

ตัวอย่างข้อสอบ

ภาค 1/2527

- คำสั่ง
- 1) ในการทดสอบสมมติฐาน ให้แสดง 6 ขั้นตอนของการทดสอบให้ครบ (การให้คะแนนจะแบ่งให้ตามขั้นต่าง ๆ)
 - 2) นักศึกษาอาจเลือกใช้สูตรที่หมำกกับข้อสอบชุดนี้ได้
 - 3) ข้อสอบชุดนี้มีทั้งหมด 5 ข้อ ข้อละ 20 คะแนน

1. จงตอบคำถามหรือเติมข้อความให้ถูกต้องและได้ความ (ไม่ต้องลอกโจทย์)
 - 1) เราใช้ตัวสถิติ F ในการทดสอบสมมติฐานเกี่ยวกับอะไรบ้าง บอกมา 2 สมมติฐาน
 - 2) ถ้า $Z_i \sim \text{NID}(0, 1)$, $i = 1, \dots, n$ แล้ว $Y = \sum_{i=1}^n Z_i^2$ จะมี การแจกแจง แบบใด ถ้ามี degrees of freedom ให้บอกด้วย
 - 3) ในการทดสอบการเท่ากันของ k binomial proportions นั้นคือทดสอบ $H_0: p_1 = p_2 = \dots = p_k$, $k > 2$ เราใช้ตัวสถิติอะไรในการทดสอบ
 - 4) ถ้า $z \sim \text{N}(0, 1)$ และ $X^2 \sim \chi_n^2$ แล้ว $Z/\sqrt{X^2/n}$ จะมี การแจกแจงแบบใด ถ้ามี degrees of freedom ให้บอกด้วย
 - 5) จาก Model 1 : $\mu_{y|x} = \beta_0 + \beta_1 x$, x และ
Model 2 : $\mu_{y|x_1, x_2, x_3} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$
เราเรียก Model 1 ว่า Simple regression model ส่วน Model 2 เราเรียกว่าอะไร
 - 6) จาก Model 1 ในข้อ 5) เราเรียก Y ว่าตัวแปรตาม (dependent variable) เราเรียก x และ β_1 ว่าอะไร
 - 7) r (sample correlation coefficient) จะเป็นตัวประมาณค่า degree ของความสัมพันธ์ของตัวแปรเชิงสุ่ม 2 ตัวคือ X และ Y ก็คือเมื่อ X และ Y มีความสัมพันธ์กันในรูป _____ และในการทดสอบ $H_0: \rho = 0$ เราใช้ตัวสถิติในการทดสอบซึ่งค่าคำนวณ

ได้จากสูตร $\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ ตัวสถิติดังกล่าวจะมี การแจกแจง เป็น __ distribution ที่มี degrees of freedom = _____ (ถ้ามี)

- 8) วิธีวิเคราะห์ทางสถิติวิธีหนึ่งซึ่งรวมวิธีวิเคราะห์ความแปรปรวนและวิธีการวิเคราะห์การถดถอยเข้าด้วยกัน วิธีดังกล่าวมีชื่อเรียกว่าอะไร
- 9) ถ้าหา 95% confidence interval ของ σ_1^2/σ_2^2 ได้ (0.295, 0.875) เราจะใช้ confidence interval นี้เพื่อช่วยสรุปผลการทดสอบ $H_0 : \sigma_1^2 = \sigma_2^2, H_1 = \sigma_1^2 \neq \sigma_2^2$ ที่ $\alpha = .05$ ว่าอย่างไร
- 10) ถ้าต้องการทดสอบว่า นิสัยการสูบบุหรี่ขึ้นอยู่กับกลุ่มอายุของผู้สูบบุหรี่หรือไม่ โดยแบ่งระดับการสูบบุหรี่เป็น 3 ระดับ และแบ่งกลุ่มอายุเป็น 4 กลุ่มอายุ ในการจำแนกข้อมูลที่ได้เก็บมาได้จากตัวอย่างเราจะสร้างตารางความถี่แบบ 2 ทางซึ่งมีขนาด _____ เราเรียกการทดสอบนี้ว่าอะไร การทดสอบนี้ใช้ตัวสถิติ chi-square ในการทดสอบ ซึ่งมี d.f. = _____

2. ให้เลือกทำ ก) หรือ ข) เพียงข้อเดียว

- ก) โยนเหรียญ 1 อันจนกว่าจะได้หัว จดจำนวนครั้ง (X) ที่ต้องโยนจนได้หัว 1 ครั้งไว้ หลังจากทำการทดลอง 256 ครั้ง ผลปรากฏดังนี้

x	1	2	3	4	5	6	7	8	รวม
จำนวนครั้ง	136	60	34	12	9	1	3	1	256 = 2^8

จงทดสอบสมมติฐานที่ $\alpha = .05$ ว่าการแจกแจงของ X เป็น Geometric distribution ที่มี $p = \frac{1}{2}$ หรือไม่ กำหนดค่าจำนวนของตัวสถิติที่ใช้ในการทดสอบ = 2.5714 ให้แสดงขั้นตอนในการทดสอบอย่างละเอียด

หมายเหตุ ถ้า X -Geometric ($p = \frac{1}{2}$)

$$\Pr [X = x] = p^{x-1} q, \quad x = 1, 2, \dots$$

$$= \left(\frac{1}{2}\right)^x = \frac{1}{2^x}, \quad x = 1, 2, \dots$$

ข) ในการศึกษาเพื่อดูว่าลักษณะการสูบบุหรี่ของนักศึกษาในปีต่าง ๆ แตกต่างกันหรือไม่ ได้สุ่มนักศึกษาปี 1 มา 50 คน ปี 2 มา 40 คน ปี 3 มา 60 คน และปี 4 มา 50 คน จัดจำแนกตามนิสัยการสูบบุหรี่ได้ผลดังนี้

ปีการศึกษา	นิสัยการสูบบุหรี่			รวม
	สูบน้อย	สูบปานกลาง	สูบบ่อย	
1	21	12	17	50
2	13	8	19	40
3	13	18	29	60
4	3	22	25	50
รวม	50	60	90	200

จงทดสอบที่ $\alpha = .05$ ว่าลักษณะการสูบบุหรี่ของนักศึกษาทั้ง 4 ปีไม่ต่างกัน กำหนดค่าจำนวนของตัวสถิติที่ใช้ในการทดสอบ = 21.1926 ให้แสดงขั้นตอนในการทดสอบอย่างละเอียด

3. นักศึกษาหญิง 16 คน ชาย 25 คน ทำข้อสอบมาตรฐานวิชาคณิตศาสตร์เบื้องต้น ผลปรากฏว่า นักศึกษาหญิงทำคะแนนเฉลี่ยได้ 78 คะแนน และ s.d. = 8 คะแนน ส่วนนักศึกษาชายทำคะแนนเฉลี่ยได้ 82 คะแนน และ s.d. = 9 คะแนน

ถ้า σ_1^2 เป็น population variance ของคะแนนของนักศึกษาชายทั้งหมดที่จะทำข้อสอบนี้ และ σ_2^2 เป็น population variance ของคะแนนของนักศึกษาหญิงทั้งหมดที่จะทำข้อสอบนี้ และถ้าทั้งสอง populations ต่างก็มีการกระจายเป็น normal

- 1) จงหา 95% confidence interval ของ σ_1^2
- 2) จงทดสอบที่ $\alpha = .05$, $H_0 : \sigma_2^2 = 49$, $H_1 : \sigma_2^2 > 49$
- 3) จงทดสอบที่ $\alpha = .02$, $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$

หมายเหตุ ให้แทนค่าสูตรในทุกขั้นตอนให้ชัดเจน ถ้ามีเวลาจึงคำนวณคำตอบของแต่ละข้อ คะแนนคำตอบสุดท้ายจะไม่มาก

4. จากข้อมูลที่กำหนดให้

x	1	1	1	2	3	3	4	5	5
y	9	7	8	10	15	12	19	24	21

$$\Sigma x_i = 25, \Sigma y_i = 125, \bar{x} = 2.78, S_{xx} = \Sigma (x_i - \bar{x})^2 = 27.56$$

$$\Sigma x_i^2 = 91, \Sigma y_i^2 = 2041, \bar{y} = 13.89, S_{yy} = \Sigma (y_i - \bar{y})^2 = 304.89$$

$$\Sigma x_i y_i = 426$$

$$S_{xy} = \Sigma (x_i - \bar{x})(y_i - \bar{y}) = 78.78$$

Assumed model : $\mu_{y|x} = \beta_0 + \beta_1 x$

1) ถ้า $b_1 = 2.86$ จงหา b_0 และเขียน prediction equation

2) กำหนด $SSR = b_1 S_{xy} = 225.31$

จงสร้างตาราง ANOVA ที่สมบูรณ์เพื่อใช้ในการทดสอบที่ $\alpha = .05$

2.1) $H_0 : \beta_1 = 0$ (กำหนดค่าของตัวสถิติในการทดสอบให้ = 19.82) และ

2.2) $H_0 : \text{ไม่มี lack of fit}$ (กำหนดค่าของตัวสถิติในการทดสอบให้ = 8.31)

หมายเหตุ ให้คิดทศนิยม 2 ตำแหน่ง

5. โรงงานผลิตสินค้าชนิดหนึ่งเข้าเครื่องจักรมาใหม่ 4 เครื่องจากโรงงานผลิตเครื่องจักร 4 โรงงาน ต้องการทราบว่าเครื่องจักรทั้ง 4 จะผลิตสินค้าได้ด้วยความเร็วเฉลี่ยแตกต่างกันหรือไม่ โรงงานจึงนับปริมาณสินค้า (ชิ้น) ที่เครื่องจักรแต่ละเครื่องผลิตได้ในช่วงเวลาครึ่งละ 1 ชั่วโมง ได้ข้อมูลดังนี้

ให้ x_{ij} เป็นปริมาณสินค้าที่เครื่องจักรที่ i ผลิตได้ในเวลาที่ j

$$i = 1, \dots, 4 \text{ และ } j = 1, \dots, 5$$

	เครื่องจักร				
	1	2	3	4	
	8	7	12	10	
	10	5	9	9	
	7	10	13	8	
	14	9	12	12	
	11	9	14	16	
Total : T_i	50	40	60	45	$G = 195$
Mean : \bar{x}_i	10	8	12	9	$\bar{x} = 9.75$
T_i^2	2500	1600	3600	2025	$\Sigma T_i^2 = 9725$
s_i^2	7.5	4	3.5	5	$\Sigma s_i^2 = 20$

$$G^2 = 38025$$

$$\sum_{i=1}^4 \sum_{j=1}^5 x_{ij}^2 = 1995$$

สมมติว่าข้อมูลที่ได้มาจากเครื่องจักรแต่ละเครื่องเป็นตัวอย่างสุ่มจาก Normal population $N(\mu_i, \sigma^2)$, $i = 1, \dots, 4$ ตามลำดับ

1) จงทดสอบที่ $\alpha = .05$ $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

กำหนดบางส่วนของตาราง ANOVA ให้ จงเติมตารางให้สมบูรณ์เพื่อใช้ในการทดสอบ

ANOVA

S.V.	d.f.	S.S.	M.S.	f
Treatments				
Error	16		3.125	
Total				

2) ถ้า reject H_0 จงใช้วิธี Duncan's new multiple range test เพื่อตรวจสอบว่ามี population mean ใดบ้างที่ต่างกันอย่างมีนัยสำคัญ โดยกำหนดค่า r_p จากตารางและค่า r_p ที่คำนวณได้ให้ดังนี้

$$\text{ที่ } \alpha = .05, (\text{d.f. ของ SSE}) = 16, n_1 = n_2 = \dots = n_4 = 5 = n$$

$$R_p = r_p \sqrt{\frac{\text{MSE}}{n}} = .79 r_p$$

p	2	3	4
r_p	2.9988	3.144	3.235
R_p	2.3688	2.484	2.556

จงเลือกใช้สูตรตามต้องการ (บางสูตรที่ต้องใช้อาจไม่มีในรายการสูตรนี้)

$$1) (1-\alpha) 100\% \text{ C.I. ของ } \left(\frac{(n-1)s^2}{\chi_{n-1, \alpha}^2}, \frac{(n-1)s^2}{\chi_{n-1, 1-\alpha/2}^2} \right)$$

$$2) (1-\alpha) 100\% \text{ C.I. ของ } \sigma_1^2/\sigma_2^2 \text{ do } \left(\frac{s_1^2}{s_2^2} \frac{1}{f_{(v_1, v_2), \frac{\alpha}{2}}}, \frac{s_1^2}{s_2^2} f_{(v_1, v_2), \frac{\alpha}{2}} \right)$$

$$3) \chi_c^2 = (n-1)s^2, \chi_c^2 = \sum_{ei} \frac{(o_i - e_i)^2}{e_i}, \chi_c^2 = \sum_{ij} \sum_{ij} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$4) SST = \sum y_i^2 - (\sum y_i)^2/n, SSR = b_1 S_{xy}$$

$$SS (p.e.) = \sum_{i=1}^k SS_i = \sum_{i=1}^k \left(\sum_{j=1}^m (y_{ij} - \bar{y}_i)^2 \right)$$

$$5) SST = \sum_{i=1}^l \sum_{j=1}^{n_i} x_{ij}^2 - \frac{G^2}{N}, N = \sum_{i=1}^l n_i$$

$$SS_{tr} = \sum_{i=1}^l \frac{T_i^2}{n_i} - \frac{G^2}{N}, R_p = r \sqrt{\frac{MSE}{1-\alpha}}$$

เฉลยข้อสอบภาค 1/2527

1.
 - 1) ก. $H_0: \mu_1 = \mu_2 = \dots = \mu_k$
U. H.: $\sigma_1^2 = \sigma_2^2$
 - 2) Chi square distribution ที่มี d.f. = n
 - 3) Chi square
 - 4) t distribution ที่มี d.f. = n
 - 5) Multiple regression model
 - 6) x คือตัวแปรอิสระ และ β_1 คือสัมประสิทธิ์การถดถอย
 - 7) เชนเส้นตรง, t distribution ที่มี d.f. = n-2
 - 8) วิเคราะห์ความแปรปรวนร่วม
 - 9) สรุปว่าปฏิเสธ H_0 ที่ $\alpha = .05$
 - 10) (3×4) , การทดสอบความเป็นอิสระ, d.f. = $(3-1)(4-1) = 6$

2.
 - ก) ดูแบบฝึกหัดข้อ 3.18 ของคู่มือนี้
 - ข) ดูแบบฝึกหัดข้อ 3.31 ของคู่มือนี้

3.
 - 1) ดูแบบฝึกหัดข้อ 2.9
 - 2) $H_0: \sigma_2^2 = 49$
 $H_1: \sigma_2^2 > 49$
 $\alpha = .05$
CR : $X^2 > \chi^2_{15, .05} = 24.996$
 $\chi_c^2 = \frac{(n_2-1)s_2^2}{49} = \frac{15 (64)}{49} = 19.59$
 $\therefore \chi_c^2 < 24.996$ เราไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = .05$

$$3) H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1 = \sigma_1^2 \neq \sigma_2^2$$

$$a = .02$$

$$CR : F > f_{(24,15),.01} = 3.21 \text{ และ } F < f_{(24,15),.99} = \frac{1}{f_{(15,24),.01}} = \frac{1}{2.89} = 1.346$$

$$fc = \frac{s_1^2}{s_2^2} = \frac{81}{64} = 1.2656$$

$\therefore fc < 3.21$ เราไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = .02$

4. ดูแบบฝึกหัดข้อ 5.4 ประกอบด้วย (ค่าที่ได้ต่างกันเพราะใช้จำนวนทศนิยมไม่เท่ากัน)

1) $b_1 = 2.86$ (กำหนดให้)

$$b_0 = \bar{y} - b_1 \bar{x} = 13.89 - (2.86)(2.78) = 5.94$$

$$\therefore \hat{y} = 5.94 + 2.86x$$

2) กำหนด $SSR = 225.31$

$$SST = S_{yy} = 304.89$$

$$\therefore SSE = 304.89 - 225.31 = 79.58$$

ทว่า $SS(p.e.) = 11$ (ดูข้อ 5.4)

$$\therefore SS(lof) = SSE - SS(p.e.) = 79.58 - 11 = 68.58$$

ANOVA

S.V.	d.f.	S.S.	MS	f
Regression	1	225.31	225.31	$f_1 = \frac{225.31}{11.37} = 19.82$
Error	7	79.58	11.37	
Lack of fit	3	68.58	22.86	$f_2 = \frac{22.86}{2.75} = 8.31$
Pure error	4	11	2.75	
Total	8	304.89		

2.1 $H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

$\alpha = .05$

CR : $F > f_{(1,7),.05} = 5.59$

$f_c = f_1 = 19.82$ (กำหนดให้)

$\therefore f_c > 5.59$ เราปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือ X เป็นเหตุของผล Y

2.2 H_0 : ไม่มี lack of fit

H_1 : มี lack of fit

$\alpha = .05$

CR : $F > f_{(3,4), .05} = 6.59$

$f_c = f_2 = 8.31$ (กำหนดให้)

$\therefore f_c > 6.59$ เราปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือเราต้องทำการปรับปรุงตัวแบบที่จะแสดงความสัมพันธ์ระหว่าง X และ Y

5. 1. $H_0: \mu_1 = \mu_2 = \dots = \mu_4$

H_1 : มี mean อย่างน้อย 1 คู่ไม่เท่ากัน

$\alpha = .05$

CR : $F > f_{(3,16), .05} = 3.24$

ANOVA

S.V.	d.f.	SS	MS	f
Treatments	3	43.75	14.5833	$f_c = 4.667$
Error	(16)	50.00	(3.125)	
Total	19	93.75		

$\therefore f_c = 4.667 > 3.24$ เราปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือเครื่องจักรทั้ง 4 มีความเร็วเฉลี่ยในการผลิตไม่เท่ากัน

2) $n_1 = \dots = n_4 = 5 = n$

d.f. ของ error = 16

$R_p = .79 r_p$

p	2	3	4
r_p	2.998	3.144	3.235
R_p	2.368	2.484	2.556

ii Duncan's new multiple range test

1) เรียงลำดับ \bar{x} จากน้อยไปหามาก

\bar{x}_2	\bar{x}_4	\bar{x}_1	\bar{x}_3
8	9	10	12

2) $\bar{x}_3 - \bar{x}_2 = 4 > R_4 = 2.556$ นั่นคือ $\mu_3 \neq \mu_2$

$\bar{x}_3 - \bar{x}_4 = 3 > R_3 = 2.484$ นั่นคือ $\mu_3 \neq \mu_4$

$\bar{x}_3 - \bar{x}_1 = 2 < R_2 = 2.368$ นั่นคือ $\mu_3 = \mu_1$

$\bar{x}_1 - \bar{x}_2 = 2 < R_3 = 2.484$ นั่นคือ $\mu_1 = \mu_2$ และ $\mu_1 = \mu_4$ ด้วย

3. สรุป

\bar{x}_2	\bar{x}_4	\bar{x}_1	\bar{x}_3
8	9	10	12

นั่นคือ $\mu_2 = \mu_4 = \mu_1, \mu_1 = \mu_3$ ส่วนคู่อื่น ๆ ต่างกันหมด