MODUL 2

Aspek Analisis Multivariat dan Vektor Random

Prof. Dr. Sri Haryatmi Kartiko, M.Sc.

PENDAHULUAN

Analisis multivariat adalah cabang dari statistika yang mempelajari variabel random yang saling berkolerasi. Jika dua variabel random berkolerasi, pengetahuan tentang kelakuan satu variabel memberi gambaran tentang kelakuan variabel lainnya.

Untuk melakukan probabilitas terjadi suatu kejadian, dalam praktik sering kali perlu diamati beberapa variabel random secara srimultan. Sebagai contoh akan dievaluasi kesuksesan suatu program sosial. Karena sukses atau tidaknya suatu program tergantung pada beberapa pengukuran yang sering berkolerasi, masalah ini harus diselesaikan dengan analisis multivariat.

Dalam modul ini diperkenalkan contoh penggunaan analisis multivariat dalam beberapa bidang kedokteran, sosiologi, pendidikan, dan sebagainya. Melalui contoh-contoh ini dapat terlihat bahwa analisis *univariat* tidak lagi cocok untuk digunakan.

Penyajian data multivariat beserta statistik deskripsinya dibicarakan dalam aspek analisis multivariat yang disajikan dalam Kegiatan Belajar 1. Kegiatan Belajar 2 mempelajari vektor random termasuk harga harapan dan variansinya yang juga telah dipelajari dalam Pengantar Statistik Matematik I. Kovariansi vektor random kovariansi kombinasi linear vektor random juga dibicarakan dalam kegiatan belajar ini.

Secara umum, setelah mempelajari modul ini Anda diharapkan dapat menjelaskan aspek analisis multivariat dan vektor random.

Secara khusus, setelah selesai mempelajari modul ini Anda diharapkan dapat:

1. menyajikan data multivariat;
2. menghitung ukuran pusat, ukuran sebaran, dan ukuran keeratan hubungan sekalihus menyajikannya dalam bentuk matriks;

3. menghitung jarak statistik dua titik tertentu;
4. menghitung harga harapan dan kovariansi vektor random; dan
5. menghitung harga harapan dan kovariansi kombinasi linear vektor random.

KEGIATAN BELAJAR 1

Aspek Analisis Multivariat

Teknik statistik adalah bagian dari pemeriksaan (pengujian) kebenaran ilmiah. Oleh karena itu, penggunaanya sangat luas. Untuk data lebih dari satu variabel yang saling berkolerasi yang banyak dijumpai dalam bidang fisika,sosial,medis,teknik statistika univariat tidak lagi sesuai digunakan. Seperti telah Andabaca dalam pengantar, diperlukan teknik statiska multivariat yang memperhitungkan korelasi antar variabel yang ada.

Dalam Kegiatan Belajar 1 diberikan contoh-contoh dari situasi-situasi di mana teknik multivariat harus digunakan. Hal ini bertujuan untuk membantu Anda mengidenfikasi masalah yang memerlukan teknik multivariat.

A. APLIKASI TEKNIK MULTIVARIAT

Berikut ini akan diberikan contoh-contoh dari situasi-situasi di mana teknik multivariat digunakan.

1. Dalam Bidang Kedokteran

Pada penderita penyakit kanker akan diselidiki reaksi yang ditimbulkan oleh radioterapi. Pengamatan diambil pada 6 variabel reaksi untuk 98 pasien.

Karena sukar untuk membuat/memberikan interpretasi pada 6 variabel reaksi secara bersama-sama, maka diperlukan 1 variabel baru yang merupakan ringkasan dari ke-6 variabel tersebut (teknik ini sering disebut pengurangan data atau data reduction). Dalam hal ini, diperlukan analisis multivariat untuk mengkontruksikan variabel ringkasan tersebut.

2. Dalam Bidang Sosiologi

Sepintas terlihat petani menggarap sawah dapat hidup “layak”.Adalah agak mengherankan bahwa petani tanpa (memiliki) tanah ini dapat hidup”cukup”. Untuk itu, diselidiki beberapa variabel yang mempengaruhi kesuksesan dalam hidupnya, misalnya tingkat pendidikan, jumlah jam kerja dan masih banyak lagi variabel yang mungkin mempengaruhi kehidupannya. Biasanya variabel-variabel tersebut saling berkorelasi satu dengan lainnya. Tentu saja untuk memahami hal ini diperlukan teknik multivariat.

3. Dalam Bidang Pendidikan

Skor SAT (*Scolitic Aptitude Test*) dan nilai sewaktu SLTA sering digunakan untuk memprediksi kesuksesan anak diperguruan tinggi. Diamati 5 variabel prediktor yang merupakan nilai sebelum masuk perguruan tinggi dan 4 variabel kriteria yang merupan nilai setelah berada diperguruan tinggi.

Akan ditentukan hubungan antara variabel prediktor dan variabel kriteria.Dari hubungan ini akan diprediksi nilai diperguruan tinggi berdasarkan nilai SLTA. Hal ini dapat diperluas untuk menentukan apakah mahasiswa akan berhasil atau tidak di perguruan tinggi berdasarkan nilai SLTA.

4. Dalam Bidang Biologi

Dua spesies tanaman sangat sukar untuk diindetifikasi. Untuk menyelidikinya, dilakukan pengukuran pada 4 variabel untuk kedua spesies tanaman tersebut. Keempat variabel ini selanjutnya digunakan untuk mengkonstruksikan fungsi yang harganya dapat memisahkan kedua spesies. Selanjutnya dengan fungsi tersebut suatu tanaman dapat dikategorikan dalam salah satu spesies.

5. Dalam Bidang Lingkungan Hidup

Dipelajari konsentrasi atmosfir dari polusi udara di suatu daerah untuk tiap hari selama kurun waktu tertentu. Diamati 7 variabel yang berhubungan dengan polusi. Ingin diketahui apakah tingkat polusi udara konstan dalam 1 minggu atau ada perbedaan antara hari kerja dan hari libur. Tujuan penyelidikan yang lain adalah mencari 1 variabel yang merupakan ringkasan ke 7 variabel tersebut. Variabel baru ini hendaknya mudah diinterpretasikan.

Ilustrasi di atas memberikan gambaran penggunaan metode multivariat dalam berbagai bidang. Meskipun dalam bentuk yang agak berbeda, masalah analisis data statistik yang telah dibicarakan adalah mirip analisis multivariat. Seperti teknik statistik yang lain, statistik multivariat dibatasi untuk bidang tertentu.

B. ORGANISASI DATA

Dalam modul ini akan dibicarakan analisis hasil pengukuran yang dibuat pada beberapa variabel atau karakteristik. Hasil pengukuran yang biasa disebut data, harus dirangkai dan disajikan dalam berbagai cara.

Sebagai contoh penyajian dalam bentuk tabel dan grafik merupakan bantuan yang penting dalam analisis data. Ringkasan numerik yang merupakan ringkasan berupa angka dari data juga merupakan deskripsi data yang sering diperlukan.

Berikut ini akan diperkenalkan langkah awal dari organisasi data.

1. Array

Data multivariat timbul bila seseorang mempelajari fenomena sosial atau, secara statistik, memilih sejumlah p variabel atau karakter dengan p ≥ 1 sebagai pengamatan.

Harga variabel ini dicatat untuk setiap item,individu,atau trial yang berbeda. Misalkan x_{ij} menunjukan harga tertentu pada variabel ke *i* item (trial) ke *j* (pengamatan ke *j*).

$$X_{ij} = \text{item ke-}j \text{ untuk variabel ke-}i$$

n pengukuran pada *p* variabel tersebut dapat ditulis sebagai berikut:

	item 1	…	item j	…	item n
Variabel	x_{11}	…	x_{1j}	…	x_{1n}
Variabel 2	x_{21}	…	x_{2j}	…	x_{2n}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
Variabel i	x_{i1}	…	x_{ij}	…	x_{in}
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
Variabel p	x_{p1}	…	x_{pj}	…	x_{pn}

Data ini biasa ditulis sebagai matrix $\underset{\sim}{X}$, dengan p baris dan n kolom,

$$\underset{p\times n}{\underset{\sim}{X}} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{bmatrix}$$

Contoh 2.1

Dari 7 orang pembeli pada sebuah toko buku dicatat total pembelian masing-masing pembeli (dalam puluhan ribu rupiah) beserta jumlah buku yang dibeli.

Misal variabel 1 adalah total pembelian dan variabel 2 adalah jumlah buku yang dibeli. Jadi ada tujuh pengukuran pada 2 variabel. Dalam bentuk tabel data disajikan sebagai

Variabel 1 (total pembelian)	42	26	52	60	48	50	58
Variabel 2 (jumlah buku)	4	2	5	6	4	5	3

Menggunakan notasi yang diterangkan di atas

$x_{11} = 42 \;\; x_{12} = 26 \;\; x_{13} = 52 \;\; x_{14} = 60 \;\; x_{15} = 48 \;\; x_{16} = 50 \;\; x_{17} = 58$

$x_{21} = 4 \;\; x_{22} = 2 \;\; x_{23} = 5 \;\; x_{24} = 6 \;\; x_{25} = 4 \;\; x_{26} = 5 \;\; x_{27} = 3$

data ini dapat ditulis dalam bentuk matriks

$$\underset{2x4}{\underset{\sim}{X}} = \begin{pmatrix} 42 & 26 & 52 & 60 & 48 & 50 & 58 \\ 4 & 2 & 5 & 6 & 4 & 5 & 3 \end{pmatrix}$$

2. Statistik Deskriptif

Seperti telah Anda ketahui untuk suatu data diperlukanringkasan angka yang disebut statistik deskriptif. Misal rata-rata adalah merupakan ukuran pusat sedang variansi merupakan ukuran sebaran. Misal $x_{11}, x_{12}, \dots x_{1n}$ adalah n pengukuran pada variabel 1.

Rata-rata pengukuran yang juga disebut *rata-rata* (mean) *sampel* ditulis dengan $\overline{x}_1$ adalah $\overline{x}_1 = \dfrac{1}{n}\sum_{j=1}^{n} x_{1j}$

Secara umum adalah mean sampel untuk variabel ke-i bila ada p variabel dan n pengukuran adalah

$$\overline{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{1j} \quad i = 1,2,\ldots\ldots, p \tag{2.1}$$

variansi sampel untuk variabel ke-i adalah

$$S_{ii} = S_i^2 = \frac{1}{n}\sum_{j-1}^{n}\left(x_{ij} - \overline{x}_i\right)^2,\ i = 1,2,\ldots\ldots, p$$

(bila digunakan pembagi n-1 sebagai pengganti n, variansi sampel merupakan penduga tak bias untuk variansi populasi).

Akar variansi sampel, $\sqrt{S_{ii}}$ adalah standart deviasi sampel (mempunyai satuan sama dengan observasi).

Kovariansi sampel untuk variabel ke-i dan k adalah

$$S_{ik} = \frac{1}{n}\sum_{j=1}^{n}\left(x_{ij} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_k\right) \quad ;\ i = 1,2,\ldots\ldots, p$$
$$k = 1,2,\ldots\ldots, p \tag{2.2}$$

Kovariansi variabel ke-i dan i adalah variansi variabel ke-i dengan $S_{ik} = S_{ki}$ untuk setiap i dan k.

Koefisien kolerasi sampel merupakan ukuran hubungan linier antara 2 variabel (tidak tergantung satuan observasi).

Koefisien kolerasi sampel untuk variabel ke- i dan k adalah

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}}\sqrt{S_{kk}}} = \frac{\sum_{j=1}^{n}\left(x_{ij} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_k\right)}{\sqrt{\sum_{j=1}^{n}\left(x_{ij} - \overline{x}_i\right)^2}\sqrt{\sum_{j=1}^{n}\left(x_{kj} - \overline{x}_k\right)^2}} \tag{2.3}$$

$i = 1,2,\ldots, p.$

$k = 1,2,\ldots, p.$

$r_{ik} = r_{ik}$ untuk setiap i dan k.

r_{ik} mempunyai harga sama bila digunakan pembagi$n$ atau$n$-1 pada S_{ii} , S_{kk} dan S_{ik} .

Koefisien kolerasi antara $\frac{x_{ij}-\overline{x}_k}{\sqrt{S_{ii}}}$ dan $\frac{x_{kj}-\overline{x}_k}{\sqrt{S_{kk}}}$ merupakan kovariansi antara x_i dan x_k .

Meskipun tanda dari r_{ik} dan S_{ik} sama, tetapi lebih mudah memberikan interpretasi pada koefisiensi kolerasi r_{ik} karena harganya terbatas. Koefisiensi kolerasi r mempunyai sifat:

a. $-1 < r < 1$.
b. r menunjukkan ukuran hubungan linier.
   1) $r = 0$ berarti tidak ada hubungan linear antara kedua komponen.
   2) $r < 0$ berarti kecenderungan 1 komponen besar bila komponen lain kecil.
   3) $r > 0$ berarti kecenderungan 1 komponen besar bila komponen lain besar
c. r_{ik} tidak berubah bila variabel ke-i di ubah menjadi $y_{ij} = ax_{ij} + b \;\; j = 1,2,...n$ dan variabel ke-k diubah menjadi $y_{kj} = cX_{kj} + d$, $j = 1,2,...n$ dengan syarat a dan c sama tanda.

S_{ik} dan r_{ik} tidak secara umum menunjukan hubungan antara 2 variabel. Hubungan nonlinear dapat terjadi dengan tidak diperlihatkan oleh kedua ukuran tersebut. Kovariansi dan kolerasi menunjukkan ukuran hubungan linear atau hubungan sepasang garis. Jadi, harganya tidak informatif untuk hubungan lain.

Disamping itu kuantitas ini sangat sensitif terhadap observasi pencilan (*outlier*),sehingga apabila hal ini terjadi kolerasi atau kovariansi mungkin menunjukkan hubungan yang sebetulnya tidak ada.

Bila observasi pencilan dilihat dahulu apakah ada kesalahan pada observasi pencilan ini, bila ada sebaiknya dikoreksi terlebih dahulu. Bila tidak ada kesalahan pada observasi pencilan koefisien kolerasi hendaknya dihitung dan tanpa observasi pencilan.

Statistik deskriptif

Mean sampel $\underset{\sim}{\bar{x}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix}$ (2.4)

Variansi dan kovariansi sampel $\underset{\sim}{S}_n = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix}$ (2.5)

Kolerasi sampel $\underset{\sim}{R} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}$ (2.6)

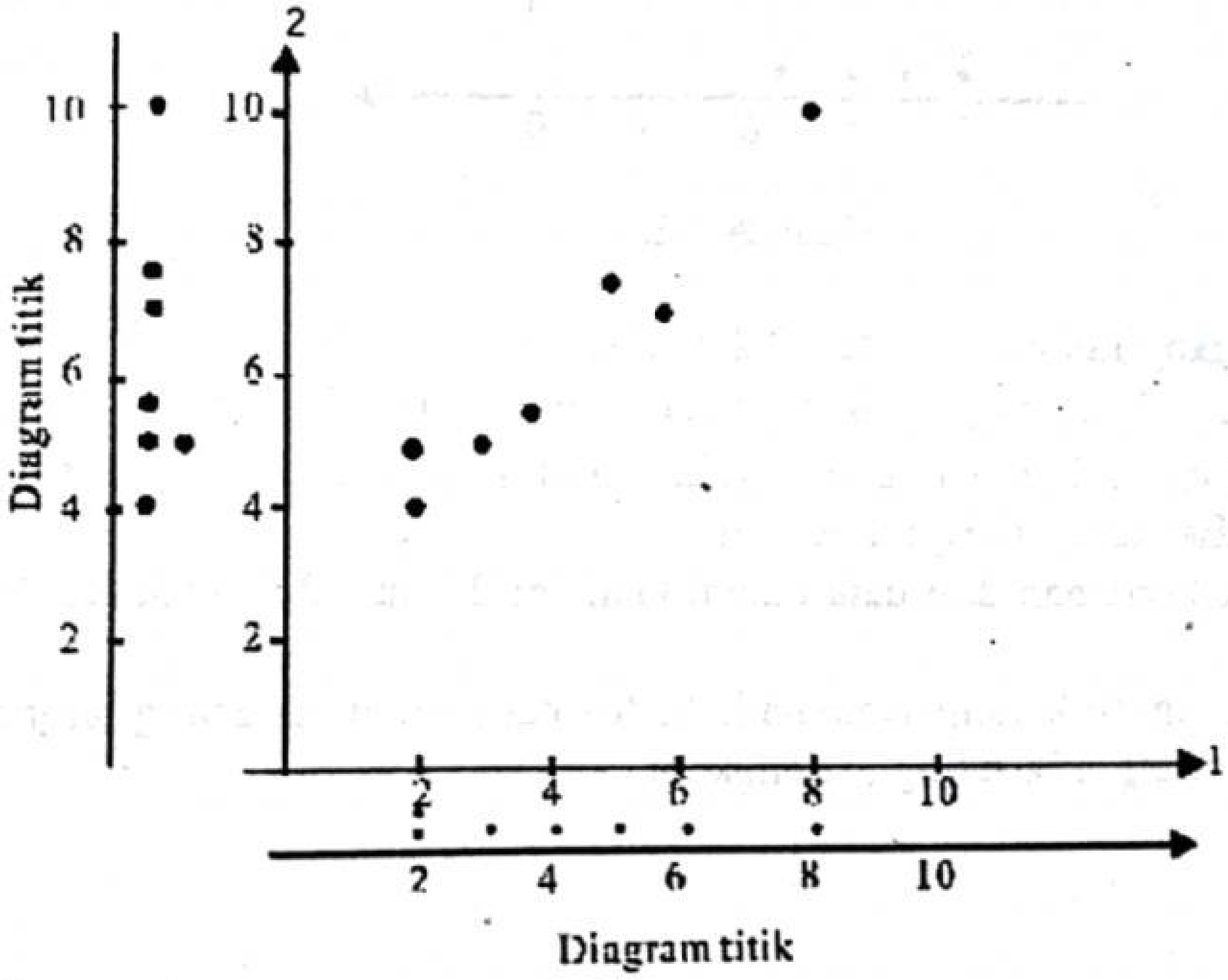
3. Teknik Grafik

Plot observasi sering diabaikan meski sebenarnya penting dan dapat membantu analisis data. Plot untuk semua pengukuran pada beberapa variabel hampir tidak mungkin dibuat. Walaupun demikian plot untuk masing-masing variabel dan pasangan variabel masih banyak memberikan informasi.

Misalkan 7 pasang pengukuran pada 2 variabel adalah:

Variabel 1 (x_1)	3	4	2	6	8	2	5
Variabel 2 (x_2)	5	5,5	4	7	10	5	7,5

Data ini diplot sebagai 7 titik pada 2 dimensi. Kooordinat titik-titik adalah (3;5), (4;5,5), … (5;7,5). Plot 2 dimensi ini disebut diagram pencar (*scatter diagram*).Dalam Gambar 2.1 disajikan plot individual variabel 1 dan 2 (diagram titik marginal variabel 1 dan 2) beserta diagram pencarnya.Diagram titik didapat dengan memproyeksikan observasi atau diagram pencar pada masing-masing sumbu.



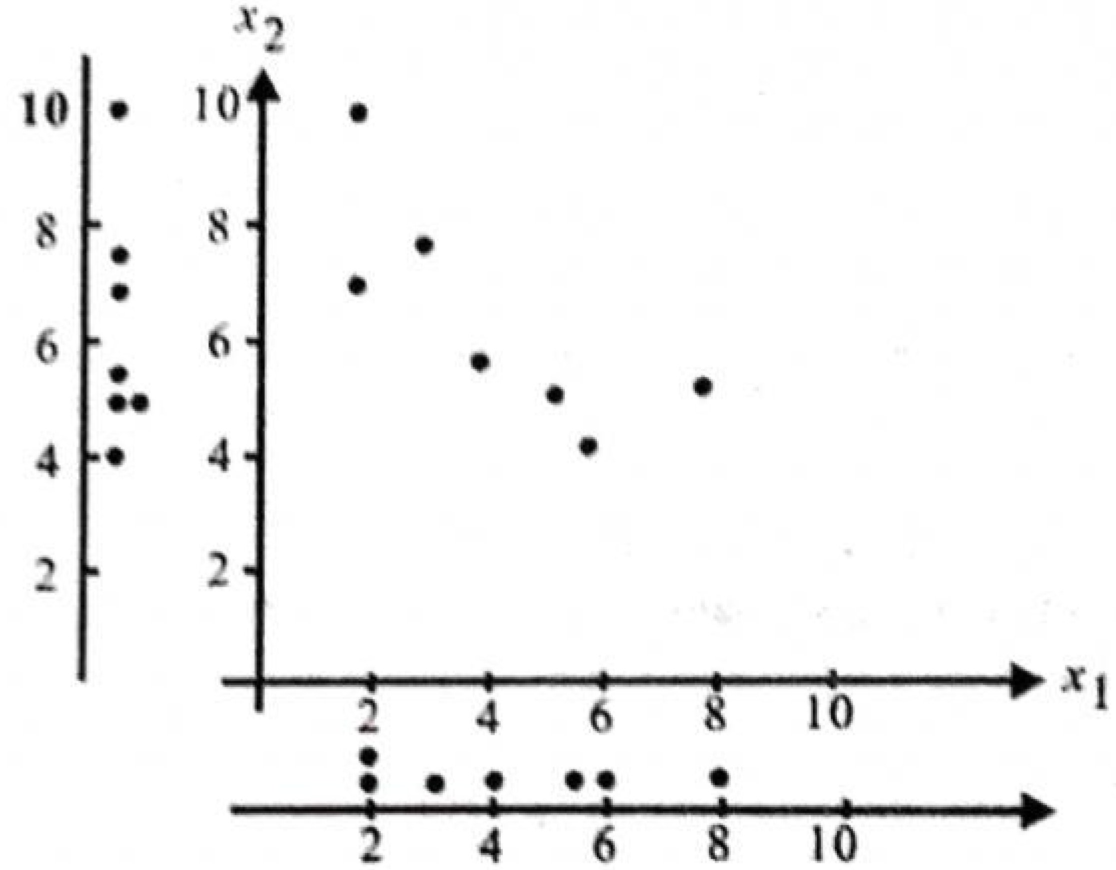
Gambar 2.1
Diagram Titik

Diagram titik dapat digunakan untuk menghitung mean sampel x_1, $x_2$dan variansi sampelS_{11} dan S_{22}. Diagram pencar dan koordinatnya dapat digunakan untuk menghitung kovariansi sampelS_{12}.Dari Gambar 2.1, $x_1$besar untuk $x_2$besar, $x_1$kecil untuk $x_2$kecil sehingga S_{12}bertanda positif.

Diagram titik dan diagram pencar memberikan informasi yang berbeda. Informasi pada diagram titik tidak cukup untuk mengkonstruksikan diagram pencar.Sebagai ilustrasi Anda perhatikan data baru

Variabel 1 (x_1)	5	4	6	2	2	8	3
Variabel 2 (x_2)	5	5,5	4	7	10	5	7,5

Diagram titik dan diagram pencar dari data disajikan dalam Gambar 2.2.



Gambar 2.2

Dengan perbandingan Gambar 2.1 dan 2.2 tampak bahwa diagram titik (marginal) yang sama memberikan diagram pencar berbeda. Dalam hal ini $\overline{x}_1, \overline{x}_2, S_{11}, S_{12}$ tidak berubah tetapi dari Gambar 2.2 tampak bahwax_1 kecil untuk x_2 besar dan x_1 besar untuk x_2 kecil sehingga S_{12} berubah tanda menjadi negatif.

Jadi, orientasi yang berlainan dari data dalam Gambar 2.1 dan Gambar 2.2 tidak terlihat dari diagram titiknya saja.Demikian juga diagram titik yang sama pada kedua kasus tidak langsung tampak dari diagram pencarnya. Kedua grafik saling melengkapi.

C. JARAK

Kebanyakan teknik multivariat berdasar pada konsep jarak. Jarak garis lurus (jarak Euclidian) telah Anda ketahui. Contohnya jarak antara

$P\left(x_{1,}x_2\right)$ dan O(0,0) adalah

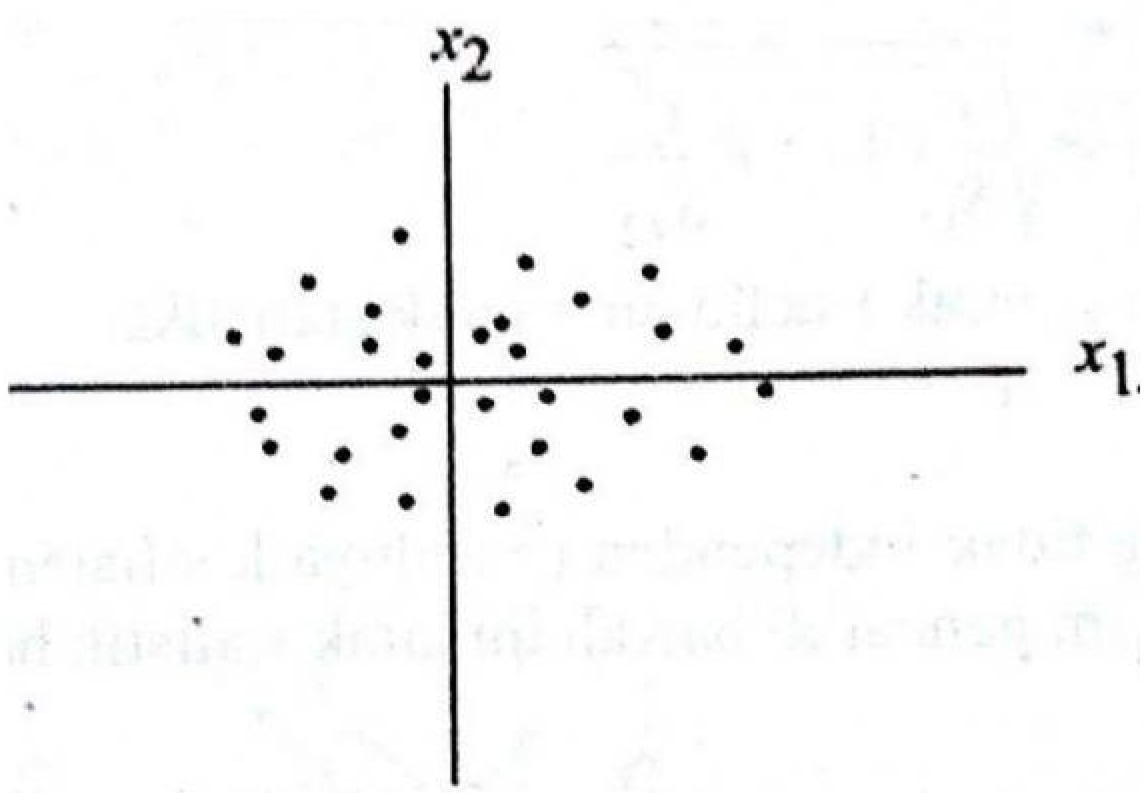
$$d_{(OP)} = \sqrt{x_1^2 + x_2^2}$$

Secara umum jarak antara $P\left(x_1,\ldots,x_p\right)$ dan $Q\left(y_{1,}\ldots,y_p\right)$ adalah

$$d(P,Q) = \sqrt{\left(x_1 - y_1\right)^2 + \ldots + \left(x_p - y_p\right)^2}$$

Jarak Euclidian ini tidak cocok digunakan untuk perhitungan dalam statistik.Hal ini disebabkan oleh kontribusi yang sama dari tiap koordinat untuk perhitungan jarak. Bila koordinat merupakan pengukuran yang mempunyai fruktuasi random dari magnitude berbeda, dibutuhkan pemberian bobot pada tiap-tiap koordinat dengan bobot besar untuk variabilitas yang kecil dan sebaliknya.Hal ini akan memberikan ukuran jarak yang berbeda dengan jarak Eulidian. Jarak yang digunakan dalam statistik diberi nama *jarak statistik.*

Untuk himpunan observasi (p variabel) dibuat diagram pencar p dimensi. Akan dicari jarak statistik $O(0, \ldots, 0)$ dan $P\left(x_1, x_2 \ldots x_p\right)$. Sebagai ilustrasi diambil$n$ titik pada 2 variabel x_1, x_2 yang *saling independen.*



Gambar 2.3

Variansi x_1 lebih besar variansi x_2 seperti tampak pada Gambar 2.3. Untuk suatu titik yang berjarak (Euclidian) sama terhadap kedua sumbu koordinat, koordinat (statistik) x_1 tidak sebesar koordinat x_2.

Jadi bobot x_1 lebih kecil untuk bobot x_2. Salah satu cara untuk memberikan bobot adalah dengan menggunakan standar deviasi sampel. Membagi koordinat dengan standar devisiasinya didapat standarisasi koordinat.

$$x_1^* = \frac{x_1}{\sqrt{S_{11}}} \text{ dan } x_2^* = \frac{x_2}{\sqrt{S_{22}}}$$

$$d(O,P)=\sqrt{\left(x_1^*\right)^2+\left(x_2^*\right)^2}=\sqrt{\left(\frac{x_1}{\sqrt{S_{11}}}\right)^2+\left(\frac{x_2}{\sqrt{S_{22}}}\right)^2}=\sqrt{\frac{x_1^2}{S_{11}}+\frac{x_2^2}{S_{22}}}$$

Perbandingan jarak Euclidian dan statistik terletak pada bobot $k_1=\frac{1}{S_{11}}$ dan $k_2=\frac{1}{S_{22}}$ pada x_1^2 dan x_2^2. Bila variansi sampel sama, $k_1=k_2$ atau x_1^2 dan x_2^2 mempunyai bobot sama maka jarak Euclidian = jarak statistik.

Semua titik dengan koordinat $\left(x_{1,}x_2\right)$ dan mempunyai jarak konstan$c$ terhadap 0 memenuhi

$\frac{x_1^2}{S_{11}}+\frac{x_2^2}{S_{22}}=c^2$ (persamaan elips)

Perluasan jarak statistik untuk p variabel adalah:

$P(x_1,x_2,\ldots,x_p)$, $Q\left(y_1,y_2,\ldots y_p\right)$

$$d\left(P,Q\right)=\sqrt{\frac{\left(x_1-y_1\right)}{S_{11}}+\frac{\left(x_2-y_2\right)^2}{S_{22}}+\ldots+\frac{\left(x_p-y_p\right)^2}{S_{pp}}} \tag{2.8}$$

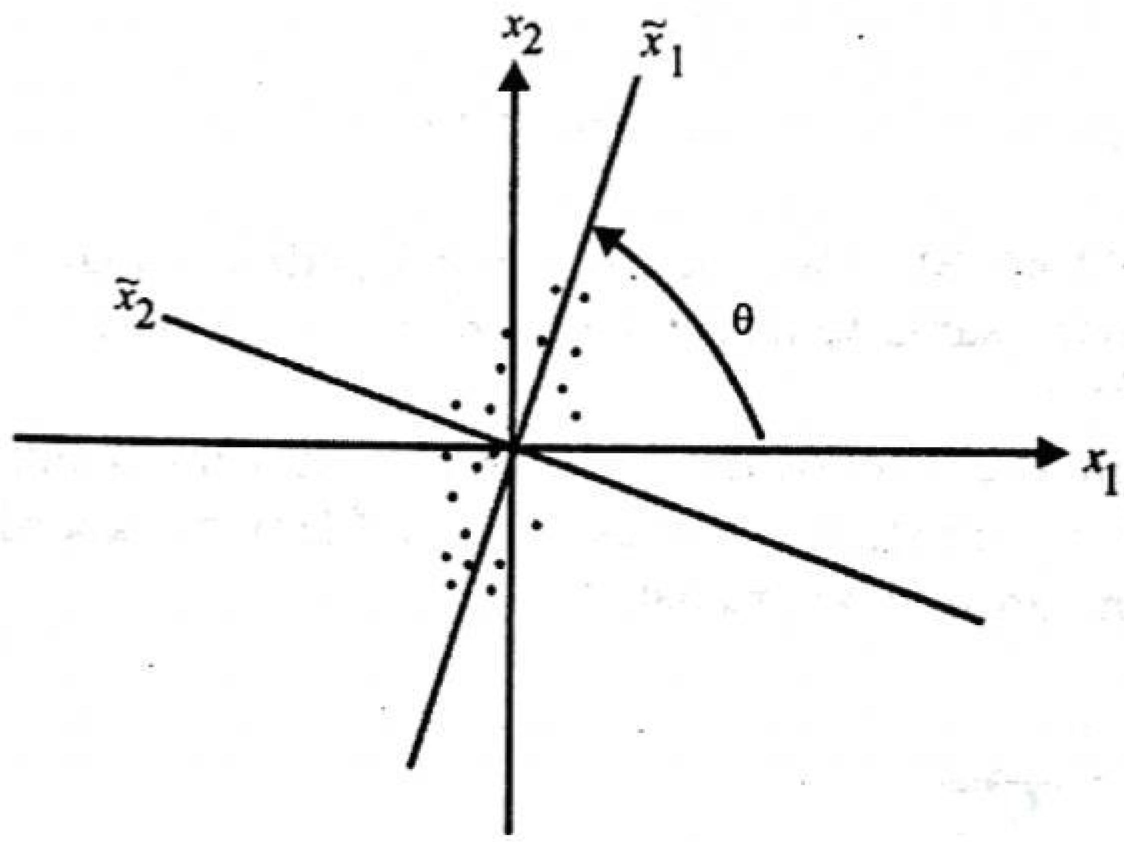
Semua titik dengan jarak tetap terhadap Q terletak pada hiper elipsolid yang berpusat di Q dengan sumbu mayor dan minor sejajar sumbu-sumbu koordinat.

a. Jarak P ke$O$, $d\left(O,P\right)=\sqrt{\frac{x_1^2}{S_{11}}+\ldots+\frac{x_2^2}{S_{22}}}$

b. Bila $S_{11}=S_{22}=\cdots=S_{pp}$, jarak Euclidian = jarak statistik.

Catatan

Untuk 2 variabel yang tidak independen (misalnya koefisien kolerasi sampel positif) seperti tampak pada diagram pencar dibawah ini jarak statistik harus dimodifikasi.



Gambar 2.4

Dari Gambar 2.4 tampak bahwa melakukan rotasi sebesar sudut θ sehingga didapat sumbu baru x_1^2 dan x_2^2 di mana diagram pencar tampak seperti diagram dalam Gambar 2.3.

Dengan sumbu baru ini $P\left(\tilde{x}_1, \tilde{x}_2\right), O(0,0)$ mempunyai

$$d(O,P)=\sqrt{\frac{\tilde{x}_1^2}{S_{11}}+\dots\frac{\tilde{x}_2^2}{S_{22}}} \tag{2.9}$$

di mana $\tilde{x}_1 = x_1 \cos\theta + x_2 \sin\theta$

$\tilde{x}_2 = -x_1 \sin\theta + x_2 \sin\theta$

dengan perhitungan aljabar didapat

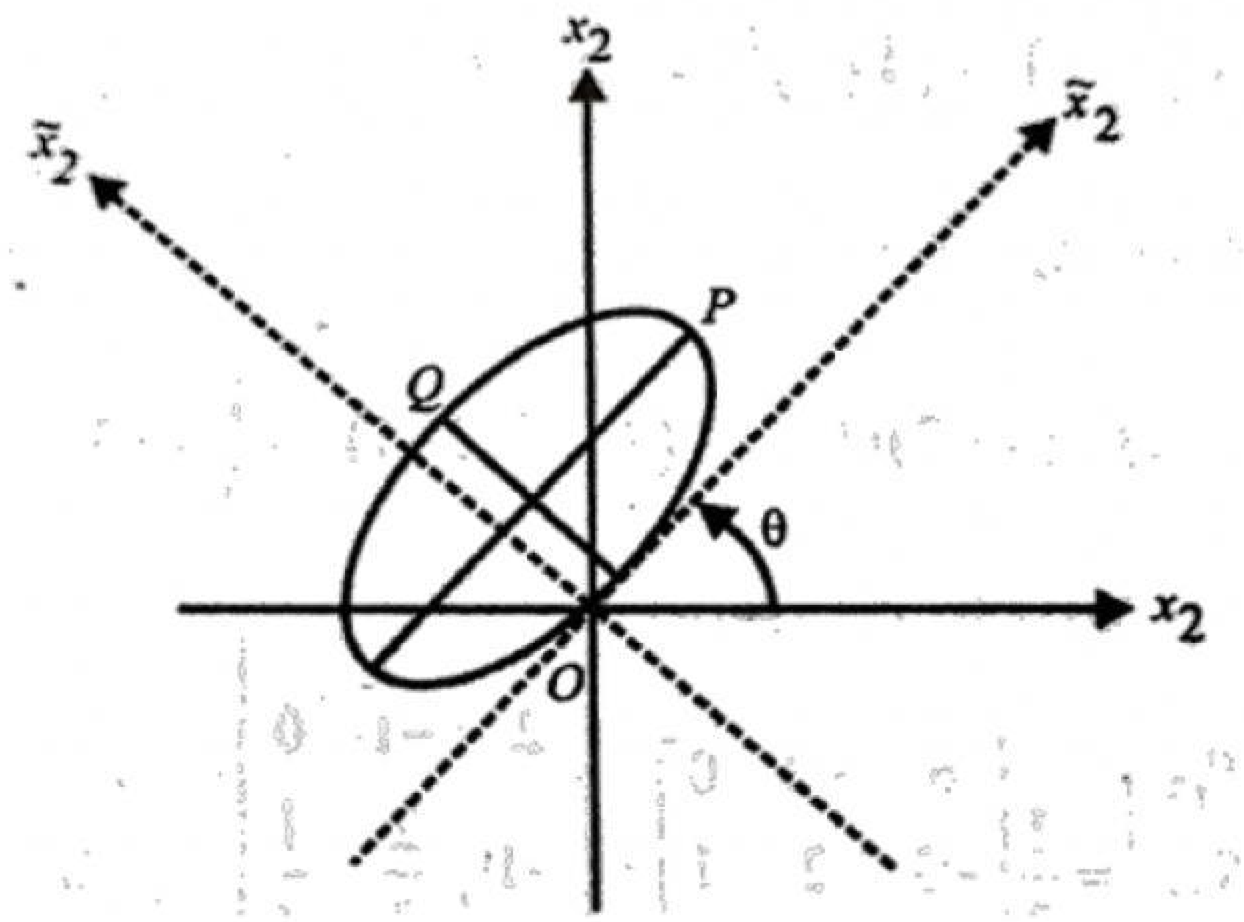
$$d(O,P)=\sqrt{a_{11}x_1^2+2a_{12}x_1x_2+a_{22}x_2^2} \tag{2.10}$$

dengan

$$a_{11}=\frac{\cos^2(\theta)}{\cos^2(\theta)S_{11}+2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{22}}+\frac{\sin^2(\theta)}{\cos^2(\theta)S_{22}-2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{11}}$$

$$a_{22}=\frac{\sin^2(\theta)}{\cos^2(\theta)S_{11}+2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{22}}+\frac{\cos^2(\theta)}{\cos^2(\theta)S_{22}-2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{11}}$$

$$a_{12}=\frac{\cos(\theta)\sin(\theta)}{\cos^2(\theta)S_{11}+2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{22}}-\frac{\cos\theta\sin\theta}{\cos^2(\theta)S_{22}-2\sin(\theta)\cos(\theta)S_{12}+\sin^2(\theta)S_{11}}$$



Gambar 2.5

Secara umum jarak $P(x_1, x_2)$ terhadap titik tetap $Q(y_1, y_2)$ bila kedua variabel berkolerasi adalah

$$d(P,Q) = \sqrt{a_{11}(x_1 - y_1)^2 + 2a_{12}(x_1 - y_1)(x_2 - y_2) + a_{22}(x_2 - y_2)^2} \quad (2.11)$$

Koordinat semua titik $P(x_1, x_2)$ yang mempunyai jarak tetap c^2 dari titik Q memenuhi

$$a_{11}(x_1 - y_1)^2 + 2a_{12}(x_1 - y_1)(x_2 - y_2) + a_{22}(x_2 - y_2)^2 = c^2$$

Contoh 2.2

1. Untuk data matriks

$$\underset{\sim}{x} = \begin{bmatrix} 1 & 4 & 4 \\ 2 & 1 & 0 \\ 5 & 6 & 4 \end{bmatrix}$$

2. Hitunglah $\overline{\underset{\sim}{x}}, \underset{\sim}{S}_n$ dan $\underset{\sim}{R}$

Penyelesaian:

$$\overline{x}_i = \sum_{j=1}^{3} x_{ij} ; i = 1,2,3$$

$$\underset{\sim}{x} = \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \overline{x}_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$$

$$S_{ik} = \frac{1}{3}\sum_{j=1}^{3}\left(x_{ij} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_k\right), i,k = 1,2,3$$

$$S_{11} = \frac{1}{3}\sum\left(x_{1j} - \overline{x}_2\right)^2$$

$$= \frac{1}{3}\left\{(1-3)^2 + (4-3)^2 + (4-3)^2\right\} = \frac{1}{3}.6 = 2$$

$$S_{12} = \frac{1}{3}\sum_{j=1}^{3}\left(x_{1j} - \overline{x}_1\right)\left(x_{2j} - \overline{x}_1\right)$$

$$= \frac{1}{3}\left\{(1-3)(2-1) + (4-3)(1-1) + (4-3)(0-1)\right\} = \frac{1}{3} = -1$$

dan seterusnya.

$$\underset{\sim}{S} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{13} & S_{23} & S_{33} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 6 & -3 & 0 \\ -3 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

$$r_{ik} = \frac{S_{ik}}{\sqrt{S_{ii}}\sqrt{S_{kk}}} \qquad i,k = 1,2,3$$

$$r_{11} = \frac{S_{11}}{\sqrt{S_{11}}\sqrt{S_{11}}} = 1$$

$$r_{12} = \frac{S_{12}}{\sqrt{S_{11}\sqrt{S_{22}}}} = \frac{-1}{\sqrt{2}\sqrt{\frac{2}{3}}} = -0,86$$

$$r_{13} = \frac{S_{13}}{\sqrt{S_{11}}\sqrt{S_{33}}} = 0$$

$$r_{23} = \frac{S_{23}}{\sqrt{S_{22}}\sqrt{S_{33}}} = \frac{\frac{1}{3}}{\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}} = \frac{1}{2}$$

$$\underset{\sim}{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{12} & r_{22} & r_{23} \\ r_{13} & r_{23} & r_{33} \end{bmatrix} = \begin{bmatrix} 1 & -0,86 & 0 \\ -0,86 & 1 & 0,5 \\ 0 & 0,5 & 1 \end{bmatrix}$$

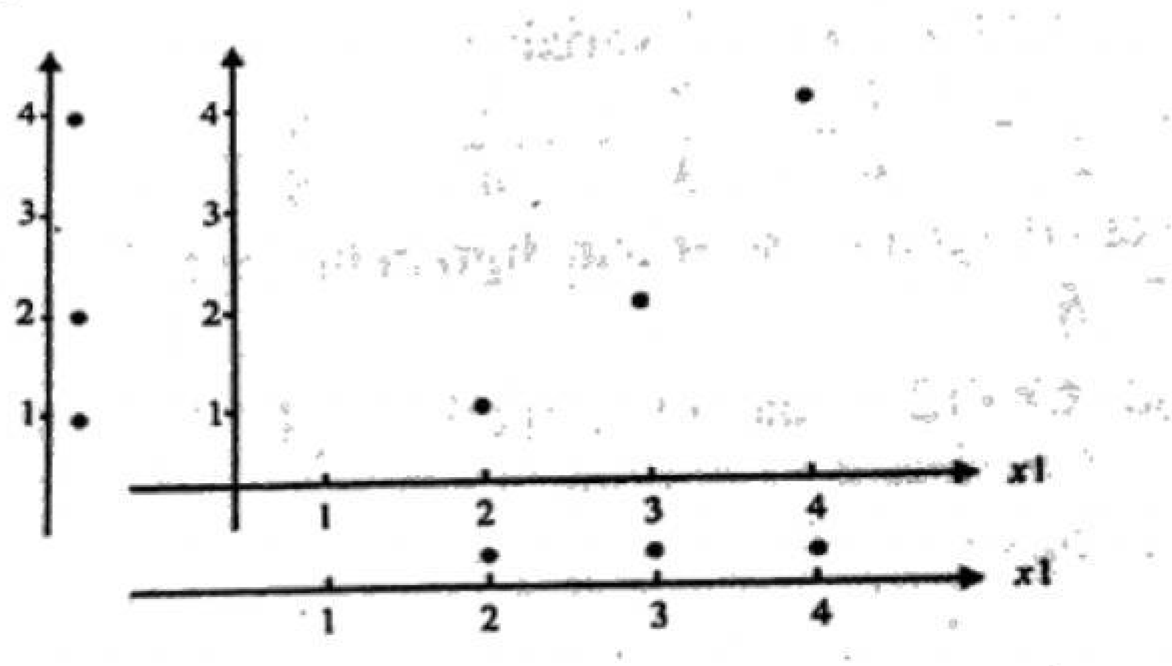
Contoh 2.3

Diberikan 3 observasi untuk data dua variabel.

Variabel 1	$x_{11} = 2$	$x_{12} = 3$	$x_{13} = 4$
Variabel 2	$x_{21} = 1$	$x_{22} = 2$	$x_{23} = 4$

Buatlah diagram pencar dan diagram titiknya.

Penyelesaiannya :



Contoh 2.4

1. Carilah jarak Euclidian!
2. Carilah jarak statistik dengan $a_{11} = \frac{1}{3}, a_{22} = \frac{4}{27}, a_{12} = \frac{1}{9}$

Penyelesaian:

1. $d(P,Q) = \sqrt{(-1-1)^2 + (-1+0)^2} = \sqrt{5}$
2. $d(P,Q) = \sqrt{a_{11}(x_1 - y_1)^2 + 2a_{12}(x_1 - y_1)(x_2 - y_2) + a_{22}(x_2 - y_2)^2}$

$$= \sqrt{\frac{1}{3}(-1-1)^2 + 2.\frac{1}{9}(-1-1)(-1-0) + \frac{4}{27}(-1-0)^2}$$

$$= \sqrt{\frac{4}{3} + \frac{4}{9} + \frac{4}{27}} = \frac{2}{3}\sqrt{\frac{13}{3}}$$

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Dalam sebuah surat kabar tertera daftar harga mobil bekas (dalam jutaan rupiah) beserta usia mobil (dalam tahun) x_1 menyatakan usia mobil (dalam tahun), x_2 menyatakan harga (dalam jutaan rupiah).

x_1	3	5	5	7	7	7	8	9	10	1
x_2	2,3	1,9	1,0	0,7	0,3	1,0	1,05	0,45	0,70	0,3

   a) Konstruksikan diagram pencar dan diagram titiknya.
   b) Cocokkan keadaan yang nampak dari diagram pencar dengan tanda dari S_{12}
   c) Hitung $\bar{x}_1, \bar{x}_2, S_{11}, S_{12}$
   d) Tuliskan matriks $\underset{\sim}{\bar{x}}, \underset{\sim}{S}_n, \underset{\sim}{R}$.

2) Diketahui 5 pengukuran untuk 3 variabel x_1, x_2, x_3 adalah sebagai berikut.

x_1	9	2	6	5	8
x_2	12	8	6	4	10
x_3	3	4	0	2	1

   Tulis matriks x, S_n, R.

3) Diberikan 6 observasi untuk $p = 2$ variabel.

x_1	2	3	4	2	3	4
x_2	1	2	4	3	4	5

   a) Konstruksikan diagram pencar dan diagram titiknya.
   b) Hitung $\underset{\sim}{\bar{x}}, \underset{\sim}{S}_n, \underset{\sim}{R}$.

4) Hitung jarak statistik $P(2, 2)$ dan O(0, 0) bila $\theta = 45^\circ$

Petunjuk Jawaban Latihan

1) a) Hitung S_{12} pandang diagram pencar.
   b) Jelas.
   c) $\overline{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$

   $$S_{ik} = \frac{1}{n}\sum_{j=1}^{n}\left(x_{ij} - \overline{x}_i\right)\left(x_{kj} - \overline{x}_k\right)$$

   d) $\underset{\sim}{\overline{x}} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}; \underset{\sim}{S}_n = \begin{bmatrix} S_{11} & \cdots & S_{pp} \\ \vdots & \ddots & \vdots \\ S_{1p} & \cdots & S_{pp} \end{bmatrix}; R = \begin{bmatrix} 1 & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{1p} & \cdots & 1 \end{bmatrix}$
2) Sama dengan nomor 1c, d.
3) Sama dengan no. 1, a-d.
4) Hitung dahulu a_{11}, a_{12}, a_{22} . Kemudian $d\left(P, O\right)$ dengan menggunakan rumus (2.11)

RANGKUMAN

1. Untuk Sampel

$$\underset{\sim}{x} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pn} \end{pmatrix}$$

Dihitung dengan pusat sebaran dan ukuran keeratan hubungan linear antara satu variabel dengan variabel lainnya, yaitu

$$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \underset{\sim}{S}_n = \begin{pmatrix} S_{11} & \cdots & S_{1p} \\ \vdots & \ddots & \vdots \\ S_{p1} & \cdots & S_{pp} \end{pmatrix} \underset{\sim}{R} = \begin{pmatrix} 1 & \cdots & r_{1p} \\ \vdots & \ddots & \vdots \\ r_{1p} & \cdots & 1 \end{pmatrix}$$

2. Untuk melihat bentuk visual data dibuat dengan diagram pencar dan diagram titik (marginal), kedua diagram ini dapat membatu analisis data.
3. Konsep jarak statistik memperhitungkan variabilitas. Misalnya untuk kedua variabel

a. bila kedua variabel independen $P(x_1, x_2), Q(y_1, y_2)$

$$d(P,Q) = \frac{(x_1 - y_1)^2}{S_{11}} + \frac{(x_2 - y_2)^2}{S_{22}}$$

b. bila kedua variabel tidak independen

$$d(P,Q) = \sqrt{a_{11}(x_1 - y_1)^2 + 2a_{12}(x_1 - y_1)(x_2 - y_2) + 2a_{22}(x_2 - y_2)^2}$$

TES FORMATIF 1

Pilihlah satu jawaban yang paling tepat!

Diketahui dua variabel x_1, x_2 dengan

x_1	42	52	48	58
x_2	4	5	4	3

1) $\underset{\sim}{S}_n = ...$

A. $\begin{bmatrix} 29 & -1 \\ -1 & 0,5 \end{bmatrix}$

B. $\begin{bmatrix} 30 & -2 \\ -2 & 4 \end{bmatrix}$

C. $\begin{bmatrix} 20 & -1 \\ -1 & 25 \end{bmatrix}$

D. A, B, C tidak benar

2) $\underset{\sim}{R} = ...$

A. $\begin{bmatrix} 1 & -0,2 \\ -0,2 & 1 \end{bmatrix}$

B. $\begin{bmatrix} 1 & -0,36 \\ -0,36 & 1 \end{bmatrix}$

C. $\begin{bmatrix} 1 & -0,26 \\ -0,26 & 1 \end{bmatrix}$

D. A, B, C tidak benar

3) Untuk n observasi, p variabel.
   A. Bila variabel-variabel saling independen jarak Euclidian = jarak Statistik.
   B. Bila kovarian setiap pasang variabel sama dengan nol, maka jarak Euclidian = jarak statistik.
   C. Bila $S_{11} = S_{22} = \ldots = S_{pp} = a^2$, maka Jarak Euclidian = Jarak Statistik.
   D. A, B, C tidak benar.

4) Jarak $P(1,1)$ dan $O(0,0)$ bila $a_{11} = \frac{1}{3}, a_{22} = \frac{4}{27}, a_{12} = \frac{1}{9}$ adalah ….
   A. $\sqrt{\frac{19}{27}}$
   B. $\frac{4}{3}\sqrt{3}$
   C. $4\sqrt{3}$
   D. A, B, C tidak benar

5) Diketahui 8 pasang pengukuran untuk x_1 dan x_2 sebagai berikut.

x_1	-6	-3	-2	1	2	5	6	8
x_2	-2	-3	1	-1	2	1	5	3

   Dianggap sumbu koordinat
   $Q = 26^o, \cos 26 = 0{,}899, \sin 26^o = 0{,}438$.
   Bila P mempunyai koordinat (4,-2) hitunglah $d(0,P)$
   A. $\sqrt{29}$
   B. $\sqrt{\frac{30}{6}}$
   C. 0,872
   D. A, B, C tidak benar

6) Diketahui observasi untuk total pembelian (x_1) dan pembelian jumlah buku (x_2) sebagai berikut.

$$\underset{\sim}{x} = \begin{bmatrix} 42 & 26 & 52 & 60 & 48 & 50 & 58 \\ 4 & 2 & 5 & 6 & 4 & 5 & 3 \end{bmatrix}$$

Data ini mempunyai mean =

A. $\begin{bmatrix} 48 \\ 4,143 \end{bmatrix}$

B. $\begin{bmatrix} 48 \\ 4,7 \end{bmatrix}$

C. $\begin{bmatrix} 47,5 \\ 4,2 \end{bmatrix}$

D. A, B, C tidak benar

7) Dari soal no. 6 matriks kovariansi adalah ….

A. $\begin{bmatrix} 1 & 0,49 \\ 0,49 & 1 \end{bmatrix}$

B. $\begin{bmatrix} 130,67 & 10,67 \\ 10,67 & 1,81 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0,69 \\ 0,69 & 1 \end{bmatrix}$

D. A, B, C tidak benar

8) Dari soal no. 6 matriks koreksi adalah ….

A. $\begin{bmatrix} 1 & 0,84 \\ 0,84 & 1 \end{bmatrix}$

B. $\begin{bmatrix} 1 & 0,27 \\ 0,27 & 1 \end{bmatrix}$

C. $\begin{bmatrix} 1 & 0,69 \\ 0,69 & 1 \end{bmatrix}$

D. A, B, C tidak benar

9) Jarak Statistik ….
   A. sama dengan Euclidian
   B. lebih besar dari jarak Euclidean
   C. lebih kecil dari jarak Euclidean
   D. keduanya tidak bisa dibandingkan

10) Jarak Statistik = 5 jarak Euclidean artinya ….
   A. Bervariasinya data pada sumbu x_1, 5 kali besarnya variasi daripada sumbu yang lain.
   B. Bervariasinya data pada sumbu x_1, $\frac{1}{5}$ kali besarnya variasi daripada sumbu yang lain.
   C. Bervariasinya data pada semua sumbu sama.
   D. A, B, C tidak benar

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 1 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 1.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
< 70% = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan Kegiatan Belajar 2. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 1, terutama bagian yang belum dikuasai.

KEGIATAN BELAJAR 2

Vektor Random dan Matrik Random

Vektor random adalah vektor yang elemen-elemennya variabel random. Matriks random adalah matriks yang elemen-elemennya variabel random. Harga harapan matriks random adalah harga harapan masing-masing elemennya.

Misal $\underset{\sim}{X} = \{x_{ij}\}$ adalah matriks random berdimensi $p \times n$. Akan diberikan definisi untuk harga harapan matriks random

Definisi 2.1

Harga harapan dari $\underset{\sim}{x}_{pxn}$ ditulis $E(\underset{\sim}{X})$ adalah

$$E(\underset{\sim}{X}) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1n}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{p1}) & E(X_{p2}) & \cdots & E(X_{pn}) \end{bmatrix} \tag{2.12}$$

di mana

$$E(x_{ij}) = \begin{cases} \int\limits_{-\infty}^{\infty} x_{ij} f_{ij}(x_{ij})\, dx_{ij} \\ \sum\limits_{semua\, x_{ij}} x_{ij} P_{ij}(x_{ij}) \end{cases} \qquad 2.13)$$

bila X_{ij} variabel random kontinu dengan fungsi densitas $f_{ij}(x_{ij})$

bila X_{ij} variabel random diskrit dengan fungsi probalitas $p_{ij}(x_{ij})$.

Teorema 2.1

Bila $\underset{\sim}{X}$ dan $\underset{\sim}{Y}$ adalah matriks random dengan dimensi sama maka $\underset{\sim}{A}$ dan $\underset{\sim}{B}$ matriks konstan yang conformabel dengan $\underset{\sim}{X}$ dan $\underset{\sim}{Y}$ maka

$$E(\underset{\sim}{X} + \underset{\sim}{Y}) = E(\underset{\sim}{X}) + E(\underset{\sim}{Y})$$

$$E(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}) = \underset{\sim}{A} E(\underset{\sim}{X}) \underset{\sim}{B} \tag{2.14}$$

A. VEKTOR MEAN DAN MATRIKS KOVARIANSI

Definisi 2.2

Mean vektor random

$\underset{\sim}{X} = \left[X_1, X_2, \ldots, X_p \right]$ adalah vektor random berdimensi p x 1

$$\mu_i = E\left(X_i\right) = \begin{cases} \int\limits_{-\infty}^{\infty} x_i f_i\left(x_i\right) dx_i; \\ \sum\limits_{semua\, x_i} x_i P_i\left(x_i\right); \end{cases} \tag{2.15}$$

bila X_i variabel random kontinu dengan fungsi densitas $f_i\left(x_i\right)$

bila X_i variabel random diskrit dengan fungsi probabilitas $p_i\left(x_i\right)$

Definisi 2.3

Ukuran keeratan hubungan linear antara variabel random X_i dan X_k oleh kovariansi σ_{ik} dengan $\sigma_{ik} = E\left(x_i - \mu_i\right)\left(x_k - \mu_k\right)$

$$= \begin{cases} \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \left(x_i - \mu_i\right)\left(x_k - \mu_k\right) f_{ik}\left(x_i, x_k\right); \\ \sum\limits_{semua\, x_i x_k} \left(x_i - \mu_i\right)\left(x_k - \mu_k\right) f_{ik}\left(x_i, x_k\right); \end{cases} \tag{2.16}$$

bila X_i, X_k variabel random kontinu dengan fungsi densitas $f_{ik}\left(x_i, x_k\right)$

bila X_i, X_k variabel random diskrit dengan fungsi probabilitas $p_{ik}\left(x_i, x_k\right)$

Bila $i = k$, kovariansi menjadi variansi.

Teorema 2.2

$$Cov\left(X_i, X_k\right) = \sigma_{ik} = E\left(X_i - \mu_i\right)\left(X_i - \mu_k\right) = E\left(X_i X_k\right) - \mu_i \mu_k \tag{2.17}$$

Bukti:

$E\left(X_i - \mu_i\right)\left(X_i - \mu_k\right)$

$$= E(X_iX_k - \mu_iX_k - \mu_kX_i + \mu_i\mu_k)$$
$$= E(X_iX_k) - \mu_iE(X_k) - \mu_kE(X_i) + \mu_i\mu_k$$
$$= E(X_iX_k) - \mu_i\mu_k$$

Teorema 2.3

$$\operatorname{var}(X_i) = \sigma_{ii} = E(X_i - \mu_i)^2 = E(X_i^2) - \mu_i^2 \tag{2.18}$$

Bukti:

$$E(X_i - \mu_i)^2 = E(X_i^2 - 2\mu_iX_i + \mu_i^2)$$
$$= E(X_i^2) - 2\mu_iE(X_i) + \mu_i^2$$
$$= E(X_i^2) - \mu_i^2$$

Teorema 2.4

Bila X_i dan X_k independen maka $\operatorname{cov}(X_i, X_k) = 0$ (2.19)

Bukti

X_i dan X_k independen

$$f_{ik}(x_i, x_k) = f(x_i)f(x_k)$$
$$E(X_i, X_k) = E(X_i)E(X_k) = \mu_i\mu_k$$
$$\operatorname{cov}(X_i, X_k) = E(X_iX_k) - \mu_i\mu_k = \mu_i\mu_k - \mu_i\mu_k = 0$$

Mean dan kovariansi vektor random $\underset{\sim}{x}_{px1}$ dapat ditulis dalam bentuk matriks.Harga harapan dalam setiap elemen termuat dalam vektor mean $\underset{\sim}{\mu} = t(\underset{\sim}{x})$. P variansi σ_{ii} dan $p(p-1)/2$ kovariansi $\sigma_{ik}\,(i<k)$ termuat dalam matriks variansi – kovariansi (yang jelas simetris), $\underset{\sim}{\Sigma} = E(\underset{\sim}{x} - \underset{\sim}{\mu})(\underset{\sim}{x} - \underset{\sim}{\mu})'$.

$$E(\underset{\sim}{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \underset{\sim}{\mu} \tag{2.20}$$

$$\underset{\sim}{\Sigma} = E(\underset{\sim}{X} - \underset{\sim}{\mu})(\underset{\sim}{X} - \underset{\sim}{\mu})'$$

$$= E\left(\begin{bmatrix} (X_1-\mu_1) \\ \vdots \\ (X_p-\mu_p) \end{bmatrix}\left[X_1-\mu_1,\ldots,X_p-\mu_p\right]\right)$$

$$= E\begin{pmatrix} (X_1-\mu_1)^2 & (X_1-\mu_1)(X_2-\mu_2) & \cdots & (X_1-\mu_1)(X_p-\mu_p) \\ (X_2-\mu_2)(X_1-\mu_1) & (X_2-\mu_2)^2 & \cdots & (X_2-\mu_2)(X_p-\mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ (X_p-\mu_p)(X_1-\mu_1) & (X_p-\mu_p)(X_2-\mu_2) & \cdots & (X_p-\mu_p)^2 \end{pmatrix}$$

$$= E\begin{pmatrix} E(X_1-\mu_1)^2 & E(X_1-\mu_1)^2(X_2-\mu_2) & \cdots & E(X_1-\mu_1)(X_p-\mu_p) \\ E(X_2-\mu_2)(X_1-\mu_1) & E(X_2-\mu_2)^2 & \cdots & (X_2-\mu_2)(X_p-\mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p-\mu_p)(X_1-\mu_1) & E(X_p-\mu_p)(X_2-\mu_2) & \cdots & E(X_p-\mu_p)^2 \end{pmatrix}$$

$$\underset{\sim}{\Sigma} = \operatorname{cov}(\underset{\sim}{X}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix} \qquad (2.21)$$

Karena $\sigma_{ik} = \sigma_{ki}$ maka

$$\underset{\sim}{\Sigma} = \operatorname{cov}(\underset{\sim}{X}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{pmatrix}$$

$\underset{\sim}{\Sigma}$ merupakan matriks simetris.

Perlu Anda ingat bahwa $\underset{\sim}{\mu}$ dan $\underset{\sim}{\Sigma}$ adalah vektor mean populasi dan matriks Varian kovarian populasi.

Catatan:

1. Vektor random $\underset{\sim}{X}' = \begin{bmatrix} X_1 & X_2 & \ldots & X_p \end{bmatrix}$ mempunyai fungsi densitas probabilitas $f\left(x_1, x_2, \ldots, x_p\right) = f\left(\underset{\sim}{x}\right)$. Dalam modul ini fungsi densitas normal multivariat banyak digunakan. Distribusi normal multivariat akan tertentu dengan lengkap bila vektor mean $\underset{\sim}{\mu}$ dan varian-kovarian matriks $\underset{\sim}{\Sigma}$ tertentu. Oleh karena itu, kuantitas ini memegang peranan penting dalam analisis multivariat.
2. Apabila $X_1, X_2, \ldots, X_p$ mutually independen maka $\operatorname{cov}\left(X_i, X_k\right) = \sigma_{ik} = 0, i \neq k.$

$$\text{Jadi, } \underset{\sim}{\Sigma} = \begin{bmatrix} \sigma_{11} & & \underset{\sim}{0} \\ & \sigma_{22} & \\ & & \ddots \\ \underset{\sim}{0} & & \sigma_{pp} \end{bmatrix} = D\left[\sigma_{11}, \sigma_{22}, \ldots, \sigma_{pp}\right]$$

Definisi 2.4

Ukuran keeratan hubungan linear antara variabel random X_i, X_k adalah koefisien kolerasi populasi, diberi notasi σ_{ik} dengan

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}} \tag{2.22}$$

Matriks koefisien kolerasi populasi adalah merupakan matriks simetris $\underset{\sim}{\rho}, p \times p$ dengan

$$\underset{\sim}{\rho} = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} \end{bmatrix}$$

$$\underset{\sim}{\rho} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{1p} & \cdots & 1 \end{bmatrix} \tag{2.23}$$

Misal matriks standar deviasi $\underset{\sim}{V}^{\frac{1}{2}}$, berukuran p x p adalah

$$\underset{\sim}{V}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{pp}} \end{bmatrix} \tag{2.24}$$

Teorema 2.5

$$\underset{\sim}{V}^{\frac{1}{2}} \underset{\sim}{\rho} \underset{\sim}{V}^{\frac{1}{2}} = \underset{\sim}{\Sigma} \text{ dan } \underset{\sim}{\rho} = \left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1} \underset{\sim}{\Sigma} \left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1} \tag{2.25}$$

Jadi, $\underset{\sim}{\Sigma}$ dapat dihitung dari $\underset{\sim}{V}^{\frac{1}{2}}$ dan $\underset{\sim}{\rho}$; dan $\underset{\sim}{\rho}$ dapat dihitung dari $\underset{\sim}{\Sigma}$.

Secara umum vektor random $\underset{\sim}{X}$, berukuran p x1 dapat dibagi menjadi 2 grup berukuran q dan $p-q$.

$\underset{\sim}{X}$ dapat ditulis:

$$\underset{\sim}{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ \cdots \\ X_{q+1} \\ \vdots \\ X_P \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ \cdots \\ X^{(2)} \end{bmatrix} \text{ dengan } \quad X^{(1)} = \begin{bmatrix} X_1 \\ \vdots \\ X_2 \end{bmatrix} \quad X^{(2)} = \begin{bmatrix} X_{q+1} \\ X_P \end{bmatrix} \tag{2.26}$$

$$\underset{\sim}{\mu} = E = \left(\underset{\sim}{X}\right)\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \\ \cdots\cdots \\ X_{q+1} \\ \vdots \\ \mu_p \end{bmatrix} = \begin{bmatrix} \mu^{(1)} \\ \cdots\cdots \\ \mu^{(2)} \end{bmatrix} \text{dengan } \mu^{(1)} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_2 \end{bmatrix}, \mu^{(2)} = \begin{bmatrix} \mu_{q+1} \\ \vdots \\ \mu_p \end{bmatrix} \tag{2.27}$$

$$\left(\underset{\sim}{X}^{(1)} - \underset{\sim}{\mu}^{(1)}\right)\left(\underset{\sim}{X}^{(2)} - \underset{\sim}{\mu}^{(2)}\right)' = \begin{bmatrix} x_1 - \mu_1 \\ x_1 - \mu_1 \\ \vdots \\ x_q - \mu_q \end{bmatrix}\left[x_{q+1} - \mu_{q+1}, x_{q+2} - \mu_{q+2}, \ldots, x_p - \mu_p\right]$$

$$= \begin{pmatrix} (x_1 - \mu_1)(x_{q+1} - \mu_{q+1}) & (x_1 - \mu_1)(x_{q+2} - \mu_{q+2}) & \cdots & (x_1 - \mu_1)(x_p - \mu_p) \\ (x_2 - \mu_2)(x_{q+1} - \mu_{q+1}) & (x_2 - \mu_2)(x_{q+2} - \mu_{q+2}) & \cdots & (x_2 - \mu_2)(x_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ (x_q - \mu_q)(x_{q+1} - \mu_{q+1}) & (x_q - \mu_q)(x_{q+2} - \mu_{q+2}) & \cdots & (x_q - \mu_q)(x_p - \mu_p) \end{pmatrix}$$

$$E\left(x^{(1)} - \underset{\sim}{\mu}^{(1)}\right)\left(x^{(2)} - \underset{\sim}{\mu}^{(2)}\right)' = \begin{pmatrix} \sigma_{1,q+1} & \sigma_{1,q+2} & \cdots & \sigma_{1p} \\ \sigma_{2,q+1} & \sigma_{2,q+2} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{q,q+1} & \sigma_{q,q+2} & \cdots & \sigma_{2p} \end{pmatrix} = \underset{\sim}{\Sigma}_{12} \tag{2.28}$$

$\underset{\sim}{\Sigma}_{12}$ mempunyai elemen-elemen kovariansi,

$$\sigma_{ij}, i = 1, 2 \ldots q, j = q+1, q+2, \ldots, p$$

antara komponen-komponen $X^{(1)}$ dan komponen-komponen $X^{(2)}$. $\underset{\sim}{\Sigma}_{12}$ tidak simetris ataupun bujur sangkar

$$\left(\underset{\sim}{x} - \underset{\sim}{\mu}\right)\left(\underset{\sim}{x} - \underset{\sim}{\mu}\right)' = \begin{vmatrix} \underset{\sim}{x}^{(1)} - \underset{\sim}{\mu}^{(1)} \\ \underset{\sim}{x}^{(2)} - \underset{\sim}{\mu}^{(2)} \end{vmatrix}\left|\left(\underset{\sim}{x}^{(1)} - \underset{\sim}{\mu}^{(1)}\right)' \quad \left(\underset{\sim}{x}^{(2)} - \underset{\sim}{\mu}^{(2)}\right)'\right|$$

$$\begin{bmatrix} \left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}^{(1)}\right)\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}_1^{(1)}\right)' & \left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}^{(1)}\right)\left(\underset{\sim}{x}^{(2)}-\underset{\sim}{\mu}_2^{(2)}\right)' \\ \scriptstyle qx1 \qquad 1xq & \scriptstyle qx1 \qquad 1x(p-q) \\ \left(\underset{\sim}{x}^{(2)}-\underset{\sim}{\mu}^{(2)}\right)\left(\underset{\sim}{x}^{(1)}-\underset{\sim}{\mu}_1^{(1)}\right)' & \left(\underset{\sim}{x}^{(2)}-\underset{\sim}{\mu}^{(2)}\right)\left(\underset{\sim}{x}^{(2)}-\underset{\sim}{\mu}_1^{(2)}\right)' \\ \scriptstyle (p-q)x1 \qquad 1xq & \scriptstyle (p-q)x1 \qquad 1x(p-q) \end{bmatrix}$$

$$\underset{\sim}{\Sigma} = E\left(\underset{\sim}{X}-\underset{\sim}{\mu}\right)\left(\underset{\sim}{X}-\underset{\sim}{\mu}\right)' = \begin{array}{c} q \\ \\ p-q \end{array}\begin{bmatrix} \overset{q}{\underset{\sim}{\Sigma}_{11}} & \vdots & \overset{p-q}{\underset{\sim}{\Sigma}_{12}} \\ \cdots & \vdots & \cdots \\ \underset{\sim}{\Sigma}_{21} & \vdots & \underset{\sim}{\Sigma}_{21} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1q} & \sigma_{1,q+1} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & \sigma_{q,q+1} & \cdots & \sigma_{qp} \\ \sigma_{q+1,1} & \cdots & \sigma_{q+1,q} & \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{p,q} & \sigma_{p,q+1} & \cdots & \sigma_{p,p} \end{bmatrix}$$

dengan

$$Kov\left(\underset{\sim}{X}^{(1)}\right) = \underset{\sim}{\Sigma}_{11} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1q} \\ \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} \end{bmatrix}$$

$$Kov\left(\underset{\sim}{X}^{(2)}\right) = \underset{\sim}{\Sigma}_{22} = \begin{bmatrix} \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & \ddots & \vdots \\ \sigma_{p,q+1} & \cdots & \sigma_{p,p} \end{bmatrix}$$

$$Kov\left(\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}\right) = \underset{\sim}{\Sigma}_{12} = \begin{bmatrix} \sigma_{1,q+1} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{q,q+1} & \cdots & \sigma_{2p} \end{bmatrix}$$

B. VEKTOR MEAN DAN MATRIKS KOVARIANSI UNTUK KOMBINASI LINEAR VARIABEL RANDOM

Teorema 2.6

Bila X_1, X_2 veriabel random dengan mean μ_1, μ_2, variansi σ_{11}, σ_{22} dan kovariansi σ_{12} maka

1. $E(cX_i) = c\mu_i$
2. $Var(cX_i) = c^2\sigma_{ii}$
3. $E(a_1X_1 + a_2X_2) = a_1\mu_1 + a_2\mu_2$
4. $Kov(a_1X_1, a_2X_2) = a_1a_2\sigma_{12}$
5. $Var(a_1X_1 + a_2X_2) = a_1^2\sigma_{11} + a_2^2\sigma_{22} + 2a_1a_2\sigma_{12}$

Bukti:

1. $E(cX_i) = cE(X_i) = c\mu_i$
2. $Var(cX_i) = E(cX_i - c\mu_i)^2 = c^2(X_i - \mu_i)^2$
$$= c^2Var(X_i) = c^2\sigma_{ii}$$
3. $E(a_1X_1 + a_2X_2) = a_1E(X_1) + a_2E(X_2) = a_1\mu_1 + a_2\mu_2$
4. $Kov(a_1X_1, a_2X_2) = E\left[(a_1X_1 - a_1\mu_1)(a_2X_2 - a_2\mu_2)\right]$
$$= a_1a_2E\left[(X_1 - \mu_1)(X_2 - \mu)\right]$$
$$= a_1a_2kov(X_1, X_2)$$
$$= a_1a_2\sigma_{12}$$
5. $Var(a_1X_1 + a_2X_2)$

$$= E\left((a_1X_1 + a_2X_2) - (a_1\mu_1 + a_2\mu_2)\right)^2$$
$$= E\left(a_1(X_1 - \mu_1) + a_2(X_2 - \mu_2)\right)^2$$
$$= E\left(a_1^2(X_1 - \mu_1)^2 + a_2^2(X_2 - \mu_2)^2 + 2a_1a_2(X_1 - \mu_1)(X_2 - \mu_2)\right)$$
$$= a_1^2E(X_1 - \mu_1)^2 + a_2^2E(X_2 - \mu_2)^2 + 2a_1a_2E\left[(X_1 - \mu_1)(X_2 - \mu_2)\right]$$
$$= a_1^2\sigma_{11} + a_2^2\sigma_{22} + 2a_1a_2\sigma_{12}$$

Kombinasi linear variabel random akan menjadi lebuh sederhana apabila ditulis dalam bentuk matriks

$$a_1X_1 + a_2X_2 = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \underset{\sim}{a}'\underset{\sim}{X} \text{ dengan } a' = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$$

Teorema 2.7

$\underset{\sim}{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ vektor random dengan mean $\underset{\sim}{\mu} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ dan matriks kovariansi

$\underset{\sim}{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$

Bila $y = a_1x_1 + a_2x_2 = \underset{\sim}{a}'\underset{\sim}{x}$ dengan $a' = \begin{bmatrix} a_1 & a_2 \end{bmatrix}$ maka

$$E\left(\underset{\sim}{a}'\underset{\sim}{X}\right) = \underset{\sim}{a}'\underset{\sim}{\mu} \tag{2.29}$$

$$Var\left(\underset{\sim}{a}'\underset{\sim}{X}\right) = \underset{\sim}{a}'\underset{\sim}{\Sigma}\underset{\sim}{a}$$

Bukti:

$$\begin{aligned} E\left(\underset{\sim}{a}'\underset{\sim}{X}\right) &= E\left(a_1X_1 + a_2X_2\right) \\ &= a_1E\left(X_1\right) + a_2E\left(X_2\right) \\ &= \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} E\left(X_1\right) \\ E\left(X_2\right) \end{bmatrix} = \underset{\sim}{a}' \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = a'\underset{\sim}{\mu} \end{aligned}$$

$$\begin{aligned} & Var\left(\underset{\sim}{a}'\underset{\sim}{X}\right) = Var\left(a_1X_1 + a_2X_2\right) \\ & = a_1^2\sigma_{11} + 2a_1a_2\sigma_{12} + a_2^2\sigma_{22} \\ & = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \underset{\sim}{\Sigma} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\ & = \underset{\sim}{a}'\underset{\sim}{\Sigma}\underset{\sim}{a} \end{aligned}$$

Secara umum dipandang q kombinasi linear p variabel random $x_1, \ldots, x_p$

$$Z_1 = c_{11}X_1 + c_{12}X_2 + \ldots + c_{1p}X_p$$

$$Z_2 = c_{21}X_1 + c_{22}X_2 + \ldots + c_{2p}X_p$$

$$\vdots$$

$$Z_q = c_{q1}X_1 + c_{q2}X_2 + \ldots + c_{qp}X_p$$

$$\underset{\sim}{Z} = \underset{qx1}{\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_q \end{bmatrix}} = \underset{qxp}{\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{q1} & c_{q2} & \cdots & c_{qp} \end{bmatrix}} \underset{px1}{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}} = \underset{\sim}{c}\underset{\sim}{X} \tag{2.30}$$

Teorema 2.8

Kombinasi linear $\underset{\sim}{Z} = \underset{\sim}{C}\underset{\sim}{X}$

$$\underset{\sim}{\mu}_Z = E\left(\underset{\sim}{Z}\right) = E\left(\underset{\sim}{C}\underset{\sim}{X}\right) = \underset{\sim}{C}\underset{\sim}{\mu}_x \tag{2.31}$$

$$\underset{\sim}{\Sigma}_z = Cov\left(\underset{\sim}{Z}\right) = Cov\left(\underset{\sim}{C}\underset{\sim}{X}\right) = \underset{\sim}{C}\underset{\sim}{\Sigma}_x\underset{\sim}{C}'$$

di mana $\underset{\sim}{\mu}_X$ dan $\underset{\sim}{\Sigma}_X$ adalah mean vektor dan kovarian matriks dari $\underset{\sim}{X}$.

C. PEMBAGIAN MEAN VEKTOR SAMPEL DAN MATRIKS KOVARIANSI SAMPEL

Misal $\vec{\underset{\sim}{x}} = \left[\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p\right]$ adalah vektor rata-rata sampel dari observasi untuk p variabel $X_1, X_2, \ldots, X_p$, dan misal

$$\underset{\sim}{S}_n = \begin{bmatrix} S_{11} & \cdots & S_{1p} \\ \vdots & \ddots & \vdots \\ S_{1p} & \cdots & S_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{3}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_1\right)^2 & \cdots & \frac{1}{n}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_1\right)\left(x_{pj} - \bar{x}_1\right) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\sum_{j=1}^{n}\left(x_{ij} - \bar{x}_1\right)\left(x_{pj} - \bar{x}_p\right) & \cdots & \frac{1}{n}\sum_{j=1}^{n}\left(x_{pj} - \bar{x}_p\right)^2 \end{bmatrix}$$

Mean vektor sampel dan matriks kovariansi dapat dipecah menurut kelompok variabelnya, misalnya

$$\underset{px1}{\underset{\sim}{\bar{x}}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_q \\ \hline \bar{x}_{q+1} \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \bar{x}^{(1)} \\ \hline \bar{x}^{(2)} \end{bmatrix} \tag{2.32}$$

$$\underset{pxp}{\underset{\sim}{S}_n} = \begin{bmatrix} S_{11} & \cdots & S_{1q} & S_{1,q+1} & \cdots & S_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ S_{q1} & \cdots & S_{qq} & S_{q,q+1} & \cdots & S_{qp} \\ S_{q+1,1} & \cdots & S_{q+1,q} & S_{q+1,q+1} & \cdots & S_{q+1,p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ S_{p1} & \cdots & S_{pq} & S_{p,q+1} & \cdots & S_{pp} \end{bmatrix} = \begin{array}{c c} & \begin{array}{cc} q & p-q \end{array} \\ \begin{array}{c} q \\ p-q \end{array} & \begin{bmatrix} \underset{\sim}{S}_{11} & \underset{\sim}{S}_{12} \\ \underset{\sim}{S}_{21} & \underset{\sim}{S}_{22} \end{bmatrix} \end{array} \tag{2.33}$$

di mana $\underset{\sim}{\bar{x}}^{(1)}$ dan $\underset{\sim}{\bar{x}}^{(2)}$ adalah mean vektor sampel dari observasi

$$x^{(1)} = \left[x_1, \ldots, x_q\right]' \text{ dan } x^{(2)} = \left[x_{q+1}, \ldots, x_{qp}\right]'$$

$\underset{\sim}{S}_{11}$ adalah matriks kovariansi sampel dari observasi dalam $\underset{\sim}{x}^{(1)}$

$\underset{\sim}{S}_{12}$ adalah matriks kovariansi sampel dari observasi dalam $\underset{\sim}{x}^{(2)}$

$\underset{\sim}{S}_{12} = \underset{\sim}{S}'_{21}$ adalah matriks kovariansi sampel dari elemen-elemen dalam $\underset{\sim}{x}^{(1)}$ dan elemen-elemen $\underset{\sim}{x}^{(2)}$.

Contoh 2.5

Diketahui vektor random $X' = \left[X_1, X_2\right]$ masing-masing komponennya diskrit dengan fungsi probabilitas.

X_1	-1	0	1
$p_1(x_1)$	0,3	0,3	0,4

X_2	0	1
$p_2(x_2)$	0,8	0,2

$$E(X_1) = \sum_{x_1} x_1 p_1(x_1) = (-1)(0,3) + (0)(0,3) + (1)(0,4) = 0,1$$

$$E(X_2) = \sum_{x_2} x_2 p_2(x_2) = (0)(0,8) + (1)(0,2) = 0,2$$

$$E(\underset{\sim}{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix} = \begin{bmatrix} 0,1 \\ 0,2 \end{bmatrix}$$

Contoh 2.6

Carilah matriks kovariansi X_1 dan X_2 dari contoh 1. Bila fungsi probabilitas bersamanya sebagai berikut:

x_1 \ x_2	0	1	$p_1(x_1)$
-1	0,24	0,06	0,3
0	0,16	0,14	0,3
1	0,40	0,00	0,4
$p_2(x_2)$	0,8	0,2	1

Dari contoh 2.5 didapat

$$\mu_1 = E(X_1) = 0,1 \qquad \mu_2 = E(X_2) = 0,2$$

$$\sigma_{11} = E(X_1 - \mu_1)^2 = \sum (X_1 - 0,1)^2 p_1(x_1)$$

$$= (-1-0,1)^2(0,3) + (0-0,1)^2(0,3) + (1-0,1)^2(0,4) = 0,69$$

atau $\sigma_{11} = E(X_1^2) - \mu_1^2$

$$= (1.0,3 + 0.0,3 + 1.0,4) - (0,1)^2$$

$$0,7 - 0,01 = 0,69$$

$$\sigma_{22} = E(X_2^2) - \mu_2^2$$

$$= 0,08 + 1.0,2 - (0,2)^2 = 0,2 - 0,04 = 16$$

$$\sigma_{12} = E(X_1X_2) = \mu_1.\mu_2$$
$$= (-1)(0)(0,24) + (-1)(1)(0,06) + (0)(0)(0,16) + (0)(1)(0,14 +$$
$$(1)(0)(0,40) + (1)(1)(0,00) - (0,1)(0,2) = -0,08$$
$$\sigma_{21} = E(X_1X_2) - \mu_1.\mu_2 = -0,08$$

Jadi, untuk $\underset{\sim}{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$$\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0,1 \\ 0,2 \end{bmatrix}$$
$$\underset{\sim}{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 0,69 & -0,08 \\ -0,08 & 0,16 \end{bmatrix}$$

Contoh 2.7

$$\underset{\sim}{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$$

Hitung $\underset{\sim}{V}^{\frac{1}{2}}$ dan $\underset{\sim}{\rho}$

Penyelesaian:

$$\underset{\sim}{V}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & 0 \\ 0 & \sqrt{\sigma_{22}} & 0 \\ 0 & 0 & \sqrt{\sigma_{33}} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$$

$$\underset{\sim}{\rho} = \left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1} \underset{\sim}{\Sigma} \left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1}$$

$$= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{1}{6} & \frac{1}{5} \\ \frac{1}{6} & 1 & -\frac{1}{5} \\ \frac{1}{5} & -\frac{1}{5} & 1 \end{bmatrix}$$

Contoh 2.8

Hitung ρ_{13} dan koefisien korelasi antara X_1 dan $\frac{1}{2}X_2 + \frac{1}{2}X_3$ dari Contoh 2.7.

Penyelesaian:

ρ_{13} elemen baris ke-1 kolom ke-3 matriks korelasi ρ

$$\rho_{13} = \frac{1}{5}$$

$$Corr\left(X_1, \frac{1}{2}X_2 + \frac{1}{2}X_3\right)$$

$$= \frac{\operatorname{cov}\left(X_1, \frac{1}{2}X_2 + \frac{1}{2}X_3\right)}{\sqrt{\operatorname{var}(X_1)}\sqrt{\operatorname{var}\left(\frac{1}{2}X_2 + \frac{1}{2}X_3\right)}}$$

$$\operatorname{var}(x_1) = 4$$

$$\operatorname{var}\left(\frac{1}{2}X_2 + \frac{1}{2}X_3\right) = \operatorname{var}\left(\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} X_2 \\ X_3 \end{bmatrix}\right),$$

ingat $\operatorname{var}(\underset{\sim}{a}'\underset{\sim}{x}) = a'\underset{\sim}{\Sigma}a$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 9 & -3 \\ -3 & 25 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = 7$$

$$\operatorname{cov}\left(X_1, \frac{1}{2} X_2 + \frac{1}{2} X_3 \right)$$

$$= \frac{1}{2}\operatorname{cov}(X_1 . X_2) + \frac{1}{2}\operatorname{cov}(X_1, X_3)$$

$$= \frac{1}{2}1 + \frac{1}{2}2 = \frac{3}{2}$$

$$Corr\left(X_1, \frac{1}{2} X_2 + \frac{1}{2} X_3 \right)$$

$$= \frac{\frac{3}{2}}{\sqrt{4\sqrt{7}}}$$

$$= \frac{3}{4\sqrt{7}}$$

LATIHAN

Untuk memperdalam pemahaman Anda mengenai materi di atas, kerjakanlah latihan berikut!

1) Buktikan teorema 2.1!

   $\underset{2x1}{\underset{\sim}{X}}$ mempunyai $E(\underset{\sim}{X}) = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \operatorname{cov}(\underset{\sim}{X}) = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$.

2) $\underset{\sim}{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \underset{\sim}{B} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

   Hitung $E(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B})$ dan $kov(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B})$

3) $\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X}^{(1)} \\ \underset{\sim}{X}^{(2)} \end{bmatrix}, \underset{\sim}{X}^{(1)} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \underset{\sim}{X}^{(2)} = \begin{bmatrix} X_3 \\ X_4 \\ X_5 \end{bmatrix}$

$\underset{\sim}{X}$ mempunyai mean $\underset{\sim}{\mu}$ kovarian matriks $\underset{\sim}{\Sigma}$.

Tulis partisi dari $\underset{\sim}{\mu}$ dan $\underset{\sim}{\Sigma}$!

Petunjuk Jawaban Latihan

1) $\underset{\sim}{X} + \underset{\sim}{Y}$ mempunyai elemen ke (i, j), yaitu $X_i + Y_{ij}$

   $E\left(X_{ij} + Y_{ij}\right) = E\left(X_{ij}\right) + \left(Y_{ij}\right)$

   Kuantitas terakhir ini merupakan elemen ke (i, j)

   dari $E\left(\underset{\sim}{X}\right) + E\left(\underset{\sim}{Y}\right)$.

   Untuk bagian kedua perhatikan elemen ke (i, j) dari $\underset{\sim}{A}X\underset{\sim}{B}$.

2) $E\left(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}\right) = \underset{\sim}{A}E\left(\underset{\sim}{X}\right)\underset{\sim}{B}$

   $$Kov\left(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}\right) = \underset{\sim}{A}\,Kov\left(\underset{\sim}{X}\underset{\sim}{B}\right)\underset{\sim}{A}'$$

   $$= \underset{\sim}{A}\underset{\sim}{B}'\,Kov\left(\underset{\sim}{X}\right)\underset{\sim}{B}\underset{\sim}{A}'$$

3) $$\underset{\sim}{\mu} = \begin{bmatrix} \mu^{(1)} \\ \cdots \\ \mu^{(2)} \end{bmatrix}$$

   $$\underset{\sim}{\mu}^{(1)} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad \underset{\sim}{\mu}^{(2)} = \begin{bmatrix} \mu_3 \\ \mu_4 \\ \mu_5 \end{bmatrix}$$

   $$\underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{bmatrix} = \left[\begin{array}{cc|ccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \hline \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} & \sigma_{45} \\ \sigma_{15} & \sigma_{25} & \sigma_{35} & \sigma_{45} & \sigma_{55} \end{array}\right]$$

   $$\underset{\sim}{\Sigma}_{11} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

   $$\underset{\sim}{\Sigma}_{12} = \underset{\sim}{\Sigma}_{21} = \begin{bmatrix} \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{23} & \sigma_{24} & \sigma_{25} \end{bmatrix} \underset{\sim}{\Sigma}_{22} = \begin{bmatrix} \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{34} & \sigma_{44} & \sigma_{35} \\ \sigma_{35} & \sigma_{45} & \sigma_{55} \end{bmatrix}$$

RANGKUMAN

1. $E\left(\underset{\sim}{X}+\underset{\sim}{Y}\right)=E\left(\underset{\sim}{X}\right)+E\left(\underset{\sim}{Y}\right)$

   $E\left(\underset{\sim}{A}\underset{\sim}{X}\underset{\sim}{B}\right)=\underset{\sim}{A}E\left(\underset{\sim}{X}\right)\underset{\sim}{B}$

   $\underset{px1}{\underset{\sim}{X}}$ mempunyai mean dan kovariansi yang dapat di tulis sebagai

   $E\left(\underset{\sim}{X}\right)=\underset{\sim}{\mu}$

   $\underset{\sim}{\Sigma}=E\left(\underset{\sim}{X}-\underset{\sim}{\mu}\right)\left(\underset{\sim}{X}-\underset{\sim}{\mu}\right)'$

$$\underset{\sim}{\rho}=\begin{bmatrix} 1 & \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \cdots & \dfrac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \dfrac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \vdots & \ddots & \vdots\in \\ \dfrac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \dfrac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} & \cdots & 1 \end{bmatrix}$$

$$\underset{\sim}{V}^{\frac{1}{2}}=\begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{pp}} \end{bmatrix}$$

$$\underset{\sim}{V}^{\frac{1}{2}}\underset{\sim}{\rho}\underset{\sim}{V}^{\frac{1}{2}}=\underset{\sim}{\Sigma}$$

$$\underset{\sim}{\rho}=\left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1}\underset{\sim}{\Sigma}\left(\underset{\sim}{V}^{\frac{1}{2}}\right)^{-1}$$

2. $$\underset{\sim}{X}=\begin{bmatrix} x_1 \\ \vdots \\ x_q \\ \cdots \\ x_{q11} \\ \vdots \\ x_p \end{bmatrix}=\begin{bmatrix} x^{(1)} \\ \cdots \\ x^{(2)} \end{bmatrix};\underset{\sim}{\mu}=\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \\ \cdots \\ \mu_{q+1} \\ \vdots \\ \mu_p \end{bmatrix}=\begin{bmatrix} \mu^{(1)} \\ \cdots \\ \mu^{(2)} \end{bmatrix}$$

dan $\underset{\sim}{\Sigma} = \begin{bmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{bmatrix}$

$$\underset{\sim}{\Sigma}_{11} = kov\left(\underset{\sim}{X}^{(1)}\right)$$

$$\underset{\sim}{\Sigma}_{12} = kov\left(\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}\right)$$

$$\underset{\sim}{\Sigma}_{22} = kov\left(\underset{\sim}{X}^{(2)}\right)$$

$\underset{px1}{\underset{\sim}{X}}$ mempunyai $E\left(\underset{\sim}{X}\right) = \underset{px1}{\underset{\sim}{\mu}}$

$$kov\left(\underset{\sim}{X}\right) = \underset{pxp}{\underset{\sim}{\Sigma}}$$

Untuk vektor konstan $\underset{px1}{\underset{\sim}{c}}$

$$E\left(\underset{\sim}{c}'\underset{\sim}{X}\right) = \underset{\sim}{c}'\underset{\sim}{\mu}$$

$$kov\left(\underset{\sim}{c}'\underset{\sim}{X}\right) = \underset{\sim}{c}'\underset{\sim}{\Sigma}c$$

Untuk vektor konstan $\underset{qxp}{\underset{\sim}{c}}$

$$E\left(\underset{\sim}{c}\underset{\sim}{X}\right) = \underset{\sim}{c}\underset{\sim}{\mu}$$

$$kov\left(\underset{\sim}{c}\underset{\sim}{X}\right) = \underset{\sim}{c}\underset{\sim}{\Sigma}\underset{\sim}{c}'$$

TES FORMATIF 2

Pilihlah satu jawaban yang paling tepat!

1) Diberikan vektor random $\underset{\sim}{X}' = \left[X_1, X_2, \ldots, X_5\right]$

$\underset{\sim}{\mu}_x = \begin{bmatrix} 2 & 4 & -1 & 3 & 0 \end{bmatrix}$

$$\underset{\sim}{\Sigma}_x = \begin{bmatrix} 4 & -1 & \frac{1}{2} & -\frac{1}{2} & 0 \\ -1 & 3 & 1 & -1 & 0 \\ \frac{1}{2} & 1 & 6 & 1 & -1 \\ -\frac{1}{2} & -1 & 1 & 4 & 0 \\ 0 & 0 & -1 & 0 & 2 \end{bmatrix}$$

Partisi $\underset{\sim}{X}$ adalah

$$\underset{\sim}{X} = \begin{bmatrix} X_1 \\ X_2 \\ \cdots \\ X_3 \\ X_4 \\ X_5 \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ \cdots \\ X^{(2)} \end{bmatrix}$$

$$\underset{\sim}{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \underset{\sim}{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

Perhatikan kombinasi linear $\underset{\sim}{A}\underset{\sim}{X}^{(1)}$ dan $\underset{\sim}{B}\underset{\sim}{X}^{(2)}$

1) $E\left(\underset{\sim}{X}^{(1)}\right) =$

A. $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$

B. $\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$

C. $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$

D. $\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}$

2) $E\left(\underset{\sim}{A}\underset{\sim}{X}^{(1)}\right) =$

A. $\begin{bmatrix} 6 \\ 6 \end{bmatrix}$

B. $\begin{bmatrix} 6 \\ -2 \end{bmatrix}$

C. $\begin{bmatrix} -2 \\ 6 \end{bmatrix}$

D. $\begin{bmatrix} -2 \\ -2 \end{bmatrix}$

3) $Kov\left(\underset{\sim}{X}^{(1)}\right) =$

A. $\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$

B. $\begin{bmatrix} 4 & -1 \\ -1 & 3 \end{bmatrix}$

C. $\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

D. Tidak ada yang benar

4) $Kov\left(\underset{\sim}{A}\underset{\sim}{X}^{(1)}\right) =$

A. $\begin{bmatrix} 1 & -5 \\ -9 & 1 \end{bmatrix}$

B. $\begin{bmatrix} -1 & -5 \\ -9 & 1 \end{bmatrix}$

C. $\begin{bmatrix} 5 & 1 \\ 1 & 9 \end{bmatrix}$

D. $\begin{bmatrix} 9 & 1 \\ 1 & 5 \end{bmatrix}$

5) $Kov\left(\underset{\sim}{B}\underset{\sim}{X}^{(2)}\right) =$

A. $\begin{bmatrix} 2 & 9 \\ 5 & 4 \end{bmatrix}$

B. $\begin{bmatrix} 4 & 9 \\ 2 & 5 \end{bmatrix}$

C. $\begin{bmatrix} 8 & 11 \\ -11 & 4 \end{bmatrix}$

D. Tidak ada yang benar

6) $kov\left(\underset{\sim}{X}^{(1)}, \underset{\sim}{X}^{(2)}\right) = \ldots.$

A. $\begin{bmatrix} \frac{1}{2} & 1 \\ -\frac{1}{2} & -1 \\ 0 & 0 \end{bmatrix}$

B. $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & -1 & 0 \end{bmatrix}$

C. $\begin{bmatrix} \frac{1}{2} & 0 \\ -1 & 0 \end{bmatrix}$

D. $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 1 & -1 & 0 \\ 6 & 1 & -1 \end{bmatrix}$

7) $Kov\left(\underset{\sim}{A}\underset{\sim}{X}^{(1)}, \underset{\sim}{B}\underset{\sim}{X}^{(2)}\right) = \ldots.$

A. $\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix}$

B. $\begin{bmatrix} 1 & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$

C. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

D. Tidak ada yang benar

Untuk soal no. 8, 9, 10 di ambil partisi sebagai berikut

$$\underset{\sim}{X} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \text{ dengan } X^{(1)} = \begin{bmatrix} X_1 \\ X_3 \end{bmatrix} \text{ dan } X^{(2)} = \begin{bmatrix} X_3 \\ X_4 \\ X_5 \end{bmatrix}$$

8) $E\left(X^{(2)}\right) = ...$

A. $\begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix}$

B. $\begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$

C. $\begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}$

D. Tidak ada yang benar

9) $Kov\left(X^{(2)}\right) = ...$

A. $\begin{bmatrix} 3 & -1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

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

C. $\begin{bmatrix} 3 & -1 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

D. Tidak ada yang benar

10) Bila $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ dan $B = \begin{bmatrix} 2 & 1 & 3 \end{bmatrix}$ maka $Kov\left(AX^{(1)}, BX^{(2)}\right) = ...$

A. $\begin{bmatrix} 3/2 \\ -5/2 \end{bmatrix}$

B. $\begin{bmatrix} 5/2 \\ -3/2 \end{bmatrix}$

C. $\begin{bmatrix} 5/2 \\ 3/2 \end{bmatrix}$

D. Tidak ada yang benar

Cocokkanlah jawaban Anda dengan Kunci Jawaban Tes Formatif 2 yang terdapat di bagian akhir modul ini. Hitunglah jawaban yang benar. Kemudian, gunakan rumus berikut untuk mengetahui tingkat penguasaan Anda terhadap materi Kegiatan Belajar 2.

$$\text{Tingkat penguasaan} = \frac{\text{Jumlah Jawaban yang Benar}}{\text{Jumlah Soal}} \times 100\%$$

Arti tingkat penguasaan: 90 - 100% = baik sekali
80 - 89% = baik
70 - 79% = cukup
$< 70\%$ = kurang

Apabila mencapai tingkat penguasaan 80% atau lebih, Anda dapat meneruskan dengan modul selanjutnya. **Bagus!** Jika masih di bawah 80%, Anda harus mengulangi materi Kegiatan Belajar 2, terutama bagian yang belum dikuasai.

Kunci Jawaban Tes Formatif

Tes Formatif 1

1) D
2) B
3) C
4) A
5) D
6) A
7) B
8) C
9) D
10) C

Tes Formatif 2

1) A
2) C
3) B
4) D
5) D
6) B
7) C
8) B
9) C
10) A

Daftar Pustaka

Johnson, R. A., Wichern, D. W. (1982). *Applied multivariate statistical analysis.* Prentice Hall Inc.

Rencher, A. C. & Christensen, W. F. (2012). *Methods of multivariate analysis*. Third Edition. New Jersey: John Wiley & Sons, Inc.

Tabachnick, B. G., and Fidell, L. S. (2007). *Using multivariate statistics* (5^{th} ed.). Boston: Pearson Education, Inc.