ตัวอย่างข้อสอบ

ภาค 1/2526

คำสั่ง ให้นักศึกษาทุกคนทำข้อ 1, 2, 3 และ 4 และเลือกทำอีก 1 ข้อจากข้อที่เหลือ ในการทดสอบสมมติฐานให้แสดง 6 ขั้นของการทดสอบสมมติฐานให้ครบ นักศึกษาอาจเลือกใช้สูตรที่ให้มากับข้อสอบชุดนี้ได้

1. จงตอบคำถามหรือเติมข้อความให้ถูกต้องและได้ความ (ไม่ต้องลอกโจทย์)

   1) ถ้า S^2 เป็น Sample variance ของตัวอย่างสุ่ม (Random sample) ขนาด n ซึ่งสุ่มจาก Normal population ซึ่งมี variance σ^2, แล้ว Random variable $X^2 = \frac{(n-1)S^2}{\sigma^2}$ จะมี การแจกแจงเป็น.................... distribution ที่มี................. .degrees of freedom (ถ้ามี)

   2) t distribution จะเป็น distribution เดียวกับ Standard Normal distribution เมื่อขนาดของตัวอย่าง (n).....,....,.......

   3) ตัวสถิติ $\hat{\Theta}$ จะเป็น Unbiased estimator ของ parameter θ ถ้า E $(\hat{\Theta})$ = ...)
   ถ้า X มี การแจกแจง เป็น Binomial (n, p) แล้ว E (X) = np จงแสดงว่า $\hat{P} = \frac{X}{n}$ เป็น Unbiased estimator ของ p

   4) นอกจากเราใช้ Chi-square distribution ในการทดสอบสมมติฐานเกี่ยวกับ variance ของ Normal population แล้ว เรายังใช้ในการทดสอบสมมติฐานเกี่ยวคับ 1.................

   2....................

   5) ในการทดสอบสมมติฐานอาจเกิดความผิดพลาดได้ 2 ชนิด do ถ้า Ho เป็นจริง แต่เรา Reject Ho เราทำความผิดที่เรียกว่า................... ส่วนถ้าH_0 ไม่เป็นจริงแต่เราไม่ Reject H_0 เราทำความผิดที่เรียกว่า...................

6) ให้ μ_1 และ μ_2 เป็นรายได้เฉลี่ยของคนที่จบปริญญาตรี และของคนที่ไม่จบปริญญาตรี ตามลำดับ ถ้าเราต้องการจะทดสอบเพื่อดูว่ารายได้เฉลี่ยของคนที่จบปริญญาตรีจะมากกว่าของคนที่ไม่จบปริญญาตรีหรือไม่ เราจะตั้ง H_0 :.................. และ H_1 : .... และการทดสอบนี้เป็นการทดสอบแบบ ....-tailed test

7) ใน Population regression equation : $\mu_{Y/X} = \beta_0 + \beta_1 X$ สัมประสิทธิ์ของความถดถอย (Regression coefficient) do................ และตัวแปรอิสระ (Independent variable) do..................

8) r(Sample correlation coefficient) จะเป็นตัวประมาณค่า degree ของความสัมพันธ์ของ 2 random variables X และ Y, ก็ต่อเมื่อ X และ Y มีความสัมพันธ์กันในรูป ....และในการทดสอบสมมติฐาน $H_0 : \rho = 0$ เราใช้ตัวสถิติ $T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ ซึ่งจะมีการแจกแจงเป็น t distribution ที่มี..................degrees of freedom ถ้า Ho เป็นจริง

9) ในการทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$ โดยที่ σ_1^2 และ σ_2^2 เป็น variances ของ 2 Normal populations เราใช้ตัวสถิติในการทดสอบ (Test statistic) $\hat{\Theta} = S_1^2/S_2^2$, ถ้า Ho เป็นจริง และเราสุ่มตัวอย่างขนาด n_1 และ n_2 จาก 2 populations ข้างต้น โดยที่ตัวอย่างทั้ง 2 เป็นอิสระต่อกัน , S_1^2/S_2^2 จะมีการแจกแจงเป็น.................... djstribution ที่มี....................degrees of freedom (ถ้ามี)

10) ถ้าหา 95% confidence interval ของ σ_1^2/σ_2^2 ได้ (.295, 2.785) จงอธิบายความหมาย และเราจะสรุปได้หรือไม่ว่า $\sigma_1^2 > \sigma_2^2$ เพราะเหตุใด

2. ในการศึกษาว่าวิธีการลดอาหารวิธีหนึ่งจะลดน้ำหนักลงได้ภายใน 2 เดือนหรือไม่ ได้สุ่มผู้หญิงที่ต้องการลดน้ำหนักมา 5 คน ชั่งน้ำหนัก (เป็นปอนด์) ไว้ แล้วทำการลดอาหารตามวิธีการนี้ หลังจาก 2 เดือนผ่านไป ได้ทำการชั่งน้ำหนักของทั้ง 5 คนอีก น้ำหนักก่อนและหลังลดของแต่ละคนเป็นดังนี้

คนที่	น้ำหนักก่อนลด	น้ำหนักหลังลด
1	129	130
2	133	'121
3	136	128
4	141	129
5	138	132

ก. สมมุติว่าน้ำหนักก่อนและหลังลดต่างมีการกระจายเป็น Normal distribution จงทดสอบที่ α = .05 ว่าน้ำหนักเฉลี่ยก่อนลดมากกว่าน้ำหนักเฉลี่ยหลังลดอาหารหรือไม่ จากผลการทดสอบเราจะสรุปได้หรือไม่ว่าวิธีการลดอาหารวิธีนี้ได้ผลใน 2 เดือน

ข. จงหา 95% confidence interval ของผลต่างที่แท้จริงของน้ำหนักเฉลี่ยก่อนลดและน้ำหนักเฉลี่ยหลังลด

3. ในการเตรียมการที่จะคัดเลือกนักศึกษาปีที่หนึ่งเพื่อเข้าเรียนวิชาเอกคณิตศาสตร์ โดยจะใช้วิธีการสอบวิชาคณิตศาสตร์พื้นฐาน (คะแนนเต็ม 10 คะแนน) เพื่อจะศึกษาดูว่าการใช้คะแนนวิชานี้จะทำนายผลสำเร็จในการเรียนวิชาเอกได้หรือไม่ วิทยาลัยแห่งหนึ่งจึงได้เอานักศึกษาที่สำเร็จการศึกษาแล้วและได้คะแนนที่วิชาพื้นฐานต่าง ๆ กันมา 8 คน แล้วจดคะแนนตอนจบปริญญาตรีวิชาเอกคณิตศาสตร์ (grade point average : GPA) ของแต่ละคนดังนี้

คนที่	1	2	3	4	5	6	7	8
x : คะแนนคณิตศาสตร์พื้นฐาน	4	6	8	6.5	3.5.	8	7	5
y : GPA (เต็ม 4.00)	2.1	3.1	3.4	2.9	2.1	3.3	2.8	2.3

กำหนดค่าที่คำนวณจากข้อมูล

$\Sigma x = 46$ $\quad \Sigma y = 22$ $\quad \Sigma xy = 137.9$

$\Sigma x^2 = 308.5$ $\quad \Sigma y^2 = 62.42$

$\bar{x} = 6$ $\quad \bar{y} = 2.75$

$S_{xx} = \Sigma(x_i - \bar{x})^2 = 20.5$, $S_{yy} = \Sigma(y_i - \bar{y})^2 = 1.92$, $S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y}) = 5.9$

สมมุติว่าความสัมพันธ์ของ X และ Y อยู่ในรูป $\mu_{Y/X} = \beta_0 + \beta_1 X$

ก. จงหา Least square estimate ของ β_0 และ β_1

ข. จงเขียน Prediction equation

ค. จงหา Point estimate ของ $\mu_{Y/7.5}$ และอธิบายความหมาย และถ้า นาย ก. ได้คะแนนวิชาพื้นฐาน 5.5 คะแนน จงประมาณว่านาย ก.จะเรียนจบด้วย GPA เท่าใด

ง. จงทดสอบ $H_0 : \beta_1 = 0$ โดยใช้ F – test กำหนด $\alpha = .01$

จ. ถ้าความสัมพันธ์ของ X และ Y อยู่ในรูปนี้จริง ค่าประมาณของ $\sigma^2_{Y/X} =$?

4. ในการทดลองเพื่อทดสอบว่าปุ๋ย 4 ชนิดจะให้ผลต่อการเจริญเติบโตของพืชชนิดหนึ่งได้เหมือนกันหรือไม่ ได้นำพืชที่มีส่วนสูงเท่า ๆ กันและลักษณะความแข็งแรงของต้นพอ ๆ กันมา 20 ต้น แบ่งออกเป็น 4 พวกอย่างสุ่ม แต่ละพวกใส่ปุ๋ยแต่ละชนิด แล้ววัดความสูง (เป็นฟุต) หลังจาก 3 เดือนไปแล้ว ได้ผลดังนี้

	ปุ๋ย				
	A	B	C	D	
	.6	1.1	2.1	.5	
	.8	1.3	2.0	.9	
	.3	1.5	1.7	1.0	
	.5	.8	1.6	.7	$\sum_{i=1}^{4} \sum_{j=1}^{5} x_{ij}^2 = 26.2$
	.6	.9	1.0	.7	
Total : T_i	2.8	5.6	8.4	3.8	G = 20.6
T_i^2	7.84	1.36	70.56	14.44	124.2
Mean : $\bar{x}_i$	.56	.12	1.68	.76	$\bar{x}$ = 1.03

สมมุติว่า Population$_i$ ของส่วนสูงของพืชที่ได้รับปุ๋ยแต่ละชนิดต่างมีการแจกแจงเป็น Normal ที่มี $\sigma^2_A = \sigma^2_B = \sigma^2_C = \sigma^2_D = \sigma^2$ จงทดสอบที่ $\alpha = .01$ ว่าส่วนสูงเฉลี่ยเมื่อใช้ปุ๋ย 4 ชนิดไม่แตกต่างกัน จากผลการทดสอบจะสรุปได้หรือไม่ว่าปุ๋ยทั้ง 4 ชนิดดีพอ ๆ กันในการช่วยการเจริญเติบโตของพืชชนิดนี้

ถ้า $L = \frac{\mu_A + \mu_B}{2} - \frac{\mu_C + \mu_D}{2}$, L เป็น Contrast หรือไม่เพราะเหตุใด จงหา $\hat{L}$ และ $\hat{\sigma}_{\hat{L}}$

5. ในการศึกษาถึงแนวโน้มว่าผู้ชายหรือผู้หญิงจะเห็นด้วยมากกว่ากัน เกี่ยวกับนโยบายหนึ่ง จึงได้ถามความเห็นของชาย 200 คน ปรากฏว่าเห็นด้วย 120 คน และถามความเห็นของหญิง 500 คน ปรากฏว่าเห็นด้วย 240 คน ท่านเห็นด้วยหรือไม่ว่าสัดส่วน (proportion) ของชายที่เห็นด้วยมากกว่าของหญิงที่เห็นด้วย ให้ทำการทดสอบสมมติฐานที่ $\alpha = .05$ เพื่อสนับสนุนความเห็นของท่าน และจงหา 95% confidence interval ของค่าความแตกต่างของสัดส่วนที่แท้จริงของชายและหญิง

6. ในการศึกษาถึงความสัมพันธ์ระหว่างระดับการศึกษาและระดับความสนใจเกี่ยวกับการเมืองในประเทศ ได้ทำการสุ่มนักเรียนในระดับต่าง ๆ มาทั้งหมด 200 คน จำแนกตามระดับการศึกษาและระดับความสนใจดังนี้

ระดับการศึกษา	ระดับความสนใจ			รวม
	ไม่สนใจเลย	สนใจปานกลาง	สนใจมาก	
มัธยม	14	37	32	83
ปริญญาตรี	19	42	17	78
สูงกว่าปริญญาตรี	12	17	10	39
รวม	45	96	59	200

จงทดสอบที่ $\alpha = .05$ ว่าระดับการศึกษาและความสนใจเกี่ยวกับการเมืองในประเทศเป็นอิสระต่อกันหรือไม่ กำหนดค่าของตัวสถิติทดสอบ (Test statistic) ให้ = 7.4646 ให้แสดงวิธีหาค่าของตัวสถิตินี้ด้วย

7. หยิบไพ่ 3 ใบจากไพ่สำรับหนึ่งซึ่งมี 52 ใบ แบบหยิบแล้วใส่คืน (with replacement) ให้ Y เป็นจำนวนไพ่โพดำที่หยิบได้ในไพ่ 3 ใบ ได้ทำการทดลองอย่างเดียวกันนี้ 128 ครั้ง (แต่ละครั้งนับจำนวนไพ่โพดำที่หยิบได้ในไพ่ 3 ใบ) ได้ผลดังนี้

y	0	1	2	3	รวม
จำนวนครั้ง	42	62	22	2	128

จงทดสอบสมมติฐานว่า การทดลองนี้ให้ตัวเลขที่มีการแจกแจงเป็น Binomial distribution ที่มี $p = \frac{1}{4}$ หรือทดสอบว่า Random variable Y มีการแจกแจงเป็น Binomial distribution ที่มี $p = \frac{1}{4}$ $(n = 3)$ ให้ทดสอบที่ $a = .05$

หมายเหตุ ถ้า Y มีการแจกแจง เป็น Binomial $(n = 3,\ p = \frac{1}{4})$ แล้ว

y	0	1	2	3	รวม
$Pr[Y = y]$	27/64	27/64	9/64	1/64	1

กำหนดสูตร ให้เลือกใช้ตามต้องการ บางสูตรที่ต้องใช้อาจไม่มีในรายการสูตรนี้

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

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

3. $z = \dfrac{(\hat{p}_1 - \hat{p}_2) - d_0}{\sqrt{\hat{p}_1\hat{q}_1/n_1 + \hat{p}_2\hat{q}_2/n_2}}$

4. $z = \dfrac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

5. $z = \dfrac{x - np_0}{\sqrt{np_0q_0}}$

6. $z = \dfrac{\hat{p} - p_0}{\sqrt{p_0q_0/n}}$

7. $t = \dfrac{\bar{d} - d_0}{s_d/\sqrt{n}}$

8 $N = \sum_{i=1}^{t} n_i$, $SST = \sum_{i=1}^{t}\sum_{j=1}^{n_i} x_{ij}^2 - \dfrac{G^2}{N}$, $SS_{t_r} = \sum_{i=1}^{t} \dfrac{T_i^2}{n_i} - \dfrac{G^2}{N}$

9. $b_0 = \bar{y} - b_1\bar{x}$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{\Sigma x_i y_i - \Sigma x_i \ \Sigma y_i/n}{\Sigma x_i^2 - (\Sigma x_i)^2/n}$$

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

$$SSR = b_1 \Sigma(x_i - \bar{x})(y_i - \bar{y}) = b_1 [\Sigma x_i y_i - \Sigma x_i \Sigma y_i/_n]$$

10. $L = \sum_{i=1}^{t} c_i\mu_i , \hat{L} = \sum_{i=1}^{t} c_i\bar{x}_i , \sigma^2_{\hat{L}} = MSE \sum_{i=1}^{t} \frac{c_i^2}{n_i}$

11. $r = \dfrac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \ \Sigma(y_i - \bar{y})^2}} = \dfrac{\Sigma x_i y_i - \Sigma x_i \Sigma y_i/n}{\sqrt{[\Sigma x_i^2 - (\Sigma x_i)^2/n][\Sigma y_i^2 - (\Sigma y_i)^2/n]}}$

12. $(1-\alpha)$ 100% Confidence interval ของ.............

$\mu_1 - \mu_2$ do $(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

$\mu_1 - \mu_2$ คือ $(\bar{x}_1 - \bar{x}_2) \pm t_{\nu, \alpha/2}\, s_p \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

$p_1 - p_1$ คือ $(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

p คือ $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

μ_D คือ $\bar{d} \pm t_{\nu, \frac{\alpha}{2}}\, s_d/\sqrt{n}$

เฉลยข้อสอบภาค 1/2526

1. 1) Chi-square, (n − 1)

2) $n \to \infty$

3) $E(\hat{\Theta}) = \theta$, $E(\hat{P}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{1}{n}(np) = p$

4) 1. goodness of fit test 2. Test for Independence 3. Test for Homogeniety

5) Type I-error, Type II-error

6) Ho : $\mu_1 = \mu_2$, H : $\mu_1 > \mu_2$, one-tailed test

7) β_1, X

8) เชิงเส้นตรง , (n -2)

9) F-distribution , $(n_1 - 1, n_2 - 1)$

10) ด้วยความเชื่อมั่น 95% อาจกล่าวได้ว่าช่วง (.295, 2.785) จะรวมค่าที่แท้จริงของ σ_1^2/σ_2^2 ไว้ด้วย เราสรุปไม่ได้ว่า $\sigma_1^2 > \sigma_2^2$ เพราะ Lower limit = .295 < 1 แสดงว่า σ_1^2 อาจน้อยกว่า σ_2^2 ได้

2. ให้ $\mu_D = \mu_1 - \mu_2$ โดยที่ μ_1 = น้ำหนักเฉลี่ยก่อนลดและ

μ_2 = น้ำหนักเฉลี่ยหลังลด

n.

1) Ho: $\mu_D = 0$

2) $H_1 : \mu_D > 0$

3) a = .05

4) CR : $T > t_{4, .05} = 2.132$

5) d_i = (น้ำหนักก่อนลดของคนที่ i − น้ำหนักหลังลดของคนที่ i) , i = 1, ..., 5

d_i : −1, 12, 8, 12, 6

$\Sigma d_i = 37$. $\Sigma d_i^2 = 389$

$\bar{d} = 37/5 = 7.4$

$$s_d^2 = \frac{1}{n-1}[\Sigma d_i^2 - (\Sigma d_i)^2/n] = \frac{1}{4}[389 - (37)^2/5] = \frac{115.2}{4} = 28.8$$

$$s_d = \sqrt{28.8} = 5.367$$

$$t_c = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{7.4}{5.367/\sqrt{5}} = 3.083$$

6. $\because$ $t_c = 3.083 > 2.132$ เราปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือสรุปได้ว่าวิธีการลดอาหารวิธีนี้ได้ผลใน 2 เดือน

ข. 95% Confidence interval ของ μ_D คือ $\bar{d} \pm t_{\nu, \alpha/2}\, s_d/\sqrt{n}$

$$= \bar{d} \pm t_{4, .025}\, s_d/\sqrt{n}$$

$$= 7.4 \pm 2.776 \left(\frac{5.367}{\sqrt{5}}\right)$$

$$= 7.4 \pm 6.663 = (.737, 14.063)$$

3. ก. $b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - x)^2} = \frac{5.9}{20.5} = .29 = \hat{\beta}_1$

$$b_0 = \bar{y} - b_1\bar{x} = 2.75 - (.29)(6) = 1.01 = \hat{\beta}_0$$

ข. $\hat{y} = 1.01 + .29x$

ค. $\hat{\mu}_{7.5} = 1.01 + .29(7.5) = 3.2045$ นั่นคือนักศึกษาที่ได้คะแนนวิชาคณิตศาสตร์พื้นฐาน 7.5 คะแนน จะได้ GPA โดยเฉลี่ย 3.18

$\hat{y}_{5.5} = 1.01 + .29(5.5) = 2.6193$

ง. $H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

$a = .01$

CR : $F > f_{(1, 6), .01} = 13.75$

หา f_c จาก ANOVA table

SST $= \Sigma(y_i - \bar{y})^2 = 1.92$

SSR $= b_1\Sigma(x_i - \bar{x})(y_i - \bar{y}) = .29(5.9) = 1.711$

ANOVA

S.V.	df	SS	MS	f
Regression	1	SSR = 1.711	1.711	$f_c = \frac{1.711}{.035} = 48.5142$
Error	6	SSE = .209	.035 = MSE	
Total	7	SST = 1.92		

$f_c = 48.5142 > 13.75$ เราปฏิเสธ H_0 ที่ $\alpha = .01$ สรุปได้ว่าคะแนนคณิต-ศาสตร์พื้นฐานสามารถใช้ทำนาย GPA ได้

จ. $\hat{\sigma}_{Y/X} = MSE = .035$

4. μ_i = ส่วนสูงเฉลี่ยเมื่อใช้ปุ๋ยที่ i

t = จำนวน treatment = 4

$n_1 = n_2 = n_3 = n_4 = 5$, $N = 20$

$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$

H_1 : มี mean อย่างน้อย 1 คู่ไม่เท่ากัน

$\alpha = .01$

CR : $F > f_{(3,\ 16),\ .01} = 5.29$

หา f_c จาก ANOVA table

$$SST = \sum_{i=1}^{4}\sum_{j=1}^{5} x_{ij}^2 - \frac{G^2}{N} = 26.2 - \frac{(20.6)^2}{20} = 26.2-21.218 = 4.982$$

$$SSt_r = \sum_{i=1}^{4} \frac{T_i^2}{n_i} - \frac{G^2}{N} = \frac{124.2}{5} - 21.218 = 24.84-21.218 = 3.622$$

ANOVA

S.V.	df	ss	MS	f
Treatments	3	**SStr** = 3.622	1.2073	$f_c = \frac{1.2073}{.085} \approx 14.2035$
Error	16	SSE = 1.36	.085 = MSE	
Total	19	SST = 4.982		

$\because f_c = 14.2035 > 5.29$ เราปฏิเสธ H_0 ที่ $\alpha = .01$ สรุปว่าปุ๋ยทั้ง 4 ชนิดดีไม่เท่ากันในการช่วยการเจริญเติบโตของพืชชนิดนี้

$L = \frac{\mu_A + \mu_B}{2} - \frac{\mu_C + \mu_D}{2}$ เป็น Contrast เพราะ $\sum_{i=1}^{4} c_i = 0$ โดยที่ $c_1 = c_2 = \frac{1}{2}, c_3 = c_4 = -\frac{1}{2}$

$$\hat{L} = \frac{\hat{\mu}_A + \hat{\mu}_B}{2} - \frac{\hat{\mu}_C + \hat{\mu}_D}{2} = \frac{\bar{x}_A + \bar{x}_B}{2} - \frac{\bar{x}_C + \bar{x}_D}{2} = \left(\frac{.56 + 1.12}{2}\right) - \left(\frac{1.68 + 76}{2}\right)$$

$$= -.38$$

$$\hat{\sigma}^2_{\hat{L}} = MSE \sum_{i=1}^{4} \frac{c_i^2}{n_i} = .085 \left(\frac{1}{5}\right) 4 \left(\frac{1}{4}\right) = .017$$

5. ให้ p_1 และ p_2 เป็นสัดส่วนของชายที่เห็นด้วยและของหญิงที่เห็นด้วยตามลำดับ

$H_0 : p_1 = p_2$

$H_1 : p_1 > p_2$

$\alpha = .05$

$CR : Z > z_{.05} = 1.645$

$$t_c = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$n_1 = 200, \quad x_1 = 120, \quad \hat{p}_1 \approx \frac{x_1}{n_1} = \frac{120}{200} = .6$

$n_2 = 500,\ x_2 = 240,\ \hat{p}_2 = \frac{x_2}{n_2} = \frac{240}{500} = .48$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{120+240}{200+500} = \frac{360}{700} = .51$$

$$\therefore\ t_c = \frac{.6 - .48}{\sqrt{(.51)(.49)\left(\frac{1}{200} + \frac{1}{500}\right)}} = \frac{.12}{.0418} = 2.869$$

$t_c = 2.869 > 1.645$ เราปฏิเสธ H_0 ที่ $\alpha = .05$ สรุปว่าสัดส่วนของชายที่เห็นด้วยมากกว่าของหญิงที่เห็นด้วย

หมายเหตุ การทดสอบสมมติฐานในข้อนี้จะใช้ $\chi^2 - \text{test}$ ไม่ได้เพราะ $H_1 : p_1 > p_2$

ถ้า $H_1 : p_1 \neq p_2$ เราอาจใช้ $t - \text{test}$ หรือ $\chi^2 - \text{test}$ ก็ได้

95% confidence interval ของ $p_1 - p_2$ คือ

$$(\hat{p}_1 - \hat{p}_2) \pm z_{.025}\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$= (.6 - .48) \pm 1.96\sqrt{\frac{(.6)(.4)}{200} + \frac{(.48)(.52)}{500}}$$

$$= (.12 \pm .0808)$$

$$= (.0392, .2008)$$

6. H_0 : ระดับการศึกษาและระดับความสนใจเป็นอิสระต่อกัน

H_1 : ระดับการศึกษาและระดับความสนใจไม่เป็นอิสระต่อกัน

$\alpha = .05$

CR : $X^2 > \chi^2_{4,\ .05} = 9.488$, df = $(3-1)(3-1) = 4$

$$\chi^2_c = \sum_i^r \sum_j^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \quad \text{โดยที่ } e_{ij} = \frac{R_iC_j}{n}$$

o_{ij} (e_{ij})				
	14 $\left(\frac{83 \times 45}{200}\right)$	37 $\left(\frac{83 \times 96}{200}\right)$	32 $\left(\frac{83 \times 59}{200}\right)$	83
	19 $\left(\frac{78 \times 45}{200}\right)$	42 $\left(\frac{78 \times 96}{200}\right)$	17 $\left(\frac{78 \times 59}{200}\right)$	78
	12 $\left(\frac{39 \times 45}{200}\right)$	17 $\left(\frac{39 \times 96}{200}\right)$	10 $\left(\frac{39 \times 59}{200}\right)$	39
	45	96	59	200

χ^2_c = 7.4646 < 9.488 เราไม่สามารถปฏิเสธ H_0 ที่ a = .05 สรุปว่าระดับการศึกษาและระดับความสนใจเป็นอิสระต่อกัน

7. $H_0 : Y \sim \text{Binomial}\ (n = 3,\ p = \frac{1}{4})$

H_1 : y ไม่มีการแจกแจงเป็น Binomial distribution

a = .05

CR : $X^2 > \chi^2_{2\ .05}$ = 5.991, df = จำนวนเทอมที่รวมกันเป็น $\chi^2_c - 1$ = 3-l = 2

$$\chi^2_c = \Sigma \frac{(o_i - e_i)^2}{E_i}$$

n = 128

Y	0	1	2	3	รวม
$\Pr\{Y = y\} = p_i$	27/64	27/64	9/64	1/64	1
$e_i = np_i$	54	54	18	2	128

e_4 = 2 < 5 ต้องรวมกับ e_3 ให้ได้อย่างน้อย 5

Y	0	1	≥ 2	รวม
o_i	42	62	24	128
e_i	54	54	20	128

$$\chi^2_c = \frac{(42-54)^2}{54} + \frac{(62-54)^2}{54} + \frac{(24-20)^2}{20}$$

$$= 2.667 + 1.185 + .8 = 4.652$$

$\because$ χ^2_c = 4.652 < 5.991 เราไม่สามารถปฏิเสธ Ho ที่ a = .05 ได้ สรุปว่า Y มีการแจกแจงเป็น Binomial distribution ที่มี $p = \frac{1}{4}$