Cara menghitung varians sampel dan varians populasi

Sample data set

	x
x_1	17
x_2	15
x_3	23
x_4	7
x_5	9
x_6	13

Sample Variance (s^2)

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

s^2 = variance

x_i = term in data set

$\bar{x}$ = Sample mean

$\sum$ = Sum

n = Sample size

Sample mean ($\bar{x}$)

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

$$\bar{x} = \frac{17 + 15 + 23 + 7 + 9 + 13}{6}$$

$$= \frac{84}{6}$$

$$= 14$$

	x	$\bar{x}$
x_1	17	
x_2	15	
x_3	23	
x_4	7	
x_5	9	
x_6	13	
Σ	84	14

- ula-mula, jumlahkan semua data: 17 + 15 + 23 + 7 + 9 + 13 = 84 Lalu, bagi jawabannya dengan jumlah data, dalam contoh ini dengan enam: 84 ÷ 6 = 14.
  Mean sampel = $\bar{x}$ = 14.
- Anda dapat menganggap mean sebagai "titik tengah" dari data. Jika data berkumpul di sekitar mean, variansnya rendah. Jika data tersebar jauh dari mean, variansnya tinggi.

$$x_i - \bar{x}$$

$x_1 - \bar{x} = 17 - 14 = 3$

$x_2 - \bar{x} = 15 - 14 = 1$

$x_3 - \bar{x} = 23 - 14 = 9$

$x_4 - \bar{x} = 7 - 14 = -7$

$x_5 - \bar{x} = 9 - 14 = -5$

$x_6 - \bar{x} = 13 - 14 = -1$

	x	$\bar{x}$	$x_i - \bar{x}$
x_1	17		3
x_2	15		1
x_3	23		9
x_4	7		-7
x_5	9		-5
x_6	13		-1
Σ	84	14	

4 **Kurangkan nilai setiap data dengan mean.** Sekarang kita menghitung x_i - $\bar{x}$, di mana x_i adalah nilai dari tiap data. Setiap hasil menggambarkan deviasi data dari mean, atau dalam bahasa sederhana, seberapa jauh data dari mean.[5] .

- **Contoh:**

  $x_1 - \bar{x} = 17 - 14 = 3$

  $x_2 - \bar{x} = 15 - 14 = 1$

  $x_3 - \bar{x} = 23 - 14 = 9$

  $x_4 - \bar{x} = 7 - 14 = -7$

  $x_5 - \bar{x} = 9 - 14 = -5$

  $x_6 - \bar{x} = 13 - 14 = -1$

$(x_i-\bar{x})^2$

$(x_1-\bar{x})^2 = 3^2 = 9$

$(x_2-\bar{x})^2 = 1^2 = 1$

$(x_3-\bar{x})^2 = 9^2 = 81$

$(x_4-\bar{x})^2 = -7^2 = 49$

$(x_5-\bar{x})^2 = -5^2 = 25$

$(x_6-\bar{x})^2 = -1^2 = 1$

	x	$\bar{x}$	$x_i-\bar{x}$	$(x_i-\bar{x})^2$
x_1	17		3	9
x_2	15		1	1
x_3	23		9	81
x_4	7		-7	49
x_5	9		-5	25
x_6	13		-1	1
Σ	84	14		

5 **Kuadratkan hasilnya.** Seperti yang telah dijelaskan sebelumnya, jumlah dari seluruh nilai deviasi (x_i - x̄) akan sama dengan nol. Ini artinya "rata-rata deviasi" akan selalu sama dengan nol, dan hal ini tidak memberikan informasi apa-apa tentang sebaran data. Untuk menyelesaikan masalah ini, kita mengkuadratkan nilai setiap deviasi. Ini akan membuat angkanya menjadi positif semua, sehingga nilai negatif dan positif tidak saling menghilangkan.[6]

- **Contoh:**
  $(x_1 - \bar{x})^2 = 3^2 = 9$
  $(x_2 - \bar{x})^2 = 1^2 = 1$
  $9^2 = 81$
  $(-7)^2 = 49$
  $(-5)^2 = 25$
  $(-1)^2 = 1$
- Sekarang Anda mendapatkan nilai $(x_i - \bar{x})^2$ untuk setiap data dari sampel.

$$\sum (x_i - \bar{x})^2$$

$$= (x_1-\bar{x})^2 + (x_2-\bar{x})^2 + (x_3-\bar{x})^2 + (x_4-\bar{x})^2 + (x_5-\bar{x})^2 + (x_6-\bar{x})^2$$

$$= 9 + 1 + 81 + 49 + 25 + 1$$

$$= 166$$

	x	$\bar{x}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
x_1	17		3	9
x_2	15		1	1
x_3	23		9	81
x_4	7		-7	49
x_5	9		-5	25
x_6	13		-1	1
Σ	84	14		166

6 **Cari jumlah dari kuadrat nilai.** Sekarang kita akan menghitung nilai seluruh pembilang di dalam rumus: $\sum[(x_i - \overline{x})^2]$. Huruf besar sigma, $\sum$, berarti jumlah dari semua nilai secara berurutan x_i. Anda telah menghitung $(x_i - \overline{x})^2$ untuk setiap nilai x_i pada sampel, jadi sekarang Anda tinggal menjumlahkannya.

- **Contoh:** 9 + 1 + 81 + 49 + 25 + 1 = **166**.

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

$$n = 6$$

$$= \frac{166}{6-1}$$

$$= \frac{166}{5}$$

$$= 33.2$$

7 **Bagi dengan n - 1, di mana n adalah jumlah data.** Dulu, para ahli statistik hanya membagi dengan n ketika menghitung varians sampel. Dengan demikian kita mendapat nilai rata-rata dari deviasi kuadrat, yang cocok untuk menghitung varians sampel tersebut. Tetapi ingatlah, sebuah sampel hanyalah estimasi dari populasi yang lebih besar. Jika kita mengambil sampel lain secara acak dan melakukan perhitungan, hasilnya akan berbeda. Tampaknya, membagi dengan n - 1 ketimbang n memberi perkiraan nilai varians yang lebih baik untuk populasi, yang sebetulnya ingin kita ketahui. Koreksi ini sudah menjadi begitu umum sehingga sekarang diterima sebagai definisi dari varians.[7]

- **Contoh:** Ada enam data di dalam contoh ini, jadi n = 6.

  Varians sampel adalah = $s^2 = \frac{166}{6-1} = \mathbf{33.2}$

Sample variance	Sample standard deviation
$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$	$s = \frac{\sum(x_i - \bar{x})^2}{n-1}$
$s^2 = 33.2$	$s = \sqrt{33.2}$
	$= 5.76$

Menghitung varians populasi

Population data set

	x
x_1	5
x_2	5
x_3	8
x_4	12
x_5	15
x_6	18

1 **Mulailah dengan sejumlah data populasi.** Istilah "populasi" mengacu pada semua pengamatan yang relevan. Misalnya, jika kita ingin meneliti tentang usia penduduk Texas, populasi yang kita gunakan adalah usia setiap orang yang tinggal di Texas. Kita mungkin butuh membuat lembar kerja (*spreadsheet*) untuk data sebesar itu, tetapi mari kita gunakan data yang lebih kecil sebagai contoh:

- **Contoh:** Ada enam buah akuarium di dalam sebuah ruangan. Keenam akuarium tersebut terisi sejumlah ikan:

  $x_1 = 5$
  $x_2 = 5$
  $x_3 = 8$
  $x_4 = 12$
  $x_5 = 15$
  $x_6 = 18$

Population Variance (σ^2)

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{n}$$

σ^2 = population variance

x_i = term in data set

$\sum$ = sum

μ = population mean

n = population size

2 **Tuliskan rumus varians populasi.** Karena populasi memiliki semua data yang kita perlukan, rumus ini bisa kita gunakan untuk menghitung secara tepat varians populasi. Untuk membedakannya dengan varians sampel (yang hanya estimasi), ahli statistik menggunakan variabel yang berbeda:[8]

- $\sigma^2 = {}^{(\sum(x_i - \mu)^2)}/_{n}$
- σ^2 = varians populasi. Simbolnya adalah sigma dalam huruf kecil, dikuadratkan. Varians diukur dalam unit kuadrat.
- x_i melambangkan entri dari setiap data.
- Entri di dalam $\sum$ dimasukkan untuk setiap nilai x_i, lalu dijumlahkan.
- μ adalah mean dari populasi.
- n adalah jumlah data dari populasi.

Population Mean (μ)

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

$$= \frac{5+5+8+12+15+18}{6}$$

$$= \frac{63}{6}$$

$$= 10.5$$

	x	μ
x_1	5	
x_2	5	
x_3	8	
x_4	12	
x_5	15	
x_6	18	
$\sum$	63	10.5

3 **Cari mean populasi.** Ketika menganalisis sebuah populasi, simbol μ ("mu") melambangkan rata-rata aritmetik. Untuk mencari mean, jumlahkan semua data, lalu bagi dengan jumlah data.

- Anda mungkin mengira bahwa mean sama dengan "rata-rata". Berhati-hatilah sebab kata itu memiliki banyak definisi dalam matematika.
- **Contoh:** mean = μ = $\frac{5+5+8+12+15+18}{6}$ = **10.5**

$(x_i - \mu)$

$(x_1 - \mu) = 5 - 10.5 = -5.5$

$(x_2 - \mu) = 5 - 10.5 = -5.5$

$(x_3 - \mu) = 8 - 10.5 = -2.5$

$(x_4 - \mu) = 12 - 10.5 = 1.5$

$(x_5 - \mu) = 15 - 10.5 = 4.5$

$(x_6 - \mu) = 18 - 10.5 = 7.5$

	X	μ	$(x_i - \mu)$
X_1	5		-5.5
X_2	5		-5.5
X_3	8		-2.5
X_4	12		1.5
X_5	15		4.5
X_6	18		7.5
Σ	63	10.5	

4 **Kurangkan setiap data dengan mean.** Data yang lebih dekat dengan mean akan menghasilkan selisih yang lebih dekat dengan nol. Ulangi pengurangan untuk setiap data, dan Anda dapat mulai mengamati seberapa tersebarnya data.

- **Contoh:**

  x_1 - μ = 5 - 10.5 = -5.5

  x_2 - μ = 5 - 10.5 = -5.5

  x_3 - μ = 8 - 10.5 = -2.5

  x_4 - μ = 12 - 10.5 = 1.5

  x_5 - μ = 15 - 10.5 = 4.5

  x_6 - μ = 18 - 10.5 = 7.5

$$(x_i - \mu)^2$$

$(x_1 - \mu)^2 = (-5.5)^2 = 30.25$

$(x_2 - \mu)^2 = (-5.5)^2 = 30.25$

$(x_3 - \mu)^2 = (-2.5)^2 = 6.25$

$(x_4 - \mu)^2 = (1.5)^2 = 2.25$

$(x_5 - \mu)^2 = (4.5)^2 = 20.25$

$(x_6 - \mu)^2 = (7.5)^2 = 56.25$

	X	μ	$(x_i - \mu)$	$(x_i - \mu)^2$
X_1	5		-5.5	30.25
X_2	5		-5.5	30.25
X_3	8		-2.5	6.25
X_4	12		1.5	2.25
X_5	15		4.5	20.25
X_6	18		7.5	56.25
Σ	63	10.5		

5 **Kuadratkan setiap hasil.** Sekarang kita bisa melihat bahwa beberapa angka negatif dihasilkan dari proses sebelumnya, dan beberapa yang lain positif. Jika Anda membayangkan data-data tersebut pada sebuah garis bilangan, kedua kategori ini mewakili data yang berada di sebelah kiri dan sebelah kanan mean. Hal ini tidak berguna dalam menghitung varians, karena kedua kelompok ini akan saling menghilangkan. Kuadratkanlah setiap angka supaya mereka menjadi positif.

- **Contoh:**
  $(x_i - \mu)^2$ untuk setiap nilai *i* dari 1 sampai 6:
  $(-5.5)^2 = 30.25$
  $(-5.5)^2 = 30.25$
  $(-2.5)^2 = 6.25$
  $(1.5)^2 = 2.25$
  $(4.5)^2 = 20.25$
  $(7.5)^2 = 56.25$

	x	μ	$(x_i-\mu)$	$(x_i-\mu)^2$
x_1	5		-5.5	30.25
x_2	5		-5.5	30.25
x_3	8		-2.5	6.25
x_4	12		1.5	2.25
x_5	15		4.5	20.25
x_6	18		7.5	56.25
Σ	63	10.5		145.5

$$\sigma^2 = \frac{30.25+30.25+6.25+2.25+20.25+56.25}{6}$$

$$= \frac{145.5}{6}$$

$$= 24.25$$

6 **Cari mean dari hasil.** Sekarang Anda telah memperoleh sebuah nilai untuk setiap data, yang berhubungan (secara tidak langsung) dengan jarak data tersebut dari mean. Cari mean dari hasil ini dengan menjumlahkan mereka semuanya, lalu dibagi dengan jumlah angka.

- **Contoh:**

  Varians dari populasi =

$$\frac{30.25 + 30.25 + 6.25 + 2.25 + 20.25 + 56.25}{6} = \frac{145.5}{6} = 24.25$$

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{n}$$

$$= \frac{(x_1 - \mu)^2 + (x_i - \mu)^2 + \ldots + (x_n - \mu)^2}{n}$$

7 Hubungan dengan rumus semula. Jika Anda ragu apakah perhitungan ini sama dengan rumus yang diberikan di awal, coba tuliskan seluruh perhitungan secara panjang:

- Setelah menemukan selisih dengan mean dan mengkuadratkannya, Anda mendapatkan hasil $(x_1 - \mu)^2$, $(x_2 - \mu)^2$, dan seterusnya $(x_n - \mu)^2$, di mana x_n adalah data terakhir.
- Carilah mean dari nilai-nilai ini, jumlahkan dan bagi dengan n: $((x_1 - \mu)^2 + (x_2 - \mu)^2 + ... + (x_n - \mu)^2) / n$
- Setelah menuliskan kembali pembilang dalam notasi sigma, Anda mendapatkan $(\Sigma(x_i - \mu)^2)/_n$, yaitu rumus dari varians

Sumber: https://id.wikihow.com/Menghitung-Variasi