BAB 4

Pemusatan dan Penyebaran Data

A. Mengukur Pemusatan Data

Rumus yang digunakan untuk mengukur pemusatan data selalu dibedakan untuk data yang tidak dikelompokkan dan data yang dikelompokkan.

1. Rerata (mean)

Rerata merupakan konsep secara awam mengenai rata-rata. Merupakan titik berat dari seperangkat data atau observasi sensitif terhadap nilai ekstrim. Digunakan terutama bila teknik statistik lain, seperti pengujian hipotesis akan dilakukan pada data.

a. Untuk data yang tidak dikelompokkan (data tunggal)

$$\bar{x} = \frac{\sum x}{n}$$

$$\mu = \frac{\sum x}{N}$$

dimana :

$\bar{x}$ = rerata

$\sum$ = huruf besar Yunani sigma, yang berarti jumlahkan

x = nilai suatu hasil pengamatan atau observasi

$\sum x$ = jumlahkan semua observasi

n = jumlah semua observasi

b. Untuk data yang dikelompokkan

$$\bar{\boldsymbol{x}} = \frac{\sum \boldsymbol{f}.\boldsymbol{xi}}{\sum \boldsymbol{f}}$$

dimana :

$\bar{x}$ = titik tengah (*mid point*) kelas interval ke I

x_i = titik tengah interval kelas

f = frekwensi observasi pada kelas interval ke i

fx = jumlahkan frekwensi tiap kelas interval

Contoh :

Data tinggi badan mahasiswa FKIP UMB- Yogyakarta diambil 50 mahasiswa secara random :

Tabel 1. Hasil Pengukuran tinggi badan

Interval Kelas	f_i
164,5 – 167,5	6
167,5 – 170,5	7
170,5 – 173,5	8
173,5 – 176,5	11
176,5 – 179,5	7
179,5 – 182,5	6
182,5 – 185,5	5
Jumlah	**50**

Jawab :

Interval Kelas	F	Xi	f*xi
164,5 – 167,5	6	166	996
167,5 – 170,5	7	169	1183
170,5 – 173,5	8	172	1376
173,5 – 176,5	11	175	1925
176,5 – 179,5	7	178	1246
179,5 – 182,5	6	181	1086
182,5 – 185,5	5	184	920
Jumlah	**50**		8732

Maka

$$\bar{x} = \frac{\sum f.x_i}{\sum f} = \frac{8732}{50} = 174,64$$

2. Median

Median merupakan nilai tengah dari sekelompok data yang nilai tiap observasi telah disusun dari yang terkecil ke terbesar. Tidak sensitif terhadap nilai ekstrim. Median digunakan untuk mengukur pemusatan kalau distribusi mencong (*skewed*) secara jelas. Dapat dihitung pada distribusi yang tidak komplit sekalipun, misalnya distribusi yang berakhir terbuka (contoh 150-169 ; 170-189; 190-209; 210+).

a. Untuk data yang tidak dikelompokkan

1) Bila jumlah observasi (=n) ganjil, maka median adalah nilai observasi ke :$\frac{n+1}{2}$ dari urutan nilai observasi kecil ke besar.

   Contoh : 5, 4, 5, 6, 7, 1, 5, 3, 4, 6, 9. Tentukan median

   Urutkan data : 1, 3, 4, 4, 5, 5, 5, 6, 6, 7, 9

   Median $(M_e) = \frac{n+1}{2} = \frac{11+1}{2} = 6$

2) Bila banyaknya observasi (=n) genap, maka median adalah nilai di antara observasi ke : $\frac{n}{2}$ dan $\frac{n}{2}+1$, diambil rata-rata.

   Contoh :

   1, 3, 4, 4, 5, 5, 5, 6, 6, 7

   $\boldsymbol{n = 10 \rightarrow \frac{n}{2} = 5\ dan\ 5 + 1 = 6}$

   $$\boldsymbol{Me = \frac{5+6}{2} = 5,5}$$

b. Untuk data yang dikelompokkan

$$\boldsymbol{Me = Lm + \frac{w\left(\frac{n}{2} - cf\right)}{f_m}}$$

dimana :

Me = median

lm = batas bawah dari kelas interval dimana median berada (kelas median)

n = banyaknya observasi

cf = frekwensi kumulatif dari kelas interval sebelum kelas median

w = lebar kelas interval dimana median berada contoh :

Tentukan median dari data kelompok dibawah ini

Jawab :

Interval Kelas	f_i
164,5 – 167,5	6
168,5 – 171,5	7
172,5 – 175,5	8
176,5 – 179,5	11
180,5 – 183,5	7
184,5 – 187,5	6
188,5 – 191,5	5
Jumlah	**50**

cf

fm

Jawab : (sebagai latihan mahasiswa)

Menentukan kelas median $=\frac{n}{2} = \frac{50}{2} = 25$

$$\boldsymbol{Me = Lm + \frac{w\left(\frac{n}{2} - cf\right)}{f_m} = 176 + \frac{7(25-21)}{11}}$$

$$\boldsymbol{= 176 + \frac{28}{11} = 178,54}$$

3. Modus (Mode)

Modus merupakan nilai yang paling sering muncul (frekuensi terbesar) dari seperangkat data atau observasi. Mencerminkan yang paling tipikal atau kasus yang paling umum. Kalau kita ingin segera mengetahui nilai pemusatan, maka kita menghitung modus. Seperangkat data dapat saja tidak memiliki modus, tetapi sebaliknya dapat pula memiliki beberapa modus. Kalau satu modus saja disebut unimodal, dua modus disebut bimodal dan kalau tanpa modus disebut nonmodal.

a. Untuk data yang tidak dikelompokkan

Modus (*crude mode*) = nilai yang paling sering muncul

Contoh : 1, 3, 4, 4, 5, 5, 5, 6, 6, 7
$M_0 = 5$

b. Untuk data yang dikelompokkan

Modus = titik tengah dari kelas interval yang memiliki frekwensi terbesar.

$$Mo = Bb + w\left(\frac{d_1}{d_1 + d_2}\right)$$

Interval Kelas	f_i
164,5 – 167,5	6
168,5 – 171,5	7
172,5 – 175,5	8
176,5 – 179,5	11
180,5 – 183,5	7
184,5 – 187,5	6
188,5 – 191,5	5
Jumlah	**50**

$$\boldsymbol{M_0 = 176 + 7\frac{11-8}{(11-8)+(11-7)}}$$

$$\boldsymbol{= 176 + 7\left(\frac{3}{7}\right)}$$

$$\boldsymbol{= 176 + 3}$$

$$\boldsymbol{= 179}$$

CONTOH :

1. Untuk data yang tidak dikelompokkan

Berikut ini data mengenai lama perawatan sepuluh penderita yang dirawat di bangsal perawatan Psikiatri dari suatu rumah sakit :

Pasien ke	Lama perawatan (hari)	Pasien ke	Lama perawatan (hari)
1	29	6	14
2	14	7	28
3	11	8	14
4	24	9	18
5	14	10	22

Hitung : rerata, median, modul lama perawatan dari pasien-pasien ini !

1. Rata-rata

$$\bar{x} = \frac{\sum x}{n} = \frac{11+14+14....+24+28+29}{10}$$

$$\bar{x} = \frac{188}{10} = 18.8 \text{ hari}$$

2. **Median**

Urutan nilai observasi adalah sebagai berikut :

11; 14; 14; 14; 14; 18; 22; 24; 28; 29

Karena banyaknya observasi genap, maka median merupakan rata-rata nilai dari observasi ke $\frac{n}{2} = \frac{10}{2} = 5$ dan $\frac{n}{2} + 1 = 6$

Jadi :

Median = $\frac{14 + 18}{2} = 16$ hari

3. **Modus**

Oleh karena 14 hari adalah nilai yang paling sering muncul, maka modus adalah 14 hari

2. Untuk data yang dikelompokkan

Dari sejumlah penderita *typhus abdominalis* yang dirawat di bangsal penyakit menular suatu Rumah Sakit, diperoleh data sebagai berikut :

Masa inkubasi (hari) dari 170 penderita *typhus abdominalis*

Masa inkubasi (hari)	Jumlah penderita
2	25
6	80
10	30
14	15
18	12
22	6
24*	2
	total = 170

* tidak ada pasien dengan masa inkubasi 30 hari atau lebih.

Hitung : rerata, median dan modus.

Masa inkubasi (hari)	Banyaknya pasien (f)	Titik tengah (x)	fx	fx^2	Frekuensi kumulatif (cf)
2- 5	25	4	100	400	25
6 -9	80	8	640	5120	105
10 - 13	30	12	360	4320	135
14 - 17	15	16	240	3840	150
18 - 21	12	20	240	4800	162
22 -25	6	24	144	3456	168
26 - 29	2	28	56	1568	170
	Total = 170		fx = 1780	23504	

1. **Rerata**

$$\bar{x} = \frac{fx}{n} = \frac{1780}{170} = 10{,}47 \text{ hari}$$

2. **Median**

$$Md = L_m + \frac{\frac{n}{2} - cf}{fm} . w$$

$n/2 = 170/2 = 85$, kelas interval dimana median berada (kelas median) adalah: 6, maka

lm = 6

cf kelas interval sebelumnya = 25

fm = 80

w = 10 - 6 = 4

$$Md = 6 + \frac{170/2 - 25}{80}$$

$$Md = 6 + \frac{60}{80} 4$$

$$Md = 6 + 3 = 9$$

3. **Modus**

Mo = 8, oleh karena frekuensi tertinggi dimiliki kelas interval 6 - dan titik tengah kelas interval ini adalah : 8.

Latihan :

Berdasarkan data pada contoh kasus 1. Tentukan nilai mean, median, modus

NILAI	SCORE	BANYAKNYA DATA (FREKUENSI)
40 – 49	IIII	4
50 – 59	IIIII I	6
60 – 69	IIIII IIIII	10
70 – 79	IIII	4
80 – 89	IIII	4
90 – 99	II	2
$\sum$		30

B. Ukuran Letak

Agar kita dapat mengetahui lebih jauh mengenai karakteristik data observasi dengan beberapa ukuran sentral, kita sebaiknya mengetahui beberapa ukuran lain, yaitu ukuran letak. Ada tiga macam ukuran letak yang akan di bahas pada bagian ini, yaitu Kuartil, Desil, dan Persentil.

1. *Kuartil*

Kuartil adalah ukuran letak yang membagi data observasi menjadi empat bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 25% data observasi. Pada satu set data observasi mempunyai tiga buah kuartil, yaitu K_1, K_2, K_3.

Untuk menentukan nilai kuartil data observasi yang tidak berkelompok (*ungrouped data*) melalui langkah-langkah sebagai berikut ini :

1) Urutkan data observasi dari kecil ke besar
2) Tentukan letak kuartilnya

   Menentukan letak K_1, K_2, K_3 dapat digunakan formulasi sebagai berikut :

   Letak $K_1 = \frac{N+1}{4}$

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

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

3) Tentukan nilai kuartilnya.

   Nilai K_1, K_2, K_3 adalah data observasi yang terletak pada letak K_1, K_2, K_3

Contoh kasus :

Berikut ini adalah data mengenai nilai 7 orang peserta ujian Statistik di UMB Yogyakarta :

78 56 66 48 80 70 76

Tentukan K_1, K_2, K_3

Jawab :

Untuk menentukan K_1, K_2, K_3, maka langkah-langkah yang digunakan adalah sebagai berikut :

- Urutkan nilai tersebut dari kecil ke besar

48 56 66 70 76 78 80

- Tentukan letak K_1, K_2, K_3 dengan formula

  Letak $K_1 = \frac{7+1}{4} = 2$

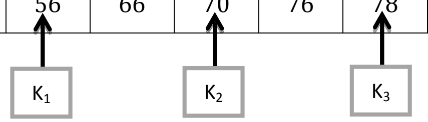
  Letak $K_2 = \frac{2(7+1)}{4} = 4$

  Letak $K_3 = \frac{3(7+1)}{4} = 6$

  Jadi letak K_1 pada urutan data ke 2, letak K_2 pada urutan data ke 4, dan letak K_3 pada urutan data ke 6

- Tentukan nilai K_1, K_2, K_3

No urut	1	2	3	4	5	6	7
nilai	48	56	66	70	76	78	80



  Nilai K_2 adalah juga merupakan median dari nilai peserta ujian tersebut. Apabila banyaknya data observasi menunjukkan bilangan genap, maka median terletak diantara dua nomor urut.

Kuartil (K_1, K_2, K_3) data observasi berkelompok dapat ditentukan dengan langkah-langkah sebagai berikut ini :

1. Tentukan kelas K_1, K_2, K_3 dengan formula :

   Kelas kuartil $1(K_1)$:

$$K_1 = \frac{N}{4}$$

   Kelas kuartil 2 (K_2) :

$$K_2 = \frac{2N}{4}$$

Kelas kuartil 3 (K_3) :

$$K_3 = \frac{3N}{4}$$

2. Tentukan K_1, K_2, K_3 dengan menggunakan formula

$$K_1 = B_{K_1} + \left(\frac{\frac{N}{4} - Cf_1}{f_{K_1}} \right) . C_i$$

Yang menyatakan bahwa :

K_1= Kuartil 1

B_{K_1}= tepi kelas bawah kelas kuartil 1

N = banyaknya data observasi ($\sum f$)

Cf_1= frekuensi kumulatif kelas sebelum kelas kuartil 1

f_{K_1}= frekuensi kumulatif kelas kuartil 1

C_i= interval kelas

$$K_2 = B_{K_2} + \left(\frac{\frac{2N}{4} - Cf_2}{f_{K_2}} \right) . C_i$$

Yang menyatakan bahwa :

K_2= Kuartil 2

B_{K_2}= tepi kelas bawah kelas kuartil 2

N = banyaknya data observasi ($\sum f$)

Cf_2= frekuensi kumulatif kelas sebelum kelas kuartil 2

f_{K_2}= frekuensi kumulatif kelas kuartil 2

C_i= interval kelas

K_2 nilainya sama dengan nilai median

$$K_3 = B_{K_3} + \left(\frac{\frac{3N}{4} - Cf_3}{f_{K_3}} \right) . C_i$$

Yang menyatakan bahwa :

K_3= Kuartil 3

B_{K_3}= tepi kelas bawah kelas kuartil 3

N = banyaknya data observasi ($\sum f$)

Cf_3= frekuensi kumulatif kelas sebelum kelas kuartil 3

f_{K_3}= frekuensi kumulatif kelas kuartil 3

C_i= interval kelas

Contoh kasus :

Tentukan K_1, K_2 dan K_3 nilai 30 peserta ujian statistik seperti yang tampak pada tabel 3.1

NILAI	FREKUENSI	TEPI KELAS	FREKUENSI KUMULATIF
40 – 49	4	39,5	4
50 – 59	6	49,5	10
60 – 69	10	59,5	20
70 – 79	4	69,5	24
80 – 89	4	79,5	28
90 – 99	2	89,5	30
$\sum$	30		

2. *Desil*

Desil adalah ukuran letak yang membagi data observasi menjadi sepuluh bagian yang sama banyak. Oleh karena itu masing-masing bagian mengandung 10% data observasi. Pada satu set data observasi mempunyai sembilan buah desil, yaitu $D_1, D_2, \dots, D_9$.

Untuk data tunggal, jika banyak data n dan D_i adalah desil ke-i, maka

Letak D_i = data ke$\frac{i(n+1)}{10}$ dengan i = 1,2,3,4,...,9

Contoh;

Tentukan D_3, dan D_5 dari ; 6, 4, 6, 4, 7, 5, 6, 5, 8, 7, 7, 7, 8, 6 !

Penyelesaian;

Data diurutkan menjadi ; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 7, 8, 8

Data	4	4	5	5	6	6	6	6	7	7	7	7	8	8
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

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

Letak $D_3 =$ data ke- $\frac{3(14+1)}{10}$

= data ke- 4,5 $(X_{4,5})$

Dengan interpolasi diperoleh :

$D_3 = X_4 + 0{,}5(X_5 - X_4)$

$= 5 + 0{,}5(6 - 5)$

$= 5,5$

Letak D_5= data ke- $\frac{5(14+1)}{10}$

= data ke- 7,5 $(X_{7,5})$

Dengan interpolasi diperoleh :

$$D_5 = X_7 + 0,5(X_8 - X_7)$$
$$= 6 + 0,5(6 - 6)$$
$$= 6$$

Desil data berkelompok dapat dihitung dengan rumus :

$$D_i = T_b + p\left(\frac{\frac{i}{10}n - F}{f}\right)$$

Dimana i= 1,2,3,4,.....,9

Dengan D_i = desil ke-i

Tb = tepi bawah interval kelas D_i
P = panjang kelas interval D_i
n = $\sum f$ = banyak data
F = frekuensi kumulatif sebelum kelas D_i
f = frekuensi pada kelas D_i

Contoh.

Hitung nilai D_5 dan D_8 dari data berdistribusi kelompok berikut :

Interval	F	F_k
21-25	3	3
26-30	9	12
31-35	4	16
36-40	10	26
41-45	3	29
46-50	11	40

Penyelesaian ;

Desil ke-5 terletak pada$\left(\frac{5}{10}.40\right) = 20$ (kelas interval 36-40)

$$D_5 = 35,5 + \frac{5(20-16)}{10}$$
$$= 37,5$$

Desil ke-8 terletak pada$\left(\frac{8}{10}.40\right) = 32$ (kelas interval 46-50)

$$D_8 = 45{,}5 + \frac{5(32-29)}{11}$$
$$= 46{,}9$$

3. Persentil

Persentil adalah ukuran letak yang membagi data observasi menjadi seratus bagian yang sama besar. Oleh karena itu masing-masing bagian mengandung 1 % data observasi. Pada satu set data observasi mempunyai 99 persentil, yaitu : $P_1, P_2, \ldots, P_{99}$.

Persentil data tunggal maka :

Letak P_i = data ke- $\left(\frac{i(n+1)}{100}\right)$, dengan i = 1,2,3,……,99

Contoh;

Tentukan P_{30}, dan P_{75} dari ; 6, 4, 6, 4, 7, 5, 6, 5, 9, 7, 10, 7, 10, 6 !

Penyelesaian;

Data diurutkan menjadi ; 4, 4, 5, 5, 6, 6, 6, 6, 7, 7, 7, 9, 10, 10

Data	4	4	5	5	6	6	6	6	7	7	7	9	10	10
Data ke-	1	2	3	4	5	6	7	8	9	10	11	12	13	14

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

Letak P_{30} = data ke- $\frac{30(14+1)}{100}$

= data ke- 4,5 $(X_{4,5})$

Dengan interpolasi diperoleh :

$P_{30} = X_4 + 0{,}5(X_5 - X_4)$

$= 5 + 0{,}25(6-5)$

$= 5{,}25$

Letak P_{75}= data ke- $\frac{75(14+1)}{100}$

= data ke- 11, 25 $(X_{11,25})$

Dengan interpolasi diperoleh :

$$P_{75} = X_{11} + 0{,}5(X_{12} - X_{11})$$
$$= 7 + 0{,}25(9-7)$$
$$= 7{,}5$$

Persentil data berkelompok dapat dihitung dengan rumus :

$$P_i = T_b + p\left(\frac{\frac{i}{100}n - F}{f}\right)$$

Dimana i= 1,2,3,4,.....,99

Dengan P_i = persentil ke-i

Tb = tepi bawah interval kelas P_i

P = panjang kelas interval P_i

n = $\sum f$ = banyak data

F = frekuensi kumulatif sebelum kelas P_i

f = frekuensi pada kelas P_i

Contoh.

Hitung nilai P_{25}dari data berdistribusi kelompok berikut :

Interval	F	F_k
21-25	3	3
26-30	9	12
31-35	4	16
36-40	10	26
41-45	3	29
46-50	11	40

Penyelesaian ;

Persentil ke-25 terletak pada$\left(\frac{25}{100}.40\right) = 10$ (kelas interval 26-30)

$$P_{25} = 25{,}5 + \frac{5(10-3)}{9}$$
$$= 29{,}4$$

C. Pengukuran Penyebaran (Dispersi)

1. Pengertian Tentang Disperse.

Digunakan untuk menunjukkan keadaan berikut :

a. Gambaran variabilitas data

Yang dimaksud dengan variabilitas data adalah suatu ukuran yang menunjukkan besar kecilnya perbedaan data dari rata-ratanya. Ukuran ini dapat juga disebutkan sebagai ukuran yang menunjukkan perbedaan antara data satu dengan yang lainnya. Ukuran pemusatan (Mean, Median, dan Modus) ini dapat kita gunakan untuk menggambarkan keadaan sekumpulan data, tetapi gambaran itu masih kurang lengkap apabila tidak disertai dengan ukuran-ukuran penyebaran. Hal ini disebabkan karena dengan ukuran gejala pusat saja mungkin beberapa kumpulan data sebenarnya berbeda dapat disimpulkan sama.

b. Perbedaan nilai satu observasi terhadap nilai observasi lainnya

Rata-rata dari serangkaian nilai-nilai observasi tidak dapat diinterpretasikan secara terpisah dengan dispersi (sebaran) nilai-nilai tersebut terhadap rata-ratanya. Jika terdapat keseragaman/kesamaan nilai-nilai observasi, X_i, maka dispersi nilai-nilai tersebut akan sama dengan nol, dan rata-ratanya akan sama dengan nilai X_i. Semakin besar variasi nilai-nilai X_i, maka rata-rata distribusi semakin kurang representatif.

Contoh:

Tabel 7-1 Rata-rata hitung hasil test mata kuliah statistik deskriptif kelompok A dan B.

kelompok	hasil test					
A	60	65	50	60	65	60
B	65	90	50	70	60	60

Mahasiswa A: $\overline{X}$ = 360/6 = 60

Mahasiswa B: $\overline{X}$ = 360/6 = 60

Rata-rata hasil test kedua mahasiswa tersebut tidak berbeda, namun dispersi hasil test mahasiswa B (30 sampai dengan 90) jauh lebih besar dari pada varisasi hasil test mahasiswa A (50 sampai dengan 65). Hal ini berarti hasil test mahasiswa A jauh lebih konsisten (stabil) dibanding mahasiswa B. Tingkat dispersi berhubungan erat dengan sifat kesamaan/kesejenisan data. Misalnya data tentang besarnya modal pedagang kaki lima khusus makanan, akan kecil variasinya jika dibandingkan dengan data seluruh pedagang kaki lima tanpa melihat jenis dagangannya. Secara umum, suatu rata-rata akan cukup representatif bagi serangkaian nilai-nilai observasi X_i bila nilai-nilai tersebut diperoleh dari data yang bersifat sejenis bagi tujuan pengamatan tertentu.

2. Pengukuran Jarak (Range)

Pengukuran jarak sebuah distribusi merupakan pengukuran dispersi yang paling sederhana. Jarak sebuah distribusi frekuensi dirumuskan sebagai "selisih atau beda antara pengukuran nilai terbesar dan nilai terkecil yang terdapat dalam sebuah distribusi frekuensi". Atau secara matematis dapat ditulis sebagai berikut:

$$R = X_u - X_i$$

Keterangan :

R = range data observasi

X_u = nilai tertinggi

X_i = nilai terindah

Beberapa Catatan Tentang Pengukuran dan Penggunaan Jarak

1) Hasil pengukuran jarak (range) sebenarnya sudah dapat menggambarkan disperse (variasi) nilai-nilai observasi dengan cara yang paling sederhana. Jika kita ingin memperoleh hasil pengukuran dispersi secara kasar dan cepat, maka ukuran range dapat digunakan.

2) Range bukan merupakan pengukuran dispersi distribusi yang memuaskan karena hasil pengukurannya jelas tergantung pada kedua nilai ekstrim tanpa mengikutsertakan pola dispersi nilai-nilai observasi $X_i$secara keseluruhan.

Contoh kasus :

Berikut ini adalah nilai ulangan harian 10 siswa mata pelajaran statistika di SMA Mercu Buana Yogyakarta:

56 66 78 94 48 82 50 76 80 70

Range nilai 10 siswa yang ikut ulangan harian statistika tersebut dapat ditentukan dengan menggunakan formula :

$$R = X_u - X_i$$

$$= 94 - 48 = 46$$

Range data observasi berkelompok (grouped data) adalah data selisih antara tepi kelas atas kelas yang terakhir dengan tepi kelas bawah kelas pertama.

Contoh kasus :

Tabel 2.1 berikut ini data mengenai nilai 30 peserta ujian Matematika di SMA Mercu Buana Yogyakarta

Tabel 2.1

NILAI	FREKUENSI (f)
40 – 49	6
50 – 59	10
60 – 69	4
70 – 79	4
80 – 89	2
90 – 99	4

Range nilai 30 peserta ujian matematika dapat ditentukan dengan menggunakan Rumus :

$$R = B_u - B_i$$

Dengan nilai-nilai

$B_u = 99{,}5$ (tepi kelas atas kelas yang terakhir)

$B_i = 39{,}5$ (tepi kelas bawah kelas yang pertama)

Sehingga besarnya Range (R)

$$R = 99{,}5 - 39{,}5 = 60$$

3. Pengukuran Deviasi Kuartil.

Nilai-nilai X_i yang ordinatnya membagi seluruh distribusi dalam 4 (empat) bagian yang sama dinamakan nilai-nilai kuartil. Q1 merupakan kuartil pertama, Q2 merupakan kuartil kedua dan sama dengan median ($Q2 = md$), sedangkan Q3 dinamakan kuartil ketiga. Dalam distribusi kuartil, 50% dari semua nilai-nilai observasi seharusnya terletak antara Q1 dan Q3. Jarak antara Q1 dan Q3 dinamakan jarak inter-kuartil (inter-quartilrange). Makin kecil jarak tersebut, maka makin tinggi tingkat konsentrasi distribusi tengah seluas 50% dari seluruh distribusi.

Secara teoritis, pengukuran deviasi kuartil sebuah sampel dapat rumuskan sebagai:

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

Selanjutnya dapat dikatakan bahwa deviasi kuartil adalah sebesar +dQ atau –dQ dari mediannya.

Pada dasarnya, pengukuran deviasi kuartil sama seperti pengukuran jarak (range). Pengukurannya didasarkan pada jarak antara Q1 dan Q3. Pengukuran tersebut tidak dipengaruhi oleh dispersi dari seluruh nilai-nilai observasi, deviasi kuartil hanya mengikutsertakan dispersi nilai-nilsi observasi X_i yang didistribusikan di tengah-tengah seluruh distribusi seluas 50% saja.

4. Pengukuran Deviasi Rata-rata(Mean Deviation)

a. *Deviasi rata-rata dari data yang belum dikelompokkan*

Dispersi serangkaian nilai-nilai observasi akan kecil bila nilai-nilai tersebut berkonsentrasi sekitar rata-ratanya. Sebaliknya, dispersinya akan besar bila nilai-nilai observasi tersebar jauh dari rata-ratanya.

Deviasi rata-rata dari seluruh nilai-nilai observasi X_i dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum(x_i - \bar{x})}{n}$$

Sedangkan pengukuran deviasi atas dasar nilai-nilai absolut dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum |x_i - \bar{x}|}{n}$$

Contoh :

Carilah deviasi rata-rata data berikut ini :

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Jawab :

Dimana i=1,2,3,4,…..,20

$$d_{\bar{x}} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{706}{20} = 35,31$$

b. Deviasi rata-rata dari data yang telah dikelompokkan

Apabila nilai-nilai observasi sudah dikelompokkan ke dalam bentuk distribusi frekuensi, maka deviasi rata-ratanya dapat dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum_{i=1}^{k} f_i |m_i - \bar{x}|}{n}$$

Dimana:

m_i = titik tengah kelas frekuensi

f_i= frekuensi dari kelas distribusi ke-i

k = jumlah kelas distribusi

Dalam beberapa kondisi tertentu, median dapat digunakan sebagai pengukuran rata-rata secara memuaskan. Deviasi rata-rata sebuah distribusi dapat juga diukur dari median distribusi yang bersangkutan seperti dirumuskan sebagai:

$$d_{\bar{x}} = \frac{\sum |x_i - md\,|}{n}$$

Atau

$$d_{\bar{x}} = \frac{\sum f_i |m_i - md\,|}{n}$$

Umumnya deviasi rata-rata merupakan pengukuran dispersi yang lebih baik jika dibandingkan dengan jarak atau deviasi kuartil. Hasil pengukuran deviasi rata-rata mencerminkan dispersi tiap-tiap nilai observasi dari rata-ratanya dan bukan hanya tergantung pada kedua nilai ekstrim.

Contoh :

Dari data tunggal dibawah ini, rubahlah menjadi data kelompok :

40	50	70	55
55	72	66	60
60	54	85	65
45	67	80	75
70	80	55	80

Dan carilah Deviasi rata-ratanya.

Jawab :

Data setelah dikelompokkan

Nilai	F	Mi
40 – 47	2	43,5
48 – 55	5	51,5
56 – 63	2	59,5
64 – 71	5	67,5
72 – 79	2	75,5
80 – 87	4	83,5
$\sum$	20	

$$Median\ (Md) = \ Mo + P\left(\frac{\frac{n}{2} - fm}{fc}\right)$$

$$= 63{,}5 + 8\left(\frac{10 - 9}{5}\right)$$

$$= 63{,}5 + 8(0{,}2)$$

$$= 63{,}5 + 1{,}6$$

$$= 65{,}1$$

$$d_{\bar{x}} = \frac{\sum f_i |x_i - md\,|}{n}$$

$$= \frac{|(43{,}5 - 65{,}1).\,2 + (51{,}5 - 65{,}1).\,5 + \cdots + (83{,}5 - 65{,}1).\,4|}{20}$$

$$= \frac{228{,}8}{20} = 11{,}44$$

5. Pengukuran Varians dan Deviasi Standar

Varians digunakan untuk melihat kehomogenan data secara kasar, dimana nilai hasil perhitungan varians sebagai titik pusat dari penyebaran data.

Contoh 1:

Seorang guru matematika melakukan tes prestasi dengan membagi siswa dalam 3 kelompok, yaitu A,B, dan C. Dalam satu kelompok terdapat 5 siswa. Walaupun dibentuk kelompok namun untuk tes dikerjakan secara individu. Didapat hasil sebagai berikut :

KELOMPOK	NILAI					$\bar{x}$
A	50	50	50	50	50	50
B	60	40	50	55	45	50
C	30	70	90	10	50	50

a. *Varians dan deviasi standar dari data yang belum dikelompokkan*

Karl Pearson merumuskan pengukuran **varians** sebagai:

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

Standarisasi unit-unit pengukuran di atas dilakukan melalui proses pengakaran, dan dinamakan **deviasi standar**, sebagai berikut:

$$s = \sqrt{s^2}$$

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

b. Variansdan deviasi standar dari data yang belum dikelompokkan

- Rumus Fisher dan Wilks
  Varians dari Fisher dan Wilks:

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

- Deviasi standar dari Fisher dan Wilks:

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

- Varians dan deviasi standar **populasi**
  Varians polupasi:

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

- Deviasi standar populasi:

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

c. ***Varians dan deviasi standar dari data yang telah dikelompokkan***

- **Varians** dari data sampel yang telah dikelompokkan:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 . f_i$$

- **Deviasi standar** dari data sampel yang telah dikelompokkan:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 . f_i}$$

dimana:

x_i= titik tengah tiap-tiap kelas

f_i = jumlah frekuensi kelas

d. *Variansi* dan deviasi standar dengan cara transformasi

Seperti halnya dengan mencari nilai mean data kelompok. Kita juga dapat mencari nilai variansi dapat dicari dengan cara transformasi.

$$u_i = x_i - a$$

Dimana :

x_i : titik tengah interval kelas ke-i

a : sembarang harga titik tengah interval kelas (**biasanya yang memiliki frekuensi terbanyak**)

sehingga rumus **VARIANSI (S^2)**adalah :

$$s^2 = c^2u^2$$

c = lebar kelas/panjang kelas

dimana :

$$s_{\bar{u}}^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i(u_i - \bar{u})^2$$

Atau dapat juga ditulis :

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} f_i u_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} f_i u_i\right)^2\right]$$

Contoh :

Dari data tinggi badan (cm) 50 mahasiswa Pendidikan Matematika FKIP Universitas Mercu Buana Yogyakarta didapat data :

Tabel 1. Perhitungan variansi data berkelompok

Interval Kelas	x_i	u_i	f_i	u_i^2	$f_i u_i$	$f_i u_i^2$
164,5 – 167,5	166	166-175=-9	6	81	6*-9=-54	6*81 =486
167,5 – 170,5	169	169-175=-6	7	36	7*-6=-42	7*36 = 252
170,5 – 173,5	172	-3	8	9	-24	72
173,5 – 176,5	175	0	11	0	0	0
176,5 – 179,5	178	178-175= 3	7	9	21	63
179,5 – 182,5	181	6	6	36	36	216
182,5 – 185,5	184	9	5	81	45	405
Jumlah			**50**		**-18**	**1494**

Berdasarkan tabel 1 dengan menggunakan rumus transormasi, maka variansinya :

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} f_i u_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} f_i u_i\right)^2\right]$$

$$= \frac{1}{50-1}\left(1494 - \frac{1}{50}(-18)^2\right) = 30{,}35$$

$$s = \sqrt{30{,}4} = 5{,}50$$

e. Beberapa catatan tentang varians dan deviasi standar dari data yang telah dikelompokkan

- Koreksi Sheppard (Sheppard's Correction): Jika distribusi frekuensi simetris atau mendekati simetris, maka hasil rata-rata hitung yang diperoleh dari distribusi frekuensi tersebut kurang lebih sama dengan hasil rata-rata yang diperoleh dari data kasar (yang belum dikelompokkan.
- Distribusi normal sebenarnya merupakan distribusi teoritis (mengikuti "hokum normal") karena pada dasarnya gejala-gejala alami tidak seluruhnya bersifat normal.

Latihan :

Dari data diabawah ini :

NO	NILAI	F
1	5 – 9,99	6
2	10 – 14,99	12
3	15 – 19,99	19
4	20 – 24,99	20
5	25 – 29,99	14
6	30 – 34,99	8
7	35 – 39,99	2
	JUMLAH	80

Maka tentukan :

1. Gambarlah diagram batang, garis
2. Tentukan Mean, median, Modus, Variansi, SD
3. Tentukan Variansi dan SD dengan cara transformasi

Latihan-latihan

1. Populasi beranggotakan orang dengan ukuran masing-masing : 4,5,6,7,10,12,14. Diambil 2 sampel ukuran dengan pengambilan sampel dilakukan tanpa pengembalian. Buatlah distribusi sampling rata-ratanya?
2. Di ketahui sebuah data tentang nilai prestasi matematika siswa kelas X SMA Mercu Buana.

Nilai (X)	f
4	2
5	3
6	8
7	4
8	5
9	3
10	2
3	3

Dari tabel diatas maka tentukan :

a. Proporsi, Kumulasi Proporsi Bawah dan Proporsi Kumulasi atas
b. Simpangan x, Jumlah Kuadrat Simpangan (JK(X)), Variansi, dan Simpangan Baku.

3. Data dikotomi tentang motivasi belajar dan prestasi belajar matematika

SISWA	MOTIVASI		PRESTASI BELAJAR	
	RENDAH (X1)	TINGGI (X2)	RENDAH (Y1)	TINGGI (Y2)
A	0	1	1	0
B	0	1	1	0
C	0	1	0	1
D	1	0	0	1
E	1	0	0	1
F	0	1	1	0
G	0	1	1	0
H	1	0	0	1
I	0	1	0	1
J	0	1	1	0
K	0	1	1	0
L	0	1	1	0
M	0	1	0	1
N	0	1	0	1
O	0	1	0	1
P	0	1	1	0
Q	0	1	0	1
R	0	1	1	0
S	1	0	0	1
T	1	0	0	1
U	1	0	0	1
V	0	1	0	1
W	1	0	0	1
X	1	0	1	0
Y	1	0	1	0
Z	1	0	0	1
AA	1	0	0	1
AB	0	1	1	0
AC	0	1	0	1
AD	0	1	0	1
AE	0	1	1	0
AF	1	0	0	1
AG	1	0	0	1
AH	1	0	0	1
AI	1	0	0	1

Dari data diatas maka tentukan : Rerata, variansi, simpangan baku untuk X1, X2, Y1 dan Y2