Kumpulan Soal & Pembahasan: Sampling & Distribusi Sampling

Latihan Terstruktur: Standard Error, Finite Population Correction (FPC), Teorema Limit Pusat (CLT), dan Distribusi Sampling Proporsi

Statistika dan Probabilitas · Pertemuan 07

Peta Topik Latihan Soal & Pembahasan

Bagian A: Standard Error & Ukuran Sampel

- **Soal 1:** Perhitungan SE dan Efek Lipat Ganda Sampel
- **Soal 2:** Ukuran Sampel Minimum untuk Target Presisi

Bagian B: Populasi Normal

- **Soal 3:** Peluang Individu vs Rata-Rata Kelompok
- **Soal 4:** Kapasitas Beban Lift (Keselamatan Rekayasa)

Bagian C: Teorema Limit Pusat (CLT)

- **Soal 5:** Durasi Panggilan Miring (**Right-Skewed**)
- **Soal 6:** Eksperimen Rata-Rata Pengundian Dadu ($n = 50$)

Bagian D & E: FPC & Proporsi

- **Soal 7:** Uji Batch Terbatas dengan Faktor FPC
- **Soal 8:** Rasio Konversi E-Commerce
- **Soal 9:** Survei Elektabilitas Pemilu (**Quick Count**)
- **Soal 10:** Uji Kontrol Cacat Mutu Semikonduktor

Bagian A — Standard Error & Ukuran Sampel

Karakteristik galat baku dan reduksi variabilitas sebanding $\sqrt{n}$

Soal 1: Perhitungan Standard Error dan Pengaruh Ukuran Sampel

KASUS

Gaji awal bulanan lulusan Program Studi Teknologi Informasi memiliki rata-rata populasi $\mu = 8{,}5$ juta rupiah dengan simpangan baku populasi $\sigma = 2{,}4$ juta rupiah.

Pertanyaan

1. Hitung Kesalahan Standar (**Standard Error** / $\sigma_{\overline{X}}$) dari rata-rata sampel jika diambil sampel acak berukuran:

- (a) $n = 16$ orang
- (b) $n = 64$ orang
- (c) $n = 400$ orang

2. Berapa kali lipat ukuran sampel harus diperbesar jika kita ingin memotong nilai **Standard Error** menjadi **sepertiganya** ($\frac{1}{3}$)?

Pembahasan 1 (Bagian 1): Standard Error untuk $n = 16$ dan $n = 64$

Rumus Dasar Standard Error: $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

Diketahui simpangan baku populasi: $\sigma = 2{,}4$ juta rupiah.

Kasus (a): Untuk ukuran sampel $n = 16$ orang

$$\sigma_{\overline{X}} = \frac{2{,}4}{\sqrt{16}} = \frac{2{,}4}{4} = \text{“0,60“ “juta rupiah“ (Rp 600.000,-)}$$

Kasus (b): Untuk ukuran sampel $n = 64$ orang

$$\sigma_{\overline{X}} = \frac{2{,}4}{\sqrt{64}} = \frac{2{,}4}{8} = \text{“0,30“ “juta rupiah“ (Rp 300.000,-)}$$

Pembahasan 1 (Bagian 2): Standard Error untuk $n = 400$ & Aturan Pangkat

Kasus (c): Untuk ukuran sampel $n = 400$ orang

$$\sigma_{\overline{X}} = \frac{2{,}4}{\sqrt{400}} = \frac{2{,}4}{20} = \text{“0,12“ “juta rupiah“ (Rp 120.000,-)}$$

Pertanyaan 2: Berapa kali lipat sampel harus diperbesar?

Karena $\sigma_{\overline{X}} \propto \frac{1}{\sqrt{n}}$, agar SE menyusut menjadi $\frac{1}{3}$ kali semula:

$$\frac{1}{\sqrt{n_{\text{baru}}}} = \frac{1}{3} \times \frac{1}{\sqrt{n_{\text{lama}}}} \Longrightarrow \sqrt{n_{\text{baru}}} = 3\sqrt{n_{\text{lama}}}$$

Kuadratkan kedua sisi:

$$n_{\text{baru}} = 3^2 \times n_{\text{lama}} = 9 \times n_{\text{lama}}$$

Ukuran sampel harus dilipatgandakan sebanyak **9 kali lipat**!

Soal 2: Penentuan Ukuran Sampel Minimum untuk Target Presisi

KASUS

Sebuah tim **Quality Assurance** pada pabrik baterai kendaraan listrik ingin mengestimasi daya tahan baterai. Diketahui simpangan baku populasi dari daya tahan baterai adalah $\sigma = 18$ bulan.

Pertanyaan

Berapa jumlah sampel baterai minimum (n) yang harus diuji secara acak agar Kesalahan Standar rata-rata (**Standard Error of the Mean**) tidak melebihi **2 bulan** ($\sigma_{\overline{X}} \leq 2$)?

Pembahasan 2: Inversi Rumus Standard Error

Solusi Langkah demi Langkah

Langkah 1: Tuliskan pertidaksamaan target presisi:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \leq 2$$

Langkah 2: Substitusikan $\sigma = 18$ bulan:

$$\frac{18}{\sqrt{n}} \leq 2$$

Langkah 3: Aljabar pertidaksamaan kuadrat:

$$18 \leq 2\sqrt{n} \Longrightarrow \sqrt{n} \geq \frac{18}{2} = 9$$

Kuadratkan kedua ruas:

$$n \geq 9^2 = 81 \text{ unit baterai}$$

Kesimpulan Rekayasa: Tim QA wajib menguji sekurang-kurangnya **81 unit baterai**.

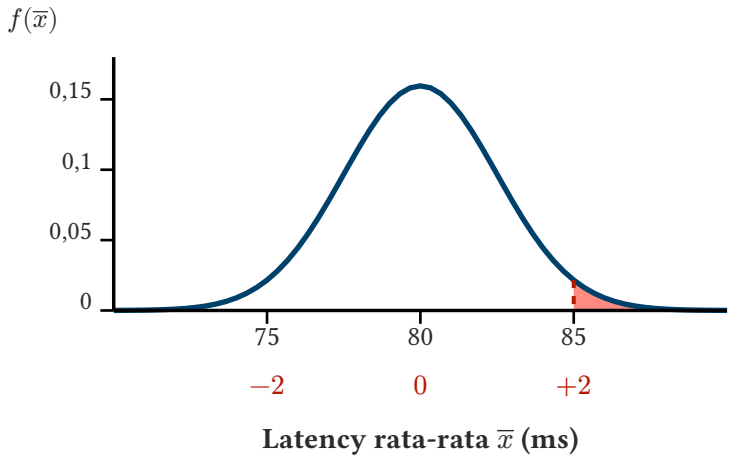
Bagian B — Distribusi Sampling Populasi Normal

Membandingkan risiko individu versus risiko rata-rata kelompok

Soal 3: Peluang Individu vs Peluang Rata-Rata Kelompok

KASUS

Waktu respons (**latency**) eksekusi kueri pada server database berdistribusi normal dengan rata-rata $\mu = 80$ ms dan simpangan baku $\sigma = 15$ ms.



Pertanyaan

1. Berapa peluang **satu kueri tunggal yang dipilih acak** memiliki latensi > 85 ms ($P(X > 85)$)?

2. Jika diambil sampel acak berukuran $n = 36$ kueri, berapa peluang **rata-rata latensi sampel ($\overline{X}$)** melebihi 85 ms ($P\left(\overline{X} > 85\right)$)?

3. Mengapa kedua peluang tersebut sangat berbeda?

Pembahasan 3 (Bagian 1): Peluang Kueri Tunggal ($n = 1$)

Menghitung Peluang Individu ($n = 1$)

Pada tingkat kueri tunggal, kita menggunakan rumus Z-Score data individu:

$$Z = \frac{X - \mu}{\sigma} = \frac{85 - 80}{15} = \frac{5}{15} = +0{,}33$$

Baca nilai dari tabel distribusi normal kumulatif:

$$\Phi(0{,}33) = 0{,}6293$$

Peluang ekor kanan (melebihi 85 ms):

$$P(X > 85) = 1 - 0{,}6293 = 0{,}3707 \quad (37{,}07\%)$$

Makna: Lebih dari sepertiga (37%) kueri individu beroperasi di atas 85 ms.

Pembahasan 3 (Bagian 2): Peluang Rata-Rata Kelompok ($n = 36$)

Menghitung Peluang Rata-Rata Sampel ($n = 36$)

Langkah 1: Hitung Standard Error terlebih dahulu:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{36}} = \frac{15}{6} = 2{,}5 \text{ ms}$$

Langkah 2: Standarisasi Skor Z untuk $\overline{X} = 85$:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{85 - 80}{2{,}5} = \frac{5}{2{,}5} = +2{,}00$$

Langkah 3: Hitung Peluang:

$$P\left(\overline{X} > 85\right) = 1 - \Phi(2{,}00) = 1 - 0{,}9772 = 0{,}0228 \quad (2{,}28\%)$$

Refleksi: Satu kueri lambat adalah hal biasa (37%). Namun rata-rata 36 kueri bersama-sama lambat adalah peristiwa sangat langka (2,28%) karena kueri cepat dan lambat saling meratakan!

Soal 4: Kapasitas Beban Maksimum Lift (Kasus Rekayasa)

KASUS

Sebuah lift gedung kantor dirancang menampung 16 orang dewasa dengan batas aman total beban 1.184 kg. Ini berarti batas beban rata-rata per penumpang adalah $\frac{1.184}{16} = 74$ kg. Berat badan karyawan berdistribusi normal dengan $\mu = 70$ kg dan $\sigma = 8$ kg.

Pertanyaan

Jika 16 karyawan acak masuk bersamaan ke dalam lift, berapakah probabilitas lift mengalami kelebihan beban (**overload**), yaitu berat rata-rata penumpang melebihi 74 kg ($P\left(\overline{X} > 74\right)$)?

Pembahasan 4: Evaluasi Keamanan Rekayasa via Distribusi Sampling

Solusi Langkah demi Langkah

1. Parameter Distribusi Sampling: Karena populasi asal berdistribusi normal, rata-rata sampel $\overline{X}$ juga berdistribusi normal tepat:

$$\mu_{\overline{X}} = 70 \text{ kg} \quad \text{dan} \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{8}{\sqrt{16}} = \frac{8}{4} = 2{,}0 \text{ kg}$$

2. Standarisasi Skor Z:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{74 - 70}{2{,}0} = \frac{4}{2{,}0} = +2{,}00$$

3. Menghitung Peluang Overload:

$$P\left(\overline{X} > 74\right) = 1 - \Phi(2{,}00) = 1 - 0{,}9772 = 0{,}0228 \quad (2{,}28\%)$$

Kesimpulan: Hanya ada risiko sekitar **2,28%** lift mengalami kelebihan beban saat diisi penuh 16 orang dewasa acak.

Bagian C — Teorema Limit Pusat (CLT)

Penerapan konvergensi normal pada populasi yang tidak normal

Soal 5: Waktu Layanan Panggilan Berdistribusi Miring

KASUS

Durasi panggilan pelanggan pada pusat layanan bantuan (**call center**) memiliki distribusi sangat miring ke kanan (**right-skewed**), dengan rata-rata $\mu = 12$ menit dan simpangan baku $\sigma = 4$ menit.

Pertanyaan

Seorang supervisor memantau sampel acak berukuran $n = 64$ panggilan.

1. Mengapa kita dapat menggunakan Distribusi Normal untuk menganalisis rata-rata sampel $\overline{X}$ padahal populasi asalnya berdistribusi miring?

2. Hitung probabilitas bahwa durasi rata-rata dari 64 panggilan tersebut berada di antara 11 hingga 13 menit ($P\left(11 \leq \overline{X} \leq 13\right)$)!

Pembahasan 5 (Bagian 1): Justifikasi Teorema Limit Pusat

Pertanyaan 1: Mengapa Pendekatan Normal Sah Digunakan?

Meskipun populasi durasi panggilan miring ke kanan, ukuran sampel yang diambil adalah:

$$n = 64 \geq 30$$

Berdasarkan **Teorema Limit Pusat (CLT)**, jika ukuran sampel $n \geq 30$, distribusi sampling dari rata-rata sampel $\overline{X}$ dijamin akan konvergen mendekati **Distribusi Normal**, terlepas dari betapa miringnya populasi asalnya!

Pembahasan 5 (Bagian 2): Standarisasi Z dan Peluang Interval

Pertanyaan 2: Menghitung $P\left(11 \leq \overline{X} \leq 13\right)$

Langkah 1: Hitung Galat Baku:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{64}} = \frac{4}{8} = 0{,}5 \text{ menit}$$

Langkah 2: Standarisasi Skor Z kedua batas:

$$z_1 = \frac{11-12}{0{,}5} = \text{-}2{,}00 \quad \text{dan} \quad z_2 = \frac{13-12}{0{,}5} = +2{,}00$$

Langkah 3: Hitung Selisih Luas Kurva Normal:

$$P\left(11 \leq \overline{X} \leq 13\right) = \Phi(2{,}00) - \Phi(\text{-}2{,}00) = 0{,}9772 - 0{,}0228 = 0{,}9544 \quad (95{,}44\%)$$

Catatan: Hasil $95{,}44\%$ ini persis konsisten dengan Aturan Empiris $\mu \pm 2\sigma_{\overline{X}}$.

Soal 6: Rata-Rata Pengundian Dadu Berulang ($n = 50$)

KASUS

Sebuah dadu bersisi enam yang seimbang dilempar berulang kali. Peubah acak mata dadu X memiliki rata-rata teoritis $\mu = 3{,}5$ dan simpangan baku $\sigma = \sqrt{\frac{35}{12}} \approx 1{,}7078$.

Pertanyaan

Jika dadu tersebut dilempar sebanyak $n = 50$ kali, berapakah probabilitas bahwa **rata-rata mata dadu yang muncul ($\overline{X}$)** bernilai lebih besar dari 3,8 ($P\left(\overline{X} > 3{,}8\right)$)?

Pembahasan 6: CLT pada Distribusi Diskrit Seragam

Solusi Langkah demi Langkah

1. Parameter Distribusi Sampling via CLT ($n = 50 > 30$):

$$\mu_{\overline{X}} = 3{,}5 \quad \text{dan} \quad \sigma_{\overline{X}} = \frac{1{,}7078}{\sqrt{50}} = \frac{1{,}7078}{7{,}0711} \approx 0{,}2415$$

2. Standarisasi Skor Z untuk $\overline{X} = 3{,}8$:

$$Z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}} = \frac{3{,}8 - 3{,}5}{0{,}2415} = \frac{0{,}3}{0{,}2415} \approx +1{,}24$$

3. Menghitung Peluang Ekor Kanan: Dari tabel normal: $\Phi(1{,}24) = 0{,}8925$.

$$P\left(\overline{X} > 3{,}8\right) = 1 - 0{,}8925 = 0{,}1075 \quad (10{,}75\%)$$

Kesimpulan: Ada peluang sekitar $10,75\%$ rata-rata 50 kali lemparan dadu melampaui 3,8.

Bagian D — Faktor Koreksi Populasi Terhingga (FPC)

Koreksi variabilitas saat ukuran sampel mencakup porsi besar populasi

Soal 7: Uji Batch Terbatas Paduan Titanium ($N = 250$)

KASUS

Sebuah pabrik kedirgantaraan memproduksi satu **batch** khusus berisi $N = 250$ lembar paduan titanium. Rata-rata ketebalan lembaran adalah $\mu = 5{,}0$ mm dengan simpangan baku $\sigma = 0{,}4$ mm. Tim inspeksi mengambil sampel acak sebanyak $n = 40$ lembar tanpa pengembalian.

Pertanyaan

1. Periksa apakah Faktor Koreksi Populasi Terhingga (**FPC**) perlu digunakan!

2. Hitung nilai **Standard Error** ($\sigma_{\overline{X}}$) dengan menyertakan faktor FPC!

3. Bandingkan hasilnya dengan nilai SE jika kita mengabaikan FPC!

Pembahasan 7: Penerapan Faktor Koreksi FPC

Solusi Langkah demi Langkah

1. Uji Fraksi Sampling:

$$\frac{n}{N} = \frac{40}{250} = 0{,}16 \quad (16\%) > 5\% \text{ ==> } \textbf{FPC Wajib Digunakan!}$$

2. Menghitung Nilai Pengali FPC:

$$\text{FPC} = \sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{250-40}{250-1}} = \sqrt{\frac{210}{249}} \approx 0{,}9184$$

3. Standard Error Terkoreksi vs Tanpa FPC:

- **Dengan FPC:** $\sigma_{\overline{X}} = \frac{0{,}4}{\sqrt{40}} \times 0{,}9184 = 0{,}0632 \times 0{,}9184 \approx$ **“0,0581“ “mm“**
- **Tanpa FPC:** $\sigma_{\overline{X}} = \frac{0{,}4}{\sqrt{40}} \approx 0{,}0632$ mm

Wawasan: Penggunaan FPC menurunkan kesalahan standar sebesar sekitar 8%.

Bagian E — Distribusi Sampling Proporsi ($\hat{p}$)

Inferensi data kategorik biner: E-Commerce, pemilu, dan kontrol cacat mutu

Soal 8: Rasio Konversi E-Commerce & Syarat CLT

KASUS

Data historis menunjukkan bahwa proporsi pengunjung web yang berbelanja (**conversion rate**) adalah $p = 0{,}08$ (8%). Seorang analis pemasaran memantau sampel acak berukuran $n = 400$ pengunjung hari ini.

Pertanyaan

1. Ujilah apakah syarat pendekatan normal CLT untuk proporsi terpenuhi!

2. Hitung rata-rata dan **Standard Error** dari proporsi sampel ($\sigma_{\hat{p}}$)!

3. Berapa probabilitas bahwa rasio konversi pada sampel tersebut mencapai **lebih dari 10%** ($P(\hat{p} > 0{,}10)$)?

Pembahasan 8 (Bagian 1): Uji Kelayakan CLT & Standard Error

Langkah 1 & 2: Syarat Normalitas & Galat Baku Proporsi

1. Uji Kelayakan Pendekatan Normal:

$$np = 400 \times 0{,}08 = 32 \geq 10 \quad \text{dan} \quad n(1-p) = 400 \times 0{,}92 = 368 \geq 10$$

Kedua syarat terpenuhi. Pendekatan kurva normal sah digunakan!

2. Rata-Rata dan Standard Error Proporsi:

- Pusat distribusi: $\mu_{\hat{p}} = p = 0{,}08$
- Galat Baku Proporsi:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}08 \times 0{,}92}{400}} = \sqrt{\frac{0{,}0736}{400}} \approx 0{,}01356$$

Pembahasan 8 (Bagian 2): Standarisasi Z dan Peluang Konversi

Langkah 3: Menghitung Peluang $P(\hat{p} > 0{,}10)$

Standarisasi Skor Z:

$$Z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0{,}10 - 0{,}08}{0{,}01356} = \frac{0{,}02}{0{,}01356} \approx +1{,}47$$

Evaluasi Tabel Normal: Dari tabel distribusi normal kumulatif: $\Phi(1{,}47) = 0{,}9292$.

$$P(\hat{p} > 0{,}10) = 1 - 0{,}9292 = 0{,}0708 \quad (7{,}08\%)$$

Kesimpulan: Ada peluang sekitar **7,08%** rasio konversi harian menembus 10%.

Soal 9: Survei Elektabilitas Pemilu (Quick Count)

KASUS

Dalam pemilihan kepala daerah, seorang kandidat memiliki elektabilitas riil di populasi sebesar $p = 0{,}52$ (52%). Sebuah lembaga survei mengambil sampel acak sebanyak $n = 1.000$ pemilih.

Pertanyaan

Berapa peluang survei tersebut menghasilkan kesimpulan yang keliru bahwa sang kandidat diprediksi kalah (yaitu proporsi suara sampel kurang dari 50%, $P(\hat{p} < 0{,}50)$)?

Pembahasan 9: Menghitung Risiko Misklaim Survei Pemilu

Solusi Langkah demi Langkah

1. Hitung Standard Error Proporsi ($n = 1.000$):

$$\sigma_{\hat{p}} = \sqrt{\frac{0{,}52 \times 0{,}48}{1.000}} = \sqrt{\frac{0{,}2496}{1.000}} \approx 0{,}0158 \quad (1{,}58\%)$$

2. Standarisasi Ambang Batas Kalah ($\hat{p} = 0{,}50$):

$$Z = \frac{0{,}50 - 0{,}52}{0{,}0158} = \frac{\text{-}0{,}02}{0{,}0158} \approx \text{-}1{,}27$$

3. Menghitung Peluang Ekor Kiri:

$$P(\hat{p} < 0{,}50) = \Phi(\text{-}1{,}27) = 1 - 0{,}8980 = 0{,}1020 \quad (10{,}20\%)$$

Wawasan: Meskipun kandidat unggul (52%), sampel 1.000 masih memiliki risiko **10,20%** salah memprediksi kekalahan karena keunggulannya tipis.

Soal 10: Uji Kontrol Cacat Mutu Semikonduktor

KASUS

Sebuah pabrik mikroprosesor mengklaim bahwa tingkat cacat produksi (**defect rate**) chip buatannya hanya $p = 0{,}02$ (2%). Dalam audit mutu independen terhadap $n = 500$ chip acak, ditemukan sebanyak 16 chip cacat ($\hat{p} = \frac{16}{500} = 0{,}032$ atau 3,2%).

Pertanyaan

1. Hitung skor Z dari hasil observasi audit mutu tersebut!

2. Berapakah peluang menemukan 16 atau lebih chip cacat jika klaim pabrik ($p = 0{,}02$) benar ($P(\hat{p} \geq 0{,}032)$)?

3. Apakah klaim produsen patut dicurigai?

Pembahasan 10: Pengantar Intuisi Uji Hipotesis

Solusi Langkah demi Langkah

1. Hitung SE dan Skor Z: Asumsikan klaim pabrik benar ($p = 0{,}02$):

$$\sigma_{\hat{p}} = \sqrt{\frac{0{,}02 \times 0{,}98}{500}} = \sqrt{\frac{0{,}0196}{500}} \approx 0{,}00626$$

$$Z = \frac{0{,}032 - 0{,}02}{0{,}00626} = \frac{0{,}012}{0{,}00626} \approx +1{,}92$$

2. Menghitung Peluang (p-value awal):

$$P(\hat{p} \geq 0{,}032) = 1 - \Phi(1{,}92) = 1 - 0{,}9726 = 0{,}0274 \quad (2{,}74\%)$$

3. Kesimpulan Ilmiah: Peluang munculnya 16 cacat jika klaim benar hanya **2,74%** ($< 5\%$). Ini bukti empiris kuat bahwa **klaim pabrik patut dicurigai** (cikal bakal Uji Hipotesis P10!).

Ringkasan Formula Penting Distribusi Sampling

Distribusi Sampling Rata-Rata ($\overline{X}$)

- **Rata-Rata:** $\mu_{\overline{X}} = \mu$ (Tak Bias)
- **Standard Error:** $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$
- **Dengan FPC ($\frac{n}{N} > 5\%$):**
$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$
- **CLT ($n \geq 30$):** $\overline{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right)$
- **Standarisasi Z:** $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$

Distribusi Sampling Proporsi ($\hat{p}$)

- **Proporsi Sampel:** $\hat{p} = \frac{X}{n}$
- **Rata-Rata:** $\mu_{\hat{p}} = p$ (Tak Bias)
- **Standard Error:** $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
- **Syarat Normal:** $np \geq 10$ & $n(1-p) \geq 10$
- **Standarisasi Z:** $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$