



Institut Informatika & Bisnis  
**DARMAJAYA**  
Keunggulan. Afiliasi. Mutu.



Kampus  
Merdeka  
INDONESIA JAYA

**MERDEKA  
BELAJAR**

# Statistika untuk Sains Data

SSD23401

PERTEMUAN 5

Egi Safitri, S.Mat., M.Si

`egisafitri@ darmajaya.ac.id`

October 26th, 2023

## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

## 1 Distribusi Data

- Pengantar Distribusi Data
- Jenis-Jenis Distribusi Data
- Distribusi Normal
- Distribusi Binomial
- Distribusi Poisson





## Distribusi Data

### Pengantar Distribusi Data

#### Jenis-Jenis Distribusi Data

##### Distribusi Normal

##### Distribusi Binomial

##### Distribusi Poisson

## Definisi

Distribusi data adalah cara data terdistribusi atau tersebar di sepanjang berbagai nilai. Ini memberikan gambaran tentang bagaimana data terbagi atau terkumpul di berbagai nilai atau interval.

Distribusi data mencakup informasi tentang frekuensi atau probabilitas munculnya berbagai nilai atau kelompok nilai dalam kumpulan data.

Dengan kata lain, itu adalah cara data didistribusikan di sepanjang sumbu nilainya. Beberapa jenis distribusi data yang umum termasuk distribusi normal, distribusi binomial, distribusi Poisson, dan banyak lainnya. Distribusi data adalah konsep penting dalam statistika untuk menganalisis dan memahami data.





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

- Distribusi Normal (Gaussian)
- Distribusi Binomial
- Distribusi Poisson
- Distribusi Eksponensial
- Dan banyak lagi...





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

## Visualisasi

- Histogram
- Diagram Pencar (Scatter Plot)
- Kurva Distribusi
- Box Plot

## Parameter

- Mean (Rata-rata)
- Variance (Varians)
- Skewness (Kemiringan)
- Kurtosis (Kurtosis)





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

## Definisi

Berdasarkan Investopedia, distribusi normal adalah sebuah distribusi probabilitas yang digambarkan dengan kurva lonceng simetris yang melandai di bagian tepi dan memuncak di bagian tengah. Tipe distribusi ini biasanya digunakan untuk menggambarkan pola penyebaran data dari variabel kontinu seperti tinggi badan, berat badan, atau tekanan darah.





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

Distribusi normal berguna untuk menggambarkan seberapa sering sesuatu terjadi dalam rentang tertentu.

## Contoh

Jika dilakukan pengukuran tinggi badan suatu populasi yang besar, kemungkinan data yang terkumpul akan mengikuti pola distribusi normal, di mana sebagian besar memiliki tinggi rata-rata, sementara jumlah orang yang sangat pendek atau sangat tinggi lebih sedikit





# Parameter Distribusi Normal

## Mean

Mean, dilambangkan dengan  $\mu$ , adalah pusat distribusi yang juga dikenal sebagai nilai rata-rata. Dalam distribusi normal, mean adalah nilai di mana penyebaran data berpusat dan berada di puncak kurva.

## Standar Deviasi

Standar deviasi, dilambangkan dengan  $\sigma$ , mengukur penyebaran data dari mean dan menunjukkan seberapa banyak data yang menyimpang dari mean.

- 1 Jika suatu data memiliki standar deviasi yang lebih kecil, berarti datanya mendekati nilai rata-rata dan bentuk kurva semakin sempit.
- 2 Sebaliknya, jika standar deviasi datanya besar, berarti datanya lebih tersebar menjauh dari rata-rata dan bentuk kurva melebar.





## Distribusi Data

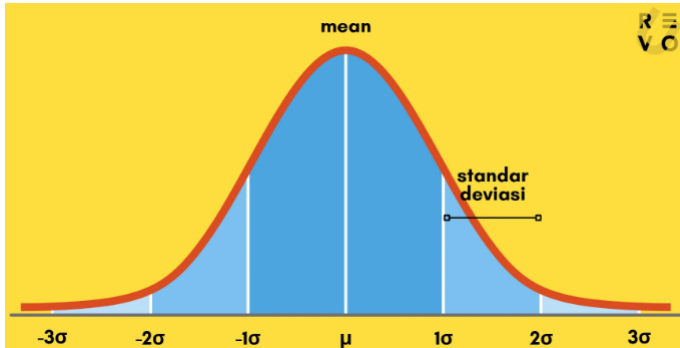
Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

### 1 Kurva berbentuk lonceng

Distribusi normal adalah jenis distribusi probabilitas yang memiliki bentuk simetris menyerupai kurva lonceng. Hal ini menunjukkan bahwa sebagian besar data mengelompok di sekitar rata-rata, dan distribusinya merata di kedua sisi.

### 2 Nilai mean, median, dan modus yang sama

Median merupakan titik tengah distribusi dan modus adalah nilai yang paling sering muncul dalam distribusi. Dalam distribusi normal, mean, median, dan modus sama nilainya karena bentuknya simetris di sekitar mean.

### 3 Tidak cocok untuk data dengan distribusi miring

Data berdistribusi miring memiliki bentuk yang tidak simetris dan tidak cocok untuk pemodelan dengan distribusi normal.





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

Distribusi miring memiliki ekor yang lebih panjang di satu sisi daripada sisi lainnya, menunjukkan bahwa datanya tidak terdistribusi secara merata. Dengan demikian, data tersebut perlu menggunakan jenis distribusi lain yang sesuai untuk memodelkan data miring.

### 4 **Cocok dengan Variabel Kontinu**

Distribusi normal cocok untuk memodelkan variabel kontinu, yang merupakan data dengan jumlah nilai yang tak terbatas. Sehingga pemanfaatannya sangat luas untuk menganalisis data di berbagai keperluan, seperti penelitian ekonomi dan ilmu sosial.





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

Aturan empiris distribusi normal, atau diketahui juga sebagai aturan 68-95-99.7, adalah sebuah cara yang bermanfaat untuk memahami distribusi data dalam distribusi normal.

Aturan ini dapat digunakan untuk memperkirakan probabilitas dan membuat prediksi berdasarkan karakteristik populasi.

Aturan ini menyatakan bahwa:

- 1 Sekitar 68% dari data termasuk dalam satu standar deviasi dari rata-rata.
- 2 Sekitar 95% data termasuk dalam dua standar deviasi dari rata-rata.
- 3 Sekitar 99,7% data termasuk dalam tiga standar deviasi dari rata-rata.





# Aturan Empiris untuk Distribusi Normal

## Distribusi Data

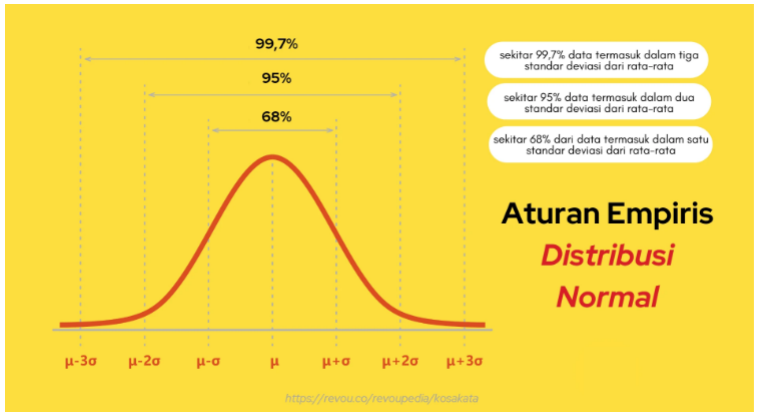
Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

## Definisi

Adalah suatu distribusi teoritis yang memakai var random diskrit, yaitu banyaknya hasil percobaan yang terjadi dalam suatu interval waktu tertentu.

## Ciri-ciri distribusi binomial:

- 1 Banyaknya hasil percobaan yang satu tidak tergantung dari banyaknya hasil percobaan yang lain.
- 2 Probabilitas hasil percobaan sebanding dengan panjang interval waktu.
- 3 Probabilitas lebih dari satu hasil percobaan yang terjadi dalam interval waktu yang singkat dalam daerah yang kecil dapat diabaikan.



## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

Berikut adalah parameter-parameter dalam distribusi binomial:

- 1  $n$  (jumlah percobaan): Parameter ini menunjukkan jumlah total percobaan atau uji coba yang akan dilakukan.
- 2  $p$  (probabilitas keberhasilan): Parameter ini mengindikasikan probabilitas keberhasilan dalam setiap percobaan tunggal. Dalam distribusi binomial, setiap percobaan harus memiliki dua kemungkinan hasil: sukses atau gagal.
- 3  $X$  (jumlah keberhasilan): Parameter ini mengacu pada jumlah keberhasilan yang Anda ingin hitung probabilitasnya dalam  $n$  percobaan.
- 4 Probabilitas kegagalan ( $q$ ): Ini adalah probabilitas kegagalan dalam setiap percobaan dan dihitung sebagai  $q = 1 - p$ . Probabilitas kegagalan digunakan dalam perhitungan distribusi binomial.



## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

Distribusi binomial memiliki beberapa karakteristik utama yang menggambarkan sifat dan perilakunya. Berikut adalah karakteristik utama dari distribusi binomial:

- 1 Jumlah Percobaan Terbatas ( $n$ )
- 2 Hasil Percobaan Bernilai Biner
- 3 Probabilitas Keberhasilan Tetap ( $p$ )
- 4 Independen dan Identik
- 5 Variabel Acak Diskrit
- 6 Parameter  $n$  dan  $p$
- 7 Rata-rata ( $n \times p$ ) dan Varians ( $n \times p \times (1 - p)$ )





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

## Definisi

Distribusi probabilitas yang menggambarkan jumlah peristiwa yang terjadi dalam interval waktu atau ruang tertentu dengan laju kejadian yang konstan, jika peristiwa-peristiwa tersebut terjadi secara acak dan independen satu sama lain.

## Parameter

Dalam distribusi Poisson, terdapat satu parameter utama, yaitu laju kejadian ( $\lambda$ ). Laju kejadian ( $\lambda$ ) mengindikasikan jumlah peristiwa yang diharapkan terjadi dalam interval yang ditentukan. Distribusi Poisson sangat berguna dalam situasi di mana peristiwa-peristiwa tersebut jarang terjadi, tetapi memiliki tingkat kejadian yang stabil.





## Distribusi Data

Pengantar Distribusi Data

Jenis-Jenis Distribusi Data

Distribusi Normal

Distribusi Binomial

Distribusi Poisson

# Karakteristik Distribusi Poisson

Karakteristik utama dari distribusi Poisson adalah sebagai berikut:

- 1 Variabel acak dalam distribusi Poisson adalah jumlah peristiwa yang terjadi dalam interval yang ditentukan.
- 2 Peristiwa-peristiwa tersebut terjadi secara acak dan independen.
- 3 Laju kejadian ( $\lambda$ ) adalah tetap dan dikenal.
- 4 Distribusi Poisson adalah distribusi probabilitas diskrit, yang berarti variabel acak dalam distribusi ini hanya dapat mengambil nilai bilangan bulat.
- 5 Distribusi Poisson digunakan untuk memodelkan situasi di mana peristiwa langka terjadi dalam interval waktu atau ruang tertentu.



Thank You.

Egi Safitri, S.Mat., M.Si

