MODUL PRAKTIKUM

STATISTIKA & PROBABILITAS

UNIVERSITAS TEKNOLOGI MATARAM

Jl. Kampus Universitas Teknologi Mataram, Kekalik, Mataram

Data dan Penyajian Data

1. Pendahuluan

- Pengertian

Setiap hari kita berhadapan dengan data dan kadang harus membuat kesimpulan/keputusan berdasarkan data tersebut. Statistik adalah disiplin ilmu yang menyediakan metode untuk membantu kita memahami data. Metode statistik yang digunakan secara cerdas, menawarkan seperangkat alat untuk mendapatkan wawasan tentang dunia di sekitar kita. Saat ini penggunaan analisis statistik telah meluas ke berbagai bidang seperti bisnis, kedokteran, pertanian, ilmu sosial, alam ilmu, dan rekayasa. Statistik mengajarkan kita bagaimana membuat keputusan cerdas terhadap variabilitas dalam sampel dan populasi. Populasi tanpa variabilitas sangat langka,variabilitas hampir universal.Variabilitas perlu dipahami untuk dapat mengumpulkan, mendeskripsikan, menganalisis, dan menarik kesimpulan dari data dalam cara yang masuk akal

Statistik adalah ilmu yang menyediakan metode untuk membantu kita memahami data.

Ada beberapa pendapat mengenai pengertian statistik :

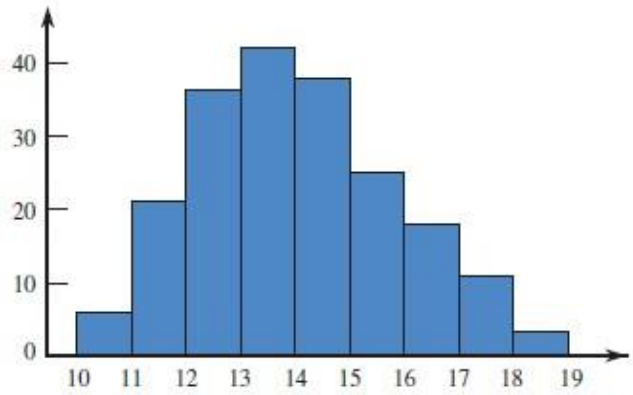
1. Statistik adalah sekumpulan cara maupun aturan-aturan yang berkaitan dengan pengumpulan, pengolahan(Analisis), penarikan kesimpulan, atas data-data yang berbentuk angka dengan menggunakan suatu asumsi-asumsi tertentu. *(Prof.Dr.H.Agus Irianto)*
2. Statistik adalah metode yang memberikan cara-cara guna menilai ketidak tentuan dari penarikan kesimpulan yang bersifat induktif. *(Stoel dan Torrie)*

Kedua pendapat diatas menyatakan statistik adalah metode untuk mendapatkan informasi yang ada pada serangkaian data.

Definisi Statistika

Statistika adalah cabang ilmu yang mempelajari tentang bagaimana mengumpulkan, menganalisis dan menginterpretasikan data. Atau dengan kata lain, statistika menjadi semacam alat dalam melakukan suatu riset empiris.

Penggunaan metode statistik yang benar dapat memberi wawasan yang lebih memadai tentang dunia di sekitar kita. Contoh, sebagai bagian dari upaya menjaga kualitas air. Dilakukan pemantauan air di lima tempat melalui alat kontrol kontaminasi air. Konsentrasi kontaminan dalam per juta (ppm) diukur dari masing-masing lokasi pemantauan.Kemudian rata-rata dari lima pengukuran tersebut dihitung. Histogram berikut merangkum nilai-nilai kontaminasi rata selama 200 hari.



Andaikan telah terjadi tumpahan bahan kimia pada salah pabrik yang berjarak 1 km dari sumur. Tidak diketahui apakah tumpahan akan mencemari tanah pada daerah tumpahan dan, jika demikian, apakah tumpahan ini juga akan mempengaruhi kualitas air sumur yang jauh dari sumber tumpahan. Satu bulan setelah tumpahan, lima spesimen air dikumpulkan dari sumur, dan kontaminasi rata-rata adalah 15,5 ppm. Jika sebelumnya tumpahan, konsentrasi kontaminan rata bervariasi dari hari ke hari dengan nilai rata-rata 15,5 maka nilai rata-rata 15,5 ppm setelah tumpahan ini mengindikasikan bahwa kualitas air sumur tidak dipengaruhi oleh tumpahan.

Dalam menganalisis data, para ilmuwan menggambarkan persepsinya tentang suatu fenomena. Pendekatan awal yang umumnya digunakan untuk menjelaskan suatu fenomena adalah *statistika deskriptif*.

- **Statistika Deskriptif**

Statistika deskriptif adalah tehnik yang digunakan untuk mensarikan data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap orang. Proses kuantifikasi ini kemudian dituangkan dalam berbagai bentuk statistik sederhana, seperti rata-rata hitung dan ditampilkan dalam bentuk tabel dan grafik.

- **Statistika Induktif**

  Pada kasus dimana tidak mungkin seluruh populasi diobservasi-karena berbagai alasan ekonomis ataupun praktis. Mengambil kesimpulan tentang suatu populasi berdasarkan data dari sampel yang terbatas merupakan tujuan dari suatu proses pengambilan keputusan *inferensial* atau *statistik induktif*

2. Data dan Menampilkan Data

- Arti data

  Data ialah bahan mentah yang perlu diolah sehingga hasilkan informasi atau keterangan, baik kualitatif maupun kuantitatif yang menunjukkan fakta. Sedangkan perolehan data seyogyanya relevan artinya data yang ada hubungannya langsung dengan masalah penelitian, *mutakhir artinya* data yang diperoleh masih hangat dibicarakan, dan diusahakan oleh orang pertama (data primer). Data yang sudah memenuhi syarat perlu diolah. Pengolahan data merupakan kegiatan terpenting dalam proses dan kegiatan penelitian. Keliruan memilih analisis dan perhitungan akan berakibat fatal pada kesimpulan, generalisasi maupun interpretasi. Hal ini perlu dikaji secara mendalam hal-hal yang menyangkut pengolahan data, supaya bisa memilih dan menentukan secara tepat dalam pengolahan data.

- Jenis data

  Data menurut jenisnya ada dua yaitu data kualitatif dan data kuantitatif:

  - ***Data Kualitatif*** yaitu data yang berhubungan dengan kategorisasi, karakteristik berwujud pertanyaan atau berupa kata-kata.

    Contohnya: Wanita itu cantik, pria itu tampan, baik, buruk, senang, sedih, harga minyak turun harga dollar naik, rumah itu besar sekali. pohon itu rindang, laut Banda dalam sekali dsb. Data ini biasanya didapat dari wawancara dan bersifat subjektif sebab data tersebut ditafsirkan lain oleh orang yang berbeda. Data kuafitatif dapat diangkakan dalam bentuk ordinal atau ranking.

- ***Data Kuantitatif*** yaitu data yang berwujud angka-angka.

  Contohnya: harga solar Rp 4.300,-/liter, yang diterima menjadi PNS 125 orang, penghasilan PT XYZ sebesar 2 milyar/tahun, Pembayaran SPP TK Harapan Rp 50.000/bulan, dan sebagainya. Data ini diperoleh dari pengukuran langsung maupun dari angka-angka yang diperoleh dengan, mengubah data kualitatif menjadi data kuantitatif. Data kuantitatif bersifat objektif dan bisa ditafsirkan sama oleh semua orang.

- Skala Pengukuran
  - **Skala Nominal** yaitu skala yang paling sederhana disusun menurut jenis (kategori) atau fungsi bilangan hanya sebagai simbol untuk membedakan sebuah karakteristik lainnya. Contoh :
    - Jenis Kulit: Hitam (1), Kuning (2), Putih (3). Angka 1, 2, 3, hanya sebagai label saja.
    - Suku Jawa (1), Madura (2), Bugis (3), Sunda (4), Batak (5), Minang (6)
  - **Skala Ordinal :** skala yang didasarkan pada rangking, diurutkan dari jenjang lebih tinggi sampai jenjang terendah atau sebaliknya. Contoh :

Nilai	100	90	80	70
Rangking	1	2	3	4

  - **Skala Interval** : skala yang menunjukkan jarak antara satu data dengan data yang lain dan mempunyai bobot yang sama.
  - **Skala ratio** skala pengukuran yang mempunyai nilai nol (0) mutlak dan mempunyai jarak yang sama.
- **Populasi dan sampel**

  Populasi merupakan objek atau subjek yang berada pada suatu wilayah dan memenuhi syarat-syarat tertentu berkaitan dengan masalah penelitian.

  Ada dua jenis populasi, yaitu:
  - **Populasi Terbatas:**

    Adalah populasi yang mempunyai sumber data yang jelas batasnya secara kuantitatif sehingga dapat dihitung jumlahnya.

- **Populasi Tak Terbatas (Tak Terhingga)**

  Populasi tak terbatas yaitu sumber datanya tidak dapat ditentukan batas-batasannya sehingga relatif tidak dapat dinyatakan dalam bentuk jumlah.

Berdasarkan sifatnya, populasi dapat digolongkan menjadi:

- *Populasi Homogen* adalah sumber data yang unsurnya memilki sifat yang sama sehingga tidak perlu mempersoalkan jumlahnya secara kuantitatif.
- *Populasi Heterogen* adalah sumber data yang unsurnya memiliki sifat atau keadaan yang berbeda (bervariasi) sehingga perlu ditetapkan batas-batasnya, baik secara kualitatif maupun kuantitatif.

Dalam melaksanakan penelitian, walaupun tersedia populasi yang terbatas dan homogen, adakalanya peneliti tidak melakukan pengumpulan data secara populasi, tetapi mengambil sebagian dari populasi yang dianggap mewakili populasi *(representatif).* Bagian dari populasi yang mempunyai ciri-ciri atau keadaan tertentu yang akan diteliti disebut sampel. Pertimbangan logis penggunaan sampel adalah kepraktisan, keterbatasan biaya, waktu, tenaga dan kadang ada percobaan yang bersifat merusak *(destruktif).*

- Pengolahan data

  Pengumpulan data dan analisis memungkinkan peneliti untuk menjawab pertanyaan berbagai fenomena.Proses analisis data dapat dilihat sebagai urutan langkah-langkah yang terarah dari rencana pengumpulan data hingga didapat kesimpulan berdasarkan informasi yang dihasilkan data. Proses ini dapat dibagi dalam enam langkah berikut:

  1. Memahami sifat dari masalah.

     Efektifitas analisis data memerlukan pemahaman tentang masalah penelitian,tujuan penelitian,apa pertanyaan yang diharapkan untuk terjawab dan memiliki arah yang jelas sebelum pengumpulan data untuk memastikan bahwa data yang dikumpulkan dapat menjawab pertanyaan yang diharapkan dapat terjawab.
  2. Memutuskan apa yang harus diukur dan bagaimana mengukurnya.

     Pada tahapan ini adalah memutuskan informasi apa yang dibutuhkan untuk menjawab pertanyaan yang menarik untuk dijawab. Misalnya, dalam sebuah studi

tentang hubungan antara berat pemain sepak bola dan posisi bermain, kita perlu untuk mengumpulkan data tentang pemain badan pemain dan posisi yang sering mereka tempati di lapangan. Pada kasus lain pilihan informasi tidak dapat diukur langsung.Miisalnya, dalam sebuah studi tentang hubungan antara gaya belajar yang disukai gaya dan intelgensi. Pada kasus ini ada masalah dengan bagaimana mengukur gaya belajar yang baik dan apa ukuran kecerdasan akan yang akan digunakan. Disini penting untuk secara hati-hati dalam menentukan variabel apa yang harus dipelajari dan metode apa tepat untuk menentukan nilainya.

3. Pengumpulan data.

   Langkah pengumpulan data sangat penting. Pada tahap ini harus diputuskan apakah sumber data yang ada saat ini telah memadai atau perlu pengumpulan data baru. Karena jenis analisis yang sesuai dan kesimpulan yang diambil tergantung pada bagaimana data dikumpulkan.

4. Data summarization dan analisis awal.

   Setelah data dikumpulkan, Langkah selanjutnya biasanya melibatkan analisis awal yang meliputi meringkas Data grafis dan numerik. Analisis awal memberikan wawasan penting karakteristik data dan dapat memberikan panduan dalam memilih yang tepat metode untuk analisa lebih lanjut.

5. Analisis data.

   Langkah analisis data dibutuhkan untuk memilih dan menerapkan metode statistik.

6. Interpretasi hasil.

   Intrerpretasi hasil adalah langkah final untuk menjawab beberapa pertanyaan harus ditangani misal

   1. Apa yang dapat kita pelajari dari data?
   2. Kesimpulan apa yang dapat yang ditarik dari analisis?
   3. Bagaimana hasil penelitian dapat jadi pemandu penelitian berikutnya?

   Interprestasi hasil sering menghasilkan pertanyaan penelitian baru, yang pada gilirannya mengarah kembali ke langkah pertama.

- **Menampilkan/ Menyajikan data**

Data populasi atau sampel yang sudah terkumpul dengan baik, apabila digunakan untuk keperluan informasi, laporan atau analisis lanjutan hendaknya diatur, disusun, dan disajikan dalam bentuk yang, jelas, rapih, serta komunikatif dengan cara menampilkan atau menyajikan data yang lebih menarik publik. Hal yang harus diperhatikan dalam penyajian data ini adalah siapa yang memerlukannya dan untuk apa. Dalam sistem manajemen kita mengenal ada tiga tingkatan manajerial dalam suatu organisasi yakni manajemen tingkat bawah, tingkat menengah dan tingkat atas/puncak. Masing-masing tingkatan ini memerlukan data sesuai dengan fungsinya masing-masing. Pada tingkat bawah data atau informasi yang dimiliki masih utuh dan lengkap seperti apa yang diperoleh di lapangan. Pada tingkatan selanjutnya data sudah disaring sedemikian rupa sehingga lebih ringkas dan padat. Dari tingkatan ini biasanya data diolah lebih lanjut untuk menghasilkan alternatif-alternatif dalam proses pengambilan keputusan oleh tingkat manajemen puncak. Adalah sesuatu yang tidak efisien bagi seorang pimpinan puncak untuk membaca seluruh data yang ada, sementara dia harus membuat keputusan yang cepat dan akurat.

Secara umum ada beberapa cara penyajian data statistik yang sering digunakan yaitu *tabel, grafk, diagram* baik untuk data ungroup dan group.

Penyajian data ungroup

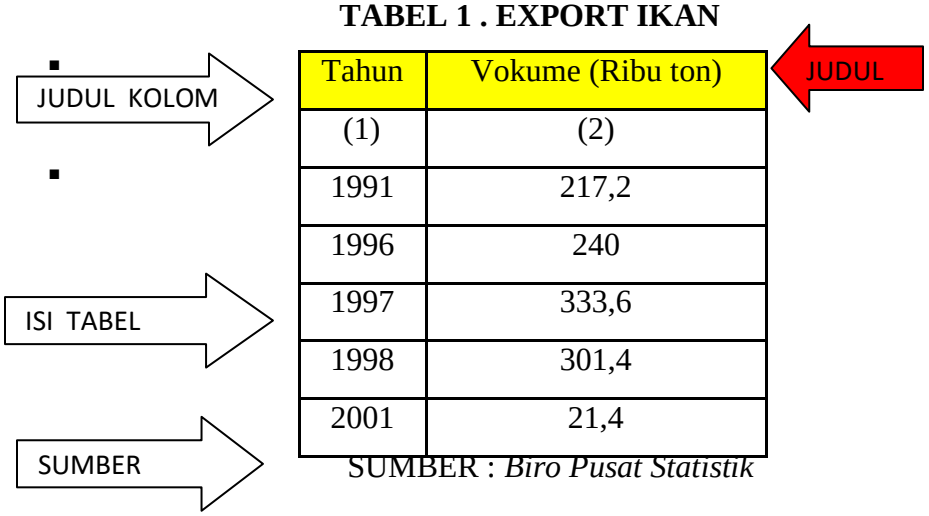
1. Tabel data

   Tabel statistik memuat semua data yang telah diringkas baik secara kuantitatif maupun kualitatif. Tabel statistik dapat dibuat dalam berbagai bentuk mulai dari tabel klasifikasi satu arah, dua arah, tiga arah dan lebih dari tiga arah. Yang mana yang akan dipilih tergantung dari ragam data yang diperoleh. Bentuk tabel biasanya akan terdiri dari komponen-komponen : judul tabel, judul kolom, isi tabel dan sumber. Berikut ini diberikan beberapa contoh tabel klasifikasi satu, dua dan tiga arah.

   - Tabel satu arah

     Adalah tabel yang hanya terdiri atas satu kategori atau karakteristik data.Contoh tabel satu arahsebagai berikut.

TABEL 1 . EXPORT IKAN



Tahun	Vokume (Ribu ton)
(1)	(2)
1991	217,2
1996	240
1997	333,6
1998	301,4
2001	21,4

SUMBER : *Biro Pusat Statistik*

- Tabel dua arah

  Adalah tabel yang terdiri atas dua kategori atau karakteristik data.Contoh tabel dua arah sebagai berikut

No	Jurusan	Asal Daerah Sekolah		Jumlah
		SMA	SMK	
(1)	(2)	(3)	(4)	(5)
1	Sistem Informasi	200	100	300
2	Teknik Informatika	240	121	361
3	Desain Grafis	155	102	257
Total		495	323	918

Sumber : Data buatan

- Tabel tiga arah

  Adalah tabel yang terdiri atas tiga kategori atau karakteristik data.Contoh tabel dua arah sebagai berikut

No	Jurusan	Asal Daerah Sekolah		Jenis Transportasi	
		SMA	SMK	Umum	Motor
(1)	(2)	(3)	(4)	(5)	(6)
1	Sistem Informasi	200	100	100	200
2	Teknik Informatika	240	121	300	61
3	Desain Grafis	155	102	100	107
Total		495	323	500	368

Sumber Data buatan

Pada prinsipnya untuk membuat tabel klasifikasi, tentukan terlebih dahulu apa yang menjadi penekanan utamanya, kemudian penekanan keduanya, ketiga dan selanjutnya, Data dengan penekanan utama ditempatkan pada kolom dan penekanan kedua pada baris. Untuk klasifikasi lebih tinggi lagi, proses ini diulang terus hingga selesai

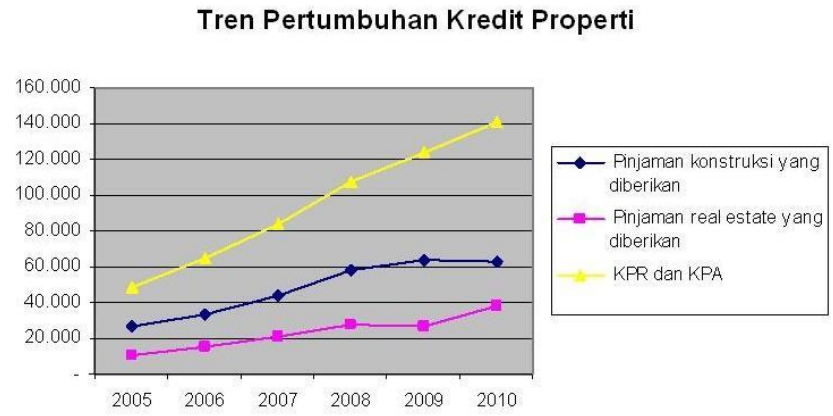
2. Grafik

Data yang tersaji dalam bentuk angka-angka kerapkali tidak begitu disukai. Metode grafik atau dikenal pula dengan istilah diagram adalah cara lain yang dapat untuk menampilkan informasi yang ringkas namun memberikan arti yang tinggi. Sebuah grafik hendaknya dibuat sesederhana mungkin dan menekankan pada aspek-aspek penting dari data yang disajikan. Kecenderungan untuk memasukkan terlalu banyak informasi dalam sebuah grafik haruslah dihindarkan karena hal ini tidak akan

memberikan informasi tambahan malah mungkin dapat membingungkan pembacanya. Berikut adalah beberapa jenis grafik yang umum digunakan untuk menampilkan data.

- Diagram garis

  Diagram garis umumnya digunakan untuk melihat perkembangan suatu gejala atau fakta yang terkait dengan dimensi waktu. Dengan diagram ini maka konsumen akan lebih mudah melihat bagaimana berfluktuasinya nilai data yang sedang diamati. Untuk menggambarkannya sumbu tegak (sumbu-*y*) digunakan untuk menyatakan data yang diamati, sedangkan sumbu mendatar (sumbu-*x*) untuk menyatakan waktu.

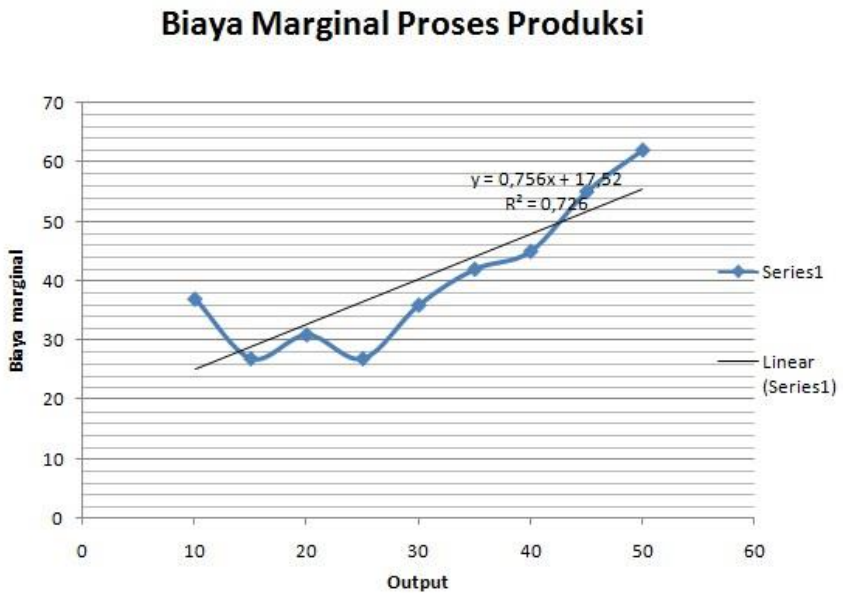


- Kurva

  Kurva sebenarnya sama saja dengan diagram garis kecuali tujuannya adalah untuk melukiskan suatu hubungan yang kontinu. Umumnya kurva digunakan untuk data hasil eksperimen dalam upaya untuk menggambarkan hubungan antara dua deret data. Kurva dua dimensi juga terdiri dari sumbu tegak dan sumbu mendatar. Setelah data diplot, kemudian dibuat kurva yang mendekati titik-titik pada grafik. Biasanya untuk menggambarkan kurva ini digunakan pendekatan melalui persamaan matematis. Contoh

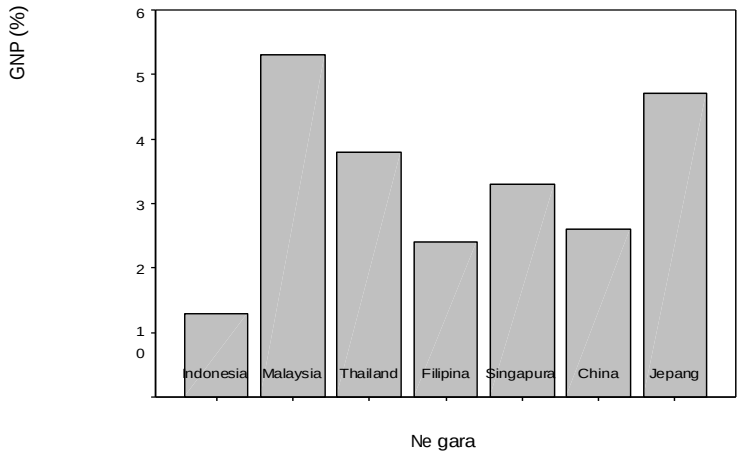
Biaya Marginal	**Output (unit)**
37	10

27	15
31	20
27	25
36	30
42	35
45	40
55	45
62	50



- Grafik/diagram batang

  Diagram batang merupakan salah satu bentuk penyajian data secara visual. Diagram batang sangat tepat digunakan jika variabel berbentuk kategori maupun atribut. Demikian pula data tahunan, asalkan jumlah tahunnya tidak terlalu banyak. Diagram batang bisa digambarkan secara vertikal (sejajar dengan sumbu-*y*) dan horizontal (sejajar dengan sumbu-*x*).



GNP beberapa negara Asia

- Grafik Lingkaran

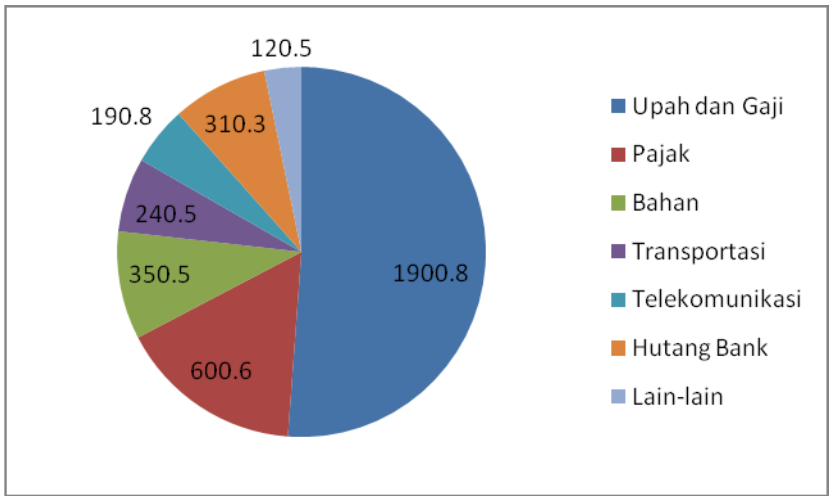
Jika diagram batang dapat digunakan untuk membandingkan sejumlah variabel, selama variabel ini bisa diukur paling sedikit satu aspek, maka diagram lingkaran dapat digunakan bilamana variabel-variabel dijumlahkan ke dalam suatu jumlah tertentu sedemikian rupa sehingga setiap variabel dapat dinyatakan sebagai persentase dari jumlah tersebut. Lingkaran yang memiliki jumlah sudut 360^0 dapat dibagi kedalam segmen-segmen untuk menunjukkan nilai penting relatif dari setiap variabel.

Pengeluaran PT. xyz tahun 2010

Jenis Pengeluaran	Jumlah (juta rupiah)
Upah dan Gaji	1900,8
Pajak	600,6
Bahan	350,5
Transportasi	240,5
Telekomunikasi	190,8
Hutang Bank	310,3
Lain-lain	120,5
Jumlah	3.714,00

SUMBER : *Databuatan*

Grafik



Penyajian data group

Pada data dengan jumlah besar sering digunakan cara pengelompokan (group) data ke dalam interval kelas yang sama panjang. Tabel pengelompokan data ini dikenal dengan nama tabel distribusi.Keuntungan dari penggunaan cara ini adalah dapat terlihatnya bentuk distribusi/penyebaran datanya. Instilah-istilah dalam tabel distrubusi.

- **Interval Kelas**
  - Tiap-tiap kelompok disebut interval kelas atau sering disebut interval atau kelas saja
- **Batas Kelas**
  - Batas bawah
  - Batas atas.
- **Tepi Kelas (Batas Nyata Kelas)**
  - Tepi bawah = batas bawah – 0,5
  - Tepi atas = batas atas + 0,5
- **Lebar kelas**
  - Lebar kelas = tepi atas – tepi bawah

- Tahapan Penyaian data group

1. Tentukan range data

Range data = $Data_{max}$ - $Data_{min}$

2. Tentukan jumlah interval

   Tentukan jumlah klas interval dengan rumus Stugges

   K= 1 + 3,3 log n (n = jumlah data)

3. Tentukan panjang klas interval p

   p = Range data / K

4. Tentukan batas bawah interval

Soal :

Berikut adalah data hasil ujian statistik mahasiswa semester 2

Hasil Ujian Statistka Dasar										
82	72	100	77	59	46	57	55	64	50	62
68	47	65	85	53	81	51	80	56	56	89
52	63	86	84	58	40	76	59	94	93	94
80	82	63	99	98	97	52	56	86	57	63
83	90	55	55	75	98	53	40	41	58	85

Dari data tersebut.

a. Buat tabel distribusi datanya
b. Tambilkan data dalam bentuk
   a. Grafik batang
   b. Grafik lingkaran

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Ukuran yang mewakili sekumpulan data tersebut dikenal dengan nama ukuran pemusatan data. Bentuk ukuran pemusatan yang umum digunakan : Rata-rata hitung (mean), Median dan Modus. Letak data dapat juga dinyatakan dalam kuartil, desil atau persentil.

Kompetensi

- Mahasiswa mampu memahami cara perhitungan rata-rata, median, modus, kwartil desil dan persentil.
- Mahasiswa mampu memahami perbedaan ukuran-ukuran yang ada pada ukuran pemusatan data

UKURAN PEMUSATAN DATA

Saat paparan dari sekumpulan data numerik hasil pengamatan, adalah umum untuk menunjukan suatu nilai yang mewakili sekumpulan data hasil pengamatan. Seperti ukuran yang menggambarkan kira-kira dimana letak data atau dimana data "berpusat" di sepanjang garis bilangan. Ukuran yang mewakili sekumpulan data tersebut dikenal dengan nama ukuran pemusatan data. Contoh, catatan jumlah penumpang kereta yang dihimpun sepanjang tahun 2011 diberikan seperti pada tabel berikut:

Jumlah Penumpang Kereta Api, 2011(000 Orang)

Bulan	Jawa			Sumatera	Total
	Jabotabek	Non Jabotabek	Jabotabek + Non Jabotabek		
(1)	(2)	(3)	(4)	(5)	(6)
Januari	10,354	6,092	16,446	445	16,891
Februari	9,270	5,249	14,519	371	14,89
Maret	10,733	5,851	16,584	394	16,978
April	10,188	5,843	16,031	410	16,441
Mei	10,513	6,505	17,018	504	17,522
Juni	10,147	6,659	16,806	459	17,265
Juli	10,749	6,883	17,632	500	18,132
Agustus	9,678	4,814	14,492	354	14,846
September	9,692	6,661	16,353	568	16,921
Oktober	10,152	5,910	16,062	399	16,461

November	9,852	5,913	15,765	414	16,179
Desember	9,777	6,556	16,333	478	16,811

Sumber : PT Kereta Api Indonesia

Andaikan paparan dari data diatas kita fokuskan pada kolom ke 2 , kita akan dapatkan nilai/ukuran pemusatan yang mewakili data yaitu:

- Rata-rata penumpang per bulan = 10092.
- Median = 10150.

Ukuran Pemusatan Data

Ukuran pemusatan adalah nilai tunggal yang mewakili suatu kumpulan data dan menunjukkan karakteristik dari data. Ukuran pemusatan menunjukkan pusat dari nilai data.

Bentuk ukuran pemusatan yang umum digunakan :

- Rata-rata hitung (mean)
- Median
- Modus

- Rata-rata hitung

  Jika

  x= variabel sampel data

  x_i = variabel sampel data ke i

  n = jumlah sampel

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

  μ = rata-rata hitung populasi

  N = jumlah populasi

  maka,

  Rata-Rata Hitung populasi dan sampel dirumuskan sebagai berikut.

  - Data non group

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \quad \text{dan} \quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Contoh :

1. Misalkan diketahui di kota "Industri" terdapat 8 pabrik masing-masing tercatat mempunyai banyak pekerja sebagai berikut :
   1850, 700, 3150, 2250, 1750, 1900,600,680.
2. Setiap 12 jam sekali bagian QC pabrik minuman ringan memeriksa 6 kaleng contoh untuk diperiksa kadar gula sintetisnya (%). Berikut adalah data 6 kaleng minuman contoh yang diperiksa :
   13.5 12.5 13 12 11.5 12.5

Soal pertama menyatakan populasi, perhitungan rata-rata adalah sebagai berikut

$$\mu = \frac{\sum_1^8 x_i}{8} = \frac{12880}{8} = 1610$$

Soal kedua menyatakan sample, perhitungan rata-rata adalah sebagai berikut

$$\bar{x} = \frac{\sum_1^6 x_i}{6} = \frac{75}{6} = 12.5\ \%$$

- Data group
  Nilainya merupakan **pendekatan**

  Biasanya berhubungan dengan rata-rata hitung sampel

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

sehingga :

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{n}$$

$\underline{x}$: rata-rata hitung sampel

n : ukuran Sampel

f_i : frekuensi di kelas ke-i

x_i : Titik Tengah Kelas ke-i

Kelas	Nilai Tengah Kelas (xi)	Frekuensi (fi)	fi xi
16-23	19.5	10	195
24-31	27.5	17	467.5
32-39	35.5	7	248.5
40-47	43.5	10	435
48-55	51.5	3	154.5
56-63	59.5	3	178.5
Jumlah ( )	dicari	50	1679

Jawab : $\bar{x}$ = $\frac{1679}{50}$ = 33.58

Selain dengan rumus tersebut, dapat dengan suatu nilai dugaan (M)

$$\bar{x} = M + \frac{\sum_{i=1}^{n} f_i d_i}{n}$$

d_i : Nilai Tengah (x_i) - M

Kelas	Nilai Tengah	M	di	Frekuensi(fi)	fi di
16-23	19.5	39.5	-20	10	-200
24-31	27.5	39.5	-12	17	-204
32-39	35.5	39.5	-4	7	-28
40-47	43.5	39.5	4	10	40
48-55	51.5	39.5	12	3	36
56-63	59.5	39.5	20	3	60

Jumlah (□)			0	50	-296

Jawab :

$$\bar{x} = M + \frac{\sum_{i=1}^{n} f_i d_i}{n} = 39.5 + \frac{-295}{50} = 39.5 - 5.92 = 33.58$$

Note : M dapat ditentukan dari Nilai Tengah Kelas (x_i) dengan ketentuan.

- jika banyak kelas (k) ganjil maka ambil (x_i) pada kelas ke $\frac{k}{2}$+1 (kelas yang di tengah-tengah)
- jika banyak kelas (k) genap maka gunakan (x_i) pada kelas ke$\frac{k}{2}$ dan kelas ke $\frac{k}{2}$ +1 selanjutnya kedua nilai (x_i) tersebut dibagi dua

- Median

  Median adalah salah satu ukuran pemusatan yang sering digunakan. Median dari gugus data yang telah diurutkan dari yang terkecil sampai yang terbesar atau dari terbesar sampai terkecil adalah data pengamatan yang tepat di tengah-tengah bila banyaknya pengamatan itu ganjil, atau rata-rata kedua pengamatan yang di tengah bila banyaknya pengamatan genap.

  Contoh :

  Berikut adalah tabel yang telah di urutkan berdasarkan jumlah penumpang

Bulan	Jumlah
Februari	9,273
Agustus	9,678
September	9,692

Desember	9,777
November	9,852
Juni	10,147
Oktober	10,152
April	10,188
Januari	10,354
Mei	10,513
Maret	10,733
Juli	10,749

Median dari data adalah = (10147+10152)/2 = 10149,5

- Modus

  Nilai yang paling sering muncul atau nilai yang frekuensinya paling tinggi

  - Modus untuk Ungrouped Data

    Contoh :

    Hasil test statistik mahasiswa 75 60 75 60 65 75 75 80 80 90 55 75

    Modus = 75
  - Modus untuk Group Data

    Kelas Modus : Kelas di mana Modus berada atau kelas dengan frekuensi tertinggi.

$$\text{TBB kelas ke-i} = \frac{\text{BB kelas ke-i + BA kelas ke (i-1)}}{2}$$

$$\text{TBA kelas ke-i} = \frac{\text{BA kelas ke-i + BB kelas ke (i+1)}}{2}$$

$$\text{Modus} = \text{TBB Kelas Modus} + i\left(\frac{d_1}{d_1 + d_2}\right)$$

Kelas	Frekuensi (fi)
16-23	10
24-31	17
32-39	7
40-47	10
48-55	3
56-63	3
Jumlah (□)	50

- Kelas Modus = 24 - 31
- TBB Kelas Modus = 23.5
- i = 8 h
- frek. kelas Modus = 17
- frek, kelas sebelum kelas Modus = 10 Modus
- frek. kelas sesudah kelas Modus = 7 Modus
- d_1 = 17 - 10 = 7
- d_2 = 17 - 7 = 10

$$\text{Modus} = 23.5 + 8\left(\frac{7}{7+10}\right) = 23.5 + 8\left(\frac{7}{17}\right) = 23.5 + 8\,(0.41176\ldots) = 23.5 + 3.2941\ldots$$

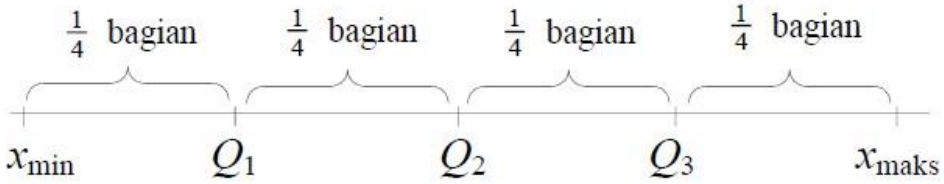
= 26.7941... ≈ 27

Ukuran Letak Data

Ukuran letak adalah ukuran yang menunjukkan pada bagian mana data tsb terletak pada suatu data yang sudah diurutkan. Bentuk ukuran letak yang umum digunakan :

- Kuartil

  Seperti yang sudah dibahas sebelumnya, bahwa median membagi data yang telah diurutkan menjadi dua bagian yang sama banyak. Kuartil membagi data yang telah diurutkan menjadi empat bagian yang sama banyak.



Dimana :

X_{min} = data terkecil

X_{max} = data terbesar

Q_1 = kuartil ke -1

Q_2 = kuartil ke -2

Q_3 = kuartil ke -3

- Quartil data nongroup

  Letak dari *Qi* dirumuskan sebagai berikut :

Q_i =data ke$\left[\frac{i(n+1)}{4}\right]$

Q_i = kuartil ke i

n = banyak data

contoh : Tentukan *Q*1, *Q*2, dan *Q*3 dari data : 3, 4, 7, 8, 7, 4, 8, 4, 9, 10, 8, 3, 7, 12.

Jawab :

Data yang telah diurutkan: 3, 3, 4, 4, 4, 7, 7, 7, 8, 8, 8, 9, 10, 12.

$Q_1 = \frac{1(14+1)}{4} = 3\frac{3}{4}$

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

$Q_2 = \frac{2(14+1)}{4} = 7\frac{1}{2}$

$= 7 + ½ (7-7) = 7$

Dengan cara yang sama didapat Q_3 = 8,25

- Quratil data group

  Menentukan letak kuartil untuk data bergolong, caranya sama dengan data tunggal.Nilai kuartil dirumuskan sebagai berikut.

$$Q_i = b_i + l\left(\frac{\frac{1}{4}N - F}{f}\right)$$

Keterangan: Qi = kuartil ke-i (1, 2, atau 3)

bi = tepi bawah kelas kuartil ke-i

N = banyaknya data

F = frekuensi kumulatif kelas sebelum kelas kuartil

l = lebar kelas

f = frekuensi kelas kuartil

Contoh :

Tentukan kuartil data berikut

Nilai	frekuensi
40 – 49	4
50 – 59	5
60 – 69	14
70 – 79	10
80 – 89	4
90 – 99	3

Penyelesaian

Nilai	frekuensi	F kumulatif
40 – 49	4	4

50 – 59	5	9
60 – 69	14	23
70 – 79	10	33
80 – 89	4	37
90 – 99	3	40

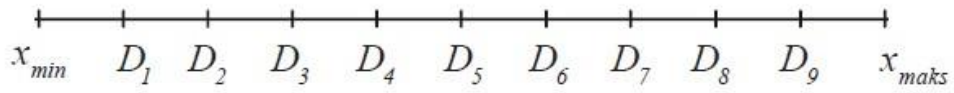
Letak *Q1* pada frekuensi = 1/4 x 40 = 10 ada di kelas 60 – 69.

$Q1 = b_i + l\left(\frac{\frac{1}{4}N - F}{f}\right) = 59.5 + 10\left(\frac{1/4(1.40)-9}{14}\right) = 59,5 + 0.07 = 59,57$

Dengan cara yang sama didapat $Q_2 = 67,36$ dan $Q_2 = 76,5$

- Desil

Jika median membagi data menjadi dua bagian dan kuartil membagi data menjadi empat bagian yang sama, maka desil membagi data menjadi sepuluh bagian yang sama besar.



Letak D_i di urutan data ke - $\frac{i(n+1)}{10}$

Di = Desil ke-*i*

i = 1, 2, 3, . . ., 99

n = banyaknya data

Contoh :

Tentukan desil ke 2 dan ke 4 data berikut

9, 10, 11, 6, 8, 7, 7, 5, 4, 5

Penyelesaian

4, 5, 5, 6, 7, 7, 8, 9, 10, 11 (data setelah diurutkan)

Letak desil ke -2 $\frac{2(10+1)}{10} = \frac{22}{10} = 2,2$

$D_2 = x_2 + 0,2(x_3 - x_2)$

$= 5 + 0,2(5-5) = 5$

Dengan cara yang sama didapat desil 4 $D_4 = 6,4$

- Persentil.

Jika data dibagi menjadi 100 bagian yang sama, maka ukuran itu disebut persentil.Letak persentil dirumuskan dengan.

Letak P_i di urutan data ke - $\frac{i(n+1)}{100}$

Pi = persentil ke-*i*

i = 1, 2, 3, . . ., 99

n = banyaknya data

Contoh :

Dengan menggunakan data pada contoh desil

Letak persentil ke -30 $\frac{30(10+1)}{100} = \frac{330}{100} = 3,3$

$P_{30} = x_3 + 0,3(x_4 - x_3) = 5 + 0,3(6-5) = 5,3$

Dengan cara yang sama didapat $P_{75} = 9,25$

Soal

1. Berdasarkan tabel “Penumpang Kereta 2011” non jabotabek (kolom 3). Tentukan.
   a. Rata-rata penumpang
   b. Mendian
   c. Modus
   d. Kuartil 1 dan 2
   e. Desil 3
   f. Persentil 60
2. Buat tebel distribusi “ Penumpang Kereta 2011” non jobtabek (kolom 3).
   Tentukan :
   a) Rata-rata penumpang
   b) Mendian
   c) Modus
   d) Kuartil 1 dan 2
   e) Desil 3
   f) Persentil 60

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Ukuran yang digunakan untuk mengetahui variasi atau dispersi data, yaitu derajat penyebaran data terhadap nilai rata-rata. Ukuran penyebaran data adalah range, rata-rata deviasi, innterquartile, dan standar deviasi. Skewness adalah derajat ketidaksimetrisan suatu distribusi.

Kompetensi

- Mahasiswa mampu memahami cara perhitungan ukuran disperse absolute dan disperse relatif.
- Mahasiswa mampu memahami perbedaan bentuk skrweness dan kurtosis

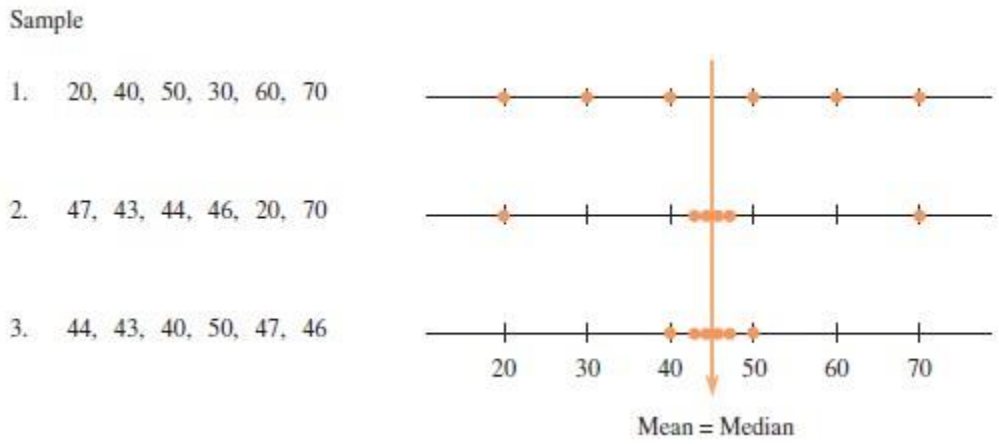
Kurtosis adalah derajat keruncingan suatu distribusi (biasanya diukur relatif terhadap distribusi normal).

UKURAN PENYEBARAN DATA

Rata-rata hitung (mean), median,modus adalah ukuran pemusatan data yang memberikan informasi tentang bagaimana data-data ini mengumpul atau memusat. Selain memusat, data juga tersebar disekitar ukuran pemusataanya. Sebagaimana ukuran pada pemusatan data, penyebaran data juga memiliki ukuran. Ukuran ini digunakan untuk mengetahui variasi atau dispersi data, yaitu derajat penyebaran data terhadap nilai rata-rata. Ukuran penyebaran data yang sering digunakan adalah range, rata-rata deviasi, innterquartile, dan standar deviasi.

A. Ukuran Penyebaran Data

Selain ukuran pemusatan data , statistika masih memiliki ukuran lain yaitu ukuran penyimpangan atau ukuran variasi atau ukuran penyebaran(disversi) data . Ukuran dispersi adalah ukuran yang menyatakan seberapa banyak nilai-nilai data yang berbeda dengan nilai pusatnya atau seberapa jauh penyimpangan nilai-nilai data tersebut dari nilai pusat. Seperti yang ditunjukan pada gambar berikut :



Ada banyak variabilitas dalam sampel pertama dibandingkan dengan sampel ketiga. Sampel kedua menunjukkan variabilitas kurang dari variabilitas pertama dan lebih dari yang ketiga, sebagian besar variabilitas dalam sampel kedua ini disebabkan oleh dua nilai ekstrim yang begitu jauh dari pusat.Ukuran ini juga menggambarkan derajat berpencarnya data kuantitatif.

Ukuran variasi pada dasarnya merupakan pelengkap dari ukuran nilai pusat dalam rangka penggambaran sekumpulan data, karena ukuran nilai pusat secara terpisah tidaklah dapat menggambarkan keadaan keseluruhan data dengan baik. Ukuran nilai pusat tersebut hanya memberikan informasi tentang sebuah nilai dimana nilai-nilai data yang lain berpencar atau dengan kata lain setengah dari keseluruhan data berada sebelum nilai pusat dan setengahnya lagi berada setelah nilai pusat.Berikut adalah beberapa fungsi atau kegunaan ukuran dispersi :

a. Ukuran penyebaran dapat digunakan untuk menentukan apakah nilai rata-ratanya benar-benar representatif atau tidak. Apabila suatu kelompok data mempunyai penyebaran yang tidak sama terhadap nilai rata-ratanya, maka dikatakan bahwa nilai rata-rata tersebut tidak representatif.
b. Ukuran penyebaran dapat digunakan untuk mengadakan perbandingan terhadap variabilitas data.
c. Ukuran penyebaran dapat membantu penggunaan ukuran statistika misal dalam pengujian hipotesis untuk menentukan apakah dua sampel berasal dari populasi yang sama atau tidak.

Jenis ukuran dispersi meliputi jangkauan (range) yang terdiri dari jangkauan antar kuartil dan interkuartil, simpangan rata-rata, simpangan baku, varians, koefisien variasi, dan angka baku.

- Ukuran Jangkauan (Range)

  Range merupakan ukuran variasi yang paling sederhana dan paling mudah ditentukan nilainya. Range R merupakan selisih nilai tertinggi Xmax data observasi dengan data terendahnyaXmin dan dirumuskan sebagai berikut.

  $$R = X_{max} - X_{min}$$

  Contoh :

Jumlah Penumpang Kereta Tahun 2011(dalam 1000)					
9273	9678	9692	9777	9852	10147
10152	10188	10354	10513	10733	10749

Sumber : PT Kereta Api Indonesia

Range R = 10749000 -9273000 = 1476000

Jika nilai-nilai observasi telah dikelompokkan ke dalam distribusi frekuensi, maka jangkauan distribusi dirumuskan sebagai beda antara pengukuran nilai titik tengah kelas pertama dan nilai titik tengah kelas terakhir.

Contoh :

Interval	Frekuensi
9273-9597	1
9597-9921	4
9922-10246	3
10247-10571	2
10572-10896	2

Titik tengah kelas pertama = (9273+9597)/2 = 9435

Titik tengah kelas terahir = (10572+10896) =10734

Range R = 10734 – 9435 = 1299

- Rata-rata simpangan

  Ukuran variabelitas yang juga banyak digunakan untuk mendeskripsikan sejauh mana sampel pengamatan menyimpang dari rata-rata sampel x adalah rata-rata penyimpangan dari mean atau rata-rata simpangan. Rata-rata simpangan untuk data tunggal dirumuskan sebagai

$$S_x = \frac{\sum_1^n |x_i - \bar{x}|}{n}$$

Untuk data kelompok dirumuskan sebagai

$$S_x = \frac{\sum_1^n f_i|x_i - \bar{x}|}{n}$$

Contoh :

Tentukan rata-rata simpangan data berikut :

6092	5249	5851	5843	6505	6659	6883	4814	6661	591 1	5913	6556

Rata-rata = 5635

x_i	$\lvert x_i - x \rvert$
6092	457
5249	-386
5851	216
5843	208
6505	870
6659	1024
6883	1248
4814	-821
6661	1026
5910	275
5913	278
6556	921
Σ	5316

$$S_x = \frac{\sum_1^n |x_i - \bar{x}|}{n} = \frac{\sum_1^{12} |x_i - \bar{x}|}{12}$$

$$= \frac{5316}{12} = 443$$

Contoh :

x_i	Frekuensi	$f_i x_i$	$\lvert x_i - x \rvert$	$f_i \lvert x_i - x \rvert$
55	1	55	-20,56	-20,56
60	4	240	-15,56	-62,22
65	4	260	-10,56	-42,22
70	6	420	-5,56	-33,33
75	5	375	-0,56	-2,78
80	3	240	4,44	13,33
85	3	255	9,44	28,33

90	2	180	14,44	28,89
100	1	100	24,44	24,44
	∑=29			∑= -66,12

Rata-rata = 75,56

$S_x = \frac{\sum_1^n f_i|x_i - \bar{x}|}{n} = \frac{-66,12}{29} = -2,28$

- Simpangan baku (deviasi standar)

  Untuk populasi yang berjumlah besar, sangat tidak mungkin untuk mendapatkan nilai rata-rata populasi μ serta deviasi standartny σ. Untuk mengestimasi (menaksir) nilai μ dan σ , diambil sampel data. Nilai μ diestimasi oleh $\bar{X}$ dan σ diestimasi oleh S .

  - Rumus Deviasi Populasi

    $$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

    $$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}\ .$$

    dengan :

    N = Jumlah observasi dalam populasi

    μ = Rata-rata populasi.

  - Sampel Deviasi Sampel

    Simpangan baku atau deviasi standar (*Standard Deviation*) merupakan ukuran penyebaran yang paling baik, karena menggambarkan besarnya penyebaran tiap-tiap unit observasi. Karl Pearson menamakannya deviasi standar dan dirumuskan sebagai :

    $$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2}\ .$$

Kuadrat dari deviasi standar dinamakan ***variansi*** :

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2 .$$

Untuk distribusi sampel dengan n < 100, Fisher, Wilks dan beberapa statistisi memberi perumusan tentang variansi dan deviasi standar sebagai berikut :

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2}$$

Contoh

Jumlah Penumpang Kereta Api, 2011(000 Orang)

Bulan	Non Jabotabek		
Januari	6092	14	196
Februari	5249	-829	687241
Maret	5851	-227	51529
April	5843	-235	55225
Mei	6505	427	182329
Juni	6659	581	337561
Juli	6883	805	648025
Agustus	4814	-1264	1597696
September	6661	583	339889
Oktober	5910	-168	28224
November	5913	-165	27225
Desember	6556	478	228484
	$x_i - \bar{x}$		
Rata-rata	6078		
Σ	4183624		

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2$$

$$= \frac{1}{12-1} 4183624 = \frac{4183624}{11} = 380329{,}5$$

$$S = \sqrt{\frac{1}{12-1} 4183624} = \sqrt{\frac{4183624}{11}} = 616{,}7086$$

- Pengukuran dispersi relatif

  Pengukuran jangkauan, deviasi kuartil, deviasi rata-rata dan deviasi standar merupakan pengukuran yang absolut. Pengukuran demikian itu sebetulnya hanya dapat digunakan bagi penggambaran dispersi nilai-nilai observasi sebuah distribusi secara definitive. Bila kita ingin melakukan perbandingan tingkat dispersi antara dua atau beberapa distribusi dan bila jumlah nilai-nilai observasi dari dua atau beberapa distribusi di atas tidak sama, maka pengukuran digunakan dispersi secara absolut sebagai metode guna membandingkan dispersi akan memperoleh hasil yang menyesatkan.

  Dalam membandingkan tingkat variasi dua atau lebih distribusi hendaknya rata-rata distribusi digunakan sebagai dasar pengukuran variasinya secara relatif dan dinamakan **ko-efisien variasi** (*co-efficient of variation*) :

  $$V = \frac{S}{\bar{X}}$$

  dimana

  S = deviasi standar sampel

  $\bar{X}$ = rata-rata hitung sampel

  Contoh :

  Tentukan koefisien variansi tabel berikut

Interval Kelas	*fi*
97-103	4
104-110	8
11-117	15
118-124	35
125-131	25
132-138	6

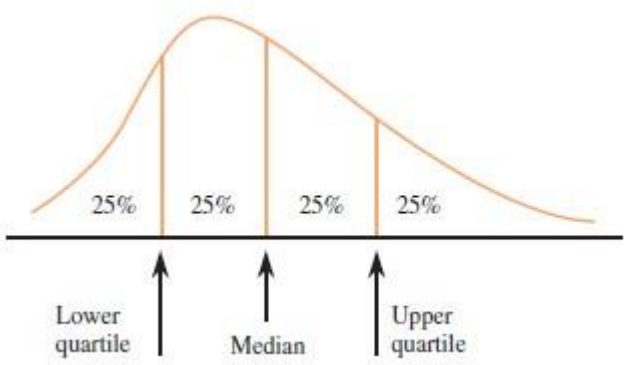
139-145	4
146-152	3
	100

Dari hasil perhitungan diperoleh :
$\overline{X} = 122{,}26$ dan $S_S = 10{,}21$.

Jadi, $V = \frac{10{,}21}{122{,}26}$ x 100% = 8,35%.

- Range Interquartil

  Median didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi dua bagian yang sama dan kuartil didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi empat bagian yang sama. Range interkuartil adalah ukuran variabilitas berdasarkan kuartil. Perhatikan gambar berikut.



Pada gambar di atas terlihat kuartil bawah memisahkan 25% kumpulan data kebawah dan kuartil atas memisahkan 25% dari kumpulan data ke atas. Kuartil tengah adalah median dan memisahkan 50% kumpulan data. Jika jumlah data n ganjil maka median bisa digantikan dengan nilai kuartil 2 Q_2.

Pada distribusi kuartil, 50% dari semua nilai observasi seharusnya terletak di antara Q_1 dan Q_3 . Jangkauan antara Q_1 dan Q_3 dinamakan Range Interquartil (*inter-quartile-range*) dan dirumuskan sebagai

Iqr= kuartir atas – kuartir bawah

atau

$$Iqr = Q_3 - Q_1$$

Pengukuran dispersi atas dasar jangkauan inter-kuartil dinamakan deviasi *kuartil* atau *simpangan kuartil* (*quartile deviation*) dan diruuskan sebagai.

$$D_{iqr} = \frac{Kuartir\ atas - kuartir\ bawah}{2}$$

atau

$$D_{iqr} = \frac{Q_3 - Q_1}{2}$$

Contoh :

Berikut adalah tabel produksi pulsa telpon di indonesia

Produksi Pulsa Telepon Indonesia 1998 - 2008

Tahun	Lokal (000 pulsa)	SLJJ (Menit)
1998	16236246427	29668416066
1999	16236724396	31021632143
2000	18516778571	34342636
2001	20227877123	38161484336
2002	19730308403	41397291119
2003	23887950222	42447349726
2004	19936304184	45215914717
2005	22920220767	57745329624
2006	23646924115	61443360381
2007	29018054840	53129188172
2008	22233240642	40706864477

Sumber : Kantor Pusat PT. TELKOM Indonesia

Median = 23887950222

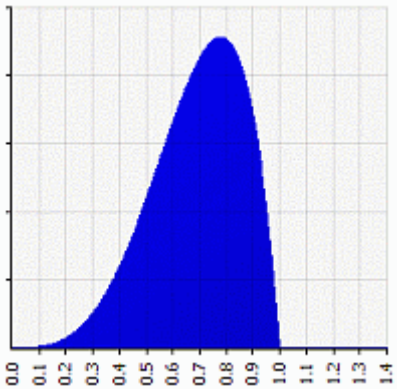
Kuartil bawah = 18516778571

Kuartil atas = 23646924115

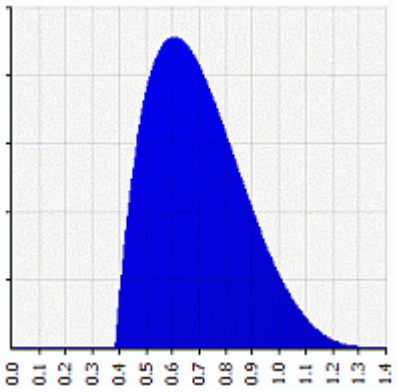
$D_{iqr} = \frac{Q_3 - Q_1}{2}$ = (23646924115- 18516778571)/2 = 2565072772

B. Skewnes dan Kutosis

Skewness adalah derajat ketidaksimetrisan suatu distribusi. Jika kurva frekuensi suatu distribusi memiliki ekor yang lebih memanjang ke kanan (dilihat dari meannya) maka dikatakan menceng kanan (positif) dan jika sebaliknya maka menceng kiri (negatif).Secara perhitungan, skewness adalah momen ketiga terhadap mean. Distribusi normal (dan distribusi simetris lainnya, misalnya distribusi t atau Cauchy) memiliki skewness 0 (nol). Perhatikan gambar berikut . Kedua gambar memiliki μ = 0.6923 and σ = 0.1685 yang sama tetapi keduanya memiliki kemencengan yang berbeda.



Beta(α=4.5, β=2)
skewness = −0.5370



1.3846 − Beta(α=4.5, β=2)
skewness = +0.5370

Rumus koefisien skewnes

$g_1 = m_3 / m_2^{3/2}$

dimana :

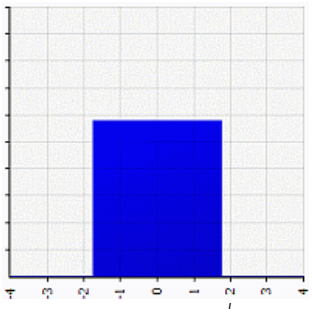
$m_2 = \sum(x-x)^2 / n$ disebut varian

$m_3 = \sum(x-x)^3 / n$ momem ke tiga

skewnes sampel data

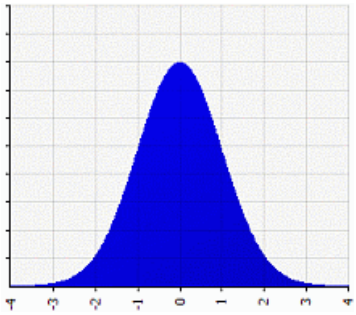
$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1$$

Kurtosis adalah derajat keruncingan suatu distribusi (biasanya diukur relatif terhadap distribusi normal). Kurva yang lebih lebih runcing dari distribusi normal dinamakan leptokurtik, yang lebih datar platikurtik dan distribusi normal disebut mesokurtik. Kurtosis dihitung dari momen keempat terhadap mean. Distribusi normal memiliki kurtosis = 3, sementara distribusi yang leptokurtik biasanya kurtosisnya >3.Visualisasi kurtosis dapat dilihat pada gambar berikut.



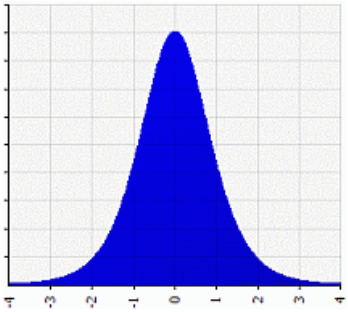
Uniform(min=−√3, max=√3)

kurtosis = 1.8, excess = −1.2



Normal(μ=0, σ=1)

kurtosis = 3, excess = 0



Logistic(α=0, β=0.55153)

kurtosis = 4.2, excess = 1.2

Rumus kurtois

- kurtosis: $a_4 = m_4 / m_2^2$
- excess kurtosis: $g_2 = a_4 - 3$

dimana :

$m_2 = \sum(x-x)^2 / n$ disebut varian

$m_4 = \sum(x-x)^4 / n$ momem ke empat

skewnes sampel data

$$G_2 = \frac{n-1}{(n-2)(n-3)}[(n+1)g_2 + 6]$$

Contoh : Diketahui data tinggi mahasiswa dalam berikut

Tinggi (inches)	Nilai Tengah x	frekuensi f
59.5–62.5	61	5
62.5–65.5	64	18

65.5–68.5	67	42
68.5–71.5	70	27
71.5–74.5	73	8

Tentukan kurtois dan excess kurtoisnya.

Nilai tengah	frekuensi f	x−x	$(x-x)^4 f$
61	5	-6.45	8653.84
64	18	-3.45	2550.05
67	42	-0.45	1.72
70	27	2.55	1141.63
73	8	5.55	7590.35
		∑	19937.60
		m_4	1.993.760

Kurtois $a_4 = m_4 / m_2^2 = 199.3760/8.5275^2 = 2.7418$

Excess kurtosis $g_2 = 2.7418-3 = -0.2582$

Karena data yang adalah sampel maka

excess kurtosis $G_2 = [99/(98\times97)]\,[101\times(-0.2582)+6)] = -0.2091$

Soal

1. Hasil survey tingkat pendidikan penduduk yang lulus sekolah menengah atas diberikan sebagai berikut.

21	27	26	19	30	35	35	26	47	26	27	30
24	29	22	24	29	20	20	27	35	38	25	31
19	24	27	27	23	34	34	25	32	26	26	24
22	28	26	30	23	25	22	25	29	33	34	30
17	25	23									

Dari data diatas

Tentukan

1. Ukuran jangkauan penyebaran data
2. Rata-rata simpangan
3. deviasi standar
4. Variansi
5. Simpangan kuartil
6. Skewnes dan kutoisnya.

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Angka indeks adalah ukuran statistik yang dirancang untuk mengekspresikan perubahan atau perbedaan variabel atau sekelompok variabel terkait. Tujuan pembuatan angka indeks adalah untuk mengukur secara kuantitatif terjadinya perubahan dalam dua waktu yang berlainan. Jenis angka indeks

Kompetensi

- Mahasiswa mampu memahami pengertian angka indeks dan jenis-jenis perhitungannya.

tergantung dari jenis variabel yang diperbandingkan.

ANGKA INDEX

A. Pendahuluan

Dalam kehidupan sehari-hari sering kita mendapatkan perbedaan/perubahan nilai dari komoditas tertentu.Mungkin kita ingin tahu bagaimana perubahan nilai pada suatu periode dan mengekpersikan dalam bentuk nilai. Untuk mengekpresikan nilai perubahan tersebut digunakan angka index. Angka indeks adalah ukuran statistik yang dirancang untuk mengekspresikan perubahan atau perbedaan variabel atau sekelompok variabel terkait.Dalam mengekpresikan suatu nilai umumnya digunakan bentuk persentase.

Ciri khas dari angka indeks ini adalah perhitungan rasio (pembagian), dimana hasil rasio tersebut selalu dikalikan dengan bilangan 100 untuk menunjukkan perubahan tersebut dalam persentase. Dengan demikian, basis dari angka indeks apapun selalu 100, (Singgih Santoso, 2003). Angka indeks adalah angka yang digunakan sebagai perangkat untuk perbandingan antara kuantitas, harga atau nilai dalam situasi yang berbeda, misalnya waktu dan di tempat tertentu.Jika perbandingan adalah dalam hal harga maka disebut indeks harga untuk kuantitas fisik disebut angka indeks kuantitas. Angka indeks lainnya didefinisikan dengan cara serupa. Angka indeks atau sering disebut indeks saja, pada dasarnya merupakan suatu angka yang dibuat sedemikian rupa sehingga dapat dipergunakan untuk melakukan perbandingan antara kegiatan yang sama (produksi, ekspor, hasil penjualan, jumlah uang beredar, dan lain sebagainya) dalam waktu yang berbeda (J. Supranto, 1990).

Contoh :

Perhitungan Angka Indeks Penjualan Laptop
Tahun 2008 - 2011 (dalam miliar rupiah)

Tahun	Jumlah Penjualan	Angka Indeks
2008	145	100%
2009	155	(145/155)*100%

2010	150	(150/145)*100%
2011	160	(160/145)*100%

Sumber : Data Buatan

Tujuan pembuatan angka indeks adalah untuk mengukur secara kuantitatif terjadinya perubahan dalam dua waktu yang berlainan. Misalnya indeks harga untuk mengukur perubahan harga (berapa kenaikannya atau penurunannya), indeks produksi untuk mengetahui perubahan yang terjadi dalam kegiatan produksi, indeks biaya hidup untuk mengukur tingkat inflasi, dll.

B. Jenis Angka Index

Angka indeks adalah angka yang digunakan dalam merepsentasikan perbandingan suatu variabel, jenis angka indeks tergantung dari jenis variabel yang diperbandingkan. Berikut adalah beberapa jenis angka index.

- Index harga konsumen

  Indeks Harga Konsumen (IHK) merupakan indeks yang memperhatikan harga-harga yang harus dibayar konsumen baik di perkotaan maupun pedesaan, (Suharyadi, Purwanto S.K, 2003). IHK mengukur rata-rata perubahan harga dari suatu paket komoditas yang dikonsumsi oleh masyarakat/rumah tangga di suatu daerah (urban) dalam kurun waktu tertentu. Persentase perubahan Indeks Harga Konsumen (IHK) bisa bernilai positif atau negatif. Bila persentase perubahan IHK positif dapat dikatakan terjadi inflasi (kenaikan harga eceran secara umum) dan sebaliknya bila persentase perubahan IHK bernilai negatif berarti terjadi deflasi (penurunan harga secara umum).

- Index Produktivitas

  Produktivitas merupakan rasio antara output atau produksi dengan input. Produktivitas input bisa mencerminkan jenisnya seperti produktivitas tenaga kerja, produktivitas modal dan produktivitas mesin. Namun demikian pada saat teknologi berkembang, sumbangan input sudah tidak dapat dipisahkan, maka sebutan produktivitas diarahkan pada produktivitas total. Apabila indeks lebih dari 100, menunjukkan bahwa produktivitas lebih baik dari tahun dasar

- Index Nilai Tukar Petani

  Nilai Tukar Petani (NTP) adalah rasio antara indeks harga yang diterima petani (IT) dengan indeks harga yang dibayar petani (IB) yang dinyatakan dalam persentase. Secara konsepsional NTP adalah pengukur kemampuan tukar

barang-barang (produk) pertanian yang dihasilkan petani dengan barang atau jasa yang diperlukan untuk konsumsi rumah tangga dan keperluan dalam memproduksi produk pertanian. Indeks harga yang diterima petani (IT) adalah indeks harga yang menunjukkan perkembangan harga produsen atas hasil produksi petani. Indeks harga yang dibayar petani (IB) adalah indeks harga yang menunjukkan perkembangan harga kebutuhan rumah tangga petani, baik kebutuhan untuk konsumsi rumah tangga maupun kebutuhan untuk proses produksi pertanian.

C. Cara Penentuan Angka Index

Dibidang ekonomi ada 3 jenis angka index yang umum digunakan.

1. Indeks Harga (*Price Index*)
   Menunjukkan perubahan harga dari satu periode ke periode lain.
2. Indeks Kuantitas (*Quantity Index*)
   Menunjukkan perubahan kuantitas (misalnya volume penjualan, jumlah produksi, dsb.) dari satu periode ke periode lain.
3. Indeks Nilai (*Value Index*)
   Menunjukkan perubahan nilai uang dari satu periode ke periode lain. Nilai ini dapat diperoleh dari hasil kali antara antara harga dan kuantitas .

Indeks harga termasuk yang paling sering digunakan dalam kehidupan sehari-hari,karena secara langsung mencerminkan pergerakan harga berbagai jenis barang. Dalam pengukurannya indeks harga bisa dicari dengan metode tak tertimbang, metode tertimbang, metode relatif.

- Metode tak terimbang

  Metode tak tertimbang adalah metode perhitungan index berdasarkan nilai uang barang tertentu.Indeks tidak berimbang dalam pembuatannya tidak memasukkan faktor yang mempengaruhi naik-turunnya angka indeks. Metode-metode yang ada dalam kelompok ini adalah :

  a. Metode Angka Relatif

b. Metode Agregate
c. Metode Rata-Rata Relatif

- Metode tertimbang

Metode tertimbang merupakan metode perhitungan harga index yang didasarkan atas satuan barang tersebut, misal berat barang (kilogram), panjang (meter) dan seterusnya.Indeks tertimbang memasukkan faktor yang mempengaruhi naik-turunnya angka indeks. Metode-metode yang ada dalam kelompok ini adalah :

a. Metode Laspeyres
b. Metode Paasche
c. Metode Drobisch
d. Metode Irving Fisher
e. Metode Marshall – Edgeworth

Pada metode tertimbang atau tak tertimbang, proses perhitungan dimulai dengan menjumlah seluruh komponen yang ada kemudian dilakukan rata-rata. Sedangkan metode relatif memulai dengan menghitung setiap indeks komponen, kemudian baru melakukan rata-rata dari semua indeks individe yang didapat.

Pada metode indeks harga tertimbang ada bobot yang digunakan untuk membedakan variabel yang satu dengan yang lain. Seperti adanya penimbang yang berupa kuantitas barang yang terjual untuk berbagai jenis barang yang berlainan harganya. Pembobobotan ini dilakukan karena pada dasarnya setiap barang dan jasa mempunyai tingkat utilitas (manfaat dan kepentingan) yang berbeda. Misal, beras mungkin dirasakan lebih penting dibandingkan jenis barang yang lain. Indeks harga tertimbang biasa digunakan untuk indeks agregat di mana ada sekian banyak jenis komoditi.

D. Perhitungan Angka Index

A. Indeks harga relatif sederhana (*simple relative price index*) ialah indeks yang terdiri dari satu macam barang saja, baik untuk indeks produksi maupun indeks harga (misalnya indeks produksi beras, indeks produksi karet, indeks produksi ikan, indeks

harga beras, indeks harga karet, indeks harga ikan, dsb). Perhitungan angka index digunakan rumus berikut.

$$IK = \frac{Qn}{Qo} x100$$

dimana Qn nilai tahun yang ditinjau dan Qo nilai tahun dasar

Contoh :

Bulan	Harga	Kuantitas	Indeks	
			Harga	Kuantitas
Januari	3500	50	100	100
Februari	3800	52	109	104
Maret	3400	56	97	112
April	4000	49	114	98
Mei	4200	51	120	102
Juni	3900	48	111	96

- Indeks harga bulan Februari dengan waktu dasar bulan Januari

$$\frac{3800}{3500} \times 100\% = 109$$

- Indeks kuantitas bulan Februari dengan waktu dasar bulan Januari

$$\frac{52}{50} \times 100\% = 104$$

B. Index Agegate

Indeks agregatif merupakan indeks yang terdiri dari beberapa barang (kelompok barang), misalnya indeks harga 9 macam bahan pokok, indeks impor Indonesia, indeks ekspor Indonesia, indeks harga bahan makanan, indeks biaya hidup, indeks hasil penjualan suatu perusahaan (lebih dari satu barang yang dijual), dll. Angk index dihitung dengan rumus :

$$IA = \frac{\sum Pn}{\sum Po} x100$$

dimana Pn periode tahun yang ditinjau, dan Po tahun dasar perhitungan.

Contoh :

Tentukan Index agregatif data berikut.

Nama Komoditi	Po	Pn
Beras	4500	5500
Minyak	7800	8500
Gula	6500	7500
Tepung	3600	4500
Jumlah	22400	26000

∑ harga tahun dasar Po = 2240

∑harga tahun Pn = 26000

$$IA = \frac{\sum Pn}{\sum Po} x100 = \frac{22400}{26000} \text{ x } 100\ \% = 8{,}6\ \%$$

C. Index Relatif Rata-rata

Untuk menghitung Index Relatif rata-rata digunakan rumus beriku :

$$IRH = \frac{\sum \frac{Pn}{Po}}{n} x100$$

dimana Pn periode tahun yang ditinjau, dan Po tahun dasar perhitungan.

Contoh :

Nama Komoditi	**Po**	**Pn**	**Pn/Po**
Beras	4500	5500	1,22
Minyak	7800	8500	1,09
Gula	6500	7500	1,15

Tepung	3600	4500	1,25
Jumlah			4,72

$$IRH = \frac{\sum \frac{Pn}{Po}}{n} x100 = \frac{4,72}{4} \text{ x } 100\ \% = 117,8\ \%$$

D. Metode Laspeyers

Perhitungan angkan index menggunakan rumus berikut

$$IL = \frac{\sum QoPn}{\sum QoPo} x100$$

Komoditas	Po	Pn	Qo	PnQo	Po Qo
Cengkeh	4959	6437	184	1184408	9124560
Kopi	5600	14595	994	14507430	14812588
Teh	7400	7900	24	189600	641424
Lada	17252	21595	420	9069900	7245840
Tembakau	34322	34650	540	18711000	18533880
			Jumlah	43662338	50358292

$$IL = \frac{\sum QoPn}{\sum QoPo} x100$$ = (43662338/50358292)*100% =100,9%

E. Indeks Drobisch

Perhitungan menggunakan rumus berikut

$$IP = \frac{\sum QnPn}{\sum QoPn} x100$$

Komoditas	Qo	Pn	Qn	PnQo	Pn Qn
Cengkeh	4959	6437	184	1184408	9124560
Kopi	5600	14595	994	14507430	14812588

Teh	7400	7900	24	189600	641424
Lada	17252	21595	420	9069900	7245840
Tembakau	34322	34650	540	18711000	18533880
			Jumlah	43662338	50358292

$$IP = \frac{\sum QnPn}{\sum QoPn} x100$$ = (50358292/43662338)*100% =115,3%

F. Index Fisher

Untuk menghitung angkan index menggunakan cara index fisher digunakan rumus berikut.

$$IF = \sqrt{ILxIP}$$

Lihat cara perhitungan IL dan IP diatas

G. Marshall'-Edgeworth

Untuk menghitung angkan index menggunakan cara index fisher digunakan rumus berikut.

$$IME = \frac{\sum Qn(Po+Pn)}{\sum Qo(Po+Pn)} x100$$

Komoditas	Po	Pn	Qo	Qn	Po +Pn	Qn(Po+Pn)	Qo(Po+Pn)
Cengkeh	4959	6437	4959	184	9124560	1678919040	45.248.693.040
Kopi	5600	14595	5600	994	14812588	14723712472	82.950.492.800
Teh	7400	7900	7400	24	641424	15394176	4.746.537.600
Lada	17252	21595	17252	420	7245840	3043252800	125.005.231.680
Tembakau	34322	34650	34322	540	18533880	10008295200	636.119.829.360
					Σ	29469573688	894.070.784.480

$$IME = \frac{\sum Qn(Po+Pn)}{\sum Qo(Po+Pn)} x100 = \frac{2946957368\,8}{8940707844\,80} x100\% = 3\%$$

Soal :

1. Berikut ini adalah harga beberapa bahan kebutuhan pokok di Jawa Tengah :

No	Bahan Pokok	2000	2001
1	Beras	Rp 4.500	Rp 5.000
2	Minyak Goreng	Rp 9.500	Rp 11.000
3	Bawang Merah	Rp 4.000	Rp 4.300
4	Bawang Putih	Rp 3.000	Rp 4.000
5	Cabe Rawit	Rp 7.000	Rp 7.500
6	Sawi	Rp 2.200	Rp 3.400
7	Ketang	Rp 4.500	Rp 5.000

   Hitunglah indeks harga untuk ke tujuh jenis bahan pokok diatas dengan menggunakan metode indeks harga tidak tertimbang agregatif dan rata-rata relatif.

2. Berikut ini adalah harga beberapa bahan kebutuhan pokok di Jawa Tengah :

No	Bahan Pokok	Harga		Kuantitas	
		2000	2001	2000	2001
1	Beras	Rp 4.500	Rp 5.000	350	500
2	Minyak Goreng	Rp 9.500	Rp 11.000	250	500
3	Bawang Merah	Rp 4.000	Rp 4.300	300	450
4	Bawang Putih	Rp 3.000	Rp 4.000	230	400
5	Cabe Rawit	Rp 7.000	Rp 7.500	140	230
6	Sawi	Rp 2.200	Rp 3.400	300	430
7	Ketang	Rp 4.500	Rp 5.000	230	300

   Bedasarkan soal di atas maka hitunglah indeks harga dengan menggunakan metode

   a. indeks harga tertimbang
   b. dengan cara Laspeyres,
   c. Cara Paasche,
   d. Cara Drobisch,
   e. Cara Fisher,

f. Cara Marshall'-Edgeworth!

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Konsep analisis regresi adalah mencari hubungan terbaik antara y dan x, mengukur kekuatan hubungannya dan metode untuk memprediksi respon nilai-nilai Y dari nilai-nilai x yang diberikan. Analisa estimasi garis regresi digunakan pendekatan analisisa varians (ANOVA). Bila bentuk hubungan antar peubah x dengan variabel y sudah dapat kita terima maka perlu pula diketahui

Kompetensi

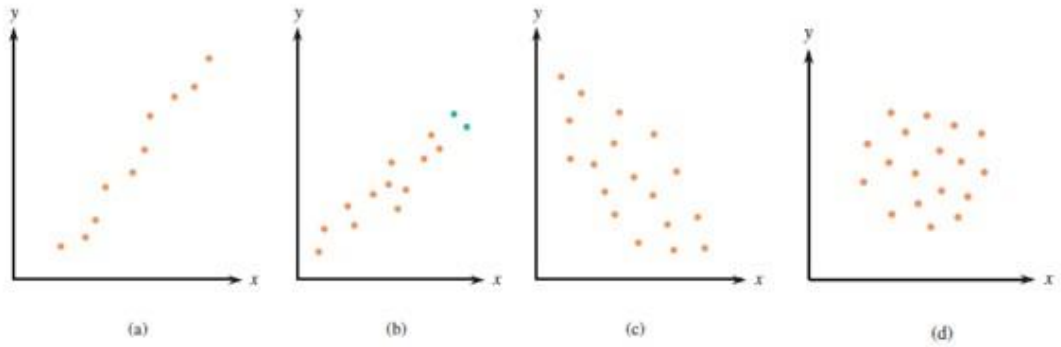
- Mahasiswa mampu untuk memahami pengertian regresi linear dalam hubungan antar variable
- Mahasiswa mampu untuk memahami pengertian nilai korelasi dalam kuat tidaknya hubungan antar variabel

seberapa besar keeratan hubungannya (korelasinya).

REGRESI LINIER

A. PENGERTIAN REGRESI

Secara umum ada dua macam hubungan antara dua variabel atau lebih, yaitu bentuk hubungan dan keeratan hubungan. Untuk mengetahui bentuk hubungan digunakan analisis regresi. Konsep analisis regresi adalah mencari hubungan terbaik antara Y dan x, mengukur kekuatan hubungan itu dan metode yang memungkinkan untuk memprediksi respon nilai-nilai Y dari nilai-nilai x yang diberikan. Untuk hubungan keeratan variabelnya dapat diketahui dengan analisis korelasi. Kekuatan hubungan keeratan variabel x dan dinyatakan dalam bentuk nilai numerik yang disebut koesien korelasi. Diagram pancar (scatter) berikut menampilkan bentuk-bentuk hubungan yang bebeda antar x dan y .



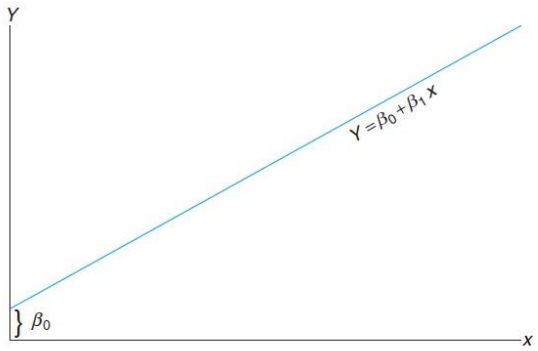
a) Menunjukkan hubungan positif yang kuat antara x dan y, karena setiap pasangan titik di plot, satu dengan nilai x lebih besar juga memiliki nilai y lebih besar. Artinya, peningkatan x dipasangkan dengan peningkatan y.
b) Menunjukkan kecenderungan kuat y tetapi tidak untuk x , walaupun ada beberapa pengecualian. Sebagai contoh, nilai-nilai x dan y dari dua titik dengan nilai-nilai x terbesar (ditampilkan dalam warna yang berbeda) masuk berlawanan arah (untuk pasangan titik, x meningkat tapi menurun y). Namun demikian, plot seperti ini masih menunjukkan hubungan positif yang cukup kuat.

c) Menunjukkan hubungan negatif antar x dan y . Nilai y cederung turun saat nilai x bertambah.menunjukkan bahwa tidak ada hubungan yang kuat antara x dan y.
d) Tidak ada kecenderungan pada y untuk bertambah atau berkurang dengan meningkatnya nilai x.

Analisis regresi digunakan untuk menelaah hubungan antara dua variabel atau lebih, terutama untuk menelusuri pola hubungan yang modelnya belum diketahui dengan sempurna, atau untuk mengetahui bagaimana variasi dari beberapa variabel independen mempengaruhi variabel dependen.Jika x adalah variabel-variabel independen dan y adalah variabel dependen, maka hubungan fungsional antara x dan y secara matematis dapat ditulis dalam bentuk persamaan berikut.

$$Y = \beta o + \beta_1 x$$

Dimana βo disebut intercept (titik potong) dan β_1 slope (kemiringan garis). Hubungan kedua variabel dapat digambarkan sebagai berikut :



Pada banyak kasus mungkin terdapat lebih dari satu independen. Misal harga rumah ditentukan oleh luas tanah dan umur bangunan. Jika Y adalah harga rumah dan x_1 variabel luas tanah dan x_2 variabel umur rumah, maka hubungan independen variabel dengan dependen variabel dapat ditulis dalam bentuk berikut :

$$Y = \beta o + \beta_1 x_1 + \beta_2 x_2$$

Bentuk hubungan diatas disebut multiple regressi (regresi berganda)

Dalam mengestimasi hubungan antara independen variabel x dengan variabel dependent menggunanakan analisis regresi, setidaknya ada empat tahapan yang harus dilakukan,diantaranya:

(1) Mengadakan estimasi terhadap parameter berdasarkan data empiris,

(2) Menguji berapa besar variasi variabel dependen dapat diterangkan oleh variasi variabel independen,

(3) Menguji apakah estimasi parameter tersebut signifikan atau tidak, dan

(4) Melihat apakah tanda dan magnitud dari estimasi parameter cocok dengan teori

(M. Nazir, 1983).

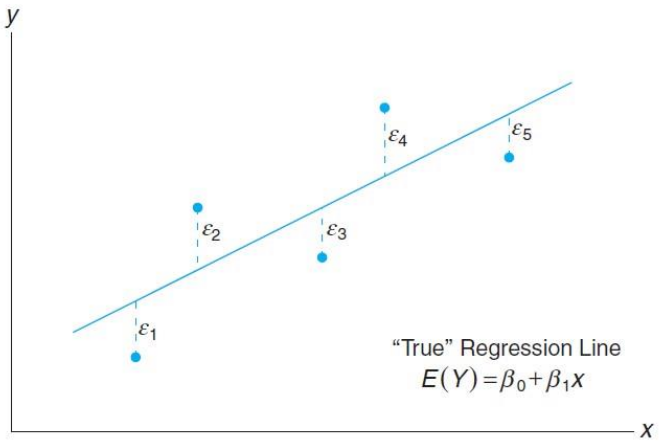
B. REGRESI LINIER SEDERHANA

Dalam menganalisa hubungan antara Y dan x diperlukan model statistik. Dimana hubungan antar variabel tidak deterministik atau dengan kata lain harus ada komponen random untuk persamaan yang menghubungkan variable.

Sebuah model sering digunakan oleh ahli statistik sebagai representasi ideal yang pada dasarnya mendefinisikan bagaimana kita memandang bahwa data yang dihasilkan oleh sistem yang bersangkutan. Model ini harus mencakup himpunan $\{(x_i, y_i); i = 1, 2, \ldots, N\}$ data yang melibatkan pasangan n(x,y) nilai. Nilai y_i tergantung pada x_i melalui struktur linier yang juga memiliki komponen acak. Model ini juga mencakup apa yang diasumsikan tentang sifat-sifat statistik dari komponen acak seperti model regresi linier sederhana yang dirumuskan berikut ini

$$Y = \beta o + \beta_1 x + \epsilon$$

Dimana βo disebut intercept (titik potong) dan β_1 slope (kemiringan garis). ϵ random variable dengan $E(\epsilon)=0$ dan dengan variansi kesalahan atau variansi residu $V(\epsilon) = \sigma^2$. Dengan melibatkan variabel random ϵ model diatas terlihat lebih alami. Hubungan antara garis regresi dengan nilai sampel dapat dilihat pada gambar berikut.



Jika b_0 dan b_1 adalah estimasi dari β_0 dan β_1, perhitungan koefisien regresi linier dengan menggunakan metode least squre diberikan sebagai berikut :

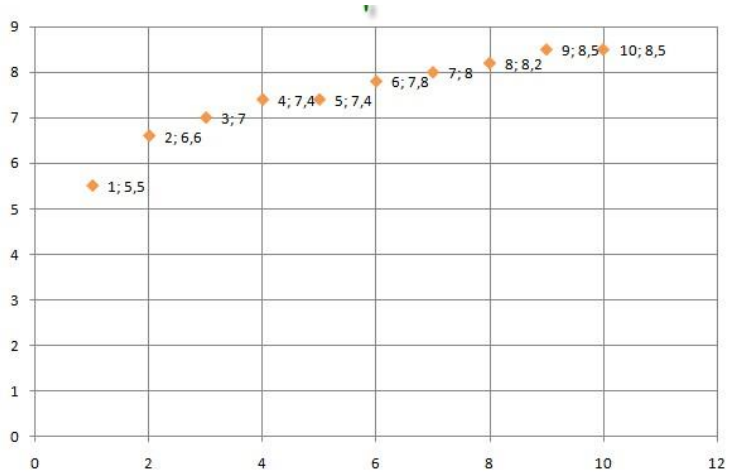
$$b_1 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{S_{XY}}{S_{XX}}$$

$$b_0 = \frac{\sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} x_i}{n} = \quad Y - b_1 X$$

Contoh :

Untuk melihat hubungan antara pemberian tugas dengan nilai ahir yang didapat seorang mahasiswa. Dibuat penelitian sederhana untuk masalah tersebut. Data yang dihimpun dari penelitian tersebut diberikan pada tabel berikut.

No.	Jumlah tugas	Nilai Ujian
1	1	5,5
2	2	6,6
3	3	7
4	4	7,4
5	5	7,4
6	6	7,8
7	7	8
8	8	8,2
9	9	8,5
10	10	8,5



Jika x = tugas dan y nilai

x	y				
1	5,5	-4,5	-1,99	20,25	8,955
2	6,6	-3,5	-0,89	12,25	3,115
3	7	-2,5	-0,49	6,25	1,225
4	7,4	-1,5	-0,09	2,25	0,135
5	7,4	-0,5	-0,09	0,25	0,045
6	7,8	0,5	0,31	0,25	0,155
7	8	1,5	0,51	2,25	0,765
8	8,2	2,5	0,71	6,25	1,775
9	8,5	3,5	1,01	12,25	3,535
10	8,5	4,5	1,01	20,25	4,545

	∑= 82,5	∑= 24,25

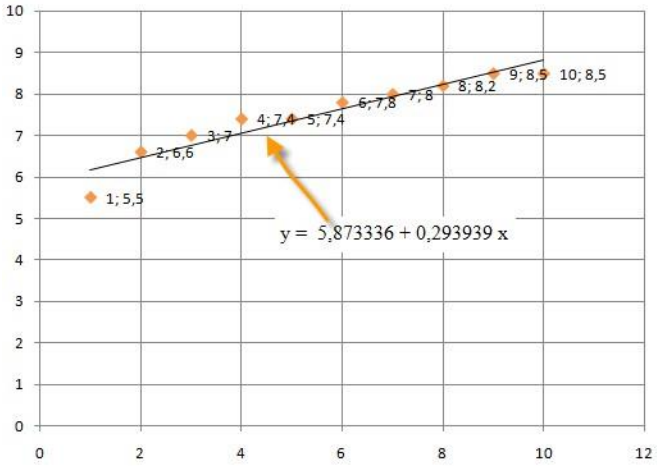
$b_1 = \frac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^n (x_i - \bar{x})} = \frac{24,25}{82,5} = 0,293939$

$b_0 = \bar{y} - b_1 \bar{x}$ =7,49 - 0,293939 . 5,5 = 5,873336

$\hat{y}$ = $b_0 + b_1 x$

$\hat{y}$ = 5,873336 + 0,293939x

$\hat{y}$ = nilai estimasi



C. *Uji Keberartian Regresi*

Seringkali masalah analisa estimasi garis regresi digunakan pendekatan analisisa varians (ANOVA). Jaitu suatu prosedur dimana variasi total (SST) dalam variabel dependent dibagi menjadi komponen yang berarti dan kemudian diamati dan diperlakukan dengan cara yang sistematis.

$$SST = SSR + SSE$$

Dimana :

$\sum_{i=1}^{n} (Y_i - \bar{Y})^2$: Jumlah Kuadrat Sekitar Rataan,

Sum of Square Total, SST

$\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$: Jumlah Kuadrat Karena Regresi

Sum of Square Regressionl, SSR

$\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$: Jumlah Kuadrat Sekitar Regresi,

atau Error, *Sum of Square Error, SSE*

Pemeriksaan keberartian regresi dilakukan melalui pengujian hipotesis nol, bahwa koefisien regresi b sama dengan nol (tidak berarti) melawan hipotesis tandingan bahwa koefisien arah regresi tidak sama dengan nol.

Pengujian koefisien regresi dapat dilakukan dengan langkah-langkah pengujian hipotesis berikut:

1. Menentukan rumusan hipotesis H_o dan H_1.

   $H_o : \beta_1 = 0$: variabel X dan Y saling independent Y.

   $H_1 : \beta_1 \neq 0$: variabel X dan Y tidak saling independent.

2. Menghitung beberapa nilai yang dibutuhkan untuk uji F

   a. Jumlah Kuadrat Regresi dengan rumus:

   $$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$$

   b. Jumlah Kuadrat Sekitar Regresi, dengan rumus:

   $$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

   c. Jumlah Kuadrat Sekitar Regresi dengan rumus:

   $$\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

   d. Kuadrat tengah dengan rumus :

   $$s^2 = \frac{SSE}{n-2}$$

   e. Mengitung F, dengan rumus:

   F=

3. Menentukan nilai kritis (α) atau nilai tabel F pada derajat bebas $db_{reg\ b/a} = 1$ dan $db_{res} = n - 2$.

4. Membandingkan nilai uji F dengan nilai tabel F, dengan kriteria uji, Apabila nilai hitung F lebih besar atau sama dengan (≥) nilai tabel F, maka H_0 ditolak.

x	71	53	82	67	56	70	64	78	55	70	53	84
y	354	313	322	334	247	377	308	340	301	349	293	368

5. Membuat kesimpulan

Langkah-langkah uji keberartian regresi di atas dapat disederhanakan dalam sebuah tabel anova, dimana tiga suku di atas akan menjadi komponen Tabel A-nalisis Variansi (ANOVA) sebagai berikut :

Tabel ANOVA

Sumber Variasi (*Source*)	Derajat Bebas (db) (*df*)	Jumlah Kuadrat (JK) (*SS*)	Kuadrat tengah (KT) = JK/db (*MS*)	F hitung
Regresi	1	$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$	SSR	$\frac{SSR}{s^2}$
Error atau Residual	*n*-2	$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	$s^2 = \frac{SSE}{n-2}$	
Total, terkoreksi	*n*-1	$\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$		

Contoh :

Berikut ini adalah tabel regangan alumunium

a. Cari persamaan garis regresi

x	y				

71	354	4,083333	28,5	16,67361	116,375
53	313	-13,9167	-12,5	193,6736	173,9583
.....	.....	.....	.....	.....	.....
70	349	3,083333	23,5	9,506944	72,45833
53	293	-13,9167	-32,5	193,6736	452,2917
84	368	17,08333	42,5	291,8403	726,0417
				1334,917	3008,5

$b_1 = \frac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^n (x_i - \bar{x})} = \frac{3008,55}{1334,9175} = 2,253699$

$b_0 = \bar{y} - b_1 \bar{x} = 174,69$

$\hat{y} = b_0 + b_1 x$

$\hat{y} = 174,69+2,253699x$ ($\hat{y}$ = nilai estimasi)

b. Pengujian hipotesis

Menentukan rumusan hipotesis H_0 dan H_1.

$H_o : \beta_1 = 0$: variabel X dan Y saling independent Y.

$H_1 : \beta_1 \neq 0$: variabel X dan Y tidak saling independent.

Menghitung beberapa nilai yang dibutuhkan untuk uji F

x	y	= 174,69+2,253699x	SST	SSR	SSE
71	354	334,7026	84,6883	372,3885	812,25
53	313	294,1360	827,8880	355,8487	98,1900
..	..	.....	.....	.....	.....

84	368	364,0007	15,9942	15,9943	0
		∑	5338,4590	7618,7474	11354,7225

Kuadrat tengah:

$$s^2 = \frac{SSE}{n-2} = \frac{11354,722}{12-2} = 1135,4722$$

f hitung :

***f* hitung =** $\frac{SSR}{s^2}$ **= (7168,75/1135,47)=** 6,313465

Nilai kritis (α) atau nilai $f_\alpha(1,n-2)$ =

$f_{0,05}(1,12-2)$ =10,044

Membandingkan nilai *f* uji dengan nilai *f*

F hitung > f tabel , maka Ho ditolak

Kesimpulan

Variabel x dan y tidak saling independent,berarti setiap perubahan x akan mempengaruhi nilai y.

D. Korelasi

Bila bentuk hubungan antar peubah x dengan variabel y sudah dapat kita terima maka perlu pula diketahui seberapa besar keeratan hubungannya(korelasinya). Walaupun bentuk hubungan antara variabel x dengan variabel y ada. Tetapi belum tentu korelasinya besar karena banyak peubah lain yang turut mempengaruhi perubahan variabel y.Besarnya perubahan peubah y yang dapat diterangkan oleh peubah x dengan menggunakan persamaan garis regresi yang diperoleh disebut koefisien determinan Koefisien determinat diberi lambing r^2 untuk bentuk persamaan garis regresi sederhana dengan rumus :

$$r = \sqrt{\frac{SSR}{S_{yy}}}$$

dimana

$$SSR = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

$$Syy = \sum_{i=1}^{n} (Y_i - \bar{Y})$$

Untuk contoh soal diatas

$$r = \sqrt{\frac{SSR}{S_{yy}}} = \sqrt{\frac{7618,7474}{14399}} = 0,529116258$$

Soal :

1. Berikut adalah tabel hubungan varibel x dan y

x	6	7	8	6	8	7	9	8	7
y	5	6	7	6	7	6	8	6	7

Dari data diatas tentukan

a. Persamaan garis regresinya
b. Nilai y jika x = 6,5

2. Hasil pengeringan suatu produk untuk berbagai tingkat/ lama pengeringan diberikan dalam tabel berikut.

x	**y (gram)**	
4,4	13.1	14,2
4,5	9,0	11,5
4,8	10,4	11,5
5,5	13,8	14 8
5,7	1	15
		12
		16

6,9	16,4	15,7
7,5	17,6	16,9
7,8	18,3	17,2

.

a. Tentukan persamaan garis linier tabel tersebut
b. Test keberartian garis linier yang didapat pada tingkat α=0,05
c. Tentukan nilai korelasinya r

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Analisis regresi ganda merupakan pengembangan dari analisis regresi sederhana, regresi ganda digunakan untuk meramalkan nilai variabel terikat (Y) apabila variabel bebasnya (X) dua atau lebih. Bentuk persamaan linier dengan k variabel independen dapat

Kompetensi

Mahasiswa mampu memahami pengertian regresi linear berganda dan regresi non linear dalam hubungan antar variable

diselesaikan dengan metode eliminasi Gauss-Yordan atau metode Cramer.

REGRESI GANDA

Dalam masalah penelitian di mana analisis regresi diterapkan, mungkin ada lebih dari satu variabel independen yang diperlukan untuk membangun model garis regresinya.Dalam masalah ini maka model regresi linier sederhana tidak dapat digunakan.

A. REGRESI GANDA

Analisis regresi ganda merupakan pengembangan dari analisis regresi sederhana. Kegunaannya yaitu untuk meramalkan nilai variabel terikat (Y) apabila variabel bebasnya (X) dua atau lebih.

Analisis regresi ganda adalah alat untuk meramalkan nilai pengaruh dua variabel bebas atau lebih terhadap satu variabel terikat (untuk membuktikan ada tidaknya hubungan fungsional atau hubungan kausal antara dua atau lebih variabel bebas X_1, X_2, …., X_i terhadap suatu variabel terikat Y.

Untuk k variabel independen x_1, x_2, x_3,....,x_k , estimasi variabel dependen dari sampe $\hat{y}$, dapat dituliskan dalam bentuk persamaan berikut ini.

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + b_k x_k$$

Dimana b_0 ..... b_k koefisien persamaan yang akan diestimasi

Dengan menggunakan metode least square untuk mengestimasi nilai b_0, ... b_k , Sum Square Erro (SSE) dimana ,

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_{1i} - b_2 x_{2i} - \cdots - b_k x_{ki})^2.$$

Dengan mendifensir SSE terhadap b_0, ... b_k , maka didapat

$$
\begin{aligned}
nb_0 + b_1 \sum_{i=1}^{n} x_{1i} \quad &+ b_2 \sum_{i=1}^{n} x_{2i} \quad + \cdots + b_k \sum_{i=1}^{n} x_{ki} \quad = \sum_{i=1}^{n} y_i \\
b_0 \sum_{i=1}^{n} x_{1i} + b_1 \sum_{i=1}^{n} x_{1i}^2 \quad &+ b_2 \sum_{i=1}^{n} x_{1i}x_{2i} + \cdots + b_k \sum_{i=1}^{n} x_{1i}x_{ki} = \sum_{i=1}^{n} x_{1i}y_i \\
\vdots \quad &\vdots \quad \vdots \quad \vdots \quad \vdots \\
b_0 \sum_{i=1}^{n} x_{ki} + b_1 \sum_{i=1}^{n} x_{ki}x_{1i} &+ b_2 \sum_{i=1}^{n} x_{ki}x_{2i} + \cdots + b_k \sum_{i=1}^{n} x_{ki}^2 \quad = \sum_{i=1}^{n} x_{ki}y_i
\end{aligned}
$$

Bentuk persamaan diatas adalah bentuk sistem persamaan linier dengan k variabel independen dan dapat diselesaikan dengan metode solusi sistem persamaan linier yang ada. Misal menggunakan metode eliminasi Gauss-Yordan atau metode Cramer.

Untuk persamaan linier ganda dengan variabel bebas k =2, maka bentuk persamaan diatas dapat ditulis sebagai :

$$nb_0 + b_1 \sum_{i=1}^{n} x_{1i} + b_2 \sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i$$

$$b_0 \sum_{i=1}^{n} x_{1i} + b_1 \sum_{i=1}^{n} x_{1i}^2 + b_2 \sum_{i=1}^{n} x_{1i} = b_0 \sum_{i=1}^{n} x_{1i}\, y_i$$

$$b_0 \sum_{i=1}^{n} x_{2i} + b_1 \sum_{i=1}^{n} x_{1i}\, x_{2i} + b_2 \sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i}\, y_i$$

Solusi persmaan linier diatas akan menghasilkan nilai b_0,b_1 dan b_2 yang kemudian disubtitusikan ke persamaan :

$$\hat{y} = b_0 + b_1\, x_1 + b_2\, x_2 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots + b_k\, x_k$$

Sehingga didapat persamaan garis regresi

$$\hat{y} = b_0 + b_1\, x_1 + b_2\, x_2$$

Contoh :

Seorang Manajer Pemasaran makanan ringan ingin mengetahui apakah promosi dan harga

berpengaruh terhadap keputusan konsumen membeli produk tersebut. Hasil survey terhadap 10 orang responden didapat data berikut.

No. Responden	**Promosi** x_1	**Harga** x_2	**Konsumen (Y)**
1	10	7	23
2	2	3	7
3	4	2	15
4	6	4	17
5	8	6	23
6	7	5	22
7	4	3	10
8	6	3	14
9	7	4	20
10	6	3	19
Jumlah	60	40	170

Tabel Pembantu

No. Resp (*i*)	x_1	x_2	y	x_1y	x_2y	x_1x_2	x_1^2	x_2^2

1	10	7	23	230	161	70	100	49
2	2	3	7	14	21	6	4	9
3	4	2	15	60	30	8	16	4
4	6	4	17	102	68	24	36	16
5	8	6	23	184	138	48	64	36
6	7	5	22	154	110	35	49	25
7	4	3	10	40	30	12	16	9
8	6	3	14	84	42	18	36	9
9	7	4	20	140	80	28	49	16
10	6	3	19	114	57	18	36	9
Jumlah	60	40	170	1122	737	267	406	182

Dengan mensubtitusi nilai-nilai dari tabel pembantu ke persamaan diatas, maka didapat sistem persamaan linier.

$10\,b_0 + 60b_1 + 40b_2 = 170$

$60b_0 + 406b_1 + 267b_2 = 1122$

$40b_0 + 267b_1 + 182b_2 = 737$

Sistem persamaan linier dengan tiga variabel b_0,b_1 dan b_2 yang tidak diketahui tersebut dapat diselesaikan dengan metode Cramer.

Berdasarkan aturan Cramer , determinan yang relevan dengan |B| dan |Bj| diperoleh sebagai :

$$|B| = \begin{bmatrix} 10 & 60 & 40 \\ 60 & 40 & 267 \\ 40 & 267 & 182 \end{bmatrix} \qquad |B_0| = \begin{bmatrix} 170 & 60 & 40 \\ 1122 & 40 & 267 \\ 4737 & 267 & 182 \end{bmatrix}$$

$$|B_1| = \begin{bmatrix} 10 & 170 & 40 \\ 60 & 1122 & 267 \\ 40 & 737 & 182 \end{bmatrix} \qquad |B_2| = \begin{bmatrix} 10 & 60 & 170 \\ 60 & 40 & 1122 \\ 40 & 267 & 737 \end{bmatrix}$$

$$b_o = \frac{\begin{bmatrix} 170 & 60 & 40 \\ 1122 & 40 & 267 \\ 4737 & 267 & 182 \end{bmatrix}}{\begin{bmatrix} 10 & 60 & 40 \\ 60 & 40 & 267 \\ 40 & 267 & 182 \end{bmatrix}} = 3{,}9186 \qquad b_1 = \frac{\begin{bmatrix} 10 & 170 & 40 \\ 60 & 1122 & 267 \\ 40 & 737 & 182 \end{bmatrix}}{\begin{bmatrix} 10 & 60 & 40 \\ 60 & 40 & 267 \\ 40 & 267 & 182 \end{bmatrix}} = 2{,}4909$$

$$b_2 = \frac{\begin{bmatrix} 10 & 60 & 170 \\ 60 & 40 & 1122 \\ 40 & 267 & 737 \end{bmatrix}}{\begin{bmatrix} 10 & 60 & 40 \\ 60 & 40 & 267 \\ 40 & 267 & 182 \end{bmatrix}} = -0{,}466$$

dengan mensubtitusi nilai b_0,b_1 dan b_2 kepersamaan $\hat{y} = b_0 + b_1 x_1 + b_2 x_2$ maka didapat persamaan garis regresi.

$\hat{y}$ = 3,9186 + 2,4909 x_1 -0,466 x_2

Kofiesien Korelasi

. Besarnya perubahan peubah y yang dapat diterangkan oleh variabel x_1 dan x_2 dengan menggunakan persamaan garis regresi yang diperoleh, disebut koefisien determinan Koefisien determinat diberi lambing r^2 untuk bentuk persamaan garis regresi sederhana dengan rumus :

$$R = \frac{b_1 \sum_{i=1}^{n} x_{1i} y_i + b_2 \sum_{i=1}^{n} x_{2i} y_i}{\sum_{i=1}^{n} y^2}$$

R = ((2,4909 * 1122)+(-0,466 * 737))/ 3162 =0,7753

B. PENGUJIAN KEBERARTIAN REGRESI GANDA

Seringkali masalah analisa estimasi garis regresi digunakan pendekatan analisisa varians (ANOVA). Jaitu suatu prosedur dimana variasi total (SST) dalam variabel dependent dibagi menjadi komponen yang berarti dan kemudian diamati dan diperlakukan dengan cara yang sistematis.

$$SST = SSR + SSE$$

Dimana :

$\sum_{i=1}^{n}(Y_i - \bar{Y})^2$: Jumlah Kuadrat Sekitar Rataan,

Sum of Square Total, SST

$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$: Jumlah Kuadrat Regresi

Sum of Square Regressionl, SSR

$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$: Jumlah Kuadrat Sekitar Regresi,

atau Error, *Sum of Square Error, SSE*

Pemeriksaan keberartian regresi dilakukan melalui pengujian hipotesis nol, bahwa koefisien regresi b sama dengan nol (tidak berarti) melawan hipotesis tandingan bahwa koefisien arah regresi tidak sama dengan nol.

Pengujian koefisien regresi dapat dilakukan dengan langkah-langkah pengujian hipotesis berikut:

6. Menentukan rumusan hipotesis H_o dan H_1.

   $H_o : \beta_1 = \beta_2 = 0$: variabel X dan Y saling independent Y.

   $H_1 : \beta_1 \neq \beta_2 \neq 0$: variabel X dan Y tidak saling independent.

7. Menghitung beberapa nilai yang dibutuhkan untuk uji F

   f. Jumlah Kuadrat Regresi SSR dengan rumus:

   $$\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$$

   g. Jumlah Kuadrat Sekitar Regresi SSE, dengan rumus:

   $$\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

   h. Kuadrat tengah dengan rumus :

   $$MSR = \frac{SSR}{k}$$

   $$MSE = \frac{SSE}{n-(k+1)}$$

   i. Mengitung F, dengan rumus:

   F=

8. Membandingkan nilai uji F dengan nilai tabel F, dengan kriteria uji, Apabila nilai hitung F lebih besar atau sama dengan (≥) nilai tabel F, maka H_0 ditolak.

9. Membuat kesimpulan

Langkah-langkah uji keberartian regresi di atas dapat disederhanakan dalam sebuah tabel anova, dimana tiga suku di atas akan menjadi komponen Tabel A-nalisis Variansi (ANOVA) sebagai berikut :

Tabel ANOVA

Sumber Variasi (*Source*)	Derajat Bebas (db)	Jumlah Kuadrat (JK) (*SSR*)(SSE)(SST)	Kuadrat tengah (*MSR*) (MSE)	F hitung

	(*df*)			
Regresi	k	$\sum_{i=1}^{n}(\hat{Y}_i - Y)^2$	MSR= $\frac{SSR}{k}$	$\frac{MSR}{MSE}$
Error atau Residual	*n*-(k+1)	$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$	$MSE = \frac{SSE}{n-(k+1)}$	
Total, terkoreksi	*n*-1	$\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$		

Contoh :

Uji keberartian koefisien regresi soal contoh diatas.

$\hat{y}$ = 3,9186 + 2,4909 x_1 -0,466 x_2

1. hipotesis H_o dan H_1.

   H_o : $\beta_1 = \beta_2 = 0$: Variabel Promosi Dan Harga Tidak Berpengaruh Signifikan terhadap Keputusan Konsumen.

   H_1 : $\beta_1 \neq \beta_2 \neq 0$: Variabel Promosi Dan Harga Berpengaruh Signifikan terhadap Keputusan Konsumen.

2. Menghitung beberapa nilai yang dibutuhkan untuk uji F

   SSR,SSE,MSR,MSE dan F dengan rumus:

   $\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$, $\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$

   $MSR = \frac{SSR}{k}$, $MSE = \frac{SSE}{n-(k+1)}$

   F=

No. Resp	x_1	x_2	y	= 3,9186 + 2,4909	-0,466	SSR	SSE	SST
1	10	7	23	25,53		72,7609	6,4009	36
2	2	3	7	7,49		90,4401	0,2401	100
3	4	2	15	12,94		16,4836	4,2436	4
4	6	4	17	16,98		0,0004	0,0004	0
5	8	6	23	21,02		16,1604	3,9204	36
6	7	5	22	19		4	9	25
7	4	3	10	12,47		20,5209	6,1009	49
8	6	3	14	17,45		0,2025	11,9025	9
9	7	4	20	19,47		6,1009	0,2809	9
10	6	3	19	17,45		0,2025	2,4025	4
∑	**60**	**40**	**170**	**169,8**		**226,872**	**44,4922**	**272**

Tabel ANOVA

Sumber Variasi (*Source*)	Derajat Bebas (db) (*df*)	Jumlah Kuadrat (JK) (*SSR*)(SSE)(SST)	Kuadrat tengah (*MSR*) (MSE)	F hitung
Regresi	2	SSR=266,872	MSR=133,4436	$\frac{MSR}{MSE}$ = 20,9949
Error atau Residual	*10*-(2+1)=7	SSE = 44,4922	MSE= 6,356	
Total, terkoreksi	*10*-1=9	SST=272		

Nilai tabel Nilai kritis (α) atau nilai $f_{\alpha}(2,7)$ =

$f_{0,05}(2,7) = 4.737$

3. F hitung > F tabel , maka H_0 ditolak.
4. Kesimpulan

   Promosi Dan Harga Berpengaruh Signifikan terhadap keputusan konsumen.

Soal

1. Naik turunnya import suatu barang dari suatu negara ditentukan antara lain oleh produksi barang tersebut dan rasio(perbandingan) tingkat harga barang import tersebut terhadap tingkat harga dalam negeri. Berikut adalah tabel index import, index produksi dan rasio index harga impor dan harga dalam negeri.

y	120	128	140	130	125	120	144	150	135
x_1	105	115	130	123	120	115	130	145	128
x_2	105	100	128	115	115	110	126	138	120

Jika :

y= indeks import kopi

x_i = index produksi kopi dalam negeri

x_2 = rasio index harga impor dan harga dalam negeri

Tentukan :

a. Persamaan garis regresi
b. Nilai y, jika $x_1 = 125$ dan $x_2 = 120$

2. Untuk melihat hubungan antara tugas,absensi dan nilai ahir matakuliah “Pemrograman Berorientasi Objek “. Dilakukan penelitian sederhana dengan mengambil sampel 14 orang mahasiswa yang mengambil matakuliah tersebut. Data nilai tugas, absen dan nilai ahir diberikan pada tabel berikut.

No	Tugas	Absensi	Nilai Ahir	No	Tugas	Absensi	Nilai Ahir

1	50	65	60	8	75	85	70
2	60	65	60	9	80	75	70
3	60	70	65	10	80	80	70
4	70	75	65	11	85	75	75
5	70	70	70	12	85	80	70
6	75	70	70	13	90	80	80
7	75	80	70	14	90	85	80

Dengan menggunakan data diatas

a. Tentukan persamaan garis regresi data diatas
b. Dengan menggunkan garis regresi tersebut, tentukan estimasi nilai ahir , jika nilai tugas = 85 dan absensi 70.
c. Tentukan koefisien korelasi garis tersebut
d. Lakukan uji keberartian garis regresi yang didapat pada tingkat α=0,05 dengan hipotesi :

$H_o : \beta_1 = \beta_2 = 0$: Variabel tugas Dan absensi tidak berpengaruh signifikan terhadap nilai ahir.

$H_1 : \beta_1 \neq \beta_2 \neq 0$: Variabel tugas Dan absensi berpengaruh signifikan terhadap Keputusan Konsumen.

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992

MODUL PERKULIAHAN

STATISTIKA DAN PROBABILITAS

DISUSUN OLEH ; Dr. H. MUHAMAD AHYAT S.E., M.M.

UNIVERSITAS TEKNOLOGI MATARAM

Abstract

Time Series (Data Berkala) adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan perkembangan suatu kegiatan atau sekumpulan hasil observasi yang diatur menurut urutan kronologis waktu. Analisis data berkala untuk mengetahui perkembangan suatu kejadian serta pengaruhnya/hubunganya terhadap kejadian lain sehingga dapat dibuat ramalan berdasarkan garis regresi atau garis trend. Pola data deret berkala dapat berbentuk trend, gerak siklis, musiman dan ireguler.

Kompetensi

- Mahasiswa mampu untuk menganalisis data berkala dari waktu kewaktu.
- Mahasiswa mampu memahami metoda untuk menentukan trend data deret berkala.

Analisa Data Berkala

A. Pengenalan

Time Series (Data Berkala /Data Deret waktu) adalah data yang dikumpulkan dari waktu ke waktu untuk menggambarkan perkembangan suatu kegiatan atau sekumpulan hasil observasi yang diatur dan didapat menurut urutan kronologis waktu, misalnya perkembangan produksi, harga barang, hasil penjualan, jumlah penduduk, dll.

Ada dua tujuan utama dari analisis data berkala:

(a) mengidentifikasi sifat dari fenomena diwakili oleh urutan pengamatan,
(b) peramalan (memprediksi nilai masa depan dari variabel time series).

Kedua tujuan mengharuskan pola data berkala yang diamati diidentifikasi terlebih dahulu. Dengan plot data ke dalam bentuk garafik dan melihat pola yang terbentuk , kita dapat menafsirkan dan kemudian menerapkan model analisis yang sesuai untuk pola data tersebut untuk memprediksi kejadian masa depan

Analisis data deret waktu pada dasarnya digunakan untuk melakukan analisis data yang mempertimbangkan pengaruh waktu. Data-data yang dikumpulkan secara periodik berdasarkan urutan waktu, bisa dalam jam, hari, minggu, bulan, kuartal dan tahun, bisa dilakukan analisis menggunakan metode analisis data deret waktu. Analisis data deret waktu tidak hanya bisa dilakukan untuk satu variabel (*Univariate*) tetapi juga bisa untuk banyak variabel (*Multivariate*). Selain itu pada analisis data deret waktu bisa dilakukan peramalan data beberapa periode ke depan yang sangat membantu dalam menyusun perencanaan ke depan.

Analisis data berkala memungkinkan kita untuk mengetahui perkembangan suatu/beberapa kejadian serta pengaruhnya/hubunganya terhadap kejadian lain dan dapat pula membuat ramalan berdasarkan garis regresi atau garis trend. Metode yang digunakan dalam analisa data berkala adalah metode kuantitatif sehingga perlu diperhatikan beberapa syarat :

1. Ketersedian informasi tentang masa lalu.
2. Informasi tersebut dapat di kuantitatifkan dalam bentuk data numerik.

3. Dapat diasumsikan bahwa beberapa aspek pola masa lalu akan terus berlanjut di masa mendatang.

Data berkala terdiri dari komponen-komponen, sehingga dengan analisis data berkala kita dapat mengetahui masing-masing komponen atau bahkan menghilangkan suatu/beberapa komponen. Karena ada pengaruh dari komponen, data berkala selalu mengalami perubahan-perubahan, sehingga apabila dibuat grafik akan menunjukkan adanya fluktuasi.

B. Pola Data Deret Berkala

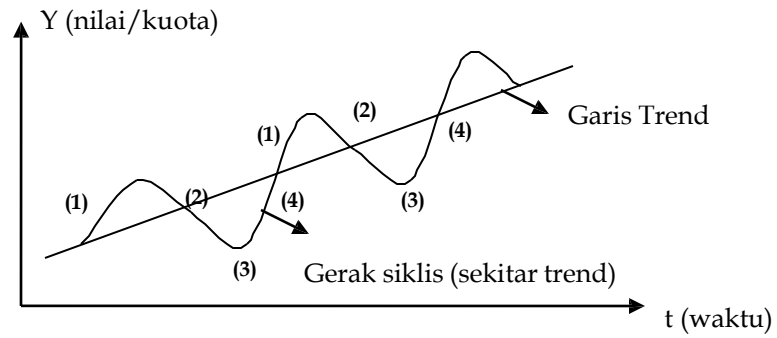
- Trend

  Trend melukiskan gerak data berkala selama jangka waktu yang panjang/cukup lama. Gerak ini mencerminkan sifat kontinuitas atau keadaan yang serba terus dari waktu ke waktu selama jangka waktu tersebut. Karena sifat kontinuitas ini, maka trend dianggap sebagai gerak stabil dan menunjukkan arah perkembangan secara umum (kecenderungan menaik/menurun). Trend sangat berguna untuk membuat peramalan *(forecasting)* yang merupakan perkiraan untuk masa depan yang diperlukanbagi perencanaan. Trend dibedakan menjadi dua jenis, yakni :

  a. Trend Linier → mengikuti pola garis lurus (Y = a + b t)
  b. Trend Non Linier → mengikuti pola lengkung (parabola, eksponensial, logaritma, dll).

- Gerak Siklis

  Gerak siklis adalah gerak/variasi jangka panjang di sekitar garis trend (temponya lebih pendek). Gerak siklis terjadi berulang-ulang namun tidak perlu periodic, artinya bisa berulang setelah jangka waktu tertentu atau bisa juga tidak berulang dalam jangka waktu yang sama.Perkembangan perekonomian yang turun naik di sekitar trend dan *"Business Cycles"* adalah contoh gerak siklis. Gerak siklis melukiskan terjadinya empat fase kejadian dalam jangka waktu tertentu, yakni kemajuan, kemunduran, depresi dan pemulihan.



Keterangan :

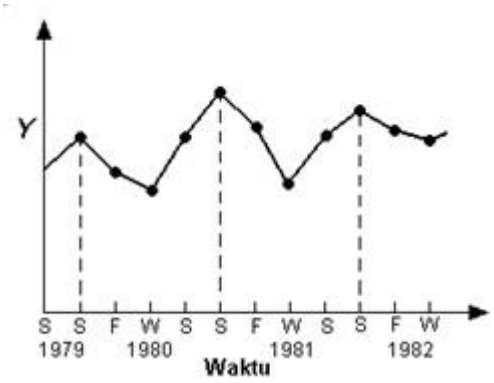
(1) Kemajuan

(2) Kemunduran

(3) Depresi

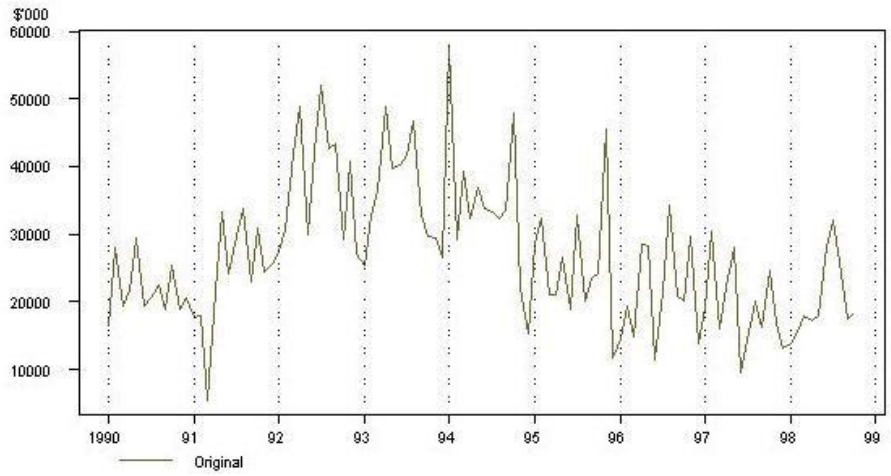
- Musiman

  Gerak musiman terjadi lebih teratur dibandingkan garak siklis dan bersifat lengkap, biasanya selama satu tahun kalender. Gerak ini berpola tetap dari waktu ke waktu. Factor utama yang menyebabkan gerak ini adalah iklim dan kebiasaan.



- Gerak Ireguler

  Gerak ini bersifat sporadis/tidak teratur dan sulit dikuasai. Gerak ini bisa diakibatkan oleh bencana alam, mogok dan kekacauan adalah beberapa faktor yang terkenal yang bisa menyebabkan gerak ini terjadi. Dengan adanya pengaruh tersebut, maka gerak ireguler sulit untuk dilukiskan dalam suatu model.



C. Metode Analisa Data Berkala

Ada empat cara yang akan dipelajari untuk menentukan persamaan trend linier.

a. metode bebas,
b. metode setengah rata-rata,
c. metode rata-rata bergerak
d. metode kuadrat terkecil

Keempat cara ini dipakai untuk menentukan bentuk umum persamaan trend linier,yaitu

$$y = a + bx$$

Dimana :

- y adalah nilai trend pada periode tertentu (variabel tak bebas)
- x adalah periode waktu (variabel bebas)
- a adalah intersep (konstanta) dari persamaan trend
- b adalah koefisien kemiringan atau gradien dari persamaan trend yang menunjukkan besarnya perubahan Y bila terjadi perubahan satu unit pada x.

1. Metode Tangan Bebas

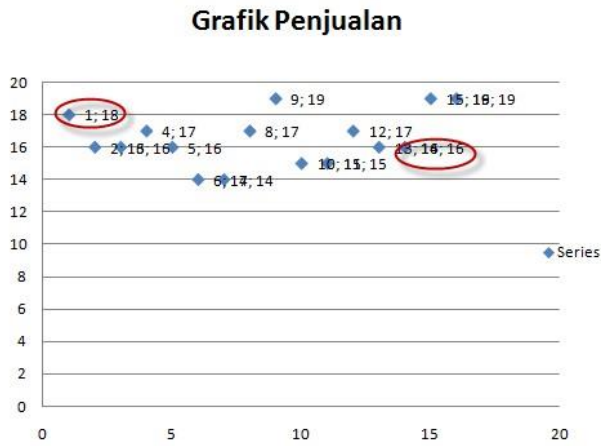
Metode tangan bebas adalah cara yang paling sederhana dalam menentukan garis tend. Cara ini bisa jadi sangat subjektif. Jika beberapa orang diminta untuk menarik garis trend tersebut maka akan ada lebih dari satu garis trend. Langkah-langkah yang diperlukan untuk menentukan persamaan trend dengan cara ini adalah sebagai berikut.

- Buatlah sumbu X dan sumbu Y dalam sistem koordinat Catesius.
- Buatlah diagram pencar (scatter diagram) dari pasangan titik (x,y) yang menyatakan kaitan antara waktu dan nilai data berkala.
- Tariklah garis linier yang arahnya mengikuti arah penyebaran nilai-nilai data barkala.
- Pilihlah dua titik sembarangan untuk menentukan persamaan trend linier, misalnya titik (x_1,y_1) dan (x_2,y_2).
- Pilih salah satu periode waktu data berkala sebagai titik asal (x=0).
- Subtitusikan nilai-nilai x dan y dari dua titik yang telah dipilih pada rumus persamaan berikut.

$$(y - y_1) = \frac{y_2 - y_1}{x_2 - x_1} . (x - x_1)$$

Contoh :

Tahun	x	Penjualan (000 kg) y
1995	1	18
1996	2	16
1997	3	16
1998	4	17
1999	5	16
2000	6	14
2001	7	14
2002	8	17
2003	9	19



Dengan menggunakan titik-titik yang ditandai lingkaran merah diatas, maka dengan menggunakan persamaan garis dua titik didapat trend

y = -0,15 x +18,15

2004	10	15
2005	11	15
2006	12	17
2007	13	16
2008	14	16
2009	15	19
2010	16	19

2. Semi Rata-Rata

Penentuan persamaan trend linier $\hat{Y} = a + bX$ dengan metode semi rata-rata dilakukan dengan tahapan-tahapan berikut.

a. Bagilah data berkala menjadi dua kelompok data yang sama banyak,

b. Jika jumlah data ganjil, hapus data yang membagi kedua kelompok tersebut.

c. Tentukan rata-rata hitung masing-masing kelompok, y_1 dan y_2.

d. Tentukan dua titik, yaitu (x_1,y_1) dari kelompok pertama dan (x_2, y_2)dari kelompok kedua. absis x_1 dan x_2 diambil dari data yang ditengah.

Misal :

Untuk kelompok data ganjil

x_1, x_2, x_3 x_4, x_5, x_6

absis $x_1 = x_2$ titik pertama $(x_2,\bar{y}_1)$

absis $x_2 = x_5$ titik pertama $(x_5,\bar{y}_2)$

Untuk kelompok data genap

x_1, x_2, x_3, x_4 x_5, x_6, x_7, x_8

absis $x_1 = (x_2 + x_3)/2$ titik pertama $((x_2 + x_3)/2,\bar{y}_1)$

absis $x_2 = (x_5 + x_6)/2$ titik pertama $((x_5 + x_6)/2,\bar{y}_2)$

e. Tentukan persamaan garis melalui dua titik tersebut.

Contoh :

Kelompok 1									Kelompok 2							
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
18	16	16	17	16	14	14	17	19	15	15	17	16	16	19	19	17
$x_1 = (x_4 + x_5)/2 =$ 16,5 $y_{rata\text{-}rata} = 16$									$x_2 = (x_{13} + x_{14})/2 = 16$ $y_{rata\text{-}rata} = 16{,}8$							

tidak digunakan

Dari data diatas didapat titik pertama (16,5,16) dan titik kedua (16,16,8)

Dengan mensubtitusi kedua titik kepersamaan garis melalui dua titik didapat persamaan garis trend

$$(y - y_1) = \frac{y_2 - y_1}{x_2 - x_1} . (x - x_1)$$

y - 16 = (0,8/-0,5)(x +16,5)

y = -1,6 x -9,6

3. Rata-Rata Bergerak (Moving Average)

Rata- rata bergerak merupakan urutan rata-rata hitung. Jika ada n data y_i, maka urutan rata-rata bergerak diberikan sebagai berikut :

$$\frac{\sum_{i=1}^{n} y_i}{n} = \frac{y_1 + y_2 + y_3 + \cdots + y_n}{n}$$

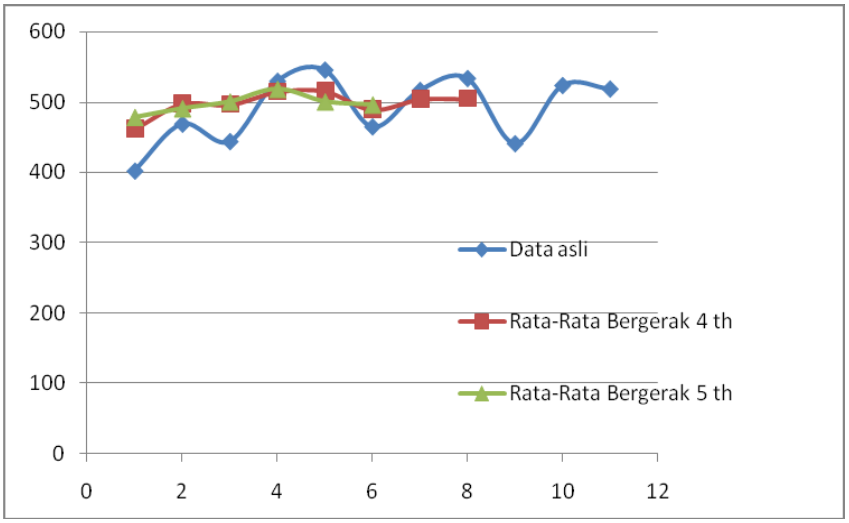
$$\frac{\sum_{i=2}^{n} y_i}{n} = \frac{y_2 + y_3 + y_4 + \cdots + y_n}{n}$$

$$\frac{\sum_{i=3}^{n} y_i}{n} = \frac{y_3 + y_4 + y_5 + \cdots + y_n}{n}$$

Contoh :

Hasil Penjualan Sepatu Plastik PT. Plastik Indah

Tahun	Penjualan (x 1000.000)
1992	426
1993	468
1994	448
1995	535
1996	535
1997	407
1998	550
1999	522
2000	544
2001	511
2002	479



Jika dibuat rata-rata bergerak 4 dan 5 lima tahun

Tahun	Penjuakan (dalam Jutaan Rupiah)	Rata-Rata bergerak	
		4 tahun	5 tahun
1992	402		
1993	469		
1994	444	461,25	
1995	530	497,25	478,2
1996	546	496,25	490,8
1997	465	514,5	500,4
1998	517	515,5	518,4
1999	534	489,25	500,6

2000	441	504	496,2
2001	524	504,5	
2002	519		

Data asli dan data rata-rata bergerak dapat dilihat pada grafik diatas. Dari grafik terlihat makin besar derajat rata-rata bergerak semakin mulus bentuk kurvanya.

4. Metode kuadrat terkecil

$$b = \frac{\sum_{i=1}^{n} t_i y_i}{\sum_{i=1}^{n} t_i^2} \qquad a = \frac{\sum_{i=1}^{n} y_i}{n}$$

Dimana t_i = waktu ke i

Contoh :

Tahun	t_i	Penjualan (000 kg) y	t_i^2	$t_i y_i$
1995	-15	18	225	-270
1996	-13	16	169	-208
1997	-11	16	121	-176
1998	-9	17	81	-153
1999	-7	16	49	-112
2000	-5	14	25	-70
2001	-3	14	9	-42
2002	-1	17	1	-17
2003	1	19	1	19
2004	3	15	9	45
2005	5	15	25	75

2006	7	17	49	119
2007	9	16	81	144
2008	11	16	121	176
2009	13	19	169	247
2010	15	19	225	285
		Σ	**1360**	**62**

$$a = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{264}{16} = 16{,}5$$

$$b = \frac{\sum_{i=1}^{n} t_i y_i}{\sum_{i=1}^{n} t_i^{\,2}} = \frac{62}{1360} = 0{,}046$$

Subtitusi nilai a, b ke bentuk umum persamaan linier y = a + bx, maka didapat persamaan trend

$$y = 16{,}5 + 0{,}046\, x$$

Berdasarkan garis trend diatas, estimasi penjualan tahun 2011

y = 16,5 + 0,046 * 17 = 17,82

Soal :

Penjualan PT. Tali Plastik selama 11 tahun diberikan sebagai berikut

Tahun	**Penjualan (dalam Jutaan Rupiah)**

1992	323
1993	257
1994	326
1995	295
1996	275
1997	275
1998	346
1999	250
2000	247
2001	266
2002	336

Berdasarkan data diatas.

a. Tentukan trend data menggunakan semi rata-rata
b. Metode kuadrat terkecil
c. Metode rata-rata bergerak 4 dan 5 tahun
d. Gambar grafik penjualan data asli, rata-rata bergerak 4 dan 5 tahun dalam satu gambar.

Daftar Pustaka

- Bambang Kustituanto dan Rudy Badrudin, Statistika I (Deskriptif), Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Haryono Subiyakto, Statistika 2, Seri Diktat Kuliah, Penerbit Gunadarma, Jakarta, 1994
- Levin, Richard I & David Rubin, Statistika for management, Prentice Hall, New Jersey, 1991
- Ronald E Walpole, Pengantar Statistika, Edisi Terjemahan, PT Gramedia, Jakarta 1992