดังนั้น $CI = 7$ ขนาดความกว้างของอันตรภาคอาจจะเท่ากันทุกชั้นหรือไม่เท่ากันก็ได้ ทั้งนี้ขึ้นอยู่กับลักษณะของข้อมูลแต่โดยส่วนมากจะมีขนาดของอันตรภาคชั้นเท่ากันทุกชั้น
8. อันตรภาคชั้นปลายเปิดควรนำมาใช้ในกรณีที่มีข้อมูลสุดโต่ง (Extreme Values) เพื่อช่วยลดความเสี่ยงความถี่เป็นศูนย์ แต่มีข้อเสียคือไม่ทราบความกว้างของอันตรภาคชั้นนั้น
9. จุดกึ่งกลางชั้น (Mid-point) คือค่าเฉลี่ยระหว่างขอบบนและขอบล่างของชั้นนั้น ๆ เช่น จุดกึ่งกลางของขนาดอันตรภาคชั้น 82–88 คือ $\frac{81.5 + 88.5}{2} = 85$ หรือเป็นค่าเฉลี่ยระหว่างขีดจำกัดบน และขีดจำกัดล่างของชั้นนั้น ๆ คือ $\frac{82 + 88}{2} = 85$
10. ความถี่สัมพัทธ์ (Relative Frequency) คืออัตราส่วนระหว่างความถี่ของค่านั้น หรือของอันตรภาคชั้นนั้นเทียบกับผลรวมของความถี่ทั้งหมด $\times$ ความถี่สัมพัทธ์อาจแสดงในรูปเศษส่วนหรือทศนิยมหรือร้อยละก็ได้ ความถี่สะสมสัมพัทธ์ (Relative Cumulative Frequency) คืออัตราส่วนระหว่างความถี่สะสมของค่านั้น หรืออันตรภาคชั้นนั้นเทียบกับผลรวมของความถี่ทั้งหมด $\times$ ความถี่สะสมสัมพัทธ์อาจแสดงในรูปเศษส่วน หรือทศนิยมหรือร้อยละก็ได้

ตัวอย่างที่ 1.3

จงสร้างตารางแจกแจงความถี่ที่มี 6 อันตรภาคชั้น พร้อมทั้งหาร้อยขีด ความถี่ ความถี่สัมพัทธ์ และร้อยละความถี่สัมพัทธ์ จากข้อมูลจำนวนลูกค้าร้านหมูกระทะแห่งหนึ่งหน้ามหาวิทยาลัยขอนแก่น ในแต่ละวันมีจำนวนลูกค้าเป็นดังนี้

61	32	18	23	19	97	30	54	46	25
28	21	15	40	84	50	26	71	22	58
41	64	39	60	28	44	31	55	61	73
56	92	83	28	33	51	87	59	79	62

และตอบคำถามต่อไปนี้



กิจกรรม
ในห้องเรียน

1. ค่าพิสัยของข้อมูลคือ
2. ความกว้างของแต่ละชั้นคือ
3. ชั้นที่มีความถี่มากที่สุดมีความถี่เท่ากับ
4. ชั้นที่มีความถี่น้อยที่สุดมีความถี่เท่ากับ
5. ข้อมูลที่มีความถี่มากที่สุด คิดเป็นร้อยละเท่าไรของข้อมูลทั้งหมด

วิธีทำ

ขีดจำกัด	รอยขีด	ความถี่	ความถี่สัมพัทธ์ หรือสัดส่วน	ร้อยละ ความถี่สัมพัทธ์
รวม				

1.7 การวัดแนวโน้มเข้าสู่ส่วนกลาง
(Measures of Central Tendency)

การวัดแนวโน้มเข้าสู่ส่วนกลาง เป็นการคำนวณหาค่ากลางของข้อมูลว่าอยู่ตรงตำแหน่งใด การศึกษาในกรณีนี้ต้องการหาค่าเพียงค่าเดียวเพื่อใช้อธิบายข้อมูลทั้งหมด จึงนิยามหาค่ากลาง ๆ ที่เป็นตัวแทนของข้อมูลทั้งหมด ประกอบด้วย ค่าเฉลี่ยเลขคณิต ค่าเฉลี่ยเรขาคณิต ค่าเฉลี่ยฮาร์โมนิก มัธยฐาน และฐานนิยม โดยมีวิธีการคำนวณหาค่าต่าง ๆ ได้ดังนี้

1.7.1 ค่าเฉลี่ยเลขคณิตหรือมัชฌิมเลขคณิต (Mean)

ค่าเฉลี่ยเลขคณิตหรือมัชฌิมเลขคณิตคือ ค่าที่ได้จากผลรวมทั้งหมดหารด้วยจำนวนคนหรือจำนวนข้อมูล การคำนวณหาค่าแบ่งเป็นการหาค่าเฉลี่ยเลขคณิตสำหรับข้อมูลที่ไม่ได้จัดเป็นกลุ่ม และการหาค่าเฉลี่ยเลขคณิตสำหรับข้อมูลที่จัดเป็นกลุ่ม ในบทนี้ขอนำเสนอเฉพาะการหาค่าเฉลี่ยเลขคณิตสำหรับข้อมูลที่ไม่ได้จัดเป็นกลุ่มดังนี้

ค่าเฉลี่ยเลขคณิตประชากร $\mu = \frac{\sum_{i=1}^n x_i}{N}$ ค่าเฉลี่ยเลขคณิตตัวอย่าง $\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$

1.7.2 ค่าเฉลี่ยเรขาคณิต (Geometric Mean; GM)

ค่าเฉลี่ยเรขาคณิตคือ ค่ากลางของข้อมูลที่เกิดจากการถอดรากที่ N ของผลคูณของข้อมูล N ค่า ในกรณีทราบค่า N ของประชากร แต่ในกรณีของตัวอย่างค่าเฉลี่ยเรขาคณิตก็จะเป็นค่ากลางของข้อมูลที่เกิดจากการถอดรากที่ n ของผลคูณของข้อมูล n ค่า ในกรณีทราบค่า n ของตัวอย่าง

■ ข้อมูลที่ไม่ได้จัดกลุ่ม (Ungrouped Data)

$$\text{GM ของประชากร GM} = \sqrt[N]{x_1 \cdot x_2 \cdot x_3 \dots x_N} \quad \text{เมื่อ } x_1, x_2, x_3, \dots, x_N$$

$$\text{GM ของตัวอย่าง GM} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \dots x_n} \quad \text{เมื่อ } x_1, x_2, x_3, \dots, x_n$$

เมื่อ N และ n คือ จำนวนข้อมูลที่เป็นค่าบวก

1.7.3 ค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean; HM)

ใช้ในกรณีที่ต้องการหาค่ากลางของข้อมูลที่มีขอบเขตจำกัด นิยมใช้สำหรับการหาอัตราเฉลี่ย หรือใช้สำหรับเฉลี่ยข้อมูล เช่น ระยะทางต่อชั่วโมง งานต่อหน่วยเวลา เป็นต้น ใช้กับข้อมูลที่ไม่ได้จัดกลุ่มหรือไม่ได้แจกแจงความถี่ (Ungrouped Data)

$$\text{HM ของประชากร HM} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}} \quad \text{เมื่อ } x_i = x_1, x_2, x_3, \dots, x_N$$

$$\text{HM ของตัวอย่าง HM} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} \quad \text{เมื่อ } x_i = x_1, x_2, x_3, \dots, x_n$$

ตัวอย่างที่ 1.4

จากข้อมูลตัวอย่างลูกค้าที่มาใช้บริการร้านป๊ายไวโนลในอำเภอวารินชำราบ 5 วัน มียอดลูกค้าเป็นดังต่อไปนี้ 2 4 8 16 32 คน จงหาค่าเฉลี่ยเลขคณิต ค่าเฉลี่ยแบบเรขาคณิต และ ค่าเฉลี่ยแบบฮาร์โมนิก จำนวนลูกค้าที่มาใช้บริการร้านป๊ายไวโนลในอำเภอแห่งนี้

วิธีทำ จากสูตรค่าเฉลี่ยเลขคณิต $\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{2 + 4 + 8 + 16 + 32}{5} = 12.4 \text{ คน}$

จากสูตรค่าเฉลี่ยแบบเรขาคณิต $\text{GM} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5}$
 $= \sqrt[5]{2 \cdot 4 \cdot 8 \cdot 16 \cdot 32} = \sqrt[5]{2^{15}} = 2^3 = 8 \text{ คน}$

จากสูตรค่าเฉลี่ยแบบฮาร์โมนิก $\text{HM} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$
 $= \frac{5}{\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32}}$
 $= 5.1613$
 $\approx 5 \text{ คน}$

ตอบ ค่าเฉลี่ยเลขคณิตคือ ลูกค้า 12 คน

ค่าเฉลี่ยแบบเรขาคณิตคือ ลูกค้า 8 คน

ค่าเฉลี่ยแบบฮาร์โมนิกคือ ลูกค้า 5 คน

ข้อดีของค่าเฉลี่ย

- การคำนวณหาไม่ยุ่งยากและสามารถใช้เครื่องคิดเลขช่วยคำนวณได้
- เป็นที่แพร่หลายและส่วนใหญ่ใช้เป็นค่ากลางของข้อมูลโดยเฉพาะค่าเฉลี่ยเลขคณิต
- ค่าเฉลี่ยเลขคณิตสามารถใช้ข้อมูลทุกตัวในการคำนวณได้

ข้อจำกัดของค่าเฉลี่ย

- ใช้ได้เฉพาะในกรณีที่ข้อมูลเป็นข้อมูลเชิงปริมาณเท่านั้น
- ค่าที่คำนวณได้ไม่จำเป็นต้องเป็นค่าของข้อมูลตัวใดตัวหนึ่งเสมอไป
- ถ้ามีข้อมูลในชุดที่แตกต่างจากตัวอื่นมาก ๆ (Outliers) จะมีผลต่อค่าเฉลี่ยเลขคณิตของข้อมูลชุดนี้

1.7.4 ค่ามัธยฐาน (Median)

ค่ามัธยฐานคือ ค่ากลางของข้อมูล หรือค่าที่แบ่งข้อมูลออกเป็น 2 ส่วนเท่า ๆ กันคือ มากกว่ามัธยฐาน 50% น้อยกว่ามัธยฐาน 50% หรือคือค่าในตำแหน่งกึ่งกลางของการแจกแจงของข้อมูล

ดังนั้นค่ามัธยฐานคือ ค่าของข้อมูล ณ ตำแหน่งที่ $\frac{n+1}{2}$ เมื่อเรียงลำดับข้อมูลจากน้อยไปมากหรือมากไปน้อยแล้ว

ในกรณีที่จำนวนข้อมูลเป็นเลขคี่ มัธยฐานคือค่าของข้อมูลที่อยู่กึ่งกลางพอดี แต่ถ้าจำนวนข้อมูลเป็นเลขคู่ ค่ามัธยฐานจะเท่ากับค่าเฉลี่ยของ 2 จำนวนตรงกลางบวกกันแล้วหาร 2 ตัวอย่างเช่น

10 13 15 16 18 19 20 ตำแหน่งมัธยฐานคือ ตำแหน่งที่ $\frac{7+1}{2} = 4^{\text{th}}$
 $\therefore$ ค่ามัธยฐานคือ 16

11 12 12 13 15 15 18 19 ตำแหน่งมัธยฐานคือ ตำแหน่งที่ $\frac{8+1}{2} = 4.5^{\text{th}}$
 $\therefore$ ค่ามัธยฐานคือ $\frac{13+15}{2} = 14$

ข้อดีของค่ามัธยฐาน

- ค่ามัธยฐานจะไม่ได้รับอิทธิพลเมื่อมีข้อมูลที่มีค่าสูงหรือต่ำผิดปกติ
- หาค่ามัธยฐานจากการนำข้อมูลทั้งหมดมาจัดลำดับจากน้อยไปมากหรือกลับกันได้ง่าย
- จะเป็นค่าของข้อมูล ถ้ามีข้อมูลเป็นจำนวนคี่

ข้อจำกัดของค่ามัธยฐาน

- ไม่ได้นำทุกค่าของข้อมูลมาคิดคำนวณ ทำให้ค่าที่คำนวณได้มีความคลาดเคลื่อน
- ใช้ได้ในกรณีที่ข้อมูลเป็นข้อมูลเชิงปริมาณเท่านั้น
- ถ้ามีข้อมูลเป็นจำนวนมาก การจัดเรียงข้อมูลจะทำได้ค่อนข้างลำบาก
- จะไม่ใช่ค่าที่แท้จริงของข้อมูล ถ้าจำนวนข้อมูลเป็นจำนวนคู่

1.7.5 ฐานนิยม (Mode)

ฐานนิยมคือ ค่าที่เกิดขึ้นมากที่สุดหรือค่าที่มีความถี่สูงสุดในชุดข้อมูลทั้งหมด สำหรับข้อมูลที่ไม่ได้จัดกลุ่มค่าฐานนิยมก็คือค่าที่มีความถี่ของค่านั้นซ้ำกันมากที่สุด

ข้อดีของค่าฐานนิยม

- ค่าฐานนิยมจะไม่ได้รับอิทธิพลเมื่อมีข้อมูลที่มีค่าสูงหรือต่ำผิดปกติ
- เป็นค่ากลางที่ใช้วัดข้อมูลเชิงคุณภาพและเชิงปริมาณ

ข้อจำกัดของค่าฐานนิยม

- ในกรณีที่ไม่มีค่าของข้อมูลที่ซ้ำกัน จะไม่มีค่าฐานนิยม
- กรณีที่ข้อมูลจัดกลุ่ม ค่าฐานนิยมจะเปลี่ยนไปถ้าจำนวนชั้นเปลี่ยนไป
- ข้อมูลบางชุดอาจมีค่าฐานนิยมมากกว่า 1 ค่า โดยที่ค่าฐานนิยมนั้นอาจแตกต่างกันมาก

ตัวอย่างที่ 1.5

สุ่มพนักงานขายสินค้าของบริษัทขายเครื่องสำอางแห่งหนึ่งมา 6 คน สอบถามจำนวนสินค้าที่ขายได้ทั้งเดือน ปรากฏดังนี้ 50 55 45 60 65 และ 60 ชิ้น ตามลำดับ จงหาค่าเฉลี่ยเลขคณิต ค่ามัธยฐาน ค่าฐานนิยม ค่าเฉลี่ยเรขาคณิต ค่าเฉลี่ยฮาร์โมนิก และหาสัดส่วนของพนักงานที่ขายเครื่องสำอางได้มากกว่า 60 ชิ้น

วิธีทำ

■ ค่าเฉลี่ย $\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{50 + 55 + 45 + 60 + 65 + 60}{6} = \frac{335}{6} = 55.8333$

ตอบ ค่าเฉลี่ย = 55.8333 ชิ้น

■ ค่าฐานนิยม คือ 60 ชิ้น

■ ค่ามัธยฐาน คือ ค่ากลางของข้อมูล หาได้ดังนี้

1. เรียงอันดับข้อมูลจากน้อยไปมาก 45 50 55 60 60 65

2. หาตำแหน่งค่ามัธยฐาน $\text{Median}^{\text{th}} = \frac{(n+1)^{\text{th}}}{2} = \frac{(6+1)^{\text{th}}}{2} = 3.5^{\text{th}}$

∴ ตำแหน่งที่ 3 คือจำนวนสินค้า 50 ชิ้น และตำแหน่งที่ 4 คือจำนวนสินค้า 60 ชิ้น

∴ $\text{Median} = \frac{(55 + 60)}{2} = 57.5$

ตอบ ค่ามัธยฐานของจำนวนสินค้าที่พนักงาน 6 คน ขายได้คือ 57.5 ชิ้น

■ ค่าเฉลี่ยเรขาคณิต $GM = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \dots x_n}$

$$= \sqrt[6]{45 \cdot 50 \cdot 55 \cdot 60 \cdot 60 \cdot 65}$$

$$= \sqrt[6]{28,957,500,000}$$

$$= 55.4149$$

ตอบ ค่าเฉลี่ยเรขาคณิต (GM) = 55.4149 ชิ้น

$$\begin{aligned}
 \blacksquare \text{ ค่าเฉลี่ยฮาร์โมนิก HM} &= \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_6}} \\
 &= \frac{6}{\frac{1}{45} + \frac{1}{50} + \frac{1}{55} + \frac{1}{60} + \frac{1}{60} + \frac{1}{65}} \\
 &= \frac{6}{0.1088} \\
 &= 55.1471
 \end{aligned}$$

ตอบ ค่าเฉลี่ยฮาร์โมนิก (HM) = 55.1471 ขึ้น

$$\begin{aligned}
 \blacksquare \text{ สัดส่วนของพนักงานที่ขายสินค้าได้มากกว่า 60 ชิ้น มี 1 คน จาก 6 คน} \\
 = P(E) = \frac{n}{N} = \frac{1}{6} = 0.1667 \text{ หรือ } 16.67\%
 \end{aligned}$$

ตอบ สัดส่วนของพนักงานที่ขายเครื่องสำอางได้มากกว่า 60 ชิ้น = 0.1667 หรือ 16.67%

1.8 การวัดการกระจายของข้อมูล (Measure of Distribution)

การวัดการกระจายของข้อมูล เป็นสถิติประเภทหนึ่งที่คำนวณออกมาเป็นตัวเลข เพื่อใช้อธิบายลักษณะการกระจายของข้อมูล การที่ข้อมูลชุดหนึ่ง ๆ ประกอบด้วยคะแนนที่มีค่าต่าง ๆ กันเราเรียกว่าเป็นข้อมูลที่มีการกระจาย ถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าต่างกันอย่างน้อยเรียกว่า เป็นข้อมูลที่มีการกระจายมาก แต่ถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าต่างกันอย่างน้อยเรียกว่า เป็นข้อมูลที่มีการกระจายน้อย และถ้าข้อมูลชุดนั้นประกอบด้วยคะแนนที่มีค่าเท่ากันหมดเรียกว่า เป็นข้อมูลที่ไม่มีการกระจาย ค่าที่ใช้การวัดการกระจายของข้อมูล ได้แก่

1.8.1 พิสัย (Range; R)

พิสัย หมายถึงการหาการกระจายของข้อมูลโดยนำข้อมูลที่มีค่าสูงที่สุดลบกับข้อมูลที่มีค่าต่ำที่สุด เพื่อให้ได้ค่าที่เป็นช่วงของการกระจาย ซึ่งสามารถบอกถึงความกว้างของข้อมูลชุดนั้น ๆ สำหรับสูตรที่ใช้ในการหาพิสัยคือ

$$\begin{aligned}
 \text{พิสัย (R)} &= \text{ค่าสูงสุด} - \text{ค่าต่ำสุด} \\
 &= X_{\max} - X_{\min} \text{ หรือ } = \text{Max} - \text{Min}
 \end{aligned}$$

1.8.2 ความแปรปรวน (Variance; V(x))

ความแปรปรวน คือค่าเฉลี่ยของกำลังสองของผลต่างระหว่างค่าของข้อมูลแต่ละค่ากับค่าเฉลี่ยหรือความแปรปรวนก็คือ กำลังสองของส่วนเบี่ยงเบนมาตรฐาน นั่นเอง

$$\text{ความแปรปรวนของประชากร } (\sigma)^2 = \frac{\sum (x_i - \mu)^2}{N}$$

$$\text{ความแปรปรวนของตัวอย่าง } S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

1.8.3 ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation; SD)

ส่วนเบี่ยงเบนมาตรฐาน คือตัวเลขหรือค่าที่แสดงว่าข้อมูลที่มีอยู่ทุกจำนวนนั้นโดยเฉลี่ยแล้ว มีค่าต่างจากค่าเฉลี่ยเท่าใด

$$\text{ส่วนเบี่ยงเบนมาตรฐานของประชากร } \sigma = \sqrt{\sigma^2}$$

$$\text{ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง } S = \sqrt{S^2}$$

1.8.4 สัมประสิทธิ์การแปรผัน (Coefficient of Variation; CV)

สัมประสิทธิ์การแปรผัน คือค่าที่ใช้วัดการกระจายของข้อมูลที่ไม่มีหน่วย ซึ่งต่างจากค่าสถิติตัวอื่นที่ใช้วัดการกระจายข้อมูล ซึ่งมีหน่วยเป็นหน่วยเดียวกับหน่วยของข้อมูล หรือหน่วยข้อมูล ยกกำลังสอง ค่าสัมประสิทธิ์ความแปรผันคือ ส่วนเบี่ยงเบนมาตรฐานหารด้วยค่าเฉลี่ย เราสามารถเปรียบเทียบการกระจายของข้อมูลเหล่านั้นได้ โดยการแปลงให้เป็นอัตราส่วนหรือร้อยละของสัดส่วนของค่าการกระจายกับค่ากลาง (Mason et al., 1991)

$$\text{สัมประสิทธิ์การแปรผันของประชากร } CV = \frac{\sigma}{\mu} \times 100$$

$$\text{สัมประสิทธิ์การแปรผันของตัวอย่าง } CV = \frac{s}{\bar{x}} \times 100$$

1.8.5 ส่วนเบี่ยงเบนควอร์ไทล์ (Quartile Deviation; QD)

ส่วนเบี่ยงเบนควอร์ไทล์ คือค่าที่ใช้วัดการกระจายของข้อมูลโดยพิจารณาจากครึ่งหนึ่งของระยะจากควอร์ไทล์ที่ 3 (Q_3) ถึงควอร์ไทล์ที่ 1 (Q_1) หรือครึ่งหนึ่งของความแตกต่างระหว่างควอร์ไทล์ที่ 3 (Q_3) และควอร์ไทล์ที่ 1 (Q_1) ของคะแนนข้อมูลชุดหนึ่ง ๆ เป็นการวัดการกระจายเพื่อวัดแนวโน้มเข้าสู่ส่วนกลางด้วยค่ามัธยฐาน

$$\text{ส่วนเบี่ยงเบนควอร์ไทล์ } QD = \frac{|Q_3 - Q_1|}{2}$$

1.8.6 ส่วนเบี่ยงเบนเฉลี่ย (Average Deviation; AD or Mean Deviation; MD)

ส่วนเบี่ยงเบนเฉลี่ย คือค่าเฉลี่ยเลขคณิตของค่าสัมบูรณ์ของผลต่างระหว่างข้อมูลแต่ละตัวกับค่าเฉลี่ยเลขคณิตของข้อมูลชุดนั้น

$$\text{ส่วนเบี่ยงเบนเฉลี่ยของประชากร } MD \text{ หรือ } AD = \frac{\sum_{i=1}^N |x_i - \mu|}{N}$$

$$\text{ส่วนเบี่ยงเบนเฉลี่ยของตัวอย่าง } MD \text{ หรือ } AD = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

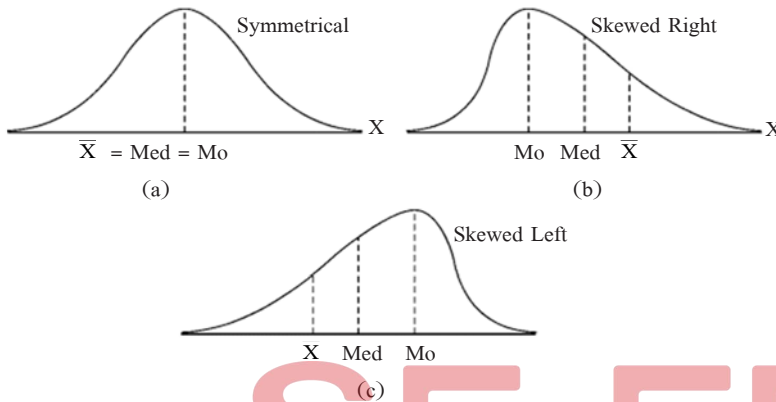
1.9 การวัดรูปร่างของข้อมูล (Measure of Shape)

การวัดรูปร่างของข้อมูล มีดังนี้

1. การวัดความเบ้ (Skewness Measurement; S_k) ความเบ้คือ ค่าที่ใช้การวัดรูปร่างการกระจายหรือการแจกแจงของข้อมูลที่บอกให้ทราบว่าข้อมูลมีรูปการแจกแจงอยู่ในรูปของเส้นโค้งปกติหรือไม่ หรือข้อมูลมีการเบ้ไปทางใด สูตรการหาเป็นดังนี้

$$\text{ความเบ้ของประชากร } S_k = \frac{3(\mu - \text{Median})}{\sigma} \quad \text{หรือ} \quad S_k = \frac{3(\mu - \text{Mode})}{\sigma} \quad (\text{Black, 2011})$$

$$\text{ความเบ้ของตัวอย่าง } S_k = \frac{3(\bar{x} - \text{Median})}{s} \quad \text{หรือ} \quad S_k = \frac{3(\bar{x} - \text{Mode})}{s}$$



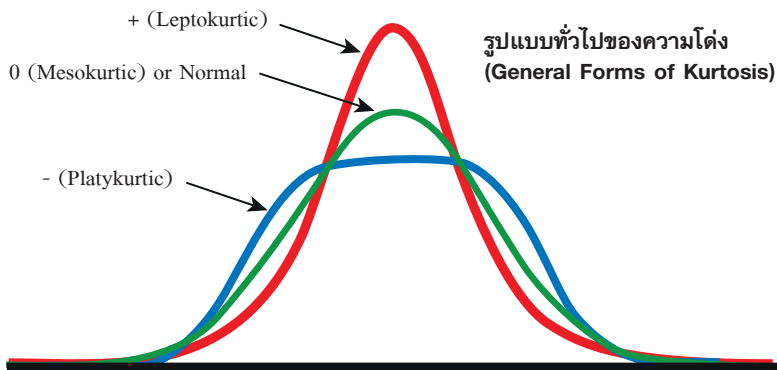
การแปลผลความเบ้ (Gravetter & Wallnau, 2014; Trochim & Donnelly, 2006) ทำได้ดังนี้

ถ้า $S_k = 0$ คือการแจกแจงของข้อมูลอยู่ในรูปของเส้นโค้งปกติหรือสมมาตร

ถ้า $S_k > 2$ คือการแจกแจงของข้อมูลอยู่ในรูปของเส้นโค้งเบ้ขวา

ถ้า $S_k < -2$ คือการแจกแจงของข้อมูลอยู่ในรูปของเส้นโค้งเบ้ซ้าย

2. การวัดความโด่ง (Kurtosis Measurement; K_u) การวัดความโด่งของการแจกแจงของข้อมูลที่อยู่ในรูปโค้งปกติสังเกตได้จากรูปการแจกแจงของข้อมูลดังนี้



กราเวตเตอร์ เอฟ. และวอลล์นัว แอล. (Gravetter, F. & Wallnau, L., 2014) และโทรchim ดับเบิลยู. เอ็ม. และดอนเนลลี เจ. พี. (Trochim, W. M. & Donnelly, J. P., 2006) ได้ให้เกณฑ์การแจกแจงของข้อมูลที่อยู่ในรูปโค้งปกติที่เป็นค่าความโด่งดังนี้

Leptokurtic คือการแจกแจงที่สูงโด่งเกินกว่าปกติ $K_u > 2$

Mesokurtic คือการแจกแจงที่สูงโด่งปกติ $K_u = 0$

Platykurtic คือการแจกแจงที่สูงโด่งต่ำกว่าปกติ $K_u < -2$

สูตรการวัดความโด่ง (EasyCalculation.com, 2016) คือ

$$\text{ประชากร } K_u = \frac{\sum (x_i - \mu)^4}{(N - 1)\sigma^4} \quad \text{ตัวอย่าง } K_u = \frac{\sum (x_i - \bar{X})^2}{(n - 1)s^4}$$

$$\text{หรือ } K_u = \frac{QD}{P_{90} - P_{10}} \quad \text{เมื่อ } QD = \frac{Q_3 - Q_1}{2}$$

ตัวอย่างที่ 1.6

สมุลูกจ้างในร้านอาหารแห่งหนึ่งมา 8 คน เพื่อสอบถามค่าใช้จ่ายรายวันเป็นดังนี้ 106 109 114 116 121 122 125 129 บาท จงหาค่าเฉลี่ยเลขคณิต พิสัย ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์การแปรผัน ส่วนเบี่ยงเบนเฉลี่ย ความโด่ง และค่าความเบ้

วิธีทำ

■ ค่าเฉลี่ยเลขคณิต

$$\begin{aligned} \bar{X} &= \frac{\sum_{i=1}^8 x_i}{8} \\ &= \frac{106 + 109 + 114 + 116 + 121 + 122 + 125 + 129}{8} \\ &= \frac{942}{8} = 117.75 \end{aligned}$$

ตอบ ค่าเฉลี่ยเลขคณิต (Mean) = 117.75 บาท

■ พิสัย

$$= \text{Max} - \text{Min} = 129 - 106 = 23$$

ตอบ

■ ความแปรปรวน

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

$$\begin{aligned}\text{วิธีที่ 1 } S^2 &= \frac{\sum (x_i - \bar{x})^2}{n - 1} \\ &= \frac{(106 - 117.75)^2 + (109 - 117.75)^2 + \dots + (129 - 117.75)^2}{8 - 1} \\ &= 62.7857\end{aligned}$$

ตอบ

วิธีที่ 2 สามารถคำนวณความแปรปรวนโดยใช้ตารางคำนวณดังนี้

x_i	$\bar{x}$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
106	117.75	- 11.75	138.0625
109	117.75	- 8.75	76.5625
114	117.75	- 3.75	14.0625
116	117.75	- 1.75	3.0625
121	117.75	3.25	10.5625
122	117.75	4.25	18.0625
125	117.75	7.25	52.5625
129	117.75	11.25	126.5625
$\sum (x_i - \bar{x})^2$			439.5

$$S^2 = \frac{439.5}{7} = 62.7857$$

ตอบ

■ ส่วนเบี่ยงเบนมาตรฐาน (S) = $\sqrt{S^2} = \sqrt{62.7857} = 7.9237$

ตอบ

■ ค่าสัมประสิทธิ์การแปรผัน (CV)

$$CV = \frac{S}{\bar{x}} \times 100$$

$$= \frac{7.9237}{117.75} \times 100$$

$$= 6.7293\%$$

ตอบ สัมประสิทธิ์การแปรผัน (CV) = 6.7293%

■ ส่วนเบี่ยงเบนเฉลี่ยของตัวอย่าง $AD = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$

$$\begin{aligned}AD &= \frac{|106 - 117.75| + |109 - 117.75| + |114 - 117.75| + \dots + |129 - 117.75|}{8} \\ &= \frac{52}{8} = 6.5\end{aligned}$$

ตอบ ส่วนเบี่ยงเบนเฉลี่ยของตัวอย่าง AD = 6.5

■ ความโด่ง (Kurtosis, Ku)

$$\text{สูตร } Ku = \frac{\sum (x_i - \bar{x})^4}{(n-1)s^4}$$

$(x_i - \bar{X})^2$	$(x_i - \bar{X})^4$
138.0625	19,061.250000
76.5625	5,861.816000
14.0625	197.753900
3.0625	9.378906
10.5625	111.566400
18.0625	326.253900
52.5625	2,762.816000
126.5625	16,018.070000
รวม	44,348.910000

จาก

$$K_u = \frac{\sum (x_i - \bar{X})^4}{(n-1)s^4} = \frac{44,348.91}{7 \times 3,942.0441} = 1.6072$$

ตอบ ความโด่ง (K_u) = 1.6072

■ ค่าความเบ้ (S_k)

$$\text{สูตร } S_k = \frac{3(\bar{x} - \text{Median})}{s}$$

จากโจทย์กำหนด $\bar{x} = 117.75$ $S = 7.9237$ หา Median = 118.5

$$S_k = \frac{3(\bar{x} - \text{Median})}{s}$$

$$\begin{aligned} \text{แทนค่า } S_k &= \frac{3(117.75 - 118.5)}{7.9237} \\ &= \frac{-2.25}{7.9237} = -0.2840 \end{aligned}$$

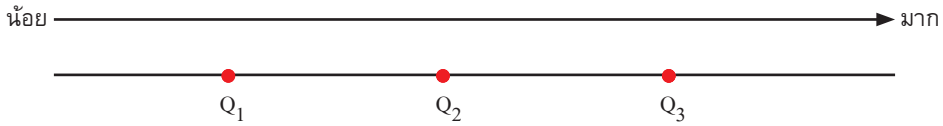
ตอบ ความเบ้ (S_k) = -0.2840

1.10 การวัดตำแหน่งของข้อมูล (Measure of Position)

โดยทั่วไป เมื่อเรากล่าวถึงตำแหน่งหรือลำดับที่ของข้อมูลชุดหนึ่ง เช่น เมื่อมีประกาศผลการสอบหรือการแข่งขันในการประกวดต่าง ๆ ของผู้เข้าสอบหรือผู้เข้าแข่งขัน คนก็มักจะถามว่าสอบได้ที่เท่าไร แข่งได้ที่เท่าไร ถ้าผู้เข้าสอบตอบว่าสอบได้ที่ 5 หรือแข่งได้ที่ 5 ถ้ามีข้อมูลเพียงแค่นี้ก็ไม่สามารถหาได้ว่าตำแหน่งที่ดีหรือไม่ดี ถ้ารู้ข้อมูลเพิ่ม เช่น จากผู้เข้าสอบทั้งหมด 50 คน ก็คงพอจะมองออกว่าเก่ง แต่ถ้าได้ที่ 5 จาก 6 คน ก็คงจะไม่เก่งนัก เพื่อขจัดปัญหาในการที่จะต้องรู้จำนวนผู้เข้าสอบทั้งหมด เราจะใช้วิธีจัดตำแหน่งของข้อมูล ซึ่งได้แก่ ค่าสถิติของข้อมูลที่เรียงลำดับจากน้อยไปหามาก โดยเป็นค่าที่แสดงตามตำแหน่งของข้อมูลซึ่งแบ่งออกเป็น 3 ชนิดคือ

1.10.1 ควอร์ไทล์ (Quartiles)

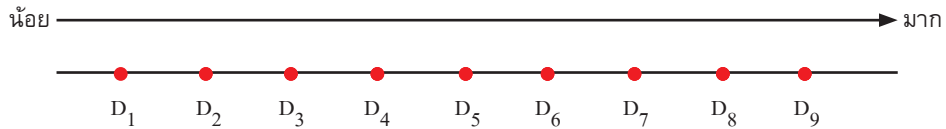
ควอร์ไทล์เป็นการแบ่งข้อมูลที่เรียงลำดับจากน้อยไปหามากออกเป็น 4 ส่วนเท่า ๆ กัน โดยมีจุดแบ่ง 3 จุด



- **ควอร์ไทล์ที่ 1** Q_1 เป็นค่าที่มีจำนวนค่าในข้อมูลซึ่งน้อยกว่าค่านี้อยู่ประมาณ $\frac{1}{4}$ ของจำนวนข้อมูลทั้งหมด (หรือ 25% ของจำนวนข้อมูลทั้งหมด)
- **ควอร์ไทล์ที่ 2** Q_2 เป็นค่าที่มีจำนวนค่าในข้อมูลซึ่งน้อยกว่าค่านี้อยู่ประมาณ $\frac{2}{4}$ ของจำนวนข้อมูลทั้งหมดซึ่งก็คือมัธยฐานนั่นเอง (หรือ 50% ของจำนวนข้อมูลทั้งหมด)
- **ควอร์ไทล์ที่ 3** Q_3 เป็นค่าที่มีจำนวนค่าในข้อมูลซึ่งน้อยกว่าค่านี้อยู่ประมาณ $\frac{3}{4}$ ของจำนวนข้อมูลทั้งหมด (หรือ 75% ของจำนวนข้อมูลทั้งหมด)

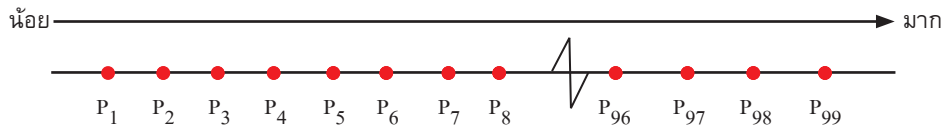
1.10.2 เดไซล์ (Deciles)

เดไซล์เป็นการแบ่งข้อมูลที่เรียงลำดับจากน้อยไปหามากออกเป็น 10 ส่วนเท่าๆ กัน มีจุดแบ่ง 9 จุดคือ $D_1, D_2, D_3, D_4, D_5, \dots, D_9$ เดไซล์ที่ k เป็นค่าที่มีจำนวนค่าในข้อมูลซึ่งน้อยกว่าค่านี้อยู่ประมาณ $\frac{k}{10}$ จากจำนวนข้อมูลทั้งหมด



1.10.3 เปอร์เซ็นไทล์ (Percentiles)

เปอร์เซ็นไทล์เป็นการแบ่งข้อมูลที่เรียงลำดับจากน้อยไปหามากออกเป็น 100 ส่วนเท่าๆ กัน มีจุดแบ่ง 99 จุดคือ $P_1, P_2, P_3, P_4, P_5, \dots, P_{99}$ เปอร์เซ็นไทล์ที่ k เป็นค่าที่มีจำนวนค่าในข้อมูลซึ่งน้อยกว่าค่านี้อยู่ประมาณ $\frac{k}{100}$ ของจำนวนข้อมูลทั้งหมด



ถ้าข้อมูลมีช่วงห่างเท่ากันทุกช่วง $Q_2 = D_5 = P_{50} = \text{มัธยฐาน}$

- การหาตำแหน่งของข้อมูลควอร์ไทล์ เดไซล์ และเปอร์เซ็นไทล์ของข้อมูลที่ไม่แจกแจงความถี่

$$Q_x = \frac{4x}{N + 1} \quad \text{เมื่อ } x \text{ คือตำแหน่งข้อมูลที่ต้องการหา}$$

$$D_x = \frac{10x}{N + 1} \quad N \text{ คือจำนวนข้อมูลทั้งหมด}$$

$$P_x = \frac{100x}{N + 1}$$

- การหาค่าของข้อมูลควอร์ไทล์ เดซิส์ และเปอร์เซ็นต์ไทล์ของข้อมูลที่ไม่แจกแจงความถี่ ถ้า N เป็นจำนวนข้อมูล หาดำแหน่ง Q_x , D_x , P_x ได้จากสูตร

$$\text{ตำแหน่งที่ของ } Q_x = \frac{x(N+1)}{4}$$

$$\text{ตำแหน่งที่ของ } D_x = \frac{x(N+1)}{10}$$

$$\text{ตำแหน่งที่ของ } P_x = \frac{x(N+1)}{100}$$

ในกรณี

- ตำแหน่ง th เป็นจำนวนเต็มให้ตอบจำนวนในตำแหน่งนั้นเลย
- ตำแหน่ง th เป็นจำนวนทศนิยม 0.5 ให้นำจำนวนในตำแหน่งที่อยู่ระหว่างตำแหน่ง th มาบวกกันแล้วหาร 2
- ตำแหน่ง th เป็นจำนวนทศนิยมอื่น พิจารณาว่า th เข้าใกล้จำนวนเต็มใด ก็ให้ตอบจำนวนนั้น หรืออาจเทียบบัญญัติไตรยางค์หาค่าเพื่อความละเอียดและความถูกต้องมากยิ่งขึ้น

ตัวอย่างที่ 1.7

ผลการทดสอบคะแนนวิชาบัญชีเบื้องต้นคะแนนเต็ม 20 คะแนน ของนักศึกษา คณะบริหารศาสตร์ชั้นปีที่ 1 โดยสุ่มมา 10 คน ปรากฏคะแนน ดังนี้ 9 11 18 10 9 13 15 9 13 14

จงหา Q_1 D_8 P_{50}

วิธีทำ จากข้อมูลมี $N = 10$

1. เรียงลำดับข้อมูลจากน้อยไปหามาก 9 9 9 10 11 13 13 14 15 18

2. หาดำแหน่ง $Q_1 = \frac{1(10+1)}{4} = 2.75^{th}$ แสดงว่าอยู่ระหว่างตำแหน่งที่ 3 กับ 2

ตำแหน่งต่างกัน $(3 - 2) = 1$ คะแนนต่างกัน $9 - 9 = 0$

ตำแหน่งต่างกัน $(2.75 - 2) = 0.75$ คะแนนต่างกัน $\frac{0.75 \times 0}{1} = 0$

คะแนน ณ ตำแหน่ง Q_1 คือ $9 + 0 = 9$ คะแนน

ดังนั้น ตำแหน่ง Q_1 คือ 9 คะแนน

ตอบ

3. หาดำแหน่ง $D_8 = \frac{8(10+1)}{10} = \frac{88}{10} = 8.8^{th} = 9^{th}$

เพราะฉะนั้นข้อมูลตำแหน่ง 9^{th} คือ $D_8 = 15$ คะแนน

ตอบ

4. หาดำแหน่ง $P_{50} = \frac{50(10+1)}{100} = \frac{550}{100} = 5.5^{th}$

แสดงว่าข้อมูลอยู่ระหว่างตำแหน่งที่ 5 กับ 6

ตำแหน่งต่างกัน $(6 - 5) = 1$

คะแนนต่างกัน $(13 - 11) = 2$ คะแนน

ตำแหน่งต่างกัน $(5.5 - 5) = 0.5$

คะแนนต่างกัน $\frac{0.50 \times 2}{1} = 1$ คะแนน

ตำแหน่ง $P_{50} = 11 + 1 = 12$ คะแนน

ดังนั้น ตำแหน่ง P_{50} คือ 12 คะแนน

ตอบ

ตัวอย่างที่ 1.8

จากข้อมูลต่อไปนี้ เป็นค่าใช้จ่ายรายวันของพนักงานโรงแรมแห่งหนึ่งในจังหวัดอุบลราชธานี 220 250 270 280 260 230 260 300 บาท จงหา P_{15} P_{20} P_{50} D_5 D_9 Q_3 Q_1 ส่วนเบี่ยงเบนควอร์ไทล์ และค่าใช้จ่ายรายวัน 280 บาท เป็นเปอร์เซ็นต์ไทล์ เดซิส์ และควอร์ไทล์ที่เท่าไร

วิธีทำ จากข้อมูลมี $N = 8$



กิจกรรม
ในห้องเรียน

1. เรียงลำดับข้อมูลจากน้อยไปหามาก

.....

2. หาดำแหน่ง $P_{15} = \frac{15(8 + 1)}{100} = \dots\dots\dots = \dots\dots\dots$

.....

3. หาดำแหน่ง $P_{20} = \dots\dots\dots$

.....

4. หาดำแหน่ง $P_{50} = \dots\dots\dots$

.....

5. หาดำแหน่ง $D_5 = \dots\dots\dots$

.....

6. หาดำแหน่ง $D_9 = \dots\dots\dots$

.....

7. หาดำแหน่ง $Q_1 = \dots\dots\dots$

$Q_3 = \dots\dots\dots$

8. ส่วนเบี่ยงเบนควอร์ไทล์.....

.....

.....

.....

9. ค่าใช้จ่ายรายวัน 280 บาท เป็นเปอร์เซ็นต์ไทล์ เดซิส์ และควอร์ไทล์ที่

.....

.....

.....

.....

.....

.....

.....

.....

.....

SE-ED

inspiration starts here

1.11 การทบทวนคณิตศาสตร์ระดับมัธยมศึกษา

ความรู้ในวิชาคณิตศาสตร์ระดับมัธยมศึกษาที่นักศึกษาจะทบทวนเพื่อที่จะนำความรู้ที่ได้มาต่อยอดเรียนเนื้อหาสถิติธุรกิจในบทถัดไปนั้น ได้แก่ แฟกทอเรียล การเรียงสับเปลี่ยน และการจัดหมู่ รายละเอียดแต่ละหัวข้อมีดังนี้

1.11.1 แฟกทอเรียล (Factorial : !)

นิยาม เมื่อ n เป็นจำนวนเต็มบวก ผลคูณของจำนวนเต็มบวกตั้งแต่ 1 ถึง n คือ แฟกทอเรียล n หรือ n แฟกทอเรียล เขียนแทนด้วยสัญลักษณ์ $n!$

ดังนั้น $n! = n \cdot (n - 1) \cdot (n - 2) \cdot (n - 3) \dots 3 \cdot 2 \cdot 1$

หรือ $n! = n \times (n - 1) \times (n - 2) \times (n - 3) \dots \times 3 \times 2 \times 1$

$n! = 1 \cdot 2 \cdot 3 \dots (n - 2) \cdot (n - 1) \cdot n$ ที่นิยมใช้คือเรียงจากมากไปน้อย

เช่น $7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

ในบางครั้งเราไม่จำเป็นต้องกระจายจนถึง 1 เช่น $9! = 9 \cdot 8 \cdot 7 \cdot 6!$

บทนิยาม เมื่อ $n = 0$ แฟกทอเรียล 0 เท่ากับ 1
 นั่นคือ $0! = 1$

ข้อควรระวัง :

(1) $5! \times 3! \neq (5 \times 3)!$

(2) $\frac{5!}{3!} \neq \left(\frac{5}{3}\right)!$

(3) $(-3)!$ ไม่นิยาม เพราะ -3 ไม่เป็นจำนวนเต็มบวก

SE-ED

inspiration starts here

ตัวอย่างที่ 1.9 การคำนวณหาค่าแฟกทอเรียล

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

$$(n + 3)! = (n + 3) \cdot (n + 2) \cdot (n + 1) \dots 3 \cdot 2 \cdot 1$$

$$(n - r)! = (n - r) \cdot (n - r - 1) \cdot (n - r - 2) \dots 3 \cdot 2 \cdot 1$$

$$X! = X \cdot (X - 1) \cdot (X - 2) \dots 3 \cdot 2 \cdot 1$$

$$n! = n \cdot (n - 1) \cdot (n - 2) \dots 3 \cdot 2 \cdot 1$$

$$5! = 5 \cdot 4 \cdot 3! = 5 \cdot 4 \cdot 6 = 120$$

$$\frac{7!}{5!} = \frac{7 \cdot 6 \cdot 5!}{5!} = 7 \cdot 6 = 42$$

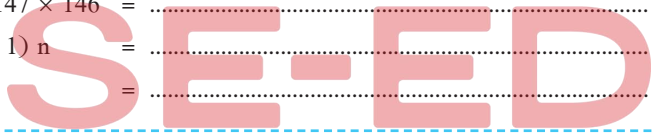
ตัวอย่างที่ 1.10 จงหาค่าของแฟกทอเรียล ดังต่อไปนี้

1. $4! = \dots\dots\dots$
2. $8! = \dots\dots\dots$
3. $6! = \dots\dots\dots$
4. $7 \cdot 9! = \dots\dots\dots$
5. $\frac{7!}{5! \cdot 3!} = \dots\dots\dots$
6. $2! \cdot 3! \cdot 6! = \dots\dots\dots$
7. $\frac{9!}{5!} = \dots\dots\dots$
8. $\frac{5! \cdot 4!}{3!} = \dots\dots\dots$
9. $0! = \dots\dots\dots$
10. $1! = \dots\dots\dots$
11. จงพิสูจน์ว่า $0! = 1!$ $\dots\dots\dots$

ตัวอย่างที่ 1.11 จงเขียนจำนวนต่อไปนี้ในรูปที่มีแฟกทอเรียล



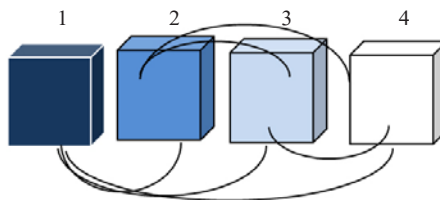
1. $1 \times 2 \times 3 \times \dots \times 9 = \dots\dots\dots$
2. $10 \times 9 \times 8 = \dots\dots\dots$
3. $30 \times 29 \times 28 \times 27 = \dots\dots\dots$
4. $150 \times 149 \times 148 \times 147 \times 146 = \dots\dots\dots$
5. $(n + 3)(n + 2)(n + 1)n = \dots\dots\dots$
6. $(n - 2)(n - 1)n = \dots\dots\dots$



inspiration starts here

1.11.2 การเรียงสับเปลี่ยน (Permutation)

ถ้ามีของ n สิ่งต่างๆ กัน จัดเรียงสับเปลี่ยนสิ่งของทีละ x สิ่ง จาก n สิ่ง มาเรียงเป็นแถวตามลำดับโดยคำนึงถึงลำดับก่อนหลัง ลำดับสลับที่กันถือว่าเป็นคนละวิธี จำนวนวิธีที่จะกระทำได้อคือ $P_{n, x} = \frac{n!}{(n - x)!}$ เช่น มีหนังสืออยู่ 4 เล่ม นำมาเรียงสับเปลี่ยนเป็นแถว แถวละ 2 เล่ม จะจัดเรียงได้กี่วิธี



จะเห็นได้ว่าเรียงสับเปลี่ยนได้ทั้งหมด 12 วิธีคือ เล่มที่ 12 13 14 23 24 34 และ 21 31 41 32 42 43 (สามารถสลับที่เล่มหนังสือได้) ถ้าใช้วิธีนับแบบนี้คงจะไม่สะดวกหากจำนวน n มีจำนวนมาก ดังนั้นการใช้สูตรในการคำนวณจึงสะดวกกว่าดังนี้ n คือจำนวนสิ่งของ x คือจำนวนที่นำมาเรียงสับเปลี่ยน เช่น มีหนังสือ 4 เล่ม เลือกหนังสือมา 2 เล่ม สามารถเรียงสับเปลี่ยนโดยวิธีคำนวณได้ดังนี้

$$P_{n, x} = P_{4, 2} = \frac{4!}{(4-2)!} = \frac{4!}{2!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{(2 \cdot 1)} = 12 \text{ วิธี}$$

หมายเหตุ :

$$P_{4, 2} = {}_4P_2$$

จงหาจำนวนวิธีเรียงสับเปลี่ยนต่อไปนี้



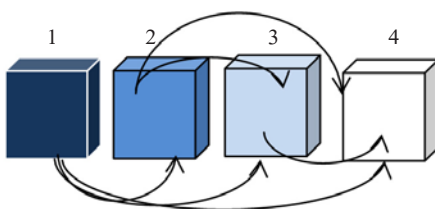
กิจกรรม
ในห้องเรียน

- ${}_7P_3 = \dots\dots\dots$
 ${}_9P_5 = \dots\dots\dots$
- ในการประกวดเขียนแผนธุรกิจระดับภูมิภาค มีผู้แข่งขันทั้งสิ้น 30 คน จะมีกี่วิธีที่ผู้แข่งขันจะได้รับเหรียญทอง เหรียญเงิน เหรียญทองแดง และชมเชย ถ้าผู้แข่งขันมีโอกาสนั้นจะชนะเท่า ๆ กันทุกคน
- คนกลุ่มหนึ่งมี 6 คน จะนำคนในกลุ่มนี้มา 3 คน เพื่อเลือกเข้าร่วมรับการอบรมภาษาอังกฤษธุรกิจ จะจัดได้กี่วิธี
- มีหนังสือต่างกัน 10 เล่ม วางอยู่บนโต๊ะ จะนำหนังสือเหล่านี้มา 5 เล่ม เพื่อนำมาจัดบนชั้นหนังสือ จะจัดได้กี่วิธี
- บริษัทแห่งหนึ่งมีตำแหน่งงานว่างอยู่ 3 ตำแหน่ง แต่มีคนมาสมัคร 15 คน ทางบริษัทมีวิธีคัดเลือกคนเข้าทำงานได้กี่วิธี

SE-ED
inspiration starts here

1.11.3 การจัดหมู่ (Combination)

ถ้ามีของ n สิ่งต่าง ๆ กัน จัดหมู่สิ่งของทีละ x สิ่ง จาก n สิ่ง มาเรียงเป็นแถวตามลำดับโดยไม่คำนึงถึงลำดับก่อนหลัง ลำดับสลับที่กันถือว่าเป็นวิธีเดียวกัน จำนวนวิธีที่จะกระทำได้อคือ

$$C_{n, x} = \frac{n!}{(n-x)! \times x!} \text{ เช่น มีหนังสืออยู่ 4 เล่ม นำมาเรียงจัดหมู่เป็นแถว แถวละ 2 เล่ม จะจัดเรียงได้กี่วิธี}$$


จะเห็นได้ว่าจัดหมู่ได้ทั้งหมด 6 วิธีคือ เล่มที่ 12 13 14 23 24 34

ถ้าใช้วิธีนับแบบนี้คงจะไม่สะดวกหากจำนวน n มีจำนวนมาก ดังนั้นการใช้สูตรในการคำนวณจึงสะดวกกว่าดังนี้ ให้ n คือ จำนวนสิ่งของ เช่น $n = 4$ และให้ x คือจำนวนที่นำมาจัดหมู่ เช่น $x = 2$ จะคำนวณได้ดังนี้

$$C_{n, x} = C_{4, 2} = \frac{4!}{(4-2)! \cdot 2!} = \frac{4!}{2! \cdot 2!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1 \cdot 2 \cdot 1} = 6 \text{ วิธี}$$

หมายเหตุ :

$$C_{4, 2} = {}_4C_2 = {}^4C_2 = \binom{4}{2}$$

จงหาจำนวนวิธีของการจัดหมู่ (Combination) ต่อไปนี้



กิจกรรม
ในข้อนี้

1. ${}_8C_4 = \dots\dots\dots$
2. ${}_8C_2 = \dots\dots\dots$
3. ${}_7C_3 = \dots\dots\dots$
4. $\binom{5}{1} = \dots\dots\dots$
5. $\binom{5}{4} = \dots\dots\dots$
6. $\binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3} = \dots\dots\dots$
7. $\binom{6}{0} + \binom{6}{1} + \binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6} = \dots\dots\dots$

SE-ED

inspiration starts here

1.12 การวิเคราะห์ค่าสถิติเบื้องต้นด้วยโปรแกรม SPSS

ในการวิเคราะห์ค่าสถิติเบื้องต้นที่เป็นสถิติเชิงพรรณนาด้วยโปรแกรม SPSS นั้น เป็นหลักการที่อธิบายหรือบรรยายลักษณะของข้อมูลที่เก็บรวบรวมมาโดยส่วนใหญ่ นิยมนำเสนอในรูปตารางและอธิบายข้อความประกอบได้ตาราง การบรรยายในส่วนผู้เขียนขอแนะนำเสนอเฉพาะสถิติเชิงบรรยายที่เป็นที่นิยมที่ใช้ในการวิเคราะห์ข้อมูลในการทำวิจัย ซึ่งเป็น 2 ตอนดังนี้

ตอนที่ 1 การวิเคราะห์ข้อมูลที่เป็นตัวแปรจัดกลุ่ม มีดังนี้

- ค่าความถี่
- ร้อยละ

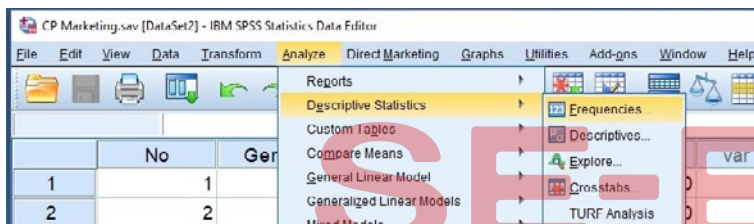
ตอนที่ 2 การวิเคราะห์ที่เป็นตัวแปรต่อเนื่อง มีดังนี้

- ค่าเฉลี่ย
- ส่วนเบี่ยงเบนมาตรฐาน
- ค่าต่ำสุด
- ค่าสูงสุด
- ความเบ้
- ความโด่ง

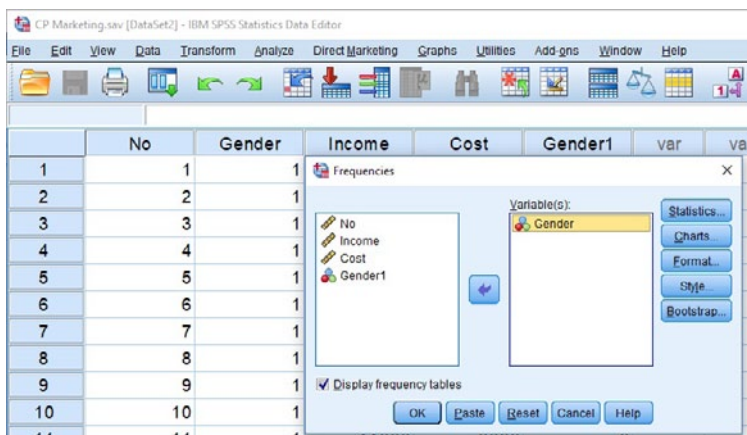
การวิเคราะห์ข้อมูลเบื้องต้น ตัวแปรจัดกลุ่มสามารถหาค่าความถี่ และร้อยละในโปรแกรม SPSS ทำได้ ดังนี้

Analyze → Descriptive Statistics → Frequencies

ดังแสดงในรูปต่อไปนี้



ใส่ตัวแปรที่ต้องการวิเคราะห์ความถี่ ร้อยละ ในช่อง Variable(s) แล้วกด OK ในกรณีตัวอย่าง ใส่ตัวแปร Gender (เพศ) ในช่อง Variable(s) ดังแสดงในรูปต่อไปนี้



ผลการวิเคราะห์จากโปรแกรม SPSS ในกรณีนี้ที่ตัวแปรจัดกลุ่มเป็นดังนี้

เพศ	ความถี่	ร้อยละ	Valid Percent	Cumulative Percent
Valid	male	5	15.6	15.6
	female	27	84.4	100.0
	Total	32	100.0	

เราสามารถจัดรูปแบบตารางให้เข้าใจง่ายและสะดวกในการแปลผลได้ดังแสดงในตารางที่ 1.2

ตารางที่ 1.2 ผลการวิเคราะห์สถิติเชิงบรรยายของตัวแปรเพศ

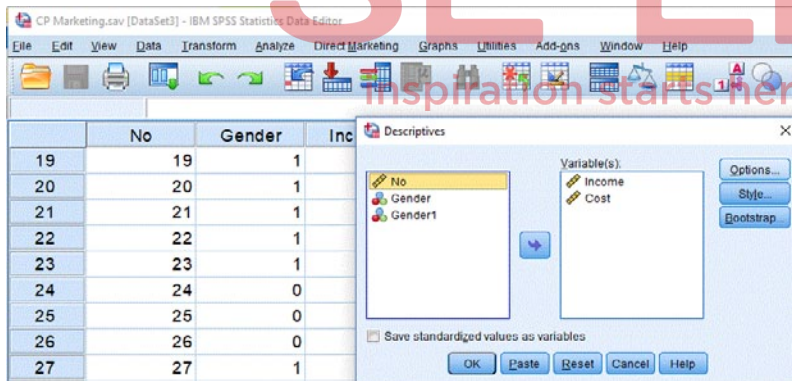
ตัวแปร	ความถี่	ร้อยละ
ชาย	5	15.6
หญิง	27	84.4
รวม	32	100.0

จากตารางที่ 1.2 พบว่ามีผู้ตอบแบบสอบถาม 32 คน ส่วนใหญ่เป็นเพศหญิง 27 คน คิดเป็นร้อยละ 84.4 และเป็นเพศชาย 5 คน คิดเป็นร้อยละ 15.6

การหาค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน ค่าต่ำสุด ค่าสูงสุด ความเบ้ ความโด่ง ของตัวแปรต่อเนื่องในโปรแกรม SPSS ทำได้ ดังนี้

Analyze → Descriptive Statistics → Descriptives

ดังแสดงในรูปต่อไปนี้



จากนั้นคลิก Options แล้วเลือก สถิติ Mean SD Min Max Skewness Kurtosis

ผลการวิเคราะห์จากโปรแกรม SPSS ในกรณีตัวแปรแบบต่อเนื่องเป็นดังนี้

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Income	32	7,000	15,000	11,218.75	1,929.974	.272	.414	.067	.809
Cost	32	6,000	15,000	9,343.75	2,168.088	1.280	.414	1.912	.809

เราสามารถจัดรูปแบบตารางให้เข้าใจง่ายและสะดวกในการแปลผลได้ดังแสดงในตารางที่ 1.3

ตารางที่ 1.3 ผลการวิเคราะห์สถิติเชิงบรรยายของตัวแปรรายได้กับค่าใช้จ่ายต่อเดือนของนักศึกษา
คณะบริหารศาสตร์ มหาวิทยาลัยอุบลราชธานี ปีที่ 4 ภาคการศึกษา 2/2560

ตัวแปร	N	Mean	SD	Min	Max	Sk	Ku
รายได้	32	11,218.75	1,929.97	7,000	15,000	0.272	0.067
ค่าใช้จ่าย	32	9,343.75	2,168.09	6,000	15,000	1.28	1.912

จากตารางที่ 1.3 พบว่ามีนักศึกษาตอบแบบสอบถาม 32 คน มีรายได้เฉลี่ยต่อเดือน 11,218.75 บาท ส่วนเบี่ยงเบนมาตรฐาน 1,929.97 บาท รายได้ต่อเดือนต่ำสุด 7,000 บาท และรายได้ต่อเดือนสูงสุด 15,000 บาท รายได้เฉลี่ยต่อเดือนมีความเบ้เท่ากับ $0.272 < 2.00$ ดังนั้นข้อมูลนี้มีรูปร่างเป็นโค้งปกติหรือไม่เบ้ ส่วนความโด่งมีค่าเท่ากับ $0.067 < 2$ ดังนั้นข้อมูลนี้มีรูปร่างเป็นโค้งปกติ และผู้ตอบแบบสอบถาม 32 คน มีค่าใช้จ่ายต่อเดือนเฉลี่ย 9,343.75 บาท ส่วนเบี่ยงเบนมาตรฐาน 2,168.09 บาท รายได้ต่อเดือนต่ำสุด 6,000 บาท และรายได้ต่อเดือนสูงสุด 15,000 บาท ค่าใช้จ่ายต่อเดือนมีความเบ้เท่ากับ $0.272 < 2.00$ ดังนั้นข้อมูลนี้มีรูปร่างเป็นโค้งปกติหรือไม่เบ้ ส่วนความโด่งมีค่าเท่ากับ $0.067 < 2$ ดังนั้นข้อมูลนี้มีรูปร่างเป็นโค้งปกติ

ในกรณีข้อมูลแบบเรียงอันดับ (Ordinal Scale)

ลักษณะข้อมูลเป็นแบบมาตราส่วนประเมินค่า (Rating Scale) ที่กำหนดมาตรวัดเป็นระดับตามวิธีของลิเคิร์ต (Likert's Scale) เช่น มี 5 ระดับ ในการแปลความหมายระดับความสำคัญนั้น มีเกณฑ์การแปลความหมายดังนี้

คะแนน

ระดับความสำคัญ

- 5 หมายความว่า ผู้ตอบให้ความสำคัญกับด้านนั้นมากที่สุด
- 4 หมายความว่า ผู้ตอบให้ความสำคัญกับด้านนั้นมาก
- 3 หมายความว่า ผู้ตอบให้ความสำคัญกับด้านนั้นปานกลาง
- 2 หมายความว่า ผู้ตอบให้ความสำคัญกับด้านนั้นน้อย
- 1 หมายความว่า ผู้ตอบให้ความสำคัญกับด้านนั้นน้อยที่สุด

เกณฑ์ในการแปลความหมายของค่าเฉลี่ย

ผู้เขียนใช้เกณฑ์ของบุญชม ศรีสะอาด และบุญส่ง นิลแก้ว (2535) ที่มีเกณฑ์การแปลความหมายดังนี้

ค่าเฉลี่ย	การแปลผล
4.51 - 5.00	อยู่ในระดับมากที่สุด
3.51 - 4.50	อยู่ในระดับมาก
2.51 - 3.50	อยู่ในระดับปานกลาง
1.51 - 2.50	อยู่ในระดับน้อย
1.00 - 1.50	อยู่ในระดับน้อยที่สุด

ผู้เขียนขอยกตัวอย่างการวิเคราะห์ข้อมูลค่าเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และการแปลผลจากงานวิจัยของละออง มังตะการ และพิชญาดา พันผา (2561) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อโอกาสการตัดสินใจเลือกใช้บริการทางการเงินธนาคารพาณิชย์ไทยของผู้ประกอบการ SMEs ในจังหวัดอุบลราชธานี โดยมีวัตถุประสงค์ในการวิจัยข้อที่ (1) เพื่อศึกษาระดับปัจจัยส่วนประสมการตลาดบริการ ปัจจัยภาพลักษณ์องค์กร ปัจจัยการบริหารลูกค้าสัมพันธ์ และการตัดสินใจเลือกใช้บริการทางการเงินธนาคารพาณิชย์ไทยของผู้ประกอบการ SMEs ในจังหวัดอุบลราชธานี กลุ่มตัวอย่างเป็นผู้ประกอบการ SMEs ในจังหวัดอุบลราชธานี จำนวน 382 คน ได้จากการสุ่มตัวอย่างแบบแบ่งพื้นที่มา 3 อำเภอ ในจังหวัดอุบลราชธานี เครื่องมือวิจัยเป็นแบบสอบถาม มีค่าสัมประสิทธิ์ความเที่ยงเท่ากับ 0.97 จากวัตถุประสงค์การวิจัยดังกล่าว ได้ผลการวิจัยดังนี้

ตารางที่ 1.4 ผลการวิเคราะห์ปัจจัยส่วนประสมการตลาดของสินเชื่อบankerพาณิชย์ไทยตามมุมมองของผู้ประกอบการ SMEs ในจังหวัดอุบลราชธานี เป็นดังนี้

ปัจจัยส่วนประสมการตลาด	Mean	SD	ระดับความสำคัญ
1. ด้านผลิตภัณฑ์	3.637	0.701	มาก
2. ด้านราคา	3.370	0.712	ปานกลาง
3. ด้านสถานที่	3.699	0.773	มาก
4. ด้านการส่งเสริมการตลาด	3.624	0.703	มาก
รวม	3.643	0.585	มาก

ผลการวิเคราะห์ข้อมูล พบว่า ปัจจัยส่วนประสมการตลาดของสินเชื่อบankerพาณิชย์ไทยตามการรับรู้ของผู้ประกอบการ SMEs ในจังหวัดอุบลราชธานี ในภาพรวมอยู่ในระดับมาก มีค่าเฉลี่ยเท่ากับ 3.643 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.585 เมื่อพิจารณาตามแต่ละด้านเรียงลำดับค่าเฉลี่ยจากมากไปหาน้อยเป็นดังนี้ ด้านสถานที่ที่มีค่าเฉลี่ยเท่ากับ 3.699 อยู่ในระดับมาก และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.773 รองลงมาคือ ด้านผลิตภัณฑ์มีค่าเฉลี่ยเท่ากับ 3.637

อยู่ในระดับมาก และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.701 ด้านการส่งเสริมการตลาดมีค่าเฉลี่ยเท่ากับ 3.624 อยู่ในระดับมาก และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.703 ส่วนด้านที่มีความสำคัญในระดับปานกลาง คือด้านราคามีค่าเฉลี่ยเท่ากับ 3.370 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.712 ตามลำดับ

ผลการวิเคราะห์ข้อมูลเกี่ยวกับระดับความสำคัญของปัจจัยภาพลักษณ์องค์กรเป็นดังนี้

ตารางที่ 1.5 ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (SD) ของปัจจัยภาพลักษณ์องค์กร

ปัจจัยภาพลักษณ์องค์กร	Mean	SD	ระดับความสำคัญ
1. ความมั่นคงและความน่าเชื่อถือของธนาคาร	3.812	0.576	มาก
2. ด้านตราสินค้า	3.438	0.872	ปานกลาง
3. ด้านความรับผิดชอบต่อสังคม	3.707	0.741	มาก
รวม	3.649	0.630	มาก

จากตารางที่ 1.5 พบว่า ผู้ตอบแบบสอบถามมีความคิดเห็นต่อปัจจัยภาพลักษณ์องค์กรในภาพรวมอยู่ในระดับมาก มีค่าเฉลี่ยเท่ากับ 3.649 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.630 เมื่อพิจารณาตามแต่ละด้าน เรียงลำดับค่าเฉลี่ยจากมากไปหาน้อยเป็นดังนี้ ด้านความมั่นคงและความน่าเชื่อถือของธนาคารมีค่าเฉลี่ยเท่ากับ 3.812 อยู่ในระดับมาก และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.576 รองลงมาคือ ด้านความรับผิดชอบต่อสังคมมีค่าเฉลี่ยเท่ากับ 3.707 อยู่ในระดับมาก และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.741 ด้านตราสินค้ามีค่าเฉลี่ยเท่ากับ 3.438 อยู่ในระดับปานกลาง และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.872 ตามลำดับ

SE-ED
inspiration starts here

1.13 สรุปเนื้อหาบทที่ 1

สถิติ หมายถึงศาสตร์ที่ศึกษาเกี่ยวกับการเก็บรวบรวมข้อมูล การวิเคราะห์ข้อมูล การแปลความหมาย การสรุปผล รวมถึงการนำเสนอข้อมูล และอีกความหมายคือค่าที่วิเคราะห์ได้จากตัวอย่าง

สถิติทางธุรกิจ หมายถึงศาสตร์ที่ศึกษาเกี่ยวกับการเก็บรวบรวมข้อมูลหรือจำนวนทางธุรกิจ เช่น กำไร ขาดทุน รายรับ รายจ่าย ยอดการลงทุน ราคาสินค้า ดัชนีราคา เป็นต้น แล้วข้อมูลทางธุรกิจที่ได้มาวิเคราะห์ แปลความหมาย สรุปผล และนำเสนอ ผลการวิเคราะห์ข้อมูลสามารถนำไปใช้ประโยชน์ต่อการดำเนินการทางธุรกิจในลักษณะต่างๆ เช่น การวางแผนการลงทุน การขยายการลงทุน การตัดสินใจซื้อสินค้า การจ้าง-ลดพนักงาน เป็นต้น

สถิติแบ่งออกเป็น 2 ประเภท

1. สถิติเชิงบรรยาย เป็นสถิติที่ใช้ในการนำเสนอลักษณะข้อมูลเบื้องต้น ใช้ในการศึกษาข้อเท็จจริงจากข้อมูลที่รวบรวมมาได้ เพื่อให้ทราบรายละเอียดเกี่ยวกับลักษณะของข้อมูลกลุ่มนั้น โดยไม่ได้สรุปอ้างอิงผลการศึกษาไปยังประชากรที่ศึกษา สถิติเชิงบรรยาย ได้แก่ การแจกแจงความถี่ การวัดตำแหน่งการเปรียบเทียบ การวัดแนวโน้มเข้าสู่ส่วนกลาง และการวัดการกระจายของข้อมูล เป็นต้น

2. สถิติเชิงอ้างอิงหรือเชิงอนุมาน เป็นสถิติที่ว่าด้วยการสรุปอ้างอิงข้อมูลเพียงบางส่วนจากตัวอย่างไปยังประชากร ตัวอย่างสถิติเชิงอ้างอิง เช่น ความน่าจะเป็น การประมาณค่า และการทดสอบสมมติฐาน ถ้าหากแบ่งกลุ่มสถิติอ้างอิงตามลักษณะพารามิเตอร์ จะแบ่งได้ 2 กลุ่ม ได้แก่

- (1) กลุ่มสถิติพาราเมตริก (Parametric Statistics) เป็นกลุ่มสถิติวิเคราะห์ที่มีข้อกำหนดหรือข้อตกลงเบื้องต้นเกี่ยวกับค่าพารามิเตอร์และการแจกแจงของสถิติทดสอบ เช่น t-test, Z-test, F-test เป็นต้น
- (2) กลุ่มสถิตินอนพาราเมตริก (Non-parametric Statistics) หรือสถิติที่ไม่มีเงื่อนไขการแจกแจง ซึ่งเป็นสถิติวิเคราะห์ที่ไม่มีข้อกำหนดเกี่ยวกับค่าพารามิเตอร์ เช่น χ^2 -test, Kolmogorov-smirnov Test, Mann-whitney U Test, Kruskal-wallis One-way ANOVA เป็นต้น

สถิติธุรกิจและการวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS



ประวัติผู้เขียน

ผศ.ดร. พิชญาดา พันผา

ตำแหน่งปัจจุบัน :

ผู้ช่วยศาสตราจารย์ สาขาการจัดการธุรกิจ คณะบริหารศาสตร์ มหาวิทยาลัยอุบลราชธานี

การศึกษา :

- พ.ย. 2554 - ก.พ. 2558 ได้รับทุนจาก The British Council ประเภททุน The Researcher Links ไปทำวิจัยที่ The University of Strathclyde ประเทศอังกฤษ
- พ.ศ. 2555 จบการศึกษาระดับปริญญาเอก จากจุฬาลงกรณ์มหาวิทยาลัย (GPA=3.92) สาขาวิชาวิทยาการวิจัยทางการศึกษา โดยได้รับทุนจาก สกอ. ประเภททุนในและต่างประเทศ
- พ.ศ. 2554 ได้รับประกาศนียบัตรทางการเรียน Social Research Methodology จาก University of California, Los Angeles (UCLA) สหรัฐอเมริกา
- พ.ศ. 2547 จบการศึกษาระดับปริญญาโท จากจุฬาลงกรณ์มหาวิทยาลัย สาขาวิชาทางการศึกษา (GPA=3.74)
- พ.ศ. 2544 จบการศึกษาระดับปริญญาตรี จากมหาวิทยาลัยขอนแก่น สาขาการสอนคณิตศาสตร์ เกียรตินิยมอันดับ 2 (GPA=3.33) โดยได้รับทุนคณิตศาสตร์ เร่งรัดพัฒนาคณิตศาสตร์ (รพค.)

ผลงานทางวิชาการ :

- ตำรา สถิติธุรกิจ (2560). โรงพิมพ์มหาวิทยาลัยอุบลราชธานี
- บทความวิชาการ / บทความวิจัยที่ได้รับการตีพิมพ์ในวารสารในประเทศและต่างประเทศมากกว่า 20 เรื่อง เช่น Journal of Business Administration, Economics and Communication (BEC Journal), Vol 18 (1) in TC11 and Asian Citation Index (ACI), RTNA Journal of Social Sciences, Humanities, and Education, Vol 9, No 1., ABAC Journal, Vol 39, Vol. 41 in SCOPUS index Q1, Frontiers in Education, Vol. 4 (6) in SCOPUS Q2, วารสารครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ปีที่ 45 ฉบับที่ 4, วารสารพัฒนาบริหารศาสตร์ (NIDA Development Journal) ปีที่ 58 ฉบับที่ 4, วารสารศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม ปีที่ 14 ฉบับที่ 2 และ 3, วารสารบริหารการศึกษา มหาวิทยาลัยขอนแก่น ปีที่ 12 ฉบับที่ 2, Asian Forum on Business Education Journal, Vol. 9 (2), วารสารสมาคมนักวิจัยแห่งประเทศไทย Vol. 23 (2) และอื่นๆ อีกมากมาย

ความเชี่ยวชาญ :

- ด้านระเบียบวิธีวิจัยและสถิติทางด้านบริหารศาสตร์ ด้านสังคมศาสตร์ คณิตศาสตร์ธุรกิจ และด้านการศึกษา สถิติขั้นสูง การวิเคราะห์โมเดล โครงสร้าง การวิเคราะห์พหุระดับ การวิเคราะห์ Survival Analysis และมีความเชี่ยวชาญด้านโปรแกรม SPSS, MPLUS, HLM และ LISREL เป็นต้น



www.se-ed.com



sbc.fans



SE-ED Publisher

พร้อมจำหน่ายในรูปแบบ

- | | |
|--|---|
| ☒ e-book (PDF) | ☐ audiobooks |
| ☐ e-book (EPUB) | ☐ audio CD / MP3 |
| ☒ ปกอ่อน | ☐ LARGE PRINT (ตัวอักษรขนาดใหญ่) |

ISBN 978-616-08-4748-8



9 786160 847488

299 บาท