3.6. NILAI SAMPEL DARI KOMBINASI LINEAR DARI VARIABEL

Kita sudah memperkenalkan Kombinasi linear p variabel di pasal 2.6. Pada kebanyakan prosedur multivariat, kita pasti dengan sendirinya menganggap kombinasi linear dari bentuk

$$c'X = c_1X_1 + c_2X_2 + \cdots + c_pX_p$$

yang memperhatikan nilai dipercobaan ke-j

$$c'x_j = c_1x_{1j} + c_2x_{2j} + \cdots + c_px_{pj} \ , \qquad j = 1,2,\ldots,n \qquad (3-31)$$

n berasal dari percobaan di 3-31 maka kita dapatkan

$$rata-rata\ sampel = \frac{(c'x_1 + c'x_2 + \cdots + c'x_n)}{n}$$

$$= c'(x_1 + x_2 + \cdots + x_n)\frac{1}{n} = c'\overline{x} \qquad (3-32)$$

Karena $\left(c'x_j - c'\overline{x}\right)^2 = \left(c'\left(x_j - \overline{x}\right)^2\right) = c'\left(x_j - \overline{x}\right)\left(x_j - \overline{x}\right)' c$, kita dapatkan

$$variansi\ sampel = \frac{(c'x_1 - c'\overline{x})^2 + (c'x_2 - c'\overline{x})^2 + \cdots + (c'x_n - c'\overline{x})^2}{n-1}$$

$$= \frac{c'(x_1 - \overline{x})(x_1 - \overline{x})'c + c'(x_2 - \overline{x})(x_2 - \overline{x})'c + \cdots + c'(x_n - \overline{x})(x_n - \overline{x})'c}{n-1}$$

$$= c'\left[\frac{(x_1 - \overline{x})(x_1 - \overline{x})' + (x_2 - \overline{x})(x_2 - \overline{x})' + \cdots + (x_n - \overline{x})(x_n - \overline{x})'}{n-1}\right]c$$

atau

$$\text{variansi sampel dari } c'X = c'Sc \qquad (3-33)$$

Persamaan (3-32) dan (3-33) adalah sampel persamaan dari (2-43). Mereka berhubungan untuk mengganti jumlah sampel $\overline{x}$ dan S untuk jumlah populasi μ dan $\sum$, secara langsung, di (2-43).

Andaikan Kombinasi linear kedua

$$b'X = b_1X_1 + b_2X_2 + \cdots + b_pX_p$$

yang memperlihatkan nilai di percobaan ke- j

$$b'x_j = b_1x_{1j} + b_2x_{2j} + \cdots + b_px_{pj} \ , \qquad j = 1,2,\ldots,n \qquad (3-34)$$

Mengikuti dari (3-32) dan (3-33) bahwa rata-rata sampel dan variansi dari derived percobaan itu

$$rata - rata\ sampel\ dari\ b'X = b'\bar{x}$$

$$variansi\ sampel\ dari\ b'X = b'Sb$$

Terlebih lagi, kovarian sampel dihitung dari pasangan percobaan pada $b'X$ dan $c'X$ adalah

$$kovarian\ sampel$$
$$= \frac{(b'x_1 - b'\bar{x})(c'x_1 - c'\bar{x}) + (b'x_2 - b'\bar{x})(c'x_2 - c'\bar{x}) + \cdots + (b'x_n - b'\bar{x})(c'x_n - c'\bar{x})}{n-1}$$

$$= \frac{b'(x_1 - \bar{x})(x_1 - \bar{x})'c + b'(x_2 - \bar{x})(x_2 - \bar{x})'c + \cdots + b'(x_n - \bar{x})(x_n - \bar{x})'c}{n-1}$$

$$= b'\left[\frac{(x_1 - \bar{x})(x_1 - \bar{x})' + (x_2 - \bar{x})(x_2 - \bar{x})' + \cdots + (x_n - \bar{x})(x_n - \bar{x})'}{n-1}\right]c$$

atau

$$\text{Kovarian sampel dari } b'X \text{ dan } c'X = b'Sc \qquad (3-35)$$

Secara ringkasnya, kita dapatkan hasil berikut

Hasil 3.5. Kombinasi linear

$$b'X = b_1X_1 + b_2X_2 + \cdots + b_pX_p$$

$$c'X = c_1X_1 + c_2X_2 + \cdots + c_pX_p$$

Kita dapatkan rata-rata sampel, varinsai, dan kovarian yang berhubungan ke $\bar{x}$ dan S dari

$$rata - rata\ sampel\ dari\ b'X = b'\bar{x}$$
$$rata - rata\ sampel\ dari\ c'X = c'\bar{x}$$

$$variansi\ sampel\ dari\ b'X = b'Sb$$

$$variansi\ sampel\ dari\ c'X = c'Sc$$

$$kovarian\ sampel\ dari\ b'X\ \ dan\ c'X = b'Sc$$

Contoh :

Kita akan menggap dua Kombinasi linear dan nilai derived untuk n = 3 percobaan diberikan di contoh 3.8 sebagai

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 4 & 4 \\ 2 & 1 & 0 \\ 5 & 6 & 4 \end{bmatrix}$$

Rata-rata, variansi, dan kovarian akan pertama kali di selesaikan secara langsung dan kemudian diselesaikan dengan (3-36).

Anggap dua kombninasi linear

$$b'X = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = 2X_1 + 2X_2 - X_3$$

dan

$$c'X = \begin{bmatrix} 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = X_1 - X_2 + 3X_3$$

Percobaan pada kombinasi linear diperoleh dengan mengganti X_1 , X_2 , dan X_3 dengan nilai percobaan. Untuk contoh, n=3 percobaan pada $b'X$ adalah

$$b'x_1 = 2x_{11} + 2x_{21} - x_{31} = 2(1) + 2(2) - (5) = 1$$

$$b'x_2 = 2x_{12} + 2x_{22} - x_{32} = 2(4) + 2(1) - (6) = 4$$

$$b'x_3 = 2x_{13} + 2x_{23} - x_{33} = 2(4) + 2(0) - (4) = 4$$

Rata-rata sampel dan variansi dari nilai tersebut adalah

$$rata - rata\ sampel = \frac{(1 + 4 + 4)}{3} = 3$$

$$variansi\ sampel = \frac{(1-3)^2 + (4-3)^2 + (4-3)^2}{3-1} = 3$$

Dengan cara yang sama , n = 3 percobaan pada $c'X$ adalah

$$c'x_1 = x_{11} - x_{21} + 3x_{31} = 1(1) - 1(2) + 3(5) = 14$$

$$c'x_2 = x_{12} - x_{22} + 3x_{32} = 1(4) - 1(1) + 3(6) = 21$$

$$c'x_3 = x_{13} - x_{23} + 3x_{33} = 1(4) - 1(0) + 3(4) = 16$$

dan

$$rata - rata\ sampel = \frac{(14 + 21 + 16)}{3} = 17$$

$$variansi\ sampel = \frac{(14-17)^2 + (21-17)^2 + (16-17)^2}{3-1} = 13$$

Terlebih lagi, kovarian sampel dihitung dari pasangan percobaan $(b'x_1, c'x_1)$, $(b'x_2, c'x_2)$, dan $(b'x_3, c'x_3)$ adalah

$$kovarian\ sampel = \frac{(1-3)(14-17) + (4-3)(21-17) + (4-3)(16-17)}{3-1} = \frac{9}{2}$$

Dengan cara lain, kita gunakan rata-rata sampel vector $\overline{x}$ dan kovarian sampel matrik S derived dari data asli data matrik X untuk menghitung rata-rata sampel, variansi, dan kovarian dari kombinasi linear. Jadi, jika hanya statistik deskriptif yang diperhatikan , kita tidak memperlukan perhitungan dengan percobaan $b'x_j$ dan $c'x_j$.

Dari contoh 3.8,

$$\overline{x} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}\ dan\ S = \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Akibatnya , menggunakan (3-36), dua rata-rat sampel dari derived percobaan adalah

$$rata - rata\ sampel\ b'X = b'\overline{x} = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} = 3$$

$$rata - rata\ sampel\ c'X = c'\overline{x} = \begin{bmatrix} 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} = 17$$

Dari (3-36) kita dapatkan

$$variansi\ sampel\ dari\ b'X = b'Sb$$

$$= \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ -\frac{3}{2} \\ 0 \end{bmatrix} = 3$$

$variansi\ sampel\ dari\ c'X = c'Sc$

$$= \begin{bmatrix} 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} \frac{9}{2} \\ -1 \\ \frac{5}{2} \end{bmatrix} = 13$$

Kovarian sampel dari $b'X\ dan\ c'X = b'Sc$

$$= \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 3 & -\frac{3}{2} & 0 \\ -\frac{3}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} \frac{9}{2} \\ -1 \\ \frac{5}{2} \end{bmatrix} = \frac{9}{2}$$

Untuk menunjukan hasil terakhir diperiksa dengan jumlah sampel yang berhubungan dihitung secara langsung dari percobaan pada kombinasi linear.

Rata-rata sampel dan hubungan kovarian pada hasil 3.5 bersinggungan untuk sembarang bilangan kombinasi linear. Anggap kombinasi linear q

$$a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p\ , i = 1,2, \ldots, q \qquad (3-37)$$

Dapat ditulis pada notasi matriks sebagai berikut

$$\begin{bmatrix} a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\ \vdots \\ a_{q1}X_1 + a_{q2}X_2 + \cdots + a_{qp}X_p \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{q1} & \cdots & a_{qp} \end{bmatrix} \begin{bmatrix} X_1 \\ \\ X_q \end{bmatrix} = \mathrm{AX} \qquad (3-38)$$

Menerima baris ke - i dari A, $a_i{}'$ sampai b' dan baris ke k dari A, $a_k{}'$, sampai c', persamaan (3-36) menyebabkan baris ke- i dari AX memeliki rata-rata $a_i{}'\overline{x}$ dan baris ke-i dan ke - k dari AX memiliki kovaris sampel $a_i{}'Sa_k$. catatan bahwa $a_i{}'Sa_k$ adalah unsur (i,k) dari ASA’.

Hasil. 3.6. kombinasi linear q dari AX di (3-38) memiliki rata-rata sampel vektor $A\bar{x}$ dan kovarian sampel matrik ASA’.

3.7. MENYELESAIKAN SEBUAH SAMPEL PERCOBAAN SEBAGAI SEBUAH POPULASI

Pada kasus tertentu, dimungkinkan untuk menyelesaikan sampel percobaan sebagai populasi dari nilai. Berguna untuk dua alsan : (1) serves sebagai sebuah cara yang ampuh untuk menyimpulkan sifat umum sampel dari sifat populasi yang berhubungan dan (2) memberikan hubungan untuk menghitung rata-rata populasi dan kovarian dari data yang dikumpulkan pada sensus yang komplit.pada situasi dimana data matrik berisi semua informasi tersedia tentang subjek untuk contoh seperti situasi yang terjadi saat perusahaan mobil ingin pada penjualan automobile tahunaan pada sebuah negara, mengumpulkan semua tokoh dari semua deller.

Misalkan n (percobaan) kolom dari p x n data matrik $X = [x_1, x_2, \dots, x_n]$ mewakili semua sensus dari pengukuran. Kita mendapatkan p variabel menarik. Misalkan kita menganggap kolom vektor sebagai “ informasi yang serupa “ dan menetapkan masing-masing bobot 1/n. setara dengan distribusi probabilitas pada cara seragam sehingga vektor percobaan x_j adalah ditetapkan probabilitas 1/n untuk masing-masing j = 1, 2, …, n.

Misalkan X melambangkan variabel vektor peubah acak maka kita asumsikan x_j dengan probabilitas 1/n. untuk distribusi ini , nilai yang diharapkan adalah

$$rata - rata = E(X) = x_1\left(\frac{1}{n}\right) + x_2\left(\frac{1}{n}\right) + \cdots + x_n\left(\frac{1}{n}\right) = \bar{x} \qquad (3-39)$$

Similarly, matrik kovarian dari X adalah

$$\text{Kovarian} = E(X - E(X))(X - E(X))' = E(X - \bar{x})(X - \bar{x})'$$

$$= (x_1 - \bar{x})(x_1 - \bar{x})'\left(\frac{1}{n}\right) + (x_2 - \bar{x})(x_2 - \bar{x})'\left(\frac{1}{n}\right) + \cdots + (x_n - \bar{x})(x_n - \bar{x})'\left(\frac{1}{n}\right) = S_n \quad (3-40)$$

Persamaan (3-39) dan (3-40) membentuk hubungan formal antara perhitungan dan S_n (lihat (1-4) dan (1-8)) sebagai ststistik deskritif dan menyelesaikannya sebagai momen populasi. Hasil 3.7 merangkum ide-ide tersebut.

Hasil 3.7, jika nilai percobaan $x_1, x_2, \dots, x_n$ diselesaikan sebagai himpunan komplit nilai-nilai untuk peubah acak dan masing-masing diselesaikan seperti ditetapkannya probabilitas 1/n, hasil distribusi didapatkan

$$rata - rata\ populasi\ vektor = \bar{x}$$

$$kovarian\ matrik\ populasi = \boldsymbol{S}_n$$

Kita simpulkan dengan mengilustrasikan bagaimana hubungan sampel melibatkan $\boldsymbol{S}_n$ disimpulkan dari hasil populasi yang diketahui dari mengkhususkan situasi yang berlaku dimana $\bar{x}$ 1 pada rata-rata populasi μ dan $\boldsymbol{S}_n$ kovarian populasi matrik $\boldsymbol{\Sigma}$.

Himpunan

$$D_n^{1/2} = \begin{bmatrix} \sqrt{s_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{s_{pp}} \end{bmatrix}$$

Dimana $s_{ii} = \frac{1}{n}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_i\right)^2$ adalah unsur diagonal dari $\boldsymbol{S}_n$. kemudian $D_n^{1/2}$ menggantikan $V^{1/2}$ di (2-37) dan R menggantikan ρ, sehingga

$$D_n^{-1/2} S_n D_n^{-1/2} = R \qquad dan \quad S_n = \ D_n^{1/2} R D_n^{1/2}$$