

Program Pasca Sarjana Terapan  
Politeknik Elektronika Negeri Surabaya



# Probability and Random Process

Topik 2. Statistik Deskriptif

Prima Kristalina

Maret 2016

# Outline

1. Penyajian Data
  - Tabel dan grafik frekuensi kemunculan, tabel dan grafik frekuensi relative
  - Pengelompokan data, histogram
2. Ukuran Pemusatan Data
  - Mean, Median, Modus
3. Ukuran Penyebaran Data
  - Range, deviasi rata-rata, varians, simpangan baku
4. Ukuran Kemiringan Data (skewness)
5. Ukuran Keruncingan Data (kurtosis)

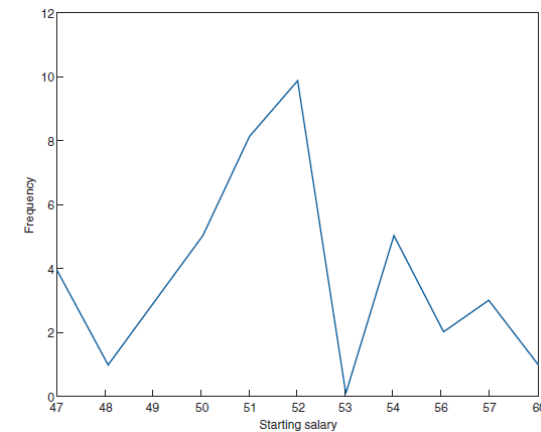
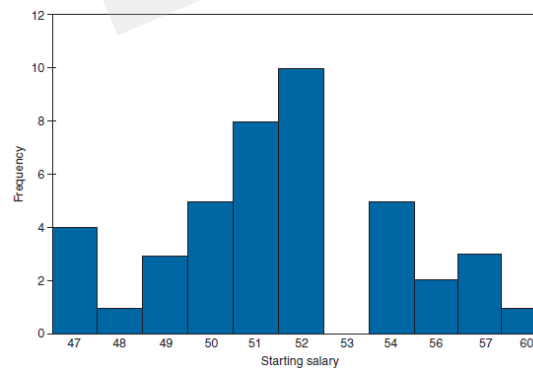
# Penyajian Data

(1/2)

## Tabel dan grafik frekuensi kemunculan

- Sekelompok data yang memiliki perbedaan nilai yang relatif kecil bisa disajikan dalam bentuk tabel frekuensi kemunculan.
- Selain itu juga bisa disajikan dalam bentuk grafik, baik grafik batang, grafik plot maupun pie chart

Starting Salary	Frequency
47	4
48	1
49	3
50	5
51	8
52	10
53	0
54	5
56	2
57	3
60	1



# Penyajian Data

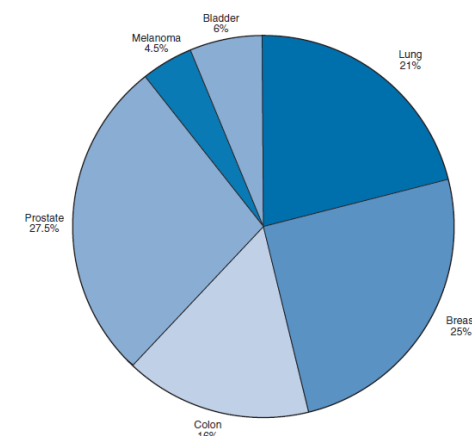
(2/2)

## Tabel dan grafik frekuensi relative

- Jika sekelompok data terdiri dari  $n$  nilai, maka frekuensi kemunculan sebuah nilai tertentu,  $f$ , adalah  $f/n$ , dinamakan sebagai frekuensi relatif.
- Jadi frekuensi relatif sebuah nilai data adalah proporsi data yang memiliki nilai tersebut.
- Frekuensi relatif data juga bisa disajikan dalam bentuk table atau grafik frekuensi relatif.
- Sekilas grafik frekuensi relatif sama dengan frekuensi kemunculan, hanya dibedakan pada *sumbu y* yang harus dibagi dengan jumlah nilai  $n$ .

Starting Salary	Frequency
47	$4/42 = .0952$
48	$1/42 = .0238$
49	$3/42$
50	$5/42$
51	$8/42$
52	$10/42$
53	0
54	$5/42$
56	$2/42$
57	$3/42$
60	$1/42$

Type of Cancer	Number of New Cases	Relative Frequency
Lung	42	.21
Breast	50	.25
Colon	32	.16
Prostate	55	.275
Melanoma	9	.045
Bladder	12	.06



# Pengelompokan Data (1/2)

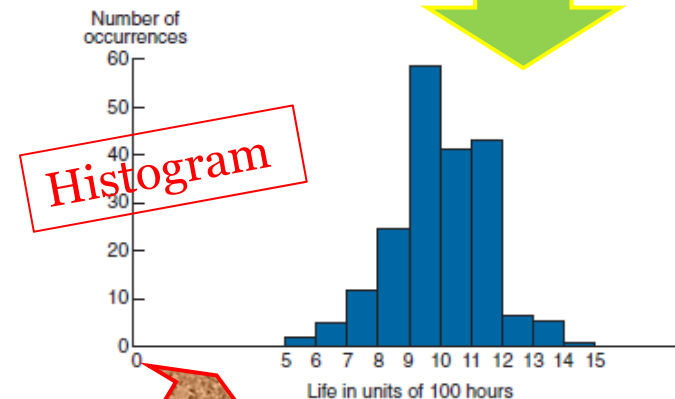
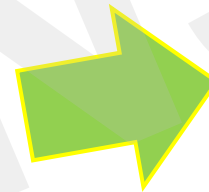
- Penyajian data dengan bentuk frekuensi keluaran dan frekuensi relative hanya cocok digunakan jika sebaran datanya tidak terlalu tinggi.
- Jika sebaran nilai pada data sangat besar, sebaiknya nilai-nilai tersebut dapat dikelompokkan menjadi beberapa sub grup, atau kelas interval.
- Data bisa di-plot berdasarkan kelas interval tersebut.

# Pengelompokan Data (2/2)

- Contoh:

Item Lifetimes									
1,067	919	1,196	785	1,126	936	918	1,156	920	948
855	1,092	1,162	1,170	929	950	905	972	1,035	1,045
1,157	1,195	1,195	1,340	1,122	938	970	1,237	956	1,102
1,022	978	832	1,009	1,157	1,151	1,009	765	958	902
923	1,333	811	1,217	1,085	896	958	1,311	1,037	702
521	933	928	1,153	946	858	1,071	1,069	830	1,063
930	807	954	1,063	1,002	909	1,077	1,021	1,062	1,157
999	932	1,035	944	1,049	940	1,122	1,115	833	1,320
901	1,324	818	1,250	1,203	1,078	890	1,303	1,011	1,102
996	780	900	1,106	704	621	854	1,178	1,138	951
1,187	1,067	1,118	1,037	958	760	1,101	949	992	966
824	653	980	935	878	934	910	1,058	730	980
844	814	1,103	1,000	788	1,143	935	1,069	1,170	1,067
1,037	1,151	863	990	1,035	1,112	931	970	932	904
1,026	1,147	883	867	990	1,258	1,192	922	1,150	1,091
1,039	1,083	1,040	1,289	699	1,083	880	1,029	658	912
1,023	984	856	924	801	1,122	1,292	1,116	880	1,173
1,134	932	938	1,078	1,180	1,106	1,184	954	824	529
998	996	1,133	765	775	1,105	1,081	1,171	705	1,425
610	916	1,001	895	709	860	1,110	1,149	972	1,002

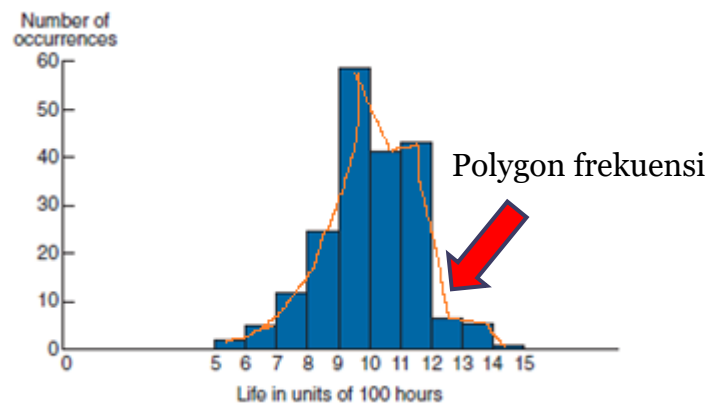
Class Interval	Frequency (Number of Data Values in the Interval)
500-600	2
600-700	5
700-800	12
800-900	25
900-1000	58
1000-1100	41
1100-1200	43
1200-1300	7
1300-1400	6
1400-1500	1



Kelas interval

# Histogram

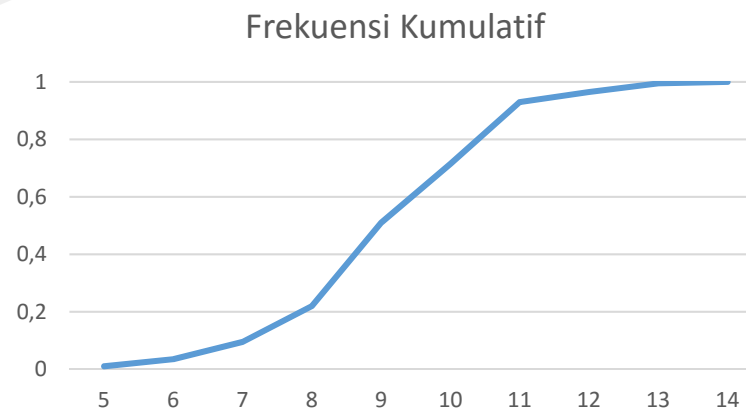
- Gambar grafik kelas interval yang diwakili dengan sebuah bar, dimana lebar dari bar tersebut menyatakan kelas interval, disebut *histogram*.
- Batas tepi kotak dari bar menunjukkan tepi bawah dan tepi atas dari kelas interval.
- Sumbu x dari grafik tersebut menunjukkan kelas intervalnya, sedangkan sumbu y menunjukkan frekuensi kemunculan atau frekuensi relative nilai pada masing-masing kelas interval.
- Jika titik tengah atas dari bar-bar tersebut saling dihubungkan, maka akan didapatkan *polygon frekuensi*.



# Kurva Frekuensi Kumulatif (1/2)

- Grafik yang disajikan berdasarkan data yang sudah tersusun dalam bentuk *tabel distribusi frekuensi kumulatif* dinamakan **Ogive**
- Grafik dari tabel yang tersusun dari frekuensi kumulatif kurang dari, berupa ogive positif, sedangkan grafik dari table dengan frekuensi kumulatif lebih dari, berupa ogive negatif

No.	Kelas Interval	Frek Kemunculan	Frek. Relatif	Frek. Kumulatif
5	500-600	2	0,01	0,01
6	600-700	5	0,025	0,035
7	700-800	12	0,06	0,095
8	800-900	25	0,125	0,22
9	900-1000	58	0,29	0,51
10	1000-1100	41	0,205	0,715
11	1100-1200	43	0,215	0,93
12	1200-1300	7	0,035	0,965
13	1300-1400	6	0,03	0,995
14	1400-1500	1	0,005	1

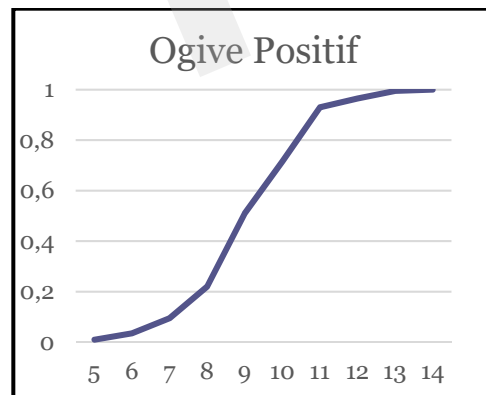




# Kurva Frekuensi Kumulatif (2/2)

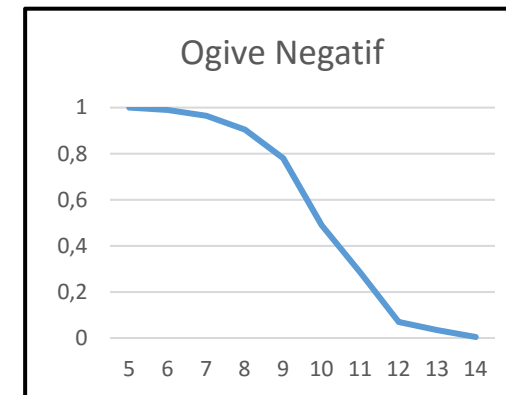
## Ogive Positif

No.	Kelas Interval	Frek Kemunculan	Frek Kemunculan kurang dari	Frek. Kumulatif Kurang Dari
5	500-600	2	2	0,01
6	600-700	5	7	0,035
7	700-800	12	19	0,095
8	800-900	25	44	0,22
9	900-1000	58	102	0,51
10	1000-1100	41	143	0,715
11	1100-1200	43	186	0,93
12	1200-1300	7	193	0,965
13	1300-1400	6	199	0,995
14	1400-1500	1	200	1



## Ogive Negatif

No.	Kelas Interval	Frek Kemunculan	Frek Kemunculan Lebih dari	Frek. Kumulatif Lebih dari
5	500-600	2	200	1
6	600-700	5	198	0,99
7	700-800	12	193	0,965
8	800-900	25	181	0,905
9	900-1000	58	156	0,78
10	1000-1100	41	98	0,49
11	1100-1200	43	57	0,285
12	1200-1300	7	14	0,07
13	1300-1400	6	7	0,035
14	1400-1500	1	1	0,005



# Ukuran Pemusatan Data (1/5)

## 1. Rata-rata (mean)

- Merupakan jumlah dari seluruh nilai data dibagi dengan banyaknya data.
- Langkah pertama sebelum menentukan rata-rata adalah mengelompokkan data dahulu.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

# Ukuran Pemusatan Data (2/5)

## 2. Median

- Merupakan nilai tengah dari nilai-nilai pengamatan yang disusun secara teratur menurut besarnya data
- Median tidak dipengaruhi nilai ekstrim (*outlier*)

$$M_e = b + P \left[ \frac{\frac{1}{2}n - F}{f} \right]$$

Me = median  
b=batas bawah kelas median  
n=banyaknya data  
P=Panjang Kelas/interval  
F=frekuensi kumulatif sebelum kelas yang memuat median  
f=frekuensi kelas yang memuat median

# Ukuran Pemusatan Data (3/5)

## 3. Modus

- Nilai yang mempunyai frekuensi terbesar, kemunculan terbanyak dalam sebuah kumpulan data.

$$M_o = b + P \frac{b_1}{b_1 + b_2}$$

Me = modus

b=batas bawah kelas modus

P=Panjang Kelas/interval

b1 = frek kelas modus dikurangi frek sebelumnya

b2= = frek kelas modus dikurangi frek berikutnya

# Ukuran Pemusatan Data (4/5)

## Contoh:

1. Diketahui data nilai 10 siswa adalah sbb:

6 5 5 7 7.5 8 6.5 5.5 6 9

- Hitung mean dari nilai ke-10 siswa tersebut
- Hitung median dari ke-10 nilai tersebut
- Hitung modus dari ke-10 nilai tersebut

## Jawab

$$\text{Mean} = \bar{x} = \frac{6+5+5+7+7.5+8+6.5+5.5+6+9}{10} = 6.55$$

Disusunurut : 5 5 5.5 6 6 6.5 7 7.5 8 9

$$\text{Median} = M_e = \frac{6+6.5}{2} = 6.25$$

Modus =  $M_o$  = 5 dan 6

# Ukuran Pemusatan Data (5/5)

2. Diketahui tabel data dari tinggi badan sekelompok siswa adalah sebagai berikut:

Hitung mean, median dan modus dari data-data tersebut

Tinggi Badan	Frekuensi	Frekuensi Kumulatif
131 - 140	2	2
141 - 150	5	7
151 - 160	12	19
161 - 170	10	29
171 - 180	6	35

$$\bar{x} = \frac{(2 \times 135.5) + (5 \times 145.5) + (12 \times 155.5) + (10 \times 165.5) + (6 \times 175.5)}{(2 + 5 + 12 + 10 + 6)}$$

$$= \frac{5572.5}{35} = 159.21$$

$$M_e = b + P \left[ \frac{\frac{1}{2}n - F}{f} \right] = 150.5 + 10 \left( \frac{((0.5 \times 35) - 7)}{12} \right) = 159.25$$

$$M_o = b + P \left[ \frac{b_1}{b_1 + b_2} \right] = 151 + 10 \left( \frac{(12 - 7)}{(12 - 7) + (12 - 10)} \right) = 158.14$$

# Ukuran Penyebaran Data (1/5)

- Kegunaan Ukuran Penyebaran Data:
  - a. Untuk menggambarkan bagaimana suatu kelompok data menyebar terhadap pusat datanya.
  - b. Untuk menentukan apakah nilai rata-ratanya benar-benar representative atau tidak
  - c. Untuk mengadakan perbandingan terhadap variabilitas data
  - d. Dapat dipakai untuk membantu ukuran statistika dalam menguji hipotesa, apakah dua sampel berasal dari populasi yang sama atau tidak

# Ukuran Penyebaran Data (2/5)

## 1. Rentang(range)

- Selisih antara nilai maksimum dengan nilai minimum dalam sebuah gugus data

$$R = X_n - X_1$$

## 2. Deviasi rata-rata

- Deviasi rata-rata adalah rata-rata penyimpangan data-data dari rata-ratanya (mean)
- Dicari harga mutlak dari selisih tiap data dengan mean-nya

$$\text{Deviasi} = \frac{\sum |x - \bar{x}|}{n}$$



$$\sigma_{xx}^2 = \frac{1}{N} \sum (X_i - \bar{X})^2$$

## Ukuran Penyebaran Data (3/5)

### 3. Ragam (varians)

- Menunjukkan keragaman nilai di dalam sebuah himpunan.
- Untuk sekumpulan data, rata-rata varians adalah jumlah kuadrat dari selisih nilai observasi terhadap nilai rata-rata elemen data himpunan tersebut dibagi dengan jumlah data.

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

### 4. Simpangan Baku (deviasi standard)

- Merupakan akar dari varians.

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

# Ukuran Penyebaran Data (4/5)

## Contoh:

1. Diketahui data nilai UAS dari dua kelas adalah sebagai berikut:

Kelas A	90	80	70	90	70	100	80	50	75	70
Kelas B	80	80	75	95	75	70	95	60	85	60

- Hitung Range dan deviasi rata-rata (DR) nya

Range kelas A adalah  $100 - 50 = 50$

Range kelas B adalah  $95 - 60 = 35$

$$\bar{X}_{\text{Kelas A}} = \frac{50 + 70 + 70 + 70 + 75 + 80 + 80 + 90 + 90 + 100}{10} = 77.5$$

$$\bar{X}_{\text{Kelas B}} = \frac{60 + 60 + 70 + 75 + 75 + 80 + 80 + 85 + 95 + 95}{10} = 77.5$$

$$\text{DR Kelas A} = \frac{|50 - 77.5| + |70 - 77.5| + |70 - 77.5| + |70 - 77.5| + |75 - 77.5| + |80 - 77.5| + |80 - 77.5| + |90 - 77.5| + |90 - 77.5| + |100 - 77.5|}{10} = 9.75$$

$$\text{DR Kelas B} = \frac{|60 - 77.5| + |60 - 77.5| + |70 - 77.5| + |75 - 77.5| + |75 - 77.5| + |80 - 77.5| + |80 - 77.5| + |85 - 77.5| + |95 - 77.5| + |95 - 77.5|}{10} = 9.25$$

# Ukuran Penyebaran Data (5/5)

2. Interval nilai matakuliah tertentu di sebuah kelas ditunjukkan pada tabel berikut:

Carilah:

- Varians
- Simpangan Baku

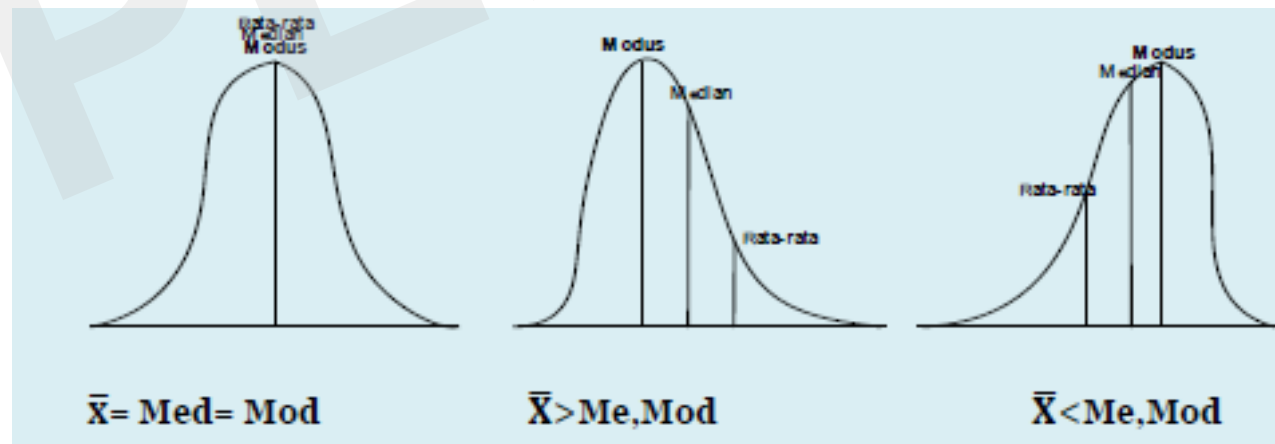
Interval	Frekuensi	$x_i$
53 – 58	2	55.5
59 – 64	12	61.5
65 – 70	10	67.5
71 – 76	23	73.5
77 – 82	14	79.5
83 – 88	10	85.5
89 – 94	5	91.5
95 – 100	4	97.5
Jumlah	80	$\mu=76.5$

## Ukuran Kemiringan Kurva (1/4)

- Kemiringan (*skewness*) adalah derajat ketidaksimetrisan atau penyimpangan dari kesimetrisan sebuah distribusi
- Distribusi yang tidak simetris akan memiliki rata-rata (mean), median dan modus yang tidak sama besarnya, sehingga distribusi data akan terkonsentrasi pada salah satu sisi kurvanya
- Untuk distribusi yang memiliki ekor lebih panjang ke kanan disebut miring ke kanan atau memiliki kemiringan positif
- Sebaliknya, jika ekor lebih panjang ke kiri disebut miring ke kiri atau memiliki kemiringan negatif

# Ukuran Kemiringan Kurva (2/4)

- Tingkat kemiringan (koefisien kemiringan) dapat didekati dengan beberapa persamaan:
- Pearson
- Bowley
- Persentil
- Momen



$\bar{X}$ =mean, Me=Median, Mod=Modus

# Ukuran Kemiringan Kurva (3/4)

- Koefisien Kemiringan Pearson

- Merupakan nilai selisih rata-rata ( $\bar{X}$ ) dengan modus (mod) dibagi simpangan baku (S)
- Dinyatakan sebagai:

$$s_k = \frac{\bar{X} - M_o}{S}$$

- Jika dinyatakan dengan median (med), maka:

$$s_k = \frac{3(\bar{X} - M_e)}{S}$$

$s_k = 0$  Kurva berbentuk simetris

$s_k > 0$  Nilai terkonsentrasi di kanan

Kurva miring ke kanan, miring positif

$s_k < 0$  Nilai terkonsentrasi di kiri

Kurva miring ke kiri, miring negatif

# Ukuran Kemiringan Kurva (4/4)

- **Koefisien Kemiringan Momen**

- Didasarkan pada perbandingan momen ke-3 dengan pangkat tiga simpangan baku
- Dinyatakan sebagai:

$$a_3 = \frac{M_3}{S^3} = \frac{\frac{1}{2} \sum (X - \bar{X})^3}{S^3}$$



Data Tunggal

- Atau

$$a_3 = \frac{M_3}{S^3} = \frac{\frac{1}{2} \sum (X - \bar{X})^3 f}{S^3}$$



Data Berkelompok

- Jika:

$0,0 \leq (s_k = a^3) < 0,1$  kurva distribusi dikatakan normal

$0,0 \leq (s_k = a^3) < 0,3$  kurva distribusi sedikit menceng (negatif ke kiri, positif ke kanan)

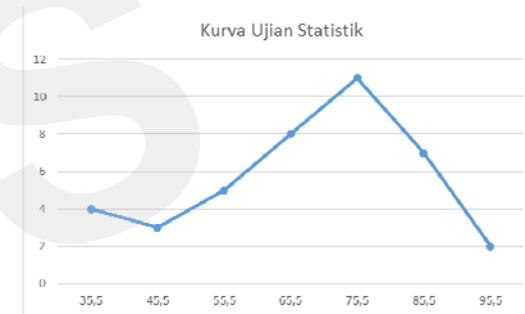
$(s_k = a^3) \geq 0,3$  kurva distribusi menceng sekali (negatif ke kiri, positif ke kanan)

Contoh:

Berikut diberikan daftar nilai statistika dari sebuah prodi di PENS

Interval Nilai	Frekuensi
31 - 40	4
41 - 50	3
51 - 60	5
61 - 70	8
71 - 80	11
81 - 90	7
91 - 100	2

- Tentukan nilai  $s_k$  dan ujlah arah kemiringan kurvanya.
- Gambarlah kurvanya



Interval Nilai	$X_r$	f	c	$c^2$	fxc	fxc <sup>2</sup>
31 - 40	35,5	4	-4	16	-16	64
41 - 50	45,5	3	-3	9	-9	27
51 - 60	55,5	5	-2	4	-10	20
61 - 70	65,5	8	-1	1	-8	8
71 - 80	75,5	11	0	0	0	0
81 - 90	85,5	7	1	1	7	7
91 - 100	95,5	2	2	4	4	8
Jumlah		40			-32	134

$$\bar{X} = X_0 + i \left( \frac{\sum f_i \cdot c_i}{\sum f_i} \right) = 75,7 + 10 \left( \frac{-32}{40} \right) = 67,5$$

$$S = i \sqrt{\frac{\sum f_i \cdot c_i^2}{n} - \left[ \frac{\sum f_i \cdot c_i}{n} \right]^2} = 10 \sqrt{\frac{134}{40} - \left[ \frac{-32}{40} \right]^2} = 16,2$$

$$M_e = B_m + P \left\{ \frac{\frac{1}{2}n - F}{f_m} \right\} = 60,5 + 10 \left\{ \frac{\frac{1}{2}40 - (4 + 3 + 5)}{8} \right\} = 60,5 + 10 = 70,5$$

$$\text{mod} = Med + i \left\{ \frac{d_1}{d_1 + d_2} \right\} s = 70,5 + 10 \frac{4}{4 + 5} = 70,5 + 4,44 = 74,94$$

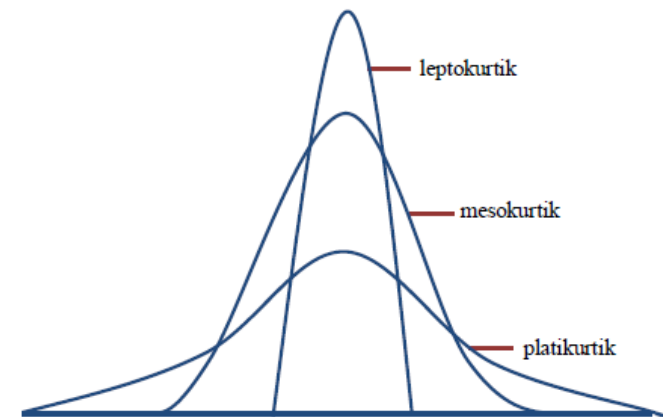
sehingga:

$$s_k = \frac{\bar{X} - \text{mod}}{S} = \frac{67,5 - 74,94}{16,2} = -0,46$$



# Ukuran Keruncingan Kurva (1/2)

- Keruncingan (kurtosis) adalah tingkat kepuncakan dari sebuah distribusi yang biasanya diambil secara relatif terhadap distribusi normal
- Berdasarkan keruncingannya dibagi atas:
- Leptokurtik
  - Merupakan distribusi yang memiliki puncak relatif tinggi
- Platikurtik
  - Merupakan distribusi yang memiliki puncak hampir mendatar
- Mesokurtik
  - Merupakan distribusi yang memiliki puncak tidak tinggi tidak rendah



## Ukuran Keruncingan Kurva (2/2)

- Koefisien keruncingan digunakan untuk mengetahui keruncingan suatu distribusi
  - Jika  $\alpha < 3$ , distribusinya platikurtik
  - Jika  $\alpha > 3$ , distribusinya leptokurtik
  - Jika  $\alpha = 3$ , distribusinya mesokurtik

$$\alpha_4 = \frac{\frac{1}{n} \sum (X - \bar{X})^4}{S^4}$$



Data Tunggal

$$\alpha_4 = \frac{\frac{1}{n} \sum (X - \bar{X})^4 f}{S^4}$$



Data Kelompok

- Contoh

- Tentukan keruncingan kurva dari data: 1 3 5 8 12

Jawab:

$$\bar{X} = \frac{1+3+5+8+12}{5} = 5,8$$

$$s = \sqrt{\frac{\sum(x_i)^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = 2,315$$

$$\alpha_4 = \frac{\frac{1}{5}(2093,78)}{(2,315)^4} = \frac{418,756}{28,72} = 14,58$$

X	X- $\bar{X}$	(X- $\bar{X}$ ) <sup>4</sup>
1	-4,8	530,842
3	-2,8	61,4656
5	-0,8	0,4096
8	2,2	23,4256
12	6,2	1477,63
Jumlah	0	2093,78

Karena nilai  $\alpha > 3$  maka distribusinya adalah leptokurtik

# Tugas

1. Kumpulan nilai semua juri ( $N=50$ ) mengenai hasil score sebuah perlombaan dinyatakan dalam tabel berikut ini

95	75	85	90	85
80	65	80	75	85
100	75	85	85	80
90	95	90	80	75
95	90	90	85	70
90	95	95	85	80
90	90	85	100	90
85	100	80	90	95
80	90	85	80	95
90	100	85	75	80

Kelompokkan nilai tersebut menjadi kelas interval dengan panjang kelas =10

Hitung:

1. Rata-rata nilai
2. Median
3. Modus
4. Variansi
5. Simpangan Baku
6. Frekuensi Kumulatif nya
7. Gambarkan histogram nya

2. Statistik di sebuah hotel dalam 1 bulan memiliki data sebagai berikut:

Jumlah hari menginap	Jumlah tamu
1 - 3	7
4 - 6	6
7 - 9	10
10 - 12	4
13 - 15	3

Hitung:

1. Rata-rata jumlah tamu menginap dalam 1 bulan
2. Modus jumlah tamu
3. Dengan frekuensi kumulatif menggunakan ogive positif, dapatkan
  - a. Berapa persen jumlah tamu menginap antara 10 sampai 12 hari ?
  - b. Berapa persen jumlah tamu menginap kurang dari 6 hari ?

3. Upah pekerja harian di sebuah perusahaan disajikan sebagai berikut:

Upah kerja (Ribuan Rp)	Jumlah pekerja harian
65 - 67	3
68 - 70	6
71 - 73	7
74 - 76	8
77 - 79	5
80 - 82	3
83 - 85	2

1. Buat tabel frekuensi kumulatif lebih dari (*ogive negative*)
2. Hitung berapa persen pekerja harian yang memiliki upah lebih dari Rp 74.000,- ?
3. Buat grafik penyebaran data jumlah pekerja harian tersebut terhadap kelas interval upah nya.
4. Dapatkan koefisien kemiringannya, dan miring ke arah mana ?
5. Dapatkan koefisien keruncingannya, dan termasuk tipe manakah keruncingan grafiknya ?