เฉลย Assignment 13
MAC1304 ความน่าเป็นและสถิติ

หัวข้อ การทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวน ค่าสัดส่วน และการถดถอยเชิงเส้นเชิงเดียว **สัปดาห์ที่** 14 **คะแนนเต็ม** 10 คะแนน
ผู้สอน ผศ.ดร.ธนัชยศ จำปาหวาย สาขาวิชาคณิตศาสตร์ คณะครุศาสตร์ มหาวิทยาลัยราชภัฏสวนสุนันทา

---

1. การศึกษาผลสัมฤทธิ์ของวิชาคณิตศาสตร์เรื่องสถิติเบื้องต้น โดยใช้การจัดการเรียนรู้แบบ 5E ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ของโรงเรียนแห่งหนึ่ง โดยเลือกสุ่มตัวอย่างแบบแบ่งกลุ่ม (cluster random sampling) จำนวน 12 คน ตั้งสมมติฐานว่านักเรียนมีผลสัมฤทธิ์หลังเรียนสูงกว่าก่อนเรียน โดยมีผลคะแนนทดสอบก่อนเรียน-หลังเรียน (คะแนนเต็ม 20) ดังนี้

นักศึกษาคนที่	คะแนนก่อนเรียน	คะแนนหลังเรียน	นักศึกษาคนที่	คะแนนก่อนเรียน	คะแนนหลังเรียน
1	10	12	7	13	15
2	9	13	8	5	18
3	13	15	9	10	19
4	12	16	10	11	18
5	11	10	11	17	20
6	13	15	12	11	11

จงทดสอบสมมติฐานของการวิจัยดังกล่าวว่าเป็นจริงหรือไม่ ที่ระดับนัยสำคัญ 0.05 สมมติประชากรมีการแจกแจงปกติ

แนวคำตอบ กำหนดสมมติฐาน

$$\begin{aligned} H_0 &: \mu_D = 0 \\ H_1 &: \mu_D > 0 \end{aligned}$$

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.05$
สุ่มตัวอย่างขนาด $n = 12$ (ข้อมูลแบบคู่) จากประชากรปกติ

นักศึกษาคนที่	คะแนนก่อนเรียน	คะแนนหลังเรียน	d_i
1	10	12	2
2	9	13	4
3	13	15	2
4	12	16	4
5	11	10	-1
6	13	15	2
7	13	15	2
8	5	18	13
9	10	19	9
10	11	18	7
11	17	20	3
12	11	11	0

จะได้ว่า $\bar{D} = 3.92$ และ $S_d = 3.96$ โดยที่ $\nu = 12 - 1 = 11$ จะได้ว่า

$$t_{\text{คำนวณ}} = \frac{\bar{D} - d_0}{\frac{S_d}{\sqrt{n}}} = \frac{3.92 - 0}{\frac{3.96}{\sqrt{12}}} = 3.43$$

จะเห็นว่า

$$\text{P-value } = P(t > 3.43) = 0.0028 < 0.05 = \alpha$$

ดังนั้นปฏิเสธ H_0

สรุปได้นักเรียนมีผลสัมฤทธิ์หลังเรียนสูงกว่าก่อนเรียน ที่ระดับนัยสำคัญ 0.05 #

2. ชาวสวนทุเรียนกล่าวอ้างว่าน้ำหนักทุเรียนในสวนมีการแจกแจงปกติ และมีส่วนเบี่ยงเบนมาตรฐาน 1.5 กิโลกรัม พ่อค้าที่จะเหมาทุเรียนในสวนดังกล่าวจึงต้องทดสอบคำกล่าวอ้างของชาวสวนจึงสุ่มตัวอย่างทุเรียน 12 ลูกมีน้ำหนัก (กิโลกรัม) ดังนี้

5.5 4.5 5.8 6.0 6.5 5.0 4.0 3.5 6.0 4.7 5.4 7.0

จงทดสอบว่าความแปรปรวนประชากรจะมากกว่า 2.25 ที่ระดับนัยสำคัญ 0.10

แนวคำตอบ กำหนดสมมติฐาน

$$\begin{aligned} H_0 &: \sigma^2 = 2.25 \\ H_1 &: \sigma^2 > 2.25 \end{aligned}$$

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.10$
สุ่มตัวอย่างขนาด $n = 12$ จะได้ว่า $S^2 = 1.06$ และ $\nu = 12 - 1 = 11$ นั่นคือ

$$\chi^2_{\text{คำนวณ}} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(12-1)1.06}{2.25} = 5.18$$

วิธีที่ 1 เนื่องจาก $\chi^2_{\alpha,\nu} = \chi^2_{0.10,11} = 17.275$ บริเวณวิกฤตคือ $\chi^2 > 17.275$
จะเห็นว่า $\chi^2_{\text{คำนวณ}} = 5.18 < 17.275$ อยู่ในบริเวณยอมรับ ดังนั้นยอมรับ H_0

วิธีที่ 2 จะเห็นว่า

$$\text{P-value } = P(\chi^2 > 5.18) = 0.9221 > 0.10 = \alpha$$

ดังนั้นยอมรับ H_0

สรุปได้ว่าส่วนเบี่ยงเบนมาตรฐานของน้ำหนักทุเรียนในสวนเท่ากับ 1.5 กิโลกรัม ที่ระดับนัยสำคัญ 0.10 #

3. สุ่มคะแนนสอบย่อยของวิชาแคลคูลัส ๒ ของนักศึกษาจำนวน 8 คน ได้คะแนนดังนี้ 2, 3, 5, 5, 5, 6, 7, 8
สุ่มคะแนนสอบย่อยของวิชาความน่าจะเป็นและสถิติ ของนักศึกษาจำนวน 10 คน ได้คะแนนดังนี้ 1, 4, 3, 5, 6, 7, 8, 9, 10, 7
สมมติประชากรทั้งสองมีการแจกแจงปกติ จงทดสอบว่าความแปรปรวนของคะแนนวิชาแคลคูลัส ๒ ไม่แตกต่างจากวิชาความน่าจะเป็นและสถิติ ที่ระดับนัยสำคัญ 0.05

แนวคำตอบ กำหนดสมมติฐาน

$$\begin{aligned} H_0 &: \sigma_1^2 = \sigma_2^2 \\ H_1 &: \sigma_1^2 \neq \sigma_2^2 \end{aligned}$$

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.05$
สุ่มตัวอย่างขนาด (วิชาแคลคูลัส ๒) $n_1 = 8$ จะได้ว่า $S_1^2 = 3.84$ และ $\nu_1 = 8 - 1 = 7$
สุ่มตัวอย่างขนาด (วิชาความน่าจะเป็นและสถิติ) $n_2 = 10$ จะได้ว่า $S_2^2 = 7.78$ และ $\nu_2 = 10 - 1 = 9$
นั่นคือ

$$F_{\text{คำนวณ}} = \frac{S_1^2}{S_2^2} = \frac{3.84}{7.78} = 0.49$$

วิธีที่ 1 เนื่องจาก $f_{\frac{\alpha}{2},(\nu_1,\nu_2)} = f_{0.025,(7,9)} = 4.20$ และ

$$f_{1-\frac{\alpha}{2},(\nu_1,\nu_2)} = f_{0.975,(7,9)} = \frac{1}{f_{0.025,(9,7)}} = \frac{1}{4.82} = 0.207$$

บริเวณวิกฤตคือ $f < 0.207$ หรือ $f > 4.20$

จะเห็นว่า $F_{\text{คำนวณ}} = 0.49$ และ $0.207 < F_{\text{คำนวณ}} < 4.20$ อยู่ในบริเวณยอมรับ ดังนั้นยอมรับ H_0

วิธีที่ 2 จะเห็นว่า

$$\text{P-value } = P(f > 0.49) = 0.8203 > 0.025 = \frac{\alpha}{2}$$

ดังนั้นยอมรับ H_0

สรุปได้ว่าความแปรปรวนของคะแนนวิชาแคลคูลัส ๒ ไม่แตกต่างจากวิชาความน่าจะเป็นและสถิติ ที่ระดับนัยสำคัญ 0.05 #

4. ในมหาวิทยาลัยแห่งหนึ่งประมาณว่า 80% ของจำนวนนักเรียนผ่านการสอบสัมภาษณ์รอบ TCAS 1 จะยืนยันสิทธิ์เพื่อเข้าศึกษาต่อ ค่าประมาณนี้ถูกต้องหรือไม่ ที่ระดับนัยสำคัญ 0.06 ถ้าสุ่มตัวอย่างนักเรียนที่ผ่านการสอบสัมภาษณ์จำนวน 200 คน ยืนยันสิทธิ์จำนวน 150 คน

แนวคำตอบ กำหนดสมมติฐาน

$$\begin{aligned} H_0 &: \quad p = 0.8 \\ H_1 &: \quad p \neq 0.8 \end{aligned}$$

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.06$

สุ่มตัวอย่างขนาด $n = 200$ และ $X = 150$ จะได้ว่า $\hat{P} = \dfrac{150}{200} = 0.75$ นั่นคือ

$$Z_{\text{คำนวณ}} = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.75 - 0.8}{\sqrt{\frac{0.8(0.2)}{200}}} = -1.77$$

จะเห็นว่า

$$\text{P-value} = P(Z > 1.77) = 1 - P(Z < 1.77) = 1 - 0.9616 = 0.0384 > 0.03 = \frac{\alpha}{2}$$

ดังนั้นยอมรับ H_0

สรุปได้ว่าจำนวนนักเรียนผ่านการสอบสัมภาษณ์รอบ TCAS 1 จะยืนยันสิทธิ์เพื่อเข้าศึกษาต่อ 80% ที่ระดับนัยสำคัญ 0.06 #

5. เพื่อต้องการทราบความรู้สึกของนักศึกษาว่าพอใจหลักสูตรครุศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์ (๔ ปี) จึงหยั่งเสียงจากนักศึกษาชาย 100 คน มี 82 คนพอใจ และหยั่งเสียงจากนักศึกษาหญิง 100 คน มี 90 คนพอใจ จึงสรุปได้ว่านักศึกษาหญิงพอใจมากกว่านักศึกษาชายอยู่ 10% คุณเห็นด้วยหรือไม่ กำหนดระดับนัยสำคัญ 0.03

แนวคำตอบ กำหนดสมมติฐาน

$$\begin{aligned} H_0 &: \quad p_1 - p_2 = 0.1 \\ H_1 &: \quad p_1 - p_2 \neq 0.1 \end{aligned}$$

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.03$
สุ่มตัวอย่างขนาด (หญิง) $n_1 = 100$ และ $X_1 = 90$ จะได้ว่า $\hat{P}_1 = \frac{90}{100} = 0.90$
สุ่มตัวอย่างขนาด (ชาย) $n_2 = 100$ และ $X_2 = 82$ จะได้ว่า $\hat{P}_2 = \frac{82}{100} = 0.82$
นั่นคือ

$$Z_{\text{คำนวณ}} = \frac{(\hat{P}_1 - \hat{P}_2) - d_0}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} = \frac{(0.90 - 0.82) - 0.1}{\sqrt{\frac{0.9(0.1)}{100} + \frac{0.82(0.18)}{100}}} = -0.41$$

จะเห็นว่า

$$\text{P-value} = P(Z > 0.41) = 1 - P(Z < 0.41) = 1 - 0.3409 = 0.6591 > 0.015 = \frac{\alpha}{2}$$

ดังนั้นยอมรับ H_0

สรุปได้ว่านักศึกษาหญิงพอใจมากกว่านักศึกษาชายอยู่ 10% ที่ระดับนัยสำคัญ 0.03 #

6. สมมติมีการจัดการเรียนรู้ต่างกัน 2 วิธี คือแบบ Oline และ Onside ตัวอย่างสุ่มของนักนักศึกษากลุ่มละ 50 คน ถูกกำหนดให้เรียนคนละแบบ เกรดคะแนนสอบครั้งสุดท้ายถูกนำมาเพื่อเปรียบเทียบผลการจัดการเรียนรู้ ได้นำมาลงตารางการณ์จรดังต่อไปนี้

	เกรด					
การจัดการเรียนรู้	A	B	C	D	F	รวม
Online	8	12	20	5	5	50
Onside	9	25	10	4	2	50
รวม	17	37	30	9	7	100

จงทดสอบสมมติฐานว่าสัดส่วนของเกรดต่าง ๆ ไม่ขึ้นกับการจัดการเรียนรู้ ที่ระดับนัยสำคัญ 0.05

แนวคำตอบ กำหนดสมมติฐาน

H_0 : การจัดการเรียนรู้ไม่มีผลต่อเกรด (อิสระต่อกัน)
H_1 : การจัดการเรียนรู้มีผลต่อเกรด (ไม่อิสระต่อกัน)

ทดสอบที่ระดับนัยสำคัญ $\alpha = 0.05$
คำนวณค่าสถิติจากตัวอย่าง

	เกรด					
การจัดการเรียนรู้	A	B	C	D	F	รวม
Online	$o_{11} = 8$	$o_{12} = 12$	$o_{13} = 20$	$o_{14} = 5$	$o_{15} = 5$	$R_1 = 50$
Onside	$o_{21} = 9$	$o_{22} = 25$	$o_{23} = 10$	$o_{24} = 4$	$o_{25} = 2$	$R_2 = 50$
รวม	$C_1 = 17$	$C_2 = 37$	$C_3 = 30$	$C_4 = 9$	$C_5 = 7$	$N = 100$

จะได้ว่า

	A	B	C	D	F
Online	$e_{11} = \frac{R_1C_1}{N} = 8.5$	$e_{12} = \frac{R_1C_2}{N} = 18.5$	$e_{13} = \frac{R_1C_3}{N} = 15$	$e_{14} = \frac{R_1C_4}{N} = 4.5$	$e_{15} = \frac{R_1C_5}{N} = 3.5$
Onside	$e_{21} = \frac{R_2C_1}{N} = 8.5$	$e_{12} = \frac{R_2C_2}{N} = 18.5$	$e_{13} = \frac{R_2C_3}{N} = 15$	$e_{14} = \frac{R_2C_4}{N} = 4.5$	$e_{15} = \frac{R_2C_5}{N} = 3.5$

โดยมี $r = 2$ และ $c = 5$ จะเห็นว่า $\nu = (2-1)(5-1) = 4$ และ

$$\begin{aligned}
\chi^2_{\text{คำนวณ}} &= \sum_{i=1}^{2}\sum_{j=1}^{5}\frac{(o_{ij}-e_{ij})^2}{e_{ij}} \\
&= \frac{(8-8.5)^2}{8.5} + \frac{(9-8.5)^2}{8.5} + \frac{(12-18.5)^2}{18.5} + \frac{(25-18.5)^2}{18.5} + \frac{(20-15)^2}{15} + \frac{(10-15)^2}{15} + \\
&\quad \frac{(5-4.5)^2}{4.5} + \frac{(4-4.5)^2}{4.5} + \frac{(5-3.5)^2}{3.5} + \frac{(2-3.5)^2}{3.5} \\
&= 9.36
\end{aligned}$$

จะเห็นว่า

$$\text{P-value } = P(\chi^2 > 9.36) = 0.0527 > 0.05 = \frac{\alpha}{2}$$

ดังนั้นยอมรับ H_0
สรุปได้ว่าสัดส่วนของเกรดต่าง ๆ ไม่ขึ้นกับการจัดการเรียนรู้ ที่ระดับนัยสำคัญ 0.05 #

7. พิจารณาข้อมูล x และ y ดังนี้

x	-3	-1	0	2	3
y	0	k	$k+3$	$k+4$	$k+6$

เมื่อ k เป็นค่าคงที่ ถ้าสัมประสิทธิ์การถดถอยของ $\hat{y} = a + bx$ เท่ากับ $b = 2.5$ จงหา a

แนวคำตอบ จะเห็นสัมประสิทธิ์การถดถอย $b = 2.5$ พิจารณาตารางต่อไปนี้

	x	y	xy	x^2
	-3	0	0	9
	-1	k	$-k$	1
	0	$k+3$	0	0
	2	$k+4$	$2k+8$	4
	3	$k+6$	$3k+18$	9
รวม	1	$4k+13$	$4k+26$	23

จะเห็นว่า $n = 5$ และ

$$
\begin{aligned}
2.5 = b = \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) - n\sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n\sum_{i=1}^{n} x_i^2} &= \frac{1(4k+13) - 5(4k+26)}{1^2 - 5(23)} = \frac{16k+117}{114} \\
2.5(114) &= 16k + 117 \\
285 &= 16k + 117 \\
168 &= 16k \\
10.5 &= k
\end{aligned}
$$

ดังนั้น

$$
a = \frac{\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i}{n} = \frac{(4k+13) - 2.5(1)}{5} = \frac{4(10.5) + 13 - 2.5}{5} = 10.5 \quad \#
$$

8. ปริมาณการส่งออกทุเรียนสดของประเทศไทย แสดงดังตาราง

ปี พ.ศ.	2554	2555	2556	2557	2558	2559	2560
ปริมาณส่งออก (แสนตัน)	2.71	3.51	3.67	3.69	3.58	4.03	k

* กรมส่งเสริมการค้าระหว่างประเทศ กระทรวงพานิชย์

ถ้าใช้สมการถดถอยจะได้ผลคาดการณ์ว่าปี 2564 ปริมาณส่งออกทุเรียนสดของประเทศไทยประมาณ 5.61 แสนตัน จงหาปริมาณส่งออกทุเรียนสดของประเทศไทยในปี 2560 หรือค่า k ในหน่วยแสนตัน

แนวคำตอบ ให้ $x = -3, -2, -1, \ldots$ แทนปี พ.ศ. 2554, 2555, 2556, ... ตามลำดับ
และ y แทนปริมาณการส่งออกทุเรียนสดของประเทศไทย หน่วยเป็นแสนตัน แสดงตารางได้ดังนี้

	x	y	xy	x^2
	-3	2.71	-8.13	9
	-2	3.51	-7.02	4
	-1	3.67	-3.67	1
	0	3.69	0	0
	1	3.58	3.58	1
	2	4.03	8.06	4
	3	k	$3k$	9
รวม	0	$21.19+k$	$-7.18+3k$	28

สมการถดถอยของความสัมพันธ์ดังกล่าวคือ $\hat{y} = a + bx$ เนื่องจากปี 2564 คือ $x = 7$ และ $\hat{y} = 5.61$ นั่นคือ

$$a + 7b = 5.61$$

จาก จะเห็นว่า $n = 7$ และ

$$b = \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) - n\sum_{i=1}^{n} x_i y_i}{\left(\sum_{i=1}^{n} x_i\right)^2 - n\sum_{i=1}^{n} x_i^2} = \frac{0(21.19+k) - 7(-7.18+3k)}{0^2 - 7(28)} = \frac{-7.18+3k}{28}$$

และ

$$a = \frac{\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i}{n} = \frac{(21.19+k) - b(0)}{7} = \frac{21.19+k}{7}$$

จะได้ว่า

$$\begin{aligned}
\left(\frac{21.19+k}{7}\right) + 7\left(\frac{-7.18+3k}{28}\right) &= 5.61 \\
4(21.19+k) + 7(-7.18+3k) &= 5.61(28) \\
84.76 + 4k - 50.26 + 21k &= 157.08 \\
25k &= 122.58 \\
k &= 4.90
\end{aligned}$$

ดังนั้นปริมาณส่งออกทุเรียนสดของประเทศไทยในปี 2560 เท่ากับ 4.90 แสนตัน #