DISTRIBUSI KONTINU

Prof. Dr. Arna Fariza, S.Kom., M.Kom.



Topik Bahasan

1. Recap: Variabel Acak Kontinu
2. Pengertian Distribusi Normal
3. Sifat-Sifat Distribusi Normal
4. Aturan Empiris (68-95-99.7)
5. Standardisasi (Z-score)
6. Penggunaan Tabel Normal Standar
7. Aplikasi di Bidang Teknik Informatika
8. Contoh Soal & Latihan
9. Kesimpulan

RECAP - VARIABEL ACAK KONTINU

Perbedaan Diskrit vs Kontinu:

Diskrit	Kontinu
Nilai terbatas / terhitung	Nilai dalam rentang tak terbatas
Diperoleh dari menghitung	Diperoleh dari mengukur
$P(X = x) > 0$	$P(X = x) = 0$
Contoh: Jumlah bug	Contoh: Waktu eksekusi (detik)

Kunci Penting:
Pada variabel acak kontinu, kita **tidak** menghitung peluang satu titik (misal $P(X = 5)$), melainkan menghitung peluang **rentang/interval** (misal $P(4 < X < 6)$).
Distribusi kontinu yang paling penting: Distribusi Normal!

PENGERTIAN DISTRIBUSI NORMAL

Definisi:
Distribusi Normal (Gaussian Distribution) adalah distribusi probabilitas kontinu yang paling banyak digunakan dalam statistika karena banyak fenomena alam dan sosial yang mengikuti pola distribusi ini.

Karakteristik Visual:
- Berbentuk **lonceng simetris** (bell curve).
- Sumbu X: Nilai variabel.
- Sumbu Y: Fungsi kepadatan probabilitas (densitas).

Notasi:

$$X \sim N(\mu, \sigma^2)$$

- μ = rata-rata (mean) populasi
- σ = standar deviasi populasi
- σ^2 = varians populasi

Fungsi Densitas:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

SIFAT-SIFAT DISTRIBUSI NORMAL

Sifat 1: Simetris

Kurva simetris terhadap garis vertikal $x = \mu$ (mean).

- 50% area berada di kiri mean.
- 50% area berada di kanan mean.

Sifat 2: Mean = Median = Modus

Ketiga ukuran pemusatan data ini berada tepat di puncak kurva.

Sifat 3: Total Area = 1

Luas seluruh area di bawah kurva adalah 1 (atau 100%). Peluang dinyatakan sebagai area di bawah kurva.

Sifat 4: Mendekati Nol di Ekor

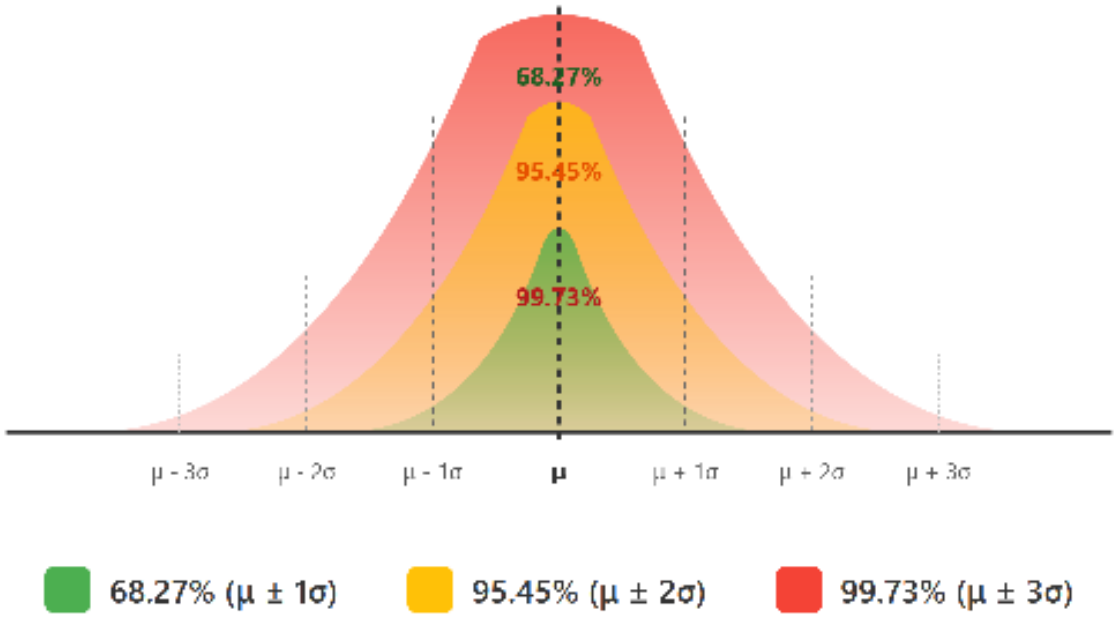
Kurva mendekati sumbu X tetapi tidak pernah menyentuhnya (asimtotik). Nilai yang sangat jauh dari mean memiliki peluang sangat kecil.

Sifat 5: Ditentukan oleh μ dan σ

- μ menentukan **lokasi** pusat kurva.
- σ menentukan **lebar/tinggi** kurva (semakin besar σ, kurva semakin lebar dan rendah).

ATURAN EMPIRIS (68-95-99.7 RULE)

Aturan Praktis pada Distribusi Normal:



Semakin jauh dari rata-rata (μ), peluang terjadinya semakin kecil

Interpretasi:

- $\mu \pm 1\sigma$ mencakup sekitar **68.27%** data.
- $\mu \pm 2\sigma$ mencakup sekitar **95.45%** data.
- $\mu \pm 3\sigma$ mencakup sekitar **99.73%** data.

Aplikasi IT:

Jika waktu respons server memiliki mean 100 ms dan standar deviasi 10 ms, maka 95% request akan merespons antara 80 ms hingga 120 ms.

MENGAPA PERLU STANDARDISASI?

Masalah:
Setiap fenomena memiliki μ(mean) dan σ(standar deviasi) yang berbeda.

- Waktu respons Server A: $\mu = 100$ms, $\sigma = 10$ms
- Waktu respons Server B: $\mu = 500$ms, $\sigma = 50$ms

Pertanyaan: Bagaimana cara membandingkan performa keduanya secara adil?

Solusi: Standardisasi (Z-score)
Mengubah distribusi normal apa pun (dengan μdan σberapa pun) menjadi **Distribusi Normal Standar** dengan $\mu = 0$ dan $\sigma = 1$.

Notasi Distribusi Normal Standar: $Z \sim N(0, 1)$

RUMUS Z-SCORE

Rumus Standardisasi:

$$Z = \frac{X - \mu}{\sigma}$$

Keterangan:

- Z= skor standar (Z-score)
- X= nilai variabel yang ingin diukur
- μ= rata-rata populasi
- σ= standar deviasi populasi

Interpretasi Z-score:

Z-score mengukur **berapa banyak standar deviasi** suatu nilai X dari rata-rata.

- $Z = 0 \rightarrow$ Nilai Xsama dengan mean.
- $Z > 0 \rightarrow$ Nilai Xdi atas mean.
- $Z < 0 \rightarrow$ Nilai Xdi bawah mean.
- $Z = 1.5 \rightarrow$ Nilai Xberjarak 1.5 standar deviasi di atas mean.

CONTOH PERHITUNGAN Z-SCORE

Kasus: Waktu Respons Server
Sebuah server memiliki waktu respons berdistribusi normal dengan:

- Rata-rata (μ) = 50 ms
- Standar deviasi (σ) = 5 ms

Hitung Z-score untuk:
a) $X = 60$ms
b) $X = 45$ms
c) $X = 50$ms

Penyelesaian:
a) $Z = \frac{60-50}{5} = \frac{10}{5} = \mathbf{2.0}$
Artinya: Waktu respons 60 ms berada **2 standar deviasi di atas** rata-rata.

b) $Z = \frac{45-50}{5} = \frac{-5}{5} = \mathbf{-1.0}$
Artinya: Waktu respons 45 ms berada **1 standar deviasi di bawah** rata-rata.

c) $Z = \frac{50-50}{5} = \frac{0}{5} = \mathbf{0}$
Artinya: Waktu respons 50 ms sama dengan rata-rata.

PENGGUNAAN TABEL NORMAL STANDAR

Tabel Z (Tabel Normal Standar)
Tabel ini memberikan nilai **luas area di bawah kurva normal standar dari** $-\infty$ **hingga** Z (peluang kumulatif $P(Z < z)$).

Cara Membaca Tabel Z:

Contoh 1: Mencari $P(Z < 1.25)$

1. Lihat baris **1.2** di kolom paling kiri.
2. Lihat kolom **0.05** di baris paling atas.
3. Pertemuan baris 1.2 dan kolom 0.05 adalah nilai **0.8944**.

Hasil: $P(Z < 1.25) = 0.8944$ atau **89.44%**
Ilustrasi Tabel Z (Potongan):

Z	0.00	0.01	...	0.05
1.2	0.8849	0.8869	...	**0.8944**
1.3	0.9032	0.9049	...	0.9115

CONTOH PENGGUNAAN TABEL (KIRI & KANAN)

1. Peluang Kurang Dari ($Z < z$)

$$P(Z < 1.50) = 0.9332$$

Artinya: Peluang nilai berada di bawah 1.50 adalah 93.32%.

2. Peluang Lebih Dari ($Z > z$)

Gunakan sifat komplemen: $P(Z > z) = 1 - P(Z < z)$

$P(Z > 1.50) = 1 - 0.9332 = \mathbf{0.0668}$(6.68%)

3. Peluang Antara Dua Nilai ($a < Z < b$)

$$P(-1.00 < Z < 1.50) = P(Z < 1.50) - P(Z < -1.00)$$

- $P(Z < 1.50) = 0.9332$
- $P(Z < -1.00) = 0.1587$
- Hasil = $0.9332 - 0.1587 = \mathbf{0.7745}$(77.45%)

4. Peluang di Luar Rentang

$$P(|\ Z\ | > 1.96) = P(Z > 1.96) + P(Z < -1.96)$$

- $P(Z < -1.96) = 0.0250$
- $P(Z > 1.96) = 1 - 0.9750 = 0.0250$
- Total = $0.0250 + 0.0250 = \mathbf{0.0500}$(5%)

📋 **Tabel Normal Standar (Potongan)**

Z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944
1.5	**0.9332**	0.9345	0.9357	0.9370	0.9382	0.9394
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798

APLIKASI DI BIDANG TEKNIK INFORMATIKA

1. Analisis Kinerja Sistem (QoS)
- Mengukur peluang waktu respons server melebihi SLA (Service Level Agreement).
- Contoh: Berapa peluang respons > 150 ms jika mean = 100 ms, σ = 20 ms?

2. Deteksi Anomali (Outlier)
- Nilai dengan Z-score > 3 atau < -3 dianggap anomali.
- Mendeteksi serangan siber atau error sistem yang tidak wajar.

3. Kualitas Jaringan
- Memodelkan latensi (latency) dan packet loss.
- Menentukan kapasitas bandwidth yang dibutuhkan.

4. Machine Learning
- Normalisasi fitur (Feature Scaling) menggunakan Z-score.
- Algoritma yang mengasumsikan data berdistribusi normal (LDA, Naive Bayes Gaussian).

STUDI KASUS TERAPAN

Kasus: Server Web dan SLA
Sebuah server memiliki waktu respons berdistribusi normal dengan:

- Mean (μ) = 120 ms
- Standar deviasi (σ) = 15 ms

Pertanyaan:
a) Berapa peluang waktu respons kurang dari 100 ms?
b) Berapa peluang waktu respons lebih dari 150 ms?
c) Jika SLA menetapkan waktu respons maksimal 150 ms, berapa persen request yang melanggar SLA?

Penyelesaian:
a) $Z = \frac{100-120}{15} = -1.33$
$P(Z < -1.33) = 0.0918$atau **9.18%**
b) $Z = \frac{150-120}{15} = 2.00$
$P(Z > 2.00) = 1 - P(Z < 2.00) = 1 - 0.9772 = 0.0228$atau **2.28%**
c) Pelanggaran SLA terjadi jika respons > 150 ms.
Berdasarkan perhitungan (b), **2.28%** request melanggar SLA.

LATIHAN SOAL

Soal 1:
Sebuah sistem memiliki waktu pemrosesan data berdistribusi normal dengan rata-rata 200 detik dan standar deviasi 25 detik.
a) Berapa peluang waktu pemrosesan kurang dari 180 detik?
b) Berapa peluang waktu pemrosesan antara 190 dan 230 detik?

Soal 2:
Nilai ujian statistika mahasiswa berdistribusi normal dengan mean 70 dan standar deviasi 8.
a) Berapa peluang mahasiswa mendapat nilai lebih dari 85?
b) Jika 10% mahasiswa teratas mendapat nilai A, berapa nilai minimum untuk mendapat A?

Soal 3 (Aplikasi IT):
Jumlah paket data yang diproses oleh router per detik berdistribusi normal dengan rata-rata 5000 paket dan standar deviasi 400 paket.
a) Berapa peluang router memproses lebih dari 5800 paket (berpotensi overload)?
b) Berapa peluang router memproses kurang dari 4500 paket?

JAWABAN LATIHAN SOAL

Jawaban Soal 1:

a) $Z = \frac{180-200}{25} = -0.80$

$P(Z < -0.80) = 0.2119$atau **21.19%**

b) $Z_1 = \frac{190-200}{25} = -0.40 \Rightarrow P(Z < -0.40) = 0.3446$

$$Z_2 = \frac{230-200}{25} = 1.20 \Rightarrow P(Z < 1.20) = 0.8849$$

$P(-0.40 < Z < 1.20) = 0.8849 - 0.3446 = 0.5403$atau **54.03%**

Jawaban Soal 2:

a) $Z = \frac{85-70}{8} = 1.875 \approx 1.88$

$P(Z > 1.88) = 1 - 0.9699 = 0.0301$atau **3.01%**

b) Mencari nilai Z untuk area 0.90 (karena 10% teratas berarti 90% di bawahnya).

$$P(Z < z) = 0.90 \Rightarrow z \approx 1.28$$

$$X = \mu + z\sigma = 70 + (1.28)(8) = 70 + 10.24 = \mathbf{80.24}$$

Nilai minimum untuk A adalah **80.24**.

Jawaban Soal 3:

a) $Z = \frac{5800-5000}{400} = \frac{800}{400} = 2.00$

$P(Z > 2.00) = 1 - 0.9772 = 0.0228$atau **2.28%**

b) $Z = \frac{4500-5000}{400} = -1.25$

$P(Z < -1.25) = 0.1056$atau **10.56%**

RINGKASAN MATERI

1. **Distribusi Normal** adalah distribusi kontinu berbentuk lonceng simetris yang paling banyak digunakan.
2. **Sifat:** Simetris terhadap mean, total area = 1, ditentukan oleh μdan σ.
3. **Aturan Empiris:** 68% data dalam 1σ, 95% dalam 2σ, 99.7% dalam 3σ.
4. **Z-score** = $(X-\mu)/\sigma\rightarrow$ mengubah distribusi apa pun menjadi Normal Standar $N(0,1)$.
5. **Tabel Normal** digunakan untuk mencari peluang kumulatif $P(Z<z)$.
6. **Aplikasi IT:** QoS, deteksi anomali, normalisasi data ML, analisis jaringan.

TUGAS

Soal 1:
Sebuah aplikasi memiliki waktu loading berdistribusi normal dengan mean 3.5 detik dan standar deviasi 0.8 detik.
a) Berapa peluang aplikasi loading kurang dari 2.5 detik?
b) Berapa peluang aplikasi loading lebih dari 5 detik?

Soal 2:
Skor ujian masuk IT berdistribusi normal dengan mean 600 dan standar deviasi 80.
a) Jika batas kelulusan adalah 500, berapa persen peserta yang lulus?
b) Jika hanya 5% peserta terbaik yang diterima, berapa skor minimum yang dibutuhkan?

Soal 3 (Aplikasi IT):
Sebuah jaringan memiliki latensi berdistribusi normal dengan rata-rata 40 ms dan standar deviasi 10 ms.
a) Berapa peluang latensi melebihi 65 ms?
b) Berapa peluang latensi antara 30 ms dan 50 ms?