ดังนั้น เราจึงปฏิเสธสมมุติฐานหลัก $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$

เมื่อ $|t_c| > t_{n+m-2, 1-\alpha/2}$

หรือปฏิเสธสมมุติฐานหลักเมื่อ

$$\left| \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{\sqrt{\frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}} \cdot \sqrt{\frac{1}{n} + \frac{1}{m}}} \right| > t_{n+m-2, 1-\alpha/2}$$

และถ้ากำหนดให้ $s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$ เราจะปฏิเสธสมมมุติฐานหลักเมื่อ

$$\left| \frac{(\bar{x} - \bar{y}) - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \right| > t_{n+m-2, 1-\alpha/2}$$

แต่ $\because \mu_x = \mu_y$ หรือ $\mu_x - \mu_y = 0$ ดังนั้นเราจึงปฏิเสธสมมุติฐานหลักเมื่อ

$$\left| \frac{\bar{x} - \bar{y}}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \right| > t_{n+m-2, 1-\alpha/2} \text{ หรือ } |\bar{x} - \bar{y}| > t_{n+m-2, 1-\alpha/2} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

ในทำนองเดียวกัน เราสามารถพิสูจน์ได้ว่า

1. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x > \mu_y$

   ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} > t_{n+m-2, 1-\alpha} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

2. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

   ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} < t_{n+m-2, \alpha} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

ขอเว้นการพิสูจน์ไว้เป็นแบบฝึกหัด

สำหรับการหา β และ π ให้ปฏิบัติเช่นเดียวกับตอน 7.11 แต่ใช้ภาพ 6.5-6.8 กล่าวคือให้ใช้ d = $\frac{|\mu_x - \mu_y|}{2\sigma}$ สำหรับกรณี two tail, d = $\frac{\mu_y - \mu_x}{2\sigma}$ สำหรับกรณีหางซ้าย และ d = $\frac{\mu_x - \mu_y}{2\sigma}$ สำหรับกรณีหางขวา ทั้งนี้ให้ใช้ขนาดตัวอย่างสำหรับเลือกโค้งเป็น n โดยที่ n = 2n − 1 = 2m − 1 และ $\hat{\sigma}^2 = s_p^2$

อนึ่ง สำหรับการทดสอบสมมุติฐาน H_0

$H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$

$H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x > \mu_y$ หรือ $\Delta > 0$

$H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$ หรือ $\Delta < 0$

สำหรับกรณีไม่ทราบค่า σ_x^2 และ σ_y^2 และไม่อาจถือได้ว่า $\sigma_x^2 = \sigma_y^2 = \sigma^2$ นั้นเราสามารถทดสอบได้โดยอาศัย t-test ซึ่งเป็นค่าประมาณที่พัฒนาขึ้นมาโดย BL. Welch เหตุที่เป็นค่าประมาณเพราะเราประมาณ df ของ t ขึ้นมา

df ดังกล่าวคือ

$$\nu = \frac{(s_x^2/n + s_y^2/m)^2}{\{(s_x^2/n)/(n+1) + (s_y^2/m)/(m+1)\}} - 2$$

ดังนั้น สำหรับ

1) $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$

จะปฏิเสธสมมติฐานหลักเมื่อ

$$|\bar{x} - \bar{y}| > t_{\nu, 1-\alpha/2}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

หรือ

$$t' = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} > t_{\nu, 1-\alpha/2}$$

2) $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x > \mu_y$

จะปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} > t_{\nu, 1-\alpha}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

หรือ $t' = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_x^2}{n} + \dfrac{s_y^2}{m}}} > t_{v,1-\alpha}$

3. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

จะปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} < t_{v,\alpha}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

หรือ $t' = \dfrac{\bar{x} - \bar{y}}{\sqrt{\dfrac{s_x^2}{n} + \dfrac{s_y^2}{m}}} < t_{v,\alpha}$

ตัวอย่าง 7.4 บริษัท ก. กำลังตัดสินใจเลือกใช้แบตเตอรี่ 2 ยี่ห้อ ทั้งนี้แบตเตอรี่ยี่ห้อ X มีราคาแพงกว่ายี่ห้อ Y เล็กน้อย และบริษัทตั้งใจจะใช้ยี่ห้อ X เว้นแต่ยี่ห้อ Y จะมีอายุใช้งานเฉลี่ยสูงกว่าอย่างมีนัยสำคัญ

เพื่อประกอบการตัดสินใจทางบริษัทได้สุ่มตัวอย่างแบตเตอรี่จากทั้งสองบริษัทมายี่ห้อละ 20 หม้อ และบันทึกอายุใช้งาน ปรากฏข้อมูลดังนี้

X	45.1	47.2	43.6	41.9	50.4	46.7	45.6	44.9	42.6
	49.1	49.1	43.7	38.5	40.9	42.7	44.7	46.1	36.9
	47.5								
Y	45.2	37.6	45.1	36.9	38.3	39.1	45.7	44.3	40.5
	43.7	35.7	40.8	38.5	41.1	39.6	42.3	41.8	38.6
	44.6								

ก) จากข้อมูลน่าจะสรุปผลอย่างไร ทั้งนี้ให้ถือว่าอายุใช้งานมีการแจกแจงแบบปกติเมื่อถือว่า $\sigma_x^2 = \sigma_y^2 = \sigma^2$

ข. เพื่อให้ $\beta = .10$ เมื่ออายุใช้งานเฉลี่ยของแบตเตอรี่ยี่ห้อ X สูงกว่ายี่ห้อ Y ถึง 5.2 ชั่วโมง อยากทราบว่าควรสุ่มตัวอย่างมาตรวจสอบเท่าไร ทั้งนี้ให้ถือว่า s_x^2 และ s_y^2 เป็นตัวประมาณค่าที่ดีของ σ_x^2 และ σ_y^2 ตามลำดับ

วิธีทำ

ก) $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x > \mu_y$ หรือ $\Delta = 0$

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$(\bar{x} - \bar{y}) > t_{n+m-2,1-\alpha}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

เมื่อ

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

จากข้อมูลพบว่า

$$s_x^2 = \frac{1}{19}\left(40379.46 - \frac{(896)^2}{20}\right) = 12.56$$

$$\bar{x} = 896/20 = 44.8$$

$$s_y^2 = \frac{1}{19}\left(34044.4 - \frac{(823)^2}{20}\right) = 9.37$$

$$\bar{y} = 823/20 = 41.15$$

$$(\bar{x} - \bar{y}) = (44.8 - 41.15)$$

$$= 3.65$$

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

$$= \frac{19(12.56) + 19(9.37)}{38}$$

$$s_p^2 = 10.965$$

$$\therefore\ s_p = 3.311$$

$$t_{n+m-2,1-\alpha}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = t_{38,.950}\, s_p \sqrt{\frac{1}{20} + \frac{1}{20}}$$

$$= (1.684)(3.311)(0.316)$$

$$= 1.763$$

จะเห็นว่า $(\bar{x} - \bar{y}) > 1.763$ ดังนั้นเราจึงปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ 5% และเชื่อว่าแบตเตอรี่ยี่ห้อ X มีอายุเฉลี่ยสูงกว่า

ข) เมื่อถือว่า s_x^2 และ s_y^2 เป็นตัวประมาณค่าที่ดีของ σ_x^2 และ σ_y^2 ตามลำดับ ดังนั้นจึงคำนวณหาค่า $\beta(\Delta)$, $\Pi(\Delta)$ และคำนวณหาขนาดตัวอย่างโดยอาศัยสมการความเสี่ยงประเภทที่ 2 ของ Normal Test ดังนี้

จาก

$$\beta(\Delta) = \Pr\{Z \leq Z_{1-\alpha} - \frac{\Delta}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}}}\}$$

ดังนั้น

$$n = m = \frac{(Z_{1-\alpha} - Z_\beta)^2(\hat{\sigma}_x^2 + \hat{\sigma}_y^2)}{\Delta^2}$$

กำหนดให้ $\alpha = .05, \beta = .100, \Delta = 5.2$

ดังนั้น

$$n = m = \frac{(1.64 + 1.28)^2(12.56 + 9.37)}{(5.2)^2} = 6.9 \simeq 7$$

ตัวอย่าง 7.5 โรงงานอุตสาหกรรมซึ่งใช้น้ำมันเชื้อเพลิงเป็นหลักในการเดินเครื่องจักรสงสัยว่าความสิ้นเปลืองเชื้อเพลิงในขณะเดินเครื่องจักรกะกลางวันน่าจะน้อยกว่ากะกลางคืน

จากการบันทึกข้อมูลการใช้จ่ายน้ำมัน 30 วัน (กะกลางวัน 06.00 น. - 18.00 น. กะกลางคืน 18.00 น. ถึง 06.00 น.) ปรากฏข้อมูลดังนี้

กะกลางวัน	กะกลางคืน
X = ความสิ้นเปลืองน้ำมันกะกลางวัน	Y = ความสิ้นเปลืองน้ำมันกะกลางคืน
$\Sigma x_i = 7500$ แกลลอน	$\Sigma y_i = 7860$ แกลลอน
$\Sigma x_i^2 = 1881320$ แกลลอน	$\Sigma y_i^2 = 2068000$ แกลลอน

ก) จงทดสอบสมมุติฐาน ณ. ระดับนัยสำคัญ 1%

$H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

ข. ถ้า $\Delta = \mu_x - \mu_y = 2$ อยากทราบว่าควรเก็บบันทึกข้อมูลกี่วันจึงจะเหมาะสมที่จะควบคุมให้ $\alpha = \beta = .05$ ให้ถือว่า s_p^2 เป็นตัวประมาณค่าที่ดีของ σ^2

วิธีทำ

ก) $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ เมื่อ

$$(\bar{x} - \bar{y}) < t_{n+m-2,\alpha}\, s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$\bar{x} = 7500/30 = 250, \bar{y} = 7860/30 = 262, \bar{x} - \bar{y} = -12$$

$$s_x^2 = \frac{1}{29}\left(1881320 - \frac{(7500)^2}{30}\right) = 217.931,\ s_x = 14.76$$

$$s_y^2 = \frac{1}{29}\left(2068000 - \frac{(7860)^2}{30}\right) = 299.31, s_y = 17.30$$

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2} = \frac{(29)(217.931) - (29)(299.31)}{58}$$

$$= 258.62$$

$$s_p = 16.08$$

$$s_p\sqrt{\frac{1}{n} + \frac{1}{m}} = 16.08\sqrt{\frac{1}{30} + \frac{1}{30}} = 4.512$$

$$t_{n+m-2,\alpha} = t_{58,.05} = -1.671$$

$$t_{n+m-2,\alpha}\, s_p\sqrt{\frac{1}{n} + \frac{1}{m}} = (-1.671)(4.152)$$

$$= -6.938$$

จะเห็นว่า $(\bar{x} - \bar{y}) < -6.938$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ 5% และเชื่อว่าความสิ้นเปลืองน้ำมันเชื้อเพลิงในกะกลางวันมีน้อยกว่ากะกลางคืน

ข) จาก Normal Test

$$\beta(\Delta) = \Pr\{Z > Z_\alpha - \frac{\Delta}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}}\}$$

$$= \Pr\{Z > Z_\alpha - \frac{\Delta}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}\}$$ เมื่อ $\sigma_x^2 = \sigma_y^2 = \sigma^2 = s_p^2$

$$\Rightarrow \qquad Z_{1-\beta} = Z_\alpha - \frac{\Delta}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

$$\Rightarrow \qquad n = m = \frac{(Z_\alpha - Z_{1-\beta})^2 s_p^2}{\Delta^2}$$

$$= \frac{(-1.64 - 1.64)^2(258.62)}{4}$$

$$= 678.72 \simeq 679$$

ตัวอย่าง 7.6 บริษัทผู้จำหน่ายน้ำมันหล่อลื่นเชื่อว่าการโฆษณาสินค้าจะมีผลให้ความต้องการซื้อสินค้าสูงขึ้น ข้อมูลต่อไปนี้แสดงปริมาณความต้องการซื้อน้ำมันหล่อลื่นรายสัปดาห์ก่อนโฆษณาและหลังโฆษณา

ความต้องการซื้อ (1000 แกลลอน)

ก่อนโฆษณา (X)	62.3	57.6	55.1	63.1	64.1	58.1	59.6	62.7	60.5	56.9
หลังโฆษณา (Y)	61.5	67.3	62.7	61.7	60.4	66.8	63.7	59.7	67.9	63.8

วิธีทำ $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} = t_{\nu,\alpha}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

$$\nu = \frac{(s_x^2/n + s_y^2/m)^2}{\{(s_x^2/n)^2/(n+1)\} + \{(s_y^2/m)^2/(m+1)\}} - 2$$

$$s_x^2 = \frac{1}{9}(36082.4 - \frac{600^2}{10}) = 9.156,\ s_x = 3.026$$

$$\bar{x} = 600/10 = 60$$

$$s_y^2 = \frac{1}{9}(41057.2 - \frac{640^2}{10}) = 10.8,\ s_y = 3.286$$

$$\bar{y} = 640/10 = 64$$

$$\bar{x} - \bar{y} = 60 - 64 = -4$$

$$\nu = \frac{(s_x^2/n + s_y^2/m)^2}{\{(s_x^2/n)/(n+1)\} + \{(s_y^2/m)/(m+1)\}} - 2$$

$$= \frac{(9.156/10 + 10.8/10)^2}{\{(9.15/10)^2/11\} + \{(10.8/10)^2/11\}} - 2$$

$$= 21.851701 - 2 = 19.851708 \simeq 20$$

$$t_{\nu,\alpha}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} = t_{20,.05}\sqrt{\frac{9.156}{10} + \frac{10.8}{10}}$$

$$= (-1.725)(1.413) = -2.437$$

จะเห็นว่า $\bar{x} - \bar{y} < -2.437$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ 5% และเชื่อว่าการโฆษณามีผลให้ความต้องการซื้อเพิ่มสูงขึ้น

7.1.3 PAIRED COMPARISON (PC)

ในกรณีที่เราต้องการเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มประชากร 2 กลุ่มนั้น ในบางครั้งเราไม่อาจถือได้ว่ากลุ่มตัวอย่างที่สุ่มมาจากประชากรทั้งสองเป็นอิสระต่อกัน ทั้งนี้เพราะมีปัจจัยร่วมหลายประการที่หน่วยตัวอย่างในกลุ่มตัวอย่างทั้งสอง เช่นการวิเคราะห์เปรียบเทียบสารที่ทำให้เนื้อเปื่อยนุ่มโดยใช้เนื้อชิ้นเดียวกัน หรือการเปรียบเทียบประสิทธิภาพของยาแก้หวัด 2 ชนิดในหนูทดลองตัวเดียวกัน (ต่างวาระกัน) เป็นต้น การกระทำดังนี้ จะถือว่ากลุ่มตัวอย่างจากการทดลองทั้งสองครั้งนี้เป็นอิสระต่อกันไม่ได้ เพราะมีความผันแปรภายในหน่วยทดลองร่วมกันอยู่ ในทำนองเดียวกัน ถ้าเราวางแผนการทดลองโดยควบคุมตัวแปรภายนอกได้ไม่ดีพอ เช่นการทดลองเปรียบเทียบคุณภาพของสารเคมี 2 ชนิด แต่ไม่ควบคุมสิ่งแวดล้อมขณะทดลองให้มีลักษณะเดียวกัน เช่นอุณหภูมิและความชื้น ก็ย่อมมีผลให้เราไม่อาจถือว่ากลุ่มตัวอย่างทั้งสองเป็นอิสระต่อกัน จนกระทั่งใช้ค่าตัวแทนคือ $\bar{X}$ และ $\bar{Y}$ มาเปรียบเทียบกันได้ ดังนั้นในกรณีเช่นนี้ เราจำเป็นต้องวางแผนการวิเคราะห์เสียใหม่ โดยแทนที่จะเปรียบเทียบค่าตัวแทนของกลุ่มโดยตรง ซึ่งอาจมีผลเสียเพราะมี error ปะปนอยู่ในระหว่างการทดลองให้ดำเนินการวิเคราะห์โดยเปรียบเทียบผลการทดลองทีละคู่ในรูปของผลต่าง $d_i = x_i - y_i$; $i = 1, 2, \ldots, n$ โดยถือว่าคู่ลำดับ (x_i, y_i) คือค่าของตัวแปรจากการทดลองทั้งสองครั้งภายใต้เงื่อนไขหรือสิ่งแวดล้อมขณะทดลองเดียวกันเช่น เนื้อชิ้นเดียวกัน หนูทดลองตัวเดียวกัน ระดับอุณหภูมิเดียวกัน ระดับความชื้นเดียวกัน ฯลฯ เป็นต้น และในที่นี้จะศึกษาเฉพาะกรณีที่ถือว่า $d_i \sim N(\mu_d, \sigma_d^2)$ ก่อนโดยถือว่า σ_d^2 เป็นพารามิเตอร์ที่เราไม่อาจทราบค่าได้ และมีสมมุติฐานดังนี้

$H_0 : \mu_x - \mu_y = \mu_d = 0$ vs $H_1 : \mu_d \neq 0$

การสร้าง test ให้พัฒนาดังนี้

จากเงื่อนไขข้างต้นแสดงว่า $\omega = \{\mu_d, \sigma_d^2; \mu_d = 0, \sigma_d^2 > 0\}$ และ

$$\Omega = \{\mu_d, \sigma_d^2; -\infty < \mu_d < \infty, \sigma_d^2 > 0\}$$

ดังนั้น เมื่อทำการทดลอง n ครั้ง เราจะพบว่า

$$L\omega = \prod_i^n f_{D_i}(d_i; \mu_d = 0, \sigma_d^2)$$

$$= \left(\frac{1}{\sigma_d\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma_d^2}\sum_i^n d_i^2\right\}$$

$$L\Omega = \prod_i^n f_{D_i}(d_i; \mu_d, \sigma_d^2)$$

$$= \left(\frac{1}{\sigma_d\sqrt{2\pi}}\right)^n \exp\left\{-\frac{1}{2\sigma_d^2}\sum_i^n (d_i - \mu_d)^2\right\}$$

$$\frac{\partial}{\partial \sigma_d} \ln L\omega = 0 \Rightarrow 0 - \frac{n}{\sigma_d} - 0 + \frac{1}{\sigma_d^3}\sum_i^n d_i^2 = 0$$

$\Rightarrow$ $\hat{\sigma}_d^2 = \frac{1}{n}\sum_i^n d_i^2$ หรือ $\hat{\sigma}_d^2 = \frac{1}{n}\sum_i^n (d_i - \mu_d)^2$ เมื่อ $\mu_d = 0$

เมื่อแทนที่ $\hat{\sigma}_d^2$ ลงใน σ_d^2 ของ $L\omega$ จะได้ Sup $L\omega$ ดังนี้

$$\hat{L}\omega = \left(\frac{n}{2\pi\Sigma d_i^2}\right)^{n/2} \exp(-n/2)$$

จาก $L\Omega$ จะพบว่า

$$\frac{\partial}{\partial \mu_d} \ln L\Omega = 0 \Rightarrow 0 + \frac{1}{\sigma_d^2}\sum_i^n (d_i - \mu_d) = 0 \Rightarrow \mu_d = \bar{d}$$

$$\frac{\partial}{\partial \sigma_d} \ln L\Omega \Rightarrow 0 - \frac{n}{\sigma_d} - 0 + \frac{1}{\sigma_d^3}\sum_i^n (d_i - \mu_d)^2 = 0$$

เมื่อแทนค่า μ_d ด้วย $\bar{x}$ จะพบว่า $\hat{\sigma}_d^2 = \frac{1}{n}\sum_i^n (d_i - \bar{d})^2$

ดังนั้น

$$\hat{L}\Omega = \left(\frac{n}{2\pi\Sigma(d_i - \bar{d})^2}\right)^{n/2} \exp(-n/2)$$

$\Rightarrow$ $\lambda = \hat{L}\omega/\hat{L}\Omega = \{\sum_i^n (d_i - \bar{d})^2/\sum_i^n d_i^2\}^{-n/2}$

$$= \left\{ \frac{\sum_{i}^{n} (d_i - \mu_d)^2}{\sum_{i}^{n} (d_i - \bar{d})^2} \right\}^{-n/2} ; \ \mu_d = 0$$

$$= \left\{ \frac{\sum_{i}^{n} (d_i - \bar{d}) + n(\bar{d} - \mu_d)^2}{\sum_{i}^{n} (d_i - \bar{d})^2} \right\}^{-n/2}$$

$$= \left\{ 1 + \frac{n(\bar{d} - \mu_d)^2}{(n-1)s_d^2} \right\}^{-n/2}$$

$$= \left\{ 1 + \frac{(\bar{d} - \mu_d)^2}{\sigma_d^2/n} \bigg/ \frac{(n-1)s_d^2}{\sigma_d^2} \right\}^{-n/2}$$

$$= \left\{ 1 + \frac{Z^2}{\chi^2_{n-1}} \right\}^{-n/2}$$

$$= \left\{ 1 + \frac{t^2}{n-1} \right\}^{-n/2}$$

โดยที่ $t = \dfrac{Z\sqrt{n-1}}{\sqrt{\chi^2_{n-1}}} = \dfrac{\sqrt{n-1}\,.\,(\bar{d} - \mu_d)}{\sigma_d/\sqrt{n}} \bigg/ \dfrac{(n-1)s_d^2}{\sigma_d^2}$

$$= \frac{\bar{d}\sqrt{n}}{s_d}$$

เราจะปฏิเสธสมมุติฐานหลัก เมื่อ $\lambda < k$ หรือ

$$\left(1 + \frac{t^2}{n-1}\right)^{-n/2} < k$$

$\Rightarrow \quad t^2 > (n-1)(k^{-2/k} - 1) = k$

$\Rightarrow \quad |t| > k$ เมื่อ $t = \dfrac{\bar{d}\sqrt{n}}{s_d} = \dfrac{\bar{d}}{s_d/\sqrt{n}}$

ดังนั้น เราจะปฏิเสธสมมุติฐานหลักเมื่อ $|t| > k$ โดยที่ k คือค่า Ordinate ใต้โค้ง t ในลักษณะที่ k สามารถปรับค่าไปมาได้จนกระทั่งทำให้พื้นที่ใต้โค้งในเขตวิกฤติเท่ากับ α

ปัญหาคือ $k = ?$

จาก $\quad \alpha = \Pr\{|t| > k \mid \mu_d = 0\}$

$\Rightarrow \quad k = t_{n-1,1-\alpha/2}$

นั่นคือ เราจะปฏิเสธสมมุติฐานหลัก เมื่อ $\left|\dfrac{\bar{d}\sqrt{n}}{s_d}\right| > t_{n-1,1-\alpha/2}$

เมื่อพิจารณาเปรียบเทียบระหว่าง test สำหรับ $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$ กรณีเมื่อไม่ทราบค่า σ_x^2, σ_y^2 แต่ถือว่า $\sigma_x^2 = \sigma_y^2 = \sigma^2$ และ $H_0 : \mu_d = \mu_x - \mu_y = 0$ vs $H_1 : \mu_d = \mu_x - \mu_y \neq 0$ เมื่อไม่ทราบค่า σ_d^2 จะพบข้อดีข้อเสียของ Paired Comparison ดังนี้

ข้อดี

1) P.C. ไม่ต้องการข้อตกลงว่า $\sigma_x^2 = \sigma_y^2 = \sigma^2$ เพราะใช้ σ_d^2 (ความจริงแล้วข้อตกลงว่า $\sigma_x^2 = \sigma_y^2$ นั้นไม่ค่อยเหมาะสมเพราะโดยปกติ σ_x^2 จะไม่เท่ากับ σ_y^2 การตกลงว่า $\sigma_x^2 = \sigma_y^2$ จึงค่อนข้างจะเป็นการบังคับสถานการณ์ให้โน้มเอียงไปสู่ลู่ทางทฤษฎีกล่าวคือ ถ้าไม่ตกลงเช่นนี้จะสร้าง test ไม่ได้ เพราะ $(n-1)s_x^2$ และ $(m-1)s_y^2$ จำเป็นต้องหารด้วย σ^2 จึงจะมีการแจกแจงแบบ χ^2 ได้ σ^2 ดังกล่าวจะได้มาก็โดยถือว่า (ตกลงว่า) σ_x^2 และ σ_y^2 มีค่าเท่ากันเท่ากับ σ^2

2) P.C. สามารถใช้ได้เสมอไม่ว่ากลุ่มตัวอย่างทั้งสองคือ $(x_1, x_2, \ldots, x_n)$ และ $(y_1, y_2, \ldots, y_n)$ จะเป็นอิสระต่อกันหรือไม่ก็ตาม

3) เมื่อมีปัจจัยภายนอกมาส่งอิทธิพลร่วม P.C. จะแก้ปัญหานี้ได้ เพราะการจับคู่เป็นการเปรียบเทียบผลต่างของตัวแปรเฉพาะในสภาพแวดล้อมหนึ่ง ๆ การจับคู่จึงช่วยขจัดหรือลดความแปรปรวนที่ปัจจัยภายนอกนำเข้ามาสู่หน่วยทดลอง ซึ่งมีผลสะท้อนให้ test มี power สูงขึ้น ถ้าเปรียบเทียบ Power Curve หรือ OC-Curve จะมองเห็นความจริงข้อนี้

จุดอ่อน

1) การใช้ P.C. จะใช้ได้เฉพาะกรณีที่ $n = m$ เท่านั้น

2) P.C. ให้ df เพียง $n - 1$ ขณะที่ Unpaired test ให้ df เป็น 2 เท่า คือ $df = n + m - 2 = 2(n - 1)$ นั่นย่อมหมายความว่า P.C. เป็น test ที่เสีย df มากกว่า ความน่าเชื่อถือจะต่ำกว่าเพราะใช้ข้อสนเทศน้อยกว่า ดังนั้นถ้าไม่จำเป็นหรือไม่มีสถานการณ์ที่บังคับให้ต้องใช้จริง ๆ เราก็ควรจะหลีกเลี่ยงการใช้ P.C.

ในทำนองเดียวกันเราสามารถพิสูจน์ได้ว่า

1) $H_0 : \mu_d = 0$ vs $H_1 : \mu_d > 0$

จะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} > t_{n-1, 1-\alpha}\, s_d/\sqrt{n}$$

2) $H_0 : \mu_d = 0$ vs $H_1 : \mu_d < 0$

จะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} < t_{n-1, \alpha}\, s_d/\sqrt{n}$$

3) $H_0 : \mu_d = \mu_0 = \Delta$ vs $H_1 : \mu_d = \mu_1 > \mu_0 = \Delta$

จะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} > \Delta + t_{n-1,1-\alpha}\, s_d/\sqrt{n}$$

4) $H_0 : \mu_d = \mu_0 = \Delta$ vs $H_1 : \mu_d = \mu_1 < \mu_0 = \Delta$

จะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} < \Delta + t_{n-1,\alpha}\, s_d/\sqrt{n}$$

5) $H_0 : \mu_d = \mu_0 = \Delta$ vs $H_1 : \mu_d \neq \mu_0 = \Delta$

จะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$|\bar{d}| > \Delta + t_{n-1,1-\alpha/2}\, s_d/\sqrt{n}$$

ตัวอย่าง 7.7 ในการทดลองเพื่อทดสอบคุณภาพของยาชนิดใหม่ว่าจะมีผลในการเพิ่มความดันหรือไม่ โดยทำการวัดความดันแก่ผู้ที่เข้ารับการทดลอง 15 คน โดยการวัดก่อนให้ยาและหลังให้ยาดังนี้

	ความดันโลหิต														
คนที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ก่อนให้ยา	116	118	120	124	128	130	131	134	136	137	137	140	143	146	147
หลังให้ยา	119	124	126	128	121	135	137	138	139	135	143	146	150	149	142

ก) จงทดสอบ ณ ระดับนัยสำคัญ 5% ว่ายามิได้มีผลให้ความดันโลหิตเพิ่มสูงขึ้น

ข) จงทดสอบสมมุติฐาน ณ ระดับนัยสำคัญ 5% สำหรับสมมุติฐาน

$H_0 : \mu_d = 5$ vs $H_1 : \mu_d < 5$

วิธีทำ $H_0 : \mu_d = 0$ vs $H_1 : \mu_d \neq 0$

ปฏิเสธสมมุติฐานหลักเมื่อ

$$|\bar{d}| > t_{n-1,1-\alpha/2}\, s_d/\sqrt{n}$$

จากข้อมูล เราสามารถหาค่าผลต่าง d ได้ดังนี้

คนที่	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ก่อนให้ยา	116	118	120	124	128	130	131	134	136	137	137	140	143	146	147
หลังให้ยา	119	124	126	128	121	135	137	138	139	135	143	146	150	149	142
ผลต่าง d	+3	+6	+6	+4	−7	+5	+6	+4	+3	−2	+6	+6	+7	+3	−5

$$s_d^2 = \frac{1}{14}\left(391 - \frac{(45)^2}{15}\right) = 18.285,\ s_d = 4.276$$

$$|\bar{d}| = 45/15 = 3$$

$$t_{n-1,1-\alpha/2}\, s_d/\sqrt{n} = t_{14,.975}\, s_d/\sqrt{n}$$

$$= \frac{(2.145)(4.276)}{\sqrt{15}} = 2.368$$

จะเห็นว่า $|\bar{d}| > 2.368$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก และเชื่อว่ายาชนิดใหม่จะมีผลให้ความดันเปลี่ยนแปลงไปจากระดับปกติ

ข) $H_0 : \mu_d = 5$ vs $H_1 : \mu_d < 5$

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} < \Delta + t_{n-1,\alpha}\, s_d/\sqrt{n}$$

$$\bar{d} = 3$$

$$\Delta + t_{n-1,\alpha}\, s_d/\sqrt{n} = 5 - (1.761)(4.276)/\sqrt{15} = 3.055$$

จะเห็นว่า $\bar{d} < 3.055$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α และเชื่อว่ายาชนิดใหม่มีผลให้ความดันเพิ่มสูงขึ้นแต่ไม่ถึง 5

ตัวอย่าง 7.8 จากการทดลองเปรียบเทียบคุณภาพในการซักล้างของผงซักฟอก 2 ชนิด โดยการทดลองซักเสื้อที่มีความสกปรกแบบเดียวกัน ข้อมูลปริมาณสิ่งสกปรกที่ถูกซักล้างออกปรากฏดังนี้

	คราบความสกปรกที่หลุดออกมา (กรัม)								
เสื้อตัวที่	1	2	3	4	5	6	7	8	9
ผงซักฟอก X	41	47	28	35	22	39	53	48	33
ผงซักฟอก Y	36	41	27	32	18	37	48	45	29

วิธีทำ

$H_0 : \mu_d = 0$ vs $H_0 : \mu_d > 0$

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\bar{d} > t_{n-1,1-\alpha}\, s_d/\sqrt{n}$$

จากข้อมูลพบว่าผลต่าง $d_i = x_i - y_i$; $i = 1, 2, ..., 9$ ปรากฏดังนี้

คือ 5, 6, 1, 3, 4, 2, 5, 3, 4

$$s_d^2 = \frac{1}{8}(141 - \frac{33^2}{9}) = 2.5,\ s_d = 1.58,\ \bar{d} = 33/9 = 3.67$$

$$t_{n-1,1-\alpha}\, s_d/\sqrt{n} = t_{8,.95}\, 1.58/\sqrt{9}$$

$$= \frac{(1.86)(1.58)}{3} = 0.98$$

จะเห็นว่า $\bar{d} > 0.98$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก และเชื่อว่าผงซักฟอก X มีคุณภาพสูงกว่า ผงซักฟอก Y

7.2 การทดสอบความแตกต่างระหว่างความแปรปรวนของกลุ่มประชากรปกติสองกลุ่ม โดยที่ถือว่าไม่ทราบค่าเฉลี่ยจริงของกลุ่มประชากรทั้งสอง [1]

$H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 \neq \sigma_y^2$

สำหรับการทดสอบสมมุติฐานนี้ถือว่า $X \sim N(\mu_x, \sigma_x^2)$, $Y \sim N(\mu_y, \sigma_y^2)$ โดยที่ไม่ทราบค่า μ_x และ μ_y

ดังนั้น $\omega = \{\mu_x, \mu_y, \sigma_x^2 = \sigma_y^2 = \sigma^2; -\infty < \mu_x <, -\infty < \mu_y < \infty, \sigma^2 > 0\}$

และ $\Omega = \{\mu_x, \mu_y, \sigma_x^2, \sigma_y^2; -\infty < \mu_x < \infty, -\infty < \mu_y < \infty, \sigma_x^2 > 0, \sigma_y^2 > 0\}$

---

[1] กรณีของ $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 \neq \sigma_y^2$ เมื่อไม่ทราบค่า μ_x และ μ_y จะขอเว้นไว้เป็นแบบฝึกหัดให้นักศึกษาสร้าง test ขึ้นเอง เพราะกระทำได้เช่นเดียวกันกับกรณีที่ไม่ทราบค่า μ_x และ μ_y

เมื่อสุ่มตัวอย่างจากกลุ่มประชากร X และ Y มา n และ m หน่วยตามลำดับ

$$L\omega = \prod_{i}^{n} f_{X_i}(x_i; \mu_x, \sigma^2) \prod_{i}^{m} f_{Y_i}(y_i : \mu_y, \sigma^2)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n+m} \exp\{-\frac{1}{2\sigma^2}\sum_{i}^{n}(x_i - \mu_x)^2\} \exp\{-\frac{1}{2\sigma^2}\sum_{i}^{m}(y_i - \mu_y)^2\}$$

$$=> \quad L\Omega = \prod_{i}^{n} f_{X_i}(x_i; \mu_x, \sigma_x^2) \prod_{i}^{m} f_{Y_i}(y_i ; \mu_y, \sigma_y^2)$$

$$= \left(\frac{1}{\sigma_x\sqrt{2\pi}}\right)^{n} \left(\frac{1}{\sigma_y\sqrt{2\pi}}\right)^{m} \exp\{-\frac{1}{2\sigma_x^2}\sum_{i}^{n}(x_i - \mu_x)^2\}$$

$$\exp\{-\frac{1}{2\sigma_y^2}\sum_{i}^{m}(y_i - \mu_y)^2\}$$

ดิฟเฟอเรนชิเอท ln $L\omega$ wrt. σ, μ_x และ μ_y จะได้ MLE ของ σ^2, μ_x และ μ_y ตามลำดับ. กล่าวคือ

$$\frac{\partial}{\partial\mu_x} \ln L\omega = 0 => 0 + \frac{1}{\sigma^2}\sum_{i}^{n}(x_i - \mu_x) + 0 = 0 => \hat{\mu}_x = \bar{x}$$

$$\frac{\partial}{\partial\mu_y} \ln L\omega = 0 \qquad 0 + 0 + \frac{1}{\sigma^2}\sum_{i}^{m}(y_i - \mu_y) = 0 => \hat{\mu}_y = \bar{y}$$

$$\frac{\partial}{\partial\sigma} \ln L\omega = 0 => 0 - \frac{n+m}{\sigma} - 0 + \frac{1}{\sigma^3}\sum_{i}^{n}(x_i - \mu_x)^2 + \frac{1}{\sigma^3}\sum_{i}^{m}(y_i - \mu_y)^2 = 0$$

$$=> \quad \sigma^2 = \frac{1}{n+m}\{\sum_{i}^{n}(x_i - \mu_x)^2 + \sum_{i}^{m}(y_i - \mu_y)^2\}$$

เมื่อแทนค่า μ_x ด้วย $\hat{\mu}_x = \bar{x}$ และแทน μ_y ด้วย $\hat{\mu}_y = \bar{y}$

$$=> \quad \sigma^2 = \frac{1}{n+m}\{(n-1)s_x^2 + (m-1)s_y^2\}$$

เมื่อแทนที่ $\hat{\mu}_x$, $\hat{\mu}_y$ และ $\hat{\sigma}^2$ ลงใน μ_x, μ_y σ^2 ของฟังก์ชั่น $L\omega$ จะได้ Sup $L\omega$ ตามต้องการ และจาก $L\Omega$

$$\frac{\partial}{\partial\mu_x} \ln L\Omega = 0 => 0 + 0 + \frac{1}{\sigma_x^2}\sum_{i}^{n}(x_i - \mu_x) - 0 = 0 => \hat{\mu}_x = \bar{x}$$

$$\frac{\partial}{\partial\mu_y} \ln L\Omega = 0 => 0 + 0 - 0 + \frac{1}{\sigma_y^2}\sum_{i}^{m}(y_i - \mu_y) = 0 => \hat{\mu}_y = \bar{y}$$

$$\frac{\partial}{\partial\sigma_x} \ln L\Omega = 0 => 0 - \frac{n}{\sigma_x} - 0 + 0 + \frac{1}{\sigma_x^3}\sum_{i}^{n}(x_i - \mu_x)^2 - 0 = 0$$

$$=> \quad \sigma_x^2 = \frac{1}{n}\sum_{i}^{n}(x_i - \mu_x)^2$$

เมื่อแทนค่า μ_x ด้วย $\bar{x}$ จะได้ MLE ของ σ_x^2 ดังนี้คือ

$$\hat{\sigma}_x^2 = \frac{1}{n} \sum_i^n (x_i - \bar{x})^2$$

$$\frac{\partial}{\partial \sigma_y} \ln L\Omega = 0 => 0 + 0 - \frac{m}{\sigma_y} - 0 - 0 + \frac{1}{\sigma_y^3} \sum_i^m (y_i - \mu_y)^2 = 0$$

=> MLE ของ σ_y^2 คือ $\hat{\sigma}_y^2 = \frac{1}{m} \sum_i^m (y_i - \bar{y})^2$

และเมื่อแทนที่ μ_x, μ_y, σ_x^2 และ σ_y^2 ด้วย $\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_x^2$ และ $\hat{\sigma}_y^2$ จะได้ Sup LΩ ตามต้องการ

โดยอาศัย MLRT เราจะปฏิเสธสมมติฐานหลักเมื่อ $= \frac{\text{Sup } L\Omega}{\text{Sup } L\Omega} < k$ และพบว่า

$$\lambda = \{\frac{m+n}{(n-1)s_x^2 + (m-1)s_y^2}\}^{(n+m)/2} / \{\frac{n}{(n-1)s_x^2}\}^{n/2} \{\frac{m}{(m-1)s_y^2}\}^{m/2}$$

$$= \{\{\frac{n+m}{(n-1)s_x^2 + (m-1)s_y^2}\}^{n/2} / \{\frac{n}{(n-1)s_x^2}\}\}^{n/2}$$

$$\{\{\frac{n+m}{(n-1)s_x^2 + (m-1)s_y^2}\}^{m/2} / \{\frac{m}{(m-1)s_y^2}\}^{m/2}\}$$

$$= \{-\frac{(n+m)(n-1)s_x^2}{n((n-1)s_x^2 + (m-1)s_y^2)}\}^{n/2} \{\frac{(n+m)(m-1)s_y^2}{m((n-1)s_x^2 + (m-1)s_y^2)}\}^{m/2}$$

$$= \{\frac{1}{(n/(n+m)) + (n(m-1)/(n+m)(n-1))s_y^2/s_x^2}\}^{n/2}$$

$$\cdot \{\frac{1}{(m(n-1)/(n+m)(m-1))s_x^2/s_y^2 + (m/(n+m))}\}^{m/2}$$

$$= \{\frac{1}{c_1 + c_2 s_y^2/s_x^2}\}^{n/2} \{\frac{1}{c_3 s_x^2/s_y^2 + c_4}\}^{m/2}$$

$$=> \quad \lambda = \{\frac{1}{c_1 + c_2/(s_x^2/s_y^2)}\}^{n/2} \{\frac{1}{c_3(s_x^2/s_y^2) + c_4}\}^{m/2} \leqslant k$$

$$U = s_x^2/s_y^2$$

$$=> \quad \lambda = (\frac{U}{c_1 U + c_2})^{n/2} (\frac{1}{c_3 U + c_4})^{m/2}$$

โดยอาศัย L'Hospital Rule เข้าช่วยเหลือจะพบว่าเมื่อ $U \to 0$ แล้ว $\lambda \to 0$ และเมื่อ $U \to \infty$ แล้ว $\lambda \to 0$ แสดงว่า λ ซึ่งเป็นฟังก์ชั่นของ U จะมีค่าน้อยลง ตรงตามนิยามซึ่งในขณะเดียวกันก็ชี้ให้เห็นว่า U จะมีค่าปรากฏในช่วงของค่าคงที่ 2 ค่า ซึ่งค่าหนึ่งมีค่าน้อยอีกค่าหนึ่งมีค่ามาก ที่เมื่อ U มีค่าใกล้ค่าทั้งสองนี้เมื่อไร เมื่อนั้น λ จะมีค่าใกล้ศูนย์ ให้ค่าคงที่ทั้งสองคือ k_1 และ k_2

ดังนั้น $\Pr\{k_1 \leqslant U \leqslant k_2 | \mu_x, \mu_y, \sigma_0^2\} = 1 - \alpha$

=> $\Pr\{k_1 \leqslant s_x^2/s_y^2 \leqslant k_2 | \mu_x, \mu_y, \sigma_0^2\} = 1 - \alpha$

$\Pr\{k_2 \leqslant s_x^2/s_y^2 < k_1 | \mu_x, \mu_y, \sigma_0^2\} = \alpha$

ดังนั้น $\Pr\{ \frac{n-1}{m-1} k_2 < \frac{(n-1)S_x^2}{(m-1)S_y^2} < \frac{n-1}{m-1} k_1 | \mu_x, \mu_y, \sigma_0^2\} = \alpha$

$$\Pr\{ \frac{n-1}{m-1} k_2 < \frac{(n-1)\, S_x^2/\sigma_0^2}{(m-1)S_y^2/\sigma_0^2} < \frac{n-1}{m-1} k_1\} = \alpha^1$$

$$\Pr\{ \frac{n-1}{m-1} k_2 < \frac{V}{W} < \frac{n-1}{m-1} k_1\} = \alpha.$$

โดยที่ $V \sim \chi^2_{(n-1)}$ และ $W \sim \chi^2_{(m-1)}$

=> $\Pr\{k_2 < \frac{V/(n-1)}{W/(m-1)} < k_1\} = \alpha$

แต่ตัวแปรสุ่ม $\frac{V/(n-1)}{W/(m-1)} \sim F_{n-1,m-1}$ ดังนั้น α จึงเป็นค่าพื้นที่ใต้โค้ง F และเพื่อความสะดวกเราจะแบ่งพื้นที่ α ออกเป็นสองส่วนเท่า ๆ กัน ส่วนละ $\alpha/2$ ดังนั้น

$$\Pr\{k_2 < \frac{V/(n-1)}{W/(m-1)} < k_1\} = \alpha$$

$$= \Pr\{F_{n-1,m-1,1-\alpha/2} < F_{n-1,m-1} < F_{n-1,m-1,\alpha/2}\}$$

=> $k_1 = F_{n-1,m-1,\alpha/2}$, $k_2 = F_{n-1,m-1,1-\alpha/2}$

ดังนั้นจากการปฏิเสธสมมุติฐานหลักเมื่อ

$$k_2 < U < k_1$$

---

[1] นำ μ_x และ μ_y ไปใช้ในฟังก์ชั่น s_x^2 และ s_y^2 แล้วโดยประมาณค่า μ_x ด้วย $\bar{x}$ และประมาณค่า μ_y ด้วย $\bar{y}$ และนำ σ_0^2 ไปใช้โดยการหาร s_x^2 และ s_y^2 ดังนั้น จึงไม่ปรากฏ "given"

นั่นคือจึงปฏิเสธสมมุติฐานหลักเมื่อ

$$F_{n-1,m-1,1-\alpha/2} < s_x^2/s_y^2 < F_{n-1,m-1,\alpha/2}$$

แต่เนื่องจากค่าของ $F_{n-1,m-1,\alpha/2}$ มีค่าน้อยมาก จนสามารถตัดทิ้งไปได้ ดังนั้น จึงปฏิเสธสมมุติฐานหลักเมื่อ

$$s_x^2/s_y^2 > F_{n-1,m-1,1-\alpha/2}$$

$$\begin{aligned}
\beta(\sigma_x^2, \sigma_y^2) &= \Pr\{s_x^2/s_y^2 < F_{n-1,m-1,1-\alpha/2} \mid \sigma_x^2, \sigma_y^2\} \\
&= \Pr\{\frac{(n-1)s_x^2/\sigma_x^2}{(m-1)s_y^2/\sigma_y^2} < \frac{(n-1)\sigma_y^2}{(m-1)\sigma_x^2} F_{n-1,m-1,1-\alpha/2}\} \\
&= \Pr\{F_{n-1,m-1} < \frac{n-1}{m-1} \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,\alpha/2}\}
\end{aligned}$$

และถ้าต้องการให้ได้ค่าที่แม่นยำให้ใช้รูปเดิม ดังนี้

จากการปฏิเสธสมมุติฐานหลัก ในระดับนัยสำคัญ α เมื่อ

$$F_{n-1,m-1,1-\alpha/2} < \frac{S_x^2}{S_y^2} < F_{n-1,m-1,\alpha/2}$$

ดังนั้นจึงยอมรับสมมุติฐานหลักเมื่อ

$$F_{n-1,m-1,\alpha/2} < \frac{S_x^2}{S_y^2} < F_{n-1,m-1,1-\alpha/2}$$

ดังนั้น

$$\begin{aligned}
\beta(\sigma_x^2, \sigma_y^2) &= \Pr\{F_{n-1,m-1,\alpha/2} < \frac{S_y^2}{S_y^2} < F_{n-1,m-1,1-\alpha/2} \mid \sigma_x^2, \sigma_y^2) \\
&= \Pr\{F_{n-1,m-1} < \frac{n-1}{m-1} \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,1-\alpha/2}\} \\
&\quad - \Pr\{F_{n-1,m-1} < \frac{n-1}{m-1} \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,\alpha/2}\}
\end{aligned}$$

ในทำนองเดียวกัน เราสามารถพิสูจน์ได้ว่า

1) $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 > \sigma_y^2$

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\frac{S_x^2}{S_y^2} < F_{n-1,m-1,1-\alpha}$$

และ

$$\beta(\sigma_x^2, \sigma_y^2) = \Pr\{F_{n-1,m-1} \leqslant \frac{n-1}{m-1} \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,1-\alpha}\}$$

2) $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 < \sigma_y^2$

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\frac{s_x^2}{s_y^2} < F_{n-1,m-1,\alpha}$$

และ

$$\beta(\sigma_x^2, \sigma_y^2) = \Pr\{F_{n-1,m-1} \geqslant \frac{n-1}{m-1} \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,\alpha}\}$$

สำหรับขนาดตัวอย่างอุตมะนั้นการคำนวณให้อาศัยสมการความเสี่ยงประเภทที่ 2 ทั้งนี้ต้องถือว่า n = m มิเช่นนั้น จะคำนวณหาค่าขนาดตัวอย่างตามต้องการได้ยาก การพิสูจน์ในส่วนนี้สามารถกระทำได้โดยง่าย จึงขอเว้นไว้เป็นแบบฝึกหัด

ตัวอย่าง 7.9 สุ่มตัวอย่างตัวต้านทานที่ผลิตโดยบริษัท X มา 30 ตัว และผลิตโดยบริษัท Y มา 40 ตัว แล้ววัดความต้านทานเพื่อเปรียบเทียบตัวต้านทานของสองบริษัทว่ามีความต้านทานแตกต่างกันหรือไม่ ปรากฏข้อมูลดังนี้

$X : \Sigma x_i = 156, \Sigma x_i^2 = 811.345, \Sigma y_i = 204, \Sigma y_i^2 = 1040.4585$

จงทดสอบ ณ ระดับนัยสำคัญ 1% สำหรับสมมุติฐานต่อไปนี้

ก. $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 \neq \sigma_y^2$

พร้อมทั้งคำนวณหาขนาดตัวอย่างที่เหมาะสมเพื่อให้ $\alpha = \beta = .01$ และเมื่อ $\sigma_x^2 = 2\sigma_y^2$

ข. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$

วิธีทำ

ก. $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 \neq \sigma_y^2$

(1) ปฏิเสธสมมุติฐานหลักเมื่อ

$$F_{n-1,m-1,1-\alpha/2} < \frac{s_x^2}{s_y^2} < F_{n-1,m-1,\alpha/2}$$

จากข้อมูล

$$s_x^2 = \frac{1}{29}(811.345 - \frac{156^2}{30}) = .005,\ s_x = .071$$

$$s_y^2 = \frac{1}{39}(1{,}040.4585 - \frac{204^2}{40}) = .0015,\ s_y = .039$$

$$s_x^2/s_y^2 = 3.333$$

$$F_{29,39,.005} \simeq .429 \quad , \quad F_{29,39,.995} = 2.20$$

จะเห็นว่า $s_x^2/s_y^2 > 2.20$ ดังนั้นจึงปฏิเสธสมมุติฐานหลักและเชื่อว่าตัวต้านทานของสองบริษัทมีความแตกต่างกันในความแปรปรวน

หมายเหตุ

1. ตาราง F มักเสนอค่าความน่าจะเป็นไว้ในรูป $F_{m,n,1-\alpha}$ ถ้าต้องการหาค่าของ $F_{n,m,\alpha}$ ให้ใช้บทแทรก 3.6 หน้า 82 ดังนี้คือ

$$F_{n,m,\alpha} = 1/F_{m,n,1-\alpha} = \frac{1}{2.33} = 0.429$$

2. ตาราง F ที่มีอยู่เสนอพื้นที่ใต้โค้งไว้เพียงสี่ระดับคือ 75% 90% 95% และ 99% มิได้เสนอระดับ 99.5% ดังนั้นเมื่อต้องการหาพื้นที่ 99.5% จึงใช้พื้นที่ 99% แทนกันได้โดยอนุโลม ส่วนพื้นที่ 25% 10% 5% และ 1% ใช้บทแทรก 3.6 เป็นเครื่องมือช่วยวิเคราะห์[1] อย่างไรก็ตามมิใช่ว่าตาราง F จะมิได้เสนอพื้นที่ที่ละเอียดกว่าที่เสนอไว้ท้ายเล่ม นักศึกษาสามารถหาตารางจากเอกสารอ้างอิงอื่น ๆ ที่เสนอพื้นที่ละเอียดกว่านี้ได้ แต่จุดอ่อนคือตารางเหล่านี้อาจเสนอ df ไว้เป็นช่วงที่กว้างมากกว่า ซึ่งย่อมชี้ให้เห็นว่าค่า F ที่ได้ยังคงเป็นค่าประมาณอยู่เช่นเดิม

(2) การคำนวณหาขนาดตัวอย่าง ให้อาศัยสมการความเสี่ยงประเภทที่ 2 ดังนี้

$$\beta(\sigma_x^2, \sigma_y^2) \simeq \Pr\{F_{n-1,m-1} < \frac{n-1}{m-1}\ \frac{\sigma_y^2}{\sigma_x^2} \cdot F_{n-1,m-1,1-\alpha/2}\}$$

$$=> \quad F_{n-1,m-1,\beta} \simeq \frac{(n-1)}{(m-1)}\ \frac{\sigma_y^2}{\sigma_x^2}\ F_{n-1,m-1,1-\alpha/2}$$

---

[1] ดูทฤษฎีและบทแทรกหน้า 81-85

ในที่นี้ต้องถือว่า n = m มิเช่นนั้นจะแก้สมการได้ยากเนื่องจากต้องเปรียบเทียบหาอัตราส่วนระหว่างค่า F ในทุกค่าทั่วตาราง

$$F_{n-1,m-1,\beta} \simeq \frac{\sigma_y^2}{\sigma_x^2} F_{n-1,m-1,1-\alpha/2}$$

ดังนั้นค่าขนาดตัวอย่างอุดมะ (n = m) สามารถคำนวณหาได้จากสมการ

$$\frac{F_{n-1,m-1,\beta}}{F_{n-1,m-1,1-\alpha/2}} \simeq \frac{\sigma_y^2}{\sigma_x^2}$$

เมื่อ $\sigma_x^2 = 2\sigma_y^2$

$$\frac{F_{n-1,m-1,\beta}}{F_{n-1,m-1,1-\alpha/2}} \simeq \frac{\sigma_y^2}{2\sigma_y^2} = \frac{1}{2}$$

จากตารางจะพบว่า

$$\frac{F_{200,200,.01}}{F_{200,200,.99}} = \frac{1/1.38}{1.38} \simeq .53$$

ดังนั้น $n - 1 \simeq 200$ หรือ $n \simeq 201$ หรือ $n \leqslant 201$

ข. $H_0\ \mu_x = \mu_y$ vs $H_1 : \mu_x \neq \mu_y$

จากตารางที่ 1 แสดงว่า $\sigma_x^2 = \sigma_y^2$ ดังนั้นการทดสอบสมมุติฐานข้อ ข. นี้จึงใช้ t-test กล่าวคือ ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$|\bar{x} - \bar{y}| > t_{\nu,1-\alpha/2}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}$$

$$\nu = \frac{(s_x^2/n + s_y^2/m)^2}{\{(s_x^2/n)/(n+1)\} + \{(s_y^2/m)/(m+1)\}} - 2$$

$$= \frac{(.005/30 + .0015/40)^2}{\{(.005/30)/31\} + \{(.0015/40)/41\}} - 2$$

$$= 42.8 \simeq 43$$

$$t_{\nu,1-\alpha/2}\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} = t_{43,.995}\sqrt{\frac{.005}{30} + \frac{.0015}{40}} = .039$$

$$|\bar{x} - \bar{y}| = |5.2 - 5.1| = 0.1$$

จะเห็นว่า $|\bar{x} - \bar{y}| > .039$ ดังนั้นจึงปฏิเสธสมมุติฐานหลักและเชื่อว่าตัวต้านทานที่ผลิตโดยบริษัททั้งสองมีคุณภาพต่างกัน

ตัวอย่าง 7.10 สหกรณ์โคนมผู้รับซื้อนมจากฟาร์มโคนมต้องการเปรียบเทียบปริมาณมันเนยในน้ำนมจากฟาร์มสองแห่ง (X และ Y) โดยสุ่มตัวอย่างน้ำนมจากฟาร์ม X มา 20 ควอท พบปริมาณน้ำมันเนยเฉลี่ย .012 ปอนด์ ส่วนเบี่ยงเบนมาตรฐาน .0015 ปอนด์ และสุ่มตัวอย่างน้ำนมจากฟาร์ม Y มา 16 ควอท พบปริมาณเนยเฉลี่ย .014 ปอนด์ ส่วนเบี่ยงเบนมาตรฐาน .0010 ปอนด์ จงทดสอบสมมุติฐาน ณ. ระดับนัยสำคัญ 5% สำหรับ

ก. $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 > \sigma_y^2$

ข. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

วิธีทำ

ก. $H_0 : \sigma_x^2 = \sigma_y^2$ vs $H_1 : \sigma_x^2 > \sigma_y^2$

ปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$s_x^2/s_y^2 \quad > \quad F_{n-1,m-1,1-\alpha}$$

$$s_x^2/s_y^2 \quad = \quad (.0015)^2/(.001)^2 = 1.44, \; F_{19,15,.95} = 2.33$$

จะเห็นว่า $s_x^2/s_y^2 < 2.33$ ดังนั้นจึงไม่อาจปฏิเสธสมมุติฐานหลักและเชื่อว่าความแปรปรวนในปริมาณเนยต่อควอทของฟาร์มทั้งสองมิได้แตกต่างกัน

ขณะเดียวกัน ถ้า $\sigma_x^2 = .07\sigma_y^2$ เราสามารถคำนวณหาความเสี่ยงที่จะปฏิเสธความจริงข้อนี้ ได้ดังนี้

$$\beta(\sigma_x^2, \sigma_y^2) \quad = \quad \Pr\{F_{n-1,m-1} \leqslant \frac{n-1}{m-1} \; \frac{\sigma_y^2}{\sigma_x^2} \cdot F_{n-1,m-1,1-\alpha}\}$$

$$= \quad \Pr\{F_{n-1,m-1} \leqslant \frac{19}{15} \; \frac{1}{0.07} \; (2.33)\}$$

$$= \quad \Pr\{F_{n-1,m-1} \leqslant 42.16\}$$

โดยอาศัย CLT; $E(F) = 20/18 = 1.111$

$$V(F) \quad = \quad \frac{2(15)^2(20+15-2)}{20(13)^2(11)} = .399, \; \sigma_F = .632$$

$$\Pr\{F_{n-1,m-1} \leqslant 42.16\} \quad \simeq \quad \Pr\{Z < \frac{42.16 - 1.11}{.632}\}$$

$$= \Pr\{Z < 64.95\} \quad \simeq \quad 1$$

ข. $H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x < \mu_y$

จากผลในข้อ ก. ยืนยันว่า $\sigma_x^2 = \sigma_y^2 = \sigma^2$ ซึ่งเราสามารถประมาณค่า σ^2 โดย s_p^2 ดังนั้นจึงปฏิเสธสมมุติฐานหลัก ณ. ระดับนัยสำคัญ α เมื่อ

$$\bar{x} - \bar{y} < t_{n+m-2,\alpha} s_p \sqrt{\frac{1}{n} + \frac{1}{m}}$$

$$\bar{x} - \bar{y} = .012 - .014 = -0.002$$

$n = 20, m = 16$

$$s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$$

$$= \frac{19(.0015)^2 + 15(.0010)^2}{34}$$

$$= .0000192$$

$$s_p = .0043817$$

$$t_{n+m-2,\alpha} s_p \sqrt{\frac{1}{n} + \frac{1}{m}} = t_{33,0.05}(.0043817)\sqrt{\frac{1}{20} + \frac{1}{16}}$$

$$= (-1.697)(.0043817)(.335419)$$

$$= -.0025$$

จะเห็นว่า $\bar{x} - \bar{y} > -0.0025$ ดังนั้นจึงไม่อาจปฏิเสธสมมุติฐานหลักและเชื่อว่าฟาร์มโคนมทั้งสองผลิตน้ำนมที่มีปริมาณเนยไม่ต่างกัน

7.3 การทดสอบสมมุติฐานเกี่ยวกับความแปรปรวนของหลายกลุ่มประชากร (Test of Homogeneity of Variances)

$H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$ vs $H_1 : \sigma_i^2; i = 1, 2, \ldots, k$ ไม่เท่ากันทั้งหมด ในกรณีนี้ถือว่า $X_{ij} \sim N(\mu_i, \sigma_i^2)$ เมื่อ X_{ij} คือตัวแปรสุ่ม ที่สุ่มมาจากกลุ่มประชากรที่ i; $j = 1, 2, \ldots, n_i$; $i = 1, 2, \ldots, k$ ทั้งนี้ถือว่า μ_i และ σ_i^2 เป็นพารามิเตอร์ที่ไม่ทราบค่า (Unknown, μ_i, σ_i^2)

$$\omega = \{\sigma^2, \mu_i; i = 1, 2, \ldots, k : \sigma^2 > 0, -\infty < \mu_i < \infty\}$$

$$\Omega = \{\sigma_i^2, \mu_i; i = 1, 2, \ldots, k : \sigma_i^2 > 0, -\infty < \mu_i < \infty\}$$

สุ่มตัวอย่างขนาด $n_1, n_2, \ldots, n_k$ จากกลุ่มประชากรที่ 1, 2, ..., k ตามลำดับ[1]

$$=> \quad L\omega = \prod_i^k \prod_j^{n_i} f_{X_{ij}}(x_{ij}; \mu_i, \sigma_i^2)$$

$$= \prod_i^k \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n_i} \exp\left(-\frac{1}{2\sigma^2}\sum_j^{n_i}(x_{ij}-\mu_i)^2\right)$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n} \exp\left(-\frac{1}{2\sigma^2}\sum_i^k\sum_j^{n_i}(x_{ij}-\mu_i)^2\right)^{2}$$

$$L\omega = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^{n} \exp\left\{-\frac{1}{2\sigma^2}\left\{\left(\sum_j^{n_1}(x_{ij}-\mu_1)^2 + \ldots + \sum_j^{n_k}(x_{kj}-\mu_i)^2\right\}\right\}\right.$$

$$\frac{\partial}{\hat{\partial}\mu_i}\ln L\omega = 0 => 0 - 0 - \ldots - 0 + \frac{1}{\sigma^2}\sum_j^{n_i}(x_{ij}-\mu_i) - 0 - 0 - \ldots = 0; \quad i = 1, 2, \ldots, k$$

$$=> \quad \hat{\mu}_i = \frac{1}{n}\sum_j^{n_i} x_{ij} = \bar{x}_i; \; i = 1, 2, \ldots, k$$

$$\frac{\partial}{\partial\sigma}\ln L\omega = 0 => -\frac{n}{\sigma} - 0 + \frac{1}{\sigma^3}\sum_i^k\sum_j^{n_i}(x_{ij}-\mu_i)^2 = 0$$

$$=> \quad \sigma^2 = \frac{1}{n}\sum_i^k\sum_j^{n_i}(x_{ij}-\mu_i)^2$$

แทนค่า μ_i ด้วย MLE $\hat{\mu}_i = \bar{x}_i$

$$=> \quad \hat{\sigma}^2 = \frac{1}{n}\sum_i^k\sum_j^{n_i}(x_{ij}-\bar{x}_i)^2 \; {}^{3}$$

$$=> \quad L\hat{\omega} = \prod_i^k\prod_j^{n_i} f_{X_{ij}}(x_{ij}, \hat{\mu}, \hat{\sigma}^2)$$

---

[1] กลุ่มตัวอย่างขนาด n_i จากกลุ่มประชากรที่ i คือ $(X_{i1}, X_{i2}, \ldots, X_{in_i})$; $i = 1, 2, \ldots, k$

[2] ให้ $n = n_1 + n_2 + \ldots + n_k$

[3] $\hat{\sigma}^2 = \frac{1}{n}\sum_i^k\sum_j^{n_i}(x_{ij}-\bar{x}_i)^2 = \frac{1}{n}\sum_i^k(n_i-1)s_i^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2 + \ldots + (n_k-1)s_k^2}{n_1 + n_2 + \ldots + n_k}$

แสดงว่า $\hat{\sigma}^2$ คือ pooled variance และเพื่อให้ง่ายเราจะใช้ $\hat{\sigma}^2 = \frac{n_1s_1^2 + n_2s_2^2 + \ldots + n_ks_k^2}{n_1 + n_2 + \ldots + n_k}$

ตามวิธี Stratified Random Sampling