

เอกสารประกอบการเรียน
รายวิชา หลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์
(Principles of Statistics for Computer Scientists)

พิชญ์สินี พุทธิทวีศรี
วท.ม. (เทคโนโลยีสารสนเทศ)
วท.บ. (วิทยาการคอมพิวเตอร์)

คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยราชภัฏสวนดุสิต
2560

คำนำ

เอกสารประกอบการเรียน รายวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์ (Principles of Statistics for Computer Scientists) รหัสวิชา 4121405 นี้ ได้เรียบเรียงขึ้นอย่างเป็นระบบ ครอบคลุมเนื้อหาสาระรายวิชา ในหมวดวิชาเฉพาะ วิชาแกน ของมหาวิทยาลัย หลักสูตรวิทยาศาสตร์บัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ เพื่อใช้เป็นเครื่องมือสำคัญของผู้สอนในการใช้ประกอบการสอนของอาจารย์ ที่มุ่งเน้นให้ผู้เรียนมีความรู้ความเข้าใจในเนื้อหา

รายวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์ เป็นวิชาที่เกี่ยวข้องกับวิธีการทางสถิติ การวิจัยเบื้องต้น สามารถนำความรู้เกี่ยวกับสถิติไปประยุกต์ใช้ในการวิเคราะห์ข้อมูล สามารถออกแบบงานวิจัย สร้างเครื่องมือที่ใช้ในการวิจัย และเลือกใช้โปรแกรมประยุกต์ทางด้านสถิติในการวิเคราะห์สถิติได้อย่างถูกต้อง สามารถแปลผลและนำเสนอข้อมูลได้อย่างเหมาะสม ผู้เรียนจำเป็นต้องเข้าฟังการบรรยาย และทำแบบฝึกหัดควบคู่กันไปทุกครั้ง เพื่อให้เกิดประสิทธิภาพในการเรียนมากยิ่งขึ้น

เอกสารเล่มนี้ได้แบ่งเนื้อหาในการเรียนการสอนไว้ 12 บท ได้แก่ ความรู้ทั่วไปเกี่ยวกับสถิติ สถิติบรรยาย ความน่าจะเป็น การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม การสุ่มกลุ่มตัวอย่าง ตัวแปรและเครื่องมือที่ใช้ในการวิจัย การจัดเตรียมข้อมูลสำหรับการประมวลผลด้วยโปรแกรมประยุกต์ทางด้านสถิติ การวิเคราะห์สถิติบรรยาย การประมาณค่าและการทดสอบสมมติฐาน การทดสอบไคสแควร์ การวิเคราะห์การถดถอยและสหสัมพันธ์อย่างง่าย และ การเขียนรายงานการวิจัย ผู้เรียนควรได้ศึกษารายละเอียดแต่ละหัวข้อเรื่องที่สอนจากเอกสาร หนังสือ ตำรา หรือสื่ออื่นๆ เพิ่มเติมอีก หวังว่าเอกสารประกอบการสอนเล่มนี้คงอำนวยความสะดวกต่อการเรียนการสอนตามสมควร หากท่านที่นำไปใช้มีข้อเสนอแนะ ผู้เขียนยินดีรับฟังข้อคิดเห็นต่างๆ และขอขอบคุณมา ณ โอกาสนี้

พิชญ์สินี พุทธิทวีศรี

4 มกราคม 2560

สารบัญ

	หน้า
คำนำ	(1)
สารบัญ	(3)
สารบัญภาพ	(7)
สารบัญตาราง	(11)
บทที่ 1 ความรู้ทั่วไปเกี่ยวกับสถิติ	1
ความหมายของสถิติ	1
ระเบียบวิธีการทางสถิติ	1
วิธีวิเคราะห์ข้อมูลเชิงสถิติ	2
ความหมายของข้อมูล	3
ประเภทของข้อมูล	3
ความหมายของการวิจัย	4
ประเภทของการวิจัย	5
ขั้นตอนการทำวิจัย	6
บทบาทของสถิติในงานวิจัย	8
การนำสถิติมาใช้ในการวิจัย	9
สรุป	11
แบบฝึกหัด	11
เอกสารอ้างอิง	12
บทที่ 2 สถิติบรรยาย	13
การแจกแจงความถี่	13
การวัดแนวโน้มเข้าสู่ส่วนกลาง	20
การวัดการกระจาย	29
การนำเสนอข้อมูล	38
สรุป	44
แบบฝึกหัด	44
เอกสารอ้างอิง	47
บทที่ 3 ความน่าจะเป็น	49
การทดลองสุ่ม	49
สเปซตัวอย่าง	49
เหตุการณ์	50
การนับจุดตัวอย่าง	51

	หน้า
ความน่าจะเป็น	57
สรุป	67
แบบฝึกหัด	67
เอกสารอ้างอิง	69
บทที่ 4 การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม	71
ตัวแปรสุ่ม	71
ชนิดของตัวแปรสุ่ม	72
การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม	74
การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง	77
การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง	85
สรุป	99
แบบฝึกหัด	99
เอกสารอ้างอิง	101
บทที่ 5 การสุ่มกลุ่มตัวอย่าง	103
ประชากรและกลุ่มตัวอย่าง	103
ลักษณะของกลุ่มตัวอย่างที่ดี	104
การกำหนดขนาดของกลุ่มตัวอย่าง	105
ประเภทของการสุ่มกลุ่มตัวอย่าง	110
สรุป	114
แบบฝึกหัด	115
เอกสารอ้างอิง	116
บทที่ 6 ตัวแปรและเครื่องมือในที่ใช้ในการวิจัย	117
ความหมายของตัวแปร	117
ประเภทของตัวแปร	117
ระดับการวัดตัวแปร	118
เครื่องมือที่ใช้ในการวิจัย	120
แนวคิดและขั้นตอนในการสร้างแบบสอบถาม	122
การตรวจสอบคุณภาพของเครื่องมือ	127
สรุป	129
แบบฝึกหัด	129
เอกสารอ้างอิง	131

	หน้า	
บทที่ 7	การจัดเตรียมข้อมูลสำหรับการประมวลผลด้วยโปรแกรมประยุกต์ทางด้านสถิติ	133
	โปรแกรมประยุกต์ทางด้านสถิติ	133
	การประมวลผลด้วยโปรแกรมประยุกต์ทางด้านสถิติ	133
	การกำหนดรหัสแทนข้อมูล	135
	การจัดทำคู่มือการลงรหัส	139
	การเริ่มต้นใช้งานโปรแกรมประยุกต์ทางด้านสถิติ	144
	การสร้างแฟ้มข้อมูล	149
	การจัดการข้อมูลในโปรแกรมประยุกต์ทางด้านสถิติ	153
	การจัดกระทำกับข้อมูลในโปรแกรมประยุกต์ทางด้านสถิติ	158
	สรุป	167
	แบบฝึกหัด	168
	เอกสารอ้างอิง	169
บทที่ 8	การวิเคราะห์สถิติบรรยาย	171
	การวิเคราะห์สถิติบรรยาย	171
	การสร้างตารางแจกแจงความถี่แบบทางเดียว	172
	การสร้างตารางแจกแจงความถี่แบบสองทาง	175
	การวิเคราะห์ค่าสถิติพื้นฐาน ด้วยคำสั่ง Descriptive	178
	การวิเคราะห์ค่าความพึงพอใจ/ทัศนคติ	183
	การวิเคราะห์คำถามแบบเลือกตอบได้หลายคำตอบ	190
	สรุปสถิติบรรยายที่ใช้ในการวิเคราะห์ข้อมูล	197
	สรุป	198
	แบบฝึกหัด	198
	เอกสารอ้างอิง	200
บทที่ 9	การประมาณค่าและการทดสอบสมมติฐาน	201
	การประมาณค่า	201
	สมมติฐาน	202
	การทดสอบสมมติฐาน	204
	ประเภทของการทดสอบสมมติฐาน	206
	ขั้นตอนการทดสอบสมมติฐาน	208
	การตรวจสอบการแจกแจงของข้อมูล	208
	สรุป	212
	แบบฝึกหัด	212
	เอกสารอ้างอิง	214

	หน้า
บทที่ 10 การทดสอบโคสแควร์	215
การทดสอบโคสแควร์กับข้อมูลจำแนกทางเดียว	215
การทดสอบโคสแควร์กับข้อมูลจำแนกสองทาง	219
การทดสอบโคสแควร์ด้วยโปรแกรมประยุกต์ทางด้านสถิติ	227
ข้อควรระวังในการทดสอบความสัมพันธ์แบบโคสแควร์	232
สรุป	233
แบบฝึกหัด	233
เอกสารอ้างอิง	235
บทที่ 11 การวิเคราะห์การถดถอยและสหสัมพันธ์อย่างง่าย	237
การวิเคราะห์สหสัมพันธ์	237
การวิเคราะห์การถดถอย	245
การวิเคราะห์การถดถอยอย่างง่าย	246
สรุป	261
แบบฝึกหัด	261
เอกสารอ้างอิง	263
บทที่ 12 การเขียนรายงานการวิจัย	265
จุดมุ่งหมายของการเขียนรายงานการวิจัย	265
หลักในการเขียนรายงานการวิจัย	265
ประเภทของรายงานทางวิชาการ	266
ส่วนประกอบของรายงานการวิจัย	267
การเขียนอ้างอิงและบรรณานุกรม	271
สรุป	274
แบบฝึกหัด	274
เอกสารอ้างอิง	275
บรรณานุกรม	277

สารบัญภาพ

ภาพที่		หน้า
2.1	ความสัมพันธ์ระหว่างค่าเฉลี่ยเลขคณิต ค่ามัธยฐาน และค่าฐานนิยม	29
2.2	แผนภูมิแท่งแสดงจำนวนของห้องพักของโรงแรมและเกสต์เฮ้าส์ พ.ศ. 2557 จำแนกตามภาค	41
2.3	แผนภูมิวงกลมแสดงร้อยละของห้องพักของโรงแรมและเกสต์เฮ้าส์ พ.ศ. 2557 จำแนกตามภาค	42
2.4	กราฟเส้นแสดงร้อยละของประชากรอายุ 6 ปีขึ้นไปที่ใช้คอมพิวเตอร์ อินเทอร์เน็ต และโทรศัพท์มือถือ พ.ศ. 2553-2557	42
2.5	แผนภาพแสดงการเติบโตของประชากรปี ค.ศ. 1650-2070	43
2.6	อินโฟกราฟิกส์แสดงความเร็วของอินเทอร์เน็ตแต่ละประเทศในอาเซียนปี 2014	43
3.1	การเรียงสิ่งของ n สิ่งที่แตกต่างกัน	54
3.2	การจัดเรียงตัวอักษร a, b, c ในแนววงกลม	55
4.1	ชนิดของตัวแปรสุ่ม	73
4.2	แผนภูมิแท่งแสดงการทอดลูกเต๋า	78
4.3	การแจกแจงแบบปกติ	87
4.4	แสดงรูปโค้งปกติเมื่อ $\mu_1 \neq \mu_2$	88
4.5	แสดงรูปโค้งปกติเมื่อ $\mu_1 = \mu_2$	89
4.6	แสดงรูปโค้งปกติเมื่อ $\mu_1 \neq \mu_2$ และ $\sigma_1 \neq \sigma_2$	89
4.7	พื้นที่ใต้โค้งปกติ แสดง $P(a < x < b)$	90
4.8	การแจกแจงโคสแควร์	95
4.9	พื้นที่ใต้โค้งโคสแควร์	98
5.1	การแบ่งกลุ่มโดยเทคนิคการสุ่มกลุ่มตัวอย่างแบบแบ่งชั้น	111
5.2	การแบ่งกลุ่มโดยเทคนิคการสุ่มกลุ่มตัวอย่างแบบกลุ่ม	111
5.3	การสุ่มกลุ่มตัวอย่างแบบหลายขั้นตอน	112
7.1	ขั้นตอนการเรียกใช้งานโปรแกรม SPSS	144
7.2	หน้าต่างเลือกวิธีการทำงาน	144
7.3	หน้าต่าง Data Editor	145
7.4	หน้าต่าง Output Window	145
7.5	ส่วนประกอบของโปรแกรม SPSS	146
7.6	ขั้นตอนการกำหนดรูปแบบการแสดงผลแบบอักษรในหน้าต่าง Data Editor	148
7.7	ขั้นตอนการกำหนดค่ามาตรฐานของแบบอักษรในโปรแกรม SPSS	149
7.8	การสร้างเพิ่มข้อมูล	149

ภาพที่		หน้า
7.9	ชนิดของตัวแปรในโปรแกรม SPSS	151
7.10	ขั้นตอนการกำหนดรหัสแทนข้อมูลและค่าที่เป็นไปได้	152
7.11	การกำหนดค่าข้อมูลขาดหาย	152
7.12	ขั้นตอนการบันทึกแฟ้มข้อมูลในครั้งแรก	154
7.13	ขั้นตอนการเปิดแฟ้มข้อมูล	154
7.14	การกรอกข้อมูลใหม่	155
7.15	การแทรกข้อมูลใหม่	155
7.16	การลบข้อมูลที่ละหลายๆ เซลล์ติดกัน	156
7.17	การลบข้อมูลทั้งชุด	156
7.18	การลบข้อมูลตัวแปรบางตัวออกทุกชุดข้อมูล	157
7.19	การสั่งพิมพ์ข้อมูล	158
7.20	ขั้นตอนการจัดเรียงข้อมูล	159
7.21	ขั้นตอนการเปลี่ยนแปลงค่าของข้อมูลแล้วทำการบันทึกไว้ในตัวแปรเดิม	161
7.22	ขั้นตอนการเปลี่ยนแปลงค่าของข้อมูลแล้วทำการบันทึกเป็นตัวแปรใหม่	162
7.23	ขั้นตอนการสร้างตัวแปรใหม่จากข้อมูลเดิม	164
7.24	ขั้นตอนการคัดเลือกข้อมูล	165
7.25	ข้อมูลที่ไม่ได้ถูกเลือก	166
8.1	ขั้นตอนการสร้างตารางแจกแจงความถี่แบบทางเดียว	172
8.2	การกำหนดค่าเพิ่มเติมในคำสั่ง Frequencies ที่ปุ่ม Statistics	173
8.3	การกำหนดค่าเพิ่มเติมในคำสั่ง Frequencies ที่ ปุ่ม Charts	173
8.4	การกำหนดค่าเพิ่มเติมในคำสั่ง Frequencies ที่ ปุ่ม Format	174
8.5	ขั้นตอนการสร้างตารางแจกแจงความถี่แบบสองทาง	176
8.6	การกำหนดค่าเพิ่มเติมในคำสั่ง Crosstabs ที่ ปุ่ม Cells	176
8.7	ขั้นตอนการวิเคราะห์ค่าสถิติพื้นฐาน ด้วยคำสั่ง Descriptives	197
8.8	ขั้นตอนการวิเคราะห์ค่าสถิติพื้นฐาน ด้วยคำสั่ง Frequencies	199
8.9	ขั้นตอนการสร้างตัวแปรความพึงพอใจรวม	202
8.10	ขั้นตอนการวิเคราะห์ความพึงพอใจของนักท่องเที่ยวในภาพรวม	203
8.11	ขั้นตอนการวิเคราะห์ความพึงพอใจของนักท่องเที่ยวจำแนกตามเพศ	205
8.12	ขั้นตอนการวิเคราะห์คำถามแบบเลือกตอบได้หลายคำตอบ ขั้นที่ 1	209
8.13	ขั้นตอนการวิเคราะห์คำถามแบบเลือกตอบได้หลายคำตอบ ขั้นที่ 2	210
8.14	ขั้นตอนการวิเคราะห์คำถามแบบเลือกตอบได้หลายคำตอบจำแนกตามตัวแปร เชิงคุณภาพ	212
9.1	ตัวอย่างการประมาณค่าแบบจุดและแบบช่วง	202
9.2	เขตการยอมรับและปฏิเสธ H_0 ภายใต้โค้งปกติ	205

ภาพที่		หน้า
9.3	ความคลาดเคลื่อนจากการทดสอบสมมติฐาน	206
9.4	การทดสอบสมมติฐานแบบทางเดียว	207
9.5	การทดสอบสมมติฐานแบบสองทาง	207
9.6	ขั้นตอนการทดสอบการแจกแจงแบบปกติของข้อมูล	210
10.1	ขั้นตอนการทดสอบความสัมพันธ์แบบไคสแควร์	228
10.2	ขั้นตอนการทดสอบระดับของความสัมพันธ์	231
11.1	ความสัมพันธ์ของค่าสัมประสิทธิ์สหสัมพันธ์	240
11.2	ขั้นตอนการทดสอบสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน	244
11.3	สมการเชิงเส้นตรง	247
11.4	เส้นถดถอยที่กระชับที่สุด	248
11.5	ค่าสัมประสิทธิ์การถดถอย	250
11.6	ขั้นตอนการวิเคราะห์การถดถอยอย่างง่าย	254
11.7	ขั้นตอนการทดสอบระดับของความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตาม	259

สารบัญตาราง

ตารางที่		หน้า
2.1	การแจกแจงความถี่ของอายุของผู้ตอบแบบสอบถาม	15
2.2	การแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	18
2.3	แสดงจำนวนและร้อยละของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	19
2.4	แสดงจำนวนกลุ่มตัวอย่างผู้ตอบแบบสอบถามจำแนกตามเพศ	19
2.5	การหาค่าเฉลี่ยจากตารางแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	22
2.6	การหาค่ามัธยฐานจากตารางแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	25
2.7	การหาค่าฐานนิยมจากตารางแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	27
2.8	การหาค่าพิสัยจากตารางแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	30
2.9	การหาค่าส่วนเบี่ยงเบนควอไทล์จากตารางแจกแจงความถี่ของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	33
2.10	การหาค่าส่วนเบี่ยงเบนมาตรฐานของคะแนนสอบวิชาหลักสถิติสำหรับนักวิทยาศาสตร์คอมพิวเตอร์	37
2.11	จำนวนและร้อยละของห้องพักของโรงแรมและเกสต์เฮ้าส์ พ.ศ. 2557 จำแนกตามภาค	39
2.12	ร้อยละของประชากรอายุ 6 ปี ขึ้นไป ที่ใช้อินเทอร์เน็ต เปรียบเทียบตามกลุ่มอายุ พ.ศ. 2553-2557	39
2.13	ร้อยละของครัวเรือนที่ใช้อินเทอร์เน็ต จำแนกตามประเภทของอินเทอร์เน็ตและภาค	40
2.14	จำนวนประชากรอายุ 15 ปีขึ้นไป จำแนกตามสถานภาพแรงงาน	40
4.1	การแทนค่าตัวแปรสุ่มในการทดสอบหลอดไฟฟ้า	72
4.2	การแทนค่าตัวแปรสุ่มในการหยายแต้มที่ได้จากการโยนลูกเต๋า 1 ลูก	72
4.3	การแจกแจงความน่าจะเป็นของผลจากการลงทุนเปิดร้านกาแฟ	74
4.4	การแจกแจงความน่าจะเป็นของการโยนเหรียญ 1 เหรียญจำนวน 3 ครั้ง	75
4.5	ตัวอย่างของตัวแปรทวินาม	79
4.6	ตารางการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน	92
4.7	ตารางการแจกแจงความน่าจะเป็นแบบไคสแควร์	97

ตารางที่		หน้า
5.1	สัญลักษณ์แทนความหมายระหว่างค่าพารามิเตอร์และค่าสถิติ	104
5.2	ขนาดของกลุ่มตัวอย่างของทาโร ยามาเน่ ที่ระดับความเชื่อมั่น 95% และความคลาดเคลื่อนต่างๆ	108
5.3	ขนาดของกลุ่มตัวอย่างของเครซีและมอร์แกน	109
8.1	แสดงจำนวนและร้อยละของผู้ตอบแบบสอบถามทางด้านเพศ	175
8.2	แสดงจำนวนและร้อยละของผู้ตอบแบบสอบถามทางด้านตามระดับการศึกษา จำแนกตามเพศ	178
8.3	แสดงค่าสถิติต่างๆ ของข้อมูลทั่วไปเกี่ยวกับผู้ตอบแบบสอบถามทางด้านอายุ	180
8.4	แสดงค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของระดับความพึงพอใจของนักท่องเที่ยว ในการเที่ยวหาดสังวาลย์ ด้านการให้บริการของเจ้าหน้าที่ภายในสถานที่ ท่องเที่ยว โดยแสดงเป็นภาพรวม	186
8.5	แสดงค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของระดับความพึงพอใจของ นักท่องเที่ยวในการเที่ยวหาดสังวาลย์ ด้านการให้บริการของเจ้าหน้าที่ ภายในสถานที่ท่องเที่ยว จำแนกตามเพศ	189
8.6	แสดงค่าความถี่และร้อยละของการรับทราบข้อมูลเกี่ยวกับการท่องเที่ยวใน หาดสังวาลย์	193
8.7	แสดงค่าความถี่และร้อยละของการรับทราบข้อมูลเกี่ยวกับการท่องเที่ยวใน หาดสังวาลย์ จำแนกตามเพศ	196
8.8	สรุปสถิติบรรยายที่ใช้ในงานวิจัย	197
10.1	แสดงผลการทดสอบความสัมพันธ์ระหว่างอาชีพกับรายได้	232
11.1	แสดงผลการทดสอบความสัมพันธ์ระหว่างเกรดเฉลี่ยของนักศึกษา กับจำนวน ชั่วโมงในการอ่านหนังสือต่อวัน	245
11.2	แสดงผลการวิเคราะห์การถดถอยเพื่อหาความสัมพันธ์ระหว่างเกรดเฉลี่ยสะสม เมื่อสำเร็จการศึกษา กับคะแนนสอบคัดเลือกก่อนเข้าเรียน	260