TO
TANTRIYA

STATISTIK DASAR DALAM PENELITIAN PENDIDIKAN

Oleh :

Dr. I Wayan Eka Mahendra, S.Pd., M.Pd.
Dra. Ni Nyoman Parmithi, MM.

Penerbit **PĀRAMITA** Surabaya

Katalog Dalam Terbitan (KDT)

**STATISTIK DASAR
DALAM PENELITIAN PENDIDIKAN**

Surabaya : Pāramita, 2015
viii + 232 hal ; 14.8 x 21 cm

ISBN : 978-602-204-547-2

STATISTIK DASAR DALAM PENELITIAN PENDIDIKAN

Oleh : Dr. I Wayan Eka Mahendra, S.Pd., M.Pd.
Dra. Ni Nyoman Parmithi, MM.
Lay Out & Cover : Udin

Penerbit & Percetakan : "PĀRAMITA"
Email:penerbitparamita@Gmail.com
http://www.penerbitparamita.com
Jl. Menanggal III No. 32 Telp. (031) 8295555, 8295500
Surabaya 60234 Fax : (031) 8295555

Pemasaran "PĀRAMITA"
Jl. Letda Made Putra 16 B Telp. (0361) 226445, 8424209
Denpasar Fax : (0361) 226445

Cetakan Pertama 2015

KATA PENGANTAR

Puji syukur penulis panjatkan kehadapan Ida Sang Hyang Widhi Wasa atas berkat dan Rahmat-Nya, buku dengan judul Statistik Dasar dalam Penelitian Pendidikan ini dapat diselesaikan pada waktunya. Bagi orang yang memiliki kemampuan dasar matematika tingkat menengah ke atas tentunya tidak akan sulit memahami konsep dasar statistik, apalagi statistik deskriptif. Tetapi, bagi sebagian orang statistik masih dianggap sebagai bacaan yang ruwet, penuh dengan angka dan simbul, serta membutuhkan ketelitian dalam analisisnya. Berdasarkan hal tersebutlah buku ini disusun. Selain itu, penulisan buku ini dilatarbelakangi oleh kepedulian dan ketertarikan penulis terhadap pembelajaran statistik dan fungsi statistik dalam penelitian, khususnya penelitian pendidikan.

Buku ini disajikan dalam bahasa yang sederhana, mudah dibaca, dan sangat mudah untuk dipahami. Buku ini bersifat midel, disajikan dengan singkat dan padat yang memuat konsep-konsep dasar statistik, khususnya satistik deskriptif. Sehingga buku ini sangat cocok sebagai buku pegangan (referensi) bagi mahasiswa maupun peneliti, dari disiplin ilmu murni maupun dari disiplin ilmu sosial.

Buku ini menguraikan konsep-konsep dasar statistik deskriptif yang disertai dengan contoh soal dan penyelesaiannya sesuai dengan dasar teori yang disajikan. Buku ini terdiri dari 7 bab, yang meliputi: Bab I Pendahuluan membahas tentang perbedaan statistik dan statistika, bab II tentang statistik deskriptif, bab III tentang pemusatan data, bab IV tentang ukuran letak dan ukuran penyebaran data, bab V tentang ukuran kemiringan (*skewness*) dan ukuran keruncingan (*kurtosis*), bab VI tentang kurva normal, dan bab VII tentang z-skor dan T-skor.

Melalui kesempatan ini kami menyampaikan ucapan terima kasih kepada pihak-pihak yang telah membatu dalam penyusunan buku ini. Tentunya buku ini jauh dari sempurna, oleh karena itu kami mengharapkan dari berbagai kalangan, khususnya yang berkecimpung dalam duania statistik agar dapat memberikan masukan yang dapat digunakan sebagai pertimbangan dalam menyempurnakan buku ini. Akhir kata penulis ucapakan semoga buku ini bermanfaat bagi kita semua.

Denpasar, Mei 2015

Penulis

DAFTAR ISI

KATA PENGANTAR v
DAFTAR ISI vii

BAB I PENDAHULUAN 1
A. Statistik dan Statistika 1
B. Fungsi Statistik dalam Penelitian 5
C. Macam-macam Statistik 6
D. Macam-macam Data 9

BAB II STATISTIK DESKRIPTIF 18
A. Tabel atau Daftar 18
1) Tabel distribusi frekuensi data tunggal......... 22
2) Tabel distribusi frekuensi data bergolong 24
B. Macam-macam Grafik/Diagram 60

BAB III PEMUSATAN DATA 84
A. Rata-rata (mean) 86
B. Nilai Tengan (median) 104
C. Modus 110
D. Hubungan Empiris antara Mean, Median, dan Modus 114

BAB IV UKURAN LETAK DAN UKURAN PENYEBARAN DATA 119
A. Ukuran Letak 120
1) Kuartil 121
2) Desil... 131
3) Persentil 141

B. Ukuran Penyebaran Data (dispersi) 152
1) Jangkauan 153
2) Simpangan kuartil 154
3) Simpangan rata-rata 157
4) Simpangan baku 159
5) Varian 161
6) Koefisien variasi 168

BAB V UKURAN KEMIRINGAN
DAN UKURAN KERUNCINGAN 172
A. Ukuran Kemiringan (*skewness*) 172
1) Koefisien kemiringan Pearson 174
2) Koefisien kemiringan Bowley 176
3) Koefisien kemiringan Persentil 176
4) Koefisien kemiringan Moment 177
B. Ukuran Keruncingan (*kurtosis*) 184
1) Koefisien keruncingan kuartil dan persentil.. 185
2) Koefisien kemiringan momen ke-4 186

BAB VI KURVA NORMAL 192

BAB VII Z-SKOR DAN T-SKOR 219
A. Z-skor.. 220
B. T-skor.. 224

DAFTAR PUSTAKA 229
CATATAN 231

BAB I
PENDAHULUAN

A. Statistik dan Statistika

Penggunaan istilah statistik berakar dari istilah-istilah dalam bahasa latin modern, yaitu *statisticum collegiums* yang berarti “dewan negara” dan bahasa Italia *statista* yang berarti “negarawan” atau “politikus”. Gottfried Achenwall (1749) menggunakan statistik dalam bahasa Jerman untuk pertama kalinya sebagai nama bagi kegiatan analisis data kenegaraan, dengan mengartikannya sebagai “ilmu tentang negara (*state*)”. Pada awal abad ke-19 telah terjadi pergeseran arti menjadi “ilmu mengenai pengumpulan dan klasifikasi data”. Sir John Sinclair memperkenalkan nama (*statistics*) ke dalam bahasa Inggris.

Jadi, statistik secara prinsip mula-mula hanya mengurus data yang dipakai lembaga-lembaga administratif dan pemerintahan. Jadi tidak mengherankan kalau pada mulanya statistik dipahami sebagai kumpulan angka-angka, tentang jumlah penduduk, angka tentang pendapatan masyarakat atau angka-angka lain yang berhubungan dengan masalah pemerintahan. Statistik dalam arti sempit diartikan sebagai data tetapi dalam arti luas diartikan sebagai alat.

Statistik pada dasarnya merupakan alat bantu untuk memberikan gambaran atas suatu kejadian melalui bentuk yang sederhana, baik berupa angka-angka ataupun grafik. Mengingat peranannya sebagai alat bantu, perlu disadari bahwa kunci keberhasilan analisis statistik masih terletak pada pemakaiannya. Statistik dapat diartikan sebagai: a) kumpulan data yang disusun dalam bentuk tabel atau diagram yang menggambarkan suatu persoalan, b) dipergunakan untuk menyatakan ukuran sebagai wakil dari kumpulan data mengenai sesuatu hal, c)suatu koleksi metode-metode yang dapat membantu seseorang dalam membuat keputusan-keputusan dari sejumlah informasi yang terbatas atau suatu alat untuk mengumpulkan, mengatur dan menganalisa data dari suatu percobaan/survei.

Berdasarkan beberapa pengertian di atas, secara umum statistik dapat diartikan sebagai sekumpulan cara atau aturan-aturan atau metode atau prosedur yang berkaitan dengan pengumpulan, analisis, penafsiran dan penarikan kesimpulan atas data-data yang berbentuk angka hasil penelitian dengan menggunakan suatu asumsi-asumsi tertentu.

Sementara itu, statistika dapat diartikan sebagai: teknik atau cara pengumpulan data, analisis data, penafsiran data dan penarikan kesimpulan. Statistika merupakan cabang dari matematika dan merupakan ilmu yang mempelajari cara-cara menentukan penduga, serta kemudian bertugas mengambil kesimpulan berdasarkan nilai pendugaan tersebut. Dengan kata lain statistika merupakan ilmu yang mempelajari statistik. Sebagian besar konsep dasar statistika mengasumsikan teori probabilitas (peluang).

Beberapa istilah statistika antara lain: populasi, sampel, unit sampel, probabilitas, dll. Dengan demikian antara istilah statistik dan statistika ada perbedaan, statistik merupakan penduga sedangkan statistika merupakan ilmu yang mempelajari penduga tersebut. Oleh karena statistika merupakan suatu metodelogi ilmiah, yang merupakan cabang dari matematika terapan, metode-metodenya adalah berbagai macam teknik mengumpulkan, mengorganiasi-kan, mentabelasi, menganalisis, menginterpretasikan, menggambarkan dan menyajikan data dalam bentuk angka-angka.

Statistika banyak diterapkan dalam berbagai disiplin ilmu, baik ilmu-ilmu alam (misalnya astronomi dan biologi maupun ilmu-ilmu sosial (termasuk sosiologi dan psikologi), maupun di bidang bisnis, ekonomi, dan industri. Statistika juga digunakan dalam pemerintahan untuk berbagai macam tujuan; sensus penduduk merupakan salah satu prosedur yang paling dikenal. Aplikasi statistika lainnya yang sekarang populer adalah prosedur jajak pendapat atau polling (misalnya dilakukan sebelum pemilihan umum), serta hitung cepat (perhitungan cepat hasil pemilu) atau *quick count*. Di bidang komputasi, statistika dapat pula diterapkan dalam pengenalan pola maupun kecerdasan buatan.

Cabang statistika yang pada saat ini sangat luas digunakan untuk mendukung metode ilmiah adalah statistika inferensial yang dikembangkan pada paruh kedua abad ke-19 dan awal abad ke-20 oleh Ronald Fisher (peletak dasar statistika inferensi), Karl Pearson (metode regresi linear), dan William Sealey Gosset (penemu uji-t). Penggunaan statistika pada masa sekarang dapat dikatakan telah menyentuh semua bidang ilmu pengetahuan, mulai dari astronomi hingga linguistika. Bidang-bidang ekonomi, biologi dan cabang-cabang terapannya, serta psikologi banyak dipengaruhi oleh statistika dalam metodologinya. Akibatnya lahirlah ilmu-ilmu gabungan seperti ekonometrika, biometrika (atau biostatistika), dan psikometrika.

Meskipun ada pihak yang menganggap statistika sebagai cabang dari matematika, tetapi sebagian pihak lainnya menganggap statistika sebagai bidang yang banyak terkait dengan matematika melihat dari sejarah dan aplikasinya. Di Indonesia, kajian statistika sebagian besar masuk dalam fakultas matematika dan ilmu pengetahuan alam, baik di dalam jurusan/fakultas tersendiri maupun tergabung dengan matematika.

B. Fungsi Statistik dalam Penelitian

Statistik memiliki fungsi yang sangat besar khususnya dalam penelitian pendidikan. Dua pekerjaan penting yang mencakup dalam statistik adalah menyajikan data dan menafsirkan data. Sehingga statistik akan dapat memberikan teknik yang tepat dalam mengumpulkan, mengklasifikasikan, dan menyajikan data sehingga hasil-hasil penelitian lebih mudah untuk ditafsirkan.

Fungsi statistik dalam penelitian pendidikan atara lain: a) membantu peneliti menghitung besarnya anggota sampel dari suatu populasi. Dengan demikian jumlah sampel yang diperlukan representatif, b) sebagai alat untuk menguji validitas dan menghitung koefisien reliabilitas instrumen, sehingga instrumen yang digunakan dalam penelitian betul-betul valid (sahih/tepat) dan reliabel (konsiten/ajeg), c) membantu peneliti membaca data yang telah terkumpul, d) membantu peneliti melihat ada tidaknya perbedaan atau hubungan antara satu variabel dengan variabel yang lain, e) membantu peneliti melakukan prediksi, dan f) membantu peneliti melakukan interprestasi atas data yang terkumpul, sehingga pada akhirnya membuat suatu kesimpulan.

C. Macam-macam Statistik

Statistik dapat dibedakan menjadi dua, yaitu statistik deskriptif (statistik dasar) dan statistik lanjut (statistik inferensial/ statistik induktif/ statistik probabilitas). **Statistik deskriptif** adalah statistik yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian yang tidak digunakan untuk membuat kesimpulan yang lebih luas (generalisasi). Statistik deskriptif merupakan metode yang mengatur, merangkum dan mempresentasikan data dengan cara yang informatif. Dengan kata lain statistik deskriptif adalah statistik yang membahas mengenai penyusunan data ke dalam daftar, grafik atau bentuk lain yang sama sekali tidak menyangkut penarikan kesimpulan. Proses statistik deskriptif dimulai dengan mengumpulkan data, pengorganisasian data, mengklasifikasikan serta penyajian dalam bentuk tabel, grafik maupun bentuk lainnya. Dengan demikian yang termasuk statistik deskriptif diantaranya:

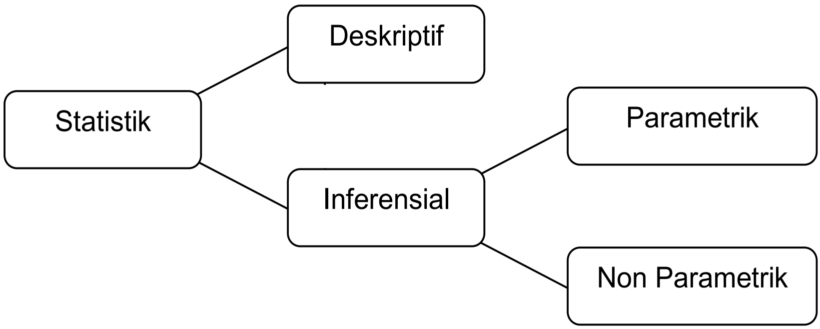
penyajian data dalam bentuk tabel maupun grafik, mean (rata-rata), median (nilai tengah), modus (nilai paling sering muncul), standar deviasi (simpangan baku), varian (ragam), kuartil, desil, maupun persentil. Berdasarkan paparan di atas, secara teknik dalam statistik deskriptif tidak ada uji hipotesis, tidak ada taraf kesalahan (taraf signifikansi) karena tidak melakukan generalisasi.

Statistik Inferensial adalah statistik yang digunakan untuk menganalisis data sampel dan hasilnya akan digeneralisasi pada populasi di mana sampel itu diambil. Bisa dikatakan bahwa statistik inferensial adalah metode yang digunakan untuk mengestimasi sifat populasi berdasarkan data sampel. Statistik ini akan cocok digunakan bila sampel diambil dari populasi yang jelas dan teknik pengambilan sampelnya akurat, diambil secara random sehingga diperoleh sampel yang representatif, sampel yang mewakili populasi. Kesimpulan yang diberlakukan terhadap populasi berdasarkan data sampel sebenarnya bersifat peluang atau *probability* oleh karena itu statistik ini disebut dengan statistik probabilitas. Suatu kesimpulan dari data sampel yang diberlakukan untuk populasi di mana sampel itu diambil memiliki peluang kesalahan dan kebenaran yang dinyatakan dalam bentuk persentase. Bila peluang kesalahannya (taraf signifikansi) 5%, maka peluang kepercayaannya (taraf signifikan) 95%.

Statistik inferensial dapat digolongkan menjadi dua, yaitu: statistik parametrik dan statistik non parametrik. Dalam bidang metodelogi parameter diartikan sebagai ciri-ciri tentang populasi. Untuk itu, statistik parametrik diartikan sebagai suatu prosedur pengambilan kesimpulan statistik yang didasarkan pada asumsi ciri-ciri populasi atau parameter. Asumsi tersebut adalah data yang diambil harus berdistribusi normal. Idealnya syarat-syarat parameter populasi haruslah berskala interval/rasio, sampel diambil dengan random, berdistribusi normal, memiliki varian yang homogen, model regresinya linier, dll. Syarat-syarat tersebut disebabkan karena dalam pengembangan rumus-rumus statistik inferensial didasarkan oleh beberapa asumsi. Oleh karenanya penggunaan statistik sebagai alat analisis tanpa diikuti dengan persyaratan yang diperlukan akan menyesatkan pemakainya.

Statistik non parametrik adalah suatu prosedur pengambilan kesimpulan statistik yang tidak didasarkan pada asumsi-asumsi parameter, artinya data yang diambil dari populasi tidak harus berdistribusi normal (bebas distribusi). Parameter populasi bebas dari syarat-syarat berskala interval/rasio, sampel diambil dengan random, berdistribusi normal, memiliki varians yang homogen, model regresinya linier, dll. Statistik non parametrik tidak menguji parameter populasi tetapi menguji distribusi populasi.

Beberapa ahli mengatakan bahwa statistik parametrik memiliki kekuatan yang lebih dibandingkan statistik non parametrik, apabila asumsi yang mendasarinya dapat dipenuhi. Bisa dikatakan statistik parametrik satu langkah lebih maju dibandingkan satatistik non parametrik. Secara skematis macam-macam statistik dapat digambarkan pada Gambar 1 berikut ini.



Gambar 1. Macam-macam Statistik

Penelitian yang dilakukan pada populasi atau tanpa mengambil sampel jelas akan menggunakan statistik deskriptif dalam analisis datanya. Tetapi apabila penelitian dilakukan pada sampel, analisis datanya dapat menggunakan statistik deskriptif maupun statistik inferensial, selama penggunaan statistik deskriptif hanya untuk mendeskripsikan sampel tanpa melakukan kesimpulan terhadap populasi dimana sampel itu diambil.

Secara sederhana untuk memahami statistik deskriptif dan statistik inferensial bisa diilustrasikan sebagai berikut. Misalnya seorang peneliti memperoleh data bahwa rata-rata hasil belajar statistik kelompok siswa yang sekolah di desa adalah 75, sedangkan rata-rata hasil belajar matematika kelompok siswa yang belajar di kota adalah 79. Jika peneliti hanya ingin mengetahui rara-rata hasil belajar matematika dari kedua kelompok tersebut melalui analisis statistik, peneliti cukup menggunakan statistik deskriptif.

Tetapi apabila peneliti ingin mengetahui lebih jauh dari rata-rata kedua kelompok tersebut, misalnya apakah rata-rata hasil belajar matematika kedua kelompok tersebut berbeda atau tidak dan kemudian ingin menarik kesimpulan dari rata-rata tersebut, maka statistik yang digunakan adalah statistik inferensial berupa uji beda rata-rata.

D. Macam-macam Data

Dalam statistik tidak bisa dilepaskan dari data, karena datalah yang nantinya yang akan dianalisis menggunakan teknik statistik, baik dengan statistik deskriptif maupun inferensial. Data dapat diartikan sebagai keterangan yang diperlukan untuk memecahkan suatu masalah. Data juga dapat diartikan sebagai catatan atas kumpulan fakta. Data merupakan bentuk jamak dari datum, berasal dari bahasa latin yang berarti "sesuatu yang diberikan". Dalam penggunaan sehari-hari data berarti suatu pernyataan yang diterima apa adanya.

Menurut sumbernya data digolongkan menjadi dua, yaitu data primer (*primary data*) dan data sekunder (*secondary data*). Data primer adalah data yang langsung diperoleh dari objeknya atau data yang diperoleh peneliti secara langsung (dari tangan pertama) tanpa perantara. Ketika seorang peneliti ingin mengetahui tingkat kecemasan siswa dalam menghadapi ujian nasional, peneliti secara langsung menyebarkan kuisioner pada responden atau juga data hasil wawancara peneliti dengan narasumber. Data-data yang diperoleh merupakan contoh data primer.

Sedangkan data sekunder adalah data yang diperoleh dalam bentuk data jadi, yang sudah dikumpulkan dan dianalisis oleh pihak lain. Data ini telah disediakan sebelumnya, sehingga kita tinggal mencari dan mengumpulkannya. Data sekunder dapat diperoleh dengan lebih mudah dibandingkan dengan data primer karena telah disediakan sebelumnya. Data sekunder umumnya berupa bukti, catatan, atau laporan yang telah tersusun dalam data dokumenter (arsip). Data sekunder biasanya digunakan sebagai pendukung data primer. Misalnya data absensi siswa selama satu semester, data tentang banyaknya siswa dalam satu kelas, dll.

Menurut sifatnya data juga digolongkan menjadi dua, yaitu data kualitatif dan data kuantitatif. Data kualitatif adalah data yang bukan merupakan bilangan, tetapi berbentuk kategori atau atribut (ciri-ciri, sifat-sifat, keadaan atau gambaran dari kualitas objek yang sedang diteliti).Data kualitatif berbentuk kata-kata atau berwujud pernyataan-pernyatan verbal yang bukan merupakan hasil pengukuran. Data kualitatif dijaring atau dikumpulkan berdasarkan cara-cara melihat suatu proses penelitian. Data ini lebih melihat proses dibandingkan hasil, karena didasarkan pada deskripsi proses, bukan didasarkan pada analisis matematis. Sebagai contoh tentang kualitas sebuah pensil, apakah baik, sedang, atau kurang.

Sedangkan data kuantitatif adalah data yang berupa bilangan yang bersifat variatif atau nilainya bisa berubah-ubah. Data ini diperoleh dari hasil pengukuran (pemberian angka pada atribut tertentu). Data semacam ini diperoleh lebih pada analisis matematis. Nilai ujian statistik Andi 75, berat badan Tono adalah 49 kg, merupakan contoh-contoh data kuantitatif.

Data kuantitatif digolongkan menjadi dua, yaitu: data diskrit (cacah) dan data kontinu (ukuran). Data diskrit adalah data yang diperoleh dengan cara membilang. Misalnya guru yang berpendidikan sarjana di SMA Sukamaju sebanyak 12 orang. Data nominal merupakan bagian dari data diskrit. Data kontinu adalah data yang diperoleh dengan cara mengukur. Misalnya tinggi rata-rata siswa perempuan di SMA Sukamaju adalah 145 cm. Data rasio, interval dan ordinal merupakan data-data yang tergolong data kontinu. Jenis-jenis data ini sangat menentukan statistik mana yang digunakan dalam analisisnya nanti. Oleh karena itu, pemahaman kita tentang berbagai jenis data mutlak diperlukan, sehingga bisa menggunakan statistik yang tepat dalam menganalisisnya. Data yang tergolong data rasio dan interval lebih tepat dalam analisisnya menggunakan statistik parametrik, sedangkan data ordinal dan data nominal lebih tepat dianalisis dengan menggunakan statistik non parametrik

a. Data Nominal

Data nominal dikatakan sebagai data kategori atau klasifikasi, yaitu data yang hanya memberikan label tanpa memberikan tingkatan apapun. Pada dasarnya mengacu pada kategori data diskrit seperti nama sekolah, jenis mobil, nama buku, dan yang lainnya. Data nominal adalah bentuk data yang paling sederhana. Merupakan bagian dari data kualitatif dan hanya bisa dianalisis dengan menggunakan statistik non parametrik. Data nominal hanya memberikan informasi yang bersifat dasar, kategori, klasifikasi, diskrit, tidak memiliki urutan sehingga tidak dapat dinotasikan dalam fungsi matematika.

Sebagai contoh pegawai negeri sipil diberi label 1, pegawai swasta diberi label 2, dan wirausaha diberi lebel 3. Pemberian label angka 1 pada pegawai negeri sipil, angka 2 pada pegawai swasta, dan angka 3 pada wirausaha tidak mengindikasikan bahwa tingkatan pegawai negeri sipil lebih tinggi dari pegawai swasta dan wirausaha. Tidak juga mengindikasikan bahwa tingkatan pegawai swasta lebih tinggi dari wirausaha, maupun sebaliknya. Posisi label data tersebut setara, tidak ada tingkatan. Label angka tersebut merupakan bilangan bulat dan bukan bilangan pecahan. Label angka tersebut juga tidak memberikan arti apa-apa jika dilakukan operasi matematika, seperti penjumlahan, pengurangan, perkalian maupun pembagian. Jika 3 - 2 = 1, maka tidak benar diartikan sebagai wirausaha-pegawai swasta = pegawai negeri sipil.

b. Data Ordinal

Data ordinal memiliki level yang lebih tinggi satu tingkat dari data nominal. Merupakan data dalam bentuk data nominal tetapi memiliki urutan. Bisa dikatakan data ordinal adalah data yang dinyatakan dalam bentuk kategori dan memiliki peringkat. Memiliki posisi yang tidak setara. Label angka yang diberikan memiliki tingkatan. Karena memiliki posisi yang tidak setara, sehingga sering digunakan untuk mengurutkan objek dari tingkatan yang paling tinggi ke tingkatan yang paling rendah atau sebaliknya. Masing-masing tingkatan tidak memiliki jarak yang pasti artinya, jika siswa yang memperoleh nilai 90 sebagai ranking I, yang memperoleh nilai 80 sebagai ranking II, maka siswa yang memperoleh ranking III nilainya tidak harus mempunyai selisi 10 dari ranking II, yaitu 70 bisa saja selisihnya lebih dari sepuluh atau bahkan kurang dari sepuluh asal nilai itu kurang dari 80 begitu seterusnya. Dari ilustrasi di atas dapat disimpulkan bahwa selisih nilai dari ranking I ke rangking II tidak sama dengan selisih nilai ranking II ke ranking III, walaupun sama-sama berbeda satu peringkat.

Untuk lebih memperjelas pemahaman kita tentang data ordinal berikut ini diberikan klasifikasi kepuasan siswa tentang cara mengajar gurunya dengan menyebarkan angket. Sangat puas diberi lebel 1, puas diberi lebel 2, cukup puas diberi lebel 3, kurang puas diberi lebel 4, dan tidak puas diberi lebel 5. Sebenarnya pemberian lebel tersebut tidak harus mulai dari 1 sampai 5, bisa saja angkanya dibalik dari 5 ke 1, atau tidak menggunakan angka 1 sampai 5, bisa angka yang lain tergantung kesepakatan. Dari kasus di atas sikap siswa yang sangat puas lebih tinggi dari sikap siswa yang puas, dan seterusnya. Tetapi berapa jarak dari sangat puas ke puas, dari puas ke cukup puas tidak diketahui dengan pasti. Data ordinal tidak mungkin dilakukannya operasi matematis, sama seperti data nominal. Tidak berlaku jika 1 + 2 = 3, maka sangat puas + puas = cukup puas. Data-data nominal ini nantinya bisa dianalisis dengan menggunakan statistik non parametrik.

c. Data Interval

Data interval adalah data yang paling sering digunakan dalam pengukuran pendidikan terutama aspek-aspek psikologi. Berbeda halnya dengan data ordinal yang memiliki jarak yang tidak pasti, data interval memiliki jarak yang sama pada setiap pengukuran sehingga dapat dibandingkan. Akan tetapi data interval tidak memiliki jumlah absolut dari objek yang diukur. Siswa yang memiliki nilai hasil belajar matematika 80 adalah dua kali dari nilai hasil belajar matematika siswa yang memiliki nilai 40, tetapi siswa yang memperoleh nilai 80 tidak dua kali lebih pintar dari siswa yang memperoleh nilai 40. Begitu juga perbedaan dari nilai 40 ke nilai 60 sama dengan perbedaan dari nilai 80 ke nilai 100, yaitu sama-sama naik 20 nilai tetapi dari segi kualitas akan lebih sulit memperoleh menaikkan nilai 80 menjadi 100 dibandingkan menaikkan nilai 40 menjadi 60. Data ini juga tidak memiliki nilai nol mutlak. Seorang siswa yang mendapat nilai ujian matematika 0 karena tidak menjawab dengan benar soal yang diberikan bukan berarti siswa itu tidak tahu materi matematika sama sekali.

Pada contoh data ordinal, yaitu angket kepuasan siswa tentang cara mengajar gurunya memang pada dasarnya data yang diperoleh melupakan data ordinal, karena merupakan tingkatan. Tetapi apabila angket yang diberikan terdiri dari beberapa butir pernyataan, data-data ordinal tersebut akan berubah menjadi data interval. Hal ini dilakukan karena jarak antar kategori dalam data ordinal tersebut diasumsikan sama sehingga skor totalnya digolongkan sebagai data interval. Diasumsikan sama karena tidak dapat diukur dengan pasti berapa jarak kepuasan siswa antara sangat puas dan puas, antara cukup puas dengan kurang puas dan seterusnya karena semua itu menyangkut perasaan seseorang. Dengan sedikit logika di atas kita bisa memahami kenapa angket yang diberikan seperti contoh di atas, yaitu angket kepuasan siswa tentang cara mengajar gurunya skor totalnya tergolong data interval.

d. **Data Rasio**

Dalam statistik data rasio adalah data yang memiliki tingkatan paling tinggi dan paling ideal. Dikatakan paling ideal karena rasio memiliki spesifikasi yang paling kuat diantara data-data lain, (data nominal, ordinal dan data interval). Data rasio juga memiliki ukuran yang paling kompleks dan memiliki sifat-sifat yang dimiliki oleh data nominal, data ordinal dan data interval serta ditambah dengan memiliki 0 mutlak (0 absolut), yaitu bilangan yang menunjukkan tidak ada gejala. Banyaknya buku Andi: jika 13, berarti ada 13 buku, jika 0, berarti tidak ada buku (absolut 0).

Selain memiliki 0 mutlak data rasio memiliki interval yang jelas, jarak antar kategori jelas (perbandingan maupun selisihnya). Misalnya, jika berat badan Eka adalah 40 kg dan berat badan Citra adalah 80 kg, maka berat badan Citra adalah dua kali berat badan Eka. Contoh lain: bila hari pertama Eka mampu meminum air putih 200 ml, hari kedua mampu minum 250 ml dalam sekali minum, sedangkan Citra mampu meminum 425 ml di hari pertama, 475 ml dalam sekali minum, jadi perubahan volume air putih yang mampu diminum oleh Eka dan Citra adalah sama yaitu 50 ml.

Alat-alat ukur seperti neraca (timbangan), meteran, termometer, barometer, dll yang digunakan dalam ilmu-ilmu fisika adalah alat ukur atau instrumen yang pengukurannya menghasilkan data rasio. Contoh-contoh data rasio adalah, ukuran panjang, ukuran berat, ketinggian, usia, dan lain sebagainya. Instrumen-instrumen dalam ilmu sosial dan humaniora tidak mampu mengukur ciri-ciri data rasio.

Setiap data dari yang paling tinggi, yaitu data rasio bisa di transformasikan menjadi data di bawahnya, yaitu data interval, data ordinal maupun, data nominal. Begitu juga data interval bisa ditransformasi menjadi data di bawahnya, yaitu data ordinal, maupun data nominal dst, tetapi tidak berlaku sebaliknya. Sebagai ilustrasi: misalnya seorang siswa yang mendapat nilai 100 akan menjadi rangking I, yang mendapat nilai 90 ranking II, dan mendapat nilai 80 menjadi ranking III, ilustrasi ini merupakan contoh transformasi dari data interval ke data ordinal. Tetapi sangat sulit mentransformasikan dari data ordinal kedata interval, misalnya bagi siswa yang memperoleh ranking satu harus diberi nilai berapa, begitu juga untuk ranking II dan ranking III.

Latihan 1

1. Jelaskanlah apa yang dimaksud dengan statistik dan statistika beserta contohnya serta sebutkan dan jelaskan bagian-bagian statistik teresbut!
2. Jelaskan kembali apa yang dimaksud dengan skala nominal, ordinal, interval dan rasio!
3. Skala apa yang digunakan dalam pengukuran-pengukuran di bawah ini.

a. suatu survei terhadap 500 orang tentang pekerjaanya yang menunjukkan bahwa 250 orang berasal dari Jakarta, 150 orang berasal dari Bali dan sisanya berasal dari Sumatra.
b. Pengukuran intelegensi (IQ) mahasiswa
c. Jarak yang ditempuh oleh mahasiswa ke kampus
d. Jumlah jam belajar mahasiswa per minggu
e. Kalasifikasi mahasiswa berdasarkan jenis kelamin

4. Berikan contoh transformasi dari skala rasio ke interval, dari skala interval ke ordinal!. Apakah mungkin melakukan transformasi dari skala interval ke skala rasio?

BAB II
STATISTIK DESKRIPTIF

A. Tabel atau Daftar

Ketika kita diberikan data tunggal dengan sederetan angka atau data dalam bentuk naskah, kemungkinan besar mengalami kesulitan untuk membaca data tersebut dan menginterpretasikannya. Oleh karena itu, dibutuhkan suatu cara agar data yang diberikan lebih mudah untuk dibaca dan diinterpretasikan. Salah satu cara tersebut adalah dengan membuat tabel atau daftar. Penyajian data dengan tabel atau daftar lebih baik daripada penyajian data dalam bentuk naskah. Tabel merupakan daftar yang berisi ikhtisar sejumlah data informasi, biasanya berupa kata-kata dan bilangan yang tersusun secara bersistem diurutkan ke bawah di lajur dan dengan deret tertentu dengan garis pembatas sehingga dapat dengan mudah dibaca.

Lajur dari atas ke bawah selanjutnya disebut dengan kolom, sedangkan dari lajur kiri ke kanan disebut dengan baris. Baris pertama dalam tabel disebut dengan kepala tabel. Tabel merupakan alat bantu visual selain grafik dan peta. Fungsi utama tabel adalah memudahkan pembaca untuk memahami isi data. Dengan kata lain fungsi tabel adalah menjelaskan suatu fakta atau informasi secara singkat, lebih menarik, dan lebih meyakinkan pembaca dibandingkan dengan kata-kata. Berikut ini disajikan beberapa jenis tabel.

a. Tabel baris dan kolom

Tabel baris dan kolom merupakan penyajian data dalam bentuk tabel dengan bentuk susunan baris dan kolom yang saling berhubungan. Berikut ini disajikan bentuk tabel baris dan kolom secara umum.

Judul Tabel

Judul Baris		sel		sel	
	sel				
			sel		
		sel			sel

→ Judul Kolom

→ Badan Tabel

Pada ilustrasi tabel di atas, dapat dilihat bagian-bagian tabel yang dapat dijelaskan sebagai berikut. Di bagian atas tabel merupakan judul tabel. Judul tabel harus disesuaikan dengan data yang akan disajikan dalam tabel. Kolom pertama pada tabel merupakan judul baris, sedangkan baris pertama pada tabel merupakan judul kolom yang sekaligus merupakan kepala tabel (lebih spesifik dapat dilihat kepala tabel adalah bagian yang diarsir). Bagian-bagian yang berisi kata “sel” merupakan badan tabel. Kita dapat membuat lebih dari satu tabel untuk sekumpulan data. Semakin banyak kategori atau klasifikasi data yang diberikan, semakin sulit tabel yang akan dibuat. Ada beberapa jenis tabel baris dan kolom, yaitu: tabel 1 arah (1 komponen), tabel 2 arah (2 komponen), tabel 3 arah (3 komponen).

Tabel satu arah adalah tabel yang memuat keterangan mengenai satu hal atau satu karakteristik (kategori) saja. Misalnya produksi padi Kabupaten Tabanan per Januari Tahun 2014 menurut varietas yang ditanam.

Tabel 2.1 Produksi Padi Kabupaten Tabanan per Januari Tahun 2014

Varietas Padi	Jumlah produksi (ton/ha)
Gogo	59
IR	35
PB	57
C4	86
Jumlah	**237**

Tabel 2.1 di atas terdiri dari empat sel baris dan dua sel kolom, sering disebut dengan tabel satu arah dengan ukuran 4 x 2. Tabel dua arah adalah adalah tabel yang menunjukkan hubungan antara dua hal atau karakteristik. Misalnya data mahasiswa yang dilihat menurut jurusan dan jenis kelaminnya, asal daerah dan agamanya, jurusan dan pekerjaan orang tua, usia dan jenis kelaminnya, dan lainnya.

Tabel 2.2 Banyaknya Mahasiswa di Suatu Universitas Negeri Dalam Satu Tahun Menurut Jurusan dan Asal

Jurusan	Bali	Lombok	NTT	NTB	Total
Matematika	45	31	24	25	125
Fisika	23	38	32	26	119
Biologi	54	24	12	25	115
Kimia	26	26	19	24	95
Total	**148**	**119**	**87**	**100**	**454**

Tabel 2.2 di atas terdiri dari empat sel baris dan empat sel kolom, sering disebut dengan tabel dua arah dengan ukuran 4 x 4. Tabel dua arah ini sering disebut tabel kontingensi,

tabel yang memiliki ciri khusus, yaitu tabel yang digunakan untuk menyajikan data yang terdiri atas dua hal atau dua faktor (karakteristik) saja. Faktor yang satu yang terdiri atas b kategori (pada baris) dan faktor yang lagi satu terdiri atas k kategori (pada kolom), dengan demikian dapat dibuat daftar tabel kontingensi berukuran b x k, dengan b menyatakan baris dan k menyatakan kolom.

Sementara itu, tabel tiga arah adalah tabel yang menunjukkan hubungan antara tiga hal atau karakteristik. Misalnya data mahasiswa yang dilihat menurut, daerah asal, jurusan dan jenis kelamin, asal daerah, pekerjaan orang tua dan agama, jurusan, agama, dan pekerjaan orang tua, usia, makanan kesukaan, dan jenis kelamin, dan lainnya.

Tabel 2.3 Banyaknya Mahasiswa di Suatu Universitas Negeri Dalam Satu Tahun Menurut Jurusan, Asal Daerah dan Jenis Kelamin

Jurusan	**Bali**		**Lombok**		**NTT**		**NTB**		**Total**
	L	**P**	**L**	**P**	**L**	**P**	**L**	**P**	
Matematika	45	22	31	22	24	11	25	12	**192**
Fisika	23	12	38	25	32	19	26	12	**187**
Biologi	54	45	24	24	12	13	25	15	**212**
Kimia	26	34	26	34	19	14	24	16	**193**
Total	**148**	**113**	**119**	**105**	**87**	**57**	**100**	**55**	**784**

b. Tabel distribusi frekuensi

Tabel distribusi frekuensi adalah penyajian data dalam bentuk tabel selain tabel kolom dan baris. Tabel ini digunakan untuk menyajikan data yang dikumpulkan dengan jumlah yang cukup banyak, sehingga dapat disajikan lebih baik dan jelas. Tabel distribusi frekuensi dibuat untuk menyederhanakan bentuk dan jumlah data, sehingga ketika disajikan lebih mudah dibaca dan dipahami.

Tabel distribusi frekuensi membagi data dalam beberapa kelas. Kelas yang dimaksud tidak hanya dalam bentuk bilangan, bisa jadi dalam bentuk kategori. Oleh karena itu, tabel distribusi frekuensi dibagi menjadi dua, yaitu tabel distribusi frekuensi *categorical* dan tabel distribusi frekuensi *numerical*. Tabel distribusi frekuensi *categorical* identik dengan tabel baris dan kolom satu arah seperti Tabel 2.1, di mana kelas-kelas dalam tabel tersebut dibagi berdasarkan macam-macam data atau golongan data yang dilakukan secara kualitatif.

Sedangkan tabel distribusi frekuensi *numerical* adalah tabel distribusi frekuensi di mana kelas-kelas dalam tabel tersebut dinyatakan dalam angka atau numerik. Tabel distribusi frekuensi *numerical* dibagi menjadi dua, yaitu tabel distribusi frekuensi data tunggal dan tabel distribusi frekuensi data bergolong/kelompok (dengan kelas interval).

1) Tabel distribusi frekuensi data tunggal

Tabel distribusi frekuensi data tunggal dibuat dengan cara menggabungkan data yang sama dalam satu kelas kemudian dihitung jumlahnya atau frekuensinya. Tabel ini digunakan untuk menyusun data yang jumlahnya relatif sedikit. Langkah-langkah yang dilakukan untuk membuat tabel distribusi data tunggal adalah: a) urutkanlah data tunggal dari nilai terbesar ke nilai terkecil atau sebaliknya, b) kelompokanlah masing masing data yang memiliki nilai yang sama, c) hitunglah banyaknya nilai pada masing-masing kelompok yang merupakan frekuensi masing-masing kelas dengan menggunakan turus (*tally)*, dan d) buat tabel distribusi frekuensinya.

Berikut ini disajikan nilai ujian statistik mahasiswa semester III jurusan pendidikan matematika pada suatu universitas.

66	70	80	78	90	75	70	64	66	80
82	78	70	75	90	64	82	82	78	82
80	66	80	70	78	70	75	66	70	64

Tentunya sangat sulit menarik suatu simpulan dari daftar data tersebut. Belum bisa menentukan berapa nilai ujian terkecil atau nilai ujian terbesar. Demikian pula, untuk mengetahui dengan tepat, berapa nilai ujian yang paling banyak atau berapa banyak mahasiswa yang mendapatkan nilai tertentu. Oleh karena itu, diperlukan analisis data tersebut terlebih dulu agar dapat memberikan gambaran atau keterangan yang lebih baik. Langkah-langkah yang ditempuh adalah sebagai berikut.

Pertama mengurutkan data tunggal tersebut dari data dengan nilai terbesar ke data dengan nilai terkecil.

90	90	82	82	82	82	80	80	80	80
78	78	78	78	75	75	75	70	70	70
70	70	70	66	66	66	66	64	64	64

Kedua, menghitung banyaknya data pada setiap kelompok dengan membuat turus (*tally*), dan *ketiga*, langsung membuat tabel distribusi frekuensi data tunggal.

Tabel 2.4 Nilai Ujian Statistik Mahasiswa Semester III Jurusan Pendidikan Matematika Pada Suatu Universitas

Nilai	Turus	Frekuensi (f)
90	//	2
82	////	4
80	////	4
78	////	4
75	///	3
70	//// /	6
66	////	4
64	///	3
Total	**-**	**30**

Dari tabel di atas dapat dilihat bahwa nilai ujian terkecil yang diperoleh mahasiswa adalah 64 dan nilai terbesarnya 90. Nilai ujian yang banyak adalah 70 yang diperoleh oleh 6 orang dari 40 orang yang mengikuti ujian. Untuk data yang sangat besar, jika menggunakan distribusi frekuensi data tunggal, maka akan diperoleh tabel yang panjang. Tentunya hal tersebut kurang efektif. Oleh karena itu, data tersebut harus dikelompokkan dalam kelas-kelas sehingga diperoleh tabel distribusi data berkelompok/bergolong.

2) Tabel distribusi frekuensi data bergolong

Tabel distribusi data berkelompok/bergolong disusun untuk data yang jumlahnya sangat banyak. Tabel ini akan membagi data menurut kelompoknya masing-masing yang selanjutnya disebut dengan kelas interval. Berikut akan diberikan bentuk umum tabel distribusi frekuensi data berkelompok sehingga nantinya lebih mudah untuk dipahami.

Tabel 2.5 Bentuk Umum Tabel Distribusi Frekuensi Data Berkelompok

No	Kelas Interval	Frekuensi
1.	a - b	f_1
2.	c - d	f_2
3.	e - f	f_3
4.	g - h	f_4
5.	i - j	f_5
6.	k - l	f_6
dst.	… - …	f_n
Total		$\sum_{i=1}^{n} f_i$

Kemungkinan beberapa orang akan mendefinisikan bentuk umum tabel distribusi frekuensi data berkelompok berbeda dengan Tabel 2.5 di atas. Hal tersebut tidak masalah selama pendefinisian tersebut sesuai dengan bentuk umumnya dan digunakan secara konsisten. Ada beberapa istilah yang harus dipahami dalam menyusun tabel distribusi frekuensi data berkelompok, istilah-istilah tersebut antara lain:

a) Nomor

Kolom pertama pada tabel bentuk umum distribusi frekuensi data berkelompok menyatakan banyak kelas interval yang dibuat. Jika nomornya dari 1 sampai dengan 6, maka banyak kelas interval data tersebut adalah 6 kelas interval. Hal ini menunjukkan bahwa data tunggal yang disusun ke dalam tabel dikelompokan menjadi 6 kelas interval.

b) Kelas interval (KI)

Kelas interval adalah selang interval tertentu yang membagi data menjadi beberapa kelompok. Selang interval inilah yang disebut dengan panjang kelas interval. Berarti di dalam kelas interval ada dua hal penting yang harus dipahami, yaitu: banyak kelas interval (k) serta panjang kelas interval (p). Pada Tabel 2.5 banyaknya kelas interval adalah 6 kelas interval. Kelas interval diberi nama dari atas ke bawah.

a - b adalah kelas interval pertama
c - d adalah kelas interval kedua
e - f adalah kelas interval ketiga
g - h adalah kelas interval keempat
i - j adalah kelas interval kelima
k - l adalah kelas interval keenam
dst.

Kelas interval juga dapat dikatakan sebagai banyaknya objek yang dikumpulkan dalam kelompok-kelompok tertentu. Kelas interval pertama yaitu a - b dimasukkan semua data yang bernilai a sampai dengan bernilai b. Kelas interval bisa diurutkan dari atas ke bawah dimulai dari data yang terkecil sampai data terbesar atau sebaliknya mulai dari data terbesar menuju ke data terkecil. Nilai awal ini disebut dengan *starting point*. Nilai a bisa diambil dari nilai data paling kecil (minimum) atau paling besar (maksimum), tergantung data yang ingin dibuat. Bisa juga diambil di bawah data paling kecil atau di atas data paling besar asal penyimpangannya tidak lebih dari 10% dari banyaknya data. Misalnya terdapat 40 buah data tunggal dengan nilai minimum 24 dan nilai maksimum 80. 10% dari 40 adalah 4, dengan demikian nilai a bisa mulai dari 20, 21, 22 atau 23 atau bila dimulai dari 84, 83, 82, atau 81. Dengan catatan seluruh data masuk ke dalam distribusi frekuensi.

c) Frekuensi (f)

Frekuensi adalah banyaknya kejadian (nilai) yang muncul pada selang kelas tertentu. Jumlah objek yang masuk dalam kelas interval tersebut, atau banyaknya data yang termasuk dalam kelompok suatu kelas interval. Frekuensi kelas dapat juga diartikan sebagai banyaknya data yang termasuk ke dalam kelas tertentu dari data acak. Frekuensi mewakili berapa kali data tersebut muncul dalam kelas interval tertentu.

d) Ujung/tepi kelas interval

Secara umum ujung/tepi kelas merupakan batas nyata suatu kelas interval. Batas kelas tidak memiliki tempat atau ruang untuk nilai-nilai antara kelas interval yang satu dengan kelas interval yang lainnya. Terdapat dua tepi kelas yang berbeda, yaitu: tepi bawah kelas dan tepi atas kelas. Tepi bawah kelas interval adalah bilangan yang terletak di ujung kiri pada masing-masing kelas interval. Dari Tabel 2.5 tepi bawah masing-masing kelas interval adalah: a, c, e, g, i, k, dst.

a merupakan tepi bawah kelas interval pertama
c merupakan tepi bawah kelas interval kedua
e merupakan tepi bawah kelas interval ketiga
g merupakan tepi bawah kelas interval keempat
i merupakan tepi bawah kelas interval kelima
dst.

Sedangkan tepi atas kelas interval adalah bilangan yang terletak di ujung kanan pada masing-masing kelas interval.

Dari Tabel 2.5 tepi atas masing-masing kelas interval adalah: b, d, f, h, j, l, dst.

b merupakan tepi atas kelas interval pertama
d merupakan tepi atas kelas interval kedua
f merupakan tepi atas kelas interval ketiga
h merupakan tepi atas kelas interval keempat
j merupakan tepi atas kelas interval kelima
dst.

Apabila kelas interval suatu data distribusi frekuensi data berkelompok dimulai dari nilai terbesar ke nilai terkecil, maka ujung/tepi bawah kelas interval terletak di ujung sebelah kanan, sedangkan tepi/ujung atas kelas interval terletak di sebelah kiri. Berlaku sebaliknya dengan ketentuan pada Tabel 2.5.

e) Batas kelas

Batas kelas merupakan limit kelas sesungguhnya. Batas kelas (*class limits*) merupakan nilai-nilai yang membatasi kelas interval yang satu dengan kelas interval yang lain. Batas kelas merupakan batas semu dari setiap kelas interval. Dikatakan semu karena di antara kelas yang satu dengan kelas yang lain masih terdapat tempat atau ruang untuk nilai-nilai tertentu yang terdapat dalam suatu kelas interval yang tidak nampak secara nyata. Misalnya antara kelas interval pertama dengan kelas interval kedua, yaitu tepi atas kelas interval pertama dengan tepi bawah kelas interval ke dua terdapat nilai atau angka yang tidak memiliki ruang atau tempat. Terdapat dua batas kelas, yaitu: batas kelas bawah (*lower class limits*) dan batas kelas atas (*upper class limits*). Hubungannya dengan tepi kelas, batas bawah merupakan tepi bawah dikurangi ketelitian yang ditentukan, dan batas atas merupakan tepi atas ditambah ketelitian yang ditentukan. Ketelitian yang digunakan tergantung dengan nilai dari tepi kelas.

(a) Jika tepi kelas merupakan bilangan bulat, maka ketelitiannya 0,5. Dengan demikian batas atas ditambah 0,5 dan batas bawah dikurangi 0,5.
(b) Jika tepi kelas merupakan bilangan desimal satu angka dibelakang koma, maka ketelitiannya 0,05. Dengan demikian batas atas ditambah 0,05 dan batas bawah dikurangi 0,05.
(c) Jika tepi kelas merupakan bilangan desimal dua angka dibelakang koma, maka ketelitiannya 0,005. Dengan demikian batas atas ditambah 0,005 dan batas bawah dikurangi 0,005. Begitu seterusnya

Dengan demikian dapat ditentukan batas atas dan batas bawah masing-masing kelas interval pada Tabel 2.5 sebagai berikut. Batas bawahnya adalah sebagai berikut.

a - 0,5 merupakan batas bawah kelas interval pertama
c - 0,5 merupakan batas bawah kelas interval kedua
e - 0,5 merupakan batas bawah kelas interval ketiga
g - 0,5 merupakan batas bawah kelas interval keempat
i - 0,5 merupakan batas bawah kelas interval kelima
dst.

Sedangkan batas atas masing-masing kelas interval adalah sebagai berikut.

b + 0,5 merupakan batas atas kelas interval pertama
d + 0,5 merupakan batas atas kelas interval kedua
f + 0,5 merupakan batas atas kelas interval ketiga
h + 0,5 merupakan batas atas kelas interval keempat
j + 0,5 merupakan batas atas kelas interval kelima
dst.

Dapat dilihat pula bahwa batas atas kelas interval pertama merupakan batas bawah kelas interval kedua. Batas atas kelas interval kedua merupakan batas bawah kelas interval ketiga. Batas atas kelas interval ketiga merupakan batas bawah kelas interval keempat, begitu seterusnya.

f) Titik Tengah (x_i)

Titik tengah suatu kelas interval sering disebut tanda kelas. Titik tengah kelas atau tanda kelas adalah angka atau nilai data yang tepat terletak di tengah suatu kelas interval. Titik tengah suatu kelas interval ditentukan dengan cara menjumlahkan ujung bawah dengan ujung atas suatu kelas interval yang hasilnya kemudian dibagi dua.

$$\text{Titik tengah} = \frac{\text{ujung bawah} + \text{ujung atas}}{2}$$

Dengan menggunakan rumus tersebut maka data pada Tabel 2.5 diperoleh titik tengah masing-masing kelas interval, yaitu :

$p = \frac{R}{k}$, adalah titik tengah kelas interval pertama

$\frac{c+d}{2}$, adalah titik tengah kelas interval kedua

$\frac{e+f}{2}$, adalah titik tengah kelas interval ketiga

$\frac{g+h}{2}$, adalah titik tengah kelas interval keempat

dst.

g) Panjang kelas interval

Panjang suatu kelas interval merupakan jarak dari ujung/tepi bawah kelas interval sampai dengan ujung/tepi atas kelas interval, yang perlu diperhatikan adalah tepi bawah kelas interval harus ikut dihitung. Misalnya kelas panjanganya suatu kelas interval dimulai dari 50 - 54, panjang kelas interval tersebut adalah 5, yaitu dimulai dari 50, 51, 52, 53 dan 54. Secara matematis panjang kelas interval adalah interval tertutup [50,54]. Untuk data yang kelas intervalnya dalam bentuk bilangan bulat sangat mudah menentukan panjang kelas intervalnya. Salah satunya dengan cara menjumlahkan ujung atas dan ujung bawah kelas interval bersangkutan kemudian ditambah satu. Seperti pada Tabel 2.5, panjang kelas intervalnya bisa kita tentukan misalnya dari kelas interval pertama, yaitu a + b + 1, atau c + d + 1, dan seterusnya dengan ujung-ujung kelas interval merupakan bilangan bulat. Ditambah 1 sebenarnya diperoleh dari dua kali ketelitian yang digunakan, di mana untuk kelas interval yang merupakan bilangan bulat ketelitiannya adalah 0,5 seperti yang diungkapkan pada bagian batas bawah dan batas atas.

Kita akan mengalami sedikit kesulitan apabila kita menemukan ujung-ujung kelas interval, baik ujung bawah maupun ujung atas merupakan bilangan desimal. Untuk itu berikut ini diberikan beberapa cara menentukan panjang suatu kelas interval. 1) panjang kelas interval diperoleh dengan cara mengurangi ujung/tepi bawah kelas interval setelahnya dengan ujung/tepi bawah kelas interval bersangkutan atau mengurangi ujung bawah kelas interval bersangkutan dengan ujung bawah kelas interval sebelumnya, 2) panjang kelas interval diperoleh dengan cara mengurangi ujung/tepi atas kelas interval setelahnya dengan ujung/tepi atas kelas interval bersangkutan atau mengurangi ujung atas kelas interval bersangkutan dengan ujung atas kelas interval sebelumnya, dan 3) panjang kelas interval diperoleh dengan cara menambahkan ujung atas dan ujung bawah kelas interval bersangkutan dan hasilnya kemudian ditambahkan dengan 2 kali ketelitian yang digunakan.

Menyusun sekumpulan data tunggal ke dalam data distribusi frekuensi bergolong dapat dilakukan dengan dua cara, yaitu dengan cara konvensional dan dengan cara aturan Sturges.

h) Cara konvensional

(a) Urutkanlah data tunggal yang diketahui dari nilai terbesar ke nilai terkecil atau sebaliknya.

(b) Tentukan jangkauan (rentangan/range)

Rentangan merupakan jarak antara data maksimum dengan data minimum.

R = nilai maksimum – nilai minimum

(c) Tentukan banyak kelas interval dan panjang kelas interval

Dengan cara konvensional peneliti bisa terlebih dahulu menentukan banyak kelas interval kemudian baru menentukan panjang kelas interval. Begitu juga sebaliknya menentukan terlebih dahulu kisaran panjang kelas interval dengan menggunakan rumus kemudian baru menentukan banyak kelas interval.

Pertama, apabila peneliti terlebih dahulu menentukan banyak kelas interval biasanya peneliti mengambil paling sedikit 4 sampai paling banyak 20 kelas interval atau ada yang mengambil 5 sampai dengan 15 tergantung peneliti. Setelah itu baru menentukan panjang kelas interval. Panjang kelas interval disesuaikan dengan pola atau banyaknya data sehingga semua data masuk ke dalam kelompok-kelompok kelas interval.

Kedua, apabila peneliti terlebih dahulu menentukan panjang kelas interval peneliti bisa menggunakan rumus kisaran panjang kelas intervalnya, yaitu :

Selang maksimum (I_{max}) = R/7
Selang minimum (I_{min}) = R/15
Dari ketentuan di atas diperoleh $I_{max} \leq p \leq I_{min}$ atau $R/_{15} \leq p \leq R/_{7}$

Dengan demikian banyak kelas interval ditentukan dari interval yang diberikan. Tentunya kedua cara ini tidak memberikan pilihan yang pasti berapa panjang kelas interval dan banyak kelas interval yang harus dipilih. Jadi mungkin saja untuk data tunggal yang sama dua orang peneliti akan menemukan tabel distribusi frekuensi data bergolong yang berbeda tergantung pilihan masing-masing peneliti.

Seperti yang terlihat pada cara kedua, banyak kelas interval ditentukan dari interval yang diberikan. Sehingga semakin panjang interval yang diberikan tentunya semakin banyak pula pilihan yang diambil di dalam menentukan banyak kelas interval. Begitu juga sebaliknya, semakin pendek intervalnya semakin sedikit pula pilihannya. Padahal tujuan dari pengelompokan data ke dalam distribusi frekuensi digunakan untuk mengungkap atau menekankan pola dari data kelompok tersebut. Sehingga terlalu sedikit atau terlalu banyak kelas interval akan mengaburkan pola data tersebut. Untuk mengatasi kelemahan tersebut, ditemukan suatu cara menentukan panjang kelas interval dan banyak kelas interval oleh H. A. Sturges pada tahun 1926 yang selanjutnya disebut dengan aturan Sturges.

i). Aturan Sturges

Menyusun sekumpulan data tunggal ke dalam data distribusi frekuensi bergolong dengan aturan Sturges dapat mengikuti langkah-langkah berikut ini.

(a) Urutkanlah data tunggal yang diketahui dari nilai terbesar ke nilai terkecil atau sebaliknya.

(b) Tentukan jangkauan (rentangan/range)

Rentangan merupakan jarak antara data maksimum dengan data minimum.

$$R = \text{nilai maksimum} - \text{nilai minimum}$$

(c) Tentukan banyak kelas interval dan panjang kelas interval

Aturan sturges digunakan untuk menentukan banyak kelas interval dengan pasti (walau dalam beberapa kasus terjadi pembulatan). Rumus yang digunakan adalah sebagai berikut.

$$k = 1 + 3{,}3 \log (n)$$

dengan
k = banyak kelas interval
n = banyaknya data

Dari rumus tersebut tentunya nilai k tidak selamanya bilangan bulat bisa terjadi nilai k merupakan bilangan desimal semua itu tergantung nilai log n. Jika nilai log n kelipatan 10, maka nilai k pasti bilangan bulat, jika tidak maka nilai k merupakan bilangan desimal. Apabila nilai k bilangan desimal, maka dilakukan pembulatan. Pembulatan bisa dilakukan ke atas atau ke bawah. Tetapi sebaiknya pembulatan dilakukan ke atas, agar kemungkinan data tidak masuk ke dalam kelas interval semakin kecil. Karena ada kemungkinan untuk data-data *out layer* tidak masuk dalam kelas interval yang telah ditentukan. Hal itu bisa dilakukan dengan cara membuang data *out layer* atau menambah banyak kelas interval.

Setelah menentukan banyak kelas interval, kemudian ditentukan panjang kelas interval dengan rumus sebagai berikut.

$$p = \frac{R}{k}$$

dengan

p = panjang kelas interval

R = rentangan/jangkauan

k = banyak kelas interval

Contoh 2.1 Berikut ini disajikan data hasil ujian akhir semester (UAS) mata kuliah statistik dasar mahasiswa suatu perguruan tinggi

50	68	73	70	96	79	65	97
86	84	79	65	78	78	73	80
67	75	88	75	82	89	67	73

73	82	87	82	73	87	75	72
57	81	68	71	74	94	75	78
88	72	90	93	62	77	95	80
78	63	55	55	54	60	70	76
75	78	60	72	82	55	54	71
90	74	56	76	74	63	80	88
75	70	63	61	66	66	75	75

Susunlah data di atas ke dalam tabel distribusi frekuensi berkelompok!

Langkah-langkah penyelesainnya

(1) Urutkan data dari nilai terkecil ke nilai terbesar

$\boxed{50}$	54	54	55	55	55	56	57	60	60
61	62	63	63	63	65	65	66	66	67
67	68	68	70	70	70	71	71	72	72
72	73	73	73	73	73	74	74	74	75
75	75	75	75	75	75	75	76	76	77
78	78	78	78	78	79	79	80	80	80
81	82	82	82	82	84	86	87	87	88
88	88	89	90	90	93	94	95	96	$\boxed{97}$

(2) Menentukan rentangan atau jangkauan (R)

R = nilai terbesar – nilai terkecil

= 97 - 50

= 47

Dengan cara tradisional

Menentukan panjang kelas interval dengan mencari selang yang diberikan.

$$I_{mak} = \frac{R}{7}$$

$$= \frac{47}{7}$$

$$= 6{,}71$$

$$I_{min} = \frac{R}{15}$$

$$= \frac{47}{15}$$

$$= 3{,}13$$

Pilih panjang kelas interval di antara 3 sampai dengan 6 (3,13 ≤ *P* ≤ 6,71), dipilih panjang kelas interval 5. Dengan melihat struktur data dipilih pula banyak kelas interval sebanyak 10. Sehingga semua data tunggal terakomodasi dalam distribusi frekuensi data bergolong.

(a) Ujung bawah kelas interval pertama diambil 50 (sesuai dengan nilai data terkecil). Berarti kelas interval pertama mulai dari 50 sampai dengan 54 atau 50 - 54, karena panjang kelas interval 5. Data-data yang masuk ke dalam kelas interval pertama adalah 50, 54, dan 54. Sehingga turusnya sebanyak /// dan frekuensinya 3.

(b) Kelas interval kedua mulai dari 55 sampai dengan 59 atau 55 - 59. Data-data yang masuk ke dalam kelas interval kedua adalah 55, 55, 55, 56, dan 57. Sehingga turusnya sebanyak ///// dan frekuensinya 5.

(c) Begitu seterusnya sampai data yang terakhir habis. Setiap memasukkan data ke dalam kelas interval tertentu sebaiknya data yang dimasukkan tersebut langsung dicoret menghindari kemungkinan dipilih lagi. Secara lengkap hasil dengan menggunakan cara tradisional dapat dilihat pada Tabel 2.6 berikut ini.

Tabel 2.6 Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Interval	*Tally*	Frekuensi
1	50 - 54	///	3
2	55 - 59	/////	5
3	60 - 64	///// //	7
4	65 - 69	///// ///	8
5	70 - 74	///// ///// ///// /	16
6	75 - 79	///// ///// ///// ///	18
7	80 - 84	///// ////	9
8	85 - 89	///// //	7
9	90 - 94	////	4
10	95 - 99	///	3
Total		**-**	**80**

Tabel distribusi frekuensi data bergolong yang diperoleh dari data tunggal dengan mengunakkan cara tradisional akan dibandingkan dengan cara dengan aturan *Sturges*.

Dengan aturan Sturges

Menentukan banyaknya kelas interval

k = 1 + 3,3 log n, dengan n = 80

= 1 + 3,3 log 80 (1,903)

$= 1 + (3,3)(1,903)$
$= 7,279 \approx 7$ (dibulatkan)

Menentukan banyaknya kelas interval

$$p = \frac{R}{k}$$
$$= \frac{47}{7}$$
$$= 6,71 \approx 7 \text{ (dibulatkan)}$$

Berdasarkan hasil analisis diambil banyaknya kelas interval tujuh dan panjang kelas interval juga tujuh.

(a) Ujung bawah kelas interval pertama diambil 50 (sesuai dengan nilai data terkecil) sama seperti cara tradisional. Berarti kelas interval pertama mulai dari 50 sampai dengan 56 atau 50 - 56, karena panjang kelas interval 7. Data-data yang masuk ke dalam kelas interval pertama adalah 50, 54, 54, 55, 55, 55 dan 56. Sehingga turusnya sebanyak ///// // dan frekuensinya 7.

(b) Kelas interval kedua mulai dari 57 sampai dengan 63 atau 57 - 63. Data-data yang masuk ke dalam kelas interval kedua adalah 57, 60, 60, 61, 62, 63, 63 dan 63. Sehingga turusnya sebanyak ///// /// dan frekuensinya 8.

(c) Begitu seterusnya sampai dengan kelas interval ketujuh. Setiap memasukkan data ke dalam kelas interval tertentu sebaiknya data yang dimasukkan tersebut langsung dicoret menghindari kemungkinan dipilih lagi. Secara lengkap hasil dengan menggunakan aturan Sturges dapat dilihat pada Tabel 2.7 berikut ini.

Tabel 2.7 Hasil Ujian Akhir Semester (Uas) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Interval	*Tally*	Frekuensi
1	50 - 56	///// //	7
2	57 - 63	///// ///	8
3	64 - 70	///// ///// /	11
4	71 - 77	///// ///// ///// ///// ////	24
5	78 - 84	///// ///// ///// /	16
6	85 - 91	///// ////	9
7	92 - 98	/////	5
Total		**-**	**80**

Sebenarnya turus (*tally*) tidak harus dibuat dalam daftar tabel distribusi frekuensi. Kegunaan turus hanya untuk menghitung satu persatu data dalam suatu kelas interval agar lebih teliti, jadi penulisan turus bisa dihilangkan. Distribusi frekuensi yang terdapat padat tabel 2.7 di atas disebut dengan distribusi frekuensi absolute, yaitu frekuensi mutlak setiap kelas interval dalam bentuk angka. Frekuensi absolute ini bisa dirubah ke dalam distribusi frekuensi relatif (f_r), yaitu distribusi frekuensi masing-masing kelas interval berbanding dengan frekuensi totalnya dikalikan 100%. Distribusi frekuensi relatif ini dalam bentuk persentase (%). Idealnya jumlah frekuensi relatif adalah 100% tetapi dalam beberapa kasus jumlah frekuensi relatif biasanya bisa kurang dari 100% atau bahkan lebih dari 100% karena terjadi pembulatan. Tetapi kekurangan atau kelebihan tersebut biasanya tidak lebih dari 1. Pada kolom frekuensi relatif tanda persentase (%) tidak perlu ditulis, cukup ditulis pada baris pertama kolom bersangkutan. Berdasarkan Tabel 2.5, dapat dibuat tabel distribusi relatif sebagai berikut.

Tabel 2.8 Bentuk Tabel Distribusi Frekuensi Relatif

No	Kelas Interval	Frekuensi Relatif (%)
1.	a - b	f_{r1}
2.	c - d	f_{r2}
3.	e - f	f_{r3}
4.	g - h	f_{r4}
5.	i - j	f_{r5}
6.	k - l	f_{r6}
dst.	… - …	f_{rn}
Total		100

i. Frekuensi relatif kelas interval pertama

$$f_{r1} = \frac{f_1}{\sum_{i=1}^{n} f_i} \times 100\%$$

ii. Frekuensi relatif kelas interval kedua

$$f_{r2} = \frac{f_2}{\sum_{i=1}^{n} f_i} \times 100\%$$

iii. Frekuensi relatif kelas interval ketiga

$$f_{r3} = \frac{f_3}{\sum_{i=1}^{n} f_i} \times 100\%$$

iv. Frekuensi relatif kelas interval keempat

$$f_{r4} = \frac{f_4}{\sum_{i=1}^{n} f_i} \times 100\%$$

v. Frekuensi relatif kelas interval kelima

$$f_{r5} = \frac{f_5}{\sum_{i=1}^{n} f_i} \times 100\%$$

vi. Begitu seterusnya

$$f_{rn} = \frac{f_n}{\sum_{i=1}^{n} f_i} \times 100\%$$

Selain distribusi frekuensi relatif terdapat juga distribusi frekuensi kumulatif, yaitu distribusi yang nilai frekuensinya (f) diperoleh dengan cara menjumlahkan frekuensi demi frekuensi atau selangkah-demi selangkah. Distribusi frekuensi kumulatif dalam bentuk angka absolute bukan persentase (%). Terdapat dua jenis distribusi frekuensi kumulatif, yaitu distribusi frekuensi kumulatif kurang dari dan distribusi frekuensi kumulatif lebih dari atau sama dengan.

Sebagai patokan dalam menentukan distribusi frekuensi kurang dari maupun lebih dari atau sama dengan adalah ujung bawah setiap kelas interval termasuk kelas interval yang tidak nampak (kelas interval semu), kelas interval di bawah kelas interval terakhir. Jika tidak dilakukan penambahan satu kelas interval, maka frekuensi kelas interval terakhir tidak akan dihitung padahal frekuensinya ada. Lihat kembali Tabel 2.5, jika kelas interval terakhir ujung bawahnya k, maka frekuensi absolute pada kelas interval terakhir yaitu kelas interval k - l tidak dihitung. Oleh karena itu diperlukan penambahan satu kelas interval semu agar frekuensi terakhir dihitung, yaitu kelas interval m - n dengan ujung bawah m.

Dengan demikian dapat dihitung distribusi frekuensi kumulatif kurang dari data pada bentuk umum distribusi frekuensi data bergolong adalah sebagai berikut.

- Frekuensi kumulatif kurang dari kelas interval pertama, yaitu kurang dari a = 0
- Frekuensi kumulatif kurang dari kelas interval kedua, yaitu kurang dari c = f_1
- Frekuensi kumulatif kurang dari kelas interval ketiga, yaitu kurang dari e = $f_1 + f_2$
- Frekuensi kumulatif kurang dari kelas interval keempat, yaitu kurang dari g = $f_1 + f_2 + f_3$
- Frekuensi kumulatif kurang dari kelas interval kelima, yaitu kurang dari i = $f_1 + f_2 + f_3 + f_4$
- Frekuensi kumulatif kurang dari kelas interval keenam, yaitu kurang dari k = $f_1 + f_2 + f_3 + f_4 + f_5$
- dst.

Jika disajikan dalam tabel, maka distribusi frekuensi kumulatif kurang dari dapat dilihat pada Tabel 2.9 di bawah ini.

Tabel 2.9 Bentuk Tabel Distribusi Frekuensi Kumulatif Kurang Dari

No	Kelas Interval	Frekuensi Kumulatif Kurang Dari
1.	Kurang dari a	0
2.	Kurang dari c	f_1
3.	Kurang dari e	$f_1 + f_2$
4.	Kurang dari g	$f_1 + f_2 + f_3$
5.	Kurang dari i	$f_1 + f_2 + f_3 + f_4$
6.	Kurang dari k	$f_1 + f_2 + f_3 + f_4 + f_5$

Sementara itu, distribusi frekuensi kumulatif lebih dari atau sama dengan data pada bentuk umum distribusi frekuensi data bergolong adalah sebagai berikut.

- Frekuensi kumulatif lebih dari atau sama dengan kelas interval pertama, yaitu lebih dari atau sama dengan a = $f_1 + f_2 + f_3 + f_4 + f_5$
- Frekuensi kumulatif lebih dari atau sama dengan kelas interval kedua, yaitu lebih dari atau sama dengan c = $f_2 + f_3 + f_4 + f_5$
- Frekuensi kumulatif lebih dari atau sama dengan kelas interval ketiga, yaitu lebih dari atau sama dengan e = $f_3 + f_4 + f_5$
- Frekuensi kumulatif lebih dari atau sama dengan kelas interval keempat, yaitu lebih dari atau sama dengan g = $f_4 + f_5$
- Frekuensi kumulatif lebih dari atau sama dengan kelas interval kelima, yaitu lebih dari atau sama dengan i = f_5
- Frekuensi kumulatif lebih dari atau sama dengan kelas interval keenam, yaitu lebih dari atau sama dengan k = 0

Jika disajikan dalam tabel, maka distribusi frekuensi kumulatif lebih atau sama dengan dapat dilihat pada Tabel 2.10 di bawah ini.

Tabel 2.10 Bentuk Tabel Distribusi Frekuensi Kumulatif Lebih dari Dari atau Sama Dengan

No	Kelas Interval	Frekuensi Kumulatif Lebih Dari
1.	Lebih dari atau sama dengan a	$f_1 + f_2 + f_3 + f_4 + f_5$
2.	Lebih dari atau sama dengan c	$f_2 + f_3 + f_4 + f_5$

3.	Lebih dari atau sama dengan e	$f_3 + f_4 + f_5$
4.	Lebih dari atau sama dengan g	$f_4 + f_5$
5.	Lebih dari atau sama dengan i	f_5
6.	Lebih dari atau sama dengan k	0

Jika mau dicari lebih lanjut tentang distribusi frekuensi kumulatif relatif, baik frekuensi kumulatif relatif kurang dari maupun frekuensi kumulatif relatif lebih dari dapat digunakan rumus berikut.

$$\text{frekuensi kumulatif relatif kelas ke} - i = \frac{\text{frekuensi komulatif kelas ke} - i}{\sum_{i=1}^{n} f_i} x100\%$$

Sebenarnya menentukan distribusi frekuensi kumulatif relatif kurang dari maupun distribusi frekuensi kumulatif relatif lebih dari atau sama dengan bisa dengan menjumlahkan satu demi satu atau selangkah demi selangkah distribusi frekuensi relatif masing-masing kelas interval. Dengan demikian dapat ditentukan distribusi frekuensi kumulatif relatif kurang dari sebagai berikut.

- Frekuensi kumulatif relatif kurang dari kelas interval pertama, yaitu kurang dari a = 0
- Frekuensi kumulatif relatif kurang dari kelas interval kedua, yaitu kurang dari c = f_{r1}
- Frekuensi kumulatif relatif kurang dari kelas interval ketiga, yaitu kurang dari e = $f_{r1} + f_{r2}$
- Frekuensi kumulatif relatif kurang dari kelas interval keempat, yaitu kurang dari g = $f_{r1} + f_{r2} + f_{r3}$
- Frekuensi kumulatif relatif kurang dari kelas interval kelima, yaitu kurang dari i = $f_{r1} + f_{r2} + f_{r3} + f_{r4}$

- Frekuensi kumulatif relatif kurang dari kelas interval keenam, yaitu kurang dari k = $f_{r1} + f_{r2} + f_{r3} + f_{r4} + f_{r5}$
- dst.

Jika disajikan dalam tabel, maka distribusi frekuensi kumulatif relatif kurang dapat dilihat pada Tabel 2.11 di bawah ini.

Tabel 2.11 Bentuk Tabel Distribusi Frekuensi Kumulatif Kurang Dari

No	Kelas Interval	Frekuensi Kumulatif Relatif Kurang Dari (%)
1.	Kurang dari a	0
2.	Kurang dari c	f_{r1}
3.	Kurang dari e	$f_{r1} + f_{r2}$
4.	Kurang dari g	$f_{r1} + f_{r2} + f_{r3}$
5.	Kurang dari i	$f_{r1} + f_{r2} + f_{r3} + f_{r4}$
6.	Kurang dari k	$f_{r1} + f_{r2} + f_{r3} + f_{r4} + f_{r5}$

Sementara itu, distribusi frekuensi kumulatif relatif lebih dari atau sama dengan adalah sebagai berikut.

- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval pertama, yaitu lebih dari atau sama dengan a = $f_{r1} + f_{r2} + f_{r3} + f_{r4} + f_{r5}$
- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval kedua, yaitu lebih dari atau sama dengan c = $f_{r2} + f_{r3} + f_{r4} + f_{r5}$
- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval ketiga, yaitu lebih dari atau sama dengan e = $f_{r3} + f_{r4} + f_{r5}$

- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval keempat, yaitu lebih dari atau sama dengan g = $f_{r4} + f_{r5}$
- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval kelima, yaitu lebih dari atau sama dengan i = f_{r5}
- Frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval kelima, yaitu lebih dari atau sama dengan k = 0

Jika disajikan dalam tabel, maka distribusi frekuensi kumulatif relatif lebih dapat dilihat pada Tabel 2.12 di bawah ini.

Tabel 2.12 Bentuk Tabel Distribusi Frekuensi Kumulatif Relatif Lebih dari Dari atau Sama Dengan

No	Kelas Interval	Frekuensi Kumulatif Relatif Lebih Dari atau Sama Dengan (%)
1.	Lebih dari atau sama dengan a	$f_{r1} + f_{r2} + f_{r3} + f_{r4} + f_{r5}$
2.	Lebih dari atau sama dengan c	$f_{r2} + f_{r3} + f_{r4} + f_{r5}$
3.	Lebih dari atau sama dengan e	$f_{r3} + f_{r4} + f_{r5}$
4.	Lebih dari atau sama dengan g	$f_{r4} + f_{r5}$
5.	Lebih dari atau sama dengan i	f_5
6.	Lebih dari atau sama dengan k	0

Ada beberapa catatan yang perlu diperhatikan, jika distribusi frekuensi relatif masing-masing kelas interval merupakan bilangan bulat, maka dalam menentukan distribusi frekuensi kumulatif relatif kurang dari maupun distribusi frekuensi kumulatif relatif lebih dari atau sama dengan bisa menggunakan rumus maupun dengan cara menjumlahkan satu demi satu atau selangkah demi selangkah distribusi frekuensi relatif masing-masing kelas interval. Namun, jika distribusi frekuensi relatif masing-masing kelas interval merupakan bilangan desimal tak terbatas atau misalnya lebih dari lima angka dibelakang koma, sebaiknya dalam menentukan distribusi frekuensi kumulatif relatif kurang dari maupun distribusi frekuensi kumulatif relatif lebih dari atau sama dengan menggunakan rumus yang telah diberikan.

Jika mencari distribusi frekuensi kumulatif relatif kurang dari maupun distribusi frekuensi kumulatif relatif lebih dari atau sama dengan dengan menjumlahkan satu demi satu distribusi frekuensi relatifnya, maka akan terjadi pembulatan berulang-ulang. Oleh karena distribusi frekuensi relatif setiap kelas interval dalam bentuk bilangan desimal biasanya merupakan hasil pembulatan, sehingga kalau itu dilakukan tentunya akan memberikan *error* yang lebih besar dibandingkan dengan menggunakan rumus.

Untuk lebih memahami konsep distribusi frekuensi relatif, distribusi frekuensi kumulatif kurang dari, distribusi frekuensi kumulatif lebih dari atau sama dengan, distribusi frekuensi kumulatif relatif kurang dari maupun distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kita lihat kembali Contoh 2.1 pada bagian sebelumnya.

Contoh 2.2 Berikut ini disajikan data hasil ujian akhir semester (UAS) mata kuliah statistik dasar mahasiswa suatu perguruan tinggi dalam bentuk distribusi frekuensi data bergolong

Tabel 2.13 Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Interval	Frekuensi	Titik Tengah
1	50 - 56	7	53
2	57 - 63	8	60
3	64 - 70	11	67
4	71 - 77	24	74
5	78 - 84	16	81
6	85 - 91	9	88
7	92 - 98	5	95
Total		**80**	**-**

Tentukanlah dan kemudian susun dalam tabel dari:

a) Distribusi frekuensi relatif
b) Distribusi frekuensi kumulatif kurang dari
c) Distribusi frekuensi kumulatif lebih dari atau sama dengan
d) Distribusi frekuensi kumulatif relatif kurang dari
e) Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan

Penyelesaian :

a) Distribusi frekuensi relatif
   Kelas interval pertama (50 - 56)

$$f_{r1} = \frac{f_1}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{7}{80} \times 100\% = 8{,}8\%$$

Kelas interval kedua (57 - 63)

$$f_{r2} = \frac{f_2}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{8}{80} \times 100\% = 10{,}0\%$$

Kelas interval ketiga (64 - 70)

$$f_{r3} = \frac{f_3}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{11}{80} \times 100\% = 13{,}8\%$$

Kelas interval keempat (71 - 77)

$$f_{r4} = \frac{f_4}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{24}{80} \times 100\% = 30{,}0\%$$

Kelas interval kelima (78 - 84)

$$f_{r5} = \frac{f_5}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{16}{80} \times 100\% = 20{,}0\%$$

Kelas interval keenam (85 - 91)

$$f_{r6} = \frac{f_6}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{9}{80} \times 100\% = 11{,}3\%$$

Kelas interval ketujuh (92 - 98)

$$f_{r7} = \frac{f_7}{\sum_{i=1}^{7} f_i} \times 100\% = \frac{5}{80} \times 100\% = 6{,}3\%$$

Jika disubstitusi ke dalam tabel, maka hasilnya dapat dilihat pada Tabel 2.14 di bawah ini.

Tabel 2.14 Distribusi Frekuensi Relatif Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Interval	Frekuensi Relatif (%)
1	50 - 56	8,8
2	57 - 63	10,0
3	64 - 70	13,8
4	71 - 77	30,0
5	78 - 84	20,0
6	85 - 91	11,3
7	92 - 98	6,3
Total		**100**

Sebenarnya nilai total pada kolom distribusi frekuensi relatif adalah 100,2. Nilainya lebih besar 0,002 dari nilai 100 hal ini terjadi karena adanya pembulatan bilangan, oleh karena itu nilai 100,2 tersebut dibulatklan menjadi 100.

b) Distribusi frekuensi kumulatif kurang dari

   i. Distribusi frekuensi kumulatif kurang dari kelas interval pertama, yaitu kurang dari 50. Nilai yang kurang dari 50 tidak ada sehingga frekuensinya 0 (tidak ada).

ii. Distribusi frekuensi kumulatif kurang dari kelas interval kedua, yaitu kurang dari 57. Nilai yang kurang dari 57, dari 50 sampai dengan 56 adalah 7.
iii. Distribusi frekuensi kumulatif kurang dari kelas interval ketiga, yaitu kurang dari 64. Nilai yang kurang dari 64, dari 50 sampai dengan 63 adalah 7 + 8 = 15.
iv. Distribusi frekuensi kumulatif kurang dari kelas interval keempat, yaitu kurang dari 71. Nilai yang kurang dari 71, dari 50 sampai dengan 70 adalah 7 + 8 + 11 = 26.
v. Distribusi frekuensi kumulatif kurang dari kelas interval kelima, yaitu kurang dari 78. Nilai yang kurang dari 78, dari 50 sampai dengan 77 adalah 7 + 8 + 11 + 24= 50.
vi. Distribusi frekuensi kumulatif kurang dari kelas interval keenam, yaitu kurang dari 85. Nilai yang kurang dari 85, dari 50 sampai dengan 84 adalah 7 + 8 + 11 + 24 + 16 = 66.
vii. Distribusi frekuensi kumulatif kurang dari kelas interval ketujuh, yaitu kurang dari 92. Nilai yang kurang dari 92, dari 50 sampai dengan 91 adalah 7 + 8 + 11 + 24 + 16 + 9 = 75.
viii. Distribusi frekuensi kumulatif kurang dari kelas interval kedelapan, yaitu kurang dari 99. Nilai yang kurang dari 99, dari 50 sampai dengan 98 adalah 7 + 8 + 11 + 24 + 16 + 9 + 5 = 80.

Jika disubstitusi ke dalam tabel, maka hasilnya dapat dilihat pada Tabel 2.15 di bawah ini.

Tabel 2.15 Distribusi Frekuensi Kumulatif Kurang Dari Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Nilai Ujian	Frekuensi Kumulatif Kurang Dari
1.	Kurang dari 50	0
2.	Kurang dari 57	7
3.	Kurang dari 64	15
4.	Kurang dari 71	26
5.	Kurang dari 78	50
6.	Kurang dari 85	66
7.	Kurang dari 92	75
8.	Kurang dari 99	80

c) Distribusi frekuensi kumulatif lebih dari atau sama dengan

i. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval pertama, yaitu lebih dari atau sama dengan 50. Nilai yang lebih dari atau sama dengan 50 sampai dengan 99 adalah 7 + 8 + 11 + 24 + 16 + 9 + 5 = 80.

ii. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval kedua, yaitu lebih dari atau sama dengan 57. Nilai yang lebih dari atau sama dengan 57 sampai dengan 99 adalah 8 + 11 + 24 + 16 + 9 + 5 = 73.

iii. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval ketiga, yaitu lebih dari atau sama dengan 64. Nilai yang lebih dari atau sama dengan 64 sampai dengan 99 adalah 11 + 24 + 16 + 9 + 5 = 65.

iv. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval keempat, yaitu lebih dari atau sama dengan 71. Nilai yang lebih dari atau sama dengan 71 sampai dengan 99 adalah 24 + 16 + 9 + 5 = 54.

v. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval kelima, yaitu lebih dari atau sama dengan 78. Nilai yang lebih dari atau sama dengan 78 sampai dengan 99 adalah 16 + 9 + 5 = 30.

vi. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval keenam, yaitu lebih dari atau sama dengan 85. Nilai yang lebih dari atau sama dengan 85 sampai dengan 99 adalah 9 + 5 = 14.

vii. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval ketujuh, yaitu lebih dari atau sama dengan 92. Nilai yang lebih dari atau sama dengan 92 sampai dengan 99 adalah 5.

viii. Distribusi frekuensi kumulatif lebih dari atau sama dengan kelas interval kedelapan, yaitu lebih dari atau sama dengan 99. Nilai yang lebih dari atau sama dengan 99 tidak ada sehingga frekuensinya 0 (tidak ada).

Jika disubtitusi ke dalam tabel, maka hasilnya dapat dilihat pada Tabel 2.16 di bawah ini.

Tabel 2.16 Distribusi Frekuensi Kumulatif Lebih Dari atau Sama Dengan Hasil Ujian Akhir Semester (Uas) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Nilai Ujian	Frekuensi Kumulatif Lebih Dari atau Sama Dengan
1.	Lebih dari atau sama dengan 50	80
2.	Lebih dari atau sama dengan 57	73
3.	Lebih dari atau sama dengan 64	65
4.	Lebih dari atau sama dengan 71	54
5.	Lebih dari atau sama dengan 78	30
6.	Lebih dari atau sama dengan 85	14
7.	Lebih dari atau sama dengan 92	5
8.	Lebih dari atau sama dengan 99	0

c) Distribusi frekuensi kumulatif relatif kurang dari

i. Distribusi frekuensi kumulatif relatif kurang dari kelas interval pertama adalah

$$fkr_1 = \frac{fr_1}{\sum_{i=1}^{7} f_i} x100\% = \frac{0}{80} x100\% = 0\%$$

ii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval kedua adalah

$$fkr_2 = \frac{fr_2}{\sum_{i=1}^{7} f_i} x100\% = \frac{7}{80} x100\% = 8{,}8\%$$

iii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval ketiga adalah

$$fkr_3 = \frac{fr_3}{\sum_{i=1}^{7} f_i} x100\% = \frac{15}{80} x100\% = 18{,}8\%$$

iv. Distribusi frekuensi kumulatif relatif kurang dari kelas interval keempat adalah

$$fkr_4 = \frac{fr_4}{\sum_{i=1}^{7} f_i} x100\% = \frac{26}{80} x100\% = 32{,}5\%$$

v. Distribusi frekuensi kumulatif relatif kurang dari kelas interval kelima adalah

$$fkr_5 = \frac{fr_5}{\sum_{i=1}^{7} f_i} x100\% = \frac{50}{80} x100\% = 62{,}5\%$$

vi. Distribusi frekuensi kumulatif relatif kurang dari kelas interval keenam adalah

$$fkr_6 = \frac{fr_6}{\sum_{i=1}^{7} f_i} x100\% = \frac{66}{80} x100\% = 82{,}5\%$$

vii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval ketujuh adalah

$$fkr_7 = \frac{fr_7}{\sum_{i=1}^{7} f_i} x100\% = \frac{75}{80} x100\% = 93{,}8\%$$

viii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval kedelapan adalah

$$fkr_8 = \frac{fr_8}{\sum_{i=1}^{7} f_i} x100\% = \frac{80}{80} x100\% = 100\%$$

Jika disubstitusi ke dalam tabel, maka hasilnya dapat dilihat pada Tabel 2.17 di bawah ini.

Tabel 2.17 Distribusi Frekuensi Kumulatif Relatif Kurang Dari Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Nilai Ujian	Frekuensi Kumulatif Relatif Kurang Dari (%)
1.	Kurang dari 50	0
2.	Kurang dari 57	8,8
3.	Kurang dari 64	18,8
4.	Kurang dari 71	32,5
5.	Kurang dari 78	62,5
6.	Kurang dari 85	82,5
7.	Kurang dari 92	93,8
8.	Kurang dari 99	100

e) Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan

i. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval pertama adalah

$$fkr_1 = \frac{fr_1}{\sum_{i=1}^{7} f_i} x100\% = \frac{80}{80} x100\% = 100\%$$

ii. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval kedua adalah

$$fkr_2 = \frac{fr_2}{\sum_{i=1}^{7} f_i} x100\% = \frac{73}{80} x100\% = 91,3\%$$

iii. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval ketiga adalah

$$fkr_3 = \frac{fr_3}{\sum_{i=1}^{7} f_i} x100\% = \frac{65}{80} x100\% = 81,3\%$$

iv. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval keempat adalah

$$fkr_4 = \frac{fr_4}{\sum_{i=1}^{7} f_i} x100\% = \frac{54}{80} x100\% = 67,5\%$$

v. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval kelima adalah

$$fkr_5 = \frac{fr_5}{\sum_{i=1}^{7} f_i} x100\% = \frac{30}{80} x100\% = 37,5\%$$

vi. Distribusi frekuensi kumulatif relatif lebih dari atau sama dengan kelas interval keenam adalah

$$fkr_6 = \frac{fr_6}{\sum_{i=1}^{7} f_i} x100\% = \frac{14}{80} x100\% = 17,5\%$$

vii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval ketujuh adalah

$$fkr_7 = \frac{fr_7}{\sum_{i=1}^{7} f_i} x100\% = \frac{5}{80} x100\% = 6,3\%$$

viii. Distribusi frekuensi kumulatif relatif kurang dari kelas interval kedelapan adalah

$$fkr_8 = \frac{fr_8}{\sum_{i=1}^{7} f_i} x100\% = \frac{0}{80} x100\% = 0\%$$

Jika disubstitusi ke dalam tabel, maka hasilnya dapat dilihat pada Tabel 2.18 di bawah ini.

Tabel 2.18 Distribusi Frekuensi Kumulatif Relatif Lebih Dari atau Sama Dengan Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Nilai Ujian	Frekuensi Kumulatif Relatif Lebih Dari atau Sama Dengan (%)
1.	Lebih dari atau sama dengan 50	100
2.	Lebih dari atau sama dengan 57	91,3
3.	Lebih dari atau sama dengan 64	81,3
4.	Lebih dari atau sama dengan 71	67,5
5.	Lebih dari atau sama dengan 78	37,5
6.	Lebih dari atau sama dengan 85	17,5
7.	Lebih dari atau sama dengan 92	6,3
8.	Lebih dari atau sama dengan 99	0

B. Macam-macam Grafik/Diagram

Data hasil penelitian yang disajikan dalam bentuk tabel kadang-kadang sulit untuk dipahami. Oleh karena itu, selain disajikan dalam bentuk tabel data juga bisa disajikan dalam bentuk grafik. Grafik sering juga disebut sebagai diagram, bagan, atau *mauoun chart*. Grafik adalah gambar yang menyajikan data secara visual yang biasanya berasal dari tabel yang telah dibuat. Grafik adalah alat penyajian statistik yang tertuang dalam bentuk lukisan, baik lukisan garis, lukisan gambar, maupun lambang. Pada dasarnya grafik berfungsi memberikan penjelasan kepada para pembaca grafik atau orang yang membutuhkan data. Grafik itu sendiri bisa memudahkan pembaca untuk mengetahui dan membaca data tanpa menggunakan kata-kata yang bertele-tele karena grafik menyajikan data dalam bentuk angka, dalam sebuah lembar kerja, dan dalam bentuk visualisasi grafik. Meskipun demikian, diagram masih memiliki kelemahan, yaitu pada umumnya diagram tidak dapat memberikan gambaran yang lebih detail.

Menyajikan data dengan grafik/diagram dapat dibagi menjadi dua jenis, yaitu grafik data tunggal dan grafik data bergolong. Data tidak bergolong umumnya merupakan data yang jumlahnya sangat kecil sehingga sulit untuk dikelompokan. Umunya data ini dalam bentuk diskrit atau terpisah-pisah sehingga ada tidak ada titik penghubung antara data yang satu dengan data yang lainnya (*disjoint*). Grafik data tunggal meliputi, grafik batang, grafik daun, grafik garis, grafik lingkaran, grafik lambang dan grafik pencar. Sedangkan yang termasuk grafik data bergolong adalah grafik histogram poligon frekuensi serta ogive.

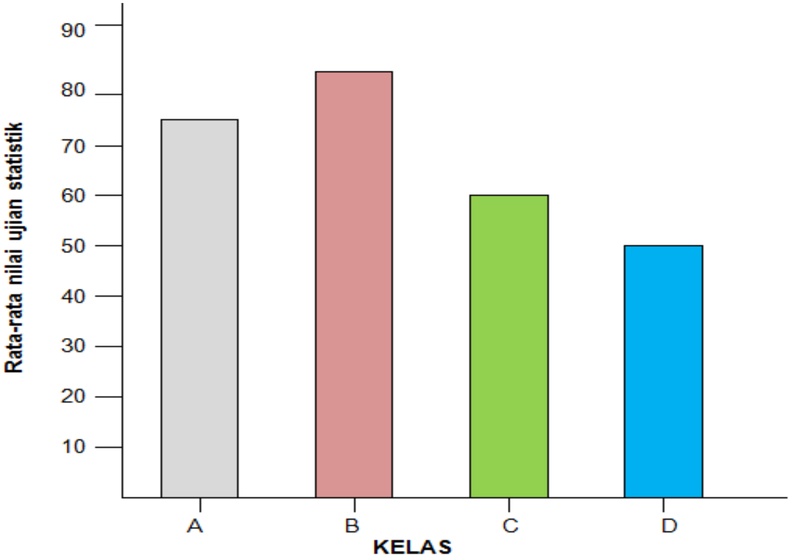
1) Grafik Data Tidak Berkelompok

a) Grafik batang

Grafik batang atau balok sering disebut *bar chart*. Grafik atau diagram batang adalah diagram berdasarkan data berbentuk kategori. Sebuah batang menggambarkan jumlah tertentu dari data. Diagram ini banyak digunakan untuk membandingkan data maupun menunjukan hubungan suatu data dengan data keseluruhan. Grafik batang setiap kategori tertentu bisa berupa batang tunggal (*single bar charts*) yang menggambarkan satu hal atau masalah atau berupa batang-batang ganda atau lebih (*multiple bar charts*) yang lebih dari satu hal atau masalah. Grafik batang biasanya berbentuk persegi panjang tetapi bisa juga berbentuk balok. Pembuatannya tergantung pada si pembuat grafik.

Sebagai contoh hasil ujian akhir semester pada mata kuliah statistik dari empat kelas pada suatu universitas diperoleh data sebagai berikut. Kelas A memiliki rata-rata nilai ujian 70,5; kelas B memiliki rata-rata nilai ujian 80,9; kelas C memiliki rata-rata nilai ujian 60,0; dan kelas D memiliki rata-rata nilai ujian 55,0. Nilai rata-rata tersebut akan sedikit sulit untuk dibandingkan apabila disajikan dalam bentuk kalimat seperti sebelumnya. Terdapat cara menyajikan data yang dapat memudahkan kita untuk memudahkan dalam membandingkan masing-masing kategori, yaitu dengan diagram batang.

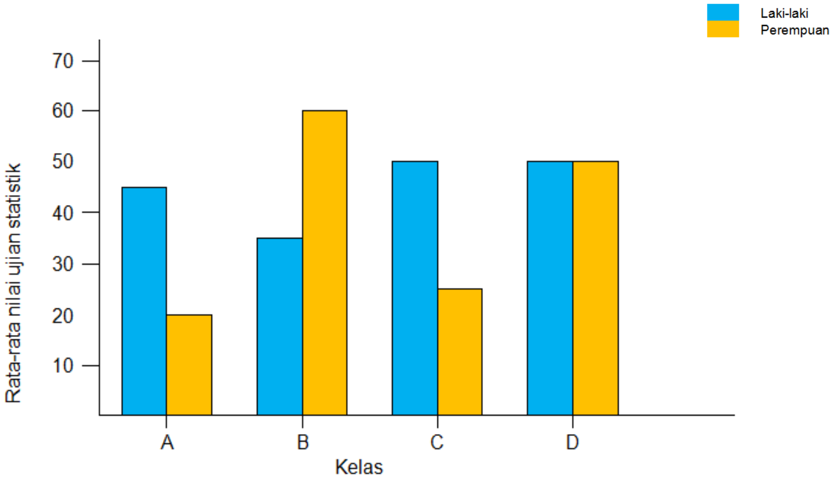
Ada beberapa langkah dilakukan untuk menyusun grafik batang, yaitu: a) buat sumbu datar dan sumbu tegak berpotongan tegak lurus. Sumbu datar merupakan sumbu x biasanya menyatakan keterangan dari data yang dibuat grafiknya, sumbu y biasanya menyatakan jumlah atau banyaknya data yang dibuat atau frekuensinya, b) bagilah sumbu datar dan tegak menjadi beberapa bagian dengan skala yang sama, dengan perbandingan skala antara sumbu tegak tidak harus sama, c) buatlah batang untuk masing-masing kategori dalam bentuk persegi panjang. Perlu diperhatikan lebar masing-masing persegi harus sama, sedangkan tingginya disesuaikan frekuensi data. Jarak antara batang yang satu dengan yang lainnya juga harus sama. Dengan mengambil contoh nilai rata-rata ujian akhir semester di atas, grafik batangnya dapat digambarkan sebagai berikut.



Gambar 1 Grafik Batang Nilai Ujian Akhir Semester Mata Kuliah Statistik (Batang Tunggal)

Dengan melihat Gambar 1 tentunya akan lebih mudah seorang peneliti atau pembaca dalam memahami perbandingan maupun hubungan antara keempat rata-rata nilai ujian akhir semester mata kuliah statistik tersebut. Grafik batang tersebut merupakan grafik batang tunggal (*single bar charts*). Jika rata-

rata ujian akhir tersebut mempertimbangakan jenis kelamin, gambar grafiknya menjadi seperti pada Gambar 2 berikut ini.



Gambar 2 Grafik Batang Nilai Ujian Akhir Semester Mata Kuliah Statistik (Batang Ganda)

b) Grafik daun

Bagian kedua yang tergolong grafik data tidak bergolong adalah grafik daun. Grafik daun sering disebut grafik batang daun(*steam and leaf diagram*). Grafik ini bisa menggambarkan kecenderungan data. Apakah data tersebut menyebar ataukah memusat pada suatu nilai tertentu. Grafik ini dengan mudah melihat nilai-nilai yang sering muncul atau yang jarang muncul.

Grafik batang daun terdiri dari bagian, yaitu batang dan daun yang dipilah menjadi dua bagian. Perlu diperhatikan batang yang dimaksud di sini adalah penempatan angka puluhan atau ratusan bukan batang berbentuk persegi panjang seperti halnya pada grafik batang. Sedangkan daunnya merupakan penempatan angka satuan, tidak sama sekali seperti daun pada pohon. Sehingga dengan demikian dapat ditentukan bahwa angka pertama ditempatkan pada bagian diagram yang disebut batang, dan angka kedua dan seterusnya (kalau ada) ditempatkan pada bagian yang disebut daun.

Jadi dalam membuat grafik batang daun perlu diperhatikan hal-hal sebagai berikut.

(a) Dalam diagram batang daun, data yang terkumpul diurutkan lebih dulu dari data nilai data terkecil sampai dengan nilai data yang terbesar atau sebaliknya.

(b) Bagian pertama dalam grafik batang daun adalah batang, sedangkan bagian kedua adalah daun. Jika data dalam satuan, maka batangnya adalah 0 dan daunnya adalah angka satuan yang dimaksud. Jika datanya puluhan, maka batangnya adalah puluhan dan daunnya adalah satuan. Jika datanya ratusan, maka batangnya bisa ratusannya dan daunnya adalah puluhan dan satuan atau bisa juga batangnya ratusan dan puluhan sedangkan daunnya satuan begitu seterusnya.

Jadi, suatu data yang merupakan suatu bilangan, misalnya 8, akan dipisahkan sebagai 0 dan 8,(0 pada batang dan 8 pada daun), 45 akan dipisahkan sebagai 4 dan 5,(4 pada batang dan 5 pada daun), sedangkan 123 akan dipisahkan sebagai 1 dan 23 ,(1 pada batang dan 23 pada daun), atau 12 dan 3,(12 pada batang dan 3 pada daun). Untuk lebih memahami diagram batang daun, perhatikan contoh berikut.

Contoh 2.3 Berikut ini disajikan data hasil ujian akhir semester (UAS) mata kuliah statistik dasar mahasiswa suatu perguruan tinggi

50	68	73	70	96	80	65	97
86	84	79	65	78	79	73	80
67	75	88	75	82	89	67	73
73	82	87	82	73	87	57	72
57	81	68	71	74	94	75	78
88	72	90	93	62	77	95	80
80	63	55	55	54	60	70	76
68	78	60	72	82	55	54	71
90	74	56	76	74	63	80	88
75	70	63	61	66	66	57	75

Susunlah data di atas ke dalam grafik batang daun!
Langkah-langkah penyelesaiannya
(1) Urutkan data dari nilai terkecil ke nilai terbesar

50	54	54	55	55	55	56	57	57	57
60	60	61	62	63	63	63	65	65	66
66	67	67	68	68	68	70	70	70	67
71	71	72	72	73	73	73	73	73	74
74	74	75	75	75	75	75	76	76	77
78	78	78	79	79	80	80	80	80	80
81	82	82	82	82	84	86	87	87	88
88	88	89	90	90	93	94	95	96	97

(2) Buatlah tabel grafik batang daun

Batang	Daun
5	0 4 4 5 5 5 6 7 7 7
6	0 0 1 2 3 3 3 5 5 6 6 7 7 7 8 8 8
7	0 0 0 1 1 2 2 3 3 3 3 3 4 4 4 5 5 5 5 5 6 6 7 8 8 8 9 9
8	0 0 0 0 0 1 2 2 2 2 4 6 7 7 8 8 8 9
9	0 0 3 4 5 6 7

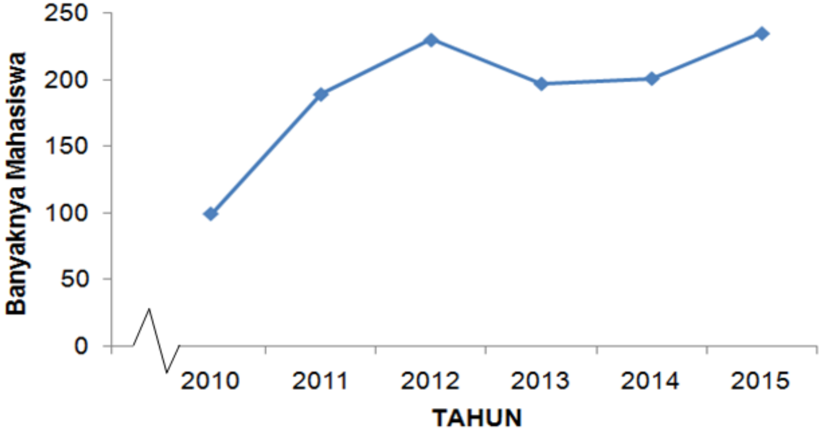
c) Grafik garis

Grafik garis (*line chart*) digunakan untuk menggambarkan keadaan yang serba terus atau berkesinambungan, misalnya banyaknya mahasiswa jurusan matematika setiap tahun pada suatu universitas, banyaknya pertumbuhan penduduk tiap tahun, banyaknya konsumsi premium kendaraan bermotor setiap tahunnya dan lain sebagainya. Seperti diagram batang, pada grafik garis kategori data disajikan pada sumbu horizontal dan nilai data berada di sumbu vertikal. Sehingga diperlukan sistem sumbu datar/sumbu x dan sumbu tegak/sumbu y yang saling tegak lurus. Berikut ini Tabel 2.19 disajikan data banyaknya mahasiswa jurusan pendidikan matematika pada suatu universitas 6 tahun terakhir.

Tabel 2.19 Banyaknya Mahasiswa Jurusan Pendidikan Matematika Pada Suatu Universitas

No	Tahun	Banyaknya Mahasiswa
1	2010	99
2	2011	189
3	2012	230
4	2013	197
5	2014	201
6	2015	235
Total		**1151**

Tabel 2.19 di atas disajikan dalam bentuk grafik garis pada Gambar 3 berikut ini.



Gambar 3 Grafik Garis Banyaknya Mahasiswa Jurusan Pendidikan Matematika Pada Suatu Universitas

d) Grafik lingkaran

Nama lain grafik lingkaran adalah pie atau pastel. Grafik lingkaran digunakan untuk mengetahui perbandingan suatu data dengan keseluruhan. Grafik lingkaran adalah gambaran grafik informasi kuantitatif menggunakan lingkaran yang dibagi menjadi beberapa sektor yang ukuran relatifnya sesuai dengan proporsi kuantitas. Sektor-sektor ini dalam matematika disebut dengan juring, yaitu daerah lingkaran yang dibatasi oleh dua buah jari-jari. Pada dasarnya, diagram ini menampilkan hubungan persentase antar bagian dibandingkan dengan keseluruhan. Besarnya persentase masing-masing sektor dihitung dengan menggunakan besarnya sudut pusat lingkaran. Berdasarkan hal tersebut kadang yang dicari adalah proporsi sudut pusat lingkaran untuk masing-masing kategori. Dengan demikian dalam grafik lingkaran yang dipakai adalah distribusi frekuensi relatif. Grafik lingkaran tidak menampilkan informasi frekuensi absolute masing-masing secara detail. Grafik lingkaran hanya menampilkan besar sudut dalam derajat atau persentase yang merupakan frekuensi relatif masing-masing data, sehingga yang nampak adalah perbandingan frekuensi data secara visual.

Untuk mencari besarnya sudut pusat untuk masing-masing kategori apabila yang diketahui frekuensi absolutnya adalah:

$$\text{sudut pusat} = \frac{f_i}{N} \text{ x } 360^0$$

Jika yang diketahui frekuensi relatif, maka untuk mencari besarnya sudut pusat masing-masing kategori adalah:

$$\text{sudut pusat} = \frac{fr_i}{100\%} \text{ x } 360^0$$

Untuk lebih memahami perhatikan contoh berikut ini.

Contoh 2.4 **Dari 720 orang mahasiswa pada sebuah universitas ternyata 120 orang mengikuti UKM Bola Voli, 135 mengikuti UKM Bulu Tangkis, 210 mengikuti UKM Tari dan sisanya mengikuti UKM Sepak Bola. Gambarkalah UKM pilihan mahasiswa tersebut dalam grafik lingkaran menurut persentasenya!**

Penyelesaian

Terlebih dahulu tentukan besarnya persentase atau frekuensi relatif masing-masing kategori kemudian tentukan sudut pusat masing-masing sektor atau UKM yang diketahui.

UKM Bola Voli $= \frac{120}{720} \times 100\% = 17\%$

UKM Bulu Tangkis $= \frac{135}{720} \times 100\% = 19\%$

UKM Tari $= \frac{210}{720} \times 100\% = 29\%$

UKM Sepak Bola $= \frac{255}{720} \times 100\% = 35\%$

Sementara itu besarnya sudut pusat untuk masing-masing sektor adalah:

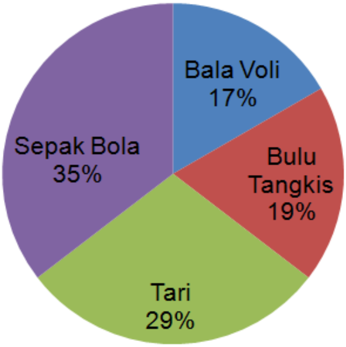
UKM Bola Voli $= \frac{17\%}{100\%} \times 360^0 = 61{,}2^0$

UKM Bulu Tangkis $= \frac{19\%}{100\%} \times 360^0 = 68{,}4^0$

UKM Tari $= \frac{29\%}{100\%} \times 360^0 = 104{,}4^0$

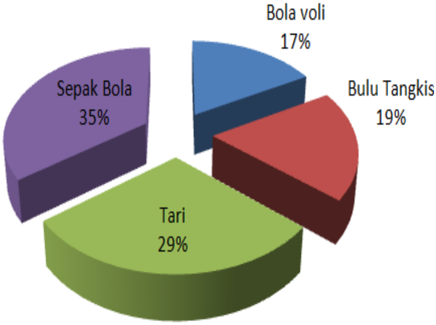
UKM Sepak Bola $= \frac{33\%}{100\%} \times 360^0 = 118{,}8^0$

Grafik lingkaran data di atas dapat dilihat pada Gambar 4 di bawah ini.



Gambar 4 Grafik Lingkaran Unit Kegiatan Mahasiswa (UKM) Pada Sebuah Universitas

Apabila Gambar 4 di atas kita gambarkan dalam lingkaran yang ada tebalnya (seperti donat), maka sering disebut dengan pastel. Di bawah ini grafik pastel Contoh 2.4.



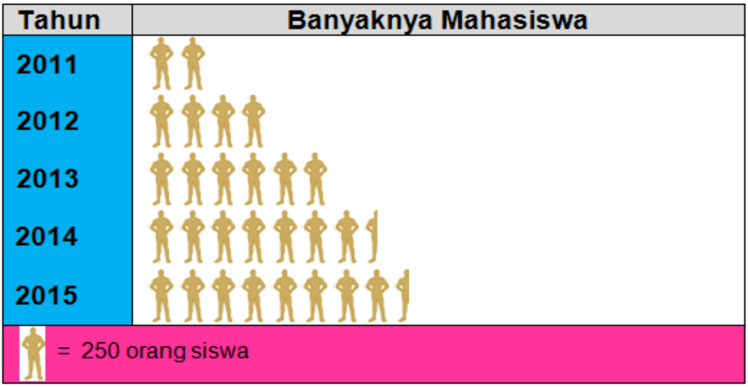
Gambar 5 Grafik Pastel Unit Kegiatan Mahasiswa (UKM) Pada Sebuah Universitas

e) Grafik lambang

Penyajian data statistik dengan menggunakan lukisan atau gambar disebut diagram lambang atau diagram gambar (*pictogram*). Dalam piktogram simbol-simbol yang digunakan sebaiknya disesuaikan dengan data atau objek-objek yang digambarkan. Misalnya data yang digambarkan banyaknya mahasiswa digunakan gambar orang, data untuk panen buah digambarkan dengan buah, data penjualan mobil digambarkan dengan mobil, dan sebagainya. Meskipun penyajian data dengan piktogram itu sederhana, akan tetapi pemakaiannya sangat terbatas. Biasanya piktogram dipakai untuk menyajikan data yang nilainya cukup besar dengan nilai-nilai data yang telah dibulatkan.

Grafik ini sering dipakai untuk mendapatkan gambaran kasar suatu hal dan sebagai alat visual bagi orang awam. Setiap satuan jumlah tertentu dibuat sebuah simbol sesuai dengan macam datanya. Kelemahan grafik ini adalah kurang efisien tempat apalagi untuk data yang cukup besar, kesulitan menggambarkan bagian simbol untuk suatu hal yang merupakan bagian tertentu dari suatu yang utuh. Misalnya menggambarkan setengah, seperempat, seperdelapan, dan yang lainnya. Gambar-gambar atau lambang-lambang yang digunakan dibuat kongruen (sama), sehingga lebih jelas dan mampu mewakili jumlah tertentu untuk satu gambar dan lambang tersebut.

Contoh 2.5 **Pada suatu universitas diketahui data banyaknya mahasiswa dari tahun 2010 sampai dengan tahun 2015. Pada tahun 2011 ada sebanyak 500 orang siswa, tahun 2012 ada 1.000 orang siswa, tahun 2013 ada 1.500 orang siswa, tahun 2014 ada 1.750 orang siswa, dan pada tahun 2015 ada 2.250 orang siswa. Sajikanlah data banyak siswa dari tahun 2010 sampai dengan tahun 2015 dengan piktogram**

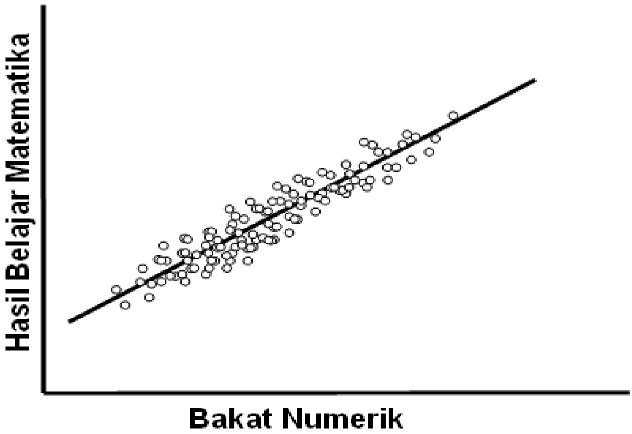


Gambar 6 Grafik Lambang Universitas Diketahui Data Banyaknya Mahasiswa Dari Tahun 2010 Sampai Dengan Tahun 2015 Pada Suatu Universitas

Pada dasarnya, penyajian data dalam bentuk grafik lambang memang menarik. Akan tetapi, penggunaan piktogram sangatlah terbatas. Misalnya pada Contoh 2.5 sangat sulit menggambarkan jika banyaknya mahasiswa Tahun 2014 misalnnya sebanyak 2.345 orang atau banyaknya mahasiswa Tahun 2015 sebanyak 2.555 orang dan sebaginya.

f) Grafik pencar

Grafik atau diagram pencar biasanya digunakan untuk mengetahui hubungan logis antar dua variabel dan menunjukkan keeratan hubungan antara dua variabel tersebut. Nama lain dari grafik ini adalah grafik tebaran atau *scatter*. Diagram pencar juga dapat digunakan untuk mengecek apakah suatu variabel dapat digunakan untuk mengganti variabel yang lain. *Scatter* sering digunakan sebagai analisis tindak lanjut untuk menentukan apakah penyebab yang ada benar-benar memberikan dampak kepada karakteristik kualitas.



Gambar 6 Grafik Pencar Hubungan antara Bakat Numerik dengan Hasil Belajar Matematika

Pada contoh terlihat *scatter* diagram yang menggambarkan plot bakat numerik dengan hasil belajar matematika yang mengindikasikan hubungan kuat positif diantara dua variabel. Jika bakat numerik mahasiswa meningkat atau semakin tinggi, maka hasil belajar matematika mahasiswa cenderung meningkat. Begitu juga sebaliknya, jika bakat numerik mahasiswa menurun, maka hasil belajar matematikanya cenderung menurun.

2) Grafik Data Bergolong

a) Poligon dan histogram frekuensi

Histogram adalah penyajian tabel distribusi frekuensi dengan menggunakan gambar berbentuk persegi panjang yang saling berhimpit. Himpitan antara batang yang satu dengan batang yang lain ditandai oleh batas kelas masing-masing kelas interval. Histogram merupakan grafik berbentuk batang, karena frekuensinya disajikan dalam bentuk balok yang digunakan untuk menggambarkan bentuk distribusi frekuensi data berkelompok. Pada sumbu horizontal atau absis (X) merupakan tepi kelas interval dan frekuensi setiap kelas pada sumbu vertikal atau ordinat (Y) merupakan frekuensi setiap kelas. Langkah-membuat histogram adalah sebagai berikut.

(1) Buat salib sumbu koordinat pada diagram kartesius. Ordinat atau sumbu Y merupakan frekuensi masing-masing kelas interval, sedangkan absis atau sumbu X merupakan batas kelas masing masing kelas interval. Sebaiknya untuk membuat histogram jangan menggunakan titik tengah, karena kalau menggunakan titik tengah kelihatan seperti grafik batang pada data tunggal. Titik tengah nantinya digunakan dalam menggambar poligon frekuensi.

(2) Apabila skala atau jarak antar titik pada sumbu Y tidak sama dengan skala pada sumbu X, maka pada sumbu X diisi tanda dimampatkan. Setelah diisi tanda dimampatkan baru menulis batas bawah kelas interval pertama (titik pertama) sampai dengan batas atas kelas interval pertama untuk titik berikutnya (titik kedua), titik ketiga merupakan batas atas kelas interval kedua, karena batas bawah kelas interval kedua merupakan batas atas kelas interval pertama, jadi cukup ditulis sekali. Titik keempat merupakan batas atas kelas interval ketiga atau batas bawah kelas interval keempat begitu seterusnya sampai dengan batas atas kelas interval terakhir.

(3) Setelah menulis skala pada sumbu Y maupun sumbu X, mulailah membuat ruas garis tegak sejajar dengan sumbu Y dengan titik pangkal batas bawah kelas interval pertama diperpanjang sesuai dengan frekuensi kelas interval pertama. Garis yang sama panjangnya juga dibuat pada titik batas atas kelas interval pertama. Setelah kedua

garis tersebut terbentuk kemudian hubungkan masing-masing titik ujungnya dengan sebuah garis lurus sehingga membentuk balok atau persegi panjang.

(4) Hal yang sama seperti langkah (3) dilakukan pada kelas interval kedua. Buat dua buah garis lurus yang masing-masing berpangkal pada batas bawah dan batas atas kelas interval kedua. Panjangnya dibuat sesuai dengan frekuensi kelas interval kedua, jika frekuensi kelas interval pertama lebih kecil dari frekuensi kelas interval kedua, maka garis lurus yang dibuat pada batas bawah kelas interval kedua tinggal memperpanjang garis lurus pada batas atas kelas interval pertama. Tetapi sebaliknya, jika frekuensi kelas interval pertama lebih tinggi atau sama dengan kelas interval kedua, maka tidak usah lagi membuat garis pada batas bawah kelas interval kedua tinggal menghubungkan kedua ujungnya dengan garis lurus. Hal ini dilakukan untuk kelas interval ketiga, keempat, kelima dan seterusnya.

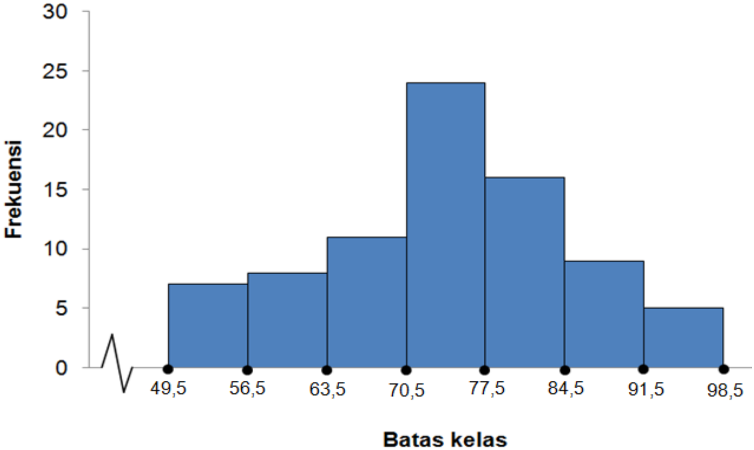
Perhatikan tabel distribusi frekuensi data berkelompok di bawah ini.

Tabel 2.20 Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Interval	Frekuensi	Titik Tengah
1	50 - 56	7	53
2	57 - 63	8	60
3	64 - 70	11	67
4	71 - 77	24	74

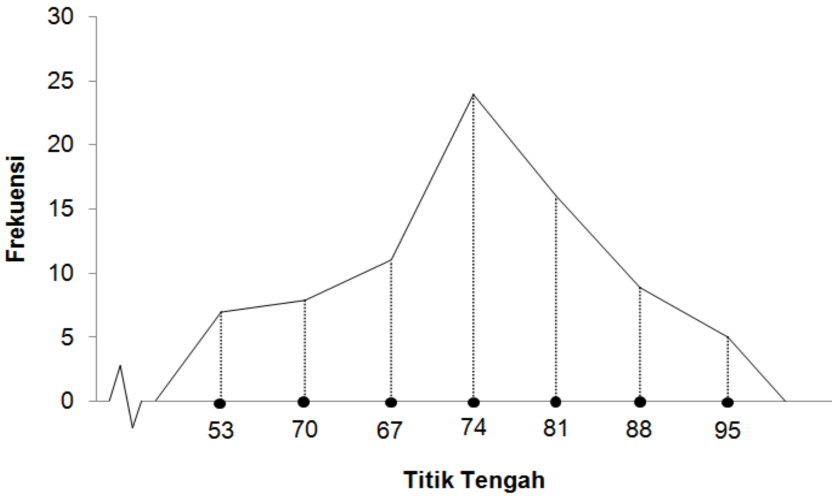
No	Interval	Frekuensi	Titik Tengah
5	78 - 84	16	81
6	85 - 91	9	88
7	92 - 98	5	95
Total		**80**	**-**

Selanjutnya kita sajikan Tabel 2.20 di atas ke dalam grafik histogram.



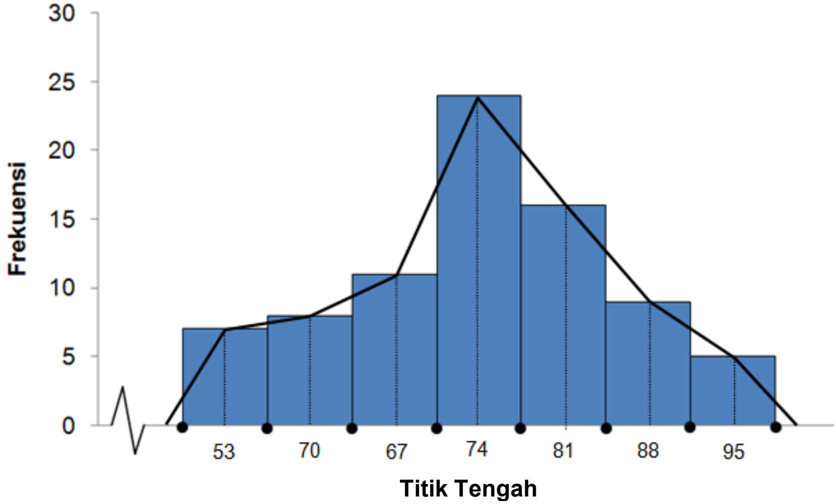
Gambar 7 Histogram Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

Jika grafik histogram pada Gambar 7 ditambah titik tengah pada sumbu X atau sumbu mendatar kemudian setiap titik tengah masing-masing kelas interval dihubungkan dengan garis, maka grafik tersebut akan menjadi grafik histogram dan poligon frekuensi. Untuk membuat grafik poligon, sebenarnya tidak ada perbedaan penting antara grafik histogram dengan grafik poligon. Grafik histogram biasanya dibuat dengan mengunakan batas atas dan batas bawah, sedangkan grafik poligon selalu menggunakan titik tengah. Grafik histogram berwujud segiempat-segiempat atau berupa batang, sedangkan grafik poligon berwujud garis-garis atau kurva. Grafik poligon selanjutnya disebut dengan grafik poligon frekuensi, dibuat dengan menghubung-hubungkan titik-titik koordinat (pertemuan titik tengah dengan frekuensi tiap kelas) secara berturut-turut. Berikut ini disajikan grafik poligon frekuensi berdasarkan grafik histogram Gambar 7.



Gambar 8 Poligon Frekuensi Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

Histogram dan poligon frekuensi bisa juga digambarkan dalam satu grafik seperti pada Gambar 9 di bawah ini.



Gambar 9 Histogram dan Poligon Frekuensi Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

b) Ogive

Grafik data bergolong yang dibentuk berdasarkan frekuensi kumulatif baik frekuensi kumulatif kurang dari maupun frekuensi kumulatif lebih dari atau sama dengan disebut dengan ogive (ozaiv). Dengan demikian grafik distribusi frekuensi kumulatif disebut ogive atau sering disebut grafik frekuensi meningkat. Ogive pada prinsipnya sama dengan grafik garis, namun untuk ogive selalu mulai dari titik 0. Pada sumbu Y merupakan frekuensi kumulatif masing-masing kelas interval sedangkan sumbu X merupakan ujung kelas masing-masing interval bukan titik tengah seperti pada poligon frekuensi. Ogive dibuat dengan cara menempatkan titik-titik limit ujung bawah kelas interval pada sumbu horizontal dan pada sumbu vertikal ditempatkan frekuensi kumulatif. Kemudian titik-titik tersebut dihubungkan sehingga kita mendapatkan kurva yang mulus yang terus meningkat.

Grafik Ogive digunakan, apabila ingin mengetahui posisi seseorang tentang sesuatu hal dalam kelompoknya sendiri, bukan pola sifat atau kecakapan kelompok seluruhnya. Oleh karena itu, banyak ditemui hasil-hasil tes bakat, tes kemampuan khusus, dan semacamnya yang dilaporkan dalam bentuk Ogive atau grafik frekuensi meningkat. Hal ini disebabkan karena nilai-nilai test semacam itu kerapkali digunakan untuk mengadakan penilaian tentang kecakapan perorangan.

Terdapat dua jenis ogive, yaitu ogive positif dan ogive negatif. Ogive positif merupakan grafik frekuensi kumulatif kurang dari yang grafiknya dimulai dari nol (karena frekuensi sebelum kelas interval pertama adalah nol) dan berakhir pada frekuensi total data. Disebut dengan ogive positif karena kemiringan atau gradiennya positif yaitu miring ke kanan. Sedangkan ogive negatif merupakan grafik frekuensi kumulatif lebih dari atau sama dengan yang grafiknya dimulai dari jumlah frekuensi total dan berakhir titik nol (sumbu Y sama dengan nol). Disebut dengan ogive negatif karena kemiringan atau gradiennya negatif yaitu miring ke kiri.

Contoh 2.6 **Lihatlah kembali Tabel 2.21 tentang frekuensi kumulatif kurang dari dan Tabel 2.22 tentang frekuensi kumulatif lebih dari atau sama dengan di bawah ini. Dari kedua tabel tersebut gambarlah grafik ogive positif dan ogive negatif dalam satu grafik!**

Tabel 2.21 Distribusi Frekuensi Kumulatif Kurang Dari Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

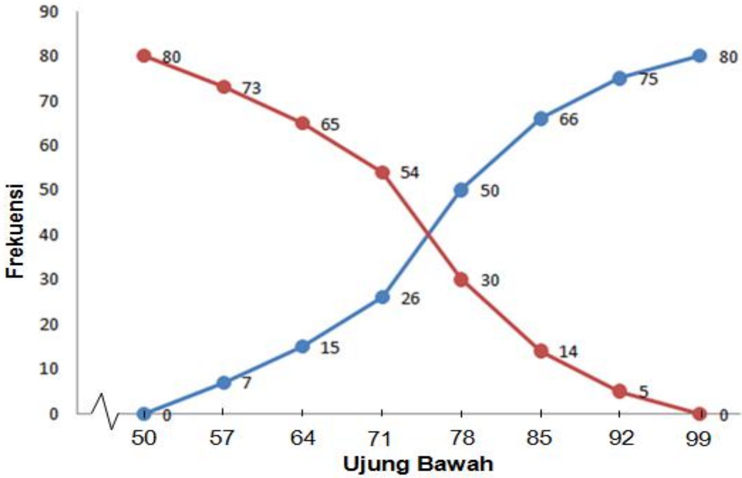
No	Nilai Ujian	Frekuensi Kumulatif Kurang Dari
1.	Kurang dari 50	0
2.	Kurang dari 57	7
3.	Kurang dari 64	15
4.	Kurang dari 71	26
5.	Kurang dari 78	50
6.	Kurang dari 85	66
7.	Kurang dari 92	75
8.	Kurang dari 99	80

Tabel 2.22 Distribusi Frekuensi Kumulatif Lebih Dari atau Sama Dengan Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

No	Nilai Ujian	Frekuensi Kumulatif Lebih Dari atau Sama Dengan
1.	Lebih dari atau sama dengan 50	80
2.	Lebih dari atau sama dengan 57	73
3.	Lebih dari atau sama dengan 64	65
4.	Lebih dari atau sama dengan 71	54
5.	Lebih dari atau sama dengan 78	30
6.	Lebih dari atau sama dengan 85	14
7.	Lebih dari atau sama dengan 92	5
8.	Lebih dari atau sama dengan 99	0

Penyelesaian

Tulislah nilai-nilai berikut pada sumbu datarnya 50, 57, 64, 71, 78, 85, 92, 99. Sumbu tegaknya bisa diisi dengan 10, 20, 30, dan seterusnya dalam skala 10. Pada sumbu datar jarak antara 0 sampai dengan 50 bisa diloncat, tetapi sumbunya harus diisi tanda dimampatkan.



Gambar 10 Ogive Hasil Ujian Akhir Semester (UAS) Mata Kuliah Statistik Dasar Mahasiswa Suatu Perguruan Tinggi

Latihan 2

1. Berikut ini diberikan data jenis pekerjaan orang tua siswa SD, SMP, dan SMA pada empat kabupaten, yaitu Kabupaten Tabanan, Kabupaten Gianyar, Kabupaten Badung, dan Kabupaten Bangli. Di Kabupaten Tabanan untuk tingkat SD pekerjaan orang tua siswa adalah TNI/

Polri/PNS 23 orang, Petani 78 orang, swasta 59 orang, untuk tingkat SMP TNI/Polri/PNS 46 orang, Petani 108 orang, swasta 79 orang, dan untuk tingkat SMA TNI/Polri/PNS 78 orang, Petani 85 orang, swasta 56 orang. Di Kabupaten Gianyar untuk tingkat SD pekerjaan orang tua siswa adalah TNI/Polri/PNS 44 orang, Petani 67 orang, swasta 78 orang, untuk tingkat SMP TNI/Polri/PNS 45 orang, Petani 78 orang, swasta 98 orang, dan untuk tingkat SMA TNI/Polri/PNS 56 orang, Petani 45 orang, swasta 89 orang. Di Kabupaten Badung untuk tingkat SD pekerjaan orang tua siswa adalah TNI/Polri/PNS 86 orang, Petani 42 orang, swasta 112 orang, untuk tingkat SMP TNI/Polri/PNS 120 orang, Petani 41 orang, swasta 93 orang, dan untuk tingkat SMA TNI/Polri/PNS 23 orang, Petani 151 orang, swasta 124 orang. Di Kabupaten Bangli untuk tingkat SD pekerjaan orang tua siswa adalah TNI/Polri/PNS 23 orang, Petani 56 orang, swasta 86 orang, untuk tingkat SMP TNI/Polri/PNS 56 orang, Petani 73 orang, swasta 120 orang, dan untuk tingkat SMA TNI/Polri/PNS 42 orang, Petani 12 orang, swasta 120 orang. Susunlah data di atas ke dalam bentuk tabel baris dan kolom!

2. Apakah yang dimaksud dengan tabel distribusi frekuensi, dan kenapa pula data tunggal perlu dibuatkan tabel distribusi frekuensi?
3. Apakah yang dimaksud dengan tabel distribusi frekuensi dan tabel distribusi frekuensi kumulatif?
4. Berikut ini diberikan data hasil ujian statistik mahasiswa.

29	34	41	35	51	36	36	38	43	40	55	41
42	42	66	42	48	46	47	50	47	56	48	48
49	52	50	51	41	52	67	52	55	42	56	28
56	56	47	57	61	61	62	62	66	42	57	34

Sususnlah data tersebut ke dalam distribusi frekuensi! Tentukan pula batas atas dan batas bawah, tepi atas dan tepi bawah masing-masing kelas interval!

5. Berdasarkan soal No. 4 buatlah histogram dan poligon frekuensi, grafik lingkaran, serta ogive!
6. Buatlah 50 buah data tunggal, sajikanlah data tersebut ke dalam distribusi frekuensi data bergolong kemudian tentukanlah unsur-unsur yang ditanyakan sesuai butir soal No. 4 dan No. 5.

BAB III
PEMUSATAN DATA

Dalam statistik ukuran data dapat digolongkan menjadi tiga, yaitu: ukuran pemusatan data (mean, median, dan modus), ukuran penyebaran data (median, kuartil, desil, persentil), dan ukuran penyebaran data (jangkauan, jangkauan antar kuartil, simpagan rata-rata, dan simpangan baku dan ragam). Pada bab ini akan dibahas tentang pemusatan data, sedangkan ukuran letak dan ukuran penyebaran akan dibahas pada bab berikutnya.

Sebelum lebih jauh membahas tentang pemusatan data alangkah baiknya kita pahami dahulu beberapa hal berikut ini. Misalnya disajikan data tunggal hasil ujian statistik mahasiswa 10, 12, 14, 14, 17, 18, 22, 22, 22, dan 25. Dari data tersebut terlihat bahwa nilai ujian paling tinggi adalah 25 dan nilai ujian paling rendah adalah 10. Dengan demikian rentangannya adalah 15. Dengan asumsi bahwa sepuluh orang yang ikut dalam ujian merupakan sebuah populasi terbatas yang sanagat kecil. Nilai atau angka 25, 10, dan 15 merupakan deskripsi tentang populasi. Nilai-nilai tersebut disebut dengan parameter. Jadi parameter adalah sebarang nilai yang mendeskripsikan atau menjelaskan ciri-ciri dari populasi. Biasanya parameter disimbolkan dalam bentuk huruf yunani. Lebih jauh dapat dilihat parameter tersebut nilainya dalam bentuk konstanta.

Apabila sepuluh orang yang ikut ujian tersebut merupakan sebagian kecil mahasiswa dari suatu universitas, maka nilai-nilai 25, 10, dan 15 tersebut tidak lagi disebut dengan parameter. Nilai-nilai tersebut mendeskripsikan bagian dari keseluruhan atau mendeskripsikan contoh. Nilai-nilai itu disebut dengan statistik. Jadi statistik adalah sebarang nilai yang menjelaskan ciri suatu contoh. Statistik disimbolkan dengan huruf kecil biasa.

Salah satu aspek yang paling penting untuk menggambarkan distribusi data adalah nilai pusat data pengamatan (*tendency central*). Ukuran kecenderungan memusat merupakan suatu bilangan yang menunjukkan tendensi (kecenderungan) memusatnya bilangan-bilangan dalam suatu distribusi. Ukuran kecenderungan memusat juga dapat digunakan untuk merangkum data dan mendeskripsikan suatu kelompok data dengan cara mencari suatu angka (indeks) yang dapat mewakili seluruh kelompok tersebut. Setiap pengukuran aritmatika yang ditujukan untuk menggambarkan suatu nilai yang mewakili nilai pusat atau nilai sentral dari suatu gugus data (himpunan pengamatan) dikenal sebagai ukuran tendensi sentral.Ukuran pemusatan adalah sebarang ukuran yang menunjukkan pusat segugus data, yang telah diurutkan dari nilai terkecil sampai dengan nilai terbesar atau sebaliknya dari nilai terbesar sampai dengan nilai terkecil. Sebagai ilustrasi setiap hari seseorang harus meminum air rata-rata sebanyak 3 liter sehari, maka nilai ini bisa dipandang sebagai nilai pusat dari beberapa nilai lainya. Bisa saja seseorang akan minum air lebih banyak ketiga berada di daerah panas atau sehabis olah raga, begitu juga seseorang akan minum lebih sedikit apabila berada di daerah dingin.

Ukuran pemusatan data digunakan agar data yang diperoleh mudah untuk dibaca dan dipahami. Ukuran pemusatan data terdiri atas mean (rata-rata), median, dan modus. Walaupun ukuran pemusatan dapat memudahkan kita dalam membaca data, tetapi ukuran-ukuran tersebut memiliki kelemahan. Kelemahan-kelemahan tersebut adalah: nilai rata-rata sangat dipengaruh oleh nilai niali-nilai ekstrim. Median terlalu bervariasi untuk dijadikan parameter populasi. Sedangkan modus hanya dapat diterapkan dalam data dengan ukuran yang besar, karena data yang berukuran kecil jarang memiliki modus.

A. Rata-rata (Mean)

Ada tiga jenis rata-rata yang dibahas pada bagian ini, yaitu rata-rata geometri (GM), rata-rata harmonik (HM), dan rata-rata aritmatika (AM). Ketiga jenis rata-rata itu sering disebut dengan rata-rata *pythagorian* dan memiliki hubungan bahwa HM≤GM≤AM. Selajutnya untuk menyatakan ketiga rata-rata tersebut merupakan simbol statistik, maka menggunakan huruf kecil. Khusus untuk rata-rata aritmatika tidak menggunakan simbol a kecil tetapi x bar ($\bar{x}$).

a. Rata-rata geometri (ukur)

Rata-rata ukur (geometri) adalah rata-rata yang diperoleh dengan mengalikan semua data dalam suatu kelompok sampel, kemudian diakarpangkatkan dengan jumlah data sampel tersebut. Rata-rata geometri disimbolkan dengan g atau u. Jika terdapat sekumpulan data tungal x_1, x_2, x_3, …, x_n, maka secara matematis rata-rata ukur (geometri) dirumuskan seperti berikut ini.

$$g = \sqrt[x]{\prod_{i=1}^{n} x_i}$$

atau

$$g = \sqrt[n]{x_1 . x_2 . x_3 . \ldots . x_n}$$

keterangan

g = rata-rata geometri

π = nilai hasil perkalian semua data

n = banyaknya data

Jika rumus di atas kedua ruas dilogkan, maka akan menjadi

$$\log(g) = \log\left(\sqrt[n]{x_1 . x_2 . x_3 . \ldots . x_n}\right)$$

$$= \frac{1}{n}\log(x_1 . x_2 . x_3 . \ldots . x_n)$$

$$= \frac{1}{n}\left(\log(x_1) + \log(x_2) + \log(x_3) + \ldots + \log(x_n)\right)$$

$$\log(g) = \frac{1}{n}\sum_{i=1}^{n}\log(x_i)$$

Contoh 3.1 **Diketahui data nilai ujian statistik sepuluh orang mahasiswa adalah sebagai berikut. 7, 7, 5, 6, 5, 6, 5, 9, 6, dan 5. Berapakah rata-rata ukur (geometri) nilai ujian statistik tersebut?**

Penyelesaian :

Rata-rata ukur (geometri) bisa dihitung dengan menggunakan rumus pertama atau rumus kedua

$$g = \sqrt[n]{x_1 . x_2 . x_3 . \ldots . x_n}$$

$$= \sqrt[10]{7 x 7 x 5 x 6 x 5 x 6 x 5 x 9 x 6 x 5}$$

$$= \sqrt[10]{59535000}$$

$$= 5{,}99$$

atau

$$\log(g) = \frac{1}{n}\sum_{i=1}^{n}\log(x_n)$$

$$= \frac{1}{10}(\log(7)+\log(7)+\log(5)+\log($$

$$+\log(9)+\log(6)+\log(5))$$

$$= \frac{1}{10}(0{,}845 + 0{,}845 + 0{,}699 + 0$$

$$+ 0{,}954 + 0{,}778 + 0{,}699)$$

$$= \frac{7{,}75}{10}$$

Log (g) = 0,775

g = antilog (0,7775)

= 5,99

Hasil yang diperoleh dengan rumus pertama atau kedua sama, atau mungkin perbedaanya dibelakang koma karena terjadi pembulatan. Apabila data dalam bentuk distribusi frekuensi data tidak bergolong maupun distribusi frekuensi data bergolong, maka rata-rata harmoniknya adalah sebagai berikut.

$$g = \sqrt[n]{x_1^{f_1}.x_2^{f_2}.x_3^{f_3}. \ldots .x_n^{f_n}}$$

Atau jika rumus di atas kedua ruas dilogkan, maka akan menjadi

$$\log(g) = \log\left(\sqrt[n]{x_1^{f_1}.x_2^{f_2}.x_3^{f_3}. \ldots .x_n^{f_3}}\right)$$

$$= \frac{1}{n}\log\left(x_1^{f_1}.x_2^{f_2}.x_3^{f_3}. \ldots .x_n^{f_n}\right)$$

$$= \frac{1}{n}\left(\log(x_1^{f_1}) + \log(x_2^{f_2}) + \log(x_3^{f_3}) + \ldots + \log(x_n^{f_n})\right)$$

$$\log(g) = \frac{1}{n}\sum_{i=1}^{n} f_i . \log(x_i)$$

Contoh 3.2 **Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-rata geometrinya!**

No	**Nilai**	**f**
1	5	3
2	6	5
3	7	7
4	8	2
Total		**17**

Penyelesaian :

No	**Nilai (x)**	**f**	**log (x)**	**f . log (x)**
1	5	3	0,699	2,097
2	6	5	0,778	3,891
3	7	7	0,845	5,916
4	8	2	0,903	1,806
Total		**17**	**-**	**13,710**

$$\log(g) = \frac{1}{n}\sum_{i=1}^{4} f_i . \log(x_i)$$

$$= \frac{1}{17}(13,710)$$

$$= 0,8064$$

Contoh 3.3 **Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-rata geometrinya!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian:

No	Kelas Interval	f	Titik tengah (x)	log (x)	f . log (x)
1	50 – 54	2	52	1,716	3,432
2	55 – 59	5	57	1,756	8,779
3	60 – 64	8	62	1,792	14,339
4	65 – 69	4	67	1,826	7,304
5	70 – 74	1	72	1,857	1,857
Total		**20**			**35,712**

$$\log(g) = \frac{1}{n}\sum_{i=1}^{5} f_i . \log(x_i)$$

$$= \frac{1}{20}(35,712)$$

$$= 1,785$$

Rata-rata geometri biasanya dipakai untuk data yang memiliki bobot/kualitas/berat (*weight*) yang berbeda di antara data-data tersebut. Umumnya data-data ini memiliki nilai batas minimum dan maksimum. Misalnya universitas A memiliki rentangan IP mahasiswa 3 - 4, sedangkan universitas B memiliki rentangan IP 2 - 3,5. Jika kita ingin mencari nilai rata-rata IP mahasiswa kedua universitas tersebut, maka rerata geometri umum digunakan karena nilainya berbobot.

b. Rata-rata harmonik

Rata-rata harmonik atau *harmonic mean* adalah rata-rata yang dihitung dengan cara mengubah semua data menjadi kebalikannya, kemudian semua bilangan tersebut dijumlahkan dan selanjutnya dijadikan sebagai pembagi jumlah data. Rata-rata harmonik ini sering disebut juga dengan kebalikan dari rata-rata hitung (aritmatik). Rata-rata harmonik disimbulkan dengan h.

Jika diketahui sekelompok data tunggal x_1, x_2, x_3, …, x_n, maka secara matematis rata-rata harmonik dirumuskan sebagai berikut.

$$h = \frac{n}{\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)}$$

$$h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + ... + \frac{1}{n_n}}$$

keterangan

h = rata-rata harmonik

n = banyaknya data

Contoh 3.4 **Lihat kembali Contoh 3.1! Diketahui data nilai ujian statistik sepuluh orang mahasiswa adalah sebagai berikut. 7, 7, 5, 6, 5, 6, 5, 9, 6, dan 5. Berapakah rata-rata harmonik nilai ujian statistik tersebut?**

Penyelesaian

$$h = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\ldots+\frac{1}{n_n}}$$

$$= \frac{10}{\frac{1}{7}+\frac{1}{7}+\frac{1}{5}+\frac{1}{6}+\frac{1}{5}+\frac{1}{6}+\frac{1}{5}+\frac{1}{9}+\frac{1}{6}+\frac{1}{5}}$$

$$= \frac{10}{0{,}143+0{,}143+0{,}2+0{,}167+0{,}2+0{,}167+0{,}2+0{,}111+0{,}167+0{,}2}$$

$$= \frac{10}{1{,}697}$$

$$= 5{,}893$$

Apabila data dalam bentuk distribusi frekuensi data tidak bergolong maupun distribusi frekuensi data bergolong, maka rata-rata harmoniknya adalah sebagai berikut.

$$h = \frac{n}{\sum_{i=1}^{n}\left(\frac{f_i}{x_i}\right)}$$

Contoh 3.5 **Lihat kembali Contoh 3.2! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-rata harmoniknya!**

No	**Nilai**	**f**
1	5	3
2	6	5
3	7	7
4	8	2
Total		**17**

Penyelesaian:

No	**Nilai (x)**	**f**	**f/x**
1	5	3	0,600
2	6	5	0,833
3	7	7	1,000
4	8	2	0,250
Total		**17**	**2,683**

$$h = \frac{n}{\sum_{i=1}^{n}\left(\frac{f_i}{x_i}\right)}$$

$$= \frac{17}{2,683}$$

$$= 6,336$$

Contoh 3.6 **Lihat kembali Contoh 3.3! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-rata harmoniknya!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian:

No	Kelas Interval	f	Titik tengah (x)	f/x
1	50 – 54	2	52	0,038
2	55 – 59	5	57	0,088
3	60 – 64	8	62	0,129
4	65 – 69	4	67	0,060
5	70 – 74	1	72	0,014
Total		**20**	**-**	**0,329**

$$h = \frac{n}{\sum_{i=1}^{n}\left(\frac{f_i}{x_i}\right)}$$

$$= \frac{20}{0,329}$$

$$= 60,790$$

Rata-rata harmonik biasanya digunakan untuk data yang berupa data rasio, seperti misalnya tekanan (gaya per luas), kecepatan (jarak per waktu), debit air (volume per waktu), dan data lainnya yang memiliki satuan rasio (perbandingan). Sebagai contoh ada dua mobil menempuh jarak yang sama yaitu dari kota X ke kota Y, tetapi dengan kecepatan yang berbeda. Mobil A menempuh kedua kota tersebut dengan kecepatan 100km/jam sedangkan mobil B menempuh dengan kecepatan 80 km/ jam. Dengan demikian rata-rata kecepatan kedua mobil tersebut adalah 88,889 km/jam. Contoh ini menunjukkan bahwa nilai yang tetap adalah jarak tempunya, yaitu dari kota A dan kota B, tetapi jika dalam kasus lain nilai yang tetap adalah waktu, maka yang digunakan untuk menghitung kecepatannya adalah rata-rata aritmatika.

c. Hitung aritmatika (hitung)

Rata-rata ini mungkin sudah diajarkan mulai sekolah dasar, tidak seperti dua rata-rata sebelumnya. Rata-rata hitung atau *arithmetic mean* atau sering disebut dengan istilah *mean* saja merupakan metode yang paling banyak digunakan untuk menggambarkan ukuran tendensi sentral. Jika hanya disebut dengan kata “rata-rata” saja, maka rata-rata yang dimaksud adalah rata-rata hitung (aritmatik). Nilai rata-rata merupakan salah satu ukuran untuk memberikan gambaran yang lebih jelas dan singkat tentang sekumpulan data mengenai suatu hal. Nilai rata-rata ini dianggap sebagai nilai yang paling dekat dengan ukuran yang sebenarnya. Untuk lebih mempermudah pemahaman, sebaiknya terlebih dahulu samakan persepsi tentang beberapa simbol berikut. x_1, x_2, x_3, …, x_n merupakan simbol dari nilai-nilai dari data kuantitatif, yang berfungsi menyatakan banyaknya data. N menyatakan banyaknya anggota populasi, n menyatakan banyaknya anggota sampel. $\bar{x}$ menyatakan rata-rata di sampel, sedangkan ì menyatakan rata-rata di populasi. Dengan demikian $\bar{x}$ menyatakan statistik dan ì menyatakan parameter.

1) Rata-rata data tunggal

Menghitung nilai rata-rata dilakukan dengan menjumlahkan seluruh nilai data suatu kelompok sampel, kemudian dibagi dengan jumlah sampel tersebut. Jadi jika diketahui sekelompok data tungal $x_1, x_2, x_3, \ldots, x_n$, maka secara matematis rata-ratanya dirumuskan sebagai berikut.

$$\bar{x} = \frac{\text{jumlah data}}{\text{banyaknya data}}$$

atau

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

atau

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

Keterangan

$\bar{x}$ = rata-rata

n = banyak data

Contoh 3.7 Lihat kembali Contoh 3.1! Diketahui data nilai ujian statistik sepuluh orang mahasiswa adalah sebagai berikut. 7, 7, 5, 6, 5, 6, 5, 9, 6, dan 5. Berapakah rata-rata harmonik nilai ujian statistik tersebut?

Penyelesaian

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

$$= \frac{7+7+5+6+5+6+5+9+6+5}{10}$$

$$= \frac{61}{10}$$

$$= 6{,}1$$

2) Rata-rata gabungan

Kadangkala pada saat menganalisis rata-rata suatu data ternyata sudah diketahui rata-rata masing-masing kelompok, kemudian disuruh menghitung rata-rata gabungannya. Jika terdapat beberapa kelompok data yang masing-masing rata-rata diketahui, maka dapat menghitung rata-rata gabungan dari kelompok-kelompok data tersebut.

$$\bar{x}_{gab} = \frac{n_1 . x_1 + n_2 . x_2 + n_3 . x_3 + ... + n_n . x_n}{n_1 + n_2 + n_3 + ... + n_n}$$

Contoh 3.8 **Rata-rata nilai ujian statistik 40 orang mahasiswa adalah 75, lima orang lainnya mengikuti ujian susulan dengan rata-rata 65. Berapakah rata-rata nilai ujian statistik seluruh mahasiswa tersebut.**

Penyelesaian

$\bar{x}_1 = 75$ $n_1 = 40$

$\bar{x}_2 = 65$ $n_2 = 5$

$$\bar{x}_{gab} = \frac{n_1 . x_1 + n_2 . x_2}{n_1 + n_2}$$

$$= \frac{(75 \times 40) + (65 \times 5)}{40 + 5}$$

$$= \frac{3000 + 325}{45}$$

$$= \frac{3325}{45}$$

$$= 73{,}889$$

3) Rata-rata data bergolong

Apabila dalam data tunggal yang cukup banyak terdapat nilai yang berulang beberapa kali, maka akan lebih mudah jika data tersebut disajikan dalan tabel distribusi frekuensi kemudian ditentukan rata-ratanya. Sehingga untuk menentukan mean atau rata-rata data bergolong/berbobot dapat dicari dengan tiga langkah, yaitu menggunakan titik tengah (secara langsung), dengan menggunakan rata-rata sementara, dan dengan menggunakan kode (*coding*).

(a) Menggunakan titik tengah

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3 + ... + f_k x_k}{n}$$

atau

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

atau

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

keterangan

$\bar{x}$ = rata-rata

x_i = nilai ke-i atau titik tengah kelas interval ke-i

k = banyaknya kelas interval

Contoh 3.9 **Lihat kembali Contoh 3.2! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-ratanya!**

No	Nilai	f
1	5	3
2	6	5
3	7	7
4	8	2
Total		**17**

Penyelesaian :

No	Nilai (x)	f	f . x
1	5	3	15
2	6	5	30
3	7	7	49
4	8	2	16
Total		**17**	**110**

$$\bar{x} = \frac{\sum_{i=1}^{4} f_i x_i}{\sum_{i=1}^{4} f_i}$$

$$= \frac{110}{17}$$

$$= 6{,}470$$

Contoh 3.10 **Lihat kembali Contoh 3.3! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah rata-ratanya!**

No	**Kelas Interval**	**f**
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian:

No	**Kelas Interval**	**f**	**Titik tengah (x)**	**f . x**
1	50 – 54	2	52	104
2	55 – 59	5	57	285
3	60 – 64	8	62	496
4	65 – 69	4	67	268
5	70 – 74	1	72	72
Total		**20**	**-**	**1225**

$$\bar{x} = \frac{\sum_{i=1}^{5} f_i\, x_i}{\sum_{i=1}^{5} f_i}$$

$$= \frac{1225}{20}$$

$$= 61{,}250$$

(b) Menggunakan rataan sementara

Sebelum menghitung rata-rata data berkelompok menggunakan simpangan rata-rata sementara, kita terlebih dahulu menetapkan rata-rata sementaranya. Perlu digaris bawahi bahwa tebakan rata-rata sementara bisa ditebak berapa saja asal masih dalam batas-batas kelas interval. Rumus yang digunakan adalah sebagai berikut.

$$\bar{x} = \bar{x}_s + \frac{\sum_{i=1}^{k} f_i\, d_i}{\sum_{i=1}^{k} f_i}$$

keterangan

$\bar{x}$ = rata-rata

$\bar{x}_s$ = rataan sementara

d = selisih antara titik tengah kelas masing-masing kelas interval dikurangi tebakan rata-rata

f = frekuensi

k = banyak kelas interval

Lihat kembali Contoh 3.10, misalkan rata-rata sementara yang di tebak adalah 65, tebakan ini tidak boleh di bawah 50 atau di atas 74. Selanjutnya kita bisa membuat tabel analisisnya sebagai berikut.

No	Kelas Interval	f	Titik tengah (x)	d (x_i - 65)	f . d
1	50 – 54	2	52	-13	-26
2	55 – 59	5	57	-8	-40
3	60 – 64	8	62	-3	-24
4	65 – 69	4	67	2	8
5	70 – 74	1	72	7	7
Total		**20**	**-**	**-**	**-75**

$$\bar{x} = \bar{x}_s + \frac{\sum_{i=1}^{5} f_i d_i}{\sum_{i=1}^{5} f_i}$$

$$= 65 + \frac{-75}{20}$$

$$= 65 - 3{,}750$$

$$= 61{,}250$$

(c) Menggunakan kode (*coding*)

Sama dengan menggunakan simpangan rata-rata sementara, sebelum menghitung rata-rata dengan cara coding, kita juga harus menetapkan rata-rata sementara. Namun rata-rata sementara yang kita tetapkan harus sama dengan salah satu nilai tengah salah satu kelas interval. Tebakannya bisa dilakukan di mana saja asal salah satu titik tengah kelas interval. Rumus yang digunakan adalah sebagai berikut.

$$\bar{x} = \bar{x}_s + \frac{\sum_{i=1}^{k} f_i c_i}{\sum_{i=1}^{k} f_i} p$$

keterangan

$\bar{x}$ = rata-rata

$\bar{x}_s$ = rataan sementara yang ditebak dari salah satu titik tengah kelas interval

c = pengkodean

f = frekuensi

p = panjang kelas interval

k = banyak kelas interval

Pengkodean dimulai dari angka 0 untuk kelas interval di mana rata-rata sementara ditetapkan. Dari kelas interval yang ditetapkan kode 0, ke arah kelas interval yang lebih kecil terus dukurangi 1 atau berturut-turut menjadi angka negatif dengan selisih (-1, -2, -3 dan seterusnya) menjauhi kelas rata-rata sementara. Berikutnya dengan kelas interval yang lebih tinggi terus ditambah 1 atau berturut-turut menjadi angka positif dengan selisih (1, 2, 3 dan seterusnya) menjauhi kelas rata-rata sementara.

Lihat kembali Contoh 3.10, misalkan rata-rata sementara yang kita tetapkan adalah 62. Selanjutnya kita bisa membuat tabel analisisnya sebagai berikut.

No	**Kelas Interval**	**f**	**Titik tengah (x)**	**c**	**f . c**
1	50 – 54	2	52	-2	-4
2	55 – 59	5	57	-1	-5
3	60 – 64	8	62	0	0

No	Kelas Interval	f	Titik tengah (x)	c	f . c
4	65 – 69	4	67	1	4
5	70 – 74	1	72	2	2
Total		**20**	**-**	**-**	**-3**

$$\bar{x} = \bar{x}_s + \frac{\sum_{i=1}^{5} f_i c_i}{\sum_{i=1}^{5} f_i} p$$

$$= 62 + \left(\frac{-3}{20} \times 5\right)$$

$$= 62 + \left(\frac{-15}{20}\right)$$

$$= 62 - 0{,}750$$

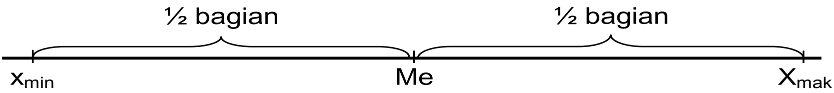
$$= 61{,}250$$

Ternyata menggunkan ketiga rumus data bergolong, yaitu dengan menggunakan titik tengah, dengan menggunakan rataan sementara, dan dengan menggunakan kode (*coding*). Untuk lebih meyakinkan bahwa hasil yang diperoleh sama atau mungkin hanya berbeda di belakang koma bisa dilakukan dengan menggunakan tebakan yang lain.

B. Median

Median atau nilai-tengah merupakan ukuran pemusatan data yang lain. Median dapat diartikan sebagai nilai tengah segugus data, yaitu: jika segugus data diurutkan dari nilai terkecil sampai nilai terbesar atau dari nilai terbesar sampai nilai terkecil, nilai pengamatan yang berada tepat di tengah-

tengah bila banyak data pengamatan ganjil, atau rata-rata kedua pengamatan yang berada di tengah bila banyak data pengamatan genap. Median tidak ditentukan oleh nilai ekstrim dari data seperti pada rata-rata yang nilainya ditentukan oleh nilai ekstrim.



Sebelum menentukan berapa nilai median dari segugusan data tunggal maupun data berkelompok, terlebih dahulu yang harus dicari adalah letak median. Adapun rumus yang digunakan untuk menentukan letak median adalah sebagai berikut.

$$Lm = \frac{(n+1)}{2}$$

keterangan
Lm = letak median
n = banyak data

1) Data tunggal

Untuk menghitung median (nilai tengah) segugus data tunggal sangat sederhana, yaitu dengan terlebih dahulu mengurutkan data dari nilai terkecil ke nilai terbesar atau sebaliknya dari nilai terbesar ke nilai terkecil, kemudian menentukan letak kelas median dan terakhir menentukan nilai mediannya. Berikut ini diberikan contoh untuk menghitung median (nilai tengah) segugus data tunggal.

Contoh 3.11 **Tentukanlah median atau nilai tengah data berikut!**

a.	22	54	36	51	72	62	56	70	46	
b.	56	68	46	35	75	25	64	72	40	28

Penyelesaian

a. Terlebih dahulu data diurutkan dari nilai terkecil ke nilai terbesar.

22 36 46 51 54 56 62 70 72

Setelah itu kita tentukan letak median

$$Lm = \frac{(n+1)}{2}$$

$$= \frac{(9+1)}{2}$$

$$= 5$$

Jadi letak median berada pada data ke-5, data kelima nilainya 54, sehingga median data tersebut adalah 54.

b. Terlebih dahulu data diurutkan dari nilai terkecil ke nilai terbesar.

25 28 35 40 46 56 64 68 72 75

Setelah itu kita tentukan letak median

$$Lm = \frac{(n+1)}{2}$$

$$= \frac{(10+1)}{2}$$

$$= 5{,}5$$

Jadi letak median berada diantara data ke-5 dan data ke-6, data kelima nilainya 46 dan data ke-6 nilainya 56, sehingga median data tersebut adalah (46+56)/2 = 102/2 = 51

2) Data bergolong

Pada data bergolong, tidak terlalu mudah untuk menentukan median. Hal ini disebabkan karena padatnya nilai-nilai serta telah terkuburnya sejumlah nilai dalam kelompok-kelompok nilai. Biasanya kesulitan ditemui apabila median data bergolong yang banyak datanya genap terletak diantara dua kelas interval yang berbeda. Misalnya median dari sekumpulan data terletak diantara data ke-12 dan data-13, sedangkan data ke-12 terletak pada kelas interval ketiga sedangkan data ke-13 terletak pada kelas interval keempat. Median data bergolong atau berbobot dapat ditentukan dengan rumus berikut ini.

$$Me \quad = Bb + \left(\frac{\frac{1}{2}n - f_k}{f_{me}} \right) p$$

keterangan

Me = median

Bb = batas bawah kelas median

f_k = frekuensi kumulatif <u>sebelum</u> kelas median

f_{me} = frekuensi kelas median

p = panjang kelas interval

n = banyak data

Contoh 3. 12 **Lihat kembali Contoh 3.10! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah mediannya!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian :

No	Kelas Interval	f	Titik tengah (x)	f_k
1	50 – 54	2	52	2
2	55 – 59	5	57	7
3	60 – 64	8	62	15
4	65 – 69	4	67	19
5	70 – 74	1	72	20
Total		**20**	**-**	

Terlebih dahulu tentukan letak median

$$Lm = \frac{(n+1)}{2}$$

$$= \frac{(20+1)}{2}$$

$$= \frac{21}{2}$$

$$= 10{,}5$$

Jadi letak median berada diantara data ke-10 dan data ke-11 kedua data tersebut terletak pada kelas interval ke-3, dengan demikian dapat ditentukan nilai dari:

$Bb = 59{,}5$
$f_k = 7$
$f_{me} = 8$
$p = 5$
$n = 20$

Sehingga Me dapat dihitung sebagai berikut.

$$Me = Bb + \left(\frac{\frac{1}{2}n - f_k}{f_{me}}\right)p$$

$$= 59{,}5 + \left(\frac{\left(\frac{1}{2}20\right) - 7}{8}\right)5$$

$$= 59{,}5 + \left(\frac{10 - 7}{8}\right)5$$

$$= 59{,}5 + \left(\frac{3}{8}\right)5$$

$$= 59{,}5 + \left(\frac{15}{8}\right)$$

$$= 59{,}5 + 1{,}875$$

$$= 61{,}375$$

C. Modus

Modus adalah nilai dari segugus data yang sering muncul atau data yang memiliki frekuensi tertinggi. Tanpa kita sadari dalam kehidupan sehari-hari kita sering mengucapkan kata yang sebenarnya merupakan modus. Kata tersebut adalah kata "sering". Misalnya seorang dosen mengatakan bahwa semester VII kelas A adalah kelas yang paling sering ribut. Ketika dosen melakukan absensi mengatakan bahwa si B paling sering tidak kuliah. Kejadian-kejadian tersebut merupakan contoh dari modus yang kita tidak sadari. Orang yang penyakit jantung dikatakan penyebabnya merokok padahal mungkin ada hal lain yang menyebabkan orang terkena penyakit jantung tetapi tidak sebanyak yang disebabkan oleh rokok. Kejadian-kejadian tersebut merupakan contoh dari ukuran pemusatan data, yaitu modus yang kita tidak sadari.

Tidak selalu segugus data memiliki modus, bisa saja segugus data tidak memiliki modus. Jika segugus data hanya memiliki satu modus, maka gugusan data tersebut disebut dengan unimodal, jika memiliki dua modus disebut dengan bimodal, atau jika mempunyai lebih dari dua modus disebut dengan multimodal.

1) Data tunggal

Gugusan data tunggal dengan mudah bisa ditentukan hanya dengan melihat data dengan frekuensi tertinggi. Menentukan modus data tunggal tidak perlu mengurutkan data, cukup dihitung frekuensi tertingginya (jika ada). Perhatikan contoh di bawah ini.

Contoh 3.13 Tentukanlah modus data berikut!

a)	5	5	5	5	5	5	5	5	5	5
b)	5	4	6	8	4	8	7	4	9	5
c)	6	7	6	8	4	7	7	9	6	9
d)	2	5	3	8	2	7	5	9	6	3

Penyelesaian

a. Gugusan data 5 5 5 5 5 5 5 5 5 5 tidak memiliki modus, 5 tidak bisa dikatakan paling sering muncul karena tidak ada data yang tidak sering muncul. Sehingga tidak ada pembandingnya untuk mengatakan bahwa 5 itu paling sering muncul atau frekuensinya paling tinggi.

b. Gugusan data 5 4 6 8 4 8 7 4 9 5 memiliki satu modus yaitu 4, karena 4 memiliki frekuensi paling tinggi, yaitu 3 dibandingkan data yang lainnya. Jadi gugusan data tersebut memiliki satu modus, yaitu 4 yang disebut dengan unimodal.

c. Gugusan data 6 7 6 8 4 7 7 9 6 9 memiliki dua modus yaitu 6 dan 7, karena 6 dan 7 memiliki frekuensi paling tinggi, yaitu 3 dibandingkan data yang lainnya. Jadi gugusan data tersebut memiliki dua modus, yaitu 6 dan 7 yang disebut dengan bimodal.

d. Gugusan data 2 5 3 8 2 7 5 9 6 3 memiliki tiga modus yaitu 2, 3 dan 5, karena 2, 3, dan 5 memiliki frekuensi paling tinggi, yaitu 2 dibandingkan data yang lainnya. Jadi gugusan data tersebut memiliki tiga modus, yaitu 2, 3 dan 5 yang disebut dengan multimodal.

2) Data bergolong

Untuk data berkelompok, dalam hal ini adalah distribusi frekuensi, modus hanya dapat diperkirakan tidak bisa dihitung dengan tepat seperti pada segugusan data tunggal. Bisa terjadi untuk data dalam bilangan bulat apabila disajikan dalam distribusi frekuensi dengan kelas interval akan ditemukan nilai modus dalam bentuk bilangan desimal. Nilai yang paling sering muncul akan berada pada kelas yang memiliki frekuensi terbesar. Kelas yang memiliki frekuensi terbesar disebut kelas modus. Modus data bergolong atau berbobot dapat ditentukan dengan rumus berikut ini.

$$Mo = Bb + \left(\frac{b_1}{b_1 + b_2} \right) p$$

keterangan

Mo = modus

Bb = batas bawah kelas median

b_1 = selisih frekuensi kelas modus dengan frekuensi kelas interval sebelumnya

b_2 = selisih frekuensi kelas modus dengan frekuensi kelas interval sesudahnya

p = panjang kelas interval

Perlu digaris bawahi bahwa kata sebelumnya dan sesudahnya bukan menyatakan urutan kelas interval. Kata "sebelumnya" mengandung makna bahwa kelas interval yang memiliki nilai lebih kecil dibandingkan kelas interval yang mengandung modus. Sedangkan kata "sesudahnya" mengandung makna bahwa kelas interval yang memiliki nilai lebih besar dibandingkan kelas interval yang mengandung modus.

Contoh 3.14 **Lihat kembali Contoh 10! Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah modusnya!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian :

No	Kelas Interval	f	Titik tengah (x)
1	50 – 54	2	52
2	55 – 59	5	57
3	60 – 64	8	62
4	65 – 69	4	67
5	70 – 74	1	72
Total		**20**	**-**

Pada tabel di atas terlihat data yang memiliki frekuensi tertinggi terletak pada kelas interval ke tiga dengan frekuensi 8, jadi modus data tersebut terletak pada kelas interval ketiga, dengan demikian dapat ditentukan nilai dari:

Bb = 59,5
b_1 = 8 - 5 = 3
b_2 = 8 - 4 = 4
p = 5

Sehingga Mo dapat dihitung sebagai berikut.

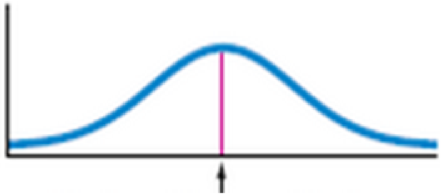
$$
\begin{aligned}
Mo &= Bb + \left(\frac{b_1}{b_1 + b_2}\right)p \\
&= 59{,}5 + \left(\frac{3}{3+4}\right)5 \\
&= 59{,}5 + \left(\frac{15}{7}\right) \\
&= 59{,}5 + 2{,}143 \\
&= 61{,}643
\end{aligned}
$$

D. Hubungan Empiris Antara Mean, Median, dan Modus

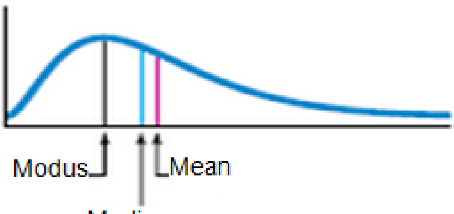
Sebelum lebih lanjut membahas hubungan empirik antara mean, median, dan modus terlebih dahulu kita bahas kesimetrisan atau kemiringan kurva distribusi data berdasarkan hubungan empiris antara nilai ketiga pemusatan data tersebut. Ukuran kemiringan atau kecondongan atau kemencengan atau juling (*skewness*) 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 atau simetris, positif, atau negatif.

1) Jika mean, median, dan modus nilainya hampir sama (berdekatan) satu sama lain atau, mean ≈ median ≈ modus maka kurva dari data tersebut akan dikatakan mendekati simetrik atau *zero skewness*. Gambar kurvanya dapat

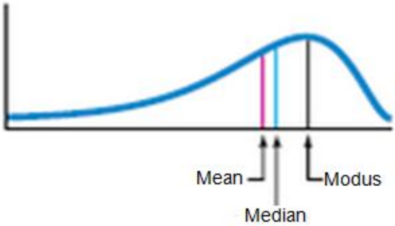
dilihat pada gambar di bawah ini. Tetapi apa bila mean data sama dengan mediannya atau hampir sama dan tidak memiliki modus maka data disebut berdistribusi *uniform*.



2) Jika nilai modus kurang dari nilai median, dan nilai median kurang dari nilai mean atau modus < median < mean, maka kurva dari distribusi data miring atau menceng ke kanan atau juling positif. Kurva yang miring ke kanan memiliki ekor yang lebih panjang ke kanan daripada yang ke kiri.



3) Jika nilai mean kurang dari nilai median, dan nilai median kurang dari nilai modus atau mean < median < modus, maka kurva dari distribusi data miring atau menceng ke kiri atau juling negatif. Kurva yang miring ke kiri memiliki ekor yang lebih panjang ke kiri daripada yang ke kanan.



Jika melihat ketiga bagian kurva tersebut maka untuk bagian kedua, yaitu nilai modus paling kecil dan nilai rata-rata hitung paling besar, sedangkan pada bagian ketiga, sebaliknya, yaitu nilai rata-rata hitung paling kecil dan nilai modus paling besar. Kurva bagian kedua dan ketiga disebut dengan kurva tidak simetris. Untuk distribusi data yang tidak simetri, yaitu juling ke kanan atau ke kiri, terdapat hubungan empiris antara mean, median dan modus sebagai berikut.

mean – modus = 3 (mean – median)

Walaupun mean, median, modus, sama-sama merupakan ukuran pemusatan data, tetapi ternyata masing-masing dari besaran mempunyai kelebihan dan kekurangan. Adapun kelebihan dan kekurangan dari masing-masing besaran tersebut disajikan pada Tabel 3.1 di bawah ini.

Tabel 3.1 Kelebihan dan Kekurangan Mean, Median, dan Modus

No	Ukuran Pemusatan	Kelebihan	Kelemahan
1	Mean	Mempertimbangkan semua nilai dalam data, mampu menggambarkan mean populasi, dan cocok untuk data homogen.	Peka atau mudah terpengaruh oleh nilai ekstrim dan kurang baik untuk data heterogen.

No	Ukuran Pemusatan	Kelebihan	Kelemahan
2	Median	Tidak terpengaruh oleh nilai ekstrim dan cocok untuk data heterogen	Tidak mempertimbangkan semua nilai dankurang dapat menggambarkan mean populasi.
3	Modus	Tidak terpengaruh oleh nilai ekstrim dan cocok untuk data homogen maupun heterogen.	Kurang menggambarkan mean populasi karena modus bisa lebih satu.

Latihan 3

1. Tentukalah mean, median, dan modus data tunggal berikut ini.
   a. 5 8 3 7 3 8 6 4 5 9 4 2 8 7 5 2 1
   b. 5 5 6 6 7 7 4 4 9 9 8 8
   c. 12 35 25 45 86 95 25 38 45 67 25 68 48 49 96 75 72
2. Sebuah tes diikuti oleh 100 siswa dari empat kelas, yaitu A, B, C, dan D. Nilai rata-rata kelas A adalah 7,5 rata-rata kelas B adalah 9 dan rata-rata kelas C adalah 8, dan rata-rata kelas D adalah 6. Banyaknya siswa kelas A, B, dan C masing-masing 25, 34, dan 27. Tentukan nilai rata-rata keempat kelas tersebut!
3. Tentukanlah mean, median, dan modus data berikut ini.

No	Kelas Interval	f
1	40 - 49	5
2	50 - 59	12
3	60 - 69	34
4	70 - 79	28
5	80 - 89	11
6	90 - 99	4
Total		**94**

4. Buatlah sebaran data tunggal sebanyak 80 data, kemudian buatlah tabel distribusi data bergolong dengan kelas interval. Dari tabel distribusi tersebut tentukanlah ukuran pemusatannya. Apa yang dapat anda simpulkan jika data tadi ukuran pemusatannya dibandingkan dengan ukuran pemusatan yang dihitung dari sebaran data tunggal.

BAB IV
UKURAN LETAK DAN UKURAN PENYEBARAN DATA

Ukuran pemusatan data mean, median, dan modus yang dibahas pada BAB III belum memberikan keterangan atau deskripsi yang mencukupi, karena besaran-besaran tersebut merupakan nilai perwakilan dari suatu distribusi frekuensi, tetapi ukuran tersebut tidak memberikan gambaran informasi yang lengkap mengenai bagaimana penyebaran data pengamatan terhadap nilai sentralnya. Dengan kata lain kita perlu mengetahui seberapa jauh pengamatan-pengamatan itu menyebar dari nilai rata-ratanya (mean). Sangat mungkin terjadi ketika dua kumpulan data hasil pengamatan atau lebih memiliki rata-rata atau median yang sama, tetapi memiliki keragaman yang berbeda. Sebagai contoh hasil ujian statistik dua kelas berikut ini.

Kelas A	:	55	56	81	65	89	66	78
Kelas B	:	70	71	70	68	74	69	68

Kita dapat melihat bahwa nilai mean hasil ujian statistik kelas A dan mean hasil ujian statistik kelas B sama, yaitu 70. Tetapi jika kita perhatikan, keragaman kedua kelompok data berbeda. Nilai ujian statistik kelas B lebih konsisten dibandingkan nilai ujian statistik kelas A. Hal ini terlihat dari data hasil ujian statistik kelas B lebih seragam dibandingkan dengan hasil ujian statistik kelas A. Nilai ujian statistik kelas B, hasilnya tidak terlalu jauh penyimpangan dari nilai rata-ratanya sebesar 70. Hasil ujian statistik kelas B, sebaran datanya sangat beragam dibandingkan nilai kelas B. Kelompok data yang seragam, penyebaran atau variasinya sangat kecil disebut data yang homogen. Sedangkan kelompok data yang kurang seragam, penyebaran atau variasinya relatif besar disebut data yang tidak homogen.

Pada contoh di atas, jelas bahwa rata-rata yang merupakan bagian dari ukuran pemusatan data, tidak cukup untuk menggambarkan distribusi frekuensi. Selain itu kita harus memiliki ukuran penyebaran data pengamatan. Ukuran penyebaran atau ukuran keragaman pengamatan dari nilai rata-ratanya disebut simpangan (*deviation*/dispersi). Terdapat beberapa ukuran untuk menentukan dispersi data pengamatan, seperti jangkauan/rentang (*range*), simpangan kuartil (*quartile deviation*), simpangan rata-rata (*mean deviation*), dan simpangan baku (*standard deviation*).

Sebelum lebih jauh kita membahas ukuran penyebaran (dispersi) terlebih dahulu kita pahami tentang ukuran letak, yaitu kuartil, desil, dan persentil yang semuanya itu merupakan pengembangan dari median yang merupakan bagian dari pemusatan data.

A. Ukuran Letak

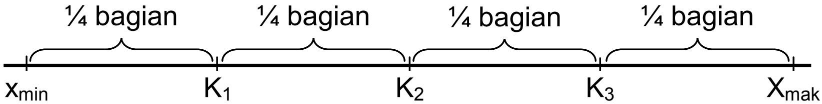
Selain ukuran pemusatan terdapat pula ukuran letak. Salah satu dari ukuran letak yang juga ukuran pemusatan data adalah median yang menunjukkan nilai tengah dalam susunan data yang diurutkan mulai dari nilai terkecil ke nilai terbesar. Dengan demikian median terletak di tengah-tengah data yang telah diurutkan dan dapat dianggap bahwa median membagi data yang telah diurutkan itu menjadi dua kelompok data yang sama banyak. Selain median ada ukuran letak lainnya, yaitu kuartil, desil, dan persentil yang diistilahkan dengan *fraktil*.

1. Kuartil

Jika median dapat dikatakan sebagai ukuran perduaan, maka kuartil disebut dengan ukuran perempatan. Kuartil dalam kehidupan sehari-hari disebut dengan kuartal dan disimbolkan dengan K atau Q. Kuartil adalah ukuran letak yang membagi distribusi data menjadi empat bagian sama besar, yaitu masing-masing $\frac{1}{4}n$ Jadi akan dijumpai tiga buah kuartil, yaitu kuartil pertama atau kuartil bawah (K_1), kuartil kedua atau kuartil tengah (K_2) yang merupakan median, dan kuartil ketiga atau kuartil atas (K_3). Menentukan kuartil pertama, kedua, dan ketiga sama seperti menentukan median, data harus diurutkan dari nilai terkecil ke nilai terbesar atau sebaliknya dari nilai terbesar ke nilai terkecil.

Kuartil pertama atau bawah (K_1) adalah suatu nilai yang membatasi 25% distribusi bagian bawah dan 75% distribusi bagian atas. Kuartil kedua atau tengah (K_2) adalah nilai yang membatasi 50% distribusi bagian bawah dan 50% distribusi bagian atas. Dalam hal ini kuartil kedua dapat diidentikkan dengan median (Me). Kuartil ketiga atau atas (K_3) adalah nilai yang membatasi 75% distribusi bagian bawah dan 25% distribusi bagian atas.

Jika data diurutkan dari nilai terkecil ke nilai terbesar, maka K_1 akan terletak di sebelah kiri K_2 dan K_3 terletak disebelah kanan K_2. Sedangkan jika data diurutkan dari nilai terbesar ke terkecil, maka K_1 akan terletak di sebelah kanan K_2 dan K_3 akan terletak di sebelah kiri K_2. Dengan posisi K_2 selalu di tengan-tengah K_1 dan K_3. Secara visualisasi dapat diilustrasikan sebagai berikut ($n \geq 4$).



Sebelum menentukan berapa nilai kuartil dari segugusan data tunggal maupun data berkelompok, terlebih dahulu yang harus dicari adalah letak kuartil. Adapun rumus yang digunakan untuk menentukan letak kuartil adalah sebagai berikut.

$$L_k = \frac{i(n+1)}{4}$$

keterangan
L_k = letak kuartil
n = banyak data
i = 1, 2, 3

a. Data tunggal

Tentunya dalam menentukan letak median data tunggal kita akan menemukan bentuk pecahan campuran yang kemungkinan pecahannya $\frac{1}{4}$, $\frac{1}{2}$, atau $\frac{3}{4}$. Sebagai ilustrasi, seandainya kita menemukan letak kuartil ke-i atau $K_i = p\frac{1}{4}$, maka k_i terletak pada data ke-p dan data ke-(p+1) sehingga nilai kuartilnya adalah nilai data ke-p + $\frac{1}{4}$(data ke-(p+1)-data ke-p). Untuk lebih jelasnya perhatikan Contoh 4.1 di bawah ini.

Contoh 4.1 Tentukalah nilai K_1, K_2, dan K_3 masing-masing data berikut!

a) 22 54 36 51 72 62 56 70 46

b) 56 68 46 35 75 25 64 72 40 28

Penyelesaian

a) Terlebih dahulu data diurutkan 22 36 46 51 54 56 62 70 72

Letak K_1 = $\frac{1.(9+1)}{4}$

= $\frac{10}{4}$

= $2\frac{1}{2}$, jadi K_1 terletak diantara data ke-2 dan ke-3

Nilai K_1 = data ke-2 + $\frac{1}{2}$(data ke-3 – data ke-2)

= 36 + $\frac{1}{2}$(46 - 36)

= 36 + $\frac{1}{2}$(10)

= 36 + 5

= 41

Letak K_2 = $\frac{2.(9+1)}{4}$

= $\frac{20}{4}$

= 5, jadi K_1 terletak di data ke-5

Nilai K_2 = data ke-5

= 54

Letak K_3 = $\frac{3.(9+1)}{4}$

= $\frac{30}{4}$

= $7\frac{1}{2}$, jadi K_3 terletak diantara data ke-7 dan ke-8

Nilai K_3 = data ke-7 + $\frac{1}{2}$(data ke-8 – data ke-7)

$= 62 + \frac{1}{2}(70 - 62)$

$= 62 + \frac{1}{2}(8)$

$= 62 + 4$

= 66

b) Terlebih dahulu data diurutkan 25 28 35 40 46 56 64 68 72 75

Letak $K_1 = \frac{1.(10+1)}{4}$

$= \frac{11}{4}$

$= 2\frac{3}{4}$, jadi K_1 terletak diantara data ke-2 dan ke-3

Nilai K_1 = data ke-2 + $\frac{3}{4}$(data ke-3 – data ke-2)

$= 28 + \frac{3}{4}(35 - 28)$

$= 28 + \frac{3}{4}(7)$

$= 28 + \frac{21}{4}$

$= 28 + 5,25$

= 33,25

$$\text{Letak } K_2 = \frac{2.(10+1)}{4}$$

$$= \frac{22}{4}$$

$= 5\frac{1}{2}$, jadi K_2 terletak diantara data ke-5 dan data

Nilai K_2 = data ke-5 + $^1\!/_2$ (data ke-6 – data ke-5)

$= 46 + {}^1\!/_2\ (56 - 46)$

$= 46 + {}^1\!/_2\,(10)$

$= 46 + {}^{10}\!/_2$

$= 46 + 5$

= 51

$$\text{Letak } K_3 = \frac{3.(10+1)}{4}$$

$$= \frac{33}{4}$$

$= 8\frac{1}{4}$, jadi K_3 terletak diantara data ke-8 dan ke-9

Nilai K_3 = data ke-8 + $^1\!/_4$ (data ke-9 – data ke-8)

$= 68 + {}^1\!/_4\ (72 - 68)$

$= 68 + \frac{1}{4} (4)$

$= 68 + 1$

$\mathbf{= 69}$

b. Data bergolong

Menentukan kuartil data bergolong sama halnya seperti menentukan median, oleh karena itu alangkah baiknya diingat kembali rumus untuk menentukan median data bergolong. Menentukan median data bergolong tergantung dengan nilai $\frac{i}{4}n$, sedangkan untuk menentukan kuartil ditentukan oleh nilai $\frac{i}{4}n$ dengan i = 1, 2, 3. Jika i diganti, maka nilai $\frac{1}{4}n$ untuk K_1, $\frac{2}{4}n$ untuk K_2, dan $\frac{3}{4}n$ untuk K_3. Dengan demikian rumus kuartil data bergolong adalah sebagai berikut.

$$Ki = Bb + \left(\frac{\frac{i}{4}n - f_k}{f} \right) p$$

keterangan

K_i = kuartil ke-i (i = 1, 2, 3)

Bb = batas bawah kelas interval yang mengandung kuartil

f_k = frekuensi kumulatif sebelum kelas interval yang mengandung kuartil

f = frekuensi kelas interval yang mengandung kuartil

p = panjang kelas interval

n = banyak data

Contoh 4.2 **Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah K_1, K_2, dan K_3!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian:

No	Kelas Interval	f	Titik tengah (x)	f_k
1	50 – 54	2	52	2
2	55 – 59	5	57	7
3	60 – 64	8	62	15
4	65 – 69	4	67	19
5	70 – 74	1	72	20
Total		**20**	**-**	

a) Menentukan nilai K_1 Terlebih dahulu tentukan letak K_1

$$\text{Letak } K_1 = \frac{i(n+1)}{4}$$

$$= \frac{1.(20+1)}{4}$$

$$= \frac{21}{4}$$

$$= 5\frac{1}{4}$$

$$= 5{,}25$$

K_1 terletak pada kelas interval ke-2, karena paling sedikit harus sama dengan 5,25. Dengan demikian dapat ditentukan nilai dari :

Bb = 54,5
f_k = 2
f = 5
p = 5
n = 20

Sehingga K_1 dapat dihitung sebagai berikut.

$$K_1 = Bb + \left(\frac{\frac{1}{4}n - f_k}{f}\right)p$$

$$= 54{,}5 + \left(\frac{\left(\frac{1}{4}.20\right) - 2}{5}\right)5$$

$$= 54{,}5 + \left(\frac{5-2}{5}\right)5$$

$$= 54{,}5 + \left(\frac{3}{5}\right)5$$

$$= 54{,}5 + 3$$

$$\mathbf{= 57{,}5}$$

b) Menentukan nilai K_2 terlebih dahulu tentukan letak K_2

$$\text{Letak } K_2 = \frac{i.(n+1)}{4}$$

$$= \frac{2.(20+1)}{4}$$

$$= \frac{42}{4}$$

$$= 10{,}5$$

K_2 terletak pada kelas interval ke-3, karena paling sedikit harus sama dengan 10,5. Dengan demikian dapat ditentukan nilai dari:

Bb = 59,5
f_k = 7
f = 8
p = 5
n = 20

Sehingga K_2 dapat dihitung sebagai berikut.

$$K_2 = Bb + \left(\frac{\frac{2}{4}n - f_k}{f}\right)p$$

$$= 59{,}5 + \left(\frac{\left(\frac{2}{4}.20\right) - 7}{8}\right)5$$

$$= 59{,}5 + \left(\frac{10-7}{8}\right)5$$

$$= 59{,}5 + \left(\frac{3}{8}\right)5$$

$$= 59{,}5 + \left(\frac{15}{8}\right)$$

$$= 59{,}5 + 1{,}875$$

$$\mathbf{= 61{,}375}$$

c) Menentukan nilai K_3 terlebih dahulu tentukan letak K_3

$$\text{Letak } K_3 = \frac{i.(n+1)}{4}$$

$$= \frac{3.(20+1)}{4}$$

$$= \frac{63}{4}$$

$$= 15\frac{3}{4}$$

$$= 15{,}75$$

K_3 terletak pada kelas interval ke-4, karena paling sedikit harus sama dengan 15,75. Dengan demikian dapat ditentukan nilai dari:

Bb = 64,5
f_k = 15
f = 4

p = 5
n = 20

Sehingga K_3 dapat dihitung sebagai berikut.

$$K_3 = Bb + \left(\frac{\frac{3}{4}n - f_k}{f} \right) p$$

$$= 64,5 + \left(\frac{\left(\frac{3}{4}.20\right) - 15}{4} \right) 5$$

$$= 64,5 + \left(\frac{15 - 15}{4} \right) 5$$

$$= 64,5 + \left(\frac{0}{4} \right) 5$$

$$= 64,5$$

2. Desil

Desil disebut sebagai persepuluh. Desil adalah ukuran letak yang membagi seluruh distribusi frekuensi dari data yang kita selidiki ke dalam 10 bagian yang sama besar, yang masing-masing sebesar $^1/_{10}$n. Dengan demikian terdapat sembilan desil yang membagi distribusi data menjadi 10 bagian yang sama, yaitu: D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, dan D_9. D_1 membatasi 10% data bagian bawah dan 90% data bagian atas, D_2 membatasi 20% data bagian bawah dan 80% data bagian atas, D_3 membatasi 20% data bagian bawah dan 80%

data bagian atas, dan seterusnya sampai D_9 yang membatasi 90% data bagian bawah dan 10% data bagian atas.

Sebelum menentukan berapa nilai desil dari segugusan data tunggal maupun data berkelompok, terlebih dahulu yang harus dicari adalah letak desil. Adapun rumus yang digunakan untuk menentukan letak desil adalah sebagai berikut.

$$L_D = \frac{i(n+1)}{10}$$

keterangan

L_D = letak desil

n = banyak data

i = 1, 2, …, 9

a. Data tunggal

Tentunya dalam menentukan letak desil data tunggal kita akan menemukan bentuk pecahan campuran yang kemungkinan pecahannya seperti $^1\!/_{10}$, $^2\!/_{10}$, $^3\!/_{10}$, $^6\!/_{10}$ atau yang lainya. Sebagai ilustrasi, seandainya kita menemukan letak desil ke-i atau $D_i = p\,^1\!/_{10}$, maka D_i terletak pada data ke-p dan data ke-(p+1) sehingga nilai desilnya adalah nilai data ke-p $+ ^1\!/_{10}$(data ke-(p+1) - data ke-p). Untuk lebih jelasnya perhatikan Contoh 4.3 di bawah ini.

Contoh 4.3 **Tentukanlah nilai D_1, D_5, dan D_9 masing-masing data berikut!**

a)	22	54	36	51	72	62	56	70	46	
b)	56	68	46	35	75	25	64	72	40	28

Penyelesaian

a) Terlebih dahulu data diurutkan 22 36 46 51 54 56 62 70 72

Letak $D_1 = \dfrac{1.(9+1)}{10}$

$= \dfrac{10}{10}$

$= 1$, jadi D_1 terletak pada data ke-1

Nilai D_1 = data ke-1

= **22**

Letak $D_5 = \dfrac{5.(9+1)}{10}$

$= \dfrac{50}{10}$

$= 5$, jadi D_5 terletak pada data ke-5

Nilai D_5 = data ke-5

= 54

Letak $D_9 = \dfrac{9.(9+1)}{10}$

$= \dfrac{90}{10}$

$= 9$, jadi D_9 terletak pada data ke-9

Nilai D_9 = data ke-9

= 72

b) Terlebih dahulu data diurutkan 25 28 35 40 46 56 64 68 72 75

Letak $D_1 = \dfrac{1.(10+1)}{10}$

$= \dfrac{11}{10}$

$= 1\dfrac{1}{10}$, jadi D_1 terletak diantara data ke-1 dan ke-2

Nilai D_1 = data ke-1 + $^1/_{10}$(data ke-2 – data ke-1)

$= 25 + {}^1/_{10}(28 - 25)$

$= 25 + {}^1/_{10}(3)$

$= 25 + {}^3/_{10}$

$= 25 + 0{,}3$

= 25,3

Letak $D_5 = \dfrac{5.(10+1)}{10}$

$= \dfrac{55}{10}$

$= 5\dfrac{1}{2}$, jadi D_5 terletak diantara data ke-5 dan data ke-6

Nilai D_5 = data ke-5 + $^1/_2$(data ke-6 – data ke-5)

$= 46 + {}^1/_2\ (56 - 46)$

$= 46 + \frac{1}{2}(10)$

$= 46 + \frac{10}{2}$

$= 46 + 5$

$\mathbf{= 51}$

Letak $D_9 = \frac{9.(10+1)}{10}$

$= \frac{99}{10}$

$= 9\frac{9}{10}$, jadi D_9 terletak diantara data ke-9 dan ke-10

Nilai D_9 = data ke-9 + $\frac{9}{10}$(data ke-10 – data ke-9)

$= 72 + \frac{9}{10}(75 - 72)$

$= 72 + \frac{9}{10}(3)$

$= 72 + 2{,}7$

$\mathbf{= 74{,}7}$

a. Data bergolong

Menentukan desil data bergolong sama halnya seperti menentukan kuartil, oleh karena itu alangkah baiknya diingat kembali rumus untuk menentukan kuartil data bergolong. Menentukan kuartil data bergolong tergantung dengan nilai $\frac{i}{4}n$, sedangkan untuk menentukan desil ditentukan oleh nilai $\frac{i}{10}n$ dengan i = 1, 2,..., 9 Jika i diganti, maka nilai $\frac{1}{10}n$

untuk D_1, $\frac{2}{10}n$ untuk D_2, $\frac{3}{10}n$ untuk D_3, dan seterusnya sampai dengan, $\frac{9}{10}n$ untuk D_9. Dengan demikian rumus desil data bergolong adalah sebagai berikut.

$$di \quad = Bb + \left(\frac{\frac{i}{10}n - f_k}{f} \right) p$$

keterangan

D_i	=	desil ke-i (i = 1, 2, 3, …, 9)
Bb	=	batas bawah kelas interval yang mengandung desil
f_k	=	frekuensi kumulatif <u>sebelum</u> kelas interval yang mengandung desil
f	=	frekuensi kelas interval yang mengandung desil
p	=	panjang kelas interval
n	=	banyak data

Contoh 4.4 **Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah D_1, D_5, dan D_8!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian :

No	Kelas Interval	f	Titik tengah (x)	f_k
1	50 – 54	2	52	2
2	55 – 59	5	57	7
3	60 – 64	8	62	15
4	65 – 69	4	67	19
5	70 – 74	1	72	20
Total		**20**	**-**	

a) Menentukan nilai D_1 Terlebih dahulu tentukan letak D_1

$$\text{Letak } D_1 = \frac{i\,(n+1)}{10}$$

$$= \frac{1.\,(20+1)}{10}$$

$$= \frac{21}{10}$$

$$= 2\frac{1}{10}$$

$$= 2{,}1$$

D_1 terletak pada kelas interval ke-2, karena paling sedikit harus sama dengan 2,1 (kelas interval pertama frekuensinya 2, jadi kurang dari 2,1) . Dengan demikian dapat ditentukan nilai dari:

Bb = 54,5
f_k = 2
f = 5
p = 5
n = 20

Sehingga D_1 dapat dihitung sebagai berikut.

$$D_1 = Bb + \left(\frac{\frac{1}{10}n - f_k}{f}\right)p$$

$$= 54{,}5 + \left(\frac{\left(\frac{1}{10}.20\right) - 2}{5}\right)5$$

$$= 54{,}5 + \left(\frac{2-2}{5}\right)5$$

$$= 54{,}5 + 0$$

$$\mathbf{= 54{,}5}$$

b) Menentukan nilai D_5 terlebih dahulu tentukan letak D_5

$$\text{Letak } D_5 = \frac{i.(n+1)}{10}$$

$$= \frac{5.(20+1)}{4}$$

$$= \frac{105}{10}$$

$$= 10{,}5$$

D_5 terletak pada kelas interval ke-3, karena paling sedikit harus sama dengan 10,5. Dengan demikian dapat ditentukan nilai dari:

Bb = 59,5
f_k = 7
f = 8
p = 5
n = 20

Sehingga D_5 dapat dihitung sebagai berikut.

$$D_5 = Bb + \left(\frac{\frac{5}{10} n - f_k}{f} \right) p$$

$$= 59,5 + \left(\frac{\left(\frac{5}{10} . 20 \right) - 7}{8} \right) 5$$

$$= 59,5 + \left(\frac{10 - 7}{8} \right) 5$$

$$= 59,5 + \left(\frac{3}{8} \right) 5$$

$$= 59,5 + \left(\frac{15}{8} \right)$$

$$= 59,5 + 1,875$$

$$\mathbf{= 61,375}$$

c) Menentukan nilai D_8 terlebih dahulu tentukan letak D_8

$$\text{Letak } D_8 = \frac{i.(n+1)}{10}$$

$$= \frac{8.(20+1)}{10}$$

$$= \frac{168}{10}$$

$$= 16\frac{8}{10}$$

$$= 16{,}8$$

D_8 terletak pada kelas interval ke-4, karena paling sedikit harus sama dengan 16,8. Dengan demikian dapat ditentukan nilai dari:

Bb = 64,5
f_k = 15
f = 4
p = 5
n = 20

Sehingga D_8 dapat dihitung sebagai berikut.

$$D_8 = Bb + \left(\frac{\frac{8}{10}n - f_k}{f}\right)p$$

$$= 64{,}5 + \left(\frac{\left(\frac{8}{10}.20\right) - 15}{4}\right)5$$

$$= 64{,}5 + \left(\frac{16-15}{4}\right)5$$

$$= 64{,}5 + \left(\frac{1}{4}\right)5$$

$$= 64{,}5 + \left(\frac{5}{4}\right)$$

$$= 64{,}5 + 1{,}25$$

$$\mathbf{= 65{,}75}$$

3. Persentil

Persentil adalah ukuran letak yang membagi data yang telah diukur atau data yang berkelompok menjadi seratus bagian yang sama besar. Persentil menentukan nilai batas tiap 1% dalam distribusi data. karena membagi data menjadi seratus bagian yang sama, maka dikenal ada 99 nilai persentil yakni: persentil ke-1 (P_1), persentil ke-2 (P_2), persentil ke-3 (P_3) dan seterusnya sampai dengan persentil ke-99 atau P_{99}. P_1 membatasi 1% data bagian bawah dan 99% data bagian atas, P_2 membatasi 2% data bagian bawah dan 98% data bagian atas, P_3 membatasi 3% data bagian bawah dan 97% data bagian atas, dan seterusnya sampai P_{99} yang membatasi 1% data bagian bawah dan 99% data bagian atas.

Sebelum menentukan berapa nilai persentil dari segugusan data tunggal maupun data berkelompok, terlebih dahulu yang harus dicari adalah letak persentil. Adapun rumus yang digunakan untuk menentukan letak persentil adalah sebagai berikut.

$$L_P = \frac{i(n+1)}{100}$$

keterangan

L_P = letak persentil

n = banyak data

i = 1, 2, 3 …, 99

a. Data tunggal

Tentunya dalam menentukan letak persentil data tunggal kita akan menemukan bentuk pecahan campuran yang kemungkinan pecahannya seperti $^1/_{100}$, $^2/_{100}$, $^3/_{100}$, $^6/_{100}$, $^9/_{100}$ atau yang lainya. Sebagai ilustrasi, seandainya kita menemukan letak persentil ke-i atau P_i = $p\frac{2}{100}$, maka P_i terletak pada data ke-p dan data ke-(p+1) sehingga nilai persentilnya adalah nilai data ke-p + $\frac{2}{100}$ (data ke-(p+1) - data ke-p). Untuk lebih jelasnya perhatikan Contoh 4.5 di bawah ini.

Contoh 4.5 **Tentukalah nilai P_{25}, P_{50}, dan P_{75} masing-masing data berikut!**

a)	22	54	36	51	72	62	56	70	46	
b)	56	68	46	35	75	25	64	72	40	28

Penyelesaian

a) Terlebih dahulu data diurutkan 22 36 46 51 54 56 62 70 72

$$\text{Letak } P_{25} = \frac{25.(9+1)}{100}$$

Letak P_{25} $= \dfrac{25.(9+1)}{100}$

$= \dfrac{250}{100}$

$= 2\dfrac{1}{2}$, jadi P_{25} terletak pada data ke-2 dan data ke-3

Nilai P_{25} = data ke-2 + $\dfrac{1}{2}$ (data ke-3 – data ke-2)

$= 36 + \dfrac{1}{2}\,(46 - 36)$

$= 36 + 5$

= 41

Letak P_{50} $= \dfrac{50.(9+1)}{100}$

$= \dfrac{500}{100}$

= 5, jadi P_{50} terletak pada data ke-5

Nilai P_{50} = data ke-5

= 54

Letak D_{75} $= \dfrac{75.(9+1)}{100}$

$= \dfrac{750}{100}$

$= 7\dfrac{1}{2}$, jadi P_{75} terletak diantara data ke-7 dan data ke-8

Nilai P_{75} $= \text{data ke-7} + \frac{1}{2} (\text{data ke-8 - data ke-7})$

$= 62 + \frac{1}{2} (70 - 62)$

$= 62 + \frac{1}{2} (8)$

$= 62 + 4$

= 66

b) Terlebih dahulu data diurutkan 25 28 35 40 46 56 64 68 72 75

Letak P_{25} $= \frac{25.(10+1)}{100}$

$= \frac{11}{4}$

$= 2\frac{3}{4}$, jadi P_{25} terletak diantara data ke-2 dan ke-3

Nilai P_{25} $= \text{data ke-2} + \frac{3}{4} (\text{data ke-3} - \text{data ke-2})$

$= 28 + \frac{3}{4} (35 - 28)$

$= 28 + \frac{3}{4} (7)$

$= 28 + \frac{21}{4}$

$= 28 + 5{,}25$

= 33,25

Letak P_{50} $= \frac{50.(10+1)}{100}$

$= \frac{11}{2}$

$= 5\frac{1}{2}$, jadi P_{50} terletak diantara data ke-5 dan data ke-6

Nilai P_{50} = data ke-5 + $\frac{1}{2}$(data ke-6 – data ke-5)

$= 46 + \frac{1}{2}\,(56 - 46)$

$= 46 + \frac{1}{2}(10)$

$= 46 + \frac{10}{2}$

$= 46 + 5$

= 51

Letak P_{75} $= \dfrac{75.(10+1)}{100}$

$= \dfrac{33}{4}$

$= 8\frac{1}{4}$, jadi P_{25} terletak diantara data ke-8 dan ke-9

Nilai P_{75} = data ke-8 + $\frac{1}{4}$(data ke-9 – data ke-8)

$= 68 + \frac{1}{4}\,(72 - 68)$

$= 68 + \frac{1}{4}\,(4)$

$= 68 + 1$

= 69

b. Data bergolong

Menentukan persentil data bergolong sama halnya seperti menentukan kuartil dan desil, oleh karena itu alangkah baiknya diingat kembali rumus untuk menentukan kuartil dan desil data bergolong. Menentukan kuartil data bergolong tergantung dengan nilai $\frac{i}{4}n$, menentukan desil data bergolong tergantung dengan nilai $\frac{i}{10}n$ sedangkan untuk menentukan persentil data bergolong ditentukan oleh nilai $\frac{i}{100}n$ dengan i = 1, 2,..., 99 Jika i diganti, maka nilai $\frac{1}{100}n$ untuk P_1, $\frac{2}{100}n$ untuk P_2, $\frac{3}{100}n$ untuk P_3, dan seterusnya sampai dengan, $\frac{9}{100}n$ untuk P_{99}. Dengan demikian rumus persentil data bergolong adalah sebagai berikut.

$$Pi = Bb + \left(\frac{\frac{i}{100}n - f_k}{f} \right) p$$

keterangan

P_i = persentil ke-i (i = 1, 2, 3, …, 9)

Bb = batas bawah kelas interval yang mengandung persentil

f_k = frekuensi kumulatif <u>sebelum</u> kelas interval yang mengandung persentil

f = frekuensi kelas interval yang mengandung persentil

p = panjang kelas interval

n = banyak data

Contoh 4.6 **Berikut ini diketahui distribusi frekuensi nilai ujian statistik mahasiswa pada suatu universitas. Tentukanlah P_{25}, P_{50}, dan P_{75}!**

No	Kelas Interval	f
1	50 – 54	2
2	55 – 59	5
3	60 – 64	8
4	65 – 69	4
5	70 – 74	1
Total		**20**

Penyelesaian:

No	Kelas Interval	f	Titik tengah (x)	f_k
1	50 – 54	2	52	2
2	55 – 59	5	57	7
3	60 – 64	8	62	15
4	65 – 69	4	67	19
5	70 – 74	1	72	20
Total		**20**	**-**	

a) Menentukan nilai P_{25} Terlebih dahulu tentukan letak P_{25}

$$\text{Letak } P_{25} = \frac{i\,(n+1)}{100}$$

$$= \frac{25\,.\,(20+1)}{100}$$

$$= \frac{21}{4}$$

$$= 5\frac{1}{4}$$

$$= 5{,}25$$

P_{25} terletak pada kelas interval ke-2, karena paling sedikit harus sama dengan 5,25. Dengan demikian dapat ditentukan nilai dari:

Bb = 54,5
f_k = 2
f = 5
p = 5
n = 20

Sehingga P_{25} dapat dihitung sebagai berikut.

$$P_{25} = Bb + \left(\frac{\frac{25}{100}n - f_k}{f} \right) p$$

$$= 54,5 + \left(\frac{\left(\frac{25}{100} . 20 \right) - 2}{5} \right) 5$$

$$= 54,5 + \left(\frac{5 - 2}{5} \right) 5$$

$$= 54,5 + \left(\frac{3}{5} \right) 5$$

$$= 54,5 + 3$$

$$\mathbf{= 57,5}$$

b) Menentukan nilai P_{50} terlebih dahulu tentukan letak P_{50}

$$\text{Letak } P_{50} = \frac{i.(n+1)}{100}$$

$$= \frac{50.(20+1)}{100}$$

$$= \frac{42}{4}$$

$$= 10{,}5$$

P_{50} terletak pada kelas interval ke-3, karena paling sedikit harus sama dengan 10,5. Dengan demikian dapat ditentukan nilai dari:

Bb = 59,5
f_k = 7
f = 8
p = 5
n = 20

Sehingga P_{50} dapat dihitung sebagai berikut.

$$P_{50} = Bb + \left(\frac{\frac{50}{100} n - f_k}{f} \right) p$$

$$= 59{,}5 + \left(\frac{\left(\frac{50}{100} . 20 \right) - 7}{8} \right) 5$$

$$= 59{,}5 + \left(\frac{10 - 7}{8} \right) 5$$

$$= 59{,}5 + \left(\frac{3}{8}\right)5$$

$$= 59{,}5 + \left(\frac{15}{8}\right)$$

$$= 59{,}5 + 1{,}875$$

$$\mathbf{= 61{,}375}$$

c) Menentukan nilai P_{75} terlebih dahulu tentukan letak P_{75}

$$\text{Letak } P_{75} = \frac{75.(n+1)}{100}$$

$$= \frac{75.(20+1)}{100}$$

$$= \frac{63}{4}$$

$$= 15\frac{3}{4}$$

$$= 15{,}75$$

P_{75} terletak pada kelas interval ke-4, karena paling sedikit harus sama dengan 15,75. Dengan demikian dapat ditentukan nilai dari:

Bb = 64,5
f_k = 15
f = 4
p = 5
n = 20

Sehingga K_3 dapat dihitung sebagai berikut.

$$P_{75} = Bb + \left(\frac{\frac{75}{100} n - f_k}{f} \right) p$$

$$= 64{,}5 + \left(\frac{\left(\frac{75}{100} .20 \right) - 15}{4} \right) 5$$

$$= 64{,}5 + \left(\frac{15 - 15}{4} \right) 5$$

$$= 64{,}5 + \left(\frac{0}{4} \right) 5$$

$$= 64{,}5$$

Dalam dunia pendidikan persentil banyak digunakan untuk mengubah skor mentah (*raw score*) ke dalam skor standar (*standard) score*). Skor mentah (*raw score*) adalah data yang belum berubah, misalnya hasil asli yang diperoleh mahasiswa dalam suatu ujian yang merupakan hasil menjawab benar, sedangkan skor standar (*standard)* adalah peringkat dalam persentil atau sejenisnya. Skala sebelas diambil dari kata "*standard eleven*" yang disingkat *Stanel* yang dipergunakan untuk mengubah skor mentah yang diperoleh siswa ke dalam 11 kelompok nilai, yaitu 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, dan 10. Skala ini paling sering digunakan oleh para guru. Di samping sudah terbiasa menggunakannya, proses perhitungannya pun mudah dan nilai tersebut bisa secara langsung mencerminkan prestasi penguasaan siswa terhadap materi tes. Mengubah dari skor mentah menjadi *stanel* dilakukan dengan jalan menghitung nilai-nilai persentil berikut. P_1- P_3- P_8- P_{21}- P_{39}- P_{61}- P_{79}- P_{92}- P_{97}- dan P_{99}.

Persentil dapat digunakan untuk menentukan kedudukan seorang mahasiswa, yaitu: pada persentil keberapakah mahasiswa itu memperoleh kedudukan ditengah-tengah kelompoknya. Persentil juga dapat digunakan sebagai alat untuk menetapkan nilai batas lulus pada tes atau seleksi. Perhatikan kembali Contoh 4.6, dari 20 orang mahasiswa akan diluluskan 5 orang saja, yaitu $\frac{5}{20}$ x 100%= 25% dan yang tidak diluluskan adalah 15 orang, yaitu $\frac{15}{20}$ x 100%=75%, hal ini berarti bahwa P_{75} adalah batas nilai kelulusan. Mereka yang nilai-nilainya berada pada P_{75} ke bawah, dinyatakan tidak lulus, sedangkan di atas P_{75} dinyatakan lulus. Dari hasil analisi Contoh 4.6 diperoleh P_{75}= 64,50; berarti yang dapat diluluskan adalah mereka yang nilainya di atas 64,50 yaitu nilai 65 ke atas.

Beberapa contoh analisis di atas juga menunjukkan hububungan ketiga ukuran letak, kuartil, desil, dan persentil, yaitu: $P_{90} = D_9$, $P_{80} = D_8$, $P_{75} = K_3$, $P_{70} = D_7$, $P_{60} = D_6$, $P_{50} = D_5 = K_2$ = Median, $P_{40} = D_4$, $P_{30} = D_3$, $P_{25} = K_1$, $P_{20} = D_2$, dan $P_{10} = D_1$.

B. Ukuran Penyebaran (Dispersi)

Telah diungkapkan di atas bahwa rata-rata dari serangkaian nilai observasi tidak dapat diinterpretasikan secara terpisah dari hasil dispersi nilai-nilai tersebut sekitar rata-ratanya. Penyebaran atau dispersi adalah pergerakan dari nilai observasi terhadap nilai rata-ratanya. Dispersi menunjukkan seberapa jauh penyimpangan nilai-nilai data dari nilai pusatnya (rata-ratanya) atau bagaimana penyebaran suatu kelompok data. Dengan demikian semakin besar dispersinya, semakin besar variasi nilainya, sehingga makin kurang representatif rata-rata distribusinya. Ukuran penyebaran penting untuk dihitung karena ukuran pemusatan yang kita ukur belum memberikan informasi yang lengkap, selain itu dispersi dapat digunakan untuk membandingkan penyebaran dua distribusi data atau lebih. Secara sederhana dapat dikatakan bahwa ukuran penyebaran dapat digunakan untuk menentukan apakah nilai rata-ratanya benar-benar representatif (mewakili) atau tidak, digunakan untuk mengadakan perbandingan terhadap variabilitas data, dan untuk menentukan apakah dua kelompok data berasal dari populasi yang homogen atau tidak.

Ada dua jenis ukuran penyebaran, yaitu ukuran penyebaran mutlak (*absolute*) merupakan ukuran penyebaran yang digunakan untuk mengetahui tingkat variasi nilai observasi pada suatu data dan ukuran penyebaran relatif merupakan ukuran penyebaran yang digunakan untuk membandingkan tingkat variasi nilai observasi pada suatu data dengan tingkat variasi nilai observasi data-data lainnya. Ukuran penyebaran mutlak (*absolute*) terdiri atas jangkauan/rentang (*range*), simpangan kuartil (*quartile deviation*), simpangan rata-rata (*mean deviation*), dan simpangan baku (standard deviation). Sedangkan yang termasuk ukuran penyebaran relatif adalah koefisien variasi (*coeficient of variation*).

1. Jangkauan/rentang (*range*)

Ukuran penyebaran yang paling sederhana adalah jangkauan/ rentang/range. Dalam sekelompok data kuantitatif akan terdapat data dengan nilai terbesar dan data dengan nilai terkecil. Range dari suatu kelompok data pengamatan adalah selisih antara nilai minimum dan maksimum.

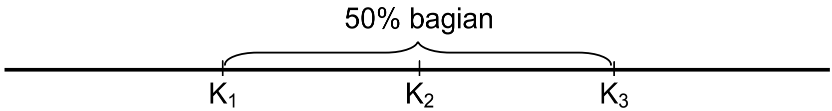
$$\mathbf{R = x_{mak} - x_{min}}$$

keterangan
R = rentangan
x_{mak} = nilai maksimum
x_{min} = nilai minimum

Semakin kecil nilai range, rangkaian data akan semakin homogen, sehingga kualitas data akan semakin baik. Kelebihan dari range adalah dapat dengan mudah dihitung serta mudah dimengerti, sedangkan kekurangannya adalah range tidak didasarkan pada seluruh nilai data tetapi hanya pada dua nilai data saja, yaitu nilai maksimum dan minimum. Range sangat dipengaruhi oleh nilai ekstrim, range sangat dipengaruhi oleh fluktuasi sampel, sehingga sangat tidak stabil atau tidak dapat diandalkan sebagai indikator dari ukuran penyebaran. Range tidak memberikan data yang cukup tentang gambaran variasi distribusi suatu data apalagi data tersebut memiliki nilai ekstrim.

2. Simpangan kuartil (*quartile deviation*)

Simpangan antar kuartil sering disebut dengan rentangan antar kuartil. Simpangan kuartil dihitung dengan cara menghapus nilai-nilai yang terletak di atas kuartil ketiga dan nilai-nilai di bawah kuartil pertama, sehingga nilai-nilai ekstrim, baik yang berada di bawah nilai maksimum ataupun di atas nilai minimum, dihilangkan. Dengan demikian nilai-nilai yang masih ada adalah nilai dari K_1 (kuartil bawah) sampai dengan nilai K_3 (kuartil atas).



Simpangan antar kuartil disimbulkan dengan "RAK" yang merupakan singkatan dari rentangan antar kuartil dan dicari dengan rumus sebagai berikut.

$$\mathbf{RAK = K_3 - K_1}$$

keterangan
RAK = rentangan antar kuartil
K_3 = Kuartil atas
K_1 = kuartil bawah

Dengan tidak mempertimbangkan nilai ekstrim, simpangan antar kuartil lebih stabil dibandingkan dengan range. Namun, simpangan antar kuartil juga tidak memperhatikan dan memperhitungkan penyimpangan semua gugus data seperti halnya range.

Selain RAK ada juga dikenal istilah simpangan semi kuartil atau rentangan semi kuartil (RSK atau SK), yang merupakan setengah dari selisih kuartil atas dengan kuartil bawah atau merupakan setengah RAK. Pagar luar, yaitu kuartil bawah dikurangi dengan semi kuartil dan pagar dalam, yaitu kuartil atas ditambah dengan semi kuartil.

$$\mathbf{SK} = \frac{K_3 - K_1}{2}$$

$$\mathbf{Pd = K_1 - SK}$$

$$\mathbf{Pl = K_3 + SK}$$

keterangan

SK = semi kuartil
K_3 = Kuartil atas
K_1 = kuartil bawah
Pl = pagar luar
Pd = pagar dalam

Contoh 4.7 **Perhatikan kembali Contoh 4.1b, dengan data tunggal 56 68 46 35 75 25 64 72 40 28. Data tunggal tersebut memiliki K_1 = 33,25 , K_2 = 51, dan K_3 = 69. Tentukalah range, simpangan antar kuartil, semi kuartil, pagar dalam dan pagar dalam.**

Penyelesaian

a) Range

$$R = x_{mak} - x_{min} = 75 - 25 = 50$$

b) Simpangan antar kuartil

$$RAK = K_3 - K_1 = 69 - 33{,}25 = 35{,}75$$

c) Semi kuartil

$$SK = \frac{K_3 - K_1}{2} = \frac{69 - 33{,}25}{2} = \frac{35{,}75}{2} = 17{,}875$$

a) Pagar dalam

Pd $= K_1 - SK$

$= 33{,}25 – 17{,}875$

$= 15{,}325$

b) Pagar luar

Pl $= K_3 + KS$

$= 69 + 17{,}875$

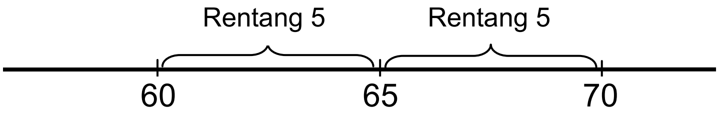
$= 86{,}875$

3. Simpangan rata-rata (*mean deviation*)

Andaikan diberikan segugus data $x_1, x_2, x_3, \ldots, x_n$ dengan rata-rata $\bar{x}$, kita dapat menentukan selisih masing-masing data dengan rata-ratanya, sehingga dapat dibentuk gugusan data baru, yaitu:

$(x_1 - \bar{x}), (x_2 - \bar{x}), (x_3 - \bar{x}), (x_4 - \bar{x}), \ldots, (x_n - \bar{x})$

Dengan melihat urutan tersebut terdapat beberapa kemungkinan selisih nilainya, untuk data yang nilainya di bawah rata-rata selisihnya berupa bilangan negatif, untuk data yang nilainya di atas rata-rata selisinya berupa bilangan positif, sedangkan untuk data yang nilainya sama dengan rata-rata maka nilainya sama dengan nol. Untuk data yang selisihnya negatif dan positif yang perlu kita perhatikan. Seandainya terdapat segugus data dengan rata-rata 65, maka data dengan nilai 60 akan memiliki selisih -5, dan data dengan nilai 70 akan memiliki selisih 5. Padahal jarak atau rentang tidak memperhatikan nilai positif atau negatif. Jadi seharusnya jarak atau rentang dari 65 ke 60 sama dengan jarak atau rentang dari 65 ke 70, yaitu 5.



Berdasarkan ilustrasi tersebut, maka nilai dari $(x_1 - \bar{x})$, $(x_2 - \bar{x})$, $(x_3 - \bar{x})$, $(x_4 - \bar{x})$, …, $(x_n - \bar{x})$ diambil harga mutlaknya, sehingga tidak ada lagi nilai atau jarak yang negatif.

$$|x_1 - \bar{x}|, |x_2 - \bar{x}|, |x_3 - \bar{x}|, ..., |x_n - \bar{x}|$$

Selisih harga mutlak ini bila dijumlahkan dan dibagi dengan banyaknya data atau n disebut dengan simpangan rata-rata (SR).

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

Contoh 4.7 **Diketahui gugusan data tunggal 5 9 4 8 4 7 6 4 8 5 tentukanlah simpangan rata-ratanya.**

Penyelesaian

$$\bar{x} = \frac{\sum_{i=1}^{10} x_1}{n}$$

$$= \frac{5+9+4+8+4+7+6+4+8+5}{10}$$

$$= \frac{60}{10}$$

$$= 6$$

Sehingga simpangan rata-ratanya adalah sebagai berikut.

$$SR = \frac{\sum_{i=1}^{10}\left|x_i - \bar{x}\right|}{n}$$

$$= \frac{|5-6|+|9-6|+|4-6|+|8-6|+|4-6|+|7-6|+|6-6|+|4-6|+|8-6|+|5-6|}{10}$$

$$= \frac{|-1|+|3|+|-2|+|2|+|-2|+|1|+|0|+|-2|+|2|+|-1|}{10}$$

$$= \frac{1+3+2+2+2+1+2+2+1}{10}$$

$$= \frac{16}{10}$$

$$= 1{,}6$$

4. Simpangan baku (*standard deviation*)

Bekerja dengan tanda mutlak atau absolut memiliki kelemahan bila dalam bentuk bilangan negatif. Bisa saja dua gugus data yang memiliki simpangan rata-rata sama tetapi memiliki range atau jangkauan yang berbeda bahkan memiliki nilai ekstrim yang berbeda. Sebagai contoh perhatikan data berikut.

$$\frac{|-1|+|3|+|-2|+|2|+|-2|}{5} = 2$$

Data di atas memiliki nilai minimum -1 dan nilai maksimum 3, dengan demikian rangenya adalah 4. Bandingkan dengan contoh di bawah ini.

$$\frac{|1|+|3|+|2|+|2|+|2|}{5} = 2$$

Data ini memiliki nilai minimum 1 dan nilai maksimum 3, dengan demikian rentangannya adalah 2. Ternyata dari dua gugus data yang memiliki simpangan rata-rata sama tetapi memiliki range yang berbeda. Dari ilustrasi di atas terlihat bahwa simpangan rata-rata tidak dapat membedakan gugusan data yang memiliki rentangan yang berbeda.

Andaikan diberikan segugus data x_1, x_2, x_3, …, x_n dengan rata-rata $\bar{x}$, kita dapat menentukan selisih masing-masing data dengan rata-ratanya kemudian dikuadratkan, sehingga dapat dibentuk gugusan data baru, yaitu:

$$(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, (x_3 - \bar{x})^2, \ldots, (x_n - \bar{x})^2$$

Melihat gugusan data di atas masalah negatif yang ditimbulkan oleh selisih suatu nilai dengan rata-ratanya tidak lagi dijadikan harga mutlak tetapi dikuadratkan. Jika selisih nilai dengan rata-ratanya tersebut dijumlahkan dan dibagi dengan banyaknya data dikurangi satu. Nilai total yang diperoleh kemudian diakarkan, nilai akhir yang diperoleh dikenal dengan simpangan baku. Simpangan baku untuk populasi disimbulkan dengan σ, sedangkan untuk sampel disimbulkan dengan s.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n - 1}$$

$$= \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n - 1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n - 1}}$$

5. Varian (ragam)

Kuadrat dari simpangan baku adalah varian atau ragam. Varians digunakan untuk mengetahui seberapa jauh persebaran nilai hasil observasi terhadap rata-rata. Varian merupakan ukuran penyebaran yang paling sering dipakai dalam statistik. Dengan demikian rumus varian adalah sebagai berikut.

$$s^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}$$

Pertanyaan yang sering muncul adalah mengapa rumus varians dibagi dengan n-1 tidak dibagi n? Ada beberapa ilustrasi dan penjelasan yang bisa diberikan.

Pertama, jika dibagi dengan n, maka rumusnya menjadi $\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n}$ dalam penerapannya nilai yang dihasilkan bisa untuk menduga varian populasi. Dengan menggunakan rumus tersebut varian populasi yang diperoleh nilainya lebih besar dibandingkan dengan nilai varian sampel, padahal harusnya sama. Oleh karena itu agar tidak bias dalam menduga varian populasi, n sebagai pembagi jumlah kuadrat diganti dengan n-1 agar nilai varian sampel mendekati nilai varian populasi. n - 1 ini selanjutnya disebut dengan derajat kebebasan (dk) derajat bebas (db) atau *degree of freedom* (df). Derajat kebebasan tersebut secara sederhana dapat diartikan sebagai jumlah anggota populasi yang masih memiliki kebebasan untuk terpilih dalam batas-batas yang telah ditentukan. Jadi derajat kebebasan tersebut berkaitan dengan peluang untuk memilih. Sebagai ilustrasi jika kita memiliki tiga buah baju A, B, dan C, maka kesempatan kita untuk memilih memakai salah satu baju itu adalah dua kali. Pertama memilih satu dari tiga baju yang tersedia, misalnya terpilih B, kedua memilih satu dari dua baju yang ada, yaitu baju A dan C, misalnya yang terpilih C, maka yang ketiga tidak ada lagi kesempatan untuk memilih karena sudah pasti yang terpilih adalah baju A. Hal ini identik dengan peluang bersyarat yaitu, pengambilan sampel tanpa pengembalian.

Kedua, dengan pembagi n-1 menunjukkan bahwa sampel yang diambil dari populasi tidak boleh sama dengan 1, karena data tunggal yang terdiri dari satu anggota tidak memilliki varian atau variasi (keragaman). Hal ini mengingat rumus varian untuk sampel penyebutnya adalah n - 1, sehingga secara matematis penyebut tidak boleh sama dengan nol atau $n-1\neq 0$, maka $n\neq 1$. Itu sebabnya penyebut dalam varian sama dengan n - 1

Selanjutnya, apabila rumus varian $\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n-1}$ dijabarkan, maka diperoleh rumus lain untuk menentukan varian data tunggal.

$$s^2 = \frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n}\left(x_i^2-2x_i\overline{x}+\overline{x}^2\right)}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} 2x_i\bar{x} + \sum_{i=1}^{n} \bar{x}^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - 2\frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i}{n} + n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n} x_i^2 - 2\frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i}{n} + n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}{n-1}$$

$$= \frac{\frac{n\sum_{i=1}^{n} x_i^2}{n} - 2\frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} + n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2}{n-1}$$

$$= \frac{n\sum_{i=1}^{n} x_i^2 - 2\left(\sum_{i=1}^{n} x_i\right)^2 + \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$$

$$s^2 = \frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)} \text{ sehingga } s = \sqrt{\frac{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}}$$

Contoh 4.8 **Diketahui gugusan data tunggal 8 6 4 7 8 4 5 tentukanlah simpangan baku dan variannya.**

Penyelesaian

$$\bar{x} = \frac{\sum_{i=1}^{7} x_1}{n}$$

$$= \frac{8+6+4+7+8+4+5}{7}$$

$$= \frac{42}{7}$$

$$= 6$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

$$= \sqrt{\frac{(8-6)^2+(6-6)^2+(4-6)^2+(7-6)^2+(8-6)^2+(4-6)^2+(5-6)^2}{7-1}}$$

$$= \sqrt{\frac{2^2+0+(-2)^2+1^2+2^2+(-2)^2+(-1)^2}{6}}$$

$$= \sqrt{\frac{4+4+1+4+4+1}{6}}$$

$$= \sqrt{\frac{18}{6}}$$

$$= \sqrt{3}$$

s = 1,732 sehingga variannya adalah $(1{,}732)^2$ = 2,99 dibulatkan menjadi 3.

Nilai varian dan standar deviasi Contoh 4.8 di atas bisa dihitung dengan menggunkan rumus yang kedua. Sebelum menghitung nilai variannya terlebih dahulu dibuat tabel kerja sebagai berikut.

No	Nilai (x)	x^2
1	8	64
2	6	36
3	4	16
4	7	49
5	8	64
6	4	16
7	5	25
Total	**42**	**270**

$$s = \sqrt{\frac{n\sum_{i=1}^{7} x_i^2 - \left(\sum_{i=1}^{7} x_i\right)^2}{n(n-1)}}$$

$$= \sqrt{\frac{(7 \times 270) - (42)^2}{7(7-1)}}$$

$$= \sqrt{\frac{1890 - 1764}{6 \times 7}}$$

$$= \sqrt{\frac{126}{42}}$$

$$= \sqrt{3}$$

s = 1,732 sehingga variannya adalah $(1{,}732)^2$ = 2,99 dibulatkan menjadi 3.

Dengan demikian hasil yang diperoleh dengan menggunakan kedua rumus di atas sama. Untuk data yang lebih banyak atau tidak dalam bilangan bulat tetapi bilangan desimal hasilnya kemungkinan akan sedikit berbeda karena akibat dari pembulatan.

Untuk data bergolong nilai varian dan standar deviasi dapat dihitung dengan menggunakan rumus berikut ini.

$$s = \sqrt{\frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^2}{n-1}}$$

atau

$$s = \sqrt{\frac{n\sum_{i=1}^{n} f_i x_i^2 - \left(\sum_{i=1}^{n} f_i x_i\right)^2}{n(n-1)}}$$

Contoh 4.9 Tentukanlah varian dan standar deviasi data berikut ini.

No	Kelas interval	Frekuensi
1	30 - 39	4
2	40 - 49	6
3	50 - 59	10
4	60 - 69	15
5	70 - 79	11
6	80 - 89	5
7	90 - 99	2
Total		**53**

Penyelesaian

Terlebih dahulu dibuat tabel kerja sebagai berikut.

No	Kelas interval	f	x_i	$f_i\,x_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
1	30-39	4	34,5	138	-28,68	822,50	3290
2	40-49	6	44,5	267	-18,68	348,91	2093,49
3	50-59	10	54,5	545	-8,68	75,33	753,29
4	60-69	15	64,5	967,5	1,32	1,74	26,17
5	70-79	11	74,5	819,5	11,32	128,16	1409,75
6	80-89	5	84,5	422,5	21,32	454,57	2272,87
7	90-99	2	94,5	189	31,32	980,99	1961,98
Total		**53**		**3348,5**			**11807,55**

$$\bar{x} = \frac{\sum_{i=1}^{7} f_i\,x_i}{\sum_{i=1}^{7} f_i}$$

$$= \frac{3348,5}{53}$$

$$= 63,18$$

$$s = \sqrt{\frac{\sum_{i=1}^{7}\left(x_i - \bar{x}\right)^2}{n-1}}$$

$$= \sqrt{\frac{11807,55}{53-1}}$$

$$= \sqrt{\frac{11807,55}{52}}$$

$$= \sqrt{227,1049}$$

= 15,07 sehingga varianya adalah $(15,07)^2$ = 227,1049

6. Koefisien variasi (*coeficient of variation*)

Koefisien variasi (CV) merupakan ukuran dispersi relatif yang digunakan untuk membandingkan variasi dua atau lebih kelompok data. Koefisien variasi merupakan standar ukuran penyebaran dari distribusi frekuensi. Nilai dari koefisien variasi dinyatakan dalam persentase. Nilai absolut atau harga mutlak dari CV ini kadang-kadang dikenal sebagai standar deviasi relatif (RSD). Koefisien variasi (CV) merupakan rasio dari standar deviasi (s) dengan rata-rata ($\bar{x}$), sehingga dapat diformulasikan sebagai berikut.

$$CV = \frac{s}{\bar{x}} \times 100\%$$

keterangan

CV = koefisien variasi

s = standar deviasi

$\bar{x}$ = rata-rata

Contoh 4.10 **Terdapat dua kelas yang mengikuti ujian statistik. Kelas A memperoleh nilai rata-rata 85 dengan standar deviasi 12, kelas B memperoleh nilai rata-rata 70 dengan standar deviasi 8. Manakah kualitas yang lebih baik apakah hasil ujian kelas A atau hasil ujian kelas B.**

Penyelesaian

Untuk menentukan mana kualitas nilai ujian yang lebih baik adalah dengan menghitung nilai koefisien variasi masing masing kelas. Kualitas yang dimaksud adalah keseragaman nilai yang diperoleh.

$$CV_A = \frac{s_A}{\bar{x}_A} x100\%$$

$$= \frac{12}{85} x\ 100\%$$

$$= 14{,}12\%$$

$$CV_B = \frac{s_B}{\bar{x}_B} x100\%$$

$$= \frac{8}{70} x\ 100\%$$

$$= 11{,}43\%$$

Nilai ujian kelas A memiliki koefisien variasi yang lebih kecil daripada nilai ujian kelas B. Dengan demikian, kualitas atau keseragaman nilai kelas B lebih baik atau lebih seragam dibandingkan nilai ujian kelas A. Dapat dikatakan bahwa kualitas nilai ujian kelas B lebih baik daripada nilai ujian kelas A. semakin kecil nilai koefisien variasi maka semakin seragam nilai yang diperoleh, begitu juga sebaliknya.

Latihan 4

1. Diketahui data tunggal sebagai berikut.

   a. 9 8 6 7 6 7 9 5
   b. 1 8 6 3 4 5 9 5 7 9 2 8
   c. 4 8 6 5 9 4

   Tentukanlah K_1, K_2, K_3, D_3, D_7, D_9, P_{11}, P_{35}, P_{78}, P_{98} dan ukuran penyebaran data tersebut
2. Dari soal No. 1 manakah kualitas nilai yang paling baik, jelaskan?
3. Terdapat dua buah jenis AC, X dan Y. AC jenis X rata-rata menyala 2000 jam dengan standar deviasi 156 dan AC jenis Y rata-rata menyala 3125 jam dengan standar deviasi 98. Tentukanlah AC mana yang memiliki kualitas lebih baik?
4. Tentukanlah ukuran letak (K_1, K_2, K_3, D_3, D_7, D_9, P_{11}, P_{35}, P_{78}, P_{98}) dan ukuran penyebaran data berikut.

No	**Kelas interval**	**f**
1	19 - 29	10
2	20 - 29	15
3	30 - 39	20
4	40 - 49	25
5	50 - 59	32
6	60 - 69	24
7	70 - 79	18
8	80 - 89	10
9	90 - 99	6
Total		**160**

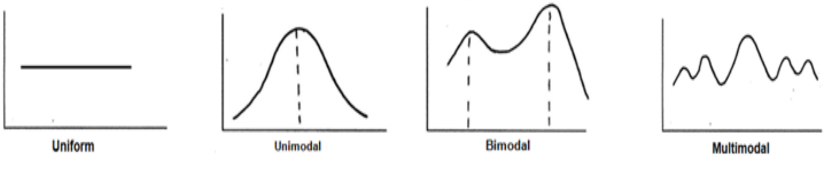
5. Buatlah sebaran data tunggal sebanyak 100 data, kemudian buatlah tabel distribusi data bergolong dengan kelas interval. Dari tabel distribusi tersebut tentukanlah ukuran letak dan ukuran penyebarannya. Apa yang dapat anda simpulkan jika data tadi ukuran letak dan ukuran penyebarannya dibandingkan dengan ukuran letak dan ukuran penyebaran yang dihitung dari sebaran data tunggal.

BAB V
UKURAN KEMIRINGAN (*SKEWNESS*) DAN UKURAN KERUNCINGAN (*KURTOSIS*)

A. Ukuran Kemiringan (*Skewness*)

Pahami kembali BAB III tentang hubungan hubungan empirik antara mean, median, dan modus. Berdasarkan hubungan empirik ketiga ukuran pemusatan tersebutlah kita akan menentukan kemiringan suatu kurva berdasarakan distribusi datanya. Ukuran kemiringan merupakan ukuran atau derajat dari ketidaksimetrisan (asimetri) suatu distribusi data.

Jika data tidak memiliki modus, maka kurvanya berbentuk garis lurus dan tidak memiliki puncak. Apabila mean data sama dengan mediannya atau hampir sama dan tidak memiliki modus maka data disebut berdistribusi *uniform*. Distribusi *uniform* secara sederhana dapat kita lihat ketika segugus data nilainya sama, sehingga mean dan mediannya sama, tetapi data tersebut tidak memiliki modus. Jika data memiliki satu modus (unimodal), maka kurvanya memiliki satu puncak puncak. Jika data memiliki dua modus (bimodal), maka kurvanya memiliki dua puncak puncak. Jika data memiliki modus lebih dari dua modus (multimodal), maka kurvanya memiliki lebih dari dua puncak puncak.



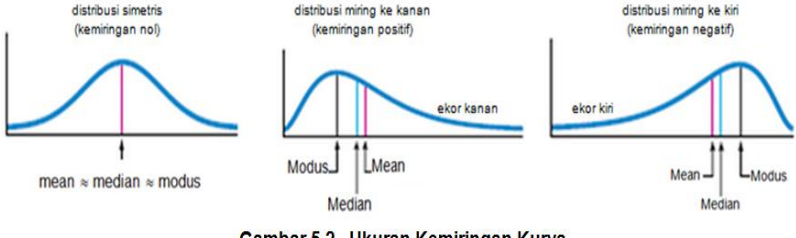
Gambar 5.1 Jenis Kurva Berdasarkan Banyaknya Modus

Selanjutnya kurva distribusi data yang unimodal atau hanya memiliki satu modus kemiringan kurvanya dapat dibagi menjadi tiga, yaitu simetris, miring ke kiri (negatif) dan miring ke kanan (positif).

Kurva simetris (kemiringan nol) menunjukkan letak nilai rata-rata, median, dan modus hampir berimpit (berkisar disatu titik), atau mungkin sama. Kurva yang simetris selanjutnya akan disebut dengan kurva normal. Kurva normal disebut dengan kurva *Gauss*, karena bentuk kurva ini dikenalkan pertama kali oleh Fredrich Gauss.

Kurva miring ke kanan (juling ke kanan) mempunyai nilai rata-rata paling besar dan modus paling kecil. Kurva ini dalam distribusi normal disebut dengan ekor kanan. Kurva yang miring ke kanan memiliki ekor yang lebih panjang ke kanan daripada yang ke ke kiri. Masa distribusi berkumpul di bagian kiri kurva artinya nilai-nilai yang lebih kecil dalam gugusan data memiliki frekuensi yang lebih banyak. Semakin besar nilainya makan semakin kecil frekuensinya.

Kurva miring ke kiri (juling ke kiri) mempunyai nilai modus paling besar dan rata-rata hitung paling kecil. Kurva ini dalam distribusi normal disebut dengan ekor kiri. Kurva yang miring ke kiri memiliki ekor yang lebih panjang ke kiri daripada yang ke kanan. Masa distribusi berkumpul di bagian kanan kurva artinya nilai-nilai yang lebih kecil dalam gugusan data memiliki frekuensi yang lebih sedikit. Semakin besar nilainya makan semakin besar pula frekuensinya.



Gambar 5.2 Ukuran Kemiringan Kurva

Untuk mengukur derajat kemiringan suatu distribusi dinyatakan dengan koefisien kemiringan (*skewness coefficient*). Ada tiga metode yang bisa digunakan untuk menghitung koefisien *skewness*, yaitu; koefisien kemiringan dari Pearson, koefisien kemiringan Bowley, koefisien kemiringan persentil, dan koefisien kemiringan moment.

1. Koefisien Kemiringan Pearson

Koefisien kemiringan dari Pearson menggunakan ukuran pemusatan data dan ukuran penyebaran data. Koefisien kemiringannya merupakan selisih antara nilai modus dan rata-rata yang hasilnya dibagi dengan standar deviasinya. Koefisien kemiringan Pearson disimbolkan dengan *sk* (*skewness coefficient)* dan dirumuskan sebagai berikut

$$sk = \frac{\bar{x} - Mo}{s}$$

Pada bagian sebelumnya telah diberikan hubungan empiris antara mean, median dan modus, yaitu:

$$\text{mean} - \text{modus} = 3\,(\text{mean} - \text{median})$$

atau

$$\bar{x} - Mo = 3\,(\bar{x} - Me)$$

Apabila Mo disubtitusi dengan $Mo = \bar{x} - 3\,(\bar{x} - Me)$, maka rumus sk akan menjadi seperti berikut ini

$$sk = \frac{3(\bar{x} - Me)}{s}$$

keterangan

sk = koefisien kemiringan

$\bar{x}$ = rata-rata

Me = median

Mo = modus
s = standar deviasi

Menurut Pearson nilai sk dapat diinterpretasikan ke dalam tiga bagian, yaitu *sk* = 0, *sk* > 0, dan *sk* < 0, titik acuan nilai *sk* adalah 0.

a) Jika *sk* = 0, maka kurva dikatakan memiliki bentuk simetris kemiringan 0. Kemungkinan ini terjadi ketika data memiliki nilai mean = modus = median.
b) Jika *sk* > 0, maka kurva dikatakan miring ke kanan (juling positif). Nilai-nilai data terkonsentrasi pada sisi sebelah kanan sehingga lebih gemuk di bagian kanan kurva, rata-rata (mean) terletak di sebelah kanan Mo dengan ketentuan median terletak di tengah-tegah, kurva memiliki ekor memanjang ke kanan.
c) Jika *sk* < 0, maka kurva dikatakan miring ke kiri (juling negatif). Nilai-nilai data terkonsentrasi pada sisi sebelah kiri sehingga lebih gemuk di bagian kiri kurva, rata-rata (mean) terletak di sebelah kiri Mo dengan ketentuan median terletak di tengah-tengah, kurva memiliki ekor memanjang ke kiri.

Dengan melihat ketentuan nilai *sk* sebenarnya rumus dari Pearson tergantung dari nilai $\bar{x}$ - Mo, apakah nilainya 0, positif, atau negatif. Nilai s tidak akan mengubah positif atau tidaknya nilai *sk*, karena nilai s selalu positif. Jadi rumus dari Pearson bisa disederhanakan lagi, jika hanya ingin melihat ukuran kemiringan apakah 0, positif, atau negatif, tetapi jika ingin melihat nilai koefisiennya atau derajat kemiringannya maka harus melibatkan nilai s.

2. Koefisien kemiringan Bowley

Koefisien kemiringan dari Pearson menggunakan ukuran pemusatan data dan ukuran penyebaran data, sedangkan koefisien kemiringan dari Bowley menggunakan nilai-nilai kuartil atas (K_3), tengah (K_2), dan bawah (K_1).

$$sk_B = \frac{(K_3 - K_2) - (K_2 - K_1)}{(K_3 - K_2) + (K_2 - K_1)} \quad \text{atau} \quad sk_B = \frac{K_3 - 2K_2 + K_1}{K_3 - K_1}$$

Keterangan

sk_B = koefisien kemiringan Bowley

K_1 = kuartil bawah

K_2 = kuartil tengah atau median

K_3 = kuartil atas

Interpretasi nilai sk_B sama halnya dengan interprestasi *sk* dari Pearson yang telah dijelaskan di atas. Ada beberapa hal yang mungkin terjadi dari analisis sk_B, yaitu: jika nilai $K_3 - K_2 > K_2 - K_1$, maka distribusi akan menceng ke kanan atau juling positif, $K_3 - K_2 < K_2 - K_1$, maka distribusi akan menceng ke kiri atau juling negatif, Jika $K_3 - K_2 = K_2 - K_1$ atau $K_3 + K_1 – 2K_2 = 0$, maka distribusi datanya simetri, Jika $K_1 = K_2$ maka distribusi datanya miring ke kanan, dan Jika $K_2 = K_3$ maka distribusi datanya miring ke kiri. Lebih jauh Bowley berpendapat bahwa $sk_B = \pm 0{,}10$ menggambarkan distribusi yang menceng secara tidak berarti. sedangkan, jika $sk_B > 0{,}30$ menggambarkan distribusi yang menceng secara berarti sekali.

3. Koefisien kemiringan persentil

Koefisien Kemencengan Persentil didasarkan atas hubungan antar persentil (P_{90}, P_{50} dan P_{10}) dari sebuah distribusi. Koefisien kemencengan dengan menggunakan persentil diformulasikan sebagai berikut.

$$sk_P = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{(P_{90} - P_{50}) + (P_{50} - P_{10})} \quad \text{atau} \quad sk_P = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}$$

keterangan

sk_P = koefisien kemiringan persentil

P_{90} = persentil ke-90

P_{50} = persentil ke-50

P_{10} = persentil ke-10

Interpretasi nilai sk_P sama halnya dengan interprestasi *sk* dari Pearson yang telah dijelaskan di atas.

4. Koefisien kemiringan moment

Rumus momen merupakan pengembangan dari rumus rata-rata. Sedangkan rumus koefisien kemiringan momen didasarkan pada perbandingan momen ke-3 dengan pangkat tiga simpang baku. Koefisien kemiringan momen dilambangkan dengan α_3. Koefisien kemiringan momen disebut juga kemencengan relatif. Sebelum melangkah lebih lanjut menghitung koefisien kemiringan momen terlebih dahulu kita lihat pengembangan rumus momen sebagai berikut.

Misalkan diberikan variabel x dengan harga-harga: x_1, x_2, x_3, …, x_n. Jika A = sebuah bilangan tetap dan r = 0, 1, 2, …, n, maka momen ke-r sekitar A, disingkat Mr didefinisikan oleh hubungan :

$$M_r = \frac{\sum_{i=1}^{n} (x_i - A)^r}{n}$$

Untuk A = 0 didapat momen ke-r sekitar nol atau disingkat momen ke-r, sehingga

$$M_r = \frac{\sum_{i=1}^{n} x_i^{\,r}}{n}$$

untuk r = 1 maka momen pertama merupakan rata-rata. Jika A diganti dengan $\bar{x}$, maka formula di atas akan menjadi sebagai berikut.

Momen data tunggal

$$M_r = \frac{\sum_{i=1}^{n} (x_i - A)^r}{n}$$

Momen data bergolong

$$M_r = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^r}{n}$$

a) Jika r = 1 yang disebut dengan momen pertama, maka M_1 merupakan rumus mean atau rata-rata
b) Jika r = 2 yang disebut dengan momen kedua, maka M_2 merupakan rumus varian populasi
c) Jika r = 3 yang disebut dengan momen ketiga, maka M_3 merupakan rumus kemiringan
d) Jika r = 4 yang disebut dengan momen keempat, maka M_4 merupakan rumus keruncingan (kurtosis) yang akan dibahas pada bagaian berikutnya.

Berdasarkan ketentuan di atas, rumus menghitung koefisien kemiringan momen adalah:

Untuk data tunggal

$$\alpha_3 = \frac{M_3}{s^3} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{ns^3}$$

Untuk data bergolong

$$\alpha_3 = \frac{M_3}{s^3} = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^3}{ns^3}$$

Jika menentukan atau menghitung koefisien kemiringan momen menggunakan coding seperti menghitung mean atau rata-rata data bergolong, maka rumusnya menjadi seperti berikut ini

$$\alpha_3 = \left(\frac{p}{s}\right)^3 \left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i^3 - 3\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i^2\right)\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i\right) + 2\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i\right)^3\right)$$

keterangan

α_3 = koefisien keruncingan
M_3 = momen ketiga
s = simpangan baku
n = banyaknya data pengamatan
k = banyaknya kelas
p = panjang kelas interval
f_i = frekuensi kelas ke-i
d_i = perbedaan kelas ke-i terhadap titik asal asumsi
$\bar{x}$ = rata-rata hitung atau mean

Contoh 5.1 Tentukanlah kemiringan kurva data berikut ini.
7 6 5 7 6 7 9 8 7 8

Penyelesaian

Terlebih dahulu data diurutkan 5 6 6 7 7 7 7 8 8 9, kemudian dihitung nilai-nilai yang diperlukan diantaranya sebagai berikut.

a. Rata-rata (mean)

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{n}$$

$$= \frac{7+6+5+7+6+7+9+8+7+8}{10}$$

$$= \frac{70}{10}$$

$= 7$

b. median (Me)

$$L_{me} = \frac{(n+1)}{2}$$

$$= \frac{(10+1)}{2}$$

$$= \frac{11}{2}$$

$= 5\frac{1}{2}$ (terletak diantara data ke-6 dan data ke-5)

Me = data ke-5 + ½ (data ke-6 – data ke-5)

= 7 + ½ (7 - 7)

= 7

c. Modus

Mo = 7

d. Kuartil Bawah (K_1)

$$L_{k1} = \frac{(n+1)}{4}$$

$$L_{k1} = \frac{(n+1)}{4}$$

$$= \frac{(10+1)}{4}$$

$$= \frac{11}{4}$$

$$= 2\frac{3}{4}$$ (terletak diantara data ke-6 dan data ke-5)

K_1 = data ke-2 + ¾ (data ke-3 – data ke-2)

= 6 + ¾ (6 - 6)

= 6

e. Kuartil Atas (K_3)

$$L_{k3} = \frac{3(n+1)}{4}$$

$$= \frac{3(10+1)}{4}$$

$$= \frac{33}{4}$$

$$= 8\frac{1}{4}$$ (terletak diantara data ke-9 dan data ke-8)

K_3 = data ke-8 + ¾ (data ke-9 – data ke-8)

= 8 + ¾ (8 - 8)

= 8

f. Standar deviasi (s)

Terlebih dahulu dibuat tabel kerja

No	Nilai (x)	$x - \bar{x}$	$(x - \bar{x})^2$
1	5	-2	4
2	6	-1	1
3	6	-1	1
4	7	0	0
5	7	0	0
6	7	0	0
7	8	1	1
8	8	1	1
9	9	2	4
10	7	0	0
Total	**70**	**-**	**12**

$$s = \sqrt{\frac{\sum_{i=1}^{10}\left(x_i - \bar{x}\right)^2}{n - 1}}$$

$$= \sqrt{\frac{12}{10 - 1}}$$

$$= \sqrt{\frac{12}{9}}$$

$$= \sqrt{1{,}33}$$

$$= 1{,}153$$

Berdasarkan nilai-nilai mean ($\overline{x}$), median (Me), Modus (Mo), kuartil bawah (K_1), kuartil atas (K_3) dan varian (s), akan dihitung koefisien kemiringan data tunggal 5 6 6 7 7 7 7 8 8 9 dengan menggunakan formula dari Pearson dan formula dari Bowley.

Formula dari Pearson

$$sk = \frac{x - Mo}{s}$$

$$= \frac{7-7}{1{,}155}$$

$$= \frac{0}{1{,}155}$$

$$= 0$$

Formula dari Bowley

$$sk_B = \frac{K_3 - 2K_2 + K_1}{K_3 - K_1}$$

$$= \frac{8-(2.7)+6}{8-6}$$

$$= \frac{8-14+6}{2}$$

$$= \frac{0}{2}$$

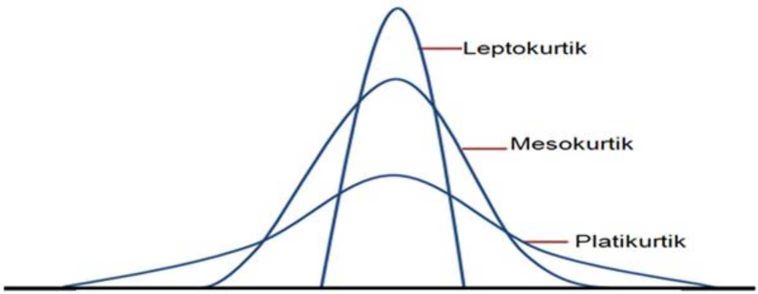
$$= 0$$

Ternyata dengan menggunakan formula dari Pearson sama hasilnya dnegan menggunakan formula dari Bowley, yaitu ditemukan nilai $sk = sk_B = 0$. Jadi, jika $sk = 0$, maka kurva dikatakan memiliki bentuk simetris (kemiringan 0). Ini terjadi ketika data memiliki nilai mean = modus = median. Untuk menghitung koefisien kemiringan dengan menggunakan formula persentil dan momen diserahkan kepada pembaca.

B. Ukuran Keruncingan (*Kurtosis*)

Selain ukuran kemiringan kurva dari gugusan data atau distribusi data dikenal juga istilah keruncingan kurva atau kurtosis. Istilah keruncingan digunakan untuk kurva-kurva yang simetris (kemiringan nol). Keruncingan erat kaitannya dengan ketinggian kurva. Keruncingan (kurtosis) adalah derajat kepuncakan (keruncingan) kurva suatu distribusi data. Ukuran atau derajat keruncingan suatu kurva ditentukan dari kurva normal standar. Derajat keruncingan tersebut relatif terhadap distribusi normal data bersangkutan. Jadi keruncingan suatu kurva distribusi data adalah derajat atau ukuran tinggi rendahnya puncak kurva terhadap distribusi normal.

Dengan melihat penjelasan di atas yang menyatakan bahwa keruncingan tersebut berkaitan dengan tinggi rendahnya kurva, kemungkinan kurva tersebut adalah menjulang tinggi (runcing), normal, datar atau tumpul. Berdasarkan keruncingannya, kurva distribusi dapat dibedakan atas tiga macam, yaitu: leptokurtik, platikurtik, dan mesokurtik. Kurva *leptokurtik* merupakan kurva simetris distribusi data yang memiliki puncak relatif tinggi, biasanya diperoleh dari distribusi data yang memiliki rentang yang kecil. Kurva *platikurtik* merupakan kurva simetris distribusi data yang memiliki puncak relatif datar (rendah). Kurva *mesokurtik* merupakan kurva distribusi data yang memiliki puncak normal (tidak terlalu tinggi dan tidak terlalu rendah). Untuk lebih memahami jenis-jenis keruncingan kurva, perhatikan gambar berikut.



Gambar 5.3 Ukuran Keruncingan Kurva

Ukuran yang sering digunakan untuk mengetahui keruncingan kurva suatu distribusi data adalah koefisien keruncingan. Untuk mengukur derajat atau koefisien keruncingan kurva suatu distribusi dinyatakan dengan koefisien kemiringan (*coefficient kurtosis*). Untuk menghitung koefisien *kurtosis* bisa menggunakan nilai-nilai kuartil dan persentil atau menggunakan formula momen ke-4, yaitu *moment coefficient of kurtosis*.

1. Koefisien keruncingan kuartil dan persentil

Untuk mengetahui kurva distribusi data apakah leptokurtik, mesokurtik, atau platikurtik dapat digunakan rumus di bawah ini.

$$kk = \frac{\frac{1}{2}(K_3 - K_1)}{P_{90} - P_{10}}$$

keterangan

kk = koefisien kurtosis

K_1 = kuartil bawah

K_3 = kuartil atas

P_{10} = persentil ke-10

P_{90} = persentil ke-90

Kreteria pengujiannya *kk* adalah: jika *kk* < 0,263; maka kurva distribusinya adalah platikurtik, jika *kk* = 0,263; maka kurva distribusinya adalah mesokurtik, dan jika *kk* > 0,263; maka kurva distribusinya adalah leptokurtik.

2. Koefisien kemiringan momen ke-4

Untuk data tunggal

$$\alpha_4 = \frac{M_4}{s^4} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4}{ns^4}$$

Untuk data bergolong

$$\alpha_4 = \frac{M_4}{s^4} = \frac{\sum_{i=1}^{k} f_i\left(x_i - \bar{x}\right)^4}{ns^4}$$

Jika menentukan atau menghitung koefisien keruncingan momen menggunakan coding seperti menghitung mean atau rata-rata data bergolong, maka rumusnya menjadi seperti berikut ini

$$\alpha_4 = \left(\frac{p}{s}\right)^4 \left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i^4 - 4\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i^3\right)\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i\right) + 6\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i^2\right)\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i\right) - 3\left(\frac{1}{n}\sum_{i=1}^{k} f_i d_i\right)^4\right)$$

keterangan

α_4 = koefisien keruncingan
M_4 = momen keempat
s = simpangan baku
n = banyaknya data pengamatan
k = banyaknya kelas
p = panjang kelas interval
f_i = frekuensi kelas ke-i
d_i = perbedaan kelas ke-i terhadap titik asal asumsi
$\bar{x}$ = rata-rata hitung atau mean

Kreteria pengujiannya α_4 adalah jika $\alpha_4 < 3$; maka kurva distribusinya adalah platikurtik, jika $\alpha_4 = 3$; maka kurva distribusinya adalah mesokurtik, dan jika $\alpha_4 > 3$; maka kurva distribusinya adalah leptokurtik.

Contoh 5.2 **Hitunglah koefisien kemiringan data** 5 6 6 7 7 7 7 8 8 9 **seperti pada Contoh 5.1.**

Penyelesaian

Dari penyelesaian Contoh 5.1 diperoleh nilai-nilai $K_1 = 6$ dan nilai $K_3 = 8$, jadi yang perlu dihitung adalah nilai P_{10} dan P_{90}.

$$L_{P10} = \frac{10(n+1)}{100}$$

$$= \frac{10(10+1)}{100}$$

$$= \frac{110}{100}$$

$= 1\frac{1}{10}$ (terletak diantara data ke-1 dan data ke-2)

P_{10} = data ke-1 + $\frac{1}{10}$ (data ke-2 – data ke-1)

$= 5 + \frac{1}{10}(6 - 5)$

$= 5\frac{1}{10}$

$= 5{,}1$

$L_{P90} = \frac{90(n+1)}{100}$

$= \frac{90(10+1)}{100}$

$= \frac{990}{100}$

$= 9\frac{9}{10}$ (terletak diantara data ke-9 dan data ke-10)

P_{90} = data ke-9 + $\frac{9}{10}$ (data ke-10 – data ke-9)

$= 8 + \frac{9}{10}(9 - 8)$

$= 8\frac{9}{10}$

$= 8{,}9$

$$kk = \frac{\frac{1}{2}\left(K_3 - K_1\right)}{P_{90} - P_{10}}$$

$$= \frac{\frac{1}{2}\left(8 - 6\right)}{8{,}9 - 5{,}1}$$

$$= \frac{\frac{1}{2}\left(2\right)}{3{,}8}$$

$$= \frac{1}{3{,}8}$$

$$= 0{,}261$$

Jadi kurva distribusi data tergolong platikurtik (0,261 < 0,263). Jika distribusi data tersebut keruncingannya dicari dengan rumus momen ke 4, maka hasilnya sebagai berikut.

No	**Nilai (x)**	**$x - \bar{x}$**	**$(x-\bar{x})^2$**	**$(x-\bar{x})^4$**
1	5	-2	4	16
2	6	-1	1	1
3	6	-1	1	1
4	7	0	0	0
5	7	0	0	0
6	7	0	0	0
7	8	1	1	0
8	8	1	1	1
9	9	2	4	1
10	7	0	0	16
Total	**70**	**-**	**12**	**36**

$$\alpha_4 = \frac{M_4}{s^4} = \frac{\sum_{i=1}^{10}\left(x_i - \bar{x}\right)^4}{ns^4}$$

$$= \frac{36}{10.(1,154)^4}$$

$$= \frac{36}{10.1,773}$$

$$= \frac{36}{17,73}$$

$= 2,030$ (termasuk platikurtik karena $\alpha_4 < 3$)

Latihan 5

1. Apa yang dimaksud dengan kemiringan dan keruncingan kurva, apa hubungan di antara kedua istilah tersebut?
2. Diketahui data 5 6 8 7 9 5 8 7 6 7 5 7 9 2 3 4. Tentukanlah momen pertama, momen kedua, momen ketiga, dan momen keempat.
3. Dari soal No. 2 tentukan kemiringan dan keruncingannya, gambarkan kurvanya!
4. Perhatikan tabel di bawah ini

No	Kelas Interval	f
1	20 - 24	4
2	25 - 29	7
3	30 - 34	11
4	35 - 39	15
5	40 - 44	10
6	45 - 49	6
7	50 - 54	3
Total		**56**

Tentukanlah momen pertama, momen kedua, momen ketiga, dan momen keempat

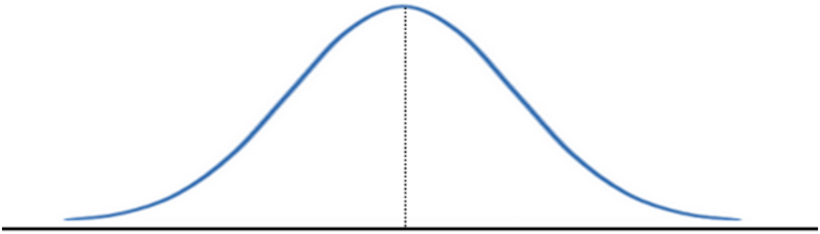
5. Dari soal No. 4 tentukan kemiringan dan keruncingannya, gambarkan kurvanya!

BAB VI
KURVA NORMAL

Carl Friedrich Gauss (1777-1855) atau yang lebih dikenal dengan sebutan *Gauss* seorang ahli matematika berkebangsaan Jerman, menemukan persamaan sebuah distribusi (sebaran) pada saat meneliti galat dalam pengukuran yang berulang-ulang mengenai bahan yang sama. Distribusi tersebut dikenal dengan nama distribusi normal. Galat yang dimaksud adalah galat pengukuran yang dilakukan oleh manusia karena tidak mampu mengukur dengan tepat apa yang diamati. Untuk menghormati jasanya, distribusi normal sering disebut sebagai distribusi galat *Gauus* (*gaussian distribution*). Sebenarnya pioneer dari distribusi normal ini adalah DeMoivre (1733) seorang ahli matematika berkebangsaan Perancis, yang berhasil menurunkan persamaan matematis bagi distribusi normal ini.

Distribusi normal merupakan salah satu distribusi yang berasal dari distribusi peubah acak kontinu. Merupakan suatu alat statistik yang sangat penting untuk menaksir dan meramalkan peristiwa-peristiwa yang lebih luas. Grafik dari distribusi normal disebut dengan kurva normal. Kata normal bukan merupakan bahasa inggris yang diartikan sebagai *ordinary*, *common*, *normal*, atau *natural.* Bukan juga dalam istilah kedokteran sebagai arti tidak gila, tidak stres atau tidak sakit, tetapi merupakan suatu model matematik yang menggambarkan penyebaran probabilitas dari pengamatan yang tidak terbatas dan diukur terus menerus. Banyak hasil dan teknik analisis statistik hanya bisa berfungsi dengan benar jika model distribusinya berupa distribusi normal, terutamanya penggunaan statistik parametrik.

Distribusi normal merupakan cerminan semua kejadian di alam ini, sehingga dikatakan bahwa distribusi normal terjadi secara alamiah. Banyak peristiwa di dunia nyata menuruti distribusi normal, seperti: orang pintar, orang kaya, berat badan, tinggi badan manusia, hewan dll. Distribusi normal merupakan model distribusi kontinu yang paling penting untuk diterapkan di berbagai bidang seperti industri dan penelitian. kurva normal dapat divisualisasikan seperti pada Gambar 6.1 berikut ini.



Gambar 6.1 Kurva Normal

Suatu peubah acak kontinu X yang memiliki distribusi seperti Gambar 6.1 di atas disebut peubah acak normal yang nilai dan distribusinya disebut dengan distribusi normal. Segugus data membentuk distribusi normal bila jumlah data di atas dan di bawah mean relatif sama. Secara matematis kurva normal dapat didefinisikan sebagai berikut. Bila X adalah suatu peubah acak normal dengan rata-rata μ dan standar deviasi σ, maka persamaan kurva normalnya adalah

$$n\,(x;\ \mu\,,\ \sigma\,) = \frac{1}{\sqrt{2\pi\sigma}}\,e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \text{ untuk } -\infty < x < \infty$$

dengan nilai π = 3,14159... merupakan sebuah konstanta yang merupakan perbandingan atau rasio antara keliling lingkaran dengan panjang diameternya dan e = 2,71828... adalah dasar dari logaritma natural. Ciri-ciri dari kurva normal adalah:

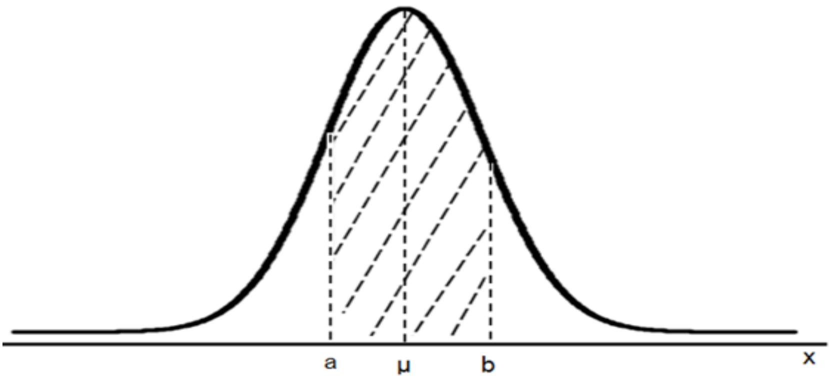
a) Bentuk kurva normal menyerupai bentuk lonceng atau genta (*bell*). Kurva normal merupakan suatu poligon yang dihaluskan atau dilicinkan dengan ordinat memuat frekuensi dan absis memuat nilai μ dan σ.
b) Selalu berada di atas sumbu X, tidak pernah momotong sumbu X karena sebagai asimtot.
c) Simetris terhadap nilai $x = \mu$, oleh karena distribusi normal bersifat simetris dan menyatakan luas daerah, maka tabel distribusi normal dibuat hanya untuk menghitung bagian sebelah kanan mean dari distribusi tersebut. Untuk menghitung nilai di sebelah kiri, nilai yang negatif dianggap sama dengan nilai positif di sebelah kanan.
d) Kurva mempunyai titik belok pada $x = \mu \pm \sigma$, cekung dari bawah bila $\mu - s < x < \mu + s$ dan cekung dari atas untuk nilai x lainnya.
e) Luas di bawah kurva normal dan di atas sumbu X merupakan satu satuan luas persegi atau 100%. Karena kurva normal simetris berbentuk lonceng dan unimodal, maka daerah di kanan dan di kiri garis tegak lurus di atas rata-rata masing-masing besarnya 0,5 atau 50%. Luas daerah di bawah kurva normal dapat dihitung dengan cara mengintegralkan persamaan di atas pada selang $-\infty$ sampai dengan $+\infty$ yang merupakan luas di bawah kurva normal.

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

Untuk menentukan luas daerah antara selang a dan b di bawah kurva normal, dilakukan dengan cara mengintegralkan yang ditafsirkan sebagai probabilitas peubah acak X antara x = a dan x = b.

$$P(a < X < b) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

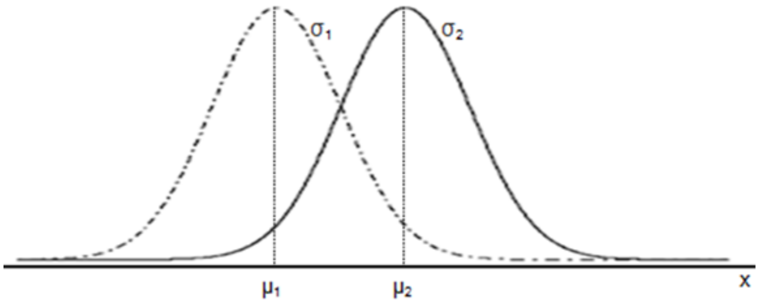
Untuk nilai-nilai μ dan σ yang diketahui, hasil integralnya berkisaran antara 0 dan 1. Mencari luas daerah di bawah kurva normal adalah dengan mengintegralkan kurva (daerah yang diarsir) sepanjang selang a dan b, seperti gambar 6.2 berikut ini.



Gambar 6.2 Luas Daerah di Bawah kurva Normal Pada Selang a dan b

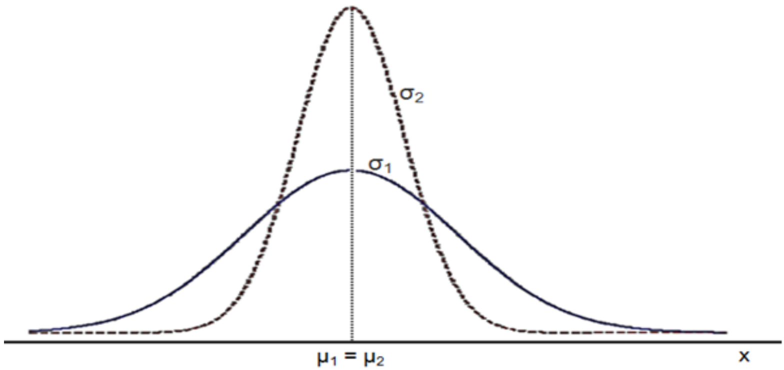
Bila nilai μ dan σ diketahui, maka kurva normal tersebut dapat diketahui dengan pasti. Bila μ = 67 dan σ =

45, maka ordinat-ordinatnya adalah n(x; 67, 45) nilai x dapat dihitung dan grafiknya dapat digambarkan. Nilai rata-rata (μ) dan standar deviasi (σ) yang disebut dengan parameter memberikan kemungkinan pembuatan kurva normal menjadi tidak terbatas, yaitu dengan menghubungkan kedua parameter ini. Dengan demikian kurva normal dengan $\mu = 67$ dan $\sigma = 45$ merupakan salah satu bentuk atau anggota keluarga dari sekian banyak pola distribusi kurva normal. Pada Gambar 6.3 di bawah ini diberikan dua buah kurva normal yang memiliki rata-rata berbeda tetapi standar deviasi yang sama.



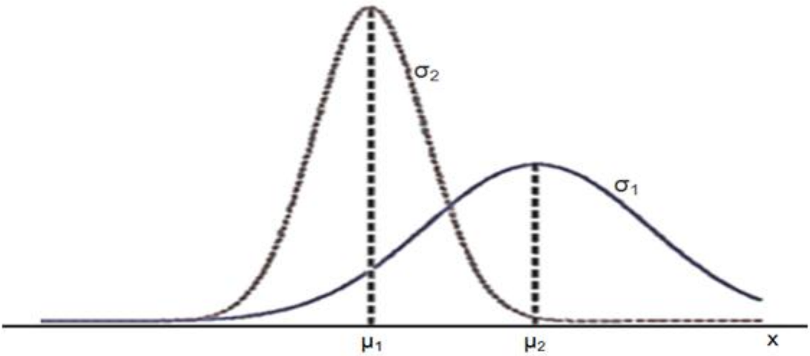
Gambar 6.3 Dua Kurva Normal Dengan $\mu_1 \neq \mu_2$ dan $\sigma_1 = \sigma_2$

Pada Gambar 6.3 terlihat bahwa kedua kurva normal memiliki bentuk yang sama, mulai dari tingginya sampai dengan lebarnya. Namun, terlihat bahwa kedua kurva tidak pernah menyentuh sumbu datar (sumbu X), perbedaanya terletak pada pusatnya, kedua kurva memiliki pusat yang berbeda pada sumbu datarnya. Selanjutnya pada Gambar 6.4 diberikan dua buah kurva normal yang memiliki rata-rata sama tetapi memiliki standar deviasi yang berbeda.



Gambar 6.4 Dua Kurva Normal Dengan $\mu_1 = \mu_2$ dan $\sigma_1 > \sigma_2$

Gambar 6.4 diberikan dua kurva normal memiliki pusat yang sama pada sumbu X karena rata-rata kedua kurva normal tersebut sama. Kurva pertama simpangan bakunya lebih besar dibandingkan dengan simpangan baku kurva kedua, sehingga kurva pertama lebih tumpul atau lebih rendah daripada dan lebih menyebar kesamping dibandingkan kurva kedua. Semakin besar nilai σ, kurva akan semakin rendah (platikurtik) dan melebar ke samping, sedangkan semakin kecil σ, kurva akan semakin runcing (leptokurtik), dan mengumpul mendekati rata-rata.



Gambar 6.5 Dua Kurva Normal Dengan $\mu_1 = \mu_2$ dan $\sigma_1 > \sigma_2$

Gambar 6.5 diberikan dua kurva normal memiliki pusat yang berbeda pada sumbu X, sehingga dapat dilihat bahwa rata-rata kedua kurva normal tersebut tidak sama. Begitu juga dengan standar deviasinya, kedua kurva normal tersebut memiliki standar deviasai yang berbeda, yaitu $\sigma_1 > \sigma_2$.

Dengan melihat Gambar 6.1 sampai dengan Gambar 6.5 maka ada beberapa hal yang harus diperhatikan. Pertama, ternyata model setiap anggota keluarga kurva normal ditentukan oleh seperangkat parameter μ dan σ, hubungan dari kedua parameter ini membuat keluarga kurva normal menjadi sangat luas yang mempunyai anggota tidak terbatas. Atas dasar itu, kurva normal diusulkan menjadi suatu model umum, karena asumsi kurva normal mampu menjelaskan sejumlah besar fenomena yang terjadi secara alami, mulai dari skor tes sampai ke fenomena alam semesta.

Kedua, membuat kurva normal umum bukanlah suatu pekerjaan yang mudah, karena persamaan kurvanya sangat sulit untuk diintegralkan. Perhatikan rumus untuk mencari fungsi densitasnya (nilai pada sumbu Y) begitu rumit. Oleh karena itu, kita akan mengalami kesulitan untuk menyusun tabel kurva normal. Juga merupakan pekerjaan yang sia-sia jika kita berusaha menyusun tabel yang berbeda untuk setiap kurva normal umum bagi setiap pasangan nilai μ dan σ yang mungkin. Tabel merupakan langkah yang paling tepat untuk jalan lain bila kita mau menghindari dari keharusan menggunakan integral yang begitu rumit.

Berdasarkan kedua hal tersebut di atas, kita bisa mentrasformasikan setiap pengamatan yang berasal dari sebarang peubah acak normal X (fungsi X) menjadi suatu nilai suatu peubah acak normal Z (fungsi Z) dengan rata-rata $\mu = 0$

dan standar deviasi σ = 1. Distribusi yang diamati distandarkan atau ditransformasikan dengan menggunakan rumus:

$$z = \frac{x-\mu}{\sigma} \text{ untuk populasi dan } z = \frac{x-\bar{x}}{s} \text{ untuk sampel}$$

Kurva hasil transformasi setiap pengamatan yang berasal dari sebarang peubah acak normal X menjadi suatu nilai suatu peubah acak normal Z disebut dengan kurva normal standar (*standart normal distribution*) atau kurva normal baku. Disebut sebagai fungsi Z, kemungkinan berasal dari kata *zirro* yang berarti nol, seperti telah dikatakan bahwa rata-rata dari Z adalah 0. Selain itu, nilai-nilai probabilitas yang terdapat dalam tabel Z adalah nilai probabilitas antara μ = 0 dan satu nilai z tertentu, bukan antara dua buah nilai z sebarang. Dengan transformasi $z = \frac{x-\mu}{\sigma}$, cukup hanya dibuat satu tabel distribusi normal standar untuk semua keluarga kurva normal.

Berdasarkan hasil transformasi, rumus dari kurva normal umum berubah menjadi rumus kurva normal standar sebagai berikut.

$$n\,(x;\,0,\,1) = \frac{1}{\sqrt{2\pi\sigma}}\,e^{-\frac{1}{2}z^2}\text{, untuk } -\infty < z < \infty$$

Dapat dibuktikan bahwa kurva normal standar memiliki rata-rata μ = 0 dan standar deviasi σ = 1 sebagai berikut.

$$\mu_z = \int_{-\infty}^{\infty} z\, f(z)\, dz$$

$$= \int_{-\infty}^{\infty} z\, \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\frac{1}{2}z^2}\, dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}z^2} \; z \, dz,$$

misal $u = ½ z^2$, maka $du = z \, dz$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u} \; du$$

$$= \left[-\frac{1}{\sqrt{2\pi}} e^{-u} \right]_{-\infty}^{\infty}$$

$$= \left[-\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \right]_{-\infty}^{\infty}$$

= 0 (terbukti $\mu = 0$)

Begitu juga dapat dibuktikan bahwa $\sigma = 1$ sebagai berikut.

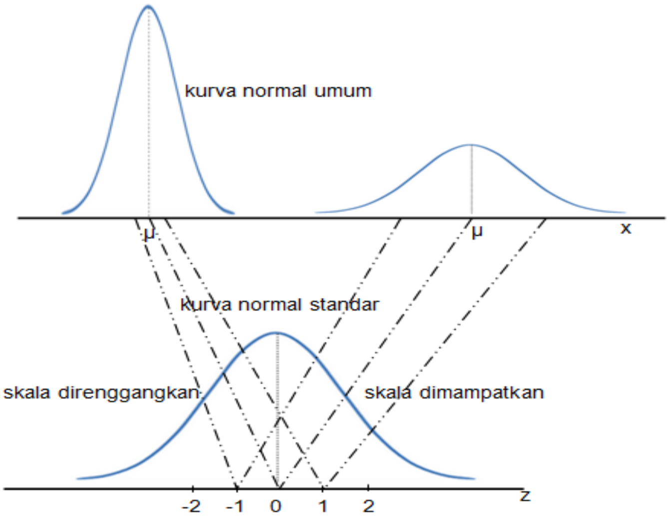
$$\sigma^2 = \int_{-\infty}^{\infty} (z - \mu_z)^2 \; f(z) \, dz$$

$$= \int_{-\infty}^{\infty} z^2 \; f(z) \, dz$$

$$= \int_{-\infty}^{\infty} \frac{z^2}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}z^2} dz$$

= 1 (terbukti $\sigma = 1$)

Seandainya terdapat dua keluarga kurva normal umum yang akan ditransforamsikan ke dalam kurva normal standar, maka dapat divisualisasikan pada Gambar 6.6 di bawah ini.

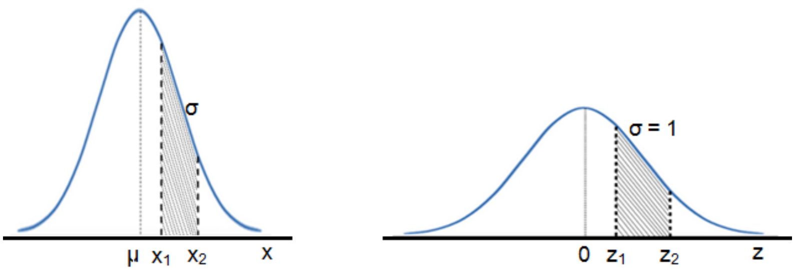


Gambar 6.6 Transformasi Kurva Normal Umum Menjadi Kurva Normal Standar

Jika peubah acak X berada di antara $x = x_1$ dan $x = x_2$, maka peubah acak Z berada di antara nilai-nilai padanannya, yaitu:

$$z_1 = \frac{x_1 - \mu}{\sigma} \text{ dan } z_2 = \frac{x_2 - \mu}{\sigma}$$

Semua nilai X yang berada di antara x_1 dan x_2, mempunyai padanan pada Z, yaitu nilai-nilai di antara z_1 dan z_2 seperti terlihat pada Gambar 6.7.



Gambar 6.7 Kurva Normal Umum dan Kurva Normal Standar

Dengan demikian luas daerah di bawah kurva normal umum (fungsi X) yang dibatasi oleh x_1 dan x_2 memiliki luas yang sama dengan luas daerah di bawah kurva normal standar (fungsi Z) yang dibatasi oleh z_1 dan z_2. Dengan demikian probabilitas peubah acak X antar x_1 dan x_2 sama dengan probabilitas peubah acak Z antar z_1 dan z_2 yang dituliskan seperti di bawah ini.

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

Sebelum lebih jauh membahas penggunaan kurva normal standar, khususnya dalam analisa data penelitian, alangkah baiknya kita perhatikan terlebih dahulu daerah di bawah kurva normal. Kurva normal standar yang merupakan hasil transformasi dari kurva normal umum mengubah nilai rata-rata dan simpangan baku fungsi X ke dalam distribusi normal standar Z. Nilai μ pada kurva normal umum diganti dengan 0 pada kurva normal standar.

Ambil nilai $x = \mu + \sigma$, nilai dari z adalah

$$\begin{aligned} z &= \frac{x-\mu}{\sigma} \\ &= \frac{\mu+\sigma-\mu}{\sigma} \\ &= \frac{\sigma}{\sigma} \\ &= 1 \end{aligned}$$

Jika titik yang diamati berada 1σ di sebelah kanan rata-rata, maka nilai z = 1

Ambil nilai $x = \mu + 2\sigma$, nilai dari z adalah

$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{\mu + 2\sigma - \mu}{\sigma}$$

$$= \frac{2\sigma}{\sigma}$$

$$= 2$$

Jika titik yang diamati berada 2σ di sebelah kanan rata-rata, maka nilai z = 2

Ambil nilai x = μ - σ, nilai dari z adalah

$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{\mu - \sigma - \mu}{\sigma}$$

$$= \frac{-\sigma}{\sigma}$$

$$= -1$$

Jika titik yang diamati berada 1σ di sebelah kiri rata-rata, maka nilai z = -1

Ambil nilai x = μ - 2σ, nilai dari z adalah

$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{\mu - 2\sigma - \mu}{\sigma}$$

$$= \frac{-2\sigma}{\sigma}$$

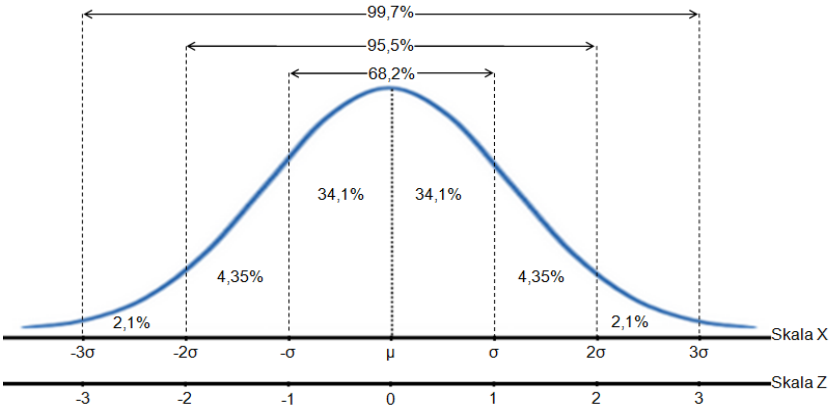
$$= -2$$

Dengan demikian nilai 1 diganti dengan $\mu + 1\sigma$, nilai 2 diganti dengan $\mu + 2\sigma$, nilai 3 diganti dengan $\mu + 3\sigma$, nilai -1 diganti dengan $\mu - 1\sigma$, nilai -2 diganti dengan $\mu - 2\sigma$, nilai -3 diganti dengan $\mu - 3\sigma$. Kurva normal adalah kurva yang simetris, yang berarti bahwa kurva ini akan membagi luas kurva menjadi 2 bagian yang sama. Seluruh luas kurva adalah 1 satuan luas persegi atau 100% dan rata-rata (μ) membagi luas kurva menjadi 2 bagian yang sama, berarti luas tiap belahan adalah 50%. Setiap penyimpangan dari rata-rata sebesar standar deviasi dapat ditentukan persentase terhadap seluruh luas kurva, yaitu:

a) Sekitar 68,2% atau 0,6820 luas daerah di bawah kurva normal yang menyimpang 1 standar deviasi dari rata-rata dan yang ditulis dengan $\mu \pm 1\sigma$. Dengan demikian luas masing-masing daerah dari μ sampai ke $\mu \pm 1\sigma$ adalah 34,1% atau 0,3410.
b) Sekitar 95,5% atau 0,9550 luas daerah di bawah kurva normal yang menyimpang 2 standar deviasi dari rata-rata dan yang ditulis dengan $\mu \pm 2\sigma$. Dengan demikian luas masing-masing daerah dari $\mu \pm 1\sigma$ sampai ke $\mu \pm 2\sigma$ adalah 4,35% atau 0,0435.
c) Sekitar 99,7% atau 0,9970 luas daerah di bawah kurva normal yang menyimpang 3 standar deviasi dari rata-rata dan yang ditulis dengan $\mu \pm 3\sigma$. Dengan demikian luas masing-masing daerah dari $\mu \pm 2\sigma$ sampai ke $\mu \pm 3\sigma$ adalah 2,1% atau 0,0210.
d) Penyimpangan lebih dari 3σ dari rata-rata biasanya tidak dihitung karena nilainya yang relatif kecil, yaitu sekitar 0,03% atau 0,0030. Oleh karena itu, tabel kurva normal

standar atau yang lebih dikenal dengan Tabel Z hanya memuat luas daerah di bawah kurva normal standar dari 0 sampai dengan 3,99.

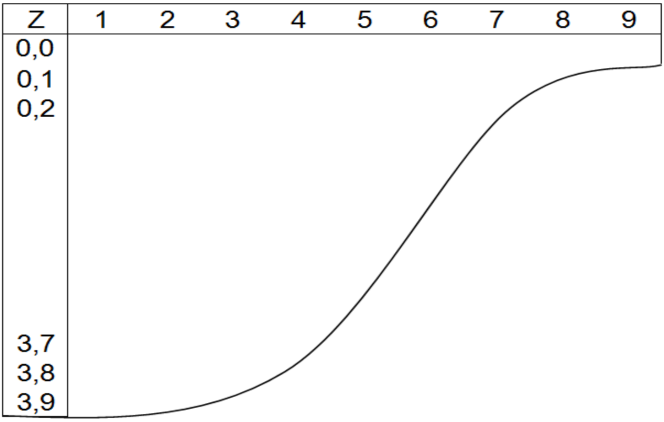
Paparan poin a) sampai dengan d) di atas secara lebih jelas dapat dilihat pada Gambar 6.7 di bawah ini.



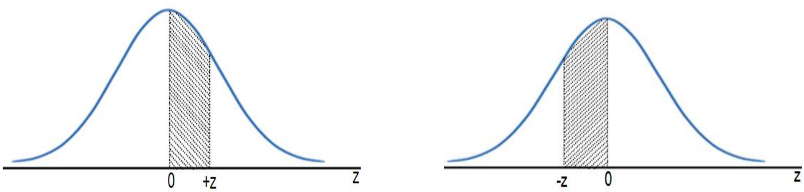
Gambar 6.8 Luas Daerah Di Bawah Kurva Normal Standar

Setelah memahami bentuk kurva normal, transformasi kurva normal umum ke dalam kurva normal standar dengan transformasi $z = \frac{x - \mu}{\sigma}$, selanjutnya adalah bagaimana caranya menghitung luas daerah di bawah kurva normal antara z_1 dan z_2. Hal pertama yang diperlukan adalah cara membaca atau menggunakan tabel distribusi normal standar atau tabel Z. berikut ini diberikan langkah-langkah menghitung luas daerah di bawah kurva normal dengan menggunakan tabel Z.

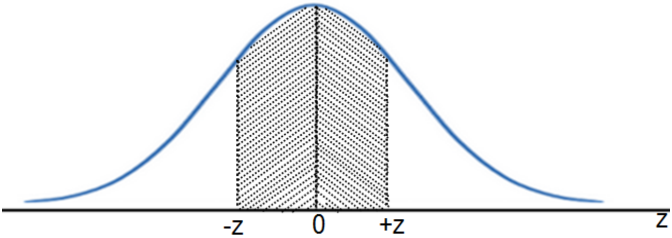
1) Hitunglah nilai z hingga dua desimal, karena dalam tabel Z umumnya memuat dua desimal atau dua angka di belakang koma, yaitu mulai dari 0,00 sampai dengan 3,99.



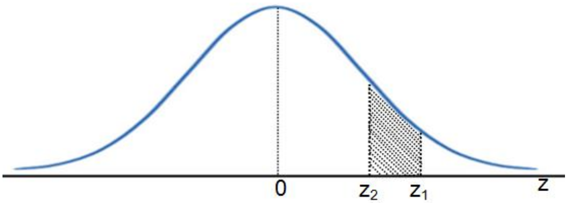
2) Sketsalah grafik kurva normal standar dengan membuat z = 0 ditengah-tengah grafik sehingga membagi grafik menjadi dua bagian yang simetris. Jika nilai z positif (+z), maka letaknya disebelah kiri z = 0 dan jika z negatif (-z), maka letaknya di sebelah kanan z = 0. Nilai z negatif atau positif hanya menentukan letak z, apakah di sebelah kanan z= 0 atau di sebelah kiri, bukan menyatakan luas daerah, karena tidak ada luas negatif. Jadi luar daerah z = +a akan sama dengan luas daerah z = -a, hanya letaknya yang berbeda.



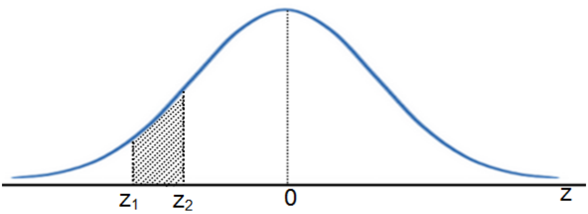
Menentukan luas daerah di bawah kurva normal yang menjadi patokan adalah z = 0. Luar daerah dari z = 0 ke +z sama dengan luas daerah dari z = 0 ke -z. Jika ada dua nilai z yang berbeda, misalnya -z dan +z, maka luar daerah di bawah kurva normal yang dibatasi oleh -z dan +z adalah jumlah luas daerah dari z = 0 ke +z dan z = 0 ke -z.



Jika dua nilai z yang berbeda sama-sama positif, misalnya z_1 dan z_2 dengan $z_1 > z_2$, maka luas daerah di bawah kurva normal yang dibatasi oleh z_1 dan z_2 adalah selisih luas daerah dari z = 0 ke z_1 dikurangi z = 0 ke z_2.



Begitu juga jika $-z_1$ dan $-z_2$ dengan $z_1 > z_2$, maka luas daerah di bawah kurva normal yang dibatasi oleh z_1 dan z_2 adalah selisih luas daerah dari z = 0 ke $-z_1$ dikurangi z = 0 ke $-z_2$.



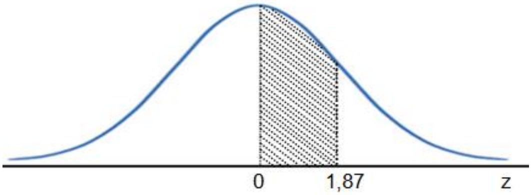
3) Luas daerah yang terdapat di dalam tabel Z adalah luas daerah dari titik z = 0 ke titik z yang diketahui. Di bawah kolom z merupakan nilai z dalam satu desimal yaitu dari nilai 0,0 sampai dengan 3,9; sedangkan desimal kedua terletak pada baris paling atas dengan nilai 0 sampai dengan 9.
4) Misalnya kita mencari luas daerah di bawah kurva normal dengan z = 0,37. Perhatikan nilai desimal yang terdapat di sebelah kiri, yaitu 0,3 maju ke kanan, dari kolom paling atas lihat nilai 7 terus turun ke bawah sampai ketemu dengan garis lurus dari titik 0,3. Pertemuan ke dua garis tersebut merupakan koordinat dari luas daerah yang dibatasi oleh z = 0,37.

z	1	2	3	4	5	6	7	8	9
0,0									
0,1									
0,2									
0,3							0,1443		
0,4									

Diperoleh luas daerah antaraa z = 0 samapai dengan z = 0,37 yang luasnya sama dengan 0,1443 atau 14,43%

Contoh 6.1 Carilah luas daerah yang dibatasi oleh z = 0 dan z = 1,87.

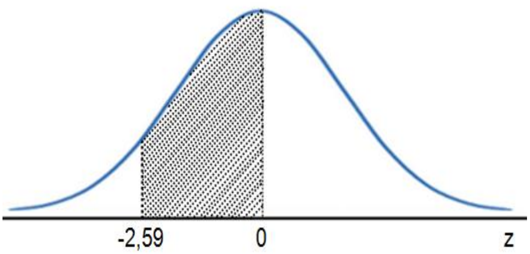
Penyelesaian



Pada tabel Z, pada kolom pertama atau kolom z cari 1,8 kemudian ke kanan sehingga bertemu dengan angka di bawah kolom angka 7.
Bilangan tersebut adalah 0,4693; luas daerah yang dicari adalah 0,4693 atau 46,93% seperti pada gambar di atas daerah yang diarsir adalah luas daerah yang dicari, yaitu luas daerah antara z = 0 sampai dengan z = 1,87.

Contoh 6.2 **Carilah luas daerah yang dibatasi oleh z = 0 dan z = -2,59.**

Penyelesaian



Pada tabel Z, pada kolom pertama atau kolom z cari 2,5 kemudian ke kanan sehingga bertemu dengan angka di bawah kolom angka 9 (ingat tanda negatif menyatakan letak di sebelah kiri z = 0). Bilangan tersebut adalah 0,4952; luas daerah yang dicari adalah 0,4952 atau 49,52% seperti pada gambar di atas daerah yang diarsir adalah luas daerah yang dicari, yaitu luas daerah antara z = 0 sampai dengan z = 2,59.

Contoh 6.3 **Untuk sebaran normal dengan μ = 60 dan σ = 25 tentukanlah probabilitas bahwa X mengambil sebuah nilai antara x_1 = 40 dan x_2 = 70.**

Penyelesaian

Nilai z_1 dan z_2 padanan dari x_1 dan x_2 adalah:

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{40-60}{25} = \frac{-20}{25} = -0{,}80$$

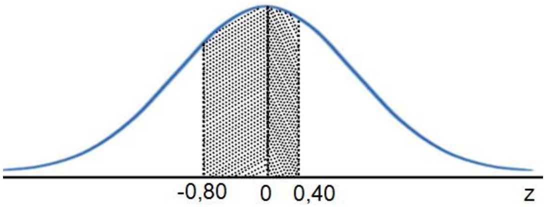
$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{70-60}{25} = \frac{10}{25} = 0{,}40$$

Sehingga,

P(40 < X < 70) = P(-0,80 < Z < 0,40)

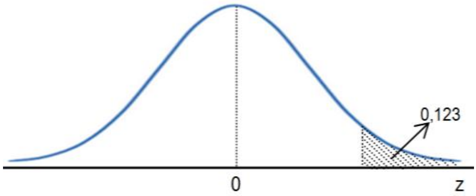
Nilai dari P(-0,80 < Z < 0,40) diberikan oleh daerah yang diarsir dalam Gambar 6.9. Luas daerah P(-0,80 < Z < 0,40) ini dapat diperoleh dengan menjumlahkan luas daerah yang dibatasi oleh z = 0 sampai dengan z = -0,80 dan z = 0 sampai dengan z = 0,40 dengan menggunakan tabel Z diperoleh

= P(40 < X < 70)
= P(-0,80 < Z < 0,40)
= P(-0,80 < Z < 0) + P(0 < Z < 0,40)
= 0,2881 + 0,1554
= 0, 4435 atau 44,35%



Contoh 6.4 **Hasil ujian statistik menunjukkan nilai rata-rata 70 dan simpangan bakunya 10. Jika 12,3% diantara peserta ujian akan diberi nilai A dan nilai ujian mengikuti sebaran normal, berapakah batas terkecil bagi nilai A dan batas terbesar bagi nilai B.**

Penyelesaian



Contoh ini mengharuskan kita terlebih dahulu menemukan nilai z-nya, dari nilai z tersebut baru ditentukan nilai x.

Berarti mencari nilai z yang luasnya 0,1230, sehingga z adalah 0,5000 – 0,1230 = 0,3770. P (Z > 1,16) = 0,3370.

Sehingga diperoleh

$$z = \frac{x - \mu}{\sigma}$$

$$\begin{aligned} x &= z.\sigma + \mu \\ &= (1,16 \,.\, 10) + 70 \\ &= 11,6 + 70 \\ &= 81,6 \end{aligned}$$

Jadi skor terendah bagi nilai A adalah 81,6 dan nilai tertinggi bagi B adalah 8,55

Contoh 6.5 **Skor rata-rata ujian masuk suatu universitas adalah 75 dengan simpangan baku 10. Jika skor ujian berdistribusi normal dan banyak calon pelamar adalah 1000 orang, tentukanlah!**

a. Berapa orang yang nilainya lebih dari 82,2?

b. Berapa calon mahasiswa yang nilainya diantara 80 dan 90?

c. Berapa orang calon yang nilainya lebih dari atau sama dengan 80?

d. Berapa orang calon yang nilainya 80?

Penyelesian

$\mu = 75$

$\sigma = 10$

x = skor ujian

a. $P\ (z > 82,2)$

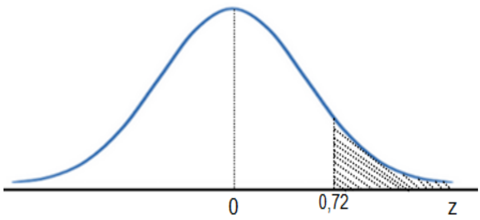
$$z = \frac{x - \mu}{\sigma}$$

$$= \frac{82,2 - 75}{10}$$

$$= \frac{7,2}{10}$$

$$= 0,72$$

Nilai yang lebih dari 82,2 berati luas daerah yang diarsir terletak di sebelah kanan z = 0,72



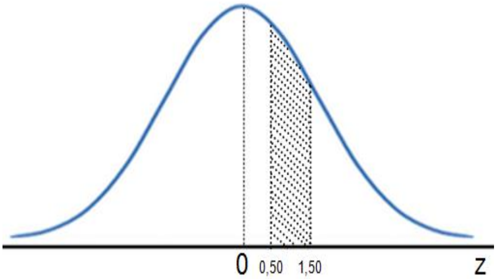
Luas daerah z = 0,72 atau $z_{0,72}$ adalah 0,2642, dengan demikian luas daerah yang lebih dari 0,72 adalah 0,5000 - 0,2642 = 0,2358. Jadi banyaknya calon mahasiswa yang skor ujiannya lebih dari 82,2 adalah 0,2358 (23,58%) atau sekitar 266 orang.

b. P (80 <X <90)

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{80-75}{10} = \frac{5}{10} = 0,50$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{90-75}{10} = \frac{15}{10} = 1,50$$

Jadi P (80 < X <90) = P (0,50 < Z < 1,50), persentase calon terletak antara nilai z_1 dan z_2.



Luas daerah z_1 adalah 0,1915 dan luas daerah z_2 adalah 0,4332. Luas daerah antara z_1 dan z_2 adalah 0,4332 - 0,1915 = 0,2417 (24,17%) atau sekitar 242 orang.

c. P (x ≥ 80)

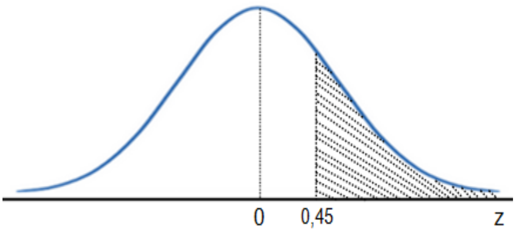
x ≥ 80 dalam hal ini berarti skor 80 juga termasuk di dalamnya, maka yang dipakai adalah batas bawah dari skor 80, yaitu 79,5. Agar skor 80 termasuk di dalam batas-batas nilai x, maka batas nilai x yang digunakan adalah 79,5.

$$z = \frac{x-\mu}{\sigma}$$

$$= \frac{79{,}5-75}{10}$$

$$= \frac{4{,}5}{10}$$

$$= 0{,}45$$

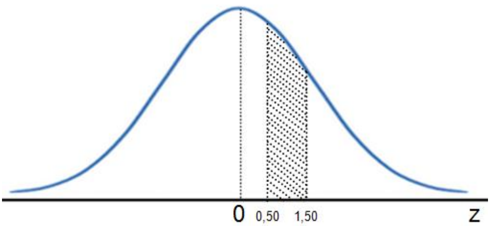


Luas daerah di bawah kurva normal dari z = 0 sampai dengan z = 0,45 adalah 0,1736. Luas daerah yang diarsir (z > 0,45) adalah 0,5000 - 0,1736 = 0,3264 (32,64%). Dengan demikian banyaknya calon pelamar yang memiliki nilai lebih dari atau sama dengan 80 adalah sekitar 226 orang.

d. Skor 80 terletak diantara batas atas dan batas bawah 80, yaitu batas bawahnya 79,5 dan batas bawahnya 80,5. Untuk mencari persentase mahasiswa yang memperoleh nilai 80 adalah terletak diantara x1 = 79,5 dan 80,5 atau P(79,5 < X < 80,5).

$$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{79{,}5 - 75}{10} = \frac{4{,}5}{10} = 0{,}45$$

$$z_2 = \frac{x_2 - \mu}{\sigma} = \frac{80{,}5 - 75}{10} = \frac{5{,}5}{10} = 0{,}55$$



Luas daerah z_1 adalah 0,1736 dan luas daerah z_2 adalah 0,2088. Luas daerah antara z_1 dan z_2 adalah 0,2088 - 0,1736 = 0,0352 (3,52%) atau sekitar 35 orang.

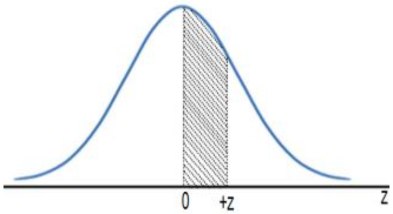
Latihan 6

1. Tentukanlah luas daerah di bawah kurva normal berikut dengan menggambar grafiknya.
   a. z = 2,57
   b. z = 0,35

c. z = -2,57
d. antara z = -0,23 dan z = 1,34
e. antara z = -1,23 dan z = -2,34
f. z > 1,35

2. Tentukanlah nilai z jika diketahui luas daerah di bawah kurva normal sebagai berikut.
   a dari z ke kiri 0,2054
   b. dari z ke kanan 0,3888
   c. dari z ke kanan 0,9988
   d. dari z ke kiri 0,9251
3. Suatu distribusi normal dengan rata-rata 40 dan simpangan baku 11 tentukanlah P (56 < X < 79).
4. Dua orang mahasiswa A dan B, masing-masing mendapatkan skor baku z_a = 1,28 dan z_b = -1,28 untuk ujian statistik dasar. Jika nilai kedua orang mahasiswa tersebut 70 dan 95 berdistribusi normal, berapakah rata-rata dan simpangan baku skor ujian tersebut.
5. Dari 3000 mahasiswa di suatu perguruan tinggi diketahui berat badan mahasiswa memiliki sebaran normal dengan rata-rata 67,5 dan simpangan baku 7,75. Tentukanlah berapa mahasiswa yang beratnya
   a. lebih dari 85
   b. lebih dari 65
   c. kurang dari 80
   d. antara 60 dan 78
   e. nilai 83
   f. nilai 60

Tabel Z
Luas Daerah Di Bawah Kurva Normal Dari 0 ke z



z	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633

z	0	1	2	3	4	5	6	7	8	9
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4493	0,4493
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

BAB VII
Z-SKOR DAN T-SKOR

Skor-skor standar yang dibahas pada bagian berikutnya berkaitan erat dengan evaluasi. Sistem evaluasi tradisional (sistem lama) dalam memberikan skor hasil ujian siswa menggunakan skor-skor standar. Jawaban siswa terhadap butir soal langsung diberikan skor standar, kemudian skor yang diperoleh dari setiap butir dijumlahkan lalu dibagi dengan jumlah butirnya. Hasil inilah yang digunakan sebagai standar hasil belajar siswa. Skor standar ini biasanya bergerak dari 0 sampai dengan 10. Supaya skor butir soal lebih bermakna dalam kaitannya dengan posisi atau letak relatif (*relative standing*) secara keseluruhan, perlu adanya skor yang dapat dibandingkan satu sama lain, skor ini disebut dengan skor standar.

Dalam sitem evaluasi modern jawaban siswa terhadap butir soal tidak langsung diberikan skor standar tetapi skor mentah (*raw score*) yang bersifat sementara. Seperti diungkapkan di atas pemberian skor standar biasanya bergerak dari dari 0 sampai dengan 10, tetapi dalam skor mentah tidak ada pembatasan yang mutlak. Bisa bergerak dari 0 sampai dengan 20, 0 sampai dengan 50, 0 sampai dengan 100, atau yang lainnya.

Tentunya skor mentah yang diperoleh siswa dalam suatu ujian atau tes memiliki kelemahan, karena belum mampu memberikan keterangan yang akurat tentang posisi atau prestasi siswa dalam tes tersebut. Sebagai contoh, seorang siswa memperoleh nilai 80 dalam suatu tes, pertanyaannya apakah siswa tersebut memperoleh hasil yang baik, sedang, atau kurang. Oleh karena itu, untuk memperoleh gambaran yang lebih akurat tentang skor yang diperoleh siswa dalam suatu tes, skor mentah yang diperoleh siswa harus diubah ke dalam skor standar (*standar score*). Skor standar adalah skor mentah yang telah ditransformasikan linier ke dalam bentuk lain yang disebut dengan skor standar.

A. Z-Skor

Jika kita mendengar atau melihat z-skor, maka secara otomatis kita mengingat distribusi normal. Pada Bab VI telah dibahas secara detail tentang distribusi normal standar Z, selanjutnya pada bagian ini kita istilahkan dengan z-skor. Z-skor merupakan skor standar yang paling sederhana yang menentukan jarak suatu skor dari mean kelompoknya dalam unit simpangan baku. Skor ini dapat berupa nilai atau dalam satuan SD. Skor ini biasanya digunakan untuk mengubah skor-skor mentah yang diperoleh dari berbagai jenis pengukuran yang berbeda-beda. Jika distribusi dua skor atau lebih mendekati normal, maka skor standar yang berasal satu distribusi mungkin dibandingkan dengan yang lainnya. Seperti telah dijelaskan sebelumnya rumus dari z-skor adalah sebagai berikut.

$$z = \frac{x - \bar{x}}{s}$$

Keterangan

z = z-skor

x = skor mentah

$\bar{x}$ = rata-rata

s = standar deviasi

Rerata standar atau yang dikenal dengan z-skor merupakan statistik yang sangat berguna karena (a) memungkinkan kita untuk menghitung probabilitas skor yang terjadi dalam distribusi normal dan (b) memungkinkan kita untuk membandingkan dua nilai yang berasal dari distribusi normal yang berbeda. Z-skor digunakan untuk mengetahui lebih detail di mana posisi suatu skor dalam suatu distribusi. Posisi dalam suatu distribusi itu sendiri ditunjukan dengan simbol positif (+) atau negatif (-). Positif berarti suatu skor berada di atas rata-rata serta kalau negatif skor tersebut berada di bawah rata-rata. Seseorang yang memiliki kemampuan lebih tinggi adalah individu yang z-skornya bertanda positif (+). Sebaliknya, yang bertanda (-) adalah individu yang memiliki kemampuan lebih lemah dari lainnya. Jika nilai yang ditunjukkan oleh z-skor bertanda positif itu makin tinggi, berarti kedudukan relatif siswa bersangkutan juga semakin tinggi dan sebaliknya, jika z-skor yang bertanda negatif itu makin besar, maka *standing position* siswa tersebut menjadi semakin rendah. Dengan demikian transformasi skor mentah menjadi z-skor ini perlu dilakukan untuk memberikan tafsiran yang lebih tepat mengenai kamampuan seorang siswa.

Contoh 7.1 **Seorang mahasiswa memperoleh nilai ujian akhir semester beberapa matakuliah sebagai berikut.**

Mata Kuliah	Skor UAS	Rata-rata Kelas	Standar Deviasi
Statistik dasar	80	74	5
Kalkulus II	75	64	8
Diskrit	85	78	6
Teori Bilangan	73	75	7

Tentukanlah kepandaian mahasiswa tersebut berdasarkan nilai UASnya!

Penyelesaian

Terlebih dahulu kita cari nilai z-skor untuk masing-masing matakuliah

a. z-skor statistik dasar

$$z_S = \frac{x - \bar{x}}{s}$$

$$= \frac{80 - 74}{5}$$

$$= \frac{6}{5}$$

$$= 1{,}20$$

jadi nilai statistik dasar mahasiswa tersebut berada 1,20s di atas rata-rata

b. z-skor kalkulus II

$$z_K = \frac{x - \bar{x}}{s}$$

$$= \frac{75 - 64}{8}$$

$$= \frac{11}{8}$$

$$= 1{,}38$$

jadi nilai kalkulus II mahasiswa tersebut berada 1,38s di atas rata-rata

c. z-skor diskrit

$$z_D = \frac{x - \bar{x}}{s}$$

$$= \frac{85 - 78}{6}$$

$$= \frac{7}{6}$$

$$= 1,17$$

jadi nilai diskrit dasar mahasiswa tersebut berada 1,17s di atas rata-rata

d. z-skor teori bilangan

$$z_T = \frac{x - \bar{x}}{s}$$

$$= \frac{73 - 75}{7}$$

$$= \frac{-2}{7}$$

$$= -0,29$$

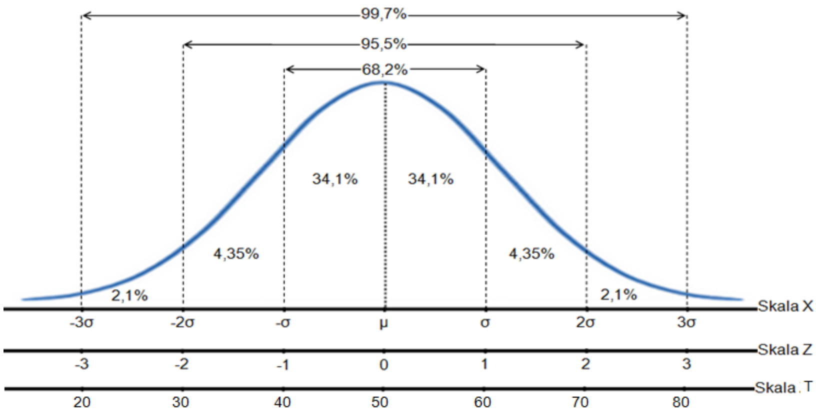
jadi nilai statistik dasar mahasiswa tersebut berada -0,29s di bawah rata-rata

Dengan melihat z-skor masing-masing nilai UAS mahasiswa tersebut, maka mahasiswa tersebut digolongkan paling bagus (pandai) pada mata kuliah kalkulus II dan kurang pandai pada mata kuliah teori bilangan.

B. T-Skor

Dalam z-skor menggunakan rata-rata 0 dan standar deviasi 1, dengan ketentuan ini kita akan bertemu dengan bilangan negatif maupun pecahan, yang mungkin nantinya kurang dipahami bagi mereka yang masih asing atau awam terhadap ukuran-ukuran statistik. Untuk menghindari hal tersebut disusunlah sebuah skor standar yang dikenal dengan T-skor. Dalam T-skor menggunakan rata-rata 50 dan jarak tiap deviasi standar 10.

Untuk memperoleh T-skor, skor standar dilipatgandakan 10 kali kemudian di tambahkan atau dikurangi 50 z-skor. Asumsi dalam teknik ini adalah bahwa secara dekat skor ini akan menjadi jarak antara lima standar deviasi dari rata-rata. Dengan demikian dalam range -3s sampai dengan +3s dalam z skor, tersebut dalam T-skor mulai dari 20 sampai dengan 80, tanpa bilangan negatif.



Gambar 7.1 Kurva Normal dalam Skala X, Skala Z, dan Skala T

Nilai -3 dalam z-skor diganti dengan 20, nilai -2 diganti dengan 30, nilai -1 diganti dengan 40, nilai 0 (rata-rata) diganti dengan 50, nilai +1 diganti dengan 60, nilai +2 diganti dengan 70, dan nilai +3 diganti dengan 80. Adapun formula dari T-skor adalah sebgai berikut.

$$T = 10\left(\frac{x - \bar{x}}{s}\right) + 50 \text{ atau } T = 10z + 50$$

keterangan
T = T-skor
z = z-skor
x = skor mentah
$\bar{x}$ = rata-rata
s = standar deviasi

Contoh 7.1 **Lihat kembali Contoh 7.1 transformasikan nilai UAS mahasiswa tersebut ke dalam T-skor.**

Penyelesaian

a. T-skor statistik dasar

$$\begin{aligned} T_S &= 10z_S + 50 \\ &= (10 \times 1{,}20) + 50 \\ &= 12 + 50 \\ &= 62 \end{aligned}$$

b. T-skor kalkulus II

$$\begin{aligned} T_k &= 10z_K + 50 \\ &= (10 \times 1{,}38) + 50 \\ &= 13{,}8 + 50 \\ &= 63{,}8 \end{aligned}$$

c. T-skor diskrit

$T_D = 10z_D + 50$
$= (10 \times 1{,}17) + 50$
$= 11{,}7 + 50$
$= 61{,}7$

d. T-skor teori bilangan

$T_k = 10z_K + 50$
$= (10 \times -0{,}29) + 50$
$= -2{,}9 + 50$
$= 47{,}10$

Latihan 7

1. Apa perbedaan skor mentah dengan skor standar?
2. Jelaskan perbedaan yang mendasar dari z-skor dan t-skor?
3. Andi mengikuti ulangan umum semester genap di sekolahnya, diketahui data hasil ulangan umum Andi sebagai berikut.

Mata Pelajaran	Skor Ulangan	Rata-rata Kelas	Standar Deviasi
Bahasa Indonesia	70	65	2
Bahasa Inggris	85	75	5
Matematika	67	66	6
PKN	82	78	3
Biologi	58	50	4
Fisika	78	79	5
Kimia	91	82	7

Unggulan dalam mata pelajaran apakah Andi, jelaskan!

4. Transformasikan nilai ulangan umum Andi ke dalam T-skor, apa yang dapat Saudara simpulkan?
5. Perhatikan hasil ujian tiga orang mahasiswa berikut ini.

Mata Pelajaran	**Skor Ujian**			**Rata-rata Kelas**	**Standar Deviasi**
	Eka	**Citra**	**Wina**		
Kewiraan	60	66	70	72	12
Kewirausahaan	82	70	76	75	14
Statistik	75	78	72	76	9
Kalkulus	81	85	78	65	15
Geometri	74	56	58	60	8

Dari ketiga orang tersebut siapakah yang paling pintar, jelaskan!

DAFTAR PUSTAKA

Budiarto, Eko. 2002. Statistik Untuk Kedokteran dan Kesehatan Masyarakat. Jakarta. EGC

Bulman, A.G. 2012. *Elementary Statistic: A Step By Approch*, Eight Edition, New York. McGraw-Hill

Bungin, Burhan. 2006. *Metode Penelitian Kuantitatif Komunikasi, Ekonomi, dan Kebijakan Publik serta Ilmu-ilmu Sosial Lainnya*. Jakarta. Prenada Media Group.

Guilford, J.P. dan Fruchter, B. 1978. *Fundamental Statistics in Psycholoy and Education*, New York. McGraw-Hill Ltd.

Hadi, S. 2000. *Statistik 1, 2, 3*, Yogyakarta. Andi Offset

Harinaldi, 2005, Prinsip-prinsip Statistik untuk Teknik dan Sains. Jakarta. Erlangga

Hasan, Iqbal. 2006. Analisis Data Penelitian dengan Statistik. Jakarta. Bumi Aksara.

Herrhyanto dan Akib Hamid, H. M. 2008. Statistik Dasar. Jakarta. Universitas Terbuka.

Irianto, Agus. 2008. Statistik Konsep Dasar dan Aplikasinya. Jakarta. Kencana.

Kadir. 2015. Statistika Terapan: Konsep dan contoh Analisis Data dengan Program SPSS/Lisrel dalam Penelitian. Jakarta. RajaGrafindo Persada.

Larson, R. dan Farber, B. 2012. *Elementary Statistic: Picturing The World, Fifth Edition,* Boston. Pearson Education.

Lind A, Marchal, and Wathen, 2008, *Statisical Techniquesin Business And Economics*. 15th Edition. Mc Graw Hill International Edition. New York.

Riduwan. 2005. *Dasar-Dasar Statistika*. Bandung. Alfabeta.

Stephen Bernstein and Ruth Bernstein. 1999. *Elements of Statistics I: Descriptive Statistics and Probability*. The McGraw-Hill Companies, Inc.

Sudjana. 2002. *Metoda Statistika edisi ke 6*. Bandung. Tarsito.

Sugiyono. 2011. *Statistika untuk Penelitian*. Bandung. Alfabeta

Suharyadi, & Purwanto S. K. 2007. Statistika: Untuk Ekonomi dan Keuangan Modern, Edisi 2. Jakarta. Salemba Empat.

Susetyo, Budi. 2010. *Statistika Untuk Analisis Data Penelitian*. Bandung. PT. Refika Aditama.

Vardiansyah, Dani. Filsafat Ilmu Komunikasi: Suatu Pengantar, Indeks, Jakarta 2008.

Walpole, R.E. 1992. *Pengantar Statistika*. Jakarta. PT Gramedia Pustaka Utama.

Wibisono, Yusuf. 2009. Metode Statistik. Yogyakarta. Gadjah Mada University Press.

CATATAN

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....