

## **BAB 2**

### **TINJAUAN PUSTAKA**

#### **2.1 Rumah Sakit**

##### **2.1.1 Pengantar Rumah Sakit**

Menurut Direktorat Pelayanan Medik Departemen Kesehatan RI dalam Rancangan Peraturan Pemerintah tentang Rumah Sakit (2004), Rumah Sakit adalah semua sarana kesehatan yang menyelenggarakan pelayanan rawat inap, rawat jalan, gawat darurat dan tindakan medik yang dilaksanakan selama 24 jam melalui upaya kesehatan perorangan.

Rumah Sakit mempunyai tugas pokok menyelenggarakan upaya kesehatan perorangan yang paripurna (promotif, preventif, kuratif dan rehabilitasi) dengan mengutamakan upaya penyembuhan penyakit dan pemulihan kesehatan tanpa mengabaikan upaya pencegahan dan peningkatan.

Fungsi Rumah Sakit adalah menyelenggarakan pelayanan medis, pelayanan keperawatan, pelayanan penunjang medis, pelayanan administrasi dan manajemen, pendidikan/pelatihan, serta penelitian dan pengembangan. Yang dimaksud dengan pelayanan medis adalah upaya kesehatan perorangan meliputi pelayanan promotif, preventif, kuratif dan rehabilitatif yang diberikan kepada pasien oleh tenaga medis sesuai dengan standar pelayanan medis dengan memanfaatkan sumber daya dan fasilitas secara optimal. Pelayanan keperawatan adalah pelayanan kesehatan yang didasarkan pada ilmu dan kiat keperawatan, yang mencakup biopsikososiospiritual yang komprehensif. Sedangkan pelayanan penunjang medis adalah kegiatan pelayanan kesehatan yang menunjang pelayanan medik yang sesuai dengan standar yang berlaku, meliputi rekam medis yaitu rekaman tentang identitas pasien, pemeriksaan, pengobatan, tindakan dan pelayanan lain kepada pasien, dapur, laundry, sterilisasi, logistic, dan sebagainya. Dan pelayanan administrasi serta manajemen merupakan kegiatan pelayanan administrasi, dan manajemen untuk mendukung penyelenggaraan pelayanan rumah sakit sesuai dengan kebutuhan perkembangan.

### 2.1.2 Klasifikasi dan Kelas Rumah Sakit

Rumah Sakit di Indonesia dapat dikategorikan menurut jenis dan pengelolannya (Departemen Kesehatan RI, 2005). Menurut jenisnya, Rumah Sakit dapat dikelompokkan menjadi :

- Rumah Sakit Umum (RSU). RSU adalah Rumah Sakit yang menyelenggarakan pelayanan kesehatan semua bidang dan jenis penyakit,
- Rumah Sakit Jiwa (RSJ),
- Rumah Sakit Khusus. Rumah Sakit Khusus adalah Rumah Sakit yang menyelenggarakan pelayanan utama pada satu atau dua bidang tertentu, berdasarkan disiplin ilmu, golongan umur, organ, jenis penyakit, atau kekhususan lainnya.

Sedangkan menurut pengelolannya, Rumah Sakit dapat dikelompokkan menjadi :

- Rumah Sakit Vertikal yang dikelola oleh Departemen Kesehatan RI,
- Rumah Sakit Provinsi yang dikelola oleh Pemerintah Daerah Tingkat I,
- Rumah Sakit Kabupaten/Kotamadya yang dikelola oleh Pemerintah Daerah Tingkat II,
- Rumah Sakit ABRI,
- Rumah Sakit Departemen Lain/BUMN,
- Rumah Sakit Swasta.

Menurut Direktorat Pelayanan Medik Departemen Kesehatan RI dalam Pedoman Penyelenggaraan Rumah Sakit (2008), rumah sakit dikelompokkan menjadi berbagai jenis pelayanan sebagai berikut :

- I. Pelayanan Umum**
  - a. Pelayanan Medik Umum,
  - b. Pelayanan Medik Gigi Dasar,
  - c. Pelayanan KIA/KB.
- II. Pelayanan Gawat Darurat**
- III. Pelayanan Spesialis Dasar**
  - a. Pelayanan Penyakit Dalam,

- b. Pelayanan Kesehatan Anak,
- c. Pelayanan Bedah,
- d. Pelayanan Obstetri dan Ginekologi.

**IV. Pelayanan Spesialistik Penunjang**

- a. Pelayanan Anestesiologi,
- b. Pelayanan Radiologi,
- c. Pelayanan Rehabilitasi Medik,
- d. Pelayanan Patologi Klinik,
- e. Pelayanan Patologi Anatomi.

**V. Pelayanan Medik Spesialistik Lain**

- a. Pelayanan Spesialis Telinga Hidung Tenggorokan,
- b. Pelayanan Spesialis Orthopaedi,
- c. Pelayanan Spesialis Kesehatan Jiwa,
- d. Pelayanan Spesialis Penyakit Saraf,
- e. Pelayanan Spesialis Penyakit Mata,
- f. Pelayanan Spesialis Penyakit Kulit dan Kelamin,
- g. Pelayanan Spesialis Jantung,
- h. Pelayanan Spesialis Paru,
- i. Pelayanan Spesialis Urologi,
- j. Pelayanan Spesialis Bedah Syaraf,
- k. Pelayanan Spesialis Lainnya.

**VI. Pelayanan Spesialistik Gigi Mulut**

- a. Pelayanan Orthodonsi,
- b. Pelayanan Prosthodonti,
- c. Pelayanan Konservasi/endodonti.

**VII. Sub Spesialis**

- a. Sub Spesialis Pelayanan Bedah,
- b. Sub Spesialis Pelayanan Penyakit Dalam,
- c. Sub Spesialis Pelayanan Kesehatan Anak,
- d. Sub Spesialis Pelayanan Kebidanan dan Penyakit kandungan,
- e. Sub Spesialis Pelayanan Mata,
- f. Sub Spesialis Pelayanan THT,

- g. Sub Spesialis Pelayanan Kulit dan Kelamin,
- h. Sub Spesialis Pelayanan Syaraf,
- i. Sub Spesialis Pelayanan Jiwa,
- j. Sub Spesialis Pelayanan Orthopedi,
- k. Sub Spesialis Pelayanan Jantung,
- l. Sub Spesialis Pelayanan Paru,
- m. Sub Spesialis Pelayanan Spesialis Gigi dan Mulut,
- n. Sub Spesialis Lainnya.

**VIII. Pelayanan Penunjang Klinik**

- a. Perawatan Intensif,
- b. Pelayanan Darah,
- c. Pelayanan Gizi,
- d. Pelayanan Farmasi,
- e. Pelayanan Sterilisasi dan Instrumen,
- f. Rekam Medik.

**IX. Pelayanan Penunjang Non Klinik**

- a. Laundry/Linen,
- b. Pelayanan Jasa Boga/Dapur,
- c. Pelayanan Teknik dan Pemeliharaan Fasilitas,
- d. Pengelolaan Limbah,
- e. Gudang,
- f. Transportasi (Ambulance),
- g. Komunikasi,
- h. Pemulasaran Jenazah,
- i. Pemadam Kebakaran,
- j. Penampungan Air Bersih.

**X. Pelayanan Administrasi**

- a. Informasi dan penerimaan pasien,
- b. Keuangan,
- c. Personalia,
- d. Keamanan,
- e. Sistem Informasi Rumah Sakit.

Sesuai dengan beban kerja dan fungsinya maka rumah sakit umum dikategorikan menjadi empat kelas :

- Rumah Sakit Kelas A, adalah Rumah Sakit Umum yang memiliki fasilitas dan kemampuan pelayanan umum (I), pelayanan gawat darurat (II), pelayanan spesialis dasar (III), pelayanan spesialistik penunjang (IV) pelayanan medik spesialistik lain (V), pelayanan penunjang klinik (VIII), pelayanan penunjang Non Klinik (IX), pelayanan administrasi (X) dan dapat ditambah dengan pelayanan spesialistik gigi dan mulut (VI), sub spesialis luas (VII),
- Rumah Sakit Kelas B, adalah Rumah Sakit Umum yang memiliki fasilitas dan kemampuan pelayanan umum (I), pelayanan gawat darurat (II), pelayanan spesialis dasar (III), pelayanan spesialistik penunjang (IV), 7 pelayanan medik spesialistik lain (V), pelayanan penunjang klinik (VIII), pelayanan penunjang non klinik (IX), pelayanan administrasi (X) dan dapat ditambah dengan pelayanan spesialistik gigi dan mulut (VI), sub spesialis terbatas (VII), serta memiliki minimal 200 fasilitas tempat tidur,
- Rumah Sakit Kelas C, adalah Rumah Sakit Umum yang memiliki fasilitas dan kemampuan pelayanan umum (I), pelayanan gawat darurat (II), pelayanan spesialis dasar (III), 4 pelayanan spesialistik penunjang (IV), pelayanan penunjang klinik (VIII), pelayanan penunjang non klinik (IX), pelayanan administrasi (X), serta memiliki minimal 100 fasilitas tempat tidur,
- Rumah Sakit Kelas D, adalah Rumah Sakit Umum yang memiliki fasilitas dan kemampuan pelayanan umum (I), pelayanan gawat darurat (II), 2 jenis pelayanan spesialis dasar atau lebih (III), pelayanan penunjang klinik (VIII) kecuali perawatan intensif, pelayanan penunjang non klinik (IX), dan pelayanan administrasi (X), serta memiliki minimal 50 fasilitas tempat tidur.

## 2.2 Indikator

### 2.2.1 Pengertian

Dalam Petunjuk Pelaksanaan Indikator Mutu Pelayanan Rumah Sakit yang dikeluarkan oleh World Health Organization Direktorat Jenderal Pelayanan Medik Departemen Kesehatan RI, 2001, Indikator adalah suatu cara untuk menilai penampilan dari suatu kegiatan dengan menggunakan instrument. Indikator merupakan variable yang digunakan untuk menilai suatu perubahan. Menurut WHO, indicator adalah variable untuk mengukur perubahan. Indikator sering digunakan terutama bila perubahan tersebut tidak dapat diukur.

Indikator yang ideal harus memiliki empat criteria, yaitu :

- Sahih (valid), yaitu benar-benar dapat dipakai untuk mengukur aspek yang akan dinilai,
- Dapat dipercaya (reliable), yaitu mampu menunjukkan hasil yang sama pada saat yang berulang kali, untuk waktu sekarang maupun yang akan datang,
- Sensitif, yaitu cukup peka untuk mengukur, sehingga jumlahnya tidak perlu banyak,
- Spesifik, yaitu memberikan gambaran perubahan ukuran yang jelas, tidak bertumpang tindih.

Indikator-indikator dalam pelayanan rumah sakit ini akan mempunyai manfaat yang sangat banyak bagi pengelola rumah sakit, terutama untuk mengukur kinerja rumah sakit itu sendiri (*self assessment*). Manfaat tersebut antara lain sebagai alat yang digunakan untuk manajemen control dan alat untuk mendukung pengambilan keputusan dalam rangka perencanaan kegiatan untuk masa yang akan datang.

### 2.2.2 Indikator Bagi Rumah Sakit

Menurut Donabedian, pengukuran mutu pelayanan kesehatan dapat diukur dengan menggunakan tiga variable (Direktorat Jenderal Pelayanan Medik Departemen Kesehatan RI, 2005), yaitu :

- Input (struktur), ialah segala sumber daya yang diperlukan untuk melakukan pelayanan kesehatan, seperti tenaga, dana, obat, fasilitas,

peralatan, bahan, teknologi, organisasi, informasi, dan lain-lain. Pelayanan kesehatan yang bermutu memerlukan dukungan input yang bermutu pula. Hubungan struktur dengan mutu pelayanan kesehatan adalah dalam perencanaan dan penggerakan pelaksanaan pelayanan kesehatan.

- Proses, ialah interaksi professional antara pemberi layanan dengan konsumen (pasien/masyarakat). Proses ini merupakan variable penilaian mutu yang paling penting.
- Output (*Outcome*), ialah hasil pelayanan kesehatan, merupakan perubahan yang terjadi pada konsumen (pasien/masyarakat), termasuk kepuasan dari konsumen tersebut.

Rumah sakit tertentu yang berstatus kepemilikannya di bawah Departemen Kesehatan, indicator yang digunakan adalah indicator mutu pelayanan rumah sakit. Proses penyusunannya dilaksanakan pada tahun 1996-1997, dan telah diuji pada 11 rumah sakit yang tersebar di Indonesia. Kegiatan penyusunan diawali dengan mencari indicator mana yang dapat dengan mudah dilaksanakan dan dilakukan inventarisasi data yang tersedia di rumah sakit yang dapat dimanfaatkan untuk diolah menjadi indicator mutu. Indikator untuk mengukur kinerja rumah sakit juga mengadopsi indicator mutu pelayanan rumah sakit. Indikator yang digunakan untuk mengukur kinerja rumah sakit maupun indicator mutu pelayanan rumah sakit tidak semuanya dapat diaplikasikan untuk semua rumah sakit pada umumnya.

### 2.2.3 Indikator Penilaian Pelayanan Rumah Sakit

Direktorat Jenderal Bina Pelayanan Medik dalam Informasi Rumah Sakit seri 1 tentang Pelayanan Rumah Sakit, 2005, untuk menilai tingkat keberhasilan atau memberikan gambaran tentang keadaan pelayanan di rumah sakit biasanya dilihat dari berbagai segi, yaitu :

- Tingkat pemanfaatan sarana pelayanan,
- Mutu pelayanan,
- Tingkat efisiensi pelayanan.

Untuk mengetahui tingkat pemanfaatan, mutu dan efisiensi pelayanan rumah sakit diperlukan beberapa indikator. Selain itu agar informasi yang ada dapat bermakna harus ada nilai parameter yang akan dipakai sebagai nilai banding antara fakta dengan standar yang diinginkan.

Terdapat banyak sekali indikator yang dipakai untuk menilai suatu rumah sakit, yang paling sering dipergunakan diantaranya adalah :

1. Bed Occupancy Rate (BOR)

Yaitu prosentase pemakaian tempat tidur pada satu satuan waktu tertentu. Indikator ini memberikan gambaran tinggi rendahnya tingkat pemanfaatan dari tempat tidur rumah sakit.

$$\text{Rumus : } \frac{\text{Jumlah hari perawatan rumah sakit}}{\text{Jumlah TT} \times \text{Jumlah hari dalam satu satuan waktu}} \times 100 \% \quad (2.1)$$

Nilai parameter dari BOR ini idealnya antara 60-85%

2. Average Length of Stay (Av LOS)

Yaitu rata-rata lama rawatan seorang pasien. Indikator ini disamping memberikan gambaran tingkat efisiensi juga dapat memberikan gambaran mutu pelayanan, apabila diterapkan pada diagnosis tertentu yang dijadikan tracer (yang perlu pengamatan lebih lanjut).

$$\text{Rumus : } \frac{\text{Jumlah hari perawatan pasien keluar}}{\text{Jumlah pasien keluar (hidup + mati)}} \quad (2.2)$$

Secara umum LOS yang ideal antara 6 - 9 hari.

3. Bed Turn Over (BTO)

Yaitu frekuensi pemakaian tempat tidur, berapa kali dalam satu satuan waktu tertentu (biasanya 1 tahun) tempat tidur rumah sakit dipakai. Indikator ini memberikan gambaran tingkat efisiensi dari pada pemakaian tempat tidur.

$$\text{Rumus : } \frac{\text{Jumlah pasien keluar (hidup + mati)}}{\text{Jumlah Tempat Tidur}} \quad (2.3)$$

Idealnya selama satu tahun, 1 tempat tidur rata-rata dipakai 40-50 kali.

## 4. Turn Over Interval (TOI)

Yaitu rata-rata hari, tempat tidur tidak ditempati dari saat terisi ke saat terisi berikutnya. Indikator ini juga memberikan gambaran tingkat efisiensi dari pada penggunaan tempat tidur.

$$\text{Rumus : } \frac{(\text{Jumlah TT x hari}) - \text{hari perawatan rumah sakit}}{\text{Jumlah pasien keluar (hidup + mati)}} \quad (2.4)$$

Idealnya tempat tidur kosong hanya dalam waktu 1 - 3 hari.

## 5. Net Death Rate (NDR)

Yaitu angka kematian > 48 jam setelah dirawat untuk tiap-tiap 1000 penderita keluar. Indikator ini dapat memberikan gambaran mutu pelayanan di rumah sakit.

$$\text{Rumus : } \frac{\text{Jumlah pasien mati < 24 jam dirawat}}{\text{Jumlah pasien keluar (hidup + mati)}} \times 1000\% \quad (2.5)$$

Nilai NDR yang dianggap masih dapat ditolerir adalah < 25 per 1000 penderita keluar.

## 6. Gross Death Rate (GDR)

Yaitu angka kematian umum untuk tiap-tiap 1000 penderita keluar.

$$\text{Rumus : } \frac{\text{Jumlah pasien mati seluruhnya}}{\text{Jumlah pasien keluar (hidup + mati)}} \times 1000\% \quad (2.6)$$

Nilai GDR ideal < 45 per 1000 penderita keluar.

## 7. Rata-rata Kunjungan Poliklinik per hari

Indikator ini dipakai untuk menilai tingkat pemanfaatan poliklinik rumah sakit. Angka rata-rata ini apabila dibandingkan dengan jumlah penduduk diwilayahnya akan memberikan gambaran cakupan pelayanan dari suatu rumah sakit.

$$\text{Rumus : } \frac{\text{Jumlah kunjungan poliklinik}}{\text{Jumlah hari buka klinik}} \quad (2.7)$$

## 8. Kegiatan Klinik Unit Darurat

Angka kematian di Klinik Unit Darurat, dibandingkan dengan jumlah pasien di Unit Darurat.

$$\text{Rumus : } \frac{\text{Jumlah kematian di klinik darurat}}{\text{Jumlah pasien di unit darurat}} \quad (2.8)$$

## 9. Kegiatan Kebidanan/Persalinan

Jumlah lahir hidup dan lahir mati dibandingkan dengan total persalinan.

$$\text{Rumus : } \frac{\text{Jumlah kelahiran (hidup+mati)}}{\text{Total persalinan}} \quad (2.9)$$

## 10. Kegiatan Pembedahan

Prosentase pembedahan menurut golongan operasi

## 11. Kegiatan Pemeriksaan Laboratorium

Prosentase pemeriksaan laboratorium menurut masing-masing kategori.

## 12. Kegiatan Keluarga Berencana Rumah Sakit

Jumlah peserta KB baru di RS, keluhan/efek samping.

## 13. Kegiatan Rujukan

Jumlah pasien rujukan dan pasien dirujuk dengan total pasien.

## 14. Cara Pembayaran

Perbandingan antara pasien yang membayar, Asuransi, keringanan, gratis dengan total pasien.

### 2.3 Analisis Multivariat

Menurut Hair, Black, Babin, Anderson, Tathan, 2006, analisis data multivariat adalah semua metode statistik yang secara serempak menganalisis berbagai pengukuran pada masing-masing individu/objek yang diinvestigasi. Hal ini dilakukan antara lain untuk pengukuran, menjelaskan dan memprediksi, serta untuk pengujian hipotesis terhadap objek yang diinvestigasi.

### 2.3.1 Skala Pengukuran

Terdapat dua macam jenis data yaitu.

- Skala Pengukuran Non Metrik

Data non metrik adalah berupa atribut, karakteristik atau kategorikal yang mengidentifikasi atau menggambarkan sebuah subjek, bisa berupa nominal ataupun ordinal. Nominal dapat diartikan sebagai jumlah angka yang mengidentifikasi objek.

- Skala pengukuran Metrik

Data metrik didefinisikan dengan angka, bisa berupa rasio atau interval. Rasio merupakan pengukuran yang lebih presisi dari interval karena mereka mengelompokkan semua dalam skala terendah dan zero poin yang absolut. Interval menyajikan tingkatan level pada pengukuran presisi yang mendekati operasi atau perhitungan matematika.

### 2.3.2 Tipe-Tipe Teknik Multivariat

Analisis multivariat merupakan seperangkat teknik pengembangan yang pernah ada untuk menganalisis data. Teknik-teknik multivariat ini terbagi menjadi dua yaitu :

1. Teknik Dependence

Sebuah atau seperangkat variabel diidentifikasi sebagai variabel dependent berdasarkan prediksi atau penjelasan dari variabel lainnya yang diketahui sebagai variabel independent. Yang termasuk dalam teknik ini adalah :

- a. Multiple Regression, digunakan untuk menguji hubungan antara sedikitnya dua variabel independent dengan skala interval dan satu variabel dependent juga dengan skala interval.
- b. Multiple Discriminant Analysis, digunakan untuk memprediksi peluang objek yang terlibat dalam dua atau lebih kategori mutually exclusive (variabel dependent) berdasarkan pada beberapa variabel independent. Atau dengan kata lain variabel dependent non metrik/kategorikal diprediksi oleh beberapa variabel independent metrik.

- c. Logit atau Logistik Regression, apabila satu variabel dependent non metrik diprediksi oleh beberapa variabel independent metrik.
- d. Manova dan Covariance merupakan teknik statistik yang menyediakan tes perbedaan rata-rata antara grup untuk dua atau lebih variabel dependent. Dimana beberapa variabel dependent metrik diprediksi oleh seperangkat variabel independent non metrik/kategorikal.
- e. Conjoint Analysis, digunakan untuk mengerti pilihan responden terhadap produk dan jasa.
- f. Canonical Correlation. Apabila terdapat dua atau lebih kriteria variabel (variabel dependent) dengan banyak variabel prediktor (variabel independent). Ini merupakan perluasan dari multiple regresi.
- g. Structural Equations Modeling (SEM) merupakan berbagai perkiraan hubungan interrelated dependence berdasarkan pada dua komponen, yaitu struktur model dan model pengukuran.

## 2. Teknik Interdependence

Teknik ini melibatkan analisa bersama dari semua variabel, tanpa membedakan variabel dependent dan independent. Yang termasuk ke dalam teknik ini adalah :

- a. Principal Component dan Common Factor Analysis, digunakan untuk menganalisis hubungan struktur antara sejumlah besar variabel untuk menentukan seperangkat dimensi umum faktor.
- b. Cluster Analysis, yaitu kelompok objek (responden, produk, dll) yang masing-masing objek sama dengan yang lainnya dalam satu kluster dan berbeda dari objek lain dalam kluster lainnya.
- c. Multidimensional Scalling (perceptual mapping), digunakan untuk mengidentifikasi dimensi-dimensi yang tidak dikenali yang mempengaruhi perilaku pembelian berdasarkan pada pertimbangan pelanggan (kesamaan atau pilihan) dan mentransformasiny ke jarak representasi sebagai peta persepsi.

- d. Correspondence Analysis, menggunakan data non metrik dan mengevaluasi hubungan linear dan non linear dalam usaha untuk mengembangkan peta perseptual yang merepresentasikan hubungan antara objek dan sebuah penjelasan karakteristik objek tersebut.

## 2.4 Asumsi-Asumsi Analisis Multivariat

Banyak asumsi yang dibutuhkan dalam teknik multivariat, empat diantaranya memberikan pengaruh yang sangat kuat dalam setiap teknik statistik univariate dan multivariat.

### 2.4.1 Normality

Asumsi mendasar dalam analisis multivariate adalah *normality*, berkaitan dengan bentuk distribusi data untuk sebuah variabel metrik dan hubungannya kepada distribusi normal. Jika variasi dari distribusi normal begitu besar, semua hasil tes statistik menjadi tidak benar, karena *normality* dibutuhkan untuk pengujian F dan t. Metode statistik univariate dan multivariate didasarkan pada asumsi *univariate normality*, dengan metode multivariate juga diasumsikan *multivariate normality*.

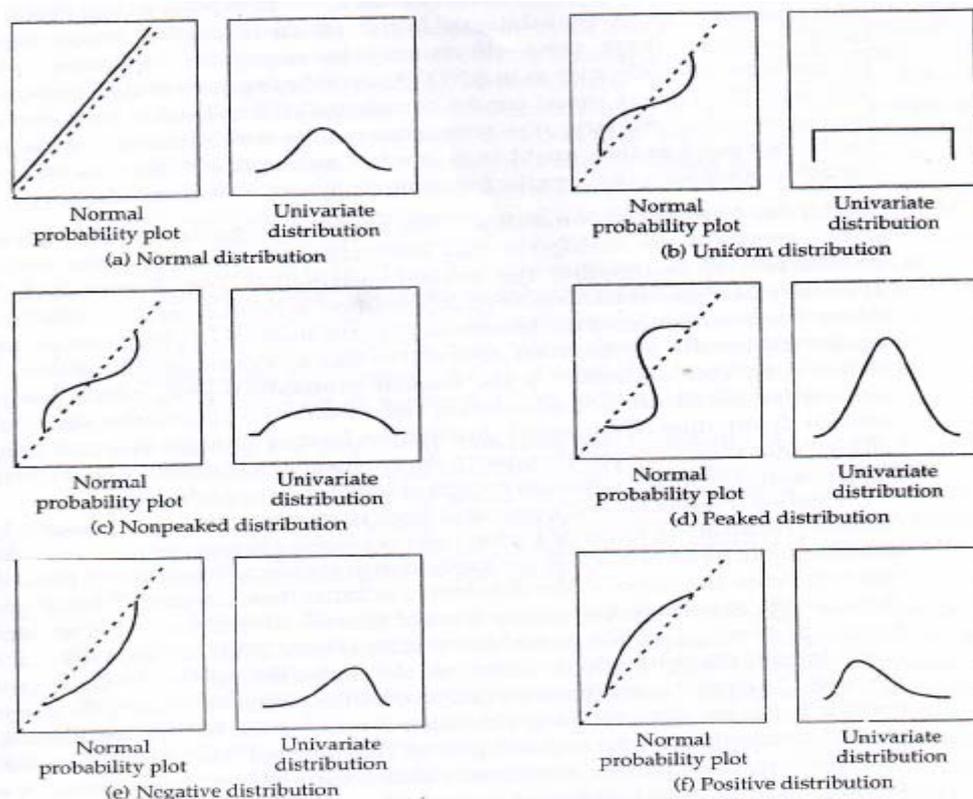
*Univariate normality* untuk satu variabel sangat mudah untuk diuji, *multivariate normality* (kombinasi dari dua atau lebih variabel) berarti bahwa variabel individu bersifat normal dalam sudut pandang *univariate* dan kombinasinya juga normal.

Penaksiran dari *nonnormality* adalah berdasarkan dua dimensi yaitu bentuk dari distribusi serta ukuran sampel. Bentuk distribusi bisa digambarkan dalam dua bentuk yaitu kurtosis dan skewness. Kurtosis berkaitan dengan puncak distribusi yang rendah (*peakness*) dan puncak distribusi yang terlalu tinggi (*flattness*) dibandingkan dengan distribusi normal. Skewness digunakan untuk menggambarkan keseimbangan distribusi, ini menjadi tidak seimbang jika bergeser ke satu sisi (kiri atau kanan). Positif skew apabila bergeser kiri, sedangkan negatif skew direfleksikan oleh pergeseran ke kanan.

Ukuran sampel mempunyai pengaruh dalam meningkatkan kekuatan statistik dengan cara mengurangi *sampling error*. Jika ukuran sampel besar akan mengurangi dampak *nonnormality*. Dalam sampel pengamatan sebanyak 50 atau

kurang, khususnya jika sampel kurang dari 30, ini akan memberikan dampak signifikan terhadap *normality*.

Sejumlah transformasi data tersedia untuk mengakomodir distribusi yang tidak normal. Bagaimanapun, ketika menguji dengan menggunakan metode multivariat, seperti *multivariate regression* atau *multivariate analysis of variance*, sebaiknya kita menggunakan *multivariate normality* (variabel dalam individu normal, begitu juga kombinasinya). Jika ketidaknormalan data telah terindikasi, ini juga akan berkontribusi terhadap pelanggaran asumsi lainnya.



**Gambar 2.1 Normal Probability Plots and Corresponding Univariate Distribution**

Sumber : Hair, Black, Babin, Anderson, Tatham, 2006

#### 2.4.2 Homoscedasticity

*Homoscedasticity* berhubungan dengan asumsi bahwa variabel dependent menunjukkan tingkat variansi yang sama terhadap rentang variabel-variabel prediktor. *Homoscedasticity* sangat diperlukan karena variansi dari variabel dependent akan dijelaskan dalam hubungan ketergantungan harus tidak terkonsentrasi hanya dalam rentang terbatas dari nilai independent. Dalam beberapa situasi, terdapat banyak perbedaan nilai dari variabel dependent pada

Universitas Indonesia

masing-masing nilai variabel independent. Untuk hubungan ini harus benar-benar diperoleh, penyimpangan (variansi) dari nilai variabel independent harus benar-benar sama dengan masing-masing nilai variabel prediktor. Jika penyimpangan ini tidak sama, hubungan ini disebut *heteroscedasticity*.

Variabel *heteroscedatic* bisa diatasi dengan transformasi data seperti distribusi yang tidak normal sebelumnya. Seperti yang dijelaskan sebelumnya, seringkali *heteroskedasticity* dihasilkan dari variabel-variabel yang tidak normal.

### 2.4.3 *Linearity*

Sebuah asumsi implisit untuk semua teknik multivariat yang berdasarkan pada pengukuran korelasi, seperti *multiple regression*, *logistic regression*, *factor analysis*, dan *structural equation modelling*, adalah *linearity*. Karena korelasi hanya akan terlihat dengan hubungan linear dari variabel-variabel, pengaruh nonlinear tidak akan merepresentasikan nilai korelasi.

Cara yang paling umum digunakan untuk melihat linearity adalah menguji scatterplot dari variabel-variabel. Jika terdeteksi ada hubungan nonlinear, pendekatan yang digunakan adalah secara langsung mentransform satu atau kedua variabel untuk mencapai *linearity*.

### 2.4.4 *Absence of Correlated Errors*

Prediksi dalam beberapa teknik dependent tidaklah sempurna. Bagaimanapun kita berusaha untuk memastikan bahwa kesalahan prediksi (eror) yang satu tidak saling berhubungan dengan yang lainnya.

Salah satu penyebab pelanggaran asumsi ini adalah karena proses pengumpulan data. Faktor yang sama akan berpengaruh kuat pada satu kelompok tetapi mungkin tidak bagi kelompok lainnya. Jika kelompok dianalisa secara terpisah, pengaruhnya adalah tetap dalam masing-masing kelompok dan tidak berdampak pada hubungan perkiraan. Tetapi jika pengamatan dari kedua kelompok dikombinasikan, kemudian perkiraan hubungan akhir harus dikompromikan antara dua hubungan aktual. Dampak dari kombinasi ini berperan penting terhadap hasil yang berat sebelah (*biased*) karena sebuah penyebab yang tak ditentukan berdampak pada hubungan perkiraan.

Sumber *correlated error* lainnya adalah data *time series*. Data untuk setiap periode sangat berhubungan dengan data periode sebelumnya atau sesudahnya. Untuk mengatasi *correlated error* ini bisa dilakukan dengan memasukkan faktor penyebab mengabaikannya dengan menyatakan alasan-alasannya ke dalam analisis multivariat.

## 2.5 Transformasi Data

Transformasi data menyediakan sebuah cara untuk memodifikasi variabel-variabel untuk satu atau dua alasan berikut : (1) untuk memperbaiki pelanggaran dari asumsi-asumsi statistik dalam teknik multivariat, atau (2) untuk meningkatkan hubungan (korelasi) di antara variabel-variabel. Transformasi data bisa didasari oleh alasan secara teoritikal (transformasi yang kepentasannya adalah berdasarkan pada sifat dari data) atau asal data (dimana transformasi sangat disarankan dengan sebuah pengujian data).

### a. Transformasi Untuk Mencapai Kenormalan dan *Homoscedasticity*

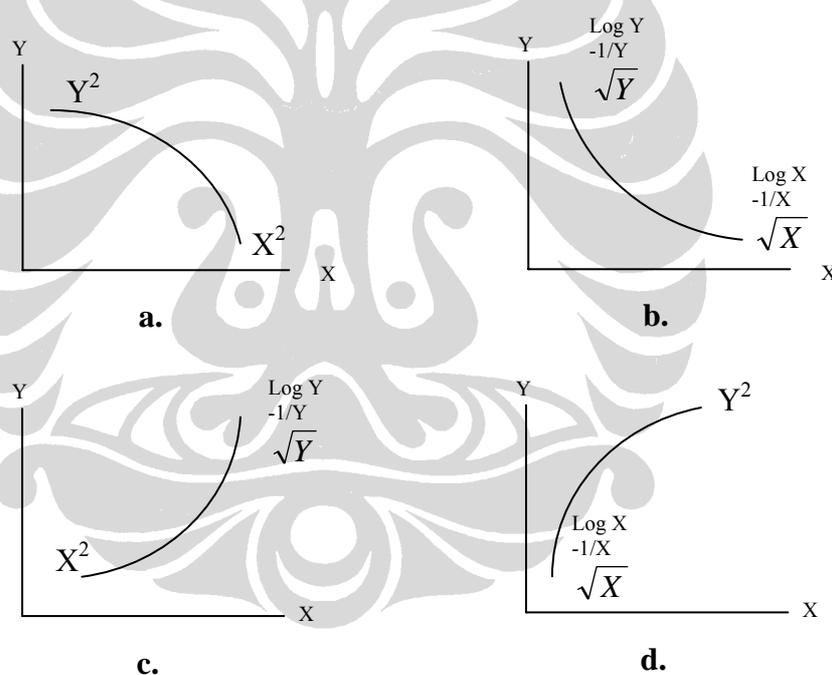
Transformasi data menyediakan cara mendasar untuk memperbaiki ketidaknormalan dan *heteroscedasticity*. Dalam hal ini, pola dari variabel-variabel menentukan transformasi secara spesifik. Untuk distribusi yang tidak normal, dua pola yang biasa terjadi adalah *flat distribution* atau *skewed distribution*. Untuk *flat distribution*, transformasi yang bisa dilakukan adalah inverse (contoh :  $1/Y$  atau  $1/X$ ). *Skewed distribution* bisa ditransformasi dengan hubungan akar pangkat dua, logaritma, kuadrat atau kubik ( $X^2$  atau  $X^3$ ) atau bahkan invers dari variabel. Biasanya, *skewed distribution* negatif sebaiknya ditransformasikan dengan menerapkan transformasi kuadrat atau kubik, dimana logaritma dan akar pangkat dua biasanya bekerja dengan baik pada positif *skewness*. Dalam beberapa hal, peneliti bisa menerapkan semua transformasi yang mungkin dan kemudian memilih variabel transformasi yang paling tepat.

*Heteroscedasticity* adalah sebuah permasalahan yang berhubungan, dan dalam beberapa hal menghilangkan masalah ini akan terselesaikan dengan menyelesaikan permasalahan *normality* secara baik. *Heteroscedasticity* juga disebabkan oleh distribusi dari variabel-variabel. Ketika memeriksa scatterplot, pola yang biasa terjadi adalah distribusi berbentuk kerucut. Jika kerucut ny

membuka ke kanan, lakukan invers; jika kerucutnya terbuka ke kiri, lakukan dengan akar pangkat dua. Beberapa transformasi bisa dihubungkan dengan tipe data tertentu. Contohnya, jumlah frekuensi disarankan dengan sebuah transformasi akar pangkat dua; proporsi sebaiknya ditransformasi dengan arcsin ( $X_{\text{new}} = 2 \arcsin \sqrt{X_{\text{old}}}$ ); dan perubahan proporsi sebaiknya ditangani dengan logaritma terhadap variabel. Dalam semua hal, sekali transformasi dilakukan, data yang telah ditransformasikan harus diuji untuk melihat apakah asumsi yang diinginkan telah dicapai.

#### b. Transformasi Untuk Mencapai *Linearity*

Sejumlah prosedur tersedia untuk mencapai *linearity* di antara dua variabel, tetapi dalam hubungan nonlinear sederhana bisa ditempatkan pada salah satu dari empat kategori seperti pada gambar 2.2.



**Gambar 2.2 Pemilihan Transformasi Untuk Mencapai *Linearity***

Sumber : Hair, Black, Babin, Anderson, Tatham, 2006

Dalam masing-masing kuadran, diberikan transformasi yang paling potensial untuk kedua variabel dependen maupun independen. Contohnya, jika hubungan terlihat seperti gambar 2.2 a, maka variabelnya bisa dikuadratkan untuk mencapai *linearity*. Ketika tersedia beberapa alternatif transformasi, mulai dari metode yang paling atas untuk masing-masing kuadran dan bergerak ke bawah sampai *linearity*

dicapai. Sebuah pendekatan alternatif adalah dengan menggunakan variabel tambahan, yang disebut *polynomial*, untuk merepresentasikan komponen nonlinear.

Terdapat beberapa poin yang perlu diingat ketika melakukan transformasi data, berikut beberapa aturan yang harus diikuti :

- Untuk menduga pengaruh potensial dari sebuah transformasi, hitung rasio mean dan standar deviasi variabel. Efeknya akan terlihat jika diperoleh rasio kurang dari empat. Jika transformasi bisa dilakukan pada kedua variabel, pilih variabel dengan rasio terkecil,
- Transformasi harus diterapkan untuk variabel independen kecuali dalam kasus *heteroscedasticity*,
- *Heteroscedasticity* bisa diselesaikan hanya dengan mentransformasi variabel dependen dalam sebuah hubungan dependen, dan jika memungkinkan harus ditransformasi pada variabel dependen,
- Transformasi bisa merubah interpretasi terhadap variabel, contohnya variabel yang ditransformasi dengan menggunakan logaritma menterjemahkan hubungannya dalam sebuah pengukuran pada perubahan proporsi (*elasticity*); selalu yakin untuk mengembangkan sepenuhnya kemungkinan interpretasi dari variabel-variabel yang ditransformasi,
- Menggunakan variabel dalam format asli (tanpa ditransformasi) ketika menginterpretasikan hasil.

## 2.6 *Multiple Regression Analysis*

*Multiple regression analysis* digunakan untuk menguji hubungan antara sedikitnya dua variabel independent dengan skala interval dan satu variabel dependent juga dengan skala interval.

Contoh model regresi bisa dinyatakan sebagai berikut :

Perkiraan jumlah Kartu kredit yang = intercept + perubahan jumlah kartu kredit karena perubahan satu unit ukuran keluarga x Ukuran keluarga

Universitas Indonesia

Atau

$$Y = b_0 + b_1 V_1$$

Interpretasi dari persamaan tersebut adalah :

1. Koefisien regresi. Perkiraan perubahan dari variabel dependen untuk setiap unit perubahan variabel independen. Jika variabel independen ditemukan secara signifikan berpengaruh (koefisien benar-benar signifikan berbeda dari nol), nilai dari koefisien regresi mengindikasikan seberapa besar hubungan variabel independent tersebut terhadap variabel dependennya.
2. Intercept. Interpretasi intercept kadang-kadang berbeda. Intercept adalah hanya sebagai nilai explanatory dalam rentang nilai variabel-variabel independen. Lebih dari itu, interpretasinya berdasarkan pada karakteristik dari variabel independen :
  - a. Dalam hubungan sederhana, intercept adalah hanya sebagai nilai interpretif jika variabel independen bernilai nol. Contoh, misalkan variabel independennya adalah biaya periklanan. Jika dalam kenyataan atau pada beberapa kondisi, tidak ada periklanan yang dilakukan, maka nilai intercept akan merepresentasikan nilai variabel dependen ketika biaya periklanan nol,
  - b. Jika nilai independen merepresentasikan sebuah pengukuran yang tidak pernah bisa mempunyai nilai sebenarnya nol (seperti tingkah laku atau persepsi), intercept membantu dalam meningkatkan proses prediksi, tetapi tidak mempunyai nilai eksplanatory.

Tahapan dalam *multiple regression analysis* :

### 2.6.1 Sasaran dari Multiple Regression

Tahap awal dalam memulai *multiple regression analysis* adalah penentuan permasalahan penelitian. Dalam memilih pengaplikasian multiple regresi, peneliti harus mempertimbangkan tiga isu utama, yaitu :

- Ketepatan permasalahan penelitian,  
Aplikasi multiple regresi terbagi menjadi dua kelas berdasarkan permasalahan penelitian, yaitu prediksi dan penjelasan. Satu tujuan fundamental dari multiple regresi adalah untuk memprediksi variabel

dependen dengan seperangkat variabel independen. Multiple regresi juga mempunyai tujuan penjelasan, yaitu memberikan arti dari derajat dan karakter hubungan antara variabel dependen dan independent dengan membentuk variate dari variabel independen dan kemudian menguji magnitude, tanda, dan signifikansi statistik dari koefisien regresi untuk masing-masing variabel independen.

- Spesifikasi sebuah hubungan statistik,  
Multiple regresi tepat jika peneliti tertarik dalam hubungan statistik, bukan hubungan fungsi. Karena dalam hubungan statistik pasti terdapat error dalam perkiraan kita.
- Pemilihan dependent dan independent variabel.  
Karena multiple regresi adalah teknik dependen, peneliti harus menentukan variabel dependent dan independent.

### **2.6.2 Rancangan Penelitian Multiple Regression Analysis**

Peneliti harus memperhatikan tiga fitur, yaitu

- Ukuran Sampel,  
Ukuran sampel mempunyai pengaruh langsung terhadap ketepatan dan kekuatan statistik dari multiple regresi. Aturan umum tentang rasio pengamatan terhadap variabel independent tidak boleh kurang dari 5:1, artinya lima pengamatan dibuat untuk masing-masing variabel independent dalam variate. Walaupun rasio minimumnya adalah 5:1, tetapi dibutuhkan 15 sampai 20 pengamatan untuk masing-masing variabel independent. Jika level ini tercapai, maka hasilnya akan mewakili secara keseluruhan karena sampel yang digunakan sudah representatif.
- Elemen Unik dari Hubungan Dependence,  
Hubungan dasar yang digambarkan dalam multipel regresi adalah liner antara metric variabel dependent dan independent berdasarkan pada korelasi produk momen. Satu permasalahan yang dihadapi oleh peneliti adalah jika data yang dimiliki adalah non mentric. Selain itu, multiple regresi tidak bisa secara langsung memodelkan hubungan nonlienear.

Dalam situasi ini, harus dihasilkan satu variabel baru dengan transformasi atau mengubah data nonmetrik dengan variabel dummy.

- Sifat dari Variabel Independent (tetap atau random).

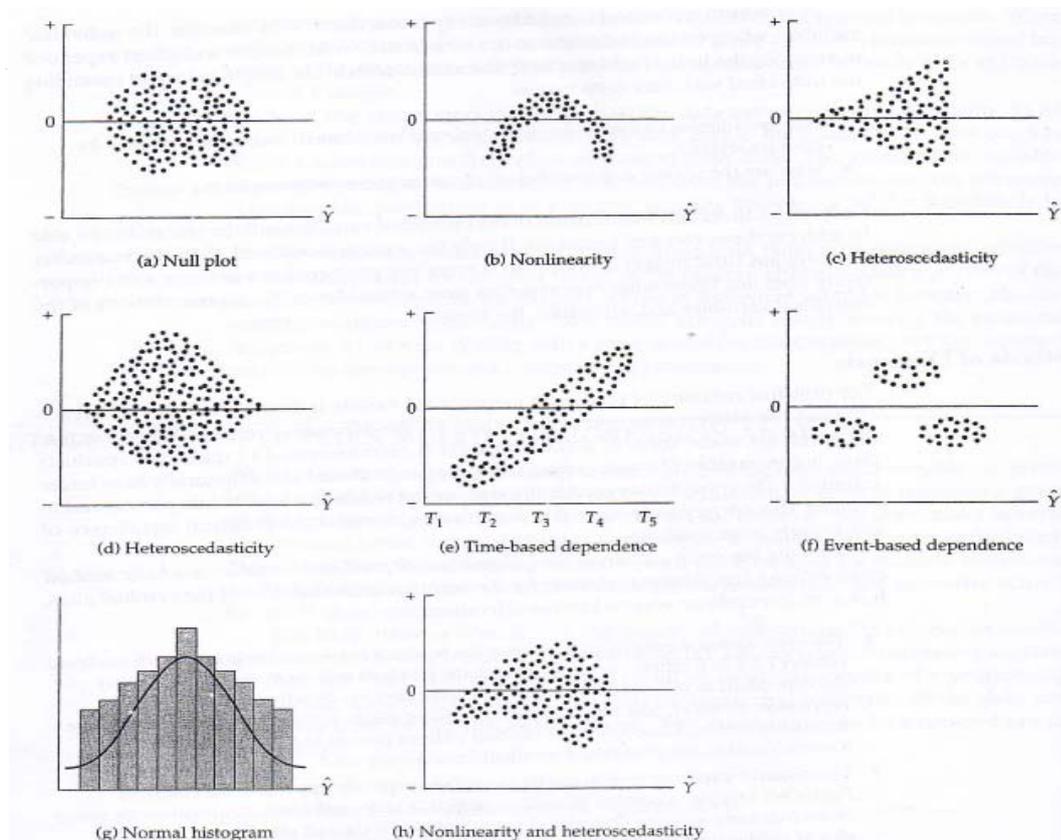
### 2.6.3 Asumsi-asumsi dalam Multiple Regression

Asumsi-asumsi yang harus diuji dalam *multiple regression analysis* yaitu (1) *linearity* dari fenomena yang diukur, (2) Variansi yang konstan dari eror (*homoscedasticity*), (3) eror yang bebas atau tidak saling berhubungan (autokorelasi), (4) *Normality* dari distribusi eror. Jika semua asumsi terpenuhi, maka plot residualnya berbentuk *null plot* seperti gambar 2.3 (a).

*Linearity* adalah hubungan antara variabel dependen dan independen merepresentasikan pada tingkat mana perubahan variabel dependen dihubungkan dengan variabel independen. Konsep korelasi adalah berdasarkan pada sebuah hubungan yang linear. *Linearity* dari beberapa hubungan bivariate secara mudah diuji melalui plot residunya. Gambar 2.3 (b) menunjukkan pola residual yang mengindikasikan hubungan nonlinear dalam model yang ada saat ini. Dalam *multiple regresi* dengan lebih dari satu variabel, sebuah pengujian residual hanya menunjukkan kombinasi dampak dari semua variabel independen, tetapi tidak bisa menguji variabel independen secara terpisah dalam sebuah plot residual. Untuk itu digunakan *parsial regression plots* yang akan menunjukkan hubungan satu variabel independen terhadap variabel dependen, mengontrol semua efek dari semua variabel-variabel independen. Oleh karena itu, *parsial regression plots* menggambarkan hubungan yang unik antara variabel dependen dan variabel independen.

Adanya variansi yang tidak sama (*heteroscedasticity*) adalah salah satu pelanggaran asumsi yang biasa terjadi. Diagnosisnya dibuat dengan plot residu atau pengujian statistik sederhana. Plot residual (*studentized*) dengan nilai dependen perkiraan dan membandingkannya pada *null plot*. Gambar 2.3 (a) menunjukkan pola yang konsisten jika variansinya tidak konstan. Pola yang mungkin terjadi adalah bentuk triangul yang ditunjukkan secara langsung pada gambar 2.3 (c). Pola berbentuk berlian gambar 2.3 (d) bisa dikira-kira persentasi dimana lebih banyak variasi pada *midrange* daripada pada ujungnya. Beberapa

kali, sebuah pelanggaran diperoleh secara bersamaan, seperti *nonlinearity* dan *heteroscedasticity* dalam gambar 2.3 (h).



**Gambar 2.3 Graphical Analysis Of Residual**

Sumber : Hair, Black, Babin, Anderson, Tatham, 2006

Dalam regresi kita berasumsi bahwa masing-masing nilai perkiraan adalah independen, yang berarti bahwa nilai yang diperkirakan tidak berhubungan terhadap prediksi lainnya. Untuk mengidentifikasinya dapat dilakukan dengan plot residual terhadap beberapa variabel yang mungkin berurutan. Jika plot residual independen, pola kelihatan acak dan sama dengan residual *nuul plot*. Pelanggaran akan bisa diketahui dengan sebuah pola residual yang konsisten. Gambar 2.3 (e) menggambarkan sebuah plot residual yang menunjukkan hubungan antara residual dan waktu. Pola lainnya ditunjukkan dalam gambar 2.3 (f). Pola ini diperoleh jika kondisi model dasar berubah tetapi tidak dimasukkan ke dalam model. Misalnya penjualan pakaian renang diukur per bulan selama 12 bulan,

dengan dua musim dingin dan satu musim panas, tidak ada faktor cuaca yang diperkirakan.

Pelanggaran asumsi yang mungkin paling sering terjadi adalah *nonnormality* dari variabel independen atau variabel dependen atau keduanya. Diagnosa yang paling sederhana adalah dengan menggunakan sebuah histogram dari residual, dengan pengujian visual untuk sebuah distribusi yang kurang lebih distribusi normal seperti ditunjukkan gambar 2.3 (g). Metode ini menjadi sulit diterapkan untuk sampel yang sedikit. Metode yang lebih baik digunakan adalah menggunakan *normal probability plots*. Ini berbeda dengan plot residual dalam *standardized residual* dibandingkan dengan distribusi normal. Distribusi normal membuat sebuah garis lurus diagonal, dan residual yang telah diplot dibandingkan dengan diagonal tersebut. Jika sebuah distribusi normal, maka garis residualnya mengikuti diagonal.

#### 2.6.4 Memperkirakan Model Regresi dan Menaksir Keseluruhan Model

Ada tiga tugas dasar yaitu :

- a. Memilih metode untuk metode regresi spesifik yang akan diestimasi

Ada tiga pendekatan yang dapat digunakan untuk menentukan model regresi. Pendekatan tersebut, (1) *Confirmatory Specification*, cara yang paling sederhana dan paling banyak diminati. Dimana peneliti menentukan secara tepat variabel-variabel independen yang akan dimasukkan dalam model regresi. Dibandingkan dengan pendekatan *sequential* atau *combinatorial* yang akan didiskusikan di bawah, pendekatan ini memungkinkan peneliti untuk mengontrol secara penuh terhadap pemilihan variabel.

(2) *Sequential Search Methods*, ada dua tipe dalam pendekatan ini, yaitu *stepwise estimation* (dimulai dengan model regresi sederhana dengan memasukkan variabel independen yang mempunyai korelasi paling tinggi  $Y = b_0 + b_1X_1$ , kemudian dilihat porsi signifikan variabel yang paling tinggi untuk kemudian dimasukkan kembali dalam persamaan regresi. Olah kembali persamaan regresi dengan dua variabel independen tersebut, jika tidak memberikan hasil yang signifikan, maka variabel tersebut dikeluarkan dari persamaan regresi. Jika memberikan nilai yang signifikan maka variabel tersebut dimasukkan dalam

persamaan regresi yang mejadi  $Y = b_0 + b_1X_1 + b_2X_2$ . Lakukan prosedur ini untuk semua variabel independen. *Forward addition* dan *backward elimination*, caranya hampir sama dengan *stepwise estimation*, tetapi *forward addition* dimulai dengan persamaan regresi dengan satu variabel independen sedangkan *backward elimination* dimulai dengan persamaan regresi dengan memasukkan seluruh variabel independen ke dalamnya, dan kemudian menghilangkan variabel independen yang tidak memberikan kontribusi yang signifikan. Perbedaan utama dari *stepwise estimation* dari *forward addition* dan *backward elimination* adalah kemampuan menambah atau menghapus variabel-variabel pada masing-masing tahap. Sekali variabel tersebut sudah ditambahkan atau dihapus dari persamaan, tidak boleh dilakukan lagi dalam tahap berikutnya. Bagi beberapa peneliti, *sequential search methods* kelihatannya merupakan solusi yang sempurna jika dihadapkan dengan pilihan pendekatan *confirmatory* untuk mencapai kekuatan prediksi maksimum dengan hanya variabel-variabel yang berkontribusi signifikan. Tetapi ada beberapa kekurangan dalam penggunaan *sequential search methods* ini, yaitu pertama, *multicollinearity* menjadi kurang dipertimbangkan dalam interpretasi model bukan hanya menguji persamaan regresi akhirnya saja, kedua, semua *sequential search methods* memberikan kontrol yang kurang bagi penelitiannya. Metode ini menjadi sering digunakan karena efisiensinya dalam memilih seperangkat variabel dependen yang memaksimalkan keakuratan prediksi. Dengan kelebihan ini menjadi potensial untuk menyesatkan dalam penjelasan hasil, dimana hanya seperangkat variabel yang berkorelasi tinggi saja yang dimasukkan dalam persamaan dan menghilangkan kontrol terhadap menentukan model.

(2) *Combinatorial approach*, menggunakan semua kombinasi variabel independent yang mungkin dalam pengujian dan seperangkat variabel yang paling tepat yang diidentifikasi. Contohnya, sebuah model dengan 10 variabel independen dengan 1024 kombinasi regresi yang mungkin.

- b. Menaksir nilai signifikan statistik dari seluruh model dalam memprediksi variabel dependent.

Dengan variabel-variabel independen yang telah dipilih dan koefisien regresi yang telah diestimasi, peneliti sekarang harus menaksir model yang telah

diperkirakan untuk menemukan asumsi-asumsi dalam multiple regresi. Pengujian yang perlu dilakukan yaitu pengujian signifikan dari keseluruhan model dengan menggunakan uji F dan juga uji t untuk pengujian signifikan terhadap koefisien model regresi.

c. Menentukan apakah ada observasi yang memberi pengaruh pada hasil.

Sekarang fokus pada indentifikasi pola umum dalam keseluruhan pengamatan. Disini kita menggeser perhatian kita pada pengamatan individu, dengan sasaran untuk menemukan pengamatan yang berada di luar pola umum data atau pengamatan yang memberikan pengaruh sangat kuat pada hasil regresi. Pengamatan ini tidak dibutuhkan dan harus dihapus. Dalam beberapa hal mereka merepresentasikan elemen tersendiri dari seperangkat data. Bagaimanapun kita harus mengidentifikasi dan menaksir dampaknya sebelum memprosesnya lebih lanjut.

### 2.6.5 Interpretasi Variabel Regresi

Pada tahap ini dilakukan interpretasi variate regresi dengan mengevaluasi koefisien regresi yang telah diestimasi untuk penjelasan terhadap variabel dependen. Selain itu peneliti tidak hanya harus mengevaluasi model regresi yang sudah diestimasi tetapi juga variabel independen potensial yang dihilangkan jika menggunakan pendekatan *sequential* ataupun *combinatorial*. Selain itu juga perlu diperhatikan *multikolinearity* yang terjadi.

Koefisien regresi yang telah diestimasi, atau disimbolkan dengan  $b$  koefisien, merepresentasikan kedua tipe hubungan (positif atau negatif) dan kekuatan hubungan antara variabel dependen dan independen dalam variate regresi. Tanda pada koefisien menunjukkan apakah hubungannya positif atau negatif, sedangkan nilai koefisiennya mengindikasikan perubahan variabel dependen setiap kali variabel independen berubah satu unit.

Dalam menggunakan koefisien regresi untuk tujuan *explanatory*, kita harus memastikan bahwa semua variabel independen berada pada skala yang bisa dibandingkan. Perbedaan variabilitas dari variabel ke variabel bisa berpengaruh kuat terhadap besarnya koefisien regresi. Untuk membuat semua variabel independen bisa dibandingkan dalam skala dan variabilitas, maka digunakan

modifikasi koefisien regresi yang disebut beta koefisien. Keuntungan penggunaan beta koefisien ini adalah bisa mengurangi permasalahan karena perbedaan unit pengukuran. Contoh untuk melihat perbedaan antara koefisien regresi ( $b$ ) dan beta ( $\beta$ ).

Dalam penggunaan kartu kredit, koefisien regresi ( $b$ ) dan beta ( $\beta$ ) untuk persamaan regresi dengan tiga variabel independen ( $V_1$ ,  $V_2$ ,  $V_3$ ) ditunjukkan sebagai berikut :

Variabel	Koefisien	
	Regresi ( $b$ )	Beta ( $\beta$ )
$V_1$ <i>Family size</i>	.635	.566
$V_2$ <i>Familiy income</i>	.200	.416
$V_3$ <i>Number of Autos</i>	.272	.108

Interpretasi menggunakan koefisien regresi dan koefisien beta ( $\beta$ ) benar-banar memberikan hasil yang berbeda. Koefisien regresi mengindikasikan  $V_1$  terlihat sekali sangat penting dibandingkan dengan  $V_2$  dan  $V_3$ , yang secara kasar disamakan. Koefisien beta memberikan cerita yang berbeda.  $V_1$  adalah masih yang paling penting, tetapi  $V_2$  adalah yang lebih penting sedangkan  $V_3$  cukup penting.

Walaupun koefisien beta merepresentasikan sebuah sasaran pengukuran kepentingan yang bisa diukur secara langsung dibandingkan, dua hal yang harus diperhatikan dalam penggunaannya, yaitu :

1. Koefisien beta digunakan sebagai panduan untuk kepentingan relatif dari variabel independen hanya jika terjadi sedikit *collinearity*. Karena *collinearity* bisa mengganggu kontribusi beberapa variabel independen jika koefisien beta yang digunakan,
2. Nilai beta bisa diartikan hanya dalam konteks variabel lainnya dalam persamaan. Contohnya, nilai beta untuk  $V_1$  *family size* menunjukkan kepentingannya hanya dalam hubungannya dengan  $V_2$  *familiy income*, tidak dalam pengertian absolut. Jika variabel independen lainnya ditambahkan dalam persamaan, koefisien beta untuk *family size* kemungkinan juga akan berubah, karena beberapa hubungan antara *family size* dan variabel baru yang mungkin terjadi.

Universitas Indonesia

Kesimpulannya penggunaan koefisien beta hanya sebagai panduan untuk melihat kepentingan relatif dari variabel-variabel independen yang termasuk dalam persamaan, dan hanya pada variabel-variabel dengan *multicollinearity* minimal.

Cara paling sederhana dan paling umum untuk mengidentifikasi *colinearity* adalah sebuah pengamatan terhadap matriks korelasi dari variabel independen. Adanya korelasi yang tinggi (umumnya 0.90 atau lebih) adalah indikasi pertama terhadap *colinearity*. Ketidakadaan korelasi yang tinggi, bagaimanapun juga tidak memastikan ketiadaan *colinearity*. *Colinearity* disebabkan karena dampak kombinasi dari dua atau lebih variabel independen (disebut *multicollinearity*).

Untuk menaksir *multicollinearity* kita harus mengukur sampai tingkat mana variabel independen dijelaskan oleh variabel-variabel independen lainnya. Dua pengukuran yang paling biasa digunakan untuk menaksir *multiple variabel collinearity* adalah *tolerance* dan inversnya, *variance inflation factor* (VIF).

*Tolerance* adalah jumlah variabilitas dari variabel independen yang tidak dijelaskan oleh variabel-variabel independen lainnya. VIF adalah  $1/\textit{tolerance}$ . Tingginya derajat *multicollinearity* direfleksikan dengan nilai *tolerance* yang kecil dan nilai VIF yang lebih tinggi. Nama VIF diambil dari fakta bahwa akar dari VIF ( $\sqrt{\textit{VIF}}$ ) adalah tingkat dimana standar eror akan meningkat karena *multicollinearity*. Ambang batas dari nilai *tolerance* adalah 0.1 dan VIF 10. Bagaimanapun, untuk ukuran sampel yang kecil, peneliti bisa membatasi karena adanya peningkatan standar eror karena *multicollinearity*.

Menurut Prof. Dr. H. Imam Ghozali, M.Com, Akt uji multikolinearitas dilakukan untuk membuktikan atau menguji ada tidaknya hubungan yang linear antara variabel bebas yang satu dengan variabel bebas lainnya. Dalam analisis multiple regresi, terdapat dua atau lebih variabel bebas atau variabel independen yang diduga akan mempengaruhi variabel tergantung atau dependennya. Pendugaan tersebut akan dapat dipertanggungjawabkan apabila tidak terjadi hubungan yang linear antara variabel-variabel independen. Adanya hubungan yang linear antar variabel independen akan menimbulkan kesulitan dalam memisahkan

pengaruh masing-masing variabel independen terhadap variabel dependennya. Pelanggaran terhadap asumsi ini akan mengakibatkan :

- Tingkat ketelitian koefisien regresi sebagai penduga sangat rendah, dengan demikian menjadi kurang akurat,
- Koefisien regresi serta ragamnya akan bersifat tidak stabil, sehingga adanya sedikit perubahan pada data mengakibatkan ragamnya berubah sangat berarti,
- Tidak dapat memisahkan pengaruh tiap-tiap variabel independen secara individu terhadap variabel dependennya.

Dalam *Multivariate Analysis*, apabila *collinearity* telah terindikasi, maka peneliti dapat mengambil beberapa pilihan (Hair, Black, Babin, Anderson, Tatham, 2006) : (1) Menghilangkan atau mengabaikan variabel independen yang mempunyai korelasi tinggi dengan memberikan alasan-alasan dan mengidentifikasi variabel-variabel independen lainnya untuk memprediksi, (2) Menggunakan model dengan variabel independen berkorelasi tinggi hanya untuk prediksi, (3) Menggunakan korelasi sederhana antara masing-masing variabel independen dan variabel dependen untuk mengerti hubungan variabel independen dan dependen, (4) Menggunakan metoda analisis yang terpercaya seperti *bayesian regression* (atau sebuah kasus tertentu – *ridge regression*) atau regresi pada komponen utama untuk memperoleh sebuah model yang lebih baik merefleksikan efek sederhana dari variabel-variabel independen.

### 2.6.6 Validasi Hasil

Pendekatan empiris validasi yang paling tepat adalah pengujian model regresi dengan sampel baru yang diambil dari populasi umum. Sampel baru itu akan memastikan kemampuan representatif dan bisa digunakan dengan beberapa cara. Pertama, model original bisa memprediksi nilai dari sampel baru, dan perkiraan yang cocok bisa dihitung. Kedua, sebuah model tersendiri bisa diestimasi dari sampel yang baru kemudian membandingkannya dengan persamaan original pada karakteristik-karakteristik seperti signifikan variabel termasuk tanda, ukuran dan tingkat kepentingan relatif dari variabel, dan keakuratan prediksi.