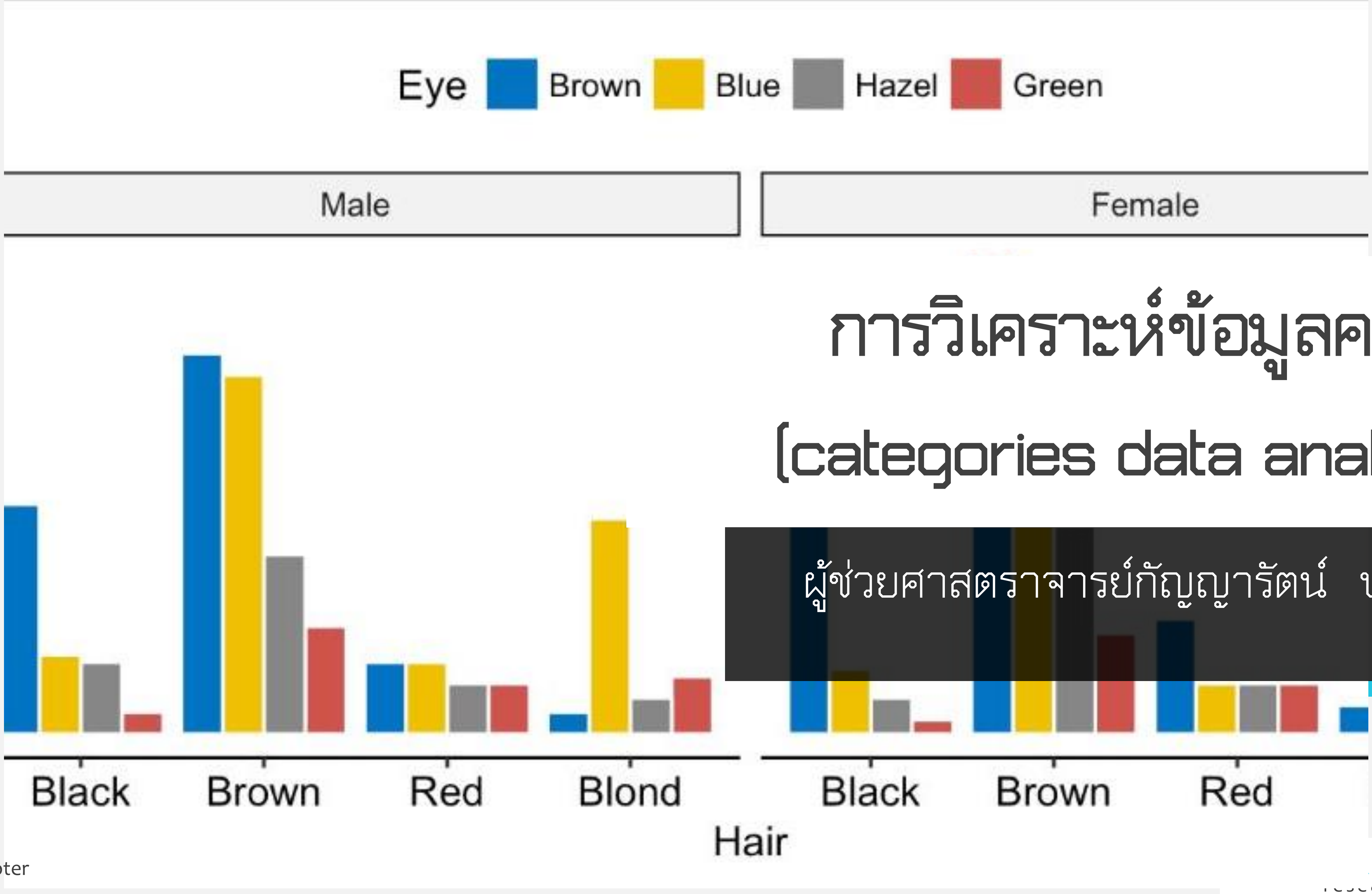

การวิเคราะห์ข้อมูลความถี่ (categories data analysis)

ผู้ช่วยศาสตราจารย์กัญญารัตน์ บุษบรรณ

การทดสอบความเป็นอิสระ

การทดสอบลักษณะที่สนใจสองลักษณะว่ามีความเป็นอิสระกันหรือไม่

- เพศ และผลการเรียนเป็นอิสระกันหรือไม่
- ระดับการศึกษาของบิดาและการสอบเข้าเรียนต่อในระดับมหาวิทยาลัยเป็นอิสระกันหรือไม่

ข้อมูลแจกแจงสองทาง

แถว	1	2	3	. . .	c	ผลรวม
1	O_{11}	O_{12}	O_{13}	. . .	O_{1c}	R_1
2	O_{21}	O_{22}	O_{23}	. . .	O_{2c}	R_2
.						
.						
.						
r	O_{r1}	O_{r2}	O_{r3}	. . .	O_{rc}	R_r
ผลรวม	C_1	C_2	C_3	. . .	C_c	n

สูตรที่ใช้ในการคำนวณ $E_{ij} = \frac{R_i C_j}{n}$

ตัวทดสอบสถิติสำหรับ
ข้อมูลแจกแจงสองทาง

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

$i = 1,2,3,\ldots,r$ และ $j = 1,2,3,\ldots,c$

เขตวิกฤต

$$\chi^2_{cal} \geq \chi^2_{\alpha,\langle (r-1)(c-1) \rangle}$$

TREY research

ขั้นตอนการทดสอบ

1. ตั้งสมมติฐาน

H_0 : ลักษณะทั้งสองเป็นอิสระต่อกัน หรือไม่มีความสัมพันธ์กัน

H_1 : ลักษณะทั้งสองไม่เป็นอิสระต่อกันหรือมีความสัมพันธ์กัน

2. บริเวณวิกฤต $\chi^2_{cal} \geq \chi^2_{\alpha,\langle(r-1)(c-1)\rangle}$

3. สถิติที่ใช้ทดสอบ $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$ $i=1,2,3,\ldots,r$ $j=1,2,3,\ldots,c$ และ $E_{ij}=\frac{R_iC_j}{n}$

4. สรุปผลการทดสอบ

TREY research

ตัวอย่างที่ 1 จากการศึกษาทัศนคติต่อรายการทีวีที่กำลังผลิตซึ่งนำเสนอความรุ่นแรง โดยสุ่มตัวอย่างที่เป็นผู้ใหญ่ 1200 คน เพื่อถามคำถาม ท่านคิดว่าความรุ่นแรงที่เสนอในทีวีกับอาชญากรรมเกี่ยวข้องกันใช่หรือไม่ได้ข้อมูลในตารางข้างล่าง จากข้อมูลสรุปได้หรือไม่ว่าทัศนคติกับเพศมีความสัมพันธ์ จงทดสอบความเป็นอิสระด้วยระดับความเชื่อมั่น 95%

เพศ	คำตอบ			รวม
	ใช่	ไม่ใช่	ไม่แน่ใจ	
ชาย	361	228	17	606
หญิง	433	141	20	594
รวม	794	369	37	1200

สมมติฐาน

H_0 : ทัศนคติกับเพศเป็นอิสระต่อกัน หรือไม่มีความสัมพันธ์กัน

H_1 : ทัศนคติกับเพศไม่เป็นอิสระต่อกันหรือมีความสัมพันธ์กัน

บริเวณวิกฤต $\chi^2_{cal} \geq \chi^2_{0.05,2}$

สถิติที่ใช้ทดสอบ $\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}} = 27.17$

เพศ	คำตอบ			รวม
	ใช่	ไม่ใช่	ไม่แน่ใจ	
ชาย	361	228	17	606
	(400.97)	(186.345)	(18.685)	
หญิง	433	141	20	594
	(393.03)	(182.655)	(18.315)	
รวม	794	369	37	1200

$\chi^2_{0.05,2} = 5.99$ และ 27.17>5.99

สรุป ปฏิเสธ H_0

นั่นคือ ทัศนคติกับเพศไม่เป็นอิสระกันหรือสัมพันธ์กัน ที่ระดับนัยสำคัญ 0.05

ตัวอย่าง 2. จากการสำรวจครอบครัวที่แต่งงานแล้ว 200 ครอบครัว แบ่งพิจารณาตามระดับการศึกษาของหัวหน้าครอบครัว และจำนวนบุตรเป็นดังนี้

ระดับการศึกษา \ จำนวนบุตร	0-1	2-3	>3	รวม
ประถมศึกษา	14	37	32	83
มัธยมศึกษา	19	42	17	78
มหาวิทยาลัย	12	17	10	39
รวม	45	96	59	200

จงทดสอบที่ระดับนัยสำคัญ 0.05 ว่าจำนวนบุตรและระดับการศึกษาของหัวหน้าครอบครัวสัมพันธ์กันหรือไม่

ขั้นตอนการทดสอบ

1.ตั้งสมมติฐาน

H_0 : จำนวนบุตรกับระดับการศึกษาของหัวหน้าครอบครัวเป็นอิสระต่อกัน หรือไม่มีความสัมพันธ์กัน

H_1 : จำนวนบุตรกับระดับการศึกษาของหัวหน้าครอบครัวไม่เป็นอิสระต่อกันหรือมีความสัมพันธ์กัน

2.บริเวณวิกฤต $\chi^2_{cal} \geq \chi^2_{0.05,4}$, $\chi^2_{0.05,4} = 9.49$

3.ตัวสถิติที่ใช้ทดสอบ

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}} = 7.46$$

4.สรุป ยอมรับ H_0 เนื่องจาก 7.46<9.49

ระดับการศึกษา \ **จำนวนบุตร**	0-1	2-3	>3	**รวม**
ประถมศึกษา	14 (18.675)	37 (39.88)	32 (24.885)	83
มัธยมศึกษา	19 (17.50)	42 (37.88)	17 (23.01)	78
มหาวิทยาลัย	12 (8.775)	17 (18.72)	10 (11.505)	39
รวม	45	96	59	200

จะได้ว่าจำนวนบุตรกับระดับการศึกษาของหัวหน้าครอบครัวเป็นอิสระกันหรือไม่สัมพันธ์ที่ระดับนัยสำคัญ 0.05

การวัดระดับความสัมพันธ์ระหว่างสองลักษณะ

เพื่อตรวจสอบดูว่าปัจจัยใดที่มีความสัมพันธ์สูงกว่ากัน

- ทัศนคติซึ่งมี 3 ระดับ มีความสัมพันธ์กับ เพศ อายุ ระดับการศึกษา ศาสนาหรือไม่
- ปัจจัยใดมีความสัมพันธ์สูงกว่ากัน

วิธีการ

1. วิธีของคราเมอร์ (Cramer)
2. วิธีของเพียร์สัน (Peason)

วิธีของคราเมอร์

สถิติที่ใช้วัด

$$v^2 = \frac{\chi^2}{n \cdot \min(r-1, c-1)}$$

วิธีของเพียร์สัน

สถิติที่ใช้วัด

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

ตัวอย่าง 3 ผู้จัดการโรงงานอุตสาหกรรมแห่งหนึ่งได้รับคำร้องเรียนจากคนงานให้โรงงานหยุดงานในวันเสาร์โดยไม่ต้องจ่ายค่าแรงให้คนงาน เพื่อให้ได้ข้อมูลประกอบในการตัดสินใจเขาได้สุ่มคนงานจำนวน 250 คน แล้วสอบถามความคิดเห็นเกี่ยวกับเรื่องการหยุดงานดังกล่าว เมื่อจำแนกตามเพศและอายุ ซึ่งแบ่งเป็น 2 ช่วงคือต่ำกว่า 35 ปีและตั้งแต่ 35 ปีขึ้นไปที่เห็นด้วย ไม่มีความเห็น และไม่เห็นด้วยกับคำร้องเรียนให้หยุดงานในวันเสาร์ปรากฏผลดังตาราง จงทดสอบว่าเพศและอายุมีอิทธิพลต่อความคิดเห็นในการหยุดงานในวันเสาร์ของคนงานหรือไม่ ถ้ามี เพศหรืออายุมีอิทธิพลต่อความคิดเห็นมากกว่ากัน

ตาราง แสดงความคิดเห็นจำแนกตามเพศ

เพศ	ความคิดเห็น			
	เห็นด้วย	ไม่มีความคิดเห็น	ไม่เห็นด้วย	รวม
ชาย	42	8	100	150
หญิง	74	6	20	100
รวม	116	14	120	250

ตาราง แสดงความคิดเห็นจำแนกตามอายุ

อายุ	ความคิดเห็น			
	เห็นด้วย	ไม่มีความคิดเห็น	ไม่เห็นด้วย	รวม
ต่ำกว่า 35	68	12	110	190
35 ปีขึ้นไป	48	2	10	60
รวม	116	14	120	250

ทดสอบความสัมพันธ์ระหว่างความคิดเห็นกับเพศ

1. ตั้งสมมติฐาน

   H_0 : ควมคิดเห็นของคนงานเกี่ยวกับการหยุดงานวันเสาร์ไม่ขึ้นกับเพศ

   H_1 : ความคิดเห็นของคนงานเกี่ยวกับการหยุดงานวันเสาร์ขึ้นกับเพศ

2.ตัวสถิติที่ใช้ทดสอบ $\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}=54.63$

3.ค่าวิกฤต $\chi^2_{0.05,2}=5.99$

4. สรุปผลการทดสอบ ปฏิเสธ H_0

5. อภิปรายผล ความคิดเห็นของคนงานเพศหญิงและชายเกี่ยวบการหยุดงานวันเสาร์แตกต่างกัน

ทดสอบความสัมพันธ์ระหว่างความคิดเห็นกับอายุ

สมมติฐาน H_0 : ความคิดเห็นของคนงานเกี่ยวกับการหยุดงานวันเสาร์ไม่ขึ้นกับอายุ

H_1 : ความคิดเห็นของคนงานเกี่ยวกับการหยุดงานวันเสาร์ขึ้นกับอายุ

ตัวสถิติที่ใช้ทดสอบ $$\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}=36.08$$

ค่าวิกฤต $\chi^2_{0.05,2}=5.99$

สรุปผลการ**ทดสอบ** ปฏิเสธ H_0

ความคิดเห็นของคนงานที่มีอายุต่ำกว่า 35 ปี และ ตั้งแต่ 35 ปีขึ้นไปเกี่ยวกับการหยุดงานวันเสาร์แตกต่างกัน

ตรวจสอบดูว่าปัจจัยใดที่มีความสัมพันธ์สูงกว่ากัน

วัดระดับความสัมพันธ์ระหว่างความคิดเห็นกับเพศ

วิธีของคราเมอร์

$$v^2 = \frac{\chi^2}{n \cdot \min(r-1, c-1)} = \frac{54.63}{250 \times 1} = 0.2185$$

วิธีของเพียร์สัน

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{54.63}{54.63 + 250}} = 0.4235$$

วัดระดับความสัมพันธ์ระหว่างความคิดเห็นกับอายุ

วิธีของคราเมอร์

$$v^2 = \frac{\chi^2}{n \cdot \min(r-1, c-1)} = \frac{36.08}{250 \times 1} = 0.1443$$

วิธีของเพียร์สัน

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}} = \sqrt{\frac{36.08}{36.08 + 250}} = 0.3551$$

ระดับความสัมพันธ์ระหว่างความคิดเห็นกับเพศสูงกว่าระดับความสัมพันธ์ระหว่างความคิดเห็นกับอายุ

เพศมีอิทธิพลมากกว่าอายุ

Thank You