ตัวอย่างข้อสอบ

ภาค 2/2526

คำสั่ง 1) ในการทดสอบสมมติฐานให้แสดง 6 ขั้นของการทดสอบให้ครบ (การให้คะแนนจะแบ่งให้ตามขั้นต่าง ๆ)

2) นักศึกษาอาจเลือกใช้สูตรที่ให้มากับข้อสอบชุดนี้ได้

3) ข้อสอบชุดนี้มีทั้งหมด 5 ข้อ ให้ทำทุกข้อ

1. จงตอบคำถามหรือเติมข้อความให้ถูกต้อง และได้ความชัดเจน (ไม่ต้องลอกโจทย์)

1) เราใช้ตัวสถิติ F ในการทดสอบสมมติฐานเกี่ยวกับอะไรบ้าง จงบอกมา 2 อย่าง

2) วิธีการประมาณค่า (Estimation) เราประมาณค่าของอะไร และการประมาณค่ามีกี่วิธีอะไรบ้าง

3) ถ้าเราใช้ $\bar{X}_1 - \bar{X}_2$ เป็นตัว estimator ของ $\mu_1 - \mu_2$ ในการประมาณค่า $\mu_1 - \mu_2$ ในรูป $a < \mu_1 - \mu_2 < b$, การหาค่า a และ b เราต้องทราบ

ก) ....ของ $\mu_1 - \mu_2$ และ

ข) ....ของ $\bar{X}_1 - \bar{X}_2$

4) การทดสอบสมมติฐานจะเป็น One-tailed test หรือ Two-tailed test เรายึด H_0 หรือ H_1 เป็นหลัก

5) ให้ μ_1 และ μ_2 เป็นรายได้ของคนที่จบปริญญาตรี และของคนที่ไม่จบปริญญาตรีตามลำดับ ถ้าเราต้องการจะทดสอบเพื่อดูว่ารายได้เฉลี่ยของคนที่จบปริญญาตรีจะมากกว่าคนที่ไม่จบปริญญาตรีหรือไม่ เราจะต้อง H_0 : ....และ H_1 : .... และเราเรียกการทดสอบนี้ว่าเป็นการทดสอบแบบ..........-tailed Test

6) ถ้าสุ่มตัวอย่างขนาด n จาก Normal population ที่มี variance $= \sigma^2$ ในการทดสอบ $H_0 : \sigma^2 = \sigma_0^2$ เราใช้ตัวสถิติทดสอบ (Test statistic) คือ................(บอกสูตรด้วย) และถ้า H_0 นี้เป็นจริง ตัวสถิตินี้จะมี การแจกแจงเป็น ....distribution ที่มี df. = ...(ถ้ามี)

7) วิธีการแบ่ง Total variation (SST) ของ variable ที่สนใจออกเป็นส่วน ๆ ตามที่มาของมัน เพื่อดูอิทธิพลของที่มาของส่วนที่สนใจ เราเรียกวิธีการนี้ว่า..........

8) ในการทดสอบสารรูปสนิทดี (Goodness of fit test) เราทดสอบเกี่ยวกับอะไรบ้าง

9) ใน Regression Analysis การทดสอบ $H_0 : \beta_1 = 0, H_1 : \beta_1 \neq 0$ เราใช้ตัวสถิติทดสอบ ซึ่งมีสูตร $b_1/s(b_1)$ ในการทดสอบ ถ้า H_0 เป็นจริง ตัวสถิตินี้จะมีการแจกแจงเป็น..........distribution ที่มี df. = ...(ถ้ามี) การทดสอบนี้เป็นการทดสอบแบบ................-tailed test

10) ถ้าต้องการทดสอบว่า นิสัยการสูบบุหรี่ขึ้นอยู่กับเพศของผู้สูบหรือไม่ โดยแบ่งระดับการสูบบุหรี่เป็น 3 ระดับ คือสูบน้อย สูบปานกลาง และสูบมาก จงสร้าง Contingency Table เพื่อใช้ในการแสดงข้อมูลที่เก็บมาได้จากตัวอย่าง และเราเรียกการทดสอบนี้ว่าอะไร

2. โรงงานประกอบรถยนต์โรงงานหนึ่ง จะต้องตัดสินใจว่าจะใช้ยางรถยนต์ชนิด A หรือชนิด B กับรถยนต์ใหม่ที่ผลิตขึ้น เพื่อช่วยการตัดสินใจ โรงงานได้ทำการทดลองโดยสุ่มยางชนิด A มา 10 เส้น ยางชนิด B มา 9 เส้นใช้จนหมดอายุ ผลปรากฏว่า

ยางชนิด A : ระยะทางที่วิ่งได้เฉลี่ย $x_1 = 23600$ ไมล์ และ $s_1 = 3200$ ไมล์

ยางชนิด B : ระยะทางที่วิ่งได้เฉลี่ย $x_2 = 24800$ ไมล์ และ $s_2 = 3700$ ไมล์

สมมุติว่าอายุการใช้งาน (คิดเป็นระยะทางที่วิ่งได้) ของยางทั้ง 2 ชนิดต่างก็มีการกระจายเป็นNormal distribution ที่มีอายุการใช้งานเฉลี่ยเป็น μ_1 และ μ_2 มี variance σ_1^2 และ σ_2^2 ตามลำดับ โรงงานต้องการทดสอบ $H_0 : \mu_1 = \mu_2$ แต่จะตรวจสอบดูก่อนว่า $\sigma_1^2 = \sigma_2^2$ หรือไม่ จึงขอให้นักศึกษาช่วยวิเคราะห์ข้อมูลโดย

1) ทดสอบที่ $\alpha = .05, H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 < \sigma_2^2$

2) จากผลการทดสอบใน 1) จงแสดงวิธีการทดสอบที่ $\alpha = .05, H_0 : \mu_1 = \mu_2, H_1 : \mu_1 \neq \mu_2$ (บอกสูตรและแทนค่าสูตรเท่าที่จะทำได้จากข้อมูลที่กำหนดให้ ไม่ต้องคำนวณค่าสำเร็จ)

3) จงหา 90% confidence interval ของ σ_1/σ_2 (แทนค่าในสูตรให้ครบไม่ต้องหาคำตอบ)

3. (ข้อสอบข้อ 7 ภาค 1/2526)

4. กำหนดข้อมูลให้ต่อไปนี้

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14
x_i	2.1	2.6	2.7	2.7	2.7	3.0	3.0	3.0	3.1	3.1	3.2	3.3	3.6	3.8
y_i	2.6	2.8	2.9	3.6	3.6	3.1	3.4	3.1	3.9	3.1	4.1	3.6	4.0	3.6

$\Sigma X_i = 41.9$ $\qquad\qquad \Sigma y_i = 41.4$

$\Sigma x_i^2 = 327.79 \qquad \Sigma y_i^2 = 163.26$

$\Sigma x_i y_i = 143.59$

$\Sigma(x_i-\bar{x})^2 = 2.3893$

$X(y_i-\bar{y})^2 = 2.7771$

$\Sigma(x_i-\bar{x})(y_i-\bar{y}) = 1.7286$

$n\bar{y}^2 = 160.4829$

Simple Regression Model : $Y = \beta_0 + \beta_1 x + E$

กำหนด $b_0 = \hat{\beta}_0 = 1.2203$ และ

$b_1 = \hat{\beta}_1 = 0.1235$

1) จงเขียน Prediction equation หรือ Fitted regression equation

2) จงเติมตารางนี้ให้สมบูรณ์

ANOVA

S.V.	df.	SS	MS	f_c
Regression Error				$f_1 = 9.8318$
Lack of tit Pure error		 0.7066		$f_2 = .8294$
Total				

3) จากตาราง ANOVA ในข้อ 2) จงทดสอบที่ $\alpha = .05$

ก) $H_0 : \beta_1 = 0$

ข) H_0 : There is no lack of fit, เราต้องปรับปรุง model ใหม่หรือไม่

5. โรงงานผลิตสินค้าชนิดหนึ่งซื้อเครื่องจักรมาใหม่ 4 เครื่องจากโรงงานผลิตเครื่องจักร 4 โรงงาน ต้องการทราบว่าเครื่องจักร 4 เครื่องจะผลิตสินค้าได้เร็วเท่า ๆ กันหรือไม่ โรงงานจึงทำการนับปริมาณสินค้า (ชิ้น) ที่เครื่องจักรแต่ละเครื่องผลิตได้ในช่วงเวลาครั้งละ 1 ชั่วโมง ได้ข้อมูลดังนี้

ให้ x_{ij} เป็นปริมาณสินค้า (ชิ้น) ที่เครื่องจักรที่ i ผลิตได้ในช่วงเวลาที่ j, $i = 1, \ldots, 4$ และ $j = 1, \ldots, 5$

	เครื่องจักร				
	1	2	3	4	
	8	7	12	10	
	10	5	9	9	
	7	10	13	8	
	14	9	12	12	
	11	9	14	6	
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	9725
s_i^2	7.5	4	3.5	5	20

$$\sum_{i=1}^{4} \sum_{j=1}^{5} x_{ij}^2 = 1995 \text{ และ } G^2 = (195)^2 = 38025$$

1) จงทดสอบที่ $a = .05$, $H_0 : \sigma_1^2 = \sigma_3^2 = \sigma_3^2 = \sigma_4^2$ โดยที่ σ_i^2 เป็น population variance ของปริมาณสินค้าที่ผลิตจากเครื่องจักรที่ i, $i = 1, \ldots, 4$ โดยใช้วิธีของ Cochran (กำหนดให้ว่าจะ Reject H_0 เมื่อค่าของตัวสถิติทดสอบ G มีค่ามากกว่า $g_{(5,\ 4),\ .05} = 0.6287$)

2) ผลการทดสอบจาก 1) ถ้าเราไม่สามารถ reject H_0 ได้ จงทดสอบที่ a = .05 ว่าปริมาณสินค้าที่ผลิตได้โดยเฉลี่ยจากเครื่องจักรทั้ง 4 เท่ากันหมด และเราสรุปได้หรือไม่ว่าเครื่องจักรที่ 4 มีประสิทธิภาพในการผลิตดีเท่า ๆ กัน

กำหนดสูตรให้เลือกใช้ตามต้องการ บางสูตรที่ต้องใช้อาจไม่มีในรายการสูตรนี้

1. $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$

2. $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$

3. $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, $s_p^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}$

4. $SST = \sum_{i=1}^{t} \sum_{j=1}^{n_i} x_{ij}^2 - \dfrac{G^2}{N}$, $N = \sum_{i=1}^{t} n_i$

$SStr = \sum_{i=1}^{t} \dfrac{T_i^2}{n_i} - \dfrac{G^2}{N}$

5. $SST = \sum_{i=}^{n} (y_i - \bar{y})^2 = \Sigma y_i^2 - (\Sigma y_i)^2/n$

$SSR = b_1 \sum_{i=}^{n} (x_i - \bar{x})(y_i - \bar{y}) = b_1 [\Sigma x_i y_i - \Sigma x_i \Sigma y_i/n]$

6. $(1 \quad \alpha)$ 100% confidence interval ของ σ_1^2/σ_2^2 คือ

$$\left(\frac{s_1^2}{s_2^2} \frac{1}{f_{(\nu_1, \nu_2), \alpha/2}}, \frac{s_1^2}{s_2^2} f_{(\nu_2, \nu_1), \alpha/2} \right)$$

7. Cochran's test statistic : $G = \dfrac{\text{largest } S_i^2}{\sum_{i=1}^{t} S_i^2}$, t = จำนวนกลุ่ม

เฉลยข้อสอบภาค 2/2526

1) ก. H_0 : a: = σ_2^2 ข. Ho : $\mu_1 = \mu_2 = \ldots = \mu_t$, t > 2

2) parameter, 2 ii คือ point estimation และ interval estimation

3) ft. point estimate ข. Sampling distribution

4) H_1 (alternative hypothesis)

5) $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 > \mu_2$, one-tailed test

6) $X^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$, Chi-square distribution, df = n - 1

7) Analysis of variance (การวิเคราะห์ความแปรปรวน)

8) ก. fit distribution ที่สนใจ

ข. fit proportions ใน multinomial distribution

9) t-distribution , df = n-2, two-tailed test

IO) ทดสอบความเป็นอิสระ (Test for Independence)

เพศ	ปริมาณที่สูบ			รวม
	น้อย	ปานกลาง	มาก	
ชาย				
หญิง				
รวม				

2. 1) $H_0 : \sigma_1^2 = \sigma_2^2$

H_1 : a: $< \sigma_2^2$

a = .05

CR : $F < f_{(9, 8), .95} = .3096$

$$\left(f_{(9,\ 8),\ .95} = \frac{1}{f_{(8,\ 9),\ .05}} = \frac{1}{3.23} = .3096\right)$$

$$f_c = \frac{s_1^2}{s_2^2} = \frac{(3200)^2}{(3700)^2} = .748$$

$\because f_c = .748 > .3096$ เราไม่สามารถปฏิเสธ H_0 ที่ $\alpha = .05$ ได้ นั่นคือยอมรับว่า $\sigma_1^2 = \sigma_2^2$

2) n_1 และ n_2 มีขนาดเล็ก, ไม่ทราบค่า σ_1^2 และ σ_2^2 แต่ $\sigma_1^2 = \sigma_2^2$ (จาก 1)
   ดังนั้นในการทดสอบ $H_0 : \mu_1 = \mu_2$ เราใช้ pooled t-test

   $H_0 : \mu_1 = \mu_2$

   $H_1 : \mu_1 \neq \mu_2$

   $\alpha = .05$

   CR : $T < -t_{17,\ .025} = -2.11$ และ $T > t_{17,\ .025} = 2.11$

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{23600 - 24800}{s_p\sqrt{\frac{1}{10}+\frac{1}{9}}}$$

$$s_p^2 = \frac{(n_1 \quad 1)\ s_1^2 + (n_2 - 1)\ s_2^2}{n_1 + n_2 - 2} = \frac{9(3200)^2 + 8(3700)^2}{17}$$

ถ้า t_c ตกอยู่ใน CR. เราจะปฏิเสธ H_0 นั่นคือไม่ยอมรับว่ายางทั้ง 2 ชนิดมีอายุการใช้งานเฉลี่ยพอ ๆ กัน

ถ้า t_c ตกอยู่นอก CR. เราจะไม่ปฏิเสธ H_0 นั่นคือยอมรับว่ายางทั้ง 2 ชนิดมีอายุการใช้งานเฉลี่ยไม่แตกต่างกัน

3) 90% confidence interval ของ σ_1/σ_2 คือ

$$\left(\sqrt{\frac{s_1^2}{s_2^2}\frac{1}{f_{(9,\ 8),\ .05}}},\ \sqrt{\frac{s_1^2}{s_2^2}f_{(8,\ 9),\ .05}}\right)$$

$$= \left(\frac{3200}{3700}\sqrt{\frac{1}{3.39}},\ \frac{3200}{3700}\sqrt{3.23}\right)$$

3. (ดูเฉลยข้อสอบข้อ 7 ภาค 1/2526)

4. 1) $\hat{y} = 1.2203 + .7235x$

2) $SST = \Sigma(y_i - \bar{y})^2 = 2.7771$

$SSR = b_1 \Sigma(x_i - \bar{x})(y_i - \bar{y}) = (.7235)(1.7286) = 1.2506$

$SSE = SST - SSR = 1.5265$

$SS(lof) = SSE - SS(p.e.) = 1.5265 - .7066 = .8199$

ANOVA

S.V.	df.	SS	MS = SS/df	f
Regression	1	SSR = 1.2506	1.2506	$f_1 = \frac{1.2506}{.1272} = 9.8318$
Error	12	SSE = 1.5265	.1272	
Lack of fit	7	SS(lof) = .8199	.1171	$f_2 = \frac{.1171}{.1413} = .8294$
Pure error	5	SS(p.e.) = .7066	.1413	
Total	13	SST = 2.7771		

3) จากตาราง ANOVA ในข้อ 2)

n. $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$a = .05$

$CR : F > f_{(1,\ 12),\ .05} = 4.75$

$f_c = f_1 = 9.8318$

$\because f_1 = 9.8318 > 4.75$ (ตกใน CR) เราปฏิเสธ H_0 ที่ $a = .05$ นั่นคือ $\beta_1 \neq 0$ หรือ Y ขึ้นกับ X

U. H_0 : There is no lack of fit

H_1 : There is a lack of fit

$a' = .05$

$CR : F > f_{(7,\ 5),\ .05} = 4.88$

$f_c = f_2 = .8294$

$\because f_2 = .8294 < 4.88$ เราไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = .05$ สรุปได้ว่าเราไม่ต้องปรับปรุง Model ใหม่

5. 1) $H_0 : \sigma_1^2 = \ldots = \sigma_4^2$

$H_1 : \sigma_i^2$ ไม่เท่ากันหมด, $i = 1, \ldots, 4$

$\alpha = .05$

$CR : G > .6287$

$$g_c = \frac{\text{largest } s_i^2}{\Sigma s_i^2} = \frac{7.5}{20} = .375$$

$\because g_c = .375 < .6287$ (ตกอยู่นอก CR) เราไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = .05$ นั่นคือยอมรับว่า $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$

2) $H_0 : \mu_1 = \mu_2 = \ldots = \mu_4$

H_1 : มี μ_i อย่างน้อย 1 คู่ไม่เท่ากัน (หรือ μ_i ไม่เท่ากันหมด)

$\alpha = .05$

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

$N = 20, \ t = 4$

$$CF = \frac{(195)^2}{20} = 1901.25$$

$$SST = 1995 - 1901.25 = 93.75$$

$$SStr = \frac{9725}{5} - 1901.25 = 1945 - 1901.25 = 43.75$$

ANOVA

S.V.	df	ss	M S	f
Treatments	3	SStr = 43.75	14.5833	$fc = \frac{14.5833}{3.125} = 4.667$
Error	16	SSE = 50	3.125	
Total	19	SST = 93.75		

$\because f_c = 4.667 > 3.24$ (ตกอยู่ใน CR.) เราปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือเราไม่สามารถสรุปได้ว่าเครื่องจักรทั้ง 4 มีประสิทธิภาพในการผลิตดีเท่ากันหมด