Estimasi Parameter & Interval Kepercayaan

Melangkah dari Statistika Deskriptif menuju Inferensial Modern: Pendugaan Titik, Margin of Error, Selang Kepercayaan (Z vs t-Student), dan Proporsi

Statistika dan Probabilitas · Pertemuan 09 (Pertemuan Pertama Pasca-UTS)

Peta Pertemuan 09 (Gerbang Statistika Inferensial)

1 **Filosofi Estimasi: Pendugaan Titik vs Interval**

Mengapa estimasi titik hampir pasti salah di data riil, sifat penaksir ideal (tak bias, efisien, konsisten).

2 **Interval Kepercayaan Rata-Rata (σ Diketahui — Uji Z)**

Konstruksi selang kepercayaan $1 - \alpha$, nilai kritis $z_{\frac{\alpha}{2}}$, dan anatomi **Margin of Error** (E).

3 **Kondisi Riil: Distribusi t-Student (σ Tidak Diketahui)**

Sejarah William Gosset (Guinness), derajat bebas ($df = n - 1$), dan ketidakpastian ekstra sampel kecil.

4 **Interval Kepercayaan Proporsi & Ukuran Sampel Minimum**

Estimasi data biner/kategorik ($\hat{p}$), penentuan n optimal untuk target presisi rekayasa tertentu.

Blok 1 — Filosofi Estimasi Parameter

Mengapa estimasi titik rentan salah dan bagaimana interval mengukur kepastian

Apa Itu Estimasi Parameter? (Jembatan Inferensial)

Hakikat Estimasi Parameter

Estimasi parameter adalah proses ilmiah menggunakan **statistik sampel** (data terbatas yang kita miliki) untuk menduga nilai **parameter populasi** (nilai sejati yang tidak diketahui).

Statistik Sampel (Diketahui)

- Dihitung dari sampel berukuran n.
- Merupakan peubah acak (bervariasi).
- **Contoh:** $\overline{x}$ (mean), s (deviasi baku), $\hat{p}$ (proporsi).

→ **Pendugaan (Estimasi)**

Parameter Populasi (Target)

- Karakteristik seluruh populasi (N).
- Bernilai tetap/konstan, tetapi misterius.
- **Contoh:** μ (mean), σ (deviasi baku), p (proporsi).

Dua Mazhab Pendugaan: Estimasi Titik (satu angka tunggal) dan Estimasi Interval (rentang nilai beserta tingkat keyakinan).

Efek Ukuran Sampel (n): Semakin Banyak, Semakin Akurat!

Hukum Bilangan Besar & Pemusatan Statistik

Ukuran sampel (n) menentukan seberapa rapat nilai statistik sampel ($\overline{X}$) berkerumun mengelilingi parameter sejati (μ).

Sampel Terlalu Sedikit ($n = 4$)

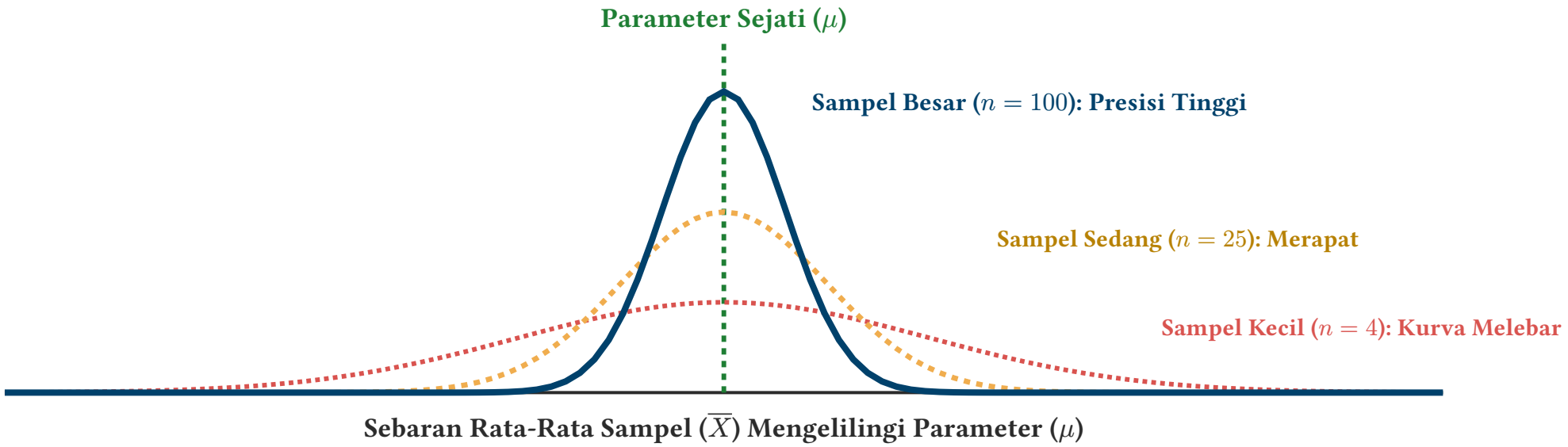
- Rentan terdistorsi oleh data ekstrem (**outlier**).
- Galat baku (SE $= \frac{\sigma}{\sqrt{n}}$) sangat lebar.
- Tebakan statistik sering meleset jauh dari nilai sejati.

Sampel Membesar ($n = 25 \rightarrow 100$)

- Nilai tinggi dan rendah saling meniadakan secara simetris.
- Kurva sebaran statistik mengerucut tajam di sekitar μ.
- **Statistik menjadi penaksir yang semakin handal dan stabil!**

Intuisi: Menanyai 4 orang teman tidak cukup untuk mengetahui preferensi seluruh kampus, tetapi menanyai 100 orang acak sudah memberikan gambaran yang sangat stabil.

Ilustrasi Visual: Efek Ukuran Sampel Terhadap Sebaran Statistik



Dilema Efisiensi: Mengapa Terlalu Banyak Sampel Tidak Efisien?

Hukum Akar Kuadrat (Square Root Law) & Diminishing Returns

Karena faktor galat baku mengandung penyebut $\sqrt{n}$, penurunan galat **tidak berbanding lurus (non-linier)** dengan jumlah sampel:

$$\text{SE} = \frac{\sigma}{\sqrt{n}} \quad \text{artinya:} \quad \text{Potong Galat Jadi } \frac{1}{2} \rightarrow n \text{ Wajib Naik } 4\times$$

Biaya Nyata di Industri

Setiap penambahan satu unit data membutuhkan: • Biaya survei / insentif responden
• Beban query & komputasi server
• Waktu pengujian laboratorium

Titik Jenuh Rekayasa (Economic Optimum)

• Menambah sampel dari $n = 16$ ke $n = 64$ (+48 data) memotong galat hingga 50%.
• Menambah sampel dari $n = 400$ ke $n = 1.000$ (+600 data!) hanya memotong galat sedikit ($< 2\%$). Biaya membengkak tanpa manfaat berarti!

Ilustrasi Visual: Kurva Diminishing Returns Galat Baku



Pemantik Nalar: Tebakan Tunggal vs Jaring Nelayan

PERTANYAAN PEMICU

- Jika seorang data scientist diminta memprediksi rata-rata waktu loading aplikasi baru kita:
- Pernyataan A: “Rata-rata latensi server adalah tepat 245,000 milidetik!”
- Pernyataan B: “Dengan tingkat keyakinan 95%, rata-rata latensi server berada di antara 240 ms hingga 250 ms.”
- Manakah dari kedua pernyataan di atas yang lebih jujur, ilmiah, dan dapat dipertanggungjawabkan secara rekayasa?

Dilema Estimasi Titik (Point Estimation)

Definisi Estimasi Titik

Menggunakan satu nilai tunggal dari sampel ($\overline{x}$ atau $\hat{p}$) sebagai tebakan langsung untuk parameter populasi (μ atau p).

Kelemahan Fatal Estimasi Titik

1. **Peluang Tepat Benar = 0:** Pada variabel kontinu, peluang bahwa rata-rata sampel $\overline{x}$ sama persis dengan parameter populasi sejati μ bernilai nol: $P\left(\overline{X} = \mu\right) = 0$.
2. **Tidak Memuat Informasi Risiko:** Angka tunggal tidak memberikan petunjuk seberapa besar variabilitas data atau seberapa besar kesalahan yang mungkin terjadi (**sampling error**).
3. **Solusi Ilmiah:** Kita membutuhkan **Interval Kepercayaan (Confidence Interval)** yang menyertakan rentang galat.

Tiga Syarat Penaksir Titik yang Ideal (Estimator Quality)

1. Tak Bias (Unbiased)

$$E\left(\hat{\theta}\right) = \theta$$

Rata-rata dari nilai penaksir di seluruh kemungkinan sampel sama persis dengan parameter populasi sejati. Contoh: $E\left(\overline{X}\right) = \mu$.

2. Efisien (Efficient)

$$\mathrm{Var}\left(\hat{\theta}_1\right) < \mathrm{Var}\left(\hat{\theta}_2\right)$$

Di antara penaksir tak bias, penaksir terbaik adalah yang memiliki ragam (varians) terkecil (paling konsisten mengerucut).

3. Konsisten (Consistent)

$$\lim_{n\to\infty} P\left(|\hat{\theta} - \theta| < \varepsilon\right) = 1$$

Ketika ukuran sampel n membesar menuju tak hingga, nilai estimasi dipastikan semakin mendekati parameter sejati.

Konsep Dasar Interval Kepercayaan (Confidence Interval)

Anatomi Selang Kepercayaan

Sebuah rentang nilai yang dihitung dari data sampel sehingga diharapkan mencakup nilai parameter populasi sejati dengan probabilitas tertentu $(1 - \alpha)$:

$$\text{Interval Kepercayaan} = \text{Estimasi Titik} \pm \text{Margin of Error } (E)$$

$$\text{Batas Bawah} = \bar{x} - E \quad \text{dan} \quad \text{Batas Atas} = \bar{x} + E$$

Margin of Error (E) merepresentasikan radius presisi maksimum dari tebakan kita. Semakin kecil nilai E, semakin presisi perkiraan interval yang dihasilkan.

Miskonsepsi Paling Fatal dalam Statistika: Makna 95%

Pernyataan SALAH (Miskonsepsi Umum)

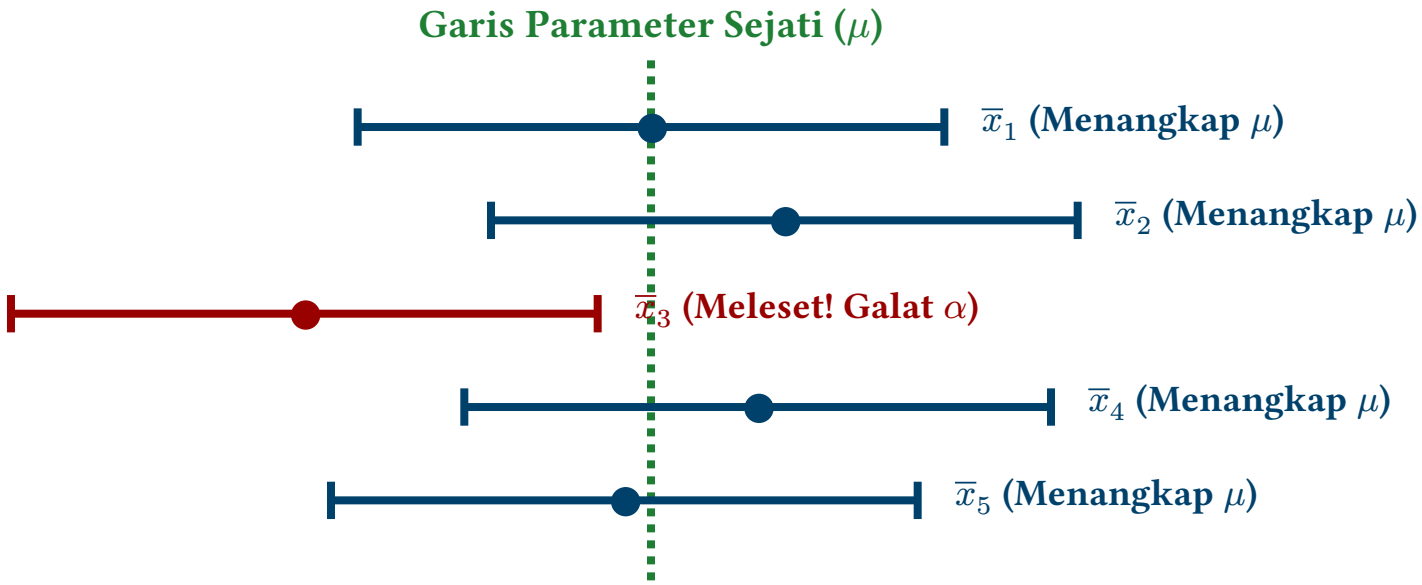
“Ada peluang 95% bahwa rata-rata populasi (μ) berada di antara 240 ms dan 250 ms.”

Mengapa Salah? Parameter populasi μ adalah konstanta tetap (bukan variabel acak). Nilai μ berada di dalam interval tersebut (peluang = 1) atau tidak berada di dalamnya (peluang = 0). Tidak ada probabilitas 95% untuk parameter tetap!

Pernyataan BENAR (Interpretasi Frequentist yang Sah)

“Jika kita mengambil 100 sampel acak berbeda dan membuat 100 selang kepercayaan 95%, maka sekitar 95 selang di antaranya akan menangkap nilai parameter μ yang sebenarnya.”

Visualisasi: Jaring Nelayan Menangkap Ikan Parameter



Inti Filosofis 95% Confidence: Nilai parameter μ (garis hijau) bernilai tetap dan tidak bergerak. Selang kepercayaan (jaring) kitalah yang berubah posisinya pada setiap sampel acak. Jika percobaan diulang 100 kali, sekitar 95 interval akan menangkap μ dan 5 interval meleset.

Studi Kasus Blok 1: Estimasi Waktu Respon Server Cloud

SKENARIO ENGINEERING SLA

Sebuah startup cloud computing ingin mempublikasikan laporan performa microservices kepada klien perbankan. Tim DevOps mengambil sampel acak $n = 100$ request pada jam sibuk dan memperoleh rata-rata waktu respon $\overline{x} = 142$ ms.

Manajer Pemasaran mengusulkan membuat klaim promosi: **"Rata-rata waktu respon server kami pasti 142 ms"**. Namun, tim Lead SRE menolak klaim tersebut dan meminta diganti menjadi format selang kepercayaan 95%: 138 ms s.d. 146 ms.

Pertanyaan Diskusi

1. Mengapa usulan Manajer Pemasaran dianggap menyesatkan secara statistika dan hukum kontrak SLA?
2. Apa arti operasional bagi klien jika sistem dijamin memiliki selang kepercayaan 95% sebesar $[138, 146]$ ms?

Pembahasan Kasus Blok 1: Keandalan Kontrak SLA Berbasis CI

1. Bahaya Klaim Tunggal (Point Estimate)

- Angka 142 ms hanyalah statistik dari satu sampel 100 request. Pada sampel 100 request berikutnya, rata-ratanya bisa bernilai 145 ms atau 139 ms akibat fluktuasi acak sampling.
- Jika klien mengaudit sistem pada detik lain dan mendapati rata-rata 144 ms, startup dapat dituntut wanprestasi karena melanggar garansi tunggal 142 ms!

2. Keunggulan Kontrak Berbasis Confidence Interval

- Selang $[138, 146]$ ms secara jujur memperhitungkan **Margin of Error** (± 4 ms).
- Klien memahami bahwa performa stabil di rentang aman SLA (misal SLA mensyaratkan < 150 ms). Seluruh rentang kepercayaan berada di bawah 150 ms, menjamin kepatuhan sistem secara ilmiah.

Blok 2 — Selang Kepercayaan Rata-Rata (σ Diketahui)

Penerapan Kurva Normal Baku (Z), Tingkat Signifikansi, dan Margin of Error

Formula Selang Kepercayaan Z untuk Rata-Rata (μ)

Syarat Penerapan Distribusi Z

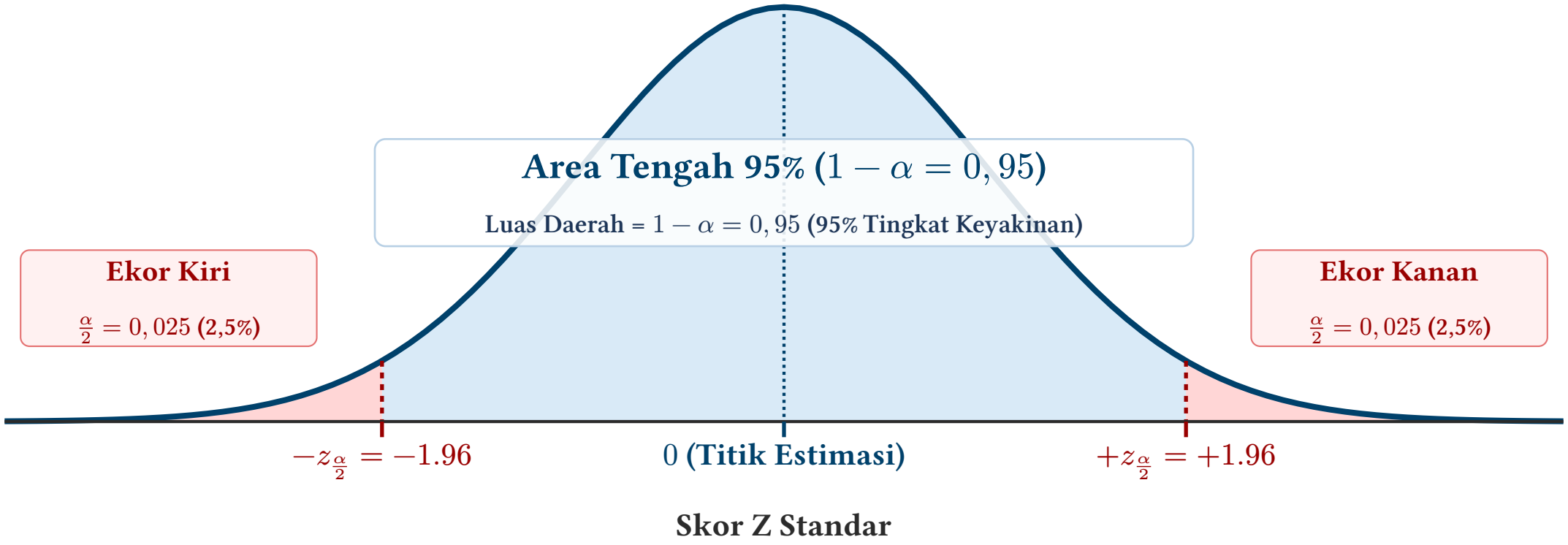
1. Simpangan baku populasi (σ) **diketahui secara pasti** (misal dari riwayat pabrik bertahun-tahun).
2. Populasi berdistribusi normal, ATAU ukuran sampel cukup besar ($n \geq 30$) berdasarkan CLT.

Formula Matematika Selang Kepercayaan

$$\overline{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\text{Margin of Error } (E) = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Visualisasi Distribusi Normal Selang Kepercayaan



Tiga Tingkat Keyakinan (Confidence Level) Paling Populer

Tingkat signifikansi α adalah peluang selang kita gagal menangkap parameter ($\alpha = 1 - \text{Confidence Level}$):

Tingkat Kepercayaan	Signifikansi (α)	Luas Tiap Ekor ($\frac{\alpha}{2}$)	Nilai Kritis ($z_{\frac{\alpha}{2}}$)
90% (0,90)	0,10	0,05	**1,645**
95% (0,95)	0,05	0,025	**1,960** (Standar Sains)
99% (0,99)	0,01	0,005	**2,576** (Standar Medis/Aero)

Perhatikan Keterkaitan: Semakin tinggi keyakinan yang kita tuntut ($90\% \to 99\%$), nilai $z_{\frac{\alpha}{2}}$ membesar ($1,645 \to 2,576$), sehingga rentang interval menjadi semakin lebar!

Trade-off Fundamental: Presisi vs Kepastian

Ingin Keyakinan Sangat Tinggi (99% / 99,9%)

- $z_{\frac{\alpha}{2}}$ sangat besar $(2,576)$.
- **Margin of Error** melebar tajam.
- Interval menjadi sangat longgar (kurang presisi).
- **Analogi:** "Saya 100% yakin nilai UTS Anda berada di antara 0 dan 100!" (Pasti benar, tetapi tidak berguna).

Ingin Presisi Sangat Ketat (Interval Ramping)

- Interval sempit menuntut $z_{\frac{\alpha}{2}}$ kecil.
- Tingkat keyakinan turun drastis (misal hanya 80% atau 70%).
- Risiko meleset dari parameter sejati meningkat pesat!

Satu-satunya cara meningkatkan presisi (mempersempit selang) TANPA menurunkan tingkat keyakinan adalah dengan MEMPERBESAR UKURAN SAMPEL (n).

Studi Kasus Blok 2: Kalibrasi Tegangan Baterai Mobil Listrik

SKENARIO PENGUJIAN MUTU MANUFAKTUR OTOMOTIF

Sebuah lini perakitan baterai kendaraan listrik (EV) memiliki standar proses dengan simpangan baku tegangan sel yang diketahui stabil pada $\sigma = 0{,}15$ Volt.

Dari batch produksi terbaru, insinyur QC mengambil sampel acak $n = 36$ unit sel baterai dan mendapati rata-rata tegangan sampel $\overline{x} = 3{,}72$ Volt.

Pertanyaan Diskusi

1. Konstruksikan selang kepercayaan 95% untuk rata-rata tegangan sel baterai populasi (μ)!
2. Jika regulasi keselamatan baterai mensyaratkan tegangan rata-rata harus berada di atas “3,65” Volt, apakah batch ini memenuhi syarat regulasi?

Pembahasan Kasus Blok 2: Perhitungan Selang Z 95%

1. Perhitungan Margin of Error & Selang Kepercayaan

Diketahui: $\overline{x} = 3{,}72$, $\sigma = 0{,}15$, $n = 36$, $z_{0{,}025} = 1{,}96$.

$$E = z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 1{,}96 \times \frac{0{,}15}{\sqrt{36}} = 1{,}96 \times \frac{0{,}15}{6} = 1{,}96 \times 0{,}025 = 0{,}049 \,\text{Volt}$$

$$\text{CI}_{95\%} = [\overline{x} - E, \overline{x} + E] = [3{,}72 - 0{,}049, 3{,}72 + 0{,}049] = [3{,}671, 3{,}769] \,\text{Volt}$$

2. Keputusan Kepatuhan Regulasi

• Batas bawah selang kepercayaan 95% adalah 3,671 Volt.

• Karena seluruh rentang (3,671 s.d. 3,769) berada di atas ambang regulasi 3,65 Volt, insinyur dapat menyatakan dengan keyakinan 95% bahwa seluruh batch produksi mematuhi regulasi keselamatan!

Blok 3 — Distribusi t-Student (σ Tidak Diketahui)

Realitas riset modern: Mengestimasi parameter ketika simpangan baku populasi misterius

Realitas Dunia Nyata: Mengapa σ Jarang Diketahui?

Paradoks Asumsi Sigma

Pikirkan secara kritis: Jika kita bahkan **tidak mengetahui rata-rata populasi** (μ) dan sedang berusaha menebaknya...

Bagaimana mungkin kita bisa mengetahui simpangan baku populasi (σ)?

Di 99% penelitian rekayasa dan sains modern, nilai σ **tidak pernah diketahui**. Kita terpaksa mengganti σ dengan simpangan baku sampel (s).

Sejarah Penemuan Distribusi t: William Sealy Gosset

Kisah Pabrik Bir Guinness (1908)

• William S. Gosset bekerja sebagai ahli kimia di pabrik bir Guinness di Dublin, Irlandia.
• Ia bereksperimen dengan batch ragi kecil ($n = 4$ atau $n = 8$).
• Ketika ia mengganti σ dengan s pada sampel kecil, kurva Normal Gauss **gagal total**—terlalu banyak kesalahan estimasi!
• Guinness melarang karyawannya mempublikasikan rahasia dagang, sehingga Gosset mempublikasikan temuannya di jurnal **Biometrika** dengan nama samaran: **"Student"**.

Bentuk Kurva t-Student

• Bentuk simetris lonceng mirip kurva normal, namun memiliki **ekor yang lebih tebal (heavier tails)**.
• Ekor tebal mengakomodasi **ketidakpastian ekstra** karena kita menaksir σ menggunakan s.
• Bentuk kurva dikendalikan oleh parameter **Derajat Bebas (Degrees of Freedom)**:

$$df = n - 1$$

Sifat-Sifat Kunci Distribusi t-Student

Karakteristik Matematis Distribusi t

1. Berpusat pada $t = 0$ dan berbentuk simetris seperti distribusi normal baku.
2. Ragam (varians) distribusi t bernilai lebih besar dari 1: $\text{Var}(T) = \frac{df}{df-2} > 1$ (karena variabilitas s).
3. Nilai kritis $t_{\frac{\alpha}{2}}$ selalu **lebih besar** daripada nilai kritis $z_{\frac{\alpha}{2}}$ untuk tingkat kepercayaan yang sama.
4. **Konvergensi Asimtotik:** Ketika ukuran sampel membesar ($n \to \infty$ atau $df \geq 30$), kurva distribusi t semakin mengerucut dan secara matematis menjadi identik persis dengan Kurva Normal Baku Z!

Formula Selang Kepercayaan t (σ Tidak Diketahui)

Rumus Estimasi Rata-Rata Sampel Kecil / Menengah

$$\overline{x} - t_{\frac{\alpha}{2}, df} \frac{s}{\sqrt{n}} \leq \mu \leq \overline{x} + t_{\frac{\alpha}{2}, df} \frac{s}{\sqrt{n}}$$

$$\text{Margin of Error } (E) = t_{\frac{\alpha}{2}, df} \frac{s}{\sqrt{n}} \quad \text{dengan} \quad df = n - 1$$

Perhatikan Dua Penggantian Krusial:

1. **Simpangan baku populasi σ digantikan oleh simpangan baku sampel s.**
2. **Nilai kritis normal $z_{\frac{\alpha}{2}}$ digantikan oleh nilai kritis tabel t pada derajat bebas $df = n - 1$.**

Studi Kasus Blok 3: Uji Waktu Booting Sistem Operasi Tertanam

SKENARIO FIRMWARE IOT / EMBEDDED SYSTEMS

Seorang perekayasa perangkat tertanam (**embedded system**) menguji optimasi kernel Linux pada perangkat router pintar baru. Karena biaya papan uji purwarupa sangat mahal, perekayasa hanya melakukan $n = 16$ kali uji **booting** independen.

Data waktu booting menghasilkan rata-rata sampel $\overline{x} = 12{,}40$ detik dengan simpangan baku sampel $s = 1{,}20$ detik. Waktu booting diasumsikan berdistribusi normal. Diketahui nilai tabel untuk $df = 15$: $t_{0{,}025,15} = 2{,}131$ dan $z_{0{,}025} = 1{,}960$.

Pertanyaan Diskusi

1. Mengapa insinyur wajib menggunakan distribusi t dan dilarang menggunakan tabel Z?
2. Hitung selang kepercayaan 95% untuk rata-rata waktu booting populasi router tersebut!

Pembahasan Kasus Blok 3: Perhitungan Selang t 95%

1. Alasan Ilmiah Pemilihan Distribusi t

• Nilai σ populasi tidak diketahui. Yang dimiliki hanyalah s dari 16 sampel.

• Ukuran sampel tergolong kecil ($n = 16 < 30$). Menggunakan Z akan meremehkan galat (**underestimating margin of error**) dan memberikan kepalsuan presisi.

2. Perhitungan Derajat Bebas & Selang Kepercayaan

Derajat bebas: $df = n - 1 = 16 - 1 = 15$. Nilai kritis tabel: $t_{0{,}025,15} = 2{,}131$.

$$E = t_{\frac{\alpha}{2},\,15} \times \frac{s}{\sqrt{n}} = 2{,}131 \times \frac{1{,}20}{\sqrt{16}} = 2{,}131 \times 0{,}30 = 0{,}639 \,\text{detik}$$

$$\text{CI}_{95\%} = [12{,}40 - 0{,}639,\, 12{,}40 + 0{,}639] = [11{,}761,\, 13{,}039] \,\text{detik}$$

Bandingkan: Jika nekat memakai Z, margin galat bernilai $E = 1{,}96 \times 0{,}30 = 0{,}588$ detik (terlalu sempit). Distribusi t memberikan selang lebih lebar ($\pm$ 0,639) untuk mengompensasi ketidakpastian estimasi s.

Blok 4 — Selang Kepercayaan Proporsi & Ukuran Sampel

Estimasi data biner/kategorik dan perencanaan ukuran sampel presisi di industri

Estimasi Proporsi Populasi (p) pada Data Biner

Karakteristik Data Kategorik (Sukses / Gagal)

Jika kita mengukur proporsi sampel: $\hat{p} = \frac{X}{n}$ (di mana X jumlah kejadian sukses):

- Rata-rata estimasi titik: $E(\hat{p}) = p$
- Kesalahan standar proporsi: $\text{SE} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Formula Selang Kepercayaan Proporsi (Pendekatan Normal Walaupun σ Tidak Ada)

$$\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Syarat Kelayakan Pendekatan Normal: Wajib memenuhi $n\hat{p} \geq 10$ dan $n(1-\hat{p}) \geq 10$.

Penentuan Ukuran Sampel Minimum untuk Mengestimasi Rata-Rata

Menjawab Pertanyaan Bisnis Terpenting: "Berapa Banyak Data yang Harus Diambil?"

Dalam rekayasa data, mengambil sampel berlebih membuang anggaran biaya komputasi/survei, sedangkan sampel terlalu sedikit membuat selang terlalu lebar (tidak berguna).

Penurunan Formula Ukuran Sampel (n) Rata-Rata

$$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad \rightarrow \quad \sqrt{n} = \frac{z_{\frac{\alpha}{2}}\, \sigma}{E} \quad \rightarrow \quad n = \left(\frac{z_{\frac{\alpha}{2}}\, \sigma}{E}\right)^2$$

Aturan Pembulatan: Nilai n selalu **dibulatkan ke atas (ceiling)** ke bilangan bulat terdekat agar nilai galat E tidak melebihi target toleransi!

Penentuan Ukuran Sampel Minimum untuk Proporsi

Dua Skenario Perhitungan Ukuran Sampel Proporsi

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{E}\right)^2 p^*(1-p^*)$$

Skenario A: Ada Estimasi Awal (p^*)

Gunakan nilai dugaan p^* dari penelitian pendahulu atau data historis (misal riset bulan lalu mencatat konversi $p^* = 0{,}15$).

Skenario B: Konservatif / Worst-Case

Jika sama sekali tidak ada informasi tentang p, asumsikan skenario paling variatif:

$$p^* = 0{,}50$$

Perkalian $p^*(1-p^*) = 0{,}25$ adalah nilai maksimum mutlak!

Studi Kasus 4A: Ukuran Sampel A/B Testing E-Commerce

SKENARIO PENGUJIAN KONVERSI PERTUMBUHAN PENGGUNA

Tim Growth Engineering pada platform e-commerce berencana meluncurkan varian banner promo baru. Mereka ingin menduga proporsi konversi klik pengguna (p).

Batasan & Target Desain Percobaan:

- **Margin of Error** yang dapat ditoleransi maksimal $E = 0{,}02$ (2%)
- Tingkat keyakinan yang diinginkan sebesar 95% ($z_{\frac{\alpha}{2}} = 1{,}96$)
- Tidak ada data riwayat atau estimasi awal mengenai tingkat konversi (p^* tidak diketahui)

Pertanyaan Diskusi Rekayasa

1. Pendekatan apa yang harus diambil jika nilai p^* sama sekali tidak diketahui?
2. Berapa jumlah minimum pengunjung (n) yang harus dialokasikan ke varian baru agar margin galat tidak melampaui 2%?

Pembahasan Kasus 4A: Komputasi Ukuran Sampel Konservatif

1. Pendekatan Konservatif (Worst-Case Scenario)

Karena nilai p^* tidak diketahui, kita memakai asumsi variabilitas maksimum $p^* = 0{,}50$ sehingga $p^*(1 - p^*) = 0{,}25$. Ini menjamin ukuran sampel yang dihasilkan aman untuk nilai proporsi berapa pun nantinya.

2. Perhitungan Formula Ukuran Sampel (n)

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{E}\right)^2 p^*(1 - p^*) = \left(\frac{1{,}96}{0{,}02}\right)^2 \times 0{,}50(1 - 0{,}50)$$

$$n = (98)^2 \times 0{,}25 = 9.604 \times 0{,}25 = 2.401 \text{ pengunjung}$$

Keputusan Praktis Tim Data: Tim wajib mengarahkan minimal 2.401 pengunjung unik ke varian banner promo. Jika alokasi lalu lintas di bawah angka ini, margin kesalahan estimasi akan membengkak di atas 2%.

Studi Kasus 4B: Audit Akurasi Model AI (OCR Document)

SKENARIO EVALUASI MODEL MACHINE LEARNING

Sebuah sistem otomatisasi perbankan menggunakan model AI untuk membaca dokumen transaksi (OCR). Untuk mengevaluasi keandalan model sebelum rilis penuh, tim MLOps menguji sampel acak sebanyak $n = 500$ dokumen faktur.

Hasil Audit Sampel:

- Ditemukan $X = 35$ dokumen mengalami kegagalan baca (salah karakter)
- Proporsi kesalahan sampel: $\hat{p} = \frac{35}{500} = 0{,}07$ (7%)
- Tingkat keyakinan yang ditetapkan manajemen adalah 95% ($z_{\frac{\alpha}{2}} = 1{,}96$)

Pertanyaan Diskusi Rekayasa

1. Apakah sampel memenuhi syarat kelayakan pendekatan distribusi normal?
2. Berapakah selang kepercayaan 95% untuk **error rate** populasi model OCR tersebut?

Pembahasan Kasus 4B: Selang Kepercayaan Proporsi Error AI

1. Uji Kelayakan Pendekatan Normal

- Jumlah sukses (gagal baca): $n\hat{p} = 500 \times 0{,}07 = 35 \geq 10$ **(Terpenuhi)**
- Jumlah gagal (sukses baca): $n(1-\hat{p}) = 500 \times 0{,}93 = 465 \geq 10$ **(Terpenuhi)**

-> Pendekatan Kurva Normal Baku (Z) valid dan sah digunakan.

2. Perhitungan Margin of Error & Interval Kepercayaan

$$E = 1{,}96 \times \sqrt{\frac{0{,}07 \times 0{,}93}{500}} = 1{,}96 \times \sqrt{0{,}0001302} = 1{,}96 \times 0{,}01141 = 0{,}0224\,(2{,}24\%)$$

$$\text{CI}_{95\%} = [0{,}07 - 0{,}0224,\, 0{,}07 + 0{,}0224] = [0{,}0476,\, 0{,}0924] \quad \rightarrow \quad [4,76\%,\, 9,24\%]$$

Kesimpulan Audit: Kita yakin 95% bahwa tingkat kesalahan model AI di seluruh populasi dokumen berada di antara **4,76%** hingga **9,24%**. Jika SLA bank mensyaratkan tingkat error $< 5\%$, model **belum siap rilis** karena batas atasnya $(9,24\%)$ masih jauh melampaui batas toleransi.

Penutup & Jembatan ke Uji Hipotesis

Keterkaitan erat Confidence Interval menuju Pengujian Hipotesis Formal (Pertemuan 10)

Rekapitulasi Pohon Keputusan Estimasi Parameter

Parameter Sasaran	Kondisi Varians	Distribusi Kritis	Formula Margin of Error (E)
Rata-rata (μ)	σ Diketahui	**Distribusi Z**	$E = z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$
Rata-rata (μ)	σ Tidak Diketahui	**Distribusi t** ($df = n - 1$)	$E = t_{\frac{\alpha}{2}, df} \frac{s}{\sqrt{n}}$
Proporsi (p)	Data Biner (0/1)	**Distribusi Z**	$E = z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Jembatan ke Pertemuan 10 (Uji Hipotesis): Jika selang kepercayaan 95% untuk waktu respon adalah $[138, 146]$ ms, maka klaim hipotesis nol $H_0 : \mu = 150$ ms dapat langsung kita TOLAK pada tingkat signifikansi 5% karena angka 150 berada di luar selang kepercayaan!