

# STATISTIK DESKRIPTIF

Dosen:        Iding Tarsidi, M. Pd.(1723)

Mata kuliah ini untuk mahasiswa Jurusan PSIKOLOGI FIP UPI.

## **Tujuan:**

Setelah mengikuti perkuliahan ini mahasiswa diharapkan memahami konsep dasar statistika deskriptif dan mampu mengaplikasikannya untuk kepentingan pengumpulan, pengolahan, dan penyajian data (hasil penelitian–pendidikan) sehingga mudah difahami oleh pembaca.

## **Materi perkuliahan statistika deskriptif mencakup:**

Konsep dasar statistika, statistik, Fungsi statistika, data statistik, sumber dan jenis data, skala pengukuran data, ukuran kecenderungan pusat/tendensi sentral (mean, median, modus), ukuran letak (kuartil, desil, persentil), penyajian data: Daftar Distribusi Frekuensi; Grafik: Poligon, Ogive; Diagram: Batang, Histogram), Populasi dan sampel, Ukuran dispersi (rentang, rerata simpangan, simpangan baku dan varians); Korelasi: Product Moment, Interpretasi koefisien korelasi.

## METODE KULIAH DAN SUMBER/BUKU

- Pendekatan:

Metode: Ceramah, tanya jawab, diskusi, latihan/tugas

Tugas: Latihan/pendalaman setiap selesai satu topik bahasan.

- Komponen Evaluasi:

UAS : 45%

UTS : 35%

Tugas : 20%

Kehadiran : Prasyarat mengikuti UAS

- **SUMBER/DAFTAR BUKU**

- Minium, E. W., King, B.M., & Bear, G. (1993). *Statistical Reasoning in Psychological and Education*. New York: John Wiley & Sons.

- Nurgiyantoro, B., Gunawan dan Marzuki. (2000). *Statistik Terapan untuk Penelitian Ilmu-Ilmu Sosial*. Yogyakarta: Gadjah Mada University Press.

- Siegel, Sidney. (1997). *Statistik Non-Parametrik untuk Ilmu-Ilmu Sosial* (Terjemahan). Jakarta: Gramedia.

- Sudjana. (1992). *Metoda Statistika*, edisi kelima. Bandung: Tarsito.

- Sudjana, N. (1991). *Tuntunan Penyusunan Karya Ilmiah, Makalah-Skripsi-Tesis-Disertasi*, Edisi kedua. Bandung: Sinar Baru.

## TOPIK-TOPIK BAHASAN STATISTIK DESKRIPTIF

1. Konsep Dasar: Pengertian, Tujuan, Fungsi dan Kegunaan Statistik Deskriptif, Hubungan Deskriptif dan Inferensial
2. Data Statistik, Mengurut dan Mengelompokkan Data, Daftar Distribusi Frekuensi (Tunggal dan Kelompok)
3. Penyajian Data (Diagram Batang, Histogram)
4. Penyajian Data (Diagram Garis/Poligon, Ogive)
5. Skala Pengukuran Data (nominal, ordinal, interval, dan rasio)
6. Ukuran Kecenderungan Pusat (Mean, Median, Modus)
7. Ukuran Letak (Kuartil, desil, persentil)
8. Ukuran Dispersi (sebaran): Rentang, Rentang Antar Kuartil, Semi Kuartil, Rerata Simpangan, Simpangan Baku dan Varians
9. Skor Baku (Z) dan Kurva Normal

## KOSEP DASAR STATISTIK, STATISTIKA, DAN CARA MEMPELAJARINYA

- **Statistik:** untuk menyatakan kumpulan data, bilangan maupun non bilangan yang disusun dalam tabel dan atau diagram, yang melukiskan atau menggambarkan suatu persoalan (misal: statistik: penduduk, kelahiran).
- Dalam konteks "sample – populasi", Statistik adalah untuk menyatakan ukuran sebagai wakil dari kumpulan data mengenai sesuatu hal (sample); sedangkan untuk menyatakan ukuran sebagai wakil dari kumpulan data (populasi) disebut parameter.
- **Statistika:** pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan atau penganalisaannya dan penarikan kesimpulan berdasarkan kumpulan data dan penganalisaan yang dilakukan.
- Ditinjau dari jalan/cara mempelajarinya statistika dapat dibedakan: (1). Statistika Matematis/Teoretis, disini dibahas secara mendasar, mendalam dan teoretis tentang: penurunan sifat, dalil, rumus, (2). Statistika Terapan, untuk penggunaan/aplikasi dalam berbagai bidang pengetahuan, yakni tentang bagaimana "metoda" statistika digunakan.  
(Sudjana, 1992: 2-4).

## STATISTIKA & JENIS-JENISNYA

- **Statistika** adalah bagian dari matematika yang secara khusus membicarakan cara-cara pengumpulan, analisis dan penafsiran data. Juga untuk menunjukkan "body of knowledge" tentang cara-cara "sampling" (pengumpulan data), analisis dan penafsiran data.

### **Jenis-Jenis Statistika dapat dibedakan/ditinjau dari:**

- **Orientasi Pembahasannya:** (1). Mathematical Statistics atau Statistika Teoretis, berorientasi kepada pemahaman model dan teknik statistika secara matematis-teoretis; (2). Applied Statistics, berorientasi kepada pemahaman intuitif atas konsep dan teknik statistika serta penggunaannya dalam berbagai bidang.
- **Tahapan atau tujuan analisisnya:** (1). Statistika Deskriptif, untuk memperoleh deskripsi tentang ukuran-ukuran data di tangan (baik sampel-statistik maupun populasi-parameter); (2). Statistika Inferensial/Induktif, yakni dari harga statistik digunakan untuk "menaksir" atau menguji hipotesis yang berlaku untuk populasi.
- **Asumsi distribusi populasi data** yang dianalisisnya: (1). Statistika Parametrik – model distribusi normal, (2). Statistika Nonparametrik – distribution free statistics.
- **Jumlah dependent variable** yang dianalisisnya: (1). Statistika Univariat, dan (2). Statistika Multivariat (dua variabel terikat atau lebih), berapapun variabel bebasnya.
- **Bidang/kajian** dimana statistika itu digunakan, misalnya "statistika" : pertanian, industri, pendidikan, ekonomi, kependudukan, "biostatistics". (Furqon, 3:2001).

# FUNGSI & KEGUNAAN STATISTIKA

**Menurut Budi Yuwono** (1987, dalam Subana, dkk., 13: 2000), fungsi statistika:

- Menggambarkan data dalam bentuk tertentu, sehingga jelas.
- Menyederhanakan data yang kompleks menjadi data yang mudah dimengerti (tabel, grafik, diagram, rata-rata, persentase, atau dalam koefisien-koefisien).
- Sebagai teknik untuk membuat perbandingan.
- Dapat memperluas pengalaman individual (dengan mempelajari kesimpulan-kesimpulan berdasarkan data yang dianalisis lainnya).
- Dapat mengukur besaran dari suatu gejala (sosial, ekonomi), dan dapat menentukan hubungan sebab akibat (untuk prediksi).

**Menurut Irianto, Agus** (1988, dalam Subana, dkk., 14:2000), kegunaan statistika:

- Membantu peneliti dalam menggunakan sampel sehingga dapat bekerja efisien dengan hasil yang sesuai dengan objek yang diteliti.
- Membantu peneliti membaca data yang terkumpul sehingga dapat mengambil kesimpulan yang tepat.
- Membantu peneliti melihat ada tidaknya perbedaan antara kelompok lainnya atas objek yang diteliti.
- Membantu peneliti melihat ada tidaknya hubungan antar variabel yang diteliti.
- Membantu peneliti memprediksi waktu yang akan datang.
- Membantu peneliti melakukan interpretasi data yang terkumpul.

Statistika Pendidikan: prinsip, metode, dan prosedur yang digunakan sebagai cara mengumpulkan, analisis, dan interpretasi data berkaitan dengan dunia pendidikan.

## DATA STATISTIK, POPULASI & SAMPEL, STATISTIK DESKRIPTIF & INFERENSIAL

- **Data/data statistik:** keterangan atau ilustrasi mengenai suatu hal, dapat berbentuk kategori (rusak, baik, gagal, puas) atau berbentuk bilangan (kuantitatif), harganya berubah-ubah atau bersifat "variabel".
- **Data kualitatif:** data yang dikategorikan menurut lukisan kualitas obyek yang dipelajari, disebut "atribut" (sakit, rusak, berhasil, dsj.).
- Dari nilainya ada dua data kuantitatif: (1). Diskrit, hasil menghitung atau membilang (3 orang, 4 buah gedung); (2). Kontinue, hasil pengukuran (tinggi, berat).
- Menurut sumbernya: (1). Data intern, bersumber dari "orang dalam", (2). Data ekstern (primer, sekunder), data dari sumber/pihak lain.
- **Populasi:** Totalitas semua nilai yang mungkin, hasil menghitung atau pengukuran, kuantitatif maupun kualitatif mengenai karakteristik tertentu dari semua anggota kumpulan yang lengkap dan jelas yang ingin dipelajari sifat-sifatnya).
- **Sampel representatif,** jika mencerminkan segala karakteristik populasi. (Sudjana, 1992: 4-6).
- **Statistika Deskriptif:** fase statistika yang hanya berusaha melukiskan dan menganalisis kelompok yang diberikan tanpa menarik kesimpulan tentang populasi atau kelompok yang lebih besar.
- **Statistika Induktif/Inferensial:** fase statistika yang berhubungan dengan kondisi-kondisi dimana kesimpulan diambil. Ini, biasanya merupakan kelanjutan statistika deskriptif. (Sudjana, 1992:7).

## SKALA PENGUKURAN DATA: NOMINAL, ORDINAL, INTERVAL, RASIO

- Jika salah satu variabel mempunyai peringkat yang berbeda, maka analisis data mengambil rumus data yang peringkatnya lebih rendah.
- Uji signifikansi untuk data nominal biasanya melalui Chi atau Kai-Kuadrat. Ini digunakan untuk mengetahui ada tidaknya perbedaan signifikan antara frekuensi harapan ( $f_e/f_h$ ) dengan frekuensi dalam kenyataan ( $f_o$ ).
- Data nominal yang "asimetrik" menggunakan Lambda (Prakiraan Guttman).
- Teknik analisis data ordinal berdasarkan teori pasangan. Skala ordinal menunjuk pada posisi relatif individu/objek. Memiliki kategori yang diurutkan/ranking posisinya berdasarkan kriteria tertentu. Mempunyai makna lebih besar dari. Jarak antara urutan 1 dengan 2 tidak bermakna sama dengan jarak 2 dan 3. Rangkaian tidak mempunyai interval yg tetap/sama.
- Hubungan yang membatasinya adalah ekuivalensi dan lebih besar dari, statistik yang cocok digunakan: persentil, median, Spearman ( $\rho$ ), dan Kendal.
- Jika kedua variabel "simetrik" gunakan Gamma, jika "asimetrik" maka gunakan "Somers'  $d_{yx}$ " (ini tidak sampai uji signifikansi). Jika hubungan simetrik berdasarkan ranking, gunakan "Spearman's  $\rho$ ". Uji signifikansinya bisa dengan Kai-Kuadrat.
- Skala interval (mempunyai rentangan konstan antara tkt satu dg lainnya, tidak mempunyai 0 mutlak), dan rasio (mempunyai 0 mutlak), hubungannya ekuivalensi, lebih besar dari, rasio sembarang dua interval diketahui. Statistika yang digunakan: Mean, SB, Variansi, Korelasi Pearson ( $r$ ), Uji-t, Uji-F, ANAVA, Regresi, dll.



## SKALA PENGUKURAN DATA: NOMINAL, ORDINAL, INTERVAL DAN RASIO

- **Skala nominal adalah paling sederhana, tidak mempunyai arti hitung, hanya mengkategorikan objek atau individu ke dalam data kualitatif, yang penting adalah kriteria untuk membedakan kategorinya (jenis kelamin, tingkat pendidikan, agama, bahasa), angka hanya simbol/label objek yang dianalisis atau identitas diri. Angka diolah dengan cara melaporkan jumlah hasil pengamatan setiap kategori.**
- **Teknik analisis data ordinal berdasarkan teori pasangan. Skala ordinal menunjuk pada posisi relatif individu/objek. Memiliki kategori yang diurutkan posisinya berdasarkan kriteria tertentu. Mempunyai makna lebih besar dari. Jarak antara urutan 1 dengan 2 tidak bermakna sama dengan jarak 2 dan 3. Rangkings tidak mempunyai interval yg tetap/sama**
- **Skala interval adalah skala yang mempunyai jarak yang sama dari suatu titik asal yang tetap. Hubungan, urutan dan jarak antara angka-angka dalam skala interval mengandung arti tersendiri. Misal, perbedaan skor siswa antara 80 dengan 90 mempunyai makna sama dengan perbedaan skor antara 30 dengan 40. Contoh, hasil tes: THB, pengukuran kecerdasan, dan pengukuran sikap.**
- **Analisis data interval (uji t dan korelasi). Uji t untuk membuktikan hipotesis komparatif atau mencari perbedaan antara dua variabel. Berfungsi menguji apakah perbedaan rerata antara dua sampel perbedaannya signifikan.**
- **Skala rasio, tertinggi sebab mempunyai titik nol sejati dan mempunyai interval yang sama. Contoh, pengukuran dengan alat ukur baku (meteran, kiloan). Semua prosedur dan analisis matematika dan statistika dapat digunakan untuk pengolahan data rasio.**

## DAFTAR DISTRIBUSI FREKUENSI

- **Peristilahan penting: Rentang (selisih skor tertinggi dan terendah), Interval, frekuensi, banyak kelas, panjang kelas, ujung kelas, batas kelas (batas nyata antara ujung atas suatu kelas/interval dengan ujung bawah kelas berikutnya (- 0,5 dan + 0,5), tanda kelas (nilai variabel antara ujung bawah dan ujung atas suatu kelas, sebagai wakil kelas).**
- **Setiap kelas (misal: 35–43) dibatasi dua buah skor, yaitu "batas bawah" (lower limit) adalah skor terendah pada kelas itu (35), dan "batas atas" (upper limit) adalah skor terbesar pada kelas itu (43).**
- **Setiap kelas juga memiliki batas nyata, yaitu "batas nyata bawah" (lower real limit) adalah batas bawah kelas itu dikurangi setengah dari satuan terkecil data itu dicatat (jika data dicatat dlm bilangan bulat, maka dikurangi dg 0,50), jika satuan terkecilnya 0,1 (data dicatat dlm satu desimal, maka dikurangi dg 0,05), sedangkan "batas nyata atas" (upper real limit) suatu kelas adalah batas atas kelas itu ditambah setengah dari satuan terkecil data yang bersangkutan dicatat. Misal:  $43+0,50 = 43,5$**
- **Titik tengah (midpoint), nilai yang membagi kelas itu menjadi dua bagian sama besar, yaitu  $\frac{1}{2}$  (batas bawah + batas atas suatu kelas). Misalnya:  $\frac{1}{2} (35 + 43) = 39$**
- **Distribusi frekuensi kumulatif adalah distribusi frekuensi dimana frekuensinya dijumlahkan secara meningkat, dan kelas intervalnya terbuka), ada "kurang dari dan lebih dari".**

## DAFTAR DISTRIBUSI FREKUENSI

- **Ditinjau dari nyata tidaknya frekuensi:**
  1. **Distribusi frekuensi absolut, yaitu: suatu jumlah bilangan yang menyatakan banyaknya data pada suatu kelompok tertentu, berdasarkan data apa adanya.**
  2. **Distribusi frekuensi relatif, yaitu; suatu jumlah persentase yang menyatakan banyaknya data pada suatu kelompok tertentu.**
- **Ditinjau dari jenisnya:**
  1. **Distribusi frekuensi numerik/kuantitatif (tunggal), yaitu distribusi frekuensi didasarkan pada data kontinum (data apa adanya).**
  2. **Distribusi frekuensi kategorikal/Kualitatif, didasarkan pada data yang terkelompok.**
- **Ditinjau dari kesatuannya:**
  1. **Distribusi frekuensi satuan, yaitu yang menunjukkan berapa banyaknya data pada kelompok tertentu (numerik maupun relatif).**
  2. **Distribusi frekuensi kumulatif, yaitu yang menunjukkan jumlah frekuensi pada sekelompok nilai/tingkat nilai tertentu mulai dari kelompok sebelumnya sampai dengan kelompok tersebut.**

**Langkah-langkahnya: (1) memilih/menentukan kelas, (2) memilih/menentukan data ke dalam kelas yang sesuai dengan tally, (3) menghitung jumlah dari setiap kelas, (4) menyajikannya dalam bentuk tabel distribusi frekuensi.**

## DAFTAR DISTRIBUSI FREKUENSI

- Raw score hasil tes kemampuan matematika sbb:

89 79 67 62 69 69 67 67 69 63 72 93 70 75 59 71 62 59 60 62  
65 36 64 65 59 56 91 85 77 70 57 67 57 54 52 73 50 50 54 72  
73 81 71 95 86 45 48 81 46 47 57 41 64 54 38 76 54 47 60 66  
66 83 77 82 41 56 43 50 55 57 72 66 68 75 63 67 70 78 56 68

- Langkah-langkah menyusun Daftar Distribusi Frekuensi
- Sebelumnya, susunlah data secara berurutan, dari terkecil ke terbesar atau sebaliknya.
- Buatlah daftar distribusi frekuensi numerik (tunggal)
  1. Menentukan rentang/range:  $St - Sr = 95 - 36 = 59$
  2. Menentukan banyak kelas:  $bk = 1 + 3,3 \log n = 1 + (3,3 \times 1,903) = 1 + 6,28 = 7,28$  dibulatkan menjadi 7
  3. Menentukan panjang kelas:  $p = R/bk = 59/7 = 8,4$  dibulatkan 9
  4. Interval kelas. Bilangan awalnya sebaiknya merupakan kelipatan "bk" dan tidak lebih kecil dari "Sr - bk". Bilangan awal harus sama dengan atau lebih kecil dari skor terkecil, yaitu "35", merupakan kelipatan "bk = 7".
  5. Menghitung frekuensi, dengan cara mentally/turus setiap data, misalnya ( // // ) = 4.

# GRAFIK

- **Grafik**, dibuat untuk merangkum dan menyederhanakan data yang kompleks menjadi suatu gambar informatif & mudah dipahami pembaca.
- **Histogram**, bentuk grafik yang menggambarkan distribusi frekuensi data (kontinu) dalam bentuk batang. Untuk data bentuk kategori (diskrit), tampilan yang serupa disebut diagram batang (bar chart). Ada sumbu datar/absis terdiri dari "batas nyata kelas", dan sumbu vertikal frekuensi data kelas tersebut. Sumbu datar dan sumbu tegak saling berpotongan tegak lurus, sehingga kaki setiap batang jatuh pada batas kelas (bawah dan atas) sehingga "titik tengah" berada di tengah kedua kaki batangnya. Disini diasumsikan skor-skor pada suatu interval kelas menyebar merata.
- **Frekuensi Poligon**, di sini skor-skor diasumsikan terpusat pada titik tengah kelasnya. Caranya dengan menarik suatu garis yang menghubungkan titik tengah setiap kelas sesuai dengan frekuensi masing-masing kelas. Kaki yang paling kiri jatuh pada titik tengah kelas di bawah kelas terkecil dan kaki paling kanan jatuh pada titik tengah kelas di atas kelas terbesar.
- **Ogif (ogive)**, poligon yang dibuat atas dasar frekuensi kumulatif seperangkat data. Disebut juga "Frekuensi poligon kumulatif" (Ferguson). Garis suatu ogif menghubungkan batas nyata atas-bawah setiap interval kelas. Menggambarkan secara visual jumlah subjek yang berada di bawah atau di atas skor tertentu. "Ozaiv" (Irianto, Agus, 19:1988).
- **Grafik lainnya**: grafik gambar (orang, binatang, dll), lingkaran, peta, dll.

## GRAFIK, UKURAN KECENDERUNGAN PUSAT DAN UKURAN LETAK

- **Histogram**, bentuk grafik yang menggambarkan distribusi frekuensi data (kontinu) dalam bentuk batang. Untuk data bentuk kategori (diskrit), tampilan yang serupa disebut diagram batang (bar chart). Ada sumbu datar/absis terdiri dari "batas nyata kelas", dan sumbu vertikal frekuensi data kelas tersebut. Setiap kaki batang jatuh pada batas kelas (bawah dan atas) sehingga "titik tengah" berada di tengah kedua kaki batangnya.
- **Mean** (rerata hitung, eks bar) data kuantitatif dalam sampel adalah hasil bagi jumlah nilai data oleh banyak data.  $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$ .
- **Modus** adalah fenomena yang paling banyak terjadi atau paling banyak terdapat. Bisa sebagai rerata data kualitatif. Untuk data kuantitatif modus ditentukan dengan jalan menentukan frekuensi terbanyak dalam data itu.
- **Median (Me)** menentukan letak data setelah data itu disusun menurut urutan nilainya. Untuk sampel genap setelah data diurutkan menurut nilainya,  $Me =$  rerata dua data tengah.
- **Kuartil**: bilangan pembagi untuk sekumpulan data yang dibagi menjadi empat bagian yang sama banyak, sesudah disusun menurut urutan nilainya. Ada tiga ( $K_1, K_2, K_3$ ).
- **Cara menentukan nilai kuartil**: 1). Susun data menurut urutan nilainya, 2). Tentukan letak kuartil, dan 3). Tentukan nilai kuartil.
- **Letak Kuartil ke-i ( $K_i$ )** = data ke  $i$  ( $(n + 1)/4$ ), dengan  $i = 1, 2, \text{ dan } 3$ .

## LANJUTAN UKURAN KECENDERUNGAN PUSAT

- $X$  (garis) =  $X$  berpalang =  $\bar{X}$  = Mean = Rerata =  $\frac{\sum X_i}{n}$  (data tunggal)
- =  $\frac{\sum f_i X_i}{\sum f_i} = \frac{899}{14} = 64,21428 = 64,21$

$X_i$	$f_i$	$f_i X_i$
70	5	350
69	6	414
45	3	135
	$\Sigma$	$\Sigma$

- Cara singkat/sandi/Code (gunakan salah satu tanda kelas,  $X_0 = 0$ )

skor	$f_i$	$X_i$	$f_i x_i$	C	$f_i C_i$
31 – 40	2	35,5	71	-3	-6
41 – 50	3	45,5	136,5	-2	-6
51 – 60	5	55,5	277,5	-1	-5
61 – 70	14	65,5	917	0	0
71 – 80	24	75,5	1810	1	24
81 – 90	20	85,5	1710	2	40
91 – 100	12	95,5	1146	3	36
	$\Sigma$			$\Sigma$	

$$\begin{aligned} \bar{X} &= \\ &Md + i \frac{(\sum f_i C_i)}{\sum f_i} \\ &65,5 + 10 \left( \frac{83}{80} \right) \\ &= 65,5 + 10,375 \\ &= 75,875 \end{aligned}$$

# MODUS

- Cara singkat/sandi (gunakan salah satu tanda kelas,  $X_0$  untuk nilai sandinya  $C = 0$ ). Untuk tanda kelas yang lebih kecil dari  $X_0$  berturut-turut diberi harga sandi  $C = -1, -2, \dots$  dst. Untuk tanda kelas yang lebih besar dari  $X_0$  berturut-turut diberi harga sandi  $C = +1, +2, \dots$ , dst. Berdasarkan contoh:
- $M_d = X_i$  (sejajar dengan  $C = 0$ ) = 65,5 ;  $i = 10$  ;  $\sum f_i C_i = 83$  ;  $\sum f_i = 80$
- Modus, fenomena yang paling banyak terjadi, dapat merupakan rata-rata data kualitatif.
- Rumus untuk data yang dikelompokkan,  $M_o = bb + p \left( \frac{b_1}{b_1 + b_2} \right)$

$bb$  = batas bawah kelas modus (kelas interval dengan  $f$  terbanyak) = 70,5

$p$  = panjang kelas = 10

Frekuensi kelas modus =  $f_i$  terbanyak = 24

$b_1$  =  $f$  kelas modus –  $f$  kelas interval sebelumnya ( $24 - 14 = 10$ )

$b_2$  =  $f$  kelas modus –  $f$  kelas interval sesudahnya ( $24 - 20 = 4$ )

$M_o = 70,5 + 10 \left( \frac{10}{10 + 4} \right) = 70,5 + 10 (0,714) = 70,5 + 7,1428$

$M_o = 77,643$



## LANJUTAN MEDIAN DAN UKURAN LETAK (KUARTIL)

- **Median, data genap setelah diurutkan merupakan rata-rata hitung dua data tengah. Median untuk data dalam daftar distribusi frekuensi, rumusnya:**

$$Me = bb + p \left( \frac{1}{2} \cdot n - F/f_i \right)$$

**n = ukuran sampel atau banyak data (80)**

**F = jumlah semua frekuensi dengan tanda kelas lebih kecil dari tanda kelas median. Contoh, berdasarkan data di atas, maka diketahui:**

$$\begin{aligned} \frac{1}{2} n &= 40 ; bb = 70,5 ; p = 10 ; f_i = 24 ; F = 2 + 3 + 5 + 14 = 24, \\ \text{maka } Me &= 70,5 + 10 \left( 40 - \frac{24}{24} \right) = 70,5 + 10 (0,666) = 70,5 + 6,666 \\ Me &= 77,1666. \end{aligned}$$

- **Cara menentukan nilai kuartil: 1) susun data menurut urutan nilainya, 2) tentukan letak kuartil, dan 3) tentukan nilai kuartil. Rumus:**

**Letak  $K_i$  = data ke  $i(n + 1)/4$  ; dimana  $i = 1, 2, 3$ . Contoh diketahui data:**

**75, 82, 66, 57, 64, 56, 92, 94, 86, 52, 60, 70. Kemudian disusun menjadi:**

**52, 56, 57, 60, 64, 66, 70, 75, 82, 86, 92, 94. Contoh, tentukan nilai  $K_3$ :**

$$\begin{aligned} \text{Letak } K_3 &= \text{data ke } \frac{3(12 + 1)}{4} = \text{data ke } 9 \frac{3}{4}, \text{ maka nilai } K_3 = \text{data ke } 9 \\ &+ \frac{3}{4}(\text{data ke } 10 - \text{data ke } 9) = 82 + \frac{3}{4}(86 - 82), \text{ maka } K_3 = 85 \end{aligned}$$

## UKURAN LETAK: KUARTIL, DESIL, PERSENTIL

- Untuk data dalam daftar distribusi frekuensi, maka rumus Kuartilnya:

- $K_i = bb + p \left( \frac{in/4 - F}{f} \right)$ , dengan  $i = 1, 2, 3$ .

bb = batas bawah kelas  $K_i$ , ialah kelas interval dimana  $K_i$  akan terletak

F = jumlah frekuensi dengan tanda kelas lebih kecil dari tanda kelas  $K_i$

Berdasarkan data, misal ingin menentukan  $K_3$ , kita perlu  $\frac{3}{4} \times 80 = 60$  data. Maka  $K_3$  terletak pada kelas interval ( $f_i = 2+3+5+14+24+20 = 60$ ), dari  $K_3$  ini didapatkan  $bb = 80,5$ ;  $p = 10$ ;  $f = 20$ ;  $F = 2+3+5+14+24 = 48$ ).

$$\begin{aligned} \text{Dengan } i = 3 \text{ dan } n = 80, \text{ maka } K_3 &= 80,5 + 10 \left( \frac{3 \times 80 / 4 - 48}{20} \right) \\ &= 80,5 + 10 \left( \frac{60 - 48}{20} \right) = 80,5 + 10 (0,6) = 86,5. \end{aligned}$$

Ini berarti ada 75% siswa yang mendapat skor paling tinggi 86,5 (misal : 86, 5; 85; 70); sedangkan 25% lagi mendapat skor paling rendah 86, 5 (misal: 87; 89, 90, dst).

- Desil ialah sekumpulan data dibagi menjadi 10 bagian yang sama (ada 9 desil,  $D_1$  s/d  $D_9$ ). Letak  $D_i = \text{data ke } i \text{ } (n + 1)/10$ . Contoh: Letak  $D_7 = \text{data ke } 7 \text{ } (12+1)/10 = 7 \times 13/10 = \text{data ke } 9,1$ . Maka nilai  $D_7 = \text{data ke } 9 + (0,1) \text{ (data ke } 10 - \text{data ke } 9)$ ; nilai  $D_7 = 82 + ((0,1) (86 - 82)) = 82 + (0, 1 \times 4) = 82,4$

## UKURAN LETAK: DESIL ( $D_i$ ), PERSENTIL ( $P_i$ )

- Ini berarti ada 70% siswa yang mendapat skor paling tinggi 82,4, sedangkan 30% lagi mendapat skor paling rendah 82,4
- Untuk data dalam daftar distribusi frekuensi, maka rumus Desil:

$$D_i = bb + p \left( \frac{in/10 - F}{fd} \right)$$

Bb = batas bawah kelas  $D_i$ , ialah kelas interval dimana  $D_i$  akan terletak

F = jumlah frekuensi dengan tanda kelas lebih kecil dari tanda kelas  $D_i$

Berdasarkan data, misal  $D_3$ , maka perlu:  $3 \times 80 / 10 = 24$  data, maka  $D_3$  terletak pada kelas interval ke 4, maka:  $bb = 60,5$ ;  $p = 10$ ;  $f = 14$ ;  $F = 2+3+5=10$ .

- Persentil, sekumpulan data dibagi menjadi 100 bagian yang sama (ada  $P_1$ – $P_{99}$ ). Maka letak  **$P_i = \text{data ke } \frac{i(n+1)}{100}$**

- Untuk data dalam daftar distribusi frekuensi, maka  **$P_i = bb + p \left( \frac{in/100 - F}{fp} \right)$**

Bb = batas bawah kelas  $P_i$ , ialah kelas interval dimana  $P_i$  terletak

F = Frekuensi kumulatif (Jumlah frekuensi dengan tanda kelas lebih kecil dari tanda kelas  $P_i$ ).

- Untuk data/sampel kecil, lebih baik gunakan data asli tidak dikelompokkan.

## UKURAN SIMPANGAN/DISPERSI/VARIASI

- Rentang (R), Rentang Antar Kuartil (RAK), Simpangan Kuartil (SK) atau Deviasi Kuartil, Rerata Simpangan (RS) atau Rerata Deviasi, Simpangan Baku (SB) atau Deviasi Standard, Varians dan Koefisien Variasi.
- Rentang: Data terbesar – Data terkecil (banyak digunakan dalam statistik industri)
- Rentang Antar Kuartil (RAK):  $K_3 - K_1$ , yaitu selisih antara  $K_3$  dan  $K_1$ . Misalnya,  $K_1 = 68$  dan  $K_3 = 90$ , maka  $RAK = 90 - 68 = 22$ . Ini ditafsirkan bahwa 50% dari data, nilainya paling rendah 68 dan paling tinggi 90 dengan perbedaan paling tinggi 22.
- SK atau Rentang Semi Antar Kuartil, harganya adalah setengah dari rentang antar kuartil.  $SK = \frac{1}{2} (K_3 - K_1)$ .
- Rata-rata Simpangan (RS), adalah jumlah harga mutlak dari selisih  $X_i$  dengan  $\bar{X}$  dibagi oleh  $n$ . Rumus  $RS = \frac{\sum |X_i - \bar{X}|}{n}$
- Contoh,  $X_i = 8, 7, 10, 11$  ;  $\bar{X} = 9$ , maka  $RS = \frac{6}{4} = 1 \frac{1}{2}$

## RATA-RATA SIMPANGAN, SIMPANGAN BAKU DAN VARIANS DATA TUNGGAL

- **Rata-rata Simpangan**

<u>Xi</u>	<u>Xi – X bar</u>	<u>  Xi – X bar  </u>	Maka RS = $\frac{\sum   Xi - X \text{ bar}  }{n}$
8	-1	1	$= \frac{6}{4}$
7	-2	2	
10	1	1	$= 1 \frac{1}{2}$
11	2	2	
n		$\Sigma$	

- **Simpangan baku untuk sampel simbolnya S (statistik), sedangkan untuk populasi simbolnya  $\sigma$  (sigma). Pangkat dua dari simpangan baku disebut Varians.**
- **Langkah-langkah mencari Varians sebagai berikut:**
  - Menghitung rerata Xbar**
  - Menentukan selisih dari Xi – Xbar**
  - Menentukan kuadrat selisih tersebut X1 – X bar, ..., Xn – X bar**
  - Kemudian kuadrat-kuadrat tersebut dijumlahkan (X1-Xbar)<sup>2</sup>, (Xn-Xbar)<sup>2</sup>**
  - Selanjutnya jumlah tersebut dibagi oleh (n – 1).**

# SIMPANGAN BAKU DAN VARIANS

- Jika ada sampel berukuran n dengan data  $X_1, X_2, \dots, X_n$ ; dan rata-rata ( $\bar{X}$ ),

A. Maka statistik  $S^2$  dihitung dengan rumus:  $S^2 = \frac{\sum(X_i - \bar{X})^2}{n - 1} = \frac{\sum x^2}{n - 1}$

- Contoh: sampel dengan data: 9, 8, 11, 12, 5.

$X_i$	$X_i - \bar{X}$	$(X_i - \bar{X})^2$
9	0	0
8	-1	1
11	2	4
12	3	9
5	-4	16

$\bar{X} = 45 : 5 = 9$   
 $\sum x^2 = 30$   
 $n - 1 = 5 - 1 = 4$   
Maka,  $S^2 = 30 : 4 = 7,5$   
Sehingga  $S = \sqrt{7,5} = 2,74$

B. Rumus Varians sampel lain (dengan nilai data asli, tanpa perlu  $\bar{X}$ )

$S^2 = \frac{n \cdot \sum X_i^2 - (\sum X_i)^2}{n(n - 1)}$  Rumus ini lebih baik, karena kekeliruannya lebih kecil.

$X_i$	$X_i^2$
9	81
8	64
11	121
12	144
5	25

$8^2 = 64; 11^2 = 121; 12^2 = 144; 5^2 = 25$ ; maka  $\sum X_i = 45; \sum X_i^2 = 354$   
maka  $S^2 = \frac{5 \times 354 - (45)^2}{5 \times 4} = \frac{1770 - 2025}{20} = \frac{150}{20} = 7,5$

## SIMPANGAN BAKU DAN VARIANS DALAM DDF

- Untuk data dalam Daftar Distribusi Frekuensi, rumus varians sbb:

$$1. S^2 = \frac{\sum f_i (X_i - \bar{X})^2}{n - 1} = \frac{\sum f_i (x)^2}{n - 1}$$

Skor	f <sub>i</sub>	X <sub>i</sub>	(X <sub>i</sub> - X̄)	(X <sub>i</sub> - X̄) <sup>2</sup>	f <sub>i</sub> (X <sub>i</sub> - X̄) <sup>2</sup>
.....	...	...	.....	.....	.....
	Σ...	-	-	-	Σ .....

$$2. S^2 = \frac{n \cdot \sum f_i X_i^2 - (\sum f_i X_i)^2}{n(n - 1)}$$

Skor	f <sub>i</sub>	X <sub>i</sub>	X <sub>i</sub> <sup>2</sup>	fX <sub>i</sub>	f <sub>i</sub> X <sub>i</sub> <sup>2</sup>
.....	...	...	....	....	.....
	Σ	-	-	Σ	Σ

- Keterangan: X<sub>i</sub> = tanda kelas ; n = Σf<sub>i</sub>  
f<sub>i</sub> = frekuensi yang sesuai tanda kelas X<sub>i</sub>

## SIMPANGAN BAKU DAN VARIANS

- **Rumus Varians dengan Cara singkat/sandi (C)**

$$S^2 = p^2 \frac{(n \cdot \sum f_i c_i^2) - (\sum f_i c_i)^2}{n(n-1)}$$

skor	f <sub>i</sub>	X <sub>i</sub>	C <sub>i</sub>	C <sub>i</sub> <sup>2</sup>	f <sub>i</sub> C <sub>i</sub>	f <sub>i</sub> C <sub>i</sub> <sup>2</sup>
31 – 40	2	35,5	-4	16	-8	32
41 – 50	3	45,5	-3	9	-9	27
51 – 60	5	55,5	-2	4	-10	20
61 – 70	14	65,5	-1	1	-14	14
71 – 80	24	75,5	0	0	0	0
81 – 90	20	85,5	1	1	20	20
91 – 100	12	95,5	2	4	24	48
Σ	80				3	161

$$S^2 = 10^2 \frac{(80 \times 161 - (3)^2)}{80 \times 79} = 100 \frac{(12880 - 9)}{6320} = 100 (2,0365) = 203,6$$

Ket: n = Σ f<sub>i</sub>;      p = panjang kelas = i (interval).



## **DISTRIBUSI NORMAL BAKU DAN DAERAH DI BAWAH KURVA NORMAL**

- Jika peubah  $X$  berdistribusi normal, dengan rerata =  $(\mu)$  dan  $S = \text{Sigma}$ . Maka jika setiap skor  $X_i$  diubah menjadi  $Z = (X_i - \text{Rerata}, \mu)/\text{Sigma}$ , maka distribusi  $Z$  akan merupakan distribusi normal baku (Freud & Walpole, 1987). Transformasi skor mentah ke skor baku ( $Z$ ) akan mengubah rerata dan varians suatu distribusi (menjadi secara berturut-turut 0 dan 1), tetapi tidak mengubah bentuk distribusi itu.
- Distribusi frekuensi skor  $Z =$  distribusi frekuensi skor mentah/aslinya.
- Distribusi normal baku dapat memecahkan permasalahan: (1). Sebagai rujukan menafsirkan data yang diperoleh; (2). Sebagai distribusi peluang, karenanya dapat digunakan menentukan besarnya peluang munculnya sst.
- Jika luas daerah distribusi normal dibagi menjadi beberapa bagian, maka dapat ditentukan frekuensi relatif (proporsi) skor yang berada pada bagian tertentu distribusi itu.
- Misalnya, lebih kurang  $1/3$  (0,3413) skor pada distribusi normal berada diantara rerata dan 1 SD di atas rerata.
- Oleh karena distribusi normal bersifat simetrik terhadap reratanya, maka kita tidak perlu menghitung luas daerah dari 0 ke  $Z$  yang bertanda negatif.