

MODUL STATISTIK 1



Dr Dhian Tyas Untari SE.,MM

KATA PENGANTAR

Statistik merupakan salah satu mata kuliah untuk mahasiswa pada Strata 1, khususnya pada Fakultas Ekonomi. Statistik nantinya akan digunakan dalam penyelesaian penelitian tugas akhir, dan hal tersebut yang mendasari penyusunan buku Statistik 1 ini. Dan alhamdulillah, penulis panjatkan puji syukur atas kehadiran Allah SWT karena Modul ini telah selesai disusun. Penulis pun menyadari jika didalam penyusunan Modul ini mempunyai kekurangan, namun penulis meyakini sepenuhnya bahwa sekecil apapun Modul ini tetap akan memberikan sebuah manfaat bagi pembaca.

Akhir kata untuk penyempurnaan Modul ini, maka kritik dan saran dari pembaca sangatlah berguna untuk penulis kedepannya.

Bekasi, Juni 2016

Penulis

DAFTAR ISI

Statistik 1
Dr Dhian Tyas Untari SE.,MM

KATA PENGANTAR

DAFTAR ISI

BAB I. Pengenalan Statistik Dasar	3
BAB II. Distribusi Frekuensi	12
BAB III. Ukuran Pemusatan Data	20
BAB IV. Ukuran Penyebaran, Dispersi, dan Variasi	33
BAB V. Ukuran Kemiringan dan Kemencengan Dta	41
BAB VI. Regresi Sederhana	48
BAB VII. Regresi Berganda	56
BAB VIII Korelasi dan Koefisien Determinasi	60

DAFTAR ISI

BAB I

Statistik 1

Dr Dhian Tyas Untari SE.,MM

PENGENALAN STATISTIKA DASAR

Tujuan pembelajaran : Mahasiswa memahami dasar dari statistic dan dapat memberi gambaran implementasi dan kebutuhan mahasiswa terhadap metode statistic dalam mendukung keberlanjutan studi.

Kita mungkin biasa mengungkapkan beberapa istilah terkait statistik namun kita kurang memahami secara jelas apa arti dan makna dari kata tersebut, sehingga hal ini akan mereduksi arti dan makna kata tersebut yang menyebabkan pemahaman menjadi bias.

Beberapa contoh kasus yang berkaitan dengan statistik dan sering kita jumpai sebagai berikut:

- Rata-rata berat badan dari mahasiswa yang mengikuti kuliah ini adalah 51 kg.
- 90% dari mahasiswa yang mengikuti kuliah ini berasal dari kota "A".
- Kecelakaan lalu lintas itu kebanyakan diakibatkan karena kecerobohan pengemudi angkutan kota.

Dan untuk memahami lebih lanjut mengenai arti statistik, berikut definisi beberapa istilah yang sering digunakan oleh masyarakat maupun akademisi terkait dengan statistik itu sendiri.

Statistik dalam arti sempit dapat dijelaskan sebagai kumpulan data yang berupa angka, seperti statistik penduduk maupun statistik ekonomi. Data yang dinyatakan dalam bentuk angka dan data tersebut dapat ditampilkan dalam bentuk deretan angka, dibuatkan tabel ataupun grafik.

Statistik dalam arti luas atau disebut sebagai **Statistika** adalah kumpulan metoda yang digunakan untuk merencanakan eksperimen, meliputi proses pengambilan/pengumpulan data, penyusunan dan peringkasan data, penyajian data, penganalisaan, penginterpretasian dan pengambilan kesimpulan yang didasarkan pada data tersebut.

Data adalah hasil observasi atau pengamatan yang telah dikumpulkan. Data dapat berupa hasil pengukuran; misalnya data tinggi dan berat badan, hasil

pengelompokan; misalnya jenis kelamin, hasil jawaban responden terhadap suatu questioner; misalnya tingkat kepuasan.

Populasi adalah koleksi lengkap semua elemen yang akan diselidiki. Suatu koleksi dikatakan lengkap jika ia memuat semua subjek yang akan diselidiki.

Sensus adalah koleksi data dari semua anggota dalam populasi.

Sampel adalah sebagian koleksi anggota yang dipilih dari populasi.

Sampling adalah bagian dari metodologi statistika yang berhubungan dengan teknik pengambilan sebagian dari populasi.

Statistika Deskriptif adalah statistika yang berkaitan dengan analisis dan deskripsi suatu grup sebagai populasinya, tanpa melakukan penarikan kesimpulan apapun untuk komunitas yang lebih luas dari grup tersebut.

Statistika Inferensi adalah statistika yang mencoba untuk membuat suatu deduksi atau kesimpulan pada populasi dengan menggunakan sampel dari populasi tersebut

Statistik Univariate adalah analisis data statistik yang dilakukan untuk satu variabel atau per variabel atau disebut juga statistik deskriptif

Statistik Bivariate adalah analisis data statistik yang dilakukan untuk menganalisis hubungan dua variabel

Statistik Multivariate adalah analisis data statistik yang dilakukan untuk menganalisis hubungan lebih dari dua variabel

Statistik Parametrik adalah analisis data statistik yang dilakukan untuk menguji parameter/berdasarkan asumsi-asumsi tertentu dan biasanya salah satu asumsinya adalah distribusi normal.

Statistik Non Parametrik adalah analisis yang dilakukan tidak untuk menguji parameter/tidak berdasarkan asumsi-asumsi tertentu.

TAHAP-TAHAP KEGIATAN STATISTIK

Statistik 1

Dr Dhian Tyas Untari SE.,MM

Setelah membahas definisi statistik sebagai metode maka sesuai dengan definisi di atas kita dapat membagi kegiatan statistik ke dalam tahap-tahap sebagai berikut:

1. Pengambilan/pengumpulan Data (*Collection of Data*)

Tahap kegiatan statistik yang pertama adalah pengumpulan data. Ada 2 cara atau metode pengumpulan data yaitu:

1. Pengumpulan data secara keseluruhan disebut metode sensus (*census*).
Misalkan : Kepala SMA "X" ingin mengetahui rata-rata tinggi badan siswa-siswa di sekolahnya yang berjumlah 600 orang. Apabila setiap siswa diukur tinggi badannya, kemudian dicatat maka cara pengumpulan data seperti ini dinamakan sensus.
2. Pengumpulan data berdasarkan sampel (*sample*) disebut metode sampel. Jadi, di sini tidak semua anggota populasi yang diteliti, tetapi hanya sebagian anggota populasi saja yang diteliti. Akan tetapi, yang sebagian itu harus menggambarkan keadaan populasi yang sebenarnya. Dengan demikian, sebagian dari anggota populasi itu dikatakan bersifat representatif.

Misalkan : Apabila jumlah siswa yang diukur tinggi badannya hanya 60 orang saja, dengan perincian:

Kelas I diambil 20 orang siswa,

Kelas II diambil 20 orang siswa,

Kelas III diambil 20 orang siswa,

Maka cara pengumpulan data seperti ini dinamakan sampling. Banyak anggota populasi biasanya dinotasikan dengan N dan banyak anggota sampel biasanya dinotasikan dengan n .

Untuk memilih sampel dari suatu populasi dapat dilakukan dalam dua cara yaitu :

- Cara acak, melalui undian ataupun menggunakan tabel bilangan acak.
- Cara tidak acak adalah cara pemilihan sejumlah anggota dari populasi dengan setiap anggotanya tidak mempunyai peluang yang sama untuk terpilih menjadi anggota sampel.

2. Penyusunan dan peringkasan Data (*Organization of Data*)

Tahap berikutnya setelah data dapat dikumpulkan adalah menyusun data dalam susunan yang teratur agar mudah dibaca dan dilihat secara visual. Kegiatan penyusunan data ini dapat dibedakan dalam tiga tahap, yaitu:

- a. Editing
Editing adalah kegiatan mendeteksi adanya kemungkinan kesalahan, ketidakkonsistenan, dan ketidakteraturan atau ketidaktepatan dari data yang telah dikumpulkan.
- b. Klasifikasi
Kegiatan klasifikasi adalah kegiatan mengelompokkan data sesuai dengan sifat-sifat yang dimiliki oleh data. Kegiatan ini dilakukan setelah kita melaksanakan editing.
- c. Tabulasi
Tabulasi adalah kegiatan untuk mengadakan pengelompokan data sesuai dengan sifat-sifat data yang telah kita tentukan dalam susunan kolom-kolom dan baris-baris, sehingga data tersebut mudah ditarik kesimpulannya.

3. Penyajian Data (Presentation of Data)

Penyajian data dimaksudkan agar data yang telah disusun dapat disebarluaskan dan mudah dilihat secara visual. Supaya data tersebut dapat dengan mudah dibaca dan dilihat secara visual maka data tersebut dibuat dalam bentuk tabel, grafik maupun diagram.

4. Analisis Data (Analysis of Data)

Analisis data adalah kegiatan menganalisis data yang sudah dikumpulkan dan telah disusun. Analisis data dilakukan dengan menggunakan metode statistik seperti: rata-rata hitung, penyimpangan, regresi maupun korelasi. Dengan melakukan analisis data, kita dapat memperoleh gambaran keseluruhan dari data yang telah dikumpulkan.

5. Interpretasi Data (Interpretation of Data)

Interpretasi data merupakan kegiatan yang paling sulit karena memerlukan keahlian tinggi, sikap hati-hati, pertimbangan yang masak, dan sikap objektif. Apabila kegiatan interpretasi data dapat dilakukan dengan baik maka akan dapat diambil suatu kesimpulan yang baik pula.

PEMBAGIAN DATA MENURUT SIFATNYA

Data merupakan kumpulan fakta atau angka atau segala sesuatu yang dapat dipercaya kebenarannya, sehingga dapat digunakan sebagai dasar menarik suatu kesimpulan.

Pada suatu penelitian, data adalah hasil observasi atau pengamatan yang telah dikumpulkan. Data dapat berupa hasil pengukuran seperti: data tinggi dan berat badan, hasil pengelompokan seperti: jenis kelamin, hasil jawaban responden terhadap suatu questioner seperti: tingkat kepuasan.

Berdasarkan sifatnya, data dapat dibagi menjadi dua:

1. **Data Kuantitatif** adalah data yang berbentuk bilangan yang menggambarkan hasil perhitungan atau hasil pengukuran.

Contoh :

- Luas bangunan hotel itu adalah 5.700 m².
- Tinggi badan sandy mencapai 170 cm.
- Banyak perguruan tinggi di kota "B" ada 4 buah.

2. **Data Kualitatif** adalah data yang berbentuk kategori atau atribut. Data kualitatif dapat dipisahkan dalam beberapa kategori atau kelompok yang dibedakan oleh karakter bukan numerik..

Contoh :

- Harga emas hari ini mengalami kenaikan.
- Sebagian dari produksi barang "A" pada perusahaan "X" rusak.

SKALA PENGUKURAN DATA

Skala pengukuran dapat dibedakan menjadi empat macam, yaitu :

1. Skala pengukuran nominal.
2. Skala pengukuran ordinal.
3. Skala pengukuran interval.
4. Skala pengukuran rasio.

1. Skala Nominal

Statistik 1

Dr Dhian Tyas Untari SE.,MM

Pada data nominal, data digabungkan pada kriteria yang jelas dan tegas serta bersifat diskrit. Kelompok yang satu dengan yang lain pada skala nominal tidak dapat dikatakan yang satu lebih tinggi dari yang lain.

Contoh:

Karyawan di suatu instansi dikelompokkan menjadi beberapa kelompok: suku Jawa, Sunda, Bali, Batak, Dayak, dan Madura.

Pada data seperti ini pengukurannya dilakukan dengan cara menjumlahkan frekuensinya.

Misalnya jumlah karyawan di instansi ini sebanyak 100 orang, angka ini diperoleh dari penjumlahan nominal: karyawan suku Jawa 30, Sunda 20, Bali 15, Batak 10, Dayak 10, dan Madura 15.

2. Skala Ordinal

Data dikumpulkan pada urutan, misalnya: tinggi, sedang, dan rendah. Pada pengelompokan ini kita tidak dapat membedakan nilai data antara kelompok yang satu dengan yang lain, sehingga tidak dapat dipergunakan dalam perhitungan.

Contoh:

Pengelompokan penghasilan masyarakat dalam kelompok: penghasilan rendah, penghasilan sedang, dan penghasilan tinggi.

Kelompok yang satu dapat dibedakan lebih rendah dan lebih tinggi dari pada kelompok yang lain, tetapi jumlah besarnya perbedaan antara kelompok yang satu dengan yang lain tidak dapat dihitung.

3. Skala Interval

Skala ini dipergunakan untuk menunjukkan adanya pengelompokan yang mempunyai besaran yang sama. Pada skala ini nilai 0, mempunyai arti yang relatif, bukan harga 0 secara mutlak. Nilai dari 0° Fahrenheit sebagai titik awal pengukuran bukan berarti tidak mempunyai nilai suhu.

Contoh:

Tahun 1000, 2000 merupakan skala interval, yang dinilai dari 0 tetapi bukan dalam arti mutlak.

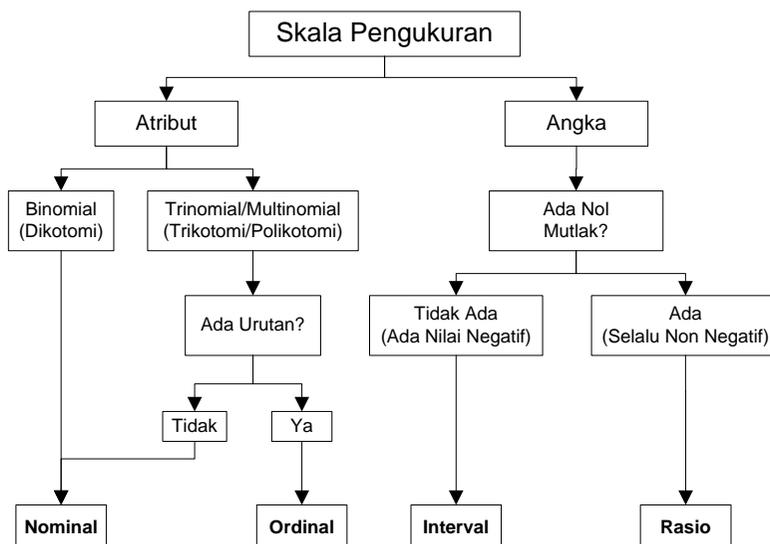
4. Skala rasio

Skala rasio memiliki skala yang hampir sama dengan skala interval. Pada skala rasio nilai 0 merupakan nilai mutlak, titik 0 pada skala panjang menunjukkan tidak ada panjang. Data pada skala rasio mempunyai kualitas bilangan riil yang dapat dijumlahkan, dikurangkan, dikalikan, dan dibagi.

Contoh:

Berat badan dalam satuan kg, si A beratnya 40 kg sedang si B beratnya 80 kg, jadi berat si B 2 kali dibanding berat si A.

Gambar 1.1 Algoritma Pemeriksaan Skala Pengukuran



DATA INTERN DAN DATA EKSTERN

Menurut sumber data yang dikumpulkan, data dapat dibagi menjadi dua yaitu data intern dan data ekstern.

Data intern adalah data yang dikumpulkan oleh suatu badan mengenai kegiatan badan itu, dan hasilnya digunakan untuk kepentingan badan itu.

Contoh :

Sebuah perusahaan mengadakan penelitian tentang produktivitas tenaga kerja di perusahaan tersebut.

Sedangkan **data ekstern** adalah data yang diperoleh dari luar badan yang memerlukannya.

Contoh :

Data tentang jumlah penganggur di negara tempat perusahaan tadi beroperasi.

Data tersebut dapat diperoleh dari hasil penelitian departemen tenaga kerja. Sehingga bagi badan tersebut data jumlah penganggur merupakan data ekstern.

DATA PRIMER DAN DATA SEKUNDER

Data ekstern dapat dibagi lagi menjadi dua yaitu data primer dan data sekunder.

Data primer adalah data ekstern yang diperoleh dari hasil penelitian sendiri, misalnya :

Suatu perusahaan ingin mengetahui preferensi konsumen terhadap produk yang dihasilkan maka perusahaan tersebut mengadakan survei pasar.

Pemerintah melalui Biro Pusat Statistik (BPS) ingin mengetahui jumlah penduduk Indonesia, maka BPS mengirimkan petugaspetugasnya untuk mendatangi secara langsung rumah tangga yang ada di Indonesia.

Perusahaan susu "SEGAR JAYA" ingin mengetahui jumlah konsumsi susu yang diminum oleh masyarakat di Kelurahan Kejaksaan maka petugas dari perusahaan tersebut secara langsung mendatangi rumah tangga yang ada di Kelurahan Kejaksaan.

Sementara itu, **data sekunder** adalah data yang diperoleh dari pihak lain atau dari hasil penelitian orang lain. Biasanya ada lembaga-lembaga tertentu yang memang bertugas untuk mengumpulkan data. Badan lain yang membutuhkan data, tinggal mengambil saja dari lembaga yang mengumpulkan data tersebut.

Jadi, badan yang menggunakan data bukan badan yang mengumpulkan data, misalnya:

Seorang mahasiswa mengumpulkan data untuk keperluan pembuatan skripsi. Ia mengambil data dari BPS dan melakukan wawancara dengan responden.

Jadi, data yang diperoleh dari BPS merupakan data sekunder. Sementara itu, data yang diperoleh dari hasil wawancara dengan responden merupakan data primer bagi mahasiswa tersebut.

BAB II

DISTRIBUSI FREKUENSI

Tujuan pembelajaran: mahasiswa dapat memahami metode pengelompokan data dan penyajian data dalam bentuk table sehingga dapat memudahkan mahasiswa dalam menginterpretasikan data secara deskriptif

PENYUSUNAN DATA

Sebelum membahas lebih lanjut tentang distribusi frekuensi, perlu dipahami tentang penyusunan data dalam urutan sistematis yang disebut dengan istilah *seriation*. Penyusunan data ini dapat dilakukan:

1. Berdasarkan waktu (*time series, chronological, historical series*).

Contoh :

Tabel. 2.1. Produksi Padi Kabupaten “X”
(dalam ribuan ton)

Tahun	Produksi Padi
2001	11
2002	13
2003	15
2004	16
2005	17
2006	19
2007	22
2008	25
2009	26

2. Berdasarkan daerah/wilayah (*geographical series, cluster*).

Contoh :

Tabel. 2.2. Produksi Padi Provinsi “A”, tahun 2009
(dalam ribuan ton)

Kabupaten	Produksi Padi
W	24
X	26
Y	30
Z	25

3. Berdasarkan keadaan/frekuensi (*frequency, conditional series*).

Penyusunan data berdasarkan kondisi fisik seperti: tinggi, berat, ataupun metode gradasi yang lain, berdasarkan banyaknya kejadian di suatu tempat tertentu dan waktu tertentu.

Tabel. 2.3. Jumlah masyarakat Desa A berdasarkan kategori usia

Umur	Frekuensi
20 – 29	5
30 – 39	10
40 – 49	15
50 – 59	13
60 – 69	7
Jumlah	50

PENYUSUNAN DATA BERDASARKAN FREKUENSI

Penyusunan data berdasarkan keadaan/frekuensi ini dapat dilakukan dengan 2 cara:

1. Secara individual (*Ungrouped Data*).

Metode ini merupakan cara menyusun data sesuai dengan hasil observasi. Data primer yang diperoleh pada suatu observasi secara langsung oleh peneliti disebut dengan data mentah (*raw data*). Sebagai contoh data mengenai nilai ujian mahasiswa yang penyajiannya masih dalam bentuk table nilai yang

biasanya hanya diurutkan berdasarkan alphabet nama mahasiswa. Data mentah yang disajikan berdasarkan urutan naik disebut ascending dan urutan turun disebut descending dan bentuk penyajian seperti ini disebut array.

Contoh:

Data mentah

2	3	3	1	4
3	4	2	2	1
1	2	1	5	3
4	6	4	3	2
2	1	5	2	5

Data setelah diurutkan :

1	2	2	3	4
1	2	2	3	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	6

2. Secara berkelompok (*Grouped Data*)

Selisih antara nilai data terbesar dan terkecil disebut rentang (range) atau jangkauan data. Untuk memudahkan bekerja jika data yang dimiliki cukup besar lebih baik disajikan dalam bentuk kelas-kelas (*group data*) berdasarkan kategori-kategori tertentu bersamaan dengan frekuensi yang bersesuaian. Suatu tabel yang menyajikan kelas-kelas data beserta frekuensinya disebut distribusi frekuensi atau tabel frekuensi. Contoh tabel frekuensi adalah sebagai berikut,

Tabel 2.4 Tabel Frekuensi Tinggi Badan Mahasiswa

No	Tinggi Badan	Frekuensi	Mid Class
1.	160 – 162	6	161
2.	163 – 165	7	164
3.	166 – 168	11	167
4.	169 – 171	4	170
Total		30	

Berdasarkan data diatas, ada beberapa istilah yang harus dipahami, yaitu :

Interval kelas adalah interval yang diberikan untuk menetapkan kelas-kelas dalam distribusi. Pada tabel 2.1, interval kelasnya adalah 160-162, 163-165, dan seterusnya. Bilangan 160 dan 162 pada kelas pertama disebut limit kelas, dimana angka 160 disebut limit kelas bawah dan angka 162 disebut limit kelas atas.

Batas kelas adalah bilangan terkecil dan terbesar sesungguhnya yang masuk dalam kelas interval tertentu. Misalnya jika dalam pengukuran tinggi badan di atas dilakukan dengan ketelitian 0.5 cm maka pada kelas pertama, batas kelas bawah adalah 159,5 dan batas kelas atas adalah 162.5.

Lebar interval kelas adalah selisih antara batas atas dan batas bawah batas kelas. Misalnya pada kelas pertama tabel 2.4 adalah $162.5 - 159.5 = 3$.

Mid class adalah titik tengah interval kelas. Ia diperoleh dengan cara membagi dua jumlah dari limit bawah dan limit atas suatu interval kelas. Contoh mid class untuk kelas pertama tabel 2.4 adalah $(160+162)/2 = 161$

Dalam pengolahan data statistik, seringkali kita dihadapkan dengan rekapitulasi *ungrouped data* yang cukup banyak, sehingga mengharuskan untuk dibuat pengelompokan data berdasarkan kelas-kelas tertentu. Berikut langkah-langkah untuk membuat tabel frekuensi:

Contoh data nilai mahasiswa :

21, 15, 15, 16,19, 20, 77, 80, 86, 90

27, 55, 46, 43, 55, 78, 80, 70, 80, 77

19, 21, 22, 47, 89, 75, 56, 60, 39, 34

78, 81, 46, 26, 18, 82, 85, 63, 67, 62

28, 28, 84, 74, 62, 90, 54, 46, 71, 69

1. Tentukan Range (R) = data terbesar – data terkecil.
Misalnya data terbesar = 90 dan data terkecil = 15
Range (R) = $90 - 15 = 75$
2. Tentukan banyak kelas interval.
 $k = 1 + 3,3 \log n$, dimana n = banyaknya data
Misalnya diketahui jumlah $n = 50$. Maka
 $k = 1 + 3,3 \log 50$
 $k = 1 + 3,3 (1,69)$
 $k = 1 + 5,61$
 $k = 6,61 \simeq 7$
Jadi banyaknya kelas yang harus dibuat adalah 7 kelas
3. Tentukan panjang interval kelas (C)
 $C = \text{Range (R)} / \text{Banyaknya kelas interval (k)}$
Sehingga $c = 75 / 7 = 10,7 \simeq 11$. Jadi, panjang interval kelas adalah 11
4. Menentukan kelas pertama. Ambil data terkecil sebagai batas bawah kelas pertama. Jumlahkan datum terkecil dengan panjang interval kelas kemudian kurangi satu (1).
Panjang interval kelas pertama = $(15 + 11) - 1 = 25$. Jadi interval kelas pertama adalah (15 – 25)

Sehingga tabel frekuensi yang didapatkan adalah sebagai berikut,

Tabel 2.5. Tabel Frekuensi Nilai Mahasiswa

No	Nilai	Frekuensi
1.	15 – 25	10
2.	26 – 36	5
3.	37 – 47	6
4.	48 – 58	4
5.	59 – 69	8
6.	70 – 80	11
7.	81 – 91	6
Total		50

Histogram dan Poligon

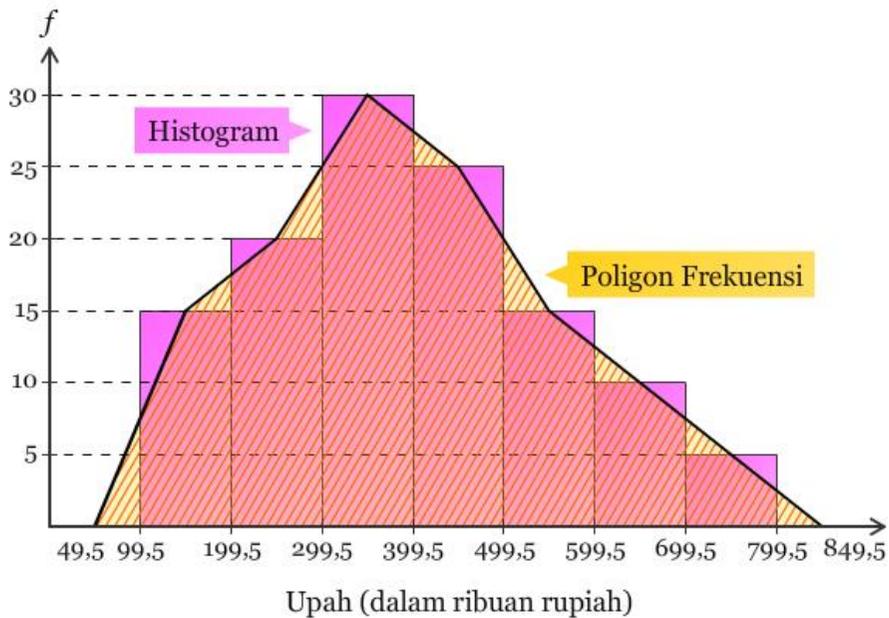
Histogram dan poligon merupakan representasi grafik untuk distribusi frekuensi. Histogram berupa sekumpulan persegi panjang dengan alas pada sumbu X, pusat alasnya adalah tanda kelas dan lebar alasnya adalah lebar kelas interval. Tinggi merupakan frekuensi pada kelas yang bersangkutan. Sedangkan poligon merupakan grafik garis yang mengaitkan frekuensi kelas dengan tanda kelas. Ia dapat digambarkan dengan menghubungkan garis lurus yang melalui titik-titik pasangan frekuensi kelas dan titik tengah (tanda) interval kelas.

Contoh,

Tabel 2.6 Distribusi Upah Perminggu dalam Rp

Interval Kelas	Frekuensi
100 – 199	15
200 – 299	20
300 – 399	30
400 – 499	25
500 – 599	15
600 – 699	10
700 – 799	5

Gambar 2.1 Grafik Histogram dan Poligon Distribusi Upah Perminggu



Ogive

Ogive adalah grafik yang digambarkan berdasarkan data yang sudah disusun dalam bentuk tabel distribusi frekuensi kumulatif. Untuk data yang disusun dalam bentuk tabel distribusi frekuensi kumulatif kurang dari, grafiknya berupa ogive positif, sedangkan untuk data yang disusun dalam bentuk tabel distribusi frekuensi kumulatif lebih dari, grafiknya berupa ogive negatif.

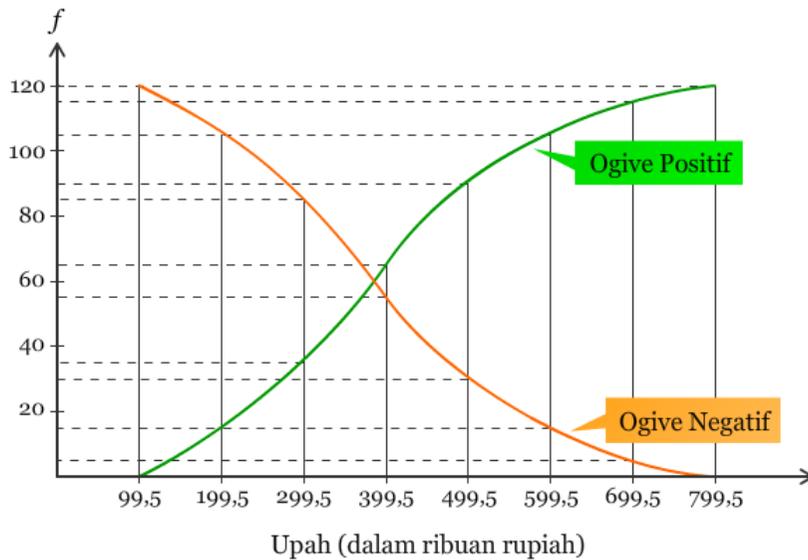
Frekuensi kumulatif kurang dari untuk suatu kelas adalah jumlah frekuensi semua kelas sebelum kelas tersebut dengan frekuensi kelas itu. Sedangkan frekuensi kumulatif lebih dari suatu kelas adalah jumlah frekuensi semua kelas sesudah kelas tersebut dengan frekuensi kelas itu.

Contoh,

Tabel 2.7 Distribusi Frekuensi Lebih Dari dan Kurang Dari

Interval Kelas	Frekuensi	Fk kurang dari		Fk lebih dari	
		\leq		\geq	
100 – 199	15	≤ 199	15	≥ 100	120
200 – 299	20	≤ 299	35	≥ 200	105
300 – 399	30	≤ 399	65	≥ 300	85
400 – 499	25	≤ 499	90	≥ 400	55
500 – 599	15	≤ 599	105	≥ 500	30
600 – 699	10	≤ 699	115	≥ 600	15
700 – 799	5	≤ 799	120	≥ 700	5

Gambar 2.2 Grafik Ogive Distribusi Upah Perminggu



BAB III

UKURAN PEMUSATAN DATA

Tujuan pembelajaran : mahasiswa dapat mempelajari beberapa ukuran yang dapat memberikan informasi tentang bagaimana data-data ini mengumpul atau memusat.

PENGERTIAN MEAN, MEDIAN, MODIS

Mean adalah nilai Rata-rata yang didapat dari hasil penjumlahan seluruh nilai dari masing-masing data, kemudian dibagi dengan banyaknya data yang ada.

Median adalah nilai tengah dari kumpulan data yang telah diurutkan, jika jumlah data Ganjil maka Nilai Median adalah satu nilai yang berada ditengah urutan, namun jika jumlah data Genap maka Mediannya adalah hasil penjumlahan dua nilai yang berada ditengah urutan data, kemudian hasilnya dibagi dua.

Modus adalah data atau nilai yang paling sering muncul atau yang memiliki jumlah frekuensi terbanyak.

MEAN

Rata-rata hitung atau Mean memiliki perhitungan dengan cara membagi jumlah nilai data dengan banyaknya data. Rata-rata hitung disebut dengan Mean.

Mean pada data yang belum dikelompokkan/ data tunggal :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

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

Contoh :

Data : 2 , 2 , 4 , 5 , 5 , 7 , 7

$$\bar{x} = \frac{2+2+4+5+5+7+7}{7} = 4,57$$

Mean pada data yang telah dikelompokkan/ disajikan dalam tabel distribusi frekuensi.

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + f_3x_3 + \dots + f_nx_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

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

Keterangan :

$f_i x_i$ = frekuensi untuk nilai x_i yang bersesuaian

x_i = data ke-i.

Contoh perhitungan Mean dalam data berkelompok

Tabel 3.1 Mean pada data nilai mahasiswa

No	Nilai	Frekuensi	x	f.x
1	15 – 25	10	20	200
2	26 – 36	5	31	155
3	37 – 47	6	41	246
4	48 – 58	4	51	204
5	59 – 69	8	61	488
6	70 – 80	11	71	781
7	81 – 91	6	81	486
	Total	50		2.560

$$\text{Mean : } \bar{x} = \frac{2.560}{50} = 51,2$$

MODUS

Modus pada data yang belum dikelompokkan.

Modus dari data yang belum dikelompokkan adalah ukuran yang memiliki frekuensi tertinggi atau nilai data yang paling sering muncul. Modus dilambangkan M_o .

Contoh:

Tabel 3.2 Modus Pada Data Tunggal

Data	Modus
2, 2, 2, 3, 4, 4, 5, 7	2
3, 4, 5, 5, 5, 6, 6, 7, 8, 8, 8	5 dan 8
1, 3, 4, 5, 7, 8, 9, 10	Tidak ada

Modus pada data yang telah dikelompokkan.

Rumus Modus dari data yang telah dikelompokkan dapat dihitung dengan menggunakan rumus:

$$M_o = L + C \frac{b_1}{b_1 + b_2}$$

Keterangan :

- M_o = Modus
- L = Tepi bawah kelas yang memiliki frekuensi tertinggi (kelas Modus)
- C = Interval kelas
- b_1 = Frekuensi kelas Modus dikurangi frekuensi kelas sebelumnya
- b_2 = Frekuensi kelas Modus dikurangi frekuensi kelas sesudahnya.

Contoh:

Tabel 3.2 Table nilai mahasiswa

No	Nilai	Frekuensi
1	15 – 25	10
2	26 – 36	5
3	37 – 47	6
4	48 – 58	4
5	59 – 69	8
6	70 – 80	11
7	81 – 91	6
Total		50

Pada tabel 3.2, kelas yang mengandung Modus adalah kelas ke 6 (frekuensi terbanyak). Sehingga diketahui $L=69,5$ $b_1=3$ dan $b_2=5$. Sehingga

$$M_o = 69,5 + 11 \times \frac{3}{3 + 5} = 69,5 + 4,125 = 73,625$$

Keunggulan Modus:

1. Dapat digunakan untuk data kualitatif maupun kuantitatif.
2. Tidak dipengaruhi oleh data ekstrim.
3. Dapat dihitung untuk data berkelompok dengan kelas terbuka.

Keunggulan Modus:

1. Dalam kasus-kasus tertentu, kumpulan data tidak memiliki Modus.
2. Jika Modus justru lebih daripada satu, tidak dapat digunakan sebagai ukuran pusat data
- 3.

MEDIAN

Median pada data yang belum dikelompokkan.

Untuk mencari nilai Median, data harus dikelompokkan terlebih dahulu dari yang terkecil sampai yang terbesar.

$$M_e = Q_2 = \begin{cases} \frac{X_{n+1}}{2} & ; \text{Jika } n \text{ ganjil} \\ \frac{\frac{X_n + X_{n+1}}{2}}{2} & ; \text{Jika } n \text{ Genap} \end{cases}$$

Contoh :

1. $n = 6 \rightarrow$ Genap:

Data : 2, 3, 5, 6, 7, 8 \rightarrow Median $\rightarrow \frac{x_3 + x_4}{2}$

$$M_e = \frac{5 + 6}{2} = 5,5$$

2. $n = 5 \rightarrow$ Ganjil

Data : 3, 5, 7, 8, 9 \rightarrow Median $\rightarrow x_{\frac{5+1}{2}} = x_3$

$$M_e = 7$$

3. Data : 2, 2, 4, 5, 5, 7, 7 \rightarrow Me = ?

$n = 7 \rightarrow$ Ganjil \rightarrow Median $\rightarrow x_{\frac{7+1}{2}} = x_4$

$$M_e = 5$$

Median pada data yang telah dikelompokkan.

Rumus Median data berkelompok sama dengan rumus Quartil, Desil dan Presentil. Pembedanya hanya pada nilai pembagi. Dimana pada:

- Median pembaginya 2
- Quartil pembaginya adalah 4, contoh $Q1 = \frac{1}{4}$
- Desil pembagiannya 10, contoh $D7 = \frac{7}{10}$
- Persentil pembaginya 100, contoh $P20 = \frac{20}{100}$.

Selebihnya rumus dan keterangannya adalah sama.

$$\text{Median : } M_e = L_2 + C \cdot \frac{\frac{n}{2} - (\Sigma f)_2}{f_{M_e}}$$

$$\text{Kuartil : } Q_i = L_i + C \cdot \frac{\frac{in}{4} - (\Sigma f)_i}{f_{Q_i}} \quad i = 1, 2, 3 \text{ atau } 4$$

$$\text{Desil : } D_i = L_i + C \cdot \frac{\frac{in}{10} - (\Sigma f)_i}{f_{D_i}} \quad i = 1, 2, 3 \dots 9$$

$$\text{Persentil : } P_i = L_i + C \cdot \frac{\frac{in}{100} - (\Sigma f)_i}{f_{P_i}} \quad i = 1, 2, 3, \dots .99$$

Keterangan

n = Banyak data

L_2 = Tepi bawah kelas Median

L_i = Tepi bawah kelas Kuartil/Desil/Persentil ke-i

C = Interval kelas

$(\Sigma f)_2$ = Frekuensi kumulatif sebelum kelas Median

$(\Sigma f)_i$ = Frekuensi kumulatif sebelum kelas Kuartil/Desil/Persentil ke-i

f_{M_e} = Frekuensi kelas Median

$f_{Q_i}/f_{D_i}/f_{P_i}$ = Frekuensi kelas Kuartil/Desil/Persentil ke-i

Contoh perhitungan Median dalam data berkelompok

Tabel 3.2 Tabel nilai mahasiswa

No	Nilai	Frekuensi
1	15 – 25	10
2	26 – 36	5
3	37 – 47	6
4	48 – 58	4
5	59 – 69	8
6	70 – 80	11
7	81 – 91	6
Total		50

Median ada di data ke : $50/2 = 25$. Data ke-25 secara kumulatif berada di kelas ke-4 (48-58).

Berdasarkan tabel 3.2 maka diketahui :

$L_2 = 47,5$; $C = 11$; $\frac{1}{2}n = 25$; $(\Sigma f)_2 = 21$ dan $fM_e = 4$.

Sehingga Median = $47,5 + 11 \frac{25-21}{4} = 58,5$.

Keunggulan Median:

1. Tidak dipengaruhi oleh data ekstrim.
2. Mudah dimengerti dan mudah dihitung, baik dari data tak-berkelompok maupun data berkelompok. Juga dapat dihitung untuk data berkelompok dengan kelas terbuka.
3. Dapat digunakan untuk data kuantitatif maupun data kualitatif.

Kelemahan Median:

1. Hanya dapat ditentukan dari data yang telah diurutkan sehingga membutuhkan waktu yang tidak sedikit
2. Dihitung bukan berdasarkan nilai data, tetapi berdasarkan jumlah data, sehingga sulit dijadikan sebagai ukuran pusat data untuk menggambarkan kumpulan datanya

RATA-RATA POLAR (ANDI SUPANGAT)

Rata-rata polar merupakan cara perhitungan rata-rata yang didasarkan pada data-data yang terpusat titik-titik sentralnya (data awal, data tengah, data akhir), setelah kumpulan data diurutkan dari data yang terkecil hingga data yang terbesar.

Rata-rata Polar pada data yang belum dikelompokkan

- Data Ganjil : $P_A = \frac{D_A + D_T + D_B}{3}$

P_A = Polar Average (Rata-rata Polar)

D_A = Data awal

D_T = Data tengah

D_B = Data akhir

Contoh :

Jika diketahui kelompok data berikut :

2, 3, 7, 7, 9, 9, 9, 11, 11, 13, 14, 16, 17, 18, 21, 23, 23, 24, 24, 26, 27

Tentukan rata-rata polarnya !

$$P_A = \frac{D_A + D_T + D_B}{3} = \frac{2 + 14 + 27}{3} = 14,33$$

- Data Genap : $P_A = \frac{D_A + (\frac{DT_1 + DT_2}{2}) + D_B}{3}$

$(DT_1 + DT_2)$ = jumlah dua data tengah

Contohnya

Jika diketahui kelompok data berikut :

12, 13, 14, 14, 14, 14, 14, 16, 16, 17, 18, 22, 23, 25, 25, 27, 27, 29, 30, 32, 34, 35

Tentukan rata-rata polarnya !

$$P_A = \frac{D_A + (\frac{DT_1 + DT_2}{2}) + D_B}{3} = \frac{12 + \frac{18+22}{2} + 35}{3} = 22,33$$

Rata-rata Polar pada data yang telah dikelompokkan

$$P_A = \frac{(M_R) \cdot (\sum F_i \cdot T_i)}{P \cdot \sum F_i}$$

P_A = Rata-rata polar

M_R = Middle range

$$MR = \frac{N_B - N_A}{2}$$

N_A = Nilai ujung bawah kelas interval pertama

N_B = Nilai ujung atas kelas interval terakhir

TI = Titik interval (jumlah ujung bawah kelas dan ujung atas kelas dibagi jumlah kelas)

$\Sigma F_i \cdot TI$ = Jumlah frekuensi x titik interval

P = Panjang kelas

ΣF_i = Jumlah data

Contoh :

Jika diketahui kelompok data berikut :

Tabel 3.3 Tabel pendapatan masyarakat kota "A"

Interval Kelas	Frekuensi
10 – 19	2
20 – 29	4
30 – 39	6
40 – 49	8
50 – 59	10
60 – 69	12
70 – 79	14
80 – 89	26
90 – 99	18
Jumlah	100

Tabel 3.4 Tabel pendapatan masyarakat kota "A" (Pengolahan)

Interval Kelas	Frekuensi i	X_i	$Fi.Xi$	Titik Interval (TI)	$Fi.TI$
10 – 19	2	14,5	29	3,22	6,44
20 – 29	4	24,5	98	5,44	21,78
30 – 39	6	34,5	207	7,67	46,00
40 – 49	8	44,5	356	9,89	79,11
50 – 59	10	54,5	545	12,11	121,11
60 – 69	12	64,5	774	14,33	172,00
70 – 79	14	74,5	1.043	16,56	231,78
80 – 89	26	84,5	2.197	18,78	488,22
90 – 99	18	94,5	1.701	21,00	378,00
Jumlah		100	6.950		1.544,44

$$MR = \frac{N_B - N_A}{2} = \frac{99 - 10}{2} = 44,5$$

$$PA = \frac{(MR)(\sum Fi.TI)}{P.\Sigma F1} = \frac{(44,5)(1.544,44)}{(10).(100)} = 68,73$$

SOAL LATIHAN:

Diperoleh nilai ujian siswa dalam satu kelas sebagai berikut :

Interval Nilai	f_i
40 – 49	1
50 – 59	4
60 – 69	8
70 – 79	14
80 – 81	10
90 – 99	3
JUMLAH	40

Tentukan Mean, Median, dan Modus dari data tersebut !

Pembahasan :

1.

Interval Nilai	f_i	x_i	$f_i \cdot x_i$
40 – 49	1	44.5	44.5
50 – 59	4	54.5	218
60 – 69	8	64.5	516
70 – 79	14	74.5	1043
80 – 81	10	84.5	845
90 – 99	3	94.5	283.5
Jumlah	40		2950

$$\text{Mean tabel distribusi frekuensi : } \bar{x} = \frac{2.950}{40} = 73,35$$

2. Berdasarkan tabel soal 1, diperoleh :

Interval Nilai	f_i	Σf_i
40 – 49	1	1
50 – 59	4	5
60 – 69	8	13
70 – 79	14	27
80 – 81	10	37
90 – 99	3	40

- Kelas Median di interval nilai 70-79 karena frekuensi kumulatif di interval tersebut sudah lebih dari $\frac{1}{2}$ frekuensi total.
- Tepi bawah, $L_2 = 69,5$
- Interval kelas, $C = 10$
- Banyak data, $n = 40$
- Frekuensi kumulatif sebelum kelas Median, $(\Sigma f)_2 = 13$
- Frekuensi kelas Median, $fM_e = 14$

Sehingga, nilai Median adalah :

$$M_e = L_2 + C \cdot \frac{\frac{1}{2}n - (\sum f)_2}{f_{M_e}} = 69,5 + 10 \cdot \frac{20 - 13}{14} = 74,5$$

3.

Interval Nilai	f_i
40 – 49	1
50 – 59	4
60 – 69	8
70 – 79	14
80 – 81	10
90 – 99	3

- Kelas Modus adapada interval 70-79
- Tepi bawah, $L = 695$
- Interval kelas, $C = 10$
- Selisih frekuensi kelas Modus dengan kelas sebelumnya,
 - $d_1 = 14 - 8 = 6$
- Selisih frekuensi kelas Modus dengan kelas sesudahnya,
 - $d_2 = 14 - 10 = 4$

Sehingga nilai Modus adalah:

$$M_o = L + C \cdot \frac{d_1}{d_1 + d_2} = 69,5 + 10 \cdot \frac{6}{6 + 4} = 75,5$$

BAB IV

UKURAN PENYEBARAN, DISPERSI DAN VARIANSI

Tujuan pembelajaran: mahasiswa memahami ukuran variansi, yang menggambarkan bagaimana berpencarnya data kuantitatif dan penggunaannya dalam studi.

Untuk itu selain ukuran gejala pusat masih ada lagi ukuran lain yang digunakan yaitu ukuran penyebaran atau ukuran dispersi. Ukuran ini dinamakan ukuran variansi, yang menggambarkan bagaimana berpencarnya data kuantitatif. Variabilitas merupakan kondisi di mana sekumpulan skor sama atau tidak. Jika sekumpulan skor itu sama maka distribusi tersebut tidak mempunyai variabilitas.

Pada bab sebelumnya kita sudah lakukan pembahasan tentang pemusatan data, namun ada beberapa hal yang mesti diperhatikan:

- Mean dan Median hanya menggambarkan pusat data dari sekelompok data, tetapi tidak menggambarkan penyebaran nilai pada data tersebut.
- Dua kelompok data dengan Mean yang sama, belum tentu memiliki penyebaran data yang sama.
- Ukuran dispersi yang kecil menunjukkan nilai data saling berdekatan (perbedaan kecil), sedangkan ukuran dispersi yang besar menunjukkan nilai data saling menyebar (perbedaan nilai masing-masing data besar).
- Ukuran dispersi digunakan untuk melengkapi perhitungan nilai pusat data.

Jenis ukuran dispersi dikelompokkan menjadi 2, yaitu dispersi mutlak dan dispersi relatif.

1. Dispersi Mutlak :

Dispersi mutlak digunakan untuk mengetahui tingkat variabilitas nilai-nilai observasi pada suatu data. servasi data lainnya. Beberapa ukuran dispersi yang terkenal antara lain rentang (*range*), rentang inter-kuartil (*inter-quartile range*), simpangan mutlak rata-rata (*Mean absolute*

deviation/MAD), simpangan kuartil (*quartile deviation*), simpangan baku (*standard deviation*).

2. Dispersi Relatif

Dispersi relatif digunakan untuk membandingkan tingkat variabilitas nilai-nilai observasi suatu data dengan tingkat variabilitas nilai-nilai observasi data lainnya.



RENTANG (*RANGE*)

Range adalah perbedaan antara skor terbesar dan skor terkecil. Sehingga rumus :

Range (Rentang) = data terbesar – data terkecil.

Contoh:

Terdapat data sebagai berikut: 5 7 4 6 9 10

Dari data diketahui skor tertinggi 10, skor terendah 4 sehingga nilai range adalah:

$$\begin{aligned} \text{Rentang (Range)} &= 10 - 4 \\ &= 6 \end{aligned}$$

RENTANG INTER-KUARTIL

Range inter-kuartil (*inter-quartile range*) adalah selisih kuartil ketiga dan kuartil pertama.

$$IQR = Q_3 - Q_1$$

Contoh:

Tabel 4.1 Data Skor Permainan

Skor	frekuensi	fk
30 – 39	5	5
40 – 49	7	12
50 – 59	10	22
60 – 69	9	31
70 – 79	5	36
80 – 89	3	39
Jumlah	39	

Untuk mencari *interquartile range* langkah-langkahnya sebagai berikut:

a) Menentukan elemen-elemen kuartil yaitu:

$$q = \frac{1}{4} \times n$$

Keterangan:

n = banyaknya responden

$$q_1 = \frac{1}{4} \times 39$$

$$= 9,75$$

$$q_2 = \frac{2}{4} \times 39$$

$$= 19,5$$

$$q_3 = \frac{3}{4} \times 39$$

$$= 29,5$$

b) Menentukan nilai kuartil yaitu:

$$Q_i = B_{bi} + C \frac{S_i}{f_{q1}}$$

Keterangan:

Q_i = kuartil ke i

B_{bi} = batas bawah kelompok yang mengandung q_i

C = interval kelompok

S_i = selisih antara q_i dengan frekuensi kumulatif sebelum kelompok yang mengandung q_i

f_{q1} = frekuensi kelompok yang mengandung q_i

q_i = batas kuartil ke i

Dari data tabel maka dapat diketahui:

$$B_{b1} = 40 \quad f_{q1} = 7$$

$$B_{b2} = 50 \quad f_{q2} = 10$$

$$B_{b3} = 60 \quad f_{q3} = 9$$

$$\begin{aligned} S_1 &= 9,75 - 5 \\ &= 4,75 \end{aligned}$$

$$\begin{aligned} S_2 &= 19,5 - 12 \\ &= 7,5 \end{aligned}$$

$$\begin{aligned} S_3 &= 29,5 - 22 \\ &= 7,5 \end{aligned}$$

Dari nilai-nilai di atas maka dapat dihitung nilai kuartil yaitu:

$$Q_1 = 40 + 10 \left(\frac{4,75}{7} \right)$$

$$= 46,78$$

$$Q1 = 50 + 10 \left(\frac{7,5}{10}\right)$$

$$= 57,5$$

$$Q1 = 60 + 10 \left(\frac{7,5}{9}\right)$$

$$= 68,33$$

Menentukan *interquartile range* yaitu:

$$\text{interquartile range} = 68,33 - 46,78$$

$$= 21,55$$

SIMPANGAN MUTLAK RATA-RATA (*MEAN ABSOLUTE DEVIATION; MAD*)

Deviasi mutlak rata-rata (*Mean absolute deviation; MAD*) adalah rata-rata nilai mutlak penyimpangan (deviasi) data terhadap reratanya.

Deviasi mutlak rata-rata untuk populasi adalah:

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

Contoh :

Data = 4, 5, 5, 6. Dari data tersebut didapatkan nilai rata-rata = 5

Maka $MAD = \frac{1}{4} (|4 - 5| + |5 - 5| + |5 - 5| + |6 - 5|) = \frac{1}{4} \cdot (2) = 1/2$

Sedangkan data berkelompok menggunakan rumus sebagai berikut:

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

Contoh :

Table 4.2. Data Usia Karyawan

No	Usia	Frekuensi	X _i	F _i ·X _i
1	21 – 23	3	22	66
2	24 – 26	5	25	125
3	27 – 29	2	28	56
Total F		10	Total FX	247
Rata-rata = 24,7				

Dari data 4.2 maka dapat diketahui

$$MAD = 1/10*(3*|22 - 24,7| + 5 * |25 - 24,7| + 2*|28 - 24,7|)$$

$$MAD = 0,1*(2*2,7 + 5*0,3 + 2*3,3) = 1,35$$

SIMPANGAN KUARTIL (*QUARTILE DEVIATION*)

Simpangan kuartil merupakan suatu ukuran dispersi yang didasarkan atas nilai kuartil yaitu kuartil pertama (K1) dan kuartil ketiga (K3). Ukuran ini juga disebut semi interquartile range yang berarti setengah jarak antara kuartil pertama hingga kuartil ketiga. Rumus yang digunakan adalah:

$$QD = \frac{Q_3 - Q_1}{2}$$

Contoh :

Data : 12 – 16 – 11 – 20 – 19 – 17 – 18 – 14 – 10 – 24

↳ 10 – 11 – 12• – 14 – 16• – 17 – 18 – 19• – 20 – 24

$$Q1 = 12$$

$$Q2 = (16+17)/2 = 16,5$$

$$Q3 = 19$$

$$QD = (19-12)/2 = 3,50$$

SIMPANGAN BAKU (*STD. DEVIATION*)

Standar deviasi atau simpangan baku merupakan salah satu cara mengukur variasi sekelompok data kuantitatif. Besar harga standar deviasi menunjukkan tingkat variansi dari sekumpulan data. Semakin besar harga standar deviasi, berarti semakin bervariasi atau heterogen angka-angka pada data kuantitatif tersebut. Standar deviasi data populasi dilambangkan dengan σ , sedangkan standar deviasi data sampel dilambangkan dengan s .

Variansi adalah rata-rata kuadrat deviasi data observasi terhadap rerata-nya. Variansi data populasi dilambangkan dengan σ^2 (dibaca: sigma kuadrat), sedangkan variansi data sampel dilambangkan dengan s^2 .

Standar Deviasi dari data tunggal $X_1, X_2, X_3, \dots, X_N$ yang berasal dari populasi. Rumus yang digunakan untuk data tidak berkelompok adalah,

$$s = \sqrt{\frac{1}{n} \sum (x_i - x)^2}$$

Contoh

Diketahui data sebagai berikut 4, 5, 5, 6. Dari data tersebut didapatkan nilai rata-rata = 5. Maka standar deviasi adalah,

$$s = \sqrt{1/4 \cdot (4-5)^2 + (5-5)^2 + (5-5)^2 + (6-5)^2}$$

$$s = \sqrt{0,25 (1+1)}$$

$$s = 0,707$$

Sedangkan untuk data berkelompok menggunakan rumus sebagai berikut

$$s = \sqrt{\frac{1}{n} \sum f_i (x_i - x)^2}$$

Contoh :

Table 4.3. Data Usia Karyawan

No	Usia	Frekuensi	X	$(X_i - \bar{X})^2$	$f(X_i - \bar{X})^2$
1	21 – 23	3	22	7,29	21,87
2	24 – 26	5	25	1,4	7
3	27 – 29	2	28	10,89	21,78
Total F		10			50,74
Rata-rata = 24,7					

Maka simpangan baku dari tabel 4.3 adalah sebagai berikut,

$$s = \sqrt{1/10 (50,74)}$$

$$s = \sqrt{5,074}$$

$$s = 2,25$$

$$\text{Varians } (s^2) = 2,25^2 = 5,074$$

BAB V

UKURAN KEMIRINGAN DAN KERUNCINGAN DATA

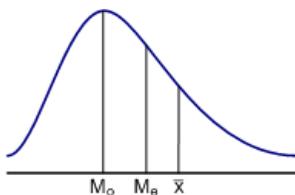
Tujuan pembelajaran : mahasiswa memahami model dan metode perhitungan kemiringan dan keruncingan pada data dan dapat mengimplemntasikan dalam studi

UKURAN KEMIRINGAN (*SKEWNESS*)

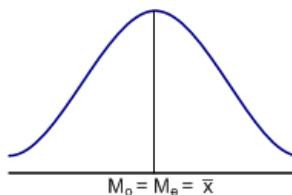
Kemiringan (*skewness*) adalah derajat ketidaksimetrisan atau penyimpangan dari kesimetrisan sebuah distribusi.

Distribusi yang tidak simetris akan memiliki rata-rata Mean, Median, dan Modus yang tidak sama besarnya.

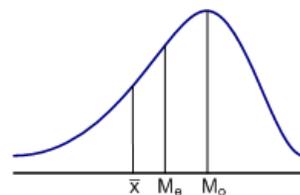
Ukuran kemiringan adalah ukuran yang menyatakan sebuah model distribusi yang mempunyai kemiringan tertentu. Apabila diketahui besarnya nilai ukuran ini maka dapat diketahui pula bagaimana model distribusinya, apakah distribusi itu simetrik, positif, atau negatif.



Distribusi Positif



Distribusi Simetrik



Distribusi Negatif

Untuk mengetahui apakah sekumpulan data mengikuti model distribusi positif, negatif, atau simetrik, hal ini dapat dilihat berdasarkan nilai koefisien kemiringannya.

KOEFISIEN KEMIRINGAN PEARSON

a. Koefisien Kemiringan (Modus).

$$\text{Koefisien Kemiringan : } S_k = \frac{\bar{x} - M_o}{s}$$

Dimana :

\bar{x} = rata-rata

M_o = Modus

S = simpangan baku

b. Koefisien Kemiringan (Median)

$$\text{Koefisien Kemiringan : } S_k = \frac{3(\bar{x} - Me)}{s}$$

Dimana :

\bar{x} = rata-rata

Me = Median

S = simpangan baku

c. Moment Matematis

$$\text{Koefisien Kemiringan : } S_k = \frac{\sum f_i (x_i - \bar{x})^2}{n \cdot s^3}$$

Dimana :

\bar{x} = rata-rata

Me = Median

S = simpangan baku

Menurut Pearson, dari hasil koefisien kemiringan diatas, ada tiga kriteria untuk mengetahui model distribusi dari sekumpulan data (baik data berkelompok maupun data tidak berkelompok), yaitu :

- Jika koefisien kemiringan < 0 , maka bentuk distribusinya negatif
- Jika koefisien kemiringan $= 0$, maka bentuk distribusinya simetrik
- Jika koefisien kemiringan > 0 , maka bentuk distribusinya positif

KOEFISIEN KEMIRINGAN BOWLEY

a. Koefisien Kemiringan Kuartil

$$\text{Koefisien Kemiringan} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

Dimana :

Q1 = Kuartil-1

Q2 = Kuartil-2

Q3 = Kuartil-3

KOEFISIEN KEMIRINGAN ANDI SUPANGAT

$$S_k = \frac{(P_{int} - M_o)}{T_k} = \frac{S_m}{T_k}$$

$$P_{int} = \frac{(N_A + N_B)}{2}$$

$$M_o = L + C \frac{b_1}{b_1 + b_2}$$

$$T_k = \frac{(P_{int} + X_m)}{2}$$

S_k : Kemiringan Kurva (Skewness)

P_{int} : Paruh Interval

M_o : Nilai Modus

T_k : Titik Tengah Kurva

S_m : Selisih Modus

X_m : Mid point yang mempunyai frekuensi tertinggi

Contoh :

Berat badan bayi (dicatat dalam Kg) yang baru lahir dirumah sakit bersalin “Bunda” dapat dilihat dalam tabel 5.1 berikut.

Table 5.1. Tabel berat badan bayi (Kg)

Berat Badan (Kg)	Banyak Bayi (f)
2,5 – 2,6	2
2,7 – 2,8	3
2,9 – 3,0	5
3,1 – 3,2	7
3,3 – 3,4	6
3,5 – 3,6	5
Jumlah	28

Hitung koefisien kemiringannya dengan menggunakan nilai kuartil !

Table 5.2. Tabel berat badan bayi (Pengolahan Skewness)

Berat Badan (Kg)	Banyak Bayi (f)	F Kum (fk)	Keterangan
2,5 – 2,6	2	2	
2,7 – 2,8	3	5	
2,9 – 3,0	5	10	Q ₁
3,1 – 3,2	7	17	Q ₂
3,3 – 3,4	6	23	Q ₃
3,5 – 3,6	5	28	
Jumlah	28		

$$Q_i = L_i + C. \frac{\frac{in}{4} - (\sum f)_i}{f Q_i}$$

- Q₁ → $\frac{1}{4} \cdot (n) = \frac{1}{4} \cdot (28) = 7$ → Kelas interval ke-3

$$Q1 = 2,85 + 0,2 ((7-5)/5) = 2,85 + 0,08 = 2,93$$

- $Q2 \rightarrow 2/4.(n) = 2/4.(28) = 14 \rightarrow$ Kelas interval ke-4

$$Q2 = 3,05 + 0,2 ((14-10)/7) = 3,05 + 0,11 = 3,16$$

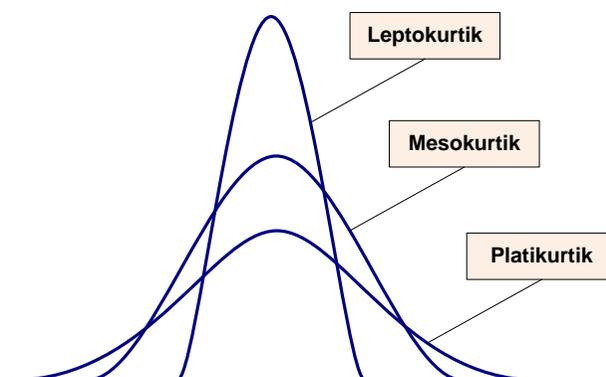
- $Q3 \rightarrow 3/4.(n) = 3/4.(28) = 21 \rightarrow$ Kelas interval ke-5

$$Q3 = 3,25 + 0,2 ((21-17)/6) = 3,25 + 0,13 = 3,38$$

- Sehingga Koefisien Kemiringan $= \frac{3,38-2(3,16)+2,93}{3,38-2,93}$
 $= 0,01/-0,45$
 $= -0,022$ (distribusi negatif)

UKURAN KERUNCINGAN (KURTOSIS)

Ukuran keruncingan adalah tingkat kepuncakan dari suatu distribusi yang biasanya diambil relatif terhadap distribusi normal. Sebuah distribusi yang mempunyai puncak relatif tinggi dinamakan leptokurtik, sebuah distribusi mempunyai puncak mendatar dinamakan platikurtik, distribusi normal yang puncaknya tidak terlalu tinggi atau tidak mendatar dinamakan mesokurtik.



Untuk mengetahui apakah sekumpulan data mengikuti distribusi leptokurtik, platikurtik, dan mesokurtik, hal ini dapat dilihat berdasarkan koefisien kurtosisnya.

Untuk menghitung koefisien kurtosis digunakan rumus koefisien kurtosis, yaitu:

$$K = \frac{\frac{1}{2}(Q3 - Q1)}{P_{90} - P_{10}}$$

Dimana:

Q1 = Kuartil pertama ; Q3 = Kuartil ketiga

P10 = Persentil ke 10 ; P90 = Persentil ke 90

Dari hasil koefisien kurtosis diatas, ada tiga criteria untuk mengetahui model distribusi dari sekumpulan data, yaitu :

- Jika koefisien kurtosisnya $< 0,263$ maka distribusinya adalah platikurtik.
- Jika koefisien kurtosisnya $= 0,263$ maka distribusinya adalah mesokurtik.
- Jika koefisien kurtosisnya $> 0,263$ maka distribusinya adalah leptokurtik.

Contoh:

Table 5.3. Tabel berat badan bayi (Pengolahan Kurtosis)

Berat Badan (Kg)	Banyak Bayi (f)	F Kum (fk)	Keterangan
2,5 – 2,6	2	2	
2,7 – 2,8	3	5	P ₁₀
2,9 – 3,0	5	10	Q ₁
3,1 – 3,2	7	17	Q ₂
3,3 – 3,4	6	23	Q ₃
3,5 – 3,6	5	28	P ₉₀
Jumlah	28		

$$P_i = L_i + C. \frac{\frac{in}{100} - (\Sigma f)_i}{f P_i}$$

- P10 → $10/100.(n) = 1/10.(28) = 2,8$ → Kelas interval ke-2

$$Q1 = 2,65 + 0,2 ((2,8-2)/3) = 2,65 + 0,05 = 2,7$$

- $P90 \rightarrow 90/100 \cdot (n) = 9/10 \cdot (28) = 25,2 \rightarrow$ Kelas interval ke-6

$$Q2 = 3,45 + 0,2 ((25,2-23)/5) = 3,45 + 0,088 = 3,54$$

- Sehingga Koefisien Kurtosisnya $= \frac{\frac{1}{2} \cdot (3,38 - 2,93)}{3,54 - 2,70}$
 $= 0,225 / 0,84$
 $= 0,268$ (Leptokurtik)

BAB VI

REGRESI LINIER SEDERHANA

Tujuan pembelajaran: mahasiswa paham cara mengitung regresi baik sederhana serta mahasiswa dapat mengimplementasikan dalam studi

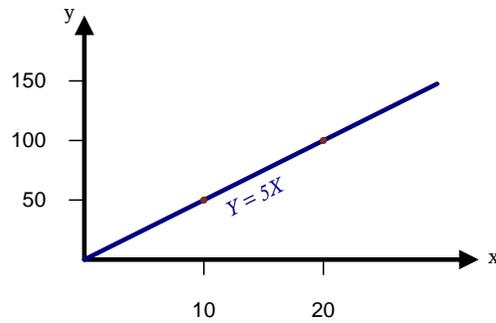
Regresi linear (linear regression) adalah teknik yang digunakan untuk memperoleh model hubungan antara 1 variabel dependen dengan 1 atau lebih variabel independen. Jika hanya digunakan 1 variabel independen dalam model, maka teknik ini disebut sebagai regresi linear sederhana (simple linear regression), sedangkan jika yang digunakan adalah beberapa variabel independen, teknik ini disebut regresi linear ganda (multiple linear regression). Variabel dependen pada regresi linear disebut juga sebagai *respons* atau *kriteria*, sedangkan variabel independen dikenal pula sebagai *prediktor* atau *regresor*.

Analisis regresi adalah analisis statistik yang mempelajari hubungan antara dua atau lebih variabel kuantitatif sehingga satu variabel dapat diramalkan (predicted) dari variabel lainnya. Hubungan antara dua variabel dapat dibedakan menjadi dua, yaitu hubungan fungsional dan hubungan statistik. Hubungan fungsional antara dua variabel dapat dinyatakan secara matematis; jika X variabel bebas (*independent variable*) dan Y variabel tak bebas (*dependent variable*), hubungan fungsional ditulis dalam bentuk :

$$Y = f(X)$$

Contoh 1:

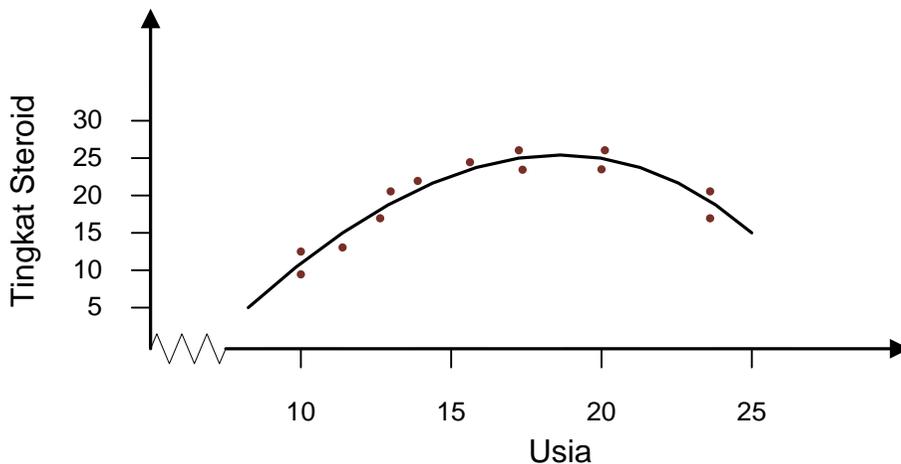
Jika diketahui nilai X tertentu, fungsi f akan memberikan nilai Y yang bersesuaian. Misalkan Y = total nilai penjualan suatu produk (dalam ribuan rupiah) dan X = jumlah unit produk yang terjual. Jika harga jual Rp 5.000,00 per unit produk maka hubungan antara Y dan X dapat dinyatakan dalam Gambar 6.1



Gambar 6.1 Hubungan jumlah unit produk terjual dengan total nilai penjualan

Contoh 2:

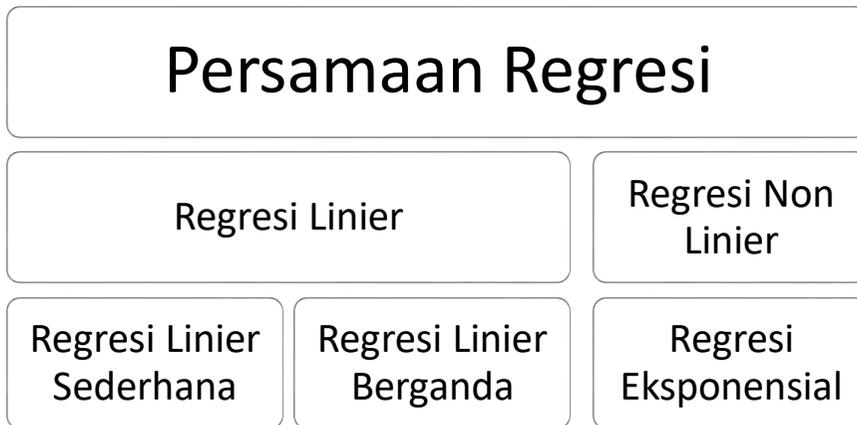
Jika hubungan antara X yang menyatakan usia (dalam tahun) dan Y yang menyatakan tingkat steroid dinyatakan dalam Gambar 6.2.



Gambar 6.2 Hubungan usia dengan tingkat steroid

Model regresi secara formal menyatakan dua hal tentang hubungan statistik, yaitu: kecenderungan variabel tak bebas (Y) berubah-ubah terhadap variabel bebas (X) dalam bentuk yang sistematis dan tersebaranya titik-titik di sekitar kurve hubungan statistik. Kedua hal tersebut dinyatakan dalam suatu model regresi yaitu: untuk setiap nilai X terdapat distribusi probabilitas dari Y dan mean dari distribusi probabilitas Y berubah-ubah secara sistematis terhadap perubahan nilai X .

Dari uraian diatas dapat dijelaskan bahwa persamaan regresi merupakan persamaan matematik yang memungkinkan peramalan nilai suatu peubah takbebas (dependent variable) dari nilai peubah bebas (independent variable).



- Bentuk Umum Regresi Linier Sederhana

$$Y = a + bX$$

Y : peubah tak bebas

X : peubah bebas

a : konstanta

b : kemiringan

- Bentuk Umum Regresi Linier Berganda

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Y : peubah tak bebas a : konstanta

X₁ : peubah bebas ke-1 b₁ : kemiringan ke-1

X₂ : peubah bebas ke-2 b₂ : kemiringan ke-2

X_n : peubah bebas ke-n b_n : kemiringan ke-n

- Bentuk umum Regresi Eksponensial

$$Y = ab^x$$

$$\log Y = \log a + (\log b) x$$

REGRESI LINIER SEDERHANA

Regresi linear sederhana adalah hubungan secara linear antara satu variabel independen (X) dengan variabel dependen (Y). Analisis ini digunakan untuk mengetahui arah hubungan antara variabel independen dengan variabel dependen apakah positif atau negatif serta untuk memprediksi nilai dari variabel dependen apabila nilai variabel independen mengalami kenaikan atau penurunan nilai. Data yang digunakan biasanya berskala interval atau rasio.

Rumus dari analisis regresi linear sederhana adalah sebagai berikut:

$$Y' = a + bX$$

Keterangan:

Y= subyek dalam variabel dependen yang diprediksi

a = harga Y ketika harga X= 0 (harga konstan)

b = angka arah atau koefisien regresi, yang menunjukkan angka peningkatan ataupun penurunan variabel dependen yang didasarkan pada perubahan variabel independen. Bila (+) arah garis naik, dan bila (-) maka arah garis turun.

X = subyek pada variabel independen yang mempunyai nilai tertentu.

Nilai-nilai a dan b dapat dihitung dengan menggunakan Rumus dibawah ini :

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

Berikut ini adalah Langkah-langkah dalam melakukan Analisis Regresi Linear Sederhana :

1. Tentukan Tujuan dari melakukan Analisis Regresi Linear Sederhana
2. Identifikasikan Variabel Faktor Penyebab (Predictor) dan Variabel Akibat (Response)
3. Lakukan Pengumpulan Data
4. Hitung X^2 , Y^2 , XY dan total dari masing-masingnya
5. Hitung a dan b berdasarkan rumus diatas.
6. Buat Model Persamaan Regresi Linear Sederhana.
7. Lakukan Prediksi atau Peramalan terhadap Variabel Faktor Penyebab atau Variabel Akibat.
8. Berikut ini adalah tabel yang telah dilakukan perhitungan X^2 , Y^2 , XY dan totalnya :

Sebagai contoh, penyelesaian persamaan regresi sebagai berikut;

Tanggal	Rata-rata Suhu Ruangan (X)	Jumlah Cacat (Y)	X^2	Y^2
1	24	10	576	100
2	22	5	484	25
3	21	6	441	36
4	20	3	400	9
5	22	6	484	36
6	19	4	361	16
7	20	5	400	25
8	23	9	529	81
9	24	11	576	121
10	25	13	625	169
11	21	7	441	49

12	20	4	400	16
13	20	6	400	36
14	19	3	361	9
15	25	12	625	144
16	27	13	729	169
17	28	16	784	256
18	25	12	625	144
19	26	14	676	196
20	24	12	576	144
21	27	16	729	256
22	23	9	529	81
23	24	13	576	169
24	23	11	529	121
25	22	7	484	49
26	21	5	441	25
27	26	12	676	144
28	25	11	625	121
29	26	13	676	169
30	27	14	729	196
Total (Σ)	699	282	16487	3112

Menghitung Konstanta (a) :

$$a = \frac{(\Sigma y) (\Sigma x^2) - (\Sigma x) (\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$a = \frac{(282) (16.487) - (699) (6.861)}{30 (16.487) - (699)^2}$$

$$a = -24,38$$

Menghitung Koefisien Regresi (b)

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{30(6.861) - (699)(282)}{30(16.487) - (699)^2}$$

$$b = 1,45$$

Sehingga model regresi yang didapat adalah

$$Y = a + bX$$

$$Y = -24,38 + 1,45X$$

SOAL LATIHAN

1. Berdasarkan data penelitian didapatkan rekapitulasi data koestioner adalah sebagai berikut untuk masing-masing variable. Buatlah model regresi sederhana dari hasil penelitian tersebut

Responden	Motivasi	Kinerja
1	23	24
2	15	18
3	26	23
4	22	19
5	32	27
6	32	26
7	23	20
8	11	27
9	23	30
10	24	37
11	32	31
12	21	18
12	22	21
13	25	19
14	34	19
15	22	30
16	18	32
17	19	32
18	22	23

19	27	22
20	30	32
21	37	32
22	31	23
23	18	11
24	21	23
25	19	24
26	19	32
27	30	21
28	24	22
29	28	25
30	29	32

BAB VII

REGRESI LINIER BERGANDA

Tujuan pembelajaran: mahasiswa paham cara mengitung regresi berganda sehingga mahasiswa dapat mengimplementasikan dalam studi

Regresi linier berganda merupakan model persamaan yang menjelaskan hubungan satu variabel tak bebas/ response (Y) dengan dua atau lebih variabel bebas/ predictor (X1, X2,...Xn). Tujuan dari uji regresi linier berganda adalah untuk memprediksi nilai variable tak bebas/ response (Y) apabila nilai-nilai variabel bebasnya/ predictor (X1, X2,..., Xn) diketahui. Disamping itu juga untuk dapat mengetahui bagaimanakah arah hubungan variabel tak bebas dengan variabel-variabel bebasnya.

Persamaan regresi linier berganda secara matematik diekspresikan oleh :

$$Y = a + b_1 X_1 + b_2 X_2 + \dots + b_n X_n$$

yang mana :

Y = variable tak bebas (nilai variabel yang akan diprediksi)

a = konstanta

b₁,b₂,..., b_n = nilai koefisien regresi

X₁,X₂,..., X_n = variable bebas

Bila terdapat 2 variable bebas, yaitu X₁ dan X₂, maka bentuk persamaan regresinya

adalah :

$$Y = a + b_1X_1 + b_2X_2$$

Keadaan-keadaan bila koefisien-koefisien regresi, yaitu b₁ dan b₂ mempunyai nilai :

- Nilai=0. Dalam hal ini variabel Y tidak dipengaruhi oleh X₁ dan X₂
- Nilainya negative. Disini terjadi hubungan dengan arah terbalik antara variabel tak bebas Y dengan variabel-variabel X₁ dan X₂

- Nilainya positif. Disini terjadi hubungan yang searah antara variabel tak bebas Y dengan variabel bebas X1 dan X2

Koefisien-koefisien regresi b1 dan b2 serta konstanta a dapat dihitung dengan menggunakan rumus :

$$a = \frac{(\sum Y) - (b_1 \times \sum x_i) - (b_2 \times \sum x_2)}{n}$$

$$b_1 = \frac{[(\sum x_2^2 \times \sum x_1 y) - (\sum x_2 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 \times x_2)^2]}$$

$$b_2 = \frac{[(\sum x_1^2 \times \sum x_2 y) - (\sum x_1 y \times \sum x_1 x_2)]}{[(\sum x_1^2 \times \sum x_2^2) - (\sum x_1 \times x_2)^2]}$$

yang mana :

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n}$$

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n}$$

$$\sum y^2 = \sum Y^2 - \frac{(\sum Y)^2}{n}$$

$$\sum x_1 y = \sum X_1 Y - \frac{\sum X_1 \sum Y}{n}$$

$$\sum x_2 y = \sum X_2 Y - \frac{\sum X_2 \sum Y}{n}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{\sum X_1 \sum X_2}{n}$$

Metode alternatif, yaitu metode matriks (metode kuadrat terkecil) dapat digunakan untuk menentukan nilai a , b_1 dan b_2 . Metode ini dilakukan dengan cara membuat dan menyusun suatu persamaan sebagai berikut :

$$\frac{a}{cA = \pi r^2 + b}$$

Latihan soal

Berdasarkan hasil penelitian didapatkan data sebagai berikut, buatlah model regresi berganda dari data tersebut

Responden	Motivasi	Supervisi	Kinerja
1	23	23	24
2	15	11	18
3	26	23	23
4	22	24	19
5	32	32	27
6	32	21	26
7	23	22	20
8	11	25	27
9	23	34	30
10	24	22	37
11	32	18	31
12	21	27	18
12	22	30	21
13	25	37	19
14	34	31	19
15	22	18	30
16	18	21	32
17	19	19	32
18	22	19	23
19	27	30	22
20	30	26	32

21	37	29	32
22	31	25	23
23	18	27	11
24	21	27	23
25	19	19	24
26	19	30	32
27	30	32	21
28	24	31	22
29	28	28	25
30	29	24	32

BAB VIII

KORELASI DAN KOEFISIEN DETERMINASI

Tujuan pembelajaran: mahasiswa paham cara mengitung korelasi dan koefisien determinasi sehingga mahasiswa dapat mengimplementasikan dalam studi

Korelasi adalah salah satu teknik analisis dalam statistik yang digunakan untuk mencari hubungan antara dua variabel yang juga bersifat kuantitatif. Sebagai contoh, kita dapat menggunakan tinggi badan dan usia siswa SD sebagai variabel dalam korelasi positif. Istilah tersebut juga dapat dikatakan istilah sebab akibat, dan istilah tersebut menjadi ciri khas tersendiri dari statistik korelasi.

Koefisien Korelasi Sederhana pada umumnya disebut juga dengan Koefisien Korelasi Pearson karena memiliki rumus perhitungan Koefisien korelasi sederhana dikemukakan oleh Karl Pearson yaitu seseorang ahli Matematika yang berasal dari Inggris. (Rumus ini disebut juga dengan Pearson product moment)

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\{n\sum x^2 - (\sum x)^2\} \{n\sum y^2 - (\sum y)^2\}}}$$

Keterangan Rumus :

n adalah Banyaknya Pasangan data X dan Y

$\sum x$ adalah Total Jumlah dari Variabel X

$\sum y$ adalah Total Jumlah dari Variabel Y

$\sum x^2$ adalah Kuadrat dari Total Jumlah Variabel X

$\sum y^2$ adalah Kuadrat dari Total Jumlah Variabel Y

$\sum xy$ adalah Hasil Perkalian dari Total Jumlah Variabel X dan Variabel Y

Bentuk Hubungan Antara 2 Variabel

Korelasi Linear Positif (+1)

Perubahan Nilai Variabel diikuti perubahan Nilai Variabel yang lainnya secara teratur dengan arah yang sama. Jika Nilai Variabel X mengalami kenaikan, maka Variabel Y juga ikut naik. Jika Nilai Variabel X mengalami penurunan, maka Variabel Y pun ikut turun. Jika Nilai Koefisien Korelasi mendekati +1 (positif Satu) berarti pasangan data Variabel X dan Y mempunyai Korelasi Linear Positif yang kuat.

Korelasi Linear Negatif (-1)

Perubahan Nilai Variabel diikuti perubahan Nilai Variabel yang lainnya secara teratur tetapi dengan arah yang berlawanan. Jika Nilai Variabel X mengalami kenaikan, maka Variabel Y akan turun. Jika Nilai Variabel X turun, maka Nilai Variabel Y mengalami kenaikan. Apabila Nilai Koefisien Korelasi mendekati -1 maka hal ini menunjukkan pasangan data Variabel X dan Variabel Y mempunyai Korelasi Linear Negatif yang kuat/erat.

Tidak berkorelasi (0)

Kenaikan Nilai Variabel yang satunya terkadang diikuti dengan penurunan Variabel yang lainnya atau terkadang diikuti dengan kenaikan Variable yang lainnya. Arah hubungannya tidak teratur, searah, dan terkadang berlawanan. Apabila Nilai Koefisien Korelasi mendekati 0 (Nol) berarti pasangan data Variabel X dan Y mempunyai korelasi yang sangat lemah atau berkemungkinan tidak berkorelasi

Pedoman umum untuk menentukan korelasi antar hubungan adalah sebagai berikut

Nilai Korelasi Sampel (r)	Interpretasinya
0,00 - 0,09	Hubungan korelasinya diabaikan
0,10 - 0,29	Hubungan korelasi rendah
0,30 - 0,49	Hubungan korelasi moderat
0,50 - 0,70	Hubungan korelasi sedang
> 0,70	Hubungan korelasi sangat kuat

KOEFISIEN DETERMINASI

Koefisien determinasi pada regresi linier sering diartikan sebagai seberapa besar kemampuan semua variabel bebas dalam menjelaskan varians dari variabel terikatnya. Secara sederhana koefisien determinasi dihitung dengan mengkuadratkan Koefisien Korelasi (R). Sebagai contoh, jika nilai R adalah sebesar 0,80 maka koefisien determinasi (R Square) adalah sebesar $0,80 \times 0,80 = 0,64$. Berarti kemampuan variabel bebas dalam menjelaskan varians dari variabel terikatnya adalah sebesar 64,0%. Berarti terdapat 36% (100%-64%) varians variabel terikat yang dijelaskan oleh faktor lain. Berdasarkan interpretasi tersebut, maka tampak bahwa nilai R Square adalah antara 0 sampai dengan 1.

Penggunaan R Square (R Kuadrat) sering menimbulkan permasalahan, yaitu bahwa nilainya akan selalu meningkat dengan adanya penambahan variabel bebas dalam suatu model. Hal ini akan menimbulkan bias, karena jika ingin memperoleh model dengan R tinggi, seorang penelitian dapat dengan sembarangan menambahkan variabel bebas dan nilai R akan meningkat, tidak tergantung apakah variabel bebas tambahan itu berhubungan dengan variabel terikat atau tidak.

Oleh karena itu, banyak peneliti yang menyarankan untuk menggunakan Adjusted R Square. Interpretasinya sama dengan R Square, akan tetapi nilai Adjusted R Square dapat naik atau turun dengan adanya penambahan variabel baru, tergantung dari korelasi antara variabel bebas tambahan tersebut dengan variabel terikatnya. Nilai Adjusted R Square dapat bernilai negatif, sehingga jika nilainya negatif, maka nilai tersebut dianggap 0, atau variabel bebas sama sekali tidak mampu menjelaskan varians dari variabel terikatnya.

DAFTAR PUSTAKA

- Budiyono. 2009. Statistika Untuk Penelitian. Surakarta: Rajawali Press.
- Hadi, Sutrisno. 2015. Statistika. Yogyakarta: Pustaka Pelajar
- Irianto, Agus. 2012. Statistik Konsep Dasar, Aplikasi, dan Pengembangannya.
Jakarta : Kencana
- Sudjana. 2002. Metoda Statistika. Bandung : Tarsito
- Sugiyono. 2016. Statistika Untuk Penelitian. Bandung: Alfabeta

BIODATA PENULIS

Penulis lahir di Purworejo, 9 April 1981. Menyelesaikan studi program S1 Manajemen dan S2 Magister Manajemen di Universitas Mercubuana, dan menyelesaikan studi program S3 Manajemen Ekowisata dan Jasa Lingkungan di Institut Pertanian Bogor.

Penulis aktif sebagai pengajar di beberapa kampus di Jakarta diantaranya Universitas Bhayangkara Jakarta Raya, Universitas Bina Nusantara, Trisakti School of Management dan Universitas M.H Thamrin. Selain itu, penulis juga aktif dalam penelitian dan publikasi baik dalam skala Nasional maupun Internasional.