BAHAN AJAR

MATA KULIAH :

Matematika Dan Statistika Arsitektur

OLEH :

JOHANSEN CRUYFF MANDEY
NIP. 197502092005011002

JURUSAN TEKNIK ARSITEKTUR
FAKULTAS TEKNIK
UNIVERSITAS SAM RATULANGI

POKOK BAHASAN:

“PENGANTAR DAN PENGERTIAN STATISTIK”

Kata “Statistik” telah digunakan untuk membatasi cara-cara ilmiah untuk mengumpulkan, menyusun, meringkas, dan menyajikan data penyelidikan. Lebih lanjut statistik merupakan cara untuk mengolah data tersebut dan menarik kesimpulan-kesimpulan yang teliti dan keputusan-keputusan yang logic dari pengolahan data tersebut.

Kata statistik juga digunakan untuk menunjuk kepada angka-angka pencataan dari suatu kejadian atau kasus tertentu seperti misalnya :

- Statistik Nikah-Talak-Rujuk : N-T-R = 46 - 21 – 34
- Statistik bentuk badan miss Universe : 38-22-36 (data-pinggang-pinggul)
- Statistik kecelakaan lalu lintas : Januari 6, Februari 38, Maret 21, ……..
- Statistik tinggi badan rata-rata : Rata-rata T = 164 cm, dan seterusnya.

1.1 KEGUNAAN STATISTIK

Dahulu statistik hanya digunakan untuk menggambarkan keadaan dan menyelesaikan persoalan-persoalan Negara seperti pencatatan banyaknya penduduk, penarikan pajak, dan semacamnya. Sekarang hampir tiada bidang yang tidak menggunakan statistik: psikologi, pendidikan, kimia, biologi, pertanian, kedokteran, hukum, politik, sosiologi, teknik dan lain-lain lapangan. Mungkin hanya ilmu-ilmu pendekatannya spekulatif yang tidak menggunakan statistik, dan ilmu-ilmu pendekatannya spekulatif itu makin lama makin sedikit jumlahnya.

Banyak buku-buku baku (*standard*) yang hanya dapat dipahami secara mendalam dengan pengetahuan tentang statistik. Sebabnya ialah sekarang lebih banyak buku-buku yang ditulis berdasarkan penelitian-penelitian ilmiah atau riset.

Kecendurungan akan *facts and figures*, akan faakta dan angka-angka menimbulkan kebutuhan yang tidak dapat dihindari akan statistik, hingga tiada suatu riset pun dewasa ini ditulis tanpa angka-angka statistik.

Kecuali segi-segi praktis itu, sebagaimana juga dalam matematik, ada unsure-unsur dalam statistik yang berfungsi sebagai sarana mengembangkan cara-cara berfikir logic. Statistik lebih dari matematik, merupakan suatu jalan berfikir, ya malahan suatu bahasa untuk berfikir secara ilniah, untuk merencanakan penyelidikan-penyelidikan, serta untuk mencapai kesimpulan-kesimpulan dan keputusan-keputusan yang mantap.karena itu tidak berlebihan kiranya jika GUILFORD, yang bukunya juga dijadikan bahan penyusunan buku ini, menyatakan bahwa ada empat alasan yang sangat sederhana, tetapi tidak dapat dibantah mengapa seorang mahasiswa harus mengambil mata pelajaran dan menguasai konsep-konsep pokok serta dasar-dasar kerja statistik. Empat alasan itu adalah :

1. Seorang mahasiswa harus mampu membaca literature-literatur professional.
2. Seorang mahasiswa harus menyusun cara-cara untuk kuliah-kuliah tingkat tinggi.
3. Statistik merupakan bagian esensial dari pada latihan esensial.
4. Statistik di mana saja merupakan landasan dari kegiatan-kegiatan riset.

1.2 STATISTIK DALAM RISET

Khusus untuk keperluan-keperluan riset, seperti yang telah beberapa kali disinggungkan di depan, fungsi dan peranan statistik di gambarkan oleh GUILFORD sebagai berikut :

1. Statistik memungkinkan pencatatan scara paling eksak data penyelidikan
2. Statistik memaksa penyelidik menganut tata-fikir dan tata-kerja yang definit dan eksak.
3. Statistik menyediakan cara-cara meringkas data ke dalam bentuk yang lebih banyak artinya dan lebih gampang mengerjakannya.
4. Statistik member dasar-dasar untuk menarik kongklusi-kongklusi melalui proses-proses yang mengikuti tata yang dapat diterima oleh ilmu pengetahuan.
5. Statistik member landasan untuk meramallkan secara ilmiah tentang bagaimana segala sesuatu akan terjadi dalam kondisi-kondisi yang telah diketahui.
6. Statistik memungkinkan penyelidik menganalisa, menguraikan sebab-akibat yang kompleks yang rumit, yang tanpa statistik akan merupakan peristiwa yang membingungkan, kejadian yang teruraikan.

1.3 BAGAIMANA MEMPELAJARI STATISTIK

Sedikit sekali orang yang dengan membuka buku ditengah-tengah dapat memahami arti dan maksud konsep dan tata-kerja Statistik. Statistik memang sukar dipelajari secara meloncat-loncat. Sebab tiap-tiap uaraian dan persoalan berikutnya selalu dilandaskan pada uraian dan persoalan sebelumnya. Jika kita merasa sebagai orang yang genius, sebaiknya kita mempunyai kesabaran dan ketekunan untuk belajar *setapak demi setapak.*

Statistik sebagai pengetahuan praktis, sebagai ilmu terpakai, *an applied science* kurang manfaatnya jika hanya dipelajari teori-teorinya tanpa mencobakannya pada kejadian atau data yang actual. Itulah sebabnya disamping buku ini, yang akan dikeluarkan porsi demi porsi yang disediakan untuk kuliah satu semester, dikeluarkan juga suatu buku kerja atau buku latihan. Buku latihan ini akan membantu pemula untuk memahami konsep-konsep dasar dan bagaimana menerapkannya pada peristiwa dan data actual. Kiranya tak usah diketengahkan bahwa kemajuan kita dalam mempelajari Statistik akan sejalan dengan penguasaan kita terhadap teori-teorinya dan kecekatan kita mengerjakan data dalam latihan itu.

Bagi mereka yang mempunyai biaya dan waktu disarankan membeli dan menyelidiki buku-buku Statistik lain yang dapat dijumpai dan atau dibeli. Perbedaan-perbedaan gaya, irama serta tekanan dalam penyajian teori dan persoalan dalam tiap-tiap buku kan menentukan pemilihan kita, buku mana yang paling cocok bagi kita masing-masing.

POKOK BAHASAN:

"DISTRIBUSI FREKUENSI"

Sebelum kita membicarakan tentang apa yang dimaksud dengan Distribusi Frekuensi, baiklah kiranya diketengahkan lebih dahulu beberapa istilah yang bersangkut paut dengan itu.

2.1 VARIABEL PENYELIDIKAN

Seorang sarjana psikologi mengadakan penyelidikan tentang kecerdasan; seorang guru menyelidiki kecakapan murid-murid dalam berhitung; seorang ahli beton mengadakan penyelidikan tentang campuran-campuran beton; seorang ahli pertanian menyelidiki benih-beni baru. Kecerdasan, Berhitung, Campuran beton, dan Benih-benih baru itu semuanya merupakan objek penyelidikan. Objek yang diselidiki itu disebut variable penyelidikan.

2.2 NILAI VARIABEL

Kalau seorang guru menyelenggarakan penyelidikan tentang angka rata-rata dari murid-muridnya dalam Berhitung, maka guru melihat nilai-nilai ujian mereka, atau nilai-nilai dalam buku nilai. Nilai dalam Berhitung itu disebut nilai variable. Koefisien Intelegensi, Kekuatan Campuran Beton, Hasil Benih Baru per hektar, itu semuanya merupakan nilai variable.

2.3 NILAI VARIABEL KONTINU DAN DISKRIT

Ada dua macam nilai variable, yaitu nilai yang bersambung atau kontinu, dan nilai yang terpisah atau diskrit. Nilai tinggi orang misalnya, adalah nilai yang kontinu, sebab bila mana kita sebutkan tinggi si A 165cm, pada hakekatnya tinggi si A tidak mutlak tepat 165 cm, melainkan misalnya 165,30cm. Pada umumnya angka 165 cm itu ditetapkan untuk mewakili tinggi orang dari 164,50 cm sampai 165,49 cm. Mereka yang tingginya 165,50 cm sampai 166,49 cm dicatat 166 cm. Dengan kata lain, angka 0.50 ke atas dibulatkan ke atas, sedang angka dibawah 0,50 dihilangkan.

Akan tetapi tidak demikian halnya bila orang menyelidiki hasil-hasil ujian yang akan dinilai benar dan salah, atau lulus dan gagal. Benar dan salah atau lulus dan gagal adalah nilai-nilai yang terpisah satu sama lain, nilai-nilai yang diskrit, sebab tidak ada nilai-nilai yang lain yang dipandang setengah benar atau setengah lulus.

Arti dari pada pengertian tentang keduan macam niali variable ini sangat penting untuk mengerjakan bahan-bahan teknik Statistik selanjutnya. Keduanya mempunyai konsekuensi pengolahan yang berlainan. Hal ini akan kita ketahui nanti setelah kita sampai pada uraian-uraian yang lebih lanjut.

2.4. DISTRIBUSI FREKUENSI TUNGGAL

Dalam suatu penyelidikan tentang kecakapan Berhitung murid-murid tiga kelas tertinggi dari sebagian Sekolah Rendah disuatu kotapraja telah dikumpulkan nilai-nilai rapot sebagai berikut :

Mata Pelajaran	: Berhitung
Murid	: Laki-laki
Jumlah	: 72 orang

NILAI-NILAI

7	6	6	6	5	7	6	5	4	6	7	7	6	7	5	6	6	7
6	6	6	6	6	5	6	6	6	7	7	5	7	7	8	5	6	5
7	7	5	6	7	7	7	7	6	6	6	6	5	5	7	7	5	7
5	6	5	6	7	6	7	8	5	6	5	7	5	6	7	8	8	6

Melihat angka-angka berderet-deret itu kita tidak dapat memperoleh gambaran apa-apa. Untuk mendapatkan gambaran dan kesimpulan sekedarnya, kita perlu mengatur angka-angka itu menjadi suatu table. Cara membuat table itu adalah sebagai berikut :

Tabel 1

Tabel Nilai-Nilai Berhitung 72 Orang Murid Sekolah Rendah di Kotapraja ***Y***

NILAI	Jari-jari	Frekuensi
8	////	4
7	~~////~~ ~~////~~ ~~////~~ ~~////~~ ///	23
6	~~////~~ ~~////~~ ~~////~~ ~~////~~ ~~////~~ ///	28
5	~~////~~ ~~////~~ ~~////~~ /	16
4	////	1
	N^{+} =	72

N+ sebuah singkatan dari kata Numbers, yang berarti jumlah frekuensi variable

Kita lihat dalam tabel didepan itu ada tiga kolom. Kolom pertama memuat “Nilai” variable; kolom kedua berisi “Jari-jari” yang diambil dari bahan-bahan yang terkumpul; kolom ketiga memuat salinan jumlah jari-jari (ini disebut “Frekuensi”) tiap-tiap variable itu menjadi angka-angka.

Dari tabel I diatas kita dapat sekedar mengambil kesimpulan. Nilai ”Enam” mendapat frekuensi yang tinggi, diikuti oleh nilai-nilai “Tujuh”, “Lima”, dan “Delapan”, dan nilai “Empat” mendapat pembagian frekuensi yang terendah yaitu hanya satu. Nilai Enam disebut “Mode”. Mode adalah nilai(variable) yang mendapat frekuensi tertinggi.

Akan tetapi tabel semacam itu bukanlah tabel yang sempurna. Tabel yang sempurna (yang akan disajikan kepada pembacanya) tidak menyebutkan jari-jari didalamnya. Jadi untuk membuat tabel yang akan kita sajikan kepada pembaca kita, jari-jari tidak dibuat dalam tabel itu, melainkan pada kertas yang tersendiri. Baru kemudian jumlah frekuensi jari-jari itu kita salin kedalam angka-angka dan kita muat dalam tabel yang sesungguhnya. Dengan demikian tabel kita akan berwujud sebagai berikut:

Tabel 2

Tabel Nilai-Nilai Berhitung 72 Orang Murid Laki-Laki

Sekolah Rendah di Kotapraja Y

NILAI (X)+	Frekuensi (f)+
8	4
7	23
6	28
5	16
4	1
Jumlah	72

+X adalah huruf singkatan yang biasa digunakan untuk nilai, sedang f untuk frekuensi.

Perhatikan betul-betul apakah singkatan-singkatan itu huruf besar ataukah huruf kecil

Perlu dicatat bahwa adanya dua kolom, yaitu kolom Nilai dan kolom Frekuensi, itu bukanlah syarat mutlak, ada tabel yang memuat dua kolom, tiga kolom, empat kolom, lima kolom atau lebih. Hal ini tergantung kepada keperluan dan maksud-maksud membuat tabel itu. Tentang hal ini kita akan saksikan nanti dalam uraian-uraian selanjutnya.

Tabel 2 tersebut diatas disebut Tabel Distribusi Frekuensi Tunggal. Istilah (Distribusi) digunakan dalam Statistik untuk menunjukan (seolah-olah) Penyebaran" nilai-nilai dengan jumlah orang yang mendapat nilai itu, sedang istilah "Tunggal" menunjukan tidak adanya pengelompokan nilai-nilai variabel dalam kolom pertama. Kita lihat dalam Tabel 2 itu ada lima baris, yaitu baris nilai delapan, baris nilai tujuh, baris nilai enam, baris nilai lima dan baris nilai empat. Baris nilai itu dapat kita kurangi menjadi tiga baris saja, dengan jalan mengelompokan angka-angka tujuh dan delapan menjadi satu baris untuk mewakili nilai-nilai atas sedang, angka-angka lima dan enam menjadi satu kelompok untuk mewakili nilai-nilai sedang, dan angka-angka empat ke bawah menjadi satu kelompok untuk mewakili nilai-nilai kurang. Tabel frekuensi yang menggunakan pengelompokan dalam kolom nilai disebut Tabel Dustribusi Frekuensi Bergolong.

2.5. DISTRIBUSI FREKUENSI BERGOLONG

Hasil-hasil Psikotest dari sebagian calon-calon mahasiswa sesuatu Fakultas untuk tahun kuliah 1960/1961 adalah sebagai berikut :

Test : Psychotest,
Subjek : Calon Mahasiswa
Tahun : 1960/1961
Jumlah : 71 orang

18	13	16	4	10	10	15	17	16	16
21	22	20	7	(23)	10	18	(3)	10	8
10	11	10	10	6	11	23	19	19	20
21	12	10	17	7	12	5	9	12	15
12	12	16	20	14	15	14	15	16	15
17	16	16	14	14	15	19	13	15	14
21	8	19	19	19	13	13	19	14	12
20									

Nilai yang tertinggi dari hasil ujian masuk itu adalah 23, sedang nilai terendah adalah tiga. Bilamana kita susun tabel distribusi tunggal, maka kita harus membuatnya sepanjang 21 baris (dari 23 – 3 plus 1). Untuk menyingkat ruangan dan menghemat tenaga, kita dapat mengadakan pengelompokan terhadap nilai-nilai terhadap nilai-nilai itu, misalnya tiap-tiap tiga nilai menjadi satu kelompok. Dengan demikian kita akan menjumpai tabel sebagai berikut :

Tabel 3

Tabel Hasil Psikotest dari Calon-calon mahasiswa Tahun kuliah 1960/1961

Kelompok Nilai	**Frekuensi (f)**
21-23	6
18-20	13
15-17	17
12-14	16
9-11	11
6-8	5
S-5	3
Jumlah	71

Jelas kiranya apa maksud penggolongan-penggolongan dalam distribusi bergolong semacam itu. Bilamana nilai variabel hanya bergerak dari 3 sampai 23 seperti contoh di atas, kiranya mengerjakan distribusi tunggal masih belum terlalu sulit dan menjemukan. Akan tetapi lain halnya bilamana kita harus menyusun tabel tentang Koefisien Intelegensi yang bergerak dari 0 sampai 200, atau tentang penghasilan rakyat yang bergerak dari nul rupiah (yang tidak berpengasilan sama sekali) sampai ratusan ribu rupiah setiap bulannya. Tidak dapat digambarkan bagaimana kita harus membuat tabel tunggal untuk keperluan itu. Mungkin satu tembok dari gedung bertingkat tujuh masih belum dapat memuat tabel tentang penghasilan itu.

2.6. BEBERAPA ISTILAH BARU DALAM DISTRIBUSI BERGOLONG

Interval Kelas. Tiap-tiap kelompok nilai variabel disebut interval kelas. Dalam Tabel 3 diatas kita jumpai ada tujuh interval kelas dengan masing-masing tiga nilai variabel. Kita lihat interval kelas yang paling atas berisi nilai-nilai 21, 22, dan 23, sesungguhnya yang ditulis hanya nilai-nilai 21 dan 23, interval yang terbawah mengandung tiga nilai, yaitu nilai-nilai 3, 4 dan 5, sesungguhnya yang dicantumkan hanya nilai-nilai 3 dan 5.

Interval kelas biasa disingkat dengan sebutan Kelas atau Interval saja. Dalam Statistik bilamana orang menyebut Kelas atau Interval, yang dimaksudkan adalah Interval Kelas.

Batas kelas. Batas kelas adalah nilai-nilai yang membatasi kelas yang satu dari kelas yang lain. Nilai-nilai 21 dan 23 pada kelas yang teratas dari Tabel 3 itu adalah nilai-nilai yang membatasi kelas itu dari kelas lainnya yang berdekatan. Sebab itu angka-angka 21 dan 23 disebut batas kelas. Demikian juga angka-angka dari 18 dan 20 pada interval kelas nomor dua dari atas. Pada kelas yang paling bawah angka-angka 3 dan 5 menjadi batas kelasnya.

Batas Atas dan Batas Bawah. Kita lihat dalam kolom nilai variabel dalam Tabel 3 itu ada dua deret angka-angka batas kelas, deret sebelah kiri dan deret sebelah kanan. Angka-angka batas dideret sebelah kiri ialah angka-angka 21, 18, 15, 12, 9, 6 dan 3. Angka-angka itu semuanya menjadi batas bawah dari masing-masing kelasnya. Sebab itu angka-angka itu disebut “Batas Bawah” (*Lower Limits).*

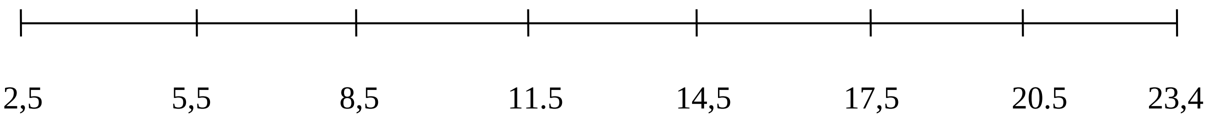
Angka-angka dideret sebelah kanan ialah angka-angka 23, 20, 17, 14, 11, 8, dan 5. Angka-angka ini masing-masing menjadi batas-batas kelasnya sendiri-sendiri. Angka-angka itu disebut “Batas Atas” (*Upper Limits).*

Nampak kepada kita bahwa dalam tabel distribusi bergolong di atas yang dicantumkan dalam kolom nilai variabel hanyalah batas-batas kelasnya, yaitu batas-batas bawah dan batas-batas atas.

Batas Semu dan Batas Nyata. Kalau kita letakkan pada garis mendatar kelas-kelas dalam Tabel 3 itu akan kelihatan sebagai berikut :

3 5 6 8 9 11 12 14 15 17 18 20 21 23

Nampak kepada kita dari pemeriksaan lukisan diatas bahwa angka 5 dan angka 6 misalnya, bukanlah batas yang nyata antara kelas yang terendah dengan kelas diatasnya. Demikian juga angka-angka 8 dan 9, angka-angka 11 dan 12, … …, angka-angka 17 dan 18 dan angka-angka 20-21. Di antara tiap-tiap rangkai angka itu ada *lubangnya.* Lubang ini akan hilang bilamana interval-interval kelas itu kita lukiskan sebagai berikut:



Dalam lukisan diatas itu angka 5,5 misalnya benar-benar merupakan batas yang nyata antara interval kelas yang terendah dengan interval kelas di atasnya. Demikian juga anka 20,5 merupakan batas yang nyata dari interval kelas yang tertinggi dengan interval kelas yang dibawahnya.

Batas- batas kelas yang nyata, yaitu yang mengandung lubang-lubang di antara kelas yang satu dengan kelas yang lainnya, disebut “Batas Semu”, sedang batas-batas kelas yang tidak megandung lubang-lubang di antara kelas yang satu dengan kelas lainnya, disebut “Batas Nyata”.

Perlu ditekankan disini bahwa pengertian akan kedua macam batas kelas ini penting sekali untuk memahami pekerjaan-pekerjaan statistik selanjutnya. Oleh sebab itu keduanya hendak dipahami benar-benar.

Dalam hubungan dengan kedua macam variabel, yaitu variabel yang kontinu dan variabel yang diskrit, perlu dicatat bahwa batas nyata hanya dimiliki oleh variabel yang kontinu. Dalam variabel yang diskrit atau terpisah, misalnya variabel jenis kelamin (laki-laki dan perempuan), pergi dan tinggal, benar dan salah, dan semacamnya, tidak ada batas nyatanya. Jenis kelamin laki-laki dan perempuan adalah terpisah satu sama lain. Tidak ada jenis lain yang dengan nyata membatasi kedua jenis kelamin itu, yaitu jenis yang setengah laki-laki dan setengah perempuan. Demikian juga keadaannya dengan variabel-variabel lain yang disebutkan diatas.

Lebar Kelas. Kita periksa kembali Tabel 3. Interval kelas yang tertinggi ditandai dengan angka-angka 21 dan 23. Kedua angka itu sebenarnya hanyalah batas kelas saja (batas semu). Antara keduanya masih ada satu angka lagi, yaitu angka 22. Demikian juga kelas yang keempat dari atas, sebenarnya mengandung angka-angka 3, 4, dan 5. Jadi tiap-tiap kelas itu sebenarnya atau terdiri atas tiga angka. Inilah yang disebut lebar kelas. ”Lebar Kelas” adalah jumlah nilai-nilai variabel dalam tiap-tiap kelas. Tabel 3 diatas mempunyai lebarr kelas tiga.

Secara matematik lebar kelas dapat didefinisikan sebagai *bayas atas nyata dikurangi batas bawah nyata* dari kelas-kelas yang bersangkutan.

Lebar kelas biasa diberi symbol “i” atau “h”. Disini kita akan memakai salah satu dari kedua sebutan yang lazim itu, yaitu singkatan “i”. Bilamana orang mengatakan “i” sama dengan tiga, ini berarti bahwa distribusi frekuensi disusun dalam tabel atau grafik yang menggunakan interval kelas dengan isi tiga angka atau nilai dalam tiap-tiap intervalnya.

Titik tengah. Yang dimaksud dengan “Titik Tengah” adalah angka atau nilai variabel yang terdapat ditengah-tengah interval kelas. Bilamana interval kelas memuat angka-angka 13, 14, dan 15, yang menjadi titik tengahnya adalah angka 14. Bilamana interval kelas memuat angka-angka 20, 21, 22, 23 dan 24, titik tengahnya adalah angka 22. Bilamana luas kelasnya genap, seperti 20, 21, 22 dan 23, titik tenaghnya adalah separuh dari jumlah angka-angka tengah, yaitu 21,5 (dari ½ x (21 ditambah 22) atau separuh dari angka-angka tengah yaitu ½ (20 + 23) = 21,5. Titik tengah kadang-kadang disebut juga tanda kelas.

Jumlah Interval. Yang disebut jumlah interval ialah banyaknya interval yang digunakan dalam penyusunan distribusi. Dalam Tabel 3 diatas jumlah intervalnya ada tujuh.

Jarak Pengukuran. Kalau kita mengukur tinggi sejumlah orang, dan kita mempunyai angka pengukuran yang tertinggi 180 cm dan angka pengukuran yang terendah, 145 cm, kita mempunyai jarak pengukuran 35 cm. (dari 180 cm dikurangi 145 cm). yang dimaksud dengan jarak pengukuran ialah angka tertinggi dari pengukuran dikurangi dengan angka terendah. Bilamana nilai-nilai Aljabar dalam suatu kelas yang tertinggi adalah delapan, dan yang terendah tiga, jaraknya lima (dari 8 dikurangi 3). Jarak pengukuran itu biasa disebut *“Range of Measurement”.* Disingkat dengan huruf R.

Sebenarnya hanya dimiliki oleh varible yang continue saja. Antara laki-laki dan perempuan tidak ada jaraknya. Demikian juga antara benar dan salah. Akibatnya ialah bahwa sebenarnya R bukanlah nilai variable tertinggi dikurangi nilai variable terendah. Kita ambil contoh tinggi orang. Angka tertinggi 180 cm. sebenarnya mencakup juga mereka yang mempunyai tinggi 180,49 cm. angka terendah 145 cm mencakup juga tinggi 144,50 cm karena itu R adalah 180,49 cm, dikurangi 144,50 cm, sama dengan kira-kira 36 cm. jadi R adalah batas nyata atas (*upper real limit)* dari nilai variable yang tertinggi dikurangi dengan batas nyata bawah (*lower real limit)* dari nilai variable yang terendah akan tetapi untuk mudahnya R adalah nilai tertinggi dikurangi dengan nilai terendah, tidak memandang batas nyatanya.

2.7. MENETAPKAN JUMLAH INTERVAL

Salah satu masalah yang kita hadapi bila kita hendak menyusun tabel dengan interval-interval adalah dengan menetapkan jumlah interval. Penetapan ini dipengaruhi oleh beberapa factor, antara lain adalah factor-faktor jumlah frekuensi (N), jarak pengukuran (R), lebar inbterval yang hendak digunakan (i) dan tujuan penyusunan distribusi itu. Pada prinsipnya jumlah interval kelas janganlah terlalu sedikit, sehingga pola-pola kelompok menjadi kabur. Kalau kita membuat tabel tentang angka-angka suatu mata pelajaran hanya menjadi dua kelompok (tabel dengan dua interval), maka kita hanya mempunyai gambaran tentang anak-anak yang mendapat angka-angka “di atas” dan “di bawah” atau anak-anka yang mendapat angka-angka dalam “baik” dengan anak-anak yang mendapat angka “kurang”. Kita tidak akan mengetahui, berapa anak-anak yang “sedang”, “cukup”, dan sebagainya. Akan tetapi jumlah interval itu juga jangan terlalu besar sehingga kita tidak dapat mendapat gambaran tentang pola kelompok. Petunjuk yang mutlak tidak ada dalam hal ini. Tetapi dalam psikologi dan pendidikan dapat kita anut kebiasaaan yang menggunakan 5 sampai 15 interval. Kalau R besar sekali, biasanya orang menggunakan 10 sampai 20 interval. Akan tetapi hal ini tidak boleh diikuti secara membabi buta.

2.8. MENENTUKAN LEBAR INTERVAL (i)

Bilamana R sudah diketahui dan jumlah interval kelas sudah ditentukan, pada dasarnya i sudah di ketemukan . rumus dari i adalah sebagai berikut:

$$i = \frac{\text{Jarak Pengukuran (R)}}{\text{Jumlah Interval}}$$

Jadi kalau misalnya hasil pengukuran kita tentang tinggi orang yang tertinggi adalah 180 cm dan yang terendah adalah 145 cm dan kita telah menetapkan interval sebanyak 9 buah, maka;

$$i = \frac{180{,}5 - 144{,}5}{9} = \frac{36}{9} = 4$$

Dalam tabel kemudian kita mencantumkan dalam kolom pertama interval-interval kelas berturut-turut dari atas kebawah sebagai berikut:

Interval Tinggi Badan
177-180
173-176
169-172
165-168
161-164
157-160
153-156
149-152
145-148

Suatu hal yang mungkin membingungkan kita untuk menetapkan i ialah bagaimana dengan rumus itu kita tidak memperoleh angka yang bulat, misalnya 4,50. Hal ini tidak mengapa, dan kita dapat bekerja terus dengan i=4,50. Akan tetapi karena kita biasanya tidak senagng bekerja dengan terlalu banyak angka-angka pecahan (juga untuk mengurangi waktu, tenaga, dan kemungkinan membuat kesalahan) maka kita dapat mengorbankan jumlah intervalnya, sehingga kita memperoleh i yang bulat. Kalau dalam contoh diatas kita sudah menetapkan jumlah interval sebanyak delapan buah, maka i-nya akan menjadi 4,50. Kita dapat mengubah jumlah intervalnya dari delapan menjadi Sembilan buah, sehingga tabel kita mempunyai pengintervalan seperti diatas. Kalau tidak maka tabel kita akan berwujud sebagai berikut:

Interval Tinggi Badan
176,50-180,00
172,00-175,50
167,50-171,00
163,00-166,50
158,50-162,00
154,00-157,50
149,50-153,00
145,00-148,00

Kemungkinan yang kedua untuk menghindari i yang berupa bilangan pecahan adalah mengubah R. Ini terjadi kalau misalnya kita, dalam contoh tinggi badan di atas, ingin mengunakan i = 5, dan batas bawahnya merupakan bilangan kelipatan angka lima [1]).

Dengan demikian interval kelas yang teratas akan bermulah dengan angka 180. Jadinya interval kelas-kelas kita adalah bagian berikut:

Interval Tinggi Badan
180-184
175-179
170-174
165-169
160-164
155-159
150-154
145-149

Ini biasa terjadi karena orang lebih suka bekerja dengan bilngan-bilangan kelipatan angka 5. Perubahan R dari (180,50-144,50) menjadi (184,50-144,50) itu mempunyai pengaruh-pengaru yang berarti untuk pengelolahan-pengelolahan selanjutnya.

Satu hal lagi yang perlu kita catat ialah kecenderungan untuk menetapkan i yang berupa bilangan gasal. Kecenderungan ini ada dasarnya juga, yaitu untuk perhitungan-perhitungan Statistik selanjutnya. Memerlukan titik tegah dari tiap-tiap interval. Hal ini akan kita mengerti setelah nanti kita mempelajari menghitung apa yang disebut Mean, Median, Standard Deviasi, dan sebagainya. Sementara ini dapat dikatakan bahwa dengan menetapkan i yang gasal kita akan meeperoleh titik tengah yng merupakan bilangan bulat. Titik tengah pecahan akan menimbulkan konsekuensi penambahan pekerjaan yang tidak sedikit untuk pengolahan-pengolahan statistik selanjutnya.

2.9. DISTRIBUSI FREKUENSI MENINGKAT (CUMULATIVE FREQUENCY DISTRIBUTION)

Penyusunan Tabel Distribusi Frekuensi Meningkat ini pada dasarnya sama saja dengan penyusunan distribusi frekuensi tunggal maupun distribusi frekuensi bergolong. Bedanya dengan penyusunan kedua distribusi itu adalah bahwa di sini kita menambahkan satu kolom lagi yang memuat frekuensi meningkat. Untuk contoh baiklah kita ambil kembali bahan-bahan dari Tabel 2.

Tabel 4

Tabel Nilai-Nilai Berhitung 72 Orang Murid Laki-Laki Sekolah Rendah di Kotapraja Y

Nilai	Frekuensi	Frekuensi Meningkat dari Bawah
8	4	72
7	23	68
6	28	45
5	16	17
4	1	1
Jumlah	72	-

Jadi di sini kita mempunyai kolom lainnya, yaitu kolom yang disediakan untuk frekuensi meningkat. Frekuensi meningkat bisa disebut dengan huruf cf, singkatan dari bahasa asing "*cumulative frequency*", yang artinya "frekuensi meningkat". Frekuensi ini diperoleh dari menjumlahkan secara meningkat frekuensi-frekuensi yang ada di kolom kedua. Angka 17 diperoleh dari menjumlahkan 1 (frekuensi nilai 4) dan 16 (frekuensi nilai 5) ; angka 45 dari angka 17 (jumlah meningkat dari frekuensi-frekuensi nilai-nilai 4 dan 5) dan 28 (frekuensi nilai-nilai 4,5,dan 6) ditambah 23 ; sedang angka 72 pada baris 8 itu adalah jumlah dari frekuensi-frekuensi nilai 4 sampai 7 (semuanya ada 68) dengan angka 4 (frekuensi nilai 8). Perlu dicatat bahwa bilamana kita mengisi kolom frekuensi meningkat dari bawah, maka jumlah frekuensi meningkat yang paling atas harus sama dengan N (dalam contoh di atas jumlah itu adalah 72, cocok dengan N yang juga 72).

Apa gambaran kasar yang dapat kita peroleh dari tabel dengan frekuensi meningkat semacam Tabel 4? Bagaimana buat sementara ini kita lepaskan dulu maksud-maksud yang jauh lebih jauh dari perbuatan distribusi frekuensi meningkat, maka secara sederhana kita dapat melihat

dengan cepat banyaknya anak yang mendapat “sesuatu nilai kebawah”. Mereka yang mendapat nilai enam kebawah misalnya, jumlahnya ada 45 orang, sedang mereka yang mendapat nilai tujuh kebawah jumlahnya ada 68 orang. Demikian seterusnya.

Contoh Tabel distribusi frekuensi dari distribusi bergolong adalah sebagai berikut :

TABEL 5

Tabel Contoh Nilai-Nilai Hasil Test A

Interval Nilai	Frekuensi (f)*	Frekuensi Meningkat dari Bawah	Frekuensi Meningkat dari Atas
70-74	1	48	1
65-69	3	47	4
60-64	4	44	8
55-59	9	40	17
50-54	9	31	26
45-49	11	22	37
40-44	5	11	42
35-39	4	6	46
30-34	2	2	48
Jumlah	48	-	-

*) f = singkatan dari frekuensi. Jumlah diberi smbol f.
Jadi $\sum f = N$

Dari contoh diatas kelihatan dengan jelas bahwa pada hakikatnya tidak ada perbedaan antara penyusunan tabel frekuensi meningkat dari distribusi tunggal dengan penyusunan tabel frekuensi meningkat dari distribusi bergolong. Dalam contoh tersebut dicantumkan kolom untuk menyebutkan frekuensi meningkat dari atas, disamping frekuensi meningkat dari bawah. Ini tidak berarti bahwa penyusunan tabel frekuensi meningkat dari distribusi berrgolong harus mencantumkan dua macam frekuensi. Hal ini bergantung pada keperluan dari penyelidikan selanjutnya.

Bilamana suatu penyelidikan membutuhkan juga gambaran frekuensi meningkat dalam persen, kita dapat menambahkan kolom lain yang mencantumkan freekuensi meningkat dalam persen dalam tabel kita. Contoh adalah sebagai berikut :

Tabel 6

Tabel Contoh Nilai-Nilai Hasil Test dengan Frekuensi Meningkat dan Persen

Interval nilai	Frekuensi (f)	Frekuensi meningkat dari bawah	Frekuensi meningkat dari bawah dalam % *)
70 – 74	1	48	100
65 – 69	3	47	98
60 – 64	4	44	92
55 - 59	9	40	83
50 – 54	9	31	65
45 – 49	11	22	46
40 – 44	5	11	23
35 – 39	4	6	13
30 - 34	2	2	4
Jumlah	48	-	-

*) Persentase dibulatkan : setengah ke atas dibulatkan menjadi satu, dan di bawah setengah dihilangkan

2.10. APAKAH DISTRIBUSI FREKUENSI ITU ?

Di muka kita belum membatasi apa yang disebut distribusi frekuensi itu. Dari uraian-uraian tersebut di atas kita dapat menyimpulkan bahwa distribusi frekuensi adalah *penyusunan bahan-bahan atas dasar nilai variabel dan frekuensi tiap-tiap nilai variabel tersebut.* Tabel untuk distribusi frekuensi disebut Tabel Distribusi Frekuensi, Tabel Distribusi atau Tabel Frekuensi saja. Distribusi tunggal adalah frekuensi yang tidak menggunakan penggolong-penggolongan. Distribusi bergolong mengginakan interval-interval kelas dalam penyusunannya. Distribusi frekuensi meningkat mencantumkan frekuensi meningkat dalam penyusunannya, dan dapat disusun baik dari distribusi tunggal maupun dari distribusi bergolong.

POKOK BAHASAN:

“MEMBUAT DAN MENYAJIKAN GRAFIK“

Diantara penyajian – penyajian statistikyang kerap kali kita jumpai dalam majalah - majalah , brosur-brosur, bulletin-buletin, monograp-monograp, dan buku-buku ilmiah adalah penyajian grafik. Bagi pekerjaan kita sehari-hari, yang tidak kurang pentingnya dari pengetahuandan kecakapan membaca grafik adalah membuat sendiri dan menyajikan grafik.

Ada banyak macam grafik. Dalam bab ini akan dibicaraka tiga diantaranya, dipilih yang penting-penting saja. Ketiga grafik itu ialah histrogram, frekuensi polygon, dan ogive.

3.1 TABEL PERSIAPAN

Untuk membuat grafik dari suatu bahan, diperlukan kita terleih dahuliu membuat tabelnya. Table semacam ini disebut Tabel persiapan, yang mungkin turut disajikan kepada pembaca tentang hasil penyelidikan kita, mungkin tidak pembuatan table persiapan ini akan sangat memudahkan pekerjaan kita dalam memuat grafik yang kita maksudkan.

3.2 LANGKAH-LANGKAH UMUM DALAM MEMBUAT GRAFIK

Karena ada langkah-langkah yang sama dalam membuat semua macam grafik, maka perlu kiranya terlebih dahulu dibicarakan langkah-langkah yang umum itu.

1. *Sumbu Absis dan ordinat.* Untuk membuat grafik, kita selalu mengunakan system sumbu, yaitu sumbu absis dan sumbu ordinat. Sumbu Absis yaitu sumbu yang mendatar, disebut Sumbu “X” (huruf X besar), sedangkan sumbu ordinat, sumbu yang menegak disebut sumbu “Y” (huruf Y besar). Sumbu X (biasanya) disediakan untuk mencantumkan Nilai, sedangkan sumbu Y untuk frekuensi.
2. *Perbandingan antara X dan Y.* Sudah menjadi kelaziman bahwa sumbu X dibuat lebih panjang dari pada sumbu Y. Perbandingan antara keduanya kira – kira adalah sepuluh dengan tujuh atau sepuluh dengan delapan, atau pada umumnya tiga banding dua. Kelaziman ini sampai sekarang masih dipertahankan, dengan maksud agar tidak memerikan gambaran yang keliru kepada pembacanya pada pembacaan

yang sepintas lalu. Hanya untuk maksud – maksud tertentu kelaziman itu dilanggar orang misalnya untu keperluan propaganda.

3. *Pemberian nama pada sumbu.* Untuk memudahkan pembacaan maka tiap – tiap sumbu diberikan nama sesuai dengan maksudnya. Sumbu X diberi nama Nilai dibawahnya tepat ditengah-tengahnya, sedangkan sumbu Y diberi nama frekuensi disebelah kirinya, juga tepat ditengah – tengahnya, atau tepat diatasnya
4. *Pemberian nama pada grafik.* Grafik yang ada namanya akan sangat membingungkan pembacanya. Sebab itu tiap – tiap grafik yang dimaksudkan untuk disajikan kepada pembaca harus diberi nama. Nama ini menurut kelaziman dicantumkan dibawah grafik, buka diatasnya (sebaliknya table mencantumkan keterangannya diatasnya, bukan dibawahnya)

Keempat hal diatas perlu kita perhatikan betul – betul dalam penyajian grafik. Hal ini ditekankan karena dalam pengalaman tidak sedikit orang yang baru pada taraf permulaan membuat table merupakan hal – hal kecil yang sangat penting itu.

3.3. HISTROGRAM

Grafik Histogra biasa disebut juga *Bar Diagram*, yaitu suatu grafik yang berbentuk eberapa segi empat.

Marilah kita mencoba membuat grafik Histogram dari ahan yang telah ada di Bab I. kita reproduksi Tabel 2 dengan menambah satu kolom untuk mencantumkan “Batas Nyata”-nya. Kolom ini sengaja dicantumkan untuk memudahkan pekerjaan kita pada taraf belajar ini. Akan tetapi untuk selanjutnya kolom ini tidak perlu kita cantumkan.

Tabel 7

Nilai nilai berhitung 72 orang murid laki – laki SD dalam kotapraja Y

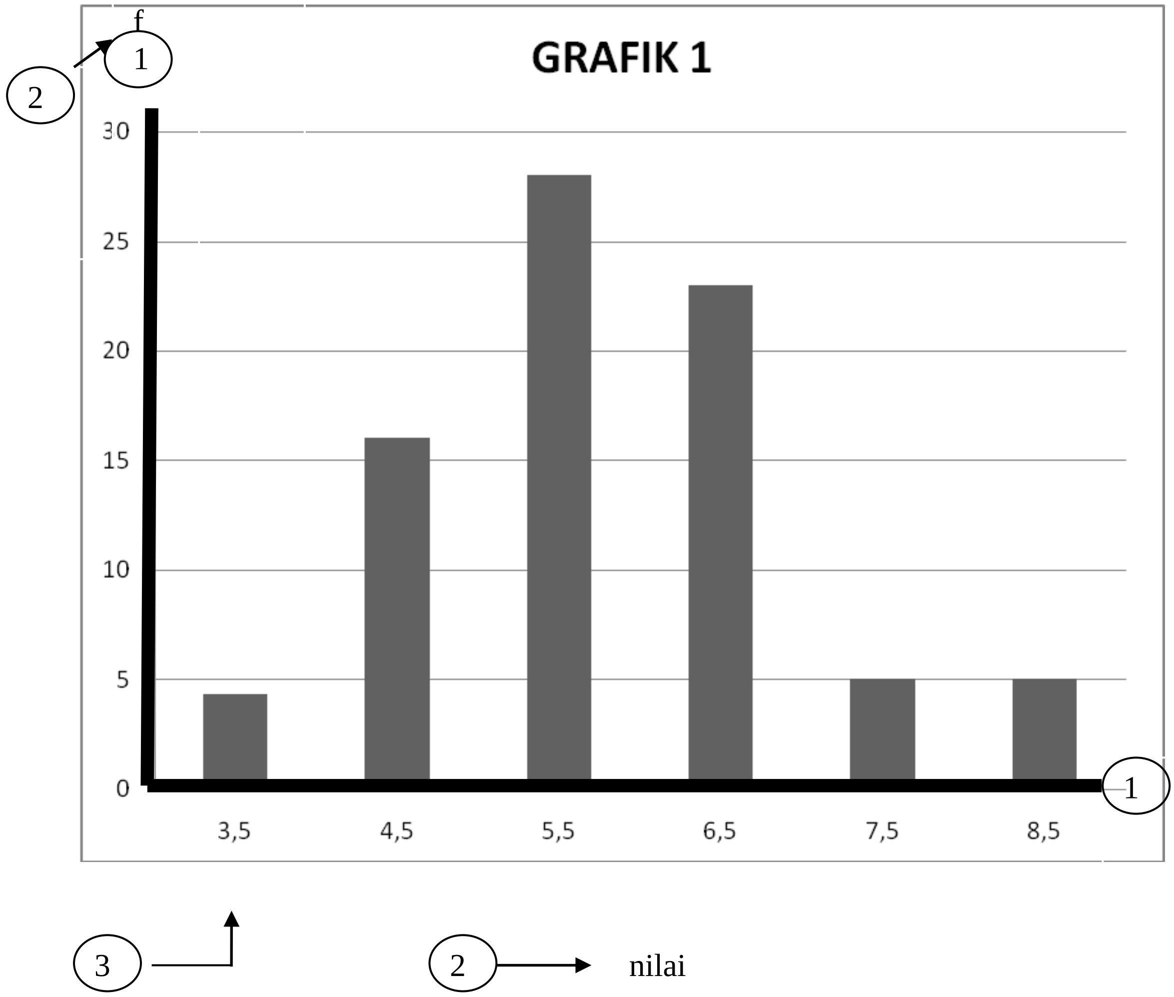
NILAI	BATAS NYATA+	FREKUENSI
	8,5	
8	7,5	4
7	6,5	23
6	5,5	28
5	4,5	16
4	3,5	1
Jumlah	-	72

+ Batas Nyata Atas dan Bawah dijadikan satu

LANGKAH-LANGKAH MEMBUAT HISTROGRAM

1. Membuat absis dan ordinat, berbanding seperti 10:7.
2. Absis kita beri nama “Nilai” dan ordinat “Frekuensi”, atau f.
3. Membuat skala pada absis dan ordinant. Perskalaan pada absis ini tidak perlu sama dengan kebutuhan perskalaan pada ordinant. Hal ini harus kita sesuaikan dengan kebutuhan kita. Yang penting adalah skala-skala pada absis harus dapat memuat semua nilai (dan oleh karena histogram dibuat atas dasar batas nyata, maka skala-skala pada ordinat harus dapat memuat frekuensi tertinggi).
4. Mendirikan segiempat-segiempat pada absis. Tinggi masing-masing segi empat harus sama dengan (sesuai dengan) frekuensi tiap-tiap nilai variabelnya. Segiempat-segiempat ini berimpit satu sama lain pada batas nyatanya.
5. Pembuatan Histogram ini kita selesaikan dengan memberi “Keterangan” selengkapnya tentang “Apa” Histogram itu kita buat dan sajikan.

Dengan memeriksa Grafik I kita mengerti untuk apa kolom batas nyata dicantumkan dalam tabel persiapan kita. Batas nyata itu diperlukan karena histogram lazimnya dibangun diatas batas nyata, sungguh pun per-cantuman itu, seperti telah telah dikatakan dimuka bukan suatu keharusan. Pada taraf belajar, sebaiknya kita tidak segan-segan menambahkan kolom batas nyata pada tabel persiapan kita.

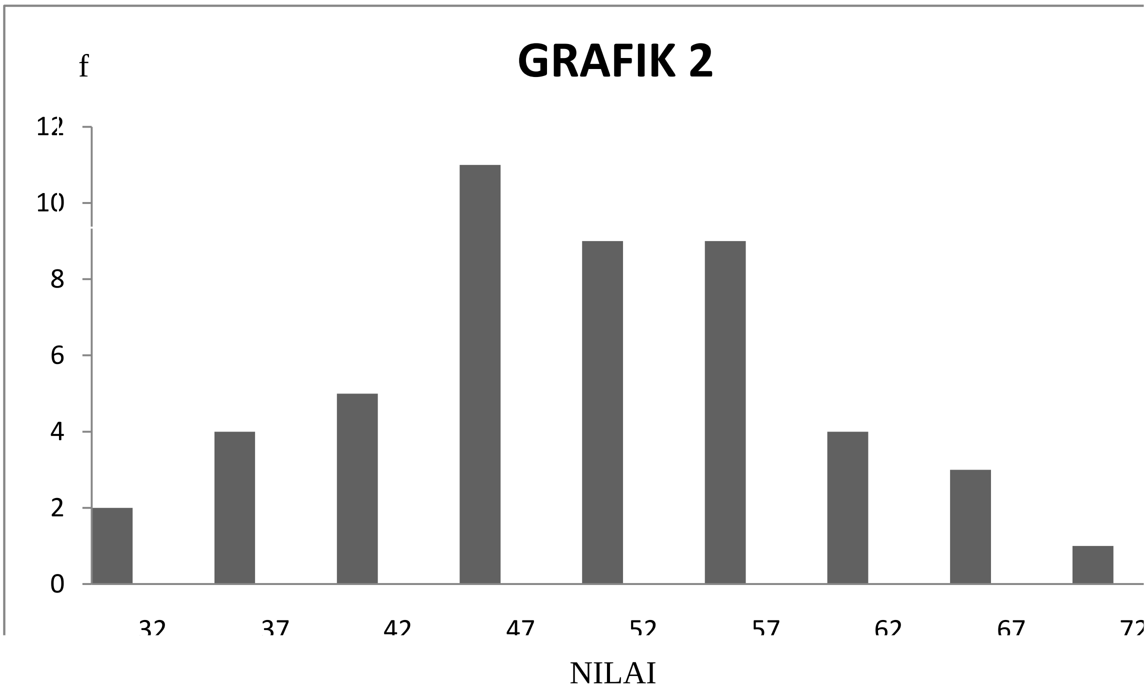


Perlu kita catat bahwa sungguhpun lazimnya histogram dibentuk diatas batas nyata, namun sekarang ada kecenderungan untuk membuat histogram dengan menggunakan titik tengah pada nilai variabelnya. Sudah barang tentu dalam hal semacam itu kita tidak perlu mempersiapkan kolom batas nyata dalam table kita. Sebagai gantinya kita cantumkan titik tengahnya. Sebagai contohnya dapat kita lihat histogram dibawah ini. Histogram inidibuat dari distribusi bergolong dari bahan tabel 5 dalam Bab I.

TABEL 8

Table contoh nilai – nilai hasil test A

Interval Nilai	Titik Tengah (x)	Frekuensi (f)
70-74	72	1
65-69	67	3
60-64	62	4
55-59	57	9
50-54	52	9
45-49	47	11
40-44	42	5
35-39	37	4
30-34	32	2
Jumlah	-	48



Histrogram menunjukan distribusi Nilai – Nilai Hasil Test A terhadap 48 orang

Jelaslah, bahwa tidak ada perbedaan pokok-pokok pembuatan histogram dengan menggunakan batas nyata dengan pembuatan histogram dengan menggunakan titik tengah. Yang berbeda hanyalah nilai-nilai yang dicantumkan pada absis; yang satu mencantumkan batas nyata, sedang yang lain mencantumkan titik tengah.

Perlu dicatat bahwa baik histogram yang menggunakan batas nyata maupun histogram yang menggunakan titik tengah, keduanya dapat dibuat dari distribusi tunggal maupun distribusi bergolong.

Untuk memahami benar-benar cara-cara menyajikan histogram, kita tinjau lagi secara sungguh-sungguh histogram diatas. Ada dua sumber : pertama sumbu mendatar dan kedua sumbu menegak. Nilai variable (dalam contoh ditas itu adalah titik-titik tengah), dicantumkan pada sumbu mendatar, sumbu absis, sedang sumbu tegaknya, sumu ordinat memuat frekuensi dari tiap-tiap interval kelas yang bersangkutan. Lebar tiap-tiap segiempat itu memanjang dari atas rendah nyata ke batas atas nyata, berimpit satu sama lain, sehingga tidak akan ada lubang diantara tiap-tiap segiempat itu.

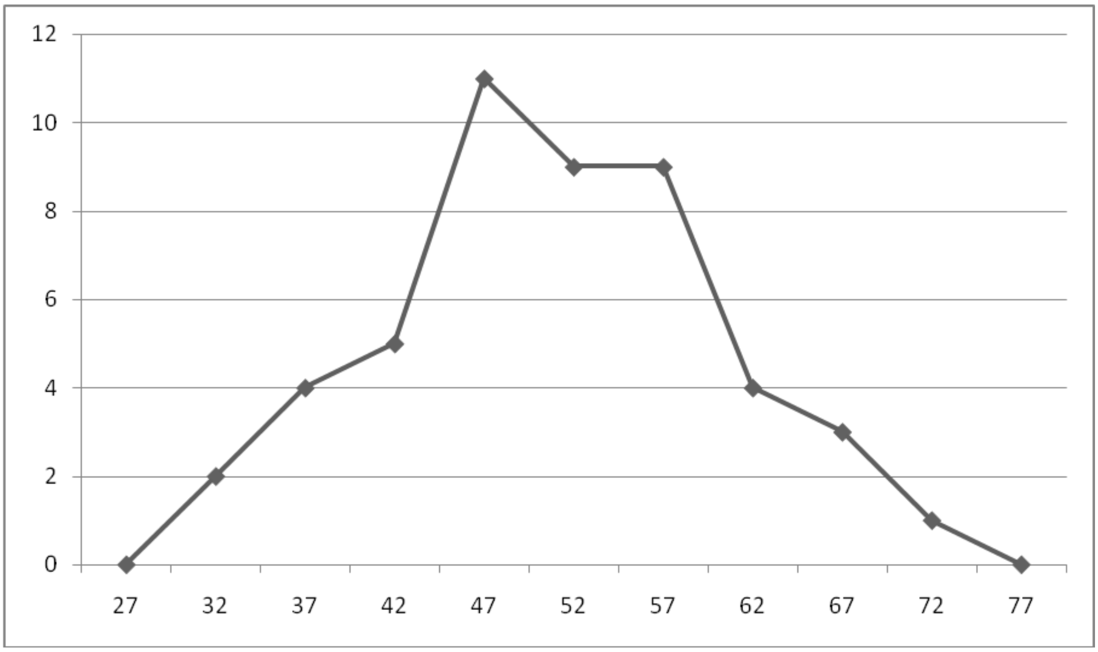
3.4. POLIGON

Pada dasarnya tidak ada perbedaan yang penting antara grafik histogram dengan grafik polygon. Perbedaannya hanyalah terletak pada :

1. Grafik hostogram ”lazimnya” dibuat dengan menggunakan batas nyata sedang grafik poligon selalu menggunakan titik tengah.
2. Grafik histogram berwujud segiempat-segiempat, sedang grafik polygon berwujud garis-garis atau kurve (garis-garis yang sudah dilicikan).

Grafik poligon yang biasa juga disebut grafik poligon frekuensi, dibuat dengan menghubung-hubungkan titik-titik tengah tiap-tiap interval kelas secara berturut-turut. Dengan menghubungkan kedua ujungnya ke titik tengah interval kelas didekatnya (di kedua ujungnya) maka akan selesailah pembuatan poligon itu. Sebagai contoh mari kita buat poligon dari bahan-bahan dalam tabel 8 (periksa grafik 3).

Akibat dari keharusan menyelesaikan poligon dengan menarik garis-garis ke absis pada kedua ujung distribusi itu ialah bahwa kita harus membagi-bagi panjang absis menjadi lebih banyak daripada jumlah interval kelas dalam distribusi. Dalam contoh diatas kita membagi-bagi absis menjadi kira-kira dua belas, tidak hanya menjadi sembilan (jumlah interval kelas ada sembilan).



Nilai

Grafik 3

Poligon menunjukan distribusi nilai-nilai hasil test A Terhadap 48 orang

Dengan grafik poligon kita dengan mudah dapat membandingkan keadaan dua distribusi, bilamana kedua distribusi itu dilukiskan dalam satu grafik. Hal ini dapat kita lihat dengan jelas pada keadaan sebagai berikut :

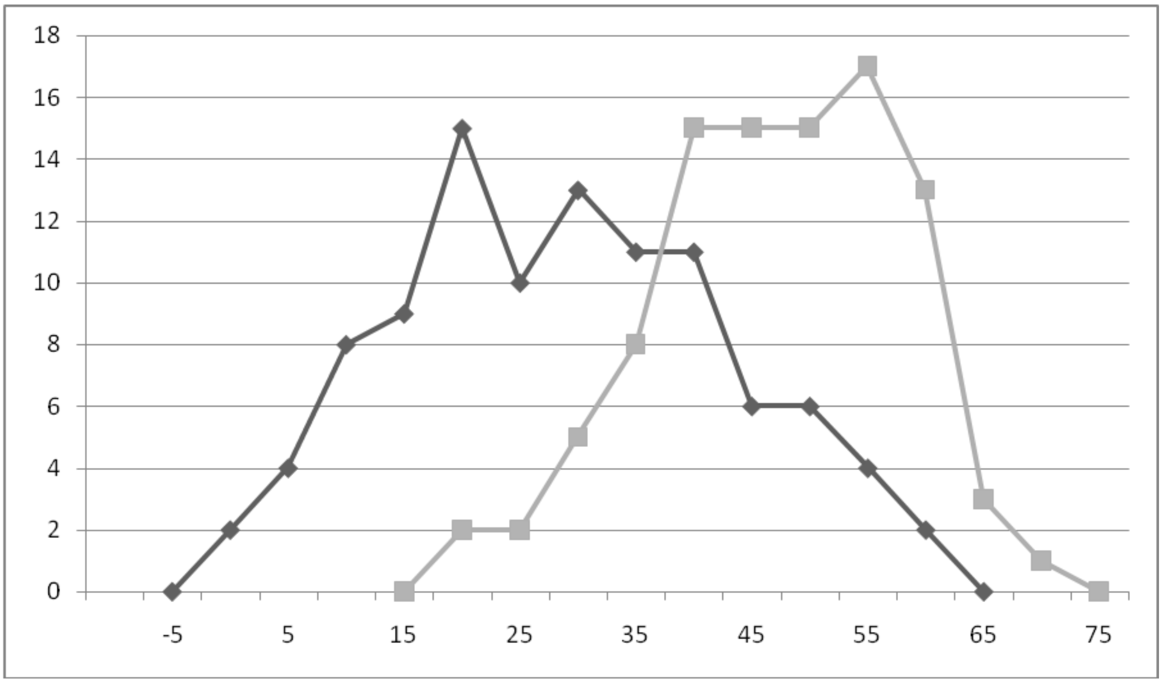
Tabel 9

Tabel distribusi frekuensi (dalam persen) tentang perhatian terhadap lapangan keinsinyuran dari para insinyur dan mahasiswa-mahasiswa tingkat I dari fakultas teknologi

Insinyur		Mahasiswa Tingkat I	
Nilai tengah Nilai perhatian	Frekuensi Dalam % x)	Nilai tengah Nilai perhatian	Frekuensi Dalam %x)
70	1	60	2
65	3	55	4
60	13	50	6
55	18	45	6
50	16	40	11
45	16	35	11
40	6	30	13
35	8	25	9
30	5	20	15
25	2	15	9
20	2	10	8
		5	4
		0	2
Jumlah :	100%	Jumlah :	100%

x) frekuensi dalam % biasa juga disebut frekuensi relatif, disingkat F%

Dengan menyajikan dua distribusi tentang suatu hal dalam satu grafik itu kita akan lebih mudah dan lebih cepat menarik beberapa kesimpulan. Kita lihat bahwa nilai-nilai perhatian keinsinyuran dan insinyur-insinyur mengelompok disekitar angka 50. Sebaliknya nilai-nilai perhatian keinsinyuran dari mahasiswa-mahasiswa tingkat I menunjukan penyebaran (biasa dalam statistik disebut “variabilitas”) yang lebih besar dan mengelompok di sekitar nilai-nilai yang lebih rendah (menunjukan rendahnya perhatian keinsinyuran).



Grafik 4

Frekuensi Poligon Bahan Tabel 9

3.5. OGIVE

Grafik ogive biasa disebut juga grafik frekuensi meningkat. Grafik macam ini agaknya tidak begitu sering kita jumpai dalam buku-buku bacaan. Akan tetapi kadang-kadang grafik ini juga ada gunanya. Seorang perancang pakaian mode-mode baru mungkin ingin mencatat dalam bentuk grafik perkembangan penjualan modenya dalam setahun dalam jumlah meningkat. Seorang perancang test ingin menyelidiki tiap-tiap soal test yang dirancangnya menurut keurutan kesuakarannya. Seorang ahli penduduk mungkin tertarik untuk mencatat secara meningkat jumlah kematian dan atau kelahuran dari tahun ke tahun. Dan sebagainya. Penyajian ogive rupa-rupanya sangat cocok untuk maksud-maksud tersebut.

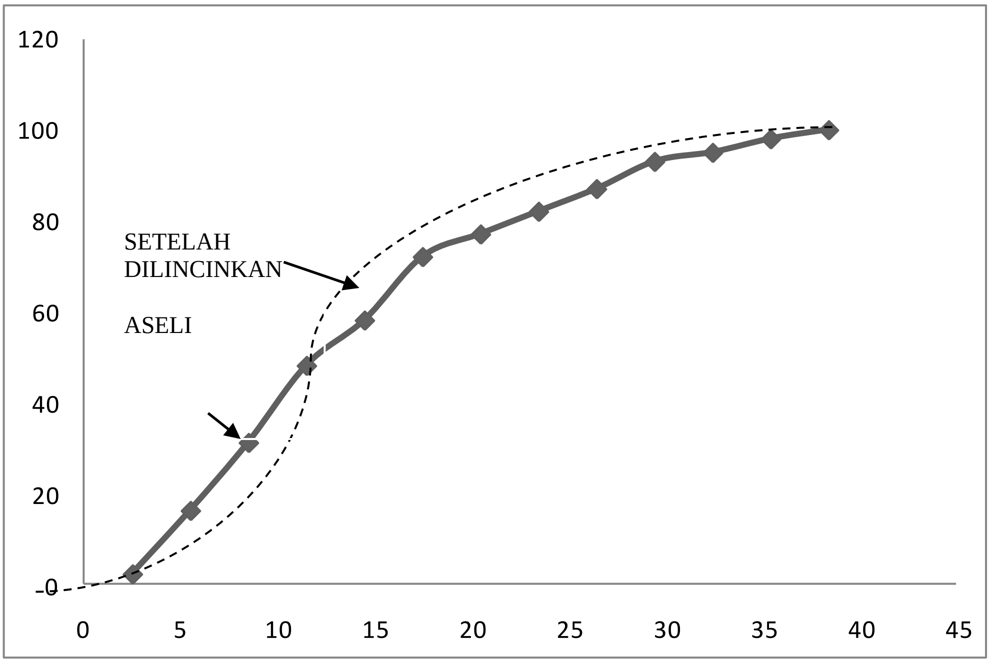
Ogive dapat dibuat baik dari distribusi tunggal maupun dari distribusi bergolong. Di bawah ini diberikan contoh untuk membuat grafik ogive dari distribusi bergolong.

Tabel 10

Tabel distribusi khayalan tentang suatu hal untuk contoh membuat grafik ogive

Interval Nilai	**Batas Nyata**	**Frekuensi**	**Frekuensi Meningkat**
36-38	38,5	2	100
33-35	35,5	3	98
30-32	32,5	2	95
27-29	29,5	6	93
24-26	26,5	5	87
21-23	23,5	5	82
18-20	20,5	5	77
15-17	17,5	14	72
12-14	14,5	10	58
9-11	11,5	17	48
6-8	8,5	15	31
3-5	5,5	14	16
0-2	2,5	2	2
	0,5		
Jumlah	N = 100		

Pembuatan ogive dimulai dengan cara-cara seperti membuat grafik lainnya : membuat sumbu absis dan ordinat, berbanding kira-kira seperti satu banding tiga perempat; memuat skala pada absis untuk mencantumkan batas-batas nyata, dan skala pada ordinat untuk mencantumkan frekuensi meningkatnya, absis kita beri nama “nilai” dan ordinat kita beri nama “frekuensi meningkat”; dengan menarik garis-garis dari batas bawah disebelah kiri berturut-turut ke batas nyata diatasnya pada ketinggian menurut frekuensi interval-interval yang bersangkutan, maka seleailah grafik ogive kita; grafik ini kita sempurnakan dengan mencantumkan “keterangan” yang diperlukan untuk penyajian itu. Jadinya grafik ogive kita adalah sebagai berikut :



Grafik 5

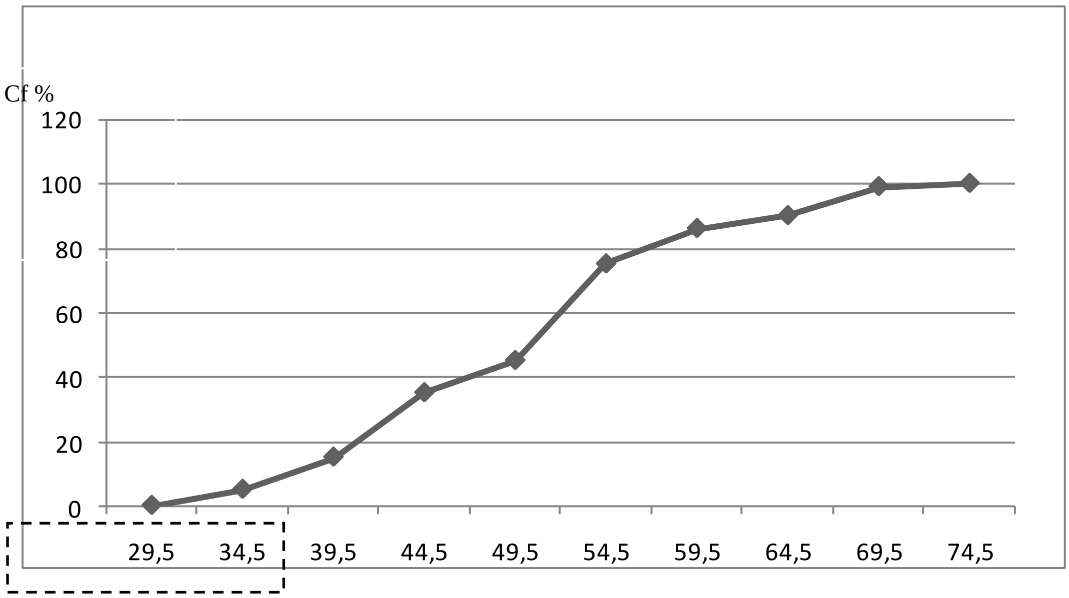
Ogive menunjukan distribusi nilai sesuatu hal dari bahan tabel 10

Perlu kita catat bahwa grafik ogive itu dibuat dengan menggunakan “batas nyata”, bukan titik tengah sebagaimana grafik poligon.

Grefik ogive dapat juga dibuat dengan frekuensi meningkat dari atas. Pokok-pokok pembuatan ogive semacam ini tidak berbeda dengan pembuatan ogive dengan frekuensi meningkat dari bawah. Tabel persiapannya tidak lagi mencantumkan frekuensi meningkat dari bawah, melainkan frekuensi meningkat dari atas dalam kolom frekuensi meningkatnya.

Bilamana kita periksa dengan seksama maka akan kita lihat bahwa pembuatan grafik ogive itu pada dasarnya tidak berbeda dengan pembuatan grafik poligon. Perbedaannya ialah: pada grafik ogive ini kita mencantumkan frekuensinya secara meningkat, sedang pada poligon kita mencantumkan frekuensi tiap-tiap nilai variabel; dan ogive selalu menggunakan batas nyata, sedang poligon selalu menggunakan titik tengah.

Kita juga dapat membuat grafik ogive dalam persen. Untuk ini yang kita cantumkan dalam grafik bukan lagi frekuensi meningkat dalam satuan angka biasa, melainkan frekuensi meningkat dalam persen. Sudah barang tentu untuk menyelesaikan grafik semacam ini kita harus mencantumkan kolom frekuensi meningkat dalam persen dalam tabel persiapan kita. Sebagai contohnya kita ambil bahan dari tabel 6.



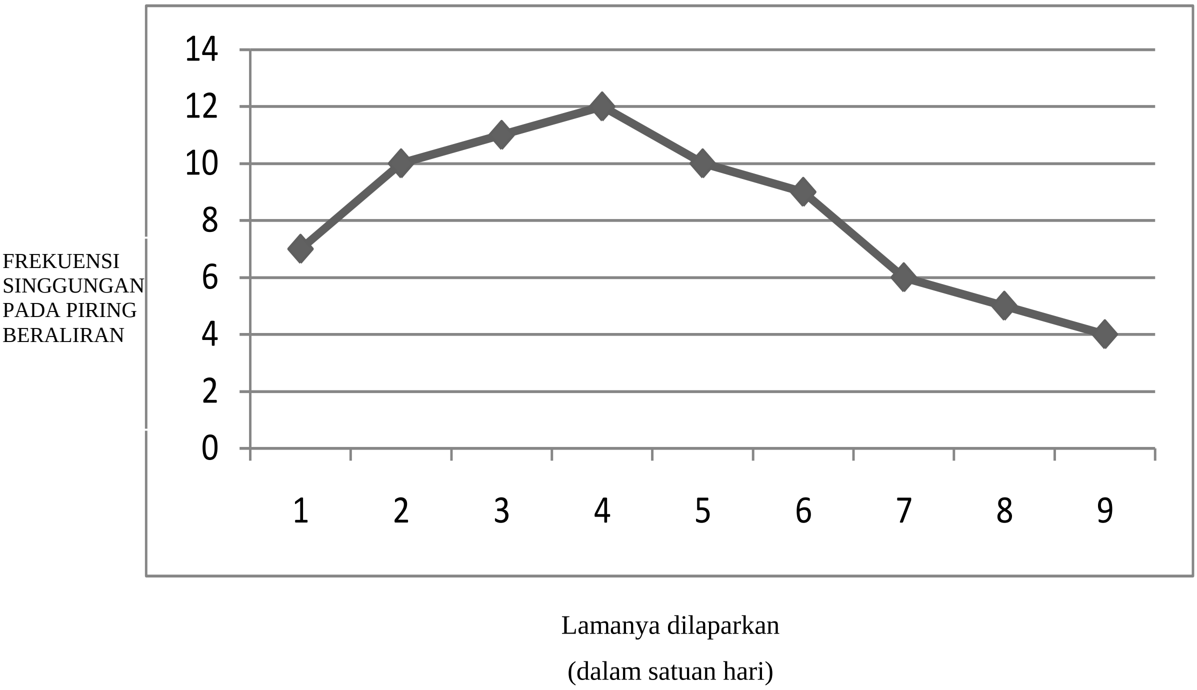
GRAFIK 6

Grafik ogive ini biasanya dipilih orang daripada hihstogram atau poligon bilamana orang ingin mengetahui “kedudukan” seseorang tentang sesuatu hal dalam kelompoknya sendiri, bukan pola sifat atau kecakapan kelompok seluruhnya. Kita ambil missal fahrudin yang hanya mendapat nilai 42 dalam test pertukangan. Kalau grafik ogive diatas menunjukan distribusi hasil-hasil test pertukangan, kita menanyakan apakah kepada fahrudin dapat kita serahkan tugas memimpin perusahaaan pertukangan yang kita dirikan? Ataukah dia kita terima untuk menjadi tukang pelitur, atau kita tolak saja lamarannya? Dengan melihat grafik diatas kita dapat mengetahui bahwa kira-kira hanya 18 persen saja orang yang mendapat nilai dibawah nilai fahrudin. Kalau hasil test itu dapat menjadi petunjuk tentang bakat dan kepandaian menukang, kita akan menasehatkan fahrudin untuk menerima tugas sebagai tukang pelitur atau mencari pekerjaan lain.

Kita akan menemui banyak hasil-hasil test bakat, test kemampuan khusus, dan semacamnya yang dilaporkan dalam bentuk ogive atau grafik frekuensi meningkat semacam itu, sebab nilai-nilai test semacam itu kerapkali digunakan untuk mengadakan penilaian tentang kecakapan perorangan.

3.6. BEBERAPA MACAM GRAFIK LAINNYA

Sampai sekian kita telah membicarakan cara-cara membuat grafik dari distribusi frekuensi. Kita sekarang akan membicarakan beberapa bahan penyelidikan yang menghendaki grafik macam lain.



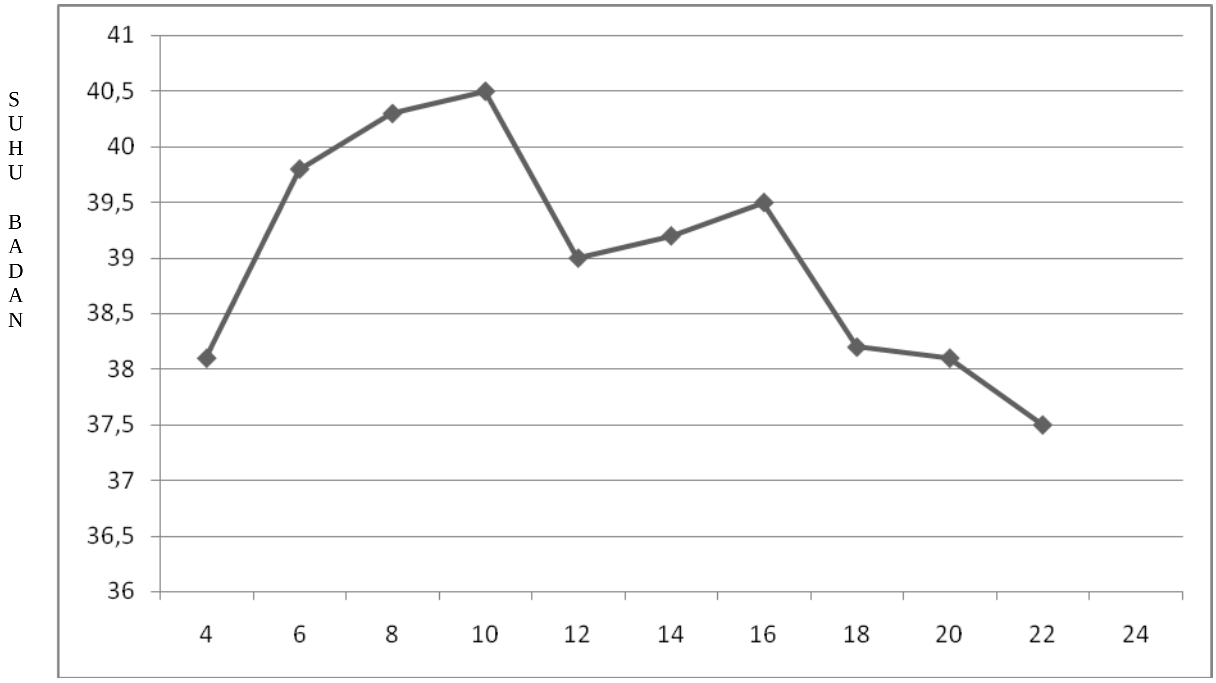
Grafik 7

Grafik Menurunkan Hubungan antara Lamanya Tikus Dilaparkan dan Frekuensi Singgungan pada Piring yang Beraliran Listrik

Dalam penyelidikan psikologi dan pendidikan kerapkali diselidiki reaksi seorang yang berulang-ulang menerima rangsang-rangsang eksperimen. Salah satu diantaranya yang paling banyak terjadi adalah penyelidikan eksperimental tentang "belajar". Dalam eksperimen semacam ini kepada orang coba diberikan kesempatan untuk mencoba mengulang-ulang suatu tugas yang kurang dikenalnya, dan kemudian dicatat kemajuan-kemajuan yang diperolehnya karena percobaan yang berulang-ulang itu. Di bawah ini adalah ilustrasi membuat grafik garis dari suatu contoh khayalan tentang bagaimana seekor tikus putihh belaraj mengenal bentuk. Makanan tikus tersebut diletakan diatas piring dari bermacam-macam bentuk. Beberapa dari piring dengan bentuk-bentuk tertentu diberi aliran listrik. Absis menunjukan lamanya si tikus dilaparkan, dan ordinat menunjukkan frekuensi tikus itu makan makanan pada piring yang diberi aliran listrik.

Dari grafik itu kurang lebih dapat diambil kesimpulan sebagai berikut : tikus yang sudah belajar mengenal bentuk bilamana dilaparkan menunjukan makin hilangnya pengenalan itu. Dalam dua haru yang pertama dari penglaparan itu kemampuan untuk membedakan bentuk dengan menyolok sekali menunjukan penurunan, penglaparan empat hari lamanya rupa-rupanya merupakan puncak dari makin hilangnya kesadaran bentuk. Akan tetapi dengan penglaparan yang lebih lama lagi kesadaran akan bentuk berangsur-angsur, tetapi secara tetap, makin kembali. Apakah hal ini disebabkan karena si tikus makin berhati-hari oleh pengalaman-pengalaman kejutan aliran listrik, ataukah dia memilih menarik diri dengan jalan mengurangi kerapnya mengunjungi sangkar-sangkar makanan, harus dipertimbangkan bahan-bahan lain yang terkumpul.

Dalam situasi lain seorang juru-rawat kesehatan ingin mencatat naik-turunnya suhu badan seorang penderita. Dibawah ini adalah contoh grafik yang memberi ilustrasi tentang naik-turunnya panas badan penderita yang diukur tiap-tiap dua jam sekali. Absis memuat saat-saat diadakannya pengukuran panas badan, sedang ordinat memuat jalan turun-naiknya panas badan penderita dalam satuan derajad celcius (grafik8).



Saat – saat pengukuran

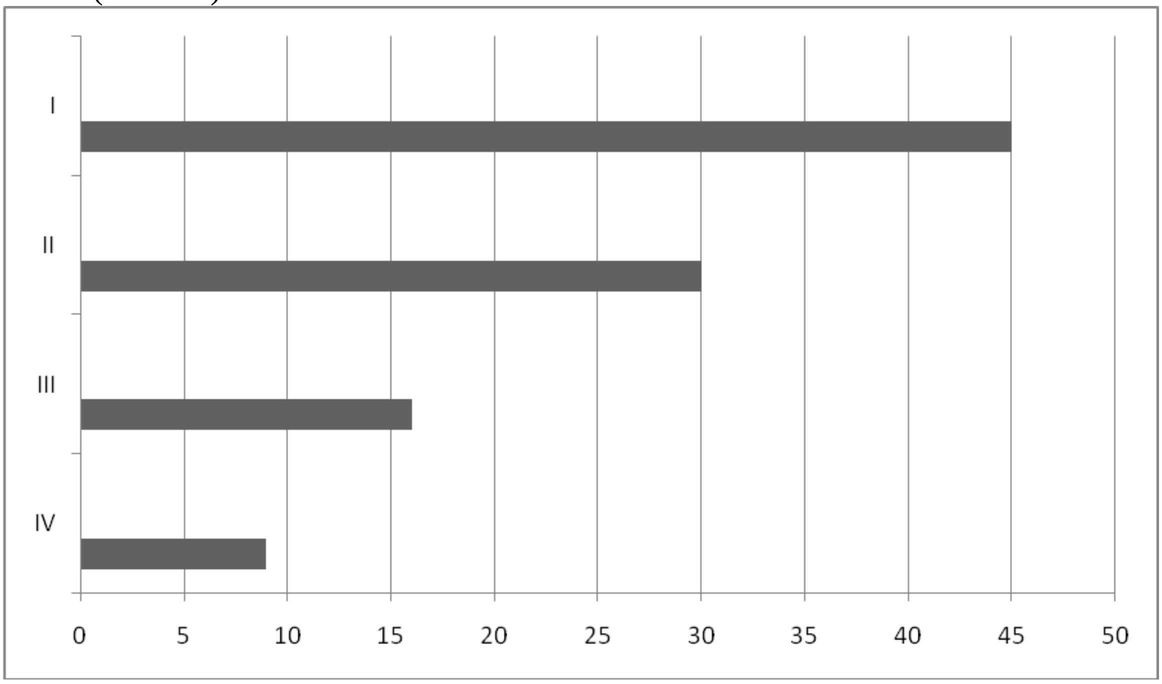
Grafik 8

Grafik menunjukan hubungan naik turunnya panas badan penderita "X" pada saat-saat tertentu

Akan tetapi tidak semua variabel yang kita selidiki merupakan variabel yang kontinu. Ukuran sepatu, macam-macam mata pelajaran sekolah, macam-macam barang dagangan, dan sebagainya, semuanya itu merupakan variabel-variabel yang diskrit, yang terpisah satu sama lain. Histogram di halaman

Rupanya merupakan contoh yang paling tepat untuk menyajikan keadaan variable yang diskrit (Grafik 9)

MAHASISWA TINGKAT



Jumlah Mahasiswa dalam %

Grafik 9

Histogram Menunjukkan Banyaknya Mahasiwa dari Empat Tingkat pada Fakultas A

Dalam grafik itu segiempat-segiempat histogramnya sengaja “ditidurkan, berbeda dengan pengetahuan kita sebelumnya. Hal ini bukan suatu kekeliruan. Atas pertimbangan-pertimbangan tertentu grafik histogram dapat berwujud semacam itu. Akan tetapi bilamana kita menghendaki histogram dengan segiempat-segiempat yang berdiri pada absis, hal inipun juga bukan suatu kekeliruan, malahan suatu kelaziman.

Untuk keperluan perbandingan, baik perbandingan bermacam-macam variable dari seseorang, maupun perbandingan variable yang sama dari sekelompok individu, disebabkan oleh sifatnya yang tidak kontinyu, kita tidak dapat menggunakan polygon atau kurve ogive. Histogram rupa-rupanya menjadi salah atu alat yang paling tepat untuk menggambarkannya.

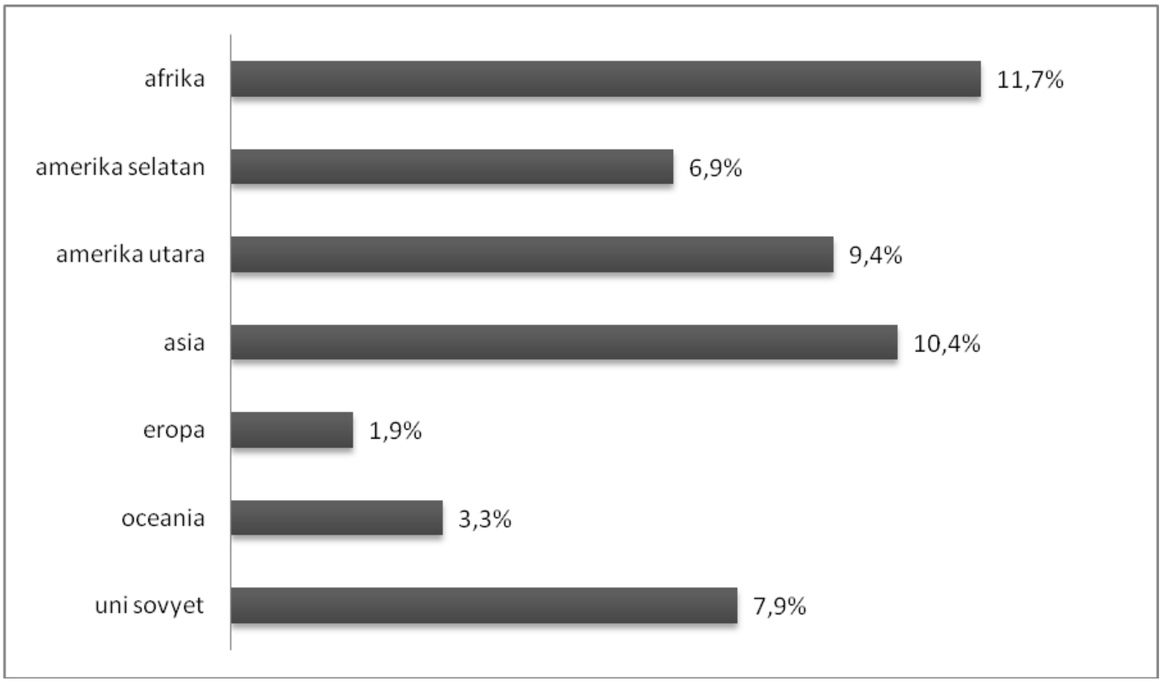
Satu macam grafik lagi yang kerap kali digunakan untuk melaporkan hasil penyelidikan adalah Grafik Serabi. Grafik ini berbentuk lingkaran (melambangkan keseluruhan) dengan jari-jari yang membagi lingkaran itu menjadi beberapa daerah yang luasnya seimbang dengan bagian-bagian gejala yang digambarkan.

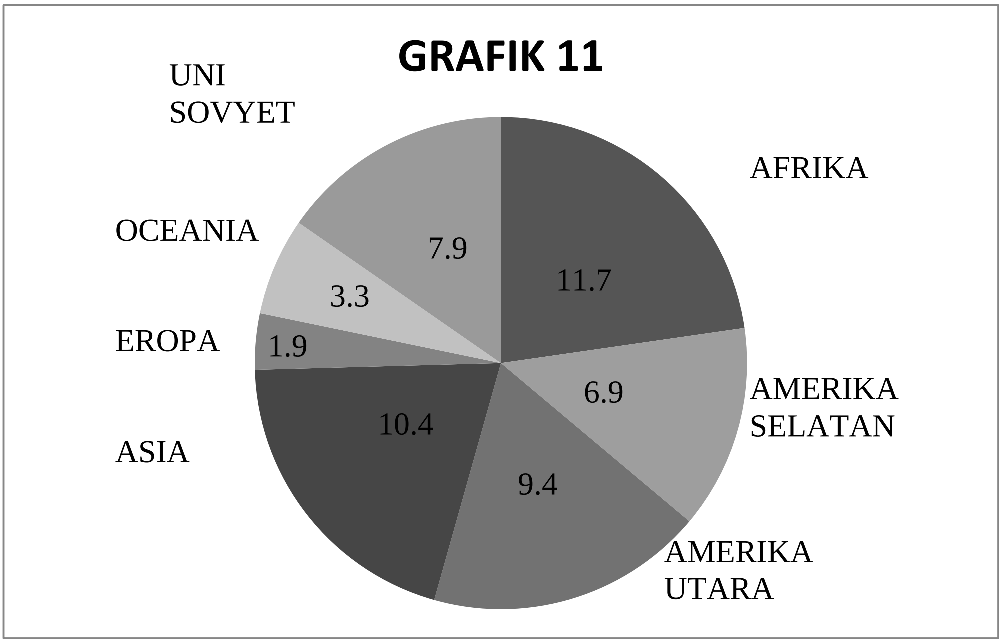
Contohnya:

Tabel 11
Luas Benua-Benua di Dunia

Benua	**Luas**[X)]
Afrika	11,7
Amerika Selatan	6,9
Amerika Utara	9,4
Asia	10,4
Eropa	1,9
Oceania	3,3
Uni Soviet	7,9
Total	51,5

[X)] *Dalam jutaan mil persegi*





Grafik 11
Luas Benua-Benua di Seluruh Dunia (Lalam Jutaan Mil Persegi)

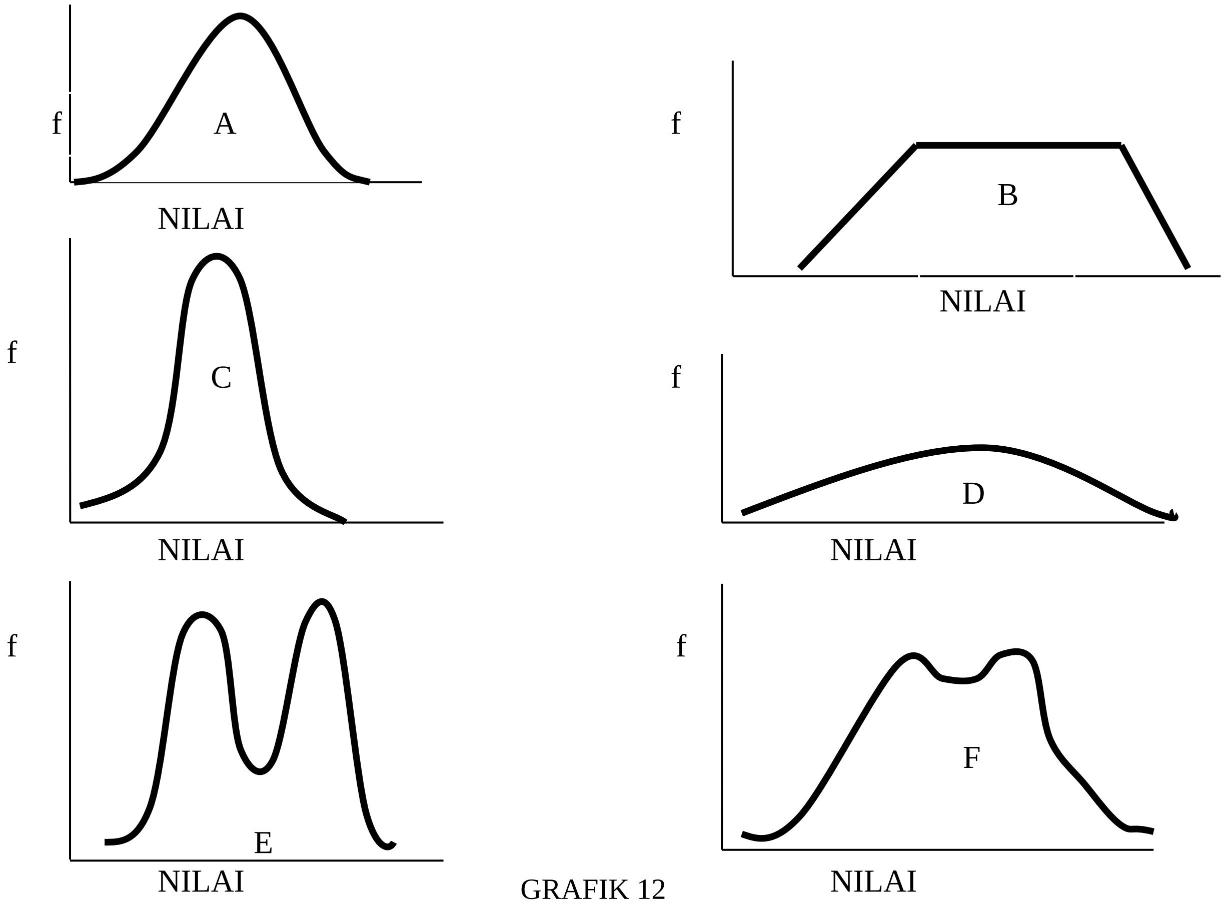
3.7. BEBERAPA BENTUK KURVA

Grafik polygon yang sudah dilicinkan disebut kurve. Kurve polygon mempunyai bentuk yang tidak terhingga banyaknya, tergantung kepada bentuk distribusinya. Oleh karena bentuk merupakan sumber keterangan yang penting sekali tentang distribusi, maka perlu bagi kita mengetahui sekedar istilah-istilah yang digunakan untuk menyebutbeberapa macam bentuk kurve.

Pada umumnya orang menggolongkan distribusi atau kurve ke dalam dua golongan besar, yaitu :

1. Distribusi atau kurve yang simetri, dan
2. Distribusi atau kurve yang a-simetri

Suatu kurve disebut simetri kalu kita lipat kurve itu tepat ditengah-tengah maka setengah lipatan bagian kiri akan menutup tepat setengah lipatan bagian kanan. Beberapa macam distribusi atau kurve yang simetri dapat dilihat pada grafik berikut :



GRAFIK 12

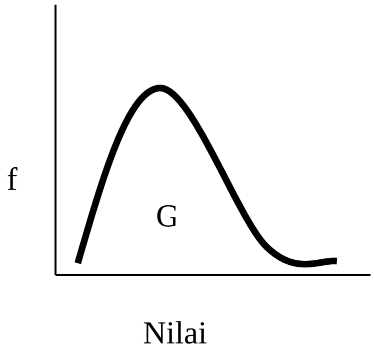
Kurve A,C dan D adalah kurve-kurve yang berbentuk seperti Bel. Kurve A, karena menunjukkan bentuk bel yang normal, disebut Kurve Normal. Kurve semacam ini akan kita jumpai berkali-kali dalam bab VI dan bab-bab selanjutnya. Kurve C disebut Kurve Bel Langsing (Leptokurtik), sedang Kurve D disebut Kurve Bel Gemuk(Platikurtik), Kurve B menunjukkan bentuk Trapesium sehingga diberi nama Kurve Trapesium (Rectangular)

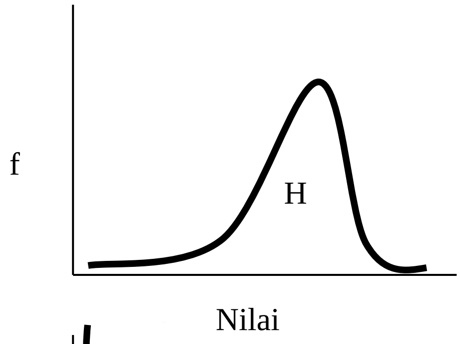
Kurve E dan Kurve F keduanya mempunyai dua puncak. Kurve simetri yang mempunyai dua puncak itu disebut Kurve Simetri Dwi mode. Kurve E hamper menyerupai bentuk huruf "U", sehingga diberi nama Kurve U.

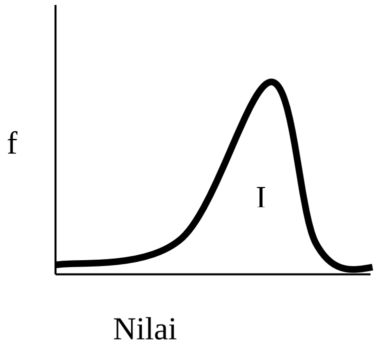
Golongan kedua adalah kurve a-simetri. Bilamana kurve a-simetri kita lipat ditengah-tengah. Maka lipatan sebelah kiri tidak akan menutup dengan tepat lipatan sebelah kanan. Di bawah ini adalah beberapa kurve yang a-simetri.

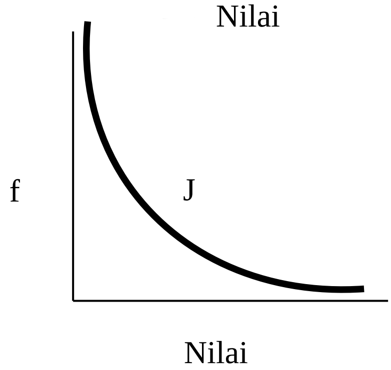
Kurve a-simetri lebih dikenal dengan nama Kurve Juling. Kejulingan ini ditunjukkan oleh arah "ekor"-nya. Bilamana suatu kurve mempunyai ekor di sebelah kanan, kurve ini disebut *Kurve Juling Positif* (lihat Kurve G). Kurve-kurve H dan I mempunyai ekor disebelah kiri, dibagian nilai-nilai yang rendah. Kurve-kurve ini disebut *Kurve Juling Negatif.*

Dengan melihat arah "kejulingan" sesuatu kurve, kita dapat dengan sepintas-lintas mengambil kesimpulan dari distribusinya. Disrtibusi yang juling positif menunjukkan bahwa sebagian besar individu memperoleh nilai dibagian bawah distribusi. Jadi kalau distribusi " Angka kecerdasan " menunjukkan orang-orang yang "kurang" kecerdasannya. Sebaliknya bilamana suatu distribusi berbentuk juling negative, nila-nilai variable di sebelah kanan atau disebelah atas disrtibusi memperoleh frekuensi yang terbanyak. Jadi bilaman hasil suatu ujian menunjukkan suatu distribusi yang juling negative, ini berarti bahwa sebagian terbesar anak-anak yang di uji mendapat angka-angka tinggi. Apakah dengan demikian kita akan menunyimpulkan bahwa soal-soal ujiannya terlalu mudah bagi anak-anak, hal ini sebenarnya bukan soal statistik. Statitik hanya menunjukksn saja kenyataanya.









Kurva J, karena bentuknya yang menyerupai huruf L, disebut Kurva L. Kurva semacam ini kita jumpai misalnya dalam distribusi “ orang – orang tua”.kalau pada absis kurva itukita cantumkan umur dari 60 tahun keatas, dan pada ordinatnya kita cantumkan banyak (frekuensi), maka akan kit abaca bahwa mereka yang mempunyai umur disekitar 60 tahun masih memperoleh frekuensi yang cukup tinggi, sedangkan mereka yang berumur makin menjauhi 60 tahun memperoleh frekuensi yang makin rendah.

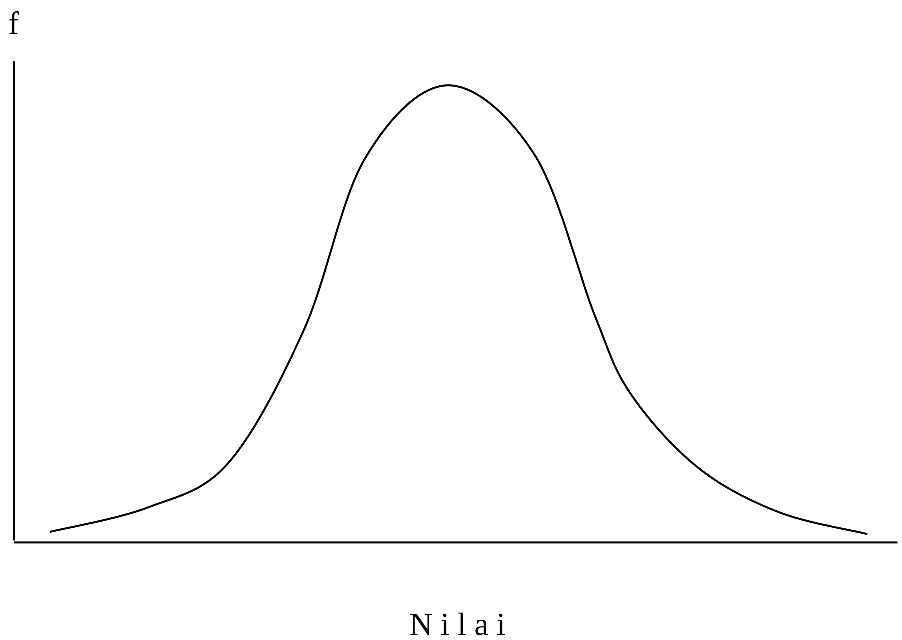
Dari semua bentuk kurve itu yang paling banyak kita jumpai adalah kurve simetri normal (periksa kembali grafik 12 A), kurve juling positif (Grafik 13G) dan kurve juling negative (grafik 13H dan I). Kurve – kurve semacam ini selalu mendapat perhatian kita disebakan karena penyelidikan – penyelidikan kita kerap kali menjumpai bentuk – bentuk kurve semacam itu. Terutama kurva Normal mempunyai arti teroretik dan pratik yang besar sekali.

POKOK BAHASAN:

“PENGUKURAN TENDENSI SENTRAL”

Pengamatan sehari-hari menunjukkan bahwa tiap orang tidak menunjukkan kesamaan dalam sesuatu hal. Kecerdasan, tinggi badan, berat badan, penghasilan dan sebagainya, bagi tiap orang kebanyakan tidaklah sama.

Bilamana sejumlah besar orang kita selidiki salah satu sifatnya dan kita buat grafik polygon dan distribusi sifat itu, maka akan kita jumpai grafik kurang lebih sebagai berikut :



Grafik 14

Kalau yang kita selidiki “kecerdasan״, maka akan kita lihat bahwa sebagian terbesar orang yang kita selidiki mempunyai kecerdasan yang “normal”. Dan bilamana kita ambil angka 100 sebagai indeks (ukuran) normalitas, maka sebagian terbesar orang yang kita selidiki akan mempunyai angka kecerdasan di sekitar 100, hanya sebagian kecil saja dari mereka yang angka kecerdasannya menyimpang jauh dari indeks normalitas kita.

Salah satu tugas dari statistik adalah mencari suatu angka di sekitar mana nilai-nilai dalam suatu distribusi memusat. Angka yang menjadi pusat sesuatu distribusi disebut ”tendensi sentral”.

Ada tiga macam tendensi sentral yang sangat penting untuk dibicarakan di sini. Ketiga tendensi sentral itu adalah Mean, Median dan Mode. Ketiganya mempunyai cara-cara menghitung yang berbeda-beda, dan mempunyai arti yang berbeda pula sebagai alat untuk mengadakan deskripsi suatu distribusi.

4.1 MEAN

Arti dari pada mean tidak lain adalah "angka rata-rata". Istilah Mean akan tetap dipakai disini oleh karena sudah lazim digunakan dalam statistik. Dari segi aritmetik Mean adalah "*Jumlah nilai-nilai dibagi dengan jumlah individu*". Penegasan ini dapat kita fahami dari contoh sebagai berikut.

Ada tiga orang berpenghasilan 10, 15, dan 20 rupiah tiap harinya. Rata-rata penghasilan mereka adalah 15 rupiah tiap harinya. Ini dicari dengan cara sebagai berikut :

$$\text{Penghasilan rata-rata} = \frac{10 + 15 + 20}{3} = \frac{45}{3} = 15$$

Dari kenyataan itu dapat dikemukakan rumus Mean sebagai berikut :

$$\text{Mean} = \frac{X_1 + X_2 + X_3 \ldots X_{n-1} + X_n}{N} \qquad (2)$$

dalam mana X_1, X_2, dan seterusnya adalah nilai-nilai individual, dan N adalah jumlah individu dalam distribusi.

Rumus itu diangkat sebagai berikut :

$$M = \frac{\sum X}{N} \qquad (3)$$

Simbol ∑ adalah huruf Yunani yang disebut 'Sigma' dan mempunyai arti jumlah.

4.2 MEAN YANG DITIMBANG

Bilamana ada empat orang yang berpenghasilan 10 rupiah, seorang yang berpenghasilan 15 rupiah, dan seorang yang berpenghasilan 20 rupiah seharinya, maka Mean daari penghasilan mereka tidak lagi 15 rupiah, melainkan 12,5 rupiah. Hal ini dapat dicari dengan tabel sebagai berikut :

Tabel 12
Tabel untuk Contoh Mencari Mean yang Diimbang

Penghasilan (X)	**Frekuensi (f)**	**fX**
20	1	20
15	1	15
10	4	40
	N = 6	∑ fX = 75

Rumusan Mean yang ditimbang sebagai berikut :

$$M = \frac{\sum X}{N} \qquad (4)$$

Diisi dengan bahan-bahan Tabel 1.2 :

$$M = \frac{75}{6} = 12{,}5$$

Tampaklah bahwa rumus Mean diatas pada dasarnya sama saja dengan rumus yang sudah dikemukakan. Mean yang ditimbang adalah Mean yang memperhitungkan frekuensi tiap-tiap nilai variabel. Selanjutnya yang dimaksudkan dengan mean adalah mean yang ditimbang.

4.3 MENGHITUNG MEAN DARI DISTRIBUSI BERGOLONG

Contoh diatas itu adalah contoh menghitung mean dari distribusi tunggal. Menghitung mean dari distribusi bergolong pada hakekatnya tidak berbeda dengan menghitung mean dari distribusi tunggal. Rumus yang sudah dikemukakan berlaku sepenuhnya disini. Hanya saja nilai X disini tidak lagi mewakili nilai variabel individual, melainkan mewakili "Titik Tengah" interval kelas. Dibawah ini adalah contoh menghitung mean dari distribusi bergolong

Sekali lagi perlu diingatkan disini bahwa X adalah mewakili "Titik Tengah" dari interval kelas dalam distribusi.

Tabel 13

Tabel untuk Contoh Menghitung Mean dari Distribusi Bergolong

Interval Nilai	Titik Tengah (X)	f	fX
148 - 149	147	1	147
140 - 144	142	3	426
135 - 139	137	5	685
130 - 134	132	8	1056
125 - 129	127	11	1397
120 - 124	122	17	2074
115 - 119	117	21	2457
110 - 114	112	22	2464
105 - 109	107	24	2568
100 - 104	102	20	2040
95 - 99	97	15	1455
90 - 94	92	12	1104
85 - 89	87	6	522
80 - 84	82	2	164
Jumlah		N = 167	∑ fX = 18559

$$M = \frac{\sum fX}{N} = \frac{18559}{167} = 111{,}13$$

4.4 MENGHITUNG MEAN DARI DISTRIBUSI BERGOLONG DENGAN RUMUS TERKAAN

Dalam menghitung, mean dari distribusi bergolong waktu kita akan sangat dihemat oleh penggunaan rumus Mean Terkaan. Dan oleh karena kita tidak terlibat pada angka-angka yang besar-besar, maka kemungkinan membuat kekeliruan adalah sedikit sekali. Rumus ini sangat praktis, terutama bilamana tidak tersedia mesin hitung.

Istilah "terkaan" jangan diartikan "raba-raba", sebab akhirnya "kesalahan" oleh terkaan itu dikoreksi kembali. Mean terkaan boleh juga disebut Mean Kerja, sebab mean terkaan itu digunakan pangkal bekerja.

Langkah-langkah untuk menghitung mean dengan mean terkaan adalah sebagai berikut :

1. Menerka sesuatu Mean. Terkaan ini boleh semau kita.
2. Mencari deviasi nilai-nilai individual dari mean terkaan itu. Deviasi-deviasi di atas mean terkaan diberi tanda plus, sedang di bawahnya diberi tanda minus.
3. Mengalikan deviasi tiap-tiap nilai itu dengan frekuensinya.
4. Menjumlahkan deviasi yang sudah dikalikan dengan frekuensi itu.
5. Mengisikan bahan-bahan yang sudah diperoleh itu ke dalam rumus.

Untuk memahami langkah-langkah itu baiklah kita lihat contoh di bawah ini sebagai berikut :

Tabel 14
Tabel untuk Menghitung dengan menggunakan Rumus Mean Terkaan

Interval Nilai	f	x^1	fx^1	
145 - 149	1	+8	+8	
140 - 144	3	+7	+21	
135 - 139	5	+6	+30	
130 - 134	8	+5	+40	
125 - 129	11	+4	+44	+258
120 - 124	17	+3	+51	
115 - 119	21	+2	+42	
110 - 114	22	+1	+22	
105 - 109	24	0	0	
100 - 104	20	-1	-20	
95 - 99	15	-2	-30	
90 - 94	12	-3	-36	-120
85 - 89	6	-4	-24	
80 - 84	2	-5	-10	
Jumlah	N = 167	-	∑ fX =	138

Rumus untuk menghitung Mean dengan Mean Terkaan adalah :

$$M = MT + \left[\frac{\sum fX}{N} \right] i \qquad (5)$$

dalam mana M adalah : Mean yang kita cari, Mean yang sebenarnya,

MT : Mean Terkaan atau Mean Kerja

∑ fX : Jumlah Deviasi atau Kesalahan akibat terkaan

i : Lebar interval

Langkah 1 : Yang kita jadikan mean terkaan dalam distribusi di atas adalah interval 105 - 109. Pada interval ini telah kita beri tanda garis tebal. Titik tengah dari interval ini adalah 107. Karena Mean harus merupakan satu angka, maka titik tengah 107 ini yang kita sebut mean terkaan.

Langkah 2 : Huruf x' yang dicantumkan dalam kolom ke tiga itu adalah deviasi dari mean terkaan. Sebab itu pada baris yang berisi mean terkaan deviasinya sama dengan nol. Selanjutnya deviasi-deviasi di bawah mean kita beri tanda negatif. Deviasi di atas mean secara berturut-turut dari bawah ke atas kita beri kode angka-angka +1, +2, +3, dan seterusnya. Deviasi di bawah mean kita beri kode dari atas ke bawah -1, -2, dan seterusnya.

Langkah 3 : Perkalian antara deviasi tiap-tiap dengan frekuensinya masing-masing kita cantumkan dalam kolom keempat.

Langkah 4 : Deviasi-deviasi yang sudah dikalikan dengan frekuensi itu kita jumlahkan. Jumlah dari deviasi-deviasi ini disebut jumlah deviasi kesalahan. Dari distribusi di atas jumlah deviasi kesalahannya ada 138.

Langkah 5 : Apa yang sudah kita ketahui dari bahan-bahan tersebut di atas adalah :

MT = 107

$\sum fx$ = 138

N = 167

i = 5

Dengan mengisikan apa yang sudah kita ketahui itu ke dalam rumusnya, maka akan kita peroleh sebagai berikut :

$$M = MT + \left[\frac{\sum fX}{N}\right] i \quad = 107 + \frac{138}{167}\ 5 =$$

$$= 107 + 0{,}826 \times 5 = 107 + 4{,}13 = 111{,}13$$

Kita lihat ada kecocokan antara hasil mean yang dikerjakan dengan menggunakan rumus biasa dengan yang menggunakan rumus mean terkaan, dari satu distribusi yang sama. Satu keuntungan yang terang dengan menggunakan rumus mean terkaan ini ialah kita tidak terlibat pada angka-angka yang besar-besar. Dengan demikian kemungkinan kekeliruan juga sedikit sekali.

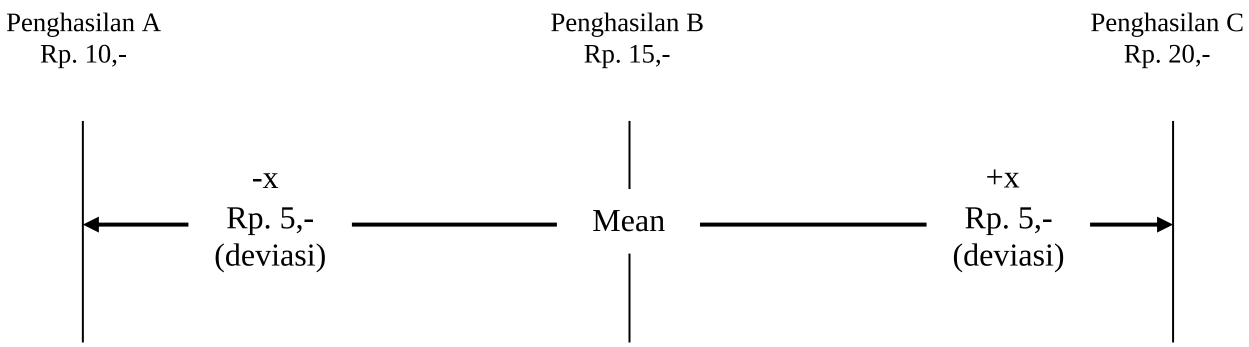
Komponen $\left[\frac{\sum fX}{N}\right] i$ dalam rumus kita itu adalah komponen koreksi.

Bilamana mean terkaan (disingkat MT) terlalu tinggi, maka komponen itu akan merupakan bilangan negatif. Sebaliknya bilamana MT kita terlalu rendah daripada mean yang sebenarnya, komponen itu akan positif. Dalam contoh tersebut di atas MT kita terlalu rendah (yang sebenarnya adalah 111,13, sedang MT kita hanya 107). Sebab itu komponen koreksinya merupakan bilangan positif terbesar 4,13, untuk menambah MT = 107 menjadi M = 111, 13.

Bilamana MT kita tepat, maka $\sum fX$ akan = 0, sehingga $\left[\frac{\sum fX}{N}\right] i = 0$. Kenyataan ini masuk akal sekali, sebab kalau mean terkaannya sama dengan mean yang sebenarnya, maka kita tidak usah kalau mean terkaannya reksi lagi.

Latar belakang rumus ini adalah sebagai berikut :

Mean adalah suatu titik, suatu nilai yang terletak tepat di tengah-tengah distribusi. Sebab itu jika deviasi-deviasi di bawah mean dijumlahkan dengan deviasi-deviasi di atas mean, hasilnya harus sama dengan nol. Ini dapat kita fahami dari contoh sederhana seperti



Mean dari penghasilan ketiga orang itu adalah Rp. 15,- (diperoleh dari Rp. 45,- : 3). Deviasi penghasilan A dari mean adalah –Rp. 5,- sedang deviasi penghasilan C dari mean adalah +Rp.5,.

Jumlah deviasi-deviasi penghasilan A dan C adalah -Rp. 5,- ditambah +Rp. 5,- sama dengan Rp. 0,-. Dengan simbol :

$$\sum x = 0 \qquad (6)$$

Karena adanya unsur "terkaan", maka jumlah deviasi-deviasi jarang sekali sama dengan nol. Oleh sebab itu harus diadakan koreksi terhadap kesalahan yang diakibatkan oleh terkaan. Arah koreksi (positif atau negatif) sangat tergantug kepada apakah terkaan kita terlalu tinggi atau terlalu rendah.

Untuk membuktikan bahwa dengan terkaan yang terlalu tinggipun akan kita peroleh mean yang sama, marilah kita ikuti contoh berikut :

Tabel 15

Reproduksi dari Tabel 13

Interval Nilai	**f**	**x¹**	**fx¹**	
145 - 149	1	+5	+5	
140 - 144	3	+4	+12	
135 - 139	5	+3	+15	+59
130 - 134	8	+2	+16	
125 - 129	11	+1	+11	
120 - 124	17	0	0	
115 - 119	21	-1	-21	
110 - 114	22	-2	-44	
105 - 109	24	-3	-72	
100 - 104	20	-4	-80	
95 - 99	15	-5	-75	-422
90 - 94	12	-6	-72	
85 - 89	6	-7	-42	
80 - 84	2	-8	-16	
Jumlah	N = 167	-	∑ fX =	-363

$$M = MT + \left[\frac{\sum fX}{N}\right] i \quad = 122 + \left[\frac{-363}{167}\right] 5 \quad =$$

$$= 107 + -2{,}174 \times 5 = 122 - 10{,}870 = 111{,}13$$

Dalam contoh di atas kita menggunakan nilai 122 (yaitu titik tengah interval yang dicoret tebal) sebagai MT kita.

Dari contoh tersebut di atas nampak dengan jelas bahwa dengan menggunakan mean terkaan di manapun, hasilnya akan sama saja. Akan tetapi sebaiknya kita tidak mencoba meletakkan mean terkaan pada bagian ujung distribusi (ujung atas atau ujung bawah), sebab dengan demikian kita akan terlibat pada angka-angka yang lebih besar. Kalau demikian keadaannya maka tidak ada artinya kita menggunakan rumus yang dimaksudkan untuk menghemat tenaga dan mencegah kemungkinan membuat kekeliruan itu.

4.5. M E D I A N

Median dapat dibatasi sebagai “suatu nilai yang membatasi 50 persen frekuensi distribusi bagian bawah dengan 50 persen frekuensi distribusi bagian atas”.

Perlu dicatat bahwa, seperti juga mean, median mungkin sekali tidak menjadi “milik” dari salah satu individu dalam distribusi. Hal ini akan menjadi jelas setelah kita bicarakan lebih lanjut.

Kita misalkan ada distribusi penghasilan dari tujuh orang seperti tersebut dalam tabel di bawah ini.

Menurut definisinya, median adalah suatu nilai yang membatasi 50 persen dari frekuensi distribusi sebelah atas dan 50 persen frekuensi distribusi sebelah bawah. Dalam distribusi tersebut di atas nilai yang dimiliki oleh individu nomor empat-lah yang menjadi batas itu. Dikatakan, median dari distribusi itu adalah Rp. 14,-. Individu nomor 4 itu membatasi separo individu di atas dan separo lagi di bawahnya.

Tabel 16

Tabel Distribusi Penghasilan Fiktif untuk Contoh Mencari Median

Individu	Penghasilan
1	Rp. 10
2	12
3	13
4	14
5	16
6	16
7	20

Dari contoh itu nampak dengan jelas bahwa median hanya tergantung pada banyaknya frekuensi, tidak tergantung kepada variasi nilai-nilai variabel. Kalau sekiranya individu nomor 4 itu berpenghasilan Rp. 16,-, maka mediannya akan menjadi Rp. 16,-.

4.6. MEDIAN PADA DISTRIBUSI DENGAN FREKUENSI GENAP

Bilamana suatu distribusi mempunyai frekuensi genap, maka median dihitung secara kompromi, yaitu dengan membagi dua nilai-nilai variabel yang ada di tengah-tengah distribusi. Misalkan ada empat orang yang masing-masing mempunyai tinggi badan 162, 162, 164, dan 166 cm, maka median tinggi badan empat orang itu ada 163 cm (diperoleh dari 162 cm ditambah 164 cm kemudian dibagi dua). Pemecahan semacam ini sama sekali tidak bertentangan dengan definisi median, sebab angka 163 cm itu sebagai batas antara tinggi 162 dan 164 cm membatasi 50 persen frekuensi variabel di bagian atas, yaitu dua orang, dan 50 persen frekuensi variabel di bagian bawah distribusi, yaitu dua orang.

4.7. MENCARI MEDIAN DARI DISTRIBUSI BERGOLONG

Rumus untuk mencari median dari distribusi bergolong adalah sebagai berikut :

$$\text{Median} = Bb + \left[\frac{\frac{1}{2} N - cf_b}{f_d} \right] i \qquad (7)$$

dalam mana :

Bb adalah batas bawah (nyata) dari interval yang mengandung median,

cf_b frekuensi kumulatif (frekuensi meningkat) di bawah interval yang mengandung median,

f_d frekuensi dalam interval yang mengandung median,

i lebar interval, dan

N jumlah frekuensi dalam distribusi.

Penggunaan rumus itu dapat kita lihat dari pekerjaan di bawah ini :

Tabel 17
Tabel Untuk Contoh Menghitung Median dari Distribusi Bergolong

Interval Nilai	f	cf
100 - 104	1	55
95 - 99	3	54
90 - 94	5	51
85 - 89	9 f_d	46
80 - 84	(13)	37
75 - 79	10	(24) ← cf_b
70 - 74	6	14
65 - 69	4	8
60 - 64	3	4
55 - 59	1	1
Jumlah	55	

Pertama harus kita ketahui bahwa untuk menghitung median kita selalu menggunakan kolom yang berisi frekuensi meningkat (frekuensi kumulatif atau cf, periksa kolom ketiga). Kolom ini diperlukan untuk mencari interval mana yang mengandung median. Hal ini dapat kita cari dengan membagi dua jumlah frekuensinya. Dalam contoh di atas, jumlah frekuensinya (atau N) ada 55. Kalau ini kita bagi dua hasilnya sama dengan 27,5 itu. Setelah ½ N ini kita ketemukan maka langkah selanjutnya adalah menemukan interval kelas yang mengandung frekuensi 27,5 itu. Interval kelas yang kita maksudkan adalah 80 – 84, sebab cf 27,5 terkandung dalam cf 37. Nah, kalau sudah kita ketemukan interval kelas yang mengandung median itu, pekerjaan kita tinggal lagi mengisi rumusnya.

Batas bawah (nyata) atau Bb dari interval yang mengandung median itu adalah 79,5. Separo dari jumlah frekuensinya, atau ½ N, adalah 55/2, sama dengan 27,5. Frekuensi kumulatif di bawah interval yang mengandung median adalah 24 (24 adalah cf di bawah 37, sedang cf 37 adalah cf yang mengandung median). Frekuensi dalam interval adalah 13, sedang lebar interval atau i-nya ada lima. Diisikan dalam rumus kita jumpai perhitungan sebagai berikut :

$$\text{Mdn} = Bb + \left[\frac{\tfrac{1}{2}N - cf_b}{f_d}\right] i = 79{,}5 + \left[\frac{27{,}5 - 24}{13}\right] 5 =$$

$$= 79{,}5 + \frac{3{,}5 \times 5}{13} = 79{,}5 + 1{,}346 = 80{,}846$$

atau 80, 85

Jadi, median dari distribusi tersebut ada 80,85. Ini adalah nilai variabel yang terdapat dalam kelas 80 – 84, dan menjadi batas antara 50 persen frekuensi di sebelah atas dengan 50 persen frekuensi di sebelah bawah distribusi. Dengan kata-kata lain, separo dari frekuensi variabel, yaitu 27,5 orang mendapat nilai di atas 80,85 dan separo lagi yaitu 27,5 orang, mendapat nilai di bawah 80,85 itu.

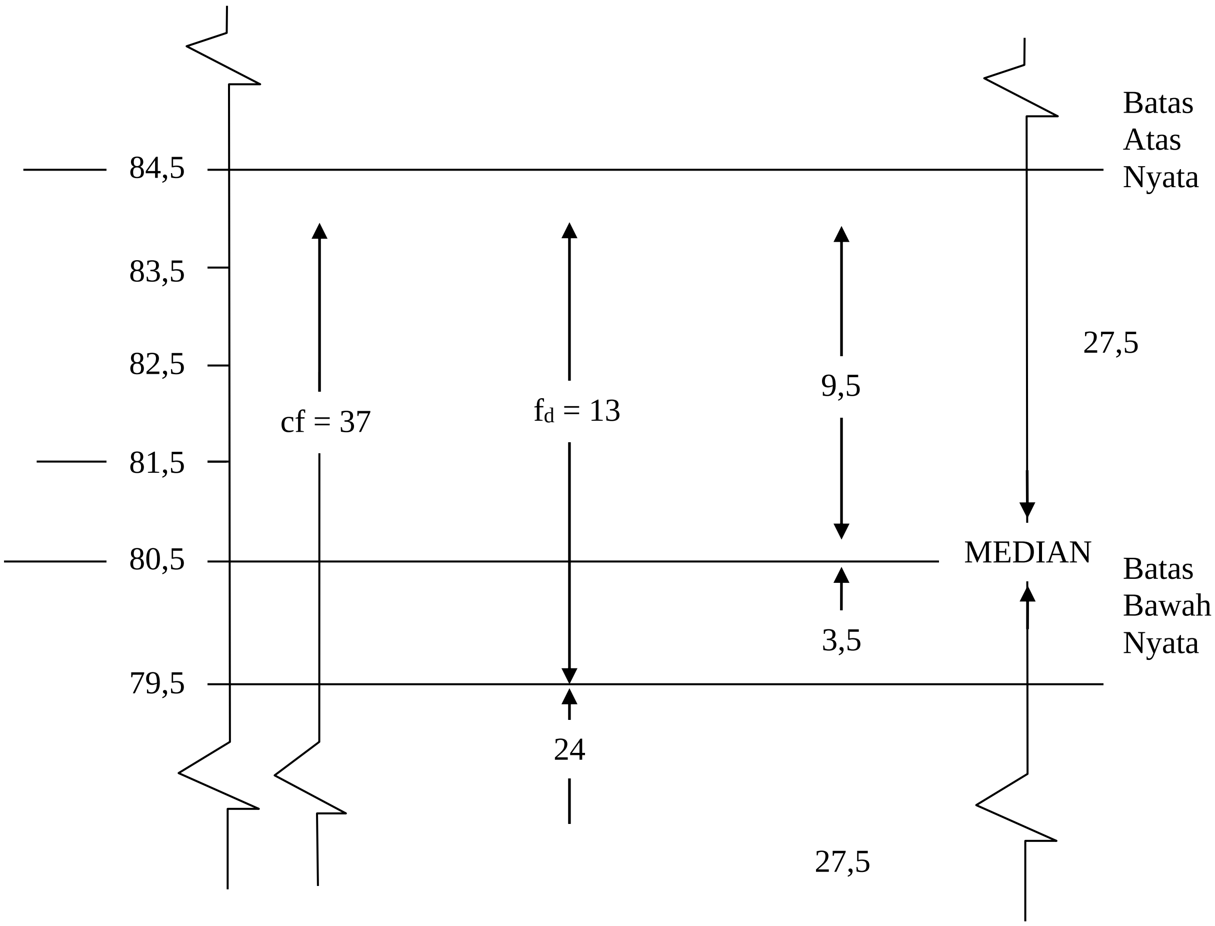
Catatlah, bahwa langkah-langkah yang paling kritik adalah mencari interval kelas mana yang mengandung median. Ini dicari dengan membagi dua jumlah frekuensi seluruhnya (atau N:2). Setelah ½ N ini kita ketemukan, kita tandai frekuensi kumulatif yang mengandung ½ N itu, dan kita tarik garis tebal pada baris yang mengandung frekuensi kumulatif itu. Dengan demikian interval kelas yang kita maksudkan telah kita ketemukan. Batas bawah (nyata) dari interval adalah separo dari batas bawah semu interval itu dengan batas atas semu interval bawahnya. Dalam contoh di atas batas bawah nyatanya adalah separo dari 80 ditambah 79, atau sama dengan 79,5. Langkah selanjutnya adalah menemukan cf_b. Ingat, cf_b frekuensi relatif di bawah interval yang mengandung Median. Ingat juga, f_d adalah frekuensi dalam interval yang mengandung media, bukan frekuensi kumulatif di dalam, di atas, atau di bawahnya.

Nah, kalau semuanya itu sudah kita ketemukan, tinggal lagi kita mengisikannya ke dalam rumusnya.

4.7. BAGAIMANA TERJADINYA RUMUS MEDIAN

Untuk mengerti asal-usulnya terjadinya rumus median perlu kita ikuti dengan teliti penjelasan di bawah ini :

Distribusi yang digunakan sebagai contoh itu kita perbesar dengan "kaca pembesar". Dengan "mikroskop" dalam khayalan ini kita periksa interval yang mengandung median. Gambarannya kira-kira adalah sebagai terlihat di bawah ini :



Seperti diterangkan di muka, median terletak dalam interval 80 – 84. Batas nyata dari interval ini di atas adalah 84,5, dan di bawah adalah 79,5. Lebar interval itu ada 5.

Frekuensi di dalam interval itu ada 13, di antaranya yang 9,5 (diperoleh dari 37 – 27,5) terletak di atas median, dan 3,5 sisanya (diperoleh dari 27,5 – 24 atau 13 + 9,5) terletak di bawah median).

Oleh karena kita tidak tahu dengan pasti penyebaran atau distribusi frekuensi sebanyak 13 itu di antara lima buah nilai (ingat i = 5) yang ada di dalam. Interval itu, maka kita mengambil dasar anggapan bahwa frekuensi interval sebanyak 13 itu dibagi rata antara kelima nilai variabel di dalamnya. Dengan demikian median terletak 3,5 / 13 dari luas kelas 5 unit di atas batas bawah (nyata) intervalnya. Jadi, mediannya ada 79,5 di tambah 3,5 / 13 x 5, sama dengan 80,846 atau 80,85.

4.8. M O D E

Mode dapat dibatasi sebagai :

1. dalam Distribusi Tunggal : nilai variabel yang mempunyai frekuensi tertinggi dalam distribusi;
2. dalam Distribusi Bergolong : titik tengah interval kelas yang mempunyai frekuensi tertinggi dalam distribusi.

4.9. MODE DALAM DISTRIBUSI TUNGGAL

Dalam serangkaian nilai-nilai 5, 6, 7, 7, 7, 8, 8, 8, 8, 9, 9, nilai yang timbul paling banyak adalah nilai 8. Nilai 8 itu disebut Mode dari distribusi nilai-nilai itu.

Tabel 18

Tabel Disediakan untuk Menaenal Mode

Nilai	Frekuensi
10	1
9	0
8	15
7	18
6	4
5	3
4	1
3	1

Kalau suatu distribusi sudah disusun dalam tabel, maka untuk mencari Mode-nya kita melihat pertama dalam kolom frekuensi. Dalam kolom frekuensi itu kita cari frekuensi yang tertinggi, kemudian kita baca nilai variabel yang sebaris dengan frekuensi tertinggi itu. Nilai itu adalah Mode dan distribusi yang telah disusun menjadi tabel itu. Untuk jelasnya periksa tabel 18.

Frekuensi yang tertinggi dari distribusi tersebut adalah 18. Nilai yang mempunyai frekuensi tertinggi itu adalah nilai 7. Jadi yang menjadi modenya adalah nilai 7.

Kalau misalnya nilai variabel yang tercantum dalam distribusi Tabel 18 itu adalah nilai-nilai suatu mata pelajaran, maka yang menjadi modenya adalah nilai 7. Artinya, dari pada nilai-nilai lainnya, sebagian terbesar adalah murid-murid memperoleh nilai 7 dalam mata pelajaran itu.

Perlu diperingatkan bahwa mode adalah *nilai*, bukan *frekuensi* yang tertinggi. Hal ini perlu ditekankan karena sedikit mahasiswa yang keliru mengartikan mode ini. Baca definisinya, mode dalam distribusi tunggal adalah *nilai* variabel yang memperoleh frekuensi terbanyak.

4.9. MODE DALAM DISTRIBUSI BERGOLONG

Bilamana kita telah memahami pengertian tentang mode dalam distribusi tunggal, tidak sukar kiranya kita memahami mode dalam distribusi bergolong. Sebagai contoh periksa distribusi dalam Tabel 19 di bawah ini :

Tabel 19
Tabel untuk Menemukan Mode dari Distribusi Bergolong

Interval Nilai	Titik Tengah (x)	Frekuensi (f)
195 - 199	197	1
190 - 194	192	2
185 - 189	187	4
180 - 184	182	5
175 - 179	177	8
170 - 174 →	(172)	← 10
165 - 169	167	6
160 - 164	162	4
155 - 159	157	4
150 - 154	152	2
145 - 149	147	3
140 - 144	142	1

Frekuensi yang tertinggi dalam distribusi itu adalah 10. Interval yang mempunyai frekuensi tertinggi itu adalah 170 - 174, dan titik tengah dari interval itu adalah 172. Jadi, yang menjadi mode dalam distribusi itu adalah nilai 172.

Definisi yang telah dikemukakan di atas adalah definisi dari apa yang disebut Mode "kasar". Akan tetapi bilamana kita menghitung mode dari distribusi frekuensi, kita membedakan antara apa yang disebut mode "asli" dari mode "kasar" itu. Mode "asli" adalah suatu suatu nilai dalam distribusi yang menjadi pemusatan dari nilai-nilai lainnya. Atau dengan kata lain, nilai yang paling banyak timbul dibandingkan dengan nilai-nilai lainnya. Bilamana skala pengukuran diperinci menjadi unit-unit nilai kecil-kecil, bilamana nilai-nilai dicatat seteliti-telitinya, dan bilamana jumlah frekuensinya besar sekali, maka mode "kasar" akan sangat mendekati mode "asli". Akan tetapi biasanya mode "kasar" hanya merupakan pendekatan saja kepada mode “asli”. Rumus untuk mencari mode yang mendekati asli, bilamana distribusinya simetri, atau setidak-tidaknya tidak sangat juling, adalah :

Mode = 3 Median - 2 Mean (8)

Bilamana kita menggunakan formula atau rumus itu untuk mengerjakan bahan dalam Tabel 19, maka modenya akan kita ketemukan 174,40. Nilai ini ternyata lebih besar sedikit dibandingkan dengan nilai mode yang diperoleh dengan rumus mode “kasar”, yaitu 172.

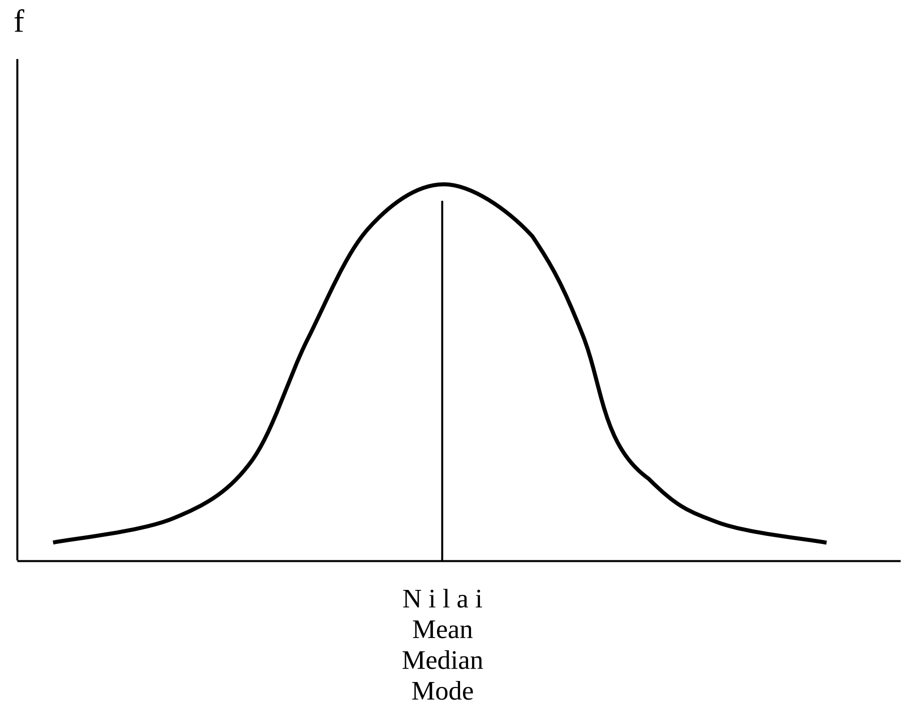
Mode kasar kadang-kadang merupakan pengukuran tendensi sentral yang kurang teliti. Akan tetapi kekurangan ini bukan merupakan kelemahan yang sangat serius seperti nampaknya sepintas lalu. Mode kasar biasanya digunakan sebagai alat pemeriksaan yang sangat sederhana untuk melihat pusat konsentrasi dalam suatu distribusi. Bilamana hanya tafsiran saja yang kita kehendaki maka kita tidak perlu menghitung Mean dan Median yang memakan banyak waktu itu.

Akhirnya perlu ditambahkan sebagai catatan disini tentang kemungkinan adanya distribusi yang mempunyai dua mode. Distribusi yang mempunyai dua mode ini, seperti telah disebutkan di muka, disebut distribusi dwimode. Suatu distribusi disebut dwimode kalau diantara kedua nilai (dalam distribusi tunggal) atau diantara kedua interval (dalam distribusi bergolong) yang mendapat frekuensi tertinggi itu ada terdapat nilai itu interval lain yang lebih rendah frekuensinya.

4.10. TEMPAT KEDUDUKAN MEAN, MEDIAN DAN MODE DALAM DISTRIBUSI

Tempat kedudukan Mean, Median dan Mode dalam satu distribusi sangat bergantung kepada bentuk distribusinya. Kita ingat kembali ada distribusi yang simetri dan ada yang juling.

Bilamana dari suatu distribusi simetri normal kita hitung mean, median dan mode-nya, maka akan kita jumpai sifat yang khas, yaitu bahwa ketiga tendensi sentral itu bersekutu satu sama lain. Hal ini mudah kita mengerti, sebab pada distribusi normal, mean membagi dua sama banyak frekuensi variabel di atas dan di bawahnya. Dengan demikian mean ini mempunyai fungsi seperti median. Karena yang menjadi mode dalam distribusi normal adalah nilai yang ada pada mean, maka dengan sendirinya mode itu bersekutu dengan mean. Jadi pada distribusi normal mean, median, dan mode ketiga-tiganya berimpit. Untuk ilustrasinya periksalah grafik 15



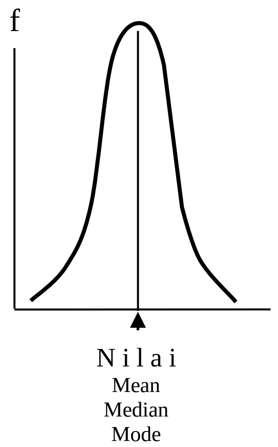
Grafik

Sungguh pun demikian janganlah disimpulkan bahwa karena pada distribusi normal mean, median dan modenya bersekutu maka bilamana mean, median dan mode dari suatu distribusi bersekutu maka distribusinya adalah normal (kalau lembu adalah binatang berkaki empat, tidak tepat menyimpulkan bahasa semua binatang berkaki empat adalah lembu). Semua distribusi normal memang mempunyai mean, median dan mode yang bersekutu; tetapi tidak semua distribusi yang mempunyai mean, median dan mode yang bersekutu adalah distribusi normal.

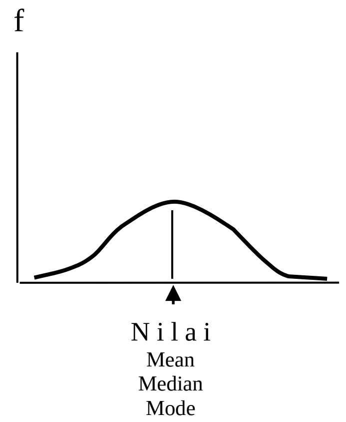
Distribusi normal adalah salah satu bentuk distribusi simetri. Distribusi simetri lainnya adalah distribusi trapesium, distribusi dwimode, dan distribusi bentuk bel yang tidak normal. Dalam distribusi terakhir ini tempat kedudukan ketiga tendensi sentral itu sama halnya dengan kedudukan mereka dalam distribusi normal.

Dalam distribusi trapesium dan distribusi dwimode, mean dan mediannya bersekutu, sedang mode-nya mempunyai kedudukan tersendiri. Pada distribusi dwimode, modenya ada di samping kiri-kanan mean dan median, sedang pada distribusi trapesium, karena tidak ada mode-nya, maka tidak dapat dipersoalkan kedudukan modenya.

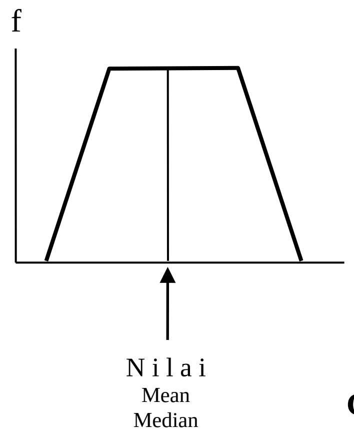
Untuk mendapat gambaran yang jelas tentang apa yang baru dibicarakan periksalah grafi-grafik di bawah ini :



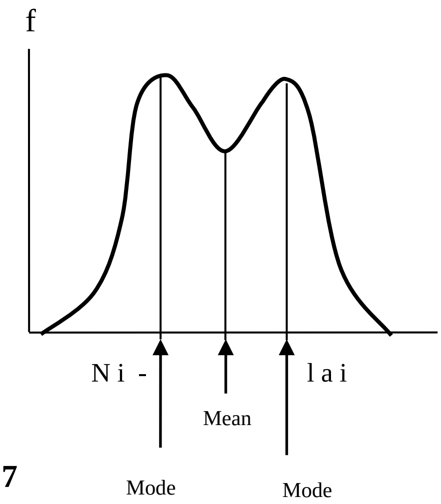
Grafik 16



Grafik Menunjukkan Letak Mean, Median, dan Mode dalam Distribusi Bel Tidak Normal



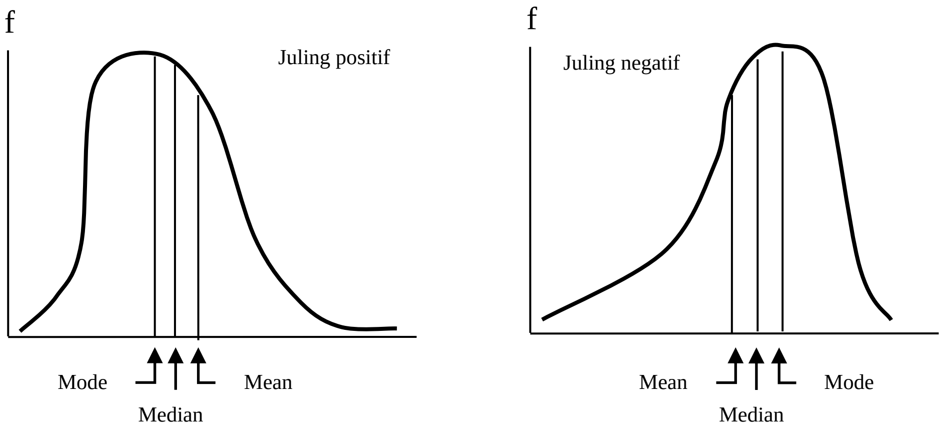
Grafik 17



Grafik Menunjukkan Letak Mean, Median, dan Mode dalam Distribusi Bentuk Trapesium dan Dwimode

Pada distribusi yang juling tempat kedudukan ketiga tendensi sentralnya terpisah satu sama lain. Bilamana distribusi juling positif, mean-nya terletak di sebelah kanan, sedang modenya di sebelah kiri. Median dari distribusi itu terletak di antara mean dan mode.

Pada distribusi juling negatif letak ketiga tendensi sentralnya secara berturut-turut dari kiri ke kanan adalah mean, median dan mode. Untuk jelasnya periksalah grafik 18.



Grafik 18

Grafik Menunjukkan Kedudukan Mean, Median, dan Mode dalam Distribusi Juling Positif dan Juling Negatif

4.11.BILAMANA MENGGUNAKAN MEAN, MEDIAN DAN MODE

Bilamana pada waktu yang bersamaan kita hanya ingin mengetahui salah satu dari ketiga referensi sentral itu, maka sebelum kita menghitungnya harus kita putuskan dahulu apakah mean, median atau mode yang akan kita pakai. Untuk memutuskan hal ini perlu kita mempertimbangkan pelbagai situasi.

1. *Waktu sangat terbatas, menggunakan Mode*

Bilamana seorang Inspektur Pendidikan hanya mempunyai waktu dua-tiga menit, tidak akan mungkin baginya menghitung mean atau median distribusi kecakapan anak-anak dalam suatu mata pelajaran. Atas dasar pertimbangan ini maka sebagai alat yang "kasar" dia harus puas dengan menggunakan mode sebagai statistik pengukuran tendensi sentral untuk mengetahui memusatnya kecakapan anak-anak.

2. *Kejadian khusus yang membutuhkan Mode*

Suatu perusahaan tidak akan suka menderita kerugian dengan memproduksi barang-barang yang tidak atau kurang laku. Sebaliknya, barang yang paling banyak dibeli konsumen-lah yang harus diproduksi dalam jumlah yang besar. Untuk itu perusahaan itu khususnya membutuhkan tendensi sentral distribusi barang-barang buatan perusahaannya yang berupa mode. Jadi misalnya, kalau suatu perusahaan sepatu tidak ingin menderita kerugian, maka ukuran sepatu yang paling lakulah yang perlu dibuat secara besar-besaran.

Demikian juga keadaannya seorang pendidik yang ingin mengetahui permainan apa yang paling digemari oleh anak-anak untuk disalurkan kea rah tujuan pendidikan.

3. *Untuk perhitungan statistik selanjutnya, kita membutuhkan Mean*

Dalam bab-bab selanjutnya akan kita ketahui bahwa Mean dapat digunakan untuk memperoleh informasi-informasi yang lebih banyak lagi. Tidak demikian halnya dengan Median dan Mode yang sekali sudah kurang sekali faedahnya. Median dan Mode adalah statistik terminal (statistik batas). Jika kita menghendaki informasi yang lebih banyak lagi, perhitungan mean menjadi keharusan.

4. *Adanya bahan-bahan yang hilang, Mean tidak dapat dihitung*

Kita misalkan ingin mengadakan testing kepada sejmlah anak-anak sekolah untuk menetapkan tendensi sentral kecerdasan anak-anak pada umur tertentu. Jarang sekali dalam keadaan demikian kita dapat mengadakan testing kepada anak-anak dari semua tingkat kecerdasan. Anak-anak dari golongan lemah kecerdasan tidak kita jumpai di sekolah-sekolah biasa. Dengan demikian bahan untuk menghitung mean kurang lengkap adanya. Mean seperti kita ketahui, sangat sensitif terhadap deviasi-deviasi nilai-nilai yang ekstrem. Perhitungan median atau mode akan memberi gambaran taksiran yang dapat memenuhi kebutuhan kita untuk menggambarkan tendensi sentral kecerdasan anak-anak itu.

5. *Distribusi dangat juling, melaporkan salah satu tendensi sentral memberi gambaran yang salah*

Jika keadaan distribusi sangat juling kita sangat mungkin kita memberi gambaran yang salah bila kita sajikan hanya salah satu tendensi sentral. Ambil sebuah contoh distribusi penghasilan rakyat atau negara. Untuk keperluan-keperluan propaganda, distribusi penghasilan yang biasa mengikuti kurva juling positif oleh Pemerintah Negara yang bersangkutan hanya dilaporkan mean-nya. Karena mean dalam distribusi juling berkedudukan di atas mode, maka orang akan mendapat gambaran yang kurang tepat tentang penghasilan rakyat negara itu. Ini dapat kita mengerti kalau kita gambarkan misalnya sebagian terbesar orang berpenghasilan Rp. 300,- sehari (jadi modenya ada Rp. 300,-), dan ada beberapa orang saja yang berpenghasilan jutaan rupiah. Dengan melaporkan mean, orang memasukkan mereka yang berpenghasilan jutaan rupiah sehari itu. Dengan demikian akan dijumpai mean, misalnya Rp. 600,-, sedang sebagian terbesar orang berpenghasilan hanya di sekitar Rp. 300,-.

Sebaliknya, negara musuh akan lebih cenderung untuk melaporkan mode penghasilan daripada mean-nya. Demikianlah negara-negara "komunis" akan lebih senang melaporkan mode penghasilan rakyat negara "kapitalis", sedang orang dari negara "kapitalis" lebih senang melaorkan mode penghasilan rakyat di negara "komunis".

Kalau toh diinginkan melaporkan salah satu tendensi sentralnya, kiranya median lebih memenuhi syarat. Kita lihat bahwa distribusi juling median terletak ditengah-tengah antara mean dan mode. Akan lebih baik kiranya bila dilaporkan ketiga-tiganya, agar pembaca dapat mengambil kesimpulan sendiri tentang kenyataannya.

6. *Dari segi stabilitas, Mean adalah tendensi sentral yang paling memuaskan*

Ditinjau dari segi stabilitasnya, mean merupakan tendensi sentral yang paling stabil, diikuti oleh median, kemudian oleh mode. Maksudnya, bila kita menyelidiki suatu kelompok, kemudian mengadakan penyelidikan berturut-turut kepada kelompok yang sejenis, dan kita menghitung ketiga tendensi sentral dari tiap-tiap kelompok itu, maka kita akan menjumpai tendensi-tendensi sentral itu yang berbeda untuk masing-masing kelompok. Akan tetapi perbedaan yang terkecil ialah perbedaan antar mean, sedang modenya akan menunjukkan perbedaan-perbedaan yang terbesar. Oleh karena kerapkali kita hanya dapat menguji sekelompok kecil anak-anak untuk menaksir kelompok anak-anak yang lebih besar jumlahnya, stabilitas ini merupakan unsur statistik yang sangat penting. Dalam hal semacam ini melaporkan atau mendasarkan diri pada mean akan lebih tepat.

Itulah beberap faktor yang mempengaruhi pemilihan tendensi sentral. Sebagai kesimpulan dapat dikatakan bahwa :

1. Mean biasanya dipilih orang sebagai pengukuran tendensi sentral, terutama bilamana distribusi mendekati normal, sebab mean mempunyai stabilitas yang terbesar dan dapat digunakan sebagai dasar perhitungan statistik selanjutnya.
2. Median adalah nilai variabel yang ditengah-tengah dan biasanya dipandang paling tepat untuk menggambarkan tendensi sentral bila distribusi menunjukkan "keistimewaan", seperti sangat juling, adanya bahan-bahan yang tidak lengkap, dan semacamnya.

3. Mode rupa-rupanya menjadi suatu alat yang paling sederhana untuk menaksir tendensi sentral dalam keadaan tergesa-gesa, atau bilamana orang mencari keadaan-keadaan yang istimewa (seperti mode ukuran sepatu, dan semacamnya)

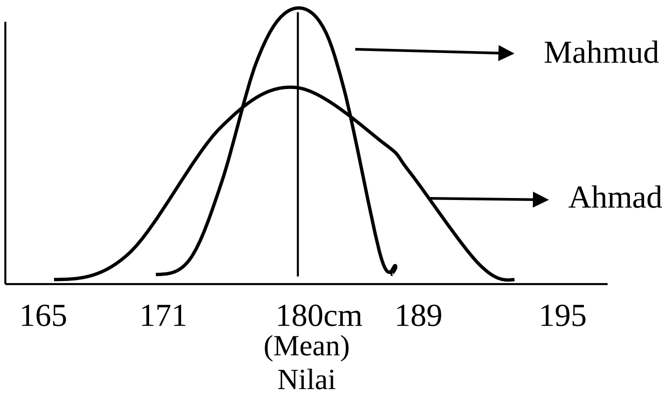
Kesimpulan-kesimpulan di atas masih belum menggambarkan semua kemungkinan. Dan perlu kita ketahui bahwa tiap-tiap bahan harus diselidiki sedemikian rupa untuk menetapkan pengukuran tendensi sentral mana yang paling tepat dipilih untuk keperluan-keperluan tertentu.

POKOK BAHASAN:

“ PENGUKURAN VARIABILITAS”

Dalam penyelidikan-penyelidikan, kerapkali kita membutuhkan informasi yang lebih banyak daripada hanya mengetahui salah satu tendensi sentral saja. Kita ingin, misalnya, mengetahui bagaimana penyebaran tiap-tiap nilai dari tendensi sentral itu. Hal inilah yang menjadi pusat perhatian kita dalam bab ini.

Kita misalkan kepada seornag pemimpin Pusat Latihan diserahkan tugas melatih dua orang peloncat tinggi untuk diajukan dalam kejuaraan nasional. Kedua peloncat itu, kita sebut saja namanya Ahmad dan Mahmud, mendapat latihan selama sebulan secara terus menerus. Dengan sangat teliti pemimpin tersebut mencatat loncatan-loncatan kedua peloncat tersebut.



Grafik 19

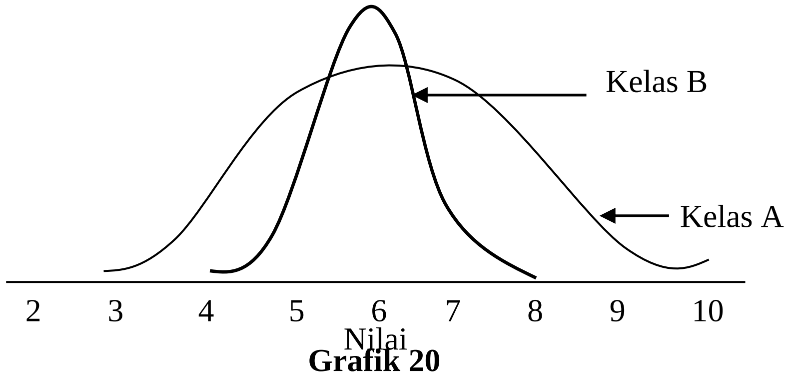
Grafik Menunjukkan Distribusi Loncatan Ahmad dan Mahmud selama satu bulan Latihan

Dengan memeriksa grafik tersebut kita dapat melihat adanya perbedaan antara loncatan kedua orang itu. Ahmad menunjukkan loncatan yang dapat dipastikan : kadang-kadang dia dapat meloncat setinggi 195, tetapi kadang-kadang dia hanya dapat meloncat setinggi 165 cm. Mahmud, sebaliknya, menunjukkan loncatan yang lebih mantap : Sungguhpun dia tidak pernah meloncat setinggi 195 cm, tetapi dia juga tidak pernah meloncat serendah 165 cm. Paling rendah loncatannya adalah 171 cm, sedang paling tinggi 189 cm.

Suatu masalah akan dihadapi oleh pemimpin pusat latihan ini bilamana dia diminta untuk memutuskan siapa yang harus dimajukan dalam perlombaan kejuaraan nasional itu apabila hanya seorang saja yang diperkenankan untuk dimajukan. Memajukan Ahmad berarti menempuh resiko yang besar, sebab sungguhpun ada kemungkinan dia meloncat setinggi 195 cm, yang berarti mungkin dia dapat menggondol juara pertama, akan tetapi mungkin sekali dia hanya dapat meloncat disekitar 165 cm, yang berarti tidak ada kejuaraan yang dapat diperolehnya. Sebaliknya memajukan Mahmud juga ada untung ruginya. Kalau dapat dikira-kirakan dengan agak pasti bahwa loncatan setinggi 175cm sampai 185cm akan menggondol kejuaraan yang kedua atau ketiga, maka dengan memajukan Mahmud pemimpin latihan itu sudah dapat memastikan bahwa piala kejuaraan nomor tiga atau nomor dua akan dapat diperolehnya. Kelemahannya ialah bahwa kejuaraan pertama sangat sedikit kemungkinannya dapat dicapai oleh Mahmud.

Keadaan semacam itu kerapkali dijumpai juga dalam bidang lain. Dua kelas anak-anak Sekolah Menengah mungkin seringkali menunjukkan nilai rata-rata yang sama dalam suatu ata ujian, katakalah saja ujian aljabar. Sekalipun nilai rata-ratanya sama, akan tetapi kelas yang satu menunjukkan *penyebaran* nilai-nilai perorangan yang lebih besar daripada kelas lainnya. Hal ini pun dapat kita mengerti dari pemeriksaan terhadap grafik tersebut halaman berikut.

Kita lihat dari grafik itu bahwa nilai rata-rata (Mean) dari kedua kelas itu adalah sama, yaitu 6. Akan tetapi nilai-nilai anak-anak dalam kelas A menunjukkan penyebaran dari angka rata-rata yang lebih besar dibandingkan dengan kelas B. Dalam kelas A ada beberapa anak yang mendapat nilai-nilai tinggi seperti 8, 9, 10. Akan tetapi nilai-nilai yang sangat rendah juga kita jumpai dalam kelas itu, yaitu nilai-nilai 2, 3, dan 4. Keadaan semacam itu tidak ada yang mendapat nilai-nilai yang sangat mencolok baiknya, juga tidak ada yang mendapat nilai-nilai yang sangat mencolok buruknya. Nilai terendah dalam kelas ini 4,5 sedang yang tertinggi adalah 7,5. Dikatakan orang bahwa nilai-nilai dalam kelas A adalah *heterogen*, sedang anak-anak dalam kelas B adalah *homogen*.



Grafik 20

Grafik menunjukkan penyebaran nilai-niali aljabar dari dua kelas B

Dengan kedua contoh diatas kita kembali kepada pokok persoalannya. Kita tengok kembali tinggi loncatan Ahmad : beberapa loncatan menyebar agak jauh dari mean loncatannya, dibandingkan dengan loncatan Mahmud. Dalam kelas A terdapat anak-anak yang kecakapannya dalam mata pelajaran Aljabar menyebar sangat jauh dari kecakapan anak arat-rata. Dengan istilah statistik dikatakan bahwa loncatan Ahmad mempunyai *variabilitas* yang lebih besar daripada loncatan Mahmud. Kelas A mempunyai *variabilitas* yang lebih besar daripada kelas B dalam soal kecakapan Aljabar.

5.1 DEFINISI

Yang dimaksud dengan ***variabilitas*** adalah *derajat penyebaran nilai-nilai variabel dari suatu tendensi sentral dalam suatu distribusi*. Bilamana dua distribusi, katakanlah distribusi A dan distribusi B dibandingkan, dan distribusi A menunjukkan penyebaran nilai-nilai variabelnya yang lebih besar dari pada distribusi B, maka dikatakan bahwa distribusi A mempunyai variablitas yang lebih besar dari distribusi B. Variabilitas ini disebut *dispersi*.

5.2 PERLUNYA INDEKS VARIABILITAS

Pengukuran tentang variabilitas termasuk dalam bidang statistik deskriptif. Dari itu mudah dimengerti bahwa pengukuran tentang variabilitas mempunyai arti praktis. Hal ini dapat kita lihat dari contoh-contoh di atas.

Disamping nilai praktis, pengukuran variabilitas mempunyai arti teoritik yang sangat penting. Hal ini akan kita ketahui setelah kita sampai pada pembicaraan-pembicaraan agak lanjut. Senetara ini marilah kita fahami beberapa cara untuk mencari variabilitas itu. Ada beberapa macam cara untuk mencari variabilitas. Disini yang akan kita bicarakan hanyalah yang penting-penting saja, yaitu Range, Mean Deviation, dan Standard Deviation.

5.3 RANGE

Range adalah pengukuran variabilitas yang paling sederhana. Seperti sudah dibicarakan dimuka, Range adalah *jarak* antara nilai yang tertinggi dengan nilai yang terendah. Dengan rumus : $R = X_t - X_r$.

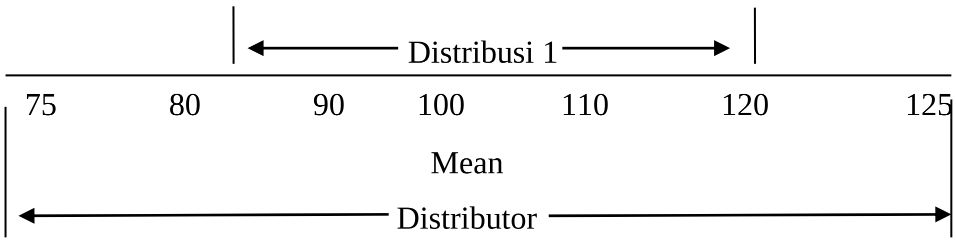
dalam mana :

R = Range

X_t = Nilai tertinggi

X_r = Nilai terendah

Dua buah distribusi tentang suatu variabel pengukuran yang sama dapat diletakkan pada satu sumbu yaitu sumbu nilai (absis). Kita misalkan variabel itu adalah nilai kecerdasan, dan diletakkan pada absis sebagai berikut :



Dari ilustrasi di atas kita dapat melihat bahwa distribusi II mempunyai variabilitas yang lebih besar daripada distribusi I. Range daripada distribusi I hanya 20 (dari 110-90), sedang range daripada distribusi II ada 50 (dari 125-75).

Akan tetapi range sebagai alat pengukuran variabilitas yang sangat sederhana ini mempunai penggunaan yang sangat terbatas. Range, sesuai dengan definisinya, sangat tergantung kepada dua nilai, yaitu nilai yang tertinggi dan nilai yang terendah. Kedua nilai ini adalah nilai-nilai yang ektrem dalam distribusi. Sebab itu range akan mempunyai fluktuasi yang sangat besar, tergantung kepada nilai-nilai ektrem itu. Hal ini dapat kita fahami daro contoh berikut :

Dalam penyelidikan-penyelidikan kerapkali kita hanya menghadapi jumlah subjek yang sangat terbatas, misalnya 20 orang atau 30 orang. Malahan tidak jarang kita hanya mempunyai 10 orang atau 10 ekor kelinci atau tikus. Kita misalkan kita hanya mempunyai 10 orang untuk diselidiki tentang kecepatannya meneliti kesalahan cetak suatu pemberitaan. Hasilnya adalah sebagai berikut : 11, 14, 14, 16, 17, 21, 21, 25, 26, 42. Angka-angka itu menunjukkan banyaknya huruf-huruf yang dapat diperiksa dalam tiap menitnya. Range dari distribusi itu adalah 42-11, atau 31. Kita misalkan sekarang orang yang ke-10 tidak turut serta dalam percobaan itu; maka R-nya menjadi 26-11=15. Peniadaan satu orang saja meniadakan separo dari range.

Dari contoh di atas nampak dengan jelas, bahwa range sangat tergantung kepada dua nilai yang ekstrem, ya malahan tergantung hanya kepada satu nilai yang ekstrem. Itulah sebenarnya kenapa range mempunyai fluktuasi yang sangat besar. Tentu saja tiap-tiap pengukuran variabilitas harus memperhitungkan tiap-tiap kejadian dalam penyelidikan. Akan tetapi alat pengukuran yang ideal kiranya tidak akan mempunyai fluktuasi yang begitu besar bilamana hanya satu orang saja ditiadakan. Tinggi loncatan Mahmud yang terkenal dapat dipastikan itu mungkin pada suatu ketika hanya setinggi lutut, disebabkan karena dia menderita sakit perut yang mendadak. Dengan demikian satu kali loncatan yang sial itu dapat dipandang sebagai cermin dari tinggi loncatan yang sebenarnya.

Kelemahan lain dari range sebagai alat pengukuran cariabilitas adalah karena range tidak memenuhi definisi untuk menjadi alat semacam itu. Seperti kita ketahui, variabilitas menunjukkan penyebaran nilai-nilai *di sekitar tendensi sentral*, dan dalam range tidak dapat kita lihat petunjuk-petunjuk dimana letak tendensi sentralnya. Dengan kata lain Range tidak menunjukkan bentuk distribusi. Sangat sukar untuk kita gambarkan bahwa suatu kelas yang mempunyai anak-anak yang berotak cemerlang dan hany mempunyai seorang saja yang bodoh dikatakan mempunyai variabilitas yang sama dengan kelas lainnya yang penghuninya sebagian terdiri dari anak-anak yang bodoh, sebagian anak-anak sedang, dan sebagian kecil saja (atau hanya seorang) yang berotak cemerlang.

Kelemahan-kelemahan yang prinsip itu, menyebabkan range dipandang orang bukan alat pengukuran variabilitas yang mantap, dan oleh karen aitu jarang digunakan orang. Bilamana range digunakan untuk mengukur variabilitas, biasanya orang menginsyafi kelemahannya, dan hanya digunakan hanya dalam keadaan-keadaan yang sangat memaksa. Bilamana hanya tersedia waktu satu atau dua menit untuk mengetahui variabilitas anak-anak dalam suatu kelas, kita dapat menggunakan range sebagai "petunjuk" semata-mata. Kebaikan lain dari range adalah tiap-tiap orang dapat mengertinya dengan mudah.

5.4. RANGE 10-90

Nilai-nilai yang ekstrem (terlalu rendah dan terlalu tinggi) adalah nila-nilai yang tidak mantap, tidak stabil. Range sangat tergantung kepada nilai-nilai ekstrem itu. Maka untuk menghindari nilai-nilai yang tidak mantap itu dapat diambil range yang lebih sempit, yaitu range antara persentil yang ke 10 dengan persentil yang ke 90. Dengan range 10-90 ini distribusi dipotong 20 persen masing-masing 10 persen pada tiap-tiap ujungnya.

Bagaimana menghitung R 10-90 itu dari suatu distribusi?

Menghitung R 10-90 dari suatu distribusi tidak akan sulit apabila kita telah memahami betul-betul cara bagaimana mencari P_{10} dan P_{90}. Ambil sebagai ilustrasi cara mencari R 10-90 dari bahan Tabel 20 yang telah dikutip lagi dalam Tabel 25 berikut ini.

Tabel 20

Tabel untuk ilustrasi mencari R 10-90

Interval Nilai		cf
195-199	1	34
190-194	5	33
185-189	8	28
180-184	10	20
175-179	6	10
170-174	3	4
165-169	1	1
Jumlah	34	-

Dari Tabel itu akan dapat kita temukan :

$$P_{90} = B_b + \frac{\frac{90}{10}N - cfb}{Fd}\ i$$

$$= 189{,}5 + \frac{30{,}6-28}{5}\ 5$$

$$= 192{,}1$$

$P_{10} \quad = B_b + \frac{\frac{1}{10}N - cfb}{Fd}\ i$

$= 169{,}5 + \frac{3{,}4 - 1}{3}\ 5$

$= 173{,}5$

Karena R 10-90 adalah P_{90}-P_{10} maka

R 10-90 = 192,1-173,5 = 18,6 (23)

Jika dihitung dengan Range penuh maka R = 199-165 = 34

Apabila kita ingin mencari R 10-90 dari bahan Tabel 21 maka perhitungannya akan seperti tercantum sebagai berikut :

$P_{90} \quad = 24 \text{ tahun} + \frac{678 - 653{,}0}{34}\ 1 \text{ tahun}$

= 24,75 tahun

$P_{10} \quad = 17 \text{ tahun} + \frac{75{,}4 - 55{,}0}{105}\ 1 \text{ tahun}$

= 17,19

Jadi R 10-90 = P_{90}-P_{10} = 24,75 tahun – 17,19 tahun

= 7,56 tahun.

Jika dihitung dengan range penuh maka akan kita jumpai

R = 30 tahun ke atas sampai 17 tahun ke bawah

Tabel 21

Tabel untuk contoh mencari range 10-90

Umur dihitung dari peringatan hari kelahiran terakhir	Frekuensi (f)	Frekuensi Kumulatif (cf)
30 tahun ke atas	23	754
29	3	731
28	13	728
27	12	715
26	16	703
25	34	687
24	27	653
23	72	626
22	92	554
21	90	462
20	104	372
19	(108)	268
18	105	(160)
17	51	55
Di bawah 17 tahun	4	4
Jumlah	754	-

Dengan kata lain R = sepanjang umur manusia, sebab 30 tahun ke atas dapat mencakup mereka yang berumur 100 tahun dan 17 tahunke bawah dapat mencakup mereka yang baru berumur 2 tahun.

Dari membandingkan hasil-hasil perhitungan R penuh dengan R 10-90 terhadap bahan yang sama seperti tersebut diatas dapat kita ambil kesimpulan :

1. R 10-90 selalu lebih kecil dari pada R penuh
2. R penuh tak dapat dihitung dari bahan distribusi terbuka.

Range 10-90 secara teoritik maupun praktis memang lebih mantap daripada range penuh. Karena itu ia merupakan alat mendeskripsi variabilitas yang lebigh baik, *a better measure of valiability*. Sungguhpun demikian kelemahan daripada range masih terkandung didalamnya, sebab ia masih tergantung kepada satu atau dua niali diujung-ujung distribusi. Tetapi kelemahan ini lebih ringan jika dibanding dengan kelemahan range penuh.

5.5. RANGE 25-75

Range 10-90 memotong 10 persen dari jumlah frekuensi pada masing-masing ujung distribusi. Range 25-75 memotong lebih banyak lagi, yaitu masing-masing 25 persen. Jadi seluruhnya range ini meniadakan 50 persen dari jumlah frekuensi yang terdapat pada ujung distribusi.

Range 25-75 disebut juga *Range antar kwartil* yang dihitung dari persamaan sebagai berikut :

$R_{25\text{-}75} = P_{75\text{-}25} = K_3 - K_1$

Dari distribusi Tabel 25 akan dapat kita temukan :

$$P_{75} = B_b + \frac{\frac{3}{4}N - cfb}{Fd}\ i$$

$$= 184{,}5 + \frac{25{,}5 - 20}{8}\ 5$$

$$= 187{,}94.$$

$$P_{25} = B_b + \frac{\frac{1}{4}N - cfb}{Fd}\ i$$

$$= 174{,}5 + \frac{8{,}5 - 4}{6}\ 5$$

$$= 178{,}25.$$

Jadi $R_{25\text{-}75}$ = 187,94 – 178,25 = 9,69.

P75 = K3 = 23,16 tahun

P25 = K1 = 19,26 tahun

Jadi

$R_{25\text{-}75}$ = 23,16 tahun – 19,26 tahun = 3,9 tahun.

Range antar kwartil ini selalu lebih kecil daripada range 10-90 dan agak lebih mantap. Namun demikian sifat-sifat dari range masih juga terdapat pada range antar kwartil ini.

5.6. RANGE SEMI ANTAR KWARTIL

Range semi antar kwartil tidak lain adalah separo dari range antar kwartil. Rumusnya adalah :

RSAK $= \frac{P75-25}{2} = {}^{1}/_{2}\ [K3 - K1]$

Jadi dari bahan Tabel 25 maka :

RSAK – nya = ${}^{1}/_{2}\ [9{,}69] = 4{,}845$

Dari Tabel 26

RSAK –nya = ${}^{1}/_{2}\ [3{,}9\ tahun] = 1{,}95\ tahun$

Range ini mempunyai sifat lebih baik daripada range-range yang dibicarakan terdahulu dan biasanya digunakan bersama-sama dengan Median. Median untuk menyelidiki tendensi sentral, sedang RSAK untuk menyelidiki variabilitasnya.

5.7. MEAN DEVIAN

Mean Deviation atau Average Deviation atau Deviasi rata-rata adalah rata-rata dari devisi nilai-nilai dari Mean dalam suatu distribusi, diambil nilainya yang absolute. Yang dimaksud dengan deviasi absolut adalah nilai-nilai yang positif.

Secara aritmetik mean deviasi dapat didefinisikan sebagai mean dari harga mutlak deviasi nilai-nilai individual.

Untuk dapat menyelesaikan pekerjaan mencari mean deviasi pertama-tama haruslah temukan mean. Kemudian ditentukan berapa besarnya penyimpangan tiap-tiap nilai dari mean itu. Misalnya, jika seorang mempunyai IQ 110, sedang mean IQ dari grupnya = 100, maka deviasi IQ orang tersebut adalah 110-100 = +10. Jika orang lain dalam grup itu mempenyuai IQ 85, maka deviasi IQ orang itu adalah 85-100 = 15. Deviasi yang bertanda plus menunjukkan deviasi diatas mean, sedang yang bertanda minus menunjukkan deviasi di bawah mean. Akan tetapi dalam perhitungan mean deviasi tanda minus ditiadakan. Dalam statistik, deviasi diberi simbul dengan huruf-huruf kecil seperti x, y, d, dan sebagainya. Rumusnya adalah x = X – M atau y – M,d = D – M, dan sebagainya.

Adapun rumus dari mean deviasi adalah :

$$MD = \frac{\Sigma|x|}{N}$$

dalam mana :

MD = Mean Deviasi

Σ|x| = Jumlah deviasi dalam harga mutlaknya,

N = Jumlah Individu/kasus

Bagaimana menerapkan rumus itu untuk memperhitungkan mean deviasi dari suatu distribusi dapat dilihat dari contoh sederhana berikut :

Tabel 22

Tabel untuk Contoh Mencari Mean Deviation

Nilai Variabel	Deviasi dai Mean dengan nilainya absolut x
19	5
18	4
17	3
16	2
15	1
14	0
13	1
12	2
11	3
10	4
9	5
-	Σ|x|=30

Dengan N = 11 dan Σ|x| = 30 maka

MD = $\frac{30}{11}$ = 2,73

Contoh lain :

Tabel 23

Tabel untuk Contoh Mencari Mean Deviasi

x	f	fX	|x|	f|x|
Rp. 13,-	1	13	1,57	1,57
Rp. 12,-	3	36	0,57	1,71
Rp. 11,-	1	11	0,43	0,43
Rp. 10,-	2	20	1,43	2,86
Total	7	80	...	6,57

Rumus untuk menyelesaikan pekerjaan ini adalah :

MD = $\frac{f|x|}{N}$

N = Σf = 7

Jadi MD = $\frac{6,57}{7}$

= 0,94

Dibandingkan dengan semua cara pengukuran variabilitasnya yang telah dibicarakan lebih dahulu mean definisi ini tetap lebih maju. Ia tidak membuang data sedikitpun. Nilai-nilai yang ektrem tetap dipakai.

Keunggulan mean deviasi terhadap pengukuran variabilitas dengan range adalah mulai dipenuhinya definisi tentang variabilitas oleh mean deviasi itu, yaiut penyebaran nilai-nilai yang ditinjau dari tendensi sentral. Seperti kita ketahui mean deviasi menggunakan mean nilai-nilai itu dicarilah deviasi-deviasinya dalam harga mutlaknya. Kemudian dicari mean dari jumlah deviasi yang telah diketemukan.

Akan tetapi mean deviasi mempunyai satu kelemahan pokok, karena cara perhitungannya mengabaikan tanda-tanda plus dan minus, sehingga oleh karenya mean deviasi tidak dikenai perhitungan-perhitungan matematik yang tetap mempertahankan nilai-nilai plus dan minus. Untuk mengatasi kelemahan itu maka timbullah cara pengukuran variabilitas lain, yaitu “Standard Deviasi”

5.8. STANDARD DEVIASI

Ini adalah alat statistik yang pertama-pertama meminta agak banyak pekerjaan untuk menyelesaikannya. Dibanding dengan mengerjakan mean deviasi, ada beberapa langkah lagi yang harus ditambahkan.

Secara matematik standard deviasi dibatasi sebaga “*akar dari jumlah deviasi kuadrad dibagi banyaknya individu*” dalam distribusi.

Dengan definisi itu kita masih sukar untuk mendapatkan gambaran yang jelas. Memang ini merupakan konsep statistik yang agak sukar dibayangkan. Suatu contoh mungkin dapat memberi penjelasan. Kita ambil kembali contoh tersebut di atas, yang kita ambil sekarang bukan nilai deviasi yang absolute, tetapi nilai deviasi yang nyata.

Untuk mencari standard deviasi pertama-tama kita harus mencari mean. Ini dapat dicari dengan rumus yang sudah kita ketahui.

$$M = \frac{X}{N} = \frac{154}{11} = 4$$

Dengan mengetahui mean ini kita dapat mencari deviasi nilai individual dari mean. Ini mencantumkan dalam kolom kedua. Jumlah deviasi dari mean ini, yaitu Σx, harus sama dengan NUL.

Tabel 24

Tabel untuk contoh mencari standard deviasi

Nama Variabel	Deviasi dari mean (x)	Deviasi dari mean kwadrat (x^2)
19	+5	25
18	+4	16
17	+3	9
16	+2	4
15	+1	1
14	0	0
13	-1	1
12	-2	44
11	-3	9
10	-4	16
9	-5	25
Jumlah : 154	$\Sigma x=0$	$\Sigma x^2=110$

Kololm ketiga disediakan untuk deviasi kuadrat yang diperoleh dengan mengkuadratkan tiap-tiap deviasi individual dan diisikan dalam kolom itu pada baris-baris yang bersangkutan. Dari ini kita memperoleh jumlah deviasi kuadrat itu, Σx^2. Dengan membagi jumlah deviasi kuadrat ini dengan jumlah individu kejadian dalam distibusi (dalam hal ini jumlah itu adalah 11), dan kemudian mengakarnya, kita akan memperoleh Standard Deviasi.

Langkah-langkah itu akan dilambangkan dengan rumus sebagai berikut :

$$SD = \sqrt{\frac{x2}{N}}$$

dalam mana :

SD = Standard Deviasi

Σx^2 = Jumlah deviasi kuadrat dan

N = Jumlah individu/kejadian dalam distribusi.

Diulangi lagi langkah-langkah mencari standard deviasi. Pertama dicari mean, kemudian dicari deviasi tiap-tiap individu dari mean yang ditemukan itu. Deviasi-deviasi itu masing-masing dikuadratkan dan kemudian dijumlahkan. Jumlah dari deviasi kuadrat ini dibagi dengan N dan kemudian diakar. Dengan demikian SD dapat dibatasi sebagai "*akar dari rata-ratanya deviasi kuadrat*". Dengan begitu maka kita memperoleh standard deviasi distribusi sebagai berikut :

$$SD = \sqrt{\frac{x2}{N}}$$

$$= \sqrt{\frac{110}{11}}$$

$$= \sqrt{10}$$

$$= 3{,}162.$$

Dasar fikiran yang ada dibalik standard deviasi adalah bahwa untuk menghitung variabilitas, tanda-tanda (negatif atau positif) harus tidak diabaikan. Sebab itu tanda-tanda itu tetap dipertahankan. Menjadi prinsip matematik bahwa baik bilangan negatif maupun positif akan menjadi positif bilamana dikuadratkan. Itulah sebabnya deviasi kuadrat yang terdapat dalam kolom ketiga itu semua merupakan bilangan positif.

Perlu kita ketahui bahwa standard deviasi memperhitungkan penyebaran nilai individual dari mean, seperti halnya perhitungan mean deviasi. Akan tetapi oleh karena dalam meperhitungkan standard deviasi ini semua tanda-tandanya (baik positif maupun negatif) dipertahankan, maka SD ini dapat digunakan untuk perhitungan-perhitungan selanjutnya yang mempergunakan prinsip-prinsip matematik.

Kuadrat dari standard deviasi disebut varius. Dengan demikian varians dapat dikatakan sebagai mean dari jumlah deviasi kuadrat atau dinyatakan dalam rumus :

$$V = SD^2 = \frac{x^2}{N}$$

Varians dari sederetan data yang berupa angka-angka mempunyai arti penting untuk menguji sesuatu hipotesis. Sifat dan penggunaan varians ini akan dibicarakan dalam bagian yang lanjut.

5.9. ARTI STANDARD DEVIASI

Sebelum kita membicarakan lebih lanjut bagaimana menghitung SD, marilah kita fahami dahulu apa arti SD itu. SD adalah suatu statistik yang digunakan untuk menggambarkan variabilitas dalam suatu distribusi maupun variabilitas beberapa distribusi. Kita peroleh SD = 3,162. Apa artinya ini?

SD dapat dipandang sebagai satuan pengukuran sepanjang absis dari suatu poligon. Untuk memahami ini, marilah kita gambarkan absis suatu poligon. Dari absis ujung sebelah kiri ke ujung sebalah kanan berturut-turut terdapat skala pengukuran dari nilai yang terendah sampai nilai yang tertinggi. Jarak antara kedua ujung itu, seperti sudah kita ketahui, disebut range. Range, dengan demikian dapat pula dipandang sebagai luas pengukuran dari suatu distribusi. SD dapat pula kita gambarkan sebagai *luas* pengukuran semacam range. Hanya saja SD *tidak seluas range*.

Dalam contoh diatas rangenya ada 10 (dari 19-9), sedang SD-nya ada 3,162. Jadi SD-nya kira-kira hanya ada sepertiga R-nya.

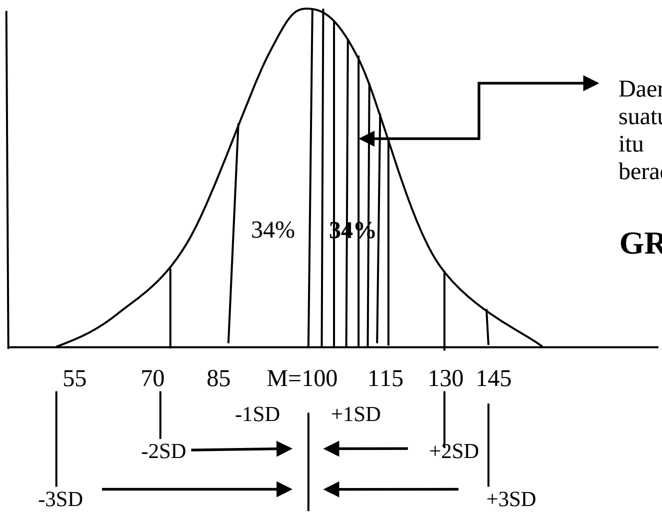
SD, seperti juga range, selalu dinyatakn dalam satuan angka kasar. Kalau satuan pengukurannya berupa cm, SD-nya juga berupa cm. Bilamana satuan pengukurannya berupa satuan mata uang, SD-nya juga berupa satuan mata uang, seperti rupiah, ringgit, dolar, yen, dsb.

Bilamana dari suatu distribusi normal kita buat grafik poligon, kita akan menjumpai suatu hal yang sangat menarik. Perlu kita ketahui bahwa suatu distribusi mempunyai ± 6 (enam) SD. Di luar 6 SD ini persentase frekuensinya sangat kecil, yaitu kurang dari 0,3 persen sehingga dalam ilmu-ilmu sosial secara praktis tidak diperhitungkan. Dalam grafik berikut kita misalkan suatu distribusi mempunyai mean 100 dan SD 15. Kita ingin membentangkan absis dalam satuan SD. Untuk ini kita mulai dengan meletakkan meannya dan menambahkan satu SD diatas mean (Ingat, SD adalah suatu pengukuran deviasi dari mean). Karena mena adalah 100 dan SD sama dengan 15, maka titik diatas mean yang kita maksudkan itu ada 115 (dari 100 + 15). Apabila kita terus menambahkan satu SD diatas titik baru ini (115), maka kita akan memperoleh titik-titik dengan nilai 130 dan 145. Hal semacam ini dapat kita kerjakan juga dengan titik-titik di bawah mean (Perlu kita ingat, bahwa deviasi itu bukan saja di atas mean, tetapi juga di bawah mean). Dengan cara itu kita akan memperoleh titik-titik dengan nilai-nilai 85 (dari 100-15), 70 (dari 100-2x15) dan 55 (dari 100-3x15).

Kita catat lagi : *jarak I* antara angka 100 dan 115 adalah *satu* SD. Jarak (satu SD) itu diisi dengan 15 angka kasar. Jarak ini biasanya disebut satu SD di atas mean, atau + 1 SD. Jarak antara 100 dan 130 ada + 2 SD, dan diantara 100 dan 145 ada + SD. Jarak antara 100 dan 85 juga ada 1 SD. Akan tetapi karena *jarahnya* ke nilai yang *lebih rendah* dari mean, jarak itu diberi tanda *negatif*. Dengan kata lain, antara 100 dan 85 adalah -1SD. Jarak antara 100 da n70 dan 55 berturut-turut adalah -2SD dan -3SD.

Kita ingat kembali bahwa kurve normal adalah kurve simetri. Oleh sebab itu tiap-tiap ordinat yang didirikan di atas titik + 1SD dan -1SD akan sma tingginya. Demikian juga ordinat-ordinat yang didirikan di atas titik +2SD dan – 2SD, dan di atas +3SD dan -3SD. Ordinat ini mewakili banyaknya frekuensi variabel pada tiap-tiap titik itu.

Adapun banyaknya frekuensi variabel *dari* mean ke titik + 1SD diwakili oleh "*daerah*" yang dibatasi oleh ordinat dari mean dan ordinat dari titik +1SD itu. Daerah ini meliputi kira-kira 34 persen dari frekuensi seluruhnya. Demikian juga daerah -1SD meliputi 34 persen dari frekuensi seluruhnya. Bilamana kedua daerah itu digabungkan akan kita peroleh frekuensi sebanyak 68 persen dari frekuensi seluruhnya. Dengan kata-kata lain : daerah satu SD *disekitar* mean mencakup krang lebih 68 persen dari seluruh frekuensi dalam distribusi. Hal ini akan dibicarakan lebih lanjut dalam bab tentang KURVE NORMAL.



Daerah Kurve : Mewakili Frekuensi dari mean ke suatu titik nilai lainnya. Dalam contoh ini daerah itu meliputi 34% dari frekuensi seluruhnya, berada dalam jarak +1 SD.

GRAFIK 21

5. 10 CARA-CARA LAIN UNTUK MENGHITUNG SD

Oleh ahli-ahli statistik sampai sekarang SD dipandang sebagai alat pengukuran variabilitas yang paling memuaskan. Pekerjaan-pekerjaan statistik yang lanjutpun banyak yang didasarkan atas perhitungan SD ini. Oleh sebab itu pemahaman konsep SD ini hendaknya mendapat perhatian yang mendalam dari kita.

Cara-cara lain untuk menghitung SD yang akan dibicarakan di bawah ini dikembangkan oleh ahli-ahli statistik untuk memberi kemudahan dan daya guna bagi mereka yang membutuhkannya. Atas dasar semacam itu penyajian ini hendaknya dipandang untuk menolong mempermudah pekerjaan kita sendiri.

Rumus untuk menghitung SD seperti yang telah dibicarakan di muka adalah rumus yang paling sederhana. Dalam banyak hal kita akan menjumpai sesederhana seperti yang diberikan dalam contoh itu. Frekuensi dari tiap-tiap nilai tidak akan satu, melainkan berbeda-beda, bergerak dari bilangan 0 ke bilangan yang tak terhingga.

Rumus untuk menghitung SD dari distribusi yang tidak sama frekuensi tiap-tiap nilai variabelnya adalah sebagai berikut :

$$SD = \sqrt{\frac{fx^2}{N}}$$

(30)

Kedua rumus yang telah kita ketahui itu disebut *rumus deviasi*. Disebut demikian karena rumus itu menggunakan deviasi dari mean sebagai salah satu komponennya. Di halaman berikut adalah contoh mencari SD dengan rumus itu.

Tabel 25

Tabel untuk menghitung SD dengan rumus deviasi

x	**f**	**fX**	**x**	**fx**	**fx^2**
10	3	30	+3,60	10,80	38,88
9	9	81	+2,60	23,40	60,84
8	13	104	+1,60	20,80	33,28
7	23	161	+0,60	13,80	8,28
6	24	144	-0,40	9,60	3,84
5	13	65	-1,40	18,20	25,48
4	10	40	-2,40	24,00	57,60
3	5	15	-3,40	17,00	57,80
	N = 100 ΣfX = 640			Σfx^2 = 286,00	

M $= \frac{fx}{N}$

$= \frac{640}{100}$

$= 6{,}40$

SD $= \sqrt{\frac{fx^2}{N}}$

$= \sqrt{\frac{286{,}00}{100}}$

$= \sqrt{2{,}86}$

$= 1{,}69$

5. 11. RUMUS ANGKA KASAR

Dari pekerjaan contoh di atas kita dapat merasakan bahwa penggunaan rumus deviasi untuk mencari SD memakan banyak tenaga dan sangat menjemukan. Bekerja dengan bilangan-bilangan decimal memang sangat sulit. Kecuali itu menghitung decimal-decimal itu kerapkali menimbulkan kesalahan. Untuk memudahkan pekerjaan kita rumus angka kasar ini mungkin lebih bermanfaat untuk dikemukakan. Rumusnya adalah sebagai berikut :

$$SD = \sqrt[2]{\frac{fX^2}{N} - \left[\frac{fx}{N}\right]} \qquad (31)$$

berikut ini adalah contoh menggunakan rumus itu :

Tabel 26

Contoh ini adalah contoh rumus angka kasar untuk mencari SD

X	f	fX	fx^2
10	3	30	300
9	9	81	729
8	13	104	832
7	23	161	1127
6	24	144	864
5	13	65	325
4	10	40	160
3	5	15	45
N = 100 ΣfX = 640		$\Sigma fX^2 = 4382$	

SD $= \sqrt[2]{\frac{fX^2}{N} - \left[\frac{fx}{N}\right]}$

$= \sqrt[2]{\frac{4382}{100} - \left[\frac{640}{100}\right]}$

$= \sqrt{43{,}82 - 40{,}96}$

$= \sqrt{2{,}86} = 1{,}69$

Ternyata perhitungan dari distribusi yang sama dengan kedua macam rumus itu memberi hasil yang sama. Hal ini dapat kita mengerti karena rumus angka kasar ini adalah penjabaran dari rumus deviasi. Bukti penjabaran ini adap diberikan seperti tersebut di bawah :

SD $= \sqrt{\frac{x^2}{N}}$

SD^2 $= \sqrt{\frac{x^2}{N}}$

$= \frac{(X-M)^2}{N}$ (x = X-M)

$= \frac{(x^2 - 2XM + M^2)}{N}$

$= \frac{X^2 - 2XM + M^2}{N}$

$= \frac{X^2}{N} - \frac{2XM}{N} + \frac{M^2}{N}$

$= \frac{X^2}{N} - \frac{2X(M)}{N} + \frac{NM^2}{N}$

$= \frac{X^2}{N} - \frac{2X}{N}\left[\frac{X}{N}\right] + M^2$

$\left(M = \frac{X}{N}\right) = \frac{X^2}{N} - 2\frac{X}{N}\left[\frac{X}{N}\right] + \left[\frac{X}{N}\right]^2$

$= \frac{X^2}{N} - 2\left[\frac{X}{N}\right]^2 + \left[\frac{X}{N}\right]^2$

SD^2 $= \frac{X^2}{N} - \left[\frac{X}{N}\right]^2$

Jadi SD $= \sqrt[2]{\frac{fX^2}{N} - \left[\frac{fx}{N}\right]}$

Dengan cara semacam ini dapat juga dibuktikan bahwa :

$\frac{fx^2}{N} = \frac{fx^2}{N} - \left[\frac{fX}{N}\right]^2$

Dan karena $\frac{fX}{N} = M$, maka rumus itu dapat juga ditulis sbb :

SD $= \sqrt{\frac{fX^2}{N} - M^2}$ (32)

Sampai sekian jauh kita bicarakan adalah mencari standars deviasi dari distribusi *tunggal*. Bagaimana hanya mencari SD dari distribusi *bergolong*?

Pada dasarnya mencari SD dari distribusi bergolong tidak ada bedanya dengan mencari SD dari distribusi tunggal. Rumus yang digunakan untuk distribusi tunggal berlaku sepenuhnya untuk mencari SD dari distribusi bergolong. Hanya saja nilai X disini bukan lagi mewakili nilai variabel individual, maelainkan mewakili *titik tengah* dari tiap-tiap interval kelas. Hal ini dapat kita periksa dari contoh berikut :

Tabel 27

Contoh untuk mencari SD dan distribusi bergolong

Interval	**Titik Tengah (X)**	**f**	**fX**	**X^2**	**fX^2**
115-119	117	1	117	13689	13689
110-114	112	0	0	12544	0000
105-109	107	11	1117	11449	125939
100-104	102	21	2142	10404	218484
95-99	97	22	2134	9409	206998
90-94	92	23	2116	8464	194672
85-89	87	14	1218	7569	105966
80-84	82	3	246	6724	20172
75-79	77	4	308	5929	23716
70-74	72	1	72	5184	5084
Jumlah	N=100	9530			

$\Sigma fX^2 = 914820$

SD $= \sqrt[2]{\frac{fX^2}{N} - \left[\frac{fx}{N}\right]}$

$= \sqrt[2]{\frac{914820}{100} - \left[\frac{9530}{100}\right]}$

$= \sqrt{9148{,}20 - 9082{,}09}$

$= \sqrt{66{,}11}$

= 8,13

Teranglah bahwa penggunaan rumus angka kasar untuk distribusi bergolong tidak hanya ada bedanya dengan penggunaannya untuk distribusi tunggaul. Perbedaannya hanya terletak disini : kalau untuk distribusi tunggal X mewakili nilai variabel individual, sedang untuk distribusi bergolong X adalah titik tengah (*midpoint*) dari interval kelas.

5.12 MENCARI SD DARI DISTRIBUSI BERGOLONG DENGAN RUMUS DEVIASI BERKODE

Untuk mencari SD dari distribusi bergolong dengan menggunakan rumus angka kasar kerapkali melihat angka-angka yang besar-besar. Untuk efisiensi dan mencegah kemungkinan kekeliruan bagi kita disediakan rumus lain, yaitu rumus deviasi berkode. Dengan rumus ini hasilnya relatif masih sama. Rumus deviasi berkode adalah sebagai berikut :

SD $= i \sqrt[2]{\frac{fX'^2}{N} - \left[\frac{fx'}{N}\right]}$

dalam mana :

i = luas interval

x’ = deviasi berkode dari mean terkaan

Di dalam halaman berikut diberikan contoh penggunaan rumus ini. Untuk memperoleh gambaran maka digunakan distribusi pada TABEL 32.

Kita lihat bahwa hasil perhitungan SD dari distribusi tergolong dengan menggunakan rumus angka kasar maupun rumus deviasi berkode secara leratif adalah sama.

Tabel 28

Untuk contoh mencari SD dengan rumus deviasi berkode

Interval	f	x’	fx’	Fx’^{2}
115-119	1	+4	+4	16
110-114	0	+3	+0	00
105-109	11	+2	+22	44
100-104	21	+1	+21	21
95-99	22	0	0	00
90-94	23	-1	-23	23
85-89	14	-2	-28	56
80-84	3	-3	-9	27
75-79	4	-4	-16	64
70-74	1	-5	-5	25
	N : 100	-	Σfx’ = -34	Σ fx’^{2} = 276

SD $= i^{2}\sqrt{\frac{fX'^{2}}{N} - \left[\frac{fx'}{N}\right]}$

$= 5^{2}\sqrt{\frac{276}{100} - \left[\frac{-34}{N100}\right]}$

$= 5\sqrt{2{,}76 - 0{,}1156}$

$= 5\sqrt{2{,}6444}$

= 5 x 1,626

= 8,13

Dengan membandingkan kedua pekerjaan itu kita dapat menarik kesimpulan bahwa rumus yang terakhir ini untuk mengerjakan distribusi bergolong sangatlah sederhana dan memungkinkan mengalami kekeliruan adalah sedikit sekali. Adapun interval yang akan kita gunkanan sebagai pangkal (yaitu interval dengan deviasi berkode nul) untuk menetapkan deviasi berkode bagi interval-interval lainnya, tergantung kepada kita sendiri. Dari manapun kita bermula dengan kode nul, hasil perhitungan yang terakhir SD-nya akan sama. Tetapi akan lebih baik jika kita mengambil interval yang ada di tengah, bukan diujung distribusi sebagai pokok pangkal pengkodean. Ini tidak berdasar sesuatu alasan pun kecuali untuk menghindarkan kita dari perhitungan dengan angka-angka yang besar-besar. Silahkan mencoba rumus itu dengan pangkal pengkodean (dengan kode deviasi nul) interval 75-79.

5.13. NILAI STANDARD

Di depan telah dibicarakan konsep SD sebagai pengukuran variabilitas. Seperti sudah kita ketahui. SD selalu dinyatakan dalam satuan angka kasar, seperti cm, rupiah, kilogram, hektar dan sebagainya, hal mana tergantung pada satuan pengukuran yang digunakan dalam distribusi.

Apa yang disebut nilai standard mempunyai satu keistimewaan, yaitu bahwa nilai standard tidak lagi tergantung kepada satuan pengukuran seperti cm, kg, dan sebagainya itu.

Nilai standard yang paling asli adalah nilai standard dalam literature statistik biasa disebut *z-score*. Menurut definisi z-score adalah suatu bilangan yang menunjukkan seberapa jauh suatu niali (angka kasar) menyimpang dari mean dalam satuan SD. Atau secara singkat, nilai standard adalah indeks deviasi sesuatu nilai.. Rumusnya adalah sebagai berikut:

$$z = \frac{X-M}{SD} \qquad (34)$$

dalam mana :

z = nilai standard

X = sesuatu berskala besar

M = Mean distribusi

SD = Standard Deviasi distribusi

Karena X-M = x, maka rumus itu dapat juga ditulis :

$$z = \frac{x}{SD} \qquad (35)$$

Dilihat dari deviasi dan rumusnya z-score dapat dipandang sebagai indeks pengukuran jarak semacam range atau SD. Bedanya dengan range dan SD ialah bahwa z-score tidak lagi menggunakan angka kasar dan satuan pengukuran, melainkan suatu jarak dalam satuan SD.

Konsep diatas dapat diterapkan dengan contoh. Kita misalkan seorang anak A mendpaat angka 70 dalam mata pelajaran sejarah. Mean dari distribusi angka sejarah dalam kelompok anak itu = 50, sedang SD-nya = 10. Kita menanyakan berap z-score dari anak itu, berapa z_a.

$$z = \frac{X-M}{SD} = \frac{70-50}{10} = \frac{+20}{10}$$

$$z = +2$$

Jadi z-score dari anak A dalam mata pelajaran sejarah = +2. Itu berarti bahwa nilai sejarah dari A ada 2 SD *di atas* mean. Dikatakan di atas karena tanda positif di depan bilangan SD-nya.

Adakah jelas bahwa bilangan 70 menyimpang 20 dari M yang besarnya 50. Karena SD-nya = 10 maka penyimpangan 20 itu dalam satuan SD sama dengan 2 SD.

Dalam rumus $z = \frac{x}{SD}$ maka bilamana dua dari sukunya telah diketahui, suku yang satu lagi dapat dicari. Dari soal di atas x telas diketahui = 20 (dari x = X-M = 70-50 = 20). Karena SD-nya juga diketahui = 10, maka

$$z = \frac{20}{10} = 2$$

Sebaliknya jika z diketahui dan SD-nya juga diketahui, maka x akan dengan mudah dapat dicari. Dari contoh di atas jika z = +2 dan SD = 10, maka

$$z = \frac{x}{SD} \longrightarrow z(SD) = x$$

$$+2(10) = x$$

$$+20 = x$$

$$x = +20$$

Karena x menurut definisi adalah deviasi suatu nilai dari M, maka x = +20 berarti suatu angka yang deviasinya = 20 point di atas mean. Selanjutnya apabila mean telah diketahui, seperti misalnya dalam contoh diatas M=50, maka angka kasarnya akan dapat diketemukan

$$x = X - M \longrightarrow x + M = X$$

$$+20 + 50 = X$$

$$X = 70$$

Pengukuran dengan z-score kita mempunyai fungsi-fungsi tertentu. Ia menjadi sumber dari apa yang disebut *weighted score* atau *scale score* yang selalu digunakan dalam proses penilaian (secara ilmiah) hasil-hasil test.

Dengan z-score kita akan dimungkinkan untuk membandingkan kecakapan seorang anak dalam bermacam-macam pelajaran. Kalau misalnya seorang pelajar P mendapat nilai 80 dalam berhitung dan nilai 50 dalam ilmu bumi, kita biasanya segera menyimpulkan bahwa kecakapan pelajar itu dalam berhitung jauh lebih baik daripada kecakapannya dalam ilmu bumi. Benarkah demikian.

Jawabnnya adalah benar jika kita menganggap misalnya nilai 60 sebagai indeks rata-rata kecakapan para pelajar dalam semua mata pelajaran. Tetapi suatu soal segera timbul, benarkah anggapan kita bahwa 60 adalah nilai rata-rata?

Untuk menjawab persoalan itu kita dapat menggunakan suatu approach, suatu pendekatan yang menggunakan *kedudukan* subyek dalam kelompok distribusinya sebagai landasan. Kalau misalnya angka rata-rata dalam berhitung terdapat 90 dan SD-nya = 10, maka kedudukan angkat rata-rata dalam berhitung adalah 1 SD di bawah mean atau -1SD. Dan bilamana M dari ilmu bumi adalah 2 SD di atas mean atau +2SD. Ditinjau dari segi itu maka justru kecakapan pelajar P dalam ilmu bumi jauh lebih baik daripada kecakapan dalam berhitung.

Seorang guru mungkin sekali sangat berkeberatan terhadap interpretasi di atas. Sebab menurut “pikiran sehat” tak dapat dimengerti bagaimana angka 50 dapat lebih baik dari angka 80.

Keberatan itu dapat diterima sekiranya kita melihat dari angka-angka sebagaiman kelihatannya. Tetapi kalau kita dapat meninjau sedikit lebih mendalam dalam system penilaian akan kita rasakan bahwa bagi guru yang “mahal” yang hanya memberi nilai 70 bagi anak yang terpandai dalam ilmu bumi, sedang bagi guru yang “murah” mau memberikan angka 100, maka sebenarnya angka 70 itu ekuivalen dengan angka 100. Karena itu kedua angka itu tak dapat dibandingkan menurut harga kelihatannya, *not matchable in terms of their face value.*

Dengan meninggalkan perbedaan tinjauan dan interpretasi, karena hal itu lebih merupakan soal pengukuran dan penilaian daripada soal statistik, akan kita lanjutkan studi kita tentang statistik menyediakan *facts and figures*. Interprerasi dari fakta-fakta dan angka-angka itu terserah kepada bidang f yang menggunakan statistik itu.

POKOK BAHASAN:

"PROBABILITAS"

Jika suatu mata uang logam yang masih baik dilemparkan secara bebas satu kali, tiap-tiap permukaannya (kita sebut saja kepala atau K dan ekor atau E) mempunyai kemungkinan muncul yang sama, yaitu satu banding satu. Dengan begitu masing-masing permukaannya, baik K ataupun E, mempunyai kemungkinan muncul separo atau 50% dari seluruh kemungkinan yang ada. Dikatakan, probabilitas munculnya K atau E adalah 1 : 2, ½ . 0,5, atau 50%.

Bilamana sebuah dadu yang bermata enam macam dilemparkan secara bebas satu kali, dan buah dadu itu masih baik, maka kemungkinan untuk memperoleh mata angka empat adalah satu diantara enam kemungkinan. Dikatakan, probabilitas untuk memperoleh angka mata empat dari sebuah dadu yang masih baik yang dilemparkan secara bebas adalah 1 : 6, atau 0,1167 , atau 15,67%.

Masalah probabilitas adalah masalah frekuensi sesuatu kejadian. Probabilitas munculnya suatu gejala atau kejadian biasa disingkat dengan huruf p dan dinyatakan bahwa sesuatu kejadian mempunyai p = 0,05 atau 5%, ini berarti bahwa kejadian itu mempunyai kemungkinan (probabilitas) muncul 5 kali diantara 100 kejadian, 1 kali diantara 200 kejadian, atau 50 kali diantara 1000 kejadian. Dalam hal mata uang tersebut diatas, p dari K adalah 0,50 atau 50%. Demikian juga p dari E adalah 0,50 atau 50%. Dalam hal buah dadu, p dari timbulnya mata angka empat adalah 0,1667 atau 16,67%.

Memahami angka-angka itu tidak akan menimbulkan sesuatu kesulitan, sebab semuanya itu didasarkan atas kenyataan-kenyataan yang paling sederhana, kita jumpai tiap-tiap hari, dan dapat kita buktikan tiap-tiap waktu. Dari angka-angka yang menunjukkan probabilitas suatu kejadian seperti dalam contoh diatas itu kita dapat mengharapkan bahwa jika suatu mata uang logam yang masih baik atau unbiased dilemparkan secara bebas 100 kali, probabilitas munculnya K adalah 50% x 100 kali = 50 kali. Jika kita lanjutkan melemparkannya 300 kali, probabilitas kita akan menyaksikan K adalah 150 kali. Demikian juga kalau ada dadu dilemparkan 100 kali, kemungkinan kita akan memperoleh mata angka empat adalah 16,67 kali atau 16 a 17 kali, kalau dilemparkan 500 kali, probabilitas timbulnya mata angka empat adalah 83 a 84 kali.

Dari itu, probabilitas suatu kejadian dapat dibatasi sebagai perbandingan *frekuensi kejadian itu dengan kejadian seluruhnya.*

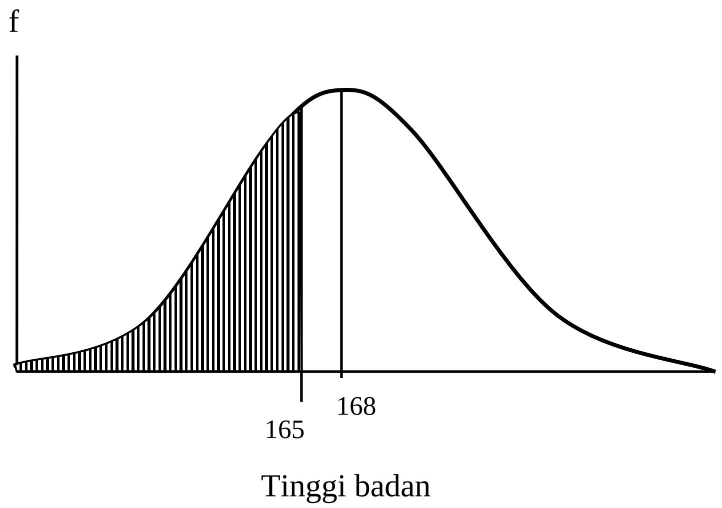
Untuk melihat bagaimana hubungan antara probabilitas dengan kurve normal adalah distribusi teoritik dari frekuensi suatu kejadian. Ini terutama dikembangkan dalam hubungan dengan perhitungan probabilitas secara matematik, dan biasa disebut kurve normal dari probabilitas. Kurve normal yang berbentuk genta itu menunjukkan bahwa makin besar deviasi sesuatu kejadian dari mean dari semua kejadian, maka kurang frekeunsi kejadian itu. Dengan kata lain makin meningkat deviasi dari mean, makin menurun probabilitasnya. Dalam contoh soal nomor 4 yang mempermasalahkan berapa tinggi loncatan yang terjadi hanya 5% atau kurang dalam distribusi loncatan yang dipersoalkan itu, sebenarnya kita menanyakan tentang tinggi loncatan yang mempunyai probabilitas 5 diantara 100, atau p = 0,05 atau 5%. Frekuensi dalam persen dari sesuatu kejadian yang kita peroleh dari tabel kurve normal adalah juga probabilitas dari suatu kejadian yang dimaksudkan.

Dengan sekedar pengertian itu mari kita pecahkan pertanyaan nomor 8 : Bagaimana probabilitas seseorang yang diambil secara sembarang dari sebuah kelompok itu dapat meloncat setinggi 190 cm atau lebih?

Langkah pertama yang harus kita tempuh adalah mencari jawaban terhadap persoalan berapa SD penyimpangan angka 190 itu. Perhitungan sebagaimana biasa memberikan hasil kepada kita suatu deviasi sebesar +2,31 SD, diperoleh dari 190 – 160 / 13. Dengan memeriksa tabel kurve normal kita dapat mengetahui ada 48,96% frekuensi loncatan antara m dan +2,31. Ini berarti ada 1,04% diperoleh dari 50,00% sampai 48,96% . Jadi, probabilitas seseorang yang dapat meloncat 190 cm adalah 1,04% atau 0,0104.

Akhirnya baiklah kita coba memecahkan persoalan ini untuk sekedar lebih memperdalam pemahaman kita tentang apa yang baru dibicarakan itu. Pesawat tempur modern adalah sedemikian penuhnya dengan alat-alat perlengkapan sehingga hanya orang-oorang yang tinggi badannya tidak lebih dari 165 cm saja yang dapat diterima dalam pasukan pesawat tempur. Pertanyaan : atau mean tinggi badan orang ada 168 cm dan SD nya ada 8 cm, sedanng dari penyelidikan- penyelidikan selalu ditunjukkan bahwa tinggi badan orang selalu mengikuti distribusi normal, berapakah probabilitas dari tiap-tiap orang pelamar akan ditolak lamarannya atas dasar tinggi badan.

Untuk menjawab pertanyaan itu baiklah kita perhatikan grafik probabilitas dibawah. Adapun perhitungannya adalah sebagai berikut. Dengan M = 168 cm dan SD = 8 cm, maka angka 165 menyimpang -0,375 SD dari M. Ini meliputi daerah seluas kira-kira 14,50%. Jadi persentase orang yang tingginya kurang dari 165cm ada 50,00% sampai 14,50% = 35,50%. Ini berarti bahwa probabilitas dari pada orang yang diterima ada 35,50% atau 0,3550 sehingga probabilitas orang yang ditolak ada 64,50%, diperoleh dari 100% - 35,50%.



Grafik 22

Bilangan-bilangan yang dikemukakan dalam contoh keluarnya kepala dan ekor dari sebuah mata uang yang dilemparkan secara bebas, maupun bilangan-bilangan yang dikemukakan sebagai jawaban pertanyaan nomor 8 dan soal tentang kemungkinan pelamar yang ditolak dan diterima berdasarkan tinggi badannya seperti yang baru dibicarakan, semua adalah bilangan-bilangan harapan, kemungkinan-kemungkinan hipotetik atau probabilitas teoritik. Kita menanyakan berapa besarnya kepercayaan yang dapat kita katakan pada bilangan-bilangan itu seandainya kita betul-betul mengadakan observasi terhadap keajadian-kejadian yang disebutkan itu.

Pertanyaan semacam itu sungguh merupakan pertanyaan yang cerdas yang tidak mungkin diberikan jawabannya dengan satu kata. Hanya dengan analisa secara cerdas pula pertanyaan itu dapat dicarikan jawabannya. Marilah kita coba!

6.1 HUBUNGAN ANTARA PROBABILITAS TEORITIK DENGAN PROBABILITAS EMPIRIK

Baiklah kita ulangi contoh tentang mata uang yang dilemparkan dan mulai kita berikan batasan-batasan secara lebih tegas dengan bahasa-bahasa yang lebih sering digunakan dalam statistik.

Dari sebuah mata uang logam yang masih baik dan dilemparkan secara bebas kemungkinannya bahwa kkita akan memperoleh kepala K atau ekor E. kemungkinan timbul atau tidak timbulnya suatu kejadian seperti K atau ekor E itu disebut probabilitas kejadian. Kemungkinan timbul disebut probabilittas sukses, sedang probabilitas tidak timbul disebut probabilitas gagal.

Jika kemungkinan sukses kita beri simbol q, dan kemungkinan timbulnya p dan q adalah sama, maka kita batasi

$$P = q = 0,50 \qquad (34)$$

Karena probabilitas selalu dihitung dari seluruh kejadian maka dari seluruh kejadian yang mungkin batasan tersebut dapat juga ditanyakan

$$Prs = p = 1 - q = Pr_G = q = 1 - p \qquad (35)$$

Dalam nama Pr_S = probabilitas sukses

Pr_G = probabilitas gagal

Jika mata uang tersebut dilemparkan 10 kali, maka menurut dasar teori probabilitas keluarnya K akan sebanyak 0,50 x 100 = 5 kali dari keluarnya E juga kan sebanyak 5,00 x 10 = 5 kali. Akan tetapi, dalam kenyataannya walaupun mata uang tersebut masih baik dan cara melemparkanya betul-betul bebas, jarang sekali dari 10 lemparan itu kita akan memperoleh 5 K dan 5 E. Umumnya ada faktor-faktor kebetulan diluar kekuasaan tangan manusia yang merubah perbandingan menurut probabilitas teoritik itu, sehingga dalam kenyataannya perbandingan antara K dan E itu mejadi 4 : 6,7 : 3 dan sebagainya. Probabilitas yang diobservasi yang dalam literature statistik berbahasa Inggris disebut *observed probability*, biasanya dinyatakan dalam bilangan pecahan, seperti 0,40, 0,60, 0,70, 0,30, dsb dengan jumlah seluruh probabilitas = 1,00.

Jika frekuensi observasi kita tambah terus menerus, misalnya mata uang tersebut tidak kita lemparkan 10 kali, maka suatu fenomim yang khas akan kita jumpai, yaitu bahwa perbedaan antara probabilitas teoritik denngan *observed probability* akan menjadi semakin kecil. Jadi jika misalnya kita lemparkan mata uang 100 kali dan keluar 57 kali K, dan kita lemparkan lagi mata uang tersebut 100 kali dan keluar 45 K maka probabilitas keluarnya K dari 200 kali lemparan bebas itu menjadi(57/100 + 45/100) : 2 = (0,57 + 0,45) : 2 = 0,51

Apa yang dimaksud dengan probabilitas empirik dari suatu kejadian tidak lain adalah probabilitas timbulnya kejadian itu dari sejumlah besar observasi. Dan jika observasi dilakukan tak terhingga, maka secara praktis dapat dikatakan bahwa probabilitas empiric akan sangat dekat atau sama dengan probabilitas teoritik. Jadi, misalnya jika kita lakuakan terus menerus melemparkan mata uang dan kita observasi keluarnya K, maka probabilitas dari K akan sangat mendekati 0,50 yaitu probabilitas teoritik dari satu kali lemparan mata uang tersebut.

Contoh tentang lemparan mata uang, suatu contoh klasik yang selalu digunakan dalam statistic untuk menerangkan konsep tentang probabilitas, seperti tersebut diatas dengan tidak dapat juga dikenakan pada kejadian kelahiran anak pria dan kelahiran anak wanita. Jika probabilitas kelahiran anak wanita, W sama dengan probabilitas kelahiran anak pria P maka kita dapat membatasi .

$Pr_W = p = Pr_P = q = 0,50$

Dari sejumlah observasi, misalnya dari 10 atau 20 kali observasi terhadap kelahiran, maka jarang sekali kita menjumpai perbandingan probabilitas kelahiran 2 jenis kelamin itu 0,50 : 0,50, melainkan 0,60 :0,40, 0,35 : 0,65 atau mungkin malahan 0,30 : 0,70. Akan tetapi jumlah observasi ditambah terus menerus maka perbandingan probabilitas antara W banding P akan makin mendekati 0,50 : 0,50 dan sudah pasti jika kita dapat mengobservasi kelahiran anak wanita dan pria, perbandingan probabiliitas antara W dan P akan menunjukkan 0,50 : 0,50 seandainya peluang kelahiran wanita dan pria adalah sama.

6.2 DISTRIBUSI PROBABILITAS GEJALA DISKRIT

Contoh- contoh tentang lemparan uang dan kelahiran tersebut diatas adalah contoh-contoh yang baik tentang probabilitas gejala diskrit. Marilah sekarang kita perluas contoh-contoh itu.

Gambarkanlah bukan sebuah mata uang tetapi dua buah mata uang logam, yang dua-duanya masih baik yang dillemparkan secara bebas bersama-sama. Kemungkinan gejala yang akan kita hadapi adalah :

1. Dari mata uang pertama keluar K dan dari mata uang yang kedua juga keluar K, atau dinyatakan dengan simbol KK.
2. Dari mata uang keluar K dan dari mata uang yang kedua keluar E dinyatakan dengan simbol KE.
3. Dari mata uang pertama keluar E dan dari mata uang yang kedua keluar K atau dinyatakan dengan simbol EK.
4. Dari mata uang pertama keluar E dan dari mata uang yang kedua juga E maka dinyatakan dengan simbol EE.

Perbandingan keluarnya KK, KE, EK dan EE. Adalah 1 : 1 : 1 : 1. Atau dinyatakan dalam bilangan probabilitas 0,25 : 0,25 : 0,25 : 0,25, dengan jumlah seluruh probabilitas = 1,00.

Oleh karena kemungkinan kedua dan ketiga yaitu keluarnya KE dan EK, tidak lain adalah keluarnya 1K dan 1E dari 2 buah mata uang yang dilemparkan bersama-sama maka kombinasi dari KE dan EK itu dalam perhitungan probabilitas dapat dikumpulkan, sehingga perbandingan probabilitasnya akan menjadi :

2K = 0,25
1K 1E = 0,25 + 0,25 = 0,50
2E = 0,25

Jika kita tambahkan lagi sebutir mata uang yang kita lemparkan, maka gejala yang kita dapati adalah KKK, EKK, KEK, EEK, EKE, KEE,dan EEE dengan perbandingan 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8. Dan jika kombinasi-kombinasi yang sama kita kumpulkan, yaitu EKK, KEK,dan KKE menjadi satu kategori lagi, kemudian probabilitas keluarnya 3K kita beri simbol p1, probabilitas keluarnya 2K1E kita beri simbol p2, probabilitas keluarnya 1K2E kita beri simbol p3 dan probabilitas keluarnya kita beri simbol p4, maka kita akan menjumpai distribusi probabilitas sebagai berikut.

Tabel 33

Gejala (G)		Probabilitas (p)	
G1	KKK	3K	p1 = 1/8 = 0,125
G2	EKK, KEK, KKE	2K1E	p2 = 1/8 + 1/8 + 1/8 = 3/8 = 0,372
G3	EEK, EKE, KEE	1K2E	
G4	EEE	3E	

Gambarkan juga sekarang bukan keluarga dengan seorang anak, tetapi keluarga dengan dua orang anak. Apabila peluang kelahiran antara anak wanita dan anak pria adalah sama, maka kemungkinan dari suatu keluarga yang mempunyai dua orang anak adalah :

1. Anak pertama wanita, anak kedua juga wanita (WW)
2. Anak pertama wanita, tetapi anak kedua pria (WP)
3. Anak pertama pria, tetapi anak kedua wanita (PW)
4. Anak pertama pria, anak kedua juga pria (PP)

dengan perbandingan WW : WP : PW : PP = 1 : 1 : 1 : 1, atau dinyatakan dalam bilangan-bilangan probabilitas 0,25 : 0,25 : 0,25 : 0,25. Jumlah seluruhnya probabilitas adalah 1,00

Jika kombinasi WP dan PW kita satukan dan kita berikan simbol-simbol masing-masing p1, p2, dan p3 untuk WW, WP dan WP, dan PP, maka perbandingan probabilitasnya akan menjadi :

WW atau 2W	p1 = 025
WP dan PW atau 1W1P	p2 = 0,25 + 0,25 = 0,50
PP atau 2P	p3 = 0,25

Bagi keluarga dengan tiga orang anak maka kemungkinan kombinasinya akan WWW, PWW, WPW, WWP, PPW, PWP, WPP, dan PPP dengan probabilitas 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8 : 1/8. Dan setelah kombinasi-kombinasi yang sama dikumpulkan, maka perbandingan probabilitasnya akan menjadi :

G1 3W p1 = 0,125

G2 3W1P p2 = 0,375

G3 1W2P p3 = 0,375

G4 3P p4 = 0,125

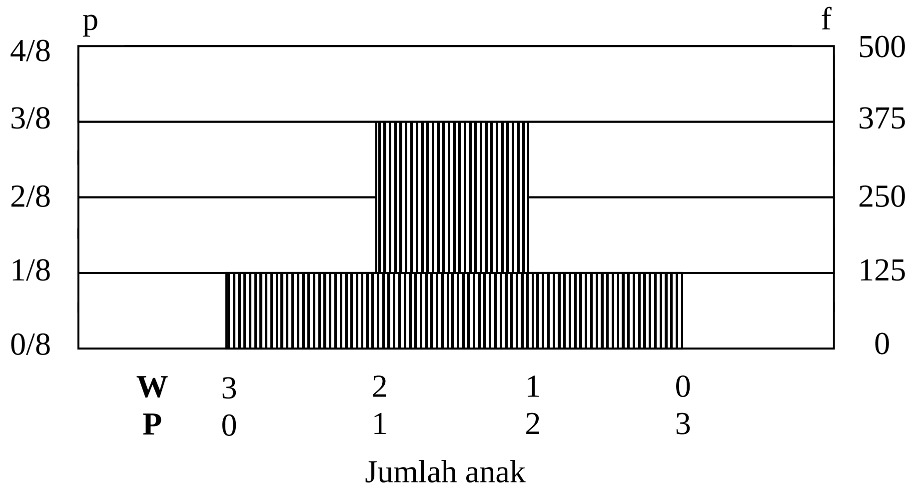
Total probabilitasnya = P = 1,00

Dari distribusi probabilitas kita dapat segera memastikan bagaimana keadaan distribusi frekuensi gejala yang diharapkan dari sejumlah observasi tertentu. Misalnya jika kita ambil kembali contoh yang baru disebutkan maka apabila peluang kelahiran anak wanita W dan anak pria P adalah sama (jadi, P = q = 0,50), maka distribusi frekuensi dari 1000 keluarga yang mempunyai 3 orang anak akan seperti berikut :

Tabel 34

G	3P	2P1W	2W1P	3W	Total
p	1/8	3/8	3/8	1/8	8/8
Np	125	375	375	125	1000

Digambarkan dalam grafik



Grafik 23

DAFTAR PUSTAKA

1. Hadi, Sutrisno, ***“Metodologi Research Jilid 1”,*** Andi Jogjakarta, 2004
2. Hadi, Sutrisno, ***“Metodologi Research Jilid 2”,*** Andi Jogjakarta, 2004
3. Hadi, Sutrisno, ***“Statistik Jilid 1”,*** Andi Jogjakarta, 2004
4. Hadi, Sutrisno, ***“Statistik Jilid 2”,*** Andi Jogjakarta, 2004
5. Harinaldi, ” ***Statistik untuk Teknik dan Sains***” Erlangga, 2005.
6. Murray R. Spieger dan Larry J. Stephens, ***“Statistik”*** Edisi ke-3, Erlangga, 2004.