

Ukuran Penyebaran Data

Setelah mengetahui nilai pusat dari suatu data melalui ukuran pemusatan data, langkah berikutnya adalah memahami seberapa jauh data menyebar dari nilai pusat tersebut. **Ukuran penyebaran data**, atau **dispersi**, memberikan informasi penting tentang variasi atau keragaman dalam kumpulan data. Dua set data bisa memiliki *mean* yang sama, tetapi penyebarannya bisa sangat berbeda, dan perbedaan ini dapat memengaruhi cara interpretasi data secara keseluruhan. Dalam sub-bab ini, akan dibahas beberapa ukuran utama penyebaran, yaitu rentang (*range*), standar deviasi, dan varians, yang membantu menggambarkan tingkat homogenitas atau heterogenitas suatu data secara lebih mendalam.

Range

Range, atau rentang data, adalah selisih antara nilai tertinggi dengan nilai terendah dalam sebuah set data. *Range* menunjukkan seberapa **lebar sebaran nilai** dalam data, namun karena hanya mempertimbangkan dua nilai ekstrem, ukuran ini sangat sensitif terhadap *outlier* dan tidak menggambarkan variasi data secara keseluruhan. Meskipun demikian, *range* tetap berguna sebagai gambaran awal tentang sebaran data.

Untuk dapat memahami fungsi *range* dalam memahami set data, mari kita pelajari kasus berikut ini. Misalnya, seorang peneliti memiliki data nilai tugas mata pelajaran Matematika dari 3 kelas yang berbeda:

Kelas A	4	4	6	6	7	4	7	3	5	4
Kelas B	2	6	6	7	7	8	4	4	3	3
Kelas C	6	5	6	4	5	6	4	4	5	5

Pertanyaan penting mengenai data tersebut adalah apakah ketiga kelas tersebut memperoleh capaian belajar yang sama. Jika hanya mengandalkan nilai rata-rata, maka ketiga kelas tersebut akan tampak sama karena memiliki nilai rata-rata yang sama, yaitu 5. Padahal jika dilihat ke perolehan nilai individual, terlihat berbeda. Nilai *range* pada data dapat membantu kita untuk menggambarkan seberapa berbeda dan bervariasi data yang dimiliki dan makna variasi tersebut dalam memahami setiap poin data. Untuk itu, kita perlu menghitung selisih skor terbesar dengan skor terkecil untuk setiap data kelas.

$$\text{Range kelas A} = \text{nilai tertinggi} - \text{nilai terendah} = 7 - 3 = 4$$

$$\text{Range kelas B} = \text{nilai tertinggi} - \text{nilai terendah} = 8 - 2 = 6$$

$$\text{Range kelas C} = \text{nilai tertinggi} - \text{nilai terendah} = 6 - 2 = 4$$

Dari hasil penghitungan, dapat dilihat bahwa kelas B memiliki lebar data yang paling besar, artinya terdapat perbedaan atau variasi yang lebih besar di dalam kelas tersebut dibandingkan

dengan kelas-kelas lainnya. Sebaliknya, kelas C memiliki rentang skor paling kecil di antara ketiganya. Dalam hal ini, satu skor yang sama (misalnya, 4) dapat dimaknai secara berbeda berdasarkan nilai *range* tiap kelas.

Satu hal yang perlu diingat adalah bahwa variasi skor yang digambarkan dalam *range* hanya menunjukkan perbedaan antara yang mendapatkan nilai paling tinggi dan yang mendapatkan nilai paling rendah. Dari nilai *range* tersebut kita tidak memiliki informasi mengenai seberapa besar ukuran variabilitas (atau seberapa bervariasi) sebuah set data. Untuk mendapatkan gambaran ini, kita menggunakan ukuran penyebaran data berikutnya, yaitu varians.

Varians

Varians (s^2) adalah ukuran penyebaran data yang menunjukkan **seberapa jauh nilai-nilai dalam suatu kumpulan data menyimpang dari nilai *mean***. Varians dihitung dengan menjumlahkan kuadrat selisih setiap nilai (X) terhadap *mean* (\bar{x}), kemudian dibagi dengan jumlah data (N) (untuk populasi) atau jumlah data dikurangi satu ($n-1$) (untuk sampel). Dapat juga ditulis sebagai:

$$s^2 = \frac{\sum (X - \bar{x})^2}{N - 1}$$

Hasil varians menunjukkan “rata-rata kuadrat penyimpangan” dari *mean*, sehingga semakin besar nilai varians, semakin besar pula variasi data di sekitar *mean*. Dengan kata lain, varians juga menggambarkan **homogenitas** data, di mana semakin kecil varians maka semakin kecil perbedaan antar poin data, dan sebaliknya semakin besar varians semakin besar pula perbedaan antar poin data.

Sebagai ilustrasi, misalnya terdapat data set skor skala sikap dari dua kelompok partisipan yang berbeda (a & b), dengan masing-masing sebaran data sebagai berikut:

Kelompok a	6	6	6	6	7	7	7	7	7	8	8	8	$\bar{x}=6.9$
Kelompok b	2	3	5	5	6	6	6	7	9	9	10	12	$\bar{x}=6.7$

Untuk dapat memperoleh nilai varians, kita perlu menghitung simpangan atau selisih dari setiap data poin terhadap mean kemudian dikuadratkan, seperti pada Table 3.

Dari hasil penghitungan tersebut, kita menemukan perbedaan varians yang cukup besar antara kedua kelompok, meskipun *mean* keduanya tidak jauh berbeda. Hal ini menandakan bahwa kelompok b memiliki sebaran data yang jauh lebih beragam (heterogen) dibandingkan kelompok a (lihat **Gambar 3.1**). Oleh karena itu, pemaknaan terhadap nilai *mean* tiap kelompok juga berbeda, relatif terhadap variansnya masing-masing.

Table 3

Kelompok a			Kelompok b		
Skor (X)	Simpangan $(X - \bar{x})$	Kuadrat simpangan $(X - \bar{x})^2$	Skor (X)	Simpangan $(X - \bar{x})$	Kuadrat simpangan $(X - \bar{x})^2$
6	-0.9	0.81	2	-4.7	22.09
6	-0.9	0.81	3	-3.7	13.69
6	-0.9	0.81	5	-1.7	2.89
6	-0.9	0.81	5	-1.7	2.89
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	7	0.3	0.09
7	0.1	0.01	9	2.3	5.29
8	1.1	1.21	9	2.3	5.29
8	1.1	1.21	10	3.3	10.89
8	1.1	1.21	12	5.3	28.09
$\Sigma(X - \bar{x})^2$		6.92	$\Sigma(X - \bar{x})^2$		92.68
$s^2 = \frac{\Sigma(X - \bar{x})^2}{N - 1}$		0.6	$s^2 = \frac{\Sigma(X - \bar{x})^2}{N - 1}$		8.4

Varians memiliki fungsi penting dalam memahami keragaman atau variasi data di sekitar nilai rata-rata. Informasi ini sangat berguna untuk a) menilai konsistensi data (misalnya, dua kelompok dengan rata-rata yang sama bisa memiliki varians yang berbeda; kelompok dengan varians kecil berarti anggotanya lebih homogen.) dan b) dasar dari analisis statistik lanjutan, di mana varians menjadi komponen penting dalam berbagai teknik analisis, seperti standar deviasi, analisis varians (*ANOVA*), uji-t, dan regresi. Dengan demikian, varians bukan hanya ukuran penyebaran, tetapi juga alat untuk mengevaluasi struktur dan kualitas data sebelum melakukan interpretasi atau pengambilan keputusan lebih lanjut.