Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin

UKURAN PENYEBARAN DATA, SKEWNESS DAN KURTOSIS (UKURAN DISPERSI)

UKURAN DISPERSI

- Ukuran dispersi adalah ukuran variasi atau seberapa jauh nilai tersebar data dengan lainnya dari gugus data.
- Aplikasi ukuran dispersi yang sering digunakan adalah standar deviasi.
- Ukuran dispersi biasanya digunakan bersamaan dengan tendensi sentral untuk mempelajari distribusi data.

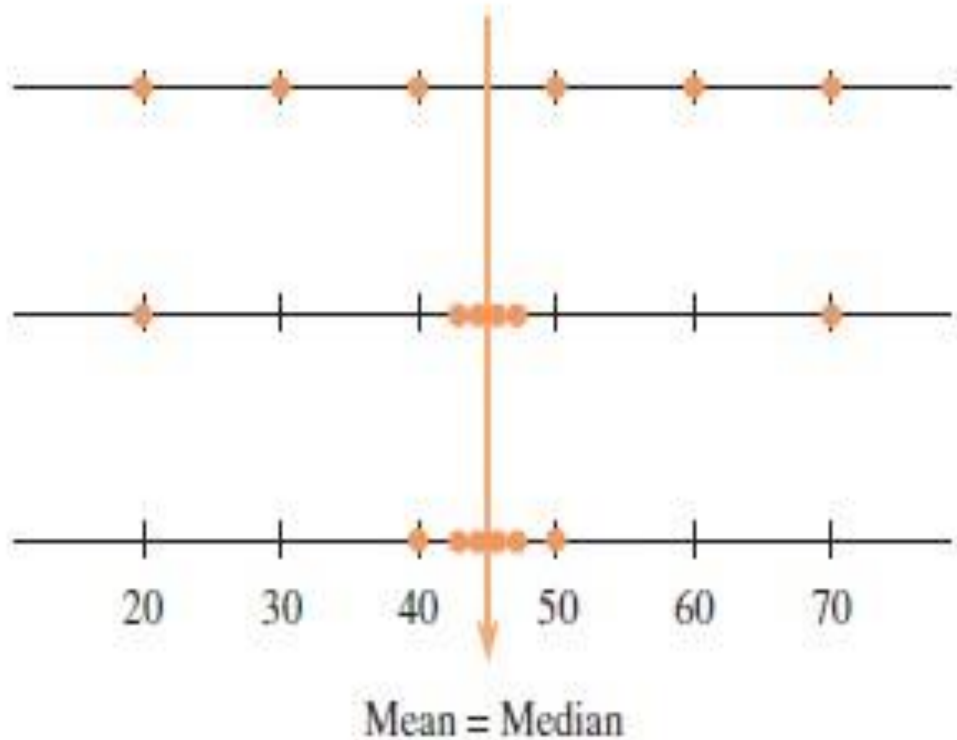
Contoh Kasus

Sample

1. 20, 40, 50, 30, 60, 70

2. 47, 43, 44, 46, 20, 70

3. 44, 43, 40, 50, 47, 46



RANGE

Ukuran dispersi yg merupakan selisih nilai maksimum (X_{max}) dan minimum (X_{min}).

Range = data terbesar – data terkecil

$$= X_{max} - X_{min}$$

26	37	39	46	49	59	69	76	83	83	83	87	87	95	95
↓ x_1														↓ x_2

$$R = 95 - 26 = 69$$

RATA-RATA SIMPANGAN

Ukuran variabelitas yang juga banyak digunakan untuk mendes- kripsikan sejauh mana sampel pengamatan menyimpang dari rata-rata sampel x adalah rata-rata penyimpangan dari mean atau rata-rata simpangan.

Simpangan untuk data tunggal dirumuskan sebagai

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

Untuk data kelompok dirumuskan sebagai

$$S_x = \frac{\sum_{1}^{n} f_i |x_i - \bar{x}|}{n}$$

RATA-RATA SIMPANGAN

Contoh : (Data tunggal)

Tentukan rata-rata simpangan data berikut :

6092	**5249**	**5851**	**5843**	**6505**	**6659**	**6883**	**4814**	**6661**	**5910**	**5913**	**6556**

Rata-rata ($\dot{x}$)= 5635

x_i	$\lvert x_i-\dot{x}\rvert$
6092	457
5249	-386
5851	216
5843	208
6505	870
6659	1024
6883	1248
4814	-821
6661	1026
5910	275
5913	278
6556	921
Σ	5316

$$S_x = \frac{\sum_1^n |x_i - \bar{x}|}{n} = \frac{\sum_1^{12} |x_i - \bar{x}|}{12}$$

$$= \frac{5316}{12} = 443$$

RATA-RATA SIMPANGAN

Contoh : (Data Berkelompok)

Tentukan rata-rata simpangan data berikut :

x_i	Frekuensi	$f_i x_i$	$\lvert x_i - \dot{x} \rvert$	$f_i \lvert x_i - \dot{x} \rvert$
55	1	55	-20,56	-20,56
60	4	240	-15,56	-62,22
65	4	260	-10,56	-42,22
70	6	420	-5,56	-33,33
75	5	375	-0,56	-2,78
80	3	240	4,44	13,33
85	3	255	9,44	28,33
90	2	180	14,44	28,89
100	1	100	24,44	24,44
	∑ = 29			∑= -66,12

Rata-rata = 75,56

$$S_x = \frac{\sum_1^n f_i \lvert x_i - \bar{x} \rvert}{n} = \frac{-66,12}{29}$$

$= -2,28$

SIMPANGAN BAKU (DEVIASI STANDAR)

Untuk populasi yang berjumlah besar, sangat tidak mungkin untuk mendapatkan nilai rata-rata populasi μ serta deviasi standartnya σ. Untuk mengestimasi (menaksir) nilai μ dan σ, diambil sampel data. Nilai μ diestimasi oleh $\bar{x}$ dan σ diestimasi oleh s .

Rumus Deviasi Populasi dengan :

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

N = Jumlah observasi dalam populasi

μ = Rata-rata populasi.

SIMPANGAN BAKU (DEVIASI STANDAR)

Standar Deviasi Sampel

Simpangan baku atau deviasi standar (*Standard Deviation*) merupakan ukuran penyebaran yang paling baik, karena menggambarkan besarnya penyebaran tiap-tiap unit observasi. Karl Pearson menamakannya deviasi standar dan dirumuskan sebagai :

$$S = \sqrt{\frac{\sum_{i=1}^{N}\left(X_i - \overline{X}\right)^2}{N-1}}$$

N = Jumlah sampel

$\overline{X}$= Rata-rata sampel.

SIMPANGAN BAKU (DEVIASI STANDAR)

Contoh :
Diberikan sample dengan data 6, 7, 8, 9, 10, Hitunglah standar deviasinya (simpangan baku)

Hitung nilai rata-rata sampel :

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

Xi	Xi-$\overline{X}$	(Xi-$\overline{X}$)^2
6	-2	4
7	-1	1
8	0	0
9	1	1
10	2	4
Jumlah (∑)		10

$$S = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \overline{X})^2}{N-1}} = \sqrt{\frac{10}{5-1}} = \sqrt{2{,}5}$$

SIMPANGAN BAKU (DEVIASI STANDAR)

Standar Deviasi dari data kelompok distribusi frekuensi yang berasal dari sampel didefinisikan :

$$S = \sqrt{\frac{\sum_{i=1}^{N} f_i \left(X_i - \overline{X}\right)^2}{N-1}}$$

N = Jumlah sampel

$\overline{X}$ = Rata-rata sampel.

f_i = Frekuensi kelas ke- i

X_i = nilai tengah kelas ke - i

SIMPANGAN BAKU (DEVIASI STANDAR)

Kelas	Frekuensi (f_i)	Nilai tengah (X_i)	f_iX_i	$X_i - \overline{X}$	$(X_i - \overline{X})^2$	$f_i (X_i - \overline{X})^2$
50-54	1	52	52	-23,375	546,3906	546,3906
55-59	2	57	114	-18,375	337,6406	675,2813
60-64	11	62	682	-13,375	178,8906	1967,797
65-69	10	67	670	-8,375	70,14063	701,4063
70-74	12	72	864	-3,375	11,39063	136,6875
75-79	21	77	1617	1,625	2,640625	55,45313
80-84	6	82	492	6,625	43,89063	263,3438
85-89	9	87	783	11,625	135,1406	1216,266
90-94	4	92	368	16,625	276,3906	1105,563
95-99	4	97	388	21,625	467,6406	1870,563
	80		6030			8538,75

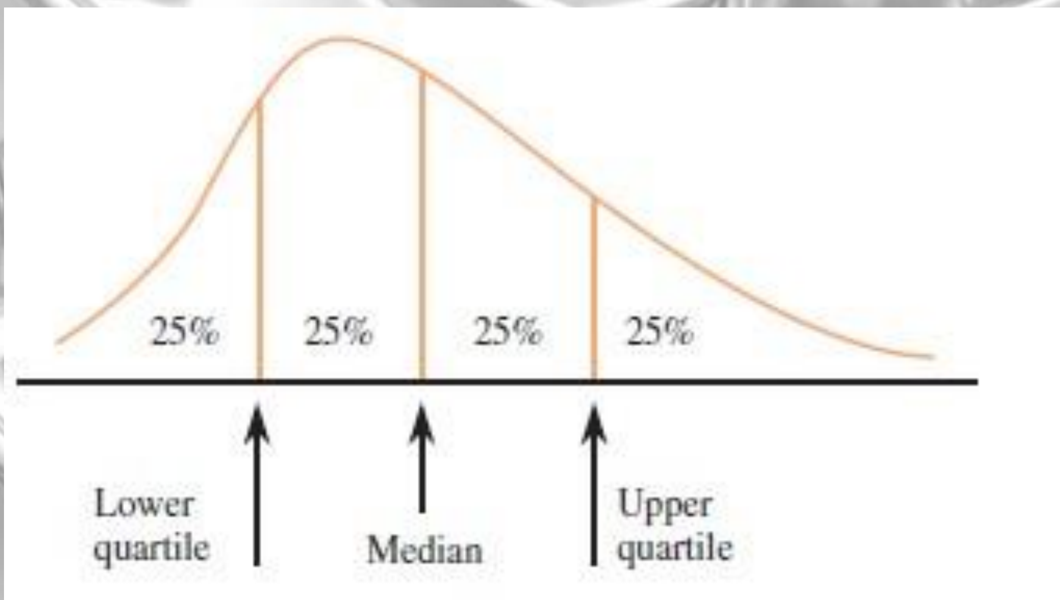
Nilai rata-rata, $\overline{X} = \frac{6030}{80} = 75{,}375$

Standar deviasi data berkelompok:

$$s = \sqrt{\frac{\sum_{i=1}^{N} f_i\left(X_i - \overline{X}\right)^2}{N-1}} = \sqrt{\frac{8538{,}75}{80-1}} = \sqrt{108{,}09} = 10{,}396$$

RANGE INTERQUARTIL

Median didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi dua bagian yang sama dan kuartil didefinisikan sebagai nilai yang membagi seluruh rentang nilai menjadi empat bagian yang sama. Rangeinterkuartil adalah ukuran variabilitas berdasarkan kuartil.



Iqr= kuartir atas - kuartir bawah

Iqr= $Q_3 - Q_1$

RANGE INTERQUARTIL

Pengukuran dispersi atas dasar jangkauan inter-kuartil dinamakan deviasi *kuartil* atau *simpangan kuartil* (*quartile deviation*) dan dirumuskan sebagai.

$$D_{iqr} = \frac{Kuartir\ atas - kuartir\ bawah}{2}$$

atau

$$D_{iqr} = \frac{Q_3 - Q_1}{2}$$

RANGE INTERQUARTIL

Berikut adalah tabel produksi pulsa telpon di Indonesia

Tahun	Lokal (000 pulsa)	SLJJ (Menit)
1998	16236246427	29668416066
1999	16236724396	31021632143
2000	18516778571	34342636
2001	20227877123	38161484336
2002	19730308403	41397291119
2003	23887950222	42447349726
2004	19936304184	45215914717
2005	22920220767	57745329624
2006	23646924115	61443360381
2007	29018054840	53129188172
2008	22233240642	40706864477

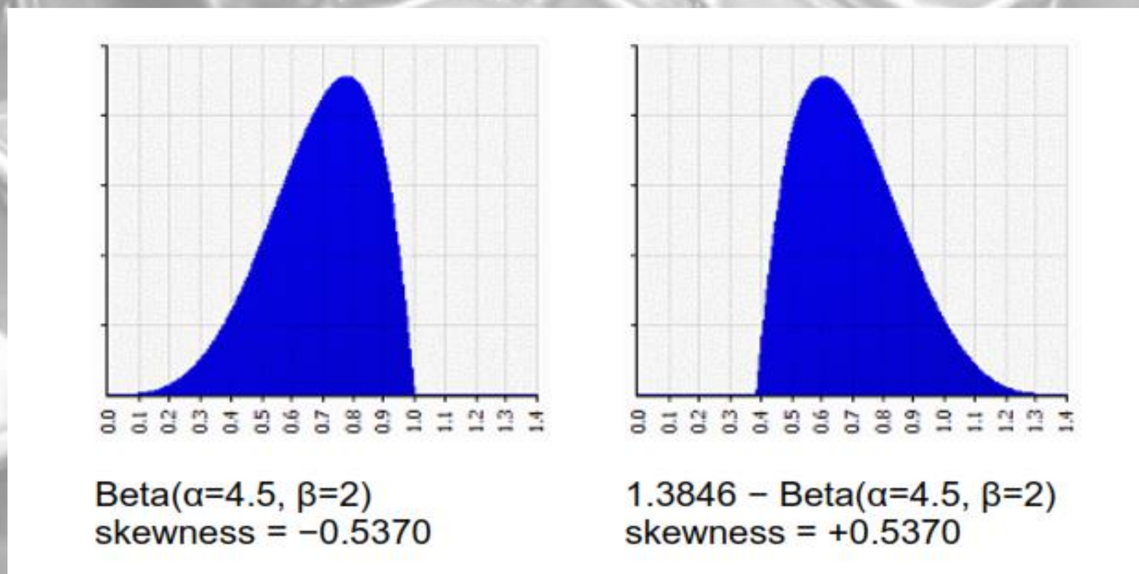
Sumber : Kantor Pusat PT. TELKOM Indonesia

Median = 23887950222
Kuartil bawah = 18516778571
Kuartil atas = 23646924115

Diqr = (23646924115- 18516778571)/2
= 2565072772

Skewnes dan Kurtosis

Skewness (kemiringan) adalah derajat ketidaksimetrisan suatu distribusi. Jika kurva frekuensi suatu distribusi memiliki ekor yang lebih memanjang ke kanan (dilihat dari meannya) maka dikatakan menceng kanan (positif) dan jika sebaliknya maka menceng kiri (negatif).Secara perhitungan, skewness adalah momen ketiga terhadap mean. Distribusi normal (dan distribusi simetris lainnya, misalnya distribusi t atau Cauchy) memiliki skewness 0 (nol). Perhatikan gambar berikut . Kedua gambar memiliki $\mu = 0.6923$ and $\sigma = 0.1685$ yang sama tetapi keduanya memiliki kemencengan yang berbeda.



SKEWNES (KEMIRINGAN)

Rumus koefisien skewnes

$$g_1 = \frac{m_3}{m_2^{3/2}}$$

dimana :

$$m_2 = \frac{\sum_i^n (x_i - \bar{x})^2}{n} \text{ disebut varian}$$

$$m_3 = \frac{\sum_i^n (x_i - \bar{x})^3}{n} \text{ Momen ketiga}$$

Skewnes sampel data

$$G_1 = \frac{\sqrt{n(n-1}}{n-1} g_1$$

Kurtosis (Keruncingan)

Rumus Kurtois

Kurtosis $\mathrm{a}_4 = \frac{m_4}{m_2^2}$

Excess Kurtosis $\mathrm{g}_2 = \mathrm{a}_4 - 3$

dimana :

$$m_2 = \frac{\sum_i^n (x_i - \bar{x})^2}{n}$$ disebut varian

$$m_4 = \frac{\sum_i^n (x_i - \bar{x})^4}{n}$$ Momen ke empat

Skewnes sampel data

$$\mathrm{G}_2 = \frac{n-1}{(n-2)(n-3)}[(n+1)\mathrm{g}_2 + 6]$$