הפרמטר ρ בדגימת משפחות

בביטוי הקודם

$$\rho = \frac{m(S_b^2/S_Y^2) - 1}{m-1}$$

כאשר

$$S_b^2 = \frac{1}{C-1} \sum_{i=1}^{C} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

$$S_Y^2 = \frac{1}{N-1} \sum_{i=1}^{C} \sum_{j=1}^{m} (Y_{ij} - \bar{Y}_{\cdot\cdot})^2$$

כאן נפתח כמה תכונות של ρ.

① נוכל לכתוב

$$(C-1) S_b^2 = \sum_{i=1}^{C} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$$

$$= \sum_{i=1}^{C} \left[\frac{1}{m} \sum_{j=1}^{m} (Y_{ij} - \bar{Y}_{\cdot\cdot}) \right]^2$$

$$= \left(\frac{1}{m}\right)^2 \sum_{i=1}^{C} \left[\sum_{j=1}^{m} (Y_{ij} - \bar{Y}_{\cdot\cdot}) \right] \left[\sum_{k=1}^{m} (Y_{ik} - \bar{Y}_{\cdot\cdot}) \right]$$

$$= \left(\frac{1}{m}\right)^2 \sum_{i=1}^{c} \sum_{j=1}^{m} \sum_{k=1}^{m} (Y_{ij} - \bar{Y}_{..})(Y_{ik} - \bar{Y}_{..})$$

$$= \left(\frac{1}{m}\right)^2 \sum_{i=1}^{c} \sum_{j=1}^{m} (Y_{ij} - \bar{Y}_{..})^2$$

$$+ \left(\frac{1}{m}\right)^2 \sum_{i=1}^{c} \sum_{\substack{j=1 \\ j \neq k}}^{m} \sum_{k=1}^{m} (Y_{ij} - \bar{Y}_{..})(Y_{ik} - \bar{Y}_{..})$$

$$= \left(\frac{1}{m^2}\right)(N-1) S_Y^2 + \frac{1}{m^2} Q$$

כאשר

$$Q = \sum_{i=1}^{c} \sum_{\substack{j=1 \\ j \neq k}}^{m} \sum_{k=1}^{m} (Y_{ij} - \bar{Y}_{..})(Y_{ik} - \bar{Y}_{..})^2$$

לכן

$$S_b^2 = \frac{N-1}{m^2(c-1)} S_Y^2 + \frac{1}{m^2(c-1)} Q$$

$$\rho = \left[m\left(S_b^2 / S_Y^2\right) - 1\right] / (m-1)$$

$$= \frac{1}{m-1}\left[\frac{N-1}{m(c-1)} - 1 + \frac{1}{m(c-1) S_Y^2} Q\right]$$

$$= \frac{1}{m-1}\left[\frac{N-1}{N-m} - 1 + \frac{1}{(N-m) S_Y^2} Q\right]$$

אם c גדול, אז:

$$\frac{N-1}{N-m} = \frac{C_{m-1}}{C_{m-m}} \doteq 1$$

ואז נוכל לכתוב

$$\hat{\rho} \doteq [(m-1)(N-1)S_Y^2]^{-1} Q$$

לאור זה, נגדיר

$$\tilde{\rho} = [(m-1)(N-1)S_Y^2]^{-1} Q$$

② טענה: נתייחס לתהליך הבא:

א. נבחר אשכול אחד באופן מקרי. נסמן את האינדקס של האשכול שבחרנו בסימן θ.

ב. מתוך האשכול שבחרנו בשלב א' לעיל, נבחר שני פרטים זה אחר זה באופן מקרי ללא החזרה, נסמן:

j_1 = האינדקס של הפרט שנבחר בבחירה הראשונה

j_2 = האינדקס של הפרט שנבחר בבחירה השנייה

נניח כי θ ו- (j_1, j_2) בלתי-תלויים.

נגדיר

$$U = Y_{\theta j_1}, \quad V = Y_{\theta j_2}$$

אזי

$$Corr(U, V) = \tilde{\rho}$$

הערה: התוצאה הזאת מחזקת את האינטרפרטציה של הפרמטר ρ (ששווה בגודלו ל- $\tilde{\rho}$) [illegible]

<u>הוכחה</u>: נזכיר כי

$$Corr(U,V) = Cov(U,V) / [Var(U) Var(V)]^{1/2}$$

כאשר $Cov(U,V) = E[(U-E[U])(V-E[V])]$

נחשב את:

$$E[U],\ E[V]$$
$$Var(U),\ Var(V)$$
$$Cov(U,V)$$

לפי התהליך המתואר, יש לנו

$$Pr(d = i) = \frac{1}{c}, \quad i = 1, \dots, c$$

$$Pr(g_1 = j, g_2 = k) = \frac{1}{m(m-1)}$$

$$j, k = 1, \dots, m \quad (j \neq k)$$

$$Pr(d = i, g_1 = j, g_2 = k) = \frac{1}{c} \frac{1}{m(m-1)}$$

$$Pr(g_1 = j) = \frac{1}{m} \qquad j = 1, \dots, m$$

$$Pr(g_2 = k) = \frac{1}{m} \qquad k = 1, \dots, m$$

$$Pr(d = i, g_1 = j) = \frac{1}{cm} = \frac{1}{N}$$

$$Pr(d = i, g_2 = k) = \frac{1}{cm} = \frac{1}{N}$$

בס״ד

כעת:

$$E[U] = E[Y_{\ell g_1}]$$
$$= \sum_{i=1}^{e} \sum_{j=1}^{m} Pr(\ell = i, g_1 = j)\, Y_{ij}$$
$$= \frac{1}{N} \sum_{i=1}^{e} \sum_{j=1}^{m} Y_{ij} = \bar{Y}_{..}$$

$$Var(U) = E[(U - E[U])^2]$$
$$= E[(Y_{\ell g_1} - \bar{Y}_{..})^2]$$
$$= \sum_{i=1}^{e} \sum_{j=1}^{m} Pr(\ell = i, g_1 = j)\,(Y_{ij} - \bar{Y}_{..})^2$$
$$= \frac{1}{N} \sum_{i=1}^{e} \sum_{j=1}^{m} (Y_{ij} - \bar{Y}_{..})^2 = \frac{N-1}{N} S_Y^2$$

מטעמי סימטריות:

$$E[V] = E[U] = \bar{Y}_{..}$$
$$Var(V) = Var(U) = \frac{N-1}{N} S_Y^2$$

כלומר:

$$Cov(U,V) = E[(U-E[U])(V-E[V])]$$

$$= E[(Y_{\vartheta g_1} - \bar{Y}_{..})(Y_{\vartheta g_2} - \bar{Y}_{..})]$$

$$= \sum_{i=1}^{c} \sum_{\substack{j=1 \\ j\neq k}}^{m} \sum_{k=1}^{m} Pr(\vartheta = i, g_1 = j, g_2 = k)\,(Y_{ij} - \bar{Y}_{..})(Y_{ik} - \bar{Y}_{..})$$

$$= \frac{1}{cm(m-1)} \sum_{i=1}^{c} \sum_{\substack{j=1 \\ j\neq k}}^{m} \sum_{k=1}^{m} (Y_{ij} - \bar{Y}_{..})(Y_{ik} - \bar{Y}_{..})$$

$$= \frac{1}{N(m-1)} Q$$

כאשר Q כמוגדר בסעיף ②.

לכן:

$$Corr(U,V)$$

$$= \left[\frac{1}{N}\,\frac{1}{m-1} Q\right] \Big/ \left[\left(\frac{N-1}{N} S_Y^2\right)\left(\frac{N-1}{N} S_Y^2\right)\right]^{1/2}$$

$$= \frac{1}{(N-1)(m-1) S_Y^2} Q = \tilde{\rho}\,.$$

☺ .ס.ה.נ