Statistika dan Probabilitas

Estimasi Parameter

Deri Siswara

deri.siswara@perbanas.id

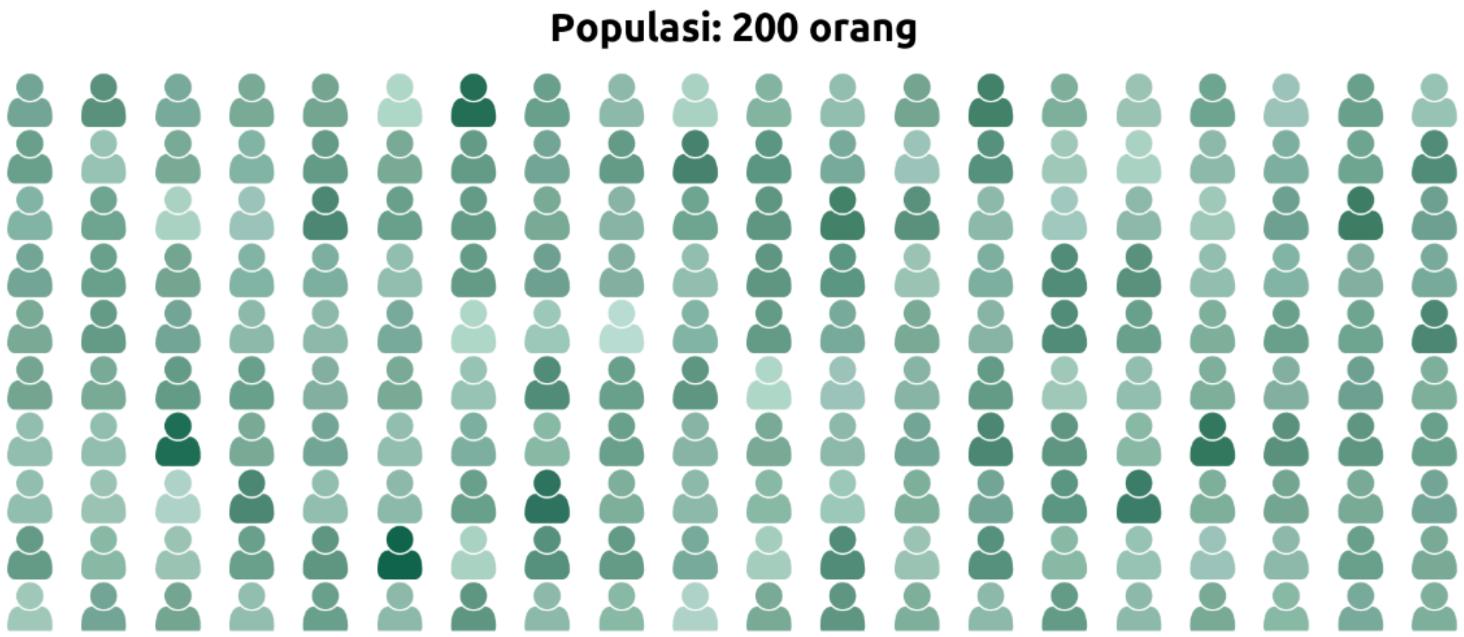
ABFI Institute Perbanas

November 22, 2025

Estimasi Parameter

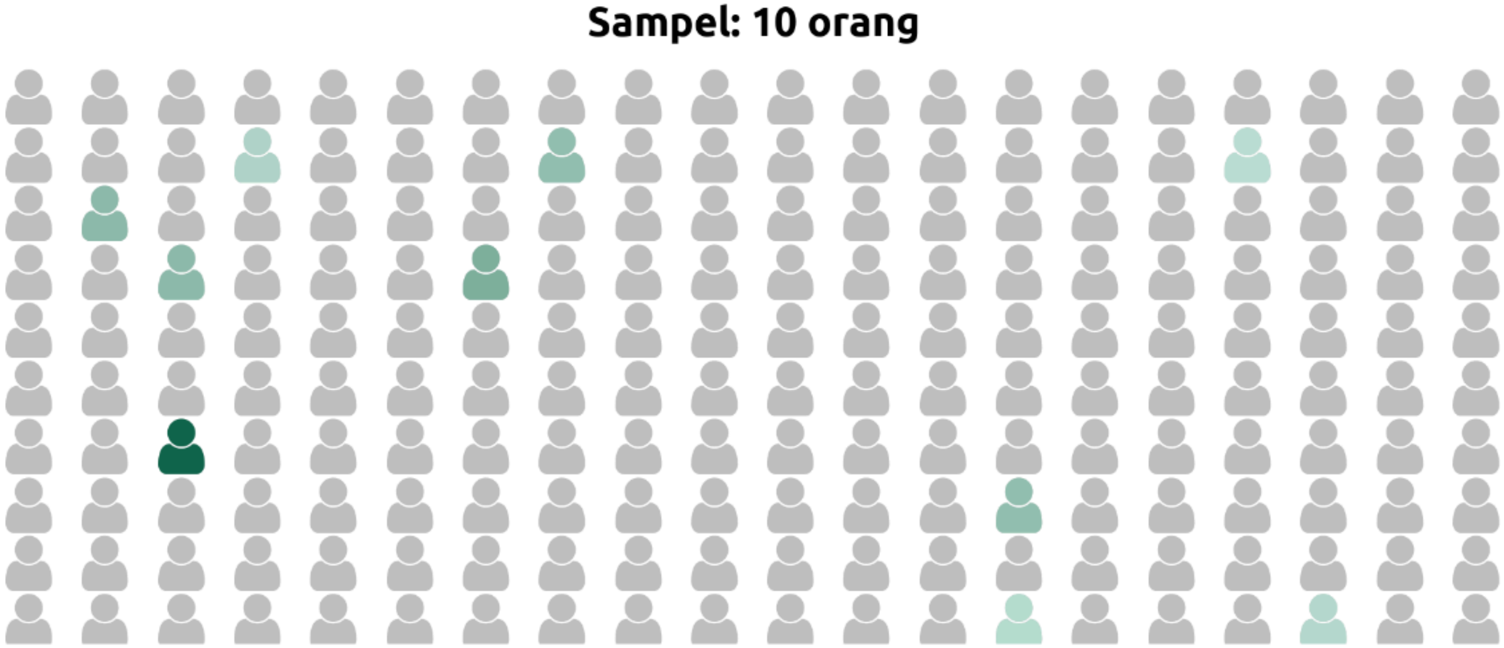
Populasi

Populasi mengacu pada seluruh kelompok individu yang ingin kita tarik kesimpulan tentangnya.

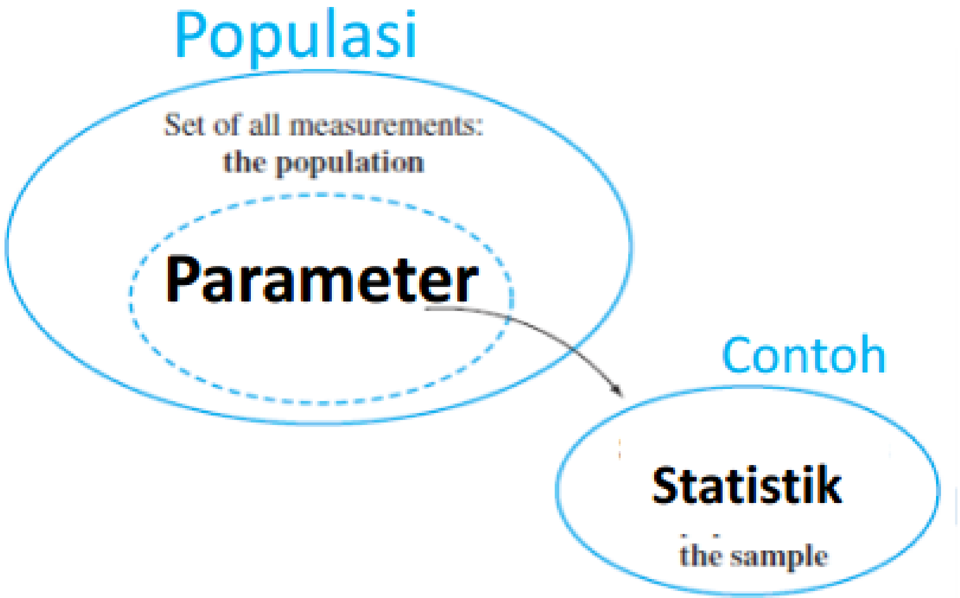


Sampel

Sampel mengacu pada kelompok orang (biasanya lebih kecil) yang telah kita kumpulkan datanya.



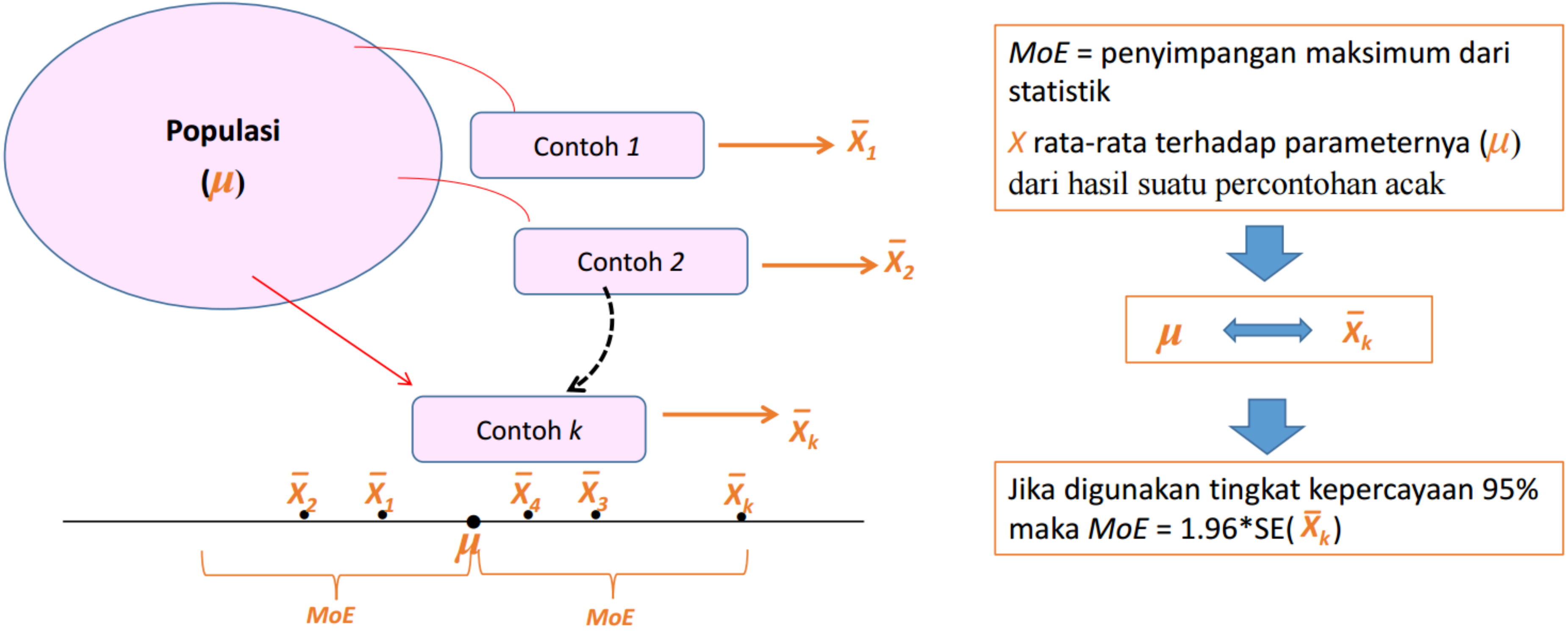
Parameter vs Statistik



	Population *Parameter*	Sample *Statistic*
Mean	μ	$\bar{x}$
Variance	σ^2	s^2
Std. Deviation	σ	s
Size	N	n

Margin of Error (MoE)

- ***Margin of error* (*MoE*) adalah ukuran keragaman hasil dugaan dari satu contoh ke contoh berikutnya.**



Statistik Sampel

Statistik	Notasi Sampel	Formula
Penduga tak bias untuk **Rata-Rata** (μ)	$\bar{x}$	$\bar{x} = \frac{\sum x_i}{n}$
Penduga tak bias untuk **Simpangan Baku** (σ)	s	$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$
Penduga tak bias untuk **Proporsi** (p)	$\hat{p}$	$\hat{p} = \frac{\text{number of successes in sample}}{n}$

Contoh 1: Proporsi Pengguna Internet

Berapa persen mahasiswa yang menggunakan internet lebih dari 3 jam per hari?

Hasil survei menunjukkan bahwa dari 7421 mahasiswa ternyata ada 2998 yang berinternet > 3 jam. Parameter p dengan statistik $\hat{p}$ dapat dihitung sebagai berikut:

$$\hat{p} = \frac{2998}{7421} = 0.4042$$

Contoh 2: Rata-rata Waktu Belajar

Berapa rata-rata waktu (jam) yang dihabiskan mahasiswa untuk belajar per hari?

Dari hasil survei terhadap 10 mahasiswa, diperoleh data berikut: 2, 3, 4, 5, 6, 4, 3, 5, 4, 4

Maka penduga tak bias untuk rata-rata (μ) adalah:

$$\bar{x} = \frac{2+3+4+5+6+4+3+5+4+4}{10} = 4.0$$

Contoh 3: Simpangan Baku Waktu Belajar

Berapa simpangan baku waktu belajar mahasiswa per hari?

Dari data yang sama seperti sebelumnya (rata-rata $\bar{x} = 4$), simpangan baku sampel dihitung dengan rumus:

$$s = \sqrt{\frac{(2-4)^2 + (3-4)^2 + ... + (4-4)^2}{10-1}} = 1.155$$

Latihan 1

Soal 1: Dalam survei terhadap 500 mahasiswa di Universitas X, ditemukan bahwa 60% mahasiswa menggunakan internet lebih dari 3 jam per hari. Apakah 60% tersebut merupakan parameter atau statistik? Jelaskan.

Jawaban

- **Statistik:** 60% $\longrightarrow$ karena dihitung dari **500 mahasiswa (sampel)**
- **Parameter:** proporsi sebenarnya seluruh mahasiswa di Universitas X yang berinternet >3 jam per hari (tidak diketahui, dilambangkan p)

Latihan 1

Soal 2: Dari sampel **50 rumah tangga** di Kota Y, rata-rata pengeluaran listrik per bulan adalah **Rp450.000**. Apakah Rp450.000 tersebut merupakan parameter atau statistik? Jelaskan.

Jawaban

- **Statistik:** Rp450.000 $\longrightarrow$ karena berasal dari **50 rumah tangga (sampel)**
- **Parameter:** rata-rata pengeluaran listrik seluruh rumah tangga di Kota Y (dilambangkan μ)

Latihan 1

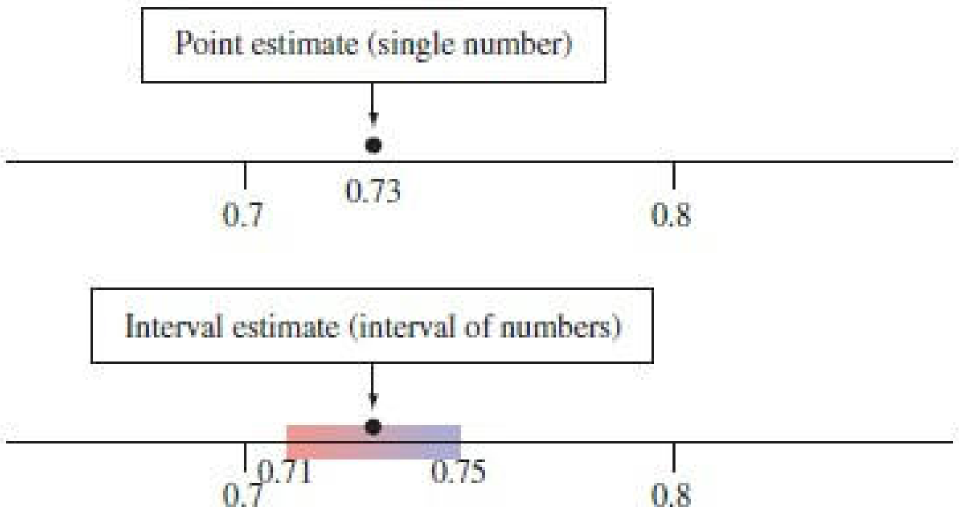
Soal 3: Diketahui bahwa rata-rata tinggi badan seluruh mahasiswa di Indonesia adalah **168 cm**, sedangkan dari survei **100 mahasiswa di kampus A** diperoleh rata-rata **170 cm**. Apakah 168 cm dan 170 cm tersebut merupakan parameter atau statistik? Jelaskan.

Jawaban

- **Parameter:** 168 cm → mewakili **populasi** (seluruh mahasiswa di Indonesia)
- **Statistik:** 170 cm → mewakili **sampel** (100 mahasiswa di kampus A)

Selang Kepercayaan

Selang Kepercayaan (Confidence Interval)



- Kita perlu memberikan penduga selang (**interval estimator bagi parameter**)
- Sebagai contoh: statistik $\bar{x}$ merupakan penduga titik untuk rata-rata dari parameter μ. Karena itu kemungkinan besar $\bar{x} \neq \mu$ walaupun $\bar{x}$ tidak bias terhadap μ.

Definisi Selang Kepercayaan

Definisi

Selang Kepercayaan adalah rentang nilai yang digunakan untuk memperkirakan nilai parameter populasi dengan tingkat kepercayaan (***confidence level***) tertentu.

Formula umum:

$$\text{Selang Kepercayaan} = \text{Penduga} \pm (\text{Nilai kritis}) \times (\text{Galat baku (SE)})$$

Nilai kritis dan galat baku tergantung pada jenis parameter yang diestimasi dan informasi yang tersedia.

Formula Selang Kepercayaan

1️⃣ **Untuk Rata-Rata (μ)** jika σ diketahui

$$\text{CI:} \quad \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

- $\bar{x}$ = rata-rata sampel
- σ = simpangan baku populasi
- $z_{\alpha/2}$ = nilai z pada distribusi normal untuk tingkat kepercayaan tertentu (misal 1.96 untuk 95%)
- n = ukuran sampel

2️⃣ **Untuk Rata-Rata (μ)** jika σ tidak diketahui

$$\text{CI:} \quad \bar{x} \pm t_{\alpha/2, df=n-1} \frac{s}{\sqrt{n}}$$

Formula Selang Kepercayaan

- Gunakan **distribusi t-Student**
- s = simpangan baku sampel
- $t_{\alpha/2,df}$ = nilai kritis t dengan derajat bebas $df = n - 1$

3️⃣ Untuk Proporsi (p)

$$\text{CI:}\quad \hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

- $\hat{p}$ = proporsi sampel
- $z_{\alpha/2}$ = nilai z untuk tingkat kepercayaan tertentu
- n = ukuran sampel

4️⃣ Untuk Simpangan Baku (σ)

Jika ingin memperkirakan simpangan baku populasi berdasarkan sampel:

Formula Selang Kepercayaan

$$\frac{(n-1)s^2}{\chi^2_{(1-\alpha/2)}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{(\alpha/2)}}$$

- Gunakan **distribusi Chi-Square** (χ^2)
- s = simpangan baku sampel
- n = ukuran sampel

Tingkat Signifikansi (α)

α disebut **tingkat signifikansi** (*level of significance*).

i Interpretasi

Besarnya **risiko (peluang)** kita **salah menolak kebenaran populasi** atau **salah memperkirakan parameter**.

Dalam konteks **selang kepercayaan (confidence interval):**

$$\text{Confidence level} = 1 - \alpha$$

📊 **Jika α = 0.05**

- Maka tingkat kepercayaan (confidence level) = **1 - 0.05 = 0.95** $\rightarrow$ **95% confidence**

Tingkat Signifikansi (α)

- Artinya: kita **percaya 95%** bahwa selang yang kita hitung mengandung nilai parameter sebenarnya, dan **ada risiko 5%** (α = 0.05) bahwa selang itu *tidak* memuat parameter tersebut.

Area di Kurva Normal

📉 **Dalam bentuk area di kurva normal:**

- Total area di bawah kurva = 1 (100%)
- Area di tengah (antara $-z$ dan $+z$) = 0.95 (95%)
- Sisa area di dua ekor = α = 0.05 $\longrightarrow$ masing-masing **0.025 di kiri dan kanan**

```
    0.025        0.95        0.025
 |---------|-------------|---------|
-1.96                   +1.96
```

Tabel nilai kritis z untuk beberapa tingkat kepercayaan:

Confidence Level	α (alpha)	Z (dua sisi)
90%	0.10	1.645
95%	0.05	1.960
99%	0.01	2.576

Latihan 2

Soal 1: Dari 7.421 mahasiswa, 2.998 orang melaporkan menggunakan internet > 3 jam/hari. Hitung CI 95% untuk proporsi mahasiswa yang berinternet >3 jam/hari.

$$\hat{p} = \frac{2998}{7421} = 0.4040$$

Galat baku (SE):

$$\mathrm{SE}\ (\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.4040(1-0.4040)}{7421}} \approx 0.00160$$

Nilai kritis untuk 95%: $z_{0.025} = 1.96$

Margin:

$$\mathrm{ME} = 1.96 \times 0.00160 \approx 0.00314$$

Latihan 2

CI 95%:

$$\hat{p} \pm \mathrm{ME} = 0.4040 \pm 0.0031 = (0.3928{,}0.4152)$$

Interpretasi: Dengan 95% kepercayaan, proporsi mahasiswa yang berinternet >3 jam/hari antara **39.28%** dan **41.52%**.

Latihan 2

Soal 2: Sampel 10 mahasiswa memberikan data jam belajar per hari: 2, 3, 4, 5, 6, 4, 3, 5, 4, 4. Hitung CI 95% untuk rata-rata jam belajar.

$$\bar{x} = \frac{2+3+4+5+6+4+3+5+4+4}{10} = 4.0$$

Simpangan baku sampel:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \approx 1.1547$$

Derajat bebas $df = n - 1 = 9$. Nilai kritis $t_{0.025,9} \approx 2.262$

Margin:

Latihan 2

$$\mathrm{ME} = t_{0.025,9} \cdot \frac{s}{\sqrt{n}} = 2.262 \cdot \frac{1.1547}{\sqrt{10}} \approx 0.8260$$

CI 95%:

$$4.0 \pm 0.826 \Rightarrow (3.17, 4.83)$$

Interpretasi: Dengan 95% kepercayaan, rata-rata jam belajar mahasiswa ≈ **3.17 — 4.83 jam/hari**.

Latihan 2

Soal 3: Sampel 100 rumah, rata-rata pengeluaran listrik = Rp50.000, simpangan baku populasi diketahui σ = Rp8.000. Hitung CI 95%.

$$\bar{x} = 50,000, \quad \sigma = 8,000, \quad n = 100$$

SE:

$$\text{SE} = \frac{\sigma}{\sqrt{n}} = \frac{8,000}{10} = 800$$

ME:

$$\text{ME} = 1.96 \times 800 = 1,568$$

CI 95%:

$$50,000 \pm 1,568 \Rightarrow (48,432{,}51,568)$$

Latihan 2

Interpretasi: Dengan 95% kepercayaan, rata-rata pengeluaran listrik antara **Rp48.432** dan **Rp51.568**.