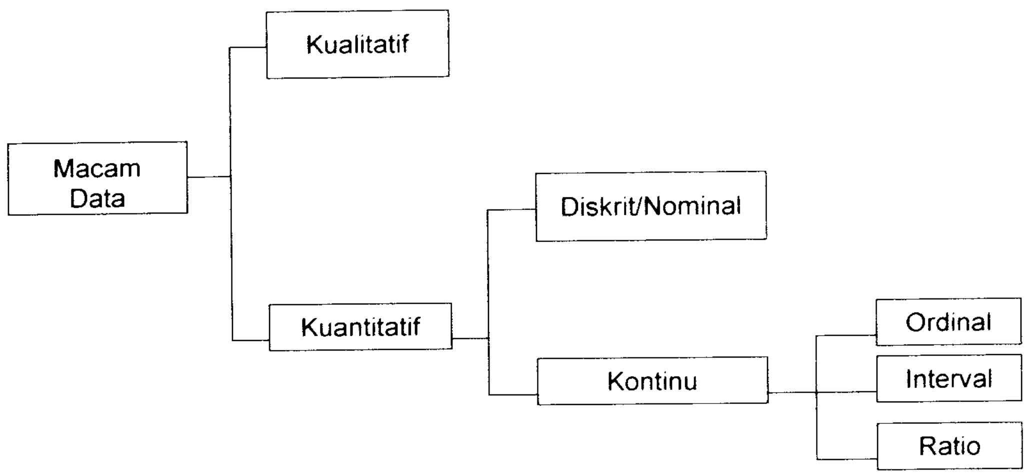
BAB II

TINJAUAN PUSTAKA

2.1 Macam-macam Data dan Skala Pengukuran dalam Penelitian

Berikut ini akan dijelaskan mengenai macam-macam data yang dapat digunakan dalam penelitian.



Gambar 2.1 Macam-Macam Data Penelitian

Berdasarkan gambar 2.1. macam data terbagi menjadi dua. yaitu:

1. Data Kualitatif. adalah data yang dinyatakan dalam bentuk kata. kalimat. dan gambar.
2. Data Kuantitatif. adalah data yang berbentuk angka. atau data kualitatif yang diangkakan (misal: baik sekali = 4. baik = 3. kurang baik = 2. tidak baik = 1).

   Data kuantitatif dibagi menjadi 2. yaitu:

   a. Data diskrit/nominal : data yang hanya dapat digolong-golongkan secara terpisah. secara diskrit atau kategori. Data diperoleh dari hasil menghitung. Contoh: dalam suatu kelompok terdapat 20 pria dan 15 wanita.

b. Data kontinu: data yang bervariasi menurut tingkatan dan diperoleh dari hasil pengukuran. Data kontinu dapat dibagi lagi menjadi 3, yaitu:

- Data Ordinal: data yang berbentuk rangking atau peringkat. Data ini bila dinyatakan dalam skala, maka jarak satu data dengan data yang lain tidak sama (jarak tidak sama). Contoh: juara I. II. III
- Data Interval: data yang jaraknya sama tetapi tidak mempunyai nilai nol (0) mutlak. Contoh: skala termometer, walaupun ada nilai 0 ^{0}C tetapi tetap ada nilainya.
- Data Rasio: data yang jaraknya sama dan mempunyai nilai nol mutlak. Contoh: data tentang berat, berat 0 kg berarti tidak ada bobotnya.

Variabel yang diukur dengan skala nominal dan ordinal, umumnya disebut variabel non-parametrik atau variabel non-metrik. Sementara, untuk variabel yang diukur dengan skala interval dan rasio disebut variabel metrik.

Skala pengukuran merupakan kesepakatan yang digunakan sebagai acuan untuk menentukan panjang pendeknya interval yang ada dalam alat ukur, sehingga alat ukur tersebut bila digunakan dalam pengukuran akan menghasilkan data kuantitatif. Menurut Sugiyono (2008:132), terdapat empat jenis skala sikap yang dapat digunakan dalam pengukuran untuk mendapatkan data interval atau rasio yaitu:

1. Skala *Likert*

Skala *Likert* digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok orang tentang fenomena sosial. Dalam penelitian, fenomena sosial ini telah ditetapkan secara spesifik oleh peneliti. Variabel yang akan diukur dijabarkan menjadi indikator variabel. Jawaban setiap item instrumen yang menggunakan skala *Likert* mempunyai gradasi dari sangat positif sampai sangat negatif, seperti: sangat setuju, setuju, ragu-ragu/netral, tidak setuju, dan sangat tidak setuju, dengan skor masing-masing adalah 5, 4, 3, 2, dan 1. Skala *Likert* dapat dibuat dalam bentuk *checklist* atau pilihan ganda.

2. Skala *Guttman*

Skala pengukuran dengan tipe ini, akan mendapatkan jawaban yang tegas, yaitu "ya-tidak", "benar-salah", "positif-negatif", "setuju-tidak setuju", dan lain-lain. Data yang diperoleh dapat berupa interval atau rasio. Penelitian menggunakan skala ini bila peneliti ingin mendapatkan jawaban yang tegas terhadap suatu permasalahan yang ditanyakan. Skala *Guttman* dapat dibuat dalam bentuk *checklist* atau pilihan ganda.

3. *Semantic Deferensial*

Skala pengukuran yang berbentuk *semantic deferensial* digunakan untuk mengukur sikap/karakteristik tertentu yang dimiliki oleh seseorang. Bentuk dari *semantic deferensial* tidak pilihan ganda atau *checklist* melainkan tersusun dalam satu garis dimana jawaban "sangat positif" terletak di bagian kanan garis dan jawaban "sangat negatif" terletak di bagian kiri garis, atau sebaliknya.

4. *Rating Scale*

Dari ketiga skala pengukuran sebelumnya, data yang diperoleh semuanya adalah data kualitatif yang kemudian dikuantitatifkan. Dalam *rating scale*, data mentah yang diperoleh berupa angka yang kemudian ditafsirkan dalam pengertian kualitatif. Penyusunan instrumen dengan *rating scale* adalah harus dapat mengartikan setiap angka yang diberikan pada alternatif jawaban pada setiap *item* instrumen. Seperti, A memilih jawaban angka 2, tetapi angka 2 yang dimaksud oleh A belum tentu sama maknanya dengan B yang juga memilih jawaban dengan angka 2.

2.2 Populasi dan Sampel

Menurut Sugiyono (2008: 115), populasi adalah wilayah generalisasi yang terdiri atas objek/ subjek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya. Sampel adalah bagian dari jumlah dan karakteristik yang dimiliki oleh populasi tersebut. Sampel yang diambil dari suatu populasi harus betul-betul *representatif* (mewakili). Ukuran sampel yang diharapkan 100% mewakili populasi adalah sama dengan jumlah anggota populasi itu sendiri.

Menurut Israel (2009:1), terdapat tiga kriteria yang biasanya dibutuhkan dalam menentukan ukuran sampel, yaitu:

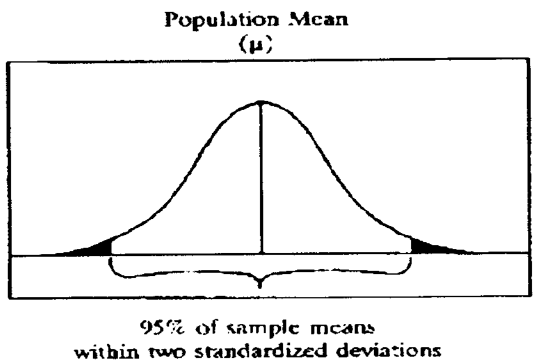
1. Tingkat ketelitian

   Tingkat ketelitian atau lebih sering disebut *sampling error* adalah cakupan dimana nilai yang benar dari populasi itu diperkirakan semestinya. Cakupan ini sering dinyatakan dalam nilai persentase (misalnya ±5%).

2. Tingkat kepercayaan atau resiko

   Tingkat kepercayaan atau resiko didapat dari suatu populasi. Ketika suatu populasi merupakan sampel pengulangan, nilai rata-rata dari atribut yang diperoleh dari sampel itu sama dengan nilai populasi yang sebenarnya. Nilai-

nilai yang diperoleh dari sampel ini berdistribusi normal, dengan beberapa sampel yang memiliki nilai tinggi memperoleh nilai yang lebih rendah dibandingkan dengan nilai populasi yang sebenarnya. Dalam distribusi normal, perkiraan 95% dari nilai sampel adalah di dalam dua simpangan baku dari nilai populasi yang sebenarnya (misalnya *mean*).



Gambar 2.2 Distribusi Normal

3. Derajat dari Variabilitas

Derajat dari variabilitas dalam atribut-atribut yang terukur mengacu pada atribut-atribut di dalam populasi. Semakin banyak variabel (heterogen) suatu populasi, semakin besar ukuran sampel yang diperlukan. Sebaliknya semakin sedikit variabel (lebih homogen) suatu populasi, semakin kecil ukuran sampel yang diperlukan.

Terdapat beberapa pendekatan yang dapat dilakukan dalam menentukan ukuran sampel, antara lain menggunakan sensus untuk ukuran populasi kecil, mengikuti ukuran sampel dari literatur yang sudah ada dan memiliki kesamaan, menggunakan tabel-tabel yang tersedia, dan menggunakan rumus-rumus untuk menghitung ukuran sampel. Berikut ini penjelasan dari masing-masing pendekatan:

1. Menggunakan sensus untuk ukuran populasi kecil.

Pendekatan ini menggunakan seluruh populasi sebagai sampel. Berdasarkan alasan pertimbangan biaya, untuk populasi yang berukuran besar penggunaan sensus dihindari. Sensus sangat baik digunakan untuk ukuran populasi yang kecil (kurang dari 200). Sensus mengeliminasi *sampling error*/galat dan menampilkan data dari semua individu di dalam populasi.

2. Mengikuti ukuran sampel dari literatur yang sudah ada dan memiliki kesamaan.

Pendekatan ini menggunakan ukuran sampel dari literatur yang sudah ada dan memiliki kesamaan dengan kasus yang akan diteliti. Tanpa meninjau ulang prosedur-prosedur pekerjaan yang sudah ada, orang yang menggunakan pendekatan ini menanggung resiko tentang pengulangan kesalahan, dimana membuat penentuan ukuran sampel untuk kasus yang lain.

3. Menggunakan tabel-tabel yang tersedia.

Pendekatan ini digunakan untuk menentukan ukuran sampel dengan menggunakan tabel-tabel yang menyediakan ukuran sampel berdasarkan ukuran-ukuran populasinya.

Tabel 2.1 Ukuran Sampel untuk Tingkat Ketelitian ±3%, ±5%, ±7%, dan ±10% dimana Tingkat Kepercayaan 95% dan P = 0.05

Ukuran Populasi	Ukuran Sampel (n) dengan ketelitian (e)			
	±3%	±5%	±7%	±10%
500	a	222	145	83
600	a	240	152	86
700	a	255	158	88
800	a	267	163	89
900	a	277	166	90
1000	a	286	169	91
2000	714	333	185	95
3000	811	353	191	97
4000	870	364	194	98
5000	909	370	196	98
6000	938	375	197	98
7000	959	378	198	99
8000	976	381	199	99
9000	989	383	200	99
10000	1000	385	200	99
15000	1034	390	201	99
20000	1053	392	202	100
25000	1064	394	202	100
50000	1087	397	203	100
100000	1099	398	204	100
>100000	1111	400	204	100
a=asumsi populasi normal adalah lemah. sehingga seluruh populasi menjadi sampel.				

(Sumber: Israel. Glenn D. Tahun 2009)

Tabel 2.2 Ukuran Sampel untuk Tingkat Ketelitian ±5%, ±7%, dan ±10% dimana Tingkat Kepercayaan 95% dan P = 0.05

Ukuran Populasi	Ukuran Sampel (n) dengan kelitian (e)		
	±5%	±7%	±10%
100	81	67	51
125	96	78	56
150	110	86	61
175	122	94	64
200	134	101	67
225	144	107	70

Lanjutan

Ukuran Populasi	Ukuran Sampel (n) dengan kelitian (e)		
	±5%	±7%	±10%
250	154	112	72
275	163	117	74
300	172	121	76
325	180	125	77
350	187	129	78
375	194	132	80
400	201	135	82
425	207	138	82
450	212	140	82

(Sumber: Israel. Glenn D. Tahun 2009)

Tabel 2.1 dan Tabel 2.2 menunjukkan ukuran sampel dari kombinasi-kombinasi ketelitian. tingkat kepercayaan. dan variabilitas.

4. Menggunakan rumus-rumus untuk menghitung ukuran sampel.

Meskipun Tabel 2.1 dan Tabel 2.2 dapat menyediakan dan menjadi panduan yang bermanfaat dalam menentukan ukuran sampel. diperlukan pula perhitungan untuk menghitung ukuran sampel yang diperlukan untuk kombinasi berbeda dari tingkat ketelitian. keyakinan. dan variabilitas. Beberapa rumus yang dapat dipergunakan. antara lain:

a. Rumus Ukuran Sampel untuk Proporsi (Populasi diketahui)

$$n = \frac{N}{1 + N(e)^2} \qquad (2.1)$$

Rumus ini digunakan untuk menghitung ukuran-ukuran sampel pada Tabel 2.1 dan Tabel 2.2. N adalah ukuran populasi. n adalah ukuran sampel. dan e adalah standar *error* yang dipergunakan.

b. Rumus Ukuran Sampel untuk Proporsi (Populasi tidak diketahui)

$$n = Z^2 \frac{p.q}{(e)^2} \quad (2.2)$$

Rumus ini dikembangkan oleh Chochran untuk menghitung ukuran sampel dengan populasi-populasi berukuran besar atau tidak diketahui. sehingga dapat menghasilkan sampel yang dapat mewakili populasi. n adalah ukuran sampel. Z adalah standar *score* untuk α yang dipilih. e adalah standar error yang dipergunakan. p = proporsi populasi yang diteliti (jika tidak dapat memperkirakan proporsi populasi maka diasumsikan $p = 0.5$). dan $q = 1-p$.

c. Koreksi Populasi Terbatas untuk Proporsi

$$n = \frac{n_0}{1 + \frac{(n_0 - 1)}{N}} \quad (2.3)$$

Rumus ini dapat digunakan jika ukuran populasi kecil. sehingga ukuran sampel dapat dikurangi sedikit. karena ukuran sampel yang diberikan menyediakan informasi yang lebih sebanding untuk populasi yang kecil dibandingkan untuk populasi yang besar. Ukuran sampel dapat disesuaikan dengan rumus ini.

d. Rumus Ukuran Sampel untuk *Mean*

Rumus ukuran sampel untuk *mean* serupa dengan proporsi. kecuali ukuran dari variabilitas. adalah:

$$n = \frac{Z^2\sigma^2}{e^2} \quad (2.4)$$

Dimana n adalah ukuran sampel. Z adalah standar *score* untuk α yang dipilih. e adalah standar *error* yang dipergunakan. dan σ^2 adalah varians atribut dalam populasi.

Berdasarkan beberapa pendekatan dalam menentukan ukuran sampel di atas, dalam penenlitian ini penentuan ukuran sampel menggunakan rumus pada persamaan 2.1.

2.3 Uji Validitas dan Reliabilitas

2.3.1 Uji Validitas

Uji validitas digunakan untuk mengukur sah atau valid tidaknya suatu kuesioner. Suatu kuesioner dikatakan valid jika pertanyaan pada kuesioner mampu untuk mengungkapkan sesuatu yang akan diukur oleh kuesioner tersebut (Ghozali.2002:135).

Ada beberapa cara untuk menguji validitas. dengan menggunakan bantuan software SPSS 15 atau cara manual dengan mengkorelasikan setiap skor butir pernyataan dengan skor total variabel. Untuk mencari nilai korelasi tersebut digunakan rumus korelasi *Pearson Product Moment* sebagai berikut:

$$r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \qquad (2.5)$$

Keterangan: r = nilai korelasi

X = skor pernyataan

Y = skor responden

2.3.2 Uji Reliabilitas

Uji reliabilitas dilakukan untuk mengetahui sejauh mana hasil alat ukur tersebut dapat diandalkan dari kesalahan pengukuran. Pengujian reliabilitas pada penelitian ini menggunakan metode *Alpha Cronbach*. Metode perhitungan reliabilitas ini merupakan metode yang dikembangkan oleh *Cronbach*. Koefisien *Alpha Cronbach* (α) dapat dihitung dengan menggunakan rumus sebagai berikut:

$$\alpha = \frac{N}{N-1}\left(\frac{S^2 - \sum S_i^2}{S^2}\right) \qquad (2.6)$$

Dimana:

α = koefisien reliabilitas *Alpha Cronbach*

N = banyaknya butir pertanyaan

S^2 = varians skor keseluruhan

S_i^2 = varians masing-masing item

Untuk menentukan keeratan hubungan digunakan kriteria Guilford :

1. $\alpha < 0.2$: Hubungan sangat kecil
2. $0.2 \leq \alpha < 0.4$: Hubungan sangat kecil (tidak erat)
3. $0.4 \leq \alpha < 0.7$: Hubungan cukup erat
4. $0.7 \leq \alpha < 0.9$: Hubungan yang erat (reliabel)
5. $0.9 \leq \alpha < 1.0$: Hubungan yang sangat erat (sangat reliabel)
6. $\alpha = 1.00$: Hubungan yang sempurna

Jika hasil uji reliabilitas kuesioner tersebut belum reliabel, maka perlu dilakukan perbaikan selanjutnya dengan penyebaran kuesioner kembali. Akan tetapi jika data sudah reliabel, maka dapat dilanjutkan ke pengolahan data.

2.4 ALJABAR MATRIKS

2.4.1 Vektor

Susunan bilangan real $(x_1, x_2, \dots, x_n)$ yang disusun dalam bentuk satu baris atau satu kolom disebut dengan vektor. Vektor ditulis dengan huruf kecil dan tebal. Vektor ini ditulis sebagai $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ atau $\boldsymbol{x}^t = [x_1, x_2, \dots, x_n]$.

Panjang dari vektor $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ sering dinamakan norm $\boldsymbol{x}$ dan dinyatakan dengan $\|\boldsymbol{x}\|$, besarnya dapat ditentukan dengan rumus:

$$L_x = \|\boldsymbol{x}\| = \sqrt{{x_1}^2 + {x_2}^2 + \cdots + {x_n}^2} \tag{2.7}$$

Teorema 1

Jika **u**, **v**, dan **w** adalah vektor-vektor di ruang-2 atau ruang-3 dan k serta l adalah skalar, maka berlaku hubungan berikut.

1. **u** + **v** = **v** + **u**
2. (**u** + **v**) + **w** = **u** + (**v** + **w**)
3. **u** + 0 = 0 + **u** = **u**
4. **u** + (-**u**) = 0
5. k (l**u**) = (kl) **u**
6. k (**u** + **v**) = k**u** + k**v**
7. (k + l) **u** = k**u** + l**u**

8. $1\mathbf{u} = \mathbf{u}$

2.4.2 Matriks

Menurut Anton (1997:22), sebuah matriks adalah susunan segiempat siku-siku dari bilangan-bilangan. Bilangan-bilangan dalam susunan tersebut dinamakan entri dalam matriks. Matriks dinotasikan dengan hutuf kapital tebal. Matriks **A** berukuran $m \times n$ dapat ditulis sebagai berikut:

$$\boldsymbol{A}_{m\times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

atau dapat ditulis pula dalam bentuk sederhana $\boldsymbol{A}_{m\times n} = \{a_{ij}\}, i = 1, 2, \dots, m \,; j = 1, 2, \dots, n$ dimana indeks i menyatakan baris dan j menyatakan kolom. Matriks $\boldsymbol{A}_{m\times 1}$ dinyatakan sebagai vektor kolom dan matriks $\boldsymbol{A}_{1\times n}$ dinyatakan sebagai vektor baris. Jika matriks **A** memiliki jumlah baris dan jumlah kolom yang sama, maka matriks tersebut dikatakan matriks persegi.

Definisi 1

Jika $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$ dan $\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$ adalah dua matriks yang ukurannya sama, maka berlaku

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

Definisi 2

Jika $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$ adalah suatu matriks dan c adalah suatu skalar, maka hasil kalinya adalah

$$c\mathbf{A} = \mathrm{c} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Definisi 3

Jika **A** adalah matriks berukuran $m \times r$ dan **B** adalah matriks berukuran $r \times n$, maka hasil kali **AB** adalah matriks berukuran $m \times n$.

Dengan kata lain, pengertian dari definisi 3 adalah perkalian matriks mengharuskan bahwa banyak kolom dari **A** harus sama dengan banyak baris pada **B** supaya membentuk hasil kali **AB**.

Definisi 4

Jika **A** adalah matriks $m \times n$ dengan sembarang elemen a_{ij}, untuk setiap i = 1, 2, …, m dan j = 1, 2, …, n maka transpos **A** dinotasikan dengan $\boldsymbol{A^t}$. $\boldsymbol{A^t}$ merupakan matriks berukuran $n \times m$ dengan elemen a_{ji} untuk setiap j = 1, 2, …, n dan i = 1, 2, …, m. Jadi, transpos dari matriks **A** diperoleh dengan menukar baris dengan kolom dan kolom dengan baris dari matriks **A**.

Teorema 2

Jika ukuran matriks seperti operasi yang diberikan dapat dilakukan, maka berlaku:

1. $(A^t)^t = A$
2. $(A + B)^t = A^t + B^t$
3. $(kA)^t = k\,A^t$, dengan k merupakan skalar
4. $(AB)^t = B^t\,A^t$

Definisi 5

Suatu matriks berukuran $m \times m$ (matriks persegi) dikatakan simetri, jika **A**=**A**^{t} atau $a_{ij} = a_{ji}$ dengan i = 1, 2, …, m dan j = 1, 2, …, m.

Definisi 6

Misalkan **A** dan **B** adalah dua buah matriks persegi sedemikian sehingga berlaku **AB** = **BA** = **I**. Matriks **B** disebut invers dari matriks **A** yang dinotasikan **A**$^{-1}$.

2.4.3 Nilai Eigen dan Vektor Eigen

Definisi 7

Misalkan **A** adalah matriks persegi berukuran $m \times m$ dan **I** adalah matriks identitas berukuran $m \times m$. Skalar $\lambda_1, \lambda_2, \ldots, \lambda_m$ yang memenuhi persamaan $|A - \lambda I| = 0$ disebut nilai eigen (akar karakteristik) dari matriks **A**. Persamaan $|A - \lambda I| = 0$ disebut persamaan karakteristik.

Definisi 8

Suatu matriks persegi **A** berukuran $m \times m$ dan λ adalah nilai eigen dari matriks **A**. Jika terdapat suatu vektor **x** tak nol sedemikian sehingga **Ax** = λ**x**, maka **x** disebut vektor eigen (vektor karakteristik) dari matriks **A** yang bersesuaian dengan nilai eigen λ.

2.4.4 Vektor Random dan Matriks Random

Vektor random adalah vektor yang elemen-elemennya merupakan variabel random. Matriks random adalah matriks yang elemen-elemennya merupakan variabel random. Nilai ekspektasi dari matriks random merupakan nilai ekspektasi dari setiap elemen-elemennya.

Jika $\mathbf{X} = \{X_{ij}\}$ merupakan matriks berukuran $p \times k$ adalah matriks random. maka nilai ekspektasi dari **X** dinotasikan dengan E(X) merupakan matriks random berukuran $p \times k$

$$E(X) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1k}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2k}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{p1}) & E(X_{p2}) & \cdots & E(X_{pk}) \end{bmatrix}$$

Jika matriks random **X** dan **Y** merupakan matriks random yang berukuran sama dan merupakan matriks konstan. maka berlaku

$$E(\boldsymbol{X} + \boldsymbol{Y}) = E(\boldsymbol{X}) + E(\boldsymbol{Y})$$

$$E(\boldsymbol{AXB}) = \boldsymbol{A}E(\boldsymbol{X})\boldsymbol{B}$$

Misalkan terdapat p buah varibel **X** yang diteliti dalam suatu pengambilan sampel berukuran n. matriks random dapat dituliskan sebagai

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

Rata-rata dari matriks di atas dapat dinyatakan dalam vektor random $\bar{\boldsymbol{x}}$ yang diperoleh dengan mencari nilai ekspektasinya

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

Untuk kepentingan pembentukan suatu fungsi diskriminan. diperlukan suatu uji statistik untuk mengetahui apakah terdapat perbedaan nilai rataan diantara kelompok-kelompok yang ada. lebih lengkapnya akan dibahas pada bab selanjutnya.

2.4.5 Matriks Varians Kovarians dan Matriks Korelasi

Dalam suatu populasi terdapat beberapa variabel yang diteliti. dalam kasus univariat dikenal besaran statistik yakni varians yang menyatakan ukuran keseragaman data. untuk kasus multivariat dari beberapa variabel yang terlibat dikenal suatu besaran statistik yang disebut matriks varians-kovarians. Matriks varians-kovarians dari suatu populasi didefinisikan sebagai

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{j1} & \sigma_{j2} & \cdots & \sigma_{jk} \end{pmatrix}$$

dimana σ_{jj} merupakan varians dari populasi ke-j dan σ_{jk} adalah kovarians dari semua pasangan populasi yang terlibat.

Matriks varians-kovarians di atas dapat diperoleh dengan mencari nilai ekspektasi

$$\Sigma = E[(X-\mu)(X-\mu)']$$

$$= E\left(\begin{bmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \vdots \\ X_p-\mu_p \end{bmatrix} [X_1-\mu_1 \quad X_2-\mu_2 \quad \cdots \quad X_p-\mu_p]\right)$$

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

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

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix}$$

Karena untuk setiap $\sigma_{jk} = \sigma_{kj}$ maka matriks varians-kovarians di atas dapat juga ditulis

$$\Sigma = E[(X-\mu)(X-\mu)'] = Cov\,(X) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{pmatrix} \qquad (2.8)$$

Jika X_j dan X_k saling bebas maka berlaku $\sigma_{jk} = Cov\,(X_j.X_k) = 0$ sehingga matriks di atas menjadi

$$\Sigma = E[(X-\mu)(X-\mu)'] = Cov\,(X) = \begin{pmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{pp} \end{pmatrix} \qquad (2.9)$$

Koefisien korelasi populasi antara X_j dan X_k dinotasikan ρ_{jk} dan didefinisikan sebagai $\rho_{jk} = \frac{\sigma_{jk}}{\sqrt{\sigma_{jj}}\sqrt{\sigma_{kk}}}$. Karena $\sigma_{jk} = \sigma_{kj}$, $j,k = 1, 2, \ldots, p$ dengan $j \neq k$, maka matriks koefisien korelasi merupakan matriks simetri. Sehingga dapat didefinisikan sebagai

$$\rho = \begin{pmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{22}}} & \cdots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{pp}}} & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}}\sqrt{\sigma_{pp}}} & \cdots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}}\sqrt{\sigma_{pp}}} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{pmatrix} \quad (2.10)$$

2.5 PENAKSIRAN PARAMETER

Menurut Herrhyanto (2003:83), terdapat beberapa sifat penaksir sebuah parameter populasi, antara lain:

1. Tak Bias

   $\hat{\theta}$ dikatakan penaksir tak bias bagi parameter θ, jika:

$$E(\hat{\theta}) = \theta \quad (2.11)$$

   Sebaliknya $\hat{\theta}$ dikatakan penaksir bias bagi parameter θ, jika:

$$E(\hat{\theta}) \neq \theta$$

   Penaksir bias ini dapat diubah menjadi penaksir tak bias dengan cara ruas kanan dikalikan atau ditambahkan dengan konstanta tertentu agar hasilnya sama dengan θ.

   Berikut ini dijelaskan beberapa penaksir tak bias berhubungan dengan perhitungan statistika multivariat.

Misalkan **X** adalah matriks data berukuran n $\times$ p dari suatu populasi yang berdistribusi normal $N_p(\mu, \Sigma)$ dengan vektor rata-rata sampel $\bar{\mathbf{x}}$, matriks kovarians sampel $\mathbf{S_n}$ dan matriks kovarians sampel gabungan adalah $\mathbf{S_{pooled}}$, sehingga berlaku $E(\bar{x}) = \mu, Cov\ (\bar{x}) = \frac{1}{n}\Sigma$. Untuk matriks kovarians $\mathbf{S_n}$, $E(S_n) = \frac{n-1}{n}\Sigma = \Sigma - \frac{1}{n}\Sigma$, sehingga $E\left(\frac{n-1}{n}S_n\right) = \Sigma$. Jadi $\frac{n-1}{n}S_n$ adalah penaksir tak bias untuk Σ. Selanjutnya akan dibuktikan bahwa:

a. $\bar{\mathbf{x}}$ penaksir tak bias untuk μ.

b. $\frac{n-1}{n}S_n$ penaksir tak bias untuk Σ.

c. Untuk kasus yang terdiri dari beberapa populasi dan memiliki matriks kovarians sama, maka $\mathbf{S_{pooled}}$ adalah penaksir tak bias dari Σ.

Bukti:

a. Akan dibuktikan $\bar{\mathbf{x}}$ penaksir tak bias untuk μ.

Untuk tiap variabel X_i, $i = 1, 2, \ldots, n$ dan $E(X_i) = \mu$.

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu \quad (2.12)$$

Jika terdapat p variabel yang terlibat maka diperoleh vektor rata-ratanya adalah $\bar{x} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{bmatrix}$. Dengan mengambil masing-masing ekspektasinya, maka diperoleh

$$E(\bar{x}) = \begin{bmatrix} E(\bar{X}_1) \\ E(\bar{X}_2) \\ \vdots \\ E(\bar{X}_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu \quad (2.13)$$

Jadi, terbukti bahwa $\bar{\mathbf{x}}$ penaksir tak bias untuk μ.

b. Akan dibuktikan $\frac{n-1}{n}S_n$ penaksir tak bias untuk Σ

Perhatikan bahwa

$$(\bar{x}-\mu)(\bar{x}-\mu)' = \left(\frac{1}{n}\sum_{j=1}^{n}(X_j-\mu)\right)\left(\frac{1}{n}\sum_{l=1}^{n}(X_l-\mu)\right)' \qquad (2.14)$$

$$= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}(X_j-\mu)(X_l-\mu)'$$

$$Cov(\bar{x}) = E(\bar{x}-\mu)(\bar{x}-\mu)' = \frac{1}{n^2}\left(\sum_{j=1}^{n}\sum_{l=1}^{n}E(X_j-\mu)(X_l-\mu)'\right) \qquad (2.15)$$

Untuk $j \neq l$. $E(X_j-\mu)(X_l-\mu)'$adalah nol karena saling bebas, sehingga diperoleh

$$Cov(\bar{x}) = \frac{1}{n^2}\left(\sum_{j=1}^{n}E(X_j-\mu)(X_j-\mu)'\right)$$

Karena $\Sigma = E(X_j-\mu)(X_j-\mu)'$ untuk setiap variabel X_j maka diperoleh

$$Cov(\bar{x}) = \frac{1}{n^2}\left(\sum_{j=1}^{n}E(X_j-\mu)(X_j-\mu)'\right)$$

$$= \frac{1}{n^2}(\Sigma+\Sigma+\cdots+\Sigma)$$

$$= \frac{1}{n^2}(n\Sigma)$$

$$= \left(\frac{1}{n}\right)\Sigma \qquad (2.16)$$

Untuk memperoleh S_n, dengan memperhatikan jumlah kuadrat dari setiap variabel maka

$$\sum_{j=1}^{n}(X_j - \bar{x})(X_j - \bar{x})' = \sum_{j=1}^{n}(X_jX_j' - X_j\bar{x}' - \bar{x}X_j' + \overline{xx}')$$

$$= \sum_{j=1}^{n} X_jX_j' - n\overline{xx}' - n\overline{xx}' + n\overline{xx}'$$

$$= \sum_{j=1}^{n} X_jX_j' - n\overline{xx}' \qquad (2.17)$$

$$E\left(\sum_{j=1}^{n}(X_j - \bar{x})(X_j - \bar{x})'\right) = E\left(\sum_{j=1}^{n} X_jX_j' - n\overline{xx}'\right)$$

$$= \sum_{j=1}^{n} E(X_jX_j') - nE(\overline{xx}')$$

Selanjutnya untuk setiap variabel acak X_j dengan $E(X_j) = \mu$ dan $Cov(X_j) = \Sigma$. maka berdasarkan sifat kovariansi $Cov(X_j) = E(X_jX_j') - \left(E(X_j)\right)^2$ diperoleh

$$E(X_jX_j') = Cov(X_j) + \left(E(X_j)\right)^2 = \Sigma + \mu\mu'$$

$$E(\overline{xx}') = Cov(\bar{x}) + \left(E(\bar{x})\right)^2 = \frac{1}{n}\Sigma + \mu\mu' \qquad (2.18)$$

$$E\left(\sum_{j=1}^{n}(X_j - \bar{x})(X_j - \bar{x})'\right) = E\left(\sum_{j=1}^{n} X_jX_j' - n\overline{xx}'\right)$$

$$= \sum_{j=1}^{n} E(X_jX_j') - nE(\overline{xx}')$$

$$= \sum_{j=1}^{n}(\Sigma + \mu\mu') - n\left(\frac{1}{n}\Sigma + \mu\mu'\right)$$

$$= n\Sigma + n\mu\mu' - \Sigma - n\mu\mu'$$

$$E\left(\sum_{j=1}^{n}(X_j-\bar{x})(X_j-\bar{x})'\right) \qquad = (n-1)\Sigma \qquad (2.19)$$

Karena

$$S_n = \frac{1}{n}\left(\sum_{j=1}^{n}(X_j-\bar{x})(X_j-\bar{x})'\right)$$

Sehingga diperoleh $E(S_n) = \frac{(n-1)}{n}\Sigma$.

c. Akan dibuktikan $\mathbf{S_{pooled}}$ adalah penaksir tak bias dari Σ

Karena diasumsikan untuk populasi yang memiliki matriks kovarians populasi sama, maka $E\left(\frac{(n_j-1)}{n_j}S_j\right) = \Sigma$, $j = 1, 2, \ldots, g$ dengan g banyaknya kelompok yang terlibat. Untuk memperoleh $\mathbf{S_{pooled}}$ dilakukan dengan menjumlahkannya dan dibagi dengan banyaknya kelompok yang terlibat, maka diperoleh

$$S_{pooled} = \frac{\left(\frac{n_1-1}{n_1}\right)S_1 + \left(\frac{n_2-1}{n_2}\right)S_2 + \cdots + \left(\frac{n_g-1}{n_g}\right)S_g}{g}$$

$$E(S_{pooled}) = E\left(\frac{\left(\frac{n_1-1}{n_1}\right)S_1 + \left(\frac{n_2-1}{n_2}\right)S_2 + \cdots + \left(\frac{n_g-1}{n_g}\right)S_g}{g}\right)$$

$$= \frac{E\left(\left(\frac{n_1-1}{n_1}\right)S_1\right) + E\left(\left(\frac{n_2-1}{n_2}\right)S_2\right) + \cdots + E\left(\left(\frac{n_g-1}{n_g}\right)S_g\right)}{g}$$

$$= \frac{\Sigma + \Sigma + \cdots + \Sigma}{g} = \frac{g\,\Sigma}{g} = \Sigma \qquad (2.20)$$

2. Variansi Minimum / Efisien

Untuk menentukan sebuah penaksir bervariansi minimum, terlebih dahulu diperiksa bahwa penaksir tersebut merupakan penaksir yang tak bias.

Sebuah penaksir tak bias akan mencapai variansi minimum diantara semua penaksir tak bias lainnya, apabila variansi dari penaksir itu minimal sama dengan batas bawah *Cramer-Rao*.

Perumusan batas bawah *Cramer-Rao* untuk variansi dari $\hat{\theta}$ adalah:

$$Var(\hat{\theta}) \geq \frac{1}{-n\,.\,E\left[\frac{d}{d\theta}\ln f(x;\theta)\right]^2} = \frac{1}{-n\,.\,E\left[\frac{d^2}{d\theta^2}\ln f(x;\theta)\right]} \qquad (2.21)$$

Penaksir disebut efisien apabila penaksir memenuhi sifat tak bias dan variansi minimum.

3. Konsisten

Jika $\hat{\theta}_n$ adalah penaksir untuk θ yang didasarkan pada sampel acak berukuran n, maka dapat dikatakan $\hat{\theta}_n$ adalah konsisten untuk θ, jika:

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < \varepsilon\right) = 1$$

Untuk membuktikan bahwa sebuah penaksir adalah konsisten bagi parameternya dapat menggunakan pertidaksamaan *Chebyshev's* dengan rumus:

$$P(|X - \mu_x| < k\sigma_{\bar{x}}) \geq 1 - \frac{1}{k^2}$$

Jika $\lim_{n\to\infty} \frac{1}{k^2} = 0$, maka dikatakan penaksir adalah konsisten bagi parameternya.

Penaksir terbaik adalah apabila ketiga sifat penaksir di atas terpenuhi.

2.6 ANALISIS MULTIVARIAT

Data multivariat adalah data yang diperoleh dengan mengukur lebih dari satu variabel kriteria pada setiap individu anggota sampel. Analisis statistika multivariat merupakan metode statistika yang memungkinkan melakukan penelitian terhadap lebih dari dua variabel secara bersamaan. Dengan menggunakan teknik analisis ini, maka dapat menganalisis pengaruh beberapa variabel terhadap variabel–variabel lainnya dalam waktu yang bersamaan. Analisis multivariat digunakan karena pada kenyataannnya masalah yang terjadi tidak dapat diselesaikan dengan hanya menghubungkan dua variabel atau melihat pengaruh satu variabel terhadap variabel lainnya.

Teknik analisis multivariat secara umum dapat diklasifikasikan menjadi dua bagian, yaitu analisis dependensi dan analisis interdependensi. Analisis dependensi berfungsi untuk menerangkan atau memprediksi variabel terikat (Y) dengan menggunakan dua atau lebih variabel bebas (X), sedangkan analisis interdependensi berfungsi untuk memberikan makna terhadap seperangkat variabel atau membuat kelompok-kelompok secara bersama-sama.

Analisis dependensi terdiri dari regresi linear berganda, analisis diskriminan, analisis konjoin, analisis varian multivariat (MANOVA), dan analisis korelasi kanonikal. Analisis independensi terdiri dari analisis faktor, analisis kluster, dan *multidimensional scaling*. Pengklasifikasian lebih sederhana dapat dilihat pada tabel berikut ini.

Tabel 2.3 Pengklasifikasian Analisis Multivariat

<table>
<tr><td rowspan="2"></td><td colspan="2" rowspan="2">Teknik yang digunakan</td><td colspan="2">Jml. Variabel</td><td colspan="2">Skala Pengukuran</td></tr>
<tr><td>Terikat</td><td>Bebas</td><td>Var. Terikat</td><td>Var. Bebas</td></tr>
<tr><td rowspan="9">Analisis Multivariat</td><td rowspan="6">Analisis Dependensi</td><td>Regresi Berganda</td><td>1</td><td>2 atau lebih</td><td>Metrik (Interval atau Rasio)</td><td>Metrik (Interval atau Rasio)</td></tr>
<tr><td>ANOVA</td><td>1</td><td>2 atau lebih</td><td>Metrik (Interval atau Rasio)</td><td>Non-metrik (Nominal atau Ordinal)</td></tr>
<tr><td>Analisis Diskriminan</td><td>1</td><td>2 atau lebih</td><td>Non-metrik (Nominal atau Ordinal)</td><td>Metrik (Interval atau Rasio)</td></tr>
<tr><td>Analisis Konjoin</td><td>1</td><td>2 atau lebih</td><td>Non-metrik (Nominal atau Ordinal)</td><td>Non-metrik (Nominal atau Ordinal)</td></tr>
<tr><td>Korelasi Kanonikal</td><td>2 atau lebih</td><td>1</td><td>Metrik (Interval atau Rasio)</td><td>Metrik (Interval atau Rasio)</td></tr>
<tr><td>MANOVA</td><td>2 atau lebih</td><td>1</td><td>Metrik (Interval atau Rasio)</td><td>Non-metrik (Nominal atau Ordinal)</td></tr>
<tr><td rowspan="3">Analisis Interdependensi</td><td>Analisis Faktor</td><td colspan="4">Metrik (Interval atau Rasio)</td></tr>
<tr><td>Analisis Kluster</td><td colspan="4">Metrik (Interval atau Rasio)</td></tr>
<tr><td>Multidimensional Scaling</td><td colspan="4">Metrik (Interval atau Rasio)</td></tr>
</table>