ตัวอย่างข้อสอบ
ภาค 2/2527

คำสั่ง 1) ในการทดสอบสมมติฐาน ให้แสดง 6 ขั้นของการทดสอบให้ครบ (การให้คะแนนจะแบ่งให้ตามขั้นต่าง ๆ)

2) ข้อสอบมี 5 ข้อ ให้ทำทุกข้อ ข้อละ 20 คะแนน

---

1. จงตอบคำถามต่อไปนี้ให้ได้ความ (ไม่ต้องลอกโจทย์)

1) ถ้า $X \sim t_v$ แล้ว $X^2 \sim$ เป็นอะไร (บอก d.f.ด้วยถ้ามี)

2) ในการวิเคราะห์การถดถอย (ซึ่งมีตัวแบบ $Y_i = \beta_o + \beta_1 X_i + E_i$) B_o และ B_1 เป็น unbiased estimator ของ β_o และ β_1 ตามลำดับ จงพิสูจน์ว่า $B_o = \bar{Y} - B_1\bar{X}$ เป็น unbiased estimator ของ β_o

3) ถ้า $\bar{X} \sim N(\mu, \sigma^2/n)$ และ $X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ แล้ว

$$\frac{Z}{\sqrt{X^2/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$ จะมีการแจกแจง เป็นแบบใด บอก d.f.ด้วย ถ้ามี

โดยที่ $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

4) t = จำนวนวิธีการ = 5 ถ้า $n_1 = n_2 = \ldots\ldots = n_5 = n$ และ

$$L_1 = 4\mu_1 - \mu_2 - \mu_3 - \mu_4 - \mu_5 = \sum_{i=1}^{5} C_i\mu_i$$

$$L_2 = \mu_2 + \mu_3 - \mu_4 - \mu_5 = \sum_{i=1}^{5} b_i\mu_i$$

L_1 และ L_2 เป็น orthogonal contrast หรือไม่ เพราะเหตุใด

5) จากข้อ 4) point estimator ของL_1 คืออะไร

6) ถ้า ρ คือ population correlation coefficient ระหว่างตัวแปรเชิงสุ่ม X และ Y ในการทดสอบ $H_o : \rho = O$ ถ้าเราไม่ปฏิเสธ H_o เราสรุปผลการทดสอบว่าอย่างไร และถ้าปฏิเสธ H_o และคำนวณค่า r เป็นลบ หมายความว่าอย่างไร

7) ในการทดสอบ $H_o: \sigma_1^2 = \sigma_2^2 = \ldots\ldots = \sigma_k^2$, $k > 2$ ถ้า n_i ไม่เท่ากันหมด เราใช้วิธีของ Bartlette ถ้า n_i เท่ากันหมดเราจะใช้วิธีอะไร (บอกแต่ชื่อ)

8) การวิเคราะห์ความแปรปรวนร่วม (Analysis of Covariance) เป็นวิธีวิเคราะห์ทางสถิติวิธีหนึ่งซึ่งรวมวิธีวิเคราะห์ทางสถิติ 2 อย่างเข้าด้วยกัน วิธีวิเคราะห์ทั้งสองวิธีดังกล่าวคืออะไร

9) ในตัวแบบของการทดสอบ (Regression model) $Y = \beta_o + \beta_1 x + E$ ตัวใดบ้างในตัวแบบเป็นตัวแปรเชิงสุ่ม และตัวใดบ้างเป็นตัวพารามิเตอร์

10) ถ้า S^2 เป็นความแปรปรวนของตัวอย่างสุ่มขนาด n ซึ่งสุ่มจากประชากรแบบปกติซึ่งมีความแปรปรวน σ^2 แล้วตัวแปรเชิงสุ่ม $X^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ จงแสดงวิธีหา $(1-\alpha)$ 100% Confidence interval ของ σ^2

2. ปริมาณสารเคมี y (กรัม) ซึ่งละลายในน้ำ 100 กรัม ณ อุณหภูมิ x (°c) เป็นดังนี้

x (°c)	y (กรัม)
0	8,6,8
15	12,10,14
30	25,21,24
45	31,33,28
60	44,39,42
75	48,51,44

สมมุติตัวแบบ : $\mu_{y/x} = \beta_0 + \beta_1 x$

ก) จงอธิบายความหมายของ β_0 และ β_1 สำหรับโจทย์ข้อนี้

ข) ถ้า $b_0 = 5.8261$ และ $b_1 = 0.5676$ จงเขียน prediction equation

ค) จงประมาณว่าสารเคมีจำนวนกี่กรัมที่สามารถละลายในน้ำ 100 กรัม ที่อุณหภูมิ 50°C

ง) จงเติมตาราง ANOVA ข้างล่างให้สมบูรณ์

ANOVA

ที่มา (S.V.)	d.f.	S.S.	M.S.	f
Regression	?	?	3805.7580	f_1 = 574.35
Error	?	?	6.6262	
Lack of fit	?	?	9.1716	f_2 = 1.5874
Pure error	?	69.3333	5.7778	
Total	?	3911.7778		

จ) จงใช้ค่าจากตารางในข้อ ง) ทดสอบที่ $\alpha = .01$

1) $H_0 : \beta_1 = 0,\ H_1 : \beta_1 \neq 0$

2) H_0 : ไม่มี lack of fit เราจะต้องปรับปรุง model ใหม่หรือไม่

ฉ) $100\ r^2\% = \frac{SSR}{SST} \times 100\% = 97.29\%$ หมายความว่าอย่างไร

3. ในการทดลองเพื่อทดสอบว่าอาหาร 3 ชนิดจะลดน้ำหนักได้เหมือนกันหรือไม่ ได้เอาคนที่ต้องการลดน้ำหนัก และมีน้ำหนักพอ ๆ กันมา 17 คน แบ่งออกเป็น 4 พวกอย่างสุ่ม โดยสุ่ม 5 คนแรกให้เป็นกลุ่มควบคุม (Control group) โดยให้ทานอาหารตามปกติ สุ่มอีก 3 คนให้ทานอาหารชนิดที่ 1 สุ่ม 4 คน ให้ทานอาหารชนิดที่ 2 ส่วน 5 คนที่เหลือให้ทานอาหารชนิดที่ 3 แล้ววัดน้ำหนัก (ก.ก.) ที่ลดลงในช่วงเวลา 2 เดือนได้ผลดังนี้

	กลุ่มควบคุม	อาหารชนิดที่ 1	ชนิดที่ 2	ชนิดที่ 3
	.6	1.1	2.1	.5
	.8	1.3	2.0	.9
	.3	1.5	1.7	1.0
	.4		1.6	.7
	.6			.7
รวม : T_i	2.8	3.9	7.4	3.8

สมมติว่าตัวอย่างทั้ง 4 มาจากประชากรแบบปกติที่มี (μ_i,σ^2), $i=1,\ldots,4$

1) จงเติมตาราง ANOVA ให้สมบูรณ์

ANOVA

ที่มา	d.f.	S.S.	M.S.	f
วิธีการ (Treatments)	?	?	1.4561	35.4282
Error	?	0.5430	.0411	
Total	?	4.9024		

2) ใช้ค่าจากตาราง ANOVA ข้างบน ทดสอบที่ $\alpha=.01$ $H_o : \mu_1=\mu_2 = \ldots = \mu_4$

3) ถ้าปฏิเสธ H_o ในข้อ 2) จงใช้วิธีของ Scheffé ทำ Pairwise Comparison ที่ $\alpha=.05$ โดยหาค่าตามตารางต่อไปนี้ (ให้คิดทศนิยม 2 ตำแหน่ง แล้วสรุปผลการทำ Pairwise Comparison ตามวิธีนี้ด้วย

Li	ii	ii ± 4.1497 $\sqrt{MSE\ \Sigma c_i^2/n_i}$
Li = μ_1-μ_2	?	? ± .61 = ?
L_2 = μ_1-μ_3	?	? ± .56 = ?
L_3 = μ_1-μ_4	?	? ± .53 = ?
L_4 = μ_2-μ_3	?	? ± .64 = ?
L_5 = μ_2-μ_4	?	? ± .61 = ?
L_6 = μ_3-μ_4	?	? ± .56 = ?

4. ก) จำนวนรถยนต์ที่วิ่งผ่านจุด ๆ หนึ่งในช่วงเวลาหนึ่งในระยะเวลา 400 วัน เป็นดังนี้

ปริมาณรถยนต์	0	1	2	3	4	5	6	รวม
จำนวนวัน	129	137	83	38	10	2	1	400

จากข้อมูลนี้จะทำให้สรุปได้ที่ $\alpha = .01$ หรือไม่ว่าปริมาณรถยนต์ที่วิ่งผ่านจุด ๆ นี้ (X) ในช่วงเวลาหนึ่งเป็นตัวแปรเชิงสุ่มที่มีการแจกแจงเป็นปัวซอง (Poisson distribution) ที่มี m = 1

ถ้า X~ Poisson (m = 1) กำหนด $p_i = P_r\,[X=x_i]$ ดังนี้

x_i	0	1	2	3	4	5	6	7
p_i	.3679	.3679	.1839	.0613	.0153	0031	.0005	.0001

และกำหนดค่าคำนวณของตัวสถิติที่ใช้ในการทดสอบให้ = 15.4014 จงแสดงรายละเอียดของการวิเคราะห์อย่างละเอียดทุกขั้นตอน

ข) ต้องการเปรียบเทียบผลการรักษาโรคไอกรนด้วยยาชนิดต่าง ๆ 5 ชนิด และดูความไร้ผลของการรักษาหลังจากรักษาแล้ว 6 วัน จากผลการทดลองในตารางข้างล่างนี้จงทดสอบที่ $\alpha = .05$ ว่า อัตราส่วนของการรักษาที่ไร้ผลจากการใช้ยาทั้ง 5 ชนิด เท่ากันหมดหรือไม่

ยาชนิดที่	จำนวนผู้ป่วยทั้งสิ้น	จำนวนผู้ป่วยที่ไม่หายจากโรค
1	45	23
2	66	26
3	27	14
4	55	23
5	49	17

กำหนดค่าคำนวณของตัวสถิติที่ใช้ในการทดสอบให้ = 3.4681 จงแสดงรายละเอียดของการวิเคราะห์อย่างละเอียดทุกขั้นตอน

ก) ในการเปรียบเทียบค่าความแปรปรวนของคะแนนสอบโดยใช้ข้อสอบ 2 ชนิด ปรากฏว่าความแปรปรวนของผลการสอบโดยใช้ข้อสอบปรนัยกับนักเรียน 21 คนมีค่าเท่ากับ 105.75 ส่วนความแปรปรวนของผลการสอบโดยใช้ข้อสอบแบบอัตนัยที่ใช้กับนักเรียน 13 คน มีค่าเท่ากับ 63.41 เราอาจสรุปได้หรือไม่ โดยอาศัยข้อเท็จจริงจากตัวอย่าง ว่าความแปรปรวนของคะแนนสอบเมื่อใช้ข้อสอบทั้ง 2 ชนิดนี้ไม่แตกต่างกัน ให้ทดสอบที่ $\alpha = .02$
(กำหนด 105.75/63.41 = 1.6677)

ข) เครื่องซักผ้าชนิดหนึ่งถูกผลิตออกมาขายในสีต่าง ๆ กัน 5 สี นักวิจัยตลาดต้องการศึกษาว่าผู้ซื้อนิยมสีต่าง ๆ พอ ๆ กันหรือไม่ ได้สุ่มตัวอย่างคือผู้ซื้อ 300 คน ผลปรากฏดังนี้

สี	เขียว	เหลือง	ขาว	นวล	น้ำตาล	รวม
จำนวนคนที่ซื้อ	88	65	52	40	55	300

เราจะสรุปได้ (ที่ $\alpha = .05$) หรือไม่ว่าผู้ซื้อนิยมสีต่าง ๆ พอ ๆ กัน กำหนดค่าคำนวณของตัวสถิติที่ใช้ในการทดสอบให้ = 21.6335 จงแสดงรายละเอียดของการวิเคราะห์อย่างละเอียดทุกขั้นตอน

เฉลยข้อสอบภาค 2/2527

1. 1) F distribution, df. = (1, v)

2) $$B_0 = \bar{Y} - B_1\bar{X}$$

$$= \frac{1}{n}\sum_{i=1}^{n} Y_i - B_1\bar{X}$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\beta_0 + \beta_1 X_i + E_i) - B_1\bar{X}$$

$$E(B_0) = \frac{1}{n}\sum_{i=1}^{n} E(\beta_0 + \beta_1 X_i + E_i) - X E(B_1)$$

$$= \frac{1}{n}\sum_{i=1}^{n} (\beta_0 + \beta_1 X_i) - \bar{X}\beta_1$$

$$= \frac{1}{n}(n\beta_0 + \beta_1 \sum_{i=1}^{n} X_i) - \bar{X}\beta_1$$

$$= \beta_0 + \beta_1\bar{X} - \beta_1\bar{X}$$

$= \beta_0$ ดังนั้น Bo เป็น unbiased estimator ของ β_0

3) t distribution, df. = n − 1

4) L_1 และ L_2 เป็น orthogonal contrast เพราะ $\sum_{i=1}^{5} c_i b_i = 0$

5) $\hat{L}_1 = 4\bar{X}_1 - \bar{X}_2 - \bar{X}_3 - \bar{X}_4 - \bar{X}_5$

6) ถ้าเราไม่ปฏิเสธ $H_0 : \rho = 0$ เราสรุปว่า X และ Y ไม่มีความสัมพันธ์กันเชิงเส้นตรง แต่อาจมีความสัมพันธ์กันในรูปอื่น หรืออาจไม่มีความสัมพันธ์กันเลย

ถ้าเราปฏิเสธ $H_0 : \rho = 0$ และ r มีค่าเป็นลบ หมายความว่าถ้าตัวแปรตัวหนึ่งมีค่าเพิ่มขึ้น ตัวแปรอีกตัวหนึ่งจะมีค่าลดลง

7) วิธีของ Cochran หรือวิธีของ Hartley

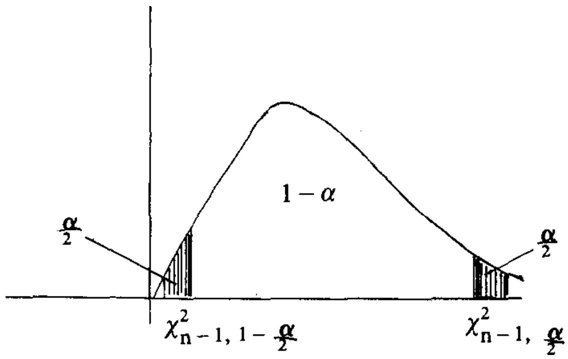
8) 1. การวิเคราะห์ความแปรปรวน (Analysis of Variance)

2. การวิเคราะห์การถดถอย (Regression Analysis)

0) ตัวแปรเชิงสุ่ม : Y และ E

ตัวพารามิเตอร์ : β_0 และ β_1

10) จากรูป



$$\Pr\left[\chi^2_{n-1,\,1-\frac{\alpha}{2}} < X^2 < \chi^2_{n-1,\,\frac{\alpha}{2}}\right] = 1-\alpha$$

$$\Pr\left[\chi^2_{n-1,\,1-\frac{\alpha}{2}} < \frac{(n-1)\,S^2}{\sigma^2} < \chi^2_{n-1,\,\frac{\alpha}{2}}\right] = 1-\alpha$$

$$\Pr\left[\frac{(n-1)\,S^2}{\chi^2_{n-1,\,\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)\,S^2}{\chi^2_{n-1,\,1-\frac{\alpha}{2}}}\right] = 1-\alpha$$

$(1-\alpha)$ **100% Confidence interval** ของ σ^2 คือ

$$\left(\frac{(n-1)\,S^2}{\chi^2_{n-1,\,\frac{\alpha}{2}}},\ \frac{(n-1)\,S^2}{\chi^2_{n-1,\,1-\frac{\alpha}{2}}}\right)$$

2. ก) β_0 เป็นปริมาณสารเคมีเฉลี่ยซึ่งละลายในน้ำ 100 กรัม ที่อุณหภูมิ 0°C

β_1 เป็นปริมาณสารเคมีซึ่งละลายในน้ำได้เพิ่มขึ้นหรือลดลง เมื่ออุณหภูมิเปลี่ยนแปลงไป 1°C

ข) $\hat{y} = 5.8261 + 0.5676\,x$

ค) $\hat{y} = 5.8261 + 0.5676\,(50) = 5.8261 + 28.38 = 34.2061$

ง) **SSR = MSR = 3805.7580**

SSE = SST-SSR = 3911.7778 − **3805.7580 = 106.0198**

SS (lof) = **SSE- SS** (p.e.) = **106.0198-69.3333 = 36.6865**

n = 18. k = 6

ANOVA

S.V.	d.f.	s s	MS	f
Regression	1	3805.7580	3805.7580	f_1 = 574.35
Error	16	106.0198	6.6262	
Lack of fit	4	36.6865	9.1716	f_2 = 1.5874
Pure error	12	69.3333	5.7778	
Total	17	3911.7778		

ข) 1. $H_o : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$a = .01$

CR : $F > f_{(1,16), .01} = 8.53$

$f_c = f_1 = 574.35$

$\because$ f_c ตกใน CR เราปฏิเสธ H_0 ที่ $\alpha = .01$ นั่นคือ X และ Y มีความสัมพันธ์กันในรูปเหตุและผล

2. H_0 : ไม่มี lack of fit

H_1 : มี lack of fit

$a = .01$

CR : $F > f_{(4, 12), .01} = 5.41$

$f_c = f_2 = 1.5874$

$\because$ f_c ไม่ตกใน CR เราไม่สามารถปฏิเสธ H_0 ที่ $a = .01$ ดังนั้นเราไม่ต้องปรับปรุง model

ฉ) $100 r^2\%$ = 97.29% หมายความว่า 97.29% ของความแปรปรวนของ Y เนื่องมาจากความสัมพันธ์เชิงเส้นตรงกับ X

3. 1) $SStr = SST - SSE$

$= 4.9024 - 0.5430 = 4.3684$

$n_1 = 5$, "2 = 3, "3 = 4, $n_5 = 5$

$$N = \sum_{i=1}^{4} n_i = 17$$

$t = 4$

ANOVA

S.V.	d.f.	ss.	MS.	f
Treatments	3	4.3594	1.4561	35.4282
Error	13	0.5430	0.0411	
Total	16	4.9024		

2) $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

H_1 : มี mean อย่างน้อย 1 คู่ที่ไม่เท่ากัน

$\alpha = .01$

CR : $F > f_{(3,\ 13),\ .01} = 5.74$

$f_c = 35.4282$

$\because$ f_c ตกใน CR เราปฏิเสธ H_0 ที่ $\alpha = .01$ นั่นคือเราไม่สามารถสรุปได้ว่าอาหารทั้ง 4 ชนิดช่วยลดน้ำหนักได้เท่ากันหมด

3)

n_i	5	3	4	5
T_i	2.8	3.9	7.4	3.8
$\overline{x}_i$	.56	1.3	1.85	.76

L_i	$\hat{L}_i$	$\hat{L}_i \pm 4.1497 \sqrt{MSE \sum_{i=1}^{4} c_i^2/n_i}$
$L_1 = \mu_1 - \mu_2$	$\bar{x}_1 - \bar{x}_2 = -.74$	$-.74 \pm .61 = (-1.35, -.13)$
$L_2 = \mu_1 - \mu_3$	$\bar{x}_1 - \bar{x}_3 = -1.29$	$-1.29 \pm .56 = (-1.85, -.73)$
$L_3 = \mu_1 - \mu_4$	$\bar{x}_1 - \bar{x}_4 = -.2$	$-.2 \pm .53 = (-.73, .33)$
$L_4 = \mu_2 - \mu_3$	$\bar{x}_2 - \bar{x}_3 = -.55$	$-.55 \pm .64 = (-1.19, .09)$
$L_5 = \mu_2 - \mu_4$	$\bar{x}_2 - \bar{x}_4 = .54$	$.54 \pm .61 = (-.07, 1.15)$
$La = \mu_3 - \mu_4$	$\bar{x}_3 - \bar{x}_4 = 1.09$	$1.09 \pm .56 = (.53, 1.65)$

สรุปผล $\quad \bar{x}_1 \quad \bar{x}_4 \quad \bar{x}_2 \quad \bar{x}_3$

$\quad .56 \quad .76 \quad 1.3 \quad 1.85$

นั่นคือ $\mu_1 = \mu_4$, $\mu_4 = \mu_2$ และ $\mu_2 = \mu_3$ ส่วนคู่อื่น ๆ ต่างกัน

4. ก) ให้ X = ปริมาณรถยนต์ที่วิ่งผ่านจุดๆ นี้

Ho : X $\sim$ Poisson (m = 1)

H_1 : X ไม่มีการแจกแจง แบบ Poisson

a = .01

CR : $X^2 > \chi^2_{4,\ .01} = 13.277$ (df. = 5 − 1 = 4)

x_i	0	1	2	3	4	5	6	7
Pi	.3679	.3679	.1839	.0613	.0153	.0031	.0005	.0001
e_i	147.16	147.16	73.56	24.52	6.12	1.24	.20	.04
					7.60			
Oi	129	137	83	38	10	2	1	0
					13			

$$\chi_c^2 = \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i}$$

$$= \frac{(129 - 147.16)^2}{147.16} + \dots + \frac{(13 - 6.88)^2}{6.88} = 15.4014$$

สรุป $\chi_c^2 = 15.4014 > 13.277$ (ตกใน CR) เราปฏิเสธ Ho ที่ $\alpha = .01$ นั่นคือ X ไม่มีการแจกแจง แบบ Poisson

ข) ให้ p_i = Pr [ผู้ป่วยไม่หายจากโรคเมื่อใช้ยาชนิดที่ i]

Ho : $p_1 = \dots = p_5$

H1 : p_i ไม่เท่ากันหมด

$\alpha = .05$

CR : $X^2 > \chi^2_{4\ .05} = 9.488$, df. = $(5-1)(2-1) = 4$

Oij

ยา	ไม่หายจากโรค	หายจากโรค	รวม : R_i
1	23	22	45
2	28	38	66
3	14	13	27
4	23	32	55
5	17	32	49
รวม : C_j	105	137	n = 242

$e_{ij} = \frac{R_i C_j}{n}$ เช่น $e_{11} = \frac{R_1 C_1}{n} = \frac{45(105)}{242}$ เป็นต้น

$$\chi_c^2 = \sum_{i=1}^{5} \sum_{j=1}^{2} \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

$$= 3.4681$$

สรุป $\chi_c^2 = 3.4681 < 9.488$ (ไม่ตกใน CR) เราไม่สามารถปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือสรุปว่ายาทั้ง 5 ชนิดมีประสิทธิภาพในการรักษาโรคไอกรนพอ ๆ กันหมด

•

5. ก) ข้อสอบปรนัย : $n_1 = 21$, $s_1^2 = 105.75$

ข้อสอบอัตนัย : $n_2 = 13$, $s_2^2 = 63.41$

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

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

a = .02

CR : $F > f_{(20,\ 12),\ .01} = 3.86$ และ $F < f_{(20,\ 12),\ .99} = \frac{1}{f_{(12,\ 20),\ .01}}$

$$= \frac{1}{3.23} = .3096$$

$$f_c = \frac{s_1^2}{s_2^2} = \frac{105.75}{63.41} = 1.6677$$

สรุป $f_c = 1.6677$ ไม่ตกใน CR เราไม่สามารถปฏิเสธ H_0 ได้ที่ $\alpha = .02$ นั่นคือความแปรปรวนของคะแนนสอบเมื่อใช้ข้อสอบ 2 ชนิดไม่ต่างกัน

ข) ให้ X = จำนวนผู้ซื้อสีต่าง ๆ

H_0 : X $\sim$ Uniform ($\frac{1}{5}$)

H_1 : X ไม่มี การแจกแจง แบบ Uniform

α = .05

CR : $X^2 > \chi^2_{4,\ .05} = 9.488$, df. = k - 1 = 5 − 1 = 4

o_i	88	65	52	40	55	300
e_i	60	60	60	60	60	300

$$\chi^2_c = \sum_{i=1}^{5} \frac{(o_i - e_i)^2}{e_i}$$

$$= \frac{(88-60)^2}{60} + \ldots + \frac{(55-60)^2}{60}$$

$$= 21.6335$$

สรุป χ^2_c = 21.6335 > 9.488 (ตกใน CR) เราปฏิเสธ H_0 ที่ α = .05 นั่นคือผู้ซื้อนิยมสีต่าง ๆ ไม่เท่ากันหมด